



Ю. М. АСТАПОВ
Д. В. ВАСИЛЬЕВ
Ю. И. ЗАЛОЖНЕВ

ТЕОРИЯ ОПТИКО- ЭЛЕКТРОННЫХ СЛЕДЯЩИХ СИСТЕМ

Spencer

ББК 32.81

А91

УДК 681.5

✓

Астапов Ю.М., Васильев Д.В., Заложнев Ю.И. Теория оптико-электронных следящих систем. — М.: Наука. Гл. ред. физ.-мат. лит., 1988 — 328 с. — ISBN 5-02-013886-X

Излагаются методы исследования систем сопровождения подвижных объектов с использованием оптико-электронной аппаратуры для выделения сигнала рассогласования. Следящие системы этого типа рассматриваются как сложные динамические устройства с нелинейными элементами при взаимном влиянии каналов управления. Анализируется точность сопровождения при наличии возмущений, исследуются периодические движения следящих координаторов, явление срыва сопровождения. Значительное место уделяется методам математического моделирования различных режимов работы. Для сокращения затрат машинного времени приводятся типы математических моделей различной сложности, основанные на теории идентификации.

Для специалистов в области управления, в частности управления движущимися объектами.

Табл. 20. Ил. 101. Библиогр. 180 назв.

Рецензент

доктор технических наук *И.Н. Белоглазов*

А $\frac{1502000000-110}{053(02)-88}$ 158-88

ISBN 5-02-013886-X

© Издательство "Наука".
Главная редакция
физико-математической
литературы, 1988

ОГЛАВЛЕНИЕ

Предисловие	7
Введение	9
<i>Глава 1. Принципы обработки входной информации в оптико-электронных слеящих системах</i>	<i>19</i>
§ 1.1. Обобщенные функциональные и структурные схемы оптико-электрон- ных слеящих систем. Примеры реализации	19
Состав оптико-электронной слеящей системы (20). Контрастные дискриминаторы в ОЭСС (20)	
§ 1.2. Корреляционные дискриминаторы сдвига изображений	25
Роль и место корреляционных дискриминаторов в ОЭСС и других слеящих автоматах (25). Функциональные признаки и структура КД в многомерном слеящем автомате (28)	
§ 1.3. Синтез корреляционного дискриминатора по алгоритму оптималь- ного многомерного оценивания	31
Постановка задачи (31). Проблема синтеза КД в свете обобщенного метода наименьших квадратов (32). Исследование обобщенного алгоритма многомерного оценивания по его представлениям (36). Использование дуальных решений в алгоритме КД (37). Об алго- ритмах КД непрерывных сигналов (40). Нормирование оценок и компенсация перекрестных связей в алгоритмах многомерных КД (41).	
§ 1.4. Двумерная корреляция и ее оценка при краевых эффектах.	44
Краевые эффекты в оценках двумерных корреляций изображений из-за ограниченности апертуры ОЭС (44). Особенности двумерных спектрально-корреляционных характеристик финитных изображе- ний (47). Действие линейных фильтров на двумерные спектраль- но-корреляционные характеристики (49)	
§ 1.5. Обобщенная корреляция изображений и ее применение	51
Групповые и обобщенные сдвиги в описаниях проективных преоб- разований плоских изображений (51). Представления ракурсного и "масштабно-кенового" сдвигов изображений в ОЭС (55). Обоб- щенные корреляционные функции изображений (56). Обобщенные свертки и спектры (58). Обобщения теоремы отсчетов, необходи- мые для измерения неаддитивных сдвигов изображений (59). При- мер обобщенной АКФ изображения (63)	

§ 1.6.	Особенности аналоговых корреляционных дискриминаторов	64
	Преобразования координат и операторов линейной фильтрации при построчной развертке изображений в ОЭС (64). О технических помехах из-за смены направлений развертки при разложении изображений (66). Сравнительный анализ вариантов одномерного аналогового КД (66). Корреляционный дискриминатор с ортогонализацией сигналов как универсальная реализация одномерного алгоритма оптимального оценивания (69). Влияние неидеальности ФЧХ каналов КД на его характеристику (72). Выбор и расчет ДХ в КД с ортогональными фильтрами первого порядка (73). "Предельная" выходная характеристика линейного КД (76)	
§ 1.7.	Нелинейные преобразования и стабилизация характеристик в корреляционных дискриминаторах	78
	Назначение и разновидности нелинейных преобразований в ОЭС (78). О применимости алгоритмов КД в системах с морфологическими преобразователями изображений (79). Дестабилизирующие факторы и общие принципы компенсации их влияния в КД (81). Алгоритм и структура КД со стабилизацией крутизны выходной характеристики (82). Оценка динамических свойств КД с АРУ, стабилизирующей крутизну выходной характеристики (84)	
§ 1.8.	Перекрестные связи в корреляционных дискриминаторах из-за анизотропности изображений.	89
	Анизотропность изображений как причина перекрестных связей в двумерном КД (89). Методы компенсации перекрестных связей в двумерных КД (92)	
§ 1.9.	Действие мешающих факторов на характеристики корреляционных дискриминаторов.	93
	Классификация основных мешающих факторов и модель их взаимодействия с сигналами в КД изображений (93). Флуктуационные характеристики одномерных КД (95). Влияние неравномерностей фона и чувствительности по полю зрения в каналах на характеристики одномерного КД изображений (98)	
Глава 2.	Аналитические методы исследования оптико-электронных следящих систем.	100
§ 2.1.	Формирование сигнала рассогласования в ОЭСС с непрерывным управлением.	100
	Анализаторы изображения с одноплощадочным приемником (100). Резонансные усилители (102). Амплитудные детекторы (103). Фазовые детекторы. Автоматическая регулировка усиления (104)	
§ 2.2.	Формирование сигнала рассогласования в ОЭСС с импульсным управлением	105
	Анализаторы изображения с многоплощадочным приемником (105)	
§ 2.3.	Математические модели исполнительных элементов ОЭСС	108
	Использование электрических микродвигателей (108). Индикаторный гирос привод (111). Гирос привод с разгрузкой от внешних возмущений (114)	
§ 2.4.	Анализ периодических режимов в ОЭСС с непрерывным управлением .	115
	Формальные условия возникновения кругового предельного цикла (116). Метод малого параметра (118). Аппроксимация нелинейной характеристики (121). Другие виды возможных движений следящего координатора (123)	

§ 2.5.	Анализ устойчивости сопровождения в ОЭСС с импульсным управлением	127
	Уравнения ОЭСС с АИМ (127). Область устойчивости ОЭСС с АИМ (130). Уравнения ОЭСС с ШИМ (132). Линеаризация уравнений (134)	
§ 2.6.	Анализ процесса сопровождения при двойном источнике излучения . .	138
	Точки равновесия в поле сил коррекции (140). Построение поля сил коррекции (141). Обобщенная модель сопровождения двойного источника (145). Нестационарное решение и оценка помехозащищенности (148). Возможное упрощение вычислений (153)	
§ 2.7.	Анализ процесса сопровождения источника с распределенным излучением	156
	Искажение формы пеленгационной характеристики (158). Спектральная плотность управляющего сигнала (159). Динамическая ошибка при наличии автоколебаний (161). Динамическая ошибка при подавлении автоколебаний (165)	
Глава 3.	Методы математического моделирования оптико-электронных следящих систем.	172
§ 3.1.	Особенности математического моделирования ОЭСС	172
	Виды моделирования (172). Типы моделей (174). Основные элементы моделей оптико-электронных систем (176)	
§ 3.2.	Математическое моделирование преобразования сигнала в оптико-электронном тракте	177
	Моделирование изображения в фокальной плоскости (177). Моделирование анализаторов изображения (186). Моделирование электрического сигнала на выходе фотоприемного устройства (196). Статистические характеристики сигнала на выходе сканирующей оптической системы (200)	
§ 3.3.	Моделирование цифровых изображений.	204
	Дискретизация и квантование (204). Дискретизация непрерывных изображений (204). Квантование дискретных отсчетов (206). Расчет количества уровней квантования по корреляционному критерию (206). Кодирование цифровых изображений в пространственной области (209). Специальные случаи кодирования (209). Дифференциальная импульсно-кодовая модуляция (211). Кодирование изображений путем преобразований (212). Обзор ортогональных преобразований (213)	
§ 3.4.	Математическое моделирование обработки сигнала в электронном тракте ОЭС.	215
	Функциональные звенья электронного тракта ОЭС (215). Моделирование линейных звеньев с использованием формул численного интегрирования (218). Замена непрерывных линейных систем эквивалентными импульсными (223). Моделирование линейных элементов ОЭСС на основе рекуррентных соотношений (231). Моделирование нелинейных звеньев ОЭС (233)	
§ 3.5.	Математическое моделирование флуктуаций сигнала оптико-электронных систем	235
	Статистические характеристики случайных сигналов, используемые при моделировании (235). Моделирование случайных величин и случайных процессов (236). Моделирование двумерных случайных полей (245)	

§ 3.6.	Построение математической модели по экспериментальным данным Оценки параметров модели (251). Планирование эксперимента для линейных моделей (260). Планирование эксперимента для квадра- тических моделей (269). Оптимальные планы (275)	251
§ 3.7.	Математическое моделирование на ЭВМ корреляционно-экстремаль- ной системы Методика последовательного построения моделей исследуемой сис- темы (279). Поэлементная модель ОКЭС и результаты моделирования (282). Регрессионная модель коррелятора (294). Аналитическая модель динамического контура (303)	279
Посл ^е с ^л овие		309
Приложение. Вычисление правых частей укороченных уравнений (2.4.32)		311
Основные обозначения		314
Список литературы		316
Предметный указатель		323
Abstracts.		326

ПРЕДИСЛОВИЕ

В монографии рассматриваются современные методы математического исследования двумерных систем автоматического сопровождения с чувствительными элементами, работающими в оптическом диапазоне электромагнитного излучения. Широкое распространение оптико-электронных следящих систем (ОЭСС) объясняется их высокой разрешающей способностью, а также возможностью создания на базе использования оптического диапазона электромагнитного излучения систем с избыточной информацией, обладающих повышенной помехоустойчивостью. Авторы не стремились к подробному описанию устройств разнообразных следящих систем этого типа и их элементов, ограничившись наиболее общими структурными схемами.

Постепенное усложнение задач, расширение круга применения привели к заметному усовершенствованию ОЭСС. Изменились и методы их математического и экспериментального исследования. В связи с известными трудностями проведения достаточно информативного натурального эксперимента обнаруживается тенденция получения подробной информации о характеристиках ОЭСС в лабораторных условиях, а также путем аналитического исследования и математического моделирования. Таким образом, непременным условием исследования ОЭСС является разработка математической модели. В книге представлены как аналитические методы расчета ОЭСС, основанные на сравнительно простом математическом описании устройств этого типа, так и методы, свободные от каких-либо допущений и предполагающие использование современных быстродействующих ЭВМ с достаточным объемом памяти. В книге выделены три типа ОЭСС по признаку способа обработки входной информации: со сканирующей техникой, с многоплощадочным приемником и с телевизионным приемником изображения.

Оптико-электронные следящие системы с телевизионным приемником изображения позволили осуществить весьма прогрессивный способ обработки входной информации с корреляционным алгоритмом, обладающим повышенной помехозащищенностью. Теория анализаторов этого типа содержится в главе 1, написанной Д.В. Васильевым.

В главе 2, написанной Ю.М. Астаповым, исследован ряд наиболее часто встречающихся задач динамики ОЭСС, которые удастся довести до конечных формальных зависимостей при некоторых допущениях о математичес-

ких моделях. В главе 3 Ю.И. Заложневым описаны современные подходы к исследованию ОЭСС различных типов методом математического моделирования. Такое моделирование принципиально может быть осуществлено на любом уровне подробностей в зависимости от располагаемого объема машинного времени и быстродействия ЭВМ.

Книга предполагает более подробное (по сравнению с обычным уровнем математической подготовки в технических вузах) знакомство читателя с теорией линейного векторного пространства, теорией марковских случайных процессов и методами исследования решений дифференциальных уравнений.

Авторы выражают признательность профессору И.Н. Белоглазову и Ю.Г. Гуревичу, которые прочли книгу в рукописи и сделали ряд ценных замечаний, способствовавших ее улучшению.

ВВЕДЕНИЕ

Оптико-электронные следящие системы с каждым годом находят все более широкое применение в различных областях техники, военном деле, астрономии, медицине, исследовании космического пространства. К ОЭСС принято относить телевизионные, лазерные, тепловые следящие пеленгационные системы активного, полуактивного и пассивного типов [8, 45, 54, 58, 65, 67, 75, 80, 121, 127, 131, 144]. Для них характерно использование электромагнитных волн оптического диапазона в качестве носителей энергии как излучаемых (для первых двух типов), так и принимаемых сигналов.

Основная специфика ОЭСС сосредоточена в приемной части системы. В ее структурной схеме (рис. 1.1) всегда можно выделить несколько характерных укрупненных звеньев, последовательное соединение которых образует замкнутый контур слежения. К ним относится приемник лучистой энергии, поступающей от пеленгуемых источников, который в совокупности с устройством анализа принятых изображений составляет дискриминатор. Его выходные сигналы, усиленные с необходимыми преобразованиями в электронно-усилительном тракте, пропускаются через какой-либо корректирующий фильтр и поступают на регулируемый орган

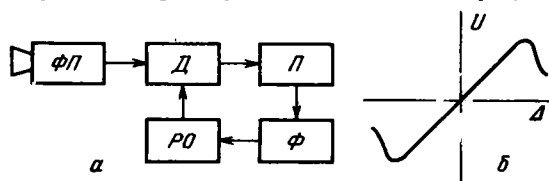


Рис. 1.1. Обобщенная функциональная схема оптико-электронной следящей системы (а) и типичная характеристика дискриминатора (б): ФП – фотоприемник; Д – дискриминатор; П – преобразователь; Ф – корректирующий фильтр; РО – регулируемый орган

который, являясь исполнительным устройством, управляет угловым положением приемника так, чтобы уменьшить рассогласование между линией визирования сопровождаемой точки в пространстве и оптической осью ОЭСС.

Объединение приемника сигналов и устройства их обработки в единый блок отражает реальные особенности, присущие многим современным

оптико-электронным головкам самонаведения, где одна часть предусмотренных алгоритмом дискриминатора операций над входными изображениями выполняется до их попадания на фоточувствительные элементы, а другая — после превращения сигналов в электрические.

В технике ОЭСС используются сотни комбинаций различных способов анализа и типов приемников изображений. Такое многообразие обусловлено обилием возможных вариантов конструктивного воплощения принципов обзора пространства и выделения информации о смещении изображений. К этим принципам относится, например, последовательное сканирование требуемой зоны обзора узким полем зрения одиночного фотоприемного элемента. Та же идея реализуется при развертке считывающего электронного пучка по фоточувствительной мишени электронно-лучевого прибора для регистрации изображений.

Альтернативным принципом является параллельное преобразование сигналов от разных участков наблюдаемой картины. Такой прием может давать ощутимый выигрыш в чувствительности ОЭСС и повышать быстродействие измерительного звена, но требует значительных аппаратных затрат и внедряется по мере совершенствования технологии матричных фотоприемников и многоканальных усилительных микросхем. В некоторых случаях параллельные фотоприемники используют в сочетании с последовательным сканированием. Так конструируют, например, компактные тепловизионные головки, где последовательный обзор за счет простого качания или вращения зеркала в одной плоскости сочетается с параллельным разложением сигналов с помощью однорядной линейки или матрицы твердотельных фотоприемников в другой плоскости [97, 114].

Третий принцип, часто используемый в дискриминаторах ОЭСС, заключается в прерывистой модуляции входного светового потока (или эквивалентного ему электронного изображения в электронно-лучевых фотоприемниках) по определенному закону, который улучшает последующее выделение сигнала из помех, а при работе по точечным источникам обеспечивает извлечение информации об отклонениях линии визирования цели от оси пеленгатора. Прерывистая модуляция является также необходимым условием нормальной работы высокочувствительных пироприемников [121, 168, 177, 179], все более широко применяемых в таких системах.

При использовании лазеров для подсвета объектов слежения в активных и полуактивных ОЭСС импульсный режим излучения часто оказывается выгоднее непрерывного. В силу этого выходные сигналы дискриминатора являются дискретизованными во времени с частотой порядка десятков герц. Аналогичная особенность возникает и в пассивных пеленгаторах в случае применения прерывистой модуляции на входе приемника.

Практически любой способ последовательного сканирования изображений, примененный в ОЭСС, влечет за собой дискретизацию выходных сигналов дискриминатора. Особенно характерен этот процесс в ТВ системах, где интервал дискретизации определяется установленной частотой следования кадров (например, 25 Гц).

Реализованные в конкретной ОЭСС принципы слежения и тип приемника излучений определяют выбор исполнительного звена, в роли которого часто используют либо прибор гироскопического типа, либо электродвигатель специального исполнения. Его быстродействие обратно пропорциональ-

но массам приводимых в движение частей приемника. В случае массивного подвижного оптического узла полоса пропускания замкнутого контура ОЭСС может оказаться существенно меньшей, чем частота дискретизации управляющих сигналов. Поэтому в ряде задач исследования дискретностью управления пренебрегают, рассматривая систему как непрерывную. Вместе с тем при оценке максимальной добротности такой ОЭСС учет дискретизации необходим. Эта необходимость становится все ощутимее и по мере повышения быстродействия пеленгаторов за счет миниатюризации узлов и совершенствования конструктивных решений.

За последние 10—15 лет опубликовано много работ, посвященных устройствам и способам приема оптических сигналов различной структуры и сопутствующим вопросам теории. В отечественной литературе достаточно подробно освещены методы пеленгации точечных источников оптического излучения (см., например, [54, 74, 121, 146, 147]), менее подробно — объектов меняющейся конфигурации. Интенсивно разрабатывается аппарат расчета наиболее современных — корреляционных ОЭСС, способных осуществлять и автоматически сохранять заданную ориентацию относительно сложного многоэлементного изображения нередко произвольной структуры.

Значительный вклад в общую теорию корреляционных автоматов внесен трудами А.А. Красовского, И.Н. Белоглазова, В.К. Баклицкого, В.П. Тарасенко и других советских ученых [1, 6, 11, 55, 64]. Ее применение выявило потребность в более детальной разработке вопросов практического построения быстродействующих корреляционных измерителей рассогласования между изображениями в ОЭСС. Эти блоки, называемые корреляционными дискриминаторами (КД), являются наиболее специфическими элементами таких систем. Реализация КД в виде компактного быстродействующего прибора иногда осложняется тем, что алгоритм корреляционного измерения сдвигов строится непосредственно на базе излишне громоздких для режима слежения процедур поиска экстремума взаимно корреляционной функции (ВКФ) сравниваемых изображений. Такие трудности возникают реже, если в ОЭСС производится преобразование оптических сигналов в электрические. Тогда для их обработки удастся использовать арсенал методов, накопленных в радиолокации и связи на случай измерения сдвигов сложных сигналов при помощи дифференциальных алгоритмов корреляционной природы.

Внедрение корреляционных методов в технику ОЭСС не только стимулировало поиск новых эффективных принципов построения измерительного звена, но и повысило практический интерес к вопросам динамики многомерных следящих систем, функционирующих при наличии случайно меняющихся и невзаимных перекрестных связей между каналами дискриминатора, выход которого может быть дискретизован во времени. Это сочетание условий, осложняющих проектирование современных ОЭСС, требует определенной реконструкции известных методов теоретического исследования двумерных систем по крайней мере при решении вопросов устойчивости.

Современная теория оптимального управления, использующая методы фильтрации в пространстве состояний, позволяет строить наилучшим образом звено Ф, которое в общем случае должно иметь переменные пара-

метры. Однако оптимальная динамическая фильтрация является дорогостоящим делом и пока применяется лишь в уникальных ОЭСС при умеренных требованиях к быстродействию. Необходимый для описания таких оптимизированных систем математический аппарат хорошо представлен в [81], а проблемы оптимальной рекуррентной фильтрации в корреляционно-экстремальных системах детально рассмотрены в фундаментальной работе [64].

Усложнение структур измерительных звеньев и фильтров в контурах управления ОЭСС, использование алгоритмов большой информационной емкости, необходимость учета различных помех, многообразие объектов слежения и других факторов, влияющих на выполнение конечной задачи, повлекли за собой разработку специализированных методов математического и смешанного моделирования таких систем и условий их эксплуатации.

Названные тенденции в развитии теории, методов расчета и моделирования ОЭСС, не нашедшие еще достаточного отражения в литературе, определили основное содержание настоящей монографии.

Материал книги распределен по трем главам. Первая посвящена дискриминаторам сдвига изображений, использующим главным образом бесприсоскобый корреляционный принцип измерения. Вторая глава освещает вопросы теории и расчета непрерывных и дискретных двумерных и многомерных динамических систем, каковыми являются современные ОЭСС. В третьей главе описаны основные принципы и специфические особенности моделирования ОЭСС, функционирующих по сложным алгоритмам и предназначенных для работы по многоэлементным изображениям.

В качестве общей алгоритмической основы, объединяющей многочисленные, нередко найденные эвристически, способы построения КД, в главе 1 рассмотрен класс известных линейных процедур измерения вектора с замороженными параметрами. Эти процедуры поддаются оптимизации в смысле обобщенного метода наименьших квадратов (ОМНК). Разные представления обобщенного алгоритма измерения порождают ряд вариантов структуры дискриминатора, не требуя обращения к исходной статистике и уточнения содержания критерия оптимальности.

В связи с привлечением аппарата двумерных корреляций и использованием КД в ОЭСС с ограниченным полем зрения в изложение включены сведения о полезных признаках симметрии корреляционных и спектральных функций, о действии операторов фильтрации и о краевых эффектах при корреляционной обработке реальных изображений. Затронуты также отдельные вопросы теории операторов обобщенного сдвига и теории непрерывных групп, без которых было бы трудно рассматривать корреляционное измерение проективных (масштабных, ракурсных и т.п.) сдвигов изображений, опирающееся на определение обобщенных авто- и взаимно корреляционных функций.

Как следует из алгоритма многомерного корреляционного измерения сдвигов, основными операциями, реализуемыми в КД, являются: ортогонализация сигналов в сочетании с их фильтрацией и нормирование выходных оценок, обеспечивающее инвариантность КД к уровню и спектральному составу сигналов, а также при необходимости компенсацию перекрестных связей между каналами оценки компонент.

Быстродействие современных цифровых вычислителей позволяет выполнять в реальном масштабе времени многие необходимые для ОЭСС корреляционные алгоритмы измерения сдвигов сложных изображений, даже если число элементов в кадре достигает сотен тысяч. Однако для сокращения затрат при построении КД часто используют аналоговую технику, сочетая ее с цифровой.

В главе 1 рассмотрена общая структура одномерного дискриминатора с ортогональными характеристиками каналов, показана принадлежность известных схем с дифференциатором или фильтром Гильберта в одном из каналов к этому типу, равно как и схем КД с устройствами задержки сигналов. В качестве примера рационального воплощения КД представлена предложенная В.И. Карандасовым [53] схема с ортогональными *RC*-фильтрами первого порядка.

Анализом общей структуры КД удастся выявить ограничения, налагаемые линейностью его трактов на форму выходной характеристики.

Использование операций взаимной ортогонализации сигналов приводит к компактным техническим решениям и при построении стабилизаторов крутизны дискриминационных характеристик в КД [25].

Для современных корреляционных автоматов характерно достаточно широкое применение нелинейных преобразований типа квантования сигналов на два или большее число уровней. Эти процедуры обязательны при использовании приемов так называемого *морфологического анализа* изображений [101–104], который выполняется обычно средствами цифровой техники. Такие преобразования допустимы и при измерении сдвигов в любых КД, если только при этом не нарушаются порядки четности взаимно корреляционных функций сравниваемых изображений.

Особенностью многомерных КД для ОЭСС являются перекрестные связи между результатами измерения отдельных компонент сдвига, обусловленные анизотропной природой реальных изображений. Поэтому, проектируя корреляционную ОЭСС, следует уделять внимание вопросам влияния таких связей на систему в целом. При необходимости высокоточного слежения за изображениями со значительным возможным разбросом показателей анизотропности могут потребоваться специальные меры по ослаблению ее вредного влияния. Они опираются либо на соответствующую фильтрацию входных сигналов в ОЭСС, либо на измерение показателей анизотропности в явной или неявной форме и последующую компенсацию искажений в оценках компонент сдвига между конкретными двумерными реализациями в соответствии с результатами измерений.

Особое место в теории корреляционных дискриминаторов для ОЭСС занимает анализ помех. Их действие специфично уже потому, что КД представляет устройство с двумя входами. Значительное влияние на дискриминационную характеристику измерительного звена в ОЭСС могут оказывать как аддитивные, так и мультипликативные мешающие факторы. К ним относятся детерминированные неравномерности в чувствительности приемника изображений по полю зрения системы, а при некоторых технических решениях и параметры запоминающих устройств для хранения эталонов, используемых в процессах корреляционного сравнения. Разделение эффектов от действия внешних и аппаратурных помех на синфазные и ортогональные составляющие по отношению к выходной характеристике КД

позволяет выделить среди мешающих факторов более опасные и подсказывает пути уменьшения их влияния на точность измерения сдвигов.

Возможность инженерного расчета таких простых показателей, как выходное отношение сигнал/шум при аддитивных флуктуациях на входе и пренебрежении остальными мешающими факторами, проиллюстрирована для КД с линейной фильтрацией в каналах и, в частности, для схемы с ортогональными RC-фильтрами первого порядка. Использование расчетных формул при оценке относительного выходного шума в условиях, типичных для корреляционного ТВ измерителя, подтверждает высокую точность такого метода измерения сдвигов между изображениями.

Оптико-электронные следящие системы, предназначенные для сопровождения подвижных объектов, представляют собой двумерные динамические системы. Некоторые из них снабжаются дополнительно третьим каналом управления, уменьшающим влияние поворота изображения вокруг оптической оси анализатора. Теоретические проблемы, связанные с разработкой таких систем, привлекают внимание многих исследователей. Достаточно указать на ранние статью [63] и монографию [51], где развит анализ двумерных систем на основе использования комплексных координат.

В главе 2 настоящей книги представлены теоретические методы анализа двумерных систем, основанные на классических приемах фазового пространства и метода малого параметра [3]. Для ОЭСС с непрерывным управлением, обладающих повышенной добротностью, характерен режим слежения, сопровождающийся автоколебаниями относительно некоторого среднего положения оптической оси. В главе 2 выделена часть проблемы, касающаяся возможности возникновения и устойчивости простых периодических движений в каналах сопровождения с выраженными резонансными свойствами.

Для ОЭСС с импульсным управлением наиболее существенной проблемой является обеспечение устойчивой работы привода при минимальной динамической ошибке слежения. В главе 2 рассмотрены линеаризованные модели ОЭСС с импульсным управлением амплитудно-импульсной (АИМ) и широтно-импульсной (ШИМ) модуляциями. Особенностью дискретных систем является зависимость их устойчивой работы от периода следования импульсов. Выведенные здесь уравнения в конечных разностях позволили решить задачу анализа устойчивости двумерных систем этого типа, а также построить границу области устойчивости в пространстве параметров ОЭСС.

Одна из важных задач, возникающих в процессе разработки ОЭСС, связана с изменениями в функционировании из-за сложной структуры поля излучения сопровождаемого источника. В главе 2 сделана попытка анализа процесса слежения за источником со сложной структурой в двух случаях: вначале рассмотрена задача сопровождения двойного источника излучения, который служит простейшей моделью распределенного поля. Актуальность исследования такой задачи диктуется необходимостью определения характеристик ОЭСС при слежении в пространстве помех в виде дополнительных источников излучения. Такое исследование предполагает вероятностный анализ процесса сопровождения. В данном случае оказалось возможным свести задачу сопровождения двойного источника к анализу срыва сопровождения при некоторых допущениях. Справедливость их

подтверждается впоследствии при математическом моделировании. Следует отметить, что построенное при этом поле сил коррекции является достаточно универсальной характеристикой ОЭСС. Так, в главе 2 показано использование этой характеристики для анализа перекрестных связей, возникающих при анизотропном поле излучения в ОЭСС с корреляционным принципом обработки входной информации.

Анализ срыва сопровождения с применением теории марковских случайных процессов проведен и для случая ОЭСС с импульсным управлением. В главе 3 рассмотрен пример такого анализа с построением нестационарной плотности распределения ошибки слежения.

Более сложная модель источника с распределенным излучением предполагает разработку обобщенной модели ОЭСС на основе предварительного цифрового моделирования анализатора изображения. Полученная при этом характеристика включена в качестве элемента в модель ОЭСС, допускающую получение конечных формальных зависимостей для оценки точности слежения. Здесь эффективно использован метод малого параметра.

Анализаторы изображения ОЭСС математически описываются дискриминационной характеристикой (ДХ), которая представляет собой зависимость сигнала на выходе анализатора от вектора рассогласования. Компонентами этого вектора будем считать углы рассогласования по двум каналам управления ОЭСС. Известно, что вид ДХ зависит как от параметров элементов анализатора (фокусного расстояния оптической системы, ее апертуры, размеров фотоприемника), так и от параметров сопровождаемого источника излучения (мощности отраженного или излучаемого сигнала, угловых размеров, распределения излучения). Как уже указывалось, аналитическое представление ДХ является весьма громоздкой задачей даже в случае точечного источника и применения раstra с простейшей геометрией в виде радиально-щелевых элементов [121, 165, 171]. Вычисление ДХ в случае протяженного источника со сложным законом распределения лучистости хотя и возможно, но приводит к совершенно неприемлемым с практической точки зрения результатам. Поэтому основным средством получения ДХ при математическом исследовании ОЭСС служит применение специально разработанных алгоритмов, выполняемых быстродействующими ЭВМ с достаточно большим объемом памяти. В главе 3 приводится методика математического моделирования различных источников излучения (в том числе фоновых поверхностей, занимающих все поле зрения анализатора изображения) и элементов анализаторов, включая оптику с учетом физики прохождения излучения и фотоприемник, преобразующий лучистую энергию в электрический сигнал.

Применение ЭВМ для моделирования анализаторов влечет за собой необходимость моделирования также и остальных блоков, которые описываются совокупностью линейных дифференциальных уравнений, эквивалентных дробно-рациональным передаточным функциям вида

$$W(s) = \frac{b_0 s^m + b_1 s^{m-1} + \dots + b_{m-1} s + b_n}{a_0 s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n},$$

и безынерционными нелинейными соотношениями (не содержащими время t в явном виде) $x(z, \dots, z^{(k)})$ между входной z и выходной x вели-

чинами. Основной проблемой, возникающей при математическом моделировании, является проблема сокращения времени счета без потери точности получаемых результатов и, разумеется, при безусловной устойчивости вычислительной процедуры. В главе 3 описаны различные приемы преодоления этих трудностей.

Большие затраты машинного времени в течение длительного периода развития методов конструирования и расчета ОЭСС служили препятствием к широкому применению цифрового моделирования в этой области. Поэтому вначале большую роль играли методы физического и полунатурного моделирования, сохранившие частично свое значение и в настоящее время. Например, считается, что физическое моделирование целесообразно использовать при разработке оптической части анализатора изображения. В ходе такого моделирования подбираются основные элементы объективов или зеркальных рефлекторов и исследуется влияние расфокусировки, абберационных искажений, исследуется взаимосвязь размеров изображения относительно характерных размеров ячеек раstra, определяются характеристики ОЭСС при различных ракурсах сопровождаемого объекта. Для физического моделирования обычно используется макетный образец конструируемой ОЭСС и какой-либо имитатор источника излучения.

Создание имитатора источника представляет собой самостоятельную проблему, которая заключается в воспроизведении поля лучистости реального источника. Сведения об исходном поле обычно получают экспериментальным путем. Для непосредственных натуральных измерений применяется радиометр с достаточно узким полем зрения и необходимой спектральной чувствительностью. Существует, однако, ситуация, когда непосредственный эксперимент становится затруднительным. Например, при необходимости моделирования излучающего летательного аппарата, движущегося в сверхзвуковом воздушном потоке. В таких случаях данные о поле излучения дополняются расчетным путем, и эта часть проблемы является общей для математического и физического моделирования.

При физических методах воспроизведения излучения обычно используют имитаторы, включающие в виде конструктивных элементов лампы накаливания, нагретые спирали, металлические или керамические пластины со специально подбираемой конфигурацией. Все эти устройства с неизбежными погрешностями воспроизводят реальное поле лучистости. Погрешности воспроизведения становятся особенно заметными при необходимости моделирования процесса слежения в условиях активных и пассивных помех, создаваемых многими распределенными в пространстве источниками при нестационарных характеристиках излучателей [150–152, 154].

Математическое моделирование, методы которого описаны в главе 3, свободно от каких-либо ограничений и поэтому позволяет осуществить исследование ОЭСС с большой экономией материальных и временных затрат.

При физическом моделировании телевизионных систем используются обычно черно-белые снимки либо макеты объекта слежения. Так как объекты слежения отличаются большим разнообразием, то возникает вопрос: насколько представительным является конкретный макет или снимок по отношению к более общей совокупности возможных реальных полей? Эта проблема представляется в данное время не решенной до конца,

и частичное ее решение заключается в статистическом цифровом моделировании. В этом случае объектом слежения служит генерируемое на ЭВМ двумерное поле с заданными статистическими характеристиками. В качестве таких характеристик в простейшем случае выступают закон распределения лучистости

$$F(e) = P\{E < e\}$$

и корреляционная функция

$$K(\xi, \eta) = \lim_{\substack{P \rightarrow \infty \\ Q \rightarrow \infty}} \frac{1}{4PQ} \int_{-P}^P \int_{-Q}^Q E(x, y) E(x + \xi, y + \eta) dx dy,$$

где x и y — пространственные координаты точки поля лучистости.

Методы моделирования двумерного поля с заданными статистическими свойствами описаны в § 3.5. Для окончательного заключения о качестве сопровождения обычно проводится дополнительный цикл моделирования по детерминированному полю, представляющему конкретный объект. Такое моделирование не вызывает трудностей методического характера и предполагает лишь аппаратные возможности автоматизированного преобразования физических полей в цифровой массив, который затем вводится в ЭВМ в качестве входных данных.

Подробная (позлементная) модель ОЭСС вместе с моделью сопровождаемого объекта и фоном является весьма совершенным инструментом для анализа характеристик системы на стадии проектирования. Она эффективно используется при улучшении характеристик системы в целом и ее отдельных блоков на стадии изготовления макетного образца и выработки заключения о реальных характеристиках системы на стадии испытаний.

Здесь следует остановиться на вопросе о достоверности математической модели. Обычно исходным материалом для первоначальной разработки модели служат сведения о номинальных параметрах и характеристиках системы, ее отдельных блоков и элементов, паспортные данные, а также чертежи и схемы конструкторской документации. Например, при разработке математической модели анализатора изображения растрового типа со сканированием используются чертеж растра, данные о фотоприемнике и характеристики оптической части. Математическая модель анализатора, построенная по этим данным, может, однако, отличаться по своим свойствам от реальных образцов анализаторов. Причинами этого являются незначительные отклонения, вносимые в геометрию растра по технологическим причинам, неучтенные изменения номиналов при массовом изготовлении в заводских условиях, неполнота документации, относящейся к фотоприемнику, изменения радиуса сканирования при настройке и т.п. Аналогичные изменения могут встречаться и при моделировании других блоков ОЭСС. В итоге появляются отличия в результатах моделирования и в результатах натурных испытаний. Для приведения математической модели в соответствие с реальной системой следует использовать специальные тесты — испытания в контролируемых условиях с одновременным проведением моделирования и подстройкой параметров модели. Одним из простейших тестов служит снятие ДХ при неподвижном источнике с контролируемыми угловыми размерами, силой излучения и спектром.

Если в данном эксперименте под ДХ понимать зависимость управляющего сигнала u от угловой ошибки ϵ при совокупности условий проведения эксперимента p , то ввиду непрерывности функции u при малом изменении вектора p управляющий сигнал меняется также незначительно. Эта идея непрерывной зависимости характеристик системы от совокупности условий проведения эксперимента позволяет ограничиться небольшим числом экспериментов, после проведения которых можно гарантировать в пределах рационально выбранного допуска совпадение математической модели с ее натурным прототипом.

Математическая модель, достоверность которой установлена с помощью контрольных тестов, допускает всестороннее исследование ОЭСС. При этом становятся доступными эксперименты, физически не осуществимые при натурных испытаниях. Приведем пример такого цифрового эксперимента.

Особый интерес вызывает ситуация, когда расстояние между приемной частью ОЭСС и сопровождаемым источником быстро уменьшается. Начиная с некоторого расстояния, качество сопровождения может резко ухудшиться. С помощью поэлементной модели можно выяснить физические причины этого явления. Достаточно, например, получить картину деформации ДХ в зависимости от параметров процесса сближения. В то же время натурный эксперимент такого рода в строго контролируемых условиях его проведения в целом ряде случаев не может быть поставлен. Приведенный пример носит общий характер для различных типов ОЭСС. В § 2.7 излагается теоретический подход к определению точности слежения, а материалы главы 3 служат для моделирования указанного процесса.

Теоретические методы, представленные в монографии, основаны на математическом анализе уравнений, описывающих ОЭСС. Такие уравнения не могут быть слишком сложными. Однако при их составлении исследователи стремятся придать им такую форму, чтобы они вобрали в себя основные характеристики реального объекта исследования. Таким образом, эти уравнения также представляют собой математическую модель ОЭСС, но эта модель носит обобщенный характер и при ее создании неизбежно возникает вопрос о необходимости учета или о возможности пренебрежения теми или иными факторами. Для объективного (научно обоснованного) решения этого вопроса служат методы идентификации моделей, развитые в § 3.7 для случая ОЭСС с корреляционно-экстремальным принципом обработки входной информации.

Процесс идентификации моделей представляет собой самостоятельную научную задачу, выполнение которой всегда ориентировано на некоторый уровень доверия, так как она носит вероятностный характер. При решении компромиссной задачи между требуемым уровнем доверия и минимальным числом экспериментов большую помощь могут оказать краткие сведения § 3.6 по современной теории эксперимента.

ПРИНЦИПЫ ОБРАБОТКИ ВХОДНОЙ ИНФОРМАЦИИ В ОПТИКО-ЭЛЕКТРОННЫХ СЛЕДЯЩИХ СИСТЕМАХ

§ 1.1. Обобщенные функциональные и структурные схемы оптико-электронных следящих систем.

Примеры реализации

Оптико-электронные следящие системы по своему назначению во многом аналогичны системам автоматического слежения за направлением прихода радиосигнала и в последние годы стали заменять их в составе локаторов, навигационных комплексов, автоматов управления движением беспилотных летательных аппаратов и т.п. Круг применений ОЭСС непрерывно растет. Расширяется диапазон электромагнитных волн (ЭМВ), регистрируемых приемниками таких систем, — от рентгеновского излучения до миллиметровых радиоволн. Современные ОЭСС наиболее приспособлены для работы в режиме пассивной или полуактивной локации, т.е. не требуют излучения собственного зондирующего сигнала, а воспринимают излучение от источника информации либо отраженную от него энергию удаленного облучателя.

Большинство оптико-электронных автоматов вне зависимости от рабочего диапазона ЭМВ отличают два общих признака: формирование входных сигналов в виде двумерных скалярных изображений трехмерного пространства по законам геометрической оптики и обработка этих сигналов преимущественно средствами электроники. В основе процесса регистрации изображения в таких системах лежит некогерентное преобразование принятого излучения в видеосигналы. Структура систем слабо связана с частотой или фазой приходящих ЭМВ, но существенно зависит от их интенсивности. Это обусловлено относительной простотой и совершенством фотоприемных элементов, реагирующих на энергию лучистого потока (видиконы, полупроводниковые фотоматрицы и т.д.). Приемник всякой ОЭСС в той или иной форме осуществляет разделение сигналов по направлению прихода, т.е. угловую селекцию. Она обеспечивается выбором параметров оптического узла (разрешающей способности объектива), размеров минимальной фоточувствительной площадки, с которой можно получить видеосигнал, не зависящий от соседних с ней участков (разрешающая способность фотоприемника), и схемы обработки, где из совокупности видеосигналов извлекается информация об угловых координатах наблюдаемого объекта.

В последние годы в связи с развитием технических средств когерентной оптики предпринимаются попытки создания когерентных оптико-локационных систем. В принципе, они позволяют реализовать более эффективную пространственно-временную селекцию сигналов, чем в радиолокации, однако ограничения, связанные с трудностями генерации строго монохро-

матического света и необходимостью охлаждения приемников, сложность и дороговизна таких систем сдерживают их внедрение.

Состав оптико-электронной следящей системы. Некогерентные оптико-электронные автоматы следящего типа являются измерителями относительного сдвига изображений. Их строят в соответствии с обобщенной функциональной схемой (см. рис. 1.1, а), характерной для многих радиотехнических следящих систем [70, 90, 91]. В ее составе обычно удается выделить фотоприемник, дискриминатор, преобразователь, фильтр управляющих сигналов и регулируемый орган, роль которого могут выполнять такие элементы, как подвижные платформы (гироплатформы), обрабатывающие нужный угол поворота, устройства управления положением развертывающего раstra, управляемые генераторы опорных сигналов и т.д. — в зависимости от выбранного технического решения. На основной вход дискриминатора поступает сигнал $y(t, \lambda)$, содержащий полезную часть, зависящую от пространственного параметра $\lambda(t)$, и помехи.

Названный параметр — скалярный или векторный — и является объектом слежения. В ОЭСС он чаще всего характеризует некоторое направление в пространстве. В плоскости изображений этот параметр преобразуется в сдвиг $\Delta(t)$, определяемый также и пространственной ориентацией приемника.

Дискриминатор ОЭСС в первом приближении можно считать безынерционным устройством по отношению к изменениям $\lambda(t)$ или $\Delta(t)$.

Типичная дискриминационная характеристика (ДХ) имеет вид нечетной функции скалярной ошибки Δ (см. рис. 1.1, б). Если Δ — вектор, то дискриминатор содержит ряд каналов с подобными ДХ на каждую измеряемую компоненту, а следящая система является многомерной.

Контрастные дискриминаторы в ОЭСС. Примером одной из реализаций структуры, приведенной на рис. 1.1, а, служит схема простейшего "контрастного" координатора, который обеспечивает автоматическую пеленгацию некоторой яркой точки на равномерном фоне (рис. 1.2, а). Четыре идентичных фоточувствительных элемента (ФЭ) размещены в плоскости изображения (ПИ) за объективом (О) так, что границы раздела между ФЭ образуют декартову систему координат (рис. 1.2, б). Световое пятно от пеленгуемого источника попадает либо во все квадранты фотоприемника, либо только в часть из них. Суммируя и вычитая видеосигналы от четырех ФЭ в соответствии с балансной схемой дискриминатора, расшифрованной на рис. 1.2, в, можно сформировать две дискриминационные характеристики $U_1(\Delta_1)$ и $U_2(\Delta_2)$, где Δ_1 и Δ_2 — взаимно ортогональные компоненты плоского сдвига центра пятна от начала координат.

Выходные сигналы дискриминатора через блок следящего привода (СП) управляют положением оптической оси ОО, благодаря чему в образовавшейся следящей системе при наличии соответствующих датчиков угла могут быть измерены перемещения источника (И) в пространстве.

В схеме предусмотрена нормировка крутизны дискриминационных характеристик по суммарному сигналу. Таким образом, выходы двумерного дискриминатора формируются согласно соотношениям

$$U_1(\Delta_1) = \frac{(a+d) - (b+c)}{a+b+c+d}, \quad U_2(\Delta_2) = \frac{(a+b) - (c+d)}{a+b+c+d},$$

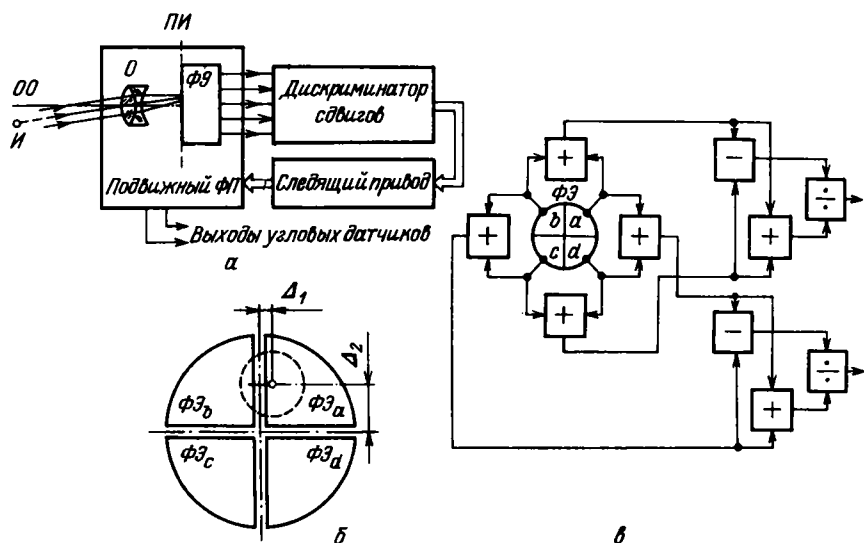


Рис. 1.2. Устройство контрастного оптико-электронного координатора: *а* – функциональная схема; *б* – четырехквадрантный фотоприемник; *в* – структура балансной схемы дискриминатора; И – источник; ОО – оптическая ось; О – объектив; ПИ – плоскость изображения; ФЭ – фоточувствительные элементы; ФП – фотоприемник

где a , b , c , d – сигналы, снимаемые с ФЭ соответствующих квадрантов и зависящие от сдвигов Δ_1 и Δ_2 .

Форма ДХ определяется распределением лучистой энергии по пятну, которое меняется в зависимости от свойств оптики, конфигурации пеленгуемого источника света и его рассеяния на пути к приемнику, а также от расстояния между ФЭ соседних квадрантов. Если от дискриминатора требуется высокая абсолютная точность измерения сдвигов Δ_1 и Δ_2 , то фоточувствительные элементы изготавливают из одной пластины полупроводника (например, кремния) с поверхностным электронно-дырочным переходом. Границы между квадрантами формируют в виде узких полос путем окисления поверхностного слоя кристалла на достаточную глубину. Использование последних достижений в технологии формирования МДП (МОП)-структур¹⁾ позволяет доводить зазоры между фоточувствительными площадками всего до нескольких микрометров [105, 122].

Фотоприемники описанного типа благодаря хорошей идентичности параметров ФЭ в квадрантах обеспечивают получение ДХ с незначительным отклонением выходного сигнала от нуля при строгом совмещении пятна с заданным центром. В целях повышения стабильности ДХ в окрестности "нуля" и снижения вредного влияния дрейфа усилителей в каналах схемы дискриминатора усиление стараются осуществлять на переменном токе, а дискриминатор в целом периодически калибруют с помощью специального источника рассеянного света, направляемого на мишень фотоприемника.

¹⁾ Общепринятое сокращение трех слов: "металл – диэлектрик (окисел) – полупроводник".

Калибровку часто автоматизируют, используя выходные сигналы дискриминатора в калибровочных интервалах для компенсации погрешностей в последующих рабочих интервалах.

В литературе [8, 28, 54, 74, 121, 146, 147] описано значительное число запатентованных дискриминаторов (координаторов), реагирующих на смещение контрастного светового пятна. Представленный выше вариант относится к классу систем без модуляции лучистого потока с неподвижным четырехплощадочным фотоприемником. Существует также обширный класс дискриминаторов с одноплощадочным фотоприемником и оптико-механическим сканированием специальной *растровой* решетки из чередующихся прозрачных и непрозрачных элементов перед плоскостью изображения. Периодическое движение раstra обеспечивает модуляцию светового потока, что позволяет отказаться от усилителей постоянного тока при формировании выходных сигналов дискриминатора.

На рис. 1.3, *а* [121] показан один из вариантов модулирующей растровой решетки, вращающейся в плоскости изображения. Размещенный позади нее фотоприемник собирает всю лучистую энергию, проходящую через растр. В зависимости от положения пеленгуемой точки на ПИ фототок приемника меняется с частотой, которая указывает на величину отклонения точки от центра мишени с точностью до некоторой ошибки дискретизации. Форма раstra такова, что получаемые сигналы имеют широтно-импульсную модуляцию, причем фаза модулирующей гармонической функции несет в себе информацию о направлении радиуса-вектора отклонения точки от центра. На рис. 1.3, *б* показаны графики модулированных видеосигналов (1, 2) и модулирующих функций (3, 4) при двух положениях радиуса-вектора пеленгуемой точки на растре (*А* и *Б*).

Множество других типов растров с вращением вокруг оптической оси координатора, так же как и описанный, обеспечивает измерение сдвига в полярных координатах. Для организации измерений в декартовых координатах удобнее пользоваться растровыми решетками, которые колеблются по взаимно ортогональным направлениям в плоскости изображения. Нужные движения светового луча часто достигаются также за счет колебания или вращения плоских зеркал и многогранных призм [121].

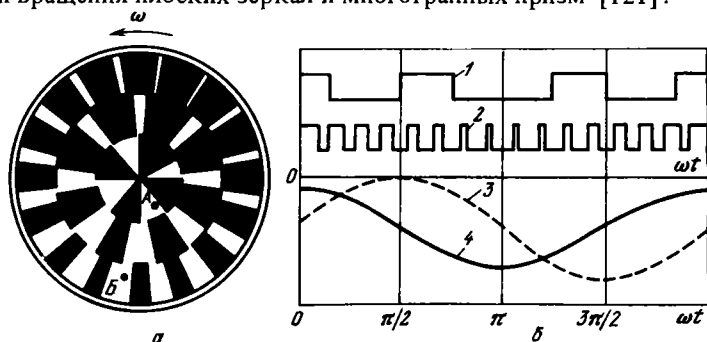


Рис. 1.3. Вариант модулирующей растровой решетки (*а*) и графики сигналов (*б*) в контрастном координаторе: 1, 2 – видеосигналы при положениях объекта соответственно в точках-А и Б (см. рис. 1.3, *а*); 3, 4 – модулирующие гармонические функции ШИМ-сигналов при разных направлениях радиуса-вектора пеленгуемой точки на вращающемся растре

Идея пространственной модуляции принимаемого изображения может быть реализована и без применения вращающихся или колеблющихся механических узлов, если в координаторе имеется вакуумный электронно-оптический преобразователь изображения (ЭОП). Оптическое или тепловое изображение наблюдаемой сцены проектируется на фотокатод 1 ЭОПа (рис. 1.4), что вызывает эмиссию потока электронов. Магнитное поле катушки 2 фокусирует полученное таким путем электронное изображение на поверхности люминофорного слоя 3, нанесенного на противоположной

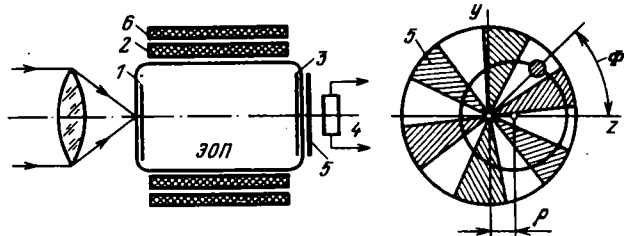


Рис. 1.4. Вариант координатора с электронно-оптическим преобразователем, сканированием переменным магнитным полем и неподвижной растровой решеткой: 1 – фотокатод; 2, 6 – катушки; 3 – люминофорный слой; 4 – фоточувствительный элемент; 5 – растровая решетка

фотокатоду торцевой поверхности преобразователя. С наружной стороны выходного торца, на котором за счет свечения люминофора образуется видимое изображение, размещен фоточувствительный элемент 4 с неподвижной растровой решеткой 5 на его поверхности.

При наличии двух пар отклоняющих катушек 6 в ЭОПе можно создать меняющееся во времени поперечное магнитное поле, с помощью которого производится круговое или более сложное сканирование электронного изображения относительно неподвижного секторного растра. При круговом сканировании, достигаемом в случае питания отклоняющих катушек сдвинутыми по фазе на 90° синусоидальными токами, точечный объект прочерчивает на поверхности ФЭ окружность, а импульсы выходного сигнала оказываются модулированными по длительности с частотой вращения электронного изображения относительно растра, причем фаза модулирующей функции указывает направление отклонения центра вращения данной точки относительно оптической оси системы ОО. Индекс модуляции (максимальное относительное изменение длительности импульсов) пропорционален величине измеряемого отклонения ρ контрастного объекта от оси координатора (Φ – угловое положение объекта). Форма выходных сигналов ФЭ u_p при нулевом (а) и ненулевом (б) отклонениях показана на рис. 1.5. Способы выделения полезной информации те же, что применяются для обработки сигналов с широтно-импульсной модуляцией (см., например, [118]).

Одна из привлекательных особенностей применения двухуровневых (черно-белых) растров для временной модуляции светового потока в контрастных дискриминаторах состоит в том, что видеосигналы, снимаемые с ФЭ, в таких системах можно подвергать двухуровневому квантованию, если только интенсивность лучистого потока на входе превосходит некоторый минимальный пороговый уровень. Благодаря этому обеспечи-

вается независимость формы и, в частности, крутизны ДХ от яркости пеленгуемой точки.

Первым важным недостатком любого контрастного дискриминатора является зависимость его ДХ от неравномерности фона, окружающего пеленгуемый объект. Вторым недостаток — зависимость формы ДХ от размеров объекта. Наличие в поле зрения координатора контрастных фрагментов,

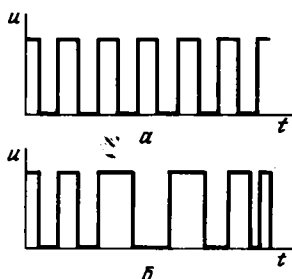


Рис. 1.5. Видеосигналы в координаторе в случае сканирования при нулевом (а) и ненулевом (б) сдвигах объекта относительно оси приемника

конкурирующих с пеленгуемой точкой, вынуждает к применению многоэлементных приемников. Сопоставляя сигналы от разных элементов наблюдаемой сцены, можно отсеять часть мешающих фрагментов по тем или иным дополнительным признакам. Такими признаками могут быть размеры и особенности формы пеленгуемых объектов: связность и другие топологические свойства их контуров, сходство контуров с некоторыми заранее известными моделями (эталоны), наличие мелких повторяющихся текстурных деталей у окружающего фона и отсутствие этих деталей у объекта и т.п. Для решения подобных задач необходимы средства телевизионной техники.

Несмотря на исключительную привлекательность идеи — наделить дискриминатор изображений способностью к распознаванию образов — реальные системы еще далеко уступают по своим возможностям хотя бы зрительному аппарату животных. В большинстве применений труднее всего выполнить требования к быстродействию анализаторов, в которых необходимо перерабатывать до сотен миллионов двоичных единиц информации в секунду. Причиной тому — огромная информационная емкость изображений. Достаточно указать, например, что количество информации в стандартном телевизионном кадре теоретически может превышать 10^7 дв. единиц. В то же время информация, необходимая для того, чтобы определить положение единственной точки в кадре с точностью порядка разрешающей способности телевидения, в сотни раз меньше указанной величины. Отсюда ясно, сколь важны поиски эффективных алгоритмов сжатия телевизионной информации с целью распознавания образов. Этим поискам посвящено много экспериментальных и теоретических исследований [77, 107], но пока лишь малая часть результатов получила внедрение, в основном в уникальных системах (например, в комплексах наземной обработки визуальной информации, поступающей со спутников Земли).

Особое место среди оптико-электронных систем, обеспечивающих ориентацию пеленгатора по отношению к сложным изображениям, заняли так называемые *корреляционные оптико-электронные автоматы*. Их рассмотрению посвящены следующие параграфы данной главы.

§ 1.2. Корреляционные дискриминаторы сдвига изображений

Роль и место корреляционных дискриминаторов в ОЭСС и других следящих автоматах. Общим для всех корреляционных измерительных систем принципом является сравнение взаимно сдвинутых, но сходных между собой сигналов, один из которых часто называют *эталонным*, а другой — *текущим*. В роли формирователей таких сигналов могут использоваться ТВ камеры, фотопроекторные устройства, датчики рельефа — высотомеры, механические профилографы, гидро- и радиолокаторы и многие другие приборы, способные воспроизводить одномерные или двумерные изображения наблюдаемых или записанных в некотором блоке запоминающих сюжетов.

В результате сравнения, производимого путем вычисления взаимно корреляционной функции сигналов, находится оценка их относительного сдвига, которая может быть использована в роли сигнала управления для приведения в действие какого-либо исполнительного элемента, парирующего измеренный сдвиг. Таким образом может быть реализован корреляционный следящий автомат, в котором роль измерительного органа выполняет специфическое устройство — корреляционный дискриминатор.

За последние 15–20 лет было изобретено множество корреляционных систем разного назначения: для измерения относительного положения движущихся объектов, их скорости, ускорения, направления движения и т.п.

Обширный обзор таких систем дан в справочном пособии С.Ф.Козубовского [55]. Авторами многих предложений являются В.П.Тарасенко и его сотрудники [11, 122, 123]. Известны многочисленные зарубежные патенты на корреляционные автоматы [2, 26, 58, 59, 116, 117, 128, 144].

К настоящему времени сформировался фундамент общей теории корреляционных дискриминаторов сдвига сложных сигналов. Эта теория развивалась в первую очередь в интересах радиолокации, где до недавних пор основными являлись задачи раздельного или совместного измерения временного (фазового) и частотного сдвигов одномерных процессов. Известно множество дискриминаторов такого назначения. Они являются обязательными элементами различных радиолокационных следящих измерителей. В разных отраслях промышленности, в военной и космической технике начали также применяться корреляционные автоматы слежения за двумерными сигналами-изображениями.

В настоящей главе изложены элементы указанной теории, развитой и дополненной применительно к дискриминаторам изображений. Потребность в таком дополнении назрела в последние годы [169] в связи с перспективностью применения корреляционных автоматов в оптико-электронных системах. Некоторая методическая особенность изложения состоит в отказе от использования байесовской процедуры на этапе обобщенного синтеза корреляционного дискриминатора. Как известно, например, из [112], результаты, достигаемые при байесовском подходе [6, 7, 64], получаются путем частной интерпретации решения задачи оценивания по обобщенному методу наименьших квадратов. Этот подход позволяет включать в критерии оптимизации различные показатели качества, которые могут быть в какой-то части и нестатистическими.

Алгоритмы оценивания сдвигов, используемые в главе 1 при синтезе дискриминаторов, находятся для замороженных параметров, постоянных на интервале наблюдения. Сглаженные оценки переменных параметров $\hat{\Delta}(t)$ получают методами динамической фильтрации с учетом последовательности наблюдений¹⁾. Этим вопросам посвящена обширная литература, освещающая и проблемы оптимизации цепей сглаживания корреляционно-экстремальных автоматов [1, 6, 11, 64].

Здесь и далее корреляционные дискриминаторы рассматриваются как звенья системы, находящейся в режиме слежения за двумерным сигналом. Примером применения подобной системы может служить комплекс управления автоматической посадки космического аппарата (КА) в заданную точку наблюдаемой поверхности с произвольно расположенными на ней ориентирами. Основная задача — направить КА в центр участка поверхности, изображение которого выбрано перед началом цикла автоматической посадки и записано в качестве эталона в запоминающее устройство (ЗУ) автомата. Эта задача решается без вмешательства человека, если КА снабжен следящей головкой самонаведения (ГСН) со специальным приемником для регистрации текущих изображений и дискриминатором для оценки сдвигов между ними и эталоном.

Физическая природа излучения, создающего изображения, безразлична, однако для определенности далее считается, что они формируются по правилам проективной геометрии и законам геометрической оптики. Это значит, что регистрируемые поля являются скалярными и из рассмотрения исключены особенности, связанные с дифракцией. Целесообразно проанализировать дискриминаторы, в которых в явном виде не используется информация о текущем расстоянии до поверхности, а метод управления системой построен на определении угловых координат, измеряемых по смещениям начальной точки прицеливания в плоскости изображений, регистрируемых приемником. Таким образом, по функциональному назначению ГСН как измерительное звено системы автоматического управления не отличается, например, от радиозвена в классической структуре комплекса радиоуправления реактивными снарядами и космическими аппаратами.

Учитывая требования оптимальности, это звено выполняют в виде следящего угломера — координатора, который поворачивает приемник ГСН так, чтобы устранить сдвиг текущего изображения относительно эталонного, и формирует выходные сигналы в опорной стабилизированной системе координат.

Качественное рассмотрение процесса движения аппарата относительно наблюдаемой поверхности показывает, что в процессе сближения с выбранным участком масштаб его изображения непрерывно укрупняется так, что при прямолинейном движении объекта ориентиры, окружающие некоторую неподвижную точку, "разбегаются" к краям поля зрения. Подобное явление можно наблюдать в кинокадрах с "наездом" на снимаемый объект. Если угол зрения и угловая разрешающая способность приемника сохраняются неизменными, то по мере сближения часть ориентиров выходит за края кадра. В то же время в изображении могут появиться новые элемен-

¹⁾ Здесь и далее символ $\hat{}$ означает оценку.

ты, угловые размеры которых при большом удалении не разрешались системой, а на меньшем расстоянии стали достаточными для различения. По тем же причинам при сближении может возникнуть "размазывание" контуров тех объектов, которые издавна наблюдались как контрастные с четкими границами между светом и тенью. Это явление при использовании простых контрастных дискриминаторов может служить причиной срыва слежения в автоматах, которые следят за выбранным объектом по признаку простого контраста с фоном.

Оптические характеристики текущих изображений при непрямолинейном движении аппарата с ГСН меняются еще более сложным образом из-за возможных вращений приемника относительно центра контролируемого участка, а также из-за изменений ракурсных углов, под которыми наблюдается поверхность. Все сказанное указывает на необходимость компенсации в процессе измерений не только плоского сдвига текущего изображения относительно эталона, но и масштабных, ракурсных и креновых взаимных деформаций сравниваемых двумерных сигналов. Очевидно, в принципе самым эффективным устройством для такой компенсации должен быть многомерный следящий измеритель всех названных возмущений. Его следовало бы устроить так, чтобы он мог автоматически согласовывать эталон и текущие изображения по масштабу, крену и ракурсу, минимизируя тем самым ошибки измерения основных параметров — компонент относительно плоского сдвига, которые обычно используют для управления рулевыми органами движущегося объекта. Структура изображений, регистрируемых приемником оптико-электронной ГСН с корреляционным дискриминатором (спектр пространственных частот, контраст ориентиров относительно фона, топологические особенности), может варьироваться очень широко из-за многообразия природных условий, состояния грунта, его влажности, условий освещения, наличия ориентиров искусственного происхождения, подвижности отдельных элементов относительно общего фона и т.п.

Потребители подобных автоматов часто требуют максимальной инвариантности их параметров к смене характеристик изображений, по которым должно осуществляться самонаведение.

Конечно, не все корреляционные автоматы должны работать в столь сложных условиях. Проще других устроен, например, оптико-электронный корреляционный автомат (рис. 1.6) для измерения скорости движения металла при его прокатке [55]. Автомат измеряет временное запаздывание τ между двумя непрерывными световыми сигналами, созданными за счет отражения некогерентного света от двух небольших освещенных площадок, которые разнесены на заданное расстояние l вдоль направления движения ленты прокатного стана. Имеющиеся на ее поверхности мелкие неоднородности регистрируются двумя фотоприемниками ФП₁ и ФП₂ измерителя как переменные слагающие входных сигналов. Регулируя дополнительную задержку τ_3 в тракте усиления одного из сигналов, можно найти значение $\tau_3 = \tau = l/V$, при котором обе переменные слагающие оказываются наиболее коррелированными, что и фиксируется по максимуму выходного напряжения коррелятора (К).

Для автоматического отыскания задержки в схему вводят какой-либо экстремальный регулятор (ЭР), который может сформировать из четной

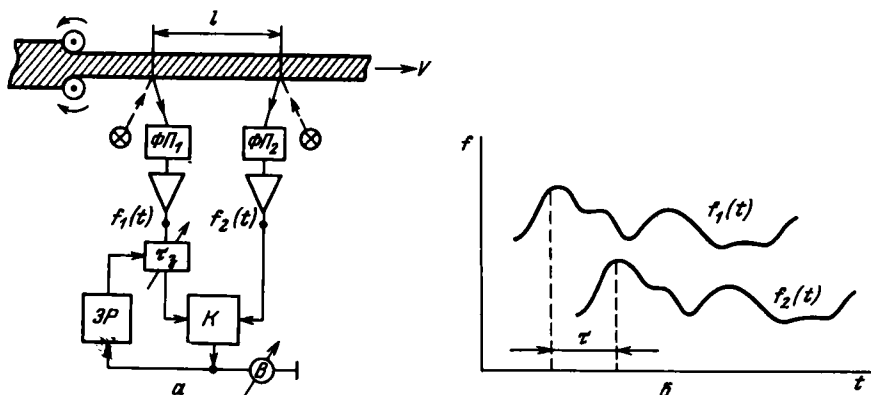


Рис. 1.6. Корреляционный автомат измерения скорости прокатки: а – функциональная схема; б – сигналы $f_1(t)$ и $f_2(t)$ на выходах фотоприемников ФП₁ и ФП₂

зависимости корреляционной функции (КФ) от сдвига нечетную дискриминационную характеристику $U_{з.р}(\tau - \tau_3)$, достаточную для того, чтобы привести в действие орган регулирования задержки τ_3 и образовать таким путем замкнутый следящий измеритель. Такие автоматы можно применять в роли датчиков отклонения от заданной скорости в составе системы авторегулирования скорости прокатки, если задавать нужное ее значение выбором фиксированных τ_3 и l . Сходное устройство имеют бесконтактные корреляционные автоматы контроля расхода или скорости движения жидкостей, состав которых неоднороден по какому-либо параметру, поддающемуся регистрации [55].

Расходомеры и скоростемеры описанного типа содержат простые одномерные корреляционные дискриминаторы, куда входят кроме коррелятора (перемножителя с интегратором) также и цепи электронного преобразования, обеспечивающие необходимую монотонную зависимость регулирующего сигнала от некомпенсированного сдвига двух колебаний одинаковой формы¹⁾. Требования к быстродействию такого дискриминатора обычно невысоки, многообразие возможных структур сигналов ограничено. Подобные автоматы могут снабжаться органами ручной подстройки коэффициентов усиления, не требуют мер по расширению дискриминационных характеристик и не содержат элементов стабилизации крутизны дискриминатора. Система в целом не имеет запоминающих элементов, что также упрощает анализ и расчет ее параметров.

Функциональные признаки и структура КД в многомерном следящем автомате. Сопоставляя разные корреляционные следящие автоматы, можно выделить ряд общих признаков, характеризующих корреляционный дискриминатор как измерительный элемент такой системы в следящем режиме.

1. На два входа КД поступают соответственно два сигнала – изображения f_z и f , различающиеся при отсутствии помех только многомерным

¹⁾ Подробнее о способах формирования нечетных ДХ рассказано в § 1.2 и 1.3.

взаимным сдвигом $\underline{\Delta} = (\Delta_1, \Delta_2, \dots, \Delta_m)$, где $\Delta_1, \dots, \Delta_m$ — составляющие сдвига. В свою очередь сдвиг можно рассматривать как разность некоторых параметров $\underline{\lambda}$ и $\underline{\lambda}_0$ в неподвижной системе координат такую, что

$$\underline{\Delta} = \underline{\lambda} - \underline{\lambda}_0, \quad \underline{\lambda} = (\lambda_1, \dots, \lambda_m), \quad \underline{\lambda}_0 = (\lambda_{10}, \dots, \lambda_{m0}).$$

Таким образом, измерение разности $\underline{\Delta}$ вместо $\underline{\lambda}$ равносильно переносу начала отсчета параметров в точку $\underline{\lambda}_0$ неподвижной координатной системы. Заметим, что в общем случае различия между двумя сигналами могут также возникать из-за действия помех и появления новых элементов в поле зрения приемника.

Размерность m пространства сдвигов, выбираемая при составлении модели дискриминатора, определяется количеством скалярных параметров, минимально необходимым для формального описания различий между коррелируемыми сигналами. Например, для модели КД в гипотетическом автомате посадки движущегося аппарата на некоторую плоскую поверхность при отсутствии вертикальной стабилизации фотоприемника $m = 5$, так как две компоненты задают плоский сдвиг, две другие — масштабы деформаций в двух плоскостях за счет поворота фотоприемника относительно взаимно перпендикулярных горизонтальных осей и изменений высоты над поверхностью и одна — угол "скручивания" изображения за счет поворота фотоприемника вокруг вертикальной оси.

Стабилизировав (например, с помощью независимой трехстепенной гироскопической платформы) положение осей фотоприемника в пространстве, можно сократить размерность возмущений до $m = 3$ (две компоненты сдвига и масштаб).

В дискриминаторе измерителя скорости прокатки $m = 1$, так как из всех составляющих измерению подлежит лишь одна — продольный сдвиг, а остальные в правильно сконструированной системе должны отсутствовать.

2. Количество выходов КД определяется числом измеряемых скалярных сдвигов, а оно в свою очередь равно числу сигналов управления, необходимых, чтобы поддерживать режим слежения за изменяющимися параметрами изображения. При поиске наилучших технических решений для следящих корреляционных автоматов следует помнить о необходимости максимального сокращения числа отслеживаемых компонент сдвига изображений. Большое количество возмущающих факторов в системе, как правило, свидетельствует о плохой ее организации, поэтому надо стремиться к использованию в КД всех априорных данных об ожидаемых рассогласованиях, а также выбирать методы фильтрации, обеспечивающие наименьшую зависимость оценок от некомпенсированных возмущений.

3. При синтезе дискриминатора постулируют, что его дискриминационные характеристики $U_i(\Delta_i)$ являются монотонными в определенной зоне, а еще лучше — нечетными функциями соответствующих компонент сдвига. Для систем с непрерывным управлением целесообразно требовать, чтобы любая i -я ДХ в некоторой окрестности нуля была линейной функцией i -й компоненты сдвига, не зависящей от других составляющих. Это требование, очевидно, соответствует режиму слежения по всем компонентам сдвига, который реализуется при наличии отдельных следящих приводов для каждой из них и после завершения переходных процессов, возникающих в момент перехода от поиска к захвату.

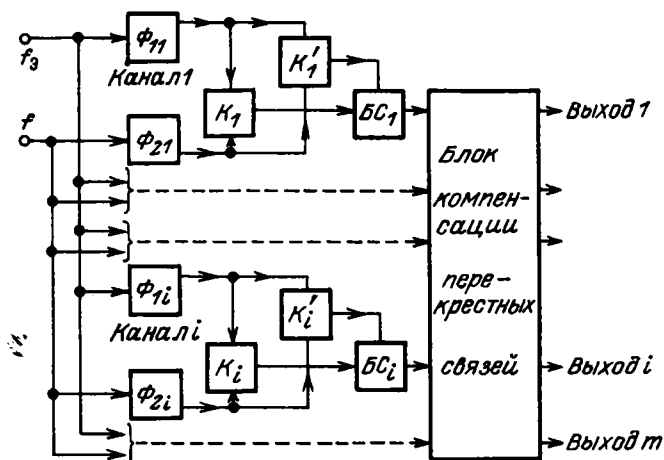


Рис. 1.7. Обобщенная структурная схема многомерного корреляционного дискриминатора

Заметим, что задачи оценки неизвестного сдвига двух сходных изображений в корреляционных системах решаются не только в режиме слежения, когда сдвиги малы, но, если нужно, и непосредственно на этапе поиска, когда сдвиги могут быть относительно велики. Соответственно этому требования к точности оценок, скорости их получения и к размерам области допустимых значений (т.е. однозначных отсчетов) начальных рассогласований могут оказаться существенно различными (например, в навигационных автоматах для "маршрутной" коррекции по картам или аэроснимкам местности [64, 77]).

4. Поскольку основной функцией измерительного органа в следящей системе является оценка некомпенсированных сдвигов между эталонным f_3 и текущим f сигналами, то к составу КД относятся все элементы, обеспечивающие независимость таких оценок от структуры изображений: стабилизаторы крутизны ДХ, компенсаторы перекрестных связей и т.д.

5. Так как спектральный состав коррелируемых сигналов прямо влияет на их взаимно корреляционную функцию ¹⁾

$$K_f(\Delta) = (\check{f}_3 \otimes f)(\Delta), \quad (1.2.1)$$

все линейные фильтры, а при необходимости и существенно нелинейные преобразователи сигналов, так же как и любые элементы, влияющие на ВКФ и, следовательно, на ДХ, относятся к составу КД.

Вместе с тем элементы, обеспечивающие нелинейное преобразование уже сформированных ДХ (ограничители, квантователи и т.п. устройства) и не участвующие в образовании нужных ВКФ, в состав КД можно не включать, учитывая их действие лишь при расчете фактической зависимости управления от рассогласований. Названные признаки многомерного КД объединяет обобщенная структурная схема (рис. 1.7), которую можно интерпре-

¹⁾ Здесь и далее $\otimes$ — символ свертки, а символ $\check{}$ над функцией означает замену знака (инверсию) ее аргумента. См. также § 1.4.

тировать и как схему алгоритма КД. В составе схемы выделены первый и i -й каналы формирования скалярных оценок компонент m -мерного сдвига между изображениями f_3 и f . Каждый i -й канал содержит пару линейных или нелинейных фильтров Φ_{1i} и Φ_{2i} перед коррелятором K_i , выходная характеристика которого в окрестности нуля представляет ненормированную оценку $\bar{\Delta}_i$ сдвига Δ_i . В ней также не учтены поправки на перекрестные связи, могущие иметь место из-за структурных особенностей входных сигналов. Процедура нормировки (стабилизации крутизны) ДХ и компенсация перекрестных связей выполняются блоками стабилизации (БС) и общим для всех каналов КД блоком компенсации. Для их работы могут быть необходимы сведения об энергетических (взаимно энергетических) параметрах сигналов в своем канале (их можно получать в корреляторах K'_i) и значения оценок на выходах соседних каналов. Необходимые связи показаны на рис. 1.7.

Приведенная структурная схема является эвристической и отражает лишь качественные рассуждения о возможностях измерения сдвигов путем корреляционного сравнения сигналов, не содержащих помех.

В следующем параграфе мы покажем, что та же обобщенная структура может быть найдена по алгоритму оптимальной системы оценивания параметров.

§ 1.3. Синтез корреляционного дискриминатора по алгоритму оптимального многомерного оценивания

Постановка задачи. Для синтеза КД, работающего в составе корреляционной ОЭСС, рассмотрим режим слежения, который характерен малыми значениями компонент сдвига $(\Delta_1, \Delta_2, \dots, \Delta_m) = \underline{\Delta}$. В целях упрощения математической модели значения вектора параметров дискретизируют, полагая их неизменными за время измерения в каждом очередном кадре наблюдения текущего изображения. Такое допущение соответствует возможностям современной телевизионной техники в большинстве систем видимого диапазона ЭМВ. Оно приемлемо и как первое приближение в описании корреляционных автоматов с меньшим быстродействием.

Указанный подход позволяет отделить задачу синтеза наилучшего фильтра в контуре следящего измерителя от задачи оптимизации дискриминатора, который должен формировать после регистрации очередного текущего изображения наиболее точную оценку $\hat{\underline{\Delta}}$ вектора постоянных параметров $\underline{\Delta} = (\Delta_1, \Delta_2, \dots, \Delta_m)$. Его компоненты суть составляющие многомерного сдвига изображений, оставшегося к моменту измерения некомпенсированным через органы управления ориентацией приемника, параметрами его оптической и развертывающей систем или ЗУ для хранения эталона. Иными словами, в нашей модели $\underline{\Delta}$ есть дискретизованная по времени ошибка слежения за вектором параметров $\underline{\lambda}(t)$ в следящем корреляционном измерителе.

Положив, что помеха n представляет случайную аддитивную ¹⁾ добавку к одному из сравниваемых сигналов, получим набор исходных условий

¹⁾ В § 1.9 будут рассмотрены модели, учитывающие и мультипликативные мешающие факторы.

для типичной задачи оценки неизвестного постоянного вектора параметров сигнала в аддитивных помехах:

$$f_3 = f_3(x_1, x_2), \quad f = f(x_1, x_2, \underline{\Delta}), \quad n = n(x_1, x_2), \quad y = f_3 + n, \quad (1.3.1)$$

где x_1, x_2 — координаты точки на плоскости в декартовой системе.

Существенно, что f здесь является нелинейной функцией вектора параметров $\underline{\Delta}$, которая носит случайный характер на множестве возможных изображений, но вполне детерминирована для каждого из них. Аналогичным образом $\underline{\Delta}$ может рассматриваться как взятое в фиксированный момент времени значение случайной функции $\underline{\Delta}(t)$, реализации которой различны для разных условий работы управляемого объекта, его динамических свойств и т.п. В то же время вектор $\underline{\Delta}$ вполне детерминирован для каждого кадра наблюдения в работающей корреляционной системе. Сделанные замечания важны для правильного выбора критериев оптимизации многомерного дискриминатора, который, очевидно, должен хорошо работать не только "в среднем", но и в каждой точке возможного диапазона значений компонент $\underline{\Delta}$ при измерении сдвига на любом из множества возможных изображений $\{f\} = \mathcal{M}_f$.

Проблема синтеза КД в свете обобщенного метода наименьших квадратов. Математический аппарат современной теории оценок содержит некоторые общие правила нахождения наилучшей оценки неизвестного вектора параметров, сохраняющие свою силу для различных вариантов критерия оптимизации и при минимальных ограничениях на статистику изображений и помех. Эти правила достаточно компактно сформулированы, в частности, в упомянутой работе [112], обобщающей ряд предыдущих результатов. В ней установлено, что каждая задача оптимальной линейной фильтрации и предсказания случайных процессов может быть сформулирована как совершенно эквивалентная ей задача оценки неслучайного вектора параметров методом наименьших квадратов. Формально оцениваемый вектор мог бы представлять даже одну из реализаций некоторого случайного процесса $\underline{\Delta}(t)$, так как последний легко описывается коэффициентами разложения по какому-либо ортонормированному базису, например Карунена—Лоэва [22]. Чтобы упростить интерпретацию обобщенного алгоритма оценивания, ограничимся случаем, когда $\underline{\Delta}$ определяется набором m чисел, а не функций времени. Положим также, что сигналы и шум представимы конечным числом перенумерованных отсчетов, взятых по определенному правилу дискретизации в M точках плоскости $\{x_1, x_2\}$, ограниченной полем зрения приемника изображений, т.е.¹⁾

$$f = \|f_1(\underline{\Delta}) \quad f_2(\underline{\Delta}) \quad \dots \quad f_M(\underline{\Delta})\|^T = \|f_\mu(\underline{\Delta})\|^T,$$

$$n = \|n_1 \quad n_2 \quad \dots \quad n_M\|^T = \|n_\mu\|^T,$$

$$f_3 = \|f_1(0) \quad f_2(0) \quad \dots \quad f_M(0)\|^T = \|f_\mu(0)\|^T \quad (\mu = 1, 2, \dots, M).$$

¹⁾ Здесь и далее двойные прямые скобки $\|\cdot\|$ означают матрицу; в данном случае она представляет собой вектор-столбец, который для удобства записан как вектор-строка, а затем транспонирован.

Это позволяет оперировать с изображениями как с конечномерными векторами, поддающимися матричным преобразованиям.

Задача, решаемая оптимальным дискриминатором, в нашей постановке заключается в нахождении такой оценки вектора параметров $\hat{\Delta}^{(1)} = \|\hat{\Delta}_1 \hat{\Delta}_2 \dots \hat{\Delta}_m\|^T = \|\hat{\Delta}_i\|^T$ ($i = 1, \dots, m$), которая минимизирует квадратичный функционал

$$\theta = \sum_{\mu, \nu=1}^M [\gamma_\mu - f_\mu(\Delta)] q_{\mu\nu} [\gamma_\nu - f_\nu(\Delta)], \quad (1.3.2)$$

где $\|q_{\mu\nu}\| = Q$ — положительно определенная симметрическая матрица, иногда называемая *весовой*.

В терминах линейной алгебры функционал θ представляет собой квадрат расстояния или, иначе, нормы разности между векторами y и f , заданными в M -мерном унитарном векторном пространстве, определяемом совокупностью M линейно независимых базисных элементов. Таким образом,

$$\theta = |y - f|^2 = (y - f)^T Q (y - f).$$

Процедура, минимизирующая (1.3.2) на множестве значений $\{\Delta\}$, характерна для обобщенного метода наименьших квадратов. Он отличается от классического метода наименьших квадратов тем, что в последнем допускаются только диагональные матрицы Q , тогда как в обобщенном достаточно обеспечить лишь их симметрию.

Выбор матрицы Q придает разное содержание понятию "близости" векторов y и f в пространстве сравниваемых сигналов. Другими словами, меняя Q , можно задавать различную метрику пространства сигналов или, что то же, придавать разную значимость корреляции элементов, находящихся на удалении друг от друга в сравниваемых изображениях. Похожий физический смысл имеет матрица Q и в случае представления сигналов набором спектральных коэффициентов вместо набора отсчетов по полю зрения приемника.

Для пояснения связи между обычным и обобщенным методами наименьших квадратов напомним, что всякую нормальную квадратную матрицу Q можно диагонализировать, подобрав новую, ортонормированную систему координат для представления векторов и операторов. Необходимое преобразование выполняется с помощью невырожденной унитарной преобразующей матрицы T , составленной с использованием собственных значений и собственных векторов матрицы Q . При этом имеет место тождество

$$(y - f)^T Q (y - f) \equiv (\tilde{y} - \tilde{f})^T \tilde{Q} (\tilde{y} - \tilde{f}),$$

где

$$\tilde{Q} = T^{-1} Q T, \quad (\tilde{y} - \tilde{f}) = T^{-1} (y - f).$$

Здесь $\tilde{Q}$ — диагональная матрица; $\tilde{y}$ и $\tilde{f}$ — векторы в новой координатной системе такой, что функционал θ выражается суммой квадратов разностей проекций $\tilde{y}$ и $\tilde{f}$ на координатные оси.

Из сказанного видно, что весовая матрица (метризующий оператор) должна иметь разную структуру в зависимости от критерия "близости"

сигналов y и f , используемого для нахождения наилучшей оценки $\hat{\Delta}^{(1)}$. Установлено [22, 112], что при нахождении оценок $\hat{\Delta}^{(1)}$ по наблюдениям y минимизацией θ за счет вариации вектора Δ при несущественных для практики ограничениях на поведение функции f справедливо тейлоровское разложение

$$\hat{\Delta}^{(1)} - \Delta = (B_f(\Delta))^{-1} \left(\frac{df(\Delta)}{d\Delta} \right)^T Qn + r_{\Delta}(k), \quad (1.3.3)$$

где $r_{\Delta}(k)$ — сумма всех членов разложения по степеням параметра Δ , начиная с $k = 2$, представляющая собой вклад, на порядок меньший первого слагаемого; $df(\Delta)/d\Delta$ — производная вектора сигнала $f(\Delta)$ по вектору параметров Δ , представляющая собой прямоугольную матрицу Якоби размера $M \times m$ и выражаемая при необходимости произведением вектора ∂ операторов дифференцирования по составляющим Δ на вектор сигнала $f(\Delta)$, так что

$$\left(\frac{\partial f(\Delta)}{d\Delta} \right)^T = \left\| \frac{\partial}{\partial \Delta_1} \dots \frac{\partial}{\partial \Delta_m} \right\|^T \|f_1(\Delta) \dots f_M(\Delta)\| = \partial f^T(\Delta), \quad (1.3.4)$$

$$B_f(\Delta) = \left(\frac{df(\Delta)}{d\Delta} \right)^T Q \left(\frac{df(\Delta)}{d\Delta} \right), \quad (1.3.5)$$

где

$$\partial = \left\| \frac{\partial}{\partial \Delta_1} \dots \frac{\partial}{\partial \Delta_m} \right\|^T; \quad f^T = \|f_1(\Delta) \dots f_M(\Delta)\|.$$

В области линейной зависимости f от Δ векторная производная $df(\Delta)/d\Delta$ и матрица $B(\Delta)$ становятся постоянными и наилучшая оценка вектора параметров отыскивается вполне точно по формуле

$$\hat{\Delta}^{(1)} = B_f^{-1} \left(\frac{df}{d\Delta} \right)^T Q(y - f), \quad (1.3.6)$$

где f — значение $f(\Delta)$ при $\Delta = 0$.

В нелинейном случае это соотношение является приближенным и входящие в (1.3.6) первые два сомножителя зависят от Δ .

Пренебрежение этой зависимостью вызывает погрешности, возрастающие по мере роста нормы вектора $|\Delta|$.

При непосредственном корреляционном сравнении неотфильтрованных мелкоструктурных изображений область линейной зависимости $\hat{\Delta}_i$ от компонент сдвига Δ_i и ширина (апертура) монотонной зоны характеристики КД по соответствующей выходной величине обычно невелики. Однако используя подходящие операторы Q , можно выделить из сигналов компоненты, для которых линейная зона характеристики увеличена. При этом расширяется апертура измерительного звена корреляционной следящей системы (КСС). В последующих параграфах эта возможность обсуждается подробнее.

Утверждения, содержащиеся в соотношениях (1.3.3) — (1.3.6), верны безотносительно к тому, рассматриваются ли векторы n как стохастические переменные или нет. Это всего лишь правила, обеспечивающие наилучшее

приближение y и f в векторном пространстве сигналов с выбранной метрикой.

Они могут быть использованы для получения известных статистических теорем, если задать статистику помехи n . Так, например, задав матрицу Q как обратную к ковариационной матрице помех R_n , можно, не делая предположений о гауссовой статистике помех, найти, что вектор $\hat{\Delta}$, вычисленный по формуле (1.3.6) при $Q = R_n^{-1}$, есть вектор оценок сдвига с асимптотически минимальной среднеквадратической погрешностью среди всех регулярных оценок [61]. Если же компоненты n являются совместно гауссовыми, то в дополнение к сказанному составляющие $\hat{\Delta}^{(1)}$ являются оценками максимального правдоподобия.

Полезно, однако, подчеркнуть важность случаев, когда $Q \neq R_n^{-1}$. Имеется ряд причин, побуждающих отказываться от непосредственного ввода обратной ковариационной матрицы помех в алгоритм оценивания при его практической реализации в корреляционном дискриминаторе. Главные из них состоят в ненадежности априорных данных о ковариационных свойствах ожидаемых помех, а также в заметной зависимости размеров области, где алгоритм оценивания сохраняет "работоспособность", от структуры оператора Q . Анализ многих эвристических вариантов построения корреляционных автоматов показывает, что имеет смысл рассматривать правило оценивания (1.3.6) как алгоритм, позволяющий непосредственно синтезировать структуру корреляционного дискриминатора¹⁾.

Состав правой части (1.3.6) таков, что в ней можно выделить: информационные члены — векторы y и f , "фильтрующие" операторы — дифференцирования $\underline{d}$ и линейного преобразования Q^2) и, наконец, нормирующий множитель — симметрическую матрицу B_f^{-1} размера $m \times m$. Ее диагональные элементы обеспечивают получение составляющих оценки $\hat{\Delta}^{(1)}$ в единицах нужной размерности для каждой из компонент³⁾, а недиагональные вводят поправки на перекрестные связи, т.е. обеспечивают независимость компонент оценки. Произведение всех сомножителей справа в (1.3.6), кроме B_f^{-1} , представляет собой ненормированную оценку многомерного сдвига $\bar{\Delta}$ с зависимыми (в общем случае) составляющими, так что

$$\hat{\Delta}^{(1)} = B_f^{-1} \bar{\Delta}, \quad (1.3.7)$$

где

$$\bar{\Delta} = \underline{d} f^T Q (y - f). \quad (1.3.8)$$

¹⁾ Такой синтез возможен без уточнения критерия оптимизации оценки, и, следовательно, синтезированные этим путем дискриминаторы правильнее было бы называть не оптимальными, а оптимизируемыми, т.е. поддающимися оптимизации путем надлежащего выбора оператора Q , соответствующего назначенному (быть может, и нестатистическому) критерию.

²⁾ В теории цепей действие линейного оператора на функцию, описывающую сигнал, отождествляется с его линейной фильтрацией.

³⁾ Следует помнить, что векторное пространство возможных сдвигов изображений не всегда является нормированным.

Исследование обобщенного алгоритма многомерного оценивания по его представлениям. Впервые синтез структуры оптимального дискриминатора по его алгоритму, полученному на основании байесовского критерия оптимизации с соответствующими ограничениями, был выполнен И.А. Большаковым и В.Г. Репиным [16]. Они назвали собственно дискриминатором блок, формирующий ненормированную оценку (1.3.8), одновременно рассмотрев схему с каналами нормировки и компенсации перекрестных связей.

По существу, устройство оценки текущего вектора параметров у И.А. Большакова и В.Г. Репина, описанное в [7, 16], представляет интересующий нас многомерный КД, который производит одновременное корреляционное измерение несольких параметров, отличающих два сравниваемых сигнала друг от друга. Четкое и сжатое изложение байесовских методов синтеза оптимальных "корреляционно-фильтровых" многомерных дискриминаторов дано в содержательной монографии Я.Д. Ширмана и В.Н. Манжоса [140].

Для разработчиков оптико-электронных систем (ОЭС) важно, какое физическое содержание могут иметь операторы фильтрации и нормировки в случаях, когда сигналы являются сложными двумерными функциями, несущими информацию о меняющихся изображениях. Остановимся на этом вопросе подробнее, ибо в названных работах он не освещен. В нашем рассмотрении критерий оптимизации КД в целях сохранения общности результатов конкретизироваться не будет. Заметим только, что матрица Q является симметрической и представляет собой линейный оператор с параметрами, постоянными в каждом очередном цикле (кадре) измерения многомерного взаимного сдвига $\underline{\Delta}$. Из теории линейных операторов известно, что всякая квадратная матрица Q с заданными свойствами может быть факторизована и представлена произведением

$$Q = LL^T, \quad (1.3.9)$$

где L — нижняя (или верхняя) треугольная матрица того же размера, что и Q . Известны и правила вычисления элементов, входящих в L [29].

Подставив (1.3.9) в (1.3.8) и раскрыв скобки, получим один из вариантов алгоритма ненормированной оценки:

$$\bar{\Delta}^{(1)} = \underline{\Delta}(L^T f)^T (L^T y) - \underline{\Delta}(L^T f)^T (L^T f). \quad (1.3.10)$$

Преобразования выявляют следующие важные особенности алгоритма.

1. Информационной частью в составе оценки $\bar{\Delta}^{(1)}$ служит первое слагаемое, которое можно трактовать либо как производную по вектору сдвига $\underline{\Delta}$ от взаимно корреляционной функции (ВКФ) профильтрованных сигналов $L^T f$ и $L^T y$, либо как ВКФ производной первого из них по $\underline{\Delta}$ и второго, недифференцированного.

2. Второе слагаемое в правой части (1.3.10) представляет собой вектор из производных по компонентам $\underline{\Delta}$ от автокорреляционной функции (АКФ) сигнала в точке, где $\Delta_i \equiv 0$ для $i = 1, 2, \dots, m$. При этом во всяком случае, когда множества значений $\{\Delta_i\}$ симметричны относительно нуля, производные по соответствующим Δ_i обращаются в нуль, что позволяет исключить это слагаемое из алгоритма (1.3.10).

Если отнести операцию дифференцирования к первому из двух "профильтрованных" сигналов-сомножителей во втором слагаемом, то вектор, полученный в результате, станет ортогональным к $L^T f$. Это также объясняет обращение второго слагаемого в нуль.

3. Линейные преобразования, необходимые для нахождения оценки сдвига, могут относиться к любому из двух сравниваемых сигналов. Например, формула (1.3.10) с учетом сказанного, а также п. 2 может быть переписана в виде

$$\bar{\Delta} = \partial(Qf)^T y. \quad (1.3.11)$$

Такое представление оценки непосредственно указывает на экономный вариант структуры корреляционного дискриминатора, в котором все операторы фильтрации использованы для преобразования только одного из сравниваемых сигналов. Им может быть, например, редко сменяемое эталонное изображение, в то время как роль другого будет выполнять текущий сигнал, обработка которого должна производиться чаще и быстрее. Выгоды указанного варианта могут оказаться ощутимыми при обработке изображений в текущем времени.

Как при подборе адекватной математической модели КД для эвристически найденного технического решения, так и при поиске новых вариантов подобных решений полезно иметь в виду возможные способы записи алгоритма многомерного оценивания. Приведем в качестве примера цепочку некоторых формул ненормированной оценки, опуская промежуточные выкладки, связанные с последовательным применением известных правил перестановки при транспонировании элементов в векторно-матричных выражениях [29]:

$$\bar{\Delta} = \partial f^T Q y = \partial y^T Q^T f = \partial [(Ly)^T (Lf)] = [\partial (Ly)^T] Lf = \partial (L_1 f)^T L_2 y = \dots, \quad (1.3.12)$$

где L_1 и L_2 — невырожденные квадратные матрицы размера $M \times M$ такие, что

$$L_1^T L_2 = Q. \quad (1.3.13)$$

Прямоугольные скобки в (1.3.12), объединяющие те или иные сомножители, определяют возможный порядок выполнения операций с сигналами в структурной схеме прибора, который реализует алгоритм КД.

Использование дуальных решений в алгоритме КД. Изучение всего многообразия инвариантных аналитических представлений оптимизируемого алгоритма корреляционного дискриминатора интересно в практическом смысле, поскольку арсенал уже готовых способов реализации КД весьма обширен [2, 13, 17, 23, 25, 36, 45, 55, 58, 59, 85, 86, 117, 121, 123, 127, 128, 142, 160, 162, 172, 174, 176], а их сравнение между собой и с вновь изобретаемыми способами при эмпирическом подходе неэффективно. Дуальность вариантов, т.е. возможность взаимной перестановки векторов f и y , использованная в (1.3.12), основана на том, что каждый элемент вектора-столбца $\bar{\Delta}$ представляет собой результат дифференцирования по компоненте Δ_i одного и того же скалярного произведения, в котором f и y коммутативны.

При рассмотрении хода теоретического решения дуальность ожидаемых алгоритмов в более общем виде выявляется еще на этапе постановки задачи. Действительно, корреляционное измерение взаимного сдвига сигналов f и y можно осуществить, компенсируя этот сдвиг изменением как первого изображения, так и второго. Следовательно, во всех формулах, приводящих в обозначениях (1.3.1) от функционала (1.3.2) к соотношениям (1.3.6) – (1.3.12), можно без ущерба для общности результатов поменять местами векторы y и f , полагая первый, а не второй зависящим от Δ . Очевидно также, что аддитивную помеху n , отнесенную нами к одному из сигналов, с равным успехом можно переносить в состав другого с противоположным знаком. В результате перестановок получится алгоритм оценки, дуальный к (1.3.6), в виде

$$\hat{\Delta}_y^{(2)} = B_y^{-1} \bar{\Delta}, \quad (1.3.14)$$

где второй множитель в правой части, как это следует из (1.3.12), тот же, что и в (1.3.7), а нормирующая матрица имеет вид

$$B_y(\Delta) = \left(\frac{dy(\Delta)}{d\Delta} \right)^T Q \left(\frac{dy(\Delta)}{d\Delta} \right).$$

Таким образом, дуальная оптимизируемая оценка вектора сдвига $\hat{\Delta}^{(2)}$ отличается от $\hat{\Delta}^{(1)}$, поскольку $B_f \neq B_y$.

Нормирующие матрицы в общем случае не совпадают потому, что зависимости $f(\Delta)$ и $y(\Delta)$ различны, хотя, может быть, и близки между собой. Учитывая (1.3.1), можно показать, что разница оценок $\hat{\Delta}^{(1)}$ и $\hat{\Delta}^{(2)}$ имеет величину порядка среднеквадратической погрешности каждой из них из-за помехи n . При этом для помех n с нулевыми средними значениями компонент легко убедиться в статистической эквивалентности дуальных оценок.

Между прочим, матрицы $R_\epsilon^{(1)}$ и $R_\epsilon^{(2)}$ ковариаций погрешностей оценок определяются в линейном приближении при заданной ковариационной матрице помехи R_n и нулевом среднем значении последней однотипными выражениями:

$$R_\epsilon^{(1)} = B_f^{-1} \partial^T f Q R_n Q \partial f^T B_f^{-1},$$

$$R_\epsilon^{(2)} = B_y^{-1} \partial^T y Q R_n Q \partial y^T B_y^{-1}.$$

В случае $Q = R_n^{-1}$ составляющие $\Delta^{(1)}$ и $\Delta^{(2)}$ становятся оценками с минимальными среднеквадратическими погрешностями (хотя бы в асимптотическом смысле), а матрицы их ковариаций приобретают вид нормирующих матриц

$$R_\epsilon^{(1)} = B_f^{-1}, \quad R_\epsilon^{(2)} = B_y^{-1}.$$

Возвращаясь к вопросу о дуальности оптимальных оценок сдвига, остановимся на их графическом представлении. Пусть Δ_0 – скалярный параметр. Если Δ_0 – истинное его значение, то оценки $\hat{\Delta}^{(1)}$ и $\hat{\Delta}^{(2)}$, полученные при конкретном измерении в одном кадре наблюдения, могут располагаться в окрестности Δ_0 , как показано на рис. 1.8, а. При этом не исключается возможность, что обе оценки, находясь поблизости от точки Δ_0 , ока-

жуются в каком-либо измерении по одну сторону от нее. В двумерном случае, очевидно, точки Δ_0 , $\hat{\Delta}^{(1)}$ и $\hat{\Delta}^{(2)}$ образуют треугольник (рис. 1.8, б), в котором положение двух вершин из трех меняется случайным образом от одного измерения к другому. Все три точки сливаются, если помеха $n \equiv 0$. Знание дуальных оценок позволяет установить некоторые границы множества допустимых решений, т.е. эквивалентных по своей оптимальности алгоритмов корреляционного дискриминатора.

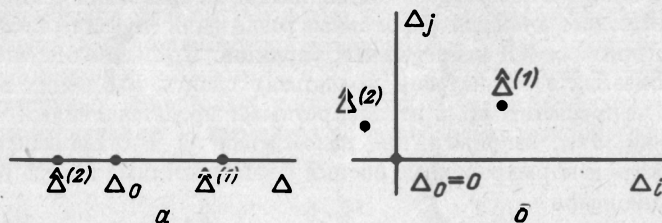


Рис. 1.8. Дуальные оптимальные оценки в пространстве одномерных (а) и двумерных (б) сдвигов

Поскольку статистика сигналов и помех нами не оговорена, обсуждение структуры множества решений беспредметно. Пытаясь нащупать некоторые границы в размытом из-за недетерминированности помех множестве эквивалентных оценок, можно только предположить, что компоненты $\hat{\Delta}_i$ оценки $\hat{\Delta}$, получаемой в дискриминаторе, вероятнее всего находятся между значениями компонент $\hat{\Delta}_i^{(1)}$ и $\hat{\Delta}_i^{(2)}$ дуальных оценок $\hat{\Delta}^{(1)}$ и $\hat{\Delta}^{(2)}$, оптимизированных в определенном смысле. Тогда многие из допустимых алгоритмов оптимального корреляционного дискриминатора удастся сконструировать, комбинируя обе оценки линейно, что равносильно комбинированию нормирующих матриц по формуле

$$B^{-1} = AB_f^{-1} + (I - A)B_y^{-1}, \quad (1.3.15)$$

где I — единичная матрица размера $m \times m$; $A = \|a_{ij}\|$ — диагональная матрица размера $m \times m$, у которой значение каждого диагонального элемента лежит в интервале от нуля до единицы.

Согласно (1.3.15), оценка $\hat{\Delta}$ отображается концом вектора, лежащим внутри m -мерного параллелепипеда, главная диагональ которого соединяет точки $\hat{\Delta}^{(1)}$ и $\hat{\Delta}^{(2)}$, а грани его устанавливают искомые границы нечеткого множества.

Таким образом, достаточно общей формой линейного алгоритма оптимального корреляционного измерения вектора сдвига путем сравнения взаимно сдвинутых M -мерных сигналов является выражение, аналогичное (1.3.7) и (1.3.14),

$$\hat{\Delta} = B^{-1} \bar{\Delta},$$

где, например, первый множитель в правой части берется с учетом (1.3.15), а второй — с использованием вариантов (1.3.12).

Заметим, что комбинации вида (1.3.15) не исчерпывают всех возможностей построения нормирующего фактора B . Их количество растет,

если искать нормирующую матрицу как смешанное произведение вида $\left(\frac{dy}{d\Delta}\right)^T Q \left(\frac{df}{d\Delta}\right)$ вместо принятого в (1.3.14).

В то же время, как можно заключить из ряда теорем, доказанных в [102], такая нормировка становится наилучшей в случае, когда энергии переменных слагающих у сравниваемых изображений различаются даже при отсутствии помех, а постоянные составляющие из сравнения исключены, т.е. под y, f и n понимаются центрированные реализации случайных полей.

Об алгоритмах КД непрерывных сигналов. При необходимости допустимо производить вычисления компонент сдвига, входящих в оценку, пользуясь непрерывными, а не дискретными представлениями сигналов-изображений. Так, например, при нахождении i -й составляющей сдвига можно искать ненормированные оценки с компонентами в виде непрерывных функционалов

$$[\partial f^T Q y]_i = \iint_R \left[\frac{\partial}{\partial \Delta_i} f(\mu, \nu) \right] q(\mu, \nu) y(\mu, \nu) d\mu d\nu, = \iint f \frac{\partial y}{\partial \Delta_i} q d\mu d\nu$$

которые являются аналогами дискретных форм

$$[\partial f^T Q y]_i = \sum_{\mu, \nu=1}^M \left(\frac{\partial}{\partial \Delta_i} f_{\mu} \right) q_{\mu\nu} y_{\nu}, = \Delta_i \left(\sum_{\mu, \nu=1}^M f_{\mu} y_{\nu} \right)$$

а элементы нормирующих матриц будут определяться интегралами вида

$$[\partial f^T Q \partial^T f]_{ij} = \iint_R \left[\frac{\partial}{\partial \Delta_i} f(\mu, \nu) \right] q(\mu, \nu) \left[\frac{\partial}{\partial \Delta_j} f(\mu, \nu) \right] d\mu d\nu,$$

$$[\partial y^T Q \partial^T y]_{ij} = \iint_R \left[\frac{\partial}{\partial \Delta_i} y(\mu, \nu) \right] q(\mu, \nu) \left[\frac{\partial}{\partial \Delta_j} y(\mu, \nu) \right] d\mu d\nu$$

с соответствующими дискретными аналогами

$$[\partial f^T Q \partial^T f]_{ij} = \sum_{\mu, \nu=1}^M \left(\frac{\partial}{\partial \Delta_i} f_{\mu} \right) q_{\mu\nu} \left(\frac{\partial}{\partial \Delta_j} f_{\nu} \right),$$

$$[\partial y^T Q \partial^T y]_{ij} = \sum_{\mu, \nu=1}^M \left(\frac{\partial}{\partial \Delta_i} y_{\mu} \right) q_{\mu\nu} \left(\frac{\partial}{\partial \Delta_j} y_{\nu} \right).$$

Здесь во всех шести выражениях индексы у квадратных скобок в левых частях означают i -ю компоненту сдвига или ij -й элемент нормирующей матрицы, R определяет двумерную замкнутую область интегрирования так, чтобы охватить все точки, где подынтегральное выражение отлично от нуля, а сочетание символов типа $\frac{\partial}{\partial \Delta_i} f_{\mu}$ означает μ -ю точку на дискретизованном поле производных непрерывного сигнала f по i -й компоненте сдвига Δ .

Приведенные интегралы имеют смысл, если функции, выбранные для описания производных от сигналов по сдвигу, являются суммируемыми [140]. Это условие всегда выполнимо для реальных изображений.

Непрерывные представления сигналов и алгоритмов их обработки удобны при исследовании аналоговых способов преобразования входной информации, тогда как дискретные более приемлемы для цифровых систем. Возвращаясь к дискретным представлениям, остановимся на структуре нормирующего множителя в алгоритме КД.

Нормирование оценок и компенсация перекрестных связей в алгоритмах многомерных КД. Для уяснения физического смысла операции нормирования рассмотрим пример алгоритма корреляционного дискриминатора, оценивающего вектор параметров размерностью $m = 2$. Это может соответствовать измерению компонент плоского сдвига-переноса изображений в прямоугольных декартовых координатах. Считая сигналы-изображения дискретизованными на M элементов по полю зрения, находим вектор-столбец оцениваемых параметров сдвига

$$\hat{\underline{\Delta}} = \|\hat{\Delta}_1 \hat{\Delta}_2\|^T = B_2^{-1} \left\| \frac{\partial}{\partial \Delta_1} \frac{\partial}{\partial \Delta_2} \right\|^T \|\varphi_1 \varphi_2 \dots \varphi_M\| \|s_1 s_2 \dots s_M\|^T = \\ = B_2^{-1} \underline{\partial \varphi}^T s,$$

где с учетом (1.3.13) введены новые обозначения для сигналов, подвергнутых фильтрации:

$$\|\varphi_1 \varphi_2 \dots \varphi_M\| = \underline{\varphi}^T = \underline{f}^T L_1, \quad \|s_1 s_2 \dots s_M\| = s = L_2^T y.$$

Примем для определенности, что нормирующий множитель формируется из сигнала f , т.е.

$$B_2 = \|\underline{\partial \varphi}^T \underline{\varphi} \underline{\partial \varphi}^T\| = \left\| \begin{array}{cc} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{array} \right\|, \quad (1.3.16)$$

где

$$\sigma_{jl} = \sum_{i=1}^M \frac{\partial \varphi_i}{\partial \Delta_j} \frac{\partial \varphi_i}{\partial \Delta_l} \quad (j, l = 1, 2, \dots). \quad (1.3.17)$$

Из (1.3.17) видим, что $\sigma_{jl} = \sigma_{lj}$, поэтому нормирующая матрица всегда симметрическая. Обычно в реальных изображениях сложной структуры корреляция между частными производными по разноименным компонентам сдвига (перекрестная корреляция) невелика. Если это так, то диагональные элементы в матрице (1.3.16) оказываются много больше недиагональных. Количественную меру ожидаемого различия можно найти, если задаться условием нормальности распределения множеств двумерных функций $\{f\}$ и $\{y\}$ и их ковариациями, как это сделано, например, в [110].

Определитель $\mathcal{D}_2$ матрицы B_2 находится, как обычно:

$$\mathcal{D}_2 = \sigma_{11} \sigma_{22} - \sigma_{12} \sigma_{21} = \sum_{i=1}^M \left(\frac{\partial \varphi_i}{\partial \Delta_1} \right)^2 \sum_{i=1}^M \left(\frac{\partial \varphi_i}{\partial \Delta_2} \right)^2 - \left(\sum_{i=1}^M \frac{\partial \varphi_i}{\partial \Delta_1} \frac{\partial \varphi_i}{\partial \Delta_2} \right)^2,$$

причем второй член много меньше первого.

Руководствуясь правилами обращения матриц и выражая в силу сказанного определитель приближенным равенством

$$\mathcal{D}_2 \approx \sigma_{11} \sigma_{22},$$

получим нормирующую матрицу в виде

$$B_2^{-1} = \frac{1}{\mathcal{D}_2} \begin{vmatrix} \sigma_{22} & -\sigma_{12} \\ -\sigma_{21} & \sigma_{11} \end{vmatrix} \approx \begin{vmatrix} \frac{1}{\sigma_{11}} & \frac{-\sigma_{12}}{\sigma_{11} \sigma_{22}} \\ \frac{-\sigma_{21}}{\sigma_{11} \sigma_{22}} & \frac{1}{\sigma_{22}} \end{vmatrix}.$$

Обозначим компоненты ненормированного вектора оценок принятыми ранее символами:

$$\bar{\Delta} = \partial \varphi^T s = \|\bar{\Delta}_1 \quad \bar{\Delta}_2\|^T.$$

После нормирования имеем

$$\hat{\Delta} = B_2^{-1} \bar{\Delta} \quad (1.3.18)$$

или в скалярном виде

$$\hat{\Delta}_1 = \frac{\sigma_{22}}{\mathcal{D}_2} \left(\bar{\Delta}_1 - \frac{\sigma_{12}}{\sigma_{22}} \bar{\Delta}_2 \right) \approx \frac{1}{\sigma_{11}} \left(\bar{\Delta}_1 - \frac{\sigma_{12}}{\sigma_{22}} \bar{\Delta}_2 \right), \quad (1.3.19)$$

$$\hat{\Delta}_2 = \frac{\sigma_{11}}{\mathcal{D}_2} \left(\bar{\Delta}_2 - \frac{\sigma_{21}}{\sigma_{11}} \bar{\Delta}_1 \right) \approx \frac{1}{\sigma_{22}} \left(\bar{\Delta}_2 - \frac{\sigma_{21}}{\sigma_{11}} \bar{\Delta}_1 \right), \quad (1.3.20)$$

где

$$\bar{\Delta}_1 = \sum_{i=1}^M \left(\frac{\partial \varphi_i}{\partial \Delta_1} s_i \right); \quad (1.3.21)$$

$$\bar{\Delta}_2 = \sum_{i=1}^M \left(\frac{\partial \varphi_i}{\partial \Delta_2} s_i \right). \quad (1.3.22)$$

Формулы (1.3.18) – (1.3.22) отражают качественное содержание алгоритма многомерного корреляционного оценивания, его отличия от одномерного случая и общие с ним признаки.

Так, из (1.3.18) видно, что вектор ненормированных оценок, состоящий из компонент, полученных одномерным дифференцированием (1.3.21) и (1.3.22), подвергается при умножении на матрицу B_2^{-1} не только нормировке по величине и размерности, но и повороту. Иными словами, как и предполагалось, оптимальный многомерный алгоритм учитывает необходимость компенсации перекрестных зависимостей между результатами измерения скалярных составляющих относительного сдвига. Это и приводит к появлению вторых членов в скобках в (1.3.19) и (1.3.20).

В случае $m = 3$ третьей составляющей сдвига изображений может являться, например, изменение масштаба или поворот. При этом нормирую-

щий множитель ищется, как обратная матрица третьего порядка¹⁾

$$B_3^{-1} = \begin{vmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{vmatrix} = \\ = \frac{1}{D_3} \begin{vmatrix} \sigma_{22}\sigma_{33} - \sigma_{23}\sigma_{32} & \sigma_{13}\sigma_{32} - \sigma_{12}\sigma_{33} & \sigma_{12}\sigma_{33} - \sigma_{13}\sigma_{22} \\ \sigma_{23}\sigma_{31} - \sigma_{21}\sigma_{33} & \sigma_{11}\sigma_{33} - \sigma_{13}\sigma_{31} & \sigma_{13}\sigma_{21} - \sigma_{11}\sigma_{23} \\ \sigma_{21}\sigma_{32} - \sigma_{22}\sigma_{31} & \sigma_{12}\sigma_{31} - \sigma_{11}\sigma_{12} & \sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21} \end{vmatrix}, \quad (1.3.23)$$

где

$$D_3 = \sigma_{11}\sigma_{22}\sigma_{33} - \sigma_{11}\sigma_{23}\sigma_{32} + \sigma_{12}\sigma_{23}\sigma_{31} - \sigma_{12}\sigma_{21}\sigma_{33} + \\ + \sigma_{13}\sigma_{21}\sigma_{32} - \sigma_{13}\sigma_{22}\sigma_{31}. \quad (1.3.24)$$

Эти выражения упрощаются, если перекрестная корреляция производных по компонентам сдвига мала. Тогда в правой части (1.3.24) можно сохранить лишь первый член, а каждый элемент матрицы (1.3.23) округлить до большего из двух слагаемых. В результате получим

$$B_3^{-1} \approx \begin{vmatrix} \sigma_{11}^{-1} & -\sigma_{12}\sigma_{11}^{-1}\sigma_{22}^{-1} & -\sigma_{13}\sigma_{11}^{-1}\sigma_{33}^{-1} \\ -\sigma_{21}\sigma_{11}^{-1}\sigma_{22}^{-1} & \sigma_{22}^{-1} & -\sigma_{23}\sigma_{22}^{-1}\sigma_{33}^{-1} \\ -\sigma_{31}\sigma_{11}^{-1}\sigma_{33}^{-1} & -\sigma_{32}\sigma_{22}^{-1}\sigma_{33}^{-1} & \sigma_{33}^{-1} \end{vmatrix}.$$

Из упрощенного выражения хорошо видно, что диагональные элементы матрицы нормируют каждую компоненту оценки $\hat{\Delta}$ по величине и размерности, а недиагональные определяют регрессию на нее со стороны других компонент.

Поскольку в рамках допущений, принятых при отыскании алгоритма дискриминатора, операция нормирования может быть только линейной, структуры оценок любой скалярной составляющей m -мерного сдвига однотипны и имеют вид

$$\hat{\Delta}_i = a_i (\bar{\Delta}_i - \sum_{\substack{j=1 \\ j \neq i}}^m b_{ij} \bar{\Delta}_j), \quad (1.3.25)$$

где a_i — нормирующий коэффициент; b_{ij} — коэффициент линейной регрессии на i -ю компоненту со стороны j -й составляющей.

В принципе с ростом абсолютной величины сдвига линейная регрессионная модель может стать слишком грубой. Для получения более точных оценок в области, где ДХ еще монотонно зависит от сдвига, возможен учет регрессий высших порядков, однако в корреляционных следящих измерителях целесообразнее минимизировать ошибки слежения, не прибегая к дальнейшим усложнениям алгоритмов оценки.

Наличие перекрестных членов в виде вторых слагаемых в составе (1.3.25) учитывает возможную анизотропность корреляционных функций двумерных скалярных полей.

¹⁾ Здесь, как и для $m = 2$, предполагается, что необращенная матрица не вырождена.

Наряду с этими особенностями при реализации алгоритмов КД нередко приходится принимать в расчет искажения оценок взаимных корреляций из-за так называемых *краевых эффектов*, обусловленных ограниченностью поля зрения приемника изображений. Следующий параграф посвящен указанным вопросам, а также краткому напоминанию понятий из аппарата обычного корреляционного анализа и теории фильтрации двумерных скалярных сигналов.

§ 1.4. Двумерная корреляция и ее оценка при краевых эффектах

Краевые эффекты в оценках двумерных корреляций изображений из-за ограниченности апертуры ОЭС. Рассмотрение корреляционных функций двумерных скалярных сигналов, т.е. изображений, и соответствующих спектрально-энергетических характеристик следует начать с определения краевого эффекта, наличие которого зависит от условий применения системы, тогда как общие свойства двумерных корреляций сохраняются независимо от них.

Пусть двумерный сигнал, регистрируемый приемником изображений, описывается скалярной функцией $f(x)$. Ее аргумент — двухкомпонентный вектор

$$x = \| x_1 \quad x_2 \|^T,$$

составляющие которого можно задать как функцию времени, если в описании сигнала надо, например, отразить процесс временной развертки. В случае когда фотоприемный тракт не является объектом специального изучения, его свойства идеализируют, полагая функцию f прямо пропорциональной значению наблюдаемого поля в каждой точке x . При необходимости, деформируя шкалу уровней сигнала, нетрудно учесть нелинейную зависимость тока фотоприемника от интенсивности падающего потока излучения, а свертывая исходное поле с апертурной функцией оптико-электронной системы (ОЭС), ввести поправки на разрешающую способность ее звеньев.

Область значений x , в пределах которой регистрируемый сигнал отличен от нуля, ограничена углами зрения оптических элементов ОЭС или размерами развертывающего телевизионного раstra, тогда как границы исходного поля произвольны. Для описания этого ограничения удобно использовать оператор "пространственного стробирования", т.е. умножить неограниченное углами зрения изображение на функцию "окна" $h_0(x)$:

$$f(x) = f_0(x) h_0(x), \quad (1.4.1)$$

где

$$h_0(x) = \begin{cases} 1 & \text{при } x \in R_x, \\ 0 & \text{при } x \notin R_x; \end{cases}$$

R_x — множество всех точек, попадающих в поле зрения системы; $f_0(x)$ — световое поле, проецируемое в плоскость приемника излучений без учета ограничений по углу зрения.

В дальнейшем будем называть *собственно изображением* функцию f , а не f_0 . Опираясь на (1.4.1) и учитывая физическую природу поля, убеж-

дается, что функция f_0 является целой и непрерывной, а функция f — непрерывной интегрируемой и имеет ограниченный носитель [137].

С другой стороны, всякое непрерывное изображение f можно дискретизировать по выбранным координатам и квантовать по уровням так, что новое изображение будет неотлично от исходного с точностью до сколь угодно малых погрешностей.

Таким образом устанавливается эквивалентность представления изображений двумерными непрерывными функциями либо двумерными матрицами конечного порядка из чисел ограниченной разрядности [109]. Матрицы, в свою очередь, можно заменять векторами, переупорядочив нумерацию элементов [100]. Это позволяет при анализе по необходимости пользоваться дискретными либо непрерывными математическими моделями изображений и их характеристик, а названные свойства функций f освобождают нас обычно от детализации вопросов сходимости и существования при интегральных преобразованиях.

Рассмотрим важную особенность оценки корреляции двух детерминированных смещенных изображений в корреляционной ОЭС с ограниченным полем зрения.

Пусть

$$f = f(x) \text{ и } f^\Delta = f(x - \underline{\Delta}) = \check{f}(\underline{\Delta} - x):$$

Двумерная АКФ обычного или "аддитивного" сдвига $\underline{\Delta}$, по определению, есть двумерная свертка f и $\check{f}$, представляемая символически в виде ¹⁾

$$K = \check{f}^\Delta \otimes f, \quad (1.4.2)$$

где $\otimes$ — символ двумерной свертки по области $\mathcal{R}_x$, охватывающей все точки x , для которых отлично от нуля подынтегральное выражение в формуле²⁾

$$K = K(\underline{\Delta}) = \int_{\mathcal{R}_x} \check{f}(\underline{\Delta} - x) f(x) dx = \int_{\mathcal{R}_x} f(x - \underline{\Delta}) f(x) dx. \quad (1.4.3)$$

Подстановка (1.4.1) в (1.4.2) дает

$$K = (\check{f}_0^\Delta \check{h}_0^\Delta) \otimes (f_0 h_0). \quad (1.4.4)$$

В корреляционной ОЭС можно реализовать лишь оценку АКФ

$$\hat{K} = (\check{f}_0^\Delta \check{h}_0^\Delta) \otimes (f_0 h_0). \quad (1.4.5)$$

¹⁾ Индекс Δ в записи часто опускают, понимая, что $K = K(\underline{\Delta})$ — функция сдвига $\underline{\Delta}$, как ее аргумента.

²⁾ Очевидно, если $\mathcal{R}_x$ — прямоугольник со сторонами, параллельными осям декартовых координат (x_1, x_2) и находящимися от них на расстояниях $\pm c_2$ и $\pm c_1$, то при сдвиге $\underline{\Delta} = \|\Delta_1 \Delta_2\|^T$ этот интеграл записывается как двумерный вида

$$K(\underline{\Delta}) = \int_{(-c_1 + \Delta_1) \dots (-c_2 + \Delta_2)}^{c_1 \dots c_2} f[(x_1 - \Delta_1), (x_2 - \Delta_2)] f(x_1, x_2) dx_2 dx_1,$$

где, учитывая финитность подынтегральных функций, можно расширить пределы интегрирования до бесконечных.

Раскрывая символическую запись (1.4.5), находим из (1.4.3) и (1.4.4) разность δK между истинной АКФ и ее оценкой:

$$\delta K(\underline{\Delta}) = K - \hat{K} = \int_{\mathbf{R}_x} \check{f}_0(\underline{\Delta} - \mathbf{x}) f_0(\mathbf{x}) [h_0(\mathbf{x} - \underline{\Delta}) - h_0(\mathbf{x})] d\mathbf{x}. \quad (1.4.6)$$

Легко видеть, что при любой функции f_0

$$\delta K(0) \equiv 0,$$

тогда как для $\underline{\Delta} \neq 0$ разность δK зависит от функции, описывающей изображение. При этом множитель в квадратных скобках под интегралом в (1.4.6) отличен от нуля только на неперекрывающихся краях взаимно сдвинутых окон h_0 и $h_0^{\underline{\Delta}}$. Поэтому функционал δK , представляющий собой погрешность в оценке истинной АКФ, можно назвать краевым эффектом, количественная мера которого определяется разностью АКФ двух участков неограниченного исходного поля f_0 , образующихся при сдвиге $\underline{\Delta}$,

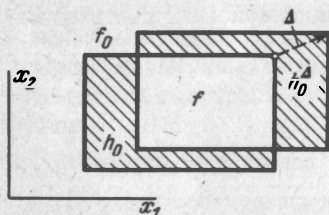


Рис. 1.9. Происхождение краевого эффекта при оценке автокорреляционной функции

как показано штриховкой на рис. 1.9. При использовании корреляционной процедуры для оценки сдвигов, отличных от нуля, краевой эффект, очевидно, также вызывает ошибки.

Между прочим, из (1.4.6) видно, что в случае статистической однородности (стационарности в широком смысле) множества реализаций функции f_0 математическое ожидание краевого эффекта равно нулю при любом $\underline{\Delta}$, откуда

$$M(\hat{K}) = M(K),$$

где $M(\cdot)$ — символ усреднения по множеству, превращающий правую часть в статистическую АКФ некоторого ансамбля случайных функций $\{f\}$, "вырезаемых" окном h_0 из реализаций $\{f_0\}$. Следовательно, $M(\hat{K})$ есть несмещенная оценка АКФ, если f_0 однородна.

Практическое значение этого вывода не слишком велико, так как при каждом конкретном измерении сдвига $\underline{\Delta}$ на вход корреляционного дискриминатора поступает вполне детерминированный, хотя, быть может, и неизвестный заранее двумерный сигнал, представляющий собой реализацию из случайного ансамбля, который в большинстве реальных случаев не является стационарным и тем более эргодическим. Последнее обстоятельство требует известной осмотрительности в использовании стохастических моделей изображений для анализа корреляционных систем, поскольку большинству таких моделей (часто без нужды) приписывается стационарность.

В связи со сделанным замечанием наиболее важными оказываются общие особенности двумерной АКФ $K(\underline{\Delta})$ всякого финитного сигнала $f(x)$, сохраняющиеся независимо от того, входит ли f в ансамбль реализаций какой-то случайной функции.

Перечислим представляющие интерес известные [107, 145] свойства двумерной АКФ и ее фурье-образа – двумерной "энергетической" спектральной плотности (СП) $S(\omega)$.

Особенности двумерных спектрально-корреляционных характеристик финитных изображений. 1. Любое вертикальное центральное сечение двумерной АКФ изображения есть четная одномерная функция своего аргумента, т.е. имеет место центральная симметрия

$$K(\underline{\Delta}) = K(-\underline{\Delta}).$$

Это свойство непосредственно выводится из (1.4.3) путем замены переменной x на $x + \Delta$.

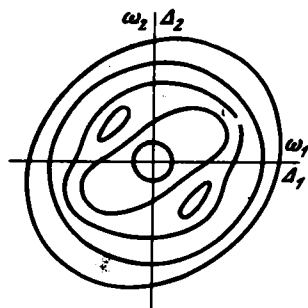
2. Глобальный максимум двумерной АКФ достигается при $\underline{\Delta} = 0$, что следует из свойств вертикальных центральных сечений, представляющих собой в любом случае одномерные АКФ сдвигов по направлению сечений.

3. Любое вертикальное центральное сечение двумерной СП изображения есть четная функция своего аргумента – двумерной пространственной частоты ω , т.е. СП, как и АКФ любого изображения, обладает центральной симметрией

$$S(\omega) = S(-\omega).$$

Центрально-симметричные двумерные функции (АКФ или СП) можно

Рис. 1.10. Типичные сечения автокорреляционной функции или спектральной плотности скалярных финитных сигналов над плоскостью сдвигов или пространственных частот соответственно



представить горизонтальными сечениями – линиями равного уровня, как показано на рис. 1.10.

4. СП и АКФ изображения f суть вещественные функции. Это следует из определения (1.4.3) и сопряженного с ним соотношения, в частности¹⁾

$$\mathcal{F}_f^* \mathcal{F}_f = S(\omega) \Leftrightarrow K(\underline{\Delta}) = \check{f} \otimes f. \quad (1.4.7)$$

Здесь и далее * – символ комплексной сопряженности, $\check{f}$ означает двумер-

¹⁾ Это выражение есть двумерный аналог (1.2.1). См. примечание на с. 30.

ное двухстороннее фурье-преобразование, а

$$\mathcal{F}_f(\underline{\omega}) = \int_{\mathbb{R}^x} f(x) \exp(j\underline{\omega}x) d x,$$

$$f(x) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \bar{\mathcal{F}}_f(\underline{\omega}) \exp(-j\underline{\omega}x) d \underline{\omega}.$$

Связь СП и АКФ в цепочке соотношений (1.4.7) устанавливается с учетом того, что

$$\mathcal{F}_f^*(\underline{\omega}) \nleftrightarrow \check{f}(x). \quad \text{— вертикальное изображение}$$

5. Вертикальное сечение двумерной АКФ по одной из координатных осей связано одномерным фурье-преобразованием с одномерной спектральной плотностью $S_0(\omega_1)$, которую находят интегрированием двумерной СП

$$K(\Delta_1, 0) \leftrightarrow S_0(\omega_1) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega_1, \omega_2) d\omega_2. \quad (1.4.8)$$

Здесь индексы 1 и 2 взаимозаменяемы, так же как и в другом, сопряженном с (1.4.16), соотношении:

$$S(\omega_1, 0) \leftrightarrow K_0(\Delta_1) = \int_{-\infty}^{\infty} K(\Delta_1, \Delta_2) d\Delta_2.$$

Существенно, что

$$S_0(\omega_1) \neq S(\omega_1, 0), \quad K_0(\Delta_1) \neq K(\Delta_1, 0).$$

Если $\Delta_2 = C \neq 0$, то вертикальное сечение АКФ смещается от начала координат на эту величину, оставаясь параллельным оси Δ_1 . Связь полученной в сечении кривой с СП в этом случае имеет вид

$$K(\Delta_1, C) \leftrightarrow \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega_1, \omega_2) e^{j\omega_2 C} d\omega_2.$$

Здесь правая часть может оказаться комплексной, в силу чего левая не будет обладать свойством четности относительно точки $\Delta_1 = 0$, которое присуще обычным АКФ одномерных сигналов и центральным сечениям двумерных АКФ.

Асимметрия нецентральных сечений АКФ (и, следовательно, СП) характеризует степень анизотропности изображения. В корреляционных дискриминаторах, служащих для покоординатного слежения в ОЭСС, анизотропность может вредить нормальной работе, если смещение поля зрения вдоль одной из координат вызывает сдвиг максимума в сечении АКФ, по которому отслеживается другая координата. Этот фактор оказывается неучтенным, если при анализе принята гипотеза о круговой симметрии АКФ, хорошо подтверждающаяся путем расчета статистической АКФ на достаточно большой совокупности реальных изображений [62].

6. Учет возможных структурных различий между изображениями достигается введением взаимных спектрально-корреляционных характеристик двух различных финитных функций $f_i(x)$ и $f_j(x)$ и соответственно их фурье-образов $\mathcal{F}_i(\omega)$ и $\mathcal{F}_j(\omega)$. Взаимная спектральная плотность (ВСП) есть в общем случае комплексная функция частоты

$$S_{ij}(\omega) = \mathcal{F}_i^*(\omega) \mathcal{F}_j(\omega),$$

а взаимно корреляционная функция связана с ВСП двумерным преобразованием Фурье и представляет собой, как и (1.4.5), свертку изображений, первое из которых повернуто на 180° инверсией аргумента: (зеркально)

$$K_{ij}(\Delta) = f_i^\Delta \otimes f_j \rightleftharpoons S_{ij}(\omega). \quad (1.4.9)$$

Заметим, что ВКФ как свертка действительных изображений всегда будет вещественной функцией сдвига Δ , легче поддающейся непосредственному измерению, нежели ВСП, которую в силу ее комплексного характера труднее регистрировать. Этим, вероятно, можно объяснить, почему изобретатели ОЭС по крайней мере до появления быстродействующих микропроцессоров предпочитали корреляционные методы спектральным способам сравнения изображений [45, 55, 82, 108].

Среди других свойств ВКФ и ВСП отметим, что взаимная корреляция двумерных функций, аргументы которых инвертированы по отношению к (1.4.9), дает ВКФ в обращенных координатах

$$K_{ji}(\Delta) = K_{ji}(\Delta_1, \Delta_2) = f_j^\Delta \otimes f_i = K_{ij}(-\Delta_1, -\Delta_2) = K_{ij}(-\Delta).$$

Соответственно для ВСП имеем

$$S_{ji}(\omega) = S_{ij}^*(\omega).$$

• Эти формулы представляют собой обобщение аналогичных одномерных зависимостей.

Центральные сечения двумерной ВКФ (равно, как и ВСП) в общем случае подобно одномерным ВКФ несимметричны, т.е.

$$K_{ij}(\Delta) \neq K_{ij}(-\Delta).$$

Поэтому глобальный максимум ВКФ не совпадает с точкой $\Delta = 0$. Именно здесь видно, что внесение отличий в одно изображение относительно другого может вызывать смещение максимума ВКФ от начала координат и порождать ошибки в работе следящей ОЭС с КД сдвигов.

Так как результаты корреляционного сравнения зависят от фильтрующих свойств звеньев, через которые оба изображения поступают на коррелятор, напомним, как действуют линейные фильтры на двумерные сигналы.

Действие линейных фильтров на двумерные спектрально-корреляционные характеристики. Линейные фильтры можно разделить на два класса по признаку однородности. Однородным называют линейный преобразователь, для которого справедлив принцип транспозиции, означающий инвариантность отклика $f_{\text{вых}}(x)$ такой системы к аддитивному сдвигу x_0

входного воздействия $f_{\text{вх}}(x)$. Это означает выполнение при любом x_0 операторного равенства

$$\mathcal{U}^{x_0} f_{\text{вх}}(x) = \mathcal{L}_x \mathcal{U}^{x_0} f_{\text{вх}}(x),$$

где $\mathcal{L}_x$ — оператор системы; $\mathcal{U}^{x_0}$ — оператор аддитивного сдвига.

В радиотехнике и автоматике однородные линейные системы называют *стационарными* и чаще всего имеют в виду фильтры электрических сигналов как функций времени. В ОЭС роль линейных преобразователей могут играть и электрические цепи, и оптические (радиооптические) элементы — линзы, зеркала и т.п. Ни один из последних не бывает строго инвариантным к аддитивному сдвигу хотя бы потому, что всегда можно указать такой сдвиг x_0 , при котором изображение выйдет из "поля зрения" элемента. Это не мешает применять стационарные математические модели оптических систем в той же степени, в какой можно использовать понятие *установившегося режима* в анализе цепей.

Пределы применимости однородной (стационарной, инвариантной) модели оптического фильтра определяются границами области, в которой его свойства практически не зависят от сдвига относительно фильтруемого изображения.

Выход $f_{\text{вых}}(x)$ стационарного фильтра для изображения $f_{\text{вх}}(x)$ представляется в виде

$$f_{\text{вых}}(x) = (f_{\text{вх}} \otimes g)(x),$$

где $g = g(x)$ — функция рассеяния точки (ФРТ), т.е. отклик элемента на двумерную δ -функцию (на однородном фоне). В математике и математической физике такой отклик называют *функцией Грина*, а в электротехнике — *импульсной характеристикой системы*. В инвариантных фильтрах, заданных откликами $g_1(x)$ и $g_2(x)$, ВКФ входных изображений $K_{ij\text{вх}}(\underline{\Delta})$ преобразуется в выходную ВКФ $K_{ij}(\underline{\Delta})$ посредством обычной двумерной свертки

$$K_{ij}(\underline{\Delta}) = (K_{ij} \otimes \varphi_{12})(\underline{\Delta}), \quad (1.4.10)$$

где

$$\varphi_{12} = \varphi_{1,2}(\underline{\Delta}) = \check{g}_1 \otimes g_2 \rightleftharpoons \check{G}_1^* \check{G}_2, \quad (1.4.11)$$

а

$$\check{G}_{1,2} = \check{G}_{1,2}(\underline{\omega}) \rightleftharpoons g_{1,2}(x).$$

Из (1.4.11) видно, что $\varphi_{12}(\underline{\Delta})$ есть ВКФ импульсных откликов.

В области двумерных пространственных частот эквивалентом для (1.4.10) будет выражение выходной ВСП профильтрованных изображений $S_{ij}(\underline{\omega})$ через их ВСП на входе $S_{ij\text{вх}}(\underline{\omega})$ и передаточные функции фильтров $\check{G}_1(\underline{\omega})$ и $\check{G}_2(\underline{\omega})$:

$$S_{ij}(\underline{\omega}) = (S_{ij\text{вх}} \check{G}_1^* \check{G}_2)(\underline{\omega}). \quad (1.4.12)$$

Из (1.4.10) и (1.4.12) видно, что подбором фильтров можно эффективно влиять на форму ВКФ изображений. Действие фильтров сказывается тем сильнее, чем "шире" функция $\varphi_{12}(\Delta)$ по сравнению с ВКФ входных сигналов $K_{ff_{вх}}(\Delta)$, а в пределе, когда $K_{ff_{вх}}(\Delta)$ задана δ -функцией, свойства фильтров целиком определяют взаимные спектрально-корреляционные характеристики выходных изображений. Такая закономерность хорошо известна и используется в теоретической и практической радиотехнике применительно к одномерным сигналам и цепям [24, 125]. Однако в ранних исследованиях корреляционных свойств изображений ролью фильтров во входных цепях коррелятора иногда пренебрегали. Это приводило к характерной ошибке — приписыванию частотных свойств анализатора свойствам изображений в области высших пространственных частот, где происходит спад передаточных функций входных каналов.

Неоднородные линейные фильтры, для которых принцип транспозиции несправедлив, в теории цепей иногда называют параметрическими в связи с тем, что в этот класс попадают одномерные фильтры с переменными параметрами для сигналов-функций времени. При нахождении откликов неоднородных фильтров на входные воздействия математический аппарат обычной "аддитивной" свертки и сопряженный с нею аппарат классического фурье-анализа оказываются малоэффективными. Более подходящими для описания нестационарных сигналов и систем являются обобщенные свертки и обобщенный гармонический анализ, которые опираются на расширенное понятие инвариантности к обобщенным сдвигам. Глубокой разработке и приложениям адекватного математического аппарата к решению линейных задач посвящены труды советского математика Я.П. Драгана и ряда других ученых украинской математической школы [41–43]. Ю.П. Пытьев с сотрудниками, специализируясь на проблемах распознавания образов, разработали математические основы морфологического анализа изображений [101–104]. Его подход допускает использование специфичных для современной цифровой техники моделей и процедур в алгоритмах идентификации плоских изображений или для выявления различий между ними. Базирующийся на понятиях абстрактной алгебры математический аппарат морфологического анализа охватывает, в частности, как методы обычной, так и обобщенной корреляции двумерных финитных функций, представляющих собой реальные изображения.

В связи с тем что процедура обобщенной корреляции может использоваться для корреляционного измерения масштабных, ракурсных и других сложных "сдвигов", отличных от обычного сдвига-переноса изображений, в следующем параграфе эти понятия рассмотрены подробнее.

§ 1.5. Обобщенная корреляция изображений и ее применение

Групповые и обобщенные сдвиги в описаниях проективных преобразований плоских изображений. Понятие обобщенной корреляции как функции, зависящей от взаимной деформации коррелируемых изображений, удобно рассматривать, основываясь на определении *группового сдвига*, который является наиболее изученным видом обобщенного сдвига, составляющего основу современного фурье-анализа [71].

Как известно (см., например, [57]), *группой* называется такое множество математических объектов (элементов), в котором определена одна операция, ставящая в соответствие каждой паре элементов некий третий элемент того же множества по правилам, именуемым *групповыми аксиомами*.

Если обозначить элементы через $a, b, c, \dots$, а символом $\circ$ — определяющую операцию, которую часто называют *групповым умножением*, то групповые аксиомы будут выражать следующие свойства:

- 1) $(a \circ b) \circ c = a \circ (b \circ c)$ — свойство ассоциативности;
- 2) для любых a, b из некоторого множества существует такой элемент c из того же множества, что

$$a \circ c = \bar{a} \cdot b,$$

и такой элемент c' из того же множества, что

$$c' \circ a = b.$$

Как следует из аксиом, в группе существует элемент e (единица группы) такой, что

$$e \circ a = a \circ e = a,$$

и всякому элементу a соответствует единственный обратный элемент a^{-1} такой, что

$$a^{-1} \circ a = a \circ a^{-1} = e. \quad (1.5.1)$$

Группа *коммутативна*, если, кроме названных, имеет место свойство

$$a \circ b = b \circ a.$$

Типичная коммутативная (иначе — абелева) группа — множество $\mathbb{R}^1$ действительных чисел с обычным сложением в качестве групповой операции. В этом случае единицу группы называют *нулем*, символ $\circ$ заменяют на $+$, обратный элемент именуют *противоположным* ($-a$) и говорят, что такая группа действительных чисел является *аддитивной*. Аналогично определяют аддитивную группу $\mathbb{C}^1$ комплексных чисел.

Совокупность $\mathbb{R}_0^1$ всех действительных чисел, отличных от нуля, образует группу, если определить в ней групповую операцию как обычное умножение. Это множество называют *мультипликативной группой*. Единицей группы является число 1, а обратным к a — число $1/a$.

Аналогично определяют мультипликативную группу $\mathbb{C}_0^1$ комплексных чисел.

В классическом гармоническом анализе одномерных сигналов-функций используется операция аддитивного сдвига-переноса на прямой. Значения сдвига берутся из множества $\mathbb{R}^1$. Такой же анализ двумерных сигналов базируется на аддитивных сдвигах $\underline{\Delta}_a$ со значениями из множества пар действительных чисел $\mathbb{R}^2$ или изоморфного ему множества комплексных чисел $\mathbb{C}^1$.

В описании двумерного аддитивного переноса оператор сдвига $\mathcal{F}^{\underline{\Delta}}$ функции $f(x)$; заданной на плоскости с координатами точек $x = \|x_1 \ x_2\|$, определяется равенством

$$\mathcal{F}^{\underline{\Delta}} f(x) = f(x + \underline{\Delta}), \quad \text{где } \underline{\Delta} = \underline{\Delta}_a = \|\Delta_1 \ \Delta_2\|^T. \quad (1.5.2)$$

Здесь метрики пространства сдвигов $\underline{\Delta}_a$ и координат x совпадают, так как сдвиг аддитивен и равносильен переносу начала отсчета на плоскости.

¹⁾ Обозначения соответствуют принятым в [83].

В более общем случае изображение, например, регистрируемое приемником ОЭС, направленным на некоторую почти плоскую поверхность, можно представить как функцию, зависящую не только от координат точек (x_1, x_2) и их аддитивного сдвига (Δ_1, Δ_2) , но также от Δ_3 — логарифма масштаба относительно центра координат, угла поворота вокруг него Δ_4 и компонент ракурсной деформации (Δ_5, Δ_6) . В роли одного из последних параметров можно использовать приращение α угла между направлением проекции и нормалью к поверхности, участок которой находится в поле зрения системы и создает изображение, а в роли другого — угол β , задающий в плоскости изображения ось, относительно которой меняется направление проекции. Масштабный сдвиг и поворот изображений сохраняют подобие фигур и не нуждаются в иллюстрациях. Характер геометрических искажений при чисто ракурсных сдвигах виден из рис. 1.11, где показаны исходная 1 и искаженные 2 проекции границ кадра, зарегистрированного приемником изображений с квадратным полем зрения.

Для простоты построение выполнено без учета перспективы, т.е. в предположении, что угол зрения оптической системы достаточно мал, чтобы центр проекции можно было считать удаленным в бесконечность. Линии, ограничивающие искаженные кадры, показывают, куда смещаются в плоскости изображений точки, находившиеся в исходном состоянии на границах поля зрения. При сделанном предположении штриховая линия под углом β к горизонтальной границе исходного кадра является осью перспективно-аффинного преобразования плоскости. С точки зрения проективной геометрии все три кадра на рис. 1.11 аффинно эквивалентны.

Ход лучей, формирующих изображение на одной из границ поля зрения системы в плоскости, инцидентной с оптической осью приемника и перпендикулярной к поверхности, показан сплошными линиями на рис. 1.12, где условия проекции исходного кадра для рис. 1.11 определяются углом α_0 ,

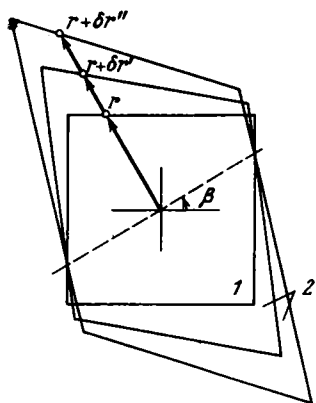


Рис. 1.11. Геометрические искажения при ракурсных сдвигах квадратного кадра без учета перспективы: 1 — исходная проекция; 2 — проекции при двух значениях ракурсного сдвига

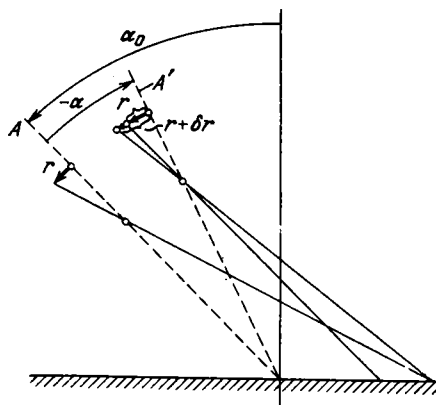


Рис. 1.12. Ход лучей от границ исходного кадра на плоскость проекции оптико-электронной системы с малым углом зрения при отсутствии (ось A) и наличии (ось A') ракурсного сдвига

а его уменьшение до $\alpha_0 - \alpha$ соответствует увеличению размера в плоскости изображений от r до $r + \delta r$.

Совместное действие на изображение всех составляющих вектора сложного сдвига

$$s = \|\Delta_1 \quad \Delta_2 \quad \Delta_3 \quad \Delta_4 \quad \Delta_5 \quad \Delta_6\|^T$$

часто можно представлять оператором сдвига $\mathcal{F}^s$, понимаемым в групповом смысле, т.е. шире, чем в (1.5.2), так, что

$$\mathcal{F}^s f(x) = f(x \odot s). \quad (1.5.3)$$

Подобное преобразование описывает, например, отображение функции f на сдвинутую, масштабированную, повернутую и деформированную систему координат плоскости (рис. 1.13). Оно сохраняет взаимно однозначное соответствие между точками и образует так называемую *аффинную группу* отображений плоскости, поскольку всякому оператору сдвига $\mathcal{F}^s$ можно противопоставить обратный ему оператор $(\mathcal{F}^s)^{-1}$, переводящий преобразованные точки на плоскости в исходные так, что

$$(\mathcal{F}^s)^{-1} \mathcal{F}^s f = \mathcal{F}^s (\mathcal{F}^s)^{-1} f = f,$$

и, следовательно, условие (1.5.1), определяющее групповую природу сдвига, выполняется. При этом единица группы есть вектор s_0 , все компоненты которого равны нулю.

Заметим, что размерности пространства сдвигов s и координат x в (1.5.3) вообще не совпадают, более того – векторное пространство сдвигов не обязательно является метрическим. Между прочим, это обстоятельство важно и для понимания обсуждавшегося в § 1.3. факта, что понятия

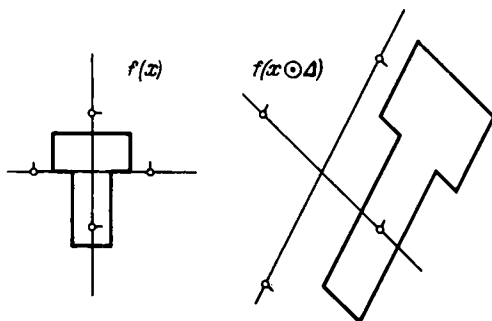


Рис. 1.13. Пример обобщенного аффинного сдвига изображения

”расстояния” и ”близости” в пространстве взаимно сдвинутых сигналов (изображений) приобретают смысл лишь после того, как назначен критерий оптимальности и таким (или иным) путем определено расстояние между точками этого пространства. При назначении метрики и критериев оптимальности в пространстве сдвигов допускается некоторая свобода выбора. Поэтому результаты многопараметрической оптимизации ОЭСС, как и всякой сложной системы, могут различаться, будучи каждый наилучшим в каком-то смысле и в рамках своих ограничений.

В практике встречаются ситуации, когда совокупность сдвигов определенного типа нельзя рассматривать как группу, поскольку не выполняется какая-либо из групповых аксиом. При этом тем не менее операторы сдвига обладают набором свойств (ассоциативность, линейность, существование

нейтрального элемента — аналога "единицы" в группе и т.д.), позволяющих построить аппарат алгебры свертки и обобщенных спектральных преобразований [42, 71, 163]. Такие операторы, следуя [163], называют *операторами обобщенного сдвига* (ООС), а соответствующие множества сдвигов — *гипергруппами* и включают в это понятие группу как частный случай. Примером ООС служат перспективно-аффинные преобразования плоскости, которые деформируют изображения при изменении ракурсных углов, т.е. направлений параллельной проекции. Эти преобразования являются элементами группы аффинных отображений, но не составляют в ней замкнутой подгруппы, ибо последовательный ракурсный сдвиг относительно двух различных осей перспективно-аффинного преобразования приводит к результату, который вообще не достигается однократным параллельным проектированием.

Аффинные преобразования составляют подгруппу в проективной группе отображений. Последние позволяют учесть перспективные искажения изображений плоскости, возникающие в реальной ОЭС при изменении точки наблюдения в пространстве, конечных значениях фокусных расстояний в оптике и дистанции между ОЭС и наблюдаемой поверхностью.

Аналитическое описание проективных преобразований несколько сложнее, чем аффинных, однако число скалярных компонент, характеризующих проективный сдвиг и смещение плоскости при постоянстве параметров оптической системы, сохраняется равным шести.

В пользу применения упрощенной аффинной модели проективных сдвигов говорит то, что в замкнутой следящей ОЭС допустимые изменения наблюдаемых изображений по необходимости малы. Эти изменения почти всегда представимы функциями, дифференцируемыми по компонентам сдвига, с якобианом, отличным от нуля. В силу этого можно утверждать, что малые проективные сдвиги всегда являются "локально аффинными", а потому первым приближением в описании характерного для ОЭС многомерного сдвига и служит именно аффинное преобразование координат.

Представления ракурсного и "масштабно-кренового" сдвигов изображений в ОЭС. Представление обобщенного сдвига зависит от выбранного способа описания координат, являющихся аргументами "сдвигаемых" функций. Например, двумерный ракурсный сдвиг можно описать матрицей R мультипликативного преобразования координатных векторов x в x' :

$$x \odot s = x' = \|x'_1 \quad x'_2\|^T = Rx = R \|x_1 \quad x_2\|^T,$$

где

$$R = \left\| \begin{array}{cc} 1 + h(\alpha, \alpha_0) \sin^2 \beta & -\frac{h(\alpha, \alpha_0)}{2} \sin 2\beta \\ -\frac{h(\alpha, \alpha_0)}{2} \sin 2\beta & 1 + h(\alpha, \alpha_0) \cos^2 \beta \end{array} \right\|; \quad (1.5.4)$$

$$h(\alpha, \alpha_0) = \operatorname{tg} \alpha_0 \sin \alpha + \cos \alpha - 1, \quad (1.5.5)$$

Обозначения углов в (1.5.4), (1.5.5) соответствуют рис. 1.11 и 1.12.

Представление масштабно-кренового сдвига (группового), соответствующего увеличению длины всех векторов x в M раз и повороту их на угол γ , задает матрица такая, что

$$x \circ s = x'' = \|x_1'' \quad x_2''\|^T = M_K x = M_K \|x_1 \quad x_2\|^T, \quad (1.5.6)$$

где

$$M_K = \exp(L) \begin{pmatrix} \cos \gamma & -\sin \gamma \\ \sin \gamma & \cos \gamma \end{pmatrix}; \quad L = \ln M. \quad (1.5.7)$$

Очевидно, (1.5.7) представляет собой изоморфизм двух групп вещественных чисел

$$M \in \mathcal{R}_0^+ \rightarrow L \in \mathcal{R}^1.$$

Результат совместного действия на координатную сетку масштабно-креновых и ракурсных сдвигов вместе со сдвигом-переносом $\underline{\Delta}_a$ можно описать одним из двух возможных способов:

$$x''' = RM_K x + \underline{\Delta}_a, \quad (1.5.8)$$

$$x''' = RM_K(x + \underline{\Delta}_a). \quad (1.5.9)$$

В обоих случаях имеют место сложные сдвиги с размерностью $m = 6$. Эффекты от преобразований (1.5.8) и (1.5.9) не совпадают, так как второе предусматривает ракурсную деформацию, масштабирование и поворот изображения в сдвинутой на $\underline{\Delta}_a$ системе координат, а первое — в исходной. Очевидно, однако, что различия результатов (1.5.8) и (1.5.9) стираются по мере уменьшения нормы вектора сдвига-переноса $|\underline{\Delta}_a|$. В силу этого при выборе простейших алгоритмов измерения отдельных компонент сложного обобщенного сдвига для следящего измерителя допустимо ценой некоторой потери точности пренебрегать всеми неизмеряемыми составляющими или их частью.

Обобщенные корреляционные функции изображений. По аналогии с АКФ и ВКФ изображений, находящимися в зависимости от обычных двумерных аддитивных сдвигов (см. § 1.4), вводятся обобщенные авто- и взаимно корреляционные функции (ОАКФ и ОВКФ) в пространстве обобщенных сдвигов с соответствующей им размерностью. Формальное определение ОВКФ (и ОАКФ) опирается на понятие сопряженного сдвига. Сопряженным ООС называют оператор $\tilde{\mathcal{F}}^s$, обеспечивающий выполнение равенства двух интегралов (в смысле Стильтьеса):

$$\int_{\mathcal{R}_x} f_1(x) \tilde{\mathcal{F}}^s f_2(x) d\mu(x) = \int_{\mathcal{R}_x} \mathcal{F}^s f_1(x) f_2(x) d\mu(x). \quad (1.5.10)$$

Обе части этого выражения представляют обобщенные скалярные произведения взаимно сдвинутых (с помощью ООС) функций f_1 и f_2 .

Сокращенная запись имеет вид

$$(f_1, \tilde{\mathcal{T}}^s f_2) = (\mathcal{T}^s f_1, f_2).$$

Входящая в (1.5.10) функция $\mu(x)$ может рассматриваться как мера на множестве $\{x\}$, инвариантная к действию ООС, такая, что

$$d\mu(x) = d(\mathcal{T}^s \mu(x)). \quad (1.5.11)$$

Представляя (1.5.11) в виде

$$d\mu(x) = \mathcal{J}(x)dx,$$

можно свести (1.5.10) к обычным интегралам (в смысле Римана), описывающим скалярное умножение взаимно сдвинутых функций f_1, f_2 с весовой (интегрирующей) функцией $\mathcal{J}$ [58].

Замена переменной $x \rightarrow z$ с использованием взаимно обратных соотношений

$$d\mu(x) = dz, \quad d\mu^{-1}(z) = dx \quad (1.5.12)$$

обеспечивает нелинейное преобразование координат, а вместе с ним — отображение обобщенного сдвига s в аддитивный сдвиг Δ_a преобразованных функций F такой, что

$$\mathcal{T}^s f(x) = f[\mu^{-1}(z + \Delta_a)] = F(z + \Delta_a),$$

$$\tilde{\mathcal{T}}^s f(x) = f[\mu^{-1}(z - \Delta_a)] = F(z - \Delta_a).$$

На языке теории множеств функция μ устанавливает изоморфизм между обобщенными сдвигами $\{s\}$ на множестве $\mathcal{M}_f$ реальных изображений f и аддитивными сдвигами $\{\Delta_a\}$ на другом множестве $\mathcal{M}_F$ изображений F , получаемых специально подобранным преобразованием координат, в которых заданы все f из $\mathcal{M}_f$. Такой подбор возможен по крайней мере при условии, что множества сдвигов $\{s\}$ и $\{\Delta_a\}$ "равномощны", другими словами, размерности сдвигов одинаковы.

Если это условие выполнено, то s и Δ_a — векторы в двумерных пространствах, а $\mathcal{J}(x)$ — квадратная матрица Якоби

$$\mathcal{J}(x) = \left\| \frac{\partial z_i}{\partial x_j} \right\|, \quad i, j \in \{1, 2\},$$

описывающая в дифференциальной форме нелинейное преобразование векторных координат $x \rightarrow z$.

Для нахождения функции μ используют условие ее инвариантности как функции меры к сдвигу (1.5.11).

В общем случае обобщенную взаимно корреляционную функцию в пространстве обобщенных сдвигов (ОВКФ) для двух детерминированных вещественных сигналов-изображений f_1 и f_2 определяют, как

$$K_{12}^0(r, s) = \int_{\mathcal{R}_x} \mathcal{T}^s f_1(x) \mathcal{T}^r f_2(x) d\mu(x). \quad (1.5.13)$$

Используя свойства ООС, можно записать (1.5.13) еще двумя способами,

а также установить связь с сопряженной ОВКФ¹⁾:

$$\begin{aligned} K_{12}^0(r, s) &= \int_{\mathcal{R}_x} \tilde{\mathcal{F}}^r \mathcal{F}^s f_1(x) f_2(x) d\mu(x) = \\ &= \int_{\mathcal{R}_x} \tilde{\mathcal{F}}^s \mathcal{F}^r f_2(x) f_1(x) d\mu(x) = K_{21}^0(s, r). \end{aligned} \quad (1.5.14)$$

Если сдвиги r и s заданы как аддитивные, то ОВКФ обращается в обычную, например, двумерную ВКФ, вычисляемую через интеграл простой свертки при $d\mu(x) = dx$:

$$\begin{aligned} K_{12}^0(r, s) &= K_{12}(r - s) = \int_{\mathcal{R}_x} f_1[x - (r - s)] f_2(x) dx = \\ &= \int_{\mathcal{R}_x} \check{f}_1[(r - s) - x] f_2(x) dx = (\check{f}_1 \odot f_2)(r - s). \end{aligned}$$

Обобщенные свертки и спектры. По аналогии с обычной сверткой в теории ООС вводят обобщенную свертку

$$(f_1 \circ f_2)(r) = \int_{\mathcal{R}_x} \tilde{\mathcal{F}}^x f_1(r) f_2(x) d\mu(x).$$

Она по определению коммутативна.

ОВКФ можно также записывать с помощью обобщенной свертки:

$$K_{12}^0(r, s) = \tilde{\mathcal{F}}^s (\check{f}_1 \circ f_2)(r),$$

где символ $\check{}$ над f_1 имеет тот же смысл, что и в § 1.4.

Обобщенная автокорреляционная функция (ОАКФ) действительных сигналов-изображений в отличие от ОВКФ обладает центральной симметрией, что прямо вытекает из (1.5.14):

$$\begin{aligned} K_f^0(r, s) &= \int_{\mathcal{R}_x} \tilde{\mathcal{F}}^r \mathcal{F}^s f(x) f(x) d\mu(x) = \\ &= \int_{\mathcal{R}_x} f(x) \tilde{\mathcal{F}}^s \mathcal{F}^r f(x) d\mu(x) = K_f^0(s, r). \end{aligned}$$

Модернизация понятия относительного сдвига влечет за собой реконструкцию спектральных представлений. Установлено, что каждый тип оператора обобщенного сдвига $\mathcal{F}^s$ связан с определенным дифференциальным оператором $\mathcal{L}_x$, которому присущ набор собственных функций $\varphi(x, \underline{\omega})$, где $\underline{\omega}$ — собственные значения. Функции φ таковы, что

$$\mathcal{L}_x \varphi = \underline{\omega} \varphi.$$

Используя эти функции в качестве ядер интегральных преобразований, можно построить систему правил для обобщенного спектрального разложения сигналов или характеристик реальных систем по таким функциям. Эти правила повторяют известные свойства обычных фурье-преобразований [40] и выводятся из соответствующих теорем при помощи

¹⁾ В более общем случае комплекснозначных функций f_1 и f_2 одна из них входит под знаки интегралов в (1.5.13) и (1.5.14) как комплексно-сопряженная.

изоморфизмов:

$$f(x+s) \rightarrow \mathcal{F}^s f(x), \quad f(x-s) \rightarrow \tilde{\mathcal{F}}^s f(x),$$

$$\frac{d}{dx} \rightarrow \mathcal{L}_x, \quad \exp(j\omega x) \rightarrow \varphi(x, \omega).$$

Указанные разложения называют еще $\mathcal{H}$ -спектрами [70] и записывают основные соотношения в виде

$$\int_{\mathcal{R}_x} f(x) \varphi^*(x, \omega) d\mu(x) = \mathcal{H}_f(\omega), \quad \int_{\mathcal{R}_\omega} \mathcal{H}_f(\omega) \varphi(x, \omega) d\nu(\omega) = f(x),$$

где $\mathcal{R}_x$, $\mathcal{R}_\omega$ — области определения подынтегральных функций; $\mu(x)$, $\nu(\omega)$ — меры на соответствующих множествах.

Формулы справедливы для всех интегрируемых с квадратом по мере $\mu(x)$ или $\nu(\omega)$ функций.

Практически важно, что функции φ инвариантны ($\mathcal{F}$ -вариантны) по отношению к оператору сдвига $\mathcal{F}^s$, порожденному дифференциальным оператором $\mathcal{L}_x$,

$$\mathcal{F}^s \varphi(x, \omega) = \varphi(s, \omega) \varphi(x, \omega),$$

так же, как инвариантны к обычному сдвигу $s = \Delta_s$ комплексные экспоненты, служащие собственными функциями оператора

$$\mathcal{F}^s e^{j\omega x} = e^{j\omega s} e^{j\omega x}.$$

Несмотря на то, что теоретический фундамент для обобщенного гармонического анализа заложен достаточно давно, области его практического применения до настоящего времени остаются ограниченными. Одна из причин кроется в том, что основное внимание чаще обращалось на задачи о детерминированных неоднородных системах с произвольным воздействием на входе. Их решение основывается на выборе простого описания системы с помощью адекватного ей оператора $\mathcal{L}_x$ [46] и последующем поиске достаточно корректного обхода трудностей, которые возникают при представлении сигналов в необычном спектральном базисе.

Отсутствие детерминированной характеристики системы, типичное для задач синтеза, делает более целесообразным согласование математических моделей сигналов с характерными для них показателями неоднородности или, как в нашем случае, непосредственно со структурой типичных сдвигов. Серия работ Я.П. Драгана [41–43], очевидно, расширит сферу применения аппарата обобщенного гармонического анализа, так как ими, в частности, создана необходимая основа для распространения результатов корреляционной теории случайных функций на $\mathcal{F}$ -вариантные стохастические процессы. Последние могут служить более точными моделями недетерминированных сигналов, нежели стационарные и так называемые квазистационарные шумы.

Обобщения теоремы отсчетов, необходимые для измерения неаддитивных сдвигов изображений. Одним из важных результатов прикладной теории $\mathcal{F}$ -вариантных процессов является обобщение теоремы отсчетов, развивающее известный подход Котельникова — Шэннона к дискретному представлению сигнала с финитным $\mathcal{H}$ -спектром. Применение обобщенной теоремы отсчетов в технике корреляционного сравнения изображений поз-

воляет производить наиболее экономную разбивку рабочей площади фотоприемника на разрешаемые элементы выбором сетки дискретных рецепторов либо разбиением сравниваемых полей в процессе дискретизации сигналов при их обработке.

Не вдаваясь в изложение вопроса в целом, остановимся на простых примерах построения одномерных и двумерных координатных сеток, с помощью которых можно производить корреляционное измерение характеристик для ОЭС сдвигов изображений.

Как известно, теоремы отсчетов устанавливают возможность дискретного представления непрерывных функций функциональными рядами, коэффициенты которых выражаются через значения разлагаемых функций. Пусть $f_{\Omega}(x)$ — детерминированная одномерная функция такая, что носитель ее $\mathcal{H}$ -спектра ограничен интервалом $[0, \Omega]$, т.е.

$$f_{\Omega}(x) = \int_0^{\Omega} \mathcal{H}(\omega) \varphi(x, \omega) d\nu(\omega).$$

Тогда по обобщенной теореме отсчетов для такой функции верно

$$f_{\Omega}(x) = \sum_{k=1}^{\infty} f_{\Omega}(x_k) a_k(x),$$

где $a_k(x)$ — некоторые обобщенные "кардинальные" функции отсчетов, характерные для данного ООС и соответствующего ему дифференциального оператора. Если этот оператор соответствует групповому сдвигу, то он — самосопряженный первого порядка, причем x_k — точки взятия отсчетов, являющиеся корнями уравнения

$$\operatorname{Im} \varphi(x, \Omega) = 0. \quad (1.5.15)$$

В случаях, когда сдвиг не является групповым, а подходит лишь под категорию обобщенного, оператор $\mathcal{L}_x$ — самосопряженный второго порядка, а точки отсчетов находятся из уравнения

$$\varphi(x, \Omega) = 0.$$

Так, в частности, одномерному аддитивному сдвигу соответствуют функции отсчетов вида

$$a_k(x) = \frac{\sin(\Omega x - k\pi)}{\Omega x - k\pi},$$

а точки взятия отсчетов можно найти, решая уравнение

$$\operatorname{Im} e^{(\epsilon + j\Omega)x} = 0.$$

Размещение точек x_k вдоль оси координат x служит решением задачи выбора наименее плотной (и реализуемой без запаса лишь теоретически) сетки отсчетов при одномерных аддитивных сдвигах.

В случае измерения двумерных аддитивных сдвигов отсчеты располагают в узлах координатной сетки, получаемой раздельной дискретизацией по двум независимым координатам. Наиболее употребительна прямоугольная декартова система (рис. 1.14, а), для которой сетку отсчетов строят, задав ограничения по частотам Ω_1 и Ω_2 во взаимно ортогональных направлениях. Возможно также использование гексагональной сетки отсчетов

(рис. 1.14, б), которые берутся в узлах косоугольной декартовой системы координат обычно с углом 60° между осями [78, 166]. Подобная топология используется при размещении чувствительных ячеек на экранах трехцветных кинескопов в телевизорах. Формально при круговом ограничении спектра гексагональная решетка экономнее квадратной примерно на 13 % с точки зрения необходимого расхода отсчетов на единицу площади изображения для обеспечения заданной детальности передачи по всему

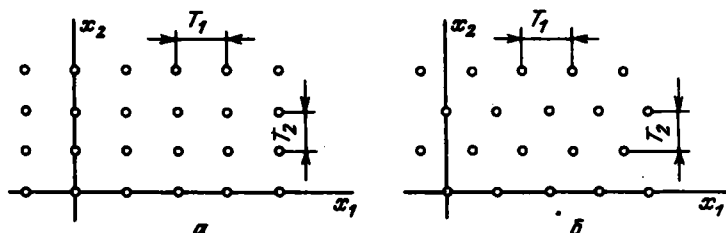


Рис. 1.14. Прямоугольная (а) и гексагональная (б) сетки отсчетов для дискретизации изображений

полю зрения [50]. Однако неудобство сопряжения таких решеток с прямоугольными развертками и специфичность алгоритмов фильтрации пока препятствует их внедрению в ОЭС, тем более, что одинаковая детальность всех участков регистрируемых изображений в таких системах не всегда нужна.

Мультипликативный масштабно-кеновый сдвиг имеет смысл представлять на плоскости комплексных координат $\{\dot{x}\}$. В этом случае

$$x \rightarrow \dot{x} = x e^{j\theta} = x_1 + jx_2, \quad M_k \rightarrow \dot{M}, \quad x \odot s \rightarrow \dot{M}\dot{x},$$

где $\dot{M} = M \exp(j\gamma) = \exp(L + j\gamma)$.

Условие инвариантности меры к сдвигу (1.5.11) удобно представить в виде

$$\mu'(x \odot s) = \frac{d(x \odot s)}{dx} \mu'(x),$$

откуда в нашем случае следует

$$\mu'(\dot{M}\dot{x}) = \dot{M}\mu'(\dot{x}).$$

Этому уравнению удовлетворяет комплексная функция

$$\mu(\dot{x}) = \ln \dot{x} = \ln x + j\theta = \dot{z} = z_1 + jz_2, \quad (1.5.16)$$

которая, согласно (1.5.12), задает отображение $\dot{x} \rightarrow \dot{z}$ такое, что мультипликативный сдвиг $\dot{M}$ преобразуется в аддитивный $\dot{\Delta}$:

$$\ln \dot{M} = L + j\gamma = \dot{\Delta}.$$

Функция (1.5.16) определяет взаимно однозначное соответствие между координатами $\{\dot{x}\}$ исходного плоского изображения $f(\dot{x})$ и его конформного отображения $F(\dot{z})$ на поверхность цилиндра во всех точках, кроме малой окрестности точки $\dot{x} = 0$.

Значения корреляционного функционала, зависящего от масштабнокренового сдвига, можно вычислять с помощью любого из двух поверхностных интегралов — по области $\mathcal{R}_x$ или $\mathcal{R}_z$. Для АКФ они имеют вид

$$K_f^0(M) = K_f^0(L, \gamma) = \int_{\mathcal{R}_x} f(\dot{x}) f(M\dot{x}) \mathcal{I}(\dot{x}) d\dot{x} = \\ = \int_{\mathcal{R}_x} f(\dot{x}) f(\dot{x} e^{-L-j\gamma}) \dot{x}^{-1} d\dot{x} = \int_{\mathcal{R}_z} F(\dot{z}) F(\dot{z} - \dot{\Delta}) d\dot{z} = K_F(\dot{\Delta}).$$

Переход от исходных координат к преобразованным иллюстрируется на рис. 1.15, который демонстрирует взаимное соответствие равномерной

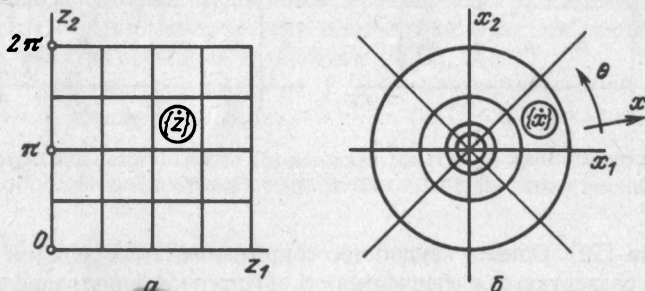


Рис. 1.15. Декартовы координаты (а), отображающие логарифмическую полярную координатную сетку (б) исходного изображения

декартовой сетки в координатах $\{z\}$ и логарифмической полярной сетки в координатах $\{x\}$. Сетки могут быть построены на основе уравнения (1.5.15), записанного относительно собственной функции дифференциального оператора, адекватного мультипликативному или аддитивному сдвигу в радиальном направлении.

Эти операторы и соответствующие им уравнения таковы:

$$\mathcal{L}_x = x \frac{d}{dx}, \quad \text{Im} \varphi(x, \Omega_x) = \text{Im}(x^{\epsilon_x + j\Omega_x}) = \text{Im}(e^{\ln x})(e^{\epsilon_x + j\Omega_x}) = 0,$$

$$\mathcal{L}_z = \frac{d}{dz}, \quad \text{Im} \varphi(z, \Omega_z) = \text{Im} e^{(\epsilon_z + j\Omega_z)z} = 0.$$

После несложных упрощений оба уравнения относительно точек отсчета x_k и z_k дают соответственно

$$\Omega_x \ln x_k = k\pi, \quad k \in \mathcal{N}, \\ \Omega_z z_k = k\pi, \quad k \in \mathcal{N}, \quad (1.5.17)$$

где $\mathcal{N}$ — множество всех целых чисел.

Очевидно также, что $\Omega_x = \Omega_z$.

Разбиение плоскости $\{z\}$ в направлении углов $\theta = z_2$ производится, как и по z_1 , с равномерным шагом, который можно связать с ограничивающим параметром Ω_θ уравнением, аналогичным (1.5.17):

$$\Omega_\theta \theta_k = k\pi,$$

где θ_k — значения полярных углов в узлах сетки отсчетов.

Поскольку значения углов θ_k образуют циклическую группу, целесообразно выбирать Ω_θ из ряда целых чисел. Тогда значения k образуют конечное множество

$$k \in \{1, 2, \dots, 2\Omega_\theta\}.$$

Пример обобщенной АКФ изображения. Пользуясь отображением фигур из пространства координат исходных изображений $\{\dot{x}\}$ с мультипликативными сдвигами в преобразованное пространство $\{\dot{z}\}$, где они становятся аддитивными, можно обычными приемами построить, например, ОАКФ двумерной функции $f(\dot{x})$ при сдвигах по масштабу и крену (L, γ) . Это сделано на рис. 1.16, б для изображения, показанного в сдвинутом и несдвинутом состояниях на рис. 1.16, а.

Для практики интересен вопрос о том, какое число независимых отсчетов m_z можно сделать на плоскости $\{\dot{z}\}$ в радиальном направлении z .

Если задать число N разрешаемых элементов на радиус (т.е. на половину ширины) поля зрения ОЭС, то величину m_z можно определить из очевидного уравнения

$$m_z = E\left(\frac{\ln N}{\ln 2}\right),$$

где оператор $E(\cdot)$ выделяет целую часть заключенного в скобках числа. Например, при $N = 256$, что является типичным для современных ТВ датчиков, имеем $m_z = 8$. Этот результат позволяет косвенно судить о величине $\delta m = m_z^{-1}$ разрешающей способности системы корреляционного измерения

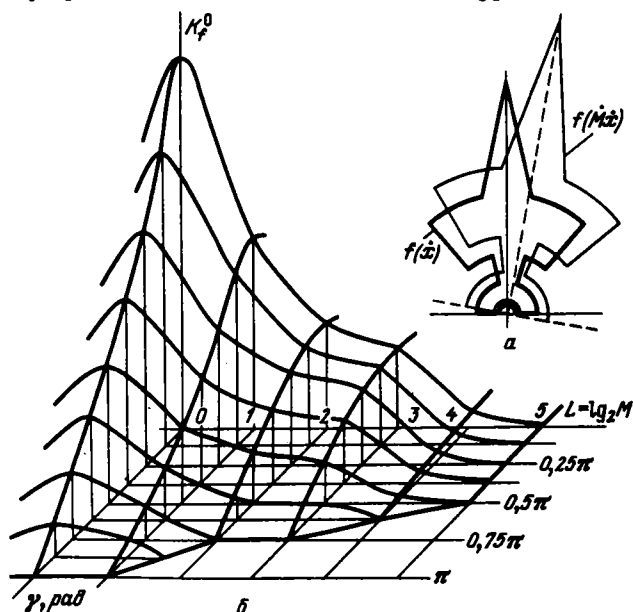


Рис. 1.16. Обобщенный масштабно-креновый сдвиг изображения (а) и его обобщенная автокорреляционная функция (б) над одним из квадрантов плоскости сдвигов $\{L, \gamma\}$

ния масштабного сдвига изображений, работающей с таким датчиком. В нашем примере $\delta M = 12,5\%$. Аналогичным образом, если X — полная ширина поля зрения ОЭС, то $\delta x = X/2N$ служит косвенной мерой разрешения при корреляционном измерении обычного аддитивного сдвига.

В обоих случаях не следует по показателям разрешения непосредственно судить об ожидаемой точности измерений, ибо ошибки при приеме сигналов с большой базой, характерной, например, для стандартного телевидения, оказываются на много порядков ниже, чем эти формальные показатели, если только помехи не сильно превышают полезные входные воздействия. В дальнейшем мы остановимся на этом вопросе подробнее.

§ 1.6. Особенности аналоговых корреляционных дискриминаторов

Преобразования координат и операторов линейной фильтрации при построчной развертке изображений в ОЭС. Процесс последовательного преобразования двумерных сигналов в одномерные путем сканирования многократно описан в литературе по телевидению, радиолокации и т.п. [111, 142]. Он является основным для различных систем обзора пространства, обеспечивающих превращение изображений в непрерывный аналоговый электрический сигнал. В связи с этим рассмотрим ряд особенностей реализации корреляционных дискриминаторов сдвига изображений средствами аналоговой техники.

Предварительно остановимся на некоторых деталях процесса преобразования входного двумерного сигнала $f(x)$ в колебание $f(t)$. В реальных системах такой процесс практически линеен. Его можно компактно представить уравнением типа обобщенной свертки исходного двумерного поля $f_0(x)$ со сканирующей по закону $x_p(t)$ апертурной функцией $g_a(x, x_p)$. В общем случае зависимость g_a от двух аргументов учитывает изменение апертурных свойств приемника ОЭС по области сканирования (например, из-за aberrаций в сканирующей оптике или из-за вариаций распределения электронов в считывающем пучке телевизионной трубки — видикона при разных отклоняющих полях). В нашем рассмотрении такой учет не обязателен. Тогда апертурная функция может считаться инвариантной к сдвигу, а двумерная свертка записываться, как обычная:

$$f(t) = \int_{\mathcal{A}_x} f_0(k) g_a[x_p(t) - k] dk = (f_0 \otimes g_a)[x_p(t)], \quad (1.6.1)$$

где $\mathcal{A}_x$ характеризует поле зрения системы.

Чтобы выразить результат преобразования в идеальном случае, достаточно представить сканирующую апертуру как неискажающую, т.е. в виде двумерной δ -функции. Тогда из (1.6.1) получим

$$f(t) = k f_0[x_p(t)],$$

где k — размерный коэффициент.

Вообще сканирующие преобразователи двумерных сигналов в одномерные могут вносить и нелинейные искажения, среди которых наиболее опасны те, что связаны с инерционными нелинейностями. Характерны в этом отношении инерционные явления в мишенях видиконов [30]. Они проявляются в "запоминании" предыдущих кадров изображения на время считывания ряда последующих. В результате возникает эффект

наложения видеосигналов от нескольких кадров, причем степень искажения текущего зависит от уровня освещенности мишени. При наличии такой инерционности движение элементов изображения в кадре зрительно воспринимается как "смазывание".

Указанные искажения в ТВ автоматах, измеряющих сдвиг изображений, в принципе могут быть причиной дополнительных ошибок, если объекты наблюдения движутся относительно ОЭС. Однако инерционность современных ТВ трубок столь мала, что вызываемые ею ошибки в худшем случае сопоставимы по величине с ошибками из-за покадровой дискретизации процесса измерения во времени, поэтому названный фактор почти всегда допустимо игнорировать.

Линейная фильтрация изображений, необходимая для реализации какого-либо алгоритма измерения сдвига, в принципе может осуществляться как до, так и после их преобразования развертывающим устройством. Поскольку электрические фильтры для многих ОЭС оказываются удобнее оптических, имеет смысл отметить некоторые особенности пересчета характеристик фильтров из временной области в пространственную при использовании наиболее употребительной строчной развертки.

Рассмотрим закон сканирования, идеализированный на участках "обратного хода" и по направлению кадровой составляющей, совместив направление строк с осью $\{x_1\}$, а кадров — с осью $\{x_2\}$:

$$x_p(t) = [V_c t - X_{M1} \sum_{i=1}^{N_c} \sigma(t - iT_c), \\ \frac{X_{M2}}{N_c} \sum_{i=1}^{N_c} \sigma(t - iT_c)], \quad t \in \{0, T_k\}, \quad (1.6.2)$$

где V_c — скорость строчной развертки; t — время; X_{M1} , X_{M2} — размах строчной и кадровой разверток, соответствующий максимальным размерам раstra в плоскости изображения; T_c — длительность строки; i — номер строки; T_k — длительность кадра (полукадра при "чересстрочной" развертке); $\sigma(t)$ — единичная ступенчатая функция (0 при $t < 0$, 1 при $t \geq 0$).

Пусть видеосигнал $f(t)$ преобразуется аналоговым фильтром с постоянными параметрами. Обозначим через $\mathcal{L}_t$ оператор линейной фильтрации как функцию дифференциального оператора d/dt . Поскольку входной сигнал f является двумерным, оператор его дифференцирования по времени при произвольной развертке $x = \|x_1(t) \ x_2(t)\|^T$ можно представить в виде

$$\frac{d}{dt} = \frac{\partial}{\partial x_1} \frac{dx_1}{dt} + \frac{\partial}{\partial x_2} \frac{dx_2}{dt}.$$

Для отдельной строки бесконечной длительности вдоль оси x имеет место равенство

$$\frac{d}{dt} = \frac{\partial}{\partial x_1} \frac{dx_1}{dt} = V_c \frac{\partial}{\partial x}, \quad (1.6.3)$$

поэтому

$$\mathcal{L}_t = \mathcal{L}\left(\frac{d}{dt}\right) = \mathcal{L}_t\left(V_c \frac{\partial}{\partial x}\right) = \mathcal{L}_{x_1}\left(\frac{\partial}{\partial x}\right), \quad (1.6.4)$$

т.е. фильтры во временной и пространственной областях эквивалентны.

О технических помехах из-за смены направлений развертки при разложении изображений. Для N_c строк конечной длительности равенство (1.6.4) строго не выполняется, потому что на выходе фильтра кроме результатов действия оператора $\mathcal{L}_{x1}$ появляются дополнительные слагаемые – отклики на скачки функции f при мгновенных "перебросах" сканирующей апертуры от конца каждой строки к началу следующей. Вследствие этого вместо (1.6.3) имеем

$$\frac{d}{dt} = V_c \frac{\partial}{\partial x_1} - X_{M1} \left[\sum_{i=1}^{N_c} \delta(t - iT_c) \right] \frac{\partial}{\partial x_1} + \frac{X_{M2}}{N_c} \left[\sum_{i=1}^{N_c} \delta(t - iT_c) \right] \frac{\partial}{\partial x_2}, \quad t \in \{0, T_K\},$$

$$\mathcal{L}_t = \mathcal{L}_{x1} + \mathcal{L}_\delta,$$

где

$$\mathcal{L}_\delta = \mathcal{L} \left[\left(\frac{X_{M2}}{N_c} \frac{\partial}{\partial x_2} - X_{M1} \frac{\partial}{\partial x_1} \right) \sum_{i=1}^{N_c} \delta(t - iT_c) \right].$$

Очевидно, при $|\mathcal{L}_\delta f| \ll |\mathcal{L}_{x1} f|$ операторы $\mathcal{L}_t$ и $\mathcal{L}_{x1}$ почти эквивалентны, т.е. с некоторым приближением верно (1.6.4). Количественной мерой ошибки приближения в условиях, когда два подвергнутых фильтрации процесса $\mathcal{L}_t f_1$ и $\mathcal{L}_t f_2$ сравниваются путем взаимной корреляции в КД, служит относительная величина разности скалярных произведений $(\mathcal{L}_{x1} f_1, \mathcal{L}_{x1} f_2)$ и $(\mathcal{L}_t f_1, \mathcal{L}_t f_2)$.

При стандартных или близких к ним ТВ развертках длительность обратных ходов в отличие от принятых в (1.6.2) конечна, однако достаточно резкие скачки в сигнале $f(t)$ на переходах от строки к строке неизбежны как из-за свойств телевизионных трубок, так и из-за обязательной для широкополосного телевидения процедуры "привязки" уровня видеосигнала к определенному потенциалу с помощью кратковременных стробирующих импульсов [142].

Априорный расчет величины технических помех в КД от переходных процессов, связанных с развертками и действием служебных импульсов, представляет трудноразрешимую задачу. При практической реализации аналоговых корреляционных дискриминаторов для ТВ изображений используют такие средства ослабления технических помех, как:

- привязку видеосигналов между строками развертки к "уровню серого" (вместо обычного "уровня черного"), что уменьшает скачки на междустрочных интервалах;
- последовательное стробирование начальных и конечных участков строчных видеосигналов в разделенных фильтрами усилительных каскадах видеотракта с целью исключения основной части переходных процессов;
- применение различных калибровочных схем и операций, обеспечивающих ручное или автоматическое "обнуление" выходов КД при подаче тождественных входных сигналов.

Сравнительный анализ вариантов одномерного аналогового КД. Описание корреляционных дискриминаторов, в которых сравниваются одномерные сигналы, получаемые линейной разверткой, неоднократно приводились под разными названиями в технической литературе [2, 13, 17, 23, 26, 36, 45, 55, 58, 59, 85, 87, 116, 122, 128, 160, 172–174]. Такие устройства, как правило, бывают выполнены в расчете на получение оценки одной компоненты сдвига, совпадающей с направлением строк. При анализе обычно предполагают, что в течение одного поля (полукадра в ТВ стандарте) развертки оба сравниваемых изображения неподвижны. Система в целом строится так, чтобы эталонное изображение было подобно текущему, т.е. формировалось путем построчного разложения.

Варианты КД рассматриваются почти всегда как часть той или иной корреляционно-экстремальной системы, но строятся без обращения к ал-

горитму оптимального измерения многомерного сдвига. В связи с этим напомним, что при непосредственном синтезе обобщенного дискриминатора в соответствии с результатами § 1.3 каждый его канал можно реализовать, исходя из задачи измерения соответствующей компоненты сдвига и считая остальные составляющие нулевыми.

Объединение нескольких одномерных КД для разных составляющих в общую схему, содержащую цепи нормировки, может с точностью до компенсации перекрестных связей соответствовать структуре обобщенного оптимизируемого алгоритма в рамках ограничений, принятых в § 1.3.

Одномерный аналоговый КД для измерения обычного сдвига сигналов, получаемых линейным построчным разложением изображений, может быть описан упрощенной структурной схемой, приведенной на рис. 1.17, где показаны блоки, образующие ненормированную оценку ошибки слежения в направлении строк развертки

$$\bar{\Delta}_t = \Delta_1 / V_c.$$

Для описания принципа действия КД достаточно рассмотреть сигналы без помех, поэтому можно считать, что текущее и эталонное изображения при отсутствии сдвига в данной схеме тождественны, а их источник является общим. Звенья 1 и 2 в составе КД включают необходимые усилители и линейные фильтры, звено 3 — коррелятор, содержащий схему перемножения и интегратор. Операции нормировки и компенсации являются функциями звена 4, выходной сигнал которого поступает на регулируемый орган, обеспечивающий слежение за сдвигом, вызванным внешним возмущением.

Наиболее известной является схема ненормированного одномерного "разностного" КД, где выходной сигнал $U(\Delta_t)$ пропорционален разности оценочных значений ВКФ $\hat{K}_{вх}$ входного процесса $f_1(t)$ и процессов

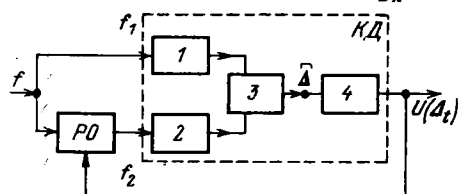


Рис. 1.17. Упрощенная структурная схема аналогового одномерного корреляционного дискриминатора для оптико-электронной следящей системы

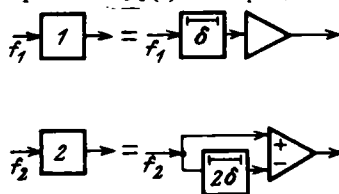


Рис. 1.18. Вариант состава блоков обработки сигналов в каналах линейного разностного корреляционного дискриминатора

$f_2(t + \Delta_t + \delta_t), f_2(t + \Delta_t - \delta_t)$. При этом δ — принудительный сдвиг, обеспечиваемый специально введенными в систему элементами перед коррелятором в соответствии с рис. 1.18. Тогда

$$U(\Delta_t) \equiv \hat{K}_{вх}(\Delta_t + \delta_t) - \hat{K}_{вх}(\Delta_t - \delta_t). \quad (1.6.5)$$

Главный недостаток схемы заключается в сильной зависимости формы дискриминационной характеристики и, главное, ее крутизны при малых Δ_t от спектрального состава сигнала. Для ослабления этой зависимости можно дополнительно регулировать сдвиги $\pm \delta_t$, однако автоматизировать эту процедуру без серьезных усложнений схемы не удастся.

В телевизионной системе алгоритм (1.6.5) нередко реализуют путем поочередного сдвига раstra фотоприемника на $\pm \delta$ в двух следующих друг за другом циклах развертки. В этом случае некоторое упрощение технического решения покупается ценой двукратного снижения быстродействия дискриминатора.

Следующим шагом в развитии разностной схемы является КД дифференциального типа, в котором выполняется предельный вариант алгоритма (1.6.5)

$$U(\Delta_t) \equiv \lim_{\delta_t \rightarrow 0} \frac{\hat{K}_{\text{вх}}(\Delta_t + \delta_t) - \hat{K}_{\text{вх}}(\Delta_t - \delta_t)}{\delta_t} = \frac{d}{d\Delta_t} \hat{K}_{\text{вх}}(\Delta_t),$$

т.е. выход пропорционален производной от оценки корреляционной функции по ее аргументу и при отсутствии мешающих факторов является нечетной функцией сдвига. Возможности дифференциальных автоматических коррелометров, отыскивающих максимум взаимной корреляции связанных процессов, использованы в [53] и изучались в [52, 132, 133].

Наиболее привлекательная особенность дифференциального алгоритма состоит в том, что для получения производной от ВКФ двух видеосигналов достаточно продифференцировать один из них. Чтобы сделать это, необходимо только включить соответствующую цепь в один из каналов перед коррелятором по схеме рис. 1.19. Она экономнее предыдущей, так как не содержит элементов, которые были бы необходимы для параллельного (или последовательного) вычисления значений ВКФ при двух разных дополнительных сдвигах.

Другим вариантом развития идеи дифференциального дискриминатора можно считать предложенный Гванеллой [167] КД с фильтром Гильберта в одном из каналов перед коррелятором вместо дифференциатора (рис. 1.20). Такой фильтр, как известно, обеспечивает фазовый сдвиг

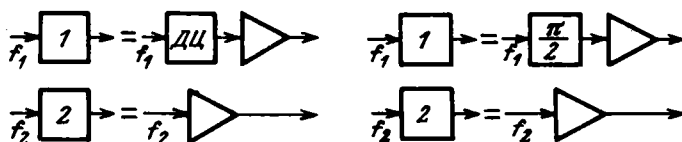


Рис. 1.19. Состав блоков обработки сигналов в каналах линейного дифференциального корреляционного дискриминатора (ДЦ – дифференцирующая цепь)

Рис. 1.20. Состав блоков обработки сигналов в каналах линейного корреляционного дискриминатора с фильтром Гильберта

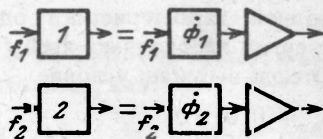
всех составляющих частотного спектра сигнала на $\pi/2$ без изменения амплитудных соотношений между ними, т.е. является идеальным фазовращателем с неограниченной полосой пропускания [24]. Результатом поворота фаз всех компонент одного из коррелируемых сигналов является (при отсутствии мешающих факторов) превращение функции корреляции в нечетную зависимость от сдвига.

Общий недостаток обеих дифференциальных схем КД в том, что ни идеальное дифференцирование, ни идеальный поворот фаз в неограниченной полосе физически неосуществимы. Для практики неудобна и прямая зави-

симость крутизны характеристики дискриминатора от распределения энергии сигналов по частотному спектру.

Корреляционный дискриминатор с ортогонализацией сигналов как универсальная реализация одномерного алгоритма оптимального оценивания. Обобщенной формой дифференциального дискриминатора, по-видимому, надо считать КД со взаимно ортогональными фильтрами для сравниваемых сигналов в обоих каналах (рис. 1.21).

Рис. 1.21. Обобщенная структура блоков обработки в каналах линейного корреляционного дискриминатора с ортогонализацией сигналов



Передаточные функции линейных фильтров $\dot{\Phi}_1$ и $\dot{\Phi}_2$ в каналах названного дискриминатора выбираются неодинаковыми, но такими, чтобы их импульсные характеристики были ортогональными и, как следствие, выход КД представлял нечетную функцию сдвига Δ_t . Это условие непосредственно вытекает из обсуждения обобщенного алгоритма ненормированной оценки сдвига (1.3.12) и преобразований, подобных (1.4.10). Оно определяет свойства каналов КД не единственным образом, в чем можно убедиться, проделав следующие выкладки.

Будем считать линейные фильтры в двух каналах произвольными и нестационарными (неоднородными) и обозначим их импульсные характеристики через $g_1(t, \tau)$ и $g_2(t, \tau)$. Некоторая непредусмотренная неинвариантность этих откликов к сдвигу может, в частности, возникнуть из-за стробирования видеосигналов в трактах обработки перед коррелятором. Без большого ущерба для общности описания можно допустить, что функции g_1 и g_2 относятся к одному классу вариантности [42]. Иными словами, зависимость названных функций от временного сдвига надо представлять как результат действия одного и того же оператора обобщенного сдвига $\mathcal{T}^\tau$ на функцию $g_{1,2}(t, 0)$:

$$\mathcal{T}^\tau g_{1,2}(t, 0) = g_{1,2}(t, \tau). \quad (1.6.6)$$

Пользуясь представлениями, рассмотренными в § 1.5, можно записывать выход КД обобщенного сдвига как ОВКФ сравниваемых сигналов, подвергнутых фильтрации перед их подачи на коррелятор. Для КД обычного сдвига результаты фильтрации в случае нестационарных фильтров с откликами типа (1.6.6) можно вычислять через обобщенные свертки, а выход, пропорциональный ОВКФ, как обычную ВКФ — через простую свертку

$$U(\Delta_t) = [(f \circ g_1) \otimes (y \circ g_2)].$$

Если сверткам внутри круглых скобок придан одинаковый смысл, это выражение можно представить в виде

$$U(\Delta_t) = \bar{K}_{\Sigma\Sigma} \otimes W_{12}, \quad (1.6.7)$$

где функцию

$$W_{12} = W_{12}(\tau, 0) = g_1 \circ g_2 = \int_{\mathcal{R}_t} \check{g}_1(t, 0) \tilde{\mathcal{T}}^\tau g_2(t, 0) dt$$

можно назвать ОБКФ импульсных реакций нестационарных фильтров, а $K_{\text{вх}}$ — ВКФ сравниваемых сигналов до их фильтрации. Здесь $\mathcal{R}_t$ — отображение $\mathcal{R}_x$ в $\{t\}$

Если f отличается от u только обычным сдвигом Δ_t , а помехи и краевые эффекты отсутствуют (например, при гладком окружающем фоне), то $K_{\text{вх}}$ обращается в АКФ сигнала $f(t)$, которая является по определению четной функцией сдвига. Следовательно, для обеспечения нечетности выходной характеристики одномерного КД согласно правилам алгебры свертки необходимо, чтобы $W(\tau, 0)$ была нечетной функцией аргумента τ . Отсюда вытекает условие

$$W(0, 0) = 0,$$

что равносильно требованию ортогональности импульсных характеристик каналов коррелятора:

$$\int_{\mathcal{R}_t} \check{g}_1(t, 0) g_2(t, 0) dt = 0. \quad (1.6.8)$$

При расширенном определении скалярного произведения функций в смысле (1.5.10) можно заменить (1.6.8) интегралом по мере $\mu(t)$.

Очевидно, что множество пар взаимно ортогональных функций g_1 и g_2 несчетно.

Остановимся на случае стационарных фильтров в каналах одномерного КД обычного сдвига. Их комплексные частотные характеристики имеют вид

$$\dot{\Phi}_{1,2} = \Phi_{1,2}(\omega) e^{j\vartheta_{1,2}(\omega)},$$

где $\Phi_{1,2}(\omega)$ — амплитудно-частотные характеристики фильтров (АЧХ); $\vartheta_{1,2}(\omega)$ — фазочастотные характеристики (ФЧХ).

Обозначим через $S_{\text{вх}}(\omega)$ спектральную плотность финитного сигнала f , понимая последний как вполне детерминированную реализацию из некоторого множества возможных реализаций. При этом, как и ранее, $f \leftrightarrow \check{G}, \check{f} \leftrightarrow \check{G}^*, S_{\text{вх}} = \check{G}^* \check{G}$. Тогда при отсутствии краевых эффектов и помех оценка сдвига становится точной и выход КД можно представлять в зависимости от сдвига Δ_t как ВКФ процессов, возникающих после фильтров, через обратное фурье-преобразование произведения спектров выходных колебаний:

$$U(\Delta_t) = \frac{1}{\pi} \int_0^\infty S_{\text{вх}}(\omega) \Phi_1(\omega) \Phi_2(\omega) \cos[\omega \Delta_t + \vartheta_1(\omega) - \vartheta_2(\omega)] d\omega. \quad (1.6.9)$$

Опустив аргументы у частотных функций, проведя замену $\vartheta = \vartheta_1 - \vartheta_2$ и несложные преобразования, упростим запись этой дискриминационной характеристики (ДХ):

$$U(\Delta_t) = \frac{1}{\pi} \left[\int_0^\infty S_{\text{вх}} \Phi_1 \Phi_2 \cos \omega \Delta_t \cos \vartheta d\omega + \int_0^\infty S_{\text{вх}} \Phi_1 \Phi_2 \sin \omega \Delta_t \sin \vartheta d\omega \right]. \quad (1.6.10)$$

Отсюда видно, что для выполнения условия нечетности ДХ в корреляционном дискриминаторе независимо от статистики и структуры входного сигнала достаточно соблюсти равенство нулю первого слагаемого в правой части, а это достигается, если

$$\vartheta = \vartheta(\omega) = \frac{\pi}{2} + n\pi, \quad n \in \mathcal{N}, \quad (1.6.11)$$

причем нечетность обеспечивается при любых АЧХ $\Phi_1(\omega)$ и $\Phi_2(\omega)$, а выходная характеристика с точностью до знака находится, как

$$U(\Delta_f) = \frac{1}{\pi} \int_0^\infty S_{\text{вх}}(\omega) \Phi_1(\omega) \Phi_2(\omega) \sin \omega \Delta_f d\omega. \quad (1.6.12)$$

Напомним, что при $\Delta_f = 0$ краевой эффект не проявляется, если нет сдвигов по другим компонентам (см. § 1.4), и, следовательно, единственным условием "строгого" нуля ДХ при выполнении (1.6.8) или вытекающего из него (1.6.11) является отсутствие помех.

Преобразовав по Фурье обе части (1.6.9), получим взаимный спектр энергии коррелируемых сигналов

$$K \leftrightarrow \dot{S} = S_{\text{вх}} \dot{\Phi}_1^* \dot{\Phi}_2,$$

где в тех же условиях, что и для (1.6.9), при выполнении (1.6.11), очевидно, имеем

$$\text{Re}(\dot{\Phi}_1^* \dot{\Phi}_2) = 0.$$

Тогда легко прийти к результату, идентичному с (1.6.8):

$$\dot{\Phi}_1^* \dot{\Phi}_2 \leftrightarrow \check{g}_1 \otimes g_2 = \varphi_{12}(\Delta_f) \equiv \varphi_{12}(-\Delta_f),$$

а отсюда $(g_1, g_2) = 0$.

Свобода выбора среди всех возможных пар взаимно ортогональных фильтров позволяет найти такую, которая обеспечит выполнение условия оптимизации КД, минимизирующего ошибку измерения в смысле метрики, назначаемой выбором матрицы Q в (1.3.3). На практике параметры фильтров определяют, сообразуясь не только с точностью оценки малых сдвигов, но учитывая требования к ширине ДХ, а также простоту конструктивного исполнения КД.

Здесь уместно заметить, что, реализуя фильтры в каналах КД с помощью цепей минимальной фазы [15], следует учитывать ограничения, вытекающие из существования однозначной связи между АЧХ и ФЧХ. Такая связь выражается дополнительным уравнением, при выводе которого используется (1.6.11):

$$\frac{2}{\pi} \int_0^\infty \frac{\omega}{\Omega^2 - \omega^2} \ln \frac{\Phi_2(\Omega)}{\Phi_1(\Omega)} d\Omega = \frac{\pi}{2} + n\pi, \quad n \in \mathcal{N}.$$

Принадлежность КД с дифференциатором или с фильтром Гильберта к более общему типу линейных дискриминаторов с ортогональной фильтрацией очевидна. Убедимся, что и разностный КД относится к тому же типу. Пусть $\Phi(\omega)$ есть АЧХ каждого из двух одинаковых каналов разностного КД. Примем условие отсутствия помех и краевых эффектов и, преоб-

разовав (1.6.5) с помощью (1.6.9), получим ДХ

$$\begin{aligned} U(\Delta_f) &= \frac{1}{\pi} \int_0^{\infty} S_{\text{вх}}(\omega) \Phi^2(\omega) [\cos \omega(\Delta_f + \delta_f) - \cos \omega(\Delta_f - \delta_f)] d\omega = \\ &= -\frac{2}{\pi} \int_0^{\infty} S_{\text{вх}}(\omega) \Phi^2(\omega) \sin \omega \delta_f \sin \omega \Delta_f d\omega, \end{aligned}$$

что с точностью до констант совпадает с (1.6.12), если

$$\Phi_1(\omega) \Phi_2(\omega) = \Phi^2(\omega) \sin \omega \delta_f. \quad (1.6.13)$$

Другими словами, процедуры относительного сдвига сигналов в разностном КД можно исключить путем введения в каналы линейных фильтров, произведение АЧХ для которых определяется формулой (1.6.13), а разность ФЧХ — формулой (1.6.11).

Практический выбор характеристик линейных трактов перед коррелятором зависит от критериев оценки пригодности реализуемых ДХ по форме, а также по стабильности и разбросу их основных параметров. Первым, хотя и не единственным, показателем качества КД служит стабильность нуля его выходной характеристики.

Влияние неидеальности ФЧХ каналов КД на его характеристику. Соотношения (1.6.9), (1.6.10) позволяют оценить, как влияют на ДХ отличия ФЧХ каналов дискриминатора от расчетных из-за действия разбросов и дестабилизирующих факторов.

Положим в (1.6.11) для определенности $n = 0$ и будем считать, что разность ФЧХ каналов КД отличается от требуемой на величину $\epsilon(\omega)$ так, что

$$\vartheta(\omega) = \frac{\pi}{2} - \epsilon(\omega), \quad (1.6.14)$$

причем для любого ω из интересующей области частот выполняется условие $\epsilon(\omega) \ll \pi/2$. Учитывая его при подстановке (1.6.14) в (1.6.10) и удерживая члены не выше второго порядка малости, получаем ДХ в виде

$$U(\Delta_f) \approx \frac{1}{\pi} \int_0^{\infty} S_{\text{вх}} \Phi_1 \Phi_2 \sin \omega \Delta_f d\omega + \int_0^{\infty} S_{\text{вх}} \Phi_1 \Phi_2 \epsilon \cos \omega \Delta_f d\omega, \quad (1.6.15)$$

где опущено обозначение частоты как аргумента всех подынтегральных функций, кроме тригонометрических.

Второе слагаемое справа есть четная функция сдвига, описывающая искажение "идеальной" нечетной ДХ, представленной первым слагаемым. Обозначив каждое из них соответственно через U_n и U_{ϵ} , можем изобразить графически оба слагаемых и сумму, как показано на рис. 1.22.

Степень искажения характеристики КД из-за нарушения фазовых соотношений удобно оценивать величиной Δ_{ϵ} сдвига нуля ДХ. Этот сдвиг, пользуясь его малостью, можно рассчитать приближенно, как видно из рис. 1.22, зная величину искажающего слагаемого $U_{\epsilon}(0)$ и крутизну ДХ в нулевой точке $U'(0)$:

$$\Delta_{\epsilon} \approx \frac{U_{\epsilon}(0)}{U'(0)} \cong \frac{\int_0^{\infty} S_{\text{вх}} \Phi_1 \Phi_2 \epsilon \cos \omega \Delta_f d\omega}{\int_0^{\infty} S_{\text{вх}} \Phi_1 \Phi_2 \omega d\omega}. \quad (1.6.16)$$

Здесь значение крутизны найдено путем дифференцирования подынтегральной функции только в первом слагаемом (1.6.15) по параметру Δ_t , поскольку вклад от второго слагаемого невелик. Выражение (1.6.16) пригодно для вычисления моментов распределения уходов нуля ДХ, если задана статистика подынтегральной функции $\epsilon(\omega)$ и форма $S_{\Sigma\chi}(\omega)$, $\Phi_1(\omega)$ и $\Phi_2(\omega)$.

Выбор и расчет ДХ в КД с ортогональными фильтрами первого порядка. Другие показатели, по которым в линейном КД можно судить о качестве

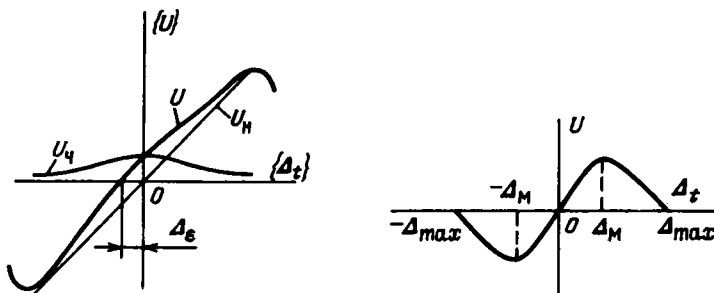


Рис. 1.22. Искажение характеристики корреляционного дискриминатора из-за нарушения условия ортогональности каналов

Рис. 1.23. Определение полной апертуры и монотонной зоны дискриминационной характеристики

ДХ и ее пригодности для работы в составе ОЭСС, также определяются частотными свойствами каналов. При инженерных расчетах, кроме параметров, необходимых при оценке помехоустойчивости (см. § 1.9), используют в роли таких показателей полную апертуру ДХ по входному воздействию $2\Delta_{max}$ и ширину монотонной зоны $2\Delta_m$ (рабочую апертуру), показанные на рис. 1.23.

От апертуры ДХ существенно зависит реакция системы на экстремальные воздействия — быстрые сдвиги изображений, максимальные уровни помех и другие факторы, способные вызывать срывы слежения. Количественные расчеты экстремальных ситуаций на ранних стадиях проектирования ОЭСС сложны и малопродуктивны. Поэтому, в частности, при выборе формы ДХ дискриминаторов руководствуются качественными требованиями, вытекающими из упрощенного анализа процессов срыва слежения.

Известно, например из [91], что ожидаемая вероятность срыва при прочих равных условиях снижается, если увеличить площадь под расчетной кривой $|U(\Delta_t)|$, описывающей обе ветви ДХ в пределах полной апертуры. Это заставляет стремиться к увеличению последней, если, конечно, в результате не ухудшаются иные свойства системы.

Принцип формирования характеристик КД для широкополосных, в частности телевизионных, сигналов с помощью фильтров предоставляет определенные возможности в части управления формой ДХ. Предварительный выбор фильтров в каналах дискриминатора ТВ сигналов можно производить без учета помех и краевых эффектов, считая главными критериями ширину ДХ и стабильность характеристики. Последующие уточнения мож-

но производить в процессе отладки системы по более сложным критериям, учитывающим динамические и флуктуационные ошибки слежения, вероятность его срывов, относительное влияние технических помех, их зависимость от настройки КД и другие факторы, не поддающиеся априорным оценкам.

Рассмотрим формирование ДХ в случае применения в каналах КД простейших взаимно ортогональных фильтров первого порядка с характеристиками

$$\dot{\Phi}_1 = \frac{\gamma}{\gamma + j\omega}, \quad \dot{\Phi}_2 = \frac{j\omega}{\gamma + j\omega}. \quad (1.6.17)$$

Как первая, так и вторая реализуются, например, RC - и CR -звеньями с одинаковыми постоянными времени $\tau = \gamma^{-1}$. Произведение $\dot{\Phi}_1^* \dot{\Phi}_2$ является чисто мнимым, а потому ВКФ импульсных реакций таких фильтров $\varphi_\gamma(\Delta_t)$ есть нечетная функция сдвига

$$\dot{\Phi}_1^* \dot{\Phi}_2 = \frac{j\gamma\omega}{\gamma^2 + \omega^2} \leftrightarrow \gamma \operatorname{sign}(\Delta_t) \exp(-\gamma|\Delta_t|) = \varphi_\gamma.$$

Чтобы оценить ожидаемую форму ДХ, аппроксимируем с точностью до константы используемую часть ожидаемого спектра энергии ТВ сигнала в одном поле квадратом АЧХ некоторого формирующего фильтра

$$\Phi_\Phi^2(\omega) = \frac{\alpha^2 \omega^2}{(\alpha^2 + \omega^2)(\beta^2 + \omega^2)}. \quad (1.6.18)$$

Такой спектр будет формироваться в однокаскадных усилителях каждого из каналов КД перед ортогональными фильтрами, если спектр входного ТВ сигнала равномерен в области, превосходящей частотный интервал от нижней границы $\omega_n = \alpha$ до верхней $\omega_b = \beta$.

Поскольку в целях снижения потерь энергии максимум произведения АЧХ ортогональных фильтров, находящийся на критической частоте $\omega_{кр} = \gamma$, следует размещать между частотами среза формирующих фильтров, в дальнейшем остановимся только на случае, когда

$$\omega_b > \omega_{кр} > \omega_n.$$

Находя взаимный энергетический спектр $\dot{S}$ сравниваемых сигналов в принятой модели, можно фурье-преобразованием определить ожидаемую ДХ $\hat{U}(\Delta_t)$ через свертку функции $\varphi_{\alpha\beta}$, которая сопряжена с (1.6.18), и функции φ_γ :

$$\dot{S}(\omega) = \frac{j\alpha^2\gamma\omega^3}{(\alpha^2 + \omega^2)(\beta^2 + \omega^2)(\gamma^2 + \omega^2)} \leftrightarrow \varphi_\gamma \otimes \varphi_{\alpha\beta} \equiv \hat{U}(\Delta_1),$$

где

$$\varphi_{\alpha\beta} = \left(1 - \frac{\alpha^2}{\beta^2}\right)^{-1} [\beta \exp(-\beta|\Delta_t|) - \alpha \exp(-\alpha|\Delta_t|)] \leftrightarrow \Phi_\Phi^2.$$

При $\gamma = 1$ результат имеет вид

$$\hat{U}(\Delta_r) \equiv \left[\frac{\beta^2}{\beta^2 - 1} (e^{-|\Delta_r|} - e^{-\beta|\Delta_r|}) - \frac{\alpha^2}{\alpha^2 - 1} (e^{-|\Delta_r|} - e^{-\alpha|\Delta_r|}) \right] \text{sign } \Delta_r. \quad (1.6.19)$$

Нарастивая число фильтрующих звеньев, можно убедиться в том, что выходная характеристика КД будет представляться последовательной сверткой нечетной функции φ_γ с соответствующим числом четных $\varphi_{\alpha\beta i}$, т.е.

$$\hat{U}(\Delta_r) = \varphi_\gamma \otimes \varphi_{\alpha\beta 1} \otimes \varphi_{\alpha\beta 2} \otimes \dots \otimes \varphi_{\alpha\beta i} \otimes \dots \quad (1.6.20)$$

Если все фильтры, обеспечивающие ортогонализацию и формирующие ДХ, являются звеньями первого порядка, то в свертке участвуют только экспоненциальные функции с показателями, линейно зависящими от сдвига Δ_r . Поэтому (1.6.20) дает в результате

$$\hat{U}(\Delta_r) \equiv \sum_{i=1}^n b_i \exp(-\alpha_i |\Delta_r|) \text{sign } \Delta_r, \quad (1.6.21)$$

где α_i — частоты среза в фильтрах; b_i — коэффициенты, являющиеся алгебраическими функциями различных α_i ; n — число свободных переменных, т.е. количество независимых фильтрующих звеньев.

Правая часть (1.6.21) позволяет построить зависимости полной апертуры и монотонной зоны ДХ от параметров формирующих и ортогонализирующих фильтров. Они находятся численным решением соответствующих уравнений

$$\hat{U}(\Delta_{\max}) = 0, \quad \hat{U}'(\Delta_m) = 0.$$

Расчеты, выполненные на ЭВМ по второй из этих формул и уравнению (1.6.19), представлены на рис. 1.24 кривыми, которые вычислены для разных соотношений между нижней $\omega_n = 2\pi F_n$ и верхней $\omega_b = 2\pi F_b$ границами полосы пропускания формирующих цепей и критической частоты $\omega_{кр} = 2\pi F_{кр}$ ортогональных фильтров в безразмерном представлении так, что

$$\alpha : \beta : \gamma = \omega_n : \omega_b : \omega_{кр}.$$

Из графиков видно, что монотонную зону ДХ $2\Delta_m$ грубо можно оценивать в секундах как величину, обратную критической частоте ортого-

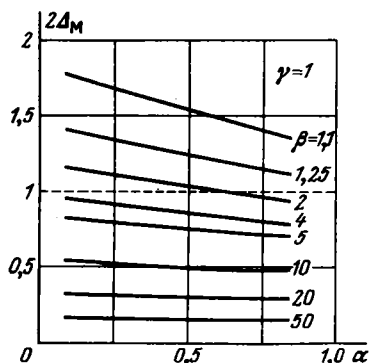


Рис. 1.24. Обобщенные зависимости монотонной зоны дискриминационной характеристики от параметров фильтров в каналах корреляционного дискриминатора

нальных фильтров $\omega_{кр}$ в радианах в секунду. Такой оценке соответствует штриховая горизонтальная прямая на уровне $2\Delta_m \gamma = 1$. При точном расчете на результате сказывается положение НЧ и ВЧ границ формирующих частотных характеристик или соответствующих точек на кривой, аппроксимирующей спектр сигнала. При $0,1 \leq \alpha/\gamma \leq 0,9$ и $1,2 \leq \beta/\gamma \leq 10$ справедливо линейное приближение

$$2\Delta_m \gamma \approx 0,4 + 0,04 \gamma \alpha^{-1} + 1,42 \gamma \beta^{-1}. \quad (1.6.22)$$

Полная апертура ДХ в указанных условиях аппроксимируется уравнением

$$2\Delta_{max} \gamma \approx 3,3 + 0,9 \gamma \alpha^{-1} + 2,45 \gamma \beta^{-1}.$$

Ход реальных ДХ в телевизионных КД со стандартным разложением сигналов при значениях критической частоты порядка $\gamma \sim 10^5 \div 10^6$ рад/с в окрестностях нуля инвариантен к смене сюжетов с точностью до крутизны, зависящей от энергии сигнала.

В монотонной зоне ДХ слабо зависит от оптической структуры изображений при наблюдении природных ландшафтов или объектов искусственного происхождения, если только в спектре регистрируемого сигнала не содержатся мощные периодические составляющие, могущие попасть в полосу рабочих частот дискриминатора. За пределами указанной зоны, как правило, становятся заметными меняющиеся от изображения к изображению отклонения от расчетных ДХ, обусловленные наличием краевых эффектов.

Учет статистики наблюдаемого поля позволяет оценить разбросы ДХ на множестве изображений, регистрируемых в ОЭСС при наличии краевых эффектов из-за ограниченности угла зрения. Для случайной модели однородного гауссова скалярного поля с заданной корреляцией такой учет выполнен в [110].

Результаты этой работы по существу показывают, что дисперсия ДХ растет пропорционально относительной мощности краевого эффекта. Более содержательные нестационарные модели полей в указанном смысле не исследовались.

"Предельная" выходная характеристика линейного КД. Рассмотрение характеристик линейных КД на простейших фильтрах позволяет заключить, что при многих характерных параметрах формирующих и ортогонализирующих цепей значения крутизны ДХ $\hat{U}'(\Delta_r)$ для всех сдвигов, отличных от нуля, по модулю не превосходят крутизны в точке $\Delta_r = 0$.

Это заключение носит достаточно общий характер, поскольку сделанный вывод можно получить из свойств АКФ всякого бесконечно дифференцируемого финитного сигнала и из алгоритма (1.6.8), описывающего одномерный линейный КД, приняв условие отсутствия помех и краевых эффектов и одно ограничение на свойства фильтров.

Запишем ДХ для КД в виде

$$U(\Delta_r) = \frac{d}{d\Delta_r} (K_{вх} \otimes \varphi), \quad (1.6.23)$$

где $K_{вх}$ — АКФ входного сигнала f ; φ — вспомогательная четная функция,

удовлетворяющая дифференциальному уравнению

$$\frac{d}{d\Delta_t} \varphi = W_{12}.$$

Ограничимся фильтрами, для которых характеристику φ можно представить как АКФ некоторой функции g :

$$\varphi = \check{g} \otimes g.$$

Используя ассоциативность и коммутативность свертки, определение АКФ сигнала и возможность ее дифференцирования путем взятия производной от одного из сомножителей [137], получаем для крутизны таких ДХ

$$-U'(\Delta_t) = [(\check{f}' \otimes \check{g}) \otimes (f' \otimes g)](\Delta_t). \quad (1.6.24)$$

Выделенные круглыми скобками результаты однократных свертки в силу дифференцируемости f и сохранения аналитичности при свертке с любой (в том числе и обобщенной) функцией можно считать новыми процессами, которые различаются лишь знаками аргументов:

$$s = f' \otimes g = s(t), \quad \check{s} = \check{f}' \otimes \check{g} = s(-t).$$

Следовательно, их свертка, выражающая с точностью до знака крутизну ДХ, представима, как АКФ K_s сигнала s :

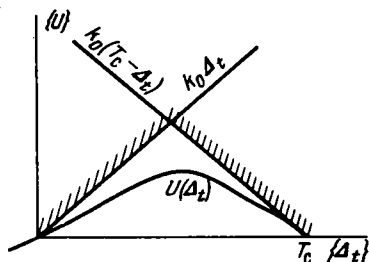
$$-U'(\Delta_t) = K_s(\Delta_t) = (\check{s} \otimes s)(\Delta_t).$$

Используя общие свойства АКФ и освобождаясь от ненужных различий в знаках, можно записать

$$|U'(\Delta_t)| \leq K_s = |U'(0)|. \quad (1.6.25)$$

Полученный результат на практике позволяет судить о том, какие ограничения на форму выходной характеристики КД накладывает линейность

Рис. 1.25. Отыскание предельной дискриминационной характеристики линейного корреляционного дискриминатора сигналов конечной длительности



трактов обработки сигналов в сочетании с условием обязательного гашения остатков переходных процессов от действия предшествующей строки развертки к началу очередной, если развертка происходит по строкам.

В телевидении указанное условие реализуется схемами привязки уровня и, строго говоря, означает, что линейные фильтры нестационарны. Это осложнение можно учесть, если ограничить длительность реакции фильтров в каналах КД временем T_c , а при рассмотрении ДХ помнить, что она не отличается от нуля вне интервала $\pm T_c$.

Пользуясь ограничениями на крутизну и ширину ДХ в линейном КД, несложно найти такую функцию $U(\Delta_f)$, которая максимизирует площадь под этой характеристикой на полуинтервале значений $\Delta_f > 0$ при граничных условиях $U(0) = 0$, $U(T_c) = 0$ и заданном значении

$$k_0 = U'(0) > 0.$$

Этой функцией, показанной на рис. 1.25 и описывающей предельную ДХ, является функция

$$U_{\max}(\Delta_f) = \begin{cases} k_0 \Delta_f & \text{при } 0 \leq \Delta_f < T_c/2, \\ k_0(T_c - \Delta_f) & \text{при } T_c/2 \leq \Delta_f < T_c, \\ 0 & \text{при } T_c \leq \Delta_f. \end{cases}$$

Найденное решение сокращает диапазон поиска наилучших фильтров для каналов линейного КД при его проектировании.

§ 1.7. Нелинейные преобразования и стабилизация характеристик в корреляционных дискриминаторах

Назначение и разновидности нелинейных преобразований в ОЭС. Данную выше интерпретацию алгоритма работы корреляционного дискриминатора можно сохранить во многих случаях, когда каналы, предшествующие коррелятору, содержат нелинейные цепи. Напомним, что общие принципы построения КД были выявлены безотносительно к спектральному составу сигналов-изображений. Как следует из (1.6.7), качественный показатель — нечетность одномерной ДХ — остается без изменений, если только используемые нелинейности (квантователи уровня, ограничители и т.п.) не нарушают четности ВКФ по измеряемому сдвигу при отсутствии помех и краевых эффектов. Очевидно, любые нелинейные преобразования сигналов, сохраняющие сдвиг между ними, можно считать допустимыми. Множество допустимых вариантов нелинейной обработки коррелируемых сигналов весьма значительно. Это нашло отражение в технической литературе по коррелометрам разного назначения [45, 55, 66, 82, 168].

Примером нелинейного преобразования из области ОЭС служит применение в телевизионном дискриминаторе операций бинарного квантования ("бинаризации") полезных переменных слагающих видеосигналов и дифференцирования одного из полученных колебаний перед подачей его на коррелятор. Такая знаковая обработка позволяет формировать входные сигналы коррелятора в виде последовательностей прямоугольных импульсов фиксированной амплитуды, поэтому его схема получается достаточно простой. Ее недостатки применительно к ОЭСС проявляются в трудностях точного центрирования переменных слагающих перед их ограничением, а погрешности выполнения этой операции сказываются на положении нуля ДХ.

Степень вредного влияния ошибок центрирования в знаковых коррелометрах изучалась в [56] для случая стационарных гауссовых процессов, однако из-за отличия распределений ТВ сигналов от стационарных и нор-

мальных достоверный расчет уходов нуля в КД такого типа затруднителен. Анализ нелинейных корреляционных дискриминаторов часто приводит к тому, что ОЭСС, построенные на их основе, приходится исследовать методом "черного ящика", подбирая адекватные математические модели лишь после обширных экспериментов с действующими макетами.

Поиск "хороших" методов нелинейного преобразования сигналов для ОЭС преследует обычно две противоположные цели: упрощение аппаратуры и улучшение изображений в смысле, зависящем от назначения системы. В примере с бинарным квантованием достигается первая цель. Она характерна для полностью автоматических ОЭСС, которые должны быть инвариантными к смене сюжетов и не решают иных задач, кроме слежения за предъявленным изображением.

Вторая цель преследуется, например, в устройствах нелинейной частотной (так называемой *гомоморфной*) фильтрации изображений [31], обычно используемой для улучшения их детальности и ослабления НЧ помех за счет подчеркивания высших частот в спектре логарифма входного сигнала. Этот прием помогает при распознавании образов и в случае визуализации видеосигналов в ОЭС, взаимодействующей с человеком.

Тому же назначению служат различные нелинейные устройства для улучшения гистограммы (распределения) мгновенных значений видеосигнала, в конечном счете повышающие различимость малоконтрастных деталей. В ТВ технике такие устройства называют *контрастерами*. Число примеров возможных преобразований видеосигналов перед их корреляционным сравнением можно было бы продолжить, пользуясь обилием описанных в литературе алгоритмов и методов "препарирования" изображений с целью подчеркивания или выделения определенных свойств [100, 148].

О применимости алгоритмов КД в системах с морфологическими преобразователями изображений. Одной из характеристик, инвариантных к колебаниям яркости элементов или освещенности объектов, составляющих изображение, служит его "форма". Сравнение "формы" изображений принято считать предметом морфологического анализа [44, 60, 101–104].

Понятие "формы" связано со ступенчатой аппроксимацией непрерывного изображения $f(x)$ некоторой конечной суммой $\varphi(x)$ из n взвешенных бинарных функций $\chi_i(x)$, служащих индикаторами непересекающихся подмножеств $\mathcal{A}_i$ поля зрения $\mathcal{X}$. При этом имеют место соотношения множеств и функций:

$$\mathcal{X} = \bigcup_{i=1}^n \mathcal{A}_i, \quad \mathcal{A}_i \cap \mathcal{A}_j = \emptyset \quad \text{при } i \neq j,$$

$$\chi_i(x) = \begin{cases} 1 & \text{при } x \in \mathcal{A}_i, \\ 0 & \text{при } x \notin \mathcal{A}_i, \end{cases}$$

$$\varphi(x) = \mathfrak{P} f(x) = \sum_{i=1}^n a_i \chi_i(x),$$

где

$$a_i = \frac{\int_{\mathcal{A}_i} f(x) d\mu(x)}{\int_{\mathcal{A}_i} d\mu(x)} = \frac{(f, \chi_i)}{|\chi_i|},$$

$\mu(x)$ — мера на множестве $\mathcal{X}$; $\mathcal{P}$ — линейный оператор "проектирования", переводящий вектор f в его проекцию φ .

Если в изображении $\varphi(x)$ все коэффициенты разложения a_i различны, то его формой будет система $\mathcal{A}(\varphi)$ из 2^n непересекающихся подмножеств поля зрения $\mathcal{X}$, в которую входят всевозможные объединения $\mathcal{A}_i$ и пустое множество ϕ . При совпадении $a_i = a_j$ объединение $\mathcal{A}_i \cup \mathcal{A}_j$ составляет одно подмножество, а форма $\mathcal{A}(\varphi)$ содержит 2^{n-1} , т.е. меньшее число множеств. Число множеств, образующих систему $\mathcal{A}(\varphi)$, может служить мерой сложности формы.

С каждым из N изображений f_α и аппроксимирующих их функций φ_α можно связать свое разбиение поля зрения $\mathcal{X}$ на подмножества $\mathcal{A}_{\alpha i}$ так, что будут иметь место соответствия (гомоморфизмы)

$$f_\alpha \rightarrow \varphi_\alpha \rightarrow \mathcal{A}_{\alpha 1} \cup \mathcal{A}_{\alpha 2} \cup \dots \cup \mathcal{A}_{\alpha i} \cup \dots = \bigcup_i \mathcal{A}_{\alpha i}, \quad \alpha \in \{1, 2, \dots, N\}.$$

Это делают, например, при разработке и исследовании алгоритмов для читающих автоматов [44].

В других случаях, когда целью морфологического анализа изображений является оценка отличий от некоторого единственного эталона [104], разбиение целесообразно выполнять один раз, обрабатывая эталон f_0 и используя для этого хотя бы процедуру нелинейного квантования уровней. Применение оператора проектирования $\mathcal{P}_0$ к любому изображению f_α , предъявленному для сравнения, приводит к разложению его по функциям $\{\chi_{i0}\}$, характеризующим форму эталона, и аппроксимации проекцией

$$\varphi_\alpha(x) = \mathcal{P}_0 f_\alpha(x),$$

которую и используют для сравнения с аппроксимированным эталонным изображением

$$\varphi_0(x) = \mathcal{P}_0 f_0(x).$$

Линейность морфологического оператора проектирования $\mathcal{P}$ позволяет применить многие результаты линейной теории сигналов и теории оценок к решению задач анализа изображений по их ступенчатым аппроксимациям, характерным, в частности, для современной цифровой техники. Строгое обоснование применимости аппарата обычного и обобщенного корреляционного сравнения к изображениям, подвергнутым такой аппроксимации, дано в упоминавшихся фундаментальных работах Ю.П. Пытьева [101, 102].

В них кроме доказательства ряда важных теорем содержатся сжатые рекомендации по использованию математических результатов для решения задач идентификации и локализации изображений путем отыскания по их проекциям максимума взаимной корреляции на множестве групповых сдвигов.

Разработка математических основ морфологического анализа явилась необходимым шагом на пути внедрения в обработку изображений цифровых методов все возрастающей сложности [44, 60, 77, 79, 96, 107, 170]. Сравнение бинарных полей яркости в современных корреляционно-экстремальных системах, обладающих умеренным быстродействием [64, 159], можно рассматривать как простой пример морфологической процедуры анализа изображений. Более широкое использование таких методов в подобных и более быстродействующих ОЭСС пока ограничивается технико-экономическими факторами. Ожидается, однако, что совершенствование алгоритмов и бурное развитие технологии производства элементов скоростной цифро-аналоговой и цифровой обработки сигналов приведут в ближайшие годы к серьезным сдвигам в этой области [93, 134, 158, 161, 175, 180].

Линейные операции проектирования, а при необходимости и нелинейные преобразования изображений, способствующие выявлению морфологических признаков, должны, очевидно, учитываться при синтезе корреляционных измерителей сдвига так же, как и любые операторы, действующие на ВКФ входных сигналов коррелятора в составе КД.

Нелинейные операции с выходными сигналами дискриминаторов сдвига, как и процессы линейной фильтрации управляющих сигналов в контуре слежения, не имеет смысла связывать с алгоритмом КД за исключением тех случаев, когда эти преобразования обеспечивают выполнение условий нормировки оценок, а также компенсации перекрестных связей, если они заметны в многомерных измерителях. Применительно к одномерным КД * собственно нормировка обеспечивает стабилизацию крутизны дискриминационной характеристики.

Дестабилизирующие факторы и общие принципы компенсации их влияния в КД. Необходимость принятия специальных мер по стабилизации чувствительности измерителя сдвига изображений в ОЭСС обусловлена следующими основными факторами.

1. Зависимость уровня входных сигналов дискриминатора в системе от условий ее работы. В частности, для ТВ устройств определяющими условиями могут быть: яркость объектов в поле зрения, светорассеяние и поглощение в пространстве распространения (воздушная, водная среда), эксплуатационные колебания чувствительности фотоприемника, деградация запоминаемых изображений в случае их хранения в аналоговых ЗУ и т.д.

2. Непостоянство спектрального состава сигналов из-за смены изображений.

3. Действие возмущений по компонентам сдвига, не подлежащих измерению в данном одномерном КД, но вызывающих частичную декорреляцию сравниваемых сигналов.

Борьба с действием факторов первой группы ведется средствами, используемыми в усилительных цепях радиоустройств для согласования

Б. Ю. М. Астапов

уровней сигналов с динамическим диапазоном тракта, т.е. прежде всего с помощью автоматической регулировки усиления (АРУ). Для той же цели применяют нелинейные элементы со специальными характеристиками, обеспечивающими сжатие динамического диапазона или, наконец, ограничители переменной составляющей, если приняты меры, достаточные для надежного центрирования сигналов.

Изменения уровня сигналов после формирующих фильтров в каналах КД можно частично парировать, охватывая петлей АРУ частотно-избирательные цепи, однако это не обеспечивает защиты от действия факторов третьей группы. Обратимся поэтому к рассмотрению стабилизаторов с использованием АРУ, охватывающей корреляционные звенья в составе дискриминатора [25].

Обозначим через $U(\Delta_r)$ нестабилизированную выходную характеристику КД и воспользуемся ее разложением в ряд по степеням сдвига Δ_r в окрестности точки $\Delta_r = 0$ при отсутствии помех и краевых эффектов:

$$U(\Delta_r) = U'_0 \Delta_r + (3!)^{-1} U'''_0 \Delta_r^3 + (5!)^{-1} U^{(V)}_0 \Delta_r^5 + \dots \quad (1.7.1)$$

Здесь под U'_0 , U'''_0 , $U^{(V)}_0$ и т.д. понимаются значения нечетных производных от ДХ при $\Delta_r = 0$.

Положим, что в системе имеется вспомогательный канал для независимого измерения первой производной от $U(\Delta_r)$. Результат измерения также представим степенным рядом

$$U'(\Delta_r) = U'_0 + (2!)^{-1} U'''_0 \Delta_r^2 + (4!)^{-1} U^{(V)}_0 \Delta_r^4 + \dots, \quad (1.7.2)$$

где в силу четности производной от ДХ КД удержаны только члены четных степеней сдвига.

Пренебрегая всеми слагаемыми, кроме первых, в выражениях (1.7.1) и (1.7.2) и поделив одно на другое, находим оценку

$$\hat{\Delta}_r \approx \frac{U(\Delta_r)}{U'(\Delta_r)}, \quad (1.7.3)$$

т.е. устанавливаем вид операции, необходимой для нормировки выхода одномерного КД в окрестности нулевого сдвига.

Знаменатель в (1.7.3) совпадает по содержанию с одномерным вариантом выражения (1.3.5), что указывает на оптимальность описанной процедуры в смысле ОМНК для случая оценки одномерного сдвига. При наличии сдвигов по другим составляющим для полной оптимизации могут потребоваться измерения и компенсация перекрестных связей.

Сомножителями для получения $U'(\Delta_r)$ должны служить сигналы, поступающие на коррелятор по двум разным каналам. Это обеспечивает нормировку в общем случае, когда разница в энергиях переменной составляющей в обоих изображениях заранее неизвестна. Если эта разница несущественна, схему нормирования можно упрощать, сообразуясь с назначением КД и оставаясь в рамках ограничений, сформулированных в § 1.3.

Алгоритм и структура КД со стабилизацией крутизны выходной характеристики. Чтобы построить алгоритм измерения $U'(\Delta_r)$, воспользуемся

(1.6.24), переписав его так, чтобы реализовать оператор дифференцирования за счет преобразования одной из импульсных характеристик каналов перед коррелятором:

$$-U'(\Delta_t) = \frac{d}{d\Delta_t} [(\check{f} \otimes \check{g}_1) \otimes (f^\Delta \otimes g_2)] = [(\check{f} \otimes \check{g}_1) \otimes (f^\Delta \otimes g'_2)]. \quad (1.7.4)$$

При тех же условиях ненормированный выход КД формируется по правилу

$$U(\Delta_t) = [(\check{f} \otimes \check{g}_1) \otimes (f^\Delta \otimes g_2)], \quad (1.7.5)$$

где g_1 и g_2 взаимно ортогональны.

Таким образом, если частотные характеристики фильтров, обеспечивающих выделение нужной части спектров и ортогонализацию сигналов перед основным коррелятором, есть

$$\dot{\Phi}_1(\omega) \leftrightarrow g_1(t), \quad \dot{\Phi}_2(\omega) \leftrightarrow g_2(t),$$

то ЧХ фильтров, предшествующих вспомогательному коррелятору, который производит оценку крутизны ДХ, могут находиться, как

$$\dot{\Phi}_1(\omega) \leftrightarrow g_1(t), \quad \dot{\Phi}_3(\omega) = j\omega \dot{\Phi}_2(\omega) \leftrightarrow g'_2(t) = g_3(t).$$

При выборе $\dot{\Phi}_1$ и $\dot{\Phi}_2$ по (1.6.17) получим, например,

$$\dot{\Phi}_1 = \dot{\Phi}_3 = \frac{j\omega}{\gamma + j\omega}, \quad \dot{\Phi}_2 = \frac{\gamma}{\gamma + j\omega}.$$

Схема, реализующая алгоритмы (1.7.3) – (1.7.5), выглядит, как показано на рис. 1.26, где K и K' – основной и вспомогательный корреляторы, а знаком $\div$ обозначен блок, в котором выполняется операция деления.

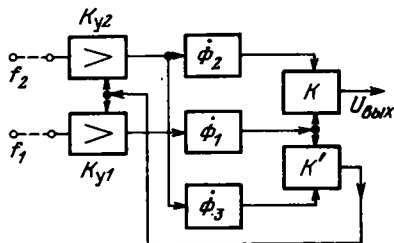
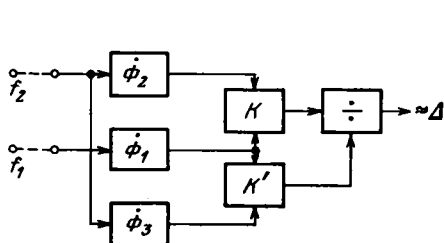


Рис. 1.26. Структура линейного корреляционного дискриминатора со стабилизацией крутизны выходной характеристики без применения обратной связи

Рис. 1.27. Структура линейного корреляционного дискриминатора со стабилизацией крутизны выходной характеристики посредством автоматической регулировки усиления

Известно, что последняя процедура достаточно точно выполняется в цифровой технике. Для аналоговых систем более подходящим является вариант с замкнутой петлей авторегулирования, подобной контуру АРУ уровня сигналов (рис. 1.27). Условимся называть ее *схемой АРУ по крутизне* характеристики КД.

Главная особенность этой схемы состоит в том, что она обеспечивает получение частного от деления U на U' с погрешностью, возрастающей по мере снижения крутизны, так как уравнение, описывающее статическую ДХ дискриминатора с АРУ, представляет собой с точностью до постоянного

множителя K_0 нелинейную зависимость вида

$$U_{\text{вых}}(\Delta_f) = \frac{K_0 U(\Delta_f)}{1 + K_0 C U'(\Delta_f)}, \quad (1.7.6)$$

где $C = C(U')$ — характеристика регулируемой схемы по регулируемому воздействию U' ; $U(\Delta_f)$ — выход КД при отсутствии стабилизации.

Как видно из (1.7.6), стабилизатор выполняет свои функции при условиях $C \approx \text{const}$ и $K_0 C U' \gg 1$. В этом можно убедиться, если представить коэффициенты передачи K_{y1} и K_{y2} усилителей в схеме на рис. 1.27 как монотонно убывающие функции регулирующего воздействия

$$K_{y1}(U') = K_{01}[1 - C_1(U')U'],$$

$$K_{y2}(U') = K_{02}[1 - C_2(U')U'],$$

где C_1 и C_2 — коэффициенты, определяющие чувствительность схемы по регулируемому воздействию. При этом принято, что

$$K_0 = K_{01} K_{02}, \quad C = C_1(U') + C_2(U') - C_1(U')C_2(U')U'.$$

В дополнение к (1.7.6) находим $U'_{\text{вых}}$ — выход вспомогательного коррелятора, вырабатывающего оценку крутизны, с учетом действия АРУ и принятых обозначений:

$$U'_{\text{вых}} = \frac{K_0 U'(\Delta_f)}{1 + K_0 C U'(\Delta_f)}.$$

При медленных отклонениях U' от номинала и таком выборе регулировочных характеристик, что $C \approx \text{const}$, для случая $K_0 C U' \gg 1$ имеем статическую ДХ

$$U_{\text{вых}}(\Delta_f) \approx \frac{U}{C U'} = C^{-1} \Delta_f.$$

Оценка динамических свойств КД с АРУ, стабилизирующей крутизну выходной характеристики. Исследование динамических свойств корреляционного дискриминатора с АРУ по крутизне представляет собой достаточно громоздкую задачу, требующую цифрового моделирования с учетом выбранных параметров фильтров в каналах КД, инерционности нагрузки, дискретности измерений сдвига и нелинейности реальных регуляторов усиления.

Некоторые оценки, позволяющие качественно прогнозировать поведение стабилизированного одномерного КД, можно сделать, приняв ряд упрощений.

Пусть регулирующие цепи в каналах КД обеспечивают в совокупности получение идеальной кусочно-линейной регулировочной характеристики

$$K(u) = \begin{cases} K_0 & \text{при } u \leq 0, \\ K_0(1 - Cu) & \text{при } 0 < u < u_3, \\ 0 & \text{при } u_3 < u, \end{cases} \quad (1.7.7)$$

где u — управляющее воздействие; u_3 — значение u , при котором КД запирается.

Рассмотрим поведение сигналов в системе регулирования по эквивалентной схеме контура стабилизации КД (рис. 1.28). Эта схема составлена с учетом того, что при покадровой съеме полезной информации для уменьшения пульсаций на выходе дискриминатора и в тракте АРУ следует использовать фиксирующую цепь (*расширитель*) P , которая расширяет отсчетные импульсы на выходе измерителя крутизны ДХ до длительности,

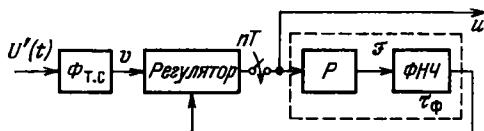


Рис. 1.28. Эквивалентная схема контура стабилизации корреляционного дискриминатора с покадровой обработкой информации при использовании автоматической регулировки усиления по крутизне

равной интервалу T между соседними отсчетами, т.е. служит интерполятором нулевого порядка.

Вслед за ним включен простой фильтр низкой частоты (ФНЧ) первого порядка с постоянной времени $\tau_{\Phi} = T/m$, необходимый, в частности, для сглаживания возможных в ТВ системе коммутационных помех.

При анализе такой замкнутой системы удобно принимать за входное воздействие ¹⁾ v для схемы, приведенной на рис. 1.28, результат сглаживания функции $U'(t)$ с помощью оператора текущего среднего, так как именно эта операция выполняется в корреляторе K' :

$$v(t) \equiv \int_{t-T}^t U'(t) dt. \quad (1.7.8)$$

Поскольку дискретизованный во времени съем информации с коррелятора соответствует действию последовательности отсчетов на входе эквивалентной схемы, ее можно рассматривать как замкнутый импульсный регулятор. Выражение, связывающее выход

$$u \equiv U'_{\text{вых}}$$

со входом, имеет вид

$$u = K_0(1 - C\mathcal{F}u)v, \quad (1.7.9)$$

где $\mathcal{F}$ — оператор, учитывающий запаздывание в расширителе P и действие ФНЧ.

Уравнение (1.7.9) характерно для АРУ многих видов. Оно является нелинейным и обычно решается приближенными методами [19, 135]. В нашем случае, однако, благодаря простоте оператора $\mathcal{F}$ оказывается возможным строго рассчитать переходный процесс при ступенчатом

¹⁾ v — телевизионный сигнал, подвергнутый фильтрации в цепях обработки $\Phi_{т.с.}$.

скачке на входе. Решение находится применением дискретного преобразования Лапласа к (1.7.9).

Рассматривая u и v как функции дискретного времени nT и обозначая через $U(q)$ и $V(q)$ их "изображения" в плоскости переменной q , получаем то же уравнение в операторной форме

$$U + K_0 C V \otimes U \Phi = K_0 V, \quad (1.7.10)$$

где $\Phi = \Phi(q)$ — дискретное преобразование Лапласа для оператора $\mathcal{F}$, а $\otimes$ — символ дискретной свертки.

Пользуясь таблицами из [135], находим для ступенчатой решетчатой функции дискретного времени

$$v(nT) = \sigma \cdot 1(nT)$$

ее изображение

$$V(q) = \sigma \frac{e^q}{e^q - 1} \rightarrow \sigma \frac{z}{z - 1}$$

Согласно [135] имеем также изображение оператора, описывающего расширитель с ФНЧ:

$$\Phi(q) = \frac{1 - e^{-m}}{e^q - e^{-m}} \rightarrow \frac{1 - e^{-m}}{z - e^{-m}}$$

Свертка в левой части (1.7.10) дает в нашем случае

$$V \otimes U \Phi = \sigma \frac{1 - e^{-m}}{e^q - e^{-m}} U, \rightarrow \sigma \frac{1 - e^{-m}}{z - e^{-m}}$$

благодаря чему получаем операторный аналог (1.7.9), в котором учтены начальные условия:

$$U + K_0 C \left(\sigma \frac{1 - e^{-m}}{e^q - e^{-m}} + v_0 \right) U + u_0 V = K_0 V. \quad (1.7.11)$$

Разрешив это уравнение относительно U и перейдя от изображений к оригиналам, получим в моменты взятия отсчетов $nT = t$ интересующую нас последовательность измеренных схемой K' в стабилизированном КД значений

$$u(nT) = \frac{K_0 v_0}{1 + \mu_0} \left\{ 1 + \frac{\sigma}{v_0} \frac{1 + \mu[(1 + \mu)e^{-m} - \mu]^n}{1 + \mu} \right\}, \quad (1.7.12)$$

где

$$\mu_0 = K_0 C v_0; \quad \mu = K_0 C (\sigma + v_0).$$

При получении этой формулы использовано соотношение

$$u(0) = \frac{K_0 v_0}{1 + K_0 C v_0}.$$

Процесс $u(nT)$ сходится к постоянной величине при

$$e^{-m} > \frac{\mu - 1}{\mu + 1}. \quad (1.7.13)$$

Это условие устанавливает границу области существования обратного преобразования Лапласа для (1.7.11).

Напомним, что в силу (1.7.7) управляющее воздействие u в регуляторе всегда положительно, а его максимальное стационарное значение характеризует чувствительность регулирующего органа

$$C^{-1} = u_{\max}.$$

Так как крутизна ДХ при действии реальных сигналов не превосходит некоторой максимальной величины $v_{\max}$, можно определить диапазон входных воздействий, включая в них и начальные условия, неравенствами

$$0 \leq \mu < \mu_{\max},$$

где $\mu_{\max} = K_0 C v_{\max}$ — максимальный нормированный фактор обратной связи.

Приравнявая нулю содержимое квадратных скобок в правой части (1.7.12), находим граничное условие, при котором переходный процесс завершается за один такт дискретизации по времени:

$$(\mu + 1)e^{-m} = \mu. \quad (1.7.14)$$

Если правая часть здесь будет больше левой, процесс установления станет колебательным, в противоположном случае — аperiodическим. Поскольку величина скачка σ и начальное условие v_0 входят в (1.7.14), заключаем, что характер переходного процесса в системе может сильно меняться, несмотря на постоянство ее параметров. Это явление типично для большинства АРУ. *назад*

При необходимости максимального увеличения быстродействия АРУ постоянную времени ФНЧ надо выбирать, обеспечивая выполнение равенства (1.7.14) при подстановке в него величины $\mu_{\max}$ вместо μ . При этом установление процесса за один такт будет достигаться при значениях *туда*

$$\mu = \frac{\mu_{\max} - 1}{2}. \quad (1.7.15)$$

Так как для успешной стабилизации крутизны КД необходимо выполнение неравенства $\mu_{\max} \gg 1$ формула (1.7.15) служит ориентиром для оценки величины скачков σ , которые могут парироваться системой АРУ, не приводя к заметному росту вероятности срыва слежения в замкнутой кор-

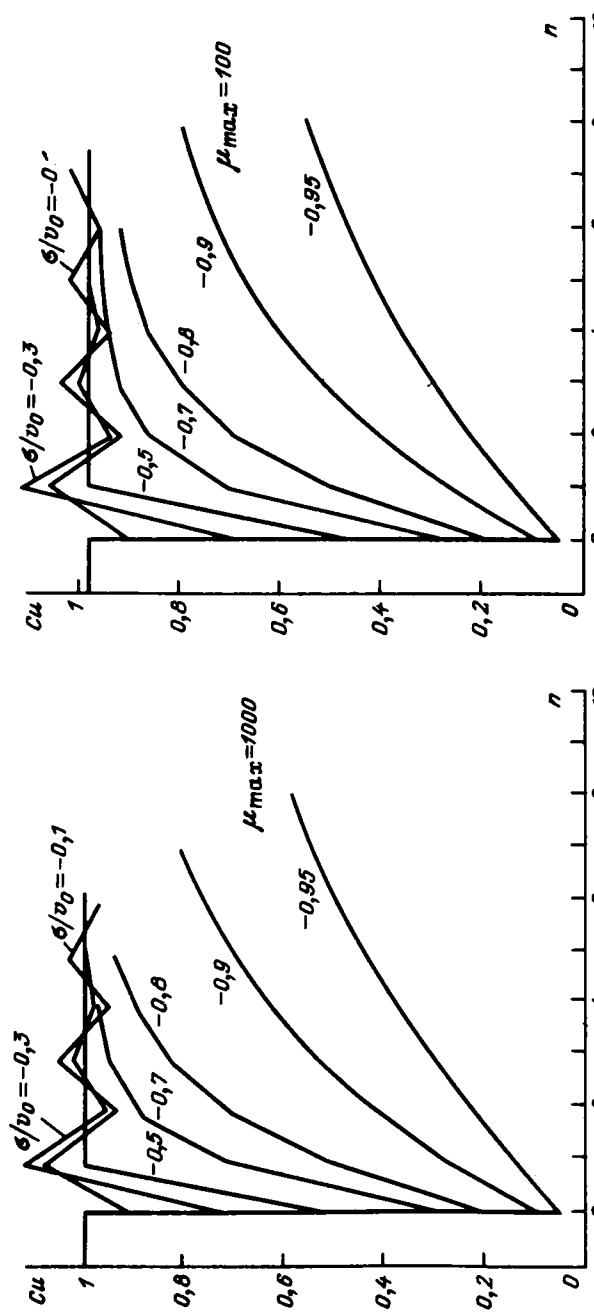


Рис. 1.29. Переходные процессы при ступенчатых возмущениях в контуре стабилизации крутизны дискриминационной характеристики линейного корреляционного дискриминатора

реляционной системе. В этом смысле наиболее опасными являются быстрые переходы от больших значений u к малым на величину, превосходящую $v_{\max}/2$, ибо это соответствует смещению рабочей точки на ДХ от середины к ее краю со снижением чувствительности КД на время до окончания переходного процесса в АРУ.

О зависимости времени установления в АРУ от величины скачка можно судить по серии кривых на рис. 1.29, построенных в соответствии с (1.7.12) для случаев, когда максимальные значения факторов обратной связи $\mu_{\max}$ достигают 100 и 1000. Сглаженные результаты расчета динамической крутизны $u(n)$ представлены в безразмерной форме. что обеспечено умножением обеих частей (1.7.12) на чувствительность регулятора S . Сопоставление кривых указывает на слабую зависимость переходных процессов от $\mu_{\max}$ при высоких значениях этого параметра. Видно также, что крутизна КД, охваченного АРУ рассмотренного типа, при начальных скачках от 1 до 0,3 стабилизируется практически за два-три такта дискретизации, но при дальнейшем увеличении размаха скачков быстродействие системы резко ухудшается.

Каждая из двух серий кривых рассчитана при своем значении параметра m :

$$m = \ln(\mu_{\max} + 1) - \ln(\mu_{\max} - 1).$$

Однако при $\mu_{\max} > 10$ допустимо приближение

$$m \approx \frac{2}{\mu_{\max}},$$

равносильное выполнению равенства

$$\tau_{\Phi} \approx \mu_{\max} \frac{T}{2}.$$

Чтобы полностью использовать на практике возможности АРУ по быстродействию, не опасаясь автоколебаний из-за превышения расчетного значения $\mu_{\max}$, в реальную схему корреляционного ТВ дискриминатора вводят нелинейные элементы, ограничивающие выход K' и регулирующее воздействие u .

§ 1.8. Перекрестные связи в корреляционных дискриминаторах из-за анизотропности изображений

Анизотропность изображений как причина перекрестных связей в двумерном КД. Опыт показывает, что реальные изображения в любом диапазоне ЭМВ редко бывают изотропными. Еще Д. Габор [164] обращал внимание на то, что естественное многоэлементное изображение практически всегда локально-анизотропно. Это свойство проявляется в структуре двумерных СП и АКФ изображений и особенно заметно, если в наблюдаемом поле зрения имеются одинаково ориентированные протяженные фрагменты, например тени от деталей рельефа, следы пахоты, параллельно расположенные печатные проводники на платах или даже всего один-два мощных элемента: прямая дорога, береговая линия, предмет вытянутой формы и т.д.

Из сказанного ясно, что всякая изотропная статистическая модель изображения служит первым приближением к реальности и годна только для анализа одномерных систем. Вторым приближением, минимально необходимым при исследовании совместного действия каналов в двумерной ОЭС, измеряющих обе компоненты плоского сдвига, является модель, обладающая эллиптическим в горизонтальных сечениях двумерным спектром и соответственно АКФ с теми же свойствами

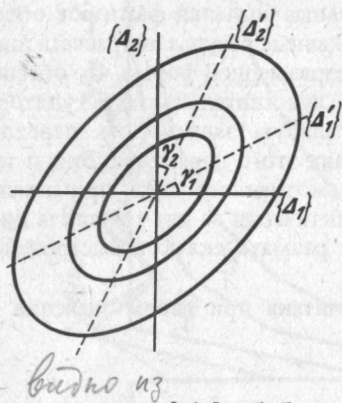


Рис. 1.30. Эллиптическая аппроксимация сечений двумерной анизотропной автокорреляционной функции

Как показано в § 1.3, обобщенный алгоритм двумерного КД предусматривает внесение в одномерные составляющие оценки поправок на влияние сдвигов по ортогональным компонентам. Поправки являются линейными в силу квадратичности меры приближения оценки.

В соответствии с (1.3.19), (1.3.20) корреляционная оценка составляющих вектора двумерного аддитивного сдвига $\|\hat{\Delta}_1 \hat{\Delta}_2\|^T$ дается соотношениями

$$\begin{aligned}\hat{\Delta}_1 &= a_1 (\bar{\Delta}_1 - b_{12} \bar{\Delta}_2), \\ \hat{\Delta}_2 &= a_2 (\bar{\Delta}_2 - b_{21} \bar{\Delta}_1),\end{aligned}\tag{1.8.1}$$

где $\bar{\Delta}_1, \bar{\Delta}_2$ — ненормированные оценки компонент сдвига, полученные с помощью одномерных КД без учета перекрестных связей.

Считая причиной указанных связей анизотропность двумерной АКФ изображения и пренебрегая помехами и краевыми эффектами, можно установить, что зависимость между корреляционными оценками компонент оказывается строго линейной лишь в тех случаях, когда все горизонтальные сечения АКФ имеют форму эллипсов с общим центром, отличающихся только масштабом, как показано на рис. 1.30.

Этим свойством обладает всякая АКФ вида $F(-\alpha_1 \Delta_1^2 - \alpha_2 \Delta_2^2)$, где α_1, α_2 — константы¹⁾. Функция F может быть любой из тех, которые пригодны для описания АКФ, например экспонентой.

При указанных условиях одномерная оценка сдвига $\bar{\Delta}_1$ определяется относительно положения максимума в вертикальном сечении АКФ, параллельном оси $\{\Delta_1\}$ и взятом на удалении $\hat{\Delta}_2$ от нее, так как максимум

¹⁾ С другой стороны, всякая АКФ реального многоэлементного изображения близка к этому виду в малой окрестности точки нулевого сдвига (см., например, [93]).

лежит на оси симметрии этого сечения. Смещение максимумов сечений при $\Delta_2 \neq 0$ равносильно сдвигу нулей ДХ в КД сдвигов Δ_1 . То же самое происходит с ДХ в КД сдвигов Δ_2 .

Анализируя рис. 1.30, можно установить, что геометрическими местами точек максимумов вертикальных сечений АКФ, параллельных оси $\{\Delta_1\}$ либо $\{\Delta_2\}$, являются сопряженные диаметры концентрических эллипсов. Их семейство характеризует двумерную АКФ отфильтрованного сигнала. Эти диаметры показаны штриховыми линиями, образующими при данном сигнале измерительную систему координат для двумерного КД, в котором не учитываются поправки на перекрестные связи.

Очевидно, если главные оси семейства (т.е. большие и малые оси эллипсов) отклоняются от направлений истинных сдвигов $\{\Delta_1\}$ и $\{\Delta_2\}$, то оси измерительной системы $\{\Delta'_1\}$ и $\{\Delta'_2\}$ перестают быть взаимно перпендикулярными.

Тангенсы углов поворота γ_1 и γ_2 диаметров, совпадающих по направлению с осями истинных сдвигов, определяют величину перекрестных связей между их оценками, полученными путем одномерных корреляционных сравнений при отсутствии помех, краевых эффектов и аппаратных погрешностей, т.е. являются коэффициентами линейной регрессии в уравнениях (1.8.1):

$$\operatorname{tg} \gamma_1 = b_{21}, \quad \operatorname{tg} \gamma_2 = b_{12}.$$

Эти показатели случайны на множестве двумерных сигналов, представляющих различные изображения.

При большой анизотропности отдельной реализации и неудачном расположении отсчетных осей $\{\Delta_1\}$, $\{\Delta_2\}$ перекрестные связи между одномерными оценками могут оказаться значительными. Так, задав меру анизотропности в виде $q = \alpha_1/\alpha_2$, можно вычислить максимальную регрессию при наихудшем в смысле перекрестной связи расположении главных осей семейства эллипсов, характеризующих двумерную структуру АКФ:

$$(b_{12})_{\max} = (b_{21})_{\max} = \frac{q^{0,5} - q^{-0,5}}{2}. \quad (1.8.2)$$

Среднеквадратическое значение регрессии на множестве эллипсов с данным параметром q при равновероятной их ориентации в плоскости сдвигов $\{\Delta_1, \Delta_2\}$ составляет

$$(b_{12})_{\text{эфф}} = (b_{21})_{\text{эфф}} = \frac{q^{0,25} - q^{-0,25}}{\sqrt{2}}. \quad (1.8.3)$$

При $q = 2$ формулы (1.8.2), (1.8.3) дают соответственно $(b_{12})_{\max} = 0,354$ и $(b_{12})_{\text{эфф}} = 0,246$, а при $q = 4$ имеем значения 0,75 и 0,5. Очевидно также, что математическое ожидание коэффициентов регрессии равно нулю.

Удачным расположением изображения относительно измерительных осей корреляционного дискриминатора следует считать такое, когда они совмещаются с главными осями семейства эллипсов, характеризующих анизотропность. Тогда имеем $\operatorname{tg} \gamma_1 = \operatorname{tg} \gamma_2 = 0$, и перекрестные связи отсутствуют при любом $q < \infty$, характеризующем квадрат отношения большой и малой осей эллипса.

Поскольку непрогнозируемая деформация измерительных осей в двумерном корреляционном дискриминаторе может вредить качеству слежения, вопросу перекрестных связей следует уделять достаточное внимание при проектировании всякой корреляционной ОЭСС.

Методы компенсации перекрестных связей в двумерных КД. В соответствии с общей структурой каналов многомерного КД можно наметить две группы методов компенсации перекрестных связей из-за анизотропности в зависимости от того, где они реализуются — до коррелятора сравниваемых сигналов или после него.

В первую группу выделим разные варианты создания *анизотропных* двумерных фильтров, которые приближают АКФ сигнала к изотропной форме или по крайней мере уменьшают наклон главных осей семейства анизотропных сечений АКФ относительно направления измерительных осей в КД.

Некоторые сведения об алгоритмах линейной анизотропной фильтрации, пригодных для реализации на больших ЭВМ, можно найти в [21]. Они не имеют прямого отношения к технике КД, но полезны при опознавании образов и могут служить отправной точкой при поиске алгоритмов для КД на базе быстродействующих цифровых процессоров.

Вторая группа методов устранения перекрестных связей между выходами корреляционного дискриминатора базируется на введении необходимых поправок в оценки компонент сдвига, полученные измерением значений одномерных ВКФ сравниваемых сигналов. Главным объектом внимания являются связи между оценками в каналах измерения составляющих плоского сдвига.

В реальных ОЭСС не исключено существование перекрестных связей и по другим каналам многомерного КД, хотя они имеют место не всегда. Так, например, структура ОАКФ в плоскости масштабно-креновых сдвигов, рассмотренных в § 1.5, не порождает подобных связей.

Относящийся ко второй группе *непрямой* метод внесения поправок в оценки плоского сдвига состоит в одновременном корреляционном измерении составляющих в двух взаимно развернутых на угол $\pi/4$ прямоугольных системах координат (Δ_1, Δ_2) и (Δ_1^*, Δ_2^*) .

Для случая линейной регрессии компонент оценки сдвигов

$$\hat{\Delta} = \|\Delta_1 \Delta_2\|^T, \quad \hat{\Delta}^* = \|\Delta_1^* \Delta_2^*\|^T,$$

измеренных в исходной и повернутой системах, можно решить совместно матричные уравнения связи с истинным сдвигом $\underline{\Delta}_n$

$$\hat{\Delta} = B \underline{\Delta}_n, \quad \hat{\Delta}^* = C \underline{\Delta}_n,$$

где B и C — квадратные матрицы вида

$$\left\| \begin{array}{cc} 1 & -b_{12}^{(*)} \\ -b_{21}^{(*)} & 1 \end{array} \right\|$$

с недиагональными элементами, представляющими собой коэффициенты перекрестных связей в соответствующих координатах. Для реализации алгоритма компенсации достаточно использовать любые два коэффициента b_{12} , b_{21} или $b_{12}^{(*)}$, $b_{21}^{(*)}$. Например, первая пара находится решением систе-

мы квадратных уравнений

$$\begin{aligned} b_{12}^{-1} b_{21}^{-1} (2\hat{\Delta}_1 - \hat{\Delta}_1^* + \hat{\Delta}_2^*) + b_{12}^{-2} \hat{\Delta}_2^* - b_{21}^{-2} \hat{\Delta}_1^* - 2b_{12}^{-1} (\hat{\Delta}_2 - \hat{\Delta}_2^*) + 2\hat{\Delta}_1^* &= 0, \\ b_{12}^{-1} b_{21}^{-1} (2\hat{\Delta}_2 - \hat{\Delta}_2^* + \hat{\Delta}_1^*) - b_{12}^{-2} \hat{\Delta}_2^* + b_{21}^{-2} \hat{\Delta}_1^* - 2b_{21}^{-1} (\hat{\Delta}_1 - \hat{\Delta}_1^*) + 2\hat{\Delta}_2^* &= 0. \end{aligned} \quad (1.8.4)$$

Применимость данного метода на практике определяется в основном техническими возможностями организации одновременной работы двух двумерных КД, измеряющих сдвиги во взаимно развернутых системах координат.

Для быстросействующих систем важно также располагать эффективным алгоритмом решения системы (1.8.4).

Прямой метод нахождения регрессионных коэффициентов основан на измерении выходных сигналов КД при специально организуемых калиброванных смещениях $\underline{\Delta}_K$ эталонного изображения, например, вдоль координатных осей:

$$\underline{\Delta}'_K = \parallel \Delta_K \ 0 \parallel^T, \quad \underline{\Delta}''_K = \parallel 0 \ \Delta_K \parallel^T.$$

Если выходы КД при первом и втором смещениях обозначить векторами

$$\underline{\hat{\Delta}}'_K = \parallel \hat{\Delta}'_1 \ \hat{\Delta}'_2 \parallel^T, \quad \underline{\hat{\Delta}}''_K = \parallel \hat{\Delta}''_1 \ \hat{\Delta}''_2 \parallel^T,$$

то искомые коэффициенты найдутся, как результаты прямых измерений

$$b_{12} = -\frac{\hat{\Delta}''_1}{\Delta_K}, \quad b_{21} = -\frac{\hat{\Delta}'_2}{\Delta_K}.$$

Прямой метод проще предыдущего, однако эффективность компенсации перекрестных связей при оценке коэффициентов по калиброванным смещениям может быть недостаточной, так как эти коэффициенты, строго говоря, зависят от величины измеряемого сдвига, а не остаются константами в соответствии с эллиптической аппроксимацией двумерной АКФ. Поскольку в непрямом методе используются оценки фактических, а не заранее назначенных пробных сдвигов, точность компенсации оказывается более высокой. Различие методов тем заметнее, чем значительнее разница между истинной формой двумерной АКФ для данного сигнала и ее эллиптической аппроксимацией. Другими словами, не прямой метод компенсации может давать желаемый эффект и для тех реализаций сигнала, АКФ которых в горизонтальных сечениях нельзя представлять семейством соосных эллипсов. В то же время, несмотря на худшую точность прямого метода компенсации эффекта от анизотропности на выходах КД, его применение может быть оправдано сравнительной простотой алгоритма.

§ 1.9. Действие мешающих факторов на характеристики корреляционных дискриминаторов

Классификация основных мешающих факторов и модель их взаимодействия с сигналами в КД изображений. Выходная характеристика дискриминатора сдвигов изображения в большинстве реальных ОЭСС представляет собой случайную функцию сдвига на множестве возможных сигналов. До сих пор при описании дискриминаторов корреляционного типа их ДХ

рассматривались как детерминированные функции без учета не только краевых эффектов, но и помех. Этот подход оправдывает себя тем, что он базируется на фундаментальных свойствах АКФ, присущих любой реализации сигнала, и позволяет, как видно из § 1.6, сформулировать общие принципы формирования одномерных ДХ в КД.

Многочисленные эксперименты с реальными изображениями, описанные в литературе или проводившиеся создателями различных КД, подтверждают, что дискриминационные характеристики ведут себя в соответствии со свойствами ВКФ ортогональных сигналов. При строгом осуществлении принципа взаимной ортогональности импульсных реакций каналов КД гарантируется получение несмещенного нуля ДХ не только в среднем, но и в каждой реализации, если на дискриминатор не действуют какие-либо помехи.

Возможные отклонения выхода КД от ожидаемого при наличии помех удобно подразделять на три вида: шумовые флуктуации во времени, имеющие место при многократном повторении корреляционной процедуры измерения одного и того же сдвига, смещения усредненного по времени нуля ДХ от точки нулевого сдвига и отклонения истинной ДХ от ожидаемой в среднем для множества сюжетов.

Последний тип искажений ДХ в основном обусловлен краевыми эффектами, которые создают разброс с дисперсией порядка отношения величины сдвига к размеру исходного изображения. Развернутая оценка этой дисперсии при довольно обширных упрощающих допущениях выполнена в [110]. При практической разработке ОЭСС целесообразно находить ожидаемые разбросы ДХ, избегая чрезмерных идеализаций законов распределения и показателей стационарности и по возможности используя для этого массивы реальных данных и технику цифрового моделирования (см. главу 3).

Анализ первых двух видов искажений ДХ из-за помех в корреляционном дискриминаторе можно проводить, опираясь на математическую модель совокупности двумерных сигналов как функций точки x на плоскости, искаженных действием аддитивных и мультипликативных факторов в виде других функций в тех же координатах.

Взаимный сдвиг полезных сигналов может быть обобщенным, однако в данном случае удобнее сразу свести его к аддитивному, пользуясь, если необходимо, представлениями, развитыми в § 1.5.

Модель имеет вид

$$y(x) = Q_1(x) + n_1(x) + [1 + Q_2(x)]f(x + \Delta), \quad (1.9.1)$$

$$f_3(x) = Q_2(x) + n_2(x) + [1 + Q_4(x)]f(x). \quad (1.9.2)$$

Здесь y и f_3 — текущий и эталонный сигналы на входе КД; $n_1(x)$ и $n_2(x)$ — случайные, независимые от сигнала, меняющиеся от кадра к кадру аддитивные помехи; $Q_1(x)$ и $Q_2(x)$ — детерминируемые аддитивные помехи; $Q_3(x)$ и $Q_4(x)$ — детерминируемые мультипликативные помехи.

Эта модель позволяет учесть большинство из мешающих факторов, которые можно обнаружить в реальной корреляционной ОЭСС при ее работе с сигналами, много большими чем квантовые шумы, и не находящейся под действием каких-либо специально организованных помех.

Так, слагаемое n_1 отображает независимые флуктуационные шумы приемника излучений и тракта усиления, преобразованные в электрическую форму. Среднее значение шумов удобно считать нулевым.

Аддитивная помеха Q_1 характерна для телевизионных или тепловизионных систем, где она представляет вклады, создаваемые детерминированным фоновым сигналом мишени фотоприемника, например видикона или фотоматрицы.

Фактор Q_2 может отражать действие аналогичного сигнала, образующегося при воспроизведении эталона из ЗУ, которое выполнено, например, на запоминающей электронно-лучевой трубке (ЗЭЛТ). Компонента Q_3 зависит от координаты x , если чувствительность приемника неравномерна по полю зрения, а фактор Q_4 учитывает возможную неравномерность того же вида в ЗУ, хранящем эталон.

Если в КД обеспечено подавление постоянных составляющих до входов коррелятора, то средние значения Q_1 и Q_2 , так же как Q_3 и Q_4 , считаются нулевыми.

Функции $Q_1(x)$, $Q_2(x)$, $Q_3(x)$ и $Q_4(x)$ могут быть детерминированы для каждого данного КД в сочетании с приемником конкретной ОЭС. В целях упрощения анализа стохастическую помеху $n_2(x)$, формально введенную в модель эталона (1.9.2), допустимо приравнять нулю, если в канале между ЗУ и коррелятором отсутствует собственный шум. При этом реализация помехи, наложившаяся на эталон при его записи, должна быть включена в состав $f_3(x)$.

Перечисленные мешающие факторы обобщенного вида позволяют учесть в модели коррелируемых сигналов и другие помехи, поддающиеся пересчету ко входам коррелятора. В частности, в $n_1(x)$, $n_2(x)$ можно учесть шумы квантования при цифровой фильтрации в трактах КД, шумы округления результатов вычислений и т.п.

Флуктуационные характеристики одномерных КД. Оценки погрешностей, возникающих из-за стохастических помех в КД, удастся довести до числовых значений, приняв гипотезу об однородности и стационарности множества сравниваемых сигналов, а также полагая, что размеры поля зрения достаточно велики, чтобы не считаться с поправками на краевые эффекты. Результаты упрощаются, если использовать еще и отсутствие корреляции сигналов с детерминируемыми мешающими факторами, например структуры сюжетов с неоднородностями мишени видикона в ТВ камере ОЭС. Перечисленные допущения позволяют пользоваться стандартным аппаратом корреляционной теории, дающим возможность находить математическое ожидание и спектрально-корреляционные характеристики процесса на выходе КД, если входные шумы являются случайными функциями времени.

Решение таких задач без учета мультипликативных искажений сигналов подготовлено работами В.В. Цветнова, В.И. Карандасова и В.Д. Пономаревой [52, 53, 132, 133].

Для примера наметим путь расчета флуктуаций на выходе аналогового КД телевизионных сигналов при независимых аддитивных шумах в каналах. Положим полезные составляющие в моделях сигналов одномерными, а все мешающие факторы, кроме названных шумов, несущественными. Приняв, что сравниваемые сигналы в КД отличаются только сдвигом,

и считая СП шумов $N(F)$ в каналах одинаковыми и стационарными, можно представить СП выходных флуктуаций $N_{\text{вых}}(F)$ в КД как функцию частоты F в герцах. Ее находят как сумму сверток спектральных плотностей, участвующих в образовании компонент типа "сигнал-шум" и "шум-шум":

$$N_{\text{вых}}(F) = \beta^2 [(S\Phi_1^2) \otimes (N\Phi_2^2) + (S\Phi_2^2) \otimes (N\Phi_1^2) + (N\Phi_1^2) \otimes (N\Phi_2^2)](F), \quad (1.9.3)$$

где $\beta = \beta(F)$ — коэффициент передачи перемножителя в КД, т.е. отношение выходного сигнала к произведению входных колебаний, зависящее от частоты в силу интегрирующих свойств нагрузки реального коррелятора; $S = S(F)$ — спектральная плотность энергии сигнала; $\Phi_{1,2}^2 = \Phi_{1,2}^2(F)$ — квадраты АЧХ каналов КД.

В реальных ОЭСС с ТВ датчиками используемая часть спектра выходного процесса в КД ограничивается частотами порядка десятка герц, тогда как спектры входных сигналов значительно шире. Поэтому для технических расчетов оказывается достаточным знать асимптотическое значение выходного шума в окрестности нулевой частоты и оперировать эффективной шумовой полосой [24] фильтра после КД $\delta F_{\text{нч}}$, полагая $\beta = \text{const}$ внутри этой полосы.

Отношение стандарта флуктуаций $\sigma_{\text{вых}}^2$ на выходе одномерного КД к k_0^2 — квадрату крутизны ДХ при $\Delta = 0$ удобно записывать, используя связь полезного выхода с входной СП и коэффициентами передачи каналов (1.6.12):

$$U(\Delta_t) = \kappa \int_0^\infty s(F) \Phi_1(F) \Phi_2(F) \sin 2\pi F \Delta_t dF,$$

где κ — произведение коэффициента передачи КД и максимума спектральной плотности энергии полезного сигнала на каждом входе КД, если $s(F)$ — нормированная на S_0 СП сигнала; Δ_t — сдвиг, преобразованный в смещение сигнала по времени за счет развертки.

Очевидно, что

$$k_0 = \int_0^\infty 2\pi F s(F) \Phi_1(F) \Phi_2(F) dF.$$

Раскрывая свертку (1.9.3) при $F = 0$, находим относительную мощность шумовых флуктуаций в выходной полосе КД

$$\frac{\sigma_{\text{вых}}^2}{k_0^2} = \frac{\delta F_{\text{нч}} \int_0^\infty [2r_{\text{вх}}^{-1} s(F) + r_{\text{вх}}^{-2} \nu(F)] \Phi_1^2(F) \Phi_2^2(F) \nu(F) dF}{[2\pi \int_0^\infty F \Phi_1(F) \Phi_2(F) s(F) dF]^2} [c^2], \quad (1.9.4)$$

где $r_{\text{вх}} = N_0/S_0$ — отношение сигнал/шум по спектральной плотности на входе КД, если $\nu(F)$ — нормированная на N_0 СП шума.

При выходе (1.9.4) выходной флуктуационный процесс в КД можно с одинаковым успехом рассматривать либо как непрерывный, получаемый непосредственным перемножением полезных видеосигналов, либо как

результат сглаживания дискретного процесса, представляющего собой последовательность отсчетов с частотой кадров (полукадров), превосходящей частоту Найквиста для выходного шума. *< добавлено при решении задачи*

При значительном превышении полезных сигналов над шумами на входах КД вторым членом в квадратных скобках числителя в (1.9.4) можно пренебречь, что дает в итоге

$$\frac{\sigma_{\text{вых}}^2}{k_0^2} \approx \frac{2\delta F_{\text{нч}} \int_0^\infty s(F) \nu(F) \Phi_1^2(F) \Phi_2^2(F) dF}{r_{\text{вх}} [2\pi \int_0^\infty F s(F) \Phi_1(F) \Phi_2(F) dF]^2}. \quad (1.9.5)$$

Чтобы оценить порядок величины выходного шума в телевизионном КД при характерных параметрах сигналов, шумов и фильтров в трактах перед коррелятором, пренебрежем неравномерностями СП сигналов и шумов, полагая, что эффективная ширина частотной функции $\Phi_1^2(F) \Phi_2^2(F)$ много меньше, чем произведение $s(F) \nu(F)$.

Нормированной СП сигнала и квадратам АЧХ каналов придадим следующую форму:

$$s(F) = \frac{F_{\text{в}}^2}{F_{\text{в}}^2 + F^2}, \quad \Phi_1^2(F) = \frac{F_{\text{кр}}^2}{F_{\text{кр}}^2 + F^2}, \quad \Phi_2^2(F) = \frac{F^2}{F_{\text{кр}}^2 + F^2}. \quad (1.9.6)$$

Напомним, что два последних выражения соответствуют паре взаимно ортогональных RC-фильтров, описанных в § 1.6.

Вычислив необходимые интегралы, получаем из (5) и (6)

$$\frac{\sigma_{\text{вых}}^2}{k_0^2} = \frac{\Delta F_{\text{нч}} (F_{\text{кр}} + F_{\text{в}})^2}{2\pi^3 r_{\text{вх}} F_{\text{в}}^4 F_{\text{кр}}}. \quad (1.9.7)$$

Используя результаты анализа КД с ортогональными RC-фильтрами из § 1.6, отыщем относительную величину эффективных шумовых флуктуаций на выходе коррелятора в долях от половины апертуры Δ_{max} ДХ, определяемой по монотонному участку приближенным равенством [см. (1.6.22)]

$$2\omega_{\text{кр}} \Delta_{\text{max}} \approx 1. \quad (1.9.8)$$

Подставив (1.9.8) в (1.9.7), получим оценку

$$\frac{1}{r_{\text{вых}}} = \frac{\sigma_{\text{вых}}}{k_0 \Delta_{\text{max}}} \approx \frac{2}{\pi F_{\text{в}}} \sqrt{\frac{\delta F_{\text{нч}} F_{\text{кр}}}{r_{\text{вх}}}}. \quad (1.9.9)$$

В этой записи учтено, что $F_{\text{кр}} \ll F_{\text{в}}$, хотя при необходимости это допущение можно снять.

Подстановка характерных для ТВ систем условий

$$r_{\text{вх}} = 100, \quad F_{\text{в}} = 5 \cdot 10^6 \text{ Гц}, \quad F_{\text{кр}} = 5 \cdot 10^5 \text{ Гц}, \quad \delta F_{\text{нч}} = 5 \text{ Гц}$$

дает после вычислений

$$\frac{1}{r_{\text{вых}}} \approx 2 \cdot 10^{-5}, \quad (\text{по интервалу } \Delta_{\text{max}} \text{ — по формуле})$$

что составляет величину, практически плохо поддающуюся измерению. Между тем, выбранное для примера отношение сигнал/шум (20 дБ) на входе хуже того, которое принято считать минимально необходимым для сохранения удовлетворительного качества стандартного ТВ изображения (25–30 дБ) при его визуальной регистрации. Полученный результат по порядку величины согласуется с упомянутой в § 1.1 оценкой информативности ТВ сигнала, которая может служить мерой достижимого улучшения отношения сигнал/шум в телевизионном КД.

Существует принципиальная возможность минимизации относительных шумов на выходе КД путем специального согласования частотных характеристик фильтров в его каналах с формой СП сигнала и шума. Полученное таким образом решение будет представлять некоторое обобщение оптимального винеровского фильтра, который, не являясь адаптивным, требует для своей реализации априорных данных о спектрах входных процессов. Применение оптимальных фильтров такого типа может оказаться полезным в ОЭСС, работающих без каналов визуализации по заранее заготовленным эталонам при малом пространственном разрешении датчиков изображений.

В ОЭСС на базе стандартного телевидения при наличии устройств визуализации текущих изображений корреляционный автомат слежения, как правило, является более устойчивым звеном, нежели система "ТВ камера + человек", обеспечивающая первоначальную ориентацию и выбор изображения для последующего отслеживания. При разработке таких ОЭСС наибольшее внимание уделяется устранению внутренних технических помех, ибо аддитивные выходные флуктуации, как видно из приведенного примера, оказываются пренебрежимо малыми и без принятия мер по оптимизации частотных характеристик.

Влияние неравномерностей фона и чувствительности по полю зрения в каналах на характеристики одномерного КД изображений. Рассматривая влияние детерминируемых мешающих факторов, необходимо оценивать прежде всего смещение нуля ДХ и ее возможную деформацию. Согласно изложенным в § 1.3 общим принципам действия одномерного КД в его каналах производится преобразование входных сигналов, обеспечивающие ортогонализацию их ВКФ. Математически эта операция для линейного КД выражается формулой (1.6.7), куда входит нечетная функция сдвига $W_{12}(\Delta_r)$, которая представляет собой ВКФ импульсных реакций фильтров в каналах КД.

Так как свертка четной ВКФ входных сигналов с нечетной функцией W_{12} образует нечетную зависимость, действие факторов Q_1 , Q_2 , Q_3 и Q_4 опасно прежде всего в той мере, в какой они могут нарушить четность входной ВКФ. Это позволяет утверждать, что одинаковые детерминируемые помехи в обоих каналах КД не вызывают смещения нуля ДХ в силу полной идентичности сравниваемых процессов и, следовательно, четности их ВКФ. Аддитивные шумы $n_{1,2}$ в этих рассуждениях не учитываются. Отсюда следует важный для практики вывод: если КД используется в составе ОЭС с ЗУ, которое не вносит дополнительных аддитивных и мультипликативных детерминируемых помех, а эталон записывается в него через тот же тракт, который служит для формирования текущих изображений, то мешающие факторы названного типа не вызывают смещения нуля ДХ.

В случае, отличном от этого, ВКФ входных сигналов КД искажается за счет добавления к ней функции $\delta K_{12}(\Delta)$ вида

$$\delta K_{12} = (\check{Q}_1 + \check{Q}_3 \check{f}^\Delta) \otimes (Q_a + Q_m f), \quad (1.9.10)$$

где

$$Q_a = Q_1 - Q_2; \quad Q_m = Q_3 - Q_4.$$

Выражение (1.9.10) находится как разность ВКФ искаженных и неискаженных детерминированными помехами сигналов. Свертка в этом уравнении может при необходимости интерпретироваться и как двумерная, и даже как обобщенная. Для оценки смещений нуля ДХ одномерного КД ее следует, очевидно, считать одномерной, полагая неучтенные компоненты сдвига нулевыми.

Из (1.9.10) видно, что поправка $\delta K_{12}(\Delta_f)$ в общем случае отлична от нуля и может содержать как четную, так и нечетную часть, поскольку она зависит от конкретных реализаций сигналов f и f^Δ . Действие детерминируемых помех на положение нуля ДХ исключается, если $Q_a = 0$ и $Q_m = 0$. Не исключено, что в случайных множествах f и f^Δ могут встретиться реализации, обеспечивающие четность $\delta K_{12}(\Delta_f)$ и при невыполнении указанных условий, однако практическая ценность этого факта невелика.

Возможность возникновения случайных сдвигов нуля ДХ из-за мультипликативных и аддитивных неравномерностей в ЗУ служит одной из причин повышенного интереса к цифровой памяти в корреляционных ОЭС. Такая память позволяет полностью избавиться от мешающих факторов Q_a и Q_m .

Кроме смещения нуля ДХ, факторы Q_1 , Q_2 , Q_3 и Q_4 , даже при условии $Q_a = 0$ и $Q_m = 0$, могут в принципе вызывать недопустимые деформации ДХ, выражающиеся в инверсии ее знака. Это возможно в случаях, когда интенсивность используемых компонент полезного сигнала меньше аналогичных спектральных составляющих во вкладах, создаваемых помехами. Такая опасность реальна в ОЭС с видиконными ТВ камерами в случае выделения из сигналов частотных компонент низших порядков и при одновременном условии недостаточной яркости входных изображений.

Причиной тому может служить типичная для видиконов неравномерность фона и чувствительности по полю, характеризующаяся убыванием этих показателей к краям фоточувствительной мишени. Эффект "опрокидывания" знака ДХ в КД иногда наблюдается при настройке реальных корреляционных ОЭС. Это, как правило, указывает на наличие интенсивных синфазных помех из-за схемных дефектов в каналах КД при недостаточном уровне входных сигналов, поступающих на коррелятор.

В телевизионной технике известны [111] и применяются различные методы компенсации неравномерностей в видеосигналах. Они достаточно сложны и чаще всего использовались в стационарных аналоговых устройствах. С внедрением цифровой техники и матричных фотоприемников, которым свойственны специфические неравномерности из-за разброса параметров отдельных элементов, системы компенсации детерминированных видеопомех стали получать все более широкое применение.

Замечание на стр. об АКОР и разбросе

ГЛАВА 2

АНАЛИТИЧЕСКИЕ МЕТОДЫ ИССЛЕДОВАНИЯ ОПТИКО-ЭЛЕКТРОННЫХ СЛЕДЯЩИХ СИСТЕМ

§ 2.1. Формирование сигнала рассогласования в ОЭСС с непрерывным управлением

Типичным примером ОЭСС с непрерывным управлением являются координаторы пространственного сопровождения, работающие в ИК-диапазоне излучения [14, 54, 65, 67, 72, 131, 144, 171]. Эти ОЭСС при слежении за точечным объектом имеют динамическую ошибку порядка 1–2 угл. мин. Специфика таких систем, обладающих высокой точностью, состоит в применении сканирующих устройств при обработке оптической информации. В настоящее время существует большое количество разнообразных технических реализаций сканирующих устройств, осуществляющих просмотр поля лучистости по заранее заданному закону. В дальнейшем для определенности остановимся на двух структурных схемах, иллюстрирующих амплитудно-фазовый и частотный принципы выделения информации в пространственных угловых координатах объекта сопровождения из сигнала анализатора.

На рис. 2.1 показана структурная схема ОЭСС с одноплощадочным приемником и частотной модуляцией потока лучистой энергии.

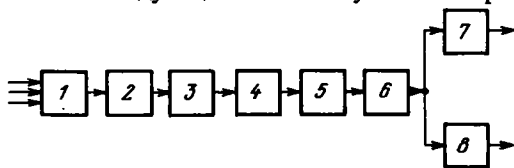


Рис. 2.1. Структурная схема оптико-электронной следящей системы с одноплощадочным приемником: 1 – фотоприемник с предусилителем и фильтром; 2, 4 – усилительные каскады с ограничением; 3 – усилитель несущей; 5 – частотный дискриминатор; 6 – усилитель огибающей; 7, 8 – фазовые детекторы

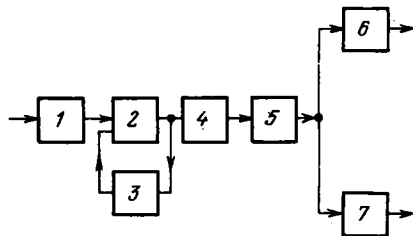
Анализаторы изображения с одноплощадочным приемником. Для создания частотной модуляции потока может применяться как вращение раstra, так и вращение изображения [14, 54]. В обоих случаях сигнал на выходе элемента 1 представляется в виде зависимости от времени t

$$u_{\text{фп}}(t) = U_0 \sin \{ [\omega_{\text{н}} + m(\Delta) \sin(\omega_0 t + \varphi)] t \}, \quad (2.1.1)$$

где U_0 – постоянная амплитуда импульсов; $\omega_{\text{н}}$ – частота несущей; ω_0 – частота огибающей; φ – фаза, соответствующая полярному углу изображения объекта в опорной системе координат; $m(\Delta)$ – коэффициент частотной модуляции, зависящий от модуля рассогласования Δ .

Каскад 2 обеспечивает работу усилителя 3 в линейном режиме при попадании в поле зрения ОЭСС источников с большой силой излучения. Каскад 4 ограничивает влияние внутренних шумов в тракте обработки информации. Об источниках этих шумов говорится в § 3.5. Частотный дискриминатор 5 благодаря резонансным свойствам своих контуров [33] имеет на выходе сигнал, амплитуда которого с точностью до постоянного коэффициента равна отклонению мгновенной частоты в выражении (2.1.1), т.е. величине $m(\Delta) \sin(\omega_0 t + \varphi)$. Этот сигнал в свою очередь поступает на

Рис. 2.2. Структурная схема оптико-электронной следящей системы с амплитудно-фазовой модуляцией: 1 — анализатор изображения, включающий фотоприемник и предусилитель; 2 — усилитель несущей; 3 — каскад автоматической регулировки усиления; 4 — амплитудный детектор; 5 — усилитель огибающей; 6, 7 — фазовые детекторы



вход усилителя огибающей 5, который представляет собой фильтр, настроенный на частоту ω_0 с шириной полосы $\Delta\omega$. Разложение сигнала управления по двум каналам осуществляется фазовыми детекторами 6 и 7 при наличии на них напряжения, поступающего с генератора опорных импульсов, сфазированного со сканирующим устройством.

Заметим, что растр здесь осуществляет не только частотную модуляцию из-за конечного размера изображения излучающего источника, которое попадает на элементы различной ширины. При затемнении части дифракционного пятна сигнал будет уменьшаться также по амплитуде.

На рис. 2.2 показана структурная схема ОЭСС с амплитудно-фазовой модуляцией. Анализаторы оптического изображения содержат растры, один из которых рассмотрен в § 3.2. Шахматная структура центральной части растра улучшает пространственную селекцию сопровождаемого объекта на фоне естественных помех [73]. Периферийная часть растра заполнена радиально-щелевыми элементами, размеры которых соизмеримы с диаметром изображения точечного источника.

Сигнал, снимаемый с элемента 1, в этом случае записывается в виде [51]

$$u_{\text{фп}}(t) = U_0 [1 + m(\Delta, \varphi) \cos(\omega_0 t + \varphi)] \cos \omega_n t, \quad (2.1.2.)$$

где U_0 — амплитуда немодулированного сигнала; $m(\Delta, \varphi)$ — коэффициент амплитудной модуляции, зависящий от модуля рассогласования и полярного угла φ . Остальные обозначения совпадают с (2.1.1).

Зависимость $m(\Delta, \varphi)$ с точностью до постоянного коэффициента представляет собой *пеленгационную характеристику*¹⁾ ОЭСС. Ее вычисление обычными математическими приемами наталкивается на значительные трудности. Поэтому основным способом вычисления характеристики $m(\Delta, \varphi)$ служит применение быстродействующей ЭВМ (см. главу 3). Огра-

¹⁾ Пеленгационная характеристика определяется свойствами дискриминаторов, рассмотренных в главе 1.

нимся здесь замечанием, что вид функции $m(\Delta, \varphi)$ оказывает существенное влияние на динамику процесса слежения и, в частности, на периодическое движение следящей системы. Подробное исследование этого вопроса помещено в § 2.4.

Резонансные усилители. В схемах, приведенных на рис. 2.1 и 2.2, присутствуют усилители, настроенные на определенную частоту ω_n или ω_0 и имеющие ограниченную полосу пропускания. Этому требованию с достаточной точностью удовлетворяет элемент с передаточной функцией вида

$$W(s) = \frac{2k\xi\omega_p s}{s^2 + 2\xi\omega_p s + \omega_p^2}. \quad (2.1.3)$$

Здесь k — коэффициент усиления каскада; ω_p — частота настройки; ξ — коэффициент демпфирования, определяющий ширину полосы пропускания.

Если условно принять за границу полосы пропускания частоту, при которой $|W(j\omega)| = 0,1 |W_{\max}|$, то ширину полосы можно вычислить по формуле, следующей из (2.1.3):

$$\frac{\Delta\omega}{\omega_p} = 2(\sqrt{1 + 99\xi^2} - 1 - \xi\sqrt{99}).$$

Так, при $\xi = 0,10$ $\Delta\omega/\omega_p = 2 \cdot 0,5843$,

при $\xi = 0,05$ $\Delta\omega/\omega_p = 2 \cdot 0,3806$,

при $\xi = 0,01$ $\Delta\omega/\omega_p = 2 \cdot 0,0946$.

Сигнал, определяемый выражением (2.1.2), можно записать в другой форме:

$$u_{\text{фп}}(t) = U_0 \cos \omega_n t + \frac{U_0 m(\Delta, \varphi)}{2} \cos [(\omega_n + \omega_0) t + \varphi] + \\ + \frac{U_0 m(\Delta, \varphi)}{2} \cos [(\omega_n - \omega_0) t + \varphi]. \quad (2.1.4)$$

При $U_0 = \text{const}$ и $m(\Delta, \varphi) = M = \text{const}$ спектральная плотность сигнала (2.1.4) выражается линейной комбинацией δ -функций

$$S_{\text{фп}}(\omega) = \pi U_0 \left[\delta(\omega - \omega_n) + \frac{M}{2} \delta(\omega - \omega_n - \omega_0) + \frac{M}{2} \delta(\omega - \omega_n + \omega_0) \right]. \quad (2.1.5)$$

При меняющихся во времени $U_0 = U_0(t)$ и $\Delta = \Delta(t)$ спектр будет иметь более "размытый" характер относительно несущей ω_n и боковых $\omega_n \pm \omega_0$.

При медленноменяющихся функциях $U_0(t)$ и $m[\Delta(t), \varphi(t)]$ сигнал на выходе усилителя несущей получается из выражений (2.1.3) и (2.1.4):

$$u_n(t) = U_0 k \cos \omega_0 t + \frac{1}{2} U_0 m(\Delta, \varphi) \lambda k \cos [(\omega_n + \omega_0) t + \varphi - \mu] + \\ + \frac{1}{2} U_0 m(\Delta, \varphi) \lambda k \cos [(\omega_n - \omega_0) t + \varphi + \mu]. \quad (2.1.6)$$

Здесь ввиду симметрии¹⁾ передаточной функции (2.1.3)

$$\lambda = \frac{2\xi\omega_n(\omega_n + \omega_0)}{(\omega_n + \omega_0)^2 + 2\xi\omega_0(\omega_n + \omega_0) + \omega_n^2},$$

$$\mu = \operatorname{arctg} \frac{2\xi\omega_n(\omega_n - \omega_0)}{\omega_n^2 - (\omega_n - \omega_0)^2} + \frac{\pi}{2}.$$

Таким образом, выражение (2.1.6) можно записать в более компактной форме

$$u_n(t) = U_0 k [1 + \lambda m(\Delta, \varphi) \cos(\omega_0 t + \varphi - \mu)] \cos \omega_n t. \quad (2.1.7)$$

Из (2.1.7) видно, что усилитель несущей вносит постоянный сдвиг по фазе μ в информацию о полярном угле сопровождаемого объекта. Этот сдвиг можно в дальнейшем компенсировать за счет соответствующего сдвига фазы генератора опорных импульсов при настройке ОЭСС.

Амплитудные детекторы. Рассмотрим математическое описание детектора с коэффициентом преобразования k_d .

В случае квадратичного детектирования сигнал на выходе детектора равен

$$u_d(t) = \frac{1}{2} k_d k^2 U_0^2 \left[1 + 2m\lambda \cos(\omega_0 t + \varphi - \mu) + \frac{\lambda^2 m^2}{2} + \frac{\lambda^2 m^2}{2} \cos 2(\omega_0 t + \varphi - \mu) \right] (1 + \cos 2\omega_n t). \quad (2.1.8)$$

При использовании линейного детектора аналогичный сигнал имеет вид

$$u_d(t) = k_d k U_0 |1 + \lambda m(\Delta, \varphi) \cos(\omega_0 t + \varphi - \mu)|. \quad (2.1.9)$$

Модуляция лучистого потока такова, что $|m(\Delta, \varphi)| \leq 1$, и в первом сомножителе (2.1.9) знак модуля можно снять. Второй сомножитель имеет разложение вида

$$|\cos x| = \frac{2}{\pi} + \sum_{k=1}^{\infty} a_k \cos kx.$$

В случае квадратичного детектирования в выражении (2.1.9) следует учесть только слагаемые с частотой ω_0 , так как постоянная составляющая и более высокие частоты $2\omega_0$ и ω_n не пропускаются усилителем огибающей. Передаточная функция усилителя огибающей описывается формулой (2.1.3), в которой следует положить $k = k_{y.o}$ и $\omega_p = \omega_0$. На выходе усилителя огибающей, следовательно, получаем сигнал вида

$$u_{y.o}(t) = k_d k^2 k_{y.o} U_0^2 \lambda m(\Delta, \varphi) \cos(\omega_0 t + \varphi - \mu). \quad (2.1.10)$$

При линейном детектировании по этой же причине

$$u_{y.o}(t) = k_d k k_{y.o} \frac{2 U_0}{\pi} \lambda m(\Delta, \varphi) \cos(\omega_0 t + \varphi - \mu). \quad (2.1.11)$$

¹⁾ Нетрудно показать, что график частотной характеристики $W(j\omega)$ в логарифмических координатах симметричен относительно частоты $\omega = \omega_p$.

Сравнение выражений (2.1.10) и (2.1.11) показывает, что они отличаются только постоянным множителем.

Фазовые детекторы. Автоматическая регулировка усиления. Дальнейшее преобразование сигнала состоит в его разложении по каналам управления с помощью фазовых детекторов (см. рис. 2.1 и 2.2). Без потери общности можно записать выражение для сигнала управления по каналам в форме

$$u_1(t) = U_{\max} m(\Delta, \varphi) \cos(\varphi + \epsilon), \quad (2.1.12)$$

$$u_2(t) = U_{\max} m(\Delta, \varphi) \sin(\varphi + \epsilon).$$

Значение амплитуды $U_{\max}$ зависит от вида схемы (способа модуляции) и включает в себя произведение коэффициентов преобразования (усиления) отдельных каскадов. Отметим, что в выражениях (2.1.12) угол расфазировки

$$\epsilon = \mu_{\text{оп}} - \mu,$$

где $\mu_{\text{оп}}$ — сдвиг фазы генератора опорных импульсов. Очевидно, что $\epsilon = 0$ при $\mu_{\text{оп}} = \mu$, и информация об угле φ будет поступать на исполнительный элемент без искажений. Однако в реальных следящих устройствах $\epsilon \neq 0$. Кроме обычных технологических допусков при настройке ОЭСС, на возникновение угла расфазировки влияет тот факт, что фазовый сдвиг пачек импульсов при модуляции потока не одинаков по всему полю раstra. Наличие угла расфазировки является весьма существенным при анализе возможных периодических режимов в ОЭСС.

На схеме, изображенной на рис. 2.2, имеется каскад АРУ. Мы не будем касаться математического описания этого устройства в этом параграфе,

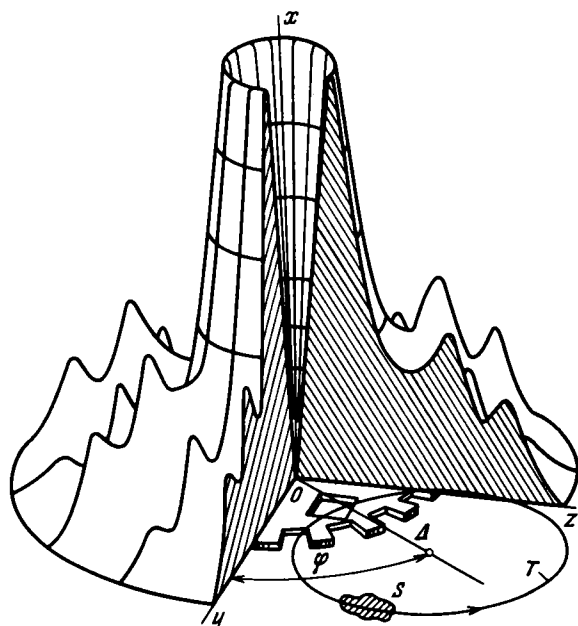


Рис. 2.3. Модуль дискриминационной характеристики: S — изображение источника; T — траектория переноса изображения

поскольку ближайшей задачей будет анализ процессов слежения при постоянных параметрах элементов ОЭСС. В первом приближении можно полагать, что работа каскада АРУ не зависит от рассогласования Δ , а связана лишь с уровнем облученности входного зрачка ОЭСС, который не меняется при слежении за неподвижным источником излучения.

Как уже упоминалось во введении (а также см. § 3.2), зависимость $m(\Delta, \varphi)$ наиболее простым путем получается на ЭВМ и поэтому в дальнейшем задается числовым массивом, который можно представить в наглядном графическом виде (рис. 2.3).

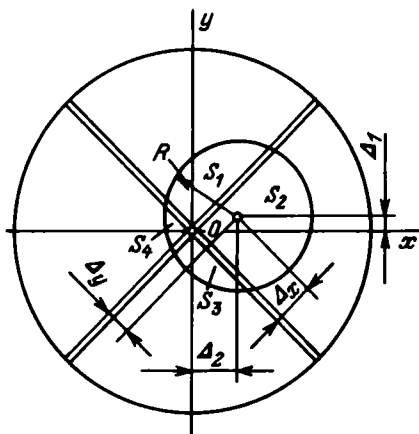
§ 2.2. Формирование сигнала рассогласования в ОЭСС с импульсным управлением

ОЭСС с импульсным управлением встречаются главным образом при использовании отраженного излучения какого-либо подсвечивающего устройства. В связи с необходимостью создания достаточно мощного излучения для сопровождения удаленных объектов в качестве подсвечивающего устройства применяются лазеры, работающие в импульсном режиме с достаточно низкой частотой посылок.

Это обстоятельство определяет схему построения импульсной ОЭСС, которая не имеет сканирующих устройств, применяемых в ОЭСС с непрерывным управлением. Для формирования сигнала ошибки в этом случае используется многоплощадочный приемник, одна из разновидностей которого показана на рис. 2.4.

Анализаторы изображения с многоплощадочным приемником. Для определенности здесь рассматривается круглое изображение объекта сопровождения. Оно, в частности, соответствует изображению бесконечно

Рис. 2.4. Четырехплощадочный приемник оптико-электронной следящей системы с импульсным управлением



удаленного (точечного) источника излучения. Сигнал ошибки по каналам управления формируется в простейшем случае как разность напряжений, поступающих с противоположащих площадок приемника. В приведенном на рис. 2.4 случае вычисление соответствующей пеленгационной характеристики не встречает затруднений при равномерном распределении облученности внутри пятна изображения. Действительно, в соответствии с

обозначениями на рис. 2.4 запишем выражение для вычисления сигналов управления для малых отклонений изображения, при которых центр перекрестия не выходит за пределы приемника:

$$\begin{aligned} S_1 + S_4 &= R^2 \arcsin \sqrt{1 - \left(\frac{\Delta_y}{R}\right)^2} - \Delta_y \sqrt{R^2 - \Delta_y^2}, \\ S_1 + S_2 &= R^2 \arcsin \sqrt{1 - \left(\frac{\Delta_x}{R}\right)^2} - \Delta_x \sqrt{R^2 - \Delta_x^2}, \\ S_3 + S_4 &= \pi R^2 - R^2 \arcsin \sqrt{1 - \left(\frac{\Delta_x^2}{R}\right)} + \Delta_x \sqrt{R^2 - \Delta_x^2}. \end{aligned} \quad (2.2.1)$$

Отсюда сразу следует

$$\begin{aligned} S_3 - S_1 &= u_x(\Delta_x, \Delta_y) = \pi R^2 - R^2 \left[\arcsin \sqrt{1 - \left(\frac{\Delta_x}{R}\right)^2} + \right. \\ &+ \left. \arcsin \sqrt{1 - \left(\frac{\Delta_y}{R}\right)^2} \right] + \Delta_x \sqrt{R^2 - \Delta_x^2} + \Delta_y \sqrt{R^2 - \Delta_y^2}, \\ S_4 - S_2 &= u_y(\Delta_x, \Delta_y) = R^2 \left[\arcsin \sqrt{1 - \left(\frac{\Delta_y}{R}\right)^2} - \arcsin \sqrt{1 - \left(\frac{\Delta_x}{R}\right)^2} \right] + \\ &+ \Delta_x \sqrt{R^2 - \Delta_x^2} - \Delta_y \sqrt{R^2 - \Delta_y^2}. \end{aligned} \quad (2.2.2)$$

К этим выражениям следует добавить формулы для пересчета компонент вектора ошибки $\underline{\Delta}$

$$\Delta_x = \frac{1}{\sqrt{2}}(\Delta_1 + \Delta_2), \quad \Delta_y = \frac{1}{\sqrt{2}}(-\Delta_1 + \Delta_2). \quad (2.2.3)$$

Подставив (2.2.3) в (2.2.2), получим сигналы управления как функции компонент вектора рассогласования $u_x(\Delta_1, \Delta_2)$ и $u_y(\Delta_1, \Delta_2)$.

Как видно, в случае многоплощадочного приемника пеленгационная характеристика выражается конечными формулами сравнительно простой структуры для математического моделирования. Следует, однако, заметить, что формулы (2.2.2) получены при допущении о равномерном распределении облученности внутри круглого пятна изображения. В действительности эти условия не выполняются. Изображение может иметь произвольную форму, определяемую индикатрисой отраженного излучения, и не обязательно заполняет односвязную область, так как поверхность отражения может давать несколько отдельных блоков.

Далее, если не пользоваться допущением о равномерном распределении облученности, то электрический сигнал ΔU на выходе каждой площадки фотоприемника при условии однородности ее свойств по полю зрения вычисляется интегрированием [76]

$$\Delta U = k_{\text{фп}} \iint_{(S)} E(x, y) dx dy, \quad (2.2.4)$$

где $E(x, y)$ — облученность фотоприемника в точке с координатами (x, y) ;
 S — площадь засвеченной области.

Крутизна преобразования $k_{\Phi\Pi}$ определяется соотношением

$$k_{\Phi\Pi} = \frac{\Delta U}{\Phi} = \frac{UF\tau\eta}{N} \frac{R_{\Phi} R_{\Pi}}{(R_{\Phi} - R_{\Pi})^2}.$$

Здесь Φ — поток лучистой энергии в засвеченной области; U — напряжение питания; $R_{\Phi\Pi}$ — темновое сопротивление (при отсутствии облученности) фотоприемника; R_{Π} — сопротивление нагрузки; F — вся приемная площадь; η — квантовая эффективность; τ — время существования носителей зарядов в зоне проводимости; N — общее число элементарных носителей.

Перечисленные факторы заставляют при математическом моделировании отказаться от использования формул (2.2.2) для вычисления сигналов управления и следовать методике моделирования пеленгационной характеристики, изложенной в § 3.2. Для исследования периодических процессов в ОЭСС с импульсным управлением достаточно рассмотреть пеленгационную характеристику при малых отклонениях от положения равновесия при сопровождении неподвижного источника. Воспользуемся разложениями

$$u_x(\Delta_x, \Delta_y) = u_x(0, 0) + \frac{\partial u_x(0, 0)}{\partial \Delta_x} \Delta_x + \frac{\partial u_x(0, 0)}{\partial \Delta_y} \Delta_y + o(\Delta_x, \Delta_y), \quad (2.2.5)$$

$$u_y(\Delta_x, \Delta_y) = u_y(0, 0) + \frac{\partial u_y(0, 0)}{\partial \Delta_x} \Delta_x + \frac{\partial u_y(0, 0)}{\partial \Delta_y} \Delta_y + o(\Delta_x, \Delta_y).$$

Дифференцируя (2.2.2) и подставляя (2.2.3), находим линеаризованную форму пеленгационной характеристики

$$u_x = 2R\sqrt{2}\Delta_2 + o(\Delta_1, \Delta_2), \quad (2.2.6)$$

$$u_y = 2R\sqrt{2}\Delta_1 + o(\Delta_1, \Delta_2).$$

Мощность импульсов, поступающих на фотоприемник, в зависимости от расстояния до подсвечиваемого предмета может меняться в широких

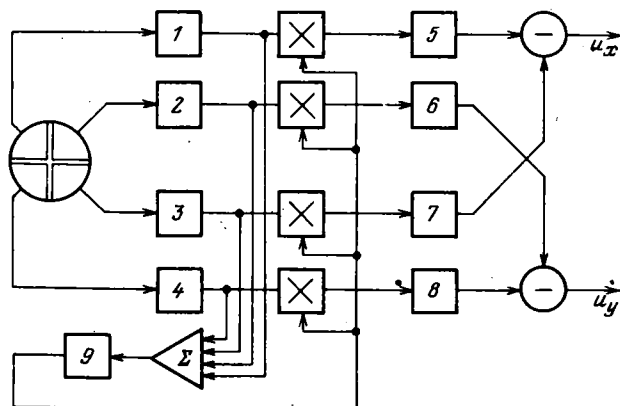


Рис. 2.5. Электронно-усилительный тракт оптико-электронной следящей системы с импульсным управлением: 1, 2, 3, 4 — пороговые устройства; 5, 6, 7, 8 — усилители; 9 — элемент с регулировочной характеристикой

пределах. Для предохранения элементов электронно-усилительного тракта от насыщения в схеме применен каскад АРУ (рис. 2.5).

Сигнал на входе АРУ не должен зависеть от рассогласования, поэтому он суммируется со всех площадок приемника и поступает на элемент 9, обладающий регулировочной характеристикой [124]. Пороговые устройства 1–4 отрезают внутренние шумы фотоприемника. Для усиления сигналы поступают на каскады 5–8 и затем на схемы вычитания, образующие управляющие напряжения u_x и u_y . Далее управляющие напряжения поступают на модулятор, работающий по схеме либо амплитудно-импульсной, либо широтно-импульсной модуляции [135], и на каналы исполнительного элемента ОЭСС.

§ 2.3. Математические модели исполнительных элементов ОЭСС

Использование электрических микродвигателей. Мы рассмотрим здесь особенности различных исполнительных элементов, применяемых в ОЭСС¹⁾. Наиболее распространенным видом исполнительного элемента приборного типа служит малогабаритный электрический двигатель. Достоинством этого элемента является сравнительная простота изготовления, не требующая повышенной точности механической обработки. Они надежны в эксплуатации, обладают вибро- и удароустойчивостью. Динамические свойства электродвигателя описываются двумя зависимостями:

$$\dot{\eta} = \dot{\eta}_0 - kM, \quad \dot{\eta}_0 = k_{дв} U. \quad (2.3.1)$$

Первая зависимость (2.3.1) называется *внешней характеристикой* электродвигателя, которая связывает значение угловой скорости вращения его ротора $\dot{\eta}$ с моментом нагрузки M . Параметр $\dot{\eta}_0$ — это угловая скорость вращения ненагруженного двигателя. Величина k определяет жесткость характеристики. Видно, что с увеличением параметра k внешняя характеристика становится мягче, так как скорость $\dot{\eta}$ сильнее падает с ростом момента M . Вторая зависимость (2.3.1) называется *регулировочной характеристикой*. Она связывает параметр $\dot{\eta}_0$ с величиной управляющего сигнала U посредством крутизны характеристики $k_{дв}$.

Кроме зависимостей (2.3.1) установившегося движения, запишем уравнение динамики

$$I\ddot{\eta} = M - M_c, \quad (2.3.2)$$

где I — момент инерции всех движущихся масс (в основном оптического блока, установленного в кардановом подвесе) ОЭСС, приведенный к оси вращения ротора; M_c — момент сопротивления движению. Если пренебречь величиной этого момента по сравнению с движущим моментом M , то из (2.3.1) и (2.3.2) вытекает соотношение

$$Ik\ddot{\eta} + \dot{\eta} = k_{дв} U, \quad (2.3.3)$$

связывающее выходную координату η исполнительного элемента с входным сигналом U . Произведение Ik имеет размерность времени и называется

¹⁾ Нередко ОЭСС называют также следящим координатором.

постоянной времени двигателя. Для ее уменьшения стремятся максимально уменьшить момент инерции I и, в частности, облегчить конструкцию ротора.

Двигатель питается от усилителя мощности, который описывается уравнением

$$\tau \dot{U} + U = k_y \Delta, \quad (2.3.4)$$

где τ — постоянная времени усилителя; k_y — его коэффициент усиления. На вход усилителя подается напряжение Δ , которое мы будем считать пропорциональным угловой ошибке сопровождения, сохранив то же обозначение. Эта ошибка равна разности между углом наклона оптической оси ОЭСС и углом наклона линии визирования на сопровождаемый объект (рис. 2.6):

$$\Delta = \eta_0 - \eta. \quad (2.3.5)$$

Последнее соотношение справедливо в линейной зоне пеленгационной характеристики и представляет собой уравнение обратной связи привода.

Объединяя (2.3.3), (2.3.4) и (2.3.5), получаем уравнение замкнутой ОЭСС для малых отклонений

$$\tau I k \eta^{(3)} + (\tau + I k) \eta^{(2)} + \eta^{(1)} + k_{дв} k_y \eta = k_{дв} k_y \eta_0. \quad (2.3.6)$$

Этому уравнению соответствует свободное движение ОЭСС, которое получается из (2.3.6) при $\eta_0 = 0$. Условием устойчивости этого движения

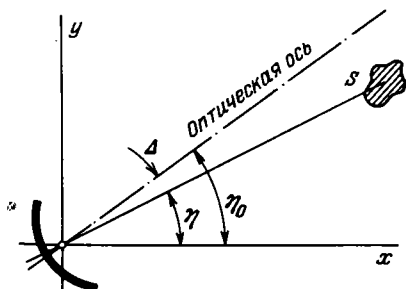


Рис. 2.6. Геометрические соотношения при сопровождении подвижного объекта S

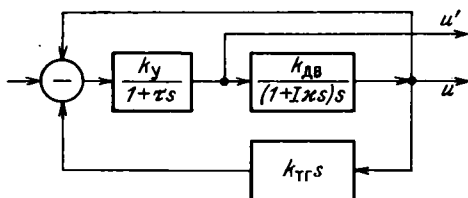


Рис. 2.7. Коррекция привода оптико-электронной следящей системы: u и u' — управляющие сигналы

служит отрицательность действительных частей корней λ , характеристического уравнения

$$\tau I k \lambda^3 + (\tau + I k) \lambda^2 + \lambda + k_{дв} k_y = 0. \quad (2.3.7)$$

Наиболее просто в данном случае условие устойчивости проверяется по критерию Гурвица [95], по которому

$$\begin{vmatrix} I k + \tau & k_{дв} k_y & 0 \\ \tau I k & 1 & 0 \\ 0 & I k + \tau & k_{дв} k_y \end{vmatrix} > 0 \quad (2.3.8)$$

и должны быть положительными оба диагональных минора определителя

(2.3.8). Таким образом, условием устойчивости свободного движения ОЭСС служит неравенство

$$k_{дв}k_y < \frac{1}{I_k} + \frac{1}{\tau}, \quad (2.3.9)$$

которое ограничивает сверху добротность замкнутого контура ОЭСС $\mu = k_{дв}k_y$. Этим условием ограничивается также и точность сопровождения перемещающегося в пространстве объекта.

Действительно, рассмотрим случай, когда линия визирования вращается в пространстве с постоянной угловой скоростью Ω , так что $\eta_0(t) = \Omega t$. Переходя в (2.3.6) к изображениям по Лапласу при нулевых начальных условиях, имеем

$$[\tau I k s^3 + (\tau + I k)s^2 + s + \mu] H(s) = \mu H_0(s). \quad (2.3.10)$$

Так как согласно (2.3.5) $\mathcal{L}\{\Delta(t)\} = H_0(s) - H(s)$, то

$$\begin{aligned} \mathcal{L}\{\Delta(t)\} &= \frac{\tau I k s^3 + (\tau + I k)s^2 + s}{\tau I k s^3 + (\tau + I k)s^2 + s + \mu} H_0(s) = \\ &= \frac{\tau I k s^2 + (\tau + I k)s + 1}{\tau I k s^3 + (\tau + I k)s^2 + s + \mu} \frac{\Omega}{s^2}. \end{aligned} \quad (2.3.11)$$

Воспользовавшись теоремой о конечном значении [38, 40], найдем установившееся значение ошибки сопровождения

$$\lim_{t \rightarrow \infty} \Delta(t) = \lim_{s \rightarrow 0} s \mathcal{L}\{\Delta(t)\} = \frac{\Omega}{\mu}. \quad (2.3.12)$$

Условия (2.3.9) и (2.3.12) иллюстрируют обычные противоречия между требованиями повышения точности сопровождения и требованием сохранения его устойчивости.

Некоторое повышение добротности привода возможно за счет применения корректирующих обратных связей. Одним из простейших способов коррекции служит использование тахогенератора с передаточной функцией $k_{тг}s$ (рис. 2.7). При наличии тахогенератора характеристическое уравнение (2.3.7) примет вид

$$\tau I k \lambda^2 + (\tau + I k)\lambda^2 + (1 + \mu k_{тг})\lambda + \mu = 0 \quad (2.3.13)$$

и условие устойчивости (2.3.9) перейдет в неравенство

$$\mu < \frac{I k + \tau}{\tau I k - k_{тг}(I k + \tau)}. \quad (2.3.14)$$

Сравнивая (2.3.9) с (2.3.14), видим, что допустимая добротность контура ОЭСС повышается за счет средств коррекции.

Помимо добротности (точности сопровождения), динамические свойства ОЭСС характеризуются быстродействием. Если параметр $k_{тг}$ выбрать слишком большим, то это приведет к затянутому аperiodическому переходному процессу. Если же $k_{тг}$ слишком мал, то получается слабо затухающий колебательный процесс. Для того чтобы выбрать оптимальное значение параметра $k_{тг}$, следует воспользоваться каким-либо критерием,

характеризующим быстродействие ОЭСС, и минимизировать его. В качестве методики, позволяющей довести вычисления до конца в общем виде, можно воспользоваться интегральной квадратической оценкой

$$I_0 = \int_0^{\infty} \Delta^2(t) dt,$$

описанной А.А. Красовским в [89]. Можно пойти и более непосредственным путем, если рассмотреть в качестве критерия быстродействия время окончания переходного процесса T , удовлетворяющее условию

$$|\Delta(t)| < 0,05 \Delta(0) \text{ при } t > T. \quad (2.3.15)$$

Однако такой критерий уже не приводит к конечным выражениям и предполагает использование ЭВМ с процедурой поиска минимума, описанной в § 3.6, посвященном идентификации математических моделей.

Заметим здесь, что быстродействие (а следовательно, и полоса пропускания замкнутого контура) ОЭСС ограничивается сравнительно небольшим крутящим моментом. Гораздо большими возможностями в этом смысле обладает гидравлический объемный привод. Высокое быстродействие его достигается благодаря высокому давлению рабочей жидкости (до 2000 кПа). При этом снижаются вес и габариты исполнительного элемента. Математическое описание гидропривода по форме не отличается от уравнений вида (2.3.1).

Индикаторный гидропривод. Обратимся теперь к наиболее распространенному виду исполнительных элементов — к гироскопам, работающим в режиме управления ОЭСС. Гироскопические исполнительные элементы обладают значительными преимуществами перед электрическими и гидравлическими: помимо малой инерционности и высокой точности работы они обладают принципиальной возможностью развязки следящего координатора по отношению к подвижному основанию и, следовательно, свойством "запоминания" положения оптической оси прибора в инерциальном пространстве. Простейшей разновидностью исполнительных элементов этого типа является трехстепенный гироскоп (рис. 2.8). Конструктивно он выполняется обычно таким образом, что его ротор 1 несет на себе и один из элементов оптической части (например, рефлектор). Эллипсоид инерции ротора сплюснут по оси x . Вместе с внутренней рамкой 2 ротор может поворачиваться вокруг оси $O\eta$. Внутренняя рамка может вращаться вместе с наружной рамкой 3 вокруг оси $O\xi$, подшипники которой жестко связаны с подвижным основанием.

Ввиду симметрии эллипсоида инерции все оси, проходящие через точку O и перпендикулярные к оси Ox , являются главными. Момент инерции относительно любой из этих осей равен I . Момент инерции относительно оси Ox обозначим через I_p . При движении гироскопа точка O остается неподвижной в системе координат $Oxyz$, и поэтому любое движение его можно рассматривать как вращение с мгновенной угловой скоростью $\underline{\omega} = (r, \omega')$. Величину $H = I_p r$ называют *собственным кинетическим моментом ротора*.

Спроектируем вектор ω' на оси y и z правой связки $Oxyz$ (см. рис. 2.8). Тогда I_p и I_q будут проекциями вектора кинетического момента $\underline{\theta}$. Угловые скорости p и q будут компонентами угловой скорости трехгран-

Рассмотрим более подробно физическую природу моментов, действующих на гироскоп, управляемый по осям $O\xi$ и $O\eta$. Заметим, что если оси Oy и Oz трехгранника $Oxyz$ совместить с осями $O\xi$ и $O\eta$ карданова подвеса, то $r' = 0$. Кроме того,

$$M_y = M_y^{\text{упр}} - M_y^{\text{тр}}, \quad M_z = M_z^{\text{упр}} - M_z^{\text{тр}}. \quad (2.3.19)$$

Здесь $M_z^{\text{упр}}$ и $M_y^{\text{упр}}$ — моменты, развиваемые датчиками под действием управляющих сигналов с каналов фазовых детекторов (см. структурные схемы на рис. 2.1 и 2.2) либо с усилителей мощности после вычитания сигналов с площадок (см. рис. 2.5); $M_y^{\text{тр}}$ и $M_z^{\text{тр}}$ — моменты сил сопротивления в подшипниках осей вращения рамок.

Для простоты исследования в пределах линейных дифференциальных уравнений, позволяющего сделать общие качественные выводы об устойчивости процесса сопровождения, положим ¹⁾

$$M_y^{\text{тр}} = \xi p, \quad M_z^{\text{тр}} = \xi q \quad (2.3.20)$$

и, считая $r' = 0$, получим уравнения движения гироскопа в виде

$$\begin{aligned} I \frac{dp}{dt} + \xi p + Hq &= M_y^{\text{упр}}, \\ I \frac{dq}{dt} + \xi q - Hp &= M_z^{\text{упр}}, \\ \frac{dH}{dt} &= M_x. \end{aligned} \quad (2.3.21)$$

Из этих трех уравнений лишь два первых относятся к системе управления. Третье уравнение характеризует вращение ротора во время раскрутки гироскопом либо его торможение под действием сил трения в подшипниках рамки 2 (см. рис. 2.8).

Моменты управления в режиме слежения зависят от углов рассогласования. В соответствии с этим уравнения (2.3.21) дополняются условиями замыкания

$$\begin{aligned} M_z^{\text{упр}} &= -k_r \Delta_y = -k_r \int_0^t q(\tau) d\tau, \\ M_y^{\text{упр}} &= k_r \Delta_z = k_r \int_0^t p(\tau) d\tau. \end{aligned} \quad (2.3.22)$$

Представляется интересным вопрос о предельно допустимой добротности k_r . Для выяснения этого значения запишем характеристическое уравнение, вытекающее из (2.3.21) и (2.3.22):

$$\begin{vmatrix} (I\lambda + \xi)\lambda & -(H\lambda + k_r) \\ H\lambda + k_r & (I\lambda + \xi)\lambda \end{vmatrix} = 0 \quad (2.3.23)$$

¹⁾ В действительности моменты сил сопротивления описываются нелинейной зависимостью. В первом приближении можно пользоваться зависимостью вида $M_y^{\text{тр}} = M_{y0} \operatorname{sign} p$ и $M_z^{\text{тр}} = M_{z0} \operatorname{sign} q$ для более точного исследования по сравнению с (2.3.20).

или

$$I^2 \lambda^4 + 2I\xi \lambda^3 + (H^2 + \xi^2)\lambda + 2k_T H\lambda + k_T^2 = 0. \quad (2.3.24)$$

Составив определитель Гурвица четвертого порядка с коэффициентами полинома (2.3.24), получаем условие, ограничивающее добротность гироскопического привода:

$$k_T < \frac{H\xi}{I}. \quad (2.3.25)$$

Так как величина ξ очень невелика, то для повышения добротности привода ОЭСС следует увеличить отношение H/I . Это достигается за счет большой скорости вращения ротора (30000 – 60000 об/мин). В свою очередь это требует повышенной частоты питающего напряжения (500 – 1000 Гц).

Гирос привод с разгрузкой от внешних возмущений. Трехступенный гироскоп не способен развивать больших управляющих моментов. В то же время в ряде конструкций ОЭСС вращающаяся часть может обладать значительной массой (например, при тяжелых телеобъективах). Внешние перегрузки со стороны подвижного основания при несбалансированности подвеса отклоняют ось оптической системы, в результате чего в управляющем сигнале появляется ложная составляющая.

В настоящее время разработаны схемы гирос привода со специально предусмотренными мерами защиты против внешних силовых возмущений. Рассмотрим двухгироскопную схему гирос привода, которая называется *гиростабилизированной платформой* (рис. 2.9). Внутренняя рамка платформы 1 жестко связана с наружными рамками 2 и 3 гироскопов Γ_1 и Γ_2 . Под действием коррекционного момента, развиваемого датчиком D_1 , платформа прецессирует вокруг оси y_K . Момент датчика D_2 заставляет платформу прецессировать вокруг оси z_K . При этом перемещаются тяги T_y и T_z , связанные с оптическим блоком, который подвешен в собственном кардановом подвесе. Таким образом осуществляется сопровождение излучающего объекта. Возмущающие внешние моменты разворачивают

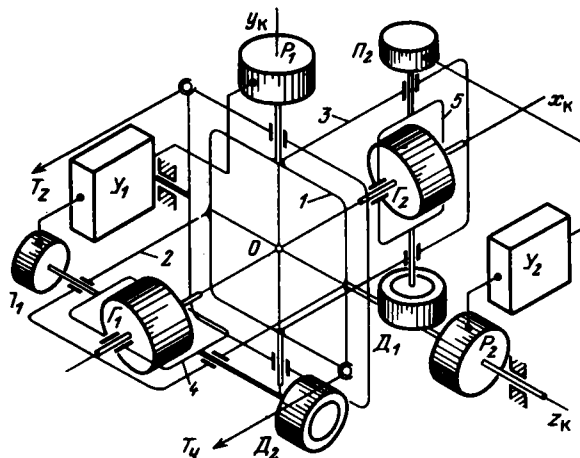


Рис. 2.9. Двухроторный гироскопический исполнительный элемент: 1 – внутренняя рамка; 2, 3 – наружные рамки; 4, 5 – внутренние рамки

гироплатформу. При этом гироскопы стремятся сохранить неизменным направление осей вращения в инерциальном пространстве, и на датчиках P_1 и P_2 углов поворота внутренних рамок 4 и 5 возникают сигналы, поступающие на усилители разгрузки Y_1 и Y_2 . Усиленные сигналы подаются на моментные датчики разгрузки P_1 и P_2 . В результате осуществляется компенсация внешних возмущений.

Работа исполнительных элементов гироскопического типа основана на принципе прецессии. При этом, как мы видели в уравнениях (2.3.21), в каналах управления появляются обратные связи. Из-за различия в знаках эти связи называются *антисимметричными* [51, 63]. В многороторных гироскопических приводах эти связи значительно ослаблены. В дальнейшем будем рассматривать общий случай гиروطвода силового типа с передаточной функцией вида

$$\frac{\varphi_z(s)}{M_z(s)} = \frac{\varphi_y(s)}{M_y(s)} = \frac{k_T}{(1 + 2\xi Ts + T^2 s^2)s} \quad (2.3.26)$$

Здесь $\varphi_z(s)$ и $\varphi_y(s)$ — изображения углов поворота оптического блока ОЭСС (под действием тяг T_z и T_y , см. рис. 2.9); $M_z(s)$ и $M_y(s)$ — изображения моментов датчиков D_1 и D_2 ; k_T — коэффициент усиления гиروطвода; ξ — коэффициент демпфирования; T — постоянная времени.

§ 2.4. Анализ периодических режимов в ОЭСС с непрерывным управлением

Современные ОЭСС представляют собой сложные технические устройства с многочисленными связями между каналами управления. Многие элементы описываются нелинейными характеристиками. Совокупность других элементов описывается линейными дифференциальными уравнениями высокого порядка. Основным методом анализа и конструирования ОЭСС служит математическое моделирование. В то же время для выявления закономерностей общего характера, определяющих динамику ОЭСС, для обзора пространства параметров элементов и возможностей выполнения требований по качеству сопровождения необходимо иметь в распоряжении аналитические методы. Анализ проводится обычно при некоторых упрощениях математической модели ОЭСС, не вносящих существенных изменений в качественную картину процесса слежения. Разработка достаточно простой модели, допускающей математическое исследование, представляет собой самостоятельную и не всегда выполнимую задачу. Поэтому часто задача исследования ограничивается созданием модели, характеристики которой совпадают с характеристиками реальной ОЭСС лишь при некоторых режимах работы.

В настоящем параграфе мы рассмотрим упрощенную модель ОЭСС с повышенной точностью слежения и с силовым гиروطводом, обладающим выраженными колебательными свойствами. Обозначим компоненты углового рассогласования $\underline{\Delta}$ через Δ_1 и Δ_2 :

$$\Delta_1 = \Delta \cos \varphi, \quad \Delta_2 = \Delta \sin \varphi. \quad (2.4.1)$$

Учитывая условия замыкания $\varphi_z = -\Delta_1$ и $\varphi_y = -\Delta_2$, из (2.3.26) получаем

$$\begin{aligned} T^2 \Delta_1^{(3)} + 2\xi T \Delta_1^{(2)} + \Delta_1^{(1)} &= -k_r M_z, \\ T^2 \Delta_2^{(3)} + 2\xi T \Delta_2^{(2)} + \Delta_2^{(1)} &= -k_r M_y. \end{aligned} \quad (2.4.2)$$

Моменты $M_z(t)$ и $M_y(t)$ в правой части (2.4.2) пропорциональны напряжениям $u_1(t)$ и $u_2(t)$ в каналах управления, которые находятся из выражений (2.1.12), т.е.

$$\begin{aligned} M_z(t) &= k_d u_1(t) = k_d U_{\max} m(\Delta, \varphi) \cos(\varphi + \epsilon), \\ M_y(t) &= k_d u_2(t) = k_d U_{\max} m(\Delta, \varphi) \sin(\varphi + \epsilon), \end{aligned} \quad (2.4.3)$$

где k_d — коэффициент преобразования моментного датчика.

Формальные условия возникновения кругового предельного цикла. Следуя [51], сделаем допущение о возможности представления функции $k_r k_d U_{\max} m(\Delta, \varphi)$ в виде произведения $F(\Delta)\Delta$. Это допущение основано на предположении о гармонических колебаниях, которые могут возникнуть в замкнутом контуре ОЭСС с непрерывным управлением. Такие колебания действительно наблюдаются при экспериментальных исследованиях и при математическом моделировании подобных систем. Поэтому в дальнейшем рассмотрим уравнения ОЭСС в виде

$$\begin{aligned} T^2 \Delta_1^{(3)} + 2\xi T \Delta_1^{(2)} + \Delta_1^{(1)} &= -F(\Delta) \Delta \cos(\varphi + \epsilon), \\ T^2 \Delta_2^{(3)} + 2\xi T \Delta_2^{(2)} + \Delta_2^{(1)} &= -F(\Delta) \Delta \sin(\varphi + \epsilon) \end{aligned} \quad (2.4.4)$$

или, учитывая обозначения (2.4.1), в виде

$$\begin{aligned} T^2 \Delta_1^{(3)} + 2\xi T \Delta_1^{(2)} + \Delta_1^{(1)} &= -F(\Delta) (\Delta_1 \cos \epsilon - \Delta_2 \sin \epsilon), \\ T^2 \Delta_2^{(3)} + 2\xi T \Delta_2^{(2)} + \Delta_2^{(1)} &= -F(\Delta) (\Delta_2 \cos \epsilon + \Delta_1 \sin \epsilon). \end{aligned} \quad (2.4.5)$$

При идеальных гармонических колебаниях

$$\Delta_1 = \Delta_0 \cos \omega_0 t, \quad \Delta_2 = \Delta_0 \sin \omega_0 t. \quad (2.4.6)$$

Здесь $\Delta_0 = \text{const}$ и $\omega_0 = \text{const}$ — амплитуда и частота этих колебаний. Заметим, что согласно (2.4.1) и (2.4.6)

$$\Delta = \sqrt{\Delta_1^2 + \Delta_2^2} = \Delta_0.$$

Поэтому в уравнениях (2.4.5) также и $F(\Delta) = F(\Delta_0) = \text{const}$.

Для формальной проверки возможности существования гармонических колебаний в ОЭСС подставим соотношения (2.4.6) в уравнения (2.4.5). При этом мы получим формальные тождества

$$\begin{aligned} \omega_0 \Delta_0 (T^2 \omega_0^2 - 1) \sin \omega_0 t - 2\xi T \omega_0^2 \Delta_0 \cos \omega_0 t &\equiv -F(\Delta_0) \Delta_0 \cos(\omega_0 t + \epsilon), \\ -\omega_0 \Delta_0 (T^2 \omega_0^2 - 1) \cos \omega_0 t - 2\xi T \omega_0^2 \Delta_0 \sin \omega_0 t &\equiv -F(\Delta_0) \Delta_0 \sin(\omega_0 t + \epsilon). \end{aligned} \quad (2.4.7)$$

Так как $\Delta_0 \neq 0$, то (2.4.7) можно сократить на Δ_0 и, умножая первое равенство на $\sin \omega_0 t$, а второе на $\cos \omega_0 t$ и вычитая, получить

$$\omega_0 (T^2 \omega_0^2 - 1) = F(\Delta_0) \sin \epsilon. \quad (2.4.8)$$

Аналогично, умножая первое соотношение на $\cos \omega_0 t$, второе на $\sin \omega_0 t$

и складывая, имеем

$$2\xi T\omega_0^2 = F(\Delta_0)\cos\epsilon. \quad (2.4.9)$$

Соотношения (2.4.8) и (2.4.9) образуют систему, из которой можно найти частоту ω_0 и амплитуду Δ_0 предполагаемого периодического движения. Действительно, при $\omega_0 \neq 0$, поделив (2.4.8) на (2.4.9), получаем

$$\frac{T^2\omega_0^2 - 1}{2\xi T\omega_0} = \operatorname{tg}\epsilon,$$

откуда находим

$$\omega_0 = \frac{1}{T} (\xi \operatorname{tg}\epsilon \pm \sqrt{\xi^2 \operatorname{tg}^2 \epsilon + 1}). \quad (2.4.10)$$

Амплитуду Δ_0 после этого можно найти либо из (2.4.8), либо из (2.4.9). Можно записать и общее выражение для определения $F(\Delta_0)$, откуда при заданной зависимости $F(\Delta)$ находится амплитуда Δ_0 . Например, умножая (2.4.8) на $\sin\epsilon$, а (2.4.9) на $\cos\epsilon$ и складывая, получаем

$$F(\Delta_0) = 2\xi T\omega_0^2 \cos\epsilon + (T^2\omega_0^2 - 1)\omega_0 \sin\epsilon. \quad (2.4.11)$$

Формулы (2.4.10) и (2.4.11) представляют собой искомые выражения для вычисления частоты и амплитуды автоколебаний. Остается выяснить, какие из значений этих параметров соответствуют реально наблюдаемым процессам.

Заметим, что устойчивое периодическое движение (автоколебания) может возникнуть лишь при неустойчивости положения равновесия, которому соответствует тривиальное решение системы (2.4.5) $\Delta_1(t) \equiv \Delta_2(t) \equiv 0$. Анализ устойчивости положения равновесия сводится к исследованию знаков корней характеристического уравнения системы (2.4.5). Наиболее просто это уравнение записывается, если в (2.4.5) перейти к изображениям. Сохраняя символы $\Delta_1(s) = L\{\Delta_1(t)\}$ и $\Delta_2(s) = L\{\Delta_2(t)\}$, запишем

$$\begin{vmatrix} [T^2 s^3 + 2\xi T s^2 + s + F(0)\cos\epsilon]\Delta_1(s) & -\Delta_2(s)F(0)\sin\epsilon \\ \Delta_1(s)F(0)\sin\epsilon & [T^2 s^3 + 2\xi T s^2 + s + F(0)\cos\epsilon]\Delta_2(s) \end{vmatrix} = 0. \quad (2.4.12)$$

Отсюда при $\lambda = s$ получаем искомое характеристическое уравнение

$$[T^2 \lambda^3 + 2\xi T \lambda^2 + F(0)\cos\epsilon]^2 + F^2(0)\sin^2\epsilon = 0. \quad (2.4.13)$$

Непосредственное применение к уравнению шестого порядка (2.4.13) процедуры Гурвица приведет к весьма громоздкой системе неравенств. Поэтому следует использовать другой прием, позволяющий понизить порядок определителя Гурвица. Запишем (2.4.13) в форме

$$(T^2 \lambda^3 + 2\xi T \lambda^2 + \lambda)^2 + 2F(0)(T^2 \lambda^3 + 2\xi T \lambda^2 + \lambda)\cos\epsilon + F^2(0) = 0. \quad (2.4.14)$$

Корни этого уравнения удовлетворяют условию

$$\begin{aligned} T^2 \lambda^3 + 2\xi T \lambda^2 + \lambda &= -F(0)\cos\epsilon \pm \sqrt{F^2(0)\cos^2\epsilon - F^2(0)} = \\ &= -F(0)\cos\epsilon \pm iF(0)\sin\epsilon. \end{aligned} \quad (2.4.15)$$

Из (2.4.15) видно, что корни λ представляют собой комплексные числа. Подставив в (2.4.15) $\lambda = \alpha + i\beta = \alpha + i\omega_0$ и приравняв вещественные части, получаем

$$T^2(\alpha^3 - 3\alpha\omega_0^2) + 2\xi T(\alpha^2 - \omega_0^2) + \alpha = -F(0)\cos\epsilon. \quad (2.4.16)$$

Очевидно, при $\alpha < 0$ положение равновесия устойчиво и автоколебаний в системе не возникает. Записав определитель Гурвица для уравнения (2.4.16), получаем условия устойчивости положения равновесия

$$\frac{2\xi T\omega_0^2}{\cos\epsilon} < \bar{F}(0) < \frac{2\xi}{T\cos\epsilon}(1 - 2T\omega_0^2). \quad (2.4.17)$$

Из (2.4.17) во всяком случае видно, что при $F(0) < 2\xi/T\cos\epsilon$ положение равновесия неустойчиво.

Рассмотрим теперь условия устойчивости в ОЭСС с непрерывным управлением кругового предельного цикла, описываемого равенствами (2.4.6). Для этого введем отклонения от этого предельного цикла

$$\delta_1 = \Delta_1 - \Delta_0 \cos \omega_0 t, \quad \delta_2 = \Delta_2 - \Delta_0 \sin \omega_0 t. \quad (2.4.18)$$

Подставив в уравнения (2.4.5) вместо Δ_1 и Δ_2 их значения из (2.4.18) и учитывая выполнение тождеств (2.4.7), получим

$$\begin{aligned} T^2 \delta_1^{(3)} + 2\xi T \delta_1^{(2)} \delta_1^{(1)} &= -F(\Delta_0)(\delta_1 \cos \epsilon - \delta_2 \sin \epsilon), \\ T^2 \delta_2^{(3)} + 2\xi T \delta_2^{(2)} + \delta_2^{(2)} &= -F(\Delta_0)(\delta_1 \sin \epsilon + \delta_2 \cos \epsilon). \end{aligned} \quad (2.4.19)$$

Эти уравнения по своей структуре не отличаются от уравнений (2.4.5), и условия устойчивости кругового предельного цикла сводятся к условиям устойчивости решения $\delta_1(t) = \delta_2(t) \equiv 0$. Поэтому на основании (2.4.17) можно сразу записать эти условия в виде

$$\frac{2\xi T\omega_0^2}{\cos\epsilon} < F(\Delta_0) < \frac{2\xi}{T\cos\epsilon}(1 - 2T\omega_0^2). \quad (2.4.20)$$

Здесь параметры ω_0 и Δ_0 предельного цикла должны удовлетворять уравнениям (2.4.7).

Метод малого параметра. Быстрые и медленные движения. Мы рассмотрели две разновидности возможных движений в системе (2.4.5). Для более широкого исследования возможных движений в ОЭСС с непрерывным управлением воспользуемся представлениями, связанными с асимптотическими методами теории нелинейных систем (см. [3], где приводится метод малого параметра для систем, описываемых уравнениями второго порядка).

При использовании асимптотических методов существенным фактором является малая величина коэффициента демпфирования ξ в передаточных функциях гиравивода (2.3.26). Идея метода Ван-дер-Поля заключается в представлении уравнения второго порядка в форме

$$\ddot{x} + x = \mu f(x, \dot{x}), \quad (2.4.21)$$

где μ — малый параметр; $f(x, \dot{x})$ — некоторая, вообще говоря, нелинейная функция своих аргументов. При $\mu = 0$ уравнение переходит в уравнение гармонического осциллятора, решение которого $x(t) = a \cos t +$

+ $b \sin t$ легко находится. При малых значениях параметра μ решение нелинейного уравнения (2.4.21) должно мало отличаться от движения гармонического осциллятора.

Следуя этому методу, запишем уравнения (2.4.5) в форме

$$\begin{aligned}\Delta_1^{(3)} + \frac{1}{T^2} \Delta_1^{(1)} &= -\frac{2\xi}{T} \Delta_1^{(2)} - \frac{F(\Delta)}{T^2} (\Delta_1 \cos \epsilon - \Delta_2 \sin \epsilon), \\ \Delta_2^{(3)} + \frac{1}{T^2} \Delta_2^{(1)} &= -\frac{2\xi}{T} \Delta_2^{(2)} - \frac{F(\Delta)}{T^2} (\Delta_1 \sin \epsilon + \Delta_2 \cos \epsilon)\end{aligned}\quad (2.4.22)$$

или, введя обозначения

$$\begin{aligned}1/T &= \omega, \\ -\frac{2\xi}{T} \Delta_1^{(2)} - \frac{F(\Delta)}{T^2} (\Delta_1 \cos \epsilon - \Delta_2 \sin \epsilon) &= \mu \Phi_1(\Delta_1, \Delta_2, \Delta_1^{(2)}), \\ -\frac{2\xi}{T} \Delta_2^{(2)} - \frac{F(\Delta)}{T^2} (\Delta_1 \sin \epsilon + \Delta_2 \cos \epsilon) &= \mu \Phi_2(\Delta_1, \Delta_2, \Delta_2^{(2)}),\end{aligned}\quad (2.4.23)$$

в форме

$$\begin{aligned}\Delta_1^{(3)} + \omega^2 \Delta_1^{(1)} &= \mu \Phi_1(\Delta_1, \Delta_2, \Delta_1^{(2)}), \\ \Delta_2^{(3)} + \omega^2 \Delta_2^{(1)} &= \mu \Phi_2(\Delta_1, \Delta_2, \Delta_2^{(2)}).\end{aligned}\quad (2.4.24)$$

При $\mu = 0$ легко находим порождающее решение уравнений (2.4.5)

$$\Delta_1 = c_1 + a_1 \cos(\omega t + \varphi_{10}), \quad \Delta_2 = c_2 + a_2 \cos(\omega t + \varphi_{20}),$$

где $c_1, c_2, a_1, a_2, \varphi_{10}, \varphi_{20}$ — постоянные, определяемые начальными условиями.

При $\mu \neq 0$ эти постоянные становятся медленноменяющимися функциями времени, скоростью изменения которых можно пренебречь при их дифференцировании. Таким образом, полагая для краткости $\omega t + \varphi_{i0} = \varphi_i(t)$, будем иметь

$$\begin{aligned}\Delta_i &= c_i(t) + a_i(t) \cos \varphi_i(t), \\ \Delta_i^{(1)} &= -a_i(t) \omega \sin \varphi_i(t), \\ \Delta_i^{(2)} &= -a_i(t) \omega^2 \cos \varphi_i(t) \quad (i = 1, 2).\end{aligned}\quad (2.4.25)$$

Оценим допущенную погрешность при записи (2.4.25). Дифференцируя формально первое равенство (2.4.25), получаем точное соотношение

$$\Delta_i^{(1)}(t) = c_i^{(1)}(t) + a_i^{(1)}(t) \cos \varphi_i(t) - a_i(t) \varphi_i^{(1)}(t) \sin \varphi_i(t).$$

Приравнявая это выражение второму условию (2.4.25), находим

$$c_i^{(1)}(t) + a_i^{(1)}(t) \cos \varphi_i(t) - a_i(t) \varphi_i^{(1)}(t) \sin \varphi_i(t) + a_i(t) \omega \sin \varphi_i(t) = 0.\quad (2.4.26)$$

Аналогично, дифференцируя формально второе равенство (2.4.25) и приравнявая результат третьему условию (2.4.25), имеем

$$-a_i^{(1)}(t) \omega \sin \varphi_i(t) - a_i(t) \varphi_i^{(1)}(t) \omega \cos \varphi_i(t) + a_i(t) \omega^2 \cos \varphi_i(t) = 0.\quad (2.4.27)$$

Наконец, дифференцируя третье равенство (2.4.25) и подставляя результат в уравнения (2.4.24) с учетом приближенного выражения второй производной $\Delta_i^{(2)}(t)$, получаем

$$-a_i^{(1)}(t) \cos \varphi_i(t) + a_i(t) \varphi_i^{(1)}(t) \sin \varphi_i(t) - a_i(t) \omega \sin \varphi_i(t) = \frac{\mu}{\omega^2} \Phi_i. \quad (2.4.28)$$

Соотношения (2.4.26) – (2.4.28) позволяют оценить скорости изменения функций $c_i(t)$, $a_i(t)$ и $\varphi_i(t)$. Действительно, из (2.4.26) и (2.4.28) получаем

$$c_i^{(1)}(t) = \frac{\mu}{\omega^2} \Phi_i(\Delta_1, \Delta_2, \Delta_i^{(2)}). \quad (2.4.29)$$

Далее, умножая (2.4.27) на $\sin \varphi_i(t)$, (2.4.28) – на $\cos \varphi_i(t)$ и складывая, видим, что

$$a_i^{(1)}(t) = -\frac{\mu}{\omega^2} \Phi_i(\Delta_1, \Delta_2, \Delta_i^{(2)}) \cos \varphi_i(t). \quad (2.4.30)$$

Умножая (2.4.27) на $\cos \varphi_i(t)$, (2.4.28) – на $\sin \varphi_i(t)$ и вычитая, получим оценку для $\varphi_i^{(1)}(t)$:

$$\varphi_i^{(1)}(t) = \omega + \frac{\mu}{\omega^2} \Phi_i(\Delta_1, \Delta_2, \Delta_i^{(2)}) \sin \varphi_i(t). \quad (2.4.31)$$

Таким образом, процессы $c_i(t)$ и $a_i(t)$ не содержат в правых частях (2.4.29) и (2.4.30) параметра ω и их следует считать медленными. Процесс $\varphi_i(t)$ следует считать быстрым с медленноменяющимся вторым слагаемым в (2.4.31).

Следуя далее методике Ван-дер-Поля, осредним правые части уравнений (2.4.29) – (2.4.31) на периоде быстроменяющихся переменных $\varphi_1(t)$ и $\varphi_2(t)$, равном $2\pi/\omega$:

$$\begin{aligned} c_i^{(1)}(t) &= \frac{\mu}{4\pi^2 \omega^2} \int_0^{2\pi} d\varphi_1 \int_0^{2\pi} \Phi_i d\varphi_2, \\ a_i^{(1)}(t) &= \frac{\mu}{4\pi^2 \omega^2} \int_0^{2\pi} d\varphi_1 \int_0^{2\pi} \Phi_i \cos \varphi_i d\varphi_2, \\ \varphi_i^{(1)}(t) &= \omega + \frac{\mu}{4\pi^2 \omega^2} \int_0^{2\pi} d\varphi_1 \int_0^{2\pi} \Phi_i \sin \varphi_i d\varphi_2. \end{aligned} \quad (2.4.32)$$

Подставив в (2.4.32) явные выражения функций Φ_i (2.4.23), получим укороченные уравнения типа Ван-дер-Поля [3].

Исследование решений системы (2.4.24) предполагает, что составляющие вектора $\underline{\Delta}$ углового рассогласования представляются в виде разложений

$$\begin{aligned} \Delta_1(t) &= c_1(t) + a_1(t) \cos \varphi_1(t) + \sum_{k=1}^{\infty} \mu^k U_k, \\ \Delta_2(t) &= c_2(t) + a_2(t) \cos \varphi_2(t) + \sum_{k=1}^{\infty} \mu^k V_k, \end{aligned} \quad (2.4.33)$$

где $U_k = U_k(a_1, a_2; c_1, c_2; \varphi_1, \varphi_2)$ и $V_k = V_k(a_1, a_2; c_1, c_2; \varphi_1, \varphi_2)$ ($k = 1, 2, \dots$) — периодические функции t с периодом $2\pi/\omega$, а функции $c_i(t), a_i(t), \varphi_i(t)$ ($i = 1, 2$) находятся из системы

$$\begin{aligned} c_i^{(1)} &= \sum_{k=1}^{\infty} \mu^k C_{ik}(a_1, a_2; c_1, c_2), \\ a_i^{(1)} &= \sum_{k=1}^{\infty} \mu^k A_{ik}(a_1, a_2; c_1, c_2), \\ \varphi_i^{(1)} &= \sum_{k=1}^{\infty} \mu^k F_{ik}(a_1, a_2; c_1, c_2). \end{aligned} \quad (2.4.34)$$

Сравнивая укороченные уравнения (2.4.32) и (2.4.34) и считая параметр μ настолько малым, что в (2.4.34) можно отбросить слагаемые с $k > 1$, найдем выражения для вычисления C_{i1}, A_{i1} и F_{i1} в явном виде:

$$\begin{aligned} C_{i1}(a_1, a_2; c_1, c_2) &= \frac{1}{4\pi^2 \omega^2} \int_0^{2\pi} \int_0^{2\pi} \Phi_i d\varphi_1 d\varphi_2, \\ A_{i1}(a_1, a_2; c_1, c_2) &= \frac{1}{4\pi^2 \omega^2} \int_0^{2\pi} \int_0^{2\pi} \Phi_i \cos \varphi_i d\varphi_1 d\varphi_2, \\ F_{i1}(a_1, a_2; c_1, c_2) &= \frac{1}{4\pi^2 \omega^2} \int_0^{2\pi} \int_0^{2\pi} \Phi_i \sin \varphi_i d\varphi_1 d\varphi_2. \end{aligned} \quad (2.4.35)$$

Аппроксимация нелинейной характеристики. Укороченные уравнения. Для дальнейших исследований необходимо записать функцию $F(\Delta)$ в аналитической форме. На рис. 2.10 представлены график функции $F(\Delta)$, полученный методом математического моделирования, и его аппроксимация полиномом четвертого порядка на интервале, поглощающем максимальную ожидаемую амплитуду колебаний $\Delta_{\max}$. Таким образом, в дальнейшем будем считать

$$F(\Delta) = f_0 + f_1 \Delta^2 + f_2 \Delta^4, \quad (2.4.36)$$

где коэффициенты f_i ($i = 0, 1, 2$) легко определяются методами интерполяции.

Записав уравнения (2.4.34) с учетом (2.4.35), в правые части которых подставлены явные выражения (2.4.23) функций Φ_i ($i = 1, 2$), получим следующую систему укороченных уравнений (см. приложение):

$$\begin{aligned} c_1^{(1)}(t) &= -c_1(f_0 + f_1 \alpha_{11} + f_2 \alpha_{12}) \cos \epsilon + c_2(f_0 + f_1 \alpha_{21} + f_2 \alpha_{22}) \sin \epsilon, \\ c_2^{(1)}(t) &= -c_1(f_0 + f_1 \alpha_{21} + f_2 \alpha_{22}) \sin \epsilon - c_2(f_0 + f_1 \alpha_{11} + f_2 \alpha_{12}) \cos \epsilon, \end{aligned} \quad (2.4.37)$$

$$a_1^{(1)}(t) = a_1 \left[-\xi \omega + \left(\frac{1}{2} f_0 + f_1 \beta_{11} + f_2 \beta_{12} \right) \cos \epsilon - c_1 c_2 (f_1 + f_2 \gamma_{12}) \sin \epsilon \right],$$

$$a_2^{(1)}(t) = a_2 \left[-\xi \omega + \left(\frac{1}{2} f_0 + f_1 \beta_{21} + f_2 \beta_{22} \right) \cos \epsilon + c_1 c_2 (f_1 + f_2 \gamma_{21}) \sin \epsilon \right],$$

$$\varphi_2^{(1)}(t) = \omega, \quad \varphi_2^{(1)}(t) = \omega.$$

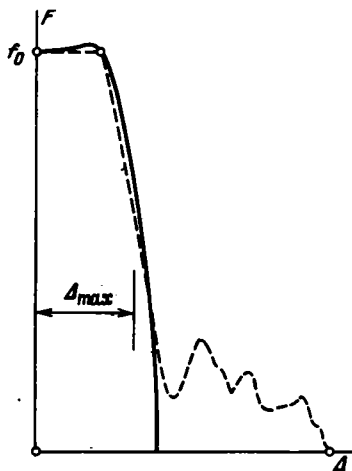


Рис. 2.10. График функции $F(\Delta)$ (штриховая линия) и ее аппроксимация полиномом четвертого порядка (сплошная линия)

Последние два уравнения легко интегрируются, и, следовательно,

$$\varphi_i(t) = \omega t + \varphi_{i0} \quad (i = 1, 2). \quad (2.4.38)$$

Остальные четыре уравнения позволяют проанализировать устойчивость периодического движения ОЭСС, описываемой уравнениями (2.4.5), найти амплитуду автоколебаний, обнаружить границы возникновения других видов возможных движений и исследовать их. Исключая время t из системы первых четырех

уравнений (2.4.37), получим соотношения вида

$$\begin{aligned} \frac{dc_1}{dc_2} &= \frac{C_1(c_1, c_2; a_1, a_2)}{C_2(c_1, c_2; a_1, a_2)}, & \frac{dc_1}{da_1} &= \frac{C_1(c_1, c_2; a_1, a_2)}{A_1(c_1, c_2; a_1, a_2)}, \\ \frac{dc_1}{da_2} &= \frac{C_1(c_1, c_2; a_1, a_2)}{A_2(c_1, c_2; a_1, a_2)}, & \frac{dc_2}{da_1} &= \frac{C_2(c_1, c_2; a_1, a_2)}{A_1(c_1, c_2; a_1, a_2)}, \\ \frac{dc_2}{da_2} &= \frac{C_2(c_1, c_2; a_1, a_2)}{A_2(c_1, c_2; a_1, a_2)}, & \frac{da_1}{da_2} &= \frac{A_1(c_1, c_2; a_1, a_2)}{A_2(c_1, c_2; a_1, a_2)}. \end{aligned} \quad (2.4.39)$$

Смысл обозначений $C_i(c_1, c_2; a_1, a_2)$ и $A_j(c_1, c_2; a_1, a_2)$ ясен из уравнений (2.4.35) и (2.4.37). Простые круговые колебания соответствуют условиям $c_j^{(1)}(t) = a_j^{(1)}(t) = 0$ ($i, j = 1, 2$). Вопрос об их устойчивости может быть решен при качественном анализе системы (2.4.39) методами теории фазового пространства [34]. Координаты особых точек фазового пространства находятся с помощью алгебраических уравнений

$$C_i(c_1, c_2; a_1, a_2) = 0, \quad A_j(c_1, c_2; a_1, a_2) = 0. \quad (2.4.40)$$

Рассмотрим окрестность точки $(0, 0; 0, 0)$ фазового пространства $\{c_1, a_1; c_2, a_2\}$. В окрестности этой точки положим $F(\Delta) = f_0 + \sigma(\Delta)$, и уравнения фазовых траекторий приобретают вид

$$\begin{aligned} \frac{dc_1}{da_1} &= \frac{2f_0}{a_1} \frac{c_2 \sin \epsilon - c_1 \cos \epsilon}{f_0 \cos \epsilon - 2\xi\omega}, & \frac{dc_2}{da_2} &= -\frac{2f_0}{a_2} \frac{c_2 \cos \epsilon + c_1 \sin \epsilon}{f_0 \cos \epsilon - 2\xi\omega}, \\ \frac{dc_2}{dc_1} &= \frac{c_1 \cos \epsilon - c_2 \sin \epsilon}{c_1 \cos \epsilon + c_1 \sin \epsilon}, & \frac{da_1}{da_2} &= \frac{a_1}{a_2}. \end{aligned} \quad (2.4.41)$$

Характеристическое уравнение для третьего соотношения системы (2.4.41) $\lambda^2 + 2\lambda \cos \epsilon + 1 = 0$ имеет корни $\lambda_{1,2} = -\cos \epsilon \pm i \sin \epsilon$ и при $\cos \epsilon > 0$ соответствует устойчивому фокусу в плоскости $\{c_1, c_2\}$, что говорит о несмещенности траекторий. Действительно, в окрестности начала координат

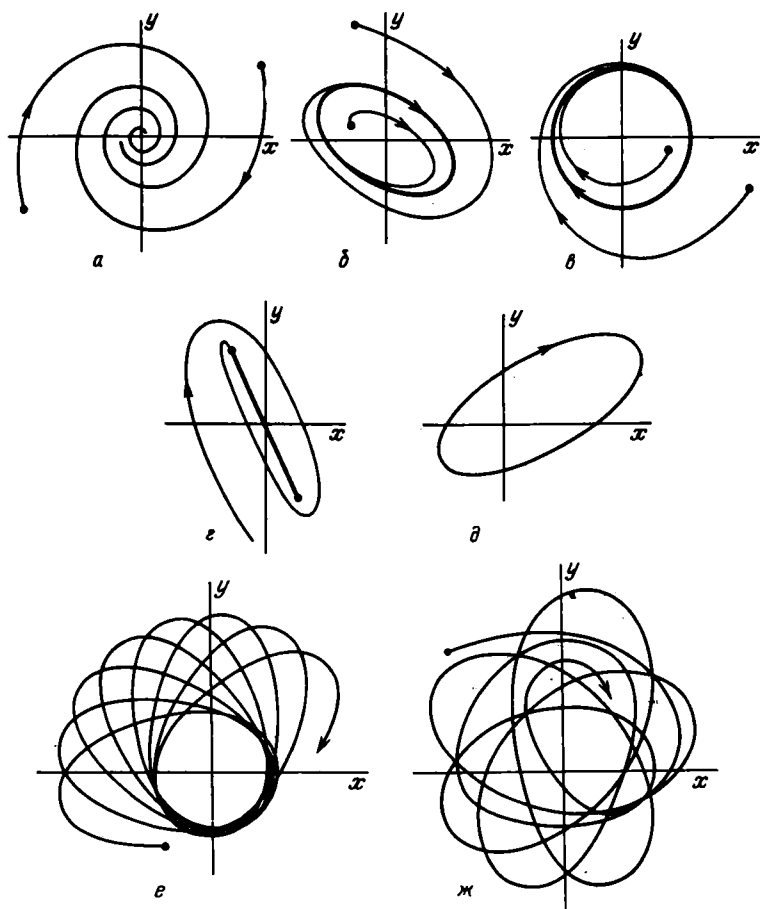


Рис. 2.11. Фазовые портреты движения оптико-электронной следящей системы с непрерывным управлением

фазовой плоскости $\{c_1, c_2\}$

$$c_1(t) = e^{-tf_0 \cos \epsilon} [c_{10} \cos(tf_0 \sin \epsilon) + c_{20} \sin(tf_0 \sin \epsilon)], \quad (2.4.42)$$

$$c_2(t) = e^{-tf_0 \cos \epsilon} [-c_{10} \sin(tf_0 \sin \epsilon) + c_{20} \cos(tf_0 \sin \epsilon)].$$

Дикритический узел в плоскости $\{a_1, a_2\}$ будет устойчивым, как это видно из третьего и четвертого уравнений (2.4.37), если $f_0 < 2\xi\omega/\cos \epsilon$, и неустойчивым при $f_0 \geq 2\xi\omega/\cos \epsilon$. В случае устойчивости узла мы имеем устойчивое положение равновесия в точке $(0, 0; 0, 0)$ фазового пространства $\{c_1, a_1; c_2, a_2\}$ (см. рис. 2.11, а). При неустойчивости положения равновесия в системе возможны другие варианты движения, для анализа которых следует учесть все выписанные члены разложения (2.4.36).

Другие виды возможных движений следящего координатора. Рассмотрим несмещенные периодические режимы, возникающие при нарушении

устойчивости тривиального решения системы (2.4.5). Для этого положим $c_1 = c_2 = 0$, $a_1 = a_{10}$, $a_2 = a_{20}$ и для определения координат особой точки воспользуемся третьим и четвертым уравнениями (2.4.37). Складывая и вычитая их почленно, получаем

$$-2\xi\omega + [f_0 + f_1(\beta_{11} + \beta_{21}) + f_2(\beta_{12} + \beta_{22})] \cos \epsilon = 0, \quad (2.4.43)$$

$$f_1(\beta_{11} - \beta_{21}) + f_2(\beta_{12} - \beta_{22}) = 0.$$

Подставив выражения многочленов β_{ij} ($i, j = 1, 2$) из приложения, запишем явные выражения (2.4.43):

$$-2\xi\omega + \left[f_0 + \frac{5}{8} f_1(a_{10}^2 + a_{20}^2) + \frac{1}{2} f_2 \left(a_{10}^4 + \frac{3}{2} a_{10}^2 a_{20}^2 + a_{20}^4 \right) \right] \cos \epsilon = 0, \quad (2.4.44)$$

$$f_1(a_{10}^2 - a_{20}^2) + f_2(a_{10}^4 - a_{20}^4) = 0.$$

Пусть колебания некруговые (см. рис. 2.11, б) т.е. $a_{10} \neq a_{20}$. Тогда, сокращая второе равенство на разность $a_{10}^2 - a_{20}^2$, находим

$$a_{10}^2 + a_{20}^2 = -\frac{f_1}{f_2}. \quad (2.4.45)$$

Очевидно, что для физически наблюдаемых колебаний ($a_{10} \neq 0$, $a_{20} \neq 0$) должно быть $f_1/f_2 < 0$. Подставив в первое уравнение (2.4.44) соотношение (2.4.45), получим биквадратные уравнения относительно a_{10} или, ввиду симметрии квадратичной формы, относительно a_{20} . Решая их, находим параметры a_{10} и a_{20} эллиптических несмещенных колебаний:

$$a_{10}^2 = -\frac{f_1}{2f_2} + \sqrt{\frac{8\xi\omega}{f_2 \cos \epsilon} - \frac{4f_0}{f_2} + \frac{3}{4} \frac{f_1^2}{f_2^2}} > 0, \quad (2.4.46)$$

$$a_{20}^2 = -\frac{f_1}{2f_2} - \sqrt{\frac{8\xi\omega}{f_2 \cos \epsilon} - \frac{4f_0}{f_2} + \frac{3}{4} \frac{f_1^2}{f_2^2}} > 0.$$

В выражениях (2.4.46) должно выполняться неравенство

$$\frac{8\xi\omega}{f_2 \cos \epsilon} - \frac{4f_0}{f_2} + \frac{3}{4} \frac{f_1^2}{f_2^2} > 0. \quad (2.4.47)$$

Объединяя неравенства (2.4.46) и (2.4.47), найдем границы изменения добротности f_0 , при которых возможны несмещенные эллиптические колебания:

$$\frac{2\xi\omega}{\cos \epsilon} + \frac{1}{8} \frac{f_1^2}{f_2} < f_0 < \frac{2\xi\omega}{\cos \epsilon} + \frac{3}{8} \frac{f_1^2}{f_2}. \quad (2.4.48)$$

При достижении параметром f_0 правой границы (2.4.48) подкоренное выражение (2.4.46) обращается в нуль и предельный цикл становится круговым с амплитудой

$$a_{10} = a_{20} = \sqrt{-\frac{f_1}{2f_2}}. \quad (2.4.49)$$

(см. рис. 2.11, в). При достижении левой границы получаем $a_{20} = 0$ и

$$a_{10} = \left| \frac{f_1}{2f_2} \right| \quad (2.4.50)$$

— колебания становятся плоскими (см. рис. 2.11, з).

Кроме круговых колебаний (2.4.49) в системе (2.4.5) возможен еще один круговой цикл. Для определения его амплитуды положим в (2.4.44) $a_{10} = a_{20}$. Тогда

$$a_{10} = a_{20} = -\frac{5}{14} \frac{f_1}{f_2} \pm \sqrt{\frac{25}{196} \frac{f_1^2}{f_2^2} - \left(f_0 - \frac{2\xi\omega}{\cos \epsilon} \right) \frac{7}{4f_2}}. \quad (2.4.51)$$

Подобно предыдущему, из (2.4.51) определяем границы добротности f_0 , при которых возможен этот круговой цикл:

$$\frac{2\xi\omega}{\cos \epsilon} < f_0 < \frac{2\xi\omega}{\cos \epsilon} + \frac{175}{784} \frac{f_1^2}{f_2}. \quad (2.4.52)$$

Кроме эллиптических колебаний (2.4.46) в системе (2.4.5) возможны более сложные виды движений. Для их выявления следует рассмотреть уравнения (2.4.37) при более общих предположениях, когда ни один из корней уравнений (2.4.40) c_{10} , c_{20} , a_{10} , a_{20} не равен нулю. В этом случае третье и четвертое уравнения системы (2.4.37) можно записать после сокращения на a_1 и a_2 в виде

$$\begin{aligned} -2\xi\omega + [f_0 + f_1(\beta_{11} + \beta_{21}) + f_2(\beta_{12} + \beta_{22})] \cos \epsilon + \\ + c_1 c_2 f_2 (\gamma_{22} - \gamma_{12}) \sin \epsilon = 0, \end{aligned} \quad (2.4.53)$$

$$[f_1(\beta_{21} - \beta_{11}) + f_2(\beta_{22} - \beta_{12})] \cos \epsilon + c_1 c_2 [2f_1 + f_2(\gamma_{22} + \gamma_{12})] \sin \epsilon = 0.$$

Из первых двух уравнений (2.4.37) при $\epsilon \neq 0$ и $\epsilon \neq \pi/2$ получаем

$$2f_0 + f_1(\alpha_{11} + \alpha_{21}) + f_2(\alpha_{12} + \alpha_{22}) = 0, \quad (2.4.54)$$

$$f_1(\alpha_{11} - \alpha_{21}) + f_2(\alpha_{12} - \alpha_{22}) = 0.$$

Подставив явные выражения полиномов α_{ij} ($i, j = 1, 2$) из приложения, второе равенство (2.4.54) запишем в виде

$$\left\{ f_1 + f_2 \left[2(c_1^2 + c_2^2) + \frac{3}{2} (a_1^2 + a_2^2) \right] \right\} (a_1^2 - a_2^2) = 0. \quad (2.4.55)$$

При круговом предельном цикле $a_1 = a_2$ и уравнение (2.4.55) переходит в тождество. Первое уравнение (2.4.54) приобретает форму

$$2f_0 + 2f_1(2a_0^2 + c_1^2 + c_2^2) + f_2 \left[\frac{15}{2} a_0^4 + 12a_0^2(c_1^2 + c_2^2) + 2(c_1^2 + c_2^2) \right] = 0. \quad (2.4.56)$$

Уравнения же (2.4.53) можно записать, как

$$\begin{aligned} -2\xi\omega + \left\{ f_0 + f_1 \left[\frac{5}{4} a_0^2 + 2(c_1^2 + c_2^2) \right] + f_2 \left[\frac{7}{4} a_0^4 + \frac{15}{2} a_0^2(c_1^2 + c_2^2) + \right. \right. \\ \left. \left. + 3(c_1^2 + c_2^2)^2 \right] \right\} \cos \epsilon = 0, \end{aligned} \quad (2.4.57)$$

$$\begin{aligned} \{ f_1(c_1^2 - c_2^2) + f_2[3a_0^2(c_1^2 - c_2^2) + 2(c_1^4 - c_2^4)] \} \cos \epsilon + \\ + c_1 c_2 \{ 2f_1 + f_2[9a_0^2 + 4(c_1^2 + c_2^2)] \} \sin \epsilon = 0. \end{aligned}$$

Соотношения (2.4.56) и (2.4.57) образуют систему относительно трех неизвестных — амплитуды a_0 кругового цикла и координат (c_1, c_2) смещения его центра. Мы не приводим здесь решения этой алгебраической системы.

Наиболее общим случаем периодического режима системы (2.4.5) являются смещенные эллиптические колебания. Их параметры находятся из системы алгебраических уравнений, которая состоит из (2.4.53) и соотношения

$$f_1 + f_2 \left[2(c_1^2 + c_2^2) + \frac{3}{2}(a_1^2 + a_2^2) \right] = 0 \quad (2.4.58)$$

вытекающего из (2.4.55) при $a_1 \neq a_2$ (см. рис. 2.11, д).

Фазовый портрет на рис. 2.11, е иллюстрирует нарушение условий устойчивости смещенных эллиптических циклов. Наконец, рис. 2.11, ж воспроизводит результат моделирования системы по методике, описанной в главе 3, свободной от каких-либо допущений, но при высокой добротности f_0 , выводящей систему за пределы устойчивости периодического движения. При моделировании возможен предельный цикл, близкий к круговому. Некоторые отклонения от этого цикла объясняются неидентичностью каналов управления и действием шумов электронного тракта. Круговой предельный цикл интересен тем, что он обладает наиболее узким спектром и получающиеся при этом периодические составляющие управляющего сигнала сравнительно легко отфильтровываются без существенного ухудшения динамических свойств ОЭСС с непрерывным управлением.

Мы рассмотрели условия возникновения кругового цикла (2.4.49) и (2.4.52). В связи со сделанным замечанием относительно свойств цикла важно исследовать устойчивость соответствующих решений системы (2.4.5). Введем обозначения: $c_1 = x$, $c_2 = y$, $a_1 = a_0 + r$, $a_2 = a_0 + s$. В пространстве $\{x, y, r, s\}$ особая точка $(0, 0; 0, 0)$ соответствует циклу радиуса a_0 с центром в начале координат (Δ_1, Δ_2) . Запишем уравнения (2.4.37) в этих обозначениях:

$$\begin{aligned} \dot{x} &= -x(f_0 + f_1 \alpha_{11} + f_2 \alpha_{12}) \cos \epsilon + y(f_0 + f_1 \alpha_{21} + f_2 \alpha_{22}) \sin \epsilon, \\ \dot{y} &= -x(f_0 + f_1 \alpha_{21} + f_2 \alpha_{22}) \sin \epsilon - y(f_0 + f_1 \alpha_{11} + f_2 \alpha_{12}) \cos \epsilon, \\ \dot{r} &= (a_0 + r) \left[-\xi \omega + \left(\frac{1}{2} f_1 \beta_{11} + f_2 \beta_{12} \right) \cos \epsilon - xy(f_1 + f_2 \gamma_{12}) \sin \epsilon \right], \\ \dot{s} &= (a_0 + s) \left[-\xi \omega + \left(\frac{1}{2} f_0 + f_1 \beta_{21} + f_2 \beta_{22} \right) \cos \epsilon + xy(f_1 + f_2 \gamma_{22}) \sin \epsilon \right]. \end{aligned} \quad (2.4.59)$$

Здесь сохранены обозначения коэффициентов полиномов при новых аргументах x, y, r, s . В окрестности особой точки $(0, 0; 0, 0)$ первые два уравнения системы (2.4.59) эквивалентны фазовой траектории, описываемой соотношением

$$\frac{dy}{dx} = \frac{-x \sin \epsilon - y \cos \epsilon}{-x \cos \epsilon + y \sin \epsilon}. \quad (2.4.60)$$

В плоскости $\{x, y\}$ эта точка — фокус [как мы видели из (2.4.41)].

Уравнение фазовой траектории относительно приращений амплитуд r и s имеет вид

$$\frac{dr}{ds} = \frac{d_1 + d_2 r + d_3 s}{e_1 + e_2 r + e_3 s}, \quad (2.4.61)$$

где согласно приложению и новым обозначениям

$$\begin{aligned} d_1 = e_2 &= a_0 \left[-\xi\omega + \left(\frac{1}{2} f_0 + \frac{5}{8} f_1 a_0^2 + \frac{7}{8} f_2 a_0^4 \right) \cos \epsilon \right], \\ d_2 = e_3 &= -\xi\omega + \left(\frac{1}{2} f_0 + \frac{11}{8} f_1 a_0^2 + \frac{23}{8} f_2 a_0^4 \right) \cos \epsilon, \\ d_3 = e_2 &= \frac{1}{2} a_0^2 (f_1 + 3f_2 a_0^2) \cos \epsilon. \end{aligned} \quad (2.4.62)$$

Если a_0 — корень первого выражения (2.4.44), то $d_1 = e_1 = 0$ и характеристическое уравнение для (2.4.61) имеет вид

$$\lambda^2 - (d_2 + e_3) \lambda + d_2 e_3 - d_3 e_2 = 0 \quad (2.4.63)$$

или в силу (2.4.62)

$$\lambda^2 - 2d_2 \lambda + d_2^2 - d_3^2 = 0. \quad (2.4.64)$$

Корни этого уравнения $\lambda_1 = d_2 - d_3$ и $\lambda_2 = d_2 + d_3$ вещественны. В случае $\lambda_1 \lambda_2 > 0$ в плоскости $\{r, s\}$ возникает особая точка типа узла, устойчивость которой зависит от знака правых частей уравнений первого приближения для третьего и четвертого соотношений (2.4.59). Если оба корня отрицательны, узел устойчив и круговые колебания с амплитудой a_0 , определяемой первым выражением (2.4.44), существуют. При $\lambda_1 \lambda_2 < 0$ корни имеют разные знаки. При этом возникает особая точка типа седла, которая всегда неустойчива. В этом случае круговой предельный цикл не наблюдается.

Таким образом, вид движения ОЭСС, описываемый уравнениями (2.4.5), определяется соотношением между добротностью системы f_0 и величиной обобщенного параметра $2\xi\omega/\cos \epsilon$. Анализ устойчивости круговых циклов требует более точного знания формы характеристики $F(\Delta)$.

§ 2.5. Анализ устойчивости сопровождения в ОЭСС с импульсным управлением

Уравнения ОЭСС с АИМ. В § 2.2 нами были получены соотношения (2.2.2), описывающие пеленгационную характеристику ОЭСС с четырехплощадочным фотоприемником. Таким образом, анализатор изображения не содержит элементов сканирующего устройства, что упрощает его конструкцию. Обычно размеры площадок совпадают с размерами изображения бесконечно удаленного источника. В связи с этим при малом числе площадок добротность ОЭСС невелика и рабочий режим слежения не должен сопровождаться автоколебаниями. Поэтому в дальнейшем ограничимся выяснением условий, при которых автоколебания отсутствуют. Рассмотрим два вида модуляции, которые используются при формировании управляющего сигнала: амплитудно-импульсную (АИМ) и широтно-импульсную (ШИМ). В качестве исполнительного элемента в обоих случаях рассмотрим

гиропривод с динамикой, описываемой уравнениями (2.3.21). При АИМ на исполнительный элемент поступают импульсы постоянной длительности T , амплитуды которых пропорциональны значениям сигналов $u_x(\Delta_x, \Delta_y)$ и $u_y(\Delta_x, \Delta_y)$ из (2.2.6). Вводя обозначения $u_1 = u_x$ и $u_2 = u_y$, запишем (2.3.21) в виде

$$\begin{aligned} I\Delta_1^{(2)} + \xi\Delta_1^{(1)} + H\Delta_2^{(1)} &= ku_1, \\ I\Delta_2^{(2)} + \xi\Delta_2^{(1)} - H\Delta_1^{(1)} &= ku_2. \end{aligned} \quad (2.5.1)$$

Здесь вместо прежних обозначений угловой скорости сопровождения положено $p = \Delta_1^{(1)}$ и $q = \Delta_2^{(1)}$.

Вводя в уравнения (2.5.1) вместо времени t безразмерный параметр $\bar{t} = t/T$, приведем их к виду [137]

$$\begin{aligned} \Delta_{1T}^{(2)}(\bar{t}) - \alpha\Delta_{1T}^{(1)}(\bar{t}) + \beta\Delta_{2T}^{(1)}(\bar{t}) &= u_{1T}(\bar{t}), \\ \Delta_{2T}^{(2)}(\bar{t}) - \alpha\Delta_{2T}^{(1)}(\bar{t}) - \beta\Delta_{1T}^{(1)}(\bar{t}) &= u_{2T}(\bar{t}), \end{aligned} \quad (2.5.2)$$

где $\alpha = -\xi T/I$; $\beta = HT/I$; $\Delta_{iT}(\bar{t}) = \Delta_i(t/T)$; $u_{iT} = (kT^2/I) u_i(t/T)$ ($i = 1, 2$).

Для исследования импульсной ОЭСС следует составить разностные уравнения. Следуя методике, указанной в [137], запишем решение системы (2.5.2) в виде

$$\begin{aligned} \Delta_{1T}(\bar{t}) &= e^{\alpha(\bar{t}-n)} [C_1 \cos \beta(\bar{t}-n) + C_2 \sin \beta(\bar{t}-n)] - \\ &- \frac{\alpha u_{1T} + \beta u_{2T}}{\alpha^2 + \beta^2} \bar{t} + B_1, \\ \Delta_{2T}(\bar{t}) &= e^{\alpha(\bar{t}-n)} [C_1 \sin \beta(\bar{t}-n) - C_2 \cos \beta(\bar{t}-n)] + \\ &+ \frac{\beta u_{1T} - \alpha u_{2T}}{\alpha^2 + \beta^2} \bar{t} + B_2 \quad (n \leq \bar{t} \leq n+1). \end{aligned} \quad (2.5.3)$$

Здесь C_i, B_j ($i, j = 1, 2$) — постоянные интегрирования, определяемые значениями, которые принимают функции $\Delta_{1T}(\bar{t})$ и $\Delta_{2T}(\bar{t})$ и их производные по $\bar{t}$ при $\bar{t} = n$ (n — целое число).

Для сокращения записей целесообразно ввести матрицы

$$G(\bar{t}) = \begin{vmatrix} \cos \beta \bar{t} & \sin \beta \bar{t} \\ \sin \beta \bar{t} & -\cos \beta \bar{t} \end{vmatrix}, \quad (2.5.4)$$

$$A = \begin{vmatrix} \alpha & \beta \\ \beta & -\alpha \end{vmatrix} \frac{1}{\alpha^2 + \beta^2} \quad (2.5.5)$$

и векторы

$$\begin{aligned} \underline{\Delta}(\bar{t}) &= \text{colon} [\Delta_{1T}(\bar{t}) \quad \Delta_{2T}(\bar{t})], \\ C &= \text{colon} [C_1 \quad C_2], \\ u &= \text{colon} [u_1 \quad u_2], \\ B &= \text{colon} [B_1 \quad B_2]. \end{aligned} \quad (2.5.6)$$

Тогда система (2.5.3) записывается в более компактной форме:

$$\underline{\Delta}(\bar{t}) = e^{\alpha(\bar{t}-n)} G(\bar{t}-n) C - G_0 A u(\bar{t}-n) + B. \quad (2.5.7)$$

Дифференцируя (2.5.3) по $\bar{t}$ и используя обозначения (2.5.4) – (2.5.6), получаем

$$\underline{\Delta}^{(1)}(\bar{t}) = e^{\alpha(\bar{t}-n)} G(\bar{t}-n) G_0 A C (\alpha^2 + \beta^2) - G_0 A u. \quad (2.5.8)$$

В уравнениях (2.5.7) и (2.5.8) $G_0 = G(0)$.

Полагая $\bar{t} = n$, найдем соотношения, определяющие постоянные интегрирования:

$$\begin{aligned} \underline{\Delta}(n) &= G_0 C + B, \\ \underline{\Delta}^{(1)}(n) &= (\alpha^2 + \beta^2) A C - G_0 A u. \end{aligned} \quad (2.5.9)$$

Заметим, что $A^{-1} = (\alpha^2 + \beta^2) A$. Поэтому из (2.5.7) – (2.5.9) следует при $\bar{t} = n + 1$

$$\underline{\Delta}(n+1) = [e^{\alpha} G(1) - G_0] A \underline{\Delta}^{(1)}(n) + \underline{\Delta}(n) + \{[e^{\alpha} G(1) - G_0] A - E\} G_0 A u \quad (2.5.10)$$

Соотношение (2.5.10) представляет собой разностное уравнение разомкнутой ОЭСС с АИМ. Оно является рекуррентным соотношением и позволяет осуществить последовательное вычисление значений вектора ошибки $\underline{\Delta}$ с произвольным шагом T по времени, что является ценным качеством при математическом моделировании, ибо приводит к экономии времени счета. Здесь мы используем соотношение (2.5.10) для решения основной задачи – исследования устойчивости работы импульсной ОЭСС. Поэтому уравнение (2.5.10) следует дополнить условием замыкания, которое можно записать в форме

$$u = \lambda G\left(\frac{\pi}{2\beta}\right) G_0 \underline{\Delta}(n), \quad (2.5.11)$$

где $\lambda = kT^2/I$.

При подстановке условия (2.5.11) в (2.5.10) полезно применить тождество

$$G(x) A G(y) \equiv G(x+y) A G_0, \quad (2.5.12)$$

в справедливости которого легко убедиться непосредственным вычислением. Таким образом, уравнения замкнутой импульсной ОЭСС с АИМ принимают вид

$$\begin{aligned} \underline{\Delta}(n+1) &= [e^{\alpha} G(1) - G_0] A \underline{\Delta}^{(1)} + \\ &+ \lambda \left[e^{\alpha} G\left(1 + \frac{\pi}{2\beta}\right) A G_0 - G\left(\frac{\pi}{2\beta}\right) \right] A \underline{\Delta}(n) + \underline{\Delta}(n), \\ \underline{\Delta}^{(1)}(n+1) &= e^{\alpha} G(1) G_0 \underline{\Delta}^{(1)}(n) + \lambda \left[e^{\alpha} G\left(1 + \frac{\pi}{2\beta}\right) - G\left(\frac{\pi}{2\beta}\right) \right] A \underline{\Delta}(n). \end{aligned} \quad (2.5.13)$$

Разностное уравнение (2.5.10) является линейным и для анализа устойчивости его тривиального решения $\underline{\Delta}(n) \equiv 0$, $\underline{\Delta}^{(1)} \equiv 0$ можно воспользо-

ваться алгебраическим критерием Гурвица. Достоинство этого критерия в данном случае состоит в том, что он позволяет обнаружить связь между параметрами гироспривода α и β , добротностью ОЭСС k и частотой следования импульсов T .

Область устойчивости ОЭСС с АИМ. Запишем характеристический многочлен (2.5.13) в скалярной форме, воспользовавшись предварительно обозначениями

$$\begin{aligned} M_{10} &= \frac{1}{\alpha^2 + \beta^2} [e^\alpha (\alpha \cos \beta + \beta \sin \beta) - \alpha], \\ M_{20} &= \frac{1}{\alpha^2 + \beta^2} [e^\alpha (\beta \cos \beta - \alpha \sin \beta) - \beta], \\ A &= \frac{1}{\alpha^2 + \beta^2} (\alpha M_{20} + \beta M_{10} - \beta), \\ B &= \frac{1}{\alpha^2 + \beta^2} (\beta M_{20} - \alpha M_{10} + \alpha). \end{aligned} \quad (2.5.14)$$

Таким образом,

$$\begin{aligned} \Delta_{1T}(n+1) &= (1 + \lambda A) \Delta_{1T}(n) + \lambda B \Delta_{2T}(n) + M_{10} \Delta_{1T}^{(1)}(n) + M_{20} \Delta_{2T}^{(1)}(n), \\ \Delta_{2T}(n+1) &= \lambda B \Delta_{1T}(n) + (1 + \lambda A) \Delta_{2T}(n) - M_{20} \Delta_{1T}^{(1)}(n) + M_{10} \Delta_{2T}^{(1)}(n), \\ \Delta_{1T}^{(1)}(n+1) &= \lambda M_{20} \Delta_{1T}(n) - \lambda M_{10} \Delta_{2T}(n) + \\ &+ e^\alpha \cos \beta \cdot \Delta_{1T}(n) - e^\alpha \sin \beta \cdot \Delta_{2T}^{(1)}(n), \\ \Delta_{2T}^{(1)}(n+1) &= \lambda M_{10} \Delta_{1T}(n) + \lambda M_{20} \Delta_{2T}(n) + \\ &+ e^\alpha \sin \beta \cdot \Delta_{1T}^{(1)}(n) + e^\alpha \cos \beta \cdot \Delta_{2T}^{(1)}(n). \end{aligned} \quad (2.5.15)$$

Характеристический многочлен относительно e^q записывается в виде определителя [135]

$$\det \begin{vmatrix} 1 + \lambda A - e^q & \lambda B & M_{10} & M_{20} \\ \lambda B & 1 + \lambda A - e^q & -M_{20} & M_{10} \\ \lambda M_{20} & -\lambda M_{10} & e^\alpha \cos \beta - e^q & -e^\alpha \sin \beta \\ \lambda M_{10} & \lambda M_{20} & e^\alpha \sin \beta & e^\alpha \cos \beta - e^q \end{vmatrix} =$$

$$= b_0 e^{4q} + b_1 e^{3q} + b_2 e^{2q} + b_3 e^q + b_4 = 0. \quad (2.5.16)$$

После вычисления определителя (2.5.16) имеем

$$\begin{aligned} b_0 &= 1, \\ b_1 &= -2(1 + \lambda A + e^\alpha \cos \beta), \\ b_2 &= e^{2\alpha} + (1 + \lambda A)^2 + \lambda^2 B^2 + 4(1 + \lambda A) e^\alpha \cos \beta - 4 \lambda M_{10} M_{20}, \\ b_3 &= -2\{e^{2\alpha}(1 + \lambda A) + [(1 + \lambda A)^2 + \lambda^2 B^2] e^\alpha \cos \beta\} + \\ &+ 4 \lambda M_{10} M_{20} (1 + \lambda A + e^\alpha \cos \beta) + 2 \lambda (\lambda B - e^\alpha \sin \beta) (M_{20}^2 - M_{10}^2), \\ b_4 &= e^{2\alpha} [(1 + \lambda A)^2 + \lambda^2 B^2] - 4 \lambda M_{10} M_{20} (1 + \lambda A) e^\alpha \cos \beta + \\ &+ 2 \lambda e^\alpha \sin \beta (M_{20}^2 - M_{10}^2) (1 + \lambda A) - 2 \lambda^2 B e^\alpha \cos \beta (M_{20}^2 - M_{10}^2) + \\ &+ \lambda^2 (M_{10}^2 + M_{20}^2). \end{aligned} \quad (2.5.17)$$

Для того чтобы воспользоваться критерием устойчивости Гурвица, необходимо предварительно выполнить дробно-линейное преобразование вида [135]

$$e^q = \frac{1+w}{1-w}.$$

Тогда характеристический многочлен (2.5.16) перейдет в многочлен относительно переменной w

$$a_0 w^4 + a_1 w^3 + a_2 w^2 + a_3 w + a_4 = 0,$$

где коэффициенты a_k ($k = 0, 1, 2, 3, 4$) вычисляются по формулам

$$\begin{aligned} a_0 &= b_0 - b_1 + b_2 - b_3 + b_4, \\ a_1 &= 4b_0 - 2b_1 + 2b_3 - 4b_4, \\ a_2 &= 6b_0 - 2b_2 + 6b_4, \\ a_3 &= 4b_0 + 2b_1 - 2b_3 - 4b_4, \\ a_4 &= b_0 + b_1 + b_2 + b_3 + b_4. \end{aligned} \quad (2.5.18)$$

При $a_0 > 0$ условия устойчивости решения разностных уравнений (2.5.15), описывающих процесс слежения импульсной ОЭСС, сводятся к выполнению системы неравенств

$$\begin{aligned} a_1 &> 0, \quad a_1 a_2 - a_0 a_3 > 0, \\ a_3(a_1 a_2 - a_0 a_3) - a_4 a_1^2 &> 0, \quad a_4 > 0. \end{aligned} \quad (2.5.19)$$

Дальнейший анализ условий устойчивости целесообразно провести численными методами, выделив в ОЭСС легко изменяемые параметры. Такими параметрами, в частности, являются добротность ОЭСС k [см. уравнения (2.5.1)] и период следования импульсов T . Параметры H , I и ξ связаны с особенностями конструктивного исполнения гиросприода и не поддаются простому изменению в процессе настройки. Для численной процедуры используются соотношения (2.5.14) и (2.5.17), в которых $\alpha = -\xi T/I$ и $\beta = HT/I$, и соотношения (2.5.18) и (2.5.19). Результат вычисления представлен на рис. 2.12 в плоскости параметров T и $\lambda = kT^2/I$. Интересно отметить, что допустимая добротность k (или пропорциональная ей величина λ) оказывается в дискретной системе выше, чем в непрерывной. Физически это объясняется несовпадением собственной частоты гироспри-

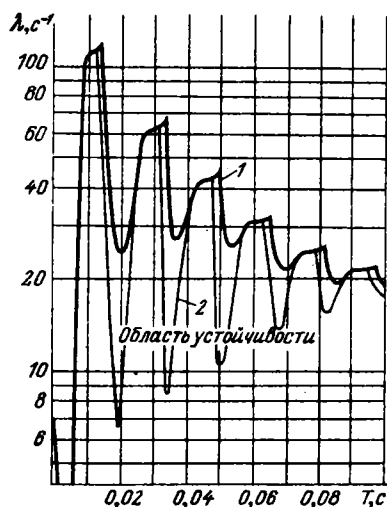


Рис. 2.12. Граница устойчивости слежения оптико-электронной следящей системы с импульсным управлением: 1 — граница с амплитудно-импульсной модуляцией; 2 — граница с широтно-импульсной модуляцией

вода и частоты следования импульсов. Подобное свойство дискретных цепей отмечалось уже в работе [130]. Из рис. 2.12 видно, что при соответствующем подборе периода следования импульсов T добротность можно существенно повысить.

При $T \rightarrow 0$ дискретная система с АИМ становится непрерывной, а система разностных уравнений (2.5.15) заменяется системой дифференциальных уравнений. Так как при $T \rightarrow 0$ параметры α и β также стремятся к нулю, то, используя разложения

$$e^\alpha = 1 + \alpha + o(\alpha),$$

$$\cos \beta = 1 - \frac{\beta^2}{2} + o(\beta^2),$$

$$\sin \beta = \beta + o(\beta),$$

можно записать

$$M_{10} = \frac{(1 + \alpha)(\alpha - \alpha\beta^2/2 + \beta^2) - \alpha + o(\alpha^2 + \beta^2)}{\alpha^2 + \beta^2} = 1 + o(\alpha, \beta),$$

$$M_{20} = \frac{(1 + \alpha)(\beta - \beta^2/2 - \alpha\beta) - \beta + o(\alpha^2 + \beta^2)}{\alpha^2 + \beta^2} = o(\alpha, \beta).$$

Из обозначений (2.5.14) видно также, что $A = o(\alpha, \beta)$ и $B = o(\alpha, \beta)$. Первые два уравнения системы (2.5.15) переходят при $T \rightarrow 0$ в тождества. Действительно, $\Delta_{1T}^{(1)}(\bar{t}) = T\Delta_{1T}^{(1)}(t)$, $\Delta_{2T}^{(1)}(\bar{t}) = T\Delta_{2T}^{(1)}(t)$ и мы приходим к соотношениям

$$\begin{aligned} \Delta_{1T}(n+1) - \Delta_{1T}(n) &= T\Delta_{1T}^{(1)}(nT) + o(T), \\ \Delta_{2T}(n+1) - \Delta_{2T}(n) &= T\Delta_{2T}^{(1)}(nT) + o(T), \end{aligned} \quad (2.5.20)$$

Обозначая $t = nT$ и переходя в (2.5.20) к пределу при $T \rightarrow 0$, получаем тождества вида $\Delta_{iT}(t) \equiv \Delta_{iT}(t)$ для любого t . Два оставшихся уравнения системы (2.5.15) можно записать в форме

$$\begin{aligned} \Delta_{1T}^{(1)}(n+1) - \Delta_{1T}^{(1)}(n) &= -\lambda\Delta_{2T}^{(1)}(n) + \alpha\Delta_{1T}^{(1)}(n) - \beta\Delta_{2T}^{(1)}(n) + o(T), \\ \Delta_{2T}^{(1)}(n+1) - \Delta_{2T}^{(1)}(n) &= \lambda\Delta_{1T}^{(1)}(n) + \beta\Delta_{1T}^{(1)}(n) + \alpha\Delta_{2T}^{(1)}(n) + o(T), \end{aligned}$$

откуда после деления на T и с учетом обозначений α , β и λ следует система дифференциальных уравнений (2.5.1), условие устойчивости решения которых было найдено в виде неравенства (2.3.25).

Уравнения ОЭСС с ШИМ. Перейдем к анализу ОЭСС с импульсным управлением по методу ШИМ. Во многих случаях этот метод модуляции по схемным и конструктивным соображениям оказывается более приемлемым. В частности, поскольку сигнал отличен от нуля лишь в части периода длительности T , нагрев деталей электронных схем получается меньше, чем при АИМ.

Выходной сигнал блока ШИМ представляет собой последовательность импульсов, относительная длительность которых γ_i зависит от уровня сигналов u_x и u_y (см. рис. 2.5), поступающих на его вход. Обозначая в дальнейшем импульсные управляющие сигналы по первому и второму каналам через $u_1(t/T)$ и $u_2(t/T)$, запишем их формальные выражения

в виде

$$\begin{aligned} u_1(t/T) &= u_1(\bar{t}) = -u_0 \operatorname{sign} u_x \sum_{\nu} [1(\bar{t} - \nu) - 1(\bar{t} - \nu - \gamma_1)], \\ u_2(t/T) &= u_2(\bar{t}) = u_0 \operatorname{sign} u_y \sum_{\nu} [1(\bar{t} - \nu) - 1(\bar{t} - \nu - \gamma_2)]. \end{aligned} \quad (2.5.21)$$

Здесь приняты обычные обозначения для единичной функции

$$1(t - t_0) = \begin{cases} 1 & \text{при } t > t_0, \\ 0 & \text{при } t \leq t_0 \end{cases}$$

и

$$\operatorname{sign} z = 2 \left[1(z) - \frac{1}{2} \right].$$

Уравнения (2.5.1) по-прежнему описывают непрерывную часть ОЭСС.

Положим без ограничения общности $\gamma_1 < \gamma_2$. Тогда на участке $n \leq \bar{t} < n + \gamma_1$ функции $u_1(\bar{t})$ и $u_2(\bar{t})$ принимают постоянные и отличные от нуля значения $-u_0$ и u_0 соответственно. Условимся их значения на участке $n + 1 \leq \bar{t} < n + \gamma_2$ обозначать $\tilde{u}_{1T}$ и $\tilde{u}_{2T}$. Очевидно, что при $n + \gamma_2 \leq \bar{t} < n + 1$ имеем $u_{1T}(\bar{t}) = u_{2T}(\bar{t}) = 0$.

На участке $n \leq \bar{t} < n + \gamma_1$ движение гиروهода описывается уравнением (2.5.10). В конце этого участка координаты системы принимают значения

$$\begin{aligned} \underline{\Delta}(n + \gamma_1) &= [e^{\alpha\gamma_1} G(\gamma_1) - G_0] A \underline{\Delta}^{(1)}(n) + \underline{\Delta}(n) + \\ &+ [e^{\alpha\gamma_1} G(\gamma_2) - G_0] A G_0 A u - G_0 A u \gamma_1, \\ \underline{\Delta}^{(1)}(n + \gamma_1) &= e^{\alpha\gamma_1} G(\gamma_1) G_0 \underline{\Delta}^{(1)}(n) + [e^{\alpha\gamma_1} G(\gamma_1) - G_0] A u. \end{aligned} \quad (2.5.22)$$

Пользуясь далее уравнением (2.5.10), запишем также выражения для координат $\underline{\Delta}(\bar{t})$ и $\underline{\Delta}^{(1)}(\bar{t})$ на участке $n + \gamma_1 \leq \bar{t} < n + \gamma_2$:

$$\begin{aligned} \underline{\Delta}(\bar{t}) &= [e^{\alpha(\bar{t} - n - \gamma_1)} G(\bar{t} - n - \gamma_1) - G_0] A \underline{\Delta}^{(1)}(n + \gamma_1) + \underline{\Delta}(n + \gamma_1) + \\ &+ [e^{\alpha(\bar{t} - n - \gamma_1)} G(\bar{t} - n - \gamma_1) - G_0] A G_0 A \tilde{u} - G_0 A \tilde{u}(\bar{t} - n - \gamma_1), \\ \underline{\Delta}^{(1)}(\bar{t}) &= e^{\alpha(\bar{t} - n - \gamma_1)} G(\bar{t} - n - \gamma_1) G_0 \underline{\Delta}^{(1)}(n + \gamma_1) + \\ &+ [e^{\alpha(\bar{t} - n - \gamma_1)} G(\bar{t} - n - \gamma_1) - G_0] A \tilde{u}. \end{aligned} \quad (2.5.23)$$

Из (2.5.23) следует

$$\begin{aligned} \underline{\Delta}(n + \gamma_2) &= [e^{\alpha(\gamma_2 - \gamma_1)} G(\gamma_2 - \gamma_1) - G_0] A \underline{\Delta}^{(1)}(n + \gamma_1) + \\ &+ \underline{\Delta}(n + \gamma_1) + [e^{\alpha(\gamma_2 - \gamma_1)} G(\gamma_2 - \gamma_1) - G_0] A G_0 A \tilde{u} - G_0 A \tilde{u}(\gamma_2 - \gamma_1), \\ \underline{\Delta}^{(1)}(n + \gamma_2) &= e^{\alpha(\gamma_2 - \gamma_1)} G(\gamma_2 - \gamma_1) G_0 \underline{\Delta}^{(1)}(n + \gamma_1) + \\ &+ [e^{\alpha(\gamma_2 - \gamma_1)} G(\gamma_2 - \gamma_1) - G_0] A \tilde{u}. \end{aligned} \quad (2.5.24)$$

На последнем участке $n + \gamma_2 \leq \bar{t} < n + 1$ имеем

$$\begin{aligned} \underline{\Delta}(\bar{t}) &= [e^{\alpha(\bar{t} - n - \gamma_2)} G(\bar{t} - n - \gamma_2) - G_0] A \underline{\Delta}^{(1)}(n + \gamma_2) + \underline{\Delta}(n + \gamma_2), \\ \underline{\Delta}^{(1)}(\bar{t}) &= e^{\alpha(\bar{t} - n - \gamma_2)} G(\bar{t} - n - \gamma_2) G_0 \underline{\Delta}^{(1)}(n + \gamma_2). \end{aligned} \quad (2.5.25)$$

Полагая в (2.5.25) $\bar{t} = n + 1$, получаем

$$\begin{aligned}\underline{\Delta}(n+1) &= [e^{\alpha(1-\gamma_2)}G(1-\gamma_2) - G_0]A\underline{\Delta}^{(1)}(n+\gamma_2) + \underline{\Delta}(n+\gamma_2), \\ \underline{\Delta}^{(1)}(n+1) &= e^{\alpha(1-\gamma_2)}G(1-\gamma_2)G_0\underline{\Delta}^{(1)}(n+\gamma_2).\end{aligned}\quad (2.5.26)$$

Разностные уравнения ОЭСС с ШИМ получаются последовательной подстановкой $\underline{\Delta}(n+\gamma_2)$ и $\underline{\Delta}^{(1)}(n+\gamma_2)$ из (2.5.24) и $\underline{\Delta}(n+\gamma_1)$ и $\underline{\Delta}^{(1)}(n+\gamma_1)$ из (2.5.22) в (2.5.26). При этом полезным средством для сокращения выкладок снова оказываются тождества вида (2.5.12). После выполнения указанных подстановок разностные уравнения разомкнутой импульсной ОЭСС с ШИМ приобретают вид

$$\begin{aligned}\underline{\Delta}(n+1) &= [e^{\alpha}G(1) - G_0]A\underline{\Delta}^{(1)}(n) + \underline{\Delta}(n) + \\ &+ \{[e^{\alpha}G(1) - e^{\alpha(1-\gamma_1)}G(1-\gamma_1)]A - \gamma_1 E\}G_0Au + \\ &+ \{[e^{\alpha(1-\gamma_1)}G(1-\gamma_1) - e^{\alpha(1-\gamma_2)}G(1-\gamma_2)]A - \\ &- (\gamma_2 - \gamma_1)E\}G_0A\tilde{u}, \\ \underline{\Delta}^{(1)}(n+1) &= e^{\alpha}G(1)G_0\underline{\Delta}^{(1)}(n) + [e^{\alpha}G(1) - e^{\alpha(1-\gamma_1)}G(1-\gamma_1)]Au + \\ &+ [e^{\alpha(1-\gamma_1)}G(1-\gamma_1) - e^{\alpha(1-\gamma_2)}G(1-\gamma_2)]A\tilde{u}.\end{aligned}\quad (2.5.27)$$

Линеаризация уравнений. В выделение области устойчивости с ШИМ. При $\gamma_1 = \gamma_2 = 1$ уравнения (2.5.27) переходят в (2.5.10), соответствующее системе с АИМ. Но в отличие от случая АИМ система с ШИМ в разомкнутом состоянии описывается нелинейными разностными уравнениями. Для анализа устойчивости положения равновесия следует осуществить линеаризацию этих уравнений. В окрестности точки равновесия относительная длительность импульсов мала и пропорциональна составляющим ошибки сопровождения импульсной ОЭСС по каналам управления.

Воспользовавшись разложениями

$$\begin{aligned}\cos \beta(1-\gamma_i) &= \cos \beta + \beta \gamma_i \sin \beta + o(\gamma_i), \\ \sin \beta(1-\gamma_i) &= \sin \beta - \beta \gamma_i \cos \beta + o(\gamma_i), \\ e^{\alpha(1-\gamma_i)} &= e^{\alpha}(1 - \alpha \gamma_i) + o(\gamma_i),\end{aligned}\quad (2.5.28)$$

находим

$$e^{\alpha(1-\gamma_i)}G(1-\gamma_i) = e^{\alpha}G(1) - \gamma_i e^{\alpha} \left[\alpha G(1) + \beta G \left(1 + \frac{\pi}{2\beta} \right) \right]. \quad (2.5.29)$$

Подставляя (2.5.29) в разностные уравнения (2.5.27), найдем их линеаризованную форму

$$\begin{aligned}\underline{\Delta}(n+1) &= [e^{\alpha}G(1) - G_0]A\underline{\Delta}^{(1)}(n) + \underline{\Delta}(n) + \\ &+ \left\{ e^{\alpha} \left[\alpha G(1) + \beta G \left(1 + \frac{\pi}{2\beta} \right) \right] A - E \right\} G_0 A [\gamma_1 u + (\gamma_2 - \gamma_1) \tilde{u}], \\ \underline{\Delta}^{(1)}(n+1) &= e^{\alpha}G(1)G_0\underline{\Delta}^{(1)}(n) + e^{\alpha} \left[\alpha G(1) + \beta G \left(1 + \frac{\pi}{2\beta} \right) \right] A [\gamma_1 u + \\ &+ (\gamma_2 - \gamma_1) \tilde{u}].\end{aligned}\quad (2.5.30)$$

В уравнениях (2.5.30) вектор-столбец $\tilde{u}$, представляющий собой воздействие на непрерывную часть ОЭСС на участке $n + \gamma_1 \leq \bar{t} < n + \gamma_2$, принимает значение либо $\tilde{u} = \text{colom}[u \ 0]$, либо $\tilde{u} = \text{colom}[0 \ u]$. Поэтому линейная форма $\gamma_1 u + (\gamma_2 - \gamma_1)\tilde{u}$ может иметь значение $\text{colom}[\gamma_2 u_1 \ \gamma_1 u_2]$ или $\text{colom}[\gamma_1 u_1 \ \gamma_2 u_2]$. Очевидно, что при малых длительностях импульсов в обоих случаях можно записать уравнение замыкания

$$\gamma_1 u + (\gamma_2 - \gamma_1)\tilde{u} = \lambda G \left(\frac{\pi}{2\beta} \right) G_0 \underline{\Delta}(n). \quad (2.5.31)$$

Таким образом, уравнения замкнутой ОЭСС с ШИМ в линеаризованном виде суть

$$\begin{aligned} \underline{\Delta}(n+1) &= [e^\alpha G(1) - G_0] A \underline{\Delta}^{(1)}(n) + \underline{\Delta}(n) + \\ &+ \lambda \left\{ e^\alpha \left[\alpha G(1) + \beta G \left(1 + \frac{\pi}{2\beta} \right) \right] A - E \right\} G_0 A G \left(\frac{\pi}{2\beta} \right) G_0 \underline{\Delta}(n), \\ \underline{\Delta}^{(1)}(n+1) &= e^\alpha G(1) G_0 \underline{\Delta}^{(1)}(n) + \lambda e^\alpha \left[\alpha G(1) + \beta G \left(1 + \frac{\pi}{2\beta} \right) \right] \times \\ &\times A G \left(\frac{\pi}{2\beta} \right) \underline{\Delta}(n). \end{aligned} \quad (2.5.32)$$

Этим уравнениям можно придать более компактную форму. Согласно тождеству (2.5.12)

$$\begin{aligned} * G_0 A G \left(\frac{\pi}{2\beta} \right) G_0 &= G \left(\frac{\pi}{2\beta} \right) A, \\ \left[\alpha G(1) + \beta G \left(1 + \frac{\pi}{2\beta} \right) \right] A G \left(\frac{\pi}{2\beta} \right) &= G \left(1 + \frac{\pi}{2\beta} \right). \end{aligned} \quad (2.5.33)$$

Поэтому уравнения замкнутой ОЭСС с ШИМ записываются в виде

$$\begin{aligned} \underline{\Delta}(n+1) &= [e^\alpha G(1) - G_0] A \underline{\Delta}^{(1)}(n) + \left\{ 1 + \lambda \left[e^\alpha G \left(1 + \frac{\pi}{2\beta} \right) - \right. \right. \\ &- \left. \left. G \left(\frac{\pi}{2\beta} \right) \right] A \right\} \underline{\Delta}(n), \\ \underline{\Delta}^{(1)}(n+1) &= e^\alpha G(1) G_0 \underline{\Delta}^{(1)}(n) + \lambda e^\alpha G \left(1 + \frac{\pi}{2\beta} \right) G_0 \underline{\Delta}(n). \end{aligned} \quad (2.5.34)$$

Для получения характеристического многочлена перейдем к скалярной форме записи уравнений (2.5.34) подобно тому, как это было выпол-

нено для случая с АИМ. Используя обозначения (2.5.14), имеем

$$\begin{aligned}
 \Delta_{1T}(n+1) &= (1 + \lambda M_{20}) \Delta_{1T}(n) - \lambda M_{10} \Delta_{2T}(n) + M_{10} \Delta_{1T}^{(1)}(n) + \\
 &+ M_{20} \Delta_{2T}^{(1)}(n), \\
 \Delta_{2T}(n+1) &= \lambda M_{10} \Delta_{1T}(n) + (1 + \lambda M_{20}) \Delta_{2T}(n) - \\
 &- M_{20} \Delta_{1T}^{(1)}(n) + M_{10} \Delta_{2T}^{(1)}(n), \\
 \Delta_{1T}^{(1)}(n+1) &= -\lambda e^\alpha \sin \beta \cdot \Delta_{1T}(n) - \lambda e^\alpha \cos \beta \cdot \Delta_{2T}(n) + \\
 &+ e^\alpha \cos \beta \cdot \Delta_{1T}^{(1)}(n) - e^\alpha \sin \beta \cdot \Delta_{2T}^{(1)}(n), \\
 \Delta_{2T}^{(1)}(n+1) &= \lambda e^\alpha \cos \beta \cdot \Delta_{1T}(n) - \lambda e^\alpha \sin \beta \cdot \Delta_{2T}(n) + \\
 &+ e^\alpha \sin \beta \cdot \Delta_{1T}^{(1)}(n) + e^\alpha \cos \beta \cdot \Delta_{2T}^{(1)}(n).
 \end{aligned} \tag{2.5.35}$$

Характеристический многочлен, вытекающий из системы (2.5.35), имеет вид

$$\det \begin{vmatrix} 1 + \lambda M_{20} - e^q & -\lambda M_{10} & M_{10} & M_{20} \\ \lambda M_{10} & 1 + \lambda M_{20} - e^q & -M_{20} & M_{10} \\ -\lambda e^\alpha \sin \beta & -\lambda e^\alpha \cos \beta & e^\alpha \cos \beta - e^q & -e^\alpha \sin \beta \\ \lambda e^\alpha \cos \beta & -\lambda e^\alpha \sin \beta & e^\alpha \sin \beta & e^\alpha \cos \beta - e^q \end{vmatrix} = \\
 = d_0 e^{4q} + d_1 e^{3q} + d_2 e^{2q} + d_3 e^q + d_4, \tag{2.5.36}$$

где

$$\begin{aligned}
 d_0 &= 1, \\
 d_1 &= -2(e^\alpha \cos \beta + 1 + \lambda M_{20}), \\
 d_2 &= (1 + \lambda M_{20})^2 + \lambda^2 M_{10}^2 + e^{2\alpha} + 4(1 + \lambda M_{20})e^\alpha \cos \beta + \\
 &+ 2\lambda(M_{10} \sin \beta - M_{20} \cos \beta)e^\alpha, \\
 d_3 &= -2[(1 + \lambda M_{20})^2 + \lambda^2 M_{10}^2]e^\alpha \cos \beta - 2(1 + \lambda M_{20})e^{2\alpha} - \\
 &- 2\lambda(1 + \lambda M_{20} + e^\alpha \cos \beta)(M_{10} \sin \beta - M_{20} \sin \beta)e^\alpha - \\
 &- 2\lambda(M_{10} \cos \beta + M_{20} \sin \beta)(e^\alpha \sin \beta - \lambda M_{10})e^\alpha, \\
 d_4 &= [(1 + \lambda M_{20})^2 + \lambda^2 M_{10}^2]e^{2\alpha} + 2\lambda(M_{10} \sin \beta - M_{20} \cos \beta)e^{2\alpha} \cos \beta - \\
 &- 2\lambda(1 + \lambda M_{20})(M_{10} \cos \beta + M_{20} \sin \beta)e^{2\alpha} \sin \beta - \\
 &- 2\lambda^2 M_{10}(M_{10} \cos \beta + M_{20} \sin \beta)e^{2\alpha} \cos \beta - \\
 &- 2\lambda^2 M_{10}(M_{10} \sin \beta - M_{20} \cos \beta)e^{2\alpha} \sin \beta + \lambda^2 e^{2\alpha}(M_{10}^2 + M_{20}^2).
 \end{aligned} \tag{2.5.37}$$

В результате вычислений можно построить границу области устойчивости в плоскости параметров $\{\lambda, T\}$ или $\{k, T\}$. Она изображена на рис. 2.12, где ее можно сравнить с ранее построенной границей в случае АИМ. Видно, что ОЭСС с ШИМ имеет более глубокие провалы в области резонансных частот, чем система с АИМ. Это объясняется тем, что короткие импульсы в ОЭСС с ШИМ порождают нутационные колебания гироспривода, в то время как маломеняющийся управляющий сигнал в ОЭСС с АИМ вызывает сравнительно плавное прецессионное движение. На рис. 2.13 приведены для сравнения фазовые портреты движения импульсной ОЭСС с обон-

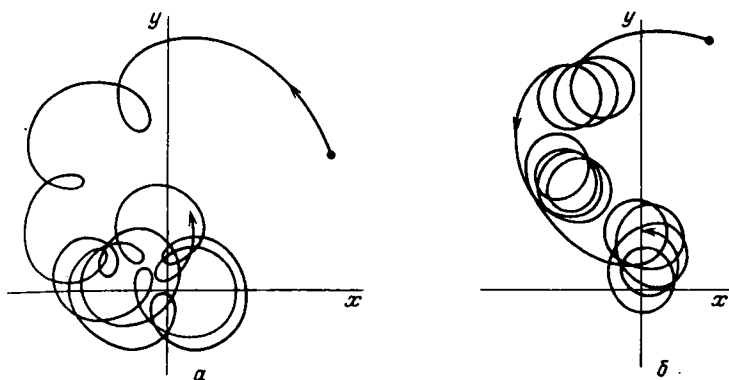


Рис. 2.13. Фазовые портреты импульсной оптико-электронной следящей системы: *а* — с амплитудно-импульсной модуляцией; *б* — с широтно-импульсной модуляцией

ми видами модуляций, на которых хорошо заметно качественное различие поведения гироспривода.

При потере устойчивости положения равновесия в импульсной ОЭСС возникают колебания, которые отличаются большим разнообразием. В случае простых автоколебаний в фазовой плоскости $\{\Delta_1, \Delta_2\}$ возникает устойчивый предельный цикл, при котором дискретные значения $[\Delta_1(n) \Delta_2(n)]$ располагаются на окружности. Радиус Δ_0 этой окружности легче всего определить при математическом моделировании, причем выражения (2.5.15) и (2.5.35) и здесь оказываются весьма полезными, так как позволяют осуществить счет на ЭВМ с произвольным шагом, например с шагом T . Без этих уравнений счет по методу Эйлера потребовал бы весьма малого шага счета. Некоторого увеличения шага можно добиться при использовании процедуры Рунге—Кутты, обращение к которой само по себе требует повышенных затрат машинного времени. В этом смысле использование дискретных выражений позволяет сократить время счета в сотни раз по сравнению с обычными процедурами.

При нарушении условий устойчивости предельного цикла колебания приобретают более сложный характер, причем в управляющем сигнале возникают низкочастотные составляющие, подавить которые в ряде случаев не представляется возможным без потери динамической точности сопровождения. Для выяснения условий существования устойчивых предельных циклов могут быть составлены разностные уравнения в отклонениях от исследуемого цикла. Однако параметры самого цикла определяются лишь численно, что ставит под сомнение целесообразность составления и исследования уравнений в отклонениях. Отметим также, что использование принципа многоплощадочного приемника в ОЭСС с дискретным управлением не позволяет добиться столь же высокой добротности следящего контура, как в ОЭСС со сканированием методом переноса изображения. Поэтому автоколебательные режимы в дискретных ОЭСС не являются рабочими, а свидетельствуют о нарушении номинальных условий сопровождения. В связи с этим исследование формы возможных автоколебаний в дискретных ОЭСС не столь актуально, как аналогичное исследование, выполненное в § 2.4.

§ 2.6. Анализ процесса сопровождения при двойном источнике излучения

Среди различных видов помех, действующих в оптическом диапазоне, следует выделить помеху в виде ложного источника излучения [147, 153, 155, 156, 171]. В качестве простой модели действия помехи рассмотрим двойной источник излучения, сопровождаемый ОЭСС с непрерывным управлением и с анализатором изображения в виде сканирующего устройства с неподвижным растром [147].

Предположим, что условия модуляции по несущей частоте (см. схему на рис. 2.2) сохраняются по всему полю раstra. Такое допущение хотя и не соответствует действительным условиям модуляции (см., например, растр на рис. 3.11), но позволяет сделать некоторые качественные выводы о процессе слежения за двойным источником. Кроме того, справедливость указанного допущения проверяется при математическом моделировании по методике, изложенной в главе 3.

Допущение о независимости условий модуляции от вектора углового рассогласования соответствует равномерному заполнению сигнала на входе усилителя несущей. Из функционального состава тракта вторичной обработки (см. рис. 2.2) следует, что управляющий сигнал на входе исполнительного элемента пропорционален амплитуде огибающей. Огибающая

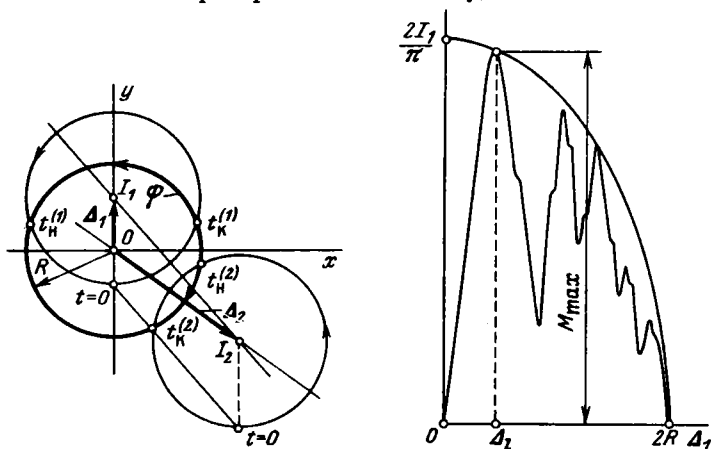


Рис. 2.14. Взаимное положение источников излучения I_1 и I_2

Рис. 2.15. Верхняя оценка пеленгационной характеристики

представляет собой при нашем допущении первую гармонику разложения в ряд Фурье последовательности прямоугольных импульсов (пачек, заполненных частотой несущей). На рис. 2.14 показано расположение двух источников излучения и следа оптической оси ОЭСС в фокальной плоскости. Для определенности положим, что радиус окружности переноса изображения совпадает с радиусом R раstra (на рис. 2.14 растр показан более жирной окружностью с центром в точке O). Угловую частоту переноса обозначим ω_0 (частота огибающей). Очевидно, что сигнал на фотоприемнике при точечном изображении сохраняет постоянное (отлич-

ное от нуля) значение при прохождении дуги окружности переноса от точки $t_{\text{н}}^{(1)}$ до точки $t_{\text{к}}^{(1)}$. На участке $[t_{\text{к}}^{(1)}, t_{\text{н}}^{(1)}]$ он равен нулю, и сигнал повторяется через $2k\pi/\omega_0$ [с] ($k = 0, 1, 2, \dots$).

Амплитуда первой гармоники ввиду выбора начала отсчета $t = 0$ находится интегрированием:

$$a_1 = \frac{\omega_0}{\pi} \int_{\omega_0 t_{\text{н}}^{(1)}}^{\omega_0 t_{\text{к}}^{(1)}} I_1 \cos \omega_0 t dt = \frac{I_1}{\pi} (\sin \omega_0 t_{\text{к}}^{(1)} - \sin \omega_0 t_{\text{н}}^{(1)}), \quad (2.6.1)$$

где I_1 — постоянная величина сигнала на участке $[t_{\text{н}}^{(1)}, t_{\text{к}}^{(1)}]$.

Ввиду симметрии относительно точки $t = 0$ можно положить $t_{\text{н}}^{(1)} = -t_1$, $t_{\text{к}}^{(1)} = t_1$ и

$$a_1 = \frac{2I_1}{\pi} \sin \omega_0 t_1. \quad (2.6.2)$$

С другой стороны, из рис. 2.14 видно, что $\Delta_1 = 2R \cos \omega_0 t_1$ и, следовательно

$$a_1(\Delta_1) = \frac{2I_1}{\pi} \sqrt{1 - \left(\frac{\Delta_1}{2R}\right)^2}, \quad |\Delta_1| \leq 2R. \quad (2.6.3)$$

Соотношение (2.6.3) представляет собой верхнюю оценку пеленгационной характеристики. На рис. 2.15 представлены реальная пеленгационная характеристика и ее верхняя оценка, приведенные к одному масштабу. Реальная характеристика совпадает со своей верхней оценкой в конце линейной зоны $\Delta_1 = \Delta_l$, где она равна M_{max} и где имеет место максимальная глубина модуляции по несущей.

Очевидно, что сигнал на выходе усилителя огибающей при наших предположениях о постоянстве условий модуляции по полю раstra и полном подавлении гармоник с частотой выше ω_0 представляет собой гармонику вида

$$u_1(t) = \frac{2I_1 k}{\pi} \sqrt{1 - \left(\frac{\Delta_1}{2R}\right)^2} \cos \omega_0 t, \quad (2.6.4)$$

где k — коэффициент усиления тракта обработки.

Рассмотрим теперь вид сигнала на выходе усилителя огибающей при наличии второго источника с постоянной величиной I_2 и вектором рассогласования Δ_2 . Легко видеть, что при неизменном расстоянии между источниками линия $I_1 I_2$ переносится в фокальной плоскости поступательно. Поэтому при $t = 0$ изображение второго источника занимает крайнее нижнее положение на своей траектории переноса. Середина пачки при этом получает сдвиг по фазе, равный ψ — углу между векторами Δ_1 и Δ_2 . Формально этот факт можно установить из соотношений, вытекающих из рис. 2.14. Действительно, для момента начала пачки $t_{\text{н}}^{(1)}$ имеем

$$\frac{\Delta_1}{2} = R \cos \omega_0 t_{\text{н}}^{(1)}. \quad (2.6.5)$$

Для момента $t_H^{(2)}$ находим

$$\begin{aligned} \frac{\Delta_2}{2} \cos \psi + \sin \psi \sqrt{R^2 - \left(\frac{\Delta_2}{2}\right)^2} &= \Delta_2 \cos \psi + R \cos \omega_0 t_H^{(2)}, \\ \frac{\Delta_2}{2} \sin \psi - \cos \psi \sqrt{R^2 - \left(\frac{\Delta_2}{2}\right)^2} &= \Delta_2 \sin \psi + R \sin \omega_0 t_H^{(2)}. \end{aligned} \quad (2.6.6)$$

Умножая первое равенство (2.6.6) на $\cos \psi$, второе на $\sin \psi$ и складывая, получаем

$$-\frac{\Delta_2}{2} = R \cos(\omega_0 t_H^{(2)} - \psi), \quad (2.6.7)$$

что и указывает на сдвиг сигнала от второго источника на угол ψ относительно сигнала от первого источника. Соответственно получаем выражение для первой гармоники от второго источника

$$u_2(t) = \frac{2I_2 k}{\pi} \sqrt{1 - \left(\frac{\Delta_2}{2R}\right)^2} \cos(\omega_0 t - \psi). \quad (2.6.8)$$

Сделаем предположение о том, что пачки не перекрываются во времени, тогда совместное действие двух источников на исполнительный элемент ОЭСС можно вычислить как сумму

$$\begin{aligned} u_{\Sigma}(t) &= a_1(\Delta_1) \cos \omega_0 t + a_1(\Delta_2) \cos(\omega_0 t - \psi) = \\ &= [a_1(\Delta_1) + a_1(\Delta_2) \cos \psi] \cos \omega_0 t + a_1(\Delta_2) \sin \psi \sin \omega_0 t. \end{aligned} \quad (2.6.9)$$

Точки равновесия в поле сил коррекции. Поставим вопрос о наличии равновесных положений оптической оси ОЭСС в плоскости Oxy . Если такие положения есть, то, как это следует из (2.6.9), в этих точках

$$[a_1(\Delta_1) + a_1(\Delta_2) \cos \psi]^2 + a_1^2(\Delta_2) \sin^2 \psi = 0. \quad (2.6.10)$$

Из (2.6.10) следует, что угол ψ должен при этом удовлетворять условию

$$\cos \psi = -\frac{a_1^2(\Delta_1) + a_1^2(\Delta_2)}{2a_1(\Delta_1)a_1(\Delta_2)}. \quad (2.6.11)$$

Но из очевидного неравенства $[a_1(\Delta_1) - a_1(\Delta_2)]^2 \geq 0$ следует, что

$$a_1^2(\Delta_1) + a_1^2(\Delta_2) \geq 2a_1(\Delta_1)a_1(\Delta_2),$$

поэтому единственным возможным решением (2.6.11) служит $\cos \psi = -1$, откуда $\psi = (2k + 1)\pi$ ($k = 0, 1, 2, \dots$). По смыслу задачи можно ограничиться значением $\psi = \pi$. Таким образом, обнаружено, что на всей плоскости Oxy есть единственная точка равновесия, которая находится на прямой, соединяющей оба источника. В фокальной плоскости это прямая, соединяющая центры окружностей переноса.

Найдем расстояния точки равновесия от обоих источников. Полагая в (2.6.10) $\psi = \pi$ и учитывая (2.6.3), получаем

$$I_1 \sqrt{1 - \left(\frac{\Delta_{10}}{2R}\right)^2} = I_2 \sqrt{1 - \left(\frac{\Delta_{20}}{2R}\right)^2}, \quad (2.6.12)$$

откуда

$$\left(\frac{I_1}{I_2}\right)^2 = \frac{4R^2 - \Delta_{20}^2}{4R^2 - \Delta_{10}^2}. \quad (2.6.13)$$

При $I_1 > I_2$ следует $4R^2 - \Delta_{20}^2 > 4R^2 - \Delta_{10}^2$, т.е. точка равновесия располагается ближе к более слабому источнику I_2 . При $I_1 < I_2$ имеем $\Delta_{10} < \Delta_{20}$.

Для определения положения точки равновесия на прямой $I_1 I_2$ введем параметр $\kappa = \Delta_{10}/\Delta_{20}$ и обозначение $\Delta_\pi = \Delta_{10} + \Delta_{20}$. Тогда $\Delta_{10} = \kappa \Delta_\pi / (1 + \kappa)$ и $\Delta_{20} = \Delta_\pi / (1 + \kappa)$, а искомый параметр

$$\kappa = \frac{\Delta_\pi \sqrt{4R^2(I_1^2 - I_2^2) + I_1^2 I_2^2 \Delta_\pi^2 - 4R^2(I_1^2 - I_2^2)}}{4R^2(I_1^2 - I_2^2) + \Delta_\pi^2 I_2^2}. \quad (2.6.14)$$

Покажем теперь, что точка равновесия, определяемая значением параметра κ (2.6.14), является неустойчивой. Для этого полезно рассматривать зависимость $a_1(\Delta_1)$ (2.6.3) и аналогичную ей зависимость $a_1(\Delta_2)$ как составляющие вектора коррекционного усилия, действующего на следящий координатор. Таким образом, при $\psi = \pi$ суммарное коррекционное усилие согласно (2.6.10) находится, как

$$f = a_1(\Delta_1) - a_1(\Delta_2) = \frac{2k}{\pi} \left[I_1 \sqrt{1 - \left(\frac{\Delta_1}{2R}\right)^2} - I_2 \sqrt{1 - \left(\frac{\Delta_2}{2R}\right)^2} \right]. \quad (2.6.15)$$

При $\Delta_1 = \Delta_{10}$, $\Delta_2 = \Delta_{20}$ имеем $f = 0$.

Положим теперь $\Delta_1 = \Delta_{10} + \delta$, $\Delta_2 = \Delta_{20} - \delta$. Тогда коррекционное усилие также получит приращение Δf :

$$\Delta f = \frac{2k}{\pi} \left(I_1 \sqrt{1 - \frac{(\Delta_{10} + \delta)^2}{4R^2}} - I_2 \sqrt{1 - \frac{(\Delta_{20} - \delta)^2}{4R^2}} \right).$$

* Составим отношение

$$\begin{aligned} \frac{\Delta f}{\delta} &= \frac{2k}{\pi \delta} \left[I_1 \sqrt{1 - \frac{(\Delta_{10} + \delta)^2}{4R^2}} - I_2 \sqrt{1 - \frac{(\Delta_{20} - \delta)^2}{4R^2}} \right] = \\ &= - \frac{k}{\pi \delta} \frac{I_1^2 \Delta_{10} \delta + I_2^2 \Delta_{20} \delta + (I_1^2 - I_2^2) \delta^2 / 2}{I_1 \sqrt{R^2 - (\Delta_{10} + \delta)^2 / 4} + I_2 \sqrt{R^2 - (\Delta_{20} - \delta)^2 / 4}}. \end{aligned} \quad (2.6.16)$$

Переходя в (2.6.16) к пределу при $\delta \rightarrow 0$ видим, что $\lim_{\delta \rightarrow 0} \frac{\Delta f}{\delta} < 0$. Это означает, что при $\delta > 0$ коррекционное усилие приобретает приращение, действующее в сторону источника I_2 , и, наоборот, при $\delta < 0$ возникает усилие, направленное в сторону источника I_1 . Итак, точка с координатами Δ_{10} и Δ_{20} , определяемыми параметром κ (2.6.14), является единственной точкой равновесия следящего координатора на плоскости Oxy , и притом неустойчивой.

Построение поля сил коррекции. Представляется интересной общая картина коррекционных усилий по всей плоскости Oxy . Для ее построения воспользуемся силовыми линиями. Силовой линией, как обычно, будем считать кривую, в каждой точке которой вектор коррекционного усилия направлен по касательной к ней. Для вывода уравнения произвольной си-

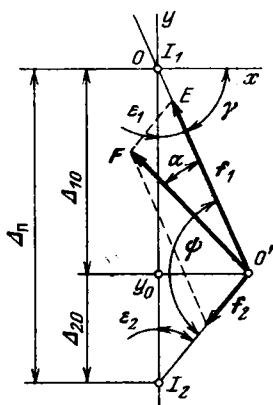


Рис. 2.16. Построение силовых линий в поле коррекционных усилий

ловой линии обратимся к рис. 2.16, на котором начало координат (x, y) помещено в точку I_1 и оси координат рассматриваются как оси управления по двум каналам ОЭСС. Текущее положение оптической оси координатора совпадает с точкой O' с координатами (x, y) . Уравнение силовой линии $y = y(x)$ по определению должно удовлетворять дифференциальному уравнению

$$y' = \operatorname{tg}(\gamma - \alpha) = \frac{\operatorname{tg} \gamma - \operatorname{tg} \alpha}{1 + \operatorname{tg} \gamma \operatorname{tg} \alpha}, \quad (2.6.17)$$

где $\operatorname{tg} \gamma = y/x$, а для определения $\operatorname{tg} \alpha$ рассмотрим треугольник EFO' , из которого следует, что

$$\frac{f_1}{\sin(\psi - \alpha)} = \frac{f_2}{\sin \alpha}. \quad (2.6.18)$$

Здесь модули векторов коррекционных усилий определяются выражениями вида (2.6.3). Таким образом, из (2.6.18) получаем

$$\operatorname{tg} \alpha = \frac{I_2 \sin \psi \sqrt{4R^2 - \Delta_2^2}}{I_2 \cos \psi \sqrt{4R^2 - \Delta_2^2} + I_1 \sqrt{4R^2 - \Delta_1^2}}. \quad (2.6.19)$$

Чтобы записать зависимость (2.6.19) в явной форме как функцию двух координат фазовой плоскости, достаточно воспользоваться очевидными соотношениями

$$\begin{aligned} \Delta_1^2 &= x^2 + y^2, \\ \Delta_2^2 &= x^2 + (\Delta_n - y)^2, \\ \cos \psi &= \frac{x^2 + y^2 - \Delta_n y}{\sqrt{(x^2 + y^2)[x^2 + (\Delta_n - y)^2]}}, \end{aligned} \quad (2.6.20)$$

вытекающими из рис. 2.16.

Обозначая для краткости $\operatorname{tg} \alpha = A(x, y)$, запишем дифференциальное уравнение силовой линии (2.6.17) в форме

$$y' = \frac{y - xA(x, y)}{x + yA(x, y)}. \quad (2.6.21)$$

Структура выражений (2.6.19) и (2.6.20) такова, что интегрирование уравнений (2.6.21) в конечной форме затруднительно, поэтому для получения искомой картины силового поля в фазовой плоскости следует использовать численные методы интегрирования. Уравнение (2.6.21), однако, допускает качественное исследование.

Наиболее интересным представляется поведение силовых линий в окрестности особых точек. Одной из таких точек является положение равновесия. При выбранной системе координат с началом в точке I_1 в точке равно-

весия положим $x = x_0 = 0$, $y = y_0$. Исследуем поведение функции $A(x, y) = \operatorname{tg} \alpha$. Так как в окрестности исследуемой точки $\psi \rightarrow \pi$, то $\sin \psi \rightarrow 0$. Из рис. 2.16 видно, что при $x \rightarrow 0$

$$\psi = \pi - \left(\frac{x}{y_0} + \frac{x}{\Delta_n - y_0} \right) + o(x) = \pi - \frac{x \Delta_n}{y_0(\Delta_n - y_0)} + o(x),$$

т. е.

$$\sin \psi = \frac{x \Delta_n}{y_0(\Delta_n - y_0)} + o(x). \quad (2.6.22)$$

Знаменатель выражения (2.6.19) уже исследован нами в форме (2.6.16). где следует δ заменить на y . Таким образом, функция $A(x, y)$ в окрестности точки равновесия может быть записана с учетом (2.6.16) и (2.6.22)

в форме $A_0 \frac{x}{y}$, и уравнение (2.6.21) в окрестности этой точки эквивалентно выражению

$$y' \sim \frac{y}{A_0 x}, \quad (2.6.23)$$

где $A_0 < 0$. Интегрируя (2.6.23), получаем семейство гипербол вида $y = Cx^{1/A_0}$, т. е. неустойчивую особую точку типа седла.

Кроме точки равновесия в фазовой плоскости Oxy имеются еще две особых точки I_1 и I_2 . Для их исследования уже нельзя пользоваться верхними оценками пеленгационных характеристик (2.6.3), так как в окрестности этих точек мы попадаем в зону линейности пеленгационной характеристики. Полагая в (2.6.21) $A(x, y) = 0$, получаем в результате интегрирования получившегося соотношения $y = Cx$ — особую точку типа дикритического узла. Легко видеть, что этот узел устойчив. На рис. 2.17 показан результат численного интегрирования уравнения (2.6.21), находящийся в соответствии с изложенным качественным исследованием. Здесь видна граница, разделяющая семейство силовых линий на два класса, каждый из которых сходится к своей устойчивой особой точке.

При использовании для пеленгационных характеристик их оценок вида (2.6.3) разделительная граница имеет форму окружности, уравнение которой получается при подстановке в (2.6.12) двух первых соотношений (2.6.20). Указанная качественная картина силового поля может быть несколько скорректирована, если при ее построении учесть угол расфазировки ϵ из уравнений (2.4.5). Наличие угла расфазировки вызывает поворот вектора коррекционного усилия в каждой точке фазовой плоскости на постоянный угол ϵ . Легко проверить, что дикритические узлы в точках I_1 и I_2 при этом переходят в устойчивые фокусы. Некоторое изменение формы претерпевает и граничная траектория, которая, в частности, теряет симметрию относительно оси Oy (см. рис. 2.17, б).

Для проверки справедливости допущений, принятых при качественном исследовании поля коррекционных усилий, следует использоватьazole-

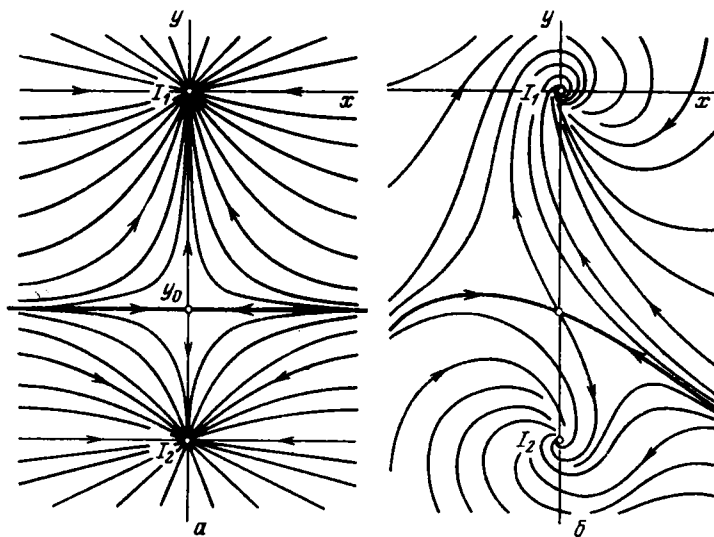


Рис. 2.17. Поле коррекционных усилий: *a* — при расфазировке $\epsilon = 0$; *б* — при расфазировке $\epsilon \neq 0$

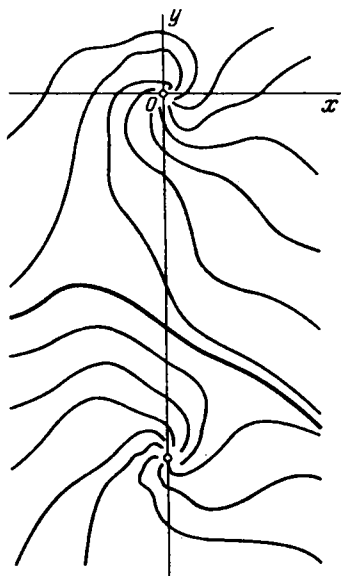


Рис. 2.18. Поле коррекционных усилий, построенное с помощью поэлементной модели оптико-электронной следящей системы

ментную модель ОЭСС по методике, описанной в главе 3. При этом результат моделирования не опирается на какие-либо допущения об условиях модуляции, неперекрывающихся во времени пачках импульсов или геометрических размерах источника. Для построения реальной картины силового поля используется разомкнутая схема (см. рис. 2.1). На вход схемы вводится рассогласование и после окончания переходных процессов в каналах обработки с выхода фазовых детекторов выводятся значения составляющих коррекционного вектора.

Для повышения точности построения поля в программе следует предусмотреть изменение величины шага по рассогласованию в зависимости от кривизны силовой линии. Построенное этим методом поле (рис. 2.18) служит объективным критерием для оценки качества первичной и вторичной обработки информации. Следует отметить, что получение такой характеристики экспериментальным путем без использования цифровой модели затруднительно. Основные трудности связаны с регистрацией угловых координат источников излучения и с фильтрацией шумов, загрязняющих сигналы в каналах управления. Большие трудности вызывает обеспечение стабильности условий проведения эксперимента.

Покажем теперь, что полученная картина силового поля играет основную роль при анализе процесса слежения ОЭСС за двойным источником излучения. При анализе периодических режимов в ОЭСС с непрерывным управлением не рассматривалось действие внутренних шумов электронного тракта. Обработка результатов моделирования показывает, что влияние шумов выражается в том, что линейчатый спектр сигналов в каналах управления, соответствующий автоколебаниям, становится непрерывным. Спектр управляющего сигнала при этом несколько расширяется. Кроме того, угловые флуктуации приобретают нерегулярный характер из-за сложной геометрии раstra анализатора (см. рис. 3.11). Совокупность этих физических факторов приводит к тому, что фазовые траектории в плоскости $\{\Delta_1, \Delta_2\}$ образуют двумерный случайный процесс, который можно рассматривать как результат наложения некоторого шума на периодические колебания. Соотношение между случайной и регулярной составляющими зависит от конкретной геометрии раstra, расфазировки и параметров электронного тракта и исполнительных элементов ОЭСС.

Обобщенная модель сопровождения двойного источника. В § 2.4 было показано, что при некотором значении обобщенного параметра $f_0 T \cos \epsilon / 2 \xi$ в ОЭСС с непрерывным управлением возможен устойчивый круговой предельный цикл. В соответствии со сделанными замечаниями будем рассматривать отклонения фазовых траекторий от этого цикла как результат случайного возмущения. На основании уравнений (2.4.5) составим математическую модель ОЭСС при действии случайных возмущений и слежении за двойным источником. Непосредственное использование уравнений вида (2.4.5) для статистического исследования ОЭСС не представляется возможным, и мы перейдем здесь к концептуальной модели слежения за двойным источником, отражающей основные особенности процесса слежения, но не повторяющей структуру уравнений (2.4.5) [143].

При анализе уравнений автоколебательной ОЭСС вида (2.4.5) было установлено, что они допускают формальное решение $\sqrt{\Delta_1^2 + \Delta_2^2} = \Delta_0 = \text{const.}$ В качестве простейшей модели, допускающей аналогичное решение, рассмотрим систему

$$\dot{\Delta} = -\lambda(\Delta - \Delta_0) + n(t), \quad \dot{\varphi} = \omega_0, \quad (2.6.24)$$

которая при $n(t) \equiv 0$ обладает решением $\Delta = \Delta_0$, $\varphi = \omega_0 t + \varphi_0$ или, при переходе к прямоугольным координатам, – решением вида (2.4.6).

В уравнениях (2.6.24) Δ и φ – полярные координаты изображающей точки в фазовой плоскости, поэтому $\Delta \geq 0$ – неотрицательная величина.

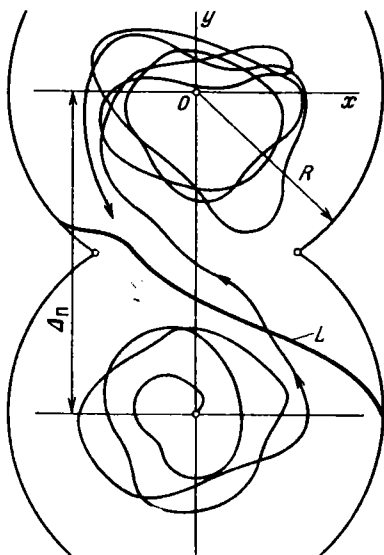


Рис. 2.19. Переход фазовой траектории через границу L

При $\Delta = 0$ изображающая точка находится в начале координат системы Oxy (рис. 2.19). В установившемся состоянии при $\dot{\Delta} = 0$ изображающая точка описывает предельный цикл радиуса Δ_0 . При наличии шума $n(t)$ фазовые траектории флуктуируют относительно этого предельного цикла. В случайные моменты времени фазовые траектории могут достигать границы L , разделяющей поле коррекционных сил, после чего слежение за источником с координатами $(0, 0)$ прекращается и ОЭСС переходит на сопровождение источника с координатами $(0, \Delta_\pi)$. При слежении за вторым источником вид уравнений (2.6.24) не меняется. Меняется лишь смысл полярного радиуса Δ , которые отсчитываются

от нового начала координат. Если $n(t)$ — белый шум, то координата $\Delta(t)$ представляет собой марковский случайный процесс, плотность распределения которого подчиняется прямому уравнению Колмогорова [5], известному также как уравнение Фоккера—Планка:

$$\frac{\partial f(t, x; \tau, y)}{\partial \tau} = - \frac{\partial}{\partial y} [a(\tau, y) f(t, x; \tau, y)] + \frac{1}{2} \frac{\partial^2}{\partial y^2} [b(\tau, y) f(t, x; \tau, y)]. \quad (2.6.25)$$

Здесь $f(t, x; \tau, y) = \frac{\partial}{\partial y} F(t, x; \tau, y)$ — плотность распределения и

$F(t, x; \tau, y) = P\{\Delta(\tau) < y \mid \Delta(t) = x\}$ — условная вероятность события $\Delta(\tau) < y$ при условии, что в момент времени $t < \tau$ имеет место равенство $\Delta(t) = x$. Коэффициенты $a(\tau, y)$ и $b(\tau, y)$ определяются из уравнений (2.6.24), первое из которых, следуя общепринятой терминологии, назовем стохастическим.

Стохастическое уравнение (2.6.24) удобно тем, что оно помимо теоретического анализа с помощью аппарата марковских случайных процессов допускает прямое моделирование по методу Монте-Карло [5] с выводом результатов на дисплей или графопостроитель для визуального контроля получаемых фазовых портретов. Плавность фазовых траекторий регулируется параметром λ , который, как это видно из первого уравнения (2.6.24), играет роль частоты среза (обратной постоянной времени линей-

ного апериодического фильтра). Спектр процесса $\Delta(t)$ равен

$$S_{\Delta}(\omega) = \frac{\sigma_0^2}{\omega^2 + \lambda^2}, \quad (2.6.26)$$

и если из экспериментальных исследований известна соответствующая характеристика при слежении за одним источником, то параметр λ можно подобрать либо по критерию ширины спектра, либо по другой процедуре идентификации динамических систем [5, 143].

Нашей ближайшей задачей является решение уравнения (2.6.25) относительно плотности распределения процесса $\Delta(t)$. При этом также ставится задача об определении времени устойчивого сопровождения каждого из источников излучения.

Вычислим коэффициенты $a(\tau, y)$ и $b(\tau, y)$, входящие в уравнение (2.6.25). Обозначая через ΔY приращение процесса Δ , т.е. $\Delta Y = \Delta(\tau) - \Delta(t)$, имеем

$$a(\tau, y) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} M[\Delta Y] = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} M\left[\int_t^{t+\Delta t} \lambda(Y - \Delta_0) d\theta\right],$$

откуда, пользуясь теоремой о среднем значении и переходя к пределу при $\Delta t = \tau - t \rightarrow 0$, получаем

$$a(\tau, y) = \lambda(y - \Delta_0). \quad (2.6.27)$$

При вычислении коэффициента $b(\tau, y)$ будем предполагать, что $n(t)$ — белый шум с корреляционной функцией $\sigma_0^2 \delta(t)$. Имеем

$$\begin{aligned} b(\tau, y) &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} M[\Delta Y^2] = \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} M\left\{\int_t^{t+\Delta t} [\lambda(u - \Delta_0) + n(u)] du \times \right. \\ &\quad \times \left. \int_t^{t+\Delta t} [\lambda(v - \Delta_0) + n(v)] dv\right\} = \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_t^{t+\Delta t} \int_t^{t+\Delta t} M[n(u)n(v)] du dv = \sigma_0^2. \end{aligned} \quad (2.6.28)$$

Таким образом, вместо (2.6.25) запишем

$$\begin{aligned} \frac{\partial f(t, x; \tau, y)}{\partial \tau} &= \frac{\partial}{\partial y} [\lambda(y - \Delta_0) f(t, x; \tau, y)] + \\ &+ \frac{\sigma_0^2}{2} \frac{\partial^2}{\partial y^2} f(t, x; \tau, y). \end{aligned} \quad (2.6.29)$$

Стационарное решение $f(y) = \lim_{\tau \rightarrow \infty} f(t, x; \tau, y)$ находится из обыкновенного дифференциального уравнения

$$\frac{\sigma_0^2}{2} \frac{df}{dy} + \lambda(y - \Delta_0)f = 0, \quad (2.6.30)$$

откуда

$$f(y) = Ce^{-\frac{\lambda(y - \Delta_0)^2}{\sigma_0^2}} \quad (y > 0). \quad (2.6.31)$$

Из условия нормировки находится постоянная интегрирования

$$C = \frac{2}{\sigma_0} \sqrt{\frac{\lambda}{\pi}} \frac{1}{1 + \Phi(\Delta_0 \sqrt{2\lambda})},$$

где

$$\Phi(u) = \frac{1}{\sigma_0 \sqrt{2\pi}} \int_0^u e^{-\frac{x^2}{2\sigma_0^2}} dx. \quad (2.6.32)$$

Полученное стационарное решение может служить также признаком для идентификации концептуальной модели (2.6.24) с полной моделью (2.4.5); а также с результатами экспериментального исследования реальной ОЭСС при сопровождении одного источника. При этом подбираются параметр Δ_0 , характеризующий положение максимума плотности $f(y)$, и параметр σ_0^2 , определяющий дисперсию найденной плотности.

Нестационарное решение и оценка помехозащищенности. Перейдем теперь к поиску нестационарного решения уравнения (2.6.29). Это уравнение записано без учета краевых условий. Одно из этих условий является условием отражения фазовых траекторий от границы $\Delta = 0$. Формально оно записывается в виде

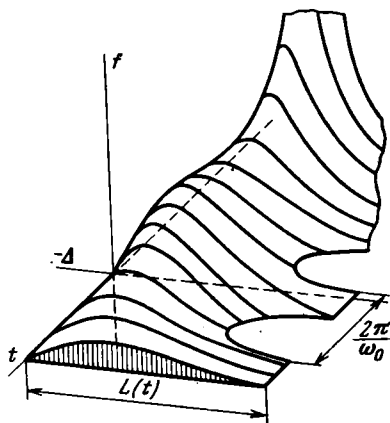
$$\left[a(t, x)f(t, x) - \frac{1}{2} b(t, x) \frac{\partial f(t, x)}{\partial x} \right] \bigg|_{x=0} = 0 \quad (2.6.33)$$

и выражает отсутствие потока фазовых траекторий через отражающую границу $\Delta = 0$. В нашем случае это условие имеет вид

$$\left[\lambda(x - \Delta_0)f(t, x) - \frac{\sigma_0^2}{2} \frac{\partial f(t, x)}{\partial x} \right] \bigg|_{x=0} = 0. \quad (2.6.34)$$

Второе краевое условие заключается в поглощении фазовых траекторий на границе области, содержащей источник излучения (рис. 2.20).

Рис. 2.20. Нестационарная плотность распределения ошибки



Граница состоит из двух участков: участка, ограничивающего поле зрения ОЭСС, и участка, разделяющего поле сил коррекции. Формально условие поглощения записывается в виде

$$f(t, x) |_{x \in L} = 0. \quad (2.6.35)$$

Решение уравнения (2.6.29) с краевыми условиями (2.6.34) и (2.6.35) аналитическими методами в настоящее время не разработано. В связи с этим воспользуемся методами вычислительной математики решения краевых задач для уравнений в частных производных [12].

Основная идея применения численных методов состоит в разбиении области $\{t, x\}$ прямыми $t_i = ih$, $x_j = jh$ ($i = 0, 1, \dots; j = 0, \pm 1, \pm 2, \dots$). Точки пересечения этих прямых называются *узлами*. Узлы, все четыре соседних узла которых принадлежат области $\{t, x\}$, называются *внутренними*. Если хотя бы один из узлов в ближайшей окрестности данного узла не принадлежит $\{t, x\}$, то такой узел называется *граничным*. Для каждого внутреннего узла составим разностное уравнение вместо соответствующего дифференциального: либо прямого уравнения (2.6.25), либо обратного уравнения Колмогорова. Об использовании последнего будет сказано в дальнейшем. Производные заменяются конечными разностными соотношениями

$$\begin{aligned} \frac{\partial f(t, x)}{\partial h} &\Rightarrow \frac{f_{i, j+1} - f_{i, j-1}}{2h}, \\ \frac{\partial^2 f(t, x)}{\partial x^2} &\Rightarrow \frac{f_{i, j+1} - 2f_{i, j} + f_{i, j-1}}{h^2}. \end{aligned} \quad (2.6.36)$$

Что касается производных по времени, то для них можно записать разностные соотношения двух типов:

$$\frac{f_{i+1, j} - f_{i, j}}{l} \quad \text{либо} \quad \frac{f_{i, j} - f_{i-1, j}}{l}. \quad (2.6.37)$$

Первое соотношение (2.6.37) приводит нас к явной разностной схеме, второе — к неявной. При явной схеме из начальной совокупности $f_{0, j}$ ($j = 0, \pm 1, \pm 2, \dots$) последовательно получаются $f_{1, j}$, $f_{2, j}$, ... При этом возникает возможность потери точности из-за возрастающих погрешностей, которые в свою очередь связаны с приближенным характером соотношений (2.6.37) и с округлением при вычислениях. Неявная разностная схема

не страдает этим недостатком, но приводит к необходимости решения больших систем линейных алгебраических уравнений.

Мы рассмотрим здесь метод прогонки [12], хорошо зарекомендовавший себя при решении краевых задач. Для решения уравнения (2.6.25) кроме краевых условий вида (2.6.33) и (2.6.35) должно быть задано также начальное распределение $f(0, x) = \tilde{f}(x)$. Для большей общности рассмотрим краевые условия в форме

$$\left[\kappa_\nu(t) \frac{\partial f(t, x)}{\partial x} + \mu_\nu(t) f(t, x) \right] \Big|_{x \in L_\nu} = 0. \quad (2.6.38)$$

При $\kappa_\nu(t) \equiv 0$ получаем условие (2.6.35), при $\kappa_\nu(t) = \frac{1}{2} b(t, L)$, $\mu_\nu(t) = a(t, L)$ — условие (2.6.33).

Итак, запишем уравнение (2.6.25) в разностной форме с неявной разностной схемой:

$$\begin{aligned} \frac{f_{i,j} - f_{i-1,j}}{l} = & - \frac{a_{i,j} f_{i,j} - a_{i,j-1} f_{i,j-1}}{h} + \\ & + \frac{1}{2} \frac{b_{i,j+1} f_{i,j+1} - 2 b_{i,j} f_{i,j} + b_{i,j-1} f_{i,j-1}}{h^2}. \end{aligned} \quad (2.6.39)$$

Группируя значения плотности $f(t, x)$ с одинаковыми индексами, находим

$$A_{i,j+1} f_{i,j+1} - 2 B_{i,j} f_{i,j} + C_{i,j-1} f_{i,j-1} = - \frac{1}{l} f_{i-1,j}, \quad (2.6.40)$$

где

$$\begin{aligned} A_{i,j+1} &= \frac{1}{2 h^2} b_{i,j+1}, \\ B_{i,j} &= \frac{h^2 + h l a_{i,j} + l b_{i,j}}{2 h^2 l}, \\ C_{i,j-1} &= \frac{2 h a_{i,j-1} + b_{i,j-1}}{2 h^2}. \end{aligned} \quad (2.6.41)$$

Левое краевое условие аппроксимируем разностной формой

$$\kappa_{1,i} \frac{f_{i,0} - f_{i,-1}}{h} - \mu_{1,i} \frac{f_{i,-1} + f_{i,0}}{2} = 0. \quad (2.6.42)$$

Эту форму можно также записать в других обозначениях:

$$f_{i,-1} = P_{i,0} f_{i,0} + Q_{i,0}, \quad (2.6.43)$$

где

$$P_{i,0} = \frac{2\kappa_{1,i} - h\mu_{1,i}}{2\kappa_{1,i} + h\mu_{1,i}}; \quad Q_{i,0} = 0.$$

Функция $Q_{i,0}$ в данном случае введена для того, чтобы иметь возможность записать соотношение (2.6.43) для других значений индекса j , т. е.

$$f_{i,j-1} = P_{i,j}f_{i,j} + Q_{i,j}. \quad (2.6.44)$$

Для вычисления функций $P_{i,j}$ и $Q_{i,j}$ в формуле (2.6.44) подставим это выражение $f_{i,j-1}$ в (2.6.40) и получим

$$f_{i,j} = \frac{A_{i,j+1}}{2B_{i,j} - C_{i,j}P_{i,j}} f_{i,j+1} + \frac{1C_{i,j-1} + f_{i-1,j}}{1(2B_{i,j} - C_{i,j}P_{i,j})}, \quad (2.6.45)$$

откуда следуют рекуррентные выражения для вычисления функций $P_{i,j}$ и $Q_{i,j}$:

$$P_{i,j+1} = \frac{A_{i,j+1}}{2B_{i,j} - C_{i,j}P_{i,j}}, \quad (2.6.46)$$

$$Q_{i,j+1} = \frac{1C_{i,j-1} + f_{i-1,j}}{1(2B_{i,j} - C_{i,j}P_{i,j})}.$$

Процедуру последовательного вычисления $P_{i,j}$ и $Q_{i,j}$ с помощью соотношений (2.6.46) называют *прогонкой* левого граничного условия (2.6.42) в прямом направлении. Вычисленные значения $P_{i,j}$ и $Q_{i,j}$ ($j = -1, 0, 1, \dots, N$) запоминаются. Из правого граничного условия известно значение $f_{i,N}$. С помощью соотношения (2.6.43) находим

$$f_{i,N-1} = P_{i,N}f_{i,N} + Q_{i,N}$$

и затем $f_{i,N-2}$, $f_{i,N-3}$, ..., $f_{i,1}$. Этот алгоритм осуществляет прогонку правого граничного условия в обратном направлении. В качестве начального распределения можно при практических вычислениях использовать любую функцию $\tilde{f}(x)$, удовлетворяющую условию нормировки и равную нулю при $x > L$. Например, такой функцией может служить $f(x_0) = f(x_0 + h) = 1/h$ и равная нулю во всех остальных узлах.

Возвратимся к стохастическому уравнению (2.6.24), описывающему процесс слежения ОЭСС за двойным источником. Левым граничным условием в данном случае служит условие отражения (2.6.34), правым — (2.6.35). Расстояние до правой границы задается (согласно рис. 2.20) условием $L(t_i) = L_i$. Это несколько усложняет процедуру прогонки, так как при желании сохранить число узловых точек по координате Δ равным $N + 2 = \text{const}$ следует пользоваться интерполяцией плотности распределения при переходе от временного слоя с индексом i к слою $i + 1$.

На рис. 2.20 показан результат применения программы прогонки. В качестве начального распределения использовано нормальное распределение с достаточно малой дисперсией. При отсутствии второго источника в системе, определяемой уравнениями (2.6.24), устанавливается почти стационарное распределение, так как угловые флуктуации с очень малой вероятностью достигают границ поля зрения ОЭСС. При появлении второго источника на фазовой плоскости появляется разделяющая граница и функция $L(t)$ становится периодической, если расстояние между источниками остается постоянным. При изменении этого расстояния периодичность пропадает.

Вероятность перехода процесса слежения p_i с одного источника на другой определяется интегрированием текущей плотности $f_i(x)$, т.е.

$$p_i = \int_0^{L_i} f_i(x) dx. \quad (2.6.47)$$

Так как $f_i(x)$ задана в узловых точках на интервале $[0, L_i]$, то для вычисления интеграла (2.6.47) целесообразно воспользоваться алгоритмом Симпсона

$$\begin{aligned} p_i &= f_{-1} + 4f_0 + 2f_1 + \dots + 4f_{N-1} + f_N = \\ &= f_{-1} + f_N + 4 \sum_{k=0}^{N-1} f_{2k} + 2 \sum_{k=1}^{N-1/2} f_{2k-1}. \end{aligned} \quad (2.6.48)$$

На рис. 2.21 показан результат применения формулы (2.6.48) для случаев постоянного углового расстояния между источниками (кривая 1) и при возрастающем во времени расстоянии (кривая 2). Здесь хорошо заметно влияние периодичности граничного условия $\Delta = L_i$, наложенное на экспоненциальный в среднем закон убывания вероятности устойчивого сопровождения одного источника. При уменьшении расстояния L_i вероятность p_i убывает быстрее. При $L_i = R$ (R — радиус поля зрения ОЭСС) p_i остается практически на постоянном уровне.

Если расстояние между источниками увеличивается, то, начиная с некоторого момента времени, величина $L_{i \min}$ (минимальное расстояние до раз-

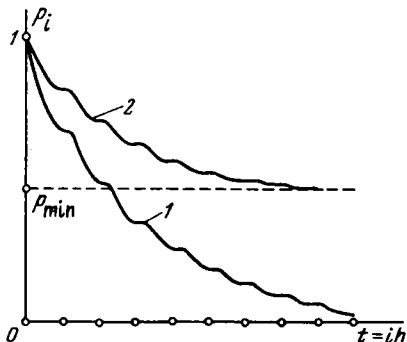


Рис. 2.21. Вероятность устойчивого сопровождения: 1 — при $\Delta_{\Pi} = \text{const}$; 2 — при увеличивающемся расстоянии Δ_{Π}

деляющей границы в поле коррекционных усилий) становится настолько большой, что вероятность слежения за одним источником остается неизменной (см. рис. 2.21).

Полученные результаты частично объясняют приемы постановки организованных помех, описанные в [147], а также приемы отстройки от помех, основанные на спектральной селекции, компенсационном методе и различных вариантах адаптации параметров ОЭСС. Мы не будем останавливаться здесь на технических подробностях указанных аспектов применения помех, отсылая читателя к книге [147] и ряду источников, приведенных в ее библиографии.

Возможное упрощение вычислений. Идентификация обобщенной модели. Заметим, что характер изменения вероятности $p(t)$ позволяет аппроксимировать эту зависимость экспонентой

$$p(t) = \frac{1}{T} e^{-\frac{t}{T}}, \quad (2.6.49)$$

где T — среднее время устойчивого сопровождения источника. Действительно, обозначая среднее время через m_T , находим

$$m_T = \int_0^{\infty} t p(t) dt = \frac{1}{T} \int_0^{\infty} t e^{-\frac{t}{T}} dt = T. \quad (2.6.50)$$

Параметр T полностью определяет характер протекания зависимости $p(t)$. Поэтому появляется возможность существенного упрощения вычислений, связанных с решением уравнения (2.6.29).

Следует обратить внимание на то, что в уравнении (2.6.25) в нашем случае коэффициенты $a(\tau, y)$ и $b(\tau, y)$ не зависят от начального момента времени t , в частности $b(\tau, y) = \sigma_0^2 = \text{const}$. В этом случае плотность распределения $f(t, x; \tau, y)$ также не будет зависеть от t , а будет зависеть от разности $u = \tau - t$, т.е., используя тот же символ f для обозначения плотности, в качестве решения уравнения (2.6.25) получим функцию $f(x; u, y)$. Дифференцируя эту зависимость, находим

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial u} \frac{\partial u}{\partial t} = - \frac{\partial f}{\partial u},$$

$$\frac{\partial f}{\partial \tau} = \frac{\partial f}{\partial u} \frac{\partial u}{\partial \tau} = \frac{\partial f}{\partial u},$$

откуда следует, что $\frac{\partial f}{\partial t} = - \frac{\partial f}{\partial \tau}$. Обращаясь теперь к первому уравнению Колмогорова [5]

$$\begin{aligned} & \frac{\partial f(t, x; \tau, y)}{\partial t} + a(t, x) \frac{\partial f(t, x; \tau, y)}{\partial x} + \\ & + \frac{1}{2} b(t, x) \frac{\partial^2 f(t, x; \tau, y)}{\partial x^2} = 0, \end{aligned}$$

запишем его в виде

$$-\frac{\partial f(x; \tau, y)}{\partial \tau} + a(x) \frac{\partial f(x; \tau, y)}{\partial x} + \frac{1}{2} b(x) \frac{\partial^2 f(x; \tau, y)}{\partial x^2} = 0. \quad (2.6.51)$$

Обозначим через θ время, в течение которого фазовая траектория стохастического уравнения (2.6.24) впервые достигнет поглощающей границы. Тогда

$$P\{\tau < \theta\} = Q(x, \tau) = \int_{L_1}^{L_2} f(x; \tau, y) dy, \quad (2.6.52)$$

и интегральный закон распределения времени θ имеет вид

$$P\{\theta < \tau\} = 1 - Q(x, \tau). \quad (2.6.53)$$

Дифференцируя выражение (2.6.53) по τ , получим плотность распределения случайной величины θ

$$q(x, \tau) = -\frac{\partial Q(x, \tau)}{\partial \tau}. \quad (2.6.54)$$

При желании вычислить среднее время достижения поглощающей границы достаточно вычислить интеграл

$$m_\theta = \int_0^\infty \tau q(x, \tau) d\tau = -\int_0^\infty \tau d\tau \int_{L_1}^{L_2} \frac{\partial f(x; \tau, y)}{\partial \tau} dy. \quad (2.6.55)$$

После изменения порядка интегрирования и имея в виду, что $f(x; \infty, y) = 0$, получаем

$$m_\theta = \int_{L_1}^{L_2} dy \int_0^\infty f(x; \tau, y) d\tau. \quad (2.6.56)$$

Проинтегрируем уравнение (2.6.51) почленно по y и τ в соответствии с результатом (2.6.56). Очевидно, что

$$\int_0^\infty \frac{\partial f(x; \tau, y)}{\partial \tau} d\tau = f(x; \infty, y) - f(x; 0, y) = -\delta(x - y).$$

Поэтому вместо первого слагаемого в уравнении (2.6.51) получаем

$$\int_{L_1}^{L_2} \delta(x - y) dy = 1.$$

Таким образом, в результате интегрирования уравнение (2.6.51) переходит в обыкновенное дифференциальное уравнение относительно среднего времени устойчивого сопровождения

$$\frac{1}{2} b(x) \frac{d^2 m_\theta}{dx^2} + a(x) \frac{dm_\theta}{dx} + 1 = 0. \quad (2.6.57)$$

В случае стохастического уравнения (2.6.24) соотношение (2.6.57) приобретает вид

$$\frac{1}{2} \sigma_0^2 \frac{d^2 m_\theta}{dx^2} - \lambda(x - \Delta_0) \frac{dm_\theta}{dx} + 1 = 0. \quad (2.6.58)$$

уравнение (2.6.58) допускает понижение порядка. Обозначая $dm_\theta/dx = z$, запишем

$$\frac{dz}{dx} - \frac{2\lambda}{\sigma_0^2} (x - \Delta_0) z + \frac{2}{\sigma_0^2} = 0. \quad (2.6.59)$$

Интегрируя линейное дифференциальное уравнение (2.6.59), находим

$$z = \frac{dm_\theta}{dx} = \left[C_1 - \frac{2}{\sigma_0^2} \int_0^x e^{-\frac{\lambda(u - \Delta_0)^2}{\sigma_0^2}} du \right] e^{\frac{\lambda(x - \Delta_0)^2}{\sigma_0^2}},$$

откуда

$$m_\theta = \int_0^x \left[C_1 - \frac{2}{\sigma_0^2} \int_0^v e^{-\frac{\lambda(u - \Delta_0)^2}{\sigma_0^2}} du \right] e^{\frac{\lambda(v - \Delta_0)^2}{\sigma_0^2}} dv + C_2. \quad (2.6.60)$$

Постоянные интегрирования C_1 и C_2 находятся из краевых условий. В данном случае из условия поглощения следует $m_\theta(L) = 0$.

Аналогично, интегрируя по y и τ условие вида (2.6.34), находим

$$\lambda \Delta_0 m_\theta(0) + \frac{\sigma_0^2}{2} \frac{dm_\theta}{dx} \Big|_{x=0} = 0. \quad (2.6.61)$$

Таким образом,

$$m_\theta = \int_L^x \left[C_1 - \frac{2}{\sigma_0^2} \int_0^v e^{-\frac{\lambda(u - \Delta_0)^2}{\sigma_0^2}} du \right] e^{\frac{\lambda(v - \Delta_0)^2}{\sigma_0^2}} dv, \quad (2.6.62)$$

где C_1 находится из соотношения

$$\begin{aligned} C_1 \left(\lambda \Delta_0 \int_0^L e^{\frac{\lambda(v - \Delta_0)^2}{\sigma_0^2}} dv - \frac{\sigma_0^2}{2} e^{-\frac{\lambda \Delta_0^2}{\sigma_0^2}} \right) = \\ = \frac{2\lambda \Delta_0}{\sigma_0^2} \int_0^L e^{\frac{\lambda(v - \Delta_0)^2}{\sigma_0^2}} dv \int_0^v e^{-\frac{\lambda(u - \Delta_0)^2}{\sigma_0^2}} du. \end{aligned} \quad (2.6.63)$$

Вычисление квадратур в равенствах (2.6.62) и (2.6.63) несравненно менее громоздкая операция, чем решение уравнения (2.6.25) методом прогонки.

Характеристика $m_\theta(x)$ может быть с успехом использована для идентификации концептуальной модели (2.6.24) с полной математической моделью ОЭСС либо с реальной ОЭСС. На рис. 2.22 показан отрезок процесса сопровождения двойного источника, полученный на полной математической модели ОЭСС, свободной от каких-либо допущений.

Математическое ожидание времени устойчивого сопровождения каждого из источников находится как статистические оценки вида

$$\hat{m}_{\theta}^{(1)} = \frac{1}{N_1} [t_1 + (t_3 - t_2) + \dots + (t_n - t_{n-1})]$$

и

$$\hat{m}_{\theta}^{(2)} = \frac{1}{N_2} [(t_2 - t_1) + (t_4 - t_3) + \dots]$$

(2.6.64)

для каждого из расстояний Δ_1 и Δ_2 до точки неустойчивого равновесия.

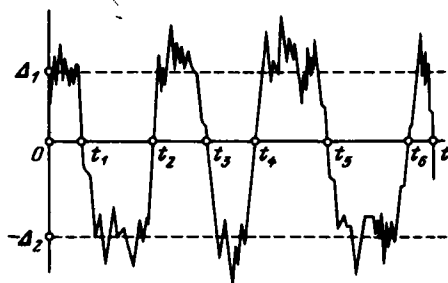


Рис. 2.22. Процесс сопровождения двойного источника при моделировании на поэлементной модели

Эти оценки и используются при достаточно длительном цифровом эксперименте при сопоставлении с результатом (2.6.62), где при заданном параметре λ подбирается уровень шума σ_0^2 .

График вида, изображенного на рис. 2.22, можно получить и при полунатурном эксперименте путем регистрации угловых флуктуаций ОЭСС при фиксированном угловом расстоянии между источниками излучения.

В заключение укажем, что модель сопровождения двойного источника допускает распространение на случай нескольких источников, на случай источника излучения с повышенными угловыми размерами, а также для исследования работы ОЭСС на пороговых уровнях сигнала. Один из примеров такого исследования будет рассмотрен в § 3.7, где в качестве аналога уравнения Фоккера — Планка использовано соответствующее интегро-разностное соотношение.

§ 2.7. Анализ процесса сопровождения источника с распределенным излучением

Одной из важных прикладных проблем является анализ работы ОЭСС в условиях сопровождения источника с большими угловыми размерами, например сравнимыми с угловыми размерами поля зрения оптической системы. Такие условия слежения возникают в системах с приемником излучения, чувствительным к длинноволновой части оптического спектра, поскольку в этом диапазоне излучают тела со сравнительно низкой температурой. Повышенной чувствительностью к низкотемпературному излучению обладают фоторезисторы с глубоким охлаждением, которое достигается дросселированием вытекающей струи газа либо с помощью термоэлектрических холодильников, работающих на основе эффекта Пельтье и Эттингсхаузена [4, 72, 76].

При математическом исследовании процесса сопровождения источника излучения с большими угловыми размерами необходимо конкретизировать модель источника, который представляется в виде двумерной функции $E_{(q)}(x, y)$ освещенности. Здесь (x, y) — координаты точки поля лучистости и символ (q) означает совокупность параметров, характеризующих условия наблюдения. Мы будем предполагать, что функция $E_{(q)}(x, y)$ относится к плоскости изображения, т.е. в ней учтены дифракционные свойства оптики. Светимость самого объекта задается функцией $R_{(q)}(x, y)$, которую можно представить в виде совокупности точек ¹⁾

$$\tilde{R}_{(q)}(x, y) = \sum_i \sum_j R_{(q)}(x_i, y_j) \delta(x - x_i, y - y_j) \Delta x_i \Delta y_j, \quad (2.7.1)$$

где $\delta(x, y)$ — двумерная δ -функция с известными свойствами:

$$\delta(x, y) = \begin{cases} 0 & \text{при } x^2 + y^2 \neq 0, \\ \infty & \text{при } x^2 + y^2 = 0, \end{cases} \quad (2.7.2)$$

$$\iint_{(D)} \delta(x, y) dx dy = 1, \quad 0 \in (D).$$

В данном случае функция $\delta(x - x_i, y - y_j)$ характеризует точечный источник с координатами (x_i, y_j) и единичной интенсивностью. В фокальной плоскости он изображается в виде кружка рассеяния с функцией распределения лучистости [98]

$$v(x, y) = v_0 \frac{J_1^2(\sqrt{x^2 + y^2})}{x^2 + y^2}. \quad (2.7.3)$$

Таким образом, совокупность точек (2.7.1) образует в плоскости фотоприемника изображение

$$E_{(q)}(x, y) = \sum_i \sum_j R_{(q)}(x_i, y_j) v(x - x_i, y - y_j) \Delta x_i \Delta y_j. \quad (2.7.4)$$

Многие реальные протяженные источники излучения можно представить в виде совокупности выпуклых множеств эллиптической формы. В случае одного такого множества в качестве зависимости $R_{(q)}(x, y)$ можно рассмотреть выражение

$$R_{(q)}(x, y) = R_{(q)\max} \exp(-\mu x^2 - \alpha y^2), \quad (2.7.5)$$

где параметры μ и α характеризуют форму линий $R_{(q)}(x, y) = \text{const}$ в координатах (x, y) .

Перейдем в выражении (2.7.4) к пределу при $\Delta x_{i\max} \rightarrow 0, \Delta y_{j\max} \rightarrow 0$. Тогда получим

$$E_{(q)}(x, y) = \iint_{(S)} R_{(q)}(\xi, \eta) v(x - \xi, y - \eta) d\xi d\eta, \quad (2.7.6)$$

где $v(x, y)$ определяется выражением (2.7.3).

¹⁾ Отдельные слагаемые в формуле (2.7.1) представляют светимости элементарных площадок поверхности протяженного источника с размерами $\Delta x_i \times \Delta y_j$.

Для приближенного вычисления двойного интеграла (2.7.6) заменим зависимость (2.7.3) близкой функцией

$$\tilde{v}(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right). \quad (2.7.7)$$

Здесь параметр σ равен $1/3$ радиуса кружка рассеяния, который определяется как первый нуль функции $J_1(\rho)$. Он равен $3,8317\dots$ и, следовательно, в (2.7.7) $\sigma = 1,2772$. Выражение (2.7.7) при подстановке в формулу (2.7.6) позволяет выполнить вычисления до конца, т.е.

$$\begin{aligned} E(q)(x, y) &= \\ &= \frac{R(q)_{\max}}{2\pi\sigma^2} \iint_{(S)} \exp\left[-\mu\xi^2 - \frac{(x-\xi)^2}{2\sigma^2}\right] \exp\left[-\alpha\eta^2 - \frac{(y-\eta)^2}{2\sigma^2}\right] d\xi d\eta = \\ &= \frac{R(q)_{\max}}{\sqrt{(1+2\alpha\sigma^2)(1+2\mu\sigma^2)}} \exp\left(-\frac{\alpha y^2}{1+2\alpha\sigma^2} - \frac{\mu x^2}{1+2\mu\sigma^2}\right). \end{aligned} \quad (2.7.8)$$

Формулы вида (2.7.8) служат для ввода информации в поэлементную модель ОЭСС по методике, описанной в [99]. Применение конечного числа этих формул позволяет описать практически любое детерминированное поле лучистости с достаточной точностью.

Проблема исследования процесса сопровождения протяженного источника излучения распадается на две части. Одна из них — анализ прохождения изображения, описываемого формулами вида (2.7.8), через анализатор изображения — встречает непреодолимые аналитические трудности и может быть решена лишь средствами цифрового моделирования. Основным результатом решения этой части проблемы является получение пеленгационных характеристик, снятых по протяженному источнику. Более общим результатом может служить поле коррекционных усилий, построенное при наличии в поле зрения ОЭСС протяженного источника при данной совокупности (q) параметров наблюдения.

Другая часть проблемы состоит из собственно анализа процесса сопровождения протяженного источника при заданной пеленгационной характеристике.

Эта часть проблемы поддается исследованию чисто аналитическими приемами и в определенной степени опирается на представления, развитые в предыдущем параграфе настоящей главы.

Искажение формы пеленгационной характеристики. На рис. 2.23 показана пеленгационная характеристика, полученная при математическом моделировании анализатора изображения и протяженного источника, линии уровня $E(x, y) = \text{const}$ которого помещены в верхней части рисунка. Из приведенной характеристики видно,

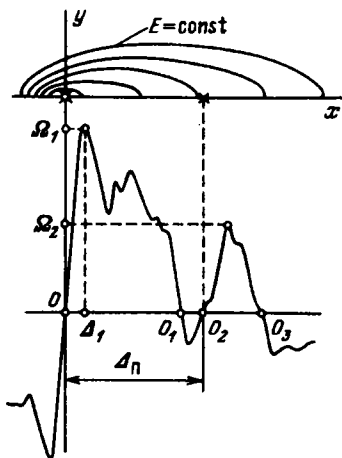


Рис. 2.23. Пеленгационная характеристика при слежении за протяженным источником

что при данных угловых размерах источника помимо точки O устойчивого сопровождения имеются еще одна точка O_2 и две неустойчивые точки O_1 и O_3 . Значения пеленгационной характеристики полезно рассматривать в данном случае как предельные значения угловой скорости вращения исполнительного элемента при сопровождении перемещающегося в пространстве источника.

Действительно, в установившемся движении скорость вращения исполнительного элемента пропорциональна напряжению, поступающему с выхода электронно-усилительного тракта, которое в свою очередь зависит от углового рассогласования по каждому из каналов управления. Если протяженный источник перемещается влево с угловой скоростью вращения линии визирования Ω_1 , то возникает динамическая ошибка сопровождения, равная Δ_1 (см. рис. 2.23). Так как привод следящего координатора не может развить скорость, большую чем Ω_1 , то эта скорость называется *располагаемой*. Моделирование показывает, что величина Ω_1 зависит от угловых размеров источника и обычно уменьшается с ростом угловых размеров сопровождаемого источника. Главным фактором, влияющим на уменьшение располагаемой скорости, служит уменьшение глубины лучистого потока.

После того как ошибка сопровождения достигнет величины Δ_1 , происходит срыв сопровождения, так как остальные пики пеленгационной характеристики не превышают величины Ω_1 . Если по каким-либо причинам начальное значение ошибки таково, что осуществляется устойчивое сопровождение точки O_2 , то для срыва этого сопровождения достаточной является скорость Ω_2 .

При неподвижном в пространстве распределенном источнике процесс сопровождения характеризуется уже рассмотренными в предыдущем параграфе случайными переходами с одной точки устойчивого сопровождения на другую и обратно. Отметим, что при допущении об идеальных условиях модуляции и неперекрывающихся пачках импульсов дополнительных устойчивых точек не возникало. Появление этих точек при отсутствии допущений объясняется изрезанностью пеленгационных характеристик.

Спектральная плотность управляющего сигнала. При слежении за протяженным источником в управляющем сигнале появляется низкочастотная составляющая, спектр которой можно оценить следующим образом. Угловые флуктуации при случайных переходах с точки O на точку O_2 (см. рис. 2.23) при расстоянии Δ_n между ними можно описать как телеграфный сигнал [68] интенсивностью перемен знаков $\lambda(t)$. Корреляционная функция такого сигнала имеет вид

$$R(t_1, t_2) = \frac{\Delta_n^2}{4} \exp \left\{ -2 \int_{t_1}^{t_2} \lambda(t) dt \right\}. \quad (2.7.9)$$

Функция $\lambda(t)$ связана с угловым расстоянием Δ_n между точками O и O_2 . При медленном изменении Δ_n во времени функция $\lambda(t)$ также медленно меняется и вместо соотношения (2.7.9) можно записать приближенное выражение

$$R(\tau) = \frac{\Delta_n^2}{4} e^{-2\lambda\tau}, \quad (2.7.10)$$

где $\tau = t_2 - t_1$.

В соответствии с (2.7.10) энергетический спектр колебаний находится преобразованием Фурье

$$S(\omega) = \frac{\Delta_{\pi}^2 \lambda}{4\lambda^2 + \omega^2}. \quad (2.7.11)$$

В действительности процесс перехода между точками O и O_2 происходит не мгновенно, а сопровождается переходным процессом в следящем координаторе. При этом происходит сглаживание телеграфного сигнала — его фильтрация с некоторой постоянной времени T_{Φ} . Кроме того, на угловые флуктуации $\eta(t)$ с амплитудой $\Delta_{\pi}/2$ наложены автоколебания $\xi(t)$, загрязненные шумом (см. рис. 2.22). Таким образом, угловые флуктуации $\xi(t)$ представляются в виде суммы

$$\xi(t) = \eta(t) + \zeta(t)$$

с энергетическим спектром

$$S_x(\omega) = S_y(\omega) + S_z(\omega) = \frac{\Delta_{\pi}^2 \lambda}{(4\lambda^2 + \omega^2)(1 + T_{\Phi}^2 \omega^2)} + S_z(\omega). \quad (2.7.12)$$

Часть спектра $S_z(\omega)$ является широкополосной, часть $S_y(\omega)$ имеет выраженный экстремум на частоте $\omega = 0$. Если считать первую производную $\xi'(t)$ с точностью до постоянного коэффициента управляющим сигналом, снимаемым с ОЭСС, то его спектр

$$S_{\Omega}(\omega) = \frac{\Delta_{\pi}^2 \lambda \omega^2}{(4\lambda^2 + \omega^2)(1 + T_{\Phi}^2 \omega^2)} + \omega^2 S_z(\omega). \quad (2.7.13)$$

Дисперсия управляющего сигнала получается интегрированием (2.7.13)

$$\sigma_{\Omega}^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{\Omega}(\omega) d\omega = \frac{\Delta_{\pi}^2 \lambda}{2T_{\Phi}(1 + 2\lambda T_{\Phi})} + \sigma_0^2. \quad (2.7.14)$$

В выражении (2.7.14) первое слагаемое возникает только при сопровождении протяженного источника. Если размеры источника достаточно малы, то дополнительных точек устойчивого сопровождения нет и $\lambda = 0$. Второе слагаемое (2.7.14) получается как результат интегрирования

$$\sigma_0^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \omega^2 S_z(\omega) d\omega \quad (2.7.15)$$

и не связано с наличием протяженного источника. Оно характеризует среднеквадратическую ошибку ОЭСС при слежении за неподвижным точечным источником и может служить критерием качества управляющего сигнала.

Таким образом, для получения текущего спектра управляющего сигнала необходимо методом математического моделирования получить последовательность пеленгационных характеристик для изменяющегося во времени протяженного источника. Затем с помощью соотношений вида (2.6.62) определить среднее время устойчивого сопровождения. Согласно определению интенсивности [68]

$$\lambda(t) = \lim_{t \rightarrow 0} \frac{1 - e^{-t/T}}{t} = \frac{1}{T} = \text{const}$$

для заданного параметра Δ_n . Это значение в дальнейшем и используется в соотношении (2.7.13) для спектра управляющего сигнала.

Динамическая ошибка при наличии автоколебаний. Перейдем теперь к определению динамической ошибки, возникающей при сопровождении перемещающегося источника излучения. Для этого вновь обратимся к модели, описываемой уравнениями (2.6.24), в которых положим: $-\lambda(\Delta - \Delta_0) = F(\Delta)$ и $n(t) = V$. Последняя величина эквивалентна детерминированному внешнему воздействию, которое заставляет оптическую ось ОЭСС отклоняться от равновесного положения, соответствующего нулевой ошибке. Таким образом, запишем

$$\frac{d\Delta}{dt} = F(\Delta) + V, \quad \frac{d\varphi}{dt} = \omega_0. \quad (2.7.16)$$

Положим, как обычно, $x = \Delta \cos \varphi$, $y = \Delta \sin \varphi$. Дифференцируя эти соотношения по t и используя (2.7.16), находим

$$\begin{aligned} \frac{dx}{dt} &= [F(\Delta) + V] \cos \varphi - \omega_0 y, \\ \frac{dy}{dt} &= [F(\Delta) + V] \sin \varphi + \omega_0 x. \end{aligned} \quad (2.7.17)$$

Заменив здесь $\cos \varphi = x/\Delta$, $\sin \varphi = y/\Delta$ и обозначив $V_x = V \cos \varphi$, $V_y = V \sin \varphi$, запишем систему (2.7.17) в виде

$$\begin{aligned} \frac{dx}{dt} &= \frac{x}{\Delta} F(\Delta) + V_x - \omega_0 y, \\ \frac{dy}{dt} &= \frac{y}{\Delta} F(\Delta) + V_y + \omega_0 x. \end{aligned} \quad (2.7.18)$$

На рис. 2.24 изображена структурная схема, соответствующая уравнениям (2.7.18). Из нее виден смысл обозначений V_x и V_y : это компоненты

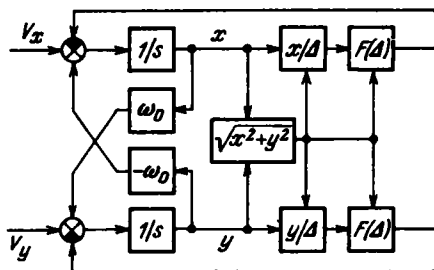


Рис. 2.24. Структурная схема автоколебательной оптико-электронной следящей системы

внешнего воздействия, приводящего к возникновению динамической ошибки в системе с круговым предельным циклом.

При $V_x = V_y = 0$ можно записать решение системы в параметрической форме

$$\Delta = \Delta_0 + [\Delta(0) - \Delta_0] e^{-\lambda t}, \quad (2.7.19)$$

$$\varphi = \varphi(0) + \omega_0 t,$$

где $\Delta(0)$ и $\varphi(0)$ — начальные координаты изображающей точки. Решение (2.7.19) показывает, что система обладает изолированной фазовой траекторией $\Delta = \Delta_0 = \text{const}$, к которой с внешней и внутренней сторон асимптотически приближаются все остальные траектории. Таким образом, система (2.7.18) обладает предельным круговым циклом. Из центра этого цикла как из особой точки начинается бесчисленное множество фазовых траекторий.

Рассмотрим теперь динамическую ошибку автоколебательной системы (2.7.18) при векторном внешнем воздействии (V_x, V_y) . При малых значениях этого воздействия система (2.7.18) по-прежнему будет обладать особым решением в виде предельного цикла. Под динамической ошибкой при этом будем понимать вектор геометрического центра этого предельного цикла. При симметричных каналах управления (см. рис. 2.24) достаточно рассмотреть случай $V_x = V = \text{const}$ и $V_y = 0$. Тогда из (2.7.18) получаем

$$\frac{d\Delta}{d\varphi} = \Delta \frac{F(\Delta) - V \cos \varphi}{V \sin \varphi - \omega_0 \Delta}; \quad (2.7.20)$$

здесь для большей общности анализа введена функция $F(\Delta)$.

Координаты особой точки внутри предельного цикла находят приравниванием нулю числителя и знаменателя правой части (2.7.20), т.е.

$$V \cos \varphi_1 = F(\Delta_1), \quad V \sin \varphi_1 = \omega_0 \Delta_1, \quad (2.7.21)$$

откуда

$$V^2 = F^2(\Delta_1) + \omega_1^2 \Delta_1^2, \quad \varphi_1 = \arctg \frac{\omega_0 \Delta_1}{F(\Delta_1)}. \quad (2.7.22)$$

Однако особая точка с координатами (Δ_1, φ_1) может не совпадать с геометрическим центром предельного цикла при $V \neq 0$. Сам цикл также может потерять первоначальную форму окружности. Поэтому следует рассмотреть совокупность фазовых траекторий уравнения (2.7.21). Следует отметить, что это уравнение не интегрируется в конечном виде при произвольных значениях входного воздействия V . Поэтому целесообразно провести вначале численное интегрирование.

Аналитическое исследование системы (2.7.18) можно осуществить методом малого параметра. Для удобства выкладок введем обозначения

$$\Delta = \Delta_0(1 + \rho), \quad \tau = \omega_0 t.$$

Тогда уравнения (2.7.18) приводятся к виду

$$\frac{d\rho}{d\tau} = -\frac{\lambda}{\omega_0} \rho + \frac{V}{\omega_0 \Delta_0} \cos \varphi, \quad \frac{d\varphi}{d\tau} = 1 - \frac{V}{\omega_0 \Delta_0} \cdot \frac{\sin \varphi}{1 + \rho}. \quad (2.7.23)$$

При малых значениях внешнего воздействия V введем малый параметр $\nu = V/\omega_0 \Delta_0$. Кроме того, обозначим $\mu = \lambda/\omega_0$. В этих обозначениях вместо (2.7.23) запишем

$$\frac{d\rho}{d\tau} = -\mu\rho + \nu \cos \varphi, \quad \frac{d\varphi}{d\tau} = 1 - \nu \frac{\sin \varphi}{1 + \rho}. \quad (2.7.24)$$

При $\nu = 0$ решение известно:

$$\rho = \rho(0) e^{-\mu\tau}, \quad \varphi = \varphi(0) + \tau. \quad (2.7.25)$$

При $\nu \neq 0$ будем искать решение системы (2.7.24) в форме

$$\rho = \sum_{k=0}^{\infty} \rho_k(\rho, \varphi) \nu^k, \quad \varphi = \sum_{k=0}^{\infty} \varphi_k(\rho, \varphi) \nu^k. \quad (2.7.26)$$

Здесь $\rho_k(\rho, \varphi)$ и $\varphi_k(\rho, \varphi)$ — пока неизвестные функции. Они находятся после подстановки (2.7.26) в исходные уравнения (2.7.24). При этом приходится так преобразовать правые части уравнений, чтобы они представляли полиномы (ряды) по степеням параметра ν .

В частности, используя разложение Тейлора, имеем

$$\begin{aligned} \cos \varphi &= \cos \left(\sum_{k=0}^{\infty} \varphi_k \nu^k \right) = \cos(\varphi_0 + \nu \sum_{k=1}^{\infty} \varphi_k \nu^{k-1}) = \\ &= \cos \varphi_0 - \nu \sin \varphi_0 \sum_{k=1}^{\infty} \varphi_k \nu^{k-1} - \frac{\nu^2}{2} \cos \varphi_0 \left(\sum_{k=1}^{\infty} \varphi_k \nu^{k-1} \right)^2 + \dots = \\ &= \sum_{j=0}^{\infty} \cos \left(\varphi_0 + j \frac{\pi}{2} \right) \frac{\nu^j}{j!} \left(\sum_{k=1}^{\infty} \varphi_k \nu^{k-1} \right)^j. \end{aligned} \quad (2.7.27)$$

Аналогично для функции двух переменных

$$\begin{aligned} \frac{\sin \varphi}{1 + \rho} &= \frac{\sin(\varphi_0 + \nu \sum_{k=1}^{\infty} \varphi_k \nu^{k-1})}{1 + \rho_0 + \nu \sum_{k=1}^{\infty} \rho_k \nu^{k-1}} = \frac{\sin \varphi_0}{1 + \rho_0} + \\ &+ \frac{\nu \cos \varphi_0}{1 + \rho_0} \sum_{k=1}^{\infty} \varphi_k \nu^{k-1} - \frac{\nu \sin \varphi_0}{(1 + \rho_0)^2} \sum_{k=1}^{\infty} \rho_k \nu^{k-1} + \\ &+ \frac{\nu^2}{2} \left[\frac{2 \sin \varphi_0}{(1 + \rho_0)^3} \left(\sum_{k=1}^{\infty} \rho_k \nu^{k-1} \right)^2 - \frac{2 \cos \varphi_0}{(1 + \rho_0)^2} \sum_{k=1}^{\infty} \rho_k \nu^{k-1} \times \right. \\ &\left. \times \sum_{k=1}^{\infty} \varphi_k \nu^{k-1} - \frac{\sin \varphi_0}{1 + \rho_0} \left(\sum_{k=1}^{\infty} \varphi_k \nu^{k-1} \right)^2 \right] + \dots \end{aligned} \quad (2.7.28)$$

Подстановка выражений (2.7.26), (2.7.27) и (2.7.28) в уравнения (2.7.24) позволяет свести их решение к вычислению рекуррентных соотношений. При этом затруднения вызывает лишь вопрос о том, на каком шаге вычислительной процедуры следует остановиться. Выход из этого затруднения заключается в параллельном численном интегрировании системы (2.7.24).

Подставив в (2.7.24) выражение (2.7.27), находим

$$\frac{d\rho_0}{d\tau} + \nu \frac{d\rho_1}{d\tau} + \nu^2 \frac{d\rho_2}{d\tau} + \dots = -\mu\rho_0 - \mu\nu\rho_1 - \mu\nu^2\rho_2 - \dots$$

$$+ \nu \cos \varphi_0 - \nu^2 \sin \varphi_0 (\varphi_1 + \nu \varphi_2 + \dots) + \dots,$$

откуда следует первая группа рекуррентных соотношений

$$\begin{aligned}\frac{d\rho_0}{d\tau} &= -\mu\rho_0, & \frac{d\rho_1}{d\tau} &= -\mu\rho_1 + \cos\varphi_0, \\ \frac{d\rho_2}{d\tau} &= -\mu\rho_2 - \varphi_1 \sin\varphi_0, \\ &\dots\dots\dots\end{aligned}\tag{2.7.29}$$

Аналогично, подставляя (2.7.28) во второе уравнение (2.7.27), получим вторую группу рекуррентных соотношений:

$$\begin{aligned}\frac{d\varphi_0}{d\tau} &= 1, \\ \frac{d\varphi_1}{d\tau} &= \frac{\sin\varphi_0}{1+\rho_0}, \\ \frac{d\varphi_2}{d\tau} &= \varphi_1 \frac{\cos\varphi_0}{1+\rho_0} - \rho_1 \frac{\sin\varphi_0}{(1+\rho_0)}, \\ &\dots\dots\dots\end{aligned}\tag{2.7.30}$$

Из первых двух уравнений (2.7.29) и (2.7.30) находим уже известные решения (2.7.25). Подставляя их во второе уравнение (2.7.29), получаем

$$\frac{d\rho_1}{d\tau} = -\mu\rho_1 + \cos[\varphi(0) + \tau].\tag{2.7.31}$$

Решение линейного уравнения (2.7.31) не представляет затруднений:

$$\rho_1(\tau) = \rho_1(0)e^{-\mu\tau} + \frac{1}{\sqrt{\mu^2 + 1}} \cos\left[\varphi(0) + \tau - \operatorname{arctg} \frac{1}{\mu}\right].\tag{2.7.32}$$

Обратимся теперь ко второму уравнению (2.7.30), записанному в виде

$$\frac{d\varphi_1}{d\tau} = \frac{\sin[\varphi(0) + \tau]}{1 + \rho(0)e^{-\mu\tau}},\tag{2.7.33}$$

интегрирование которого сводится к квадратуре

$$\varphi_1(\tau) = \int \frac{\sin[\varphi(0) + \tau]}{1 + \rho(0)e^{-\mu\tau}} d\tau.\tag{2.7.34}$$

Вычисление ее в конечном виде невозможно. Однако если поставить задачу вычисления установившегося решения, то экспоненту в знаменателе подынтегрального выражения можно не рассматривать и тогда

$$\varphi_1(\tau) = -\cos[\varphi(0) + \tau] + \varphi_1(0).\tag{2.7.35}$$

Ограничимся вычислением $\rho_2(\tau)$ и $\varphi_2(\tau)$. Из третьего уравнения (2.7.29) и уравнений (2.7.25), (2.7.34) находим

$$\rho_2(\tau) = \rho_2(0)e^{-\mu\tau} + \varphi_1(0) \cos[\varphi(0) + \tau] + \frac{1}{2} \sin^2[\varphi(0) + \tau].\tag{2.7.36}$$

Аналогично из третьего уравнения (2.7.30) и уравнений (2.7.34), (2.7.32) запишем периодическую часть функции $\varphi_2(\tau)$:

$$\begin{aligned} \varphi_2(\tau) = & \varphi_1(0) \sin [\varphi(0) + \tau] - \frac{1}{\sqrt{\mu^2 + 1}} \sin \left[\varphi(0) - \operatorname{arctg} \frac{1}{\mu} + \tau \right] - \\ & - [\varphi(0) + \tau] - \frac{1}{4} \sin 2[\varphi(0) + \tau]. \end{aligned} \quad (2.7.37)$$

Процесс последовательного вычисления функций $\rho_k(\tau)$ и $\varphi_k(\tau)$, как видим, не представляет принципиальных затруднений. Подставляя полученные функции в (2.7.26) и ограничиваясь конечным числом слагаемых, можно получить приближенное решение системы (2.7.23) и, следовательно, исходной системы (2.7.18).

На рис. 2.25 показаны решения системы (2.7.18) при возрастающих значениях параметра ν . Из графиков предельных циклов видно, что при малых значениях ν циклы мало отличаются от окружностей. Это позволяет оценить модуль ошибки как смещение центра окружности. Если радиус окружности равен Δ_{cp} , то

$$\Delta_{\max} = \Delta_0 + \epsilon, \quad \Delta_{\min} = \Delta_0 - \epsilon,$$

откуда $\epsilon = \frac{1}{2} (\Delta_{\max} - \Delta_{\min})$. Обращаясь к соотношениям (2.7.26) и удерживая в них слагаемые до порядка ν включительно, найдем

$$\epsilon = \frac{\nu}{\sqrt{\mu^2 + 1}}. \quad (2.7.38)$$

Подставляя сюда значения параметров ν и μ , принятые при выводе уравнений (2.7.24), запишем окончательно для малых значений параметра ν

$$\epsilon^* = \frac{V}{\Delta_0 \sqrt{\lambda^2 + \omega_0^2}}. \quad (2.7.39)$$

Динамическая ошибка при подавлении автоколебаний. Рассмотрим поведение решения системы (2.7.18) при больших значениях возмущающего воздействия. Для этого снова обратимся к уравнению фазовых траекторий (2.7.20). Из этого уравнения можно найти установившиеся значения координат Δ и φ , не зависящие от времени t . Нас в первую очередь будут интересовать значения $\Delta = \text{const}$, отличные от нуля. Поэтому достаточно рассмотреть систему равенства

$$F(\Delta_1) - V \cos \varphi_1 = 0, \quad \omega_0 \Delta_1 - V \sin \varphi_1 = 0. \quad (2.7.40)$$

Пользуясь в дальнейшем для определенности выражением $F(\Delta) = -\lambda(\Delta - \Delta_0)$, находим из (2.7.40)

$$V^2 = (\lambda^2 + \omega_0^2) \Delta_1^2 + \lambda^2 \Delta_0^2 - 2\lambda \Delta_0 \Delta_1,$$

откуда получаем два возможных значения координаты Δ_1 особой точки

$$\Delta_1^{(1,2)} = \frac{\lambda \Delta_0}{\lambda^2 + \omega_0^2} \pm \sqrt{\frac{\lambda^2 \Delta_0^2}{(\lambda^2 + \omega_0^2)^2} - \frac{\lambda^2 \Delta_0^2 - V^2}{\lambda^2 + \omega_0^2}}. \quad (2.7.41)$$

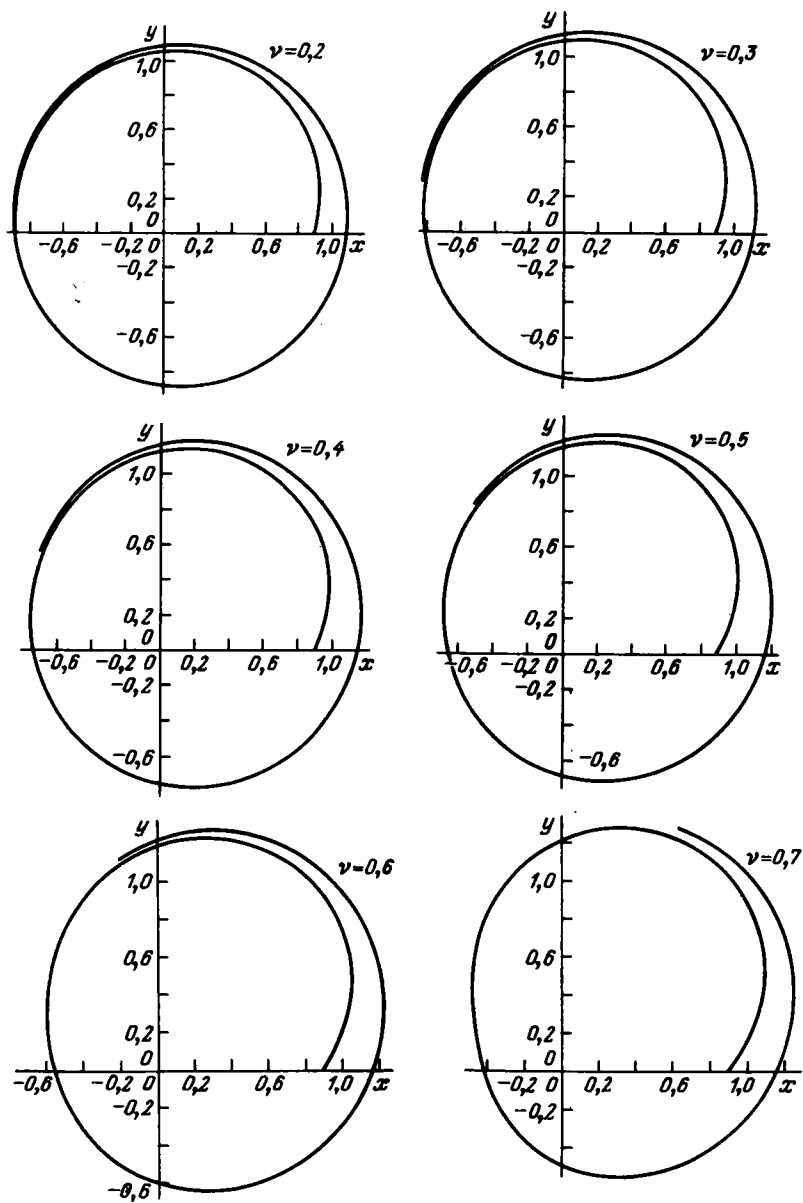


Рис. 2.25. Предельные циклы автоколебательной оптико-электронной следящей системы при возрастающих значениях параметра ν

Вещественные значения Δ_1 получаются лишь при условии

$$\frac{\lambda^2 \Delta_0^2}{\lambda^2 + \omega_0^2} \geq \lambda^2 \Delta_0^2 - V^2. \quad (2.7.42)$$

При выполнении (2.7.42) из (2.7.40) находится вторая координата каждой из двух особых точек

$$\varphi_1^{(1,2)} = \pi + \operatorname{arctg} \frac{\omega_0 \Delta_1^{(1,2)}}{\lambda(\Delta_0 - \Delta_1^{(1,2)})}. \quad (2.7.43)$$

Для того чтобы решения $\Delta_1^{(1,2)}$ и $\varphi_1^{(1,2)}$ были физически наблюдаемы, необходимо, чтобы они соответствовали устойчивым особым точкам. Для анализа устойчивости их введем безразмерную координату $\bar{\Delta} = \Delta/\Delta_0$. Тогда, используя обозначения μ и ν параметров уравнений (2.7.24), запишем вместо (2.7.20)

$$\frac{d\bar{\Delta}}{d\varphi} = \bar{\Delta} \frac{\mu(1 - \bar{\Delta}) + \nu \cos \varphi}{\bar{\Delta} - \nu \cos \varphi}. \quad (2.7.44)$$

Координаты особой точки находятся приравниванием нулю числителя и знаменателя (2.7.44):

$$\mu(1 - \bar{\Delta}^*) + \nu \cos \varphi^* = 0, \quad \bar{\Delta}^* - \nu \sin \varphi^* = 0. \quad (2.7.45)$$

В малой окрестности особой точки $(\bar{\Delta}^*, \varphi^*)$ введем переменные

$$r = \bar{\Delta} - \bar{\Delta}^*, \quad \psi = \varphi - \varphi^*. \quad (2.7.46)$$

Тогда вместо (2.7.44) получаем

$$\frac{dr}{d\psi} = (r + \bar{\Delta}^*) \frac{-\mu r - \nu \psi \sin \varphi^*}{r - \nu \psi \cos \varphi^*}. \quad (2.7.47)$$

Соответствующее характеристическое уравнение имеет вид

$$*\lambda^2 + \lambda(\mu \bar{\Delta}^* + \nu \cos \varphi^*) + \nu \bar{\Delta}^*(\mu \cos \varphi^* + \sin \varphi^*) = 0 \quad (2.7.48)$$

или, учитывая (2.7.45), более компактную форму

$$\lambda^2 + \lambda\mu(2\bar{\Delta}^* - 1) + \mu^2 \bar{\Delta}^*(\bar{\Delta}^* - 1) + (\bar{\Delta}^*)^2 = 0. \quad (2.7.49)$$

Отсюда получаются корни

$$\lambda_{1,2} = -\mu \left(\bar{\Delta}^* - \frac{1}{2} \right) \pm \sqrt{\frac{\mu^2}{4} - (\bar{\Delta}^*)^2}, \quad (2.7.50)$$

определяющие поведение фазовых траекторий в малой окрестности особых точек. Могут представиться следующие случаи.

1. $(\bar{\Delta}^*)^2 < \mu^2/4$ — оба корня вещественны. При этом если

$$\mu^2/4 - (\bar{\Delta}^*)^2 < \mu^2 \left(\bar{\Delta}^* - \frac{1}{2} \right)^2,$$

то корни одного знака и мы имеем особую точку типа узла. Узел устойчив при $\bar{\Delta}^* > \frac{1}{2}$. При $\mu^2/4 - (\bar{\Delta}^*)^2 > \mu^2 \left(\bar{\Delta}^* - \frac{1}{2} \right)^2$ корни имеют разные знаки и возникает особая точка типа седла, которое всегда неустойчиво.

2. $(\bar{\Delta}^*)^2 > \mu^2/4$ – особая точка типа фокуса, устойчивого при $\bar{\Delta}^* > 1/2$ и неустойчивого при $\bar{\Delta}^* < 1/2$.

Рассмотрим некоторые конкретные примеры. Пусть $\mu = 1$. Тогда из (2.7.45) находим

$$\bar{\Delta}^* = \frac{1}{2} \pm \sqrt{\frac{\nu^2}{2} - \frac{1}{4}}.$$

При $\nu < 1/\sqrt{2}$ вещественных значений $\bar{\Delta}^*$ не наблюдается и особых точек внутри предельного цикла не содержится. Впервые появляется единственная особая точка при $\nu = 1/\sqrt{2}$. Ее координата $\bar{\Delta}^* = 1/2$. Согласно (2.7.45)

$$\sin \varphi^* = \frac{\bar{\Delta}^*}{\nu} = \frac{1}{\sqrt{2}},$$

$$\cos \varphi^* = \frac{\bar{\Delta}^* - 1}{\nu} = -\frac{1}{\sqrt{2}},$$

т.е. $\varphi^* = 3\pi/4$.

Согласно (2.7.50) при $\mu = 1$ и $\bar{\Delta}^* = 1/2$ имеет место $\lambda_1 = \lambda_2 = 0$, т.е. особая точка соответствует положению безразличного равновесия. Это состояние наблюдается в некоторой окрестности точки $\bar{\Delta}^* = 1/2, \varphi^* = 3\pi/4$. При подстановке этих значений

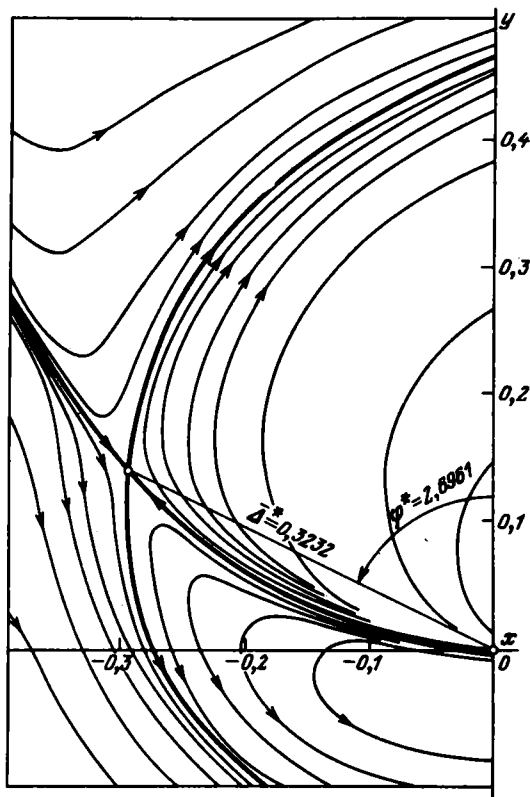


Рис. 2.26. Фазовые траектории в окрестности неустойчивой особой точки

в (2.7.47) получаем уравнение фазовой траектории в виде

$$r = -\frac{\psi}{2}.$$

С увеличением ν особая точка "раздваивается", так как получаются два значения $\bar{\Delta}^*$. Например, при $\nu = 0,75$ получаем $\bar{\Delta}_1^*(1) = 0,6768$ и $\bar{\Delta}_1^*(2) = 0,3232$. Соответствующие значения $\varphi_1^*(1) = 2,0163$ и $\varphi_1^*(2) = 2,6961$. Найдем значения корней характеристического уравнения. Согласно формуле (2.7.50) $\lambda_1^{(1)} = -0,1768 + j0,4561$, $\lambda_2^{(1)} = -0,1768 - j0,4561$, т.е. первая точка – устойчивый фокус. Для второй точки $\lambda_1^{(2)} = 0,5583$, $\lambda_2^{(2)} = -0,2047$, т.е. вторая точка оказалась седлом, которое характеризуется двумя особыми решениями вида $r = k\psi$, где k – постоянный коэффициент. Значения этого коэффициента определяются подстановкой особых решений в (2.7.47):

$$k = \bar{\Delta}^* = \frac{-k - \nu \sin \psi^*}{k - \nu \cos \varphi^*} = \bar{\Delta}^* \frac{-k - \bar{\Delta}^*}{k + 1 - \bar{\Delta}^*}.$$

Отсюда $k_1 = 0,8815$, $k_2 = 0,1185$.

Линейные функции $r = k\psi$ не сохраняют свою линейную геометрическую форму в координатах (x, y) . Для их построения можно воспользоваться формулами (2.7.46) и соотношениями вида $x = \bar{\Delta} \cos \varphi$, $y = \bar{\Delta} \sin \varphi$. Таким образом, вместо прямых $r = k\psi$ в декартовых координатах (x, y) в качестве особых решений уравнения (2.7.47) появляются параметрически заданные кривые (рис. 2.26)

$$x = (\bar{\Delta}^* + k_i \psi) \cos(\varphi^* + \psi),$$

$$y = (\bar{\Delta}^* + k_i \psi) \sin(\varphi^* + \psi) \quad (i = 1, 2).$$

Решение $\bar{\Delta}^* = 0,3232$, $\varphi^* = 2,6961$ хотя и является неустойчивым, но оказывает влияние на форму предельного цикла в окрестности этой точки благодаря характерному течению интегральных кривых (рис. 2.27).

Помимо выявленных точек типа устойчивого фокуса и седла существует еще одна особая точка $x = 0$, $y = 0$, так как в ней нарушаются условия единственности решения дифференциального уравнения фазовых траекторий (2.7.44). Эта особая точка не принадлежит к линеаризованному виду 0:0 и для ее качественного анализа необходимо проинтегрировать уравнение (2.7.44), приобретающее при $\bar{\Delta} \rightarrow 0$ форму

$$\frac{d\bar{\Delta}}{d\varphi} = \bar{\Delta} \frac{\mu + \Delta \cos \varphi}{-\nu \sin \varphi}. \quad (2.7.51)$$

Интегрируя, получаем

$$\bar{\Delta} = C \left(\operatorname{tg} \frac{\varphi}{2} \right)^{-\frac{\mu}{\nu}} (\sin \varphi)^{-1} = \frac{C \left(\cos \frac{\varphi}{2} \right)^{\frac{\mu}{\nu}-1}}{2 \left(\sin \frac{\varphi}{2} \right)^{1-\frac{\mu}{\nu}}}. \quad (2.7.52)$$

Для построения фазового портрета в координатах (x, y) из (2.7.52) находим параметрически заданное семейство фазовых траекторий при $\nu = 0,75$, $\mu = 1$:

$$\begin{aligned} x &= \frac{C}{2} \left(\cos \frac{\varphi}{2} \right)^{\frac{\mu}{\nu}-1} \left(\sin \frac{\varphi}{2} \right)^{-\frac{\mu}{\nu}-1} \cos \varphi, \\ y &= \frac{C}{2} \left(\cos \frac{\varphi}{2} \right)^{\frac{\mu}{\nu}-1} \left(\sin \frac{\varphi}{2} \right)^{-\frac{\mu}{\nu}-1} \sin \varphi. \end{aligned} \quad (2.7.53)$$

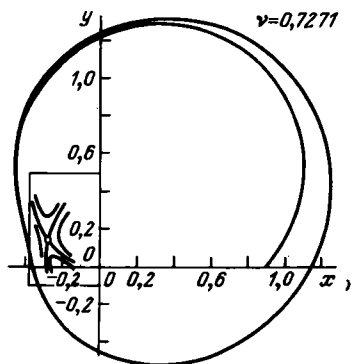


Рис. 2.27.

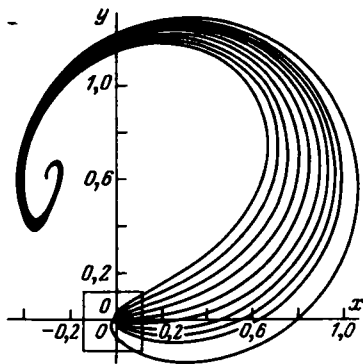


Рис. 2.28.

Рис. 2.27. Влияние неустойчивой особой точки на форму предельного цикла

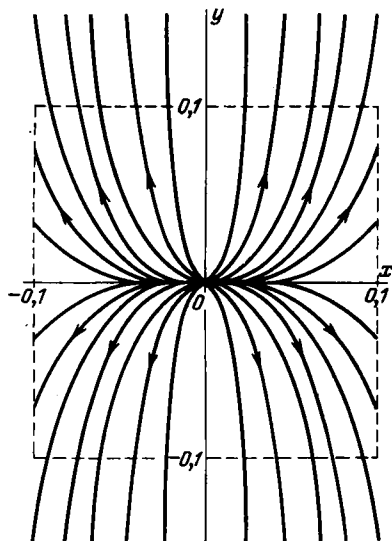


Рис. 2.29

Рис. 2.28. Поведение фазовых траекторий, выходящих из неустойчивого узла (0,0)

Рис. 2.29. Фрагмент рис. 2.28 в окрестности начала координат

Из (2.7.53) видно, что $x \rightarrow 0$ и $y \rightarrow 0$ при $\varphi \rightarrow \pi$. Качественный анализ поведения фазовых траекторий получается при замене $\varphi = \pi - \epsilon$, где $\epsilon \rightarrow 0$. Таким образом, вместо (2.7.53) запишем

$$x \sim C \left(\frac{\epsilon}{2} \right)^{\frac{\mu}{\nu} - 1}, \quad y \sim C \left(\frac{\epsilon}{2} \right)^{\frac{\mu}{\nu} - 1} \epsilon,$$

откуда

$$|y| = C_1 |x|^{\frac{\mu}{\mu - \nu}} \quad (2.7.54)$$

При $\nu < \mu$ (2.7.54) описывает поведение фазовых траекторий в окрестности узла. При $\nu = 0$ узел становится дикритическим. На рис. 2.28 видны траектории, выходящие из точки $x = 0$, $y = 0$ и стремящиеся к устойчивому фокусу. Явная форма траекторий в окрестности седла находится интегрированием уравнения (2.7.47). Полагая $r = u\psi$,

находим при $\mu = 1$

$$u + \psi \frac{du}{d\psi} = - \frac{\bar{\Delta}^* u + (\bar{\Delta}^*)^2}{u + 1 - \bar{\Delta}^*}, \quad (2.7.55)$$

откуда

$$\frac{d\psi}{\psi} = - \frac{1}{2} \frac{(2u+1) du}{u^2 + u + (\bar{\Delta}^*)^2} = \frac{1}{2} \frac{1 - 2\bar{\Delta}^*}{\lambda_1 - \lambda_2} \left(\frac{du}{u - \lambda_1} - \frac{du}{u - \lambda_2} \right)$$

и, следовательно,

$$\psi = \frac{C}{\sqrt{|u^2 + u + \bar{\Delta}^*|}} \left| \frac{u - \lambda_2}{u - \lambda_1} \right|^{\frac{1}{2} \frac{1 - 2\bar{\Delta}^*}{\lambda_1 - \lambda_2}}, \quad (2.7.56)$$

где $\lambda_1 = -1/2 + \sqrt{1/4 - \bar{\Delta}^*}$ и $\lambda_2 = -1/2 - \sqrt{1/4 - \bar{\Delta}^*}$.

Из (2.7.56) следуют параметрические уравнения фазовых траекторий вида $x = \Delta \cos \varphi$, $y = \Delta \sin \varphi$, где

$$\Delta = \bar{\Delta}^* + \frac{C}{\sqrt{|u^2 + u + \bar{\Delta}^*|}} \left| \frac{u - \lambda_1}{u - \lambda_2} \right|^{\frac{1}{2} \frac{1 - 2\bar{\Delta}^*}{\lambda_1 - \lambda_2}},$$

$$\varphi = \varphi^* + \frac{Cu}{\sqrt{|u^2 + u + \bar{\Delta}^*|}} \left| \frac{u - \lambda_1}{u - \lambda_2} \right|^{\frac{1}{2} \frac{1 - 2\bar{\Delta}^*}{\lambda_1 - \lambda_2}}. \quad (2.7.57)$$

На рис. 2.29 показан фрагмент фазового портрета, построенного по соотношениям (2.7.57). На рис. 2.28 этот же фрагмент выделен прямоугольником.

Если при моделировании на поэлементной модели срыв слежения наблюдается при значении $\nu = \hat{\nu}$, то нелинейную функцию $F(\Delta)$ в (2.7.16) следует ограничить, пользуясь приближенной формулой (2.7.38). Можно, например, положить

$$F(\Delta) = \begin{cases} \lambda(\Delta - \Delta_0) & \text{при } 0 \leq \Delta < \frac{\hat{\nu}}{\sqrt{1 + \mu^2}}, \\ 0 & \text{при } \Delta \geq \frac{\hat{\nu}}{\sqrt{1 + \mu^2}}. \end{cases}$$

При такой нелинейной характеристике срыв слежения будет наступать раньше срыва автоколебаний.

Модель слежения за распределенным источником может быть дополнена введением шумового воздействия $n(t)$. В отличие от задачи § 2.6, в данном случае уравнение Фоккера–Планка должно быть записано для стохастического уравнения второго порядка, которое вытекает из системы (2.7.16). Здесь мы не рассматриваем этого случая. Методика получения расчетных формул для вероятности срыва моделей второго порядка приводится, например, в [88].

МЕТОДЫ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ ОПТИКО-ЭЛЕКТРОННЫХ СЛЕДЯЩИХ СИСТЕМ

§ 3.1. Особенности математического моделирования ОЭСС

Виды моделирования. Существующие методы научных исследований можно разделить на три большие группы: экспериментальные, теоретические и моделирование.

В настоящее время при решении многих задач науки и техники применяется математическое моделирование на универсальных вычислительных машинах. Это обусловлено постоянно возрастающей сложностью решаемых задач, когда точное аналитическое решение практически невозможно ввиду значительных математических трудностей, а проведение экспериментальных исследований требует больших временных и экономических затрат или вообще бывает невозможным. В этих условиях математическое моделирование имеет ряд преимуществ перед другими методами исследований и позволяет в значительной мере решить одну из основных проблем современной науки — проблему сложности.

Применение математического моделирования для исследования систем автоматического управления проводится уже давно, сначала на аналоговых вычислительных машинах, затем на цифровых, которые в дальнейшем будем называть ЭВМ. Но только появление быстродействующих универсальных ЭВМ создало условия для широкого применения математического моделирования систем управления.

Начало использования ЭВМ для моделирования оптико-электронных систем в нашей стране относится к середине 60-х годов. Сейчас этот метод исследования применяется при разработке практически всех новых оптико-электронных систем управления. Это нашло свое отражение в том, что в литературе появилось много статей и даже монографий по отдельным вопросам математического моделирования на ЭВМ. Однако в этих работах рассматриваются в основном методы моделирования элементов систем автоматического управления и значительно меньше внимания уделено моделированию входных сигналов и построению модели системы в целом. Имеется несколько работ по моделированию радиосистем, и практически совсем нет работ по математическому моделированию оптико-электронных систем управления. В данной работе делается попытка несколько восполнить этот пробел.

Ввиду большого разнообразия оптико-электронных систем целесообразно выделить общие элементы различных систем, а затем дать их математическое описание и алгоритмы моделирования. На этом принципе и основано

изложение материала в этой части книги. Математические модели рассматриваются на уровне вычислительных алгоритмов, не касаясь конкретных языков программирования, так как по этому вопросу существует обширная литература.

Рассмотрим основные виды моделирования (рис. 3.1):

При *физическом* моделировании создается уменьшенная модель объекта или действующий макет системы управления. При *полунатурном* моделировании обычно исследуется готовый образец изделия с использованием

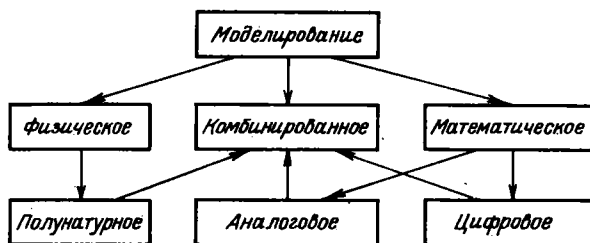


Рис. 3.1. Виды моделирования

различных имитаторов входных сигналов. При *математическом* моделировании поведение объекта описывается системой уравнений, которые решаются на аналоговых или цифровых вычислительных машинах. Часто применяется комбинированное моделирование; полунатурное с математическим.

Физическое моделирование применяется в основном на первом этапе разработки для выявления закономерностей протекания исследуемых процессов. Полунатурное моделирование используется чаще всего на последнем этапе создания системы как средство проверки функционирования изделия в условиях, близких к естественным. Наиболее широкое применение находит математическое моделирование, которое используется на всех этапах создания изделия, начиная от проектирования и кончая испытаниями. Использование математического моделирования позволяет значительно сократить сроки разработки, уменьшить стоимость, получить оценку работы системы в широких условиях применения.

Наиболее универсальным является математическое моделирование на цифровых вычислительных машинах. Основные преимущества цифрового моделирования:

- численные эксперименты на ЭВМ точно воспроизводимы, поэтому можно повторять эксперимент в одних и тех же условиях;
- точно известны все условия эксперимента и параметры модели, что невозможно при других видах моделирования;
- возможно проведение "чистого" эксперимента, т.е. исследования влияния отдельных факторов, не зашумленного посторонними воздействиями;
- цифровое моделирование позволяет воспроизводить широкий диапазон внешних условий, трудно реализуемый в физическом эксперименте и при полунатурном моделировании;
- эксперименты на ЭВМ экономически выгоднее физических экспериментов;

– на первой стадии разработки моделирование на ЭВМ является единственным средством получения достоверной информации о работе проектируемой системы.

Недостатком моделирования на ЭВМ по сравнению с аналоговым и полунатурным моделированием является невозможность моделирования сложных систем в реальном масштабе времени, однако этот недостаток постепенно смягчается по мере повышения быстродействия ЭВМ.

В дальнейшем рассматривается только математическое моделирование на цифровых вычислительных машинах.

Типы моделей. Основным этапом математического моделирования является создание математической модели исследуемой системы, т.е. математическое описание процесса функционирования с учетом внешних воздействий.

В зависимости от априорной информации о механизме функционирования исследуемой системы математические модели могут быть различной степени сложности. Один крайний случай составляют модели типа "черного ящика", когда ничего не известно о внутреннем устройстве объекта и модель создается по результатам наблюдений над входными и выходными сигналами. Другой крайний случай составляют модели систем, принцип функционирования и структура которых известны полностью. В этом случае структура модели повторяет структуру объекта, т.е. создается так называемая *поэлементная* модель или модель, топологически идентичная с изделием. Схематически эти два случая изображены на рис. 3.2.

Модели первого типа отличаются простотой и могут быть использованы для моделирования в реальном масштабе времени. Модели второго типа дают наиболее ценные результаты, так как позволяют исследовать поведение объекта в различных условиях и проводить оптимизацию параметров и структуры по определенным критериям. Основное внимание в книге уделено моделям этого типа, т.е. поэлементным моделям, и только в § 3.6 рассматривается построение моделей типа "черного ящика" по экспериментальным данным.

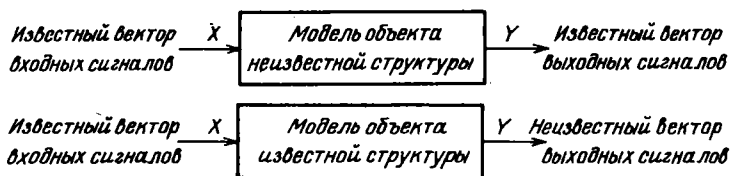


Рис. 3.2. Типы моделей

При использовании любых математических моделей будет рассматриваться так называемое *имитационное* моделирование, т.е. модель должна давать выходной сигнал, достаточно близкий к выходному сигналу изделия. Степень близости модели к изделию может быть оценена по различным критериям. Для моделей типа "черного ящика" наиболее часто используется среднеквадратический критерий, т.е. добиваются минимума функции

$$Q = \frac{1}{T} \int_0^T (Y_m - Y)^2 dt, \quad (3.1.1)$$

где Y_m — выход модели; Y — выход объекта при одинаковом входном воздействии.

Для поэлементных моделей критерием близости может быть сравнение различных характеристик — дискриминационных, пеленгационных, частотных, переходных процессов и др.

Создание поэлементных моделей основывается на принципе *функционального моделирования*, согласно которому модель воспроизводит функцию оригинала, например алгоритм преобразования входных сигналов в выходные, при этом модель и объект не подобны в целом, так как при моделировании опускаются несущественные с информационной точки зрения подробности выполнения отдельных функциональных блоков (ламповые, транзисторные, интегральные схемы и т.п.). Такая модель состоит из блоков, описываемых некоторыми обобщенными характеристиками. Более подробное моделирование, когда рассматриваются схемы построения отдельных блоков, называется *моделированием на уровне принципиальных схем*. В дальнейшем будет рассматриваться только функциональное моделирование. Такие модели являются эффективными для оценки алгоритмов обработки сигнала, оценки помехозащищенности, оценки точности слежения, оптимизации по определенным критериям и т.д.

В задачу математического моделирования входит не только математическое описание, т.е. создание математической модели объекта, но и разработка вычислительных алгоритмов, т.е. методов решения поставленной математической задачи с учетом устойчивости, точности и быстроты вычислений. Алгоритм моделирования является готовым материалом для написания программы на конкретном языке программирования.

Таким образом, в дальнейшем будет рассматриваться имитационное функциональное математическое моделирование на ЭВМ. Основные задачи при этом: построение поэлементных математических моделей и вычислительных алгоритмов.

* Какие требования предъявляются к ЭВМ, используемым для математического моделирования? Для этой цели наиболее подходят универсальные цифровые вычислительные машины типа БЭСМ-6 и серии ЕС. Эти машины отличаются большим быстродействием и большой оперативной памятью. Объем программы математической модели оптико-электронной системы с моделью входных сигналов может достигать 30 кбайт. Существенно, чтобы вся программа была расположена в оперативной памяти, так как обмен с внешними устройствами требует много времени. Скорость счета больших универсальных машин доходит до 1 млн. операций в секунду. Но даже и при таком быстродействии моделирование сложных систем в реальном времени невозможно. Изменение масштаба времени характеризуется коэффициентом масштаба

$$k_t = \frac{t_m}{t}, \quad (3.1.2)$$

где t — время работы исследуемой системы; t_m — время счета на ЭВМ.

Например, для сложных моделей коэффициент масштаба может достигать 1000, т.е. моделирование одной секунды работы изделия занимает примерно 17 минут машинного времени. Отсюда ясно, какое большое значение для моделирования имеют быстродействие и память ЭВМ.

ЭВМ должна быть оснащена широким набором вводных и выводных устройств. Данные в машину могут вводиться с перфокарт, перфолент, с магнитной ленты, с телеграфа и фототелеграфа, с дисплея. Выходная информация может выводиться на узкую печать, широкую печать, на перфокарты и перфоленты, на магнитную ленту, в виде таблиц на АЦПУ, на экран дисплея, в виде графиков на графопостроителе. Последнее представление наиболее наглядно.

Основные элементы моделей оптико-электронных систем. Для целей математического моделирования удобно классифицировать оптико-электронные системы по следующим признакам:

- тип источника излучения;
- диапазон оптического излучения;
- тип анализатора изображения;
- метод выделения координат объекта.

Тип источника излучения (излучающий объект, отражающий объект) определяет тип оптико-электронной системы. Используемый оптический диапазон совместно с типом системы определяет метод расчета входного сигнала. Тип анализатора изображения задает методику моделирования преобразования оптического сигнала в электрический. Метод выделения координат объекта задает структуру электронно-усилительного тракта. Все эти признаки представлены в табл. 3.1.

Пассивная система воспринимает оптическое излучение самоизлучающих объектов или объектов, подсвеченных естественным освещением. *Полуактивная* система воспринимает оптическое излучение искусственно подсвеченных объектов со стороны. *Активная* система воспринимает излучение объектов, подсвеченных излучателем, расположенным в самой оптико-электронной системе.

Анализаторы изображения (АИ) делятся по степени разложения изображения на отдельные элементы. *Модулирующий* анализатор изображения осуществляет модуляцию оптического излучения перед фотоприемником, интегрально воспринимающим оптический поток. *Матричный* анализатор производит разложение изображения на ограниченное число элементов. *Растровый* анализатор разлагает изображение на большое число элементов и в пределе производит непрерывную развертку (например, телевизионный анализатор изображения).

Выделение объекта из окружающего фона может производиться по его яркости, контрасту относительно фона или корреляционными методами.

Обобщенная схема оптико-электронной системы самонаведения в общем случае воздействия помех [67] приведена на рис. 3.3. Вне зависимости

Таблица 3.1

Тип системы	Оптический диапазон	Анализатор изображения	Метод выделения координат
Пассивная	Видимый	Модулирующий	Яркостный
Полуактивная	Инфракрасный	Матричный	Контрастный
Активная		Растровый	Корреляционный

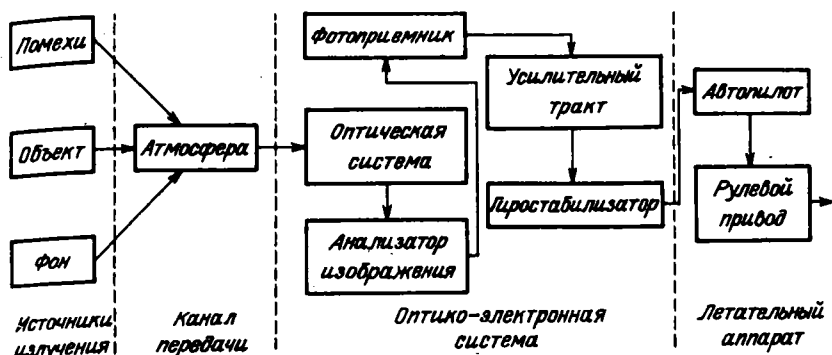


Рис. 3.3. Обобщенная схема оптико-электронной следящей системы

от типа конкретной системы можно выделить следующие основные элементы: объект, фон, помехи, атмосфера, объектив, фотоприемник и анализатор изображения, усилительный тракт, гиростабилизатор оптической системы в пространстве. В связи с этим полные математические модели оптико-электронных систем с разными принципами действия содержат общие основные блоки:

- модель объекта;
- модель атмосферы,
- модель оптической системы,
- модель анализатора изображения и фотоприемника,
- модель электронно-усилительного тракта,
- модель гиростабилизатора (исполнительного элемента),
- модель фонов и флуктуаций сигналов.

Получив математическое описание и алгоритмы моделирования рассмотренных основных элементов, нетрудно составить полную поэлементную модель любой оптико-электронной системы.

Ниже даются методы моделирования отдельных элементов оптико-электронных систем, а в § 3.7 рассмотрен пример построения поэлементной модели системы. При этом учитывается следующее. Построение модели объекта и атмосферы основывается на законах излучения, отражения и распространения оптического излучения в атмосфере и достаточно подробно рассмотрено в специальной литературе (см., например, [48, 49]). Поэтому в данной работе рассматривается методика моделирования преобразований сигнала оптико-электронной системы, начиная с ее оптического входа.

§ 3.2. Математическое моделирование преобразования сигнала в оптико-электронном тракте

Моделирование изображения в фокальной плоскости. Построение изображения точечного объекта. Рассмотрим преобразование оптической системой распределения яркости в пространстве предметов в функцию распределения освещенности в плоскости изображения, схематически представленное на рис. 3.4. Объект расположен в плоскости

X_g, O_g, Y_g , изображение объекта образуется в плоскости XOY . Оптическая система (ОС) условно показана в виде одной линзы; f — фокусное расстояние объектива; α — апертурный угол; L — расстояние до плоскости предметов. Каждой точке объекта с координатами (x_g, y_g) соответствует точка изображения с координатами (x, y) .

Рассмотрим точечный объект; расположенный в начале координат O_g . Изображение этого объекта по многочисленным причинам (абберации, дифракция и т.д.) занимает некоторую область вокруг точки O , образуя

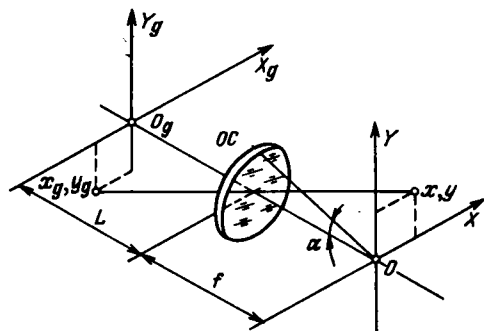


Рис. 3.4. Геометрия оптического изображения

так называемый *кружок рассеяния* или *функцию рассеяния точки* (ФРТ). Распределение освещенности внутри кружка рассеяния, являющееся функцией координат, обозначим $g(x, y)$. Если представить точечный объект как двумерную δ -функцию $\delta(x, y)$, то функция рассеяния точки $g(x, y)$ есть не что иное, как двумерная импульсная характеристика оптической системы.

Если объект находится в точке с координатами (x_g, y_g) , то его изображение имеет распределение освещенности, зависящее от координат объекта $g(x, y, x', y')$, где $x' = Mx_g, y' = My_g$, а M — передаточный коэффициент оптической системы по трансформации координат, характеризующий масштаб изображения и равный

$$M = \frac{f}{L}. \quad (3.2.1)$$

С учетом переверота изображения оптической системой коэффициент M имеет знак "минус", однако в большинстве случаев математического моделирования это может не учитываться.

Функция рассеяния точки или импульсная характеристика $g(x, y)$ является одной из важнейших характеристик оптической системы. При математическом моделировании ОЭС в соответствии с принципом функционального моделирования не рассматривается конструкция конкретных объективов, а оптическая система задается ее ФРТ. Величина кружка рассеяния определяет разрешающую способность оптической системы, т.е. минимальное расстояние между двумя точечными объектами, когда они еще различаются отдельно. Для математического моделирования величина кружка рассеяния (обычно в линейных размерах) задается разработчиком оптической системы. На предварительных этапах моделирования качество оптической системы для видимого диапазона может быть оценено через

диаметр входного зрачка по формулам [10]

$$r_0 = f \operatorname{tg} \psi, \quad \psi = \frac{122}{D}, \quad (3.2.2)$$

где r_0 — радиус кружка рассеяния; ψ — угловой размер кружка рассеяния в угловых минутах; D — диаметр входного зрачка в миллиметрах.

Если задана разрешающая способность ψ_n в линиях на миллиметр, то величина кружка рассеяния определяется просто:

$$r_0 = \frac{1}{2\psi_n}. \quad (3.2.3)$$

Импульсная характеристика оптической системы зависит от направления падения лучей на оптическую систему и является несимметричной функцией относительно положения точечного излучателя. В общем виде ФРТ могут быть следующих типов в порядке возрастания сложности [87]:

а) разделимая пространственно-инвариантная ФРТ

$$g(x, y, x', y') = g_x(x - x') g_y(y - y'); \quad (3.2.4)$$

б) разделимая пространственно-зависимая ФРТ

$$g(x, y, x', y') = g_x(x, x') g_y(y, y'); \quad (3.2.5)$$

в) неразделимая пространственно-инвариантная ФРТ

$$g(x, y, x', y') = g(x - x', y - y'); \quad (3.2.6)$$

г) неразделимая пространственно-зависимая ФРТ

$$g(x, y, x', y'). \quad (3.2.7)$$

Практически, учитывая небольшое поле зрения большинства ОЭС автоматического управления, удастся хорошо описать ФРТ выражениями типа (3.2.4) и (3.2.6), но в ряде случаев для учета неравномерности разрешения по полю зрения приходится использовать и пространственно-зависимые ФРТ.

А п р о к с и м и р у ю щ и е в ы р а ж е н и я д л я Ф Р Т. Рассмотрим выражения, описывающие распределение освещенности в пятне рассеяния. Теоретически ФРТ описывается только для идеального объектива, в котором отсутствуют aberrации. Вследствие дифракции даже идеальный объектив дает изображение изолированной точки в виде пятна конечных размеров, распределение освещенности в котором описывается функцией [81]

$$g(x, y) = \left[\frac{2J_1\left(\frac{2\pi\alpha}{\lambda}\sqrt{x^2 + y^2}\right)}{\frac{2\pi\alpha}{\lambda}\sqrt{x^2 + y^2}} \right]^2, \quad (3.2.8)$$

где λ — длина волны; α — апертурный угол в радианах.

Функция Бесселя первого порядка $J_1(z) = 0$ при $z = 3,83$. Принимая координату первого нуля за границу кружка рассеяния, получаем

$$\frac{2\pi\alpha}{\lambda} r_0 = 3,83, \quad (3.2.9)$$

откуда радиус кружка рассеяния равен

$$r_0 = \frac{3,83\lambda}{2\pi\alpha}. \quad (3.2.10)$$

Для реальных оптических систем импульсную характеристику невозможно записать в виде точной формулы, поэтому она описывается приближенно.

При математическом моделировании используются обычно следующие допущения.

1. Функция рассеяния точки принимается симметричной относительно начала координат и относительно положения точечного излучателя в силу малости углов поля зрения оптической системы и отклонения объекта от оптической оси в системах автоматического управления.

2. Функция рассеяния точки аппроксимируется простыми и удобными для вычисления выражениями, не содержащими специальных функций.

Наиболее часто используются следующие аппроксимирующие формулы.

Аппроксимация функцией Гаусса. Распределение освещенности в кружке рассеяния задается формулой

$$g(x, y) = A e^{-\frac{x^2 + y^2}{c^2}}. \quad (3.2.11)$$

Константа c выбирается из условия, чтобы радиус горизонтального сечения гауссоиды на заданном уровне от максимального (0,02–0,05) был равен радиусу кружка рассеяния r_0 , т.е.

$$e^{-\frac{r_0^2}{c^2}} = 0,02. \quad (3.2.12)$$

Тогда $c = r_0/2$.

Константа A выбирается из условия нормировки

$$\iint_{-\infty}^{\infty} A e^{-\frac{x^2 + y^2}{c^2}} dx dy = 1. \quad (3.2.13)$$

Такая нормировка удобна для введения в математическую модель абсолютного уровня сигнала, для чего достаточно вычислить величину сигнала на входе оптической системы и умножить на коэффициент пропускания оптики. С учетом (3.2.12) получим

$$A = \frac{1}{\pi c^2} = \frac{1,25}{r_0^2}. \quad (3.2.14)$$

Окончательно аппроксимирующая формула имеет вид

$$g(x, y) = 1,25 r_0^{-2} \exp \{-4 r_0^{-2} (x^2 + y^2)\}. \quad (3.2.15)$$

Аппроксимация четной степенью косинуса. Распределение освещенности задается формулой

$$g(x, y) = \begin{cases} A \cos^{2n} \left(\frac{\pi}{c} \sqrt{x^2 + y^2} \right) & \text{при } z = \sqrt{x^2 + y^2} \leq r_0, \\ 0 & \text{при } z > r_0. \end{cases} \quad (3.2.16)$$

Из условия

$$\frac{\pi}{c} r_0 = \frac{\pi}{2}$$

определяется константа c :

$$c = 2r_0. \quad (3.2.17)$$

Константа A для случая $n = 1$ вычисляется из условия

$$\iint_S A \cos^2 \left(\frac{\pi}{2r_0} \sqrt{x^2 + y^2} \right) dx dy = 1 \quad (3.2.18)$$

и равна

$$A = 1,06 r_0^{-2}. \quad (3.2.19)$$

Теперь аппроксимирующая формула запишется в виде

$$g(x, y) = 1,06 r_0^{-2} \cos^2 \left(\frac{\pi}{2r_0} \sqrt{x^2 + y^2} \right). \quad (3.2.20)$$

Аппроксимация произведением косинусов четной степени. Распределение освещенности в кружке рассеяния задается формулой

$$g(x, y) = \begin{cases} A \cos^{2n} \left(\frac{\pi x}{2r_0} \right) \cos^{2n} \left(\frac{\pi y}{2r_0} \right) & \text{при } |x| \leq r_0, |y| \leq r_0, \\ 0 & \text{при } |x| > r_0, |y| > r_0. \end{cases} \quad (3.2.21)$$

В отличие от (3.2.16) сечение ФРТ в данном случае имеет вид квадрата со сглаженными углами. Величина константы A зависит от показателя степени n и для различных n приведена в [81]. Например, для $n = 1$ имеем $A = r_0^{-2}$. Таким образом, аппроксимация ФРТ формулами (3.2.15), (3.2.20) и (3.2.21) дают близкий результат, что иллюстрируется рис. 3.5, где приведены значения аппроксимирующих выражений $g\left(\frac{x}{r_0}, 0\right)$.

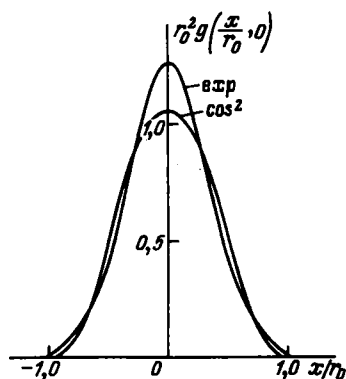


Рис. 3.5. Аппроксимация функции рассеяния точки

Аппроксимация кружком рассеяния равномерной освещенности. Распределение освещенности задается выражением

$$g(x, y) = \begin{cases} A & \text{при } z = \sqrt{x^2 + y^2} \leq r_0, \\ 0 & \text{при } z > r_0, \end{cases} \quad (3.2.22)$$

т.е. ФРТ аппроксимируется цилиндром. Условие нормировки дает

$$A \pi r_0^2 = 1,$$

откуда

$$A = \frac{1}{\pi r_0^2} = 0,318 r_0^{-2}. \quad (3.2.23)$$

Аппроксимация равномерной освещенностью используется для оптических систем с пониженным разрешением или с искусственно созданной расфокусировкой.

Наличие аберрации приводит к тому, что пятно рассеяния на краю поля зрения объектива больше, чем в центре, т.е. разрешающая способность на краю меньше. При моделировании ОЭС это явление приходится учитывать, чтобы получить совпадение реальных характеристик с полученными при моделировании. В этом случае радиус пятна рассеяния r_0 делают переменным в зависимости от величины смещения от оптической оси $z = \sqrt{x^2 + y^2}$. Хорошие результаты дает описание размера пятна рассеяния линейной зависимостью

$$r_0 = r_{0 \min} \left(1 + \frac{\beta}{z_{\max}} \sqrt{x^2 + y^2} \right), \quad (3.2.24)$$

где $z_{\max} = f \operatorname{tg} \alpha$ — максимальное смещение в плоскости изображения; β — параметр, определяемый качеством объектива.

Если заданы размеры кружка рассеяния в центре $r_{0 \min}$ и на краю $r_{0 \max}$ (при $z = z_{\max}$), то

$$\beta = \frac{r_{0 \max}}{r_{0 \min}} - 1. \quad (3.2.25)$$

Если разрешающая способность задана в линиях на миллиметр, то

$$\beta = \frac{\psi_{\text{л max}}}{\psi_{\text{л min}}} - 1. \quad (3.2.26)$$

Далее значение кружка рассеяния r_0 из (3.2.24) подставляется в аппроксимирующее выражение для ФРТ.

Описание оптической системы импульсной характеристикой $g(x, y)$ означает описание ее как пространственного двумерного фильтра. Поэтому при теоретическом исследовании и моделировании можно применять методы, разработанные для одномерных временных фильтров с учетом следующих особенностей. Импульсная характеристика временных фильтров несимметрична и отлична от нуля только при $t \geq 0$, так как отклик не может появиться раньше приложенного воздействия. Импульсная характеристика пространственного фильтра является четной функцией и су-

существует для всех значений аргументов $-\infty < x < \infty$, $-\infty < y < \infty$. Это свойство используется в § 3.5 для моделирования случайных полей. Кроме того, сигнал на выходе временного фильтра запаздывает по отношению к входному; в пространственных фильтрах такого запаздывания (в пространстве) не происходит.

Последовательное изложение теории оптических передаточных функций содержится в монографии [81].

Построение изображения протяженного источника. Импульсная характеристика $g(x, y)$ оптической системы описывает распределение освещенности в изображении изолированной точки. Такими точками можно считать малоразмерные объекты на большой дальности, например звезды в системах астроориентации. В большинстве случаев объекты приходится рассматривать как размерные и для вычисления сигнала на выходе оптической системы использовать двумерный интеграл свертки

$$B_{\text{и}}(x, y) = \iint_{-\infty}^{\infty} B_0(x', y') g(x, y, x', y') dx' dy', \quad (3.2.27)$$

где $B_0(x', y')$ — распределение яркости объекта с учетом масштабного преобразования; $B_{\text{и}}(x, y)$ — распределение яркости изображения в фокальной плоскости оптической системы. Здесь интеграл свертки записан в общем виде для неразделимой пространственно-зависимой функции рассеяния точки.

Как будет подробно рассмотрено в § 3.3, при математическом моделировании изображения задается в дискретных точках x_i, y_i с шагом $\Delta x = \Delta y = \Delta$. Тогда интеграл свертки представляется двойной суммой, которая для наиболее часто встречающейся неразделимой пространственно-инвариантной ФРТ запишется в виде

$$\begin{aligned} *B_{\text{и}}(x_i, y_j) &= \sum_{\alpha} \sum_{\beta} B_0(x_{\alpha}, y_{\beta}) g(x_i - x_{\alpha}, y_j - y_{\beta}) \Delta x \Delta y = \\ &= \sum_{\alpha} \sum_{\beta} B_0(x_i - x_{\alpha}, y_j - y_{\beta}) g(x_{\alpha}, y_{\beta}) \Delta x \Delta y, \end{aligned} \quad (3.2.28)$$

где пределы суммирования α и β определяются шириной импульсной характеристики оптической системы, т.е. размером кружка рассеяния r_0 , следующим образом:

$$\alpha_{\min} = \beta_{\min} = -\frac{r_0}{\Delta}, \quad \alpha_{\max} = \beta_{\max} = \frac{r_0}{\Delta}. \quad (3.2.29)$$

Практически это сводится к суммированию яркостей точек объекта, "накрываемых" импульсной характеристикой, с весами, пропорциональными значению ФРТ в соответствующей точке.

Представление преобразования оптического сигнала формулой (3.2.28) соответствует описанию оптической системы как двумерного нерекурсивного цифрового фильтра. Преобразование изображения цифровым фильтром можно записать в матричной форме [87]. Распределение яркости объекта, заданное в дискретных точках, представим двумерной матрицей B_0 . Импульсная характеристика двумерного фильтра $g(x, y, x', y')$, задан-

ная также в дискретных точках, может быть представлена матрицей G , называемой *пакетным оператором*. Тогда действие оптической системы можно свести к операции перемножения матриц

$$B_n = GB_0, \quad (3.2.30)$$

где B_n — матрица, представляющая собой цифровое изображение объекта и фокальной плоскости оптической системы.

В случае разделимой ФРТ матрица G представляется в виде кронекера произведения матриц

$$G = G_y [G_x]^T, \quad (3.2.31)$$

где G_x, G_y — матрицы преобразования по осям x, y соответственно. При этом элементы матрицы G имеют вид

$$g_{ij, \alpha\beta} = g_{i\alpha} g_{j\beta}, \quad (3.2.32)$$

и изображение на выходе оптической системы записывается следующим образом:

$$B_n = G_y B_0 G_x. \quad (3.2.33)$$

В этом случае изображение обрабатывается отдельно по строкам и столбцам, что приводит к экономии вычислений.

Импульсные характеристики оптической системы обычно имеют центральную симметрию: $G_x = G_y = G$. Тогда элементы матрицы выходного изображения можно представить в виде [9]

$$g_{ij} = \sum_{\beta} \lambda_{\beta} Q_i^{\beta} Q_j^{\beta}, \quad (3.2.34)$$

где $Q = GV$, а V^{β} и λ_{β} — собственные векторы и собственные числа матрицы цифрового изображения B_0 . Таким образом, представляется возможность использовать не сами цифровые изображения в виде матрицы B_0 , а работать с собственными векторами и собственными числами изображения, рассматривая его как оператор. В ряде случаев такой подход может привести к повышению быстродействия программы и экономии памяти ЭВМ. Подробнее кодирование изображений путем преобразований рассмотрено в § 3.3.

Дискретные значения ФРТ обычно выбирают в количестве 25–30, так что общее машинное время при вычислении двумерной свертки получается довольно большим. Поэтому формулу (3.2.28) используют только при построении изображений объектов с переменной яркостью, например случайных полей, описывающих облачный или земной фон. Для объектов заданной геометрической формы и с равномерной яркостью отдельных участков изображение строят с помощью масштабного преобразования (3.2.1), не вычисляя свертки, а ФРТ учитывают только на краях изображения в виде краевого следа, т.е. размытого изображения резкой границы. В наиболее простом случае равномерной освещенности в кружке рассеяния (3.2.22) изображение объекта просто увеличивается на величину кружка рассеяния. Другой крайний случай — моделирование оптических систем с высокой разрешающей способностью. В этом случае размер кружка рассеяния меньше самых мелких существенных деталей изображения и ФРТ можно не моделировать (она вырождается в двумерную δ -функцию) и проводить только масштабное преобразование.

Это можно показать на примере преобразования оптической системой двумерного случайного поля яркости. Пусть весовая функция объектива задана выражением

$$g(x, y) = \dot{g}(z) = A e^{-2\gamma(x^2 + y^2)} = A e^{-2\gamma z^2},$$

где $z = \sqrt{x^2 + y^2}$ — модуль смещения. Соответствующая частотная характеристика определяется преобразованием Ганкеля [81]

$$\begin{aligned} G(k) &= 2\pi \int_0^\infty g(z) J_0(kz) z dz = \\ &= 2\pi A \int_0^\infty z e^{-2\gamma z^2} J_0(kz) dz = \frac{A\pi}{2\gamma} e^{-\frac{k^2}{8\gamma}}. \end{aligned}$$

Здесь $k = \sqrt{u^2 + v^2}$ — модуль пространственной частоты; u, v — пространственные частоты по координатам x, y .

Зададим корреляционную функцию изотропного случайного поля выражением

$$R_0(z) = e^{-\alpha z^2},$$

при этом двумерный спектр будет

$$\begin{aligned} S_0(k) &= 2\pi \int_0^\infty R_0(z) J_0(kz) z dz = \\ &= 2\pi \int_0^\infty e^{-\alpha z^2} J_0(kz) z dz = \frac{\pi}{\alpha} e^{-\frac{k^2}{4\alpha}}. \end{aligned}$$

Как известно, спектр случайного сигнала на выходе линейной системы определяется квадратом модуля ее частотной характеристики, т.е. спектр изображения в фокальной плоскости будет

$$\begin{aligned} S_H(k) &= S_0(k) G^2(k) = \\ &= \frac{\pi}{\alpha} e^{-\frac{k^2}{4\alpha}} \frac{A^2 \pi^2}{4\gamma^2} e^{-\frac{k^2}{4\gamma}} = \frac{A^2 \pi^3}{4\alpha\gamma^2} e^{-\frac{k^2}{4\delta}}, \end{aligned}$$

$$\text{где } \delta = \frac{\alpha\gamma}{\alpha + \gamma}.$$

Соответствующая корреляционная функция равна

$$\begin{aligned} R_H(z) &= \frac{1}{2\pi} \int_0^\infty S_H(k) J_0(kz) k dk = \\ &= \frac{A^2 \pi^2}{8\alpha\gamma^2} \int_0^\infty e^{-\frac{k^2}{4\delta}} J_0(kz) k dk = \frac{A^2 \pi^2 \delta}{4\alpha\gamma^2} e^{-\delta z^2}. \end{aligned} \quad (3.2.35)$$

Сравним ширину корреляционной функции исходного сигнала и преобразованного изображения в зависимости от качества оптической системы.

При $\gamma \gg \delta$, что соответствует хорошо сфокусированной системе, получим $\delta \approx \alpha$ и $R_{\text{и}}(z) = R_0(z)$. В этом случае влияние оптической системы на изображение не существенно и ее учет при моделировании производится только масштабным преобразованием. При $\gamma \ll \alpha$, что соответствует сильно расфокусированной системе, $\delta \approx \gamma$, корреляционная функция изображения не зависит от входного сигнала и имеет место полное размытие изображения.

Необходимо отметить следующее. Если плоскость объекта и плоскость изображения не параллельны друг другу, т.е. объект визируется под углом, то в изображении возникают ракурсные искажения, которые должны обязательно учитываться. При этом наряду с масштабным преобразованием нужно провести преобразования, связанные с переходом в повернутую систему координат. Эти преобразования широко известны (см., например, [34]) и здесь не рассматриваются.

Моделирование анализаторов изображения. Общий метод моделирования анализаторов изображения. Для определения координат объекта изображение, создаваемое оптической системой, подвергается анализу с помощью специальных устройств, называемых, как известно, анализаторами изображения. Анализаторы изображения (АИ) располагаются в фокальной плоскости оптической системы и соответствующим образом перемещаются относительно изображения (или изображение относительно их). При этом световой поток модулируется или развертывается во времени. Часто анализатор изображения одновременно является и фотоприемником.

Анализаторы изображения решают три основные задачи.

1. Преобразование пространственного распределения светового потока в функцию времени (электрический сигнал) для дальнейшего усиления и преобразования средствами электроники.

2. Выделение полезного сигнала от объекта из фоновых помех, т.е. осуществление пространственной фильтрации, в результате чего повышается отношение сигнал/шум.

3. Определение координат объекта в системе координат оптико-электронной системы, т.е. внесение в электрический сигнал определенным образом закодированной информации о положении. Это дает возможность осуществить автоматическое слежение за объектом.

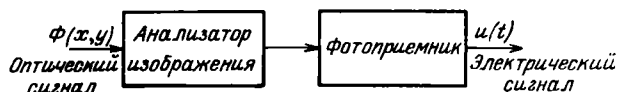


Рис. 3.6. Преобразование сигнала анализатором изображения

В общем виде преобразование сигнала анализатором изображения и фотоприемником показано на рис. 3.6.

Различные типы анализаторов изображения рассмотрены в [14, 67, 121]. В соответствии с табл. 3.1 по степени разложения оптического изображения на отдельные элементы анализаторы изображения можно разделить на следующие группы.

1. Модулирующие АИ без разложения изображения с интегральным фотоприемником.

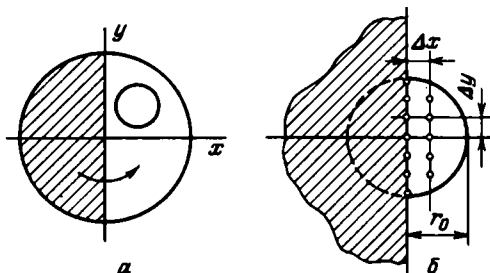
2. Матричные АИ, простейшим примером которых является четырех-площадочный анализатор изображения.

3. Растровые АИ, производящие развертку изображения, частным случаем которых является телевизионный анализатор изображения.

В дальнейшем будут рассмотрены методы моделирования анализаторов каждой группы.

Моделирование анализаторов изображения является центральным моментом в модели оптико-электронной системы. Отметим, что этот вопрос

Рис. 3.7. Оценка точности моделирования анализатора изображения



практически не освещен в литературе, за исключением небольшого раздела в [99]. Основная трудность заключается в вычислении доли сигнала, прошедшего через анализатор изображения сложной формы на фотоприемник.

Электрический сигнал на выходе фотоприемника при условии постоянства его чувствительности S по полю зрения равен

$$u = S \iint_D E_n(x, y) dx dy, \quad (3.2.36)$$

где $E_n(x, y)$ — освещенность в точке фотоприемника с координатами (x, y) ; D — площадь засвеченной части фотоприемника или прозрачной части анализатора изображения.

При сложной форме анализатора изображения интеграл (3.2.36) аналитически не вычисляется, поэтому его приходится находить численным методом, записав выражение (3.2.36) в виде двойной суммы

$$u = S \sum_{i=1}^n \sum_{j=1}^l E_n(x_i, y_j) \Delta x \Delta y. \quad (3.2.37)$$

Здесь $E_n(x_i, y_j)$ — изображение объекта в дискретной форме с шагом $\Delta x = \Delta y = \Delta$, т.е. цифровое изображение.

В этом случае для каждой точки изображения с координатами (x_i, y_j) производится проверка попадания ее в область D и суммирование выполняется только для значений изображений $E_n(x_i, y_j)$, попавших в эту область. Границы области D задаются аналитическими выражениями, вид которых зависит от конструкции анализатора изображения.

Очевидно, что точность выражения (3.2.37) возрастает с увеличением n и l , т.е. с уменьшением шага дискретизации Δ . Однако при этом существенно возрастают вычислительные затраты. Оценим точность этого метода в зависимости от шага дискретизации. Рассмотрим простейший анализатор изображения в виде вращающегося полудиска, приведенный на рис. 3.7, а. Изображение точечного объекта примем в виде кружка рассея-

ния радиуса r_0 с равномерным распределением освещенности. На рис. 3.7, б в более крупном масштабе показан момент прохождения края модулятора через изображение объекта, заданное в дискретных точках с шагом $\Delta x = \Delta y = \Delta$. При использовании формулы (3.2.37) максимальная ошибка получается, когда элементарная полоса шириной Δx расположена параллельно краю модулятора. Эта ошибка пропорциональна площади полосы

$$\Delta u_{\max} = 2Sr_0\Delta. \quad (3.2.38)$$

Амплитуда сигнала на выходе модулятора пропорциональна площади пятна рассеяния

$$u_{\max} = S\pi r_0^2. \quad (3.2.39)$$

Тогда максимальная относительная ошибка равна

$$\delta_{\max} = \frac{\Delta u_{\max}}{u_{\max}} = \frac{2}{\pi} \frac{\Delta}{r_0}. \quad (3.2.40)$$

Минимальная ошибка, очевидно, равна нулю.

Поскольку ориентация изображения объекта относительно края модулятора произвольная, примем распределение ошибки равномерным. Тогда среднее значение ошибки равно

$$m_\delta = \frac{\delta_{\max} - \delta_{\min}}{2} = \frac{1}{\pi} \frac{\Delta}{r_0}, \quad (3.2.41)$$

а дисперсия

$$\sigma_\delta^2 = \frac{(\delta_{\max} + \delta_{\min})^2}{12} = \frac{1}{3\pi^2} \left(\frac{\Delta}{r_0} \right)^2. \quad (3.2.42)$$

Среднеквадратическое значение ошибки равно

$$\sigma_\delta = \frac{1}{4\sqrt{3}} \frac{\Delta}{r_0}. \quad (3.2.43)$$

Выразив радиус пятна в относительных единицах r_0/Δ , найдем количество точек изображения в кружке рассеяния

$$N = \pi \left(\frac{r_0}{\Delta} \right)^2. \quad (3.2.44)$$

Например, для $r_0/\Delta = 3$ получим

$$m_\delta = 0,106, \quad \sigma_\delta = 0,061, \quad N = 28.$$

Таким образом, при количестве точек в кружке рассеяния около 30 получаем среднюю ошибку порядка 10%. Повышение точности вычислений достигается увеличением отношения r_0/Δ , однако при этом, как видно из (3.2.43) и (3.2.44), точность вычислений повышается пропорционально

первой степени этого отношения, а количество точек, т.е. объем вычислений, пропорционально квадрату отношения. Следовательно, повышение точности вычислений влечет за собой резкое возрастание вычислительных затрат. Поэтому при математическом моделировании всегда нужно находить разумный компромисс между точностью и объемом вычислений.

Рассмотрим алгоритмы моделирования некоторых наиболее распространенных анализаторов изображения.

Вращающийся полудиск. На рис. 3.8 приведен анализатор изображения в виде вращающегося непрозрачного полудиска. Такой анализатор изображения определяет координаты объекта в поле зрения оптико-электронной системы по сравнению фазы модулированного сигнала с фазой опорного [67].

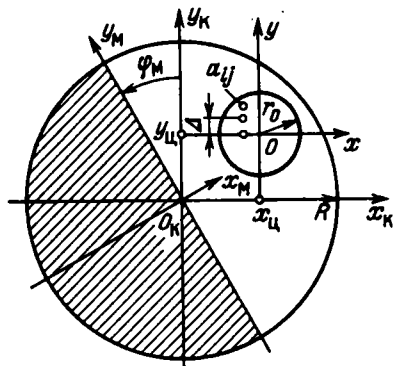


Рис. 3.8. Вращающийся полудиск

Введем системы координат: $O_k x_k y_k$ — система координат, связанная с оптической системой (система координатора), оптическая ось перпендикулярна к плоскости рисунка; $O_k x_m y_m$ — подвижная система координат, связанная с вращающимся модулятором; $O x y$ — система координат, связанная с центром изображения объекта. Для определенности будем рассматривать в качестве объекта точечный источник, тогда его изображение будет кружком рассеяния радиуса r_0 , освещенность внутри которого $g(x, y)$ задается в дискретных точках с шагом Δ . Модулятор вращается с угловой скоростью ω_m .

При моделировании прохождения оптического сигнала через АИ нужно для каждого момента времени t определить положение модулятора и затем определить точки изображения, попавшие на прозрачную часть. В этом случае алгоритм выглядит следующим образом.

Для известных координат центра $x_ц(t)$, $y_ц(t)$, вычисляемых по отдельной программе, находим координаты точки a_{ij} изображения в системе $O_k x_k y_k$:

$$x_i = x_ц(t) + i\Delta, \quad y_j = y_ц(t) + j\Delta. \quad (3.2.45)$$

В подвижной системе модулятора координаты этой точки будут

$$\begin{aligned} x_{mi} &= x_i \cos \omega_m t + y_j \sin \omega_m t, \\ y_{mj} &= y_j \cos \omega_m t - x_i \sin \omega_m t. \end{aligned} \quad (3.2.46)$$

Легко видеть, что точка изображения a_{ij} только тогда находится на

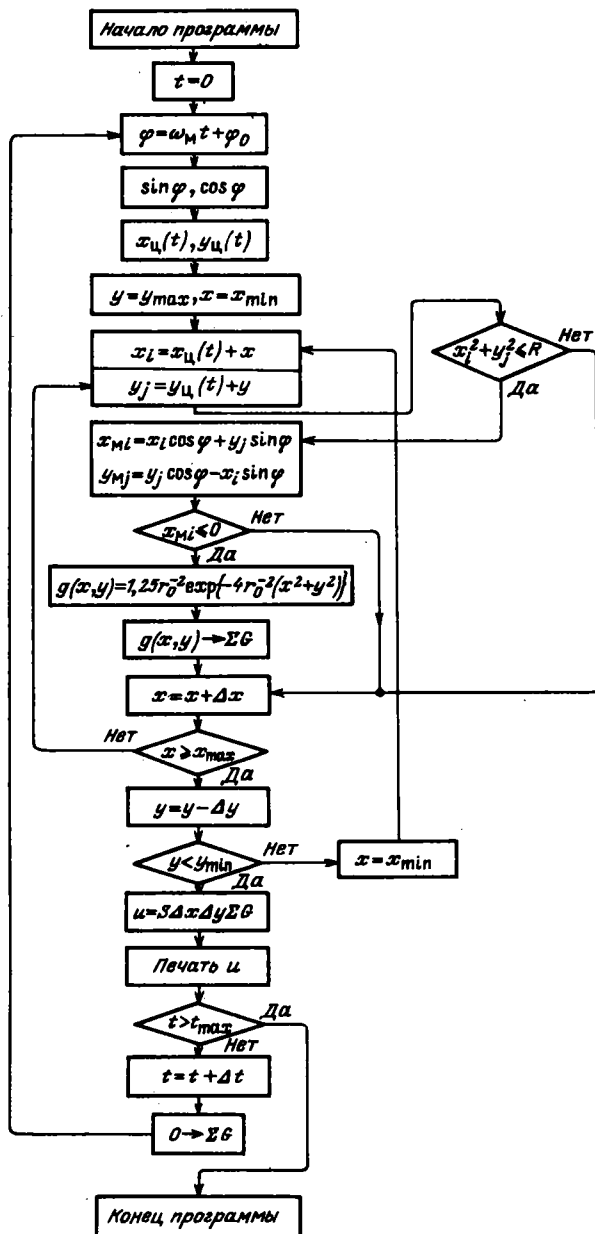


Рис. 3.9. Блок-схема моделирования полудискового анализатора изображения

прозрачной части модулятора, когда выполняется условие

$$x_{mi} \geq 0. \quad (3.2.47)$$

В этом случае вычисляется сигнал от данной точки $g(x_i, y_j)$ в соответствии с принятой аппроксимацией ФРТ и засылается в сумматор, моделирующий выражение (3.2.37). В противном случае, при $x_{mi} < 0$, точка находится на непрозрачной части модулятора и сигнал от нее равен нулю (не вычисляется).

Таким образом проводятся вычисления для всех точек изображения в каждый момент времени t . Аналогичные вычисления повторяются для следующего момента времени с шагом Δt . Индексы i, j в (3.2.45) определяются размерами изображения. Для объектов заданной геометрической формы удобнее производить перебор точек, изменяя координаты x, y от $x_{\min}$ до $x_{\max}$ и от $y_{\min}$ до $y_{\max}$ с шагом Δ . Например, для круглого изображения

$$\begin{aligned} y_{\max} &= r_0, & x_{\max} &= \sqrt{r_0^2 - y^2}, \\ y_{\min} &= -r_0, & x_{\min} &= -\sqrt{r_0^2 - y^2}. \end{aligned} \quad (3.2.48)$$

Подробнее рассмотренный алгоритм представлен на рис. 3.9. Для определенности распределение освещенности внутри кружка рассеяния аппроксимируется гауссоидой (3.2.16). Как видно из блок-схемы, программа состоит из ряда проверок попадания точек изображения за границу поля зрения, границу изображения, границу модулятора. Аналогичные алгоритмы получаются для моделирования анализаторов изображения других типов, поэтому в дальнейшем не будем приводить полную блок-схему, а ограничимся рассмотрением основных моментов при моделировании анализаторов других типов.

Вращающийся радиально-щелевой модулятор. Анализатор изображения данного типа приведен на рис. 3.10. Здесь производится модуляция потока с двойной частотой (двойная амплитудная модуляция), что позволяет произвести пространственную селекцию точечного объекта [67].

Алгоритм проверки попадания точки изображения a_{ij} на прозрачную или непрозрачную часть модулятора распадается на две задачи. Во-первых,

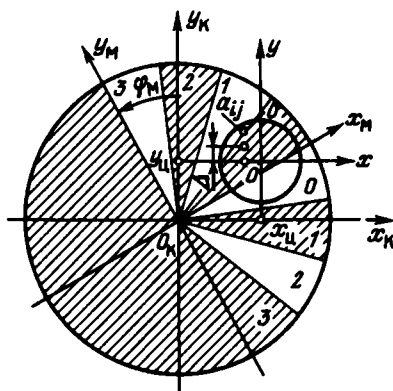


Рис. 3.10. Радиально-щелевой модулятор

Решение первой задачи производится по алгоритму, рассмотренному выше. Вторая задача решается путем сравнения угла α_{mij} в подвижной системе модулятора с величиной углового размера сектора. Величина угла равна

Пронумеруем секторы от 0 до n (на рис. 3.9 $n = 3$) в каждой из четвертей полупрозрачной половины.

$$n = \text{int} \left(\left| \frac{\alpha_{Mij}}{\alpha_0} \right| \right) \quad (3.2.50)$$
$$y_{mj} \geq 0, \quad n \text{ нечетное,} \quad (3.2.51)$$
$$u = S \Delta x \Delta y \sum g(x_i, y_j). \quad (3.2.52)$$

переносится по плоскости неподвижного модулятора [73]. Если объект находится на оптической оси, центр окружности переноса радиуса $R_{\text{п}}$ совпадает с центром модулятора, а смещение объекта вызывает смещение центра $O_{\text{м}}$ окружности переноса.

192

перейти от подвижной системы координат $O_M x_M y_M$ к неподвижной $O_K x_K y_K$. В произвольный момент времени координаты центра изображения в подвижной системе равны

$$x_{ц} = R_{п} \cos \omega_M t, \quad y_{ц} = R_{п} \sin \omega_M t. \quad (3.2.53)$$

Координаты центра изображения в системе координатора $O_K x_K y_K$

$$x_K = x_{ц} + x_M, \quad y_K = y_{ц} + y_M, \quad (3.2.54)$$

где x_M, y_M — координаты центра окружности переноса при смещении объекта от оптической оси.

Координаты точки изображения a_{ij} в системе $O_K x_K y_K$

$$x_{ki} = x_K + i\Delta, \quad y_{kj} = y_K + j\Delta. \quad (3.2.55)$$

Как видно из рис. 3.11, плоскость АИ можно разбить на внешнюю, кольцевую, зону и центральную, границей между которыми является окружность радиуса l_3 . Поэтому анализ положения точки a_{ij} состоит из трех проверок: попадания в плоскость АИ, попадания в кольцевую зону и попадания в центральную зону АИ. Попадание в плоскость АИ задается условием

$$l = \sqrt{x^2 + y^2} \leq l_4, \quad (3.2.56)$$

где для сокращения записи обозначено $x_{ki} = x, y_{kj} = y$. Попадание в кольцевую зону происходит при выполнении неравенств

$$l_3 < l \leq l_4,$$

а попадание в центральную зону происходит при $l < l_3$.

Если точка изображения a_{ij} попадает в кольцевую зону, то производится сравнение ее углового положения

$$|\alpha_{ij}| = \arctg \left| \frac{y}{x} \right| \quad (3.2.57)$$

последовательно с углами $\varphi_1, \varphi_2, \varphi_3, \varphi_4$. Выполнение условия $|\alpha_{ij}| < \varphi_n$ определяет номер кольцевого сектора, который имеет прозрачность 0 или 1. При попадании точки на прозрачный сектор вычисляется сигнал $g(x, y)$ (здесь x, y — координаты в системе Oxy), который засылается в сумматор. Далее производится переход к следующей точке изображения.

При попадании точки в центральную часть АИ производится проверка координат x_{ki}, x_{kj} на величину отклонений l_1 и l_2 . Точка попадает на прозрачную часть, если одновременно выполняются условия

$$l_1 < |x_{ki}| \leq l_2, \quad l_1 < |y_{kj}| \leq l_2. \quad (3.2.58)$$

Таким образом, все проверки можно разбить на две группы: проверка по углу и проверка по величине отклонения. В связи с симметричностью рисунка модулятора все проверки можно производить для одной четверти, взяв все углы и координаты по абсолютной величине.

Матричный анализатор изображения. Анализатор изображения, выполненный в виде матрицы фотоприемника, приведен на 13. Ю.М. Астапов

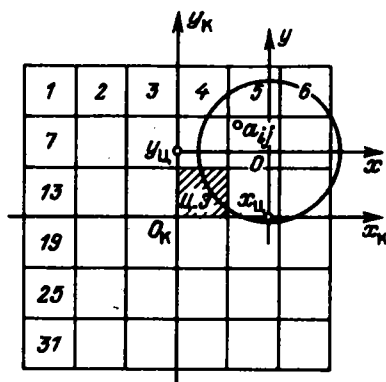


Рис. 3.12. Матричный анализатор изображения

рис. 3.12. Для определенности изображена матрица размера 6×6 элементов. Частным случаем такого фотоприемника является четырехплощадочный фотоприемник. Изображение, как и ранее, рассматривается в виде кружка рассеяния; координаты элементарной точки суть $x_i = x_{\text{ц}} + i\Delta$, $y_j = y_{\text{ц}} + j\Delta$.

Можно построить алгоритм последовательной проверки попадания точки изображения в элементы матрицы,

однако такой алгоритм содержит большое число вычислений, особенно при большой размерности матрицы. Поэтому здесь целесообразно построить алгоритм, позволяющий сразу указать номер элемента, в который попадает точка изображения.

Назовем один из элементов матрицы, примыкающий к началу координат O_k , *центральным*. Тогда номер элемента, в который попадает точка изображения a_{ij} с координатами x_i, y_j , определяется выражением

$$N = N_{\text{ц.э}} + N_y + NN_x. \quad (3.2.59)$$

Здесь $N_{\text{ц.э}}$ — номер центрального элемента; N — количество элементов в строке; N_x, N_y — количество целых элементов, укладывающихся на соответствующих отрезках:

$$N_x = \text{int} \left(\frac{x_i}{\Delta_m} \right), \quad N_y = \text{int} \left(\frac{y_j}{\Delta_m} \right), \quad (3.2.60)$$

где Δ_m — размер элемента матрицы.

При использовании этого алгоритма необходимо учитывать, что центральный элемент примыкает к началу координат справа и сверху. Поэтому при положительных координатах ($x_i > 0, y_j > 0$) в (3.2.59) подставляются N_x и N_y непосредственно из (3.2.60), а при отрицательных координатах

$$N_x = \text{int} \left(\frac{x_i}{\Delta_m} \right) - 1, \quad N_y = \text{int} \left(\frac{y_j}{\Delta_m} \right) - 1. \quad (3.2.61)$$

Такой алгоритм дает большую экономию вычислений по сравнению с последовательным просмотром. На каждый элемент матрицы отводится свой сумматор $\Sigma g(x_i, y_j)$, в который заносится значение яркости попавшей на этот элемент точки.

Телевизионный анализатор изображения. В данном анализаторе производится последовательный просмотр всего поля зрения по определенной траектории сканирования. При использовании обычного телевизионного стандарта это линейная развертка. В связи с большим

объемом информации, содержащейся в телевизионном кадре, моделирование полного телевизионного сигнала весьма затруднительно в связи с необходимостью больших вычислительных затрат. Поэтому одной из основных задач при математическом моделировании является воспроизведение телевизионного сигнала с разумным ограничением избыточной информации.

Свяжем количество строк телевизионной развертки при моделировании с верхней воспроизводимой частотой телевизионного сигнала. Длительность одной строки равна

$$T_c = \frac{T_k}{n}; \quad (3.2.62)$$

где n — количество строк разложения в модели; T_k — период кадровой развертки.

При одинаковом количестве элементов разложения по строкам и столбцам шаг по времени между двумя элементами разложения

$$\Delta t_p = \frac{T_c}{n} = \frac{T_k}{n^2}. \quad (3.2.63)$$

По теореме Котельникова [135] верхняя частота сигнала связана с шагом счета выражением

$$f_b = \frac{1}{2 \Delta t_p} = \frac{n^2}{2 T_k}. \quad (3.2.64)$$

Отсюда получаем необходимое количество элементов разложения в строке при заданной верхней частоте телевизионного сигнала

$$n = \sqrt{\frac{2 f_b}{F_k}}, \quad (3.2.65)$$

где $F_k = 1/T_k$ — кадровая частота.

Обычно в оптико-электронных системах не используется полная ширина спектра телевизионного сигнала, поэтому размер $n \times n$ матрицы разложения кадра получается в разумных пределах. Например, при $f_b = 300$ кГц, $F_k = 25$ Гц получим $n = 155$. Тогда шаг дискретизации по координатам равен

$$\Delta x_p = \Delta y_p = \Delta t_p V_p = \frac{L}{n}, \quad (3.2.66)$$

где V_p — скорость развертки; L — размер мишени телевизионного анализатора. Фильтрующее свойство развертывающего элемента характеризуется его импульсной характеристикой $g_p(x, y)$, а траектория движения задается координатами $x_p(t)$, $y_p(t)$. Если форма развертывающего элемента не изменяется в процессе развертки (пространственно-инвариантная импульсная характеристика), то сигнал на выходе телевизионного анализатора описы-

вается выражением

$$\begin{aligned} u(t) &= u[x(t), y(t)] = \\ &= \iint B_n(x, y) g_p[x - x_p(t), y - y_p(t)] dx dy, \end{aligned} \quad (3.2.67)$$

где $B_n(x, y)$ — изображение, сформированное в фокальной плоскости оптической системы.

Для перехода к дискретной форме положим

$$\begin{aligned} x &= i\Delta x, \quad y = j\Delta y, \quad x_p = \alpha\Delta x_p = kV_{px}\Delta t, \\ y_p &= \beta\Delta y_p = kV_{py}\Delta t. \end{aligned} \quad (3.2.68)$$

Тогда можно записать в матричной форме

$$u(k\Delta t) = \sum_{\alpha} \sum_{\beta} g_{pi\alpha, j\beta} b_{nij}, \quad (3.2.69)$$

где $g_{pi\alpha, j\beta}$, b_{nij} — элементы матриц развертки изображения.

Здесь необходимо отметить два обстоятельства. Во-первых, полезно сравнить ширину импульсных характеристик объектива и развертывающего элемента. Если разрешение одного из рассмотренных элементов преобразования сигнала значительно выше, чем другого, то сигнал на выходе будет определяться в основном тем элементом, разрешающая способность которого ниже. Поэтому при моделировании можно учитывать только этот элемент. Например, при хорошей оптике и грубой развертке (например, мозаичный фотоприемник) объектив можно считать идеальным и не учитывать его фильтрующие свойства. В этом случае математическая модель становится более простой.

Во-вторых, шаги дискретизации Δx , Δy , Δx_p , Δy_p изображения и развертки в общем случае не равны друг другу (обычно $\Delta_p < \Delta$). Поэтому при моделировании телевизионного анализатора приходится применять интерполяцию функции двух переменных для нахождения промежуточных значений. Если изображение задано в дискретных точках

$$B_n(x_0 + i\Delta x, y_0 + k\Delta y) = f_{ik} \quad (i, k = 0, \pm 1, \dots),$$

а изменения аргумента равны

$$\alpha = \frac{x - x_0}{\Delta x}, \quad \beta = \frac{y - y_0}{\Delta y}, \quad (3.2.70)$$

то, например, формула билинейной интерполяции имеет вид [148]

$$f(x, y) = [f_{00}(1 - \alpha) + \alpha f_{10}](1 - \beta) + [f_{01}(1 - \alpha) + \alpha f_{11}]\beta. \quad (3.2.71)$$

Таким образом, в результате рассмотренных преобразований получаем модель телевизионного сигнала (видеосигнала) в дискретные моменты времени.

Моделирование электрического сигнала на выходе фотоприемного устройства. В рассмотренные анализаторы изображения входит как составная часть приемник излучения, а в некоторых случаях, например в телевизионных системах, сам фотоприемник (телевизионная трубка) является

анализатором изображения. Поэтому при моделировании необходимо учитывать характеристики фотоприемников. На выходе приемника излучения получается электрический сигнал, а в математической модели — одномерная функция времени, которая с учетом шумов фотоприемника является случайным процессом.

Х а р а к т е р и с т и к и п р и е м н и к о в и з л у ч е н и я. Свойства приемника излучения наиболее полно могут быть описаны системой характеристик, выражающих зависимость сигнала и шума, вырабатываемых приемником, от различных факторов: величины, спектрального состава и частоты модуляции излучения, падающего на приемник, температуры окружающей среды, полосы пропускания усилителя и т.д.

Наиболее распространенными характеристиками приемника излучения являются: амплитудная, спектральная, частотная, шумовая [81]. При проведении математического моделирования не всегда имеется полный набор характеристик, поэтому во многих случаях характеристики заменяются числовыми параметрами, выражающими свойства приемника для определенных условий. В соответствии с принципом функционального моделирования будем рассматривать только те характеристики и параметры фотоприемников, которые позволяют описать его как преобразователь потока излучения в электрический сигнал. С этой точки зрения приемник можно рассматривать как четырехполюсник, имеющий крутизну преобразования S , которая определяет величину сигнала, вырабатываемого приемником и приходящегося на единицу падающего на него потока излучения. Крутизна преобразования обычно называется *чувствительностью* и в зависимости от типа фотоприемника выражается в различных единицах (вольт на ватт, ом на ватт, ампер на люкс и т.д.). Рассмотрим подробнее важнейшие характеристики.

Амплитудная характеристика (энергетическая или световая) дает зависимость сигнала от величины потока излучения, падающего на приемник, т.е. $u = f(\Phi)$. Эта зависимость является нелинейной, поэтому описать амплитудную характеристику одним значением крутизны преобразования невозможно. Различают три значения крутизны (чувствительности).

Статическая крутизна или крутизна для немодулированного сигнала

$$S_{ст} = \frac{u}{\Phi} . \quad (3.2.72)$$

Дифференциальная крутизна или чувствительность для модулированного сигнала

$$S_d = \frac{du}{d\Phi} . \quad (3.2.73)$$

Средняя крутизна $S_{ср}$ определяется отношением амплитуды первой гармоники сигнала с фотоприемника к амплитуде синусоидально модулированного потока излучения

$$S_{ср} = \frac{Au_{\sim}}{A\Phi_{\sim}} . \quad (3.2.74)$$

Значение крутизны преобразования зависит от уровня фоновой засветки (немодулированного потока излучения). При низких уровнях постоянной составляющей, т.е. при работе на линейном участке амплитудной характеристики, все три типа крутизны преобразования совпадают. Крутизна преобразования измеряется при интегральном облучении от абсолютно черного тела или другого эталонного источника излучения, поэтому является интегральным параметром.

Для описания зависимости сигнала от длины волны падающего излучения вводится *спектральная крутизна*

$$S(\lambda) = \frac{du}{\Phi_{\lambda} d\lambda}, \quad (3.2.75)$$

где λ — длина волны; Φ_{λ} — спектральная плотность потока излучения.

Интегральная и спектральная крутизна связаны между собой соотношением

$$S = \frac{\int_0^{\infty} \Phi_{\lambda} S(\lambda) d\lambda}{\int_0^{\infty} \Phi_{\lambda} d\lambda}. \quad (3.2.76)$$

Формула (3.2.76) дает значение крутизны преобразования приемника при облучении его потоком сложного спектрального состава.

Частотная характеристика показывает зависимость дифференциальной крутизны от частоты модуляции потока излучения $S_d = \varphi(f_M)$. Эта характеристика отражает динамические свойства приемника — его способность реагировать на быстрые изменения потока излучения. В наиболее простом случае динамическим эквивалентом приемника является аperiodическое звено с частотной характеристикой

$$W(j\omega) = \frac{1}{1 + j\omega T}, \quad (3.2.77)$$

где T — постоянная времени фотоприемника.

Вообще говоря, процессы, происходящие в реальных фотоприемниках, более сложны, и дальнейшее развитие модели (3.2.77) при функциональном моделировании заключается в введении двух постоянных времени T_1 и T_2 для возрастания и убывания потока излучения.

Шумы приемников излучения. Одной из важнейших характеристик шумов фотоприемников является энергетический спектр шума. Однако в большинстве случаев математического моделирования эта характеристика отсутствует, имеются сведения лишь о среднеквадратической величине шума в определенной полосе частот. Поэтому о спектре шума судят зачастую на основании общих сведений о его природе, позволяющих указать распределение мощности шума по частотам для приемника того или иного типа с учетом условий эксплуатации.

Основными видами шумов приемников излучения являются: тепловой, дробовой, токовый, генерационно-рекомбинационный. Рассмотрим математическое описание перечисленных шумов [39, 65, 81, 119].

Тепловой шум обусловлен хаотическим движением носителей заряда и имеет равномерный спектр до частот порядка 10^{11} Гц. Спектральная плотность теплового шума описывается формулой

$$N_T(f) = 4 kTR \text{ В}^2/\text{Гц}, \quad (3.2.78)$$

где R — сопротивление, ом; T — абсолютная температура, К; $k = 1,38 \cdot 10^{-23}$ Дж·К $^{-1}$ — постоянная Больцмана.

ЭДС эквивалентного генератора шума в полосе частот Δf равна

$$e_T = \sqrt{N_T(f) \Delta f}. \quad (3.2.79)$$

Дробовой шум определяется тем, что электрический ток представляет собой поток дискретных частиц, заряд которых равен или кратен заряду электрона. Этот шум проявляется в основном в электровакуумных приборах. Спектральная плотность тока постоянна и определяется формулой

$$J_{др}(f) = 2 e J_0 \text{ а}^2/\text{Гц}, \quad (3.2.80)$$

где e — заряд электрона; J_0 — уровень постоянной составляющей.

Величина тока эквивалентного генератора шума равна

$$i_{др} = \sqrt{2 e J_0 \Delta f}. \quad (3.2.81)$$

Токовый шум зависит от протекающего через полупроводник тока (природа возникновения этого шума не совсем ясна). Спектральная плотность выражается формулой

$$N_{тк}(f) = \frac{\text{const}}{f^\beta}, \quad (3.2.82)$$

где параметр $\beta = 0,8 \div 1,5$ и определяется экспериментально для каждого типа фотоприемника. Токовый шум сказывается в основном на низких частотах. ЭДС эквивалентного генератора шума равна

$$e_{тк} = \left[\int_{f_1}^{f_2} N_{тк}(f) df \right]^{1/2}. \quad (3.2.83)$$

Генерационно-рекомбинационный шум связан с флуктуациями процессов генерации и рекомбинации носителей зарядов в полупроводниках. Он аналогичен дробовому шуму в вакуумных приборах. Спектральная плотность равномерна в широкой полосе частот и определяется опытным путем для каждого типа полупроводника.

Из рассмотренных шумов для полупроводниковых фотоприемников на низких частотах преобладает токовый шум, на средних — генерационно-рекомбинационный, на высших частотах — тепловой шум. Суммарная спектральная плотность приведена на рис. 3.13.

Обычно спектральная плотность шумов фотоприемника неизвестна или задана довольно приблизительно. Поэтому при моделировании поступают следующим образом. Для реального фотоприемника измеряется средне-квадратическое значение шумов на выходе усилителя с полосой Δf при подключенном на входе фотоприемнике. Полученное значение пересчитывается ко входу усилителя по формуле

$$u_{ш.вх} = \frac{u_{ш.вых}}{k}. \quad (3.2.84)$$

Таким образом учитываются также шумы первых каскадов усилителя. Величина $u_{ш}^2$ пропорциональна площади под кривой спектра в полосе пропускания усилителя (см. рис. 3.13). В математической модели ОЭС шумы фотоприемника вводятся аддитивно, т.е. прибавляются к сигналу, полученному после анализатора изображения с учетом крутизны преобразования S . Шумы задаются с датчика случайных чисел с нормальным законом распределения, равномерным спектром в полосе частот Δf (дискретный белый шум) и дисперсией $\sigma^2 = u_{ш.вх}^2$. Методика генерации на ЭВМ случайных процессов рассмотрена в § 3.5.

В ряде случаев приходится учитывать неравномерность чувствительности по полю фотоприемника и задавать ее как функцию координат $S(x, y)$. Обычно чувствительность является случайной функцией координат, т.е. случайным полем. Методы генерации случайных полей на ЭВМ также рассмотрены в § 3.5. Здесь отметим только, что неравномерность чувствительности по полю эквивалентна мультипликативному шуму, так как сигнал умножается на случайную крутизну преобразования.

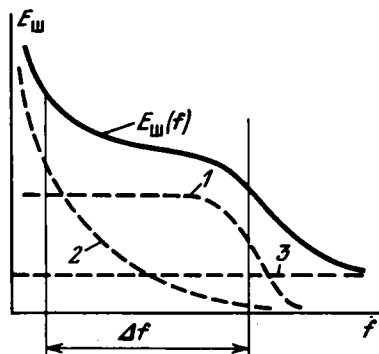


Рис. 3.13. Шумы приемников излучения: 1 - генерационно-рекомбинационный; 2 - токовый; 3 - тепловой; $E_{ш}$ - спектральная плотность шума

Статистические характеристики сигнала на выходе сканирующей оптической системы. Как отмечалось выше, теоретически вычислить сигнал на выходе анализатора изображения, как правило, не удается и приходится применять численные методы. Однако в ряде случаев можно определить аналитически статистические характеристики сигнала на выходе анализатора изображения и, таким образом, избежать прямого моделирования АИ. При необходимости можно затем произвести генерацию случайного процесса по известным статистическим характеристикам.

Найдем корреляционную функцию выходного сигнала сканирующей оптической системы с интегральным приемником. В качестве такого приемника может рассматриваться элементарная площадка в матричном анализаторе изображения или развертывающий элемент в телевизионном анализаторе. Изображение объекта считается случайным полем и задается своей двумерной корреляционной функцией. В работе [139] исследовано равномерное прямолинейное движение приемника относительно двумерного поля, а в [47] рассмотрен общий случай, когда траектория движения может быть криволинейной или даже случайной.

Рассмотрим фотоприемник, интегрально реагирующий на облученность, т.е. с точностью до постоянного множителя реакция приемника равна

$$u = \iint_S B_n(x, y) dx dy, \quad (3.2.85)$$

где $B_n(x, y)$ — распределение облученности в плоскости приемника; S — площадь поверхности фотоприемника.

Введем неподвижную систему координат, связанную с приемником, и подвижную, связанную с полем. Распределение облученности, создаваемой движущимся полем, в системе координат приемника описывается выражением

$$B_n(x, y, t) = B_n[x - \xi(t), y - \eta(t)], \quad (3.2.86)$$

где $\xi(t)$, $\eta(t)$ — функции, характеризующие движение поля относительно приемника. Тогда выходной сигнал также будет функцией времени

$$u = \iint_S B_n[x - \xi(t), y - \eta(t)] dx dy. \quad (3.2.87)$$

* Корреляционная функция интеграла связана с корреляционной функцией поля $B_n(x, y, t)$ соотношением [68]

$$K_n(t_1, t_2) = \iiint_S K_1(t_1, x_1, y_1, t_2, x_2, y_2) dx_1 dy_1 dx_2 dy_2, \quad (3.2.88)$$

где K_1 — корреляционная функция движущегося поля, записанная в неподвижной системе координат, связанной с приемником.

Таким образом, для вычисления корреляционной функции сигнала необходимо найти корреляционную функцию движущегося поля по заданной корреляционной функции $K(x, y)$ неподвижного поля.

Пусть параметры $\xi(t)$ и $\eta(t)$ являются случайными стационарными процессами, статистически не зависящими от случайного поля. Тогда выражение (3.2.86) описывает поле, прошедшее через линейную систему, осуществляющую случайную задержку по пространству. Частотная характеристика такой системы описывается выражением [68]

$$W(u, v, t) = \exp \{-i[u\xi(t) + v\eta(t)]\}, \quad (3.2.89)$$

где u и v — пространственные частоты.

Определим корреляционную функцию стационарной линейной системы из соотношения

$$K_W(u, v, \tau) = M\{W(u, v, t) W(-u, -v, t + \tau)\}. \quad (3.2.90)$$

Пусть процессы $\xi(t)$ и $\eta(t)$ статистически независимы. Тогда с учетом (3.2.89) получим

$$\begin{aligned} K_W(u, v, \tau) &= M\{\exp[iu(\xi(t + \tau) - \xi(t))]\} \times \\ &\times M\{\exp[iv(\eta(t + \tau) - \eta(t))]\} = \\ &= \theta_{2\xi}(-u, u, \tau) \theta_{2\eta}(-v, v, \tau). \end{aligned} \quad (3.2.91)$$

Здесь $\theta_{2\xi}$ и $\theta_{2\eta}$ — двумерные характеристические функции процессов $\xi(t)$ и $\eta(t)$. Для нормальных процессов корреляционная функция (3.2.91) случайной линии задержки имеет вид

$$K_W(u, v, \tau) = \exp\{-\sigma_\xi^2[1 - R_\xi(\tau)]u^2 - \sigma_\eta^2[1 - R_\eta(\tau)]v^2\}, \quad (3.2.92)$$

где $R_\xi(\tau)$, $R_\eta(\tau)$, σ_ξ^2 , σ_η^2 — нормированные корреляционные функции и дисперсии процессов ξ и η .

Используя выражения, полученные для случайных процессов [68], найдем корреляционную функцию поля после случайной задержки

$$\begin{aligned} K_1(x_1, x_2, y_1, y_2, \tau) &= \\ &= \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K_W(u, v, \tau) S(u, v) \exp\{iu(x_2 - x_1) + iv(y_2 - y_1)\} du dv, \end{aligned} \quad (3.2.93)$$

где $S(u, v)$ — пространственный спектр неподвижного поля. Используя (3.2.88) и (3.2.93), получим окончательное выражение для корреляционной функции сигнала фотоприемника

$$\begin{aligned} K_u(\tau) &= \frac{1}{4\pi^2} \underbrace{\int \int \int \int}_S \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(u, v) \times \\ &\times \exp\{-\sigma_\xi^2[1 - R_\xi(\tau)]u^2 - \sigma_\eta^2[1 - R_\eta(\tau)]v^2 + \\ &+ iu(x_2 - x_1) + iv(y_2 - y_1)\} du dv dx_1 dy_1 dx_2 dy_2. \end{aligned} \quad (3.2.94)$$

В качестве примера рассмотрим случайное движение квадратного приемника со стороны 2а относительно поля с корреляционной функцией

$$K(x, y) = \exp\{-\alpha^2(x^2 + y^2)\}.$$

Статистические характеристики движения приемника по осям x и y считаем одинаковыми и равными $R(\tau)$ и σ^2 .

Пространственный спектр поля описывается выражением

$$S(u, v) = \frac{\pi}{\alpha^2} \exp\left\{-\frac{u^2 + v^2}{4\alpha^2}\right\}.$$

Используя формулу (3.2.94), получаем

$$\begin{aligned}
 K_u(\tau) &= \frac{1}{4\pi\alpha^2} \underbrace{\int \int \int \int}_S \int_{-\infty}^{\infty} \exp \left\{ -\frac{u^2}{4\alpha^2} - \sigma^2[1 - R(\tau)]u^2 + \right. \\
 &+ iu(x_2 - x_1) \left. \right\} \exp \left\{ -\frac{v^2}{4\alpha^2} - \sigma^2[1 - R(\tau)]v^2 + \right. \\
 &+ iv(y_2 - y_1) \left. \right\} du dv dx_1 dy_1 dx_2 dy_2 = \\
 &= \frac{1}{4\pi\alpha^2} \left(\int_{-a}^a \int_{-\infty}^{\infty} \exp \left\{ -\frac{u^2}{4\alpha^2} - \sigma^2[1 - R(\tau)]u^2 + \right. \right. \\
 &+ iu(x_2 - x_1) \left. \right\} du dx_1 dx_2 \left. \right)^2.
 \end{aligned}$$

Здесь учтено, что характеристики и поля и движения в ортогональных направлениях одинаковы.

Интегрирование по u дает

$$K_u(\tau) = A \left(\int_{-a}^a dx_1 \int_{-a}^a \exp \{ -B(x_2 - x_1)^2 \} dx_2 \right)^2,$$

где

$$A = \frac{1}{1 + 4\alpha^2\sigma^2[1 - R(\tau)]}, \quad B = \frac{\alpha^2}{1 + 4\alpha^2\sigma^2[1 - R(\tau)]}.$$

После интегрирования по координатам получаем

$$K_u(\tau) = \frac{A\pi}{B^2} \left[2a\sqrt{B} \Phi(2a\sqrt{B}) + \frac{1}{\sqrt{\pi}} (e^{-4a^2B} - 1) \right]^2.$$

Для реальных полей и малых размеров фотоприемника обычно выполняется условие

$$2a\sqrt{B} \ll 1.$$

Используя разложение интеграла вероятностей и экспоненты в ряд в окрестности нуля, получим

$$K_u(\tau) = \frac{16a^4}{1 + 4\alpha^2\sigma^2[1 - R(\tau)]}.$$

Конкретный вид корреляционной функции сигнала с выхода фотоприемника определяется нормированной корреляционной функцией $R(\tau)$, характеризующей движение приемника.

Полученная корреляционная функция может быть использована непосредственно для анализа работы оптико-корреляционной системы или для синтеза формирующего фильтра с целью генерации случайного процесса методом, изложенным в § 3.5, и преобразования его в математической модели электронно-усилительного тракта.

§ 3.3. Моделирование цифровых изображений

Дискретизация и квантование. Рассмотрим подробнее некоторые вопросы, связанные с моделированием протяженных источников излучения или полей яркости. Поля яркости для математического моделирования могут быть получены двумя путями: фотометрированием аэроснимков реальных полей яркости и генерацией на ЭВМ двумерных полей с заданными статистическими характеристиками. Методика аэрофотосъемки и фотометрирования снимков выходит за рамки данной работы, она изложена, например, в работах [35, 115, 138, 149]. Генерирование двумерных полей на ЭВМ дано в § 3.5. Ниже рассматриваются особенности представления изображений в цифровой форме, т.е. цифровых изображений, для моделирования на ЭВМ.

Дискретизация непрерывных изображений. При формировании цифровых изображений возникает задача представления двумерных полей в дискретной форме, т.е. преобразование двумерной непрерывной функции распределения яркости $B(x, y)$ в функцию дискретного аргумента $B_{kl} = B(k\Delta x, l\Delta y)$, где $\Delta x, \Delta y$ — шаги дискретизации поля в направлении осей x и y . Если размер поля $M \times N$, то количество дискретных отчетов равно $m \times n$, причем $m = M/\Delta x, n = N/\Delta y$. Тогда $k = 0, 1, 2, \dots, m-1; l = 0, 1, 2, \dots, n-1$.

Для хранения в памяти ЭВМ сигнал подвергается квантованию по уровню, в результате его амплитуда принимает конечное множество значений. Полученный цифровой массив можно представить в виде матрицы B размера $m \times n$, которая получается из исходного непрерывного поля последовательным применением операторов дискретизации и квантования

$$B = P_{\text{кв}} P_{\text{д}} B(x, y), \quad (3.3.1)$$

из которых первый $P_{\text{д}}$ является линейным, а второй $P_{\text{кв}}$ — нелинейным оператором. В литературе матрица B , содержащая квантованные дискретные отчеты непрерывного двумерного поля, называется *цифровым изображением* [148].

При создании цифровых изображений встает вопрос о выборе шага дискретизации, количестве уровней квантования и размерах цифрового изображения. Выбор шага дискретизации обычно основывается на теореме отсчетов Котельникова, обобщенной на двумерные сигналы [50]. Если спектр функции $B(x, y)$ ограничен в спектральном базисе величинами

$$|\omega_x| \leq \Omega_x = 2\pi F_x, \quad |\omega_y| \leq \Omega_y = 2\pi F_y, \quad (3.3.2)$$

то ее можно воспроизвести по дискретным значениям, взятым через интервалы

$$\Delta x = \frac{\pi}{\Omega_x} = \frac{1}{2F_x}, \quad \Delta y = \frac{\pi}{\Omega_y} = \frac{1}{2F_y}. \quad (3.3.3)$$

Тогда двумерная функция представляется в виде ряда

$$B(x, y) = \frac{1}{4F_x F_y} \sum_{k,l=-\infty}^{k,l=\infty} B(k\Delta x, l\Delta y) g_0(x - k\Delta x, y - l\Delta y), \quad (3.3.4)$$

где k, l — целые числа, а функция $g_0(x, y)$ равна

$$g_0(x, y) = \frac{\sin x \Omega_x}{\pi x} \frac{\sin y \Omega_y}{\pi y}. \quad (3.3.5)$$

Формула (3.3.4) описывает двумерную функцию, восстановленную из дискретных отсчетов фильтром с прямоугольной частотной характеристикой.

Двумерный спектр дискретной функции является периодическим. Количество отсчетов изображения на единицу площади при такой дискретизации равно $4F_x F_y$, т.е. площади прямоугольника, ограничивающего пространственный спектр изображения. Чтобы это число было минимальным, необходимо уменьшать период повторения спектра $2F_x, 2F_y$ до тех пор, пока соседние периоды не начнут соприкасаться, так как более плотной "упаковке" составляющих спектра при его периодическом продолжении на плоскости пространственных частот соответствует меньшее число отсчетов изображения.

Теорема отсчетов справедлива для сигналов, спектр которых точно равен нулю вне некоторого интервала частот. Пространственный спектр реальных изображений не ограничен, поэтому при дискретизации периодические компоненты спектра начинают перекрываться, что ведет к ошибке восстановления сигнала по его дискретным отсчетам. Неточность восстановления сигнала называют *строб-эффектом*, так как в результате наложения спектров компоненты исходного сигнала с частотой $F + \Delta$ появляются в восстановительном сигнале с более низкой частотой $F - \Delta$. Поэтому дискретизации реальных изображений должна предшествовать двумерная фильтрация, ограничивающая их пространственный спектр. При фотометрировании аэроснимков это достигается усреднением сигнала апертурой сканирующего луча. При генерации случайных полей на ЭВМ ограничение спектра производится частотной характеристикой формирующего фильтра, которую для общности также будем называть апертурой. В обоих случаях форма границы пространственного спектра определяется формой апертуры фильтра.

Рассмотренной дискретизации в прямоугольной системе координат соответствует наиболее плотная упаковка спектров для прямоугольной апертуры фильтра. Однако чаще апертура фильтра имеет круговую симметрию (апертура электронного луча, апертура объектива), при этом оптимальным будет расположение отсчетов в узлах правильной шестиугольной решетки (гексагональный растр), что дает примерно 15 % экономии количества отсчетов [148]. Практически при обработке изображений в основном используется простая дискретизация по прямоугольному растру как наиболее универсальная и обеспечивающая более простые (следовательно, более быстродействующие) программы при математическом моделировании ОЭС.

В цифровых моделях оптико-электронных систем обычно выбирают одинаковый шаг дискретизации $\Delta x = \Delta y = \Delta$ во взаимно перпендикулярных направлениях. При этом исходят из величины радиуса корреляции изображения

$$r_k = \int_0^{\infty} R(x) dx, \quad (3.3.6)$$

где $R(x)$ — нормированная корреляционная функция, и выбирают шаг дискретизации из соотношения [136]

$$\Delta = (0,25 \div 0,33)r_k. \quad (3.3.7)$$

Квантование дискретных отсчетов. Вторым этапом получения цифрового сигнала из непрерывного после дискретизации является квантование сигнала по уровню. Весь диапазон изменения яркости $B_{\max} - B_{\min}$ разбивается на конечное число уровней (интервалов квантования), внутри каждого из которых величине сигнала присваивается постоянное значение. Квантование амплитуды уменьшает размерность пространства сигналов, что значительно облегчает задачу хранения больших массивов цифровых изображений.

При квантовании возникает ошибка квантования

$$\Delta B = B - B_{kv}, \quad (3.3.8)$$

где B — значение яркости неквантованного изображения; B_{kv} — значение яркости квантованного изображения.

Ошибка квантования является периодической функцией яркости и зависит от количества уровней квантования и закона распределения сигнала. Требования к точности квантования формулируются путем ограничений, накладываемых на ΔB . Наиболее общий подход [148] состоит в том, что величины ΔB и B считают случайными и вводят некоторую меру $D(\Delta B)$ отличия сигнала B от его квантованного значения B_{kv} . При таком подходе точность квантования Q характеризуют средним по распределению вероятностей $f(B_{kv})$ значением расстояния $D(\Delta B')$:

$$Q = \sum_{r=1}^{M-1} \int_{B_{kv}^r}^{B_{kv}^{r+1}} f(B_{kv}) D(\Delta B') dB_{kv}, \quad (3.3.9)$$

где B_{kv}^r, B_{kv}^{r+1} — границы r -го интервала; M — число уровней квантования.

Выбор меры зависит от конкретной задачи, но в целом этот подход типичен для техники передачи изображений. Эксперименты с использованием телевизионных устройств воспроизведения изображений показывают, что достаточно от 64 до 256 уровней квантования. При математическом моделировании ОЭС не требуется восстанавливать исходное непрерывное изображение, так как полученное цифровое изображение является тем объектом, по которому "работает" математическая модель системы. В этом случае величину интервала квантования выбирают, исходя из величины пороговой чувствительности и логики обработки сигнала в электронном тракте. Количество уровней квантования изображения для разных ОЭС может изменяться в пределах от 2 до 1000, т.е. от двухуровневого квантования до весьма точного преобразования непрерывного поля в цифровое изображение. Наиболее часто используется от 64 до 128 уровней квантования.

Расчет количества уровней квантования по корреляционному критерию. Одним из путей определения необходимого количества уровней квантования цифрового изображения при математическом моделировании ОЭС является использование корреляционного критерия, при котором сравнивается вероятность правильного обнаружения сигнала для квантованного

и неквантованного сигналов. Количество уровней квантования выбирается таким, чтобы обеспечить одинаковую вероятность $P_{\text{п.о}}$ в обоих случаях.

Рассмотрим обнаружение сигнала на фоне аддитивной нормальной помехи, не коррелированной с сигналом. Производится n некоррелированных отсчетов и используется критерий обнаружения Неймана – Пирсона, обеспечивающий максимум вероятности правильного обнаружения $P_{\text{п.о}}$ при заданной вероятности ложных тревог $P_{\text{л.т}} = \alpha$.

Аналоговый сигнал. Для аналогового коррелятора задача обнаружения сигнала формулируется как проверка гипотезы $r = 0$ против альтернативы $r > 0$, где r – коэффициент корреляции сигналов x и y . Оптимальное несмещенное правило обнаружения заключается в сравнении с величиной порога оценки коэффициента корреляции [69]

$$\hat{r} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\left[\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2 \right]^{1/2}}. \quad (3.3.10)$$

Статистика $\hat{r}$ асимптотически нормальна со средним r и дисперсией $(1 - r^2)^2/n$.

Величина порога C в алгоритме обнаружения

$$\hat{r} \geq C \quad (3.3.11)$$

при $n \gg 1$ и фиксированной вероятности ложных тревог равна

$$C = x_\alpha / \sqrt{n}, \quad (3.3.12)$$

где x_α – процентная точка нормального распределения; n – длина выборки. Вероятность пропуска сигнала находится из соотношения процентных точек

$$x_{1-\beta} = x_\alpha - r\sqrt{n}. \quad (3.3.13)$$

Расчеты порога и вероятности правильного обнаружения при вероятности ложной тревоги $P_{\text{л.т}} = 0,01$ приведены в таб. 3.2.

Удовлетворительные характеристики обнаружения для аналогового сигнала получаются при длине выборки в 64 отсчета.

Квантованный сигнал. Рассмотрим обнаружение сигнала при минимальном количестве уровней квантования, т.е. при квантовании сигнала на два уровня: 0 и 1. Алгоритм обнаружения в этом случае записывается в виде

$$\sum_{i=1}^n \text{sign} x_i \cdot \text{sign} y_i \geq C. \quad (3.3.14)$$

При заданной вероятности $P_{\text{л.т}} = \alpha$ порог C определяется выражением [69]

$$C = (x_\alpha \sqrt{n} + n)/2. \quad (3.3.15)$$

Таблица 3.2

n	32	64	128	256	512
C	0,41	0,29	0,205	0,144	0,102
$P_{\text{п.о}}$	0,7	0,95	1,0	1,0	1,0

Таблица 3.3

n	32	64	128	256	512
$\tilde{C}$	0,705	0,645	0,602	0,573	0,551

В табл. 3.3 приведена величина нормированного порога $\tilde{C} = C/n$ для $\alpha = 0,01$.

Сравнение табл. 3.2 и 3.3 показывает, что при заданной вероятности $P_{л.т}$ для двухуровневого квантования сигнала приходится повышать величину порога по сравнению с аналоговым, что снижает вероятность правильного обнаружения.

Алгоритмы обнаружения можно сравнивать по коэффициенту асимптотической относительной эффективности [69]

$$\rho = n_1/n_2, \quad (3.3.16)$$

где n_1 — длина выборки аналогового коррелятора; n_2 — длина выборки при двухуровневом квантовании для обеспечения одинаковой $P_{л.о}$ при заданной $P_{л.т}$. При нормальной помехе

$$\rho = \frac{4}{\pi^2} = 0,405.$$

Таким образом, эффективность обнаружения при двухуровневом квантовании составляет 40 % по сравнению с аналоговым сигналом.

Увеличение количества уровней квантования ведет к повышению коэффициента эффективности. Кроме того, ρ зависит от характеристики квантователя (линейный и нелинейный) и закона распределения сигнала. Для оптимального в смысле максимума информации квантователя [69] и релевского распределения амплитуды сигнала коэффициент эффективности равен

$$\rho = 1 - \frac{2,25}{m^2}, \quad (3.3.17)$$

где m — число уровней квантования. Расчеты приведены в табл. 3.4, где дана величина $1/\rho = n_2/n_1$, показывающая, во сколько раз нужно увеличить

Таблица 3.4

m	2	4	8
$1/\rho$	2,29	1,16	1,04

длину выборки квантованного сигнала по сравнению с неквантованным для получения одинаковой вероятности правильного обнаружения.

Таким образом, уже при квантовании на восемь уровней, т.е. при представлении числа тремя двоичными разрядами, эффективность квантованного и неквантованного алгоритмов становится практически одинаковой.

Математическое моделирование с квантованием сигнала по равной вероятности попадания в разряды квантования (вероятностный квантователь) и с равномерным квантованием дает близкие результаты по эффективности, что указывает на слабое влияние характеристики квантователя. Поэтому минимальное количество уровней квантования цифрового изображения для математического моделирования можно принять равным восьми. В зависимости от конкретного построения ОЭС количество уровней квантования может быть увеличено.

Кодирование цифровых изображений в пространственной области. С величиной шага дискретизации Δ и количеством уровней квантования тесно связан вопрос о необходимой памяти ЭВМ для размещения цифрового изображения. При математическом моделировании размеры изображения должны превышать поле зрения оптико-электронной системы, типовая величина которого равна $2 - 5^\circ$. Тогда необходимый размер изображения в пространстве предметов на дальности, например, $L = 10$ км составит порядка 1 км. Для типовых радиусов корреляции оптических полей $r_k = 5 \div 100$ м минимальный шаг дискретизации по формуле (3.3.7) получается около 1 м. Следовательно, цифровое изображение такого поля будет содержать 10^6 отсчетов.

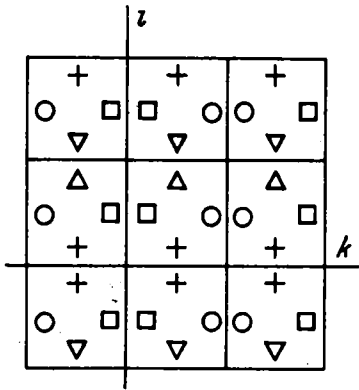
Размещение такого цифрового массива требует большого объема памяти ЭВМ. Для сокращения требуемого объема памяти цифровой массив подвергается различной обработке, которую будем называть *кодированием*. При этом стремятся сократить объем информации для описания изображения при сохранении определенного уровня верности кодированного изображения по отношению к исходному. Конкретный критерий величины различия исходного и кодированного изображений определяется методом кодирования и существом рассматриваемой задачи. Дискретизация и квантование непрерывного сигнала при переходе к цифровому изображению являются простейшими методами кодирования, позволяющими уменьшить объем информации при хранении изображения в ЭВМ.

Все методы кодирования можно разделить на кодирование в исходной (пространственной) области представления изображения и кодирование в преобразованной области (кодирование изображения посредством преобразований). Рассмотрим кодирование цифровых изображений в пространственной области.

Специальные случаи кодирования. Одним из простых методов кодирования является упаковка нескольких чисел в одну ячейку памяти ЭВМ, т.е. представление чисел в форме с плавающей запятой с меньшим количеством разрядов в мантиссе. Тогда при использовании, например, ЭВМ БЭСМ-6 удастся упаковать в одну ячейку три числа при количестве уровней квантования 1000 и порядка 40 чисел при двухуровневом кодировании. Использование этого метода позволяет значительно уменьшить объем необходимой памяти, но все-таки в большинстве случаев не дает возможности расположить все цифровое изображение в оперативной памяти ЭВМ, что необходимо для быстрой работы программы.

Экономия оперативной памяти можно получить, используя цифровые поля меньшей размерности, а при моделировании по мере необходимости периодически продолжать поле путем его переворота, т.е. производить так называемое *четное продолжение* [148] в соответствии с рис. 3.14. Эле-

Рис. 3.14. Четное продолжение изображения



менты $\tilde{b}_{k,l}$ цифрового поля связаны с элементами $b_{k,l}$ исходного поля, примыкающего справа к началу координат, следующими отношениями:

$$\tilde{b}_{k,l} = \begin{cases} b_{2m-k-2, |l|}, & m \leq k \leq 2m-2, \quad -(n-1) \leq l \leq -1; \\ b_{|k|, l}, & -(m-1) \leq k \leq -1, \quad -(n-1) \leq l \leq -1; \\ b_{k, |l|}, & 0 \leq k \leq m-1, \quad -(n-1) \leq l \leq -1; \\ b_{|k|, l}, & -(m-1) \leq k \leq -1, \quad 0 \leq l \leq n-1; \\ b_{k, l}, & 0 \leq k \leq m-1, \quad 0 \leq l \leq n-1; \\ b_{2m-k-2, l}, & m \leq k \leq 2m-2, \quad 0 \leq l \leq n-1; \\ b_{k, 2m-l-2}, & 0 \leq k \leq m-1, \quad n \leq l \leq 2n-2; \\ b_{|k|, 2n-l-2}, & -(m-1) \leq k \leq -1, \quad n \leq l \leq 2n-2; \\ b_{2m-k-2, 2n-l-2}, & m \leq k \leq 2m-2, \quad n \leq l \leq 2n-2. \end{cases} \quad (3.3.18)$$

Для сокращения количества отсчетов цифрового изображения можно использовать дискретизацию с переменным шагом по полю: центр поля дискретизируется с мелким шагом, а по мере приближения к краю шаг дискретизации увеличивается. Этот метод удобен для моделирования работы ОЭС на различных дальностях до объекта, так как элемент разрешения и поле зрения на местности прямо пропорциональны дальности. Цифровое изображение с неравномерной дискретизацией получается обработкой на ЭВМ по специальной программе цифрового изображения с равномерной дискретизацией, в этом случае удобно увеличивать вдвое шаг дискретизации при увеличении линейных размеров поля также в два раза.

Наряду с дискретизацией неравномерным может быть и квантование дискретных отсчетов. Часто используются логарифмические и вероятностные квантователи. Практически процедура создания цифрового поля сводится к предварительному нелинейному преобразованию сигнала и затем линейному квантованию. В логарифмическом квантователе сигнал преобразуется по формуле

$$B' = \ln B, \quad (3.3.19)$$

причем для исключения отрицательных значений логарифма минимальное

значение яркости B устанавливают равным единице. В вероятностном квантователе сигнал преобразуется по формуле нормального распределения

$$B' = \frac{1}{\sigma_B \sqrt{2\pi}} \int_{-\infty}^B e^{-\frac{(B - m_B)^2}{2\sigma_B^2}} dB, \quad (3.3.20)$$

где σ_B^2 , m_B — дисперсия и среднее значение исходного поля. Практически в программе для ЭВМ функция (3.3.20) задается таблично.

При использовании таких полей в математической модели ОЭС сигнал подвергается обратному нелинейному преобразованию. Этот метод дает экономию памяти ЭВМ, так как сжатие диапазона значений сигнала путем нелинейного преобразования позволяет использовать меньшее количество уровней квантования. Кроме того, здесь уменьшаются ошибки, свойственные линейному квантователю, за счет ограничения сигнала, имеющего редкие большие выбросы.

Дифференциальная импульсно-кодовая модуляция. Уменьшение количества квантования может быть достигнуто применением разностных методов, широко используемых в технике передачи изображений. Наиболее распространенным является метод дифференциальной импульсно-кодовой модуляции (ДИКМ). В основе разностных методов кодирования сигнала лежит формирование разностей между текущим отсчетом и предсказанной величиной, которая определяется по ряду предыдущих отсчетов специальным устройством — *предсказателем* [134]. Этот разностный сигнал имеет более узкую полосу и поэтому может быть проквантован на меньшее количество уровней по амплитуде, чем и достигается экономия памяти.

Частным случаем ДИКМ является дельта-модуляция, при которой разностный сигнал квантуется только на два уровня. Как правило, при квантовании разностного сигнала используют неравномерную шкалу квантования.

Для двумерных изображений и разделимой по строкам и столбцам ДИКМ разностное изображение определяется выражением [87]

$$\begin{aligned} \Delta_{k,l} = & b_{k,l} - \sum_{i=1}^n p_i b_{k-i,l} - \sum_{j=1}^n q_j b_{k,l-j} + \\ & + \sum_{i=1}^n \sum_{j=1}^n p_i q_j b_{k-i,l-j}, \end{aligned} \quad (3.3.21)$$

где $b_{k,l}$ — элементы исходного изображения. Здесь используются $2n$ коэффициентов предсказателя p_i, q_j .

Если изображение имеет разделимые статистические характеристики, то предсказатель описывается двумя треугольными матрицами

$$Q = \begin{pmatrix} 1 & & & & \\ -q_1 & 1 & & & \\ \dots & \dots & \dots & \dots & \\ -q_n & \dots & q_1 & 1 & \end{pmatrix}, \quad P^T = \begin{pmatrix} 1 - p_1 & \dots & -p_n & & \\ \dots & \dots & \dots & \dots & \\ & & & 1 - p_1 & \\ & & & & 1 \end{pmatrix} \quad (3.3.22)$$

для предсказания по столбцам и строкам. В этом случае матрица разност-

ного изображения записывается в виде

$$\underline{\Delta} = \underline{Q} \underline{B} \underline{P}^T, \quad (3.3.23)$$

где $\underline{B}$ — матрица исходного изображения.

Если производить кодирование только построчно, то разностный сигнал определяется выражением

$$\Delta_k = b_k - \sum_{i=1}^n p_i b_{k-i} \quad (3.3.24)$$

и зависит от n коэффициентов.

Восстановление исходного сигнала производится по формуле

$$b_k = \Delta_k + \sum_{i=1}^n p_i b_{k-i}. \quad (3.3.25)$$

Коэффициенты предсказателя p_i , q_j выбираются по различным критериям в зависимости от статистических характеристик изображения. Оптимальный выбор этих коэффициентов и их количества (порядок предсказателя) является самостоятельной сложной задачей. В [100] показано, что повышение порядка предсказателя выше третьего не дает дальнейшего ощутимого снижения ошибки кодирования. Для двумерного поля с экспоненциальной корреляционной функцией вычислены коэффициенты по минимуму среднеквадратической ошибки предсказания, равные (для двумерного предсказателя первого порядка)

$$p_1 = 0,628; \quad q_1 = 0,443.$$

При этом обеспечивается среднеквадратическая ошибка предсказания 4,3%.

Практические работы с ДИКМ показывают, что разностный сигнал нужно представлять числом, содержащим три-четыре двоичных разряда. Поэтому ДИКМ может дать выигрыш в памяти ЭВМ только при кодировании цифровых изображений с большим количеством уровней квантования (восемь-десять двоичных разрядов).

Кодирование изображений путем преобразований. Разложение в ортогональном базисе. При использовании этого метода изображение переводится в некоторую обратимую форму, более удобную для дальнейшего использования. Обычно применяется разложение по системе дискретных ортогональных функций.

Матрицы исходного изображения $\underline{B}$ размера $N \times N$ и преобразованного $\underline{\alpha}$ задаются соотношениями [87]

$$\underline{B} = \underline{U} \underline{\alpha} \underline{V}^T, \quad \underline{\alpha} = \underline{U}^T \underline{B} \underline{V}. \quad (3.3.26)$$

Здесь

$$\underline{U} = \underline{u}_1, \underline{u}_2, \dots, \underline{u}_N; \quad \underline{V} = \underline{v}_1, \underline{v}_2, \dots, \underline{v}_N$$

суть унитарные операторы, а $\underline{u}_i$, $\underline{v}_j$ — векторы, образованные из столбцов $\underline{U}$ и $\underline{V}$.

Матрицу изображения $\underline{B}$ можно представить в форме

$$\underline{B} = \sum_{i=1}^n \sum_{j=1}^n \alpha_{ij} \underline{u}_i \underline{v}_j^T. \quad (3.3.27)$$

Произведения $u_i v_j^T$ являются двумерными ортогональными функциями. Таким образом, выражение (3.3.27) задает разложение изображения по системе дискретных двумерных ортогональных базисных функций.

Выбор преобразований, выражаемых матрицами U и V , произволен. Эти матрицы могут выбираться как из одинаковых, так и из различных базисных функций. Всего в разложении будет N^2 членов. Возможны различные сочетания ортогональных систем, при которых $U \neq V$.

Энергия изображения сосредоточена в основном в нескольких наиболее значимых коэффициентах разложения. Следовательно, для сохранения определенной верности закодированного изображения по отношению к исходному нужно запомнить лишь несколько коэффициентов из общего количества N^2 . Таким образом можно достичь значительной экономии памяти для хранения изображения. Полученное закодированное изображение с использованием усеченного разложения представляется в форме

$$B_k = \sum_i \sum_{j \in \{k\}} \alpha_{ij} u_i v_j^T, \quad (3.3.28)$$

где $\{k\}$ — некоторое подмножество коэффициентов из общего числа N^2 . Получающаяся при этом ошибка определяется нормой разности

$$\|B - B_k\| = \sum_{i,j \in \{N^2 - k\}} \alpha_{ij} u_i v_j^T. \quad (3.3.29)$$

Как показывают работы по цифровому кодированию изображений [87], использование ортогональных преобразований позволяет уменьшить требуемую емкость памяти приблизительно на порядок.

Обзор ортогональных преобразований. Для кодирования изображений могут быть использованы различные ортогональные преобразования [87]. Наиболее распространенными являются преобразования Фурье, Уолша, Хаара, косинусное, синусное, наклонное, Карунена—Лозва, сингулярное. Наиболее оптимальным преобразованием в смысле уменьшения нормы разности является сингулярное разложение по собственным векторам матрицы B . При этом достигается наибольшая декорреляция полученных коэффициентов, так как матрица α получается диагональной (все α_{ij} , для которых $i \neq j$, равны нулю). Но это преобразование требует наибольшего количества вычислений и не имеет быстрого алгоритма.

Преобразование Карунена—Лозва по собственным векторам ковариационной матрицы изображения использует статистические характеристики двумерного поля и минимизирует средний квадрат нормы разности (3.3.29). Это преобразование оптимально в статистическом смысле, но не имеет быстрого алгоритма.

Ближкие результаты дают дискретное преобразование Фурье, косинусное, синусное, наклонное. Кроме того, все они имеют быстрые преобразования и, следовательно, не требуют большого количества вычислений.

Наиболее удобны для кодирования изображений при математическом моделировании преобразования Уолша и Хаара, имеющие быстрые алгоритмы. Наименьшее количество вычислительных операций, равное $2(N - 1)$, требует преобразование Хаара.

В большинстве случаев кодирование изображения производится только по столбцам или строкам, т.е. используются только одномерные базисные

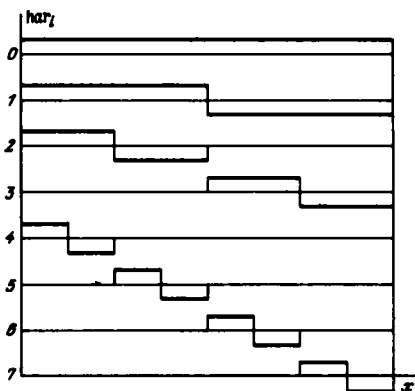


Рис. 3.15

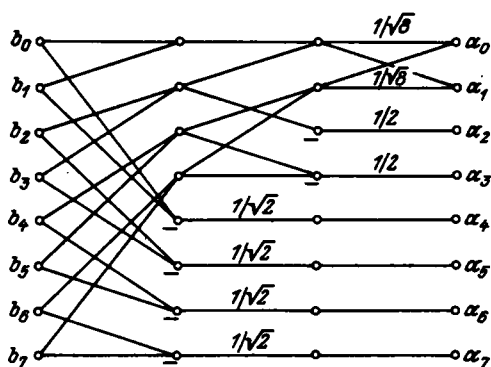


Рис. 3.16

```

SUBROUTINE FHART (X,Y N,N1)
REAL X(N),Y(N1)
W=2.0**0.5
L=N1
GO TO 3
1 L=L/2
DO 2 I=L1,N
2 X(I)=X(I)*W
3 J=1
DO 4 I=1,L
A1=X(J)
A2=X(J+1)
X(I)=A1+A2
Y(I)=A1-A2
4 J=J+2
DO 5 I=1,L
5 X(I+L)=Y(I)
IF(L.EQ.1) GO TO 6
L1=L+1
GO TO 1
6 CONTINUE
RETURN
END

```

Рис. 3.17

Рис. 3.15. Функции Хаара

Рис. 3.16. Граф быстрого преобразования Хаара

Рис. 3.17. Программа быстрого преобразования Хаара

функции U или V . Тогда каждая i -я строка изображения может быть представлена в виде

$$B_i = \sum_{j=1}^N \alpha_{ij} v_j, \quad (3.3.30)$$

т.е. определяется N коэффициентами разложения α_{ij} по базисным функциям v_j , заданным N значениями. Коэффициенты разложения для каждой строки определяются по формуле

$$\alpha_{ij} = \sum_{j=1}^N B_i v_j. \quad (3.3.31)$$

Использование усеченного разложения для строки

$$B_i = \sum_{j=1}^k \alpha_{ij} v_j, \quad (3.3.32)$$

где $k < N$, позволяет достичь значительной экономии памяти ЭВМ.

На рис. 3.15 приведены первые восемь функций Хаара, а на рис. 3.16 и 3.17 — граф и программа быстрого преобразования Хаара на языке ФОРТРАН.

Представление изображения коэффициентами разложения в ортогональном базисе одновременно с экономией памяти позволяет провести фильтрацию, так как исключение определенных коэффициентов разложения означает исключение из сигнала соответствующих частот.

Возникает вопрос о необходимом количестве членов в разложении (3.3.32). Эта величина зависит от структуры изображения и типа моделируемой ОЭС. В большинстве случаев можно получить коэффициент сжатия информации, равный $N/k = 5 \div 10$.

Дальнейшее уменьшение необходимой памяти может быть осуществлено применением гибридного метода кодирования, когда полученное усеченное разложение по строкам кодируется по столбцам с использованием ДИКМ. При этом может быть достигнут коэффициент сжатия, равный 25–50, но за счет большой сложности программ кодирования и обратного декодирования.

§ 3.4. Математическое моделирование обработки сигнала в электронном тракте ОЭС

Функциональные звенья электронного тракта ОЭС. Характеристики линейных звеньев. При построении элементарных математических моделей электронный тракт оптико-электронных систем представляется в виде совокупности отдельных функциональных звеньев (каждое со своим входом и выходом), не оказывающих влияния друг на друга. Различают линейные и нелинейные звенья. Звено называется *линейным*, если для него справедлив принцип суперпозиции, т.е. выходной эффект от суммы входных сигналов равен сумме реакций на каждый сигнал. Пусть звено производит некоторую операцию A , так что $y = A(x)$. Тогда для линейного звена выполняются соотношения

$$A(x_1 + x_2) = A(x_1) + A(x_2), \quad A(cx(t)) = cAx(t). \quad (3.4.1)$$

В дальнейшем будем рассматривать линейные динамические звенья с постоянными параметрами, т.е. описываемые обыкновенными дифференциальными уравнениями с постоянными коэффициентами.

Линейные звенья могут быть заданы различными характеристиками, связанными между собой [95].

Дифференциальное уравнение дает зависимость между входом и выходом звена и их производными и в общем случае имеет вид

$$\begin{aligned} A_0 x + A_1 \frac{dx}{dt} + A_2 \frac{d^2 x}{dt^2} + \dots + A_l \frac{d^l x}{dt^l} = \\ = B_0 y + B_1 \frac{dy}{dt} + B_2 \frac{d^2 y}{dt^2} + \dots + B_m \frac{d^m y}{dt^m}. \end{aligned} \quad (3.4.2)$$

Передаточная функция звена определяется как отношение преобразованных по Лапласу входного и выходного сигналов (при нулевых начальных условиях)

$$W(s) = \frac{y(s)}{x(s)}, \quad (3.4.3)$$

где

$$\begin{aligned} x(s) &= \int_{-\infty}^{\infty} x(t) e^{-st} dt, \\ y(s) &= \int_{-\infty}^{\infty} y(t) e^{-st} dt, \\ s &= c + j\omega. \end{aligned} \quad (3.4.4)$$

Формально передаточная функция может быть получена из дифференциального уравнения заменой операторов $(d/dt) \rightarrow s$. Тогда с учетом (3.4.2) имеем

$$W(s) = \frac{A_0 + A_1 s + A_2 s^2 + \dots + A_l s^l}{B_0 + B_1 s + B_2 s^2 + \dots + B_m s^m} = \frac{y(s)}{x(s)}. \quad (3.4.5)$$

Для физически реализуемых систем выполняется условие $l \leq m$.

Частотная характеристика описывает реакцию звена на гармонический входной сигнал и определяется как отношение преобразований Фурье входного и выходного сигналов

$$W(j\omega) = \frac{y(j\omega)}{x(j\omega)}, \quad (3.4.6)$$

где

$$x(j\omega) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt, \quad y(j\omega) = \int_{-\infty}^{\infty} y(t) e^{-j\omega t} dt. \quad (3.4.7)$$

Частотная характеристика получается из передаточной функции заменой операторов $s \rightarrow j\omega$.

Импульсная характеристика $g(t)$ звена есть реакция на δ -функцию. Частотная и импульсная характеристики связаны между собой парой преобразований Фурье

$$W(j\omega) = \int_{-\infty}^{\infty} g(t) e^{-j\omega t} dt, \quad g(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} W(j\omega) e^{j\omega t} d\omega. \quad (3.4.8)$$

Для физически реализуемых систем $g(t) \equiv 0$ при $t \leq 0$.

Различные характеристики линейного звена, по существу, описывают его реакцию на различное тестовое воздействие: гармонический сигнал, δ -функцию и т.д. В общем случае для произвольного входного сигнала связь между входом и выходом звена описывается интегралом свертки

$$y(t) = \int_{-\infty}^{\infty} x(\tau) g(t - \tau) d\tau = \int_{-\infty}^{\infty} x(t - \tau) g(\tau) d\tau, \quad (3.4.9)$$

где $g(t)$ — импульсная характеристика.

Для поэлементного функционального моделирования наиболее удобно характеризовать звено передаточной функцией и импульсной характеристикой.

Типовые линейные звенья электронного тракта О. Э. С. Наиболее часто встречающиеся линейные звенья электронного тракта описываются следующими передаточными функциями.

Интегрирующее звено с передаточной функцией

$$W(s) = \frac{1}{s} . \quad (3.4.10)$$

Апериодическое звено:

$$W(s) = \frac{1}{1 + sT} , \quad (3.4.11)$$

где T — постоянная времени звена.

Дифференцирующее (неидеальное) звено:

$$W(s) = \frac{sT}{1 + sT} . \quad (3.4.12)$$

Колебательное звено:

$$W(s) = \frac{1}{T^2 s^2 + 2T\zeta s + 1} , \quad (3.4.13)$$

где T — период собственных колебаний при отсутствии затухания; ζ — коэффициент затухания.

Колебательный контур:

$$W(s) = \frac{s}{T^2 s^2 + 2T\zeta s + 1} . \quad (3.4.14)$$

В ОЭС апериодическое звено описывает гиросtabilизатор, двигатель и интеграторы в системе автоматического регулирования. Апериодическое и дифференцирующее звенья описывают частотную характеристику полосовых усилителей и фильтров, колебательное звено описывает колебательные свойства гиросtabilизатора, колебательный контур характеризует свойства резонансных усилителей.

Для дальнейшего изложения необходимо рассмотреть запаздывающее звено, поэтому приведем его характеристики.

Запаздывающее звено описывается уравнением

$$y(t) = x(t - \tau), \quad (3.4.15)$$

т.е. выходной сигнал, не изменяя формы, отстает от входного на время τ .

Передаточная функция описывается выражением

$$W(s) = e^{-s\tau}, \quad (3.4.16)$$

а частотная характеристика — выражением

$$W(j\omega) = e^{-j\omega\tau}. \quad (3.4.17)$$

На комплексной плоскости частотная характеристика имеет вид окружности единичного радиуса, по которой конец вектора $W(j\omega)$ многократно

проходит по часовой стрелке с изменением частоты от нуля до бесконечности с периодом $2\pi/\tau$. Моделируется запаздывающее звено следующим образом. Если время запаздывания $\tau = n \Delta t$, где Δt — шаг счета модели по времени, то вводится n промежуточных ячеек памяти и входной сигнал по последовательно, на каждом такте счета, пересылается из ячейки в ячейку. Через n шагов сигнал появится на выходе, т.е. будет задержан на время τ .

Рассмотрим методы моделирования линейных динамических звеньев.

Моделирование линейных звеньев с использованием формул численного интегрирования. Представление звена в виде рекурсивного фильтра. В зависимости от используемой характеристики линейного звена получаются разные математические алгоритмы вычисления сигнала на выходе линейного фильтра. Рассмотрим случай, когда звено задано передаточной функцией $W(s)$. При математическом моделировании операция дифференцирования производится с большими погрешностями, чем операция интегрирования. Поэтому линейное звено представляют в виде фильтра, содержащего только интеграторы и не содержащего звеньев чистого дифференцирования. Поскольку передаточная функция физически реализуемых систем всегда имеет вид дробно-рационального выражения (3.4.5), то выделение интеграторов производится путем деления числителя и знаменателя этого выражения на s^m :

$$W(s) = \frac{\frac{A_0}{s^m} + \frac{A_1}{s^{m-1}} + \frac{A_2}{s^{m-2}} + \dots + \frac{A_l}{s^{m-l}}}{\frac{B_0}{s^m} + \frac{B_1}{s^{m-1}} + \frac{B_2}{s^{m-2}} + \dots + B_m} = \frac{y(s)}{x(s)}, \quad (3.4.18)$$

где $1/s^m$ — интегрирующее звено m -го порядка. Разделив числитель и знаменатель (3.4.18) на B_m , получим операторное уравнение

$$\begin{aligned} x(s) &= \left(\frac{A_0}{B_m} \frac{1}{s^m} + \frac{A_1}{B_m} \frac{1}{s^{m-1}} + \dots + \frac{A_l}{B_m} \frac{1}{s^{m-l}} \right) = \\ &= y(s) \left(\frac{B_0}{B_m} \frac{1}{s^m} + \frac{B_1}{B_m} \frac{1}{s^{m-1}} + \dots + \frac{B_{m-1}}{B_m} \frac{1}{s} + 1 \right), \end{aligned} \quad (3.4.19)$$

которое решается относительно $y(s)$:

$$\begin{aligned} y(s) &= x(s) \left(\frac{A_0}{B_m} \frac{1}{s^m} + \frac{A_1}{B_m} \frac{1}{s^{m-1}} + \dots + \frac{A_l}{B_m} \frac{1}{s^{m-l}} \right) - \\ &- y(s) \left(\frac{B_0}{B_m} \frac{1}{s^m} + \frac{B_1}{B_m} \frac{1}{s^{m-1}} + \dots + \frac{B_{m-1}}{B_m} \frac{1}{s} \right). \end{aligned} \quad (3.4.20)$$

Выражению (3.4.20) соответствует структура фильтра, приведенного на рис. 3.18. Такой фильтр называют рекурсивным, так как при расчете выходного сигнала наряду со значениями входного сигнала используются значения и выходного сигнала в предыдущие моменты времени [20].

Рис. 3.18. Рекурсивный линейный фильтр

Из общей структуры рекурсивного фильтра получаются частные случаи для типовых линейных звеньев электронного тракта ОЭС, описываемых передаточными функциями (3.4.11)–(3.4.14). Структурные схемы этих звеньев приведены на рис. 3.19.

Итак, типовые линейные звенья могут быть представлены в виде структуры с обратной связью, содержащей интеграторы. Точность моделирования и устойчивость счета в значительной мере зависят от того метода, каким производится интегрирование. Рассмотрим наиболее распространенные методы численного интегрирования [12], основанные на замене интеграла конечной суммой.

Метод прямоугольников:

$$\int_a^b y \, dt = h(y_0 + y_1 + \dots + y_{n-1}), \quad (3.4.21)$$

где $h = (b - a)/n$ — шаг интегрирования; y_i — значения интегрируемой функции в дискретных точках.

Метод трапеций:

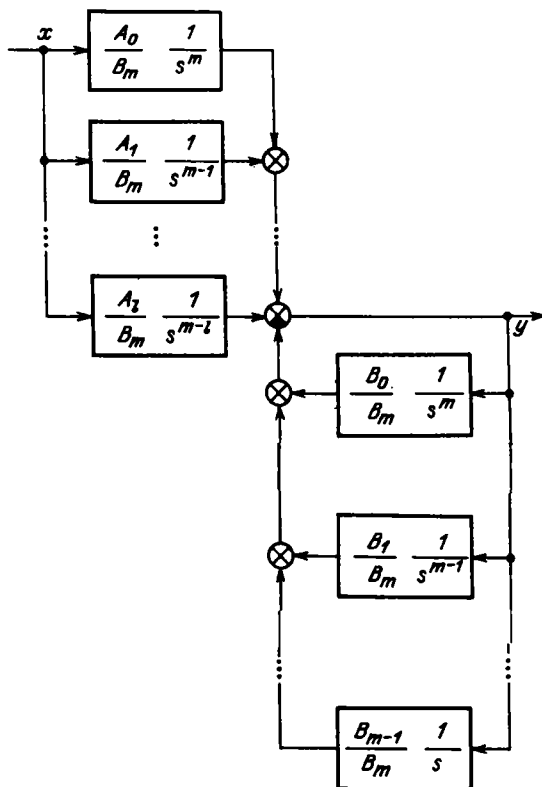
$$\int_a^b y \, dt = \frac{h}{2} (y_0 + 2y_1 + 2y_2 + \dots + 2y_{n-1} + y_n). \quad (3.4.22)$$

Метод парабол (Симпсона):

$$\int_a^b y \, dt = \frac{h}{3} (y_0 + 4y_1 + 2y_2 + 4y_3 + \dots + 4y_{n-1} + y_n). \quad (3.4.23)$$

Эти методы дают хорошие результаты при моделировании обработки сигнала в электронном тракте.

Точность интегрирования зависит от шага $h = \Delta t$ и вида функции $y(t)$. Аналитические выражения этой зависимости довольно сложны. Обычно требуемая величина шага интегрирования подбирается опытным путем (сравнением между собой результатов счета при разных Δt). Точность вычислений всего звена, как системы с обратной связью, зависит от вида дифференциального уравнения. Математическое моделирование линейных



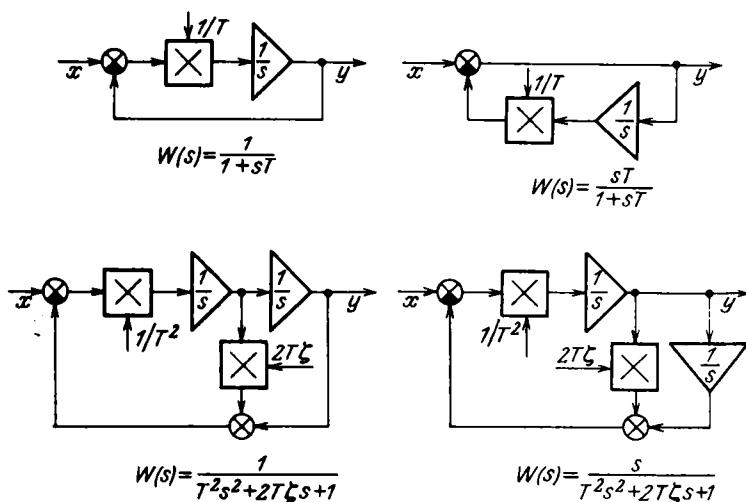


Рис. 3.19. Структурные схемы типовых звеньев

динамических звеньев означает, по существу, численное решение дифференциальных уравнений. Тогда метод трапеций соответствует методу Эйлера решения дифференциальных уравнений, а метод парабол — методу Симпсона.

При моделировании динамических звеньев очень важен вопрос об устойчивости счета. Это значит, что ошибки вычислений не должны накапливаться, иначе процесс будет расходящимся. Устойчивость счета зависит также от шага интегрирования. В некоторых случаях для обеспечения устойчивости счета приходится выбирать шаг интегрирования гораздо меньшим, чем для получения необходимой точности. Отметим, что звенья первого порядка устойчивы при любом разумном шаге счета (здесь Δt определяется требуемой точностью вычислений).

Рассмотрим метод определения необходимого шага счета для устойчивости колебательного звена, вытекающий из теории импульсных систем [135]. Звено устойчиво, если корни знаменателя передаточной функции, умноженные на Δt , расположены внутри окружности радиуса $R = 1$ с центром в точке $(-1, 0)$. Обозначим корни знаменателя $p_{1,2}$. Тогда условие устойчивости запишется в виде

$$\begin{aligned}
 [1 + \operatorname{Re}(\Delta t \cdot p_1)]^2 + \operatorname{Im}^2(\Delta t \cdot p_1) &< 1, \\
 [1 + \operatorname{Re}(\Delta t \cdot p_2)]^2 + \operatorname{Im}^2(\Delta t \cdot p_2) &< 1,
 \end{aligned}
 \tag{3.4.24}$$

где $\operatorname{Re} z$ и $\operatorname{Im} z$ — действительная и мнимая части корней. Поскольку звено второго порядка имеет комплексно-сопряженные корни, можно использовать одно из неравенств.

В качестве примера найдем необходимый шаг счета для моделирования резонансного усилителя с передаточной функцией (3.4.14). Корни

знаменателя равны

$$p_{1,2} = \frac{-\zeta \pm \sqrt{\zeta^2 - 1}}{T} \approx -\frac{\zeta}{T} \pm j \frac{1}{T},$$

поскольку $\zeta^2 \ll 1$. Введем обозначения, применяемые в радиотехнике:

$1/T = \omega_0$ — резонансная частота;

$\zeta/T = \Delta\omega$ — полуширина полосы пропускания на уровне 0,707.

Тогда корни знаменателя будут

$$p_{1,2} = 2\pi(-\Delta f \pm j f_0), \quad (3.4.25)$$

т.е. корни определяются резонансной частотой и полосой пропускания усилителя. Используя условие устойчивости (3.4.24) и выражение для корней (3.4.25), получаем значение шага счета

$$\Delta t_{\max} \approx \frac{1}{\pi} \frac{\Delta f}{f_0^2} \quad (3.4.26)$$

при условии $\Delta f^2 \ll f_0^2$. Например, для усилителя с параметрами $f_0 = 1500$ Гц, $\Delta f = 150$ Гц находим $\Delta t_{\max} = 2 \cdot 10^{-5}$ с. Для обеспечения запаса по устойчивости примем $\Delta t = 1 \cdot 10^{-5}$ с.

Сравним полученный шаг счета с периодом резонансной частоты

$$T_0 = 1/f_0 = 0,667 \cdot 10^{-3} \text{ с.}$$

На один период приходится N точек счета

$$N = T_0/\Delta t \approx 67,$$

что совершенно излишне с точки зрения точности моделирования (достаточно 10 точек на период), но необходимо для устойчивости счета. В подобных случаях с целью увеличения шага счета и повышения скорости моделирования необходимо переходить к разностным методам, рассмотренным ниже.

Представление звена в виде не рекурсивного фильтра. Рассмотрим моделирование линейных динамических звеньев, когда в качестве характеристик звена используется импульсная характеристика $g(t)$. Выходной сигнал описывается интегралом свертки

$$y(t) = \int_{-\infty}^{\infty} g(\tau) x(t - \tau) d\tau. \quad (3.4.27)$$

Учитывая, что $g(t) \equiv 0$ при $t \leq 0$, а также, что, начиная с некоторого $t \geq T$, можно считать $g(t) = 0$, интеграл (3.4.27) запишем с конечными пределами

$$y(t) = \int_0^T g(\tau) x(t - \tau) d\tau. \quad (3.4.28)$$

Выражение (3.4.28) является непрерывной математической моделью линейного звена. Для получения цифровой модели перейдем к соответствующему дискретному эквиваленту, заменив интеграл суммой с использованием методов численного интегрирования.

Введем обозначения для дискретных моментов времени:

$$t_n = n \Delta t, \quad \tau_k = k \Delta t, \quad N = \frac{T}{\Delta t}. \quad (3.4.29)$$

Тогда (3.4.28) запишется в виде

$$y(t_n) = \Delta t \sum_{k=0}^T g(k \Delta t) x(n \Delta t - k \Delta t). \quad (3.4.30)$$

Здесь переменными являются n и k , поэтому сумму можно представить в другой форме, обозначив

$$g(k) = g(k \Delta t), \quad x(n - k) = x(n \Delta t - k \Delta t).$$

Получим

$$y(n) = \sum_{k=0}^N C(k) g(k) x(n - k), \quad (3.4.31)$$

где коэффициенты $C(k)$ определяются принятым способом численного интегрирования. В методе трапеций

$$C(k) = \frac{\Delta t}{2} C_0(k), \quad (3.4.32)$$

где $C_0(k)$ в соответствии с (3.4.22) равны

$$C_0(k) = 1, 2, 2, \dots, 2, 2, 1.$$

В методе парабол

$$C(k) = \frac{\Delta t}{3} C_0(k), \quad (3.4.33)$$

и в соответствии с (3.4.23)

$$C_0(k) = 1, 4, 2, 4, \dots, 2, 4, 1.$$

Замена интеграла суммой, по существу, означает переход от непрерывной системы к дискретной. Математическим аппаратом для описания дискретных (импульсных) систем является z -преобразование

$$x(z) = \sum_{k=0}^{\infty} x(k) z^{-k}, \quad (3.4.34)$$

где $z = e^{s \Delta t}$ — оператор преобразования.

Рассмотрим оператор

$$z^{-1} = e^{-s \Delta t},$$

совпадающий по форме с передаточной функцией (3.4.16) запаздывающего звена при $\tau = \Delta t$, т.е. производящего задержку на один шаг счета. Тогда z^{-k} можно рассматривать как оператор, который производит задержку сигнала на k шагов. В связи с этим формулу (3.4.31) можно записать в другом виде, представив символически

$$x(n - k) = x(n) z^{-k}.$$

Тогда

$$y(n) = \sum_{k=0}^N C(k) g(k) x(n) z^{-k}. \quad (3.4.35)$$

Дискретная передаточная функция определяется, как отношение z -преобразования выходного дискретного сигнала к z -преобразованию входного, и формально получается из (3.4.35) делением обеих частей равенства на $x(n)$:

$$W(z) = \frac{y(n)}{x(n)} = \sum_{k=0}^N C(k)g(k)z^{-k} = \sum_{k=0}^N a(k)z^{-k}, \quad (3.4.36)$$

где

$$a(k) = C(k)g(k).$$

Структурная схема дискретного фильтра, соответствующего передаточной функции (3.4.36), приведена на рис. 3.20. Здесь последовательность

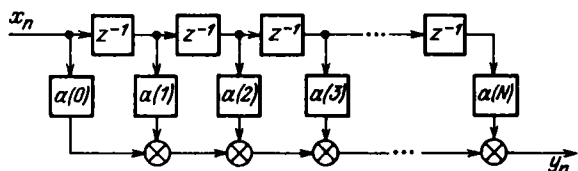


Рис. 3.20. Нерекурсивный линейный фильтр

дискретных значений входного сигнала поступает на линию задержки с N отводами, задержка между которыми равна Δt . К отводам подключены весовые усилители с коэффициентами усиления $a(k)$. Выходной сигнал равен сумме сигналов усилителей. Такой фильтр является нерекурсивным, так как для расчета выходного сигнала используются дискретные отсчеты входного сигнала. В связи с этим не возникает вопроса об устойчивости — нерекурсивный фильтр устойчив всегда. Кроме того, этот метод может быть использован для моделирования систем с переменными параметрами, когда импульсная характеристика зависит от двух переменных $g(t, \tau)$. Тогда формула (3.4.31) запишется в виде

$$y(n) = \sum_{k=0}^N C(k)g(n, k)x(n)z^{-k}. \quad (3.4.37)$$

Недостатком метода дискретной свертки является большой объем вычислений по сравнению с предыдущим методом рекурсивного фильтра, пропорциональный ширине импульсной характеристики.

Замена непрерывных линейных систем эквивалентными импульсными. Дискретная аппроксимация. Для линейных систем с постоянными параметрами значительную экономию вычислений дает применение разностных методов. Сущность этих методов состоит в замене непрерывных линейных систем эквивалентными импульсными, поведение которых можно описать простыми рекуррентными соотношениями.

Метод дискретной аппроксимации [20] состоит в замене операторов непрерывного интегрирования операторами дискретного интегрирования. Для этого методом, изложенным выше, звено приводится к структуре рекурсивного фильтра, показанного на рис. 3.18 и описываемого формулой (3.4.20).

Рассмотрим метод Тастина представления операторов дискретного интегрирования. Интегрирующее звено первого порядка описывается уравнением

$$y(t) = \int_0^t x(\tau) d\tau. \quad (3.4.38)$$

Для дискретных моментов времени можно записать

$$y_n = \int_0^{n\Delta t} x(\tau) d\tau = \int_0^{(n-1)\Delta t} x(\tau) d\tau + \int_{(n-1)\Delta t}^{n\Delta t} x(\tau) d\tau = y_{n-1} + J(n), \quad (3.4.39)$$

где y_{n-1} — значение интеграла на предыдущем шаге счета; $J(n)$ — приращение интеграла на последнем шаге счета.

Вычислим значение $J(n)$ как площадь трапеции с основаниями x_n, x_{n-1} и высотой Δt :

$$J(n) = \int_{(n-1)\Delta t}^{n\Delta t} x(\tau) d\tau = \frac{\Delta t}{2} (x_n + x_{n-1}). \quad (3.4.40)$$

Тогда для вычисления y_n получим рекуррентное соотношение

$$y_n = y_{n-1} + \frac{\Delta t}{2} (x_n + x_{n-1}), \quad (3.4.41)$$

которое можно записать в другой форме, введя оператор задержки на один шаг счета z^{-1} :

$$y = yz^{-1} + \frac{\Delta t}{2} (x + xz^{-1}). \quad (3.4.42)$$

Отсюда получаем передаточную функцию оператора дискретного интегрирования

$$W(z) = \frac{y(z)}{x(z)} = \frac{\Delta t}{2} \frac{1 + z^{-1}}{1 - z^{-1}}. \quad (3.4.43)$$

Интегрирующее звено m -го порядка представляется в методе Тастина как последовательное соединение m интегрирующих звеньев первого порядка

$$W^{(m)}(z) = \left(\frac{\Delta t}{2} \frac{1 + z^{-1}}{1 - z^{-1}} \right)^m \quad (3.4.44)$$

Более точным является метод Рагазини—Бергена [20], в котором операторы интегрирования разного порядка имеют разные передаточные функции. Приведем без вывода дискретные передаточные функции для интеграторов первых трех порядков:

$$\begin{aligned} \frac{1}{s} &\rightarrow \frac{\Delta t}{2} \frac{1 + z^{-1}}{1 - z^{-1}}, \\ \frac{1}{s^2} &\rightarrow \frac{\Delta t^2}{6} \frac{1 + 4z^{-1} + z^{-2}}{(1 - z^{-1})^2}, \\ \frac{1}{s^3} &\rightarrow \frac{\Delta t^3}{24} \frac{1 + 11z^{-1} + 11z^{-2} + z^{-3}}{(1 - z^{-1})^3}. \end{aligned} \quad (3.4.45)$$

Представления операторов непрерывного интегрирования операторами дискретного интегрирования называются *z-формами*. Много различных *z-форм* рассмотрено в [20] и [126].

Заменяя в (3.4.20) операторы интегрирования соответствующими *z-формами*, получаем рекуррентные разностные соотношения, которым в общем случае соответствует дискретная передаточная функция вида

$$W(z) = \frac{a_0 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_l z^{-l}}{1 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_m z^{-m}} = \frac{y(z)}{x(z)}. \quad (3.4.46)$$

Передаточная функция (3.4.46) описывает рекурсивный фильтр, структура которого приведена на рис. 3.21.

Рассмотренный метод дискретной аппроксимации означает, по существу, замену непрерывной системы импульсной со многими импульсными элементами (по одному на каждое интегрирующее звено). Такое многократное прерывание и сглаживание сигнала при прохождении его через систему создает погрешности вычислений. Поэтому надо стремиться представлять структуру линейной системы в виде совокупности простых типовых звеньев, чтобы не вводить интеграторы высокого порядка.

z-преобразование. Рассмотренные методы моделирования динамических звеньев в итоге сводятся к методам дискретизации непрерывных интеграторов путем использования формул численного интегрирования или *z-форм*. Погрешность вычисления зависит от выбранного типа дискретизации, т.е. в конечном счете от типа аппроксимации сигнала между дискретными отсчетами: прямоугольниками, трапециями, параболками. Такой подход типичен для теории импульсных систем, содержащих непрерывную часть, поскольку необходимо знать значение сигнала в промежутках между отсчетами.

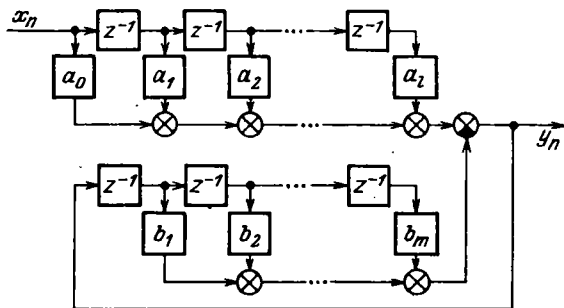


Рис. 3.21. Рекурсивный цифровой фильтр

Если не интересоваться значениями сигнала в этих промежутках, а такая картина наблюдается при математическом моделировании, когда система считается в дискретные моменты с шагом Δt , то можно использовать специальные точные методы, созданные для синтеза цифровых фильтров. Используя эти методы, можно создавать цифровые фильтры, не имеющие соответствующих аналогов в непрерывных фильтрах. В настоящее время цифровые фильтры используются не только для целей моделирования на

ЭВМ, но и в реальной аппаратуре в связи с проникновением вычислительной техники во многие системы (использование микропроцессоров, мини-ЭВМ и пр.). Рассмотрим два метода синтеза цифровых фильтров.

В методе z -преобразования синтезируется цифровой фильтр, импульсная характеристика которого в дискретные моменты времени совпадает с импульсной характеристикой непрерывного фильтра [31]. Найдем связь между передаточной функцией непрерывного фильтра и дискретной передаточной функцией цифрового фильтра.

Пусть передаточная функция непрерывного фильтра имеет вид

$$W(s) = \sum_{i=1}^m \frac{A_i}{s + a_i}, \quad (3.4.47)$$

т.е. задана в виде суммы элементарных передаточных функций. Импульсная характеристика такого фильтра находится как обратное преобразование Лапласа передаточной функции

$$g(t) = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} \left\{ \sum_{i=1}^m \frac{A_i}{s + a_i} \right\} e^{st} dt = \sum_{i=1}^m A_i e^{-a_i t}. \quad (3.4.48)$$

В дискретные моменты времени $n\Delta t$ импульсная характеристика соответствующего цифрового фильтра записывается в виде

$$g(n\Delta t) = \sum_{i=1}^m A_i e^{-a_i n\Delta t}, \quad (3.4.49)$$

так как в дискретные моменты обе импульсные характеристики должны совпадать.

Передаточная функция импульсного фильтра получается путем z -преобразования импульсной дискретной характеристики [31]

$$\begin{aligned} W(z) &= \sum_{n=0}^{\infty} g(n\Delta t) z^{-n} = \sum_{i=1}^m A_i \sum_{n=0}^{\infty} e^{-a_i n\Delta t} z^{-n} = \\ &= \sum_{i=1}^m \frac{A_i}{1 - e^{-a_i \Delta t} z^{-1}}. \end{aligned} \quad (3.4.50)$$

Здесь использована формула для суммы убывающей геометрической прогрессии

$$s = a_1 / (1 - q) \quad (3.4.51)$$

с первым членом $a_1 = 1$ и знаменателем $q = e^{-a_i \Delta t} z^{-1}$.

Таким образом, если передаточная функция непрерывного фильтра задана в виде (3.4.47), то передаточная дискретная функция соответствующего цифрового фильтра сразу может быть записана в виде

$$W(z) = \sum_{i=1}^m \frac{A_i}{1 - b_i z^{-1}}, \quad (3.4.52)$$

где

$$b_i = e^{-a_i \Delta t}.$$

Полученный цифровой фильтр имеет структуру, приведенную на рис. 3.21.

При синтезе цифрового фильтра необходимо иметь в виду следующее. Поскольку коэффициенты фильтра (3.4.52) зависят от шага счета Δt , то

такой фильтр имеет коэффициент передачи, зависящий от шага счета, в то время как непрерывный фильтр имеет постоянный коэффициент передачи, равный A_i при $\omega = 0$. Чтобы цифровой фильтр при любом шаге счета имел постоянный коэффициент передачи, необходимо вводить нормировочный множитель. Для выбора этого множителя будем считать, что цифровой и аналоговый фильтры должны иметь одинаковые коэффициенты передачи на нулевой частоте.

Учитывая, что

$z = e^{s\Delta t} = 1$ при $s = 0$,
можно записать соответствие

$$\begin{matrix} W(s) \rightarrow W(z). \\ s=0 \quad \quad z=1 \end{matrix} \quad (3.4.53)$$

Отсюда получаем соотношение для вычисления нормирующего множителя

$$W(z=1) = K \frac{A_i}{1-b} = A_i,$$

следовательно,

$$K = 1 - b = 1 - e^{-a_i \Delta t}. \quad (3.4.54)$$

Представление передаточной функции в виде (3.4.47) необходимо для метода z -преобразования, иначе бесконечная сумма в (3.4.50) не сворачивается в конечное выражение. Если условие (3.4.47) не выполняется, например для фильтра с передаточной функцией (3.4.12), то для синтеза цифрового фильтра можно использовать метод, изложенный ниже.

Практически не обязательно каждый раз производить разложение передаточной функции непрерывного фильтра на элементарные дроби. Можно использовать готовые таблицы z -преобразования. Очень хорошие таблицы соответствия передаточных функций $W(s)$, $W(z)$ и импульсных характеристик $g(t)$ приведены в [126].

Билинейное преобразование передаточной функции. Пусть имеется аналоговый фильтр с передаточной функцией $W(s)$. Его частотная характеристика $W(j\omega)$ находится путем вычисления $W(s)$ в точках на мнимой оси плоскости s . Если в функции $W(s)$ заменить оператор s рациональной функцией от z , которая отображает мнимую ось в s -плоскости на единичную окружность в z -плоскости, то полученная в результате функция $W'(z)$, вычисленная вдоль единичной окружности, примет те же значения, что и функция $w(s)$ при вычислении вдоль мнимой оси [31]. Напомним, что окружность единичного радиуса есть частотная характеристика оператора запаздывания z^{-1} .

Простейшее рациональное преобразование, отображающее ось $j\omega$ на единичную окружность, имеет вид

$$s \rightarrow \frac{z-1}{z+1}. \quad (3.4.55)$$

Обозначим аналоговую частоту через ω_a , цифровую через ω_d . Тогда при

$$\omega_a = \operatorname{tg} \frac{\omega_d \Delta t}{2} \quad (3.4.56)$$

функции $W(\omega_a)$ и $W'(\omega_d \Delta t)$ принимают одинаковые значения.

Рассмотрим это подробнее. Покажем, что функция (3.4.55) действительно отображает мнимую ось на единичную окружность. Пусть

$$s = \frac{z - 1}{z + 1},$$

откуда

$$z = \frac{1 + s}{1 - s}.$$

Подставив $s = j\omega$, получим

$$z = \frac{1 + j\omega}{1 - j\omega} = \frac{1 - \omega^2}{1 + \omega^2} + j \frac{2\omega}{1 + \omega^2} = U + jV. \quad (3.4.57)$$

Легко проверить, что соотношение (3.4.57) задает уравнение окружности единичного радиуса, так как

$$U^2 + V^2 = 1.$$

Рассмотрим привязку частотных шкал. Частотная шкала из плоскости s в плоскость z трансформируется следующим образом:

$$\text{при } j\omega = \infty \quad z = -1,$$

$$\text{при } j\omega = 0 \quad z = +1,$$

$$\text{при } j\omega = -\infty \quad z = -1,$$

что следует из (3.4.57)

Угол, который образует радиус при повороте вдоль единичной окружности, равен

$$\operatorname{tg} \varphi = \frac{V}{U} = \frac{2\omega_a}{1 - \omega_a^2}. \quad (3.4.58)$$

Запишем это уравнение относительно аналоговой частоты:

$$\omega_a^2 + \frac{2}{\operatorname{tg} \varphi} \omega_a - 1 = 0, \quad (3.4.59)$$

откуда

$$\omega_a = \frac{1 - \cos \varphi}{\sin \varphi} = \operatorname{tg} \frac{\varphi}{2} \quad (3.4.60)$$

(второе решение дает $\operatorname{ctg} \frac{\varphi}{2}$ для отрицательных частот). Угол поворота в дискретном фильтре $\omega_d \Delta t$. Следовательно, выражение

$$\omega_a = \operatorname{tg} \frac{\omega_d \Delta t}{2}$$

дает привязку частотных шкал. Отметим, что шкала частот аналогового фильтра линейная, цифрового — нелинейная.

Последовательность действий при использовании этого метода следующая.

1. Задают критические частоты требуемого цифрового фильтра ω_d .
2. Вычисляют соответствующие аналоговые частоты по формуле (3.4.56)
3. Рассчитывают передаточную функцию $W'(s)$, имеющую на новых частотах свойства цифрового фильтра.
4. Производят подстановку

$$s = \frac{z - 1}{z + 1} = \frac{1 - z^{-1}}{1 + z^{-1}}. \quad (3.4.61)$$

В качестве примера рассмотрим построение цифрового фильтра для дифференцирующего звена с передаточной функцией

$$W(s) = \frac{Ts}{1 + Ts}.$$

В данном случае одна критическая частота — частота среза, равная $\omega_d = \frac{1}{T}$.

Соответствующая аналоговая частота $\omega_a = \text{tg} \frac{\Delta t}{2T}$. Новая постоянная времени равна $T_a = \frac{1}{\omega_a} = \text{ctg} \frac{\Delta t}{2T}$, а передаточная функция

$$W'(s) = \frac{\text{ctg} \frac{\Delta t}{2T} \cdot s}{1 + \text{ctg} \frac{\Delta t}{2T} \cdot s} = \frac{as}{1 + as}.$$

Производим подстановку (3.4.61):

$$W(z) = \frac{a \frac{1 - z^{-1}}{1 + z^{-1}}}{1 + a \frac{1 - z^{-1}}{1 + z^{-1}}} = K \frac{1 - z^{-1}}{1 - bz^{-1}},$$

где

$$b = \frac{1 - \text{tg} \frac{\Delta t}{2T}}{1 + \text{tg} \frac{\Delta t}{2T}},$$

$$K = \frac{1}{1 + \text{tg} \frac{\Delta t}{2T}}.$$

Для этого фильтра выполняется нормировка на высоких частотах

$$\begin{matrix} W(s) \rightarrow W(z). \\ s \rightarrow \infty & z \rightarrow -1 \end{matrix}$$

Данный метод дает совпадение частотных характеристик аналогового и цифрового фильтров для критических частот и не дает совпадения импульсных характеристик.

Общие замечания по цифровой фильтрации. Сравним рассмотренные методы. Наиболее простым методом, требующим небольшой подготовительной работы по расчету коэффициентов, является представление звена в виде рекурсивного фильтра с использованием формул численного интегрирования. Этот метод целесообразно использовать во всех случаях, где не требуется очень высокой точности, особенно для моделирования звеньев первого порядка. Шаг счета Δt обычно выбирается из условия

$$T = (5 \div 10) \Delta t, \quad (3.4.62)$$

где T — постоянная времени фильтра. При моделировании звеньев второго порядка и выше возникает вопрос об устойчивости счета. Для обеспечения приемлемого шага Δt необходимо переходить к методам дискретных фильтров.

Применение метода дискретной свертки с представлением звена в виде нерекурсивного фильтра позволяет обойти проблему устойчивости и, кроме того, моделировать звенья с переменными параметрами. Шаг счета также выбирается из условия (3.4.62). Недостатком метода является большой объем вычислений на одно значение выходного сигнала.

Метод дискретной аппроксимации позволяет уменьшить шаг счета при моделировании колебательных систем без потери устойчивости, что позволяет значительно сократить время счета. Здесь требуется подготовительная работа по расчету коэффициентов. Частотные и фазовые характеристики цифровых фильтров получаются близкими к соответствующим аналоговым, что важно при моделировании. Шаг счета выбирается из соотношения (3.4.62).

Методы z -преобразования и билинейного преобразования дают наибольшую экономию вычислений, так как позволяют моделировать практически с любым шагом без потери устойчивости. Обычно шаг счета выбирается из соотношения (3.4.62) для колебательных систем и из соотношения

$$T = (1 \div 2) \Delta t \quad (3.4.63)$$

для звеньев первого порядка. Этими методами можно синтезировать цифровые фильтры, не имеющие непрерывных аналогов, а также использовать их для построения различных цифровых резонаторов и фильтров с резким спадом на границе полосы пропускания. Недостатком является большое отличие частотных и особенно фазовых характеристик от соответствующих непрерывных.

Необходимо отметить, что при использовании любого метода частотные характеристики цифровых фильтров и аналоговых совпадают только в определенном диапазоне частот, так как частотные характеристики цифровых фильтров периодические с периодом $\Omega = \omega \Delta t = 2\pi$, что является следствием периодичности оператора $z = e^{s \Delta t}$. Рабочая область частот цифрового фильтра равна $\Omega = 0 \div \pi$; хорошее совпадение частотных характеристик аналоговых и непрерывных получается в диапазоне $\Omega = 0 \div \frac{\pi}{2}$.

В заключение еще раз остановимся на выборе шага счета Δt . Обычно считается, что уменьшение Δt повышает точность вычислений. Это утверждение справедливо до определенной степени в силу следующих причин.

Во-первых, коэффициенты в цифровых фильтрах вычисляются с определенной погрешностью вследствие округления в цифровом процессоре ЭВМ. Как показано в [31], неточность вычисления коэффициентов тем сильнее сказывается на точности моделирования, чем меньше шаг счета Δt . Для меньшей чувствительности фильтра к неточности вычисления коэффициентов нужно всегда стремиться представить систему высокого порядка как совокупность последовательного или параллельного соединения звеньев низкого порядка (первого или второго).

Во-вторых, ошибки округления, помимо влияния через коэффициенты, участвуя в дальнейших вычислениях, могут усиливаться в зависимости от структуры фильтра.

Ошибки округления в меньшей степени сказываются при использовании универсальных ЭВМ с большой длиной слова и могут быть значительными в специализированных вычислителях. Поэтому при расчете специализированной цифровой аппаратуры основным условием для цепей фильтрации является правильный выбор требуемой длины машинного слова.

Моделирование линейных элементов ОЭСС на основе рекуррентных соотношений. Рекуррентные соотношения, используемые в этом методе моделирования, являются разновидностью разностных уравнений и получаются на основе известных решений линейных дифференциальных уравнений, описывающих некоторые элементы тракта ОЭСС, а также исполнительные элементы, работающие в линейном режиме. Рассмотрим применение метода на примере апериодического фильтра с передаточной функцией (3.4.11). Соответствующее дифференциальное уравнение имеет вид

$$T \frac{dy}{dt} + y = x(t), \quad (3.4.64)$$

где $x(t)$ — входная (управляющая) функция; $y(t)$ — выходная величина.

Согласно известному соотношению В.А. Котельникова [135] входная функция $x(t)$, обладающая финитным спектром, может быть представлена своими значениями в дискретных точках t_n ($n = 0, 1, \dots$). Обозначая $x(t_n) = x_n$, получаем вместо (3.4.64)

$$T \frac{dy}{dt} + y = x_n, \quad (3.4.65)$$

и, следовательно, при $t_n \leq t < t_{n+1}$

$$y(t) = x_n + Ce^{-\frac{t}{T}}, \quad (3.4.66)$$

откуда при $t = t_n$

$$y_n = x_n + Ce^{-\frac{t_n}{T}}. \quad (3.4.67)$$

Определяя постоянную интегрирования из (3.4.67) и подставляя ее значение в (3.4.66), запишем решение уравнения (3.4.65). При $t = t_{n+1}$ получаем

рекуррентное соотношение

$$y_{n+1} = x_n + (y_n - x_n)e^{-\frac{h}{T}}, \quad (3.4.68)$$

в котором $h = t_{n+1} - t_n$ ($n = 0, 1, 2, \dots$).

Формула (3.4.68) удобна для программирования при постоянном шаге вычисления входной функции $x_n = x(nh)$.

Аналогичными приемами находится рекуррентное соотношение для колебательного звена, описываемого передаточной функцией (3.4.13). Дифференциальное уравнение в этом случае имеет вид

$$T^2 \frac{d^2 y}{dt^2} + 2\zeta T \frac{dy}{dt} + y = x(t), \quad (3.4.69)$$

и на интервале $t_n \leq t < t_{n+1}$ решение уравнения записывается в форме

$$y(t) = x_n + C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}. \quad (3.4.70)$$

Для вычисления постоянных интегрирования C_1 и C_2 используется соотношение, вытекающее из (3.4.70):

$$\dot{y}(t) = \lambda_1 C_1 e^{\lambda_1 t} + \lambda_2 C_2 e^{\lambda_2 t}. \quad (3.4.71)$$

При $t = t_n$ имеем

$$\begin{aligned} y_n &= x_n + C_1 e^{\lambda_1 t_n} + C_2 e^{\lambda_2 t_n}, \\ \dot{y}_n &= \lambda_1 C_1 e^{\lambda_1 t_n} + \lambda_2 C_2 e^{\lambda_2 t_n}. \end{aligned} \quad (3.4.72)$$

В результате подстановки (3.4.72) в (3.4.70) получаем рекуррентные соотношения

$$\begin{aligned} y_{n+1} &= x_n - \frac{\lambda_2(y_n - x_n) - \dot{y}_n}{\lambda_1 - \lambda_2} e^{\lambda_1 h} + \frac{\lambda_1(y_n - x_n) - \dot{y}_n}{\lambda_1 - \lambda_2} e^{\lambda_2 h}, \\ \dot{y}_{n+1} &= - \frac{\lambda_1 \lambda_2 (y_n - x_n) - \lambda_1 \dot{y}_n}{\lambda_1 - \lambda_2} e^{\lambda_1 h} + \frac{\lambda_1 \lambda_2 (y_n - x_n) - \lambda_2 \dot{y}_n}{\lambda_1 - \lambda_2} e^{\lambda_2 h}. \end{aligned} \quad (3.4.73)$$

Практически интересен случай $\zeta < 1$, при котором корни характеристического уравнения λ_1 и λ_2 являются комплексными и сопряженными:

$$\lambda_1 = \frac{1}{T}(-\zeta + j\sqrt{1-\zeta^2}), \quad \lambda_2 = \frac{1}{T}(-\zeta - j\sqrt{1-\zeta^2}). \quad (3.4.74)$$

Обозначая $\alpha = -\zeta/T$, $\beta = \frac{\sqrt{1-\zeta^2}}{T}$, имеем $\lambda_1 - \lambda_2 = 2j\beta$, $\lambda_1 \lambda_2 = \alpha^2 + \beta^2$.

Кроме того, используя известные соотношения вида

$$\frac{e^{j\beta h} + e^{-j\beta h}}{2j} = \cos \beta h, \quad \frac{e^{j\beta h} - e^{-j\beta h}}{2j} = \sin \beta h, \quad (3.4.75)$$

находим

$$\begin{aligned} y_{n+1} &= x_n - (y_n - x_n)e^{\alpha h} \left(\frac{\alpha}{\beta} \sin \beta h - \cos \beta h \right) + \dot{y}_n \frac{e^{\alpha h}}{\beta} \sin \beta h, \\ \dot{y}_{n+1} &= - \frac{\alpha^2 + \beta^2}{\beta} (y_n - x_n)e^{\alpha h} \sin \beta h + \dot{y}_n e^{\alpha h} \left(\frac{\alpha}{\beta} \sin \beta h + \cos \beta h \right). \end{aligned} \quad (3.4.76)$$

Заметим, что в формулах (3.4.68) и (3.4.76) шаг счета определяется не параметрами T и ξ , а скоростью изменения входной функции $x(t)$. В частности, при $x(t) = \text{const}$ эти формулы являются точными.

При моделировании систем с импульсным управлением формулы (3.4.68) и (3.4.76) обеспечивают точное решение при шаге счета h , равном периоду следования импульсов. Это свойство было нами использовано при анализе дискретных ОЭСС в § 2.5.

Метод рекуррентных соотношений, представленный здесь, достаточно универсален. Укажем, например, на его применение для элементов ОЭСС с передаточной функцией (3.4.12). В этом случае дифференциальное уравнение звена имеет вид

$$T \frac{dy}{dt} + y = T \frac{dx}{dt}. \quad (3.4.77)$$

Обозначим разность $y(t) - x(t) = z(t)$. Тогда вместо (3.4.77) получим

$$T \frac{dz}{dt} + z = -x. \quad (3.4.78)$$

Здесь записана формула (3.4.64) в новых обозначениях. Отсюда непосредственно следует рекуррентное соотношение

$$z_{n+1} = -x_n + (z_n + x_n)e^{-\frac{h}{T}} \quad (3.4.79)$$

Возвращаясь к прежней переменной y_n , получаем окончательно

$$y_{n+1} = x_{n+1} - x_n + y_n e^{-\frac{h}{T}}. \quad (3.4.80)$$

Моделирование нелинейных звеньев ОЭС. Нелинейные безынерционные звенья систем автоматического управления задаются своей амплитудной характеристикой, представляющей зависимость выходной величины от входной. Амплитудная характеристика при моделировании может быть задана в различной форме: аналитически, графически, таблично. В первом случае характеристика описывается некоторой формальной зависимостью, во втором представляется графиком, в третьем — таблицей соответствий значений входных и выходных величин.

При аналитическом задании характеристики выходной сигнал нелинейного элемента при моделировании на ЭВМ вычисляется по соответствующей формуле.

При табличном представлении амплитудной характеристики возникает задача интерполяции между узловыми точками таблицы. Для этого обычно используется линейная или параболическая интерполяция, программы которых имеются в математическом обеспечении современных ЭВМ и нами рассматриваться не будут. В третьем случае график амплитудной характеристики описывается каким-либо подходящим аналитическим выражением (может быть, несколькими на разных участках), и дальше расчет сигнала проводится по соответствующей формуле или переводится в таблицу с определенным шагом дискретности и далее используются программы интерполяции.

Рассмотрим типовые нелинейные элементы опико-электронных систем, заданные аналитическими выражениями.

Квадратичный детектор описывает работу амплитудного детектора при малых уровнях сигнала и задается выражениями

$$y = \begin{cases} x^2 & \text{при } x \geq 0, \\ 0 & \text{при } x < 0. \end{cases} \quad (3.4.81)$$

Линейный детектор описывает работу амплитудного детектора при больших уровнях сигнала

$$y = \begin{cases} k(x - x_0) & \text{при } x \geq x_0, \\ 0 & \text{при } x < x_0. \end{cases} \quad (3.4.82)$$

Релейный элемент описывает работу пороговых схем и задается выражениями

$$y = \text{sign } x = \begin{cases} 1 & \text{при } x \geq 0, \\ -1 & \text{при } x < 0. \end{cases} \quad (3.4.83)$$

Нелинейный элемент с зоной нечувствительности описывается выражениями

$$y = \begin{cases} k(|x| - |x_0|) \text{sign } x & \text{при } |x| \geq x_0, \\ 0 & \text{при } |x| < x_0. \end{cases} \quad (3.4.84)$$

Усилитель с ограничением:

$$y = \begin{cases} kx & \text{при } |x| \leq x_0, \\ y_0 \text{sign } x & \text{при } |x| > x_0. \end{cases} \quad (3.4.85)$$

Усилитель с плавным насыщением может быть задан какой-либо гладкой функцией, например

$$y = A \arctg kx. \quad (3.4.86)$$

Квантователь по уровню с зоной нечувствительности:

$$y = y_0 E\left(\frac{x}{x_0}\right), \quad (3.4.87)$$

где функция $E(\cdot)$ означает выделение целой части со знаком, а x_0, y_0 — величины разряда квантования входного и выходного сигналов.

Фазовый детектор производит разложение сигнала по каналам управления в ОЭС с модуляцией и задается выражениями

$$u_1 = u \cos \omega_m t, \quad u_2 = u \sin \omega_m t, \quad (3.4.88)$$

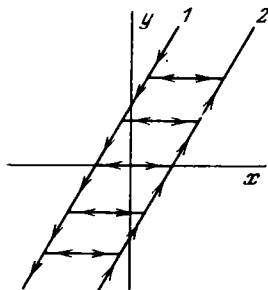
где u — сигнал в одноканальном тракте; ω_m — частота модуляции.

В рассмотренных примерах выходной сигнал нелинейного элемента зависит от одной входной величины. Встречаются нелинейные элементы, выходной сигнал которых зависит от двух переменных. Одним из таких элементов является люфт в виде петли гистерезиса в механических соединениях исполнительных элементов (рис. 3.22). Здесь выходной сигнал зависит не только от величины смещения, но и от направления смещения, т.е. амплитудная характеристика задается выражением $y = f(x, \dot{x})$. Для моделирования таких нелинейностей наряду с аналитическими зависимостями приходится использовать логические соотношения.

При движении по прямой 1 вниз (см. рис. 3.22) выходной сигнал определяется выражением $y = x + x_0$, при движении по прямой 2 вверх выходной сигнал равен $y = x - x_0$, а при нахождении в зоне люфта выходной сигнал равен некоторому постоянному значению $y = \text{const}$ в соответствии с последним значением на прямой 1 или 2.

Рассмотренные нелинейные элементы позволяют моделировать большинство нелинейных звеньев систем автоматического управления. Полностью электронный тракт ОЭС описывается комбинацией линейных динамических звеньев и нелинейных элементов. Аналогичным образом могут

Рис. 3.22. Амплитудная характеристика люфта



быть представлены модели исполнительных элементов, поэтому здесь они специально не рассматриваются.

§ 3.5. Математическое моделирование флуктуаций сигнала оптико-электронных систем

Статистические характеристики случайных сигналов, используемые при моделировании. При моделировании флуктуаций сигнала ОЭС возникает необходимость в трех статистических объектах: случайных величинах, случайных процессах и случайных полях. Случайными величинами являются параметры элементов ОЭС, случайными процессами — шумы фотоприемника и флуктуации входного сигнала, случайными полями описывается распределение яркости объектов и фонов. Поэтому возникает задача генерирования на ЭВМ перечисленных статистических объектов с заданными статистическими характеристиками.

Наиболее просто решается задача для случайных величин, так как для них исчерпывающей характеристикой является закон распределения или плотность распределения. Поэтому при генерировании на ЭВМ случайных величин необходимо обеспечить заданный закон распределения. В практике математического моделирования наиболее часто возникает необходимость генерирования случайных величин с равномерным, нормальным, релеевским и пуассоновским законами распределения. Ниже будут рассмотрены методы генерации случайных величин с перечисленными законами, а также с произвольным заданным законом распределения.

Для случайных процессов $\xi(t)$ полной характеристикой является многомерный закон распределения при стремлении размерности к бесконечности. Практически известны способы генерирования случайных процессов по заданным многомерным законам [20], однако они довольно трудоемки,

и, кроме того, обычно отсутствуют параметры этих законов для реальных физических процессов, протекающих в ОЭС. Поэтому при моделировании случайных процессов часто ограничиваются рассмотрением моментов первого и второго порядков и характеризуют процессы одномерным законом распределения и корреляционной функцией (или энергетическим спектром). Для наиболее часто встречающихся случайных процессов с нормальным законом распределения эти две величины полностью характеризуют случайный процесс, так как по ним могут быть построены законы распределения любого порядка.

Все сказанное для случайных процессов справедливо и для случайных полей, но задача здесь усложняется тем, что заданные статистические характеристики нужно обеспечить для пространства нескольких измерений. В общем случае поля яркости являются функциями трех переменных $B(x, y, t)$ — координат и времени, но во многих случаях моделирования ОЭС можно считать поле неизменным во времени, статическим, и генерировать на ЭВМ двумерные поля яркости $B(x, y)$.

Закон распределения и корреляционная функция случайного процесса определяются существом рассматриваемой задачи. Наиболее часто при моделировании ОЭС используются случайные процессы с релевским, нормальным и логарифмически нормальным законами распределения. Аппроксимирующие выражения для корреляционных функций реальных процессов могут быть выбраны по-разному, но при этом обычно исходят из следующих положений [113].

Математическое описание должно удовлетворять общим свойствам корреляционной функции и отображать характерные особенности экспериментальных результатов. Аппроксимирующее выражение может быть найдено обычными методами приближения функций с большой точностью. Однако большая точность приближения к найденным экспериментальным кривым часто бывает не только не нужна для данной практической задачи, но и может быть нежелательна, так как воспроизведение различных особенностей экспериментальных кривых, возникших из-за ограниченной точности обработки, может только исказить физическую сущность рассматриваемого явления.

При выборе аппроксимирующего выражения полезно учитывать механизм возникновения моделируемого случайного процесса. Например, большинство случайных процессов, встречающихся в технических приложениях, можно рассматривать как результат прохождения белого шума через линейную динамическую систему с частотной характеристикой $W(j\omega)$. В этом случае энергетический спектр процесса описывается выражением

$$S(\omega) = \frac{P_{2l}(\omega)}{Q_{2m}(\omega)} = |W(j\omega)|^2, \quad (3.5.1)$$

где $P_{2l}(\omega)$, $Q_{2m}(\omega)$ — полиномы степеней $2l$ и $2m$ относительно ω .

Подробные сведения о статистических характеристиках реальных сигналов приведены в [48, 115].

Моделирование случайных величин и случайных процессов. Случайные величины с заданным законом распределения. Часто исходным материалом для генерирования случайных величин с заданным законом распределения служат случайные числа, равномерно

распределенные на интервале $[0, 1]$. Поэтому рассмотрим генерирование на ЭВМ случайных величин с плотностью распределения

$$f(x) = \text{const} = 1. \quad (3.5.2)$$

Наиболее распространенным методом получения случайных чисел с равномерным законом распределения является метод выделения остатка [32] при делении числа на некоторый модуль m :

$$x_{i+1} = (ax_i + c) \bmod m, \quad (3.5.3)$$

где символ $(z) \bmod m$ означает остаток от деления числа z на число m . Очевидно, что остаток всегда заключен в интервале $[0, 1]$.

Получаемые числа являются псевдослучайными, так как через некоторый период начинают повторяться. Период повторения соизмерим с числом m , выбранным за модуль. Поэтому для получения возможно большего периода случайной последовательности число m выбирают близким к максимально возможному в данной ЭВМ. Для многих ЭВМ можно принять $m = 2^{40}$. В качестве константы выбирают целое число, часто равное модулю, т.е. $a = m$. Константу c выбирают некратной модулю, например 5^{17} . Начальное значение x_0 выбирают любым в диапазоне $[0, 1]$. Таким образом, конкретный алгоритм можно записать в виде

$$x_{i+1} = (2^{40}x_i + 5^{17}) \bmod 2^{40}. \quad (3.5.4)$$

Псевдослучайность получаемых чисел заключается также и в том, что при заданном x_0 на ЭВМ каждый раз получается одна и та же последовательность случайных чисел в отличие от реальных шумов, реализации которых различны при любом включении прибора. Если это может повлиять на результаты моделирования исследуемой системы, то в каждом новом варианте задачи нужно вводить новое x_0 или пропускать "вхолостую" некоторый интервал случайных чисел.

Получаемые случайные числа являются некоррелированными. Сказанное здесь о псевдослучайности и некоррелированности справедливо и в случае генерации случайных чисел с другими законами распределения.

Рассмотрим метод генерирования случайных чисел с произвольным заданным законом распределения. Эта задача решается путем нелинейного преобразования случайной последовательности x с плотностью распределения $f_1(x)$ в случайную последовательность y с плотностью распределения $f_2(y)$. Приведем связь [68] между законами распределения на входе и выходе нелинейного элемента с амплитудной характеристикой $y = \varphi(x)$. Пусть существует обратная однозначная функция $x = \psi(y) = \varphi^{-1}(y)$. Тогда из неравенства

$$x_0 < x \leq x_0 + dx$$

следует неравенство [для монотонно возрастающей функции $\varphi(x)$]

$$y_0 < y \leq y_0 + dy,$$

где $y_0 = \varphi(x_0)$.

При однозначном соответствии вероятности выполнения этих неравенств равны, и при достаточно малых dx и dy можно записать, опуская индекс "0",

$$f_1(x) dx = f_2(y) dy, \quad (3.5.5)$$

откуда

$$f_2(y) = f_1(x) \frac{dx}{dy}. \quad (3.5.6)$$

Так как всегда $f_1(x) \geq 0$ и $f_2(y) \geq 0$ независимо от вида функции $y = \varphi(x)$, то в (3.5.6) нужно поставить знак модуля, т.е.

$$f_2(y) = f_1(x) \left| \frac{dx}{dy} \right|, \quad (3.5.7)$$

где $dx/dy = \psi'(y)$ – производная от обратной функции.

Пусть теперь входная последовательность x равномерно распределена на интервале $[0, 1]$, тогда $f_1(x) = 1$. Плотность распределения на выходе нелинейности будет

$$f_2(y) = \frac{dx}{dy} = \psi'(y) = [\varphi^{-1}(y)]'. \quad (3.5.8)$$

Здесь знак модуля опущен, ибо рассматриваем монотонно возрастающую функцию ψ , т.е. $\frac{dx}{dy} > 0$. Но $f_2(y) = F_2'(y)$, где $F_2(y)$ – закон распределения случайной величины y . Следовательно,

$$F_2'(y) = [\varphi^{-1}(y)]' \quad (3.5.9)$$

или с точностью до постоянного слагаемого

$$F_2(y) = \varphi^{-1}(y). \quad (3.5.10)$$

Формулу (3.5.10) можно записать в другом виде:

$$\varphi(x) = F_2^{-1}(x). \quad (3.5.11)$$

Таким образом, для получения случайных чисел с законом распределения $F_2(y)$ нужно пропустить последовательность случайных чисел с равномерным распределением на интервале $[0, 1]$ через нелинейный элемент с характеристикой $\varphi(x) = F_2^{-1}(x)$, равной обратной функции требуемого закона распределения.

Этот метод широко используется для получения случайных чисел с заданным законом распределения. Например, для получения случайных чисел с релеевским законом распределения, плотность которого имеет вид

$$f(y) = \frac{y}{\sigma^2} e^{-\frac{y^2}{2\sigma^2}}, \quad (3.5.12)$$

а закон распределения – вид

$$F(y) = 1 - e^{-\frac{y^2}{2\sigma^2}}, \quad (3.5.13)$$

нужно равномерно распределенные числа пропустить через нелинейный элемент с характеристикой

$$\varphi(x) = \sigma \sqrt{-2 \ln x} = \sigma \sqrt{-2 \ln(1 - x)}. \quad (3.5.14)$$

Здесь учтено, что x и $1 - x$ имеют одинаковое равномерное распределение на интервале $[0, 1]$.

Данный метод является универсальным и точным, если нелинейное преобразование выражается аналитически. Это можно сделать не всегда. Так, для получения этим методом случайных чисел с нормальным законом распределения характеристика нелинейного элемента имеет вид

$$\varphi(x) = \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt \right)^{-1}, \quad (3.5.15)$$

не выражающийся аналитически, и ее приходится задавать таблично, используя таблицы нормального закона (поменяв местами аргумент и функцию). Поэтому для генерирования случайных чисел с нормальным законом распределения применяют специальные методы, один из которых состоит в следующем.

Известно, что модуль вектора, координаты которого распределены по нормальному закону, имеет релеевское распределение. Тогда, используя (3.5.14), генерируют релеевские случайные числа u и затем производят разложение на две составляющие

$$y_1 = u \cos \varphi, \quad y_2 = u \sin \varphi, \quad (3.5.16)$$

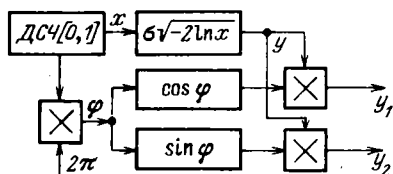
где фаза φ распределена равномерно на интервале $[0, 2\pi]$ и получается из равномерно распределенных на интервале $[0, 1]$ чисел умножением на 2π . Таким образом, здесь сразу получаются два нормально распределенных числа. Блок-схема генерирования случайных чисел с нормальным законом распределения приведена на рис. 3.23, где два датчика случайных чисел для x и φ изображены в виде одного (ДСЧ).

В технических приложениях часто используются случайные числа с пуассоновским законом распределения. Распределение Пуассона описывается формулой

$$P_m = \frac{a^m}{m!} e^{-a}, \quad (3.5.17)$$

которая задает вероятность того, что дискретная положительная целочисленная случайная величина примет значение, равное m . Здесь a — параметр закона, равный его среднему значению.

Рис. 3.23. Блок-схема генерирования случайных чисел с нормальным законом распределения



При математическом моделировании используется свойство закона Пуассона как закона редких событий [32]. Выбирают такое m , чтобы величина $p = a/m$ была достаточно малым числом, например $p = 0,05$. Затем моделируют серии по $m = pa$ независимых испытаний, в каждом из которых некоторое событие имеет вероятность p . Это можно получить, используя датчик случайных чисел с равномерным законом распределения на интервале $[0, 1]$. Тогда вероятность попадания, например, в интервал $[0,5-0,55]$ равна 0,05, так как вероятность равна длине интерва-

ла при равномерном распределении. Количество случаев появления этого события в каждой серии опытов есть дискретная случайная величина, распределенная по закону Пуассона.

Таким образом, во всех рассмотренных случаях исходным является датчик случайных чисел с равномерным законом распределения.

Нормальные стационарные случайные процессы. В зависимости от вида закона распределения методы генерирования случайных процессов можно разделить на две группы: методы моделирования случайных процессов с нормальным законом распределения и методы генерирования случайных процессов с произвольным законом распределения амплитуд.

Генерирование стационарных нормальных случайных процессов наиболее просто производится методом формирующего фильтра. При этом используется известная связь между спектрами на входе и выходе линейного фильтра [68]

$$S_y(\omega) = |W(j\omega)|^2 S_x(\omega), \quad (3.5.18)$$

где $S_x(\omega)$ – энергетический спектр случайного процесса на входе фильтра; $S_y(\omega)$ – энергетический спектр случайного процесса на выходе фильтра; $W(j\omega)$ – частотная характеристика фильтра.

Если на вход фильтра подается белый шум, у которого

$$S_x(\omega) = N_0 = 1,$$

то

$$S_y(\omega) = |W(j\omega)|^2. \quad (3.5.19)$$

Таким образом, спектр случайного процесса на выходе линейного фильтра, когда на его вход подан белый шум, равен квадрату модуля частотной характеристики.

Выражение (3.5.19) используется для вычисления частотной характеристики фильтра. При этом учитывают то обстоятельство, что для физически реализуемых систем частотная характеристика является дробно-рациональной функцией вида (3.4.5), а спектр описывается отношением полиномов четных степеней (3.5.1), которое можно разложить на множители:

$$S_y(\omega) = W(j\omega)W(-j\omega) = \frac{W_1(j\omega)W_1(-j\omega)}{W_2(j\omega)W_2(-j\omega)} = \frac{S_1(\omega)}{S_2(\omega)}. \quad (3.5.20)$$

Далее находятся корни числителя $j\omega_{1k}$ и знаменателя $j\omega_{2k}$ в выражении для спектра и берутся только те из них, которые имеют отрицательную мнимую часть. Тогда частотная характеристика фильтра может быть представлена в виде (метод факторизации) [20]

$$W(j\omega) = c \frac{\prod_{k=1}^l (j\omega - \omega_{1k})}{\prod_{k=1}^m (j\omega - \omega_{2k})}, \quad (3.5.21)$$

где константа c выбирается из условия (3.5.19).

Перечень типовых случайных процессов и соответствующих формирующих фильтров приведен в [125]. Реализация формирующего фильтра на ЭВМ производится одним из методов, рассмотренных в § 3.4.

Необходимо отметить следующее относительно входного случайного процесса. Непрерывный белый шум, т.е. случайный процесс с равномерным спектром на всех частотах, является математической абстракцией и не существует в природе, так как его дисперсия (площадь под спектром) бесконечна, а следовательно, бесконечна и мощность. Практически при реализации метода формирующего фильтра нужно подавать на вход случайный процесс с равномерным спектром только в диапазоне частот пропускания частотной характеристики $W(j\omega)$.

При цифровом моделировании на вход фильтра подаются некоррелированные случайные величины с нормальным законом распределения, т.е. дискретный белый шум, ширина полосы которого зависит от шага дискретизации по времени Δt и определяется теоремой В.А. Котельникова:

$$f_{\text{в}} = \frac{1}{2\Delta t}. \quad (3.5.22)$$

При моделировании шаг Δt выбирается таким образом, чтобы ширина спектра входного процесса была много больше полосы пропускания фильтра. Тогда такой процесс воспринимается фильтром как белый шум, что эквивалентно условию

$$\Delta t \ll T_{\text{ф}}, \quad (3.5.23)$$

т.е. шаг счета должен быть существенно меньше постоянной времени фильтра (это же условие отмечалось в § 3.4).

Для облегчения нормировки дисперсия входного процесса выбирается из условия $N_0 = 1$. Запишем

$$\sigma_x^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_x(\omega) d\omega = \frac{1}{2\pi} N_0 2\omega_{\text{в}} = \frac{N_0 \omega_{\text{в}}}{\pi}. \quad (3.5.24)$$

Подставив в (3.5.24) значение $\omega_{\text{в}}$ из (3.5.22)

$$\omega_{\text{в}} = 2\pi f_{\text{в}} = \frac{\pi}{\Delta t},$$

получим

$$\sigma_x^2 = \frac{N_0}{\Delta t}.$$

При $N_0 = 1$ между дисперсией входного случайного процесса и шагом счета имеет место соотношение

$$\sigma_x^2 \Delta t = 1, \quad (3.5.25)$$

обеспечивающее нормировку выражения (3.5.19).

Стационарные случайные процессы с произвольным законом распределения. Рассмотренный метод использует свойство линейного фильтра не менять нормальность распределения случайного процесса, т.е. поданный на вход нормальный процесс на выходе также остается нормальным. Использование этого метода для генерирования случайных процессов с законами, отличающимися от нормальных, в прямом виде невозможно, так как выходной процесс нормализуется и тем сильнее, чем больше инерционность фильтра [125]. Поэтому на вход фильтра нужно по-прежнему подавать нормальный случайный процесс,

а выходной процесс подвергать нелинейному преобразованию для получения требуемого закона распределения. При этом необходимо учитывать, что нелинейное преобразование изменяет корреляционную функцию (или спектр) случайного процесса, и вносить соответствующую поправку в частотную характеристику фильтра. Блок-схема генерирования случайного процесса с произвольным законом распределения и заданной корреляционной функцией приведена на рис. 3.24.

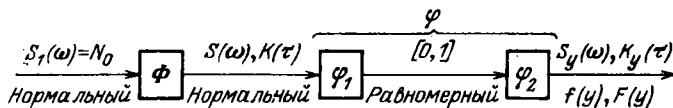


Рис. 3.24. Блок-схема генерирования случайного процесса с произвольным законом распределения

На вход фильтра Φ поступает некоррелированный нормальный случайный процесс. На выходе фильтра получаем нормальный случайный процесс с корреляционной функцией $K(\tau)$ и спектром $S(\omega)$ в соответствии с частотной характеристикой фильтра. Нелинейный элемент φ_1 преобразовывает случайный процесс с нормальным распределением в случайный процесс с равномерным распределением, нелинейный элемент φ_2 преобразовывает равномерное распределение в заданное $f(y)$, но при этом изменяется и спектральный состав случайного процесса. Задача состоит в том, чтобы для заданных характеристик $K_Y(\tau)$, $f(y)$ получить вид нелинейного преобразования, рассчитать промежуточные характеристики $S(\omega)$, $K(\tau)$ и найти частотную характеристику фильтра.

Найдем необходимое нелинейное преобразование. Если на вход второго нелинейного элемента подать случайный процесс с равномерным распределением, то согласно (3.5.11) его амплитудная характеристика задается выражением

$$\varphi_2(x) = F^{-1}(x). \quad (3.5.26)$$

Задача первого нелинейного элемента — преобразовать нормальное распределение после фильтра в равномерное на входе второго нелинейного элемента. Амплитудная характеристика в этом случае имеет вид функции распределения нормального закона

$$\varphi_1(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt \quad (3.5.27)$$

и задается таблично. На выходе элемента φ_1 получаем равномерное распределение на интервале $[0, 1]$. Оба нелинейных преобразования могут быть заменены одним

$$\varphi(x) = \varphi_2[\varphi_1(x)]. \quad (3.5.28)$$

Теперь по известному нелинейному преобразованию $\varphi(x)$ и заданным статистическим характеристикам на выходе нужно определить статистические характеристики на входе (спектр и корреляционную функцию). Для нормальных случайных процессов корреляционные функции на входе

и выходе нелинейного элемента связаны соотношением [68]

$$R_y(\tau) = \sum_{n=0}^{\infty} c_n^2 \frac{R^n(\tau)}{n!}, \quad (3.5.29)$$

где $R(\tau)$, $R_y(\tau)$ – нормированные по дисперсии корреляционные функции, а коэффициенты c_n вычисляются по формуле

$$c_n = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \varphi(x) H_n(x) e^{-\frac{x^2}{2}} dx, \quad (3.5.30)$$

где $H_n(x)$ – полиномы Эрмита. Сумма всех коэффициентов при степенях корреляционной функции равна единице, т.е.

$$\sum_{n=0}^{\infty} \frac{c_n^2}{n!} = 1. \quad (3.5.31)$$

Коэффициенты c_n в общем случае не вычисляются аналитически и находятся численным интегрированием, при этом бесконечная сумма (3.5.29) заменяется конечной суммой. В качестве номера n , на котором следует остановить процесс вычисления коэффициентов, можно взять то значение $n = k$, при котором выполняется неравенство

$$1 - \sum_{n=0}^k \frac{c_n^2}{n!} \leq \epsilon, \quad (3.5.32)$$

где ϵ – достаточно малая величина, например 0,05.

Таким образом, для нахождения промежуточной корреляционной функции $R(\tau)$ получаем следующее уравнение:

$$R_y(\tau) = c_0^2 + c_1 R(\tau) + \frac{c_2^2}{2} R^2(\tau) + \dots + \frac{c_k^2}{k!} R^k(\tau). \quad (3.5.33)$$

Решив это уравнение относительно $R(\tau)$, найдем корреляционную функцию на выходе фильтра. Решение производится на ЭВМ для последовательных значений $\tau = 0, \Delta t, 2\Delta t, \dots$, при этом получаем обычное алгебраическое уравнение степени k . Как показывают практические вычисления, коэффициенты в (3.5.29) быстро убывают и в большинстве случаев бывает достаточно ограничиться $k = 3$. Ввиду численного решения функция $R(\tau)$ получается в виде таблицы или графика, который может быть аппроксимирован подходящим аналитическим выражением.

Энергетический спектр процесса на выходе фильтра находится преобразованием Фурье корреляционной функции

$$S(\omega) = \int_{-\infty}^{\infty} R(\tau) e^{-j\omega\tau} d\tau. \quad (3.5.34)$$

Теперь можно методом факторизации (3.5.21) вычислить частотную характеристику фильтра и провести генерирование случайного процесса по блок-схеме, приведенной на рис. 3.24.

Функция распределения случайного процесса, который необходимо получить, может быть задана различными аналитическими выражениями. Если плотность распределения не очень сильно отличается от нормальной (является одновершинной и быстро спадает к нулю по обе стороны от

вершины), то такую плотность распределения удобно аппроксимировать рядом по производным функции распределения нормального закона, так называемым *рядом Эджворта* [125]:

$$\begin{aligned} f(y) &= \Phi^{(1)}(y) - \frac{\gamma_1}{3!} \Phi^{(4)}(y) + \frac{\gamma_3^2}{4!} \Phi^{(5)}(y) + \dots, \\ F(y) &= \Phi(y) - \frac{\gamma_1}{3!} \Phi^{(3)}(y) + \frac{\gamma_3^2}{4!} \Phi^{(4)}(y) + \dots, \end{aligned} \quad (3.5.35)$$

где $\Phi(y)$ — функция распределения нормального закона; γ_1 — коэффициент асимметрии; γ_3 — коэффициент эксцесса.

Таблицы производных $\Phi^{(n)}(y)$ приведены, например, в [125]. В этом случае оба нелинейных преобразования φ_1 и φ_2 задаются таблично и аппроксимируются подходящим аналитическим выражением.

В качестве примера рассмотрим генерирование на ЭВМ случайного процесса с корреляционной функцией вида

$$R_y(\tau) = e^{-\alpha \tau^3} \quad (3.5.36)$$

и законом распределения, заданным выражением (3.5.35), с параметрами $\alpha = 4$, $\gamma_1 = 2$, $\gamma_3 = 0$.

Корреляционная функция вида (3.5.36) соответствует физически нереализуемому так называемому *сингулярному процессу*, но часто используется как аппроксимация корреляционной функции бесконечно дифференцируемых случайных процессов. Закон распределения вида (3.5.35) описывает распределение яркости природных ландшафтов.

Построенная таблично, нелинейность $y = \varphi(x)$ аппроксимировалась аналитическим выражением

$$y = 0,35e^{0,53(x+3)} - 2,$$

справедливым в пределах, где плотность $f(y)$ заметно отличается от нуля.

После вычисления на ЭВМ по формуле (3.5.30) коэффициентов разложения получено уравнение для корреляционной функции

$$R_y(\tau) = 0,33 \cdot 10^{-4} + 0,65R(\tau) + 0,17R^2(\tau) + 0,57 \cdot 10^{-3}R^3(\tau) + \dots$$

Ограничиваясь только двумя членами с $R(\tau)$ и $R^2(\tau)$, получаем квадратное уравнение

$$0,17R^2(\tau) + 0,65R(\tau) - R_y(\tau) = 0.$$

Численное решение этого уравнения с шагом Δt дает график корреляционной функции, хорошо аппроксимируемый выражением

$$R(\tau) = e^{-\alpha_1 \tau^3} \quad (\alpha_1 = 3,5).$$

Таким образом, на выходе фильтра нужно получить также гауссов процесс. Соответствующий энергетический спектр равен

$$S(\omega) = \int_{-\infty}^{\infty} R(\tau) e^{-j\omega\tau} d\tau = \int_{-\infty}^{\infty} e^{-\alpha_1 \tau^3} e^{-j\omega\tau} d\tau = \sqrt{\frac{\pi}{\alpha_1}} \exp\left(-\frac{\omega^2}{4\alpha_1}\right).$$

В данном случае спектр не является дробно-рациональной функцией частоты, что соответствует физически нереализуемому фильтру. Частотная

характеристика такого фильтра находится на основании (3.5.19)

$$W(j\omega) = \sqrt{S(\omega)} = \sqrt[4]{\frac{\pi}{\alpha_1}} \exp\left(-\frac{\omega^2}{8\alpha_1}\right). \quad (3.5.37)$$

Реализация на ЭВМ такого фильтра возможна методом дискретной свертки.

Моделирование двумерных случайных полей. Многоуровневые случайные поля. Генерирование на ЭВМ двумерных случайных полей может быть произведено теми же методами, что и для случайных процессов. Задача усложняется тем, что формирующий фильтр в данном случае является двумерным. Рассмотрим определение импульсной характеристики двумерного фильтра [141]. Для изотропного поля корреляционная функция и спектр являются функциями одной переменной:

$$K(r) = K(\sqrt{x^2 + y^2}), \quad S(k) = S(\sqrt{u^2 + v^2}), \quad (3.5.38)$$

где r — модуль сдвига; k — модуль пространственной частоты; u, v — пространственные частоты по координатам x, y . Тогда выражение для спектра на выходе фильтра (3.5.19) запишется в виде

$$S(k) = W^2(k), \quad (3.5.39)$$

т.е. спектр поля на выходе двумерного фильтра, когда на его входе подается белый шум, равен квадрату частотной характеристики [поскольку импульсная характеристика двумерного фильтра является четной функцией, см. (3.2.8)]. Поэтому если задан спектр изотропного поля, то частотная характеристика двумерного фильтра находится просто, без разложения на множители:

$$W(k) = \sqrt{S(k)}, \quad (3.5.40)$$

т.е. так же, как для гауссова фильтра в предыдущем примере.

- Импульсная характеристика фильтра связана с частотной характеристикой преобразованием Фурье (3.4.8), которая для двумерных изотропных полей переходит в преобразование Ганкеля [81]

$$g(r) = \frac{1}{2\pi} \int_0^\infty \sqrt{S(k)} J_0(kr) k dk \quad (3.5.41)$$

Моделирование случайного поля лучше всего проводить методом свертки двумерного белого шума с импульсной характеристикой фильтра аналогично преобразованию оптической системой размерного источника (3.2.27)

$$\xi_2(x, y) = \iint_{-\infty}^{\infty} \xi_1(x', y') g(x - x', y - y') dx' dy', \quad (3.5.42)$$

где $\xi_1(x, y)$ — двумерный белый шум на входе фильтра; $\xi_2(x, y)$ — "окрашенное" двумерное изотропное поле.

При расчете на ЭВМ интеграл заменяется суммой, аналогичной (3.2.28):

$$\xi_2(x_i, y_j) = \sum_{\alpha} \sum_{\beta} \xi_1(x_{\alpha}, y_{\beta}) g(x_i - x_{\alpha}, y_j - y_{\beta}) \Delta x \Delta y, \quad (3.5.43)$$

где пределы суммирования α и β определяются шириной импульсной ха-

рактеристики фильтра $r_{\max}$, т.е.

$$\alpha_{\min} = \beta_{\min} = -\frac{r_{\max}}{\Delta}, \quad \alpha_{\max} = \beta_{\max} = \frac{r_{\max}}{\Delta}.$$

Здесь $\xi_1(x_\alpha, y_\beta)$ — матрица некоррелированных случайных чисел с нормальным законом распределения. В результате получается изотропное случайное поле с нормальным законом распределения амплитуд. Для получения анизотропного поля найденную реализацию можно растянуть по одной из осей x или y изменением масштаба.

Найдем вид импульсной характеристики формирующего фильтра для некоторых полей. Наиболее часто при математическом моделировании используют описание корреляционных функций изотропных полей следующими выражениями:

$$K(r) = e^{-\alpha r^2}, \quad (3.5.44)$$

$$K(r) = e^{-\beta r}. \quad (3.5.45)$$

Используя преобразование Ганкеля, получим соответствующие энергетические спектры

$$S(k) = \exp\left(-\frac{k^2}{4\alpha}\right), \quad (3.5.46)$$

$$S(k) = (\beta^2 + k^2)^{-3/2}. \quad (3.5.47)$$

По формуле (3.5.41) находим импульсные характеристики соответствующих формирующих фильтров

$$g(r) = \exp(-2\alpha r^2), \quad (3.5.48)$$

$$g(r) = r^{-1/4} K_{1/4}(\beta r), \quad (3.5.49)$$

где $K_{1/4}(\beta r)$ — цилиндрическая функция порядка $1/4$.

Моделирование фильтра с импульсной характеристикой вида (3.5.48) производится просто с использованием метода дискретной свертки. При вычислении на ЭВМ выражение (3.5.49) аппроксимируется простой зависимостью, чтобы не вычислять специальную функцию. Для этого (3.5.49) представим в виде

$$g(r) = \beta^{-1/4} r^{-1/4} \beta^{1/4} K_{1/4}(\beta r) = \beta^{1/4} z^{-1/4} K_{1/4}(z), \quad (3.5.50)$$

где $z = \beta r$.

Функция

$$g_1(z) = z^{-1/4} K_{1/4}(z) \quad (3.5.51)$$

хорошо аппроксимируется выражением

$$g_1(z) = \frac{1}{z + 0,05} - 0,4. \quad (3.5.52)$$

до $z_{\max} = 2$. Тогда максимальная ширина импульсной характеристики при вычислении суммы (3.5.43) равна $r_{\max} = 2/\beta$.

Рассмотренный метод позволяет получить двумерное поле с нормальным законом распределения и произвольной корреляционной функцией. Генерирование полей с произвольным законом распределения амплитуд может быть осуществлено методом нелинейного преобразования, изло-

женным выше, учитывая, что амплитуда поля — скаляр. Существуют методы, позволяющие генерировать поля с произвольным законом распределения и экспоненциальной корреляционной функцией [20]. Эти методы основаны на описании полей двумерными марковскими процессами.

Изотропное поле при моделировании работы ОЭСС может быть преобразовано в анизотропное изменением масштаба по координатам x и y . Существуют более сложные методы генерирования неизотропных двумерных полей с заданными корреляционными свойствами [178].

Использование метода дискретной свертки для генерирования двумерных полей означает, по существу, представление формирующего фильтра в виде нерекурсивного двумерного фильтра. Недостатком нерекурсивных фильтров, как уже отмечалось в § 3.4, является большой объем вычисления. В принципе для генерации полей возможно использование рекурсивных двумерных фильтров [87], но при этом возникает проблема обеспечения устойчивости, которая решается гораздо сложнее, чем для одномерных рекурсивных фильтров. Отметим, что вычисление двумерной дискретной свертки может быть произведено с использованием алгоритмов двумерного быстрого преобразования Фурье [148], но при этом значительно усложняется программа генерации поля и требуется большой объем оперативной памяти ЭВМ.

Двухуровневые случайные поля. В ряде случаев на предварительных этапах исследования оптико-электронных систем можно ограничиться описанием входных сигналов более простыми полями, например двухуровневыми, принимающими значения 0 и 1 или -1 и $+1$. Метод генерирования двумерных случайных полей с корреляционной функцией вида

$$K(r) = e^{-\beta r} \quad (3.5.53)$$

впервые был предложен в работе [37].

Рассмотрим случайную функцию $\xi(x)$, образованную сечением поля по координате x . Эта функция принимает только два значения $+1$ и -1 , а моменты перемены знака случайные (так называемый *телеграфный сигнал* [68]). Пусть случайное число перемен знака на интервале $\Delta x = x_2 - x_1$ распределено по закону Пуассона

$$P_m = \frac{a^m}{m!} e^{-a}, \quad (3.5.54)$$

где a — среднее число перемен знака на этом интервале, зависящее от его длины, т.е. $a = a(\Delta x)$.

Найдем корреляционную функцию $\xi(x)$ как среднее значение произведения

$$K(\Delta x) = M \{ \xi(x_1) \xi(x_2) \}. \quad (3.5.55)$$

Произведение $\xi(x_1)\xi(x_2)$ может быть равно либо 1, либо -1 с вероятностями P_1 и P_2 в зависимости от того, что

$$\xi(x_1) = \xi(x_2) \quad \text{или} \quad \xi(x_1) = -\xi(x_2). \quad (3.5.56)$$

Первое равенство означает, что на интервале Δx произошло четное число перемен знаков, т.е. $m \neq 0, 2, 4, \dots$. Вероятность этого события равна

$$P \{ \xi(x_1) = \xi(x_2) \} = \sum_{m=0}^{\infty} P_{2m} = e^{-a} \sum_{m=0}^{\infty} \frac{a^{2m}}{(2m)!} = e^{-a} \operatorname{ch} a. \quad (3.5.57)$$

Второе равенство (3.5.56) означает, что на интервале Δx произошло нечетное число перемен знаков, т.е. $m = 1, 3, \dots$. Вероятность этого события равна

$$P = \{ \xi(x_1) = -\xi(x_2) \} = \sum_{m=1}^{\infty} P_{2m-1} = e^{-a} \sum_{m=1}^{\infty} \frac{a^{2m-1}}{(2m-1)!} = e^{-a} \operatorname{sh} a. \quad (3.5.58)$$

Следовательно, корреляционная функция (3.5.55) равна

$$K(x) = 1 \cdot P_1 + (-1)P_2 = e^{-a}(\operatorname{ch} a - \operatorname{sh} a) = e^{-a}e^{-a} = e^{-2a}. \quad (3.5.59)$$

Введем интенсивность перемен знака на единицу длины, равную λ , тогда среднее число перемен знака на отрезке $[0, x]$ будет

$$a = \lambda \cdot \Delta x. \quad (3.5.60)$$

Но так как $a > 0$ всегда, то можно записать (обозначив $\Delta x = x$)

$$a = \lambda |x|. \quad (3.5.61)$$

Тогда корреляционная функция телеграфного сигнала равна

$$K(x) = e^{-2\lambda|x|} = e^{-\beta|x|}, \quad (3.5.62)$$

где $\beta = 2\lambda$.

Аналогичная корреляционная функция описывает сечение поля по координате y , т.е.

$$K(y) = e^{-2\lambda|y|} = e^{-\beta|y|}. \quad (3.5.63)$$

Двумерная корреляционная функция равна

$$K(x, y) = e^{-\beta_1|x|}e^{-\beta_2|y|}, \quad (3.5.64)$$

так как в общем случае поле может иметь разные свойства в направлениях x и y .

В данном случае получается неизотропное случайное поле даже при $\beta_1 = \beta_2 = \beta$, а заданные статистические свойства получаются лишь в направлении координатных осей.

Генерирование случайного поля производится следующим образом. С датчика случайных чисел получают числа ξ_1 и ξ_2 , распределенные по закону Пуассона с параметром a . Эти случайные величины показывают, на какое количество интервалов разбивается каждая сторона поля в направлении координат x и y (т.е. количество перемен знака по каждой координате).

Для определения координат смены знаков используется свойство закона Пуассона, по которому вероятность попадания числа точек на отрезок l зависит только от длины этого отрезка, но не от его положения на числовой оси. Иными словами, точки располагаются с одинаковой средней плотностью. (Это свойство использовалось для генерации случайных величин с пуассоновским распределением.) Поэтому далее обращаются к датчику случайных чисел с равномерным распределением в интервале $[0, 1]$ и получают две серии чисел длиной ξ_1 и ξ_2 . Числа каждой серии, умноженные на размер поля по каждой стороне Δx и Δy , дают координаты точек смены знака в направлении x и y .

Через полученные точки проводятся линии, параллельные координатным осям; таким образом производится разбивка поля на прямоугольники, которые закрашиваются в шахматном порядке. При этом четным интервалам присваивается значение +1, нечетным —1. Прямоугольники, соответствующие интервалам одного знака, считают с яркостью, равной единице, с разными знаками — с яркостью, равной нулю.

Такое двухуровневое поле, конечно, беднее многоуровневого, но может быть использовано при решении ряда задач. Его преимущество состоит в быстром методе генерирования и небольшой памяти для хранения, так как здесь не нужно помнить яркости всех точек поля, достаточно записать координаты точек смены знака на каждой из осей.

Визуализация двумерных полей. В практике математического моделирования часто требуется наглядное представление полученной реализации случайного поля, т.е. его визуализация. Такая необходимость возникает при трактовке результатов моделирования, для наглядного представления работы ОЭС, для оформления отчетных материалов. Существуют специализированные вычислительные комплексы обработки изображений, позволяющие выводить двумерные поля на дисплеи или представлять их в графической форме. Однако универсальные ЭВМ, наиболее приспособленные для математического моделирования, такими устройствами не оснащаются. Поэтому приходится пользоваться штатными устройствами вывода информации. При этом используются разные методы.

Двухуровневое поле просто рисуют по известным координатам точек перемены знака, для чего достаточно вывести на печать случайные числа по осям x и y .

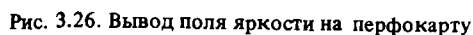
Многоуровневое поле может быть распечатано на АЦПУ символами различной степени черноты. При этом полный размах амплитуд поля разбивается на несколько градаций по яркости и каждая градация печатается определенным символом. Проведенное фотометрирование символов АЦПУ дает значения, приведенные в табл. 3.5. Можно использовать и другие симво-

Таблица 3.5

Нормированная яркость	$1,0 \div 0,8$	$0,8 \div 0,7$	$0,7 \div 0,6$	$0,6 \div 0,5$	$0,5 \div 0,4$
Символ	.	,	-	=	

Нормированная яркость	$0,4 \div 0,3$	$0,3 \div 0,2$	$0,2 \div 0,1$	$0,1 \div 0$
Символ	≠	И	Ы	Ж

Рис. 3.25. Вывод поля яркости на алфавитно-цифровое печатающее устройство



лы, а также комбинацию из нескольких символов, т.е. печатать на одном месте несколько символов, но это сильно усложняет программы вывода на печать. На рис. 3.25 приведена распечатка многоуровневого нормального поля.

В другом методе многоуровневое поле выводится на перфокарты. Здесь полный диапазон яркостей разбивается на семь интервалов, минимальной яркости соответствует отсутствие пробивки, максимальной — шесть отверстий перфорации. Каждому элементу поля соответствует определенный участок перфокарты. При большом размере поля оно выводится на несколько перфокарт, которые затем склеиваются и фотографируются. Полученное таким методом изображение двумерного поля приведено на рис. 3.26.

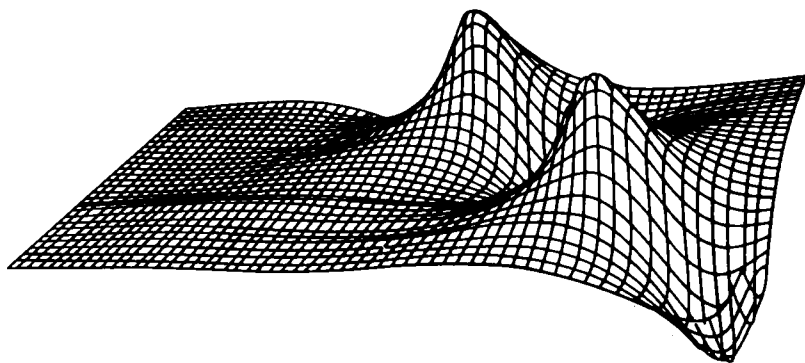


Рис. 3.27. Построение поля яркости в изометрии

Третий метод состоит в использовании графопостроителя для построения рисунка амплитуд поля в изометрии. Однако этот метод для оптических полей используется довольно редко в связи со сложностью программ и менее наглядным представлением поля оптической яркости в виде пространственного рельефа. На рис. 3.27 представлено изображение двумерного поля яркости, полученное на графопостроителе.

Наиболее распространенным методом в настоящее время является печать двумерного поля на АЦПУ.

Таким образом, рассмотренные в этом разделе методы позволяют моделировать любые флуктуации сигнала, встречающиеся при исследовании оптико-электронных систем.

§ 3.6. Построение математической модели по экспериментальным данным

Оценки параметров модели. В предыдущих параграфах рассмотрено построение математических моделей оптико-электронных систем на основании известных принципов их функционирования. Таким образом строятся поэлементные функциональные модели. В практике часто бывают такие ситуации, когда функциональное устройство объекта неизвестно или очень сложно. В таких случаях объект рассматривают как "черный ящик", у которого доступны наблюдению только входные и выходные сигналы. По результатам наблюдений этих сигналов нужно построить математиче-

скую модель объекта, т.е. дать его математическое описание, не касаясь механизма его функционирования. Кроме того, экспериментальные наблюдения над объектом нужно провести таким образом, чтобы при минимуме опытов получить максимум информации, т.е. описать объект с наименьшей возможной погрешностью.

В качестве объекта может быть использована поэлементная модель оптико-электронной системы. Тогда проводится так называемый математический эксперимент для получения интересующих зависимостей.

Рассмотрим объект, выход которого зависит от ряда величин $x_1, \dots, x_n$. Для оптико-электронных систем это могут быть коэффициенты усиления, постоянные времени, внешние возмущения. Выходной сигнал y может быть выходным сигналом по угловой скорости, или быстродействием системы, или какой-либо другой физической величиной в зависимости от целей исследования.

В дальнейшем переменные $x_1, \dots, x_n$ будем называть *входными контролируемыми* или *независимыми* переменными. Выходную переменную y будем называть *зависимой переменной* или *выходом объекта*, даже если она не обозначает буквально выход системы. В общем случае можно сказать, что между независимыми переменными и выходом объекта существует функциональная зависимость

$$y = y(x), \quad (3.6.1)$$

где $x = (x_1, x_2, \dots, x_n)^T$ — вектор-столбец значений независимых переменных; t — символ транспонирования.

Зависимость $y = y(x)$ неизвестна, ее нужно найти путем обработки экспериментальных данных. Так как всякий эксперимент связан с появлением случайных ошибок, то при построении математической модели на основе экспериментальных данных необходимо использовать методы математической статистики.

Наиболее часто при решении этой задачи применяют метод наименьших квадратов, позволяющий построить оптимальную модель объекта в смысле минимума ошибки оцениваемых параметров [84, 92]. Отметим, что экспериментальные данные не могут дать ответ на вопрос о виде модели, поэтому математическая модель объекта обычно задается в виде полинома, вид которого выбирается из имеющихся физических представлений об объекте. Например, линейная модель имеет вид

$$y(a, x) = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_n x_n, \quad (3.6.2)$$

где a_i являются неизвестными параметрами, оценки которых требуется найти путем обработки экспериментальных данных. В более сложных случаях используются квадратичные модели

$$y(a, x) = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_n x_n + a_{11} x_1^2 + a_{22} x_2^2 + \dots + a_{nn} x_n^2 + \\ + a_{12} x_1 x_2 + a_{13} x_1 x_3 + \dots + a_{kn} x_{n-1} x_n, \quad (3.6.3)$$

где

$$k = \frac{(n+2)(n+1)}{2}.$$

Реже применяются модели более высокого порядка. Модели полиномиального вида имеют большое значение в связи с тем, что с их помощью любая аналитическая функция $y(a, x)$ может быть описана как угодно точно.

Кратко модель будем записывать в виде $y(a, x)$, где $a = (a_0, a_1, \dots, a_k)^T$ — вектор-столбец параметров модели. В дальнейшем будем рассматривать модели, линейные по параметрам a_i , т.е. модели, которые можно записать в виде

$$y(a, x) = a_0 f_0(x) + a_1 f_1(x) + \dots + a_k f_k(x). \quad (3.6.4)$$

Здесь $f_i(x)$ — известные функции, являющиеся компонентами вектора

$$f(x) = [f_0(x), f_1(x), \dots, f_k(x)]^T. \quad (3.6.5)$$

Используя матричную форму записи, вместо (3.6.4) можно записать

$$y = a^T f(x) = f^T(x) a. \quad (3.6.6)$$

В линейной модели вида (3.6.2) получаем следующие значения компонент вектора $f(x)$:

$$f(x) = (1, x_1, x_2, \dots, x_n)^T, \quad (3.6.7)$$

а в квадратичной модели — значения

$$f(x) = (1, x_1, x_2, \dots, x_n, x_1^2, x_2^2, \dots, x_n^2, x_1 x_2, x_1 x_3, \dots, x_n x_{n-1})^T. \quad (3.6.8)$$

Для компонент вектора a будем различать действительные значения, которые мы не знаем, и оценки этих коэффициентов, которые будем искать, используя результаты эксперимента. Действительные значения обозначим $\bar{a}$, оценки — $\hat{a}$. В результате опытов мы не можем найти действительные значения, а получаем только оценки, так как всегда имеются ошибки эксперимента. Тогда для оценки вектора выхода можно написать уравнение, аналогичное (3.6.6):

$$\hat{y} = \hat{a}^T f(x) = f^T(x) \hat{a}. \quad (3.6.9)$$

Эксперименты проводятся в N экспериментальных точках, т.е., другими словами, проводится N различных экспериментов. Опытные точки будем обозначать верхними индексами: $x^1, x^2, \dots, x^N$, т.е. это суть значения вектора x в 1, 2, ..., N опытах. Каждая точка имеет координаты

$$x^i = (x_1^i, x_2^i, \dots, x_k^i)^T, \quad (3.6.10)$$

задающие вектор-столбец независимых переменных в i -м опыте. Количество координат k равно размерности уравнения модели или размерности пространства, в котором задано уравнение модели. Результаты наблюдений в точках x^i будем обозначать $\tilde{y}^i$, тогда получим вектор наблюдений

$$\tilde{Y} = (\tilde{y}^1, \tilde{y}^2, \dots, \tilde{y}^N)^T. \quad (3.6.11)$$

Для повышения точности в каждой точке x^i может быть поставлено ν параллельных опытов. В этом случае в качестве выхода в каждой точке используется среднее значение

$$\tilde{y}^i = \frac{1}{\nu} (\tilde{y}^{i1} + \tilde{y}^{i2} + \dots + \tilde{y}^{i\nu}). \quad (3.6.12)$$

Задача состоит в том, чтобы на основе наблюдений (3.6.11) найти наилучшие в определенном смысле оценки параметров модели $\hat{a}$ и выхода $\hat{y}$.

Рассмотрим, что следует понимать под наилучшими оценками. Для этого сопоставим друг с другом экспериментальные результаты (3.6.11) и значения

$$\hat{Y}(\hat{y}^1, \hat{y}^2, \dots, \hat{y}^N)^T, \quad (3.6.13)$$

рассчитанные с помощью оценок $\hat{a}$ и представляющие модель (3.6.9). Имеем для i -го опыта

$$\hat{y}^i = \hat{a}^T f(x^i) = f^T(x^i) \hat{a} \quad (3.6.14)$$

или короче

$$\hat{Y} = F \hat{a}, \quad (3.6.15)$$

где матрица F определяется следующим образом:

$$F = \|f_j(x^i)\| = \begin{vmatrix} f_0(x^1) & f_1(x^1) & \dots & f_k(x^1) \\ f_0(x^2) & f_1(x^2) & \dots & f_k(x^2) \\ \dots & \dots & \dots & \dots \\ f_0(x^N) & f_1(x^N) & \dots & f_k(x^N) \end{vmatrix} \quad (3.6.16)$$

Результат наблюдений $\tilde{y}^i$ в некоторой точке зависит от случайной ошибки $\tilde{e}^i$, равной

$$\tilde{e}^i = \tilde{y}^i - \bar{y}^i. \quad (3.6.17)$$

Множество значений ошибок в N экспериментальных точках может быть представлено вектором

$$\tilde{e} = \tilde{Y} - \bar{Y}. \quad (3.6.18)$$

Здесь черта сверху, как и выше, означает истинное значение соответствующей переменной.

Пусть результаты наблюдений удовлетворяют следующим условиям, которые, как правило, выполняются.

1. Результаты эксперимента не содержат систематических ошибок, т.е. математическое ожидание величины $\tilde{y}^i$ равно действительному

значению $\bar{y}^i$:

$$M\{\tilde{Y}\} = \bar{Y}. \quad (3.6.19)$$

Это эквивалентно требованию равенства нулю среднего значения ошибок измерений:

$$M\{\tilde{e}\} = 0. \quad (3.6.20)$$

2. Результат наблюдений в точке x^j не зависит от результата в точке x^i :

$$M\{(\tilde{y}^i - \bar{y}^i)(\tilde{y}^j - \bar{y}^j)\} = 0 \quad \text{для } i \neq j \quad (3.6.21)$$

или

$$M\{\tilde{e}^j \tilde{e}^i\} = 0 \quad \text{для } i \neq j. \quad (3.6.22)$$

3. Дисперсия результатов наблюдений во всех точках одинакова, т.е.

$$D\{\tilde{y}^i\} = \sigma^2 \quad \text{для всех } i \quad (3.6.23)$$

или

$$D\{\tilde{e}^i\} = \sigma^2 \quad \text{для всех } i. \quad (3.6.24)$$

Условия 2 и 3 выполняются, если

$$M\{\tilde{e} \tilde{e}^T\} = \sigma^2 I, \quad (3.6.25)$$

где I — единичная матрица.

Наложим еще два условия на оценки $\hat{a}$.

- 1. Оценка не должна содержать систематических ошибок, т.е. должна быть несмещенной:

$$M\{\hat{a}\} = \bar{a}. \quad (3.6.26)$$

2. Дисперсия оценки должна быть минимальной:

$$M\{(\hat{a}_i - \bar{a}_i)^2\} = \sigma_i^2 = \min. \quad (3.6.27)$$

Могут быть сформулированы и другие требования к оценкам, отличные от (3.6.26) и (3.6.27), но условие (3.6.27) является наиболее приемлемым, так как оно приводит к методу наименьших квадратов и при нормально распределенных ошибках наблюдений позволяет провести статистический анализ полученных оценок и проверить адекватность принятой модели.

Оценка $\hat{a}$ удовлетворяет условиям (3.6.26) и (3.6.27), если сумма

$$S_R = \sum_{i=1}^N (\tilde{y}^i - \hat{y}^i)^2 = |\tilde{Y} - \hat{Y}|^2 = (\tilde{Y} - \hat{Y})^T (\tilde{Y} - \hat{Y}) = \min. \quad (3.6.28)$$

Минимум выражения (3.6.28) находится путем приравнивания нулю част-

ных производных по параметрам $\hat{a}_i$ и решения системы нормальных уравнений. В матричной форме с учетом (3.6.15) можно записать

$$S_R = (\tilde{Y} - F\hat{a})^T(\tilde{Y} - F\hat{a}). \quad (3.6.29)$$

Выражение (3.6.29) имеет единственный минимум при

$$\hat{a} = (F^T F)^{-1} F^T \tilde{Y}, \quad (3.6.30)$$

что и является решением системы нормальных уравнений. Для существования решения необходимо существование обратной матрицы $(F^T F)^{-1}$, что имеет место, если матрица $F^T F$ не вырождена, т.е. ее определитель не равен нулю:

$$|F^T F| \neq 0. \quad (3.6.31)$$

Для этого число опытов должно быть не менее $k + 1$. Матрица $M = F^T F$ называется *информационной*, а матрица $C = M^{-1}$ — *дисперсионной*.

О ш и б к и о ц е н и в а н и я. Оценки $\hat{a}$, рассчитанные по (3.6.30), отличаются от истинных значений коэффициентов, причем ошибка тем больше, чем больше дисперсия наблюдений. Показателями точности оценок $\hat{a}_i$ и величины $\hat{y}$, рассчитанной по уравнению модели, являются дисперсии коэффициентов σ_i^2 и дисперсия рассчитанного выхода σ_y^2 . Эти дисперсии зависят не только от дисперсии ошибок наблюдений σ^2 , но и от выбранной структуры модели и точек постановки опытов, т.е. от матрицы F . При этом предполагается, что значения контролируемых величин x_i устанавливаются точно, т.е. $\sigma_{x_i}^2 \ll \sigma_y^2$.

Для определения этой зависимости найдем ковариационную матрицу

$$\text{cov}(\hat{a}) = M\{(\hat{a} - \bar{a})(\hat{a} - \bar{a})^T\}. \quad (3.6.32)$$

Из (3.6.30), (3.6.26), (3.6.20) следуют выражения

$$\hat{a} = CF^T \tilde{Y}, \quad (3.6.33)$$

$$\bar{a} = M\{\hat{a}\} = CF^T M\{\tilde{Y}\} = CF^T \bar{Y}. \quad (3.6.34)$$

Используя (3.6.34), получаем

$$\begin{aligned} M\{(\hat{a} - \bar{a})(\hat{a} - \bar{a})^T\} &= M\{CF^T(\tilde{Y} - \bar{Y})(\tilde{Y} - \bar{Y})^T FC^T\} = \\ &= CF^T M\{(\tilde{Y} - \bar{Y})(\tilde{Y} - \bar{Y})^T\} FC^T. \end{aligned} \quad (3.6.35)$$

В силу (3.6.18), (3.6.25) имеем из (3.6.35)

$$\text{cov}(\hat{a}) = CF^T FC^T \sigma^2. \quad (3.6.36)$$

Так как матрица $F^T F$ симметрическая, то

$$CF^T FC^T = (F^T F)^{-1} F^T F (F^T F)^{-1} = C$$

и, следовательно,

$$\text{cov}(\hat{a}) = C\sigma^2. \quad (3.6.37)$$

При этом, в частности, для дисперсии σ_i^2 оценки $\hat{a}_i$ имеем

$$\sigma_i^2 = c_{ii}\sigma^2, \quad (3.6.38)$$

где c_{ii} — диагональные элементы матрицы C .

Дисперсию σ_y^2 получаем с помощью следующих преобразований:

$$\begin{aligned} \sigma_y^2 &= M\{(\hat{y} - \bar{y})^2\} = M\{[(\hat{a} - \bar{a})f(x)]^2\} = \\ &= M\{f^T(x)(\hat{a} - \bar{a})(\hat{a} - \bar{a})^T f(x)\} = \\ &= f^T(x) M\{(\hat{a} - \bar{a})(\hat{a} - \bar{a})^T\} f(x) = \\ &= f^T(x) \text{cov}(\hat{a}) f(x) = f^T(x) C\sigma^2 f(x). \end{aligned} \quad (3.6.39)$$

Полученные результаты для дисперсий σ_i^2 и σ_y^2 поясняют, почему матрица C называется *дисперсионной*.

Входящая в формулы (3.6.38), (3.6.39) дисперсия ошибок наблюдений σ^2 обычно бывает неизвестна и должна быть оценена с помощью полученных экспериментальных данных. Для этого используется остаточная сумма квадратов (3.6.28), которая имеет

$$\varphi = N - (k + 1) \quad (3.6.40)$$

степеней свободы (N слагаемых, между которыми существует $k + 1$ линейная связь, определяемая системой нормальных уравнений). Величина

$$s^2 = \frac{S_R}{\varphi} \quad (3.6.41)$$

является несмещенной оценкой дисперсии ошибок наблюдений. Тогда формула (3.6.38) примет вид

$$\sigma_i^2 = c_{ii}s^2 = s_i^2. \quad (3.6.42)$$

При большой дисперсии ошибок наблюдений или слабом влиянии какой-либо контролируемой переменной некоторые оцениваемые коэффициенты модели могут оказаться сравнимыми с ошибками наблюдений, или, как говорят, *статистически незначимыми*. Проверка статистической значимости коэффициентов проводится следующим образом [92].

Рассмотрим величину

$$t = \frac{\hat{a}_i - \bar{a}_i}{s_i}, \quad (3.6.43)$$

распределенную по закону Стьюдента с φ степенями свободы. Эта величина

характеризует ошибку определения коэффициентов по отношению к дисперсии опыта. Далее принимается нуль-гипотеза о том, что $\bar{a}_i = 0$. Тогда из (3.6.43) получим

$$|\hat{a}_i| = t s_i = t \sqrt{c_{ii}} s. \quad (3.6.44)$$

Знак модуля введен, чтобы рассматривать отклонения в обе стороны от нулевого значения. Далее задаемся вероятностью P принятия нуль-гипотезы и по таблицам распределения Стьюдента находим соответствующее значение t , обозначив его $t_{кр}$. Если выполняется условие

$$|\hat{a}_i| < t_{кр} \sqrt{c_{ii}} s, \quad (3.6.45)$$

то принимается нуль-гипотеза и соответствующий коэффициент считается статистически незначимым. Обычно задаются не вероятностью P , а уровнем значимости $\alpha = 1 - P$ и выбирают его в пределах $\alpha = 0,05 \div 0,1$. Величина α означает величину вероятности отвергнуть нуль-гипотезу, когда она на самом деле верна, т.е. принять незначимый коэффициент за значимый.

Если при проверке значимости какой-либо коэффициент оказался незначимым, то он исключается из уравнения модели. Но при этом необходимо, используя прежние результаты наблюдений, пересчитать все остальные оценки коэффициентов, так как они взаимосвязаны, поскольку определяются системой уравнений. При этом будут новые матрицы $f(x)$, F , S .

Проверка адекватности модели. Выше считалось, что вид модели (линейная, квадратичная и т.д.) выбран правильно. При неверном выборе модели оценки коэффициентов, рассчитанные по методу наименьших квадратов, в общем случае оказываются смещенными, т.е. среднее значение коэффициентов не равно истинному значению:

$$M\{\hat{a}_i\} \neq \bar{a}_i. \quad (3.6.46)$$

Такая модель называется *неадекватной* и должна быть изменена.

Для проверки гипотезы об адекватности модели нужно сопоставить достигнутую точность с ошибками наблюдений. Если ошибки модели превосходят ошибки наблюдений, то модель является неадекватной. В этом случае уже нельзя оценивать ошибку наблюдений по разности между результатом наблюдений выходной переменной и результатом ее расчета по модели, так как оценка выхода оказывается также смещенной:

$$M\{\hat{y}^i\} \neq \bar{y}^i. \quad (3.6.47)$$

Поэтому дисперсия ошибок наблюдений может быть оценена лишь по результатам нескольких параллельных опытов, проведенных в каждой экспериментальной точке. Расчет коэффициентов производится с использованием средних значений для каждой экспериментальной точки согласно (3.6.12).

Для проверки гипотезы адекватности сравнивают две суммы квадратов [92].

1. Сумма квадратов, характеризующая неадекватность модели:

$$S_D = \sum_{i=1}^N \nu (\tilde{y}^i - \hat{y}^i)^2 = \nu S_R. \quad (3.6.48)$$

Эта сумма зависит от разности между наблюдаемыми и рассчитанными по модели значениями выходной переменной.

2. Сумма квадратов, характеризующая ошибки наблюдений:

$$S_e = \sum_{i=1}^N \sum_{j=1}^{\nu} (\tilde{y}^{ij} - \tilde{y}^i)^2. \quad (3.6.49)$$

Здесь $\tilde{y}^{ij}$ — значение наблюдаемого выхода в каждом параллельном опыте для точки x^i ; $\tilde{y}^i$ — среднее по ν опытам для каждой точки. Суммирование проводится по всем опытам x^i .

Сумма S_D состоит из N слагаемых, между которыми существует $k + 1$ линейная связь, определяемая системой нормальных уравнений (3.6.30). Поэтому S_D имеет

$$\varphi_1 = N - (k + 1) \quad (3.6.50)$$

степеней свободы, а оценка дисперсии неадекватности определяется выражением

$$s_1^2 = \frac{S_D}{\varphi_1}. \quad (3.6.51)$$

Сумма S_e состоит из νN слагаемых, между которыми существует только N линейных связей с учетом (3.6.12). Поэтому S_e имеет

$$\varphi_2 = N\nu - N = N(\nu - 1) \quad (3.6.52)$$

степеней свободы, а оценка дисперсии единичного наблюдения определяется выражением

$$s_2^2 = \frac{S_e}{\varphi_2}. \quad (3.6.53)$$

Дисперсия наблюдений по ν параллельным опытам в соответствующее число раз меньше дисперсии единичного наблюдения, т.е.

$$s^2 = \frac{S_e}{\nu \varphi_2}. \quad (3.6.54)$$

Проверка адекватности проводится с помощью отношения дисперсий

$$F = \frac{s_1^2}{s_2^2} = \frac{S_D / \varphi_1}{S_e / \varphi_2}, \quad (3.6.55)$$

которое является случайной величиной, имеющей распределение Фишера,

или F -распределение, со степенями свободы φ_1 и φ_2 . Далее поступают так же, как при проверке значимости коэффициентов. Задаются уровнем значимости $\alpha = 0,05 \div 0,1$, вычисляют вероятность $P = 1 - \alpha$ и по таблицам F -распределения находят соответствующее критическое значение $F_{кр}$. Если выполняется условие

$$F > F_{кр}, \quad (3.6.56)$$

то модель считается адекватной, или, точнее, гипотеза об адекватности модели не противоречит опытным данным. Заметим, что неравенство (3.6.56) имеет противоположный знак по сравнению с (3.6.45), так как здесь проверяется основная гипотеза, гипотеза об адекватности, а там — противоположная, нуль-гипотеза.

Планирование эксперимента для линейных моделей. Полные факторные планы. Метод наименьших квадратов является оптимальным для заданной выборки наблюдений, т.е. дисперсия оценки коэффициентов модели является минимальной. Однако этот минимум зависит от выбора экспериментальных точек, так как изменяется дисперсионная матрица C . Поэтому возникает задача: найти такое расположение экспериментальных точек в пространстве контролируемых переменных, чтобы получаемая точность оценок была максимальной. В этом и заключается основное содержание теории планирования эксперимента.

Начало работ по планированию экспериментов было положено шотландским математиком Р. Фишером в 30-х годах нашего века для решения задач агробиологии. В настоящее время эта ветвь математической статистики находит широкое применение в различных областях науки и техники.

Применение методов планирования эксперимента требует проведения опытов в заданных исследователем условиях (что бывает не всегда возможно). Такой эксперимент называется *активным*. Активный эксперимент в сочетании с методами планирования позволяет получить требуемые результаты с минимальными затратами.

Центральным моментом планирования эксперимента для линейных моделей является *полный факторный эксперимент*, идея которого заключается в изменении в каждом опыте нескольких независимых переменных, в противоположность однофакторному эксперименту, где в каждом опыте изменяется только одна переменная (фактор).

В планах экспериментов для построения линейных моделей контролируемые переменные варьируются на двух уровнях $X_{\max}$ и $X_{\min}$. Обычно вводят нормированные переменные

$$x_i = \frac{X_i - M X_i}{\Delta X_i}, \quad (3.6.57)$$

где

$$M X_i = \frac{X_{i \max} + X_{i \min}}{2}$$

— среднее значение;

$$\Delta X_i = \frac{X_{i \max} - X_{i \min}}{2}$$

— интервал варьирования. При этом нормированные переменные изменяются в пределах $x_i = \pm 1$.

В полном факторном эксперименте используются все возможные комбинации переменных на двух уровнях, поэтому полное число экспериментов равно

$$N = 2^n \quad (3.6.58)$$

(n — число контролируемых переменных), такие планы называют планами типа 2^n .

Для $n = 2$ полный факторный план имеет вид

$$X(2) = \begin{array}{cc} x_1 & x_2 \\ \begin{bmatrix} - & - \\ + & - \\ - & + \\ + & + \end{bmatrix} \end{array} \quad (3.6.59)$$

Здесь знаками "плюс" и "минус" обозначено нахождение фактора на верхнем или нижнем уровне. Так, в первом опыте оба фактора находятся на нижнем уровне, в четвертом — оба на верхнем.

• Пользуясь таким планированием, можно определить коэффициенты в следующей модели:

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2. \quad (3.6.60)$$

Если обозначить $x_1 x_2 = x_3$, то размерность модели $k = 3$ и число опытов $N = k + 1 = 4$, что требуется для существования обратной матрицы C . Матрица F для модели (3.6.60) имеет вид

$$F = \begin{array}{cc} & \begin{array}{cccc} x_0 & x_1 & x_2 & x_1 x_2 \end{array} \\ \begin{bmatrix} +1 & -1 & -1 & +1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & +1 & +1 \end{bmatrix} \end{array} \quad (3.6.61)$$

Матрица F получается из плана эксперимента путем добавления столбца для свободного члена и столбца для взаимодействий $x_1 x_2$, который получается перемножением элементов столбцов x_1 и x_2 .

Для трех независимых переменных $N = 2^3 = 8$ и полный факторный план имеет вид

$$X(3) = \begin{matrix} & x_1 & x_2 & x_3 \\ \begin{bmatrix} - & - & - \\ + & - & - \\ - & + & - \\ + & + & - \\ - & - & + \\ + & - & + \\ - & + & + \\ + & + & + \end{bmatrix} & & \end{matrix} \quad (3.6.62)$$

Легко заметить, что план для трех переменных (3.6.62) получается из плана для двух переменных (3.6.59), повторенного дважды с включением третьей переменной сначала на нижнем уровне, затем на верхнем. В общем случае план $X(n+1)$ может быть получен из плана $X(n)$ по формуле

$$X(n+1) = \begin{bmatrix} X(n) & l \\ X(n) & -l \end{bmatrix}, \quad (3.6.63)$$

где $l = (1, 1, \dots, 1)^T$ — вектор-столбец, содержащий n единиц.

Такое планирование позволяет оценить коэффициенты в модели

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3 + a_{123}x_1x_2x_3. \quad (3.6.64)$$

Для этой модели матрица F имеет вид

	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$	Код
$F =$	+	-	-	-	+	+	+	-	(1)
	+	+	-	-	-	-	-	+	a
	+	-	+	-	-	+	-	+	b
	+	+	+	-	+	-	-	-	ab
	+	-	-	+	+	-	-	+	c
	+	+	-	+	-	+	-	-	ac
	+	-	+	+	-	-	+	-	bc
	+	+	+	+	+	+	+	+	abc

(3.6.65)

Здесь столбцы для взаимодействий получаются путем перемножения элементов столбцов соответствующих переменных. Для упрощения вместо ± 1 записано $\pm$.

Можно, конечно, использовать план 2^3 для оценки коэффициентов модели вида

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3. \quad (3.6.66)$$

Тогда лишние опыты можно использовать как параллельные для проверки адекватности линейной модели. Проверить адекватность модели (3.6.64) при восьми опытах нельзя, так как не хватает степеней свободы: $\varphi_1 = N - (k + 1) = 0$. Для проверки адекватности модели (3.6.64) нужно ставить дополнительные опыты.

План эксперимента можно записать в компактной форме в виде кодовой строки. Для этого обозначим переменные x_1, x_2, x_3 буквами a, b, c и в обозначении каждого эксперимента будем ставить соответствующую букву там, где данный фактор был на верхнем уровне. Кодовое обозначение показано в последнем столбце (3.6.65). Если все факторы находятся на нижнем уровне, то пишется (1). Тогда план полного факторного эксперимента для трех независимых переменных запишется строкой

$$(1), a, b, ab, c, ac, bc, abc,$$

а для двух переменных —

$$(1), a, b, ab.$$

Отсюда видно, что в плане для трех переменных план для двух переменных повторяется дважды с добавлением фактора c .

Пользуясь этими обозначениями и общим методом построения полных факторных планов, легко записать кодовую строку для плана при четырех независимых переменных:

$$(1), a, b, ab, c, ac, bc, abc, d, ad, bd, abd, cd, acd, bcd, abcd.$$

Теперь можно записать матрицу плана, поставив знак "плюс" там, где в кодовой строке стоят соответствующие факторы.

- Аналогичным образом строится полный факторный эксперимент типа 2^n для любого количества независимых переменных. При увеличении числа переменных количество опытов растет как показательная функция $N = 2^n$. Однако при этом повышается и точность в определении коэффициентов модели, так как при многофакторном планировании все опыты используются для оценки каждого из коэффициентов, как это будет показано ниже.

Дробные реплики. Если можно ограничиться линейной моделью без взаимодействий, то число опытов можно резко снизить, используя не полный факторный эксперимент, а только его часть, так называемую *дробную реплику*. Пусть нужно получить линейную модель для трех независимых переменных. Полный факторный эксперимент содержит восемь опытов, но для решения этой задачи можно ограничиться четырьмя опытами, если в плане 2^2 произведение x_1x_2 приравнять третьему фактору x_3 . Получим матрицу планирования

$$X_1(3) = \begin{matrix} & \begin{matrix} x_1 & x_2 & x_3 \end{matrix} \\ \begin{bmatrix} - & - & + \\ + & - & - \\ - & + & - \\ + & + & + \end{bmatrix}, & (3.6.67) \end{matrix}$$

содержащую четыре опыта, или полуреплику от полного факторного эксперимента. Пользуясь этим планом, можно оценить коэффициенты модели

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3. \quad (3.6.68)$$

Если парные взаимодействия все же имеются, то в полученной модели коэффициенты будут оценками для совместных эффектов

$$\begin{aligned} \hat{a}'_1 &= \hat{a}_1 + \hat{a}_{23}, \\ \hat{a}'_2 &= \hat{a}_2 + \hat{a}_{13}, \\ \hat{a}'_3 &= \hat{a}_3 + \hat{a}_{12}. \end{aligned} \quad (3.6.69)$$

Эти эффекты не могут быть отдельно оценены в плане из четырех опытов, так как здесь неразличимы столбцы для линейных членов и парных взаимодействий. Например, элементы столбца для произведения x_1x_3 равны элементам столбца x_2 .

Если после постановки первых четырех опытов возникнут сомнения, что парные взаимодействия равны нулю, то можно провести еще четыре опыта, приравняв $x_3 = -x_1x_2$. Получим вторую полуреплику от полного факторного эксперимента

$$X_2(3) = \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \\ \begin{bmatrix} - & - & - \\ + & - & + \\ - & + & + \\ + & + & - \end{bmatrix} \end{array} \end{array}. \quad (3.6.70)$$

Здесь коэффициенты модели (3.6.68) будут совместными оценками следующих эффектов:

$$\begin{aligned} \hat{a}''_1 &= \hat{a}_1 - \hat{a}_{23}, \\ \hat{a}''_2 &= \hat{a}_2 - \hat{a}_{13}, \\ \hat{a}''_3 &= \hat{a}_3 - \hat{a}_{12}. \end{aligned} \quad (3.6.71)$$

Теперь можно отдельно оценить линейные эффекты и парные взаимодействия, например:

$$\hat{a}_1 = \frac{\hat{a}'_1 + \hat{a}''_1}{2}, \quad \hat{a}_{23} = \frac{\hat{a}'_1 - \hat{a}''_1}{2}. \quad (3.6.72)$$

Объединив две полуреплики, получим полный факторный эксперимент, позволяющий отдельно оценить линейные эффекты и взаимодействия.

Обратим внимание, что разбиение полного факторного плана нельзя производить произвольно на две группы. В первую полуреплику отбираются строки с нечетным числом букв в обозначении строк

$c, a, b, abc,$

что соответствует требованию $x_3 = x_1x_2$. Здесь третья переменная попадает на верхний уровень только в тех строках, где две другие переменные находятся одновременно на верхнем или нижнем уровне. Во вторую полу-

реплику берутся строки с четным числом букв

(1), ac, bc, ab

в соответствии с требованием $x_3 = -x_1x_2$.

В задаче с четырьмя независимыми переменными можно в плане 2^3 приравнять тройное взаимодействие к четвертому фактору $x_4 = x_1x_2x_3$, предположив тем самым, что $a_{123} = 0$. Получим полуреплику от полного факторного эксперимента 2^4 , которая задается строкой

(1), $ad, bd, ab, cd, ac, bc, abcd$.

Здесь все строки четные. Эта строка получается из строки для полного факторного эксперимента 2^3 путем умножения на букву d нечетного сочетания букв соответственно требованию $x_4 = x_1x_2x_3$. Здесь четвертый фактор находится на верхнем уровне, когда на верхнем уровне находятся один или три других фактора.

Вторую полуреплику получим, приравняв $x_4 = -x_1x_2x_3$. Ее план задается строкой, содержащей нечетное сочетание букв

$d, a, b, abd, c, acd, bcd, abc$.

Эта строка получена умножением на букву d четных комбинаций букв в матрице 2^3 . Объединив обе полуреплики, получим матрицу планирования для полного факторного эксперимента 2^4 . Число четных строк в полном факторном эксперименте всегда равно числу нечетных строк.

Можно построить дробные реплики высокой степени дробности: четвертьреплики, восьмые и т.д. Так, например, для оценки влияния семи переменных в линейном приближении можно ограничиться восемью опытами ($N = k + 1$). Но при этом не оцениваются эффекты взаимодействия, т.е. считается, что они малы. В этом случае можно взять полный эксперимент 2^3 и положить

$$x_4 = x_1x_2, \quad x_5 = x_1x_3, \quad x_6 = x_2x_3, \quad x_7 = x_1x_2x_3.$$

Получим $1/16$ от полного факторного эксперимента. Планы, в которых все взаимодействия заменены линейными эффектами, называются *насыщенными*. Их можно строить для числа переменных $n+1$, равного степени 2, т.е. $n = 3, 7, 15, \dots$

Для обозначения дробных реплик, в которых p линейных эффектов приравнены к эффектам взаимодействия, используют обозначение 2^{n-p} . В последнем примере рассмотрена дробная реплика $2^{7-4} = 2^7 \cdot 2^{-4} = 2^7 \cdot \frac{1}{16}$.

Такой способ обозначения не полностью характеризует свойства реплики. Дробные реплики можно получить, приравняв основные эффекты к различным взаимодействиям. Например, в полуреплике 2^{4-1} можно положить $x_4 = x_1x_j$, т.е. приравнять к одному из парных взаимодействий. При этом изменится система совместных оценок, т.е. линейные эффекты будут смешаны с другими взаимодействиями.

Разрешающая способность дробных реплик. Рассмотрим, когда и какие эффекты определяются совместно в дробных репликах.

Способ задания дробной реплики называется *генерирующим отношением* или *генератором*. Например, планирование 2^{3-1} представляется дву-

мя полурепликами с генераторами

$$x_3 = x_1 x_2, \quad x_3 = -x_1 x_2. \quad (3.6.73)$$

Будем обозначать элементы первого столбца матрицы F буквой I (они всегда равны $+1$). Найдем соотношения, определяющие элементы первого столбца для каждой из полуреplik. Умножая левые и правые части генераторов (3.6.73) на x_3 , получаем

$$x_3^2 = I_1 = x_1 x_2 x_3, \quad x_3^2 = I_2 = -x_1 x_2 x_3, \quad (3.6.74)$$

так как всегда $x_i^2 = I$ ($x_i = \pm 1$).

Соотношения, задающие элементы первого столбца, называются *определяющими контрастами* или просто *контрастами*.

Зная определяющие контрасты, легко найти соотношения, задающие все совместные оценки. Для этого нужно последовательно помножить независимые переменные на контраст, учитывая, что $x_i^2 = 1$. В случае полуреplik 2^{3-1} совместные оценки будут задаваться соотношениями

$$\begin{aligned} x_1 &= x_1 \quad I_1 = x_2 x_3, & x_1 &= x_1 \quad I_2 = -x_2 x_3, \\ x_2 &= x_2 \quad I_1 = x_1 x_3, & x_2 &= x_2 \quad I_2 = -x_1 x_3, \\ x_3 &= x_3 \quad I_1 = x_1 x_2, & x_3 &= x_3 \quad I_2 = -x_1 x_2. \end{aligned} \quad (3.6.75)$$

Это значит, что коэффициенты модели в каждой из полуреplik будут смешанными оценками

$$\begin{aligned} a'_1 &= a_1 + a_{23}, & a''_1 &= a_1 - a_{23}, \\ a'_2 &= a_2 + a_{13}, & a''_2 &= a_2 - a_{13}, \\ a'_3 &= a_3 + a_{12}, & a''_3 &= a_3 - a_{12}, \end{aligned} \quad (3.6.76)$$

как показано в разобранным раньше примере.

Рассмотрим планирование 2^{4-1} . Две полуреplik могут быть заданы генераторами

$$x_4 = x_1 x_2 x_3, \quad x_4 = -x_1 x_2 x_3; \quad (3.6.77)$$

контраст получим умножением на x_4 :

$$I_1 = x_4^2 = x_1 x_2 x_3 x_4, \quad I_2 = x_4^2 = -x_1 x_2 x_3 x_4. \quad (3.6.78)$$

Совместные оценки определяются соотношениями

$$\begin{aligned} x_1 &= x_2 x_3 x_4, & x_1 &= -x_2 x_3 x_4, \\ x_2 &= x_1 x_3 x_4, & x_2 &= -x_1 x_3 x_4, \\ x_3 &= x_1 x_2 x_4, & x_3 &= -x_1 x_2 x_4, \\ x_4 &= x_1 x_2 x_3, & x_4 &= -x_1 x_2 x_3, \\ x_1 x_2 &= x_3 x_4, & x_1 x_2 &= -x_3 x_4, \\ x_1 x_3 &= x_2 x_4, & x_1 x_3 &= -x_2 x_4, \\ x_1 x_4 &= x_2 x_3, & x_1 x_4 &= -x_2 x_3. \end{aligned} \quad (3.6.79)$$

Обычно взаимодействия высших порядков невелики — начиная с тройных их можно считать равными нулю. Тогда в планировании 2^{4-1} можно получить отдельные оценки для линейных членов и смешанные оценки

для парных взаимодействий. Таким образом, здесь, по сравнению с планированием 2^{3-1} , можно оценить отдельно линейные эффекты.

Если рассмотреть полуреплику 2^{5-1} для пяти переменных, то окажется, что линейные эффекты смешаны с четверными взаимодействиями, а парные — с тройными. Но так как тройные и выше взаимодействия можно считать равными нулю, то в полуреплике 2^{5-1} оцениваются отдельно линейные эффекты и парные взаимодействия.

Таким образом, с ростом числа независимых переменных разрешающая способность полуреplik возрастает.

Разрешающая способность полуреplik зависит от генераторов. Для планирования 2^{4-1} можно получить полуреплику с меньшей разрешающей способностью, если задаться, например, следующими генерирующими соотношениями:

$$x_4 = x_1 x_2, \quad x_4 = -x_1 x_2. \quad (3.6.80)$$

Тогда контрасты задаются выражениями

$$I_1 = x_1 x_2 x_4, \quad I_2 = -x_1 x_2 x_4, \quad (3.6.81)$$

а система совместных оценок имеет вид

$$\begin{aligned} x_1 &= x_2 x_4, & x_1 &= -x_2 x_4, \\ x_2 &= x_1 x_4, & x_2 &= -x_1 x_4, \\ x_3 &= x_1 x_2 x_3 x_4, & x_3 &= -x_1 x_2 x_3 x_4, \\ x_4 &= x_1 x_2, & x_4 &= -x_1 x_2, \\ x_1 x_3 &= x_2 x_3 x_4, & x_1 x_3 &= -x_2 x_3 x_4, \\ x_2 x_3 &= x_1 x_3 x_4, & x_2 x_3 &= -x_1 x_3 x_4, \\ x_3 x_4 &= x_1 x_2 x_3, & x_3 x_4 &= -x_1 x_2 x_3. \end{aligned} \quad (3.6.82)$$

Здесь три линейных эффекта определяются совместно с парными взаимодействиями. Разрешающая способность этих полуреplik ниже, чем у полуреplik с контрастом $I = \pm x_1 x_2 x_3 x_4$, так как там линейные эффекты определяются независимо от парных взаимодействий. Однако такие полуреплики используются в специальных случаях, когда надо оценить именно парные взаимодействия, так как здесь три парных взаимодействия смешаны с тройными, которыми можно пренебречь. Эти полуреплики задаются строками

$d, a, b, abd, cd, ac, bc, abcd;$

(1), $ad, bd, ab, c, acd, bcd, abc.$

Эти строки получаются из плана 2^3 , если обозначения строк умножить на букву d : один раз, когда $x_1 x_2 = 1$, другой раз, когда $x_1 x_2 = -1$. В этих полурепликах, в отличие от полуреplik с контрастом $I = \pm x_1 x_2 x_3 x_4$, в каждой строке есть четные и нечетные комбинации букв.

Полуреплики, имеющие только четные или нечетные комбинации букв, называются *главными полурепликами*, так как они имеют более высокую разрешающую способность.

Мы не будем рассматривать разрешение более дробных реплик. Отметим только, что с увеличением степени дробности разрешающая способ-

ность реплик падает, так как все больше эффектов смешивается вместе. Самая низкая разрешающая способность у насыщенных планов, когда все взаимодействия заменены линейными членами. Например, для рассмотренной $1/16$ реплики 2^{7-4} оценка линейного эффекта является совместной оценкой следующих эффектов:

$$a'_1 = a_{24} + a_{235} + a_{1345}.$$

Таким образом, все линейные эффекты смешаны с парными.

Вычисление коэффициентов модели. Вычисление коэффициентов модели может быть произведено по общей формуле (3.6.30) решения системы нормальных уравнений. Однако полный факторный эксперимент и дробные реплики обладают рядом полезных свойств, которые позволяют значительно упростить вычисление коэффициентов.

Полный факторный эксперимент и дробные реплики являются ортогональными планами:

$$\sum_{i=1}^N x_k^i x_l^i = \begin{cases} N, & \text{если } k = l, \\ 0, & \text{если } k \neq l, \end{cases} \quad (3.6.83)$$

т.е. скалярное произведение разных вектор-столбцов равно нулю. Информационная матрица ортогонального плана диагональна, причем все диагональные элементы равны N :

$$M = F^T F = \begin{vmatrix} N & 0 & \dots & 0 \\ 0 & N & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & N \end{vmatrix} = N I_{k+1}, \quad (3.6.84)$$

где I_{k+1} — единичная матрица размера $k+1$.

Обратная матрица при этом равна

$$M^{-1} = C = \frac{1}{N} I_{k+1}, \quad (3.6.85)$$

а все ее диагональные элементы равны $1/N$.

Ортогональность — очень важное свойство плана, так как при этом система нормальных уравнений распадается на отдельные уравнения первого порядка для каждого параметра модели, что значительно упрощает вычисления.

Вычисление коэффициентов модели в полном факторном эксперименте и дробных репликах производится по формуле

$$\hat{a}_k = \frac{\sum_{i=1}^N x_k^i \tilde{y}^i}{N}, \quad (3.6.86)$$

т.е. необходимо умножить значение независимой k -й переменной в каждом i -м опыте на значение выхода, просуммировать и разделить на количество опытов. Учитывая, что нормированные переменные принимают только значения ± 1 , при использовании формулы (3.6.86) нужно просто вычислить алгебраическую сумму выходов во всех опытах, беря знаки $\pm$ в соответствии с матрицей планирования.

Из ортогональности плана следуют свойства, используемые для проверки правильности составленной матрицы планирования. Сумма квадратов элементов каждого столбца равна количеству опытов

$$\sum_{i=1}^N (x_k^i)^2 = N, \quad (3.6.87)$$

а сумма элементов каждого столбца равна нулю (т.е. план является центральным — все экспериментальные точки расположены симметрично):

$$\sum_{i=1}^N x_k^i = 0.$$

Таким образом, оценки всех коэффициентов в факторном эксперименте вычисляются независимо друг от друга. Если какой-либо из коэффициентов окажется незначимым, то его можно отбросить, не пересчитывая другие коэффициенты.

Дисперсия коэффициентов регрессии определяется по формуле

$$s_i^2 = c_{ii} s^2 = \frac{s^2}{N}, \quad (3.6.88)$$

где s^2 определяется по формуле (3.6.41).

Простота вычислений коэффициентов при использовании ортогональных планов открыла регрессионному анализу широкую дорогу в различные области науки и техники. Основное преимущество факторного эксперимента заключается в том, что здесь одновременно варьируются все переменные. Поэтому каждый коэффициент вычисляется по результатам всех N экспериментов, и, следовательно, дисперсия оценки коэффициентов оказывается в N раз меньше дисперсии ошибки опытов (3.6.88).

Факторные планы для линейных моделей без взаимодействий, обладают еще одним полезным свойством — *ротатабельностью*. Это значит, что получаемая информация равномерно распределена по всем направлениям пространства планирования, иными словами, дисперсия рассчитанной по уравнению модели оценки выхода $\hat{y}$ в точке x зависит только от расстояния этой точки от центра плана, но не зависит от направления (аналогия — изотропные поля). Таким образом, ротатабельные планы обеспечивают постоянное значение дисперсии оценки выхода во всех точках, равноудаленных от центра плана.

Планирование эксперимента для квадратичных моделей. Ортогональные центральные композиционные планы. Для создания квадратичных моделей вида (3.6.3) применяются планы второго порядка. Нахождение оценок коэффициентов квадратичных моделей требует, чтобы каждая независимая переменная принимала три разных значения. Этого можно достичь, проведя эксперименты по плану 3^n , т.е. перебрав все возможные сочетания переменных на трех уровнях. Однако при этом получается большое количество экспериментов ($N = 3^n$), кроме того, такие планы не оптимальны. Поэтому для построения планов второго порядка был предложен ряд методов, позволяющих сократить число экспериментов. Наибольшее применение нашел метод композиционного планирования.

Композиционное планирование заключается в том, что к линейному плану типа 2^n или 2^{n-p} добавляют несколько специальным образом выбранных так называемых *звездных точек*. При этом линейная часть плана называется *ядром плана*. Кроме того, добавляется еще точка в центре плана. Координаты звездных точек равны $(\pm \alpha, 0, \dots, 0)$, $(0, \pm \alpha, 0, \dots, 0)$ и т.д., т.е. звездные точки расположены на координатных осях. Например,

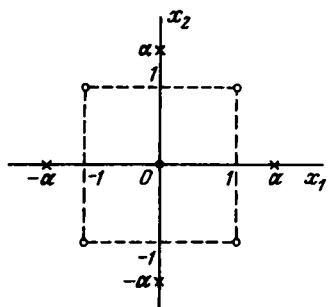


Рис. 3.28. Расположение точек на плоскости планирования

для уравнения модели второго порядка с двумя независимыми переменными получим расположение точек на плоскости планирования, приведенное на рис. 3.28. Здесь белыми кружками обозначены точки плана 2^2 с координатами

$$\begin{aligned} &(-1, -1), \\ &(+1, -1), \\ &(-1, +1), \\ &(+1, +1), \end{aligned}$$

крестиками — звездные точки с координатами

$$\begin{aligned} &(\alpha, 0), \\ &(-\alpha, 0), \\ &(0, \alpha), \\ &(0, -\alpha), \end{aligned}$$

черный кружок в начале координат — центр плана.

В зависимости от величины плеча α и числа точек в центре плана получают планы с разными свойствами — ортогональные или ротатабельные (но не одновременно, как в линейных планах). Наибольшее применение находят ортогональные композиционные планы в связи с простотой вычисления коэффициентов модели.

Как отмечено выше, свойство ортогональности плана математически означает, что скалярное произведение разных вектор-столбцов равно нулю [см. (3.6.83)]. Поскольку в матрице F для квадратичных моделей имеются столбцы для квадратов независимых переменных x_{ik}^2 , то не все скалярные произведения будут равны нулю:

$$\sum_{i=1}^N x_{0i} x_{ki}^2 \neq 0, \quad \sum_{i=1}^N x_{ik}^2 x_{il}^2 \neq 0, \quad (3.6.89)$$

так как $x_{0i} = 1$, а $x_{ki}^2 \geq 0$. Здесь x_{ki} — k -я независимая переменная в i -м опыте.

Для ортогональности плана необходимо преобразовать модель (3.6.3) следующим образом:

$$y = b_0 + a_1 x_1 + \dots + a_n x_n + a_{11}(x_1^2 - \beta) + \dots + a_{nn}(x_n^2 - \beta) + a_{12} x_1 x_2 + \dots + a_{kn} x_k x_n - 1. \quad (3.6.90)$$

Здесь

$$\beta = \frac{\sum_{i=1}^N x_{ki}^2}{N} \quad (3.6.91)$$

есть средний квадрат независимых переменных. Чтобы перейти от модели (3.6.90) к модели (3.6.3), нужно вычислить свободный член

$$a_0 = b_0 - \beta \sum_{k=1}^n a_{kk}. \quad (3.6.92)$$

В общем случае матрица F для центрального ортогонального композиционного плана представлена в табл. 3.6. Легко видеть, что общее число точек плана равно

$$N = 2^{n-p} + 2n + 1. \quad (3.6.93)$$

Это меньше, чем при планировании 3^n . Например, для $n = 3$ при использовании в качестве ядра плана полного факторного эксперимента 2^3 получим $N = 15$, тогда как планирование 3^n требует 27 опытов. Заметим, что в качестве ядра плана можно использовать дробные реплики 2^{n-p} , только начиная с пяти независимых переменных, иначе в матрице F будут совпадающие столбцы и план не удастся сделать ортогональным.

- Рассмотрим выбор плеча α ортогональных планов. Из матрицы планирования видно, что скалярные произведения столбцов при линейных членах и парных взаимодействиях равны нулю при любом выборе α (не зависят от α). От величины α зависят столбцы при членах $x_i^2 - \beta$, их попарные произведения не равны нулю в общем случае. Приравняв нулю сумму произведений элементов двух таких столбцов, получаем условие выбора α , обеспечивающее ортогональность плана:

$$2^{n-p}(1 - \beta)^2 - 4\beta(\alpha^2 - \beta) + \beta^2(2n - 4) + \beta^2 = 0. \quad (3.6.94)$$

Подставляя в (3.6.94) значение β , равное

$$\beta = \frac{2^{n-p} + 2\alpha^2}{N} = \frac{\sum_{i=1}^N x_{ki}^2}{N}, \quad (3.6.95)$$

получаем

$$2^{n-p} \left(1 - \frac{2^{n-p} + 2\alpha^2}{N} \right)^2 - 4 \frac{2^{n-p} + 2\alpha^2}{N} \left(\alpha^2 - \frac{2^{n-p} + 2\alpha^2}{N} \right) + (2n - 3) \left(\frac{2^{n-p} + 2\alpha^2}{N} \right)^2 = 0. \quad (3.6.96)$$

Таблица 3.6

	Номер опыта i	Матрица F			
		x_0	План x_n		
			x_1	x_2	
Ядро плана	1	1	+1	+1	
	2	1	-1	+1	
	3	1	+1	-1	
	4	1	-1	-1	
	5	1	+1	+1	
	.	.	.	.	
	2^{n-p}	1	.	.	
Звездные точки	$2^{n-p} + 1$	1	$+\alpha$	0	
	$2^{n-p} + 2$	1	$-\alpha$	0	
	.	1	0	$+\alpha$	
	.	1	0	$-\alpha$	
	.	.	.	.	
	.	.	.	.	
	$2^{n-p} + 2n$	1	0	0	
Центр плана	$N = 2^{n-p} + 2n + 1$	1	0	0	

Решая (3.6.96) относительно α , получаем величину плеча, обеспечивающую ортогональность плана:

$$\alpha = \sqrt{2^{\frac{n-p}{2}-1} (\sqrt{N} - 2^{\frac{n-p}{2}})}. \quad (3.6.97)$$

В табл. 3.7 приведены параметры ортогональных центральных композиционных планов для различного числа независимых переменных.

Вычисление коэффициентов модели при ортогональном планировании. Информационная матрица ортогонального плана второго порядка имеет вид

$$M = \begin{vmatrix} m_0 & 0 & 0 & 0 \\ 0 & m_1 I_n & 0 & 0 \\ 0 & 0 & m_2 I_n & 0 \\ 0 & 0 & 0 & m_3 I_{\binom{n}{2}} \end{vmatrix}, \quad (3.6.98)$$

Матрица F							
План x_n		$x_1^2 - \beta$		$x_n^2 - \beta$	$x_1 x_2$	...	$x_n x_{n-1}$
...	x_n						
	+1	$1 - \beta$		$1 - \beta$	+1		1
	+1	$1 - \beta$		$1 - \beta$	-1		1
	+1	$1 - \beta$		$1 - \beta$	-1		.
	.	.		.	+1		.
	.	.		.			.
	-1	.		.			.
	-1	.		.			.
...	-1	$1 - \beta$	...	$1 - \beta$	...	...	.
	0	$\alpha^2 - \beta$		$-\beta$	0		0
	0	$\alpha^2 - \beta$		$-\beta$	0		0
	0	$-\beta$		.	.		.
	0	$-\beta$		.	.		.
	.	.		.	.		.
	+ α	$-\beta$		$\alpha^2 - \beta$	0		0
...	- α	$-\beta$	...	$\alpha^2 - \beta$	0	...	0
	0	$-\beta$		$-\beta$	0		0

• где

$$m_0 = N = 2^n - p + 2n + 1;$$

$$m_1 = 2^n - p + 2\alpha^2;$$

$$m_2 = 2^n - p(1 - \beta)^2 + 2(\alpha^2 - \beta)^2 + (2n - 1)\beta^2; \quad (3.6.99)$$

$$m_3 = 2^n - p;$$

I_n — единичная матрица размера n ; $\binom{n}{2}$ — число сочетаний из n по 2. Отсюда

Таблица 3.7

n	Ядро	N	α	β	n	Ядро	N	α	β
2	2^2	9	1	0,6667	6	2^{6-1}	45	1,722	0,843
3	2^3	15	1,215	0,73	7	2^{7-1}	79	1,885	0,9
4	2^4	25	1,414	0,8	8	2^{8-2}	81	2,001	0,8889
5	$2^5 - 1$	27	1,547	0,77					

Таблица 3.8

n	c_0	c_1	c_2	c_3
2	0,1111	0,1667	0,5	0,25
3	0,0667	0,0913	0,2298	0,125
4	0,04	0,05	0,125	0,0625
5	0,03704	0,0481	0,0871	0,0625
6	0,0222	0,0264	0,0564	0,03125
7	0,0127	0,0141	0,0389	0,0156
8	0,0123	0,0139	0,0312	0,0156

получаем выражение для дисперсионной матрицы

$$C = M^{-1} = \begin{pmatrix} c_0 & 0 & 0 & 0 \\ 0 & c_1 I_n & 0 & 0 \\ 0 & 0 & c_2 I_n & 0 \\ 0 & 0 & 0 & c_3 I_{\binom{n}{2}} \end{pmatrix}, \quad (3.6.100)$$

$$\text{где } c_i = m_i^{-1}. \quad (3.6.101)$$

Значения элементов дисперсионной матрицы для различных n приведены в табл. 3.8.

Оценки коэффициентов модели рассчитываются независимо друг от друга, но поскольку элементы дисперсионной матрицы различны, то коэффициенты при разных факторах вычисляются по разным формулам:

$$\begin{aligned} \hat{a}_k &= c_1 \sum_{i=1}^N x_{ik} \tilde{y}^i, \\ \hat{a}_{kk} &= c_2 \sum_{i=1}^N (x_{ik}^2 - \beta) \tilde{y}^i, \\ \hat{a}_{kl} &= c_3 \sum_{i=1}^N x_{ik} x_{il} \tilde{y}^i, \\ \hat{b}_0 &= c_0 \sum_{i=1}^N \tilde{y}^i, \\ \hat{a}_0 &= \hat{b}_0 - \beta \sum_{k=1}^n a_{kk}. \end{aligned} \quad (3.6.102)$$

Оценки дисперсий коэффициентов определяются по формулам

$$s_i^2 = \begin{cases} s^2 c_0, & k=0 \text{ (свободный член)}, \\ s^2 c_1, & k=1, \dots, n \text{ (линейные члены)}, \\ s^2 c_2, & k=n+1, \dots, 2n \text{ (квадраты)}, \\ s^2 c_3, & k=2n+1, \dots \text{ (парные взаимодействия)}. \end{cases} \quad (3.6.103)$$

Оценка $\hat{a}_0$ имеет дисперсию

$$s_{\hat{a}_0}^2 = s^2(c_0 + n\beta^2 c_2). \quad (3.6.104)$$

Таким образом, в квадратичных моделях коэффициенты вычисляют с различной точностью, причем наибольшей погрешностью обладают коэффициенты при квадратах переменных.

Оценка значимости коэффициентов и проверка адекватности модели производится так же, как и при использовании линейных моделей.

Оптимальные планы. Критерии оптимальности. Будем считать, что вид модели задан, а обработка экспериментальных данных проводится методом наименьших квадратов. Рассмотрим качество плана с точки зрения достигнутой точности модели. Как видно из (3.6.38) и (3.6.39), точность оценок коэффициентов и выхода модели определяется матрицей S , которая при заданном виде модели зависит от расположения точек в факторном пространстве.

Задача планирования эксперимента заключается в том, чтобы наилучшим образом расположить точки в заданной области с целью выполнения определенного критерия оптимальности. При этом результат зависит от вида области. Естественно, что найти заранее планы для всех возможных областей нельзя, поэтому теория планирования эксперимента занимается планами для специальных областей (например, для гиперкуба), а также способами нахождения планов для произвольных заданных областей [129].

В зависимости от конкретной задачи могут быть использованы различные критерии оптимальности плана. Два критерия оптимальности уже встречались выше.

Критерий ортогональности плана. Экспериментальные точки выбираются таким образом, чтобы информационная матрица $M = F^T F$ была диагональной. Ортогональными планами являются планы полного факторного эксперимента и дробные реплики, а также специальным образом построенные планы второго порядка, как показано выше. Использование критерия оптимальности позволяет упростить вычисления (не нужно решать систему уравнений) и получить независимые оценки коэффициентов. Это значит, что замена любого коэффициента модели нулем не изменит оценок остальных коэффициентов. Такое свойство ортогональных планов очень полезно, когда точный вид модели неизвестен и необходимо отобрать переменные, существенно влияющие на выходную величину.

Критерий ротатабельности плана требует такого расположения точек в области планирования, при котором дисперсия σ_y^2 оценки выхода в точке x зависит только от расстояния этой точки до центра плана. Такой критерий дает равнозначность всех направлений в области планирования от центра плана. Условие ротатабельности плана получается из формулы (3.6.39):

$$f^T(x) C f(x) = \text{const}. \quad (3.6.105)$$

Ротатабельными являются линейные планы для моделей без взаимодействий и специальным образом построенные планы второго порядка.

Рассмотренные критерии обеспечивают полезные и удобные свойства оценок коэффициентов, однако они никак не связаны с требованием максимальной точности построения модели. В связи с развитием вычислительной техники стали применяться критерии, требующие более трудоемких вычислений, но обеспечивающие оптимальность планов с точки зрения точности оценок параметров модели или выходной переменной, рассчитанной по модели. Рассмотрим некоторые из них.

К р и т е р и й *А-оптимальности* требует такого выбора плана X_n , при котором матрица C имеет минимальный след, т.е. сумма диагональных элементов минимальна. Так как диагональные элементы c_{ii} пропорциональны дисперсии оценок коэффициентов, то критерий *А-оптимальности*, по существу, требует минимизации средней дисперсии оценок коэффициентов модели. Математически условие *А-оптимальности* записывается в виде

$$\text{Sp } C(X_n^*) = \min_{X_n \in \Omega_x} \text{Sp } C(X_n), \quad (3.6.106)$$

где X_n — произвольный план; X_n^* — оптимальный план; Ω_x — область планирования.

К р и т е р и й *Д-оптимальности* требует такого расположения точек в области планирования, при котором определитель матрицы C минимален. *Д-оптимальный* план минимизирует объем эллипсоида рассеяния оценок коэффициентов, так как геометрический смысл определителя — объем в n -мерном пространстве. Математически критерий *Д-оптимальности* записывается в виде

$$|C(X_n^*)| = \min_{X_n \in \Omega_x} |C(X_n)| = \min_{X_n \in \Omega_x} |(F^T F)^{-1}|. \quad (3.6.107)$$

К р и т е р и й *Г-оптимальности* требует такого расположения точек в области факторного пространства, при котором достигается наименьшая величина максимальной дисперсии оценки выходной переменной, т.е.

$$\max_y \sigma_y^2(X_n^*) = \min_{X_n \in \Omega_x} \{ \max_y \sigma_y^2(X_n) \}. \quad (3.6.108)$$

В отличие от критериев *А-* и *Д-оптимальности*, связанных с точностью коэффициентов модели, критерий *Г-оптимальности* требует максимальной точности оценки выхода.

Наибольшее распространение получил критерий *Д-оптимальности*, так как такие планы при весьма общих допущениях являются одновременно *А-* и *Г-оптимальными*. Заметим, что планы полного и дробного факторных экспериментов, используемые для построения линейных моделей с взаимодействиями любого порядка, являются одновременно *Д-, Г- и А-оптимальными*, если область планирования есть гиперкуб с вершинами в точках ± 1 , а матрица F не содержит одинаковых столбцов (т.е. оценки коэффициентов не смешаны). Это еще одно важное свойство факторного эксперимента.

О п т и м а л ь н ы е **п л а н ы** **д л я** **к в а д р а т и ч н ы х** **м о д е л е й.** Рассмотрим построение *Д-оптимальных* планов для квадратичных моделей. Если область планирования есть гиперкуб, то такие планы могут быть построены аналитически при условии, что полное количество опытов не задается, а задается частота каждой экспериментальной точки относительно

Таблица 3.9

n	E_0		E_1		E_2		r
	α	2^n	β	$n2^{n-1}$	γ	$n(n-1)2^{n-3}$	
1	0,6666	2	0,3333	1	0	0	3
2	0,5832	4	0,3206	4	0,0962	1	9
3	0,5758	8	0,2274	12	0,1968	6	26
4	0,5928	16	0,1228	32	0,2844	24	72
5	0,6170	32	0,0255	80	0,358	80	192

общего числа опытов. Такие планы называются *непрерывными*. Одной из возможных реализаций таких планов являются планы Кифера [92]. Эксперименты ставятся в следующих точках n -мерного гиперкуба: в вершинах — множество точек E_0 , в серединах ребер — множество точек E_1 , в центрах двумерных граней — множество точек E_2 . Обозначим через α частоту точек множества E_0 , через β — частоту точек множества E_1 и через γ — частоту точек множества E_2 . Количество соответствующих точек и их частоты приведены в табл. 3.9. В последнем столбце таблицы приведено общее число r точек плана. Планы Кифера строятся только для размерности факторного пространства $n \leq 5$.

D -оптимальные планы для квадратичных моделей не обладают свойством ортогональности, поэтому вычисление оценок коэффициентов модели проводится по общей формуле (3.6.30).

Аналитический путь построения непрерывных D -оптимальных планов возможен лишь в простейших случаях. На практике могут встречаться задачи, когда область планирования не имеет правильной геометрической формы. В этих случаях для построения оптимальных планов используют численные методы, добиваясь минимума дисперсионной матрицы, т.е. непосредственно используют определение D -оптимальности. Расчет производится на ЭВМ методом поиска. Сейчас в литературе по планированию эксперимента вычислено большое число непрерывных планов.

Непрерывный план не связан с определенным количеством наблюдений, а лишь задает соотношение числа опытов в различных точках области планирования. При проведении экспериментов необходимо от частот перейти к конкретному количеству опытов N , т.е. получить так называемый *точный план*. Если непрерывный план размещается в небольшом количестве точек r , то получить точный план, близкий к D -оптимальному, можно просто округлением частот (так, чтобы каждая частота задавалась правильной дробью). Тогда общий знаменатель всех частот дает общее количество опытов N , а числитель каждой дроби покажет количество опытов в данной точке плана. Такой метод дает планы с количеством экспериментов $N \geq r$.

Точные планы, содержащие заданное количество опытов, причем $N < r$, могут быть непосредственно вычислены на ЭВМ путем прямого поиска без округления непрерывных планов. Для сокращения поиска за начальный план берутся точки какого-либо непрерывного оптимального плана. Таким методом можно получить насыщенные планы с числом опытов $N = k + 1$.

В качестве примера в табл. 3.10 приведен насыщенный D -оптимальный план при $m = 3$.

Пользуясь этим планом, можно вычислить коэффициенты модели

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3. \quad (3.6.109)$$

Здесь $k + 1 = 10$, т.е. количество опытов равно размерности модели, план является насыщенным. Этот план также не ортогональный.

Таблица 3.10

i	x_1	x_2	x_3	i	x_1	x_2	x_3
1	+1	-1	+1	6	-1	0	1
2	+1	+1	-1	7	-1	0	-1
3	+1	-1	-1	8	-1	-1	0
4	+1	+1	+1	9	0	-1	1
5	-1	+1	+1	10	0	0	0

В насыщенных планах не остается степеней свободы для проверки значимости коэффициентов. Поэтому были рассчитаны точные планы, близкие к D -оптимальным, так называемые планы типа B_n . Эти планы содержат вершины n -мерного куба с координатами ± 1 . Эти точки образуют полный факторный эксперимент, их число 2^n . Кроме того, в план B_n входят центры $(n - 1)$ -мерных граней — это точки с координатами $(0, 0, \dots, \pm 1, 0, \dots, 0)$.

Таблица 3.11

i	x_1	x_2	x_3	i	x_1	x_2	x_3
1	-1	-1	-1	8	+1	+1	+1
2	+1	-1	-1	9	-1	0	0
3	-1	+1	-1	10	-1	0	0
4	+1	+1	-1	11	0	-1	0
5	-1	-1	+1	12	0	-1	0
6	+1	-1	+1	13	0	0	-1
7	-1	+1	+1	14	0	0	-1

По аналогии с ортогональными планами эти точки можно назвать звездными с величиной плеча, равной единице (при этом ортогональность, конечно, не сохраняется). Число таких точек $2n$. Общее число точек плана $N = 2^n + 2n > k + 1$, таким образом, здесь остаются степени свободы для проверки значимости коэффициентов.

В качестве примера в табл. 3.11 приведен план B_3 для определения коэффициентов в модели с тремя независимыми переменными.

Приведем краткое заключение. Для построения модели по экспериментальным данным используется метод наименьших квадратов, обеспечивающий минимум дисперсии оценок по сравнению с другими методами. При этом ошибки эксперимента должны быть нормальными независимыми случайными величинами с нулевым математическим ожиданием и стационарной дисперсией. В общем случае для определения оценок параметров модели необходимо решать систему нормальных уравнений.

Вид уравнения модели выбирается из имеющихся априорных сведений об объекте. Из экспериментального материала при любой обширности данных вид модели вывести нельзя. Сначала нужно выдвинуть гипотезу о виде зависимости $y = f(x)$, а затем, пользуясь экспериментальным материалом, проверить ее адекватность.

Если априорные сведения отсутствуют, то нужно начинать с простейших линейных моделей. Для проведения эксперимента следует использовать планы полного факторного эксперимента или дробные реплики, которые являются ортогональными и одновременно оптимальными по ряду критериев (ротатабельность, A -, G -, D -оптимальность). Ортогональность позволяет отбрасывать незначимые коэффициенты, не пересчитывая остальные.

Если линейная модель оказывается неадекватной, переходят к модели второго порядка, для начала также используя ортогональные планы. После проверки адекватности модели и значимости коэффициентов можно перейти к D -оптимальным планам для повышения точности оценок.

Таким образом, при построении модели проводится последовательный переход от простой модели к более сложным.

§ 3.7. Математическое моделирование на ЭВМ корреляционно-экстремальной системы

Методика последовательного построения моделей исследуемой системы.

- Рассмотрим общий путь построения математических моделей. В § 3.2 — 3.5 рассмотрены элементарные функциональные модели, повторяющие структуру объекта, в § 3.6 — регрессионные модели, описывающие зависимость между входом и выходом выбранной математической зависимостью. Первые отражают физику функционирования системы, вторые дают функциональную зависимость, не привязываясь к механизму функционирования. В практике анализа и синтеза систем автоматического управления используются также модели третьего типа (назовем их *аналитическими*), которые отражают наиболее важные стороны функционирования реальной системы, но в отличие от элементарных моделей описываются меньшим количеством функциональных зависимостей.

Рассмотренные математические модели образуют систему или иерархию математических моделей. При построении моделей различного уровня иерархии естественный путь состоит в переходе от более простых моделей к более сложным.

Что считать более сложной и более простой моделью? Казалось бы, естественно считать более сложной ту модель, которая содержит большее количество функциональных зависимостей, логических соотношений, т.е. описывается большим количеством операторов при составлении программы для ЭВМ. С этой точки зрения наиболее сложными будут эле-

ментные модели, а наиболее простыми — аналитические, которые часто описываются дифференциальными уравнениями первого или второго порядка.

На самом деле сложность поэлементной модели заключается только в ее громоздкости, в то время как для создания аналитической модели необходимо иметь общие закономерности, описывающие поведение объекта. Поэлементная модель получается просто последовательным математическим описанием отдельных функциональных звеньев объекта, т.е. тех основных элементов, которые выделены в § 3.1 и подробно рассмотрены в § 3.2 — 3.5. Получение же общих закономерностей для создания аналитических моделей чисто теоретическим путем часто бывает невозможно, поэтому приходится использовать результаты численного эксперимента на поэлементной модели, создавать регрессионную модель на основе обработки экспериментальных данных, а уже затем конструировать аналитическую модель.

Иными словами, если рассматривать модели по степени обобщения, абстракции основных свойств объекта, то наиболее простой окажется поэлементная модель, а наиболее сложной — аналитическая, отличающаяся высшей степенью абстракции и отражающая в наиболее сжатой форме основные свойства объекта.

Поэтому основной путь познания — от простого к сложному — применительно к математическому моделированию состоит в последовательном переходе от поэлементных функциональных моделей к регрессионным и аналитическим.

Сказанное иллюстрируется иерархией математических моделей, приведенной на рис. 3.29. На первом этапе разработки системы имеются только некоторые начальные сведения о структуре и параметрах ОЭС, полученные путем эскизной проработки. На основании этой информации строится поэлементная функциональная математическая модель, которая на этом этапе является единственным инструментом, позволяющим получить сведения о качестве работы проектируемой ОЭС. На поэлементной модели проводится исследование статических и динамических характеристик (пеленгационных, дискриминационных, переходных процессов) и уточнение структуры и параметров ОЭС для получения желаемых характеристик. Моделирование проводится по цифровым изображениям, которые на данном этапе обычно получают путем генерации на ЭВМ двумерных случайных полей. После достижения приемлемых характеристик проводится оценка работы ОЭС в различных условиях.

Второй этап начинается с момента появления первых образцов разрабатываемой системы. Целью этого этапа является идентификация поэлементной модели с объектом при одинаковых воздействиях путем настройки параметров модели. Моделирование проводится в основном по цифровым изображениям, полученным фотометрированием фотоснимков. Уточненная модель используется для сопровождающего математического моделирования, т.е. моделирования совместно с физическими испытаниями ОЭС, а также для моделирования в широких условиях применения.

Одновременно с этими двумя этапами разрабатываются модели более общего типа. С использованием методов теории планирования эксперимента строится регрессионная модель выходных сигналов ОЭС, применяемая

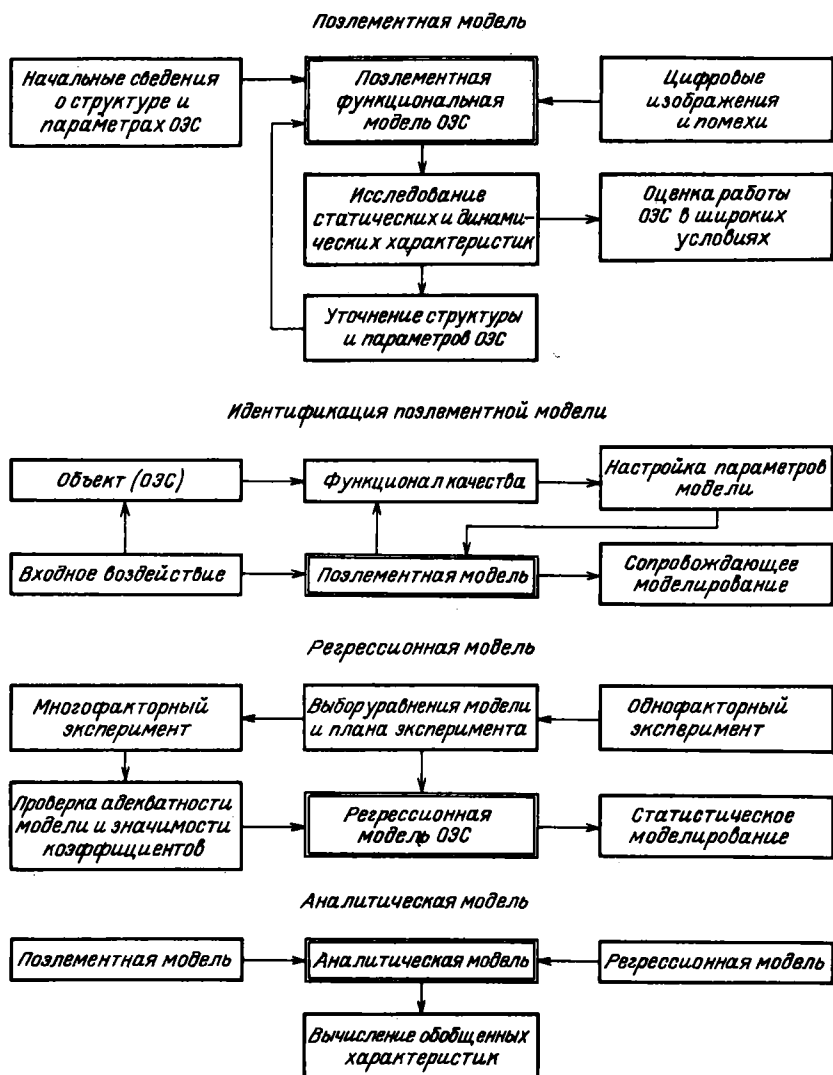


Рис. 3.29. Иерархия математических моделей

для статистического моделирования. Эта модель описывается меньшим количеством операторов по сравнению сazoleментной, поэтому позволяет провести большой объем моделирования, т.е. набрать большую статистику.

Основные уравненияazoleментной модели и регрессионные зависимости дают возможность сконструировать аналитическую модель, которая используется для вычисления обобщенных характеристик работы ОЭС, например вероятности слежения.

Далее на примере оптической корреляционно-экстремальной системы (ОКЭС) приведены рассмотренные математические модели и результаты моделирования.

Позлементная модель ОКЭС и результаты моделирования. Позэлементная функциональная модель ОКЭС. На рис. 3.30 приведена блок-схема модели оптической корреляционно-экстремальной системы с дифференциальным коррелятором, рассмотренным в [11]. Оптико-электронная система состоит из датчика изображения, усилителя сигнала изображения, коррелятора и блока памяти. Позлементная математическая модель содержит такие же функциональные блоки с добавлением цифрового поля, характеризующего распределение яркости в плоскости предметов.

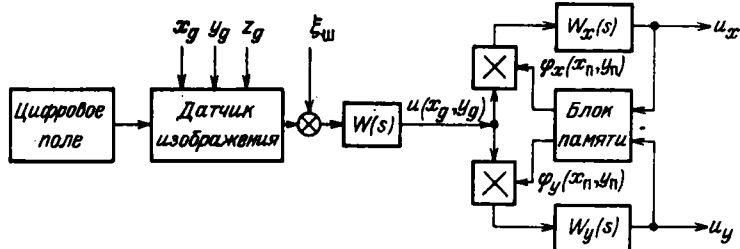


Рис. 3.30. Блок-схема оптической корреляционно-экстремальной системы

Рассмотрим подробнее математическое описание каждого из блоков модели.

Входным сигналом модели является цифровое поле, полученное методами, изложенными в § 3.3. Для моделирования использовались цифровые поля, получаемые генерированием на ЭВМ по заданным статистическим характеристикам, а также оцифрованные фотоснимки реальных ландшафтов. Моделирование по фотоснимкам необходимо в связи с тем, что получаемые генерированные изображения более гладкие по сравнению с реальными, которые обычно локально нестационарны и неизотропны, а качество работы оптико-электронной системы часто сильно зависит от степени неоднородности поля.

Цифровые поля задавались с шагом дискретности по координатам $\Delta x = \Delta y = 1,5$ м и квантовались по амплитуде на 128 уровней, т.е. представлялись семью двоичными разрядами. Пересчет прозрачности негатива в яркость поля производился по формуле

$$B(x, y) = B_{\min} \left[\frac{T_{\max}}{T(x, y)} \right]^{1/\gamma}, \quad (3.7.1)$$

где $T(x, y)$ — поле прозрачности негатива; $B(x, y)$ — поле яркости; $T_{\max}$ — максимальная прозрачность, измеренная денситометром; $B_{\min}$ — минимальная яркость, соответствующая известным условиям съемки; γ — коэффициент контрастности негатива.

Полученный цифровой массив подвергался кодированию в пространственной области путем упаковки по четыре числа в одно машинное слово и хранился на магнитной ленте. Для экономии оперативной памяти при моделировании в ОЗУ переписывалась только необходимая в данный момент часть цифрового изображения, а при выходе за пределы изображения использовалось четное продолжение.

На рис. 3.31 приведена оценка корреляционной функции использованного изображения, которая хорошо аппроксимируется выражением

$$R(x) = e^{-\frac{|x|}{r_k}}, \quad (3.7.2)$$

где радиус корреляции $r_k = 4,5$ м. Шаг дискретизации в данном примере равен $\Delta = 0,33 r_k$, что соответствует соотношению (3.3.7).

Датчик изображения состоит из оптической системы и фотоприемника. Поскольку ориентация датчика изображения относительно плоскости

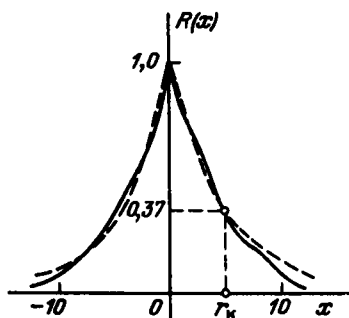


Рис. 3.31. Корреляционная функция изображения

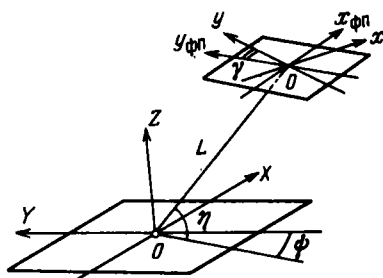


Рис. 3.32. Геометрия пространственной задачи

предметов (яркостного поля) может быть произвольной, прежде всего необходимо поставить в соответствие точки в плоскости фотоприемника с точками в плоскости предметов, т.е. учесть ракурсные искажения и масштабное преобразование. Геометрически задача иллюстрируется рис. 3.32.

- Пересчет координат из плоскости поля яркости (система координат OXY) в плоскость фотоприемника (система координат $OX_{фп} Y_{фп}$) производится по формулам, применяемым, в частности, в фотограмметрии:

$$x_{фп} = -f \frac{a_1 x + a_2 y}{L - c_1 x - c_2 y}, \quad (3.7.3)$$

$$y_{фп} = -f \frac{b_1 x + b_2 y}{L - c_1 x - c_2 y}. \quad (3.7.4)$$

Здесь f — фокусное расстояние объектива, а

$$\begin{aligned} a_1 &= \cos \gamma \cos \psi - \sin \gamma \sin \eta \sin \psi, & a_2 &= \sin \gamma \cos \eta, \\ b_1 &= -\sin \gamma \cos \psi - \cos \gamma \sin \eta \sin \psi, & b_2 &= \cos \gamma \cos \eta, \\ c_1 &= \cos \eta \sin \psi, & c_2 &= \sin \eta. \end{aligned} \quad (3.7.5)$$

Оптическая система моделировалась разделимой пространственно-инвариантной ФРТ типа (3.2.4) с аппроксимацией гауссоидой вращения (3.2.15)

$$g(x, y) = 1,25 r_0^{-2} \exp \{-4r_0^{-2}(x^2 + y^2)\}. \quad (3.7.6)$$

Изображение на выходе оптической системы вычислялось путем дискретной двумерной свертки (3.2.28)

$$B_{\text{и}}(x_i, y_j) = \sum_{\alpha} \sum_{\beta} B_0(x_{\alpha}, y_{\beta}) g(x_i - x_{\alpha}, y_j - y_{\beta}) \Delta x \Delta y$$

или с учетом разделимости в матричной форме (3.2.33)

$$B_{\text{и}} = G_y B_0 G_x.$$

Шум датчика изображения $\xi_{\text{ш}}$ моделировался как нормальный стационарный случайный процесс с корреляционной функцией вида

$$K_{\text{ш}}(\tau) = \sigma_{\text{ш}}^2 e^{-\alpha_{\text{ш}} |\tau|}, \quad (3.7.7)$$

который соответствует формирующий фильтр с передаточной функцией

$$W_{\text{ш}}(s) = \frac{k}{1 + sT_{\text{ш}}}, \quad T_{\text{ш}} = \frac{1}{\alpha_{\text{ш}}}. \quad (3.7.8)$$

Усилитель сигнала изображения обычно содержит полосовой фильтр с передаточной функцией

$$W(s) = W_1(s) W_2(s) = \frac{T_{\text{н}} s}{1 + T_{\text{н}} s} \frac{k_y}{1 + T_{\text{в}} s}, \quad (3.7.9)$$

обеспечивающей пропускание частот в диапазоне $\Delta \omega = \omega_{\text{в}} - \omega_{\text{н}}$, где

$$\omega_{\text{в}} = \frac{1}{T_{\text{в}}}, \quad \omega_{\text{н}} = \frac{1}{T_{\text{н}}}.$$

Кроме того, для компенсации изменения уровня сигнала при перепадах яркости изображения усилитель видеосигнала имеет систему автоматической регулировки усиления (АРУ), которая состоит из нелинейного звена в контуре с обратной связью. Методика моделирования на ЭВМ усилителя с АРУ рассмотрена в [99].

При проектировании системы АРУ задается амплитудная характеристика регулируемого участка, т.е. зависимость амплитуды выходного напряжения $U_{\text{вых}}$ от амплитуды входного $U_{\text{вх}}$, структурная схема и уравнения элементов АРУ. Для обеспечения заданной амплитудной характеристики необходимо определить регулировочную характеристику, т.е. зависимость коэффициента усиления регулируемого участка тракта от величины сигнала регулирования. Использование ЭВМ в качестве инструмента исследования дает возможность синтезировать регулировочную характеристику по заданной амплитудной.

В общем случае структурная схема автоматической регулировки усиления может быть представлена в виде, изображенном на рис. 3.33. В этой схеме элемент с регулируемым коэффициентом усиления k_v присоединен к цепочке линейных фильтров $\Phi_1(\omega), \dots, \Phi_n(\omega)$, охваченных нелинейной обратной связью. Типовая зависимость $U_{\text{вых}}(U_{\text{вх}})$ представляет собой кривую с линейной зоной и участком насыщения. Для ввода в ЭВМ эта характеристика задается в виде таблицы, из которой выбирают пару $(U_{\text{вх}i}, U_{\text{вых}i})$ и вычисляют $k_{vi} = U_{\text{вых}i}/U_{\text{вх}i}$. Это значение заносится в ячейку памяти. Затем на вход АРУ подается тест-сигнал с амплитудой $U_{\text{вх}i}$ и производится вычисление переходного процесса до установления на его выходе $U_{\text{вых}i} =$

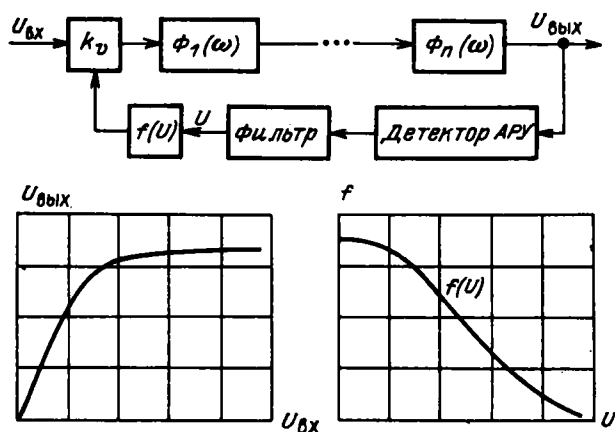


Рис. 3.33. Схема системы автоматической регулировки усиления

$= \text{const}$. В ОЗУ при этом заносится значение U_i , соответствующее k_{vi} . После того как все заданные пары $(U_{вх\ i}, U_{вых\ i})$ будут исчерпаны, в ОЗУ будет содержаться соответствующее число пар условных значений (U_i, k_{vi}) регулировочной характеристики $f(U)$.

В дальнейшем, используя стандартную программу аппроксимации, например по методу Лагранжа, получают непрерывную характеристику, которая в заданном диапазоне регулирования применяется при математическом моделировании. Рассмотренный метод справедлив для каскадов АРУ с задержкой и без задержки при произвольном числе фильтров с любыми характеристиками детектора.

Моделирование линейных звеньев в усилителе сигнала изображения, АРУ и корреляторе производится методом дискретной аппроксимации с использованием z -форм Рагазини—Бергена (3.4.45). При этом фильтры с передаточными функциями $W_1(s)$ и $W_2(s)$ в (3.7.9) заменяются цифровыми фильтрами с дискретными передаточными функциями

$$W_1(z) = k_1 \frac{1 - z^{-1}}{1 - b_1 z^{-1}}, \quad W_2(z) = k_2 \frac{1 + z^{-1}}{1 - b_2 z^{-1}}, \quad (3.7.10)$$

где

$$k_1 = \frac{2T_H}{2T_H + \Delta t}; \quad b_1 = \frac{2T_H - \Delta t}{2T_H + \Delta t};$$

$$k_2 = \frac{1}{1 + \frac{2T_B}{\Delta t}}; \quad b_2 = \frac{2T_B - \Delta t}{2T_B + \Delta t}.$$

Используя передаточные функции (3.7.10), получаем рекуррентные соотношения для вычисления сигнала на выходе фильтров, учитывая, что опе-

ратор z^{-1} обозначает задержку на один шаг:

$$\begin{aligned} y_n &= k_1 x_n - k_1 x_{n-1} + b_1 y_{n-1}, \\ y_n &= k_2 x_n + k_2 x_{n-1} + b_2 y_{n-1}. \end{aligned} \quad (3.7.11)$$

Коррелятор, работающий по дифференциальной схеме, состоит из перемножителей, линейной части с передаточными функциями $W_x(s)$, $W_y(s)$ и блока памяти, в котором находится цифровое изображение контрольного участка цифрового поля. В дифференциальной схеме сигналы в блоке памяти выбираются в четырех точках, отстоящих от точки (x_n, y_n) на некоторое фиксированное расстояние $\pm l$, а затем образуются разности этих сигналов:

$$\begin{aligned} \varphi_x(x_n, y_n) &= f(x_n + l, y_n) - f(x_n - l, y_n), \\ \varphi_y(x_n, y_n) &= f(x_n, y_n + l) - f(x_n, y_n - l). \end{aligned} \quad (3.7.12)$$

Дифференциальный коррелятор относится к классу беспойсковых корреляционно-экстремальных систем и работает при начальных рассогласованиях между точками (x_g, y_g) и (x_n, y_n) в пределах интервала корреляции поля яркости.

Теоретический анализ работы дифференциального коррелятора проведен в [11] для стационарных нормальных полей. В данном параграфе приводятся некоторые результаты математического моделирования по цифровым изображениям.

Идентификация поэлементной модели. Как показано выше, один из этапов разработки математической модели оптико-электронной системы является ее идентификация с реальным изделием. Задача идентификации формулируется следующим образом: по результатам наблюдений над входными и выходными сигналами объекта должна быть построена оптимальная в некотором смысле модель. В дальнейшем будем рассматривать идентификацию в узком смысле [143], так как структура объекта при построении поэлементной модели известна и необходимо произвести оценку параметров (иногда при этом оценивается и вектор состояния системы).

Задаче оценивания параметров можно дать следующее схематическое описание. На объект и модель действует один и тот же входной сигнал. Сравниваются выходные сигналы объекта и модели. Параметры объекта непосредственному измерению недоступны. Таким образом, критерием выбора оптимальной оценки должен быть функционал от входных сигналов, который связан с численными значениями параметров. В качестве критерия наиболее часто выбирается скалярная функция ошибки, в дальнейшем называемая *функцией качества*:

$$Q = \int_0^T [y(b, t) - y_m(\underline{\beta}, t)]^2 dt, \quad (3.7.13)$$

где y — выход объекта; y_m — выход модели; b — вектор параметров объекта; $\underline{\beta}$ — вектор параметров модели.

Положив

$$\frac{\partial Q}{\partial \beta} = 0, \quad (3.7.14)$$

получим условие минимума функции качества, из которого можно найти оценки β_i неизвестных параметров b_i . Важным свойством формулы (3.7.13) является то, что оценка по методу наименьших квадратов единственная [120].

В сложных моделях не удастся найти аналитическое выражение функции качества в зависимости от параметров, поэтому приходится вычислять функцию качества путем моделирования на ЭВМ и проводить перебор параметров по какому-либо алгоритму, т.е. производить поисковую многопараметрическую оптимизацию. В отличие от точного решения (3.7.14), в данном случае производная функции качества стремится к нулю, т.е.

$$\frac{\partial Q}{\partial \beta} \rightarrow 0. \quad (3.7.15)$$

Такой метод называется *идентификацией с использованием настраиваемой модели*. Методы поиска могут быть регулярными и случайными. Если учесть, что исследуемые поэлементные модели являются в большинстве случаев нелинейными по динамике и параметрам, то наиболее эффективным оказывается случайный поиск [105].

В дальнейшем рассматривается применение случайного поиска для идентификации поэлементной модели. Используется алгоритм случайного поиска по наилучшей пробе. Из исходной точки $B(\beta_{10}, \beta_{20}, \dots, \beta_{k0})$ делается m случайных независимых шагов (проб) в пространстве параметров модели и запоминается тот шаг, который привел к наибольшему снижению функции качества. Рабочий шаг делается в этом направлении. С увеличением числа проб выбранное направление все более приближается к направлению, обратному градиентному, и в пределе при $m \rightarrow \infty$ совпадает с ним.

Блок-схема идентификации с настраиваемой моделью приведена на рис. 3.29.

Для отработки методики вместо изделия использовалась также математическая модель с известными параметрами, а значения ее состояний использовались для вычисления функции качества как состояния объекта. В дальнейшем эту модель будем называть *эталонной*.

По степени сложности идентификацию можно разделить на идентификацию по статическим характеристикам и по динамическим процессам. В качестве статической характеристики рассматривалась дискриминационная характеристика коррелятора, т.е. зависимость выходного сигнала от сдвига между текущим и эталонным изображениями. Идентификация проводилась с целью настройки модели по заданной дискриминационной характеристике объекта, зависящей от пяти параметров (два коэффициента усиления и три постоянных времени). Функция качества определялась следующим образом:

$$Q = \frac{1}{n} \sum_{i=1}^n [u(x_i) - u_m(x_i)]^2, \quad (3.7.16)$$

где $u(x_i)$ — выход объекта; $u_m(x_i)$ — выход модели; x_i — дискретные значения координаты, по которой снимается характеристика.

Программа поиска работала по пяти пробам, останов — после пяти подряд идущих неудачных шагов. В процессе поиска было сделано 15 удачных шагов ($N_{уд}$) и 6 неудачных ($N_{неуд}$); уменьшение функции качества со-

Таблица 3.12

β_i	β_1	β_2	β_3	β_4	β_5
β_{i0}	2,1	1,5	2,5	1,5	1,6
$\Delta\beta_i$	0,09	-0,04	-0,22	0,19	0,48

ставило $\mu = Q_n/Q_k = 67$, где Q_n — начальное значение функции (3.7.16), а Q_k — ее конечное значение.

В табл. 3.12 приведены начальные значения параметров β_{i0} в настраиваемой модели и отклонения $\Delta\beta_i$ полученных экстремальных значений от эталонных. Как показали дальнейшие исследования, дискриминационная характеристика наиболее чувствительна к первым двум параметрам, что и проявилось в наиболее точном определении их значений. Дискриминационные характеристики приведены на рис. 3.34.

В качестве динамического процесса при идентификации поэлементной модели рассматривался переходный процесс, получаемый при постоянном входном воздействии. Реакция модели зависела от восьми параметров (четыре коэффициента усиления и четыре постоянных времени), начальные значения параметров отличались от эталонных в 2–3 раза. Программа поиска работала по пяти пробам, конечный останов — по 50 неудачным шагам.

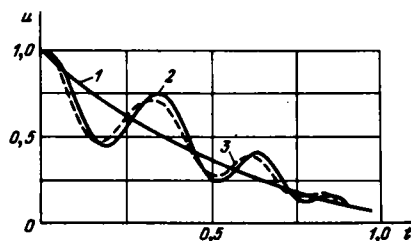
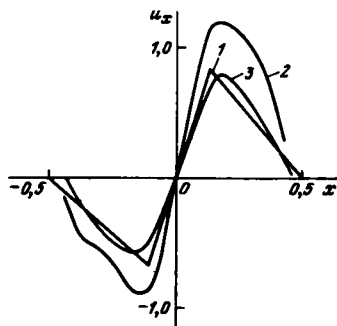


Рис. 3.34. Идентификация по дискриминационным характеристикам: 1 — заданная дискриминационная характеристика; 2 — соответствующая начальным параметрам β_{i0} в схеме поиска; 3 — полученная дискриминационная характеристика

Рис. 3.35. Идентификация по переходному процессу

Функция качества формировалась двумя способами. В первом использовался только выходной сигнал u в соответствии с формулой (3.7.16). Получены следующие параметры поиска: $N_{уд} = 83$; $N_{неуд} = 263$; $\mu = 1,1 \cdot 10^3$; среднее квадратическое отклонение по всем параметрам от эталонных равно 1,27. Сигналы в настраиваемой и эталонной моделях сильно различаются (соответственно кривые 1 и 3 на рис. 3.35); поиск неудачен.

Идентификация по одному сигналу тесно связана с понятиями наблюдаемости и идентифицируемости. В сложном контуре с обратными связями

не все координаты могут быть определены по измерению только выходного сигнала, а значит, и не все параметры идентифицируются. Поэтому желательно проводить идентификацию по возможно большему числу сигналов (переменных состояния).

Во втором способе функция качества формировалась по восьми сигналам:

$$Q = \sum_{i=1}^n \sum_{j=1}^8 a_j (u_{ij} - u_{ijm})^2. \quad (3.7.17)$$

Весовые коэффициенты каждого сигнала a_j выбраны из условия приведения сигналов к одному уровню. Здесь использовано условие независимости оптимальной оценки от весовой матрицы [120]. Настройка модели проводилась в два этапа: с произвольными начальными условиями и повышенной дисперсией поиска, затем поиск продолжался с уменьшенной дисперсией. Отклонения полученных значений параметров от эталонных приведены в табл. 3.13. Характеристики поиска следующие: $N_{уд} = 106$; $N_{неуд} = 271$; $\mu = 8,5 \cdot 10^4$. Выходной сигнал приведен на рис. 3.35 (кривая 2). Совпадение сигналов в эталонной и настраиваемой моделях хорошее, поэтому можно сделать вывод, что параметры, наиболее сильно влияющие на переходный процесс, определены точно.

В приведенных примерах рассмотрены самые сложные случаи, когда все параметры предполагались неизвестными, поэтому поиск начинался с произвольных начальных условий. На практике обычно известны номинальные значения параметров объекта, поэтому ставится задача оценить отличие реальных параметров от номинальных. Часто ставится еще более узкая задача: найти параметр, сильно отличающийся от номинального, т.е. произвести поиск неисправного параметра. Здесь поиск начинается с номинальных значений параметров. Проведенное моделирование с целью поиска параметров, отличающихся от номинала в 20 и 6 раз, позволило уверенно идентифицировать эти параметры как неисправные (отклонение найденных значений от действительных величин составило 3% и 4%).

Таким образом, использование случайного поиска позволяет успешно решать задачу идентификации поэлементной функциональной математической модели.

Исследование перекрестных связей. В двухканальных корреляторах при работе по двумерному полю возникают перекрестные связи, т.е. сигнал в каждом канале зависит от сигнала в ортогональном канале. Возникновение перекрестных связей объясняется следующим образом. Рассмотрим нормированную анизотропную двумерную корреляционную функцию, оси симметрии которой не совпадают с осями управления коррелятора x и y . Пусть для определенности аналитическое описание

Таблица 3.13

β_i	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8
$\Delta\beta$	0,15	0,06	0,03	0,47	0,37	0,49	1,05	0,925

такой корреляционной функции имеет вид экспоненты с эллиптической анизотропией

$$R(x, y) = \exp \{ -(a_{11}x^2 + 2a_{12}xy + a_{22}y^2) \}. \quad (3.7.18)$$

Выходной сигнал дифференциального коррелятора можно записать как производную от корреляционной функции двумерного поля

$$u_x = \frac{\partial R(x, y)}{\partial x} = -2(a_{11}x + a_{12}y) R(x, y), \quad (3.7.19)$$

$$u_y = \frac{\partial R(x, y)}{\partial y} = -2(a_{12}x + a_{22}y) R(x, y).$$

Следовательно, сигнал в каждом канале управления зависит также и от величины сдвига в ортогональном канале. Действительно, если имеется сдвиг y_0 , то выходной сигнал по каналу x равен нулю в точке

$$x = -y_0 \frac{a_{12}}{a_{11}}. \quad (3.7.20)$$

Аналогично для канала y имеем

$$y = -x_0 \frac{a_{12}}{a_{22}}, \quad (3.7.21)$$

т.е. получается сдвиг нуля дискриминационной характеристики в каждом из каналов. Таким образом, двумерность поля приводит к возникновению перекрестных связей, величина которых зависит от угла поворота ψ осей симметрии двумерной корреляционной функции относительно каналов управления:

$$\operatorname{tg} 2\psi = \frac{2a_{12}}{a_{11} - a_{22}}. \quad (3.7.22)$$

При $a_{12} = 0$ оси симметрии корреляционной функции совпадают с каналами управления и перекрестные связи отсутствуют.

На поэлементной математической модели исследовалась величина перекрестных связей и их влияние на работу замкнутой системы автоматического управления. Величина перекрестных связей исследовалась путем снятия дискриминационных характеристик, т.е. напряжения на входе коррелятора при сдвиге эталонного и текущего изображений.

На рис. 3.36 приведены дискриминационные характеристики при работе по двум изображениям. По оси x отложена величина рассогласования в относительных единицах по отношению к полю зрения, по оси ординат — напряжение на выходе каналов x и y . Сплошными кривыми показаны дискриминационные характеристики: u_{xx} — в канале x при рассогласовании по x , u_{yy} — в канале y при рассогласовании по y . Пунктирными кривыми даны сигналы перекрестных связей: u_{yx} — сигналы в канале y при рассогласовании по x , u_{xy} — сигнал в канале x при рассогласовании по y . На графиках рис. 3.36, a сигнал перекрестных связей меньше основного сигнала, осо-

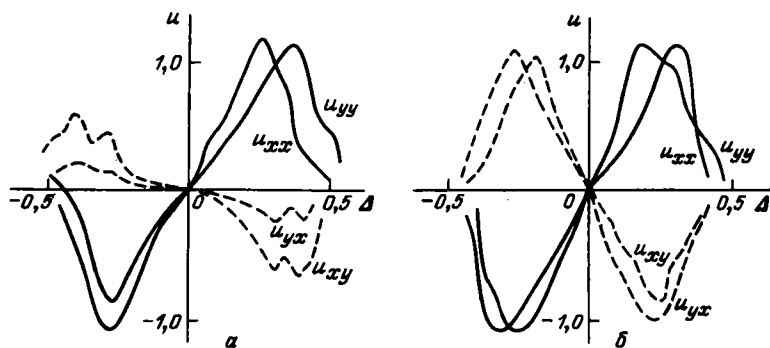


Рис. 3.36. Дискриминационные характеристики с перекрестными связями

бенно при небольших рассогласованиях; на графиках рис. 3.36, б перекрестные связи одного порядка с основным сигналом.

Включение коррелятора с перекрестными связями в систему автоматического управления может привести к потере устойчивости. Рассмотрим в общем виде вопрос об устойчивости двумерной следящей системы, не используя свойство эллиптической анизотропии поля. Единственное требование, которому должна удовлетворять двумерная корреляционная функция стационарного поля, — наличие центральной симметрии, т.е. любое проходящее через начало координат сечение должно быть четной функцией аргумента. Это свойство вытекает из свойств одномерной функции корреляции.

В соответствии с (3.7.19) сигналы управления по каналам равны соответственно $R'_x(x, y)$ и $R'_y(x, y)$. Если ввести в рассмотрение поле сил коррекции, то наклон силовой линии (т.е. направление регулирования) равен

$$\operatorname{tg} \theta = \frac{dy}{dx} = \frac{R'_y(x, y)}{R'_x(x, y)}. \quad (3.7.23)$$

Найдем наклон касательной к линии уровня $R(x, y) = \text{const}$. Так как при этом

$$dR(x, y) = R'_x(x, y)dx + R'_y(x, y)dy = 0,$$

то

$$\operatorname{tg} \nu = \frac{dy}{dx} = -\frac{R'_x(x, y)}{R'_y(x, y)}. \quad (3.7.24)$$

Сравнивая (3.7.23) и (3.7.24), убеждаемся, что $\operatorname{tg} \nu \operatorname{tg} \theta = -1$, т.е. $|\theta - \nu| = 90^\circ$, следовательно, корреляционное усилие действует внутрь линии уровня по нормали к ней.

Разложим функции $R'_x(x, y)$ и $R'_y(x, y)$ в окрестности начала координат:

$$R'_x(x, y) = R'_x(0, 0) + R''_{xx}(0, 0)x + R''_{xy}(0, 0)y + r_x(x, y),$$

$$R'_y(x, y) = R'_y(0, 0) + R''_{yx}(0, 0)x + R''_{yy}(0, 0)y + r_y(x, y). \quad (3.7.25)$$

Здесь остатки ряда удовлетворяют условию

$$\lim_{x^2+y^2 \rightarrow 0} \frac{r_x(x, y)}{\sqrt{x^2+y^2}} = 0, \quad \lim_{x^2+y^2 \rightarrow 0} \frac{r_y(x, y)}{\sqrt{x^2+y^2}} = 0, \quad (3.7.26)$$

так как для четной двумерной корреляционной функции справедливо соотношение $R'_x(0, 0) = R'_y(0, 0) = 0$.

Обозначим

$$\begin{aligned} R''_{xy}(0, 0) &= k_1, & R''_{yx}(0, 0) &= k_2, \\ R''_{xx}(0, 0) &= \mu_1, & R''_{yy}(0, 0) &= \mu_2. \end{aligned} \quad (3.7.27)$$

Тогда с точностью до слагаемых r_x , r_y можно записать

$$R'_x(x, y) = \mu_1 x + k_1 y, \quad R'_y(x, y) = \mu_2 y + k_2 x. \quad (3.7.28)$$

Структура линейной следящей двумерной системы с перекрестными связями приведена на рис. 3.37.

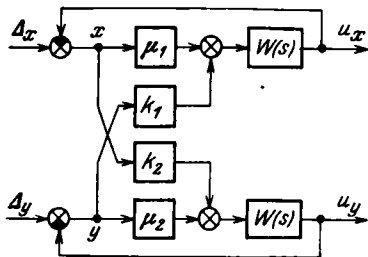


Рис. 3.37. Структурная схема следящей системы с перекрестными связями

Положим для определенности $W(s) = 1/s$. Тогда уравнения системы в операторной форме запишутся в виде

$$(s + \mu_1)X + k_1 Y = 0, \quad k_2 X + (s + \mu_2)Y = 0, \quad (3.7.29)$$

а характеристическое уравнение примет форму

$$\lambda^2 + (\mu_1 + \mu_2)\lambda + \mu_1\mu_2 - k_1k_2 = 0. \quad (3.7.30)$$

Корни характеристического уравнения равны

$$\lambda_{1,2} = -\frac{b}{2} \pm \sqrt{D}, \quad (3.7.31)$$

где

$$b = \mu_1 + \mu_2; \quad D = \sqrt{\frac{b^2}{4} - c}; \quad c = \mu_1\mu_2 - k_1k_2. \quad (3.7.32)$$

Система устойчива, если корни характеристического уравнения имеют отрицательную вещественную часть. Вещественные корни имеют место при выполнении условия $D \geq 0$, или с учетом (3.7.31), (3.7.32)

$$k_1k_2 > -\frac{(\mu_1 - \mu_2)^2}{4}. \quad (3.7.33)$$

Это условие выполняется всегда, так как $k_1k_2 > 0$. Далее, если $\frac{b}{2} > \sqrt{D}$,

то корни характеристического уравнения отрицательные. Последнее неравенство эквивалентно условию

$$k_1 k_2 < \mu_1 \mu_2. \quad (3.7.34)$$

Таким образом, выражение (3.7.34) задает ограничение на параметры корреляционной функции двумерного поля, при которых система управления является устойчивой.

Аналогичный результат можно получить из рассмотрения поля сил коррекции, не вводя в явном виде передаточную функцию $W(s)$. Ограничиваясь выражениями (3.7.28), запишем согласно (3.7.23)

$$\frac{dy}{dx} = \frac{k_2 x + \mu_2 y}{\mu_1 x + k_1 y}. \quad (3.7.35)$$

Вид фазовой траектории, а следовательно, и устойчивость системы зависят от типа особой точки $x = y = 0$. Тип особой точки определяется корнями характеристического уравнения [94]

$$\lambda^2 - (\mu_1 + \mu_2)\lambda + \mu_1 \mu_2 - k_1 k_2 = 0, \quad (3.7.36)$$

отличающегося от (3.7.30) только знаком второго члена. Поэтому из (3.7.33) следует, что корни уравнения (3.7.36) всегда действительные. Далее, если корни разных знаков, то особая точка — седло и система неустойчива; если же корни одинаковых знаков, то особая точка — узел, траектории проходят через начало координат, начальное рассогласование отрабатывается следящей системой. Условие одинаковых знаков корней характеристического уравнения (3.7.36) обеспечивается при выполнении неравенства (3.7.34), которое является, таким образом, необходимым условием для устойчивости двухканального коррелятора.

Рассмотренный анализ влияния перекрестных связей не учитывает локальных неоднородностей поля. Для учета последних было проведено прямое математическое моделирование с использованием поэлементной модели ОЭС. Моделирование проводилось по двум полям с различной

Таблица 3.14

Поле	k_1	k_2	μ_1	μ_2
1	0,35	0,14	0,88	0,7
2	0,9	1,1	1,0	0,7

степенью анизотропии. В разомкнутой системе по дискриминационным характеристикам (см. рис. 3.36) вычислены коэффициенты k и μ , приведенные в табл. 3.14. Проверка условия устойчивости (3.7.34) дает для первого поля

$$k_1 k_2 = 0,049 < 0,616 = \mu_1 \mu_2$$

и для второго

$$k_1 k_2 = 0,99 > 0,7 = \mu_1 \mu_2.$$

Для второго поля условие устойчивости не выполняется. Моделирование замкнутой системы показало, что при работе по второму полю начальное рассогласование не отрабатывается, происходит срыв слежения. Это иллюстрируется рис. 3.38, где приведены фазовые траектории слежения. Начальное рассогласование (точка A), одинаковое для обоих полей, отрабатывается на первом поле (траектория приходит в начало координат), и процесс является расходящимся для второго поля.

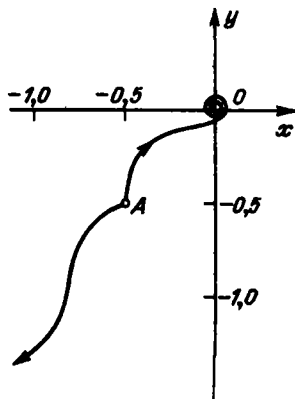


Рис. 3.38. Фазовые траектории слежения

Здесь рассмотрена устойчивость двумерной системы при начальных ошибках, не превышающих величины линейной зоны дискриминационной характеристики. Срыв слежения при больших отклонениях рассмотрен ниже при исследовании аналитической модели.

Регрессионная модель коррелятора. Построение линейной модели для пяти переменных. Поэлементная математическая модель может быть успешно использована для исследования влияния декоррелирующих факторов на точность работы коррелятора. *Декоррелирующими факторами* будем называть все внешние возмущения, приводящие к искажению текущего изображения относительно эталонного. К ним относятся прежде всего размасштабирование ΔM , поворот на угол крена γ , сдвиг в ортогональном канале Δu , ракурсные искажения — повороты на углы визирования η и курса ψ . Наличие перечисленных рассогласований текущего изображения относительно эталонного приводит к падению крутизны и сдвигу нуля дискриминационной характеристики и даже к ее полному разрушению при большой величине искажений. Поэтому оценка влияния декоррелирующих факторов и ее математическое описание являются одной из типичных задач математического моделирования на поэлементной модели ОЭС.

Предварительная оценка проводится обычно в однофакторном эксперименте при варьировании какого-либо одного из декоррелирующих факторов. Целью такого моделирования является выявление качественного вида экспериментальных зависимостей для их дальнейшего уточнения в многофакторном эксперименте. На рис. 3.39 приведены графики изменения крутизны S и сдвига нуля x_c дискриминационной характеристики при варьировании размасштабирования ΔM . Под размасштабированием

понималась величина

$$\Delta M = \frac{M_T}{M_0} - 1, \quad (3.7.37)$$

где M_T , M_0 — масштабы текущего и эталонного изображений.

Результаты однофакторного эксперимента показали, что при малых величинах декоррелирующих факторов можно ограничиться линейным

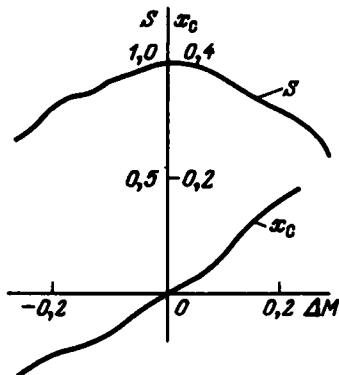


Рис. 3.39. Влияние размасштабирования на параметры дискриминационной характеристики

описанием экспериментальных зависимостей, при больших рассогласованиях необходимо использовать уравнения регрессии более высокого порядка.

Рассмотрим построение линейной регрессионной модели изменения крутизны характеристики для пяти независимых переменных: $x_1 = \Delta M$ — рассогласование по масштабу; $x_2 = \Delta \gamma$ — рассогласование по крену; $x_3 = \Delta \psi$ — рассогласование по курсу; $x_4 = \Delta \eta$ — рассогласование по углу визирования; $x_5 = \Delta y$ — сдвиг в ортогональном канале. В уравнениях регрессии наряду с линейными членами учтем парные взаимодействия, т.е. выберем модель вида

$$\begin{aligned} \tilde{S} = & a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 + a_5 x_5 + a_{12} x_1 x_2 + \\ & + a_{13} x_1 x_3 + a_{14} x_1 x_4 + a_{15} x_1 x_5 + \\ & + a_{23} x_2 x_3 + a_{24} x_2 x_4 + a_{25} x_2 x_5 + \\ & + a_{34} x_3 x_4 + a_{35} x_3 x_5 + a_{45} x_4 x_5. \end{aligned} \quad (3.7.38)$$

Здесь $\tilde{S} = S/S_{\max}$ — нормированное значение крутизны; $S_{\max}$ — крутизна при отсутствии декоррелирующих факторов. Модель (3.7.38) содержит свободный член, который априори равен единице, пять линейных членов и десять парных взаимодействий. Полный факторный эксперимент содержит $N = 2^5 = 32$ опыта. Для сокращения количества опытов была использована главная полуреплика 2^{5-1} , которая содержала 16 опытов. Генератор задавался соотношением

$$x_5 = x_1 x_2 x_3 x_4, \quad (3.7.39)$$

откуда умножением на x_5 , получен контраст

$$I = x_1 x_2 x_3 x_4 x_5. \quad (3.7.40)$$

Найдем совместные оценки, получаемые в таком плане, путем умножения контраста (3.7.40) на переменные модели (3.7.38):

$$\begin{aligned}x_1 &= x_2 x_3 x_4 x_5, & x_1 x_2 &= x_3 x_4 x_5, \\x_2 &= x_1 x_3 x_4 x_5, & x_1 x_3 &= x_2 x_4 x_5, \\x_3 &= x_1 x_2 x_4 x_5, & x_1 x_4 &= x_2 x_3 x_5, \\x_4 &= x_1 x_2 x_3 x_5, & x_1 x_5 &= x_2 x_3 x_4, \\x_5 &= x_1 x_2 x_3 x_4, & x_2 x_3 &= x_1 x_4 x_5,\end{aligned}\tag{3.7.41}$$

Здесь линейные эффекты смешаны с четверными взаимодействиями, а парные взаимодействия — с тройными. Но так как тройные взаимодействия и выше можно считать равными нулю, то в этой полуреплике линейные эффекты и парные взаимодействия оцениваются раздельно, что и надо для нашей модели.

Составим полуреплику 2^{5-1} . Для этого используем полный факторный эксперимент 2^4 , задаваемый строкой

$$(1), a, b, ab, c, ac, bc, abc, d, ad, bd, abd, cd, acd, bcd, abcd.$$

Главная полуреплика содержит строки только с четными или нечетными сочетаниями букв. Построим полуреплику с четными сочетаниями, для чего строку плана 2^4 умножим на букву e в тех сочетаниях, где количество букв нечетное. Получим одну из главных полуреплик 2^{5-1}

$$(1), ae, be, ab, ce, ac, bc, abce, de, ad, bd, abde, cd, acde, bcde, abcd.$$

Используя кодовую строку, построим матрицу планирования, проставив в ней плюсы там, где соответствующий фактор находится на верхнем уровне, т.е. имеется в сочетаниях кодовой строки. Матрица планирования приведена в табл. 3.15. Чтобы получить матрицу F , надо матрицу планирования

Таблица 3.15

i	x_1	x_2	x_3	x_4	x_5	Код строки	$\tilde{y}^i$
1	—	—	—	—	—	(1)	1,0
2	+	—	—	—	+	<i>ae</i>	0,6
3	—	+	—	—	+	<i>be</i>	0,7
4	+	+	—	—	—	<i>ab</i>	0,7
5	—	—	+	—	+	<i>ce</i>	0,7
6	+	—	+	—	—	<i>ac</i>	0,6
7	—	+	+	—	—	<i>bc</i>	0,8
8	+	+	+	—	+	<i>abce</i>	0,5
9	—	—	—	+	+	<i>de</i>	0,7
10	+	—	—	+	—	<i>ad</i>	0,8
11	—	+	—	+	—	<i>bd</i>	0,9
12	+	+	—	+	+	<i>abde</i>	0,6
13	—	—	+	+	—	<i>cd</i>	0,8
14	+	—	+	+	+	<i>acde</i>	0,6
15	—	+	+	+	+	<i>bcde</i>	0,7
16	+	+	+	+	—	<i>abcd</i>	0,7

Таблица 3.16

Уровень \ Фактор	ΔM	$\Delta \gamma$	$\Delta \psi$	$\Delta \eta$	Δy
Верхний уровень (+1)	0,10	2,0	2,0	4,0	0,04
Нижний уровень (-1)	0	0	0	0	0
Среднее значение (m_i)	0,05	1,0	1,0	2,0	0,02
Интервал варьирования (Δ_i)	0,05	1,0	1,0	2,0	0,02

дополнить слева столбцом для свободного члена, состоящим из одних +1, а справа десятью столбцами для парных взаимодействий $x_k x_l$ (значения каждого столбца получаются путем перемножения соответствующих элементов столбцов x_k и x_l). Для исключения ошибки проверим полученный план на ортогональность и симметричность:

$$\sum_{i=1}^{16} x_k^i x_l^i = 0, \quad \sum_{i=1}^{16} x_k^i = 0. \quad (3.7.42)$$

Теперь можно приступить к проведению опытов. Но так как в матрице планирования независимые переменные заданы в нормированном виде, нужно перейти к действительным переменным, которые задаются в элементной модели и являются декоррелирующими факторами. Поскольку строится линейная модель крутизны, то область факторного пространства нужно выбирать небольшой, иначе линейная модель окажется неадекватной. В табл. 3.16 приведены значения действительных переменных в относительных единицах по отношению к полю зрения оптической системы. Переход от действительных переменных к нормированным производится по формуле (3.6.57).

Блок-схема проведения эксперимента приведена на рис. 3.40. Проведение эксперимента полностью автоматизировано, для чего составляется управляющая программа, которая по заданной матрице планирования задает уровень действительных переменных в каждом из шестнадцати опы-



Рис. 3.40. Блок-схема проведения эксперимента

тов, производит вычисление дискриминационной характеристики, ее линеаризацию и нормировку, накопление массива значений крутизны S , вычисление нормированных и действительных коэффициентов регрессии КР с проверкой их значимости.

В результате проведения 16 опытов на поэлементной модели были получены значения крутизны дискриминационной характеристики (выход модели $\tilde{y}^i$), приведенные в последнем столбце табл. 3.15. Коэффициенты регрессии вычислялись по формуле (3.6.86)

$$\hat{a}_k = \frac{\sum_{i=1}^{16} x_k^i \tilde{y}^i}{16}. \quad (3.7.43)$$

Значения оценок нормированных коэффициентов регрессии приведены в табл. 3.17 (вторая строка). Эти оценки являются коэффициентами уравнения регрессии для нормированных переменных, т.е. уравнения вида

$$\begin{aligned} \tilde{S} &= \hat{a}'_0 + \hat{a}'_1 \frac{x_1 - m_1}{\Delta_1} + \hat{a}'_2 \frac{x_2 - m_2}{\Delta_2} + \dots = \\ &= \hat{a}'_0 - \frac{\hat{a}'_1}{\Delta_1} m_1 - \frac{\hat{a}'_2}{\Delta_2} m_2 - \dots + \frac{\hat{a}'_1}{\Delta_1} x_1 + \frac{\hat{a}'_2}{\Delta_2} x_2 + \dots \end{aligned} \quad (3.7.44)$$

Отсюда видно, что коэффициенты регрессии при переходе к действительным переменным вычисляются по формуле

$$\hat{a}_i = \frac{\hat{a}'_i}{\Delta_i}, \quad (3.7.45)$$

а свободный член равен

$$\hat{a}_0 = \hat{a}'_0 - \hat{a}'_1 \frac{m_1}{\Delta_1} - \hat{a}'_2 \frac{m_2}{\Delta_2} - \dots = 1. \quad (3.7.46)$$

Коэффициенты регрессии для действительных переменных с учетом табл. 3.16 приведены в третьей строке табл. 3.17. Таким образом, зависимость крутизны от декоррелирующих факторов в линейном приближении может быть описана следующей регрессионной моделью:

$$\tilde{S} = 1 - 1,6 \Delta M - 0,03 \Delta \gamma - 0,035 \Delta \psi - 0,005 \Delta \eta - 3,62 \Delta u + \dots \quad (3.7.47)$$

Как видно из уравнения, все линейные эффекты уменьшают крутизну

Таблица 3.17

КР $\hat{a}_i$	a_1	a_2	a_3	a_4	a_5	$a_{1,2}$	$a_{1,3}$
$\hat{a}'_i$	-0,08	-0,03	-0,035	-0,01	-0,073	0,003	0,003
$\hat{a}_i$	-1,6	-0,03	-0,035	-0,005	-3,62	0,062	0,062

дискриминационной характеристики, однако некоторые парные взаимодействия ее увеличивают. Наибольшее влияние оказывают разномасштабирование и перекрестная связь, наименьшее — угол визирования.

Проверить значимость коэффициентов и адекватность модели по формулам § 3.6 в этом эксперименте нельзя, так как нет достаточного количества степеней свободы:

$$\varphi = N - k - 1 = 16 - 15 - 1 = 0. \quad (3.7.48)$$

Если реализовать вторую полуреплику 2^{5-1} , выбрав генератор $x_5 = -x_1x_2x_3x_4$, и провести еще 16 опытов, тогда можно будет проверить значимость коэффициентов ($\varphi = 32 - 16 = 16$).

В заключение отметим, что коэффициенты регрессии можно было бы вычислить и обычным однофакторным экспериментом, варьируя по очереди все факторы на двух уровнях. Для этого также потребовалось бы 16 опытов (первый опыт общий при отсутствии рассогласований). Но точность оценок здесь была бы низкая, так как каждый коэффициент определялся бы по результату только одного опыта. В многофакторном эксперименте каждый коэффициент определяется по результатам всех 16 опытов, т.е. дисперсия оценки в 16 раз меньше [см. (3.6.88)]:

$$s_i = \frac{s^2}{16}. \quad (3.7.49)$$

Для достижения такой точности в однофакторном эксперименте пришлось бы каждую зависимость снимать по 16 опытам, т.е. всего провести $16 \cdot 16 = 256$ опытов. Таким образом, многофакторный эксперимент позволяет значительно сократить объем экспериментальных исследований.

Построение квадратичной модели для трех переменных. Исследование зависимости крутизны дискриминационной характеристики от декоррелирующих факторов в линейном приближении показало, что основное влияние оказывают разномасштабирование, крен и ортогональный сдвиг (курсовой угол эквивалентен и углу крена). Поэтому для дальнейших исследований после корректировки параметров поэлементной модели были оставлены три фактора: $x_1 = \Delta M$, $x_2 = \Delta \gamma$, $x_3 = \Delta y$. С целью получения регрессионной зависимости в более широкой области факторного пространства была выбрана модель второго порядка

$$\begin{aligned} \tilde{S} = & a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + \\ & + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3. \end{aligned} \quad (3.7.50)$$

	a_{14}	a_{15}	a_{23}	a_{24}	a_{25}	a_{34}	a_{35}	a_{45}
	-0,023	0,016	0,009	-0,025	-0,003	-0,015	0,009	-0,0001
	-0,225	15,75	0,009	-0,013	-0,156	-0,008	0,47	-0,003

Таблица 3.18

Уровень \ Фактор	ΔM	$\Delta \gamma$	Δy
Верхний уровень (+1)	0,2	2	0,1
Нижний уровень (-1)	-0,2	-2	-0,1
Среднее значение (m_i)	0	0	0
Интервал варьирования (Δ_i)	0,2	2	0,1

Для проведения эксперимента был выбран точный план, близкий к D -оптимальному, типа B_3 , матрица планирования которого приведена в табл. 3.11. Поскольку план не ортогональный, оценки коэффициентов необходимо вычислять по общей формуле решения системы нормальных уравнений

$$\hat{a} = CF^T Y. \quad (3.7.51)$$

Основной операцией при этом является вычисление дисперсионной матрицы C

$$C = M^{-1} = (F^T F)^{-1}. \quad (3.7.52)$$

Матрица F размера 14×10 получается из матрицы плана путем добавления слева столбца для свободного члена, состоящего из единиц, и справа — столбцов для квадратов и взаимодействий, получающихся перемножением соответствующих столбцов. Информационная матрица M размера 10×10 имеет вид

$$M = \begin{pmatrix} 14 & 0 & 0 & 0 & 10 & 10 & 10 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 10 & 0 & 0 & 0 & 0 & 0 & 0 \\ 10 & 0 & 0 & 0 & 10 & 8 & 8 & 0 & 0 & 0 \\ 10 & 0 & 0 & 0 & 0 & 10 & 8 & 0 & 0 & 0 \\ 10 & 0 & 0 & 0 & 8 & 8 & 10 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 \end{pmatrix}. \quad (3.7.53)$$

Дисперсионная матрица размера также 10×10 равна

Таблица 3.19

i	1	2	3	4	5	6	7	
$\tilde{y}^i$	0,995	0,833	0,783	0,880	0,859	0,859	0,877	

$$C = \begin{pmatrix} 0,406 & 0 & 0 & 0 & -0,156 & -0,156 & -0,156 & 0 & 0 & 0 \\ 0 & 0,1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0,1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0,156 & 0 & 0 & 0 & 0,406 & -0,094 & -0,094 & 0 & 0 & 0 \\ -0,156 & 0 & 0 & 0 & -0,094 & 0,406 & -0,094 & 0 & 0 & 0 \\ -0,156 & 0 & 0 & 0 & -0,094 & -0,094 & 0,406 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,125 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,125 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,125 \end{pmatrix}. \quad (3.7.54)$$

Матрица плана (см. табл. 3.11), как обычно, написана для нормированных переменных, принимающих три значения: +1, -1, 0. Переход от действительных значений факторов к нормированным производится по формуле (3.6.57). Уровни действительных переменных, использованных в элементной математической модели ОЭС, задавались табл. 3.18.

В результате проведенных экспериментов был получен вектор наблюдений $\tilde{Y}$, представленный в табл. 3.19.

Расчет оценок коэффициентов по формуле (3.7.51) дает значения, приведенные в табл. 3.20 (первая строка — коэффициенты для нормированных переменных, вторая — для действительных).

Для проверки значимости коэффициентов находим остаточную сумму квадратов

$$S_R = \sum_{i=1}^{14} (\tilde{y}^i - \hat{y}^i) = (\tilde{Y} - \hat{Y})^T (\tilde{Y} - \hat{Y}) \quad (3.7.55)$$

с числом степеней свободы

$$\varphi = N - k - 1 = 14 - 10 = 4. \quad (3.7.56)$$

Оценка дисперсии ошибок наблюдений равна

$$s^2 = \frac{S_R}{\varphi} = 0,486 \cdot 10^{-2}. \quad (3.7.57)$$

Дисперсии оценок коэффициентов пропорциональны диагональным элементам матрицы C

$$s_i^2 = c_{ii} s^2. \quad (3.7.58)$$

Соответственно для каждой группы коэффициентов имеем:

$$s_0^2 = c_0 s^2 = 0,406 \cdot s^2 \quad \text{для свободного члена,}$$

$$s_i^2 = c_1 s^2 = 0,1 \cdot s^2 \quad \text{для линейных членов,}$$

8	9	10	11	12	13	14
0,93	1,06	0,929	0,873	0,955	0,975	1,07

Таблица 3.20

$\hat{a}_i \backslash \text{КР}$	a_0	a_1	a_2	a_3	a_{11}
$\hat{a}'_1$	1,028	$-0,146 \cdot 10^{-1}$	$0,774 \cdot 10^{-3}$	$0,134 \cdot 10^{-1}$	$-0,327 \cdot 10^{-1}$
$\hat{a}_1$	1,028	-0,073	$0,387 \cdot 10^{-3}$	0,134	-0,818

$\hat{a}_i \backslash \text{КР}$	a_{22}	a_{33}	a_{12}	a_{13}	a_{23}
$\hat{a}'_1$	-0,114	$-0,393 \cdot 10^{-2}$	$0,389 \cdot 10^{-1}$	$0,318 \cdot 10^{-1}$	$0,148 \cdot 10^{-10}$
$\hat{a}_1$	$-0,285 \cdot 10^{-1}$	-0,393	$0,973 \cdot 10^{-1}$	1,59	0,074

$$s_{ii}^2 = c_2 s^2 = 0,406 \cdot s^2 \quad \text{для квадратов,} \quad (3.7.59)$$

$$s_{ij}^2 = c_3 s^2 = 0,125 \cdot s^2 \quad \text{для взаимодействий.}$$

Проверка значимости производится по формуле (3.6.45)

$$|\hat{a}'_i| < t_{\text{кр}} s_i.$$

Для уровня значимости $\alpha = 0,1$ и числа степеней свободы $\varphi = 4$ из таблицы распределения Стьюдента имеем $t_{\text{кр}} = 2,132$. В данном случае значимыми оказались только свободный член a_0 и коэффициент a_{22} . Поэтому изменение крутизны в данном случае может быть описано более простым уравнением

$$\tilde{S} = a_0 + a_{22} x_2^2. \quad (3.7.60)$$

Но так как план B_3 не ортогональный, то коэффициенты a_0 и a_{22} должны быть вычислены заново по тем же экспериментальным данным, но с новой матрицей $f(x)$, имеющей вид

$$f(x) = (1 \ x^2)^T. \quad (3.7.61)$$

Для уравнения с одной переменной план имеет вид

$$X(1) = \begin{vmatrix} -1 \\ +1 \\ 0 \end{vmatrix} \quad (3.7.62)$$

и содержит три опыта. Поскольку проведено 14 опытов, то добавим еще один опыт в центре плана и сгруппируем по 5 опытов. Матрица F для это-

го случая имеет вид

$$F = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad (3.7.63)$$

а информационная матрица

$$M = \begin{pmatrix} 3 & 2 \\ 2 & 2 \end{pmatrix}. \quad (3.7.64)$$

Обратная матрица находится элементарно:

$$C = \begin{pmatrix} 1 & -1 \\ -1 & 1,5 \end{pmatrix}. \quad (3.7.65)$$

Расчет коэффициентов по формуле (3.7.51) дает следующие оценки коэффициентов для действительных переменных:

$$a_0 = 0,975, \quad a_{22} = -0,019.$$

Таким образом, изменение крутизны описывается моделью

$$\tilde{S} = 1 - 0,019 \Delta \gamma^2. \quad (3.7.66)$$

Все вычисления проводились на ЭВМ по блок-схеме, приведенной на рис. 3.40, с использованием стандартных программ матричной алгебры.

Аналитическая модель динамического контура. Рассмотрим пример использования аналитической модели для анализа динамики замкнутого контура. В качестве характеристики будем рассматривать плотность ошибки слежения в зависимости от времени. Эта характеристика наиболее полно описывает поведение динамической системы в вероятностном смысле. Аналитическая модель должна быть достаточно простой, чтобы можно

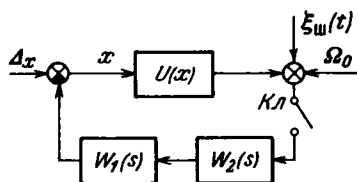


Рис. 3.41. Схема аналитической модели

было получить решение в явном виде, поэтому обычно ограничиваются системой первого или второго порядка. Схема такой модели приведена на рис. 3.41, где $U(x)$ — дискриминационная характеристика, описываемая следующей зависимостью:

$$U(x) = \begin{cases} k \cdot \sin \frac{\pi x}{\beta} & \text{при } |x| < \beta; \\ 0 & \text{при } |x| \geq \beta. \end{cases} \quad (3.7.67)$$

Гиростабилизатор описывается интегрирующим звеном с передаточной функцией

$$W_1(s) = \frac{1}{s}. \quad (3.7.68)$$

Дискретность поступления информации, например в телевизионных ОЭС, задается ключом Кл, формирующим мгновенные импульсы с периодом T_0 , и фиксирующим звеном нулевого порядка с передаточной функцией

$$W_2(s) = \frac{1 - e^{-sT_0}}{s}. \quad (3.7.69)$$

Внутренние шумы задаются случайным процессом $\xi_{ш}(t)$, спектр которого обычно много шире полосы пропускания замкнутого контура, поэтому его можно считать белым шумом. Скорость движения изображения задается угловой скоростью Ω_0 .

Очень часто коэффициент усиления k зависит от уровня входного сигнала, в этом случае амплитудные флуктуации сигнала, например за счет турбулентности атмосферы, соответствуют случайному коэффициенту усиления в контуре ОЭС. Представим коэффициент усиления в виде

$$k = k_0 - k(t), \quad (3.7.70)$$

где k_0 — среднее значение; $k(t)$ — случайная составляющая. Тогда из (3.7.67) получим

$$U(x) = [k_0 - k(t)] \sin \frac{\pi x}{\beta} = k_0 \sin \frac{\pi x}{\beta} - \xi_k(t) k_0 \sin \frac{\pi x}{\beta}. \quad (3.7.71)$$

Здесь $\xi_k(t) = \frac{k(t)}{k_0}$ — случайная составляющая коэффициента усиления,

выраженная в относительных единицах. Таким образом, случайную составляющую коэффициента усиления можно представить в виде дополнительного шума, уровень которого зависит от ошибки x .

Для решения задачи о слежении в нелинейных системах с учетом нестационарности рассматриваемого явления удобен аппарат марковских процессов, широко используемый при исследовании непрерывных систем автоматического регулирования. Однако применению этого метода для исследования импульсных систем посвящено сравнительно небольшое число работ. На основе интегро-разностных уравнений в [27] найдены рекуррентные соотношения, связывающие плотности распределения вероятностей ошибки в дискретные моменты времени $t_n = nT_0$ и $t_{n+1} = nT_0 + T_0$. Полученные результаты справедливы для нелинейных импульсных систем, которые описываются разностными уравнениями с аддитивными дискретными во времени белыми шумами постоянной

интенсивности. Случайный коэффициент усиления приводит к появлению мультипликативного дискретного шума, зависящего от ошибки слежения.

Интегро-разностное уравнение для плотности распределения вероятностей имеет вид [27]

$$\Delta f_n(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\lambda \int_{-\infty}^{\infty} e^{i\lambda(y-x)} \Phi_n(\lambda, y) f_n(y) dy, \quad (3.7.72)$$

где $f_n(x)$ – плотность вероятностей случайного процесса в момент времени t_n ($n = 0, 1, 2, \dots$); $\Delta f_n(x)$ – первая разность плотности случайного процесса.

Функция

$$\Phi_n(\lambda, x) = \Phi(\lambda, x, nT_0) = M[e^{i\lambda \Delta x_n} | x] - 1 \quad (3.7.73)$$

имеет смысл условной характеристической функции (без единицы) приращения Δx_n случайного процесса x_n в точке t_n при заданном его значении x в тот же момент времени. Уравнение (3.7.72) справедливо для любого дискретного случайного процесса, для которого существует функция $\Phi_n(\lambda, x)$, и в общем случае определяет лишь одномерный закон распределения значений процесса x_n в любой заданный дискретный момент времени. В частном случае марковских последовательностей уравнение (3.7.72) справедливо для переходной плотности вероятностей при заданном значении в момент времени $t = 0$.

Динамические свойства системы, приведенной на рис. 3.41, с учетом (3.7.67) и (3.7.71) описываются разностным уравнением

$$\Delta x_n = -T_0 k_0 \sin \frac{\pi x_n}{\beta} - T_0 \Omega_0 + \xi_{шn} T_0 + \xi_{kn} T_0 k_0 \sin \frac{\pi x_n}{\beta}. \quad (3.7.74)$$

Здесь $\xi_{шn}$ и ξ_{kn} – значения шумов $\xi_{ш}(t)$ и $\xi_k(t)$ в дискретные моменты времени t_n .

Часто спектр амплитудных флуктуаций сигнала много шире полосы пропускания системы, поэтому его также можно считать белым шумом. В этом случае шумы $\xi_{шn}$ и ξ_{kn} являются дискретными белыми шумами, а уравнение (3.7.74) описывает дискретную во времени марковскую последовательность x_n .

Используя (3.7.73) и (3.7.74) при $x_n = x$, получаем

$$\begin{aligned} \Phi_n(\lambda, x) = & \exp \left[i\lambda \left(-T_0 k_0 \sin \frac{\pi x}{\beta} - T_0 \Omega_0 \right) \right] \times \\ & \times M \left\{ \exp \left[i\lambda \left(\xi_{шn} T_0 + \xi_{kn} T_0 k_0 \sin \frac{\pi x}{\beta} \right) \right] \right\} - 1. \end{aligned} \quad (3.7.75)$$

Подставив (3.7.75) в (3.7.72), приведем его к виду

$$f_{n+1}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\lambda \int_{-\infty}^{\infty} \exp \left[i\lambda \left(y - x - T_0 k_0 \sin \frac{\pi}{\beta} y - T_0 \Omega_0 \right) \right] \times \\ \times M \left\{ \exp \left[i\lambda \left(\xi_{шн} T_0 + \xi_{kn} T_0 k_0 \sin \frac{\pi}{\beta} y \right) \right] \right\} f_n(y) dy. \quad (3.7.76)$$

Уравнение (3.7.76) является рекуррентным соотношением, связывающим плотности вероятностей в моменты времени t_n и t_{n+1} . Найдем в явном виде значение математического ожидания $M\{\cdot\}$. При независимых внутренних шумах и флуктуациях сигнала можно записать

$$M \left\{ \exp \left[i\lambda \left(\xi_{шн} T_0 + \xi_{kn} T_0 k_0 \sin \frac{\pi}{\beta} y \right) \right] \right\} = \\ = M \left[\exp (i\lambda \xi_{шн} T_0) \right] \cdot M \left[\exp \left(i\lambda \xi_{kn} T_0 k_0 \sin \frac{\pi}{\beta} y \right) \right]. \quad (3.7.77)$$

Математическое ожидание случайного процесса y при экспоненциальном преобразовании $y = e^{kz}$ равно [68]

$$M\{y\} = e^{-\frac{k^2 \sigma^2}{2}},$$

где σ^2 — дисперсия случайного процесса z . Так как дисперсия дискретного белого шума интенсивностью C_0 конечна и равна C_0/T_0 , то первый сомножитель в (3.7.77) можно записать в виде

$$M \left[\exp (i\lambda \xi_{шн} T_0) \right] = \exp \left(-\frac{1}{2} \lambda^2 T_0 C_{0ш} \right). \quad (3.7.78)$$

Для вычисления математического ожидания второго сомножителя найдем дисперсию выражения $\xi_{kn} \sin \frac{\pi}{\beta} y$, считая процессы ξ_{kn} и y независимыми. Дисперсия произведения независимых случайных процессов равна

$$\sigma_2^2 = D\{\xi_{kn}\} D\left\{ \sin \frac{\pi}{\beta} y \right\} + M^2\{\xi_{kn}\} D\left\{ \sin \frac{\pi}{\beta} y \right\} + \\ + M^2\left\{ \sin \frac{\pi}{\beta} y \right\} D\{\xi_{kn}\}. \quad (3.7.79)$$

Учитывая, что среднее значение шума ξ_k равно нулю, получаем

$$\begin{aligned} \sigma_2^2 &= D\{\xi_{kn}\} \left[M^2 \left\{ \sin \frac{\pi}{\beta} y \right\} + D \left\{ \sin \frac{\pi}{\beta} y \right\} \right] = \\ &= D\{\xi_{kn}\} M_2 \left\{ \sin \frac{\pi}{\beta} y \right\}, \end{aligned} \quad (3.7.80)$$

где $M_2\{\cdot\}$ – второй начальный момент.

В выражении (3.7.76) y есть значение процесса на предыдущем шаге, плотность вероятности которого $f_n(y)$ известна. Поэтому второй начальный момент равен

$$M_{2y} = \int_{-\infty}^{\infty} \sin^2 \frac{\pi}{\beta} y f_n(y) dy, \quad (3.7.81)$$

а математическое ожидание второго сомножителя в (3.7.77) равно

$$M\{\cdot\} = \exp \left(-\frac{1}{2} \lambda^2 T_0 k_0^2 M_{2y} C_{0k} \right). \quad (3.7.82)$$

Подставляя (3.7.78) и (3.7.82) в (3.7.76) и меняя порядок интегрирования, получаем

$$\begin{aligned} f_{n+1}(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f_n(y) dy \int_{-\infty}^{\infty} \exp \left[-\frac{1}{2} \lambda^2 T_0 \times \right. \\ &\times \left. (C_{0ш} + k_0^2 M_{2y} C_{0k}) \right] e^{-i\lambda z} d\lambda, \end{aligned} \quad (3.7.83)$$

где

$$z = -\left(y - x - T_0 k_0 \sin \frac{\pi}{\beta} y - T_0 \Omega_0 \right). \quad (3.7.84)$$

При интегрировании по λ первый сомножитель в подынтегральном выражении (3.7.83) можно рассматривать как характеристическую функцию

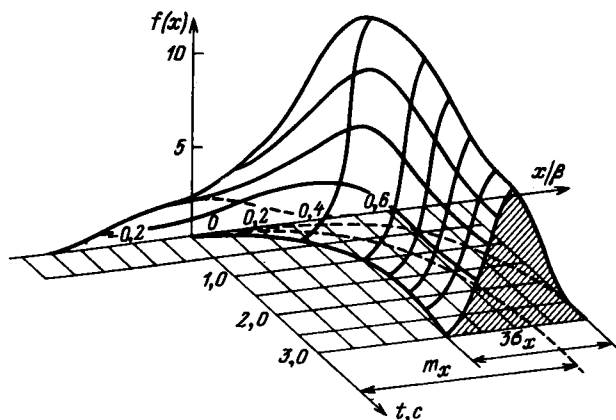


Рис. 3.42. Плотность распределения ошибки слежения

нормального процесса z с нулевым математическим ожиданием и дисперсией $T_0(C_{0ш} + k_0^2 M_{2y} C_{0к})$, а процедуру вычисления интеграла — как нахождение плотности вероятностей процесса z , которая при подобной интерпретации подынтегрального выражения также будет нормальной [27]. Окончательно получим

$$f_{n+1}(x) = \frac{1}{\sqrt{2\pi(C_{0ш} + k_0^2 M_{2y} C_{0к})T_0}} \times \\ \times \int_{-\infty}^{\infty} \exp \left\{ - \frac{(y - x - T_0 k_0 \sin \frac{\pi}{\beta} y - T_0 \Omega_0)^2}{2 T_0(C_{0ш} + k_0^2 M_{2y} C_{0к})} \right\} f_n(y) dy, \quad (3.7.85)$$

где M_{2y} определяется из (3.7.81).

Выражение (3.7.85) позволяет при известном начальном распределении $f_0(x)$ вычислять последовательно плотности распределения ошибки слежения для моментов времени t_n . Обычно начальное распределение определяется ошибкой передачи координат и может считаться нормальным с нулевым средним и дисперсией σ^2 .

По формуле (3.7.85) на ЭВМ рассчитано изменение плотности ошибки слежения во времени при нормальном начальном распределении. Результаты расчетов приведены на рис. 3.42.

В первые моменты времени система быстро отрабатывает начальную ошибку, но в дальнейшем стационарное распределение не устанавливается, что обусловлено конечностью апертуры β дискриминационной характеристики $U(x)$. При малых уровнях шумов плотность распределения ошибки близка к стационарной, при больших шумах происходит срыв слежения. Влияние начальной ошибки мало сказывается на вероятности слежения, если ее величина не превосходит $\beta/2$.

Таким образом, последовательное применение математических моделей различных уровней иерархии позволяет на разных этапах проектирования получить исчерпывающую информацию о работе оптико-электронной системы.

ПОСЛЕСЛОВИЕ

С появлением возможностей математического моделирования сложных динамических систем неизмеримо расширился круг решаемых задач. В самих методах аналитических исследований также произошли изменения: здесь стали играть заметную роль новые методы вычислительной математики, появилась теория программирования, алгоритмических языков и т.п. Многие аналитические методы под влиянием "машинной эйфории" оказались отодвинутыми на второй план. Но вскоре выяснилось, что многие теоретические и прикладные задачи не поддаются прямому решению с помощью ЭВМ. Некоторые примеры таких задач приводятся в главе 2 монографии. Среди них в первую очередь следует отметить асимптотические методы, позволившие разбить возможные периодические режимы ОЭСС на непересекающиеся классы. Наиболее плодотворные результаты получаются при совместном использовании аналитических и численных методов, причем численному решению задач обычно предшествует их качественный анализ.

При идентификации математических моделей ОЭСС по данным натуральных испытаний возникает проблема поиска глобального экстремума в пространстве параметров со сложной топологией. Без рационального аналитического подхода прямое применение известных алгоритмов поиска экстремума не приводит к положительному результату.

Читатель, ориентирующийся в современных методах обработки изображений, быть может, ощутит некоторую неудовлетворенность, не найдя в первых двух главах сведений об аппаратной реализации дискретных преобразований, без которых трудно представить современные и перспективные ОЭСС, содержащие цифровые устройства. Это недостаток частично восполняется в главе 3. Кроме того, появилась обширная литература, посвященная этому вопросу. Здесь можно упомянуть монографии А. Розенфельда, У. Прэтта, Л.П. Ярославского, Г.И. Василенко, Т. Хуанга и др.

Необходимо иметь в виду, что функции измерителей многомерного сдвига изображений состоят в том, чтобы извлекать из сигналов с большим числом степеней свободы информацию о составляющих сдвига по малому числу компонент. Благодаря такой узкой специализации, в системах технического зрения, не производящих визуализацию сопровождаемого сюжета, требования к составу операторов фильтрации оказываются совсем иными, чем в системах, обеспечивающих наилучшее воспроизведение деталей изображения для их кодирования или распознавания несмотря на иска-

жающие факторы. Указанное обстоятельство не всегда улавливается разработчиками, которые вместо измерения сдвигов принимаются решать более сложную задачу наилучшего воспроизведения и совмещения деталей двух изображений.

Не вдаваясь в далеко уводящие подробности, заметим, что продуктивным способом реализации преобразований, предусмотренных обобщенным алгоритмом дискриминатора сдвигов, является фильтрация двумерных сигналов в пространстве спектров. При разумном выборе ядра двумерного (не обязательно унитарного) преобразования дискретизованных изображений, которые в дальнейшем не восстанавливаются, удается достичь значительной экономии как в числе воспроизводимых коэффициентов разложения, так и в алгоритмах окончательного расчета нормированных оценок сдвига. Общий выигрыш в аппаратных и вычислительных затратах особенно заметен, если обрабатываются не исходные изображения, а лишь их яркостные срезы, т.е. их морфологические аналоги.

ПРИЛОЖЕНИЕ

ВЫЧИСЛЕНИЕ ПРАВЫХ ЧАСТЕЙ УКОРОЧЕННЫХ УРАВНЕНИЙ (2.4.32)

Найдем средние значения интегралов вида

$$I_m = \frac{1}{2\pi} \int_0^{2\pi} \cos^m \varphi d\varphi.$$

Очевидно,

$$I_m = I_{m-2} - \frac{1}{2\pi} \int_0^{2\pi} \cos^{m-2} \varphi \sin^2 \varphi d\varphi. \quad (\text{П1})$$

Вычисляя второе слагаемое в (П1) по частям, находим рекуррентное соотношение

$$I_m = \frac{m-1}{m} I_{m-2} \quad (m = 2, 3, 4, \dots). \quad (\text{П2})$$

Таким образом, $I_{2k+1} = 0$ ($k = 0, 1, 2, \dots$) и $I_0 = 1$, $I_2 = \frac{1}{2}$, $I_4 = \frac{3}{8}$, $I_6 = \frac{5}{16}, \dots$

Воспользуемся полученным результатом для осреднения следующих выражений:

$$\mu\Phi_1(\Delta_1, \Delta_2, \Delta_1^{(2)}) = -2\xi\omega\Delta_1^{(2)} - \omega^2 F(\Delta) (\Delta_1 \cos \epsilon - \Delta_2 \sin \epsilon),$$

$$\mu\Phi_2(\Delta_1, \Delta_2, \Delta_2^{(2)}) = -2\xi\omega\Delta_2^{(2)} - \omega^2 F(\Delta) (\Delta_1 \sin \epsilon + \Delta_2 \cos \epsilon).$$

Запишем в явной форме величины, входящие в (2.4.36):

$$\begin{aligned} \Delta^2 &= a_1^2 \cos^2 \varphi_1 + 2a_1 c_1 \cos \varphi_1 + a_2^2 \cos^2 \varphi_2 + 2a_2 c_2 \cos \varphi_2 + c_1^2 + c_2^2, \\ \Delta^4 &= a_1^4 \cos^4 \varphi_1 + 4a_1^3 c_1 \cos^3 \varphi_1 + 2a_1^2 a_2^2 \cos^2 \varphi_1 \cos^2 \varphi_2 + 4a_1^2 a_2 c_2 \cos^2 \varphi_1 \cos \varphi_2 + \\ &+ 6a_1^2 c_1^2 \cos^2 \varphi_1 + 2a_1^2 c_2^2 \cos^2 \varphi_1 + 4a_1 a_2^2 c_1 \cos \varphi_1 \cos^2 \varphi_2 + 8a_1 a_2 c_1 c_2 \cos \varphi_1 \cos \varphi_2 + \\ &+ 4a_1 c_1^3 \cos \varphi_1 + 4c_1^2 c_2^2 \cos \varphi_1 + a_2^4 \cos^4 \varphi_2 + 4a_2^3 c_2 \cos^3 \varphi_2 + 2a_2^2 c_1^2 \cos^2 \varphi_2 + \\ &+ 6a_2^2 c_2^2 \cos^2 \varphi_2 + 4a_2 c_1^2 c_2 \cos \varphi_2 + 4a_2 c_2^3 \cos \varphi_2 + c_1^4 + 2c_1^2 c_2^2 + c_2^4. \end{aligned}$$

При дальнейшем перемножении в произведениях вида $\Delta^2 \Delta_i$, $\Delta^4 \Delta_i$ ($i = 1, 2$) следует удерживать только слагаемые с четными степенями косинусов. Таким образом, в результате осреднения получаем

$$\begin{aligned} &\frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} F(\Delta) \Delta_1 d\varphi_1 d\varphi_2 = \\ &= \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} (f_0 + f_1 \Delta^2 + f_2 \Delta^4) (c_1 + a_1 \cos \varphi_1) d\varphi_1 d\varphi_2 = \\ &= c_1 \left[f_0 + f_1 \left(\frac{3}{2} a_1^2 + \frac{1}{2} a_2^2 + c_1^2 + c_2^2 \right) + \right. \end{aligned}$$

$$+ f_2 \left(\frac{15}{8} a_1^4 + \frac{3}{2} a_1^2 a_2^2 + 5 a_1^2 c_1^2 + 3 a_1^2 c_2^2 + \frac{3}{8} a_2^4 + a_2^2 c_1^2 + 3 a_2^2 c_2^2 + c_1^4 + 2 c_1^2 c_2^2 + c_2^4 \right) \Bigg], \quad (\text{ПЗ})$$

$$\begin{aligned} \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} F(\Delta) \Delta_2 d\varphi_1 d\varphi_2 &= \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} (f_0 + f_1 \Delta^2 + f_2 \Delta^4) (c_2 + a_2 \cos \varphi_2) d\varphi_1 d\varphi_2 = \\ &= c_2 \left[f_0 + f_1 \left(\frac{1}{2} a_1^2 + \frac{3}{2} a_2^2 + c_1^2 + c_2^2 \right) + \right. \\ &+ f_2 \left(\frac{3}{8} a_1^4 + \frac{3}{2} a_1^2 a_2^2 + 3 a_1^2 c_1^2 + a_1^2 c_2^2 + \frac{15}{8} a_2^4 + \right. \\ &\left. + 3 a_2^2 c_1^2 + 5 a_2^2 c_2^2 + c_1^4 + 2 c_1^2 c_2^2 + c_2^4 \right) \Bigg]. \quad (\text{П4}) \end{aligned}$$

При осреднении выражений, содержащих $\cos \varphi_i$, аналогично предыдущему находим

$$\begin{aligned} \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} (f_0 + f_1 \Delta^2 + f_2 \Delta^4) (c_1 + a_1 \cos \varphi_1) d\varphi_1 d\varphi_2 = \\ = a_1 \left[\frac{1}{2} f_0 + f_1 \left(\frac{3}{8} a_1^2 + \frac{1}{4} a_2^2 + \frac{3}{2} c_1^2 + \frac{1}{2} c_2^2 \right) + \right. \\ + f_2 \left(\frac{5}{16} a_1^4 + \frac{3}{8} a_1^2 a_2^2 + \frac{15}{4} a_1^2 c_1^2 + \frac{3}{4} a_1^2 c_2^2 + \frac{3}{16} a_2^4 + \right. \\ \left. + \frac{3}{2} a_2^2 c_1^2 + \frac{3}{2} a_2^2 c_2^2 + \frac{5}{2} c_1^4 + 3 c_1^2 c_2^2 + \frac{1}{2} c_2^4 \right) \Bigg]. \quad (\text{П5}) \end{aligned}$$

$$\begin{aligned} \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} (f_0 + f_1 \Delta^2 + f_2 \Delta^4) (c_2 + a_2 \cos \varphi_2) \cos \varphi_1 d\varphi_1 d\varphi_2 = \\ = a_1 c_1 c_2 \left[f_1 + f_2 \left(\frac{3}{2} a_1^2 + 3 a_2^2 + 2 c_1^2 + 2 c_2^2 \right) \right]. \quad (\text{П6}) \end{aligned}$$

$$\begin{aligned} \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} (f_0 + f_1 \Delta^2 + f_2 \Delta^4) (c_1 + a_2 \cos \varphi_2) \cos \varphi_2 d\varphi_1 d\varphi_2 = \\ = a_2 \left[\frac{1}{2} f_0 + f_1 \left(\frac{1}{4} a_1^2 + \frac{3}{8} a_2^2 + \frac{1}{2} c_1^2 + \frac{3}{2} c_2^2 \right) + \right. \\ + f_2 \left(\frac{3}{16} a_1^4 + \frac{3}{8} a_1^2 a_2^2 + \frac{3}{2} a_1^2 c_1^2 + \frac{3}{2} a_1^2 c_2^2 + \frac{5}{16} a_2^4 + \frac{3}{4} a_2^2 c_1^2 + \right. \\ \left. + \frac{15}{4} a_2^2 c_2^2 + \frac{1}{2} c_1^4 + 3 c_1^2 c_2^2 + \frac{5}{2} c_2^4 \right) \Bigg]. \quad (\text{П7}) \end{aligned}$$

$$\begin{aligned} \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} (f_0 + f_1 \Delta^2 + f_2 \Delta^4) (c_1 + a_1 \cos \varphi_1) \cos \varphi_2 d\varphi_1 d\varphi_2 = \\ = a_2 c_1 c_2 \left[f_1 + f_2 \left(3 a_1^2 + \frac{3}{2} a_2^2 + 2 c_1^2 + 2 c_2^2 \right) \right]. \quad (\text{П8}) \end{aligned}$$

Для краткости введем обозначения

$$\begin{aligned}
 \alpha_{11} &= \frac{3}{2} a_1^3 + \frac{1}{2} a_2^3 + c_1^3 + c_2^3, \\
 \alpha_{12} &= \frac{15}{8} a_1^4 + \frac{3}{2} a_1^2 a_2^2 + 5 a_1^2 c_1^2 + 3 a_1^2 c_2^2 + \frac{3}{8} a_2^4 + a_2^2 c_1^2 + \\
 &+ 3 a_2^2 c_2^2 + c_1^4 + 2 c_1^2 c_2^2 + c_2^4, \\
 \alpha_{21} &= \frac{1}{2} a_1^3 + \frac{3}{2} a_2^3 + c_1^3 + c_2^3, \\
 \alpha_{22} &= \frac{3}{8} a_1^4 + \frac{3}{2} a_1^2 a_2^2 + 3 a_1^2 c_1^2 + a_1^2 c_2^2 + \frac{15}{8} a_2^4 + 3 a_2^2 c_1^2 + \\
 &+ 5 a_2^2 c_2^2 + c_1^4 + 2 c_1^2 c_2^2 + c_2^4, \\
 \beta_{11} &= \frac{3}{8} a_1^3 + \frac{1}{4} a_2^3 + \frac{3}{2} c_1^3 + \frac{1}{2} c_2^3, \\
 \beta_{12} &= \frac{5}{16} a_1^4 + \frac{3}{8} a_1^2 a_2^2 + \frac{15}{4} a_1^2 c_1^2 + \frac{3}{4} a_1^2 c_2^2 + \frac{3}{16} a_2^4 + \\
 &+ \frac{3}{2} a_2^2 c_1^2 + \frac{3}{2} a_2^2 c_2^2 + \frac{5}{2} c_1^4 + 3 c_1^2 c_2^2 + \frac{1}{2} c_2^4, \\
 \beta_{21} &= \frac{1}{4} a_1^3 + \frac{3}{8} a_2^3 + \frac{1}{2} c_1^3 + \frac{3}{2} c_2^3, \\
 \beta_{22} &= \frac{3}{16} a_1^4 + \frac{3}{8} a_1^2 a_2^2 + \frac{3}{2} a_1^2 c_1^2 + \frac{3}{2} a_1^2 c_2^2 + \frac{5}{16} a_2^4 + \frac{3}{4} a_2^2 c_1^2 + \\
 &+ \frac{15}{4} a_2^2 c_2^2 + \frac{1}{2} c_1^4 + 3 c_1^2 c_2^2 + \frac{5}{2} c_2^4, \\
 \gamma_{12} &= \frac{3}{2} a_1^3 + 3 a_2^3 + 2 c_1^3 + 2 c_2^3, \\
 \gamma_{21} &= 3 a_1^3 + \frac{3}{2} a_2^3 + 2 c_1^3 + 2 c_2^3.
 \end{aligned} \tag{П9}$$

Эти обозначения использованы в системе укороченных уравнений (2.4.37).

ОСНОВНЫЕ ОБОЗНАЧЕНИЯ

$C(k)$	пространственно-частотная характеристика	u	управляющий сигнал
C	дисперсионная матрица	$\{x\}$	множество элементов x
E	освещенность	$\mathcal{F}_f$	фурье-образ для оригинала $f(t)$
H	кинетический момент	$\mathcal{L}$	оператор
$H_n(x)$	полином Эрмита	$\mathcal{M}$	множество
I	момент инерции	$\mathcal{P}$	оператор формы
J	матрица Якоби	$\mathcal{J}^\Delta, \mathcal{J}^s$	операторы обобщенного сдвига
J_1	функция Бесселя первого рода первого порядка	$\tilde{\mathcal{T}}$	сопряженный оператор обобщенного сдвига
K	корреляционная функция	$\mathcal{U}$	оператор аддитивного сдвига
K^0	обобщенная автокорреляционная функция	Δ	сдвиг; ошибка сопровождения
M	масштаб изображения; момент силы	Δt	шаг дискретизации
M	математическое ожидание	$\Phi(u)$	функция Лапласа
$N(F)$	спектральная плотность шума	Ω	угловая скорость вращения линии визирования
P	вероятность	γ	угол крена
Q	функция качества	$\delta(t)$	дельта-функция Дирака
Q	весовая матрица	$\theta(v)$	характеристическая функция
$R(x)$	нормированная корреляционная функция	λ	длина волны излучения
R	матрица мультипликативного преобразования	μ	коэффициент усиления
$S(k)$	пространственный спектр	$\mu(x)$	функция меры на множестве $\{x\}$
$S(\lambda)$	спектральная характеристика источника излучения	ξ	коэффициент демпфирования
T	постоянная времени	$\xi(x, y)$	двумерная случайная функция
$W(s)$	передаточная функция	σ	среднеквадратическое отклонение
$W(u, v)$	пространственная частотная характеристика	τ	сдвиг изображения
$W(z)$	дискретная передаточная функция	φ	фазовый угол
$W(j\omega)$	частотная характеристика	$\varphi(x, \omega)$	собственная функция оператора
$X(n)$	план эксперимента	χ	функция-индикатор
f	фокусное расстояние	ψ	угол курса
$g(t)$	импульсная характеристика	ω	круговая частота
$g(x, y)$	функция рассеяния точки	$1(t)$	единичный скачок
k	коэффициент усиления	$\in$	принадлежность к множеству
$m(\epsilon)$	коэффициент модуляции	$\bar{\in}$	символ, противоположный символу $\in$
$n(t)$	случайный процесс; шум	$\cup$	объединение множеств
r	коэффициент корреляции	$\cap$	пересечение множеств
$s(F)$	нормированная спектральная плотность сигнала	$\bullet$	символ свертки

- ⊙ — символ двумерной свертки
 ○ — символ обобщенной свертки

- ⊗ — символ группового сдвига
 ~ — символ сопряженного оператора

Индексы нижние

- | | | | |
|----|---------------------------------------|-----|--------------------------|
| а | — аддитивная; аналоговая | нч | — низкочастотная |
| в | — верхняя | оп | — опорное |
| г | — гироскопический | п | — помеха |
| д | — дискретный; датчик | р | — резонансное; развертка |
| дв | — двигатель | с | — сопротивление; сдвиг |
| др | — дробовой | ср | — среднее |
| з | — задержка; запирающее | ст | — статическая |
| и | — истинный; изображение | т | — тепловой; текущее |
| к | — конец; координатор; корре-
ляция | у | — усилитель |
| кв | — квантование | ф | — фильтр |
| кр | — критическое | фп | — фотоприемник |
| м | — модулятор; матрица; модель | ц | — центр |
| н | — нижняя; несущая; начальное | ш | — шум |
| | | эфф | — эффективный |

Индексы верхние

- т — транспонирование
 тр — трение
 упр — управляющий

Аббревиатуры

- | | | | |
|------|---|------|--|
| АИ | — анализатор изображения | КСС | — корреляционная следящая система |
| АИМ | — амплитудно-импульсная моду-
ляция | КФ | — корреляционная функция |
| АКФ | — автокорреляционная функция | ОАКФ | — обобщенная автокорреляцион-
ная функция |
| АРУ | — автоматическая регулировка
усиления | ОВКФ | — обобщенная взаимно корре-
ляционная функция |
| АЦПУ | — алфавитно-цифровое печатаю-
щее устройство | ОЗУ | — оперативное запоминающее
устройство |
| АЧХ | — амплитудно-частотная харак-
теристика | ОКЭС | — оптическая корреляционно-
экстремальная система |
| ВКФ | — взаимно корреляционная функ-
ция | ОМНК | — обобщенный метод наимень-
ших квадратов |
| ВСП | — взаимная спектральная плот-
ность | ООС | — оператор обобщенного сдвига |
| ГСН | — головка самонаведения | ОС | — оптическая система |
| ДИКМ | — дифференциальная импульс-
но-кодовая модуляция | ОЭС | — оптико-электронная система |
| ДХ | — дискриминационная характе-
ристика | ОЭСС | — оптико-электронная следящая
система |
| ЗУ | — запоминающее устройство | ПИ | — плоскость изображения |
| ЗЭЛТ | — запоминающая электронно-лу-
чевая трубка | ФП | — фотоприемник |
| КА | — космический аппарат | ФЭ | — фоточувствительный элемент |
| КД | — корреляционный дискримина-
тор | ФРТ | — функция рассеяния точки |
| | | ФЧХ | — фазочастотная характеристика |
| | | ЧХ | — частотная характеристика |
| | | ЭР | — экстремальный регулятор |

СПИСОК ЛИТЕРАТУРЫ

1. Адаптивные автоматические системы / Под ред. Г.А. Медведева. — М.: Советское радио, 1972.
2. Аналоговый видеокоррелятор для определения положения самолета. — Патент США, кл. 178 — 6.8, № 3.555.179, 1971.
3. Андронов А.А., Витт А.А., Хайкин С.Э. Теория колебаний. — М.: Физматгиз, 1959.
4. Антонов Е.И., Ильин В.Е., Коленко Е.А. и др. Устройства для охлаждения приемников излучения. — Л.: Машиностроение, 1972.
5. Астапов Ю.М., Медведев В.С. Статистическая теория систем автоматического регулирования и управления. — М.: Наука, 1982.
- ✓ 6. Баклицкий В.К., Юрьев А.Н. Корреляционно-экстремальные методы навигации. — М.: Радио и связь, 1982.
7. Бакут П.А., Большаков И.А., Герасимов Б.М. и др. Вопросы статистической радиолокации. Т. 2. — М.: Советское радио, 1964.
- ✚ 8. Барсуков Ф.И., Величкин А.И., Сухарев А.В. Телевизионные системы летательных аппаратов. — М.: Советское радио, 1979.
9. Баталов Ю.В., Мирошников Н.М. Методы представления изображений и их обработка // Оптико-механическая промышленность. — 1977. — № 7.
10. Бегунов Б.Н., Закаэнов Н.П. Теория оптических систем. — М.: Машиностроение, 1979.
- ✓ 11. Белоглазов И.Н., Тарасенко В.П. Корреляционно-экстремальные системы. — М.: Советское радио, 1974.
12. Березин И.С., Жидков Н.П. Методы вычислений. Т. 2. — М.: Физматгиз, 1962.
13. Березин С.Я., Каратаев Е.Г. Корреляционные измерительные устройства в автоматике. — М.: Энергия, 1976.
14. Биберман Л.М. Растры в электро-оптических устройствах. — М.: Энергия, 1969.
15. Бодс Г. Теория цепей и проектирование усилителей с обратной связью / Пер. с англ. — М.: Изд-во иностранной литературы, 1948.
16. Большаков И.А., Репин В.Г. Вопросы нелинейной фильтрации // Автоматика и телемеханика. — 1961. — № 4.
- ✚ 17. Бочкарев А.М. Корреляционно-экстремальные системы навигации // Зарубежная радиоэлектроника. — 1981. — № 9.
- 18. Булгаков Б.В. Прикладная теория гироскопов. — М.: Изд-во МГУ, 1976.
19. Бутицкий И.М. Устройства АРУ многоканальных систем связи. — М.: Связь, 1970.
- 20. Быков В.В. Цифровое моделирование в статистической радиотехнике. — М.: Советское радио, 1971.
21. Вайнштейн Г.Г. Преобразование изображений анизотропными фильтрами // Иконика. Цифровая голография. Обработка изображений. — М.: Наука, 1975.
22. Ван Трис Г. Теория обнаружения, оценок и модуляции. — М.: Советское радио, 1972.
23. Варенцов Г.А. Устройство для определения величины линейного сдвига изображения. — Авторское свидетельство № 1019666 // Бюллетень изобретений. — 1983. — № 19.

24. *Васильев Д.В., Витоль М.Р., Горшенков Ю.Н. и др.* Радиотехнические цепи и сигналы. — М.: Радио и связь, 1982.
25. *Васильев Д.В., Муранов Л.Н., Тимонин В.А.* Дискриминатор. — Авторское свидетельство № 885918 // Бюллетень изобретений. — 1981. — № 44.
26. Видеокорреляционное следящее устройство, использующее процессор для адаптивного обновления информации. — Патент США, кл. 358/126, № 4.227.212, 1980.
27. *Волосов И.Д.* Применение теории марковских процессов к определению статистических характеристик нелинейных АИС // Автоматика и телемеханика. — 1970. — № 2.
28. *Волохатюк В.А., Кочетков В.М., Красовский Р.Р.* Вопросы оптической локации. — М.: Советское радио, 1971.
29. *Гантмахер Ф.Р.* Теория матриц. — М.: Наука, 1966.
30. *Гершберг А.Е.* Передающие телевизионные трубки с внутренним фотоэффектом. — Л.: Энергия, 1973.
31. *Голд Б., Рэyder Ч.* Цифровая обработка сигналов / Пер. с англ. — М.: Советское радио, 1973.
32. *Голенко Д.И.* Моделирование и статистический анализ псевдослучайных чисел на ЭВМ. — М.: Наука, 1965.
33. *Гоноровский И.С.* Радиотехнические цепи и сигналы. — М.: Советское радио, 1977.
34. *Горбатенко С.А., Макашов Э.М.* Механика полета: Общие сведения. Уравнения движения // Инженерный справочник. — М.: Машиностроение, 1969.
35. *Гришин М.П., Курбанов Ш.М., Маркелов В.М.* Автоматический ввод и обработка фотографических изображений на ЭВМ. — М.: Энергия, 1976.
36. *Денисов В.С., Парамонов А.А., Сизов В.П.* Телевизионный координатный дискриминатор со стробом размером $M \times N$ элементов изображения. — Авторское свидетельство № 987853 // Бюллетень изобретений. — 1983. — № 1.
37. *Девычиков К.К.* Имитатор ИК-фонов // Изв. вузов, серия "Приборостроение". — 1969. — № 10.
38. *Дёч Г.* Руководство к практическому применению преобразования Лапласа и z -преобразования / Пер. с нем. — М.: Наука, 1971.
39. *Джеямисон Дж.Э., Мак-Фи Р.Х., Пласс Дж.Н.* Физика и техника инфракрасного излучения / Пер. с англ. — М.: Советское радио, 1965.
40. *Диткин В.А., Прудников А.П.* Интегральные преобразования и операционное исчисление. — М.: Физматгиз, 1961.
41. *Драган Я.П., Михайловский В.Н.* О развитии теории сигналов. Ч. I—III // Отбор и передача информации. Вып. 26—28. — Киев: Наукова думка, 1970, 1971.
42. *Драган Я.П.* Модели сигналов в линейных системах. — Киев: Наукова думка, 1972.
43. *Драган Я.П.* Структура и представления моделей стохастических сигналов. — Киев: Наукова думка, 1980.
44. *Ермолаев А.Г.* Оптимизация морфологических алгоритмов идентификации изображений // Автометрия. — 1982. — № 4.
45. *Жовинский В.Н., Арховский В.Е.* Корреляционные устройства. — М.: Энергия, 1974.
46. *Заде Л., Дезоер Ч.* Теория линейных систем. Метод пространства состояний / Пер. с англ. — М.: Наука, 1970.
47. *Заложнев Ю.И.* Корреляционная функция выходного сигнала сканирующей оптической системы // Оптико-механическая промышленность. — 1973. — № 2.
48. *Зуев В.Е.* Распространение видимых и инфракрасных волн в атмосфере. — М.: Советское радио, 1970.
49. *Зуев В.Е., Кабанов М.В.* Перенос оптических сигналов в земной атмосфере. — М.: Советское радио, 1977.
50. *Игнатьев Н.К.* Оптимальная дискретизация двумерных сообщений // Изв. вузов, серия "Радиотехника". — 1961. — № 6.
51. *Казмаров А.А., Палатник А.М., Роднянский Л.О.* Динамика двумерных систем автоматического регулирования. — М.: Наука, 1967.
52. *Карандасов В.И.* О влиянии приемной аппаратуры при корреляционном методе измерения временного сдвига случайных сигналов // Вопросы радиоэлектроники, серия XII общетехническая. — 1961. — Вып. 25.

53. *Карандасов В.И.* Устройство для автоматического измерения скорости летательного аппарата. — Авторское свидетельство № 145805 // Бюллетень изобретений. — 1962. — № 6.
- 54. *Катыс Г.П.* Оптико-электронная обработка информации. — М.: Машиностроение, 1973.
55. *Козубовский С.Ф.* Корреляционно-экстремальные системы. — Киев: Наукова думка, 1973.
56. *Козубовский С.Ф.* Общая теория квантования по уровню и ее применение для определения корреляции // Автоматика. — 1963. — № 1.
57. *Корн Г., Корн Т.* Справочник по математике для научных работников и инженеров. Определения. Теоремы. Формулы / Пер. с англ. — М.: Наука, 1974.
58. Корреляционная следящая ТВ-система и метод ее создания. — Патент США, № 4.270.143 // Реферативный журнал "Радиотехника" № 5Г145П. — 1982.
59. Корреляционный прибор сопровождения цели. — Патент Швеции, № 411.400 // Реферативный журнал "Радиотехника" № 8Г17П. — 1980.
60. *Косых В.П., Пустовских А.И., Яковенко Н.С.* Программная реализация морфологического процесса // Автометрия. — 1982. — № 4.
61. *Крамер Г.* Математические методы статистики / Пер. с англ. — М.: Мир, 1975.
62. *Красильников Н.Н.* Статистическая теория передачи изображений. — М.: Связь, 1976.
63. *Красовский А.А.* О двухканальных системах автоматического регулирования с антисимметричными связями // Автоматика и телемеханика. — 1957. — № 2.
64. *Красовский А.А., Белоглазов И.Н., Чигин Г.П.* Теория корреляционно-экстремальных навигационных систем. — М.: Наука, 1979.
65. *Круз П., Макглоулин Л., Макквистан Р.* Основы инфракрасной техники / Пер. с англ. — М.: Воениздат, 1964.
66. *Курочкин С.С.* Многоканальные счетные системы и коррелометры. — М.: Энергия, 1972.
67. *Лазарев Л.П.* Инфракрасные и световые приборы самонаведения и наведения летательных аппаратов. — М.: Машиностроение, 1970.
68. *Левин Б.Р.* Теоретические основы статистической радиотехники. Кн. I. — М.: Советское радио, 1966.
69. *Левин Б.Р.* Теоретические основы статистической радиотехники. Кн. III. — М.: Советское радио, 1976.
70. *Левитан Б.М.* Нормированные кольца, порожденные операторами обобщенного сдвига // Докл. АН СССР. — 1945. — Т. 47, № 1.
71. *Левитан Б.М.* Теория операторов обобщенного сдвига. — М.: Наука, 1973.
72. *Левитин И.Б.* Инфракрасная техника. — М.: Энергия, 1973.
73. *Левшин В.Л.* Обработка информации в оптических системах пеленгации. — М.: Машиностроение, 1978.
74. *Левшин В.Л.* Пространственная фильтрация в оптических системах пеленгации. — М.: Советское радио, 1971.
75. *Максимов М.В., Горгонов Г.И.* Радиоэлектронные системы самонаведения. — М.: Радио и связь, 1982.
76. *Марков М.Н.* Приемники инфракрасного излучения. — М.: Наука, 1968.
77. Межотраслевая обработка данных со спутников: Обзор // Зарубежная радиоэлектроника. — 1983. — № 8.
78. *Мерсеро Р.М.* Обработка двумерных сигналов при дискретизации по гексагональному растру // Труды Института инженеров по электротехнике и радиоэлектронике. — 1979. — Т. 67, № 6.
79. *Метлицкий Е.А., Смирнов А.В.* Обработка изображений матричными процессорами // Цифровые методы оптимальной обработки сигналов. — Новосибирск: Изд. Новосибирского электротехнического института, 1982.
80. Метод и устройство для разложения и электронной обработки термических изображений. — Патент ФРГ, кл. НО 1 N 533, № 2. 727. 811, 1979.
81. *Мирошников М.М.* Теоретические основы оптико-электронных приборов. — Л.: Машиностроение, 1977.
82. *Мирский Г.Я.* Аппаратурное определение характеристик случайных процессов. — М.: Энергия, 1972.

83. *Наймарк М.А.* Теория представлений групп. — М.: Наука, 1976.
84. *Налимов В.В., Чернова Н.А.* Статистические методы планирования экстремальных экспериментов. — М.: Наука, 1965.
85. *Носов О.Г.* Блок корреляции ТВ аналоговой корреляционно-экстремальной системы линейного позиционирования / Депонировано в Центре научно-технической информации "Информсвязь". — М.: 1980. — № 33, 7Г163 Деп.
86. *Носов О.Г.* Обеспечение требуемой зоны захвата в ТВ корреляционно-экстремальной системе / Депонировано в Центре научно-технической информации "Информсвязь". — М.: 1980. — № 32, 7Г190 Деп.
87. Обработка изображений и цифровая фильтрация / Под ред. Т. Хуанга; пер. с англ. — М.: Мир, 1979.
88. *Обрезков Г.В., Разевиг В.Д.* Методы анализа срыва слежения. — М.: Советское радио, 1972.
89. Основы автоматического регулирования: Теория / Под ред. В.В. Солодовникова. — М.: Машгиз, 1954.
90. *Первачев С.В.* Радиоавтоматика. — М.: Радио и связь, 1982.
91. *Первачев С.В., Валуев А.А., Чуликин В.М.* Статистическая динамика радиотехнических следящих систем. — М.: Советское радио, 1973.
92. Планирование эксперимента в исследовании технологических процессов / Под ред. Э.К. Лецкого; пер с нем. — М.: Мир, 1977.
93. *Полонников Р.И., Костюк В.И., Краскевич В.Е.* Матричные методы обработки сигналов. — Киев: Техника, 1977.
94. *Понтрягин Л.С.* Обыкновенные дифференциальные уравнения. — М.: Наука, 1965.
95. *Попов Е.П.* Теория линейных систем автоматического регулирования и управления. — М.: Наука, 1978.
96. Представление данных в обработке изображений: Обзор // Зарубежная радиоэлектроника. — 1983. — № 8.
97. Приборы с зарядовой связью / Под ред. Д.Ф. Барба; пер. с англ. — М.: Мир, 1982.
98. Проектирование оптических систем / Под ред. Р. Шеннона; пер. с англ. — М.: Мир, 1983.
99. Проектирование следящих систем с помощью ЭВМ / Под ред. В.С. Медведева. — М.: Машиностроение, 1979.
100. *Прэйт У.* Цифровая обработка изображений. Т. 1 / Пер. с англ. — М.: Мир, 1982.
101. *Пытьев Ю.П.* Морфологический анализ изображений // Докл. АН СССР. — 1975. — Т. 224, № 6.
102. *Пытьев Ю.П.* Проекционный анализ изображений // Кибернетика. — 1975. — № 3.
103. *Пытьев Ю.П., Текин В.В., Терентьев Е.Н.* Сравнительный анализ некоторых решающих алгоритмов на ЦВМ // Изв. АН СССР. Техническая кибернетика. — 1976. — № 4.
104. *Пытьев Ю.П., Терентьев Е.Н., Задорожный С.С.* Морфологический анализ изображений интегральных схем // Вестник МГУ, серия "Физика, астрономия". — 1976. — № 6.
105. *Растринин Л.А.* Статистические методы поиска. — М.: Наука, 1968.
106. *Розенфельд А.* Распознавание изображений // Труды Института инженеров по электротехнике и радиоэлектронике. — 1981. — Т. 69, № 5.
107. *Розенфельд А.* Распознавание и обработка изображений. — М.: Мир, 1972.
108. *Романенко А.Ф., Сергеев Г.А.* Вопросы прикладного анализа случайных процессов. — М.: Советское радио, 1968.
109. *Роуз А.* Зрение человека и электронное зрение / Пер. с англ. — М.: Мир, 1977.
110. *Русаков А.В., Васильев Д.В.* Ошибки дискретного вычисления функций автокорреляции случайного поля. Вып. 80 // Труды Московского института радиотехники, электроники и автоматики. — М.: 1975.
111. *Рыфтин Я.А.* Телевизионная система. — М.: Советское радио, 1967.
112. *Сверлинг П.* Современные методы оценки состояния с точки зрения метода наименьших квадратов // Зарубежная радиоэлектроника. — 1972. — № 9.
113. *Свешников А.А.* Прикладные методы теории случайных функций. — М.: Наука, 1968.
114. *Секен К., Томпсет М.* Приборы с переносом заряда / Пер. с англ. — М.: Мир, 1978.

115. *Сергеев Г.А., Януш Д.А.* Статистические методы исследования природных объектов. — Л.: Гидрометеиздат, 1973.
116. Система корреляции изображений. — Патент США, кл. 178 — 65, № 3. 617.025, 1971.
117. Система регистрации сдвоенных изображений. — Патент США, кл. 178 — 68, № 3.636.254, 1972.
118. *Смирнов В.А.* Основы радиосвязи на УКВ. — М.: Связьиздат, 1957.
119. *Смит Р., Джонс Ф., Чесмер Р.* Обнаружение и измерение инфракрасного излучения / Пер. с англ. — М.: Изд-во иностранной литературы, 1959.
120. *Спили К., Браун Р., Гудвин Дж.* Теория управления. Идентификация и оптимальное управление / Пер. с англ. — М.: Мир, 1973.
121. Справочник по приборам инфракрасной техники / Под ред. Л.З. Криксунова. — Киев: Техника, 1980.
122. *Тарасенко В.П., Раводин О.М.* Корреляционно-экспериментальный координатор для речных и морских судов // Поиск экстремума. — Томск: Изд-во Томского университета, 1969.
123. *Тарасенко В.П.* Применение оптических функциональных преобразователей для целей распознавания двумерных геометрических образов // Труды Сибирского физико-технологического института. Вып. 44. — Томск: Изд-во Томского университета, 1964.
124. *Тартаковский Г.П.* Динамика систем автоматической регулировки усиления. — М.-Л.: Госэнергиздат, 1957.
125. *Тихонов В.И.* Статистическая радиотехника. — М.: Радио и связь, 1982.
126. *Ту Ю.* Цифровые и импульсные системы автоматического управления / Пер. с англ. — М.: Машиностроение, 1964.
127. Усовершенствование системы сопровождения цели для головок самонаведения ракет. — Патент Великобритании, № 1.567.010 // Реферативный журнал "Радиотехника" № 11Г43П. — 1980.
128. Устройство для геометрической корреляции между двумя изображениями одного объекта. — Заявка Франции, кл. G06K15/20, № 2.267.591, 1975.
129. *Федоров В.В.* Теория оптимального эксперимента (планирование регрессионных экспериментов). — М.: Наука, 1971.
130. *Федоров С.М., Лучко С.В.* Некоторые особенности цифрового (дискретного) управления объектами с резко выраженными резонансными свойствами // Прикладные задачи технической кибернетики / Под ред. А.С. Гурина, И.А. Ушакова, Б.С. Флейшмана. — М.: Советское радио, 1966.
131. *Хадсон Р.* Инфракрасные системы / Пер. с англ. — М.: Мир, 1972.
132. *Цветков В.В., Карандасов В.И.* Флюктуационные ошибки корреляционных измерителей // Радиотехника. — 1965. — № 7.
133. *Цветков В.В., Пономарева В.Д.* Влияние n -кратного дифференцирования на флюктуации ошибок корреляционных измерителей // Радиотехника. — 1969. — № 4.
134. Цифровое телевидение / Под ред. М.И. Кривошеева. — М.: Связь, 1980.
135. *Цыпкин Я.З.* Теория линейных импульсных систем. — М.: Физматгиз, 1963.
136. *Чигин Г.П.* Проблемы моделирования корреляционно-экстремальных навигационных систем // Корреляционно-экстремальные системы управления. — Томск: Изд-во Томского университета, 1979.
137. *Шварц Л.* Математические методы для физических наук // Пер. с фр. — М.: Мир, 1965.
138. *Шершень А.И.* Аэрофотосъемка. Летно-съемочный процесс. — М.: Геодезиздат, 1949.
139. *Шестов Н.С.* Выделение оптических сигналов на фоне случайных помех. — М.: Советское радио, 1967.
140. *Ширман Я.Д., Манжос В.Н.* Теория и техника обработки радиолокационной информации на фоне помех. — М.: Радио и связь, 1981.
141. *Шкурский Б.И.* Метод математического моделирования двумерных случайных полей // Изв. АН СССР. Техническая кибернетика. — 1969. — № 6.
142. *Шмаков П.В.* Телевидение. — М.: Связь, 1979.
143. *Эйххофф П.* Основы идентификации системы управления: Оценка параметров и состояния / Пер. с англ. — М.: Мир, 1975.

144. Электронно-оптическая система наведения. — Патент США, кл. 244—3.17, № 3.794.272, 1967.
145. Эндрюс Г. Применение вычислительной машины для обработки изображений / Пер. с англ. — М.: Энергия, 1977.
146. Якушенков Ю.Г. Основы оптико-электронного приборостроения. — М.: Советское радио, 1977.
147. Якушенков Ю.Г., Луканцев В.Н., Колосов М.П. Методы борьбы с помехами в оптико-электронных приборах. — М.: Радио и связь, 1981.
148. Ярославский Л.П. Введение в цифровую обработку изображений. — М.: Советское радио, 1979.
149. Ярославский Л.П. Устройство ввода-вывода изображений для ЭЦВМ. — М.: Энергия, 1968.
150. Aerospace Daily. — 1974. — V. 66, No. 8, p. 57, 58.
151. Aviation Week. — 1983. — V. 118, No. 8, p. 139.
152. Aviation Week. — 1984. — V. 120, No. 7, p. 86.
153. Aviation Week. — 1984. — V. 120, No. 9, p. 22.
154. Aviation Week & Space Technology. — 1984. — V. 121, No. 11, p. 77.
155. Aviation Week & Space Technology. — 1984. — V. 121, No. 12, p. 66.
156. Aviation Week & Space Technology. — 1985. — V. 122, No. 6, p. 125.
157. Bergemann H. Wärmebild-Fernsehkameras mit Pyricous // Elektronik. — 1981. — B. 30, Nr. 17.
158. Berry J.E., Yoo J.K. Preprocessing for advanced image matching techniques // Proceedings IEEE-NAECON 79, Dayton. — N.Y.: 1979. — V. 1.
159. Binary Correlation video tracker. — Pat. USA, cl. 3.581.125, No. 4.364.089, 1982.
160. Boland J., Peters E., Pinson L., Malkolm W. Automatic correlation of non-compatible imaging systems // Proceedings SOUTHEASTCON 79, Reg. 3 Conf. Roanoke, 1979. — N.Y.: 1979.
161. Choate W., Boyd W. Correlation tracking concepts for THASSID // Proceedings IEEE-NAECON 79, Dayton. — N.Y.: 1979, V. 1.
162. Cunningham R.C. Sensors for second generation FLTR's // Elect-Opt. Syst. Des. — 1979. — V. 11, No. 12.
163. Delsarte J. Sur une extension de la formule de Taylor // Journ. Math. Pures et Appl. — 1938. — V. 17.
164. Gabor D. The smoothing and filtering of twodimensional imaging // Progress in Biocybernetics. V. 2. — Amsterdam: 1965.
165. Gedance A.R. Comparison of infrared tracking systems // J. Opt. Soc. Am. — 1961. — V. 51, No. 1127.
166. Colay M. Hexagonal parallel pattern transformations // IEEE Trans. Comput. — 1969. — V.C. — 19.
167. Cuanella G. Einige Anwendungen der Korrelationsmethode beim Schwingungsempfang // Nachr. Techn. Fachber (Beiheft der NTZ). B. 3. Informationstheorie. — Braunschweig; Verlag F. Viewe, 1956.
168. Martineau D., Gene R., Wong T. HgCdTe: number one night vision // Opt. Spectra. — 1979. — V. 13, No. 7.
169. Meyr H. Zur Synthese der Diskriminatorkennlinie in Tracking-systemen // Scieles. — 1977. — B. 23, Nr. 3.
170. Mostafavi H., Smith F.W. Image correlation with geometric distortion // IEEE Trans. on Aerosp. & El. Syst. — 1978. — V.AES — 14, No. 3.
171. Nicholson A.F. Error signals and discrimination on optical trackers that see several sources // Proceedings IEEE. — 1965. — V. 53, No. 1.
172. Novak L. Correlation algorithms for radar map matching // Proceedings of the 1976 IEEE Conference of Decision & Control, December, 1976.
173. Ormsby C. Advanced scene matching techniques // Proceedings IEEE Nat. Aerospace & Elektron. Conf. — NAECON 1979. — Dayton: 1979, V. 1.
174. Pratt W. Correlation techniques of image registration // IEEE Trans. on Aerosp. & El. Syst. — 1974. — V.AGS-10, No. 4.
175. Raina J. Real time measurement of TV signal statistics // Proceedings Conf. Telev. Meas. — London: 1979.
176. Reed C.G., Hogan I.I. Range correlation guidance for cruise missiles // IEEE Trans. on Aerosp. & El. Syst. — 1979. — V. 15, No. 4.

177. *Regelson E.* Impact of divices and aircraft design trends // *Astronautics & Aeronautics*. – 1967. – V. 5, No. 8.
178. *Shinazuka M.* Simulation of multivariable and multidimensional random processes // *Institute Acoust. America*. – 1971. – V. 49, No. 1.
179. *Sundaram G.* The french modular thermal-imaging system // *Int. Defense Review*. – 1982. – V. 15, No. 4.
180. *Tisdale G.E.* Digital image processors for automatic target cueing // *EASCON 80 Rec.* Arlington. 1980. – N.Y.: 1980.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Автоколебания 161
Автомат корреляционный 24
Аксиомы групповые 52
Алгоритм одномерный 69
Анализ морфологический 13, 79
– устойчивости 127
Анализатор изображения 100
– матричный 176
– – модулирующий 176
– – растровый 176
– – телевизионный 194
Анизотропность изображений 89
Аппроксимация дискретная 223
Ассоциативность 52

Базис Карунена – Лозва 32

Визуализация полей 249

- Гипергруппа 55
- Гиросплатформа 115
- Гироспривод индикаторный 111
- Гироскоп трехстепенный 111
- Головка самонаведения 26
- Гомоморфизм 80
- Группа 52
 - абелева 52
 - аддитивная 52
 - аффинная 54
- Группа коммутативная 52
 - мультипликативная 52

Дельта-модуляция 211
Детектор амплитудный 103

- фазовый 104

Деформация ракурсная 53
Дискриминатор 9

- контрастный 20
- корреляционный 11
- – аналоговый 64
- – дифференциальный 68

Дуальность вариантов 37

Единица группы 52

Звенья линейные 217

Идентификация модели 153
Излучение распределенное 156
Изображение непрерывное 204

- цифровое 204

Инверсия знака 99
Индекс модуляции 23
Источник излучения двойной 138

- – перемещающийся 161

Квантователь вероятностный 209

- логарифмический 210

Кодирование 209
Коммутативность 52
Контраст определяющий 266
Контрастер 22
Координатор 22

- следящий 108

Корреляция двумерная 44

- обобщенная 51
- перекрестная 41

Критерии оптимальности 275
Критерий корреляционный 206
Кружок рассеяния 178
Крутизна преобразования 197

Линия силовая 141

Матрица весовая 33

- дисперсионная 256
- информационная 256

Метод прогонки 150
Методы компенсации 92
Микродвигатель 108
Модели математические 108
Моделирование, виды 172
Модель неадекватная 258

- обобщенная 145
- поэлементная 174
- регрессионная 294
- эталонная 287

Модулятор 191

Модуляция амплитудно-импульсная 127
– импульсно-кодовая 211
– широтно-импульсная 132
Момент кинетический 111

Неравномерности фона 98
Нормировка 41

Область устойчивости 130
Оператор метризирующий 33
– обобщенного сдвига 55
– пакетный 184
Ортогонализация 69
Отношение генерирующее 265
Оценивание оптимальное 69
Оценка помехозащищенности 148
Ошибка динамическая 161

План Кифера 277
– композиционный 269
– насыщенный 265
– непрерывный 277
– оптимальный 275
– ортогональный 269
– точный 277
– факторный 260
– центральный 269
Планирование композиционное 270
Платформа гиросtabilизированная 114
Плотность спектральная взаимная 49
Полудиск вращающийся 189
Полуребрика 264
– главная 267
Поля случайные 245
Помехозащищенность 148
Предсказатель 211
Преобразование билинейное 227
– Ганкеля 245
– нелинейное 78
– проективное 51
Преобразования ортогональные 213
Преобразователь морфологический 79
– однородный 49
Приемник многоплощадочный 57
– одноплощадочный 100
Принцип транспозиции 49
Прогонка 151
Продолжение четное 209
Процесс сингулярный 244
– сопровождения 138

Развертка построчная 64
Расширитель 85
Регулировка усиления 82
Регулятор экстремальный 27
Режим установившийся 50
Реплика дробная 263
Решение нестационарное 148
Решения дуальные 37
Решетка растровая 22

324

Рототабельность 269
Ряд Эджворта 244

Свертка обобщенная 58
Свойство ассоциативности 52
– коммутативности 52
Связи антисимметричные 115
Сдвиг 29
– групповой 51
– масштабно-креновый 55
– обобщенный 51
– ракурсный 55
– сопряженный 56
Селекция угловая 19
Сигнал телеграфный 247
Силы коррекции 140
Система активная 176
– пассивная 176
– полуактивная 176
– сканирующая 200
– стационарная 50
Смещение нуля 98
Сопровождение 138
Спектр обобщенный 58
Строб-эффект 205

Точки звездные 270
– особые 142
– равновесия 140
Тракт оптико-электронный 177
– электронный 215
Транспозиция 49

Угломер следящий 26
Угол апертурный 178
Узел внутренний 149
– граничный 149
Умножение групповое 52
Управление импульсное 105
– непрерывное 100
Усилитель резонансный 102
Устойчивость сопровождения 127

Факторы декоррелирующие 294
– дестабилизирующие 81
Фильтр анизотропный 92
– линейный 49
– неоднородный 51
– нерекурсивный 221
– ортогональный 73
– параметрический 51
Фильтрация гомоморфная 79
– линейная 64
– цифровая 230
Функция автокорреляционная обобщенная 58
– взаимно корреляционная обобщенная 57
– Грина 50
– качества 286

Функция корреляционная обобщенная
56

– рассеяния точки 50, 178

Характеристика внешняя 108

– дискриминационная 15

– импульсная 50

– пеленгационная 101

– регулировочная 108

– спектрально-корреляционная 49

Характеристики флуктуационные 95

Цепь фиксирующая 85

Цикл предельный 116

Шум генерационно-рекомбинационный
199

Шум дробовой 199

– тепловой 199

– токовый 199

Шумы приемников 198

Эксперимент активный 260

– факторный полный 260

Элемент исполнительный 108

– противоположный 52

Эффекты краевые 44

Ядро плана 270

z -преобразование 225

z -формы 225

ABSTRACTS

Astapov Yu.M., Vasiljev D.V., Zaloghnev Yu.I. **Theory of electro-optical tracking systems.** — Moscow. Nauka, 1988.

The monography reflects some modern investigation methods for optical tracking electronics systems. These are represented as multidimensional dynamic systems with nonlinear characteristics. The problem of periodic motions is solved.

There are considered the general algorithms for measuring blocks of image trackers. The conception is based on the correlation methods of vector parameter estimation and on the elements of generalized translation theory. The applied results are oriented to TV and other scanning systems for tracking of pictures with arbitrary statistics.

Tracking precision in the presence of random disturbances is analysed. In particular the solution of the problem of tracking process failure in terms of Fokker-Plank-Kolmogorov's equation is adduced.

Considerable attention is paid to simulation methods of these complicated systems. Mathematical models are based on identification theory. Special attention to the reducing of the machine time is spared.

The authors are well-known scientists. They are dealing with different problems of theoretical radiotechnique and automatic control.

The book is intended for scientists and engineers engaged in the multidimensional tracking systems and image processing.

*Астапов Юрий Михайлович
Васильев Дмитрий Валерьевич
Заложнев Юрий Иванович*

ТЕОРИЯ ОПТИКО-ЭЛЕКТРОННЫХ СЛЕДЯЩИХ СИСТЕМ

Редактор *Ю.Г. Гуревич*
Художественный редактор *Т.Н. Кольченко*
Технические редакторы: *С.В. Геворкян, С.Н. Баронина*
Корректоры: *Л.И. Назарова, Е.Ю. Рычагова, Н.П. Круглова*

Набор осуществлен в издательстве
на наборно-печатающих автоматах

ИБ № 12411

Сдано в набор 11.01.88. Подписано к печати 15.04.88. Т — 09563
Формат 60 X 90/16. Бумага для множительных аппаратов
Гарнитура Пресс-Роман. Печать офсетная
Усл.печл. 20,5. Усл.кр.-отт. 20,5. Уч.-изд.л. 22,19
Тираж 2700 экз. Тип. зак. 33. Цена 4 р. 50 к.

Ордена Трудового Красного Знамени
издательство "Наука"

Главная редакция физико-математической литературы
117071 Москва В-71, Ленинский проспект, 15

Четвертая типография издательства "Наука"
630077 г. Новосибирск-77, ул. Станиславского, 25

ИЗДАТЕЛЬСТВО "НАУКА"

ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ

117071 Москва В-71, Ленинский проспект, 15

ГОТОВЯТСЯ К ИЗДАНИЮ:

Ким Д.П. Методы поиска и преследования подвижных объектов. — 20 л. (Аннотированный план на 1989 г., п. 141)

Большая часть книги посвящена задачам поиска подвижных объектов. Процесс поиска определяется как управление сближением по неполной априорной информации, обнаружение — как выполнение заданных терминальных условий. Поиск формализуется как статистические и игровые задачи управления (принятия решений). В последней главе рассматриваются различные алгоритмы преследования — классические (эвристические) и полученные на основе теории игр и оптимального управления.

Для специалистов в области теории управления (в частности, управления подвижными объектами), моделирования, прикладной математики.

Федосов Е.А., Инсаров В.В., Селивохин О.С. Системы управления конечным положением в условиях противодействия среды. — 20 л. (Аннотированный план на 1989 г., п. 149)

Рассматривается совокупность теоретических и прикладных методов исследования динамики нелинейных нестационарных систем управления подвижными объектами, основная цель которых — достижение конечного положения в заданном пространстве состояний. Системы рассматриваемого класса характеризуются наличием в структуре непрерывных, импульсных, цифровых и нелинейных звеньев, а также возможностью случайного изменения самой структуры системы в процессе функционирования под воздействием внешней среды и в режимах адаптации к таким воздействиям. На конкретных примерах непрерывных, дискретных и дискретно-непрерывных систем иллюстрируется применимость инженерных методов анализа и синтеза, доведенных до вычислительных алгоритмов.

Для специалистов в области теории управления, в частности управления движущимися объектами.

