3 р. 50 к.

ВИДЕОИНФОРМАТИКА

Видеоинформатика — раздел информатики, связанный с передачей и компьютерной обработкой изображений и определяющий границы возможного при решении этих задач.
В предлагаемой книге с позиций обработки изображений развивается теория информации, вводится принцип идеального линейного кодирования и понятие шумовой информации.
Рассматриваются компьютеризированные системы обработки изображений в реальном времени на приборах с зарядовой связью, в которых используются новые способы считывания видеоинформации.

«РАДИО И СВЯЗЬ»

Л.И.ХРОМОВ, А.К.ЦЫЦУЛИН
А.Н.КУЛИКОВ

ВИДЕОИНФОРМАТИКА

ПЕРЕДАЧА И КОМПЬЮТЕРНАЯ ОБРАБОТКА ВИДЕОИНФОРМАЦИИ

«РАДИО И СВЯЗЬ»

Л.И.ХРОМОВ, А.К.ЦЫЦУЛИН
А.Н.КУЛИКОВ

ВИДЕОИНФОРМАТИКА
ПЕРЕДАЧА И КОМПЬЮТЕРНАЯ ОБРАБОТКА ВИДЕОИНФОРМАЦИИ

Уважаемому Владимиру Борисовичу от авторов.
Здесь кадры сняты, выходные в сборнике "кадров" порядке обсуждаются.
Жилин
25.04.91.

Москва
"РАДИО И СВЯЗЬ"
1991

УДК 621.391:681.3

Хромов Л. И. и др. Видеоинформатика. Передача и компьютерная обработка видеоинформации/Л. И. Хромов, А. К. Цыцулин, А. Н. Куликов. — М.: Радио и связь, 1991. — 192 с.: ил. — **ISBN 5-256-00892-7.**

Рассмотрен раздел информатики, изучающий передачу и алгоритмическую переработку изображений для принятия решений — видеоинформатика. В целях оптимизации переработки видеоинформации в реальном времени развита концепция идеального линейного безблокового кодирования; доказана теорема об идеальном линейном кодировании для гауссовских каналов при квадратичном критерии точности; введено понятие о шумовой информации и определены пути ее уменьшения; получено информационное соотношение неопределенности, выражающее закон, по которому уменьшение шумовой информации достигается ценой роста потери полезной информации; разработаны нетрадиционные способы считывания видеоинформации с матричных приборов с зарядовой связью: с фазных электронов и переменной по полю четкости; введены меры для ошибок неинвариантности к сдвигу изображений, лимитирующих достоверность классификации наряду с внешним фоном, шумом и доступной сложностью видеокомпьютера. Приложение основ проектирования информационно-решающих систем дано на примере телевизионных роботов, спецификой которых является необходимость экстраполяции положения подвижных объектов.

Для научных работников в области компьютерной переработки видеоинформации: в телевидении, оптике и радиолокации; может быть рекомендована инженерам.

Табл. 1. Ил. 58. Библиогр. 118 назв.

Рецензент доктор техн. наук Ю. Б. ЗУБАРЕВ

Редакция литературы по информатике и вычислительной технике

Х $\frac{1404000000\text{-}015}{046(01)\text{-}91}$ 30—91

ISBN 5-256-00892-7

© Хромов Л. И., Цыцулин А. К., Куликов А. Н., 1991

ПРЕДИСЛОВИЕ

Пути вычислительной техники и техники телевидения пересеклись в начале 70-х годов, когда благодаря успехам физики твердого тела и микроэлектроники почти одновременно появились микропроцессор и прибор с зарядовой связью (ПЗС). На новом этапе вычислительная и телевизионная техника получила возможность развиваться на основе единой кремниевой технологии. Однако быстрому соединению ПЗС и микропроцессора препятствовали технологические трудности: недостаточная четкость формируемых ПЗС изображений и малое быстродействие микропроцессоров. Понадобилось более десятилетия для того, чтобы ПЗС удовлетворили требованиям современного телевизионного вещания и стали сегодня наиболее перспективными для телевидения высокой четкости. В 80-е годы наблюдался интенсивный рост быстродействия микропроцессоров и емкости полупроводниковой оперативной памяти, который позволил проектировать в виде единого комплекса приемлемых габаритных размеров твердотельную телевизионную систему «от света до принятия решения».

Объединение телевизионной и вычислительной техники, выполненной на основе кремниевой технологии, открывает новые возможности в физике и астрономии для решения проблемы автоматизации наблюдения изображений, недоступных глазу человека. Рассмотрению этой проблемы научного приборостроения и посвящена данная работа. В отличие от предыдущей книги авторов («Твердотельное телевидение». — М.: Радио и связь, 1986), которая подводила итоги начального этапа развития твердотельного телевидения, в новой книге рассматриваются проблемы, полная реализация которых ожидается в 90-х годах. Важнейшая из них — проектирование телевизионных роботов, способных благодаря совместному использованию телекамеры и компьютера оценивать быстропротекающие процессы и принимать решения со скоростью, недоступной человеку. Такие телевизионные роботы являются необходимым элементом в создании систем адаптивной оптики, а также в астрономии для автоматизированного наблюдения, например, за быстропротекающими процессами на Солнце.

Авторы, ведущие специалисты в области прикладного телевидения, излагают материал, интересующий специалистов различных отраслей науки и техники (оптики, телевидения, радиолокации и др.), ранее изучавших процессы кодирования изображений. Понятия об источнике информации, о каналах передачи и о процессе кодирования сигналов применимы, по существу,

к модели любого научного прибора, с помощью которого ученый получает нужную ему информацию. Работа иллюстрирует творческое применение теории передачи информации к важной практической задаче: проектированию нового поколения роботов, снабженных телевизионным зрением и компьютером — классификатором изображений.

Книга, несомненно, будет способствовать интеграции знаний в области кодирования и алгоритмической переработки оптических и радиолокационных изображений в интересах развития отечественного научного приборостроения.

Академик *А. А. Боярчук*

ОТ АВТОРОВ

Видеоинформатика определена нами как раздел информатики, изучающий передачу, хранение и использование информации, содержащейся в изображениях. Ее цель — ускорить интеграцию достижений специалистов разных областей науки и техники (оптики, телевидения, радиолокации) для создания новых компьютеризованных систем.

Книга представляет собой обобщающее изложение результатов работы авторов, опубликованных ранее в научных статьях и авторских свидетельствах на изобретения и отражающих следующий этап в развитии прикладного телевидения [1, 2]. Она отличается от книг отечественных и зарубежных авторов по цифровой обработке сигналов [54—56, 58—65] тем, что развивает теорию информации, базируясь на работе Шеннона 1959 г. [3, с. 587] применительно к аналоговому и цифровому кодированию в реальном времени. Теоретические результаты используются при проектировании информационно-решающих систем на ПЗС.

Введение, гл. 1, 2 и раздел 3.1 написаны Л. И. Хромовым, гл. 4, 5 и разделы 3.2, 3.8 — А. К. Цыцулиным, разделы 3.3, 3.5, 3.7 — А. Н. Куликовым, разделы 3.4, 3.6 — А. Н. Куликовым и А. К. Цыцулиным совместно. При изложении математических аспектов теории информации авторы с благодарностью воспользовались советами А. Б. Ковригина. Авторы также благодарят докторов техн. наук М. А. Грудзинского и Ю. Б. Зубарева и канд. техн. наук В. С. Нощенко, Л. А. Литвинчука, А. В. Мартынихина за высказанные критические замечания и поддержку работ по видеоинформатике.

ВВЕДЕНИЕ

В.1. КОМПЬЮТЕРИЗАЦИЯ И ВИДЕОИНФОРМАТИКА

Компьютеризация определяет сегодня уровень научно-технического прогресса, охватывает все области науки и техники. Компьютеризация стала возможной благодаря достижениям физики твердого тела, приведшим к развитию твердотельной электроники. Современный этап компьютеризации связан с созданием сверхбыстродействующих и сверхбольших интегральных схем. На очереди использование еще более революционных технологических достижений: приборов, применяющих высокотемпературную сверхпроводимость, и оптических компьютеров. Овладение потенциальными возможностями параллельного метода переработки информации позволит радикально увеличить быстродействие компьютеров.

Компьютер является устройством, перерабатывающим информацию по тому или иному алгоритму с принятием решения, и вместе с устройством, поставляющим на его вход информацию, образует единую систему, которая должна иметь общие принципы проектирования. Поэтому термины «информатика» и «вычислительная техника» тесно связаны.

Появление информатики отражает тенденцию к интегрированию разных областей знаний на основе представлений теории информации и применения компьютеров. «Информация» стало одним из самых модных слов в лексиконе современного человека. Однако случилось это отчасти потому, что слово «информация» имеет неоднозначное и широкое смысловое значение. В искусстве связывают с этим понятием передачу эмоций. Его трактуют и как новости в науках или в быту. В таких эмоциональных и смысловых трактовках, как еще заметил Хартли [7], информация недоступна числовой оценке. Математическая теория информации была развита только применительно к весьма узкому, специфическому понятию об информации — несемантической информации. А. Н. Колмогоров классифицировал три подхода к понятию «количество информации»: чисто комбинаторный, введенный Р. Хартли в 1928 г. [7], вероятностный подход Шеннона [3] и алгоритмический подход, развитый А. Н. Колмогоровым [4].

Термин «информатика» также может иметь широкое смысловое значение, включая все виды сообщений, получаемых человеком. То же справедливо и для предлагаемого термина «видеоинформатика». Видеоинформатика является одним из разделов информатики, оперирующим с изображениями. Мы ограничимся

достаточно узкой трактовкой видеоинформатики, рассматривая ее как технический раздел информатики, изучающий передачу (или хранение) несемантической видеоинформации (изображений) и ее алгоритмическую переработку с принятием решения.

Проблемами передачи и переработки изображений занимаются в различных областях: в оптике, телевидении и радиолокации. В оптике всегда изучалось качество формируемых объективами изображений и методы его улучшения. Развитие этой тематики привело к широкому использованию радиоэлектронных устройств обработки изображений, в том числе и ЭВМ. В телевидении и радиолокации всегда занимались проблемами передачи и кодирования изображений. Компьютеризация процессов переработки изображений заставляет объединить усилия специалистов для решения общих информационных проблем.

Для принятия компьютером решения необходимо формировать поток полезной информации в минимально достаточном количестве. Согласование информационного потока с возможностями компьютера становится важнейшим требованием, аналогичным согласованию источника с каналом в теории информации.

В.2. НЕКОТОРЫЕ НАПРАВЛЕНИЯ РАЗВИТИЯ ВИДЕОИНФОРМАТИКИ

В настоящее время рано говорить об информатике и ее разделе — видеоинформатике, как сложившемся научном направлении, обладающем методологическим единством. Скорее эти термины отражают желаемую тенденцию к становлению нового направления. Здесь нельзя повторять историю увлечения широтой кибернетики, которая, несмотря на богатство идей Винера, была лишена единства как математическая дисциплина [4, с. 57]. С самого начала видеоинформатика должна избежать опасности чрезмерного расширения своих границ и расплывчатости основных понятий. Для этого в первую очередь необходимо сформулировать направления развития видеоинформатики как научной дисциплины, обладающей единством методов. Эти направления должны содержать теоретическую и техническую части. Успех развития видеоинформатики будет зависеть от того, будут ли найдены общие принципы, законы или уравнения, которые могли бы составить теоретический фундамент видеоинформатики. Нам кажется очевидным, что построение теоретических основ видеоинформатики должно базироваться на теории информации [3] и ее новой ветви — алгоритмической теории информации (теории алгоритмов) [4]. Вторым «блоком» теоретического фундамента служит теория решений [5], относящаяся, как и теория информации, к разделам математической статистики.

Для создания теоретических основ видеоинформатики нужны совместные усилия математиков и представителей отраслевых наук: оптики, телевидения и радиолокации. Эти науки находятся на разных этапах теоретического развития. Передовое место занимает радиолокация благодаря многолетнему опыту применения в своих разработках теории решений [8]. Телевидение позже стало на путь практического приложения теории решений. Только твердотельное телевидение использовало опыт радиолокации, развив его методы борьбы с фоновым сигналом [2]. К сожалению, даже радиолокация не может «похвастаться» успешным применением теории информации. Первая попытка в этом направлении [9], хотя и была интересной, но не имела инженерного приложения. Не лучше обстояли дела с применением теории информации в телевидении и оптике. Одну из причин сложившегося здесь неблагополучного положения мы видим в разрыве между математиками, развивающими теорию информации, и отраслевыми научными работниками, применяющими теорию информации. Приведем лишь один пример, ярко, по нашему мнению, иллюстрирующий разрыв: до сих пор обобщенная теорема кодирования, сформулированная Шенноном в 1959 г. [3, с. 587], остается «вещью в себе» для инженеров. Их «теоретический» багаж ограничивается хотя и основополагающей, но не исчерпывающей работой Шеннона 1948 г. [3, с. 243]. А важные вопросы идеального кодирования, ради решения которых Шеннон потратил время, разделяющее эти две работы, остаются вне поля зрения прикладников. Виной этому отсутствие работ математиков, аналогичных докладу А. Н. Колмогорова по теории Шеннона 1948 г. [4]. Инженеры не остались в долгу и «скрыли» от математиков важность учета шумов источников информации.

Ликвидация создавшегося положения с применением теории информации входит первым пунктом в теоретическую часть программы развития видеоинформатики.

Техническая часть программы охватывает вопросы построения конкретных систем видеоинформатики. Общие черты этих систем должна отражать их математическая модель. К таким моделям предъявляют противоречивые требования: они должны быть просты и в то же время правильно отражать многообразие практических систем. Обратимся к модели Шеннона систем передачи информации (рис. В.1) [3, с. 245]. Модель Шеннона подчеркивает роль канала с его источником шума. Источник информации в модели Шеннона не имеет шума. Кодер и декодер выполняют функции согласования канала с источником и получателем информации. Подходит ли схема на рис. В.1 для моделирования систем в видеоинформатике? Несмотря на многообразие систем видеоинформатики, в них можно выделить три общие части: информационную подсистему, устройство первичной цифро-

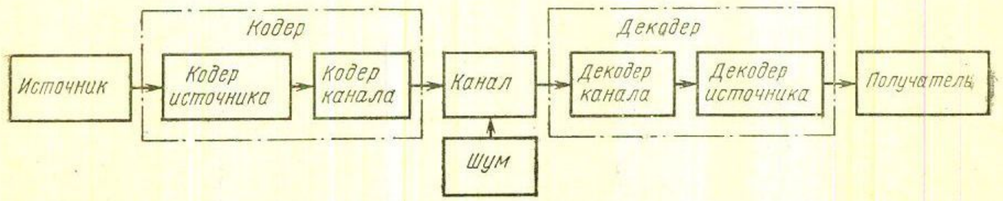


Рис. В.1. Модель системы связи по Шеннону

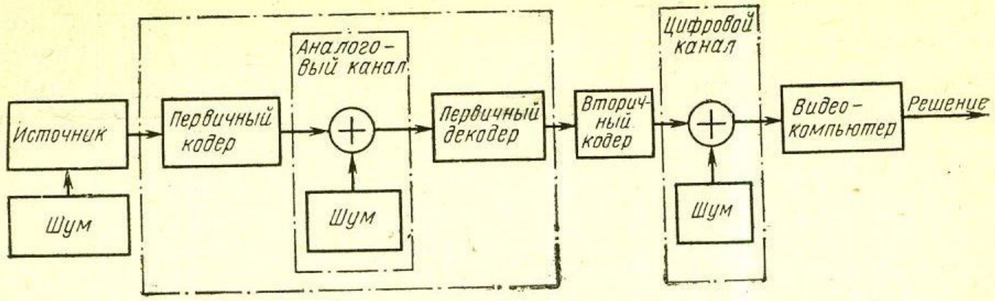


Рис. В.2. Модель информационно-решающей системы твердотельного телевидения со смешанным типом кодирования

вой обработки и видеокомпьютер. Передача информации в системе осуществляется с помощью каналов двух типов: непрерывных (аналоговых) и дискретных (цифровых). Соответственно и кодирование будет смешанным: и аналоговым, и цифровым. Для конкретизации сказанного обратимся к модели информационно-решающей системы в твердотельном телевидении, представленной на рис. В.2 [2]. Информацию об изучаемых природных объектах несет электромагнитное поле, которое служит природным источником полезной информации на фоне шума поля (фотонного и теплового). Первичное кодирование, которое макроскопически является непрерывным, осуществляется в информационной подсистеме — телекамере. Сигнал с выхода ПЗС в телекамере поступает во вторичный кодер, который инженерно реализуется с помощью аналого-цифрового преобразователя (АЦП) и устройства предварительной дискретно-аналоговой обработки. Вторичный кодер работает на цифровой канал, ведущий к видеокомпьютеру. В отсутствие специализированного видеокомпьютера его функции выполняет обычно универсальная мини-ЭВМ с частичной доработкой, например с увеличением емкости оперативной памяти на несколько телевизионных кадров.

Сравним теперь схемы на рис. В.1 и В.2. Видно, что возможно двухэтапное применение модели Шеннона (рис. В.1) к схеме на рис. В.2: сначала к непрерывному каналу, а затем к следующему за ним дискретному. Но и при таком применении модель на рис. В.1 должна быть изменена. Необходимо включить в нее шум источника информации (рис. В.3) в целях более полного описания специфики наблюдения природных источников (на-

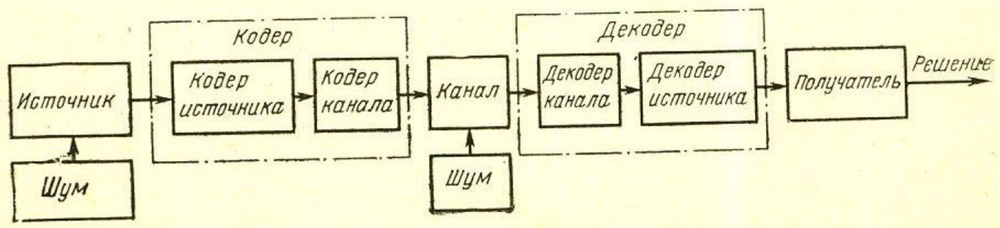


Рис. В.3. Обобщенная модель системы связи

пример, звезд в астрономии). Важно уточнить функцию получателя информации. Шеннон допускал в модели на рис. В.1 два типа получателя: человек и автомат [3, с. 246]. Видеокомпьютер относится к типу получатель-автомат. Назначение видеокомпьютера заключается в алгоритмической переработке информации с выдачей решения. Проблема алгоритмической переработки информации в решение [4] ранее не рассматривалась в теории Шеннона [3] и потому выход получателя в схеме на рис. В.1 не выдает решения для системы управления.

Специфика схемы на рис. В.2 состоит в чередовании аналогового и цифрового каналов. Такое чередование типов каналов в сложных системах, содержащих много последовательно соединенных каналов, может осуществляться даже неоднократно. Поэтому цифровые и аналоговые сигналы, цифровые и аналоговые кодеки и каналы не конкурируют, а дополняют друг друга при передаче и переработке информации. И тем не менее большинство специалистов отдает предпочтение цифровому кодированию, относя аналоговое к устаревшему, которое излагается в виде фильтрации сигналов, обычно в теории оптимального приема [84] раздельно от теории передачи информации. (Фильтрация является важным, но не единственным этапом кодирования.) Такая разобщенность «передачи» и «приема» закреплена в радиотехнике дифференциацией учебных курсов и специальностей в вузах. Более того, о линейной фильтрации сложилось мнение, что это не только полностью изученная область, но и к тому же мало перспективная, ибо нелинейная фильтрация всегда лучше линейной. В 1952 г. появилась статья Костаса [10], в которой он впервые развил представление о линейном кодировании применительно к схеме Шеннона (см. рис. В.1), в которой функции кодера и декодера выполняют линейные фильтры. Сразу возник вопрос о сравнении линейного и нелинейного способов кодирования. Ответ на этот вопрос базировался на утверждении, что нелинейное кодирование во всех случаях лучше линейного. Поэтому оставалось выяснить, насколько ошибка при оптимальном линейном кодировании Костаса больше ошибки при нелинейном кодировании Шеннона. Однако такая концепция ошибочна. На самом деле, хотя Шеннон дает *универсальный* нелинейный способ кодирования блоками — наилучший среди других способов при любых законах распределения вероятностей и любых типов каналов (дискретных и непрерывных), это не исключает того, что в некоторых частных случаях не существует другого способа кодирования, лучшего, чем способ Шеннона.

В модели на рис. В.1 учитывается только один шум канала, а сигнал, поступающий в кодер, считается бесшумным. Однако на практике приходится кодировать зашумленные сигналы, что отражено в новой модели на рис. В.3. Можно ли свести новую модель (см. рис. В.3) к старой модели (см. рис. В.1)? Поиск ответа на этот вопрос привел к введению понятия о шумовой информации, вычисляемой как средняя взаимная информация ме-

жду шумовыми процессами. Источник шума не только уменьшает полезную информацию, но создает шумовую информацию, которая не позволяет свести новую модель к старой.

Чтобы не «утонуть» в многообразии систем видеоинформатики при переходе от теоретической части книги к практической и не просто перечислить типы разработанных систем, для описания в гл. 3—5 выбран только один класс систем, наиболее полно отражающий специфику нового направления. С нашей точки зрения таким классом систем являются телевизионные роботы — продукт совместного использования телевидения, вычислительной техники и автоматизации. Информационно-решающая система телевизионного робота автоматически перерабатывает информацию, извлекаемую из электромагнитного поля, в решения, которые управляют поведением робота в реальном времени. Это означает, что кодирование и принятие решения должны осуществляться с частотой телевизионных кадров. Телевизионные роботы делятся на три подкласса:

1) роботы с кадровой частотой малокадрового телевидения;
2) роботы с кадровой частотой вещательного телевидения, равной 25 Гц;
3) роботы с кадровой частотой выше вещательной.

С каждым телевизионным кадром телекамера обрушивает поток информации в миллионы двоичных единиц на видеокомпьютер, переработка которой в реальном времени требует от вычислительной техники наибольшего быстродействия и наибольшей емкости оперативной памяти. Поэтому увеличение кадровой частоты телевизионных роботов представляет сложнейшую научную и технологическую проблемы. Научная проблема — необходимость способов кодирования и алгоритмов принятия решения, наиболее экономных с точки зрения теории информации и алгоритмов. Технологическая проблема — необходимость соответствующей элементной базы.

В цифровом вещательном телевидении решается проблема экономного кодирования после телекамеры для доставки цветного изображения на экраны телезрителей. При этом телекамеру рассматривают как заданный источник информации. В отличие от этого традиционного подхода в гл. 3 изложено аналоговое и цифровое кодирование, под которым понимается преобразование «от света до цифры» в телекамере на ПЗС. Кодирование в телекамере должно сжать кодовую информацию до значения, приемлемого для видеокомпьютера. В [2] был описан первый этап развития твердотельного компьютерного телевидения, когда основной интерес вызывали малые массы и габаритные размеры телекамер и лучшая способность к измерению оптических изображений (линейная световая характеристика ПЗС, жесткий растр и малые собственные шумы). На первом этапе решалась проблема улучшения технологии изготовления ПЗС, так как формируемые ими телевизионные изображения из-за дефектов и малой четкости не соответствовали требованиям вещательного телевидения. В настоящее время эта проблема успешно решена, и развитие вещательного телевидения высокой четкости ориентируется на использование ПЗС. На втором этапе развития твердотельного компьютерного телевидения интерес вызывают новые функциональные возможности ПЗС как кодеров. Эти возможности показаны в гл. 3 на примере существующих матриц ПЗС. Более глубокое их раскрытие требует разработки новых специализированных ПЗС.

Цифровой сигнал с выхода телекамеры на ПЗС, содержащий полезную и шумовую информации в сочетании с информацией о фоне, поступает в видеокомпьютер, реализующий алгоритм принятия решения. Ограничение массы и габаритных размеров телевизионных роботов не позволяет применять универсальные ЭВМ и диктует обращение к мини-, а лучше микрокопьютерам. Поэтому при высокой частоте телевизионных кадров вряд ли целесообразно в настоящее время ставить перед видеокомпьютером столь сложные задачи, как опознавание семантических образов. В гл. 4 описано устройство «от цифры до решения», реализующее не слишком сложный алгоритм классификации изображений по простым формализуемым признакам. Формализация задачи переработки информации и использование простых моделей сигналов позволяют оптимизировать устройства на основе простых мер качества — среднеквадратической ошибки и вероятности ошибки. Тем не менее, несмотря на простоту используемых моделей, сложность переработки информации в робототехнических системах оказывается высокой. Она связана со сложностью фоновой обстановки, необходимостью учета подвижности изображений и обеспечением стабильности работы классификаторов изображений в условиях сдвигов, поворотов и масштабирования изображений. Высокая сложность видеокомпьютера вызвана необходимостью использования многокадровой памяти и быстродействующих микропроцессоров, реализующих операции над матрицами изображений (обращение, умножение и т. п.).

В телевизионных роботах, которые должны превзойти человека по быстродействию и которые осуществляют переработку информации в решение и поведение, при едином информационном подходе согласовывают быстродействие и сложность информационно-решающей, управляющей и механической подсистем.

Глава 1.

ЭЛЕМЕНТЫ ТЕОРИИ ФИЛЬТРАЦИИ И ТЕОРИИ ИНФОРМАЦИИ

Теория фильтрации и теория информации обычно излагаются независимо. Фильтрацию относят к оптимальному приему сигналов, тогда как теория Шеннона — это «теория каналов» и, следовательно, относится к передаче сигналов. Ниже предпринята попытка изложить некоторые основные результаты теории линейной фильтрации и теории информации с позиции неразрывности передачи и приема. Математически единство изложения обусловлено спектральным методом описания случайных процессов и фильтров с применением разложения Карунена — Лоэва. Ключевым вопросом является обобщенная теорема кодирования, изложенная Шенноном в работе 1959 г. [3, с. 587]. В качестве частного случая изложен способ Костаса оптимального кодирования с помощью линейных фильтров. В заключение сформулированы две теоретические задачи развития теории информации.

1.1. СПЕКТРАЛЬНЫЙ МЕТОД ОПИСАНИЯ СИГНАЛОВ И ФИЛЬТРОВ

Обратимся к схеме на рис. В.3 с двумя шумами. Шумы могут смешиваться с сигналом сложным образом. Для упрощения описания преобразователей в канале обычно делают допущение об аддитивности сигнала и шума. Сам канал математически моделируют сумматором с шумом. Аналогичное допущение мы сделаем и для шума источника. При этом модель системы упрощается (рис. 1.1). Схема содержит два сумматора с шумом. Источник и канал в схеме на рис. 1.1 могут быть дискретными и непрерывными. Схема с дискретным источником и дискретным ка-

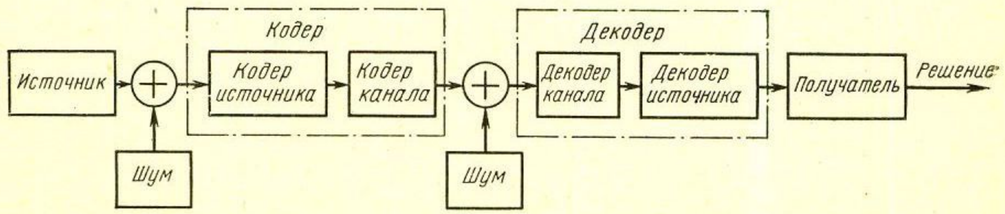


Рис. 1.1. Модель системы связи с аддитивными шумами

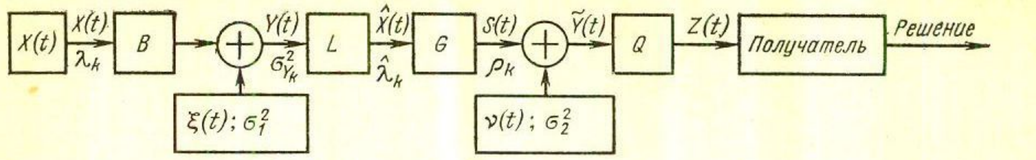


Рис. 1.2. Модель системы связи с линейным кодированием

налом является классической в теории информации и наиболее подробно изучена. Поэтому сосредоточим внимание на варианте модели системы с непрерывным источником и непрерывным каналом. В этом случае можно применить кодер и декодер наиболее простой конструкции — линейные фильтры. Схема с линейным кодированием представлена на рис. 1.2. В схеме имеются источники двух случайных процессов: полезного сигнального процесса $X(t)$ и мешающего шумового процесса $\xi(t)$. Процесс трактуется как ансамбль или класс индивидуальных функций (реализаций) с заданными вероятностями распределения. Для описания случайных процессов с непрерывным аргументом широко используется их разложение в обобщенный ряд Фурье с ортонормированным базисом Карунена — Лоэва. Метод разложения Карунена — Лоэва подробно излагается во многих учебниках (см., например, [11]). Приведем кратко основные формулы в применении к сигнальному случайному процессу $X(t)$. Допустим, что процесс имеет нулевое математическое ожидание и известную корреляционную функцию $K(t, \tau)$. Рассмотрим интегральное уравнение

$$\lambda_k \varphi_k(t) = \int_{-T/2}^{T/2} K(t, \tau)\, \varphi_k(\tau)\, d\tau.$$

Из уравнения находим собственные функции $\varphi_k(t)$ и собственные значения λ_k. По методу Карунена — Лоэва собственные функции используются как ортонормированный базис для представления случайного процесса на конечном интервале времени $[-T/2, T/2]$ обобщенным рядом Фурье с независимыми отсчетами

$$X(t) = \sum_{k=1}^{\infty} X_k \varphi_k(t); \qquad (1.1)$$

$$X_k = \int_{-T/2}^{T/2} X(t)\, \varphi_k(t)\, dt.$$

При описании сигналов мы пользовались обозначениями из учебного пособия [11]. Прописными буквами X_k обозначены непрерывные случайные величины — коэффициенты ряда Фурье. При этом непрерывная действительная случайная величина X_k обозначается так же, как и ансамбль числовых значений (числовое множество) [11]. Ряд Фурье (1.1) допускает геометрическую трактовку процесса как бесконечномерного случайного вектора в пространстве с ортами $\varphi_k(t)$:

$$\mathbf{X}=(X_1, X_2, \ldots).$$

Дисперсии случайных величин X_k равны собственным значениям интегрального уравнения:

$$\lambda_k=\mathbf{M}(X^2_k);\ \mathbf{M}(X_k)=0,$$

где $\mathbf{M}$ — обозначение операции усреднения, получения математического ожидания.

Информационные сигналы, моделируемые функциями $X(t)$, в отличие от их физических переносчиков не имеют размерности. Поэтому сигнал можно характеризовать конечной «безразмерной энергией»:

$$E_X=\int_{-T/2}^{T/2} X^2(t)\,dt=\sum_{k=1}^{\infty}\lambda_k.$$

Дисперсии λ_k коэффициентов разложения характеризуют распределение безразмерной энергии сигнала по k-м степеням свободы. Поэтому совокупность величин λ_k описывает энергетический спектр сигнала, представленного обобщенным рядом Фурье (1.1). Ниже мы будем рассматривать полезные сигналы $X(t)$, имеющие монотонно убывающие спектры:

$$\lambda_1\geqslant\lambda_2\geqslant\lambda_3\geqslant\ldots$$

Базис Карунена — Лоэва отличается от гармонического базиса классического ряда Фурье, который более привычен инженерам в радиотехнике и оптике. Различие заключается в том, что гармонический базис одинаков для разных разлагаемых сигналов, в то время как базис Карунена — Лоэва адаптивен к корреляционной функции сигнала. Зависимость ортонормированных функций $\varphi_k(t)$ от вида корреляционной функции создает, конечно, повышенные трудности для аппаратурной реализации. Но это та цена, которую необходимо платить за получение большого удобства — взаимной статистической независимости коэффициентов разложения Карунена — Лоэва. Благодаря этому источник сигнала $X(t)$ можно рассматривать как произведение источников, характеризуемое сложением энтропии «одномерных» источников:

$$H[X(t)]=\sum_{k=1}^{\infty}H(X_k).$$

Указывая определенную функцию — индивидуальную реализацию $x(t)$ сигнала $X(t)$, мы «снимаем» эту энтропию (меру неопределенности), сообщая равное ей количество информации. В силу статистической независимости коэффициентов разложения Карунена — Лоэва особенно просто описываются гауссовские случайные процессы. Совместная плотность вероятностей коэффициентов $X_1, \ldots, X_N$ ряда (1.1) описывается простым выражением [11]:

$$p(x_1, \ldots, x_N)=\prod_{k=1}^{N}\frac{1}{\sqrt{2\pi\lambda_k}}\exp\left(\frac{x_k^2}{2\lambda_k}\right),\quad N=1, 2, \ldots$$

Моделирование сигналов и шумов гауссовскими процессами важно потому, что только в этом случае удастся получить явные формулы для информационных характеристик. Хотя случайный сигнальный процесс может быть представлен геометрически бесконечномерным вектором, решение вариационных задач в теории фильтрации и теории информации всегда оказывается вырожденным. Поэтому получателю доступно лишь N-мерное приближение сигнала $X(t)$ с ошибкой, не меньшей

$$\Delta\varepsilon=\sum_{k=N+1}^{\infty}\lambda_k.$$

Для упрощения анализа сделаем допущение о том, что шумовые случайные процессы в схеме на рис. 1.2 являются белыми шумами. Такое допущение позволяет применить сигнальный базис $\{\varphi_k(t)\}$ также и к шумовым процессам, не нарушив статистической независимости коэффициентов:

$$\xi(t)=\sum_{k=1}^{\infty}\xi_k\varphi_k(t);$$
$$\nu(t)=\sum_{k=1}^{\infty}\nu_k\varphi_k(t). \qquad (1.2)$$

Коэффициенты разложения белых шумов при любом ортогональном базисе взаимно статистически независимы и имеют постоянные дисперсии

$$\sigma^2{}_1=\mathbf{M}(\xi^2{}_k);\ \sigma^2{}_2=\mathbf{M}(\nu^2{}_k).$$

Фильтры, используемые в схеме на рис. 1.2, также опишем с помощью базиса Карунена — Лоэва для полезного сигнала. Линейные фильтры описываются интегральными операторами [11]. По теореме Фредгольма для описания ядра оператора Гильберта — Шмидта можно использовать полный базис $\{\varphi_k(t)\}$. Ниже будут рассмотрены линейные фильтры, задаваемые совокупностями спектральных элементов, которые для удобства мы будем обозначать прописными буквами B, R, L, G, Q. Ядра линейных преобразований представим в виде:

$$B(t, \tau)=\sum_{k=1}^{\infty}\beta_k\varphi_k(t)\varphi_k(\tau);$$
$$R(t, \tau)=\sum_{k=1}^{\infty}r_k\varphi_k(t)\varphi_k(\tau);$$
$$L(t, \tau)=\sum_{k=1}^{\infty}l_k\varphi_k(t)\varphi_k(\tau); \qquad (1.3)$$
$$G(t, \tau)=\sum_{k=1}^{\infty}g_k\varphi_k(t)\varphi_k(\tau);$$
$$Q(t, \tau)=\sum_{k=1}^{\infty}q_k\varphi_k(t)\varphi_k(\tau).$$

При прохождении сигнала представленного разложения Карунена — Лоэва (1.1) через любой из фильтров (1.3) сигнал на выходе имеет коэффициенты разложения, которые равны произведению спектральных коэффициентов сигнала и ядра. В силу линейности схемы на рис. 1.2 суммирование процессов заключается в суммировании коэффициентов. Источник выдает два случайных процесса: сигнал и шум. Выходной сигнал источника в схеме на рис. 1.2 на основании вышеизложенного имеет простое представление:

$$Y(t)=BX(t)+\xi(t)=\sum_{k=1}^{\infty} Y_k\varphi_k(t)=\sum_{k=1}^{\infty}(\beta_k X_k+\xi_k)\varphi_k(t).$$

Непрерывный источник выдает на вход кодера случайный процесс $Y(t)$, являющийся бесконечномерным случайным вектором

$$\mathbf{Y}=(Y_1, Y_2, \ldots).$$

Кодированию доступен только сигнал в *смеси с шумом,* т. е. искаженный неоднородным, необратимым линейным преобразованием:

$$Y_k=\beta_k X_k+\xi_k;\ k\in[1, \infty];$$
$$Y_k=\int_{-T/2}^{T/2}[BX(t)+\xi(t)]\varphi_k(t)\,dt. \qquad (1.4)$$

Разложения полезного сигнала (1.1), шумов (1.2) и передаточных характеристик фильтров (1.3) позволяют весьма просто описать преобразования сигнала при линейном кодировании. Сигнал, поступающий после линейного кодера в канал, равен

$$S(t)=RY(t)=\sum_{k=1}^{\infty} S_k\varphi_k(t),$$

где $S_k=r_kY_k=r_k(\beta_kX_k+\xi_k)$; $r_k=l_kg_k$.

В канале сигнал $S(t)$ суммируется с шумом:

$$\tilde{Y}(t)=S(t)+\nu(t)=\sum_{k=1}^{\infty}\tilde{Y}_k\varphi_k(t),$$

где $\tilde{Y}_k=S_k+\nu_k$.

На выходе линейного декодирующего фильтра формируется сигнал

$$Z(t)=\sum_{k=1}^{\infty} Z_k\varphi_k(t), \qquad (1.5)$$

где $Z_k=q_k(r_kY_k+\nu_k)$.

Линейный код (1.5) строится по зашумленному сигналу источника с помощью неоднородного, необратимого линейного преобразования:

$$Z_k=q_kg_kl_kY_k+q_k\nu_k;\ k\in[1, \infty]. \qquad (1.6)$$

Два неоднородных линейных преобразования (1.4) и (1.6) полностью описывают линейное кодирование в схеме на рис. 1.2.

Теперь видно, что разложение Карунена — Лоэва благодаря независимости коэффициентов позволило свести построение линейного кода к множеству параллельных неоднородных линейных преобразований (1.6).

Сигнал (1.5) с выхода декодера поступает получателю-автомату, который выносит то или иное решение, например по классификации полезных сигналов. Достоверность принятия решения будет зависеть от степени приближения выходного сигнала $Z(t)$ к полезному сигналу $X(t)$. Критерием качества восстановления служит среднеквадратическая ошибка

$$\varepsilon=\mathbf{M}[X(t)-Z(t)]^2=\sum_{k=1}^{\infty}\mathbf{M}(X_k-Z_k)^2.$$

В гл. 1 и 2 мы будем пользоваться только среднеквадратическим критерием качества точности восстановления полезного сигнала.

Естественно возникает вариационная задача минимизации среднеквадратической ошибки с помощью выбора фильтров. Такая задача решается в теории линейной фильтрации.

1.2. ВАРИАЦИОННАЯ ЗАДАЧА В ТЕОРИИ ЛИНЕЙНОЙ ФИЛЬТРАЦИИ

А. Н. Колмогоровым была создана спектральная теория стационарных процессов, их фильтрации и экстраполяции [4, с. 255]. В рамках этой теории Н. Винером был найден оптимальный фильтр, обеспечивающий выделение полезного случайного сигнала на фоне шума с минимальным значением среднеквадратической ошибки. Этот тип оптимального фильтра получил название фильтра Винера. Его импульсная характеристика находилась решением уравнения Винера — Хопфа. Упрощенный спектральный метод нахождения оптимального фильтра с помощью вариационного исчисления был предложен в работе Боде и Шеннона [3, с. 687]. В теории фильтрации используется частный вид схемы, представленной на рис. 1.3, когда кодирующий фильтр G отсутствует ($G=1$) и на вход варьируемого декодирующего фильтра поступает зашумленный сигнал $X(t)+\nu(t)$. Оптимальный фильтр Винера, минимизирующий среднеквадратическую ошибку, является декодером:

$$q_{k\,\text{опт}}=\frac{\lambda_k}{\lambda_k+\sigma_2^2};\quad k\in[1, \infty].$$

Этот результат легко обобщить и на случай, когда фильтр G

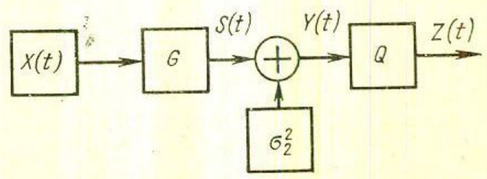


Рис. 1.3. Модель системы связи с линейным кодированием по Костасу

присутствует в схеме на рис. 1.3, но считается заданным, априори известным. Спектральная характеристика фильтра Винера имеет вид [14]:

$$q_{k\,\text{опт}} = \frac{g_k \lambda_k}{g_k^2 \lambda_k + \sigma_2^2}; \quad k \in [1, \infty].$$

Фильтр Винера формирует в схеме на рис. 1.3 по зашумленному сигналу $GX(t)+\nu(t)$ оптимальную оценку $\hat{X}(t)$ полезного сигнала $X(t)$, т. е. оценку, при которой достигается минимальная ошибка воспроизведения:

$$\varepsilon\left(g_k^2\right) = \mathbf{M}\left[X(t) - \hat{X}(t)\right]^2 = \sum_{k=1}^{\infty} \frac{\lambda_k \sigma_2^2}{g_k^2 \lambda_k + \sigma_2^2}.$$

Костас в статье 1952 г. [10], подробно изложенной в [14], воспользовался схемой на рис. 1.3, но дал новую постановку вариационной задачи: оба фильтра варьировались в целях минимизации среднеквадратической ошибки. Если фильтр G в схеме на рис. 1.3 задан, то сумматор с шумом нельзя считать каналом. Такая трактовка допускается при оптимальном приеме, когда варьируется фильтр Q. Заслуга Костаса заключается в том, что он первый увидел в схеме на рис. 1.3 линейный вариант схемы Шеннона (см. рис. 1.1), в которой функции кодера и декодера выполняют кодирующий G и декодирующий Q линейные фильтры. Поэтому статья Костаса [10] явилась пионерской работой по линейному кодированию. Для схемы на рис. 1.3 запишем спектральное выражение для вычисления ошибки:

$$\varepsilon = \sum_{k=1}^{\infty} \mathbf{M}(X_k - Z_k)^2 = \sum_{k=1}^{\infty} [\lambda_k - 2q_k g_k \lambda_k + q^2{}_k (g^2{}_k \lambda_k + \sigma^2{}_2)]. \qquad (1.7)$$

Варьируя коэффициенты фильтров g_k и q_k, необходимо найти минимум ошибки (1.7). Решение этой вариационной задачи можно упростить, преобразовав выражение (1.7) к виду, содержащему полный квадрат относительно коэффициентов q_k декодирующего фильтра [15]:

$$\varepsilon = \sum_{k=1}^{\infty} \left[\lambda_k + \left(q_k - \frac{g_k \lambda_k}{g_k^2 \lambda_k + \sigma_2^2} \right)^2 \left(g_k^2 \lambda_k + \sigma_2^2 \right) - \frac{g_k^2 \lambda_k^2}{g_k^2 \lambda_k + \sigma_2^2} \right]. \qquad (1.8)$$

Вариации коэффициентов q_k могут влиять только на второй член суммы и сделать его равным нулю для минимизации ошибки (1.8):

$$\left(q_k - \frac{g_k \lambda_k}{g_k^2 \lambda_k + \sigma_2^2} \right)^2 = 0.$$

Отсюда следует вывод, что оптимальный декодирующий фильтр будет фильтром Винера:

$$q_{k\,\text{опт}} = \frac{g_k \lambda_k}{g_k^2 \lambda_k + \sigma_2^2}. \qquad (1.9)$$

Вывод об оптимальности фильтра Винера в качестве декодирующего фильтра сохраняет свою справедливость при любом типе кодирующего фильтра в схеме на рис. 1.3. Декодер — фильтр Винера оптимален и тогда, когда кодирующий фильтр задан, и тогда, когда ищется оптимальный кодирующий фильтр для вариационной задачи. Сделанный вывод позволяет упростить решение вариационной задачи Костаса, сведя ее к нахождению только кодирующего фильтра, который обеспечивает минимум ошибки (1.8) при декодирующем фильтре Винера (1.9):

$$\varepsilon\left(q_k^2\right) = \sum_{k=1}^{\infty} \left[\lambda_k - \frac{g_k^2 \lambda_k}{g_k^2 \lambda_k + \sigma_2^2} \right]. \qquad (1.10)$$

Задача оптимизации кодирующего фильтра имеет смысл только при наличии ограничения того или иного типа. В отсутствие ограничения тривиальным решением является фильтр с бесконечным усилением $g^2{}_k = \infty$, сводящий задачу к бесшумовому случаю. При выборе типа ограничения Костас опять же исходит из аналогии со схемой Шеннона. Безразмерная энергия сигнала $S(t)$, поступающего на вход сумматора с шумом (модели канала), равна

$$\int_{-T/2}^{T/2} S^2(t)\, dt = \sum_{k=1}^{\infty} \mathbf{M}\left(S_k^2\right) = \sum_{k=1}^{\infty} g_k^2 \lambda_k.$$

Шеннон вывел свою знаменитую формулу (см. (1.18)) для пропускной способности гауссовского канала при ограничении «энергии» (или «мощности») сигнала на входе канала [3, с. 454]. Костас использовал аналогичное ограничение [10, 14]:

$$\sum_{k=1}^{\infty} g^2{}_k \lambda_k \leqslant E_s. \qquad (1.11)$$

Варьируя коэффициенты g_k кодирующего фильтра при ограничении «энергии» (1.11), найдем минимум ошибки Винера (1.10):

$$\min_{\{g_k\}} \varepsilon\left(g_k^2\right) = \min_{\{g_k\}} \mathbf{M}\left[X(t) - \hat{X}(t)\right]^2 = \varepsilon_s.$$

Обратим внимание на то, что в класс кодирующих фильтров, среди которых ищется оптимальный, входит и неискажающий фильтр с бесконечным числом степеней свободы:

$$g_k = 1; \quad k \in [1, \infty].$$

Это замечание справедливо потому, что неискажающий кодирующий фильтр удовлетворяет ограничению (1.11). В качестве примера кодирующего фильтра, не удовлетворяющего условию (1.11), можно привести выбеливающий фильтр с бесконечным числом степеней свободы:

$$g_k^2 = \frac{1}{\lambda_k}; \quad k \in [1, \infty].$$

Следуя методу Лагранжа, из вариационного исчисления построим функцию Лагранжа:

$$\Phi(g^2_k)=\varepsilon(g^2_k)+\alpha_1\sum_{k=1}^{\infty}g^2_k\lambda_k,$$

где α_1 — постоянная Лагранжа.

Уравнение Эйлера имеет вид:

$$\frac{\partial\Phi(g_k^2)}{\partial g_k^2}=0.$$

Решение уравнения Эйлера дает оптимальный кодирующий фильтр — фильтр Костаса:

$$g^2_{k\,\text{опт}}=\begin{cases}\sqrt{\dfrac{\sigma_2^2}{\alpha_1\lambda_k}}-\dfrac{\sigma_2^2}{\lambda_k} & \text{при } \lambda_k>\alpha_1\sigma_2^2,\quad k<N.\\ 0 & \text{при } \lambda_k\leqslant\alpha_1\sigma_2^2,\quad k\geqslant N.\end{cases}\qquad(1.12)$$

Видно, что решение (1.12) вариационной задачи Костаса оказалось вырожденным. Его фильтр является конечномерным, т. е. имеет конечное число степеней свободы, определяемое условием

$\lambda_k=\alpha_1\sigma^2{}_2$ при $k=N$.

Сравним в схеме на рис. 1.3 кодек — фильтр Костаса (1.12) с фильтром Винера (1.9) и кодек — неискажающий кодирующий фильтр ($g_k=1$) со своим декодирующим фильтром Винера. Несмотря на то, что первая пара фильтров имеет конечное число степеней свободы, а вторая — бесконечное, среднеквадратическая ошибка при ограничении (1.11) в первом случае будет меньше:

$$\sum_{k=1}^{N}\frac{\lambda_k\sigma_2^2}{g^2_{k\,\text{опт}}\lambda_k+\sigma_2^2}=\varepsilon_s<\sum_{k=1}^{\infty}\frac{\lambda_k\sigma_2^2}{\lambda_k+\sigma_2^2}.$$

Хотя вариационная задача Костаса и использовала схему и ограничения из теории Шеннона, она, по существу, оставалась в рамках теории фильтрации, опираясь на критерий среднеквадратической ошибки и не используя понятия энтропии источника и пропускной способности канала. Заметим, что источник в схеме Костаса (см. рис. 1.3) был непрерывным, а теорема Шеннона, охватывающая кодирование непрерывных источников [3, с. 587], была доказана лишь спустя 7 лет после выхода статьи Костаса [10]. Обратимся к этой важнейшей работе Шеннона.

1.3. ПРИНЦИП БЛОКОВОГО КОДИРОВАНИЯ

Теория Шеннона использует модель систем передачи информации, представленную схемой на рис. 1.1. Согласно теории Шеннона приспособленность канала или запоминающего устройства к передаче и хранению информации с достаточной полнотой характеризуется единственной величиной C — пропускной способ-

ностью. Приспособленность информации, создаваемой источником к передаче или хранению, тоже характеризуется единственной величиной H — энтропией источника [4, с. 38]. Для схемы с дискретным источником без шума и с любым каналом, работающим с шумом, Шеннон сформулировал в 1948 г. универсальную теорему, называемую основной теоремой кодирования [3, с. 243]. Этим была создана основа современной теории информации как дисциплины, способной к систематическому развитию [4, с. 36]. В основе теоремы лежит идея Шеннона о блоковом (нелинейном) кодировании как универсальном способе передачи информации. Принцип блокового кодирования явился крупным открытием. Как во всяком крупном открытии, в нем имелись элементы неожиданности. Неожиданной оказалась возможность передачи информации от нешумящего дискретного источника по зашумленному каналу при конечной пропускной способности с «нулевой» ошибкой (т. е. со сколь угодно малой вероятностью ошибки) [4, с. 37]. Иначе говоря, неожиданной оказалась столь большая эффективность блокового кодирования в смысле устранения ошибки при передаче информации по зашумленному каналу. Такая передача была возможна при условии согласования источника с каналом:

$$H\leqslant C.$$

Это условие согласования всегда нарушается для непрерывных источников из-за их бесконечной энтропии. Кроме того, и дискретный источник может быть не согласован с каналом из-за избытка выдаваемой информации. Теорема Шеннона 1948 г. не давала ответа на вопрос о том, как осуществлять передачу информации в таких случаях рассогласования источника и канала. Без ответа на этот вопрос нельзя было распространить теорию информации на системы с непрерывным источником. Для решения этой актуальной проблемы Шеннон еще в работе 1948 г. ввел понятие ε-энтропии источника [3, с. 318]. Термин «ε-энтропия» был предложен А. Н. Колмогоровым в работе 1956 г., чтобы подчеркнуть основную роль этого понятия в развитии теории информации [4, с. 57]. В этой же работе была дана формула для вычисления ε-энтропии источника гауссовского процесса со спектром Карунена — Лоэва (1.1) [4]:

$$H(\varepsilon)=\frac{1}{2}\sum_{k=1}^{N}\log\frac{\lambda_k}{\mu}\quad\text{при}\quad\lambda_k>\mu;\quad\lambda_k=\mu\quad\text{при}\quad k=N,\qquad(1.13)$$

где μ — выбеленная часть спектра заданной ошибки ε.

С увеличением задаваемой ошибки ε-энтропия источника монотонно падает, уменьшаясь относительно энтропии источника:

$$0\leqslant H(\varepsilon)\leqslant H;\quad\lim_{\varepsilon\to 0}H(\varepsilon)=H.$$

Решение указанной проблемы было дано Шенноном в работе 1959 г. [3, с. 587]. В ней идея универсального блокового кодирования получила дальнейшее развитие с применением ε-энтро-

пии источника. Шеннон обобщил основную теорему кодирования 1948 г., включив в нее кодирование для источника. Обобщенная теорема 1959 г. строится Шенноном как бы из двух теорем: новой теоремы о кодировании для источника, реализующем ε-энтропию, и «старой» теоремы 1948 г. о кодировании для канала [3, с. 607]. Поэтому кодер делится на два звена: кодер источника и кодер канала, как показано на рис. 1.1. Аналогично разбивается и декодер. Обобщенная теорема Шеннона охватывает оптимальное кодирование источника и канала. Принцип кодирования источника и канала по Шеннону — кодирование блоками (группами).

Незашумленный дискретный или непрерывный источник создает обычно такое количество полезной сигнальной информации, которое не удается передать по каналу. Функцией кодера источника является создание блоков из букв (знаков, символов) источника в целях выделения из информации, создаваемой источником, лишь минимального количества кодовой информации, равной ε-энтропии, которая достаточна для обеспечения требуемого качества (точности). Критерием требуемого качества служит задаваемая априори ошибка передачи и восстановления сигнала. Блок знаков или символов источника формируется из n_1 букв. Число n_1 характеризует длину блока источника. Общее число блоков, формируемых кодером Шеннона для источника, определяется длиной блоков n_1 и ε-энтропией источника:

$$N_и = 2^{n_1 H(\varepsilon)}.$$

Здесь ε-энтропия, оцениваемая средним числом двоичных единиц $H(\varepsilon)$ на букву источника, является функцией ошибки ε, которая не обязана быть среднеквадратической. Число блоков $N_и$ источника длиной n_1 букв достаточно для представления любого сигнала источника с искажением ε. Блок канала формируется из n_2 букв. Алфавит канала может отличаться от алфавита букв источника. При пропускной способности канала, равной C, через канал можно передать число различных блоков

$$N_C = 2^{n_2 C},$$

где C — среднее число двоичных единиц на букву в канале. Между блоками источника и канала должно быть соответствие, устанавливаемое механизмом перекодировки блоков. Когда число блоков $N_и$ источника равно числу блоков N_C канала, выполняется равенство ε-энтропии источника и пропускной способности канала, измеряемых в двоичных единицах на букву источника:

$$n_1 H(\varepsilon) = n_2 C$$

или

$$H(\varepsilon) = C n_2 / n_1.$$

Шеннон применил приведенное равенство для определения *идеального кодирования,* при котором достигается минимальная ошибка ε*, являющаяся решением уравнения [3, с. 597]:

$$H(\varepsilon^*) = C n_2 / n_1. \qquad (1.14)$$

Уравнение идеального кодирования (1.14) в общем случае не имеет решения в явной форме. Но это решение всегда можно найти расчетным путем с помощью функции, обратной ε-энтропии. В символическом виде решение уравнения идеального кодирования можно записать следующим образом:

$$\varepsilon^* = H^{-1}(C n_2 / n_1). \qquad (1.15)$$

Обратим внимание на то, что ошибка передачи ε* согласно выражению (1.15) определяется не только пропускной способностью C канала, но и отношением n_2/n_1 длин блоков источника и канала. При ограниченной пропускной способности достижение ошибки, близкой к нулю, возможно при идеальном кодировании за счет большого увеличения отношения n_2/n_1, приводящего к увеличению задержки передачи. Только в этом смысле можно говорить о так называемой «нулевой» ошибке передачи информации по каналу в теории Шеннона.

Не существует никакого способа кодирования, который смог бы обеспечить ошибку, меньшую значения (1.15). Важное значение уравнения идеального кодирования (1.14) заключается в том, что оно определяет границу возможного и невозможного в восстановлении незашумленного сигнала, переданного по каналу. Оптимальный блоковый код не является, строго говоря, идеальным кодом, удовлетворяющим условию (1.14). Но оптимальный блоковый код приближается к идеальному весьма близко [3, с. 600]. Доказательство существования близкого к идеальному блокового кодирования и составляет содержание обобщенной теоремы Шеннона [3, с. 587]. Только принцип блокового кодирования является *универсальным средством,* позволяющим приблизиться к идеальному кодированию при любых законах распределения вероятностей и для любых, дискретных или непрерывных, источников и каналов. Для существования оптимального блокового кодирования, обеспечивающего передачу без систематически возрастающего запаздывания, должно выполняться условие

$$H(\varepsilon) \leqslant C n_2 / n_1. \qquad (1.16)$$

Сближение оптимального кода Шеннона с идеальным требует, чтобы длина блоков источника и канала неограниченно увеличивалась при сохранении выбранного отношения n_2/n_1. Увеличение длины блоков сопровождается увеличением «блоковой» задержки. Из условия (1.16) видно, что обобщенная теорема Шеннона основывается на понятии не энтропии источника, а ε-энтропии, которая конечна не только для дискретных, но и для непрерывных источников. Тем самым обобщенная теорема впервые открыла путь применения теории информации к системам с непрерывными источниками.

В частном случае, когда ε-энтропия дискретного источника достигает значения энтропии ($\varepsilon=0$), условие (1.16) совпадает с условием, содержащимся в работе Шеннона 1948 г.:

$$H \leqslant Cn_2/n_1.$$

Число n_2 в неравенстве (1.16) можно трактовать как число использований канала, необходимых для передачи сигнала источника, состоящего из n_1 букв, каждая длительностью τ_s секунд:

$$n_2 = n_1\tau_s/\tau_C,$$

где τ_C — время использования канала на передачу одной буквы канала. Подставив значение n_2 в условие (1.16), получим оценку скорости передачи букв источника:

$$v_s \leqslant \frac{1}{\tau_C}\frac{C}{H(\varepsilon)},$$

где $v_s = \frac{1}{\tau_s}$ — число букв источника в секунду. При допущении Шеннона [3, с. 597] $\tau_C = 1$ скорость передачи букв источника лимитируется сверху отношением пропускной способности к ε-энтропии:

$$v_s \leqslant C/H(\varepsilon).$$

В обобщенной теореме Шеннона получило дальнейшее развитие установление связи между двумя понятиями: информации и качества (точности) воспроизведения сигнала. Начало было положено введением понятия об ε-энтропии и получением формулы (1.13) для ее вычисления в частном, гауссовском случае. Теперь показано, что для идеального кодирования существует минимальная ошибка передачи (1.15), которая лимитируется пропускной способностью канала.

1.4. НЕКОТОРЫЕ ЗАДАЧИ РАЗВИТИЯ ТЕОРИИ ИНФОРМАЦИИ

Обобщенная теорема Шеннона 1959 г., распространив принцип блокового кодирования с учетом точности воспроизведения на любые, дискретные и непрерывные, источники, улучшила перспективы практического применения теории информации. Все же и после этого теория информации продолжала оставаться сферой интересов математиков, занятых расширением ее математического фундамента [4, с. 254]. Инженеры и научные работники, проектирующие системы в телевидении, оптике, радиолокации, хотели найти в теоремах Шеннона описание конструкции блокового кодера и декодера, дающего хорошее приближение к идеальному коду. Не находя желаемых конструкций, инженеры «впали в разочарование», виня в этом теорию информации. Они забыли, что сила теоремы Шеннона в ее универсальности, в формулировке принципа блокового кодирования как универсального средства достижения идеального кодирования. Желанные же для инженеров конструкции кодеров всегда имеют частный характер, поскольку инженеру надо оптимально решить одну из множества частных задач с присущим только ей набором специфических начальных условий. Многообразию частных случаев кодирования соответствует многообразие частных подходов к идеальному кодированию. Причем в каждом частном случае может существовать способ кодирования, лучше приближающийся к идеальному, чем универсальный метод Шеннона. Теорема Шеннона не запрещает такую возможность. Его теорема — не справочник по выбору частных конструкций кодеров и декодеров, а «компас» для поиска необходимых конструкций. Это особенно наглядно можно проиллюстрировать на примере линейного кодирования, описанного в § 1.2. Уже отмечалось, что схема Костаса на рис. 1.3 является частным случаем схемы Шеннона на рис. 1.1, когда источник и канал непрерывны, а кодер и декодер — линейные фильтры. Отсюда вытекало желание оценить близость линейного кодирования к идеальному, определяемому условием (1.14). Как стало ясно после работы Шеннона 1959 г., для этого достаточно сравнить ошибку ε_s линейного кодирования Костаса с ошибкой идеального кодирования (1.15). Линейному кодированию присуща «единичная длина» блоков:

$$n_2 = n_1 = 1.$$

С учетом этого обстоятельства уравнение идеального линейного кодирования является частным видом уравнения (1.14):

$$H(\varepsilon^*) = C,\ n_2/n_1 = 1, \tag{1.17}$$

где $\varepsilon^* = H^{-1}(C)$ — решение уравнения.

Так как линейное кодирование Костаса не может быть лучше идеального, сразу очевидна справедливость следующих оценок:

$$\varepsilon_s \geqslant \varepsilon^*;\ H(\varepsilon_s) \leqslant C.$$

Сравнение линейного кодирования с идеальным целесообразно проводить для гауссовского случая (гауссовского закона распределения вероятностей), так как только для него имеются явные формулы. Гауссовский источник в схеме на рис. 1.3 имеет ε-энтропию, вычисляемую по формуле Колмогорова (1.13). Пропускная способность гауссовского канала в двоичных единицах в схеме на рис. 1.3 вычисляется при ограничении (1.11) «энергии» сигнала на его входе по формуле Шеннона [3, с. 453]:

$$C(E_s,\ N_s) = \frac{N_s}{2}\log\left(1+\frac{E_s}{N_s\sigma_2^2}\right); \tag{1.18}$$

$$C(E_s) = \lim_{N_s\to\infty}\frac{N_s}{2}\log\left(1+\frac{E_s}{N_s\sigma_2^2}\right) = \frac{E_s}{2\sigma_2^2}\log e,$$

где N_s — число степеней свободы сигнала $S(t)$ с энергией E_s на входе канала, которое Шеннон считал известным.

В работе Б. С. Цыбакова [16], посвященной передаче гауссовских сигналов по гауссовским каналам, подтверждено, что линейное кодирование Костаса дает ошибку ε_s, бóльшую, чем результат решения уравнения (1.17):

$$\varepsilon_s > \varepsilon^* \text{ или } H(\varepsilon_s) < C(E_s, N_s).$$

Этот результат [16], опубликованный в [17], совпадает с утверждением Галлагера о том, что в так называемом тест-канале (системе линейного кодирования) не достигается ошибка из уравнения (1.17) [18]. Поэтому авторы [16—18] получили тот результат, который ожидали: оптимальное линейное кодирование всегда дает бóльшую ошибку, чем нелинейное кодирование Шеннона, и подтвердили свою гипотезу. Такое утверждение не способствовало объединению теории фильтрации и теории информации, наметившемуся в работе Костаса [10]. Использование Костасом в схеме на рис. 1.3 двух варьируемых фильтров расширило рамки теории фильтрации, включив в нее вопросы и передачи и приема. Однако, как мы видим, этого оказалось еще недостаточно для объединения теории фильтрации и теории Шеннона. Актуальность этой проблемы подтверждается следующим моментом. Шеннон отмечал, что его теория имеет дело с преобразователями (кодер и декодер), в которых допускаются неограниченные временны́е задержки на формирование достаточно длинных блоков [3, с. 462]. Говоря о некоторых задачах теории информации, Шеннон считал, что теория преобразователей без задержек изучена хуже, хотя соответствующие задачи более важны для практики, и имеет приложение к проблеме фильтрации, оптимальной по квадратическому критерию [3].

Линейное кодирование является, по существу, *безблочным кодированием* ($n_1=1$; $n_2=1$) *и не требует временно́й задержки на формирование блоков.*

Поэтому такое кодирование можно осуществить в реальном времени, что крайне важно для практики. Актуальность этой практической задачи требует продолжить поиск для схемы на рис. 1.3 оптимального кодирующего фильтра, отличного от фильтра Костаса, который обеспечил бы ошибку, удовлетворяющую уравнению (1.17). Поиск такого идеального линейного кодирования был бы полезен инженерам и в методологическом аспекте. Во-первых, инженеры привыкли к широкому использованию линейных фильтров. Во-вторых, сам факт существования простой конструкции идеального кодирования, пусть даже и в частном случае, развеял бы страх инженеров перед математической сложностью книг по теории информации. В теоретическом же плане нахождение идеального линейного кодирования является не просто интересной задачей теории информации, но и способствовало бы снятию разобщенности между теорией информации и теорией фильтрации.

Перед теорией информации стояла и другая актуальная задача — учет влияния шума источника. Во введении уже отмечалась необходимость перехода от схемы на рис. В.1 к схеме на рис. В.3 в интересах более полного отражения систем, получающих информацию от электромагнитного поля. Необходимо было выяснить, что нового может дать включение шума источника в схему Шеннона. Кодер Шеннона одинаково перерабатывает и чистый сигнал $X(t)$ источника в схеме на рис. В.1 и зашумленный сигнал $Y(t)$ в схеме на рис. 1.1; одинаково в том смысле, что принцип блокового кодирования не меняется. Поэтому наличие шума источника не может поколебать правильность теории Шеннона. Речь может идти только о расширении сферы действия теоремы Шеннона и возникающей при таком расширении специфике. Первая попытка учета шума источника в статье Р. Л. Добрушина и Б. С. Цыбакова была предпринята «под флагом» сведения «новой» схемы на рис. 1.1 к «старой» схеме Шеннона (рис. В.1) [19]. Авторы показали, что, как следовало ожидать, обобщенная теорема Шеннона сохраняет свою силу и для схемы на рис. 1.1 с двумя шумами. Вся же новизна от учета дополнительного шума свелась авторами в гауссовском случае к замене чистого сигнала $X(t)$ на условное математическое ожидание:

$$\hat{X}(t) = \mathbf{M}\{X(t) \mid Y(t)\}.$$

Такая замена отразилась на формуле вычисления ε-энтропии источника гауссовского процесса в схеме на рис. 1.1. Этот результат был позднее использован Б. С. Цыбаковым для информационной оценки линейного кодирования Костаса с учетом входного шума [16].

Однако в работах [16, 19] не была вскрыта главная новизна, рождаемая наличием шума источника. Она заключается в том, что источник создает два случайных процесса (полезный сигнальный $X(t)$ и шумовой $\xi(t)$), каждый из которых характеризуется своей энергией. Поэтому источник уже нельзя, как ранее, задать единственным числом — ε-энтропией. Формализация этой новизны, ведущей к понятию о шумовой информации, и составляла вторую актуальную задачу в теории информации. Обе указанные задачи, как оказалось взаимосвязанные, рассмотрены в следующей главе.

Глава 2.

КОДИРОВАНИЕ В РЕАЛЬНОМ ВРЕМЕНИ ПРИ КВАДРАТИЧЕСКОМ КРИТЕРИИ ТОЧНОСТИ

Идеальное по Шеннону кодирование имеет смысл только при задании математического критерия точности. Выбор *квадратического критерия* позволяет получить наглядные результаты. Одним из главных требований информационно-решающих систем

является кодирование в реальном времени, так как решение, принятое с большой временно́й задержкой, теряет смысл. Примером кодирования без задержки на формирование блоков служит линейное кодирование, оптимизация которого рассмотрена при ограничении безразмерной энергии передаточного фильтра. Решение этой вариационной задачи дает оптимальный кодирующий фильтр, отличающийся от фильтра Костаса. Доказано, что *оптимальное линейное кодирование при некоторых условиях является абсолютно оптимальным, т. е. обеспечивает минимальную ошибку, совпадающую с получаемой при решении уравнения идеального кодирования Шеннона.* Этот вывод составляет содержание доказанной теоремы об идеальном линейном кодировании.

Идеальный по Шенному код в гауссовском случае реализуется с помощью фильтров двух типов (оптимального передающего фильтра и фильтра Винера) и не требует временно́й задержки на формирование блоков. Использование простых явных формул для информационных характеристик позволило обнаружить, что шум источника рождает шумовую информацию, которая в сумме с полезной информацией загружает канал.

2.1. ЛИНЕЙНОЕ КОДИРОВАНИЕ ПРИ ОГРАНИЧЕНИИ НА ФИЛЬТР

В отличие от вариационной задачи Костаса, решение которой было приведено в 1.2, мы заменим условие ограничения. Костас выбрал условие, которое ранее было использовано Шенноном в вариационной задаче определения пропускной способности [3, с. 454]. Это условие ограничивает безразмерную энергию сигнала (1.11) на входе канала — сумматора с шумом. В [21] была решена вариационная задача оптимизации кодека в схеме рис. 1.2 при другом типе ограничения — ограничении на безразмерную энергию (или мощность) усиления варьируемого фильтра

$$\sum_{k=1}^{\infty} g^2{}_k \leqslant E_g.$$

Для того чтобы наглядно проиллюстрировать то новое, что включает замена ограничения (1.11) на ограничение E_g, сначала приведем решение вариационной задачи для схемы на рис. 1.3 с одним шумом, а затем обобщим результат на схему с двумя шумами (см. рис. 1.2).

В схеме на рис. 1.3 над полезным сигналом $X(t)$ осуществляется необратимое линейное преобразование:

$$Y_k = g_k X_k + \nu_k;\ k \in [1, \infty].$$

Искаженный и зашумленный сигнал $Y(t)$ несет в себе оптимальную оценку $\hat{X}(t)$ полезного сигнала $X(t)$, которая является его условным математическим ожиданием:

$$\hat{X}(t) = \mathbf{M}[X(t) \mid Y(t)].$$

Эта запись выражает математическое ожидание полезного сигнала $X(t)$ при известном фиксированном значении процесса $Y(t)$. Извлечь из смеси $Y(t)$ наиболее существенную ее часть $\hat{X}(t)$ можно, пропустив $Y(t)$ через фильтр Винера:

$$\hat{X}_k = Y_k \frac{g_k^2 \lambda_k}{g_k^2 \lambda_k + \sigma_2^2};\quad k \in [1, \infty].$$

Оптимальная оценка $\hat{X}(t)$ отображает полезный сигнал с винеровской ошибкой

$$\varepsilon\left(g_k^2\right) = \mathbf{M}[X(t) - \hat{X}(t)]^2 = \sum_{k=1}^{\infty} \frac{\lambda_k \sigma_2^2}{g_k^2 \lambda_k + \sigma_2^2}, \tag{2.1}$$

где $g^2{}_k$ — заданные значения передающего фильтра G.

При геометрической трактовке процессов $[X(t) - \hat{X}(t)]$ и $Y(t)$ как векторов, теория утверждает, что вектора ортогональны:

$$[X(t) - \hat{X}(t)] \perp Y(t).$$

Условное математическое ожидание $\hat{X}(t)$ в гауссовском случае является минимальной достаточной статистикой и, образно говоря, его можно назвать «экстрактом» полезного сигнала, содержащимся в процессе $Y(t)$. Он содержит в «сжатом» виде ту же информацию о полезном сигнале, что и процесс $Y(t)$:

$$I[X(t), Y(t)] = I[X(t), \hat{X}(t)].$$

Выше, в разделе 1.2 было доказано, что в схеме на рис. 1.3 оптимальным декодирующим фильтром Q является фильтр Винера и при ограничении Костаса (1.9) и при ограничении E_g. Поэтому рассматриваемая вариационная задача сводится к минимизации винеровской ошибки $\varepsilon(g^2{}_k)$ с помощью вариации распределением безразмерной энергии усиления E_g между спектральными коэффициентами g_k варьируемого кодирующего фильтра

$$\min_{\{g_k^2\}} \varepsilon\left(g_k^2\right) = \mathbf{M}[X(t) - \hat{Z}(t)] = \varepsilon_{\text{мин}},$$

где $\hat{Z}(t)$ — сигнал на выходе декодирующего фильтра после выбора оптимального кодирующего фильтра при ограничении E_g. Теперь уже в класс варьируемых кодирующих фильтров G, среди которых ищется оптимум, не будет входить неискажающий фильтр с бесконечным числом степеней свободы ($g_k = 1$), как в задаче Костаса, так как он имеет бесконечную «энергию» усиления ($E_g = \infty$).

Построим функцию Лагранжа при ограничении E_g:

$$\Phi(g^2{}_k) = \varepsilon(g^2{}_k) + \Theta_2 \sum_{k=1}^{\infty} g^2{}_k,$$

где θ_2 — множитель Лагранжа.

Уравнение Эйлера определит искомую спектральную характеристику оптимального кодирующего фильтра:

$$g^2_{k\,\text{опт}} = \begin{cases} a\,(1-\sigma^2_2/a^2\,\lambda_k) & \text{при } k<N, \\ 0 & \text{при } k\geqslant N, \end{cases} \qquad (2.2)$$

где $a^2=\sqrt{\sigma^2{}_2/\theta_2}$.
В фильтре (2.2) величина a трактуется как коэффициент усиления, определяемый «энергией» усиления (2.1):

$$a^2 \sum_{k=1}^{\infty}(1-\sigma^2{}_2/a^2\lambda_k) = E_g.$$

Декодирующий фильтр Винера (1.9) при значениях коэффициентов $g_{k\,\text{опт}}$ (2.2) принимает вид:

$$q^2_{k\,\text{опт}} = \begin{cases} \dfrac{1}{a^2}\,(1-\sigma^2_2/a^2\,\lambda_k) & \text{при } k<N, \\ 0 & \text{при } k\geqslant N. \end{cases} \qquad (2.3)$$

Из сравнения выражений (2.2) и (2.3) видно, что спектральные характеристики *оптимальных кодирующего и декодирующего фильтров совпадают с точностью до коэффициента усиления:*

$$g_{k\,\text{опт}}/q_{k\,\text{опт}} = a^2.$$

Оптимальный кодирующий и декодирующий фильтры имеют конечное N число степеней свободы, определяемое из уравнения $a^2\lambda_k=\sigma^2{}_2$ при $k=N$. Подставив значения коэффициентов (2.3) в общую формулу (1.10) для вычисления среднеквадратической ошибки, получим простое выражение для ее минимума при ограничении (2.1):

$$\varepsilon_g = N\,\frac{\sigma^2_2}{a^2} + \sum_{k=N+1}^{\infty}\lambda_k = \sum_{k=1}^{\infty}\frac{\lambda_k\,\sigma^2_2}{g^2_{k\,\text{опт}}\,\lambda_k+\sigma^2_2}. \qquad (2.4)$$

Первое слагаемое минимальной ошибки (2.4) «выбелено» по передаваемым N степеням свободы. Второе слагаемое (2.4) отражает усечение ряда (1.1), вызываемое конечномерным кодирующим фильтром (2.2).

Проанализируем полученное решение вариационной задачи. Теперь стало видно, что нового несет замена условия ограничения для оптимального линейного кодирования: она дала новый тип оптимального кодирующего фильтра [12, 20] (2.2), отличающийся от фильтра Костаса. Для краткости назовем новый фильтр (2.2) передаточным фильтром, так как его назначение — оптимальная передача полезного сигнала $X(t)$ по каналу с шумом.

Оптимальный передающий фильтр (2.2) не только усиливает сигнал в a раз, но и предыскажает его с усечением числа степеней свободы в зависимости от спектрального отношения усиленного сигнала к шуму [1, 2]:

$$\Psi^2{}_k = a^2\lambda_k/\sigma^2{}_2.$$

В отличие от фильтра Костаса (1.12) оптимальный передающий фильтр (2.2) описывается весьма простой формулой, показывающей его связь со спектральным отношением сигнал-шум и коэффициентом усиления:

$$g^2_{k\,\text{опт}} = \begin{cases} a^2\,(\Psi^2_k-1)/\Psi^2_k & \text{при } \Psi^2_k>1; \quad k<N. \\ 0 & \text{при } \Psi^2_k\leqslant 1; \quad k\geqslant N. \end{cases}$$

Конечный запас «энергии» фильтра E_g расходуется только на усиление первых N членов разложения Карунена — Лоэва полезного сигнала (1.1), которые обладают отношением сигнал-шум, превышающим пороговое значение, равное единице:

$$g^2{}_k>0 \text{ для } \Psi^2{}_k>1;\ k<N.$$

Подчеркнем, что спектральная характеристика оптимального передаточного фильтра имеет форму, близкую к прямоугольной. Значения спектральной характеристики уменьшаются всего в два раза при снижении отношения сигнал-шум до двух:

$$g^2{}_{k\,\text{опт}} = a^2/2 \text{ при } \Psi^2{}_k = 2.$$

В узком диапазоне значений отношения сигнал-шум от двух до единицы спектральная характеристика фильтра (2.3) круто падает, обращаясь в нуль.

«Энергия» усиления E_g не расходуется на усиление «хвостовых» членов ряда (1.1), обладающих спектральным отношением сигнал-шум ниже порогового:

$$g^2{}_k = 0 \text{ для } \Psi^2{}_k\leqslant 1;\ k\geqslant N.$$

Такой расход «энергии» усиления E_g в передаточном фильтре является причиной вырождения решения вариационной задачи: сигнал $S(t)$ на входе зашумленного канала несет информацию только об N членах ряда $\{S_k\}$, описывающего искаженный фильтром G полезный сигнал:

$$S_k = \begin{cases} a^2\,\dfrac{\Psi^2_k-1}{\Psi^2_k}\,X_k & \text{при } k<N, \\ 0 & \text{при } k\geqslant N. \end{cases}$$

Если израсходовать «энергию» усиления фильтра E_g на передачу бесконечного числа членов обобщенного ряда Фурье (1.1), то на каждый член придется бесконечно малая доля «энергии» E_g. Усиление одного члена ряда в среднем ослабнет, и это приведет к более сильному поражению членов ряда (1.1) шумом и возрастанию ошибки воспроизведения относительно $\varepsilon_{\text{мин}}$. Поэтому любой бесконечномерный кодирующий фильтр G будет хуже в смысле квадратической ошибки, чем конечномерный передаточный фильтр (2.2). Ограничение «энергии» усиления E_g кодирующего фильтра является причиной конечного числа степеней свободы сигнала N, восстанавливаемого с минимальной ошибкой (2.4).

Ошибка (2.4) состоит из двух составляющих. Второе слагаемое в сумме (2.4) хорошо известно как ошибка усечения ряда (1.1). Новым моментом является учет в формуле (2.4) шумовой составляющей: на каждую из N степеней свободы приходится шумовая ошибка σ^2_2/a^2. Это означает, что шумовая составляющая ошибки (2.4) распределена равномерно по степеням свободы.

Порог в передаточном фильтре (2.2), который усекает ряд (1.1), описывающий полезный сигнал, определяется ослабленной дисперсией шума:

$$\lambda_{\text{пор}}=\sigma^2_2/a^2.$$

Хвостовые члены ряда (1.1) с индексом $k>N$ имеют «энергию» членов ниже порога ($\lambda_k<\lambda_{\text{пор}}$). Образно говоря, эти члены «утонули» в шуме и не дадут информации о полезном сигнале [1]. Только в абстрактном случае отсутствия шума ($\sigma^2_2=0$) или при допущении бесконечного усиления ($E_g=\infty$ и $a=\infty$) порог равен нулю и не происходит усечения числа степеней свободы входного бесконечномерного сигнала.

Мы рассмотрели однокаскадную схему на рис. 1.3, однако все практические схемы обычно содержат много каскадов. Поэтому полученные результаты важно распространить на многокаскадные схемы. Для упрощения изложения мы рассмотрим двухкаскадную схему (см. рис. 1.2), которая характеризуется двумя необратимыми линейными преобразованиями:

$$Y_k=\beta_k X_k+\xi_k;\ \tilde{Y}_k=r_k Y_k+\nu_k;\ r_k=l_k g_k.$$

Зашумленный сигнал $Y(t)$ после прохождения декодирующего фильтра Q становится доступным получателю в форме сигнала $Z(t)$:

$$Z_k=q_k\tilde{Y}_k=q_k(r_k Y_k+\nu_k);\quad k\in[1,\ \infty]. \tag{2.5}$$

Условные математические ожидания полезного сигнала $X(t)$ при фиксированных процессах $Y(t)$ и $\tilde{Y}(t)$ определяются следующим образом:

$$\hat{X}(t)=\mathbf{M}[X(t)\,|\,Y(t)];\ \tilde{X}(t)=\mathbf{M}[X(t)\,|\,\tilde{Y}(t)].$$

Подобно тому, как оптимальная оценка $\hat{X}(t)$ является наиболее существенной частью смеси $Y(t)$, так и процесс $\tilde{X}(t)$ представляет оптимальную оценку полезного сигнала, содержащуюся в «сжатом» виде в смеси $\tilde{Y}(t)$. Для выделения из $\tilde{Y}(t)$ оптимальной оценки $\tilde{X}(t)$ необходимо в схеме на рис. 1.2 выбрать в качестве декодирующего фильтра Q фильтр Винера. На вход декодирующего фильтра Винера Q поступает смесь $\tilde{Y}(t)$, полученная необратимым преобразованием $Y(t)$ с содержащимся в нем условным математическим ожиданием $\hat{X}(t)$, на его выходе сигнал $Z(t)$ будет равен:

$$\tilde{X}(t)=\mathbf{M}[X(t)\,|\,\tilde{Y}(t)]=\mathbf{M}[\hat{X}(t)\,|\,\tilde{Y}(t)].$$

Оптимальная оценка $\hat{X}(t)$ в первом каскаде в отличие от сигнала $Z(t)=\tilde{X}(t)$ не доступна получателю, но фильтр R всегда можно

умозрительно разбить на два звена (фильтры L и G) и выделить первое звено в качестве фильтра Винера:

$$l_k=\hat{l}_k=\frac{\beta_k\lambda_k}{\beta_k^2\lambda_k+\sigma_1^2};\quad k\in[1,\ \infty]. \tag{2.6}$$

Тогда первое звено L фильтра R способно выдать оптимальную оценку $\hat{X}(t)$, которая воспроизводит входной сигнал $X(t)$ с винеровской ошибкой

$$\varepsilon_B=\mathbf{M}[X(t)-\hat{X}(t)]^2=\sum_{k=1}^{\infty}\frac{\lambda_k\sigma_1^2}{\beta_k^2\lambda_k+\sigma_1^2}.$$

Из определений условных математических ожиданий следует ортогональность векторов, представляющих разностные процессы

$$[X(t)-\hat{X}(t)]\perp[\hat{X}(t)-\tilde{X}(t)].$$

Отсюда мы можем сформулировать *правило покаскадного суммирования винеровских ошибок в многокаскадной схеме:*

$$\mathbf{M}[X(t)-\tilde{X}(t)]^2=\mathbf{M}[X(t)-\hat{X}(t)]^2+\mathbf{M}[\hat{X}(t)-\tilde{X}(t)]^2. \tag{2.7}$$

Хотя мы получили это правило для двухкаскадной схемы (см. рис. 1.2), оно может быть обобщено методом индукции на любое число каскадов.

Оценка полезного сигнала $Z(t)=\tilde{X}(t)$ является оптимальной, т. е. обеспечивающей суммарную винеровскую ошибку (2.7) с помощью декодирующего фильтра Винера Q в схеме на рис. 1.2 при заданных передающих фильтрах B и R. Будем теперь считать, что передающий фильтр B первого каскада остается заданным, а передающий фильтр в последнем (втором) каскаде можно варьировать. В этом случае для двухкаскадной схемы на рис. 1.2 возникает вариационная задача, более общая, чем вариационная задача, решенная выше для однокаскадной схемы на рис. 1.3. Располагая возможностью варьировать передаточным фильтром R второго каскада, мы можем найти оценку $\tilde{Z}(t)$, лучшую, чем оценка $\tilde{X}(t)$:

$$\varepsilon_{\text{мин}}=\mathbf{M}[X(t)-\hat{Z}(t)]=\min_{\{r_k\}}\mathbf{M}[X(t)-\tilde{X}(t)]^2=$$

$$=\varepsilon_B-\min_{\{g_k\}}\mathbf{M}[\hat{X}(t)-\tilde{X}(t)]. \tag{2.8}$$

Решим эту вариационную задачу спектральным методом. Для этого запишем спектральную формулу для вычисления среднеквадратической ошибки в схеме на рис. 1.2, выделив в ней полный квадрат относительно коэффициентов q_k [15]:

$$\varepsilon=\mathbf{M}[X(t)-Z(t)]^2=$$

$$=\sum_{k=1}^{\infty}\left\{\lambda_k-\left[q_k-\frac{r_k\beta_k\lambda_k}{r_k^2(\beta_k^2\lambda_k+\sigma_1^2)+\sigma_2^2}\right]^2[r_k^2(\beta_k^2\lambda_k-\sigma_1^2)+\sigma_2^2]-\right.$$

$$\left.-\frac{r_k^2\beta_k^2\lambda_k^2}{r_k^2(\beta_k^2\lambda_k+\sigma_1^2)+\sigma_2^2}\right\}. \tag{2.9}$$

Анализ спектрального выражения (2.9), обобщающего прежнее выражение (1.8), подтверждает справедливость вывода о том, что оптимальный декодирующий фильтр в схеме на рис. 1.2 относится к типу фильтров Винера:

$$q_{k\,\text{опт}} = \frac{\beta_k r_k \lambda_k}{r_k^2 (\beta_k^2 \lambda_k + \sigma_1^2) + \sigma_2^2}. \qquad (2.10)$$

Фильтр (2.10) является фильтром Винера, обобщенным с учетом заданного фильтра B и двух шумов с дисперсиями $\sigma^2{}_1$ и $\sigma^2{}_2$. Ошибка (2.9) при декодирующем фильтре Винера (2.10) равна

$$\varepsilon = \mathbf{M}[X(t) - \hat{X}(t)]^2 = \frac{\lambda_k (r_k^2 \sigma_1^2 + \sigma_2^2)}{r_k^2 (\beta_k^2 \lambda_k + \sigma_1^2) + \sigma_2^2}.$$

Используя правило сложения винеровских ошибок (2.7), можно заметить, что эта ошибка является винеровской ошибкой двух каскадов и потому состоит из двух слагаемых:

$$\varepsilon = \varepsilon_B + \varepsilon(g_k^2) = \sum_{k=1}^{\infty} \left[\frac{\lambda_k \sigma_1^2}{\beta_k^2 \lambda_k + \sigma_1^2} + \sum_{k=1}^{\infty} \frac{\hat{\lambda}_k \sigma_2^2}{g_k^2 \hat{\lambda}_k + \sigma_2^2} \right], \qquad (2.11)$$

где $g_k^2 = \dfrac{r_k^2}{\hat{l}_k}$; $\hat{\lambda}_k = (\beta_k \lambda_k + \sigma_1^2)\,\hat{l}_k = \dfrac{\beta_k^2 \lambda_k^2}{\beta_k^2 \lambda_k + \sigma_1^2}$.

Первое слагаемое (2.11) зависит от передающего фильтра B первого каскада, а второе слагаемое — от передающего фильтра G. Будем считать, что первый каскад в схеме на рис. 1.2 задан, т. е. не доступен нашему воздействию. Следовательно, задана винеровская ошибка ε_B первого каскада. Остается возможность уменьшить второе слагаемое ошибки (2.11). Сравнив ее с ошибкой (2.1), полученной в схеме на рис. 1.3, видим, что одна получается из другой заменой спектра λ_k сигнала $X(t)$ на спектр $\hat{\lambda}_k$ оптимальной оценки $\hat{X}(t)$. Поэтому решение вариационной задачи для однокаскадной схемы (см. рис. 1.3) легко распространить на случай, когда варьируется кодирующий фильтр R в двухкаскадной схеме (см. рис. 1.2). Для этого достаточно решить вариационную задачу

$$\min_{\{g_k^2\}} \varepsilon(g_k^2) = \varepsilon_g, \quad \sum_{k=1}^{\infty} g_k^2 \leqslant E_g. \qquad (2.12)$$

Для решения задачи (2.12) при ограничении E_g составим функцию Лагранжа

$$\Phi(g^2{}_k) = \varepsilon(g^2{}_k) + \Theta_3 \sum_{k=1}^{\infty} g^2{}_k, \qquad (2.13)$$

где Θ_3 — множитель Лагранжа.

Решая уравнение Эйлера для функции (2.13), получим коэффициенты оптимального кодирующего двухзвенного фильтра в схеме на рис. 1.2 [15]

$$r_{k\,\text{опт}}^2 = \hat{l}_{k\,\text{опт}}^2 g_{k\,\text{опт}}^2 = \frac{\beta_k \lambda_k}{\beta_k^2 \lambda_k + \sigma_1^2} g_{k\,\text{опт}}^2, \qquad (2.14)$$

где $g_{k\,\text{опт}}^2 = \begin{cases} a^2 (1 - \sigma_2^2 / a^2 \hat{\lambda}_k) & \text{при } k < N, \\ 0 & \text{при } k \geqslant N. \end{cases}$

Первое звено L оптимального двузвенного фильтра R относится к типу фильтров Винера с коэффициентами (2.6), а второе звено является передаточным фильтром (2.2), обобщенным с учетом шума источника. Коэффициент усиления передаточного фильтра определяется имеющимся запасом «энергии»:

$$a^2 \sum_{k=1}^{\infty} (1 - \sigma^2{}_2 / a^2 \hat{\lambda}_k)^2 = E_g.$$

Подставляя коэффициенты (2.14) в выражение для декодирующего фильтра Винера (2.10), легко убедиться, что оптимальные передающий G и декодирующий Q фильтры отличаются лишь усилением

$$a^2 q_{k\,\text{опт}} = g_{k\,\text{опт}}.$$

Оптимизация фильтров определила правило построения оптимального линейного кода $\hat{Z}(t)$:

$$\hat{Z}_k = q_{k\,\text{опт}} g_{k\,\text{опт}} \hat{X}_k + q_{k\,\text{опт}} v_k.$$

Подстановка в (2.11) значения оптимальных коэффициентов (2.14) дает ее минимальное значение:

$$\varepsilon_{\text{мин}} = \varepsilon_B + \varepsilon_g = \varepsilon_B + N \frac{\sigma_2^2}{a^2} + \sum_{k=N+1}^{\infty} \hat{\lambda}_k.$$

В соответствии с (2.8) вторая составляющая минимальной ошибки характеризует ошибку, с которой оптимальный линейный код $\hat{Z}(t)$ восстанавливает оценку $\hat{X}(t)$:

$$\varepsilon_g = \mathbf{M}[\hat{X}(t) - \hat{Z}(t)]^2 = N \sigma^2{}_2 / a^2 + \sum_{k=N+1}^{\infty} \hat{\lambda}_k. \qquad (2.15)$$

С минимальной ошибкой (2.15) передаточный фильтр $g_{k\,\text{опт}}$ передает оценку $\hat{X}(t)$ через зашумленный канал. По аналогии с термином «ошибка фильтра Винера» можно назвать ошибку (2.15) «ошибкой передаточного фильтра».

Вырождение решения обобщенной вариационной задачи вызывается конечным ресурсом «энергии» усиления E_g передаточного фильтра. Воспользуемся понятием о спектральном энергетическом отношении сигнал-шум:

$$\Psi^2{}_k = a^2 \lambda_k / \sigma^2{}_2.$$

Первые N членов разложения Карунена — Лоэва сигнала на выходе фильтра L типа фильтра Винера формируются с отношением сигнал-шум $\Psi^2_k > 1$. «Хвостовые» члены разложения, для которых отношение сигнал-шум $\Psi^2_k < 1$, усекаются передаточным фильтром. Спектральное отношение сигнал-шум определяет и форму спектральной характеристики передаточного фильтра (2.14):

$$g^2_{k\,\text{опт}} = \begin{cases} a^2\,(\Psi^2_k - 1)/\Psi^2_k & \text{при } \Psi^2_k > 1;\ k < N, \\ 0 & \text{при } \Psi^2_k \leqslant 1;\ k \geqslant N. \end{cases} \tag{2.16}$$

Этот передаточный фильтр имеет пороговое значение спектрального отношения сигнал-шум, равное единице ($\Psi^2_{\text{пор}} = 1$). Его спектральная характеристика монотонно падает и близка к постоянной усиления ($g^2_k \approx a^2$) в области значений отношений сигнал-шум, существенно превышающих пороговое значение ($\Psi^2_k \gg \gg 1$).

Сравним теперь оптимальные кодирующие фильтры (2.2) и (2.14). Фильтр (2.2) был найден нами в однокаскадной схеме на рис. 1.3. Фильтр (2.14) получен для двухкаскадной схемы на рис. 1.2 в том случае, когда передающий фильтр первого каскада считался заданным и мы могли выбирать лишь в последнем (втором) каскаде кодирующий фильтр с «энергией» усиления E_g. На вход кодирующего фильтра R поступает смесь $Y(t)$ сигнала и шума, содержащая оценку $\hat{X}(t)$. С выхода кодирующего фильтра сигнал $S(t)$ поступает в шумящий канал, представленный на рис. 1.2 вторым сумматором с шумом. Поэтому кодирующий фильтр (2.14) в отличие от фильтра (2.2) выполняет не одну, а две функции: выделение из смеси $Y(t)$ ее «экстракта» $\hat{X}(t)$ и усиление его для передачи по каналу с шумом. Вполне естественно, что фильтр (2.14) состоит из двух звеньев: фильтра Винера $\{l_k\}$, формирующего оценку $\hat{X}(t)$, и передаточного фильтра (2.16), оптимально использующего отведенную «энергию» усиления E_g для передачи по каналу. Становится ясным, что фильтру (2.2) надо сопоставлять не весь фильтр (2.14), а его второе звено (2.16). Тогда видно, что передаточный фильтр (2.16) можно получить формально из фильтра (2.2) простой заменой дисперсии λ_k полезного сигнала $X(t)$ на дисперсию $\hat{\lambda}_k$ оптимальной оценки $\hat{X}(t)$. Для передаточного фильтра (2.16) второго каскада полезным сигналом служит не процесс $X(t)$, а его условное математическое ожидание $\hat{X}(t)$. В какой степени оптимальный кодирующий фильтр (2.14), введенный нами во второй каскад, способен компенсировать искажения, вызываемые заданным фильтром B первого каскада? Такая компенсация возможна лишь частично для наименее зашумленных степеней свободы, где фильтр (2.14) становится обратным искажающему фильтру B («компенсатором» [14]):

$$r_{k\,\text{опт}} \approx \frac{1}{\beta_k} \text{ при } \beta^2_k \lambda_k \gg \sigma^2_1;\ a^2 \hat{\lambda}_k \gg \sigma^2_2.$$

Очевидно, что если бы в двухкаскадной схеме на рис. 1.2 предоставила возможность выбора передающего фильтра не только во втором каскаде, как мы рассмотрели выше, но и в первом каскаде, то мы достигли бы лучших результатов в компенсации искажений. Однако на практике чаще встречаются ситуации, когда компенсатору предшествует не один заданный каскад, как мы рассмотрели выше, а несколько каскадов, каждый со своим заданным передающим фильтром. В этих ситуациях можно воспользоваться теми же решениями, которые изложены для двухкаскадной схемы на рис. 1.2. При этом ясно, что «энергию» усиления E_g фильтра желательно расходовать в каскаде, наиболее близком к первому.

Хотя, исследуя линейное кодирование в схеме на рис. 1.2, мы и пользовались такими терминами из теории информации, как «код», «источник», «канал», однако, по существу, решили вариационную задачу и нашли оптимальный двузвенный кодирующий фильтр, не пользуясь понятиями об энтропии источника и пропускной способности канала. Для достижения оптимальности линейного кодирования нам, как и Костасу, не понадобилась теория Шеннона. Однако теперь, после вычисления ошибки (2.15), присущей оптимальному линейному кодированию, мы поставим вопрос о сравнении ее с минимальной ошибкой из уравнения идеального кодирования Шеннона (1.17). Иначе говоря, мы хотим выяснить, даст ли нелинейный блоковый код Шеннона при $n_2/n_1 = 1$ меньшую ошибку, чем ошибка (2.15) оптимального линейного кода. Ответ на этот вопрос будет дан в 2.5, а предварительно рассмотрим эффект расщепления ε-энтропии источника из-за шумовой информации.

2.2. РАЗДВОЕНИЕ ε-ЭНТРОПИИ ИЗ-ЗА ШУМОВОЙ ИНФОРМАЦИИ (ОДНОМЕРНЫЙ СЛУЧАЙ)

Найдем информационные характеристики линейной схемы с двумя шумами (см. рис. 1.2) в частном случае, когда в ней действуют не случайные процессы (бесконечномерные векторы), а непрерывные случайные величины: сигналы X, Y, Z и шумы ξ и ν. Источник, порождающий гауссовскую случайную величину, в литературе часто называют дискретным по времени гауссовским источником [18]. Соответственно канал в схеме на рис. 1.2 называют «дискретный по времени канал с аддитивным гауссовским шумом» [18]. Мы же будем называть такого типа непрерывный источник и канал просто одномерными, имея в виду отличие случайных величин от случайных процессов, представленных многомерными векторами.

В данном разделе рассмотрим информационную характеристику источника, имеющего собственный шум. Учет влияния шума источника заставляет пересмотреть одно из важных положений теории информации о том, что приспособленность информации источника к передаче характеризуется единственной вели-

чиной — энтропией или ε-энтропией [4, с. 38]. Для того чтобы наглядно это продемонстрировать, обратимся к источнику гауссовского сигнала X (непрерывная случайная величина), который имеет плотность вероятности

$$p(x)=[\exp(-x^2/2\lambda)]/\sqrt{2\pi\lambda},$$

где $\lambda=\mathbf{M}(X^2)$.

Одномерный источник создает смесь двух сигналов: полезного X и шумового ξ, которые поступают на вход кодера в линейной комбинации:

$$Y=\beta X+\xi,$$

где β — коэффициент одномерного фильтра B. Рассмотрим последовательность трех сигналов: X, Y, Z. Кодирование является процессом построения выходного сигнала Z по зашумленному сигналу Y источника. Этот процесс характеризуется условной или переходной плотностью вероятности

$$p(z|y, x).$$

Переходная плотность вероятности описывает распределение значений сигнала Z при заданных сигналах Y и X.

Мы будем искать оптимальный код $\hat{Z}$ среди тех кодов Z, для которых выполняется следующее условие:

$$p(z|y, x)=p(z|y).$$

При этом средняя взаимная информация между кодом Z и парой сигналов (X, Y) равна

$$I[Z, (X, Y)]=I(Z, Y).$$

Выделенное условие означает, что для построения кода Z никакой другой информацией о сигнале X, кроме той ограниченной информации, которая содержится в зашумленном сигнале Y, мы не располагаем. Среднюю взаимную информацию между кодом Z и зашумленным сигналом Y, которая используется при построении кода, назовем кодовой. Для гауссовских сигналов кодовая информация равна:

$$I(Y, Z)=\frac{1}{2}\log\left[1+\frac{r^2(\beta^2\lambda+\sigma_1^2)}{\sigma_2^2}\right]. \qquad (2.17)$$

Так же как зашумленный сигнал источника Y, выходной сигнал состоит из двух составляющих: полезной сигнальной и шумовой:

$$Z=qrY+q\nu=qgl\beta X+\mathcal{U},$$

где $\mathcal{U}=q(gl\xi+\nu)$.

Выходной сигнал Z содержит шум $\mathcal{U}$, который представляет сумму шумов источника и канала на выходе декодера q и имеет дисперсию

$$\mathbf{M}(\mathcal{U}^2)=(r^2\sigma^2{}_1+\sigma^2{}_2)q^2.$$

Заметим, что выходной шум равен сигналу на выходе декодера, когда «отключен» полезный сигнал:

$$\mathcal{U}=Z \text{ при } X=0.$$

Разделение сигнала Z на выходе декодера на две составляющих, из которых одна — полезная, а другая — шум $\mathcal{U}$, является следствием того, что источник создает сумму полезного сигнала βX и шума ξ. Учет этого очевидного факта позволяет сделать важный вывод о том, что кроме кодовой взаимной информации (2.17) можно определить среднюю взаимную информацию между парой сигналов X и Z и между парой шумов ξ и $\mathcal{U}$. Среднюю взаимную информацию между парой сигналов X и Z назовем сигнальной, поскольку она дает информацию о полезном сигнале:

$$I(X, Z)=\frac{1}{2}\log\left(1+\frac{r^2\beta^2\lambda}{r^2\sigma_1^2+\sigma_2^2}\right). \qquad (2.18)$$

Среднюю взаимную информацию между входным и выходным шумами, понятие о которой введено в [22], естественно назвать шумовой информацией:

$$I(\xi, \mathcal{U})=\frac{1}{2}\log\left(1+\frac{r^2\sigma_1^2}{\sigma_2^2}\right). \qquad (2.19)$$

Применим понятие о коэффициенте шума к системе на рис. 1.2:

$$b=\frac{r^2\sigma_1^2+\sigma_2^2}{r^2\sigma_1^2}.$$

Шумовую информацию (2.19) можно вычислить с помощью коэффициента шума:

$$I(\xi, \mathcal{U})=\frac{1}{2}\log\frac{b}{b-1}.$$

Видно, что если коэффициент шума системы на рис. 1.2 стремится к единице ($b\to 1$), то шумовая информация будет неограниченно расти.

В формулы (2.17), (2.18) и (2.19), полученные нами в гауссовском случае, вошел только коэффициент r кодирующего фильтра, так как информация не зависит от обратимого линейного декодирующего фильтра q.

Анализируя (2.17), (2.18) и (2.19), можно обнаружить, что величины, определяемые этими формулами, удовлетворяют следующему равенству:

$$I(Y, Z)=I(X, Z)+I(\xi, \mathcal{U}). \qquad (2.20)$$

Оно означает, что *кодовая информация равна сумме полезной сигнальной информации и шумовой информации.* Вывод (2.20) содержит элемент неожиданности. Он заключается в следующем. Обычно пользуются такой картиной прохождения информации в

системе. Источник смеси полезного сигнала X и шума ξ потенциально создает на входе системы информацию конечной величины:

$$I(X, Y) = I(X, \hat{X}) = \frac{1}{2} \log \left(\frac{\beta^2 \lambda}{\sigma_1^2} + 1 \right).$$

Полезная информация $I(X, Y)$ рассматривается как входная и учитывает лимитирующее влияние шума источника. От входа системы к выходу системы полезная информация только теряется:

$$I(X, Z) < I(X, Y).$$

В этой картине, казалось бы, уже учтено полностью лимитирующее влияние шума источника. Однако оказалось, что такого учета недостаточно. Шум источника не только лимитирует полезную сигнальную информацию, но и *сам является источником шумовой («ложной») информации.* Это означает, что на входе системы имеется не только чисто полезная информация, но и шумовая. Оба вида информации сливаются в единый поток кодовой информации и поступают в систему.

Равенство (2.20) позволяет трактовать шумовую информацию как кодовую информацию, когда сигнальная равна нулю, т. е. когда «отключен» полезный сигнал:

$$I(\xi, \mathcal{U}) = I(Y, Z) \text{ при } I(X, Z) = 0;\ X = 0.$$

Новизна (2.20) проявляется в том, что оно учитывает информационно две стороны влияния шума источника. С одной стороны, шум источника ограничивает доступ ко всей полезной информации $I(X, X)$ величиной $I(X, Y)$, содержащейся в зашумленном сигнале. С другой стороны, шум источника загружает кодер и канал шумовой информации $I(\xi, \mathcal{U})$. Кодер обязан перерабатывать не только доступную сигнальную информацию, но и неустранимую шумовую. Поэтому источник информации из-за наличия шума уже нельзя характеризовать одной ε-энтропией. Рассмотрим возникающую специфику.

Эпсилон-энтропия вычисляется при заданном значении критерия качества (точности) восстановления сигнала. Мы будем пользоваться только среднеквадратической ошибкой восстановления — ошибкой ε, геометрически трактуемой как расстояние между векторами X и Z:

$$\varepsilon = \mathbf{M}(X - Z)^2 = d(X, Z).$$

Запишем общее выражение для среднеквадратической ошибки через совместную плотность вероятностей $p(x, z)$:

$$d(X, Z) = \int_X \int_Y \int_Z d(x, z)\, p(x, y, z)\, dydxdz, \tag{2.21}$$

где $d(x, z) = (x - z)^2$.

В выражении (2.21) совместная плотность вероятностей равна:

$$p(x, y, z) = p(z|y, x)\, p(y, x) = p(z|y)\, p(y)\, p(x|y).$$

Подставляя эту плотность вероятностей в выражение для ошибки (2.21), преобразуем его к виду

$$d(X, Z) = \int_Z \int_Y p(z|y)\, p(y)\, dydz \int_X (x - z)^2\, p(x|y)\, dx =$$

$$= \int_Z \int_Y p(z|y)\, p(y)\, dydz \left\{ \int_X x^2\, p(x|y)\, dx - \right.$$

$$\left. - \int_X 2zxp(x|y)\, dx + \int_X z^2\, p(x|y)\, dx \right\}. \tag{2.22}$$

Рассмотрим второй интеграл в сумме, выделенной фигурными скобками. Согласно условию, которое наложено нами на код Z, он строится только по сигналу Y, не располагая сигналом X. Поэтому можно величину z вынести за знак интеграла:

$$z \int_X xp(x|y)\, dx = z\hat{x}.$$

Первый интеграл в сумме, выделенной фигурными скобками, равен:

$$\int_X x^2\, p(x|y)\, dx = \int_X (x - \hat{x})^2\, p(x|y)\, dx + \int_X \hat{x}^2\, p(x|y)\, dx.$$

Подставим значения двух интегралов в выражение (2.22)

$$d(X, Z) = \int_Z p(z|y)\, dz \int_Y \int_X (x - \hat{x})^2\, p(x, y)\, dxdy +$$

$$+ \int_Z \int_Y (\hat{x} - z)^2\, p(y, z)\, dydz. \tag{2.23}$$

В равенство (2.23) вошло $\hat{x}$ — условное математическое ожидание $\hat{x} = \mathbf{M}(x|Y)$ сигнала X при фиксированном сигнале Y. В гауссовском случае $\hat{X}$ является оптимальной оценкой полезного сигнала X, которая получается на выходе фильтра Винера по входному зашумленному сигналу Y с минимальной ошибкой

$$\varepsilon_B = d(X, \hat{X}) = \int_X \int_Y (x - \hat{x})^2\, p(x, y)\, dxdy.$$

Равенство (2.23) представляет ошибку $d(X, Z)$ в виде суммы двух ошибок: ошибки фильтра Винера и ошибки отображения сигнала $\hat{X}$ кодом Z:

$$d(\hat{X}, Z) = \int_Z \int_Y (\hat{x} - z)^2\, p(y, z)\, dydz.$$

Вывод равенства (2.23) служит доказательством справедливости равенства (2.7), использованного в разделе 2.1, так как вывод обобщается и на многомерный случай.

Оптимальная оценка $\hat{X}$ строится по сигналу Y и является его «экстрактом», содержащим всю доступную информацию о полезном сигнале X, содержащуюся в сигнале Y:

$$I(X, Y) = I(X, \hat{X}). \tag{2.24}$$

С учетом сказанного выше среднеквадратическая ошибка равна:

$$d(X, Z) = d(X, \hat{X}) + d(\hat{X}, Z). \tag{2.25}$$

Подчеркнем, что при выводе равенства (2.25) мы не пользовались допущением о гауссовском законе распределения вероятностей. Поэтому равенство справедливо при любом непрерывном распределении вероятностей. Более того, его можно распространить и на дискретный случай, если используется квадратический критерий качества. В гауссовском случае первый член суммы (2.25) является ошибкой, присущей получению с помощью фильтра Винера оптимальной оценки $\hat{X}$ полезного сигнала по зашумленному сигналу Y:

$$\varepsilon_B = d(X, \hat{X}) = \mathbf{M}(X - \hat{X})^2 = \frac{\lambda\sigma_1^2}{\beta^2\lambda + \sigma_1^2}.$$

Равенство (2.25) показывает, что задача минимизации ошибки $d(X, Z)$ сводится к минимизации только ошибки $d(\hat{X}, Z)$:

$$\min d(X, Z) = d(X, \hat{X}) + \min d(\hat{X}, Z). \tag{2.26}$$

Кодер источника, участвующий в построении оптимального кода $\hat{Z}$, выделяет из кодовой информации $I(Y, Z)$ источника лишь то минимальное количество информации, равное ε-энтропии, которое достаточно для обеспечения заданной ошибки ε:

$$H_Y(\varepsilon) = I(Y, \hat{Z}) = \min_{p(z|y):\, d(X,Z) \leqslant \varepsilon} I(Y, Z).$$

Эпсилон-энтропию источника, определяемую по кодовой информации решением вариационной задачи, естественно называть кодовой. Код $\hat{Z}$, реализующий кодовую ε-энтропию, связан с оценкой соотношением, аналогичным ранее использованному при вычислении ε-энтропии без шума источника [3, с. 614, . с. 48]:

$$\hat{X} = \hat{Z} + V. \tag{2.27}$$

Гауссовская величина V, не зависящая от $\hat{X}$, выражает ошибку передачи оценки $\hat{X}$ кодом $\hat{Z}$:

$$\mathbf{M}(V^2) = \mathbf{M}(\hat{X} - \hat{Z})^2 = d(\hat{X}, \hat{Z}).$$

При определении кодовой ε-энтропии минимум разыскивается по всем переходным плотностям вероятностей $p(z|y)$, таким, что ошибка не превосходит заданного значения, т. е. выполняется условие

$$d(X, Z) \leqslant \varepsilon.$$

Как показывает равенство (2.25), это условие эквивалентно требованию, чтобы ошибка $d(X, Z)$ передачи сигнала не превосходила заданного значения ε_V:

$$d(X, Z) \leqslant \varepsilon_V, \text{ где } \varepsilon = \varepsilon_B + \varepsilon_V.$$

Равенство (2.25), заменяя первое условие на второе, позволяет свести вычисление кодовой ε-энтропии к вариационной задаче вычисления ε-энтропии, решенной в работах Колмогорова [4, с. 48] и Шеннона [3, с. 614]:

$$H_Y(\varepsilon_V) = I(Y, \hat{Z}) = \min_{p(z|y):\, d(\hat{X}, Z) \leqslant \varepsilon_V} I(Y, Z). \tag{2.28}$$

Решение вариационной задачи (2.28) в гауссовском случае дает простую формулу вычисления кодовой ε-энтропии источника:

$$H_Y(\varepsilon_V) = I(Y, \hat{Z}) = \begin{cases} \frac{1}{2}\log\frac{\hat{\lambda}}{\varepsilon_V} & \text{при } \varepsilon_V < \hat{\lambda}, \\ 0 & \text{при } \varepsilon_V \geqslant \hat{\lambda}. \end{cases} \tag{2.29}$$

Тот факт, что кодовая информация (2.20) состоит из сигнальной и шумовой составляющих, не позволяет характеризовать источник одной кодовой ε-энтропии. Введем вторую ε-энтропию, которую мы назовем сигнальной и которая определяется следующей вариационной задачей [22]:

$$H_X(\varepsilon) = I(X, \hat{Z}) = \min_{p(z|y)\,;\, d(X,Z) \leqslant \varepsilon} I(X, Z). \tag{2.30}$$

Для случая гауссовского распределения вероятностей сигнальная ε-энтропия (2.30) вычисляется с помощью простой формулы:

$$H_X(\varepsilon) = I(X, \hat{Z}) = \begin{cases} \frac{1}{2}\log\frac{\lambda}{\varepsilon} & \text{при } \varepsilon_B \leqslant \varepsilon < \lambda, \\ 0 & \text{при } \varepsilon \geqslant \lambda. \end{cases} \tag{2.31}$$

Превышение кодовой ε-энтропии (2.29) над сигнальной (2.31) или величины $I(Y, \hat{Z})$, над $I(X, \hat{Z})$ объясняется естественным образом. «Качество» чистой, не зашумленной информации лучше «качества» зашумленной информации. Поэтому надо брать меньше «чистых» двоичных единиц из сигнала X, чем «зашумленных» двоичных единиц из смеси сигнала и шума Y для обеспечения заданной ошибки восстановления. «Зашумленные» двоичные единицы кодовой ε-энтропии будут содержать не только полезную сигнальную информацию, но и шумовую. Это и служит естественной причиной распадания кодовой ε-энтропии на две части [22].

$$H_Y(\varepsilon_V) = H_X(\varepsilon) + I(\xi, \hat{\mathcal{U}}) \tag{2.32}$$

или $I(Y, \hat{Z}) = I(X, \hat{Z}) + I(\xi, \hat{\mathcal{U}})$, где $\hat{\mathcal{U}} = \hat{Z}$ при $X = 0$.
Равенство (2.32) вытекает из равенства (2.20), которое должно выполняться при любых Z и, в частности, при оптимальном коде $\hat{Z}$.

Когда шумовая информация равна нулю из-за гипотетического отсутствия шума источника ($\xi = 0$), формулы (2.29) и (2.31)

сливаются в одную «старую» формулу, приведенную в работе [3, с. 6.16]:

$$H(\varepsilon)=\begin{cases}\frac{1}{2}\log\frac{\lambda}{\varepsilon} & \text{при } 0\leqslant\varepsilon<\lambda;\\ 0 & \text{при } \varepsilon\geqslant\lambda,\end{cases} \tag{2.33}$$

где $\xi=0$, $\varepsilon_B=0$, $\lambda=\hat{\lambda}$.

Зная потенциально существующее количество сигнальной информации $I(X, Y)$ и сигнальную ε-энтропию (2.30), легко вычислить потерю сигнальной информации при оптимальном кодировании:

$$\Delta I=I(X, Y)-I(X, \hat{Z})=H_X(\varepsilon_B)-H_X(\varepsilon)=\frac{1}{2}\log\frac{\varepsilon}{\varepsilon_B}, \tag{2.34}$$

где $H_X(\varepsilon_B)=I(X, Y)=\frac{1}{2}\log\frac{\lambda}{\varepsilon_B}=\frac{1}{2}\log\left(1+\frac{\beta^2\lambda}{\sigma_1^2}\right)$.

Шумовую информацию можно вычислить как разность между кодовой и сигнальной ε-энтропиями [22]:

$$I(\xi, \hat{\mathcal{U}})=H_Y(\varepsilon_V)-H_X(\varepsilon)=\frac{1}{2}\log\frac{\varepsilon\hat{l}}{\varepsilon-\varepsilon_B}, \tag{2.35}$$

где $\hat{l}=\frac{\beta\lambda}{\beta^2\lambda+\sigma_1^2}$ — коэффициент фильтра Винера.

Подчеркнем, что все явные формулы (2.29), (2.31) и (2.35) получены спектральным методом только для гауссовского случая, но исходные определения ε-энтропии (2.28) и (2.30) справедливы для любых дискретных и непрерывных распределений вероятностей. Для иллюстрации указанных формул на рис. 2.1 приведены зависимости двух ε-энтропий от ошибки ε. Нижняя кривая на рис. 2.1 в диапазоне значений ошибки $\varepsilon_B\leqslant\varepsilon\leqslant\lambda$ отражает сигнальную ε-энтропию (2.31). Верхняя кривая на рис. 2.1 соответствует кодовой ε-энтропии (2.29). Вся нижняя кривая вместе с ее продолжением в область ошибок $0\leqslant\varepsilon\leqslant\varepsilon_B$ (показана штриховой линией) описывает графически формулу (2.33). По существу, можно сказать, что рис. 2.1 иллюстрирует эффект «раздвоенной» «старой» ε-энтропии на кодовую и сигнальную под влиянием шумовой информации. В ходе «раздвоения» исходная кривая, отражающая формулу (2.33), усекается в области $\varepsilon<\varepsilon_B$, давая нижнюю кривую, и сдвигается на ε_B с заменой дисперсии λ на $\hat{\lambda}$, давая верхнюю кривую (2.1).

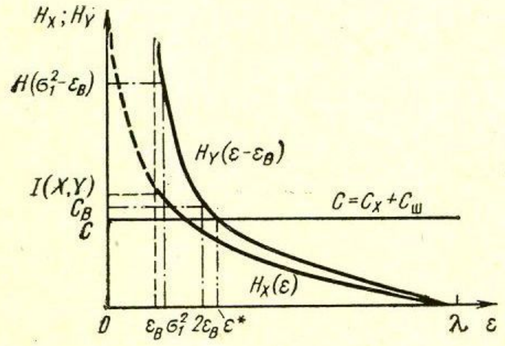


Рис. 2.1. Зависимость кодовой $H_Y(\varepsilon-\varepsilon_B)$ и сигнальной $H_X(\varepsilon)$-энтропий от ошибки гауссовского канала с пропускной способностью $C=C_x+C_ш$

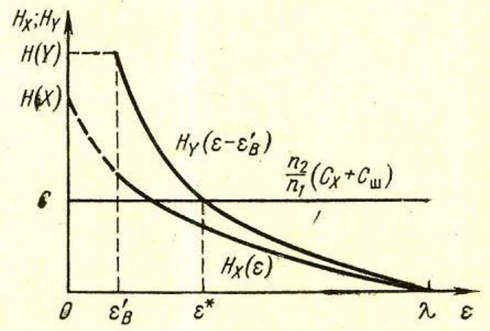


Рис. 2.2. Зависимость кодовой $H_Y(\varepsilon-\varepsilon'_B)$ и сигнальной $H_X(\varepsilon)$-энтропий от ошибки дискретного канала с пропускной способностью канала $\frac{n_2}{p}C_x+C_ш$

Эффект «раздвоения» ε-энтропии для гауссовского источника, выражаемый с помощью явных формул, наблюдается не только в гауссовском случае. Общность определений ε-энтропий (2.28) и (2.30) и равенства (2.25) позволяет утверждать, что *аналогичный эффект, вызываемый шумовой информацией, наблюдается всегда.* В качестве исходной кривой можно взять график, которым Шеннон характеризует дискретный источник [3, с. 593]. Так как мы применяем и в дискретном случае квадратический критерий точности, то можно воспользоваться дискретным аналогом равенства (2.26) с нижней границей ошибки:

$$\varepsilon'_B=d(X, \hat{X}),$$

где X, $\hat{X}$ — дискретные случайные величины.

На рис. 2.2 показан результат «раздвоения» графика Шеннона путем усечения участка в области $\varepsilon<\varepsilon'_B$ (нижняя кривая) и сдвига его на ε'_B (верхняя кривая). В отличие от непрерывного источника дискретный имеет всюду конечные значения кодовой и сигнальной ε-энтропий.

В этом разделе мы считали, что оптимальный код $\hat{Z}$, при котором реализуется кодовая ε-энтропия, существует. Теперь найдем этот код и докажем, что он является идеальным по Шеннону кодом.

2.3. ИДЕАЛЬНЫЙ КОД С УЧЕТОМ ШУМОВОЙ ИНФОРМАЦИИ (ОДНОМЕРНЫЙ СЛУЧАЙ)

Найдем оптимальный код $\hat{Z}$, который реализует в гауссовском случае кодовую ε-энтропию (2.29) и сигнальную (2.31). Применение равенства (2.25) позволяет строить оптимальный код по правилу

$$\hat{Z}=\gamma\hat{X}+W, \tag{2.36}$$

где γ — коэффициент одномерного фильтра; W — гауссовская случайная величина, не зависимая от случайных величин $\hat{X}$.

Дисперсии случайных величин, входящих в равенство (2.27), связаны между собой следующими очевидными соотношениями:

$$\mathbf{M}(\hat{Z}^2)=\mathbf{M}(\hat{X}^2)-\mathbf{M}(V^2)=\hat{\lambda}-\varepsilon_V;$$

$$\mathbf{M}(\hat{X}-\hat{Z})^2=\mathbf{M}(V^2)=\varepsilon_V.$$

Подставляя в эти соотношения $\hat{Z}$ в виде (2.36), получаем систему из двух уравнений для нахождения коэффициента γ и дисперсии σ^2_W величины W:

$$\gamma^2\hat{\lambda}+\sigma^2_W=\hat{\lambda}-\varepsilon_V;$$
$$(\gamma-1)\hat{\lambda}+\sigma^2_W=\varepsilon_V. \qquad (2.37)$$

Решения системы уравнений (2.37) дают параметры оптимального кода (2.36):

$$\gamma=1-\frac{\varepsilon_V}{\hat{\lambda}}=\frac{\mathbf{M}(\hat{Z}^2)}{\mathbf{M}(\hat{X}^2)};$$
$$\sigma^2_W=\gamma\varepsilon_V.$$

Полученные значения параметров показывают, что случайная величина W имеет дисперсию σ^2_W, определяемую величинуй γ и заданной ошибкой ε_V. Это мешает расшифровке линейной конструкции кода (2.36). Чтобы устранить затруднение, приведем преобразование (2.36) к следующему виду:

$$\hat{Z}=\frac{\sqrt{\gamma}}{a}(a\sqrt{\gamma}\hat{X}+\nu), \qquad (2.38)$$

где a — постоянная величина; $\nu=aW/\sqrt{\gamma}$ — случайная величина с дисперсией $\sigma^2_2=\mathbf{M}(\nu^2)=\frac{a^2\sigma^2_W}{\gamma}=a^2\varepsilon_V$.

Дисперсия $\sigma^2{}_2$ не зависит от параметра γ фильтра, что позволяет трактовать ν как шум в канале. Тогда код (2.38) характеризуется двумя каскадами необратимого линейного преобразования:

$$\hat{X}=\hat{l}Y=\hat{l}(\beta X+\xi);\ \hat{Z}=q(g\hat{X}+\nu), \qquad (2.39)$$

где $g=a\sqrt{\gamma}=a\sqrt{1-\sigma^2{}_2/a^2\hat{\lambda}}$; $q=\sqrt{\gamma}/a=g/a^2$.

В (2.39) оптимальный код совпадает с линейным кодированием (2.15) из 2.1, если число степеней свободы вырождается до единицы ($N=1$), т. е. в одномерном случае. Код (2.39) реализуется двузвенным кодером r, состоящим из фильтра Винера $\hat{l}$ и оптимального передаточного фильтра g (одномерного варианта фильтра (2.14)). «Энергия» усиления передаточного фильтра определяет его коэффициент усиления:

$$g^2=a^2\gamma=E_g.$$

Декодирующий фильтр в (2.39) совпадает с передаточным фильтром с точностью до коэффициента усиления.

Расшифровав конструкцию оптимального кода $\hat{Z}$, перейдем к его сопоставлению с идеальным по Шеннону кодом. Для этого надо сравнить кодовую ε-энтропию (2.29) источника с пропускной способностью канала. «Энергия» сигнала на входе канала в схеме на рис. 1.2 ограничена и равна

$$E_s=(\beta^2\lambda+\sigma^2{}_1)r^2=g^2\hat{\lambda},$$

где $\hat{\lambda}=(\beta^2\lambda+\sigma^2{}_1)\hat{l}^2$.

В рассматриваемом одномерном случае «энергия» сигнала на входе канала и «энергия» усиления фильтра g различаются только на фиксированную величину

$$E_s/E_g=\hat{\lambda}.$$

Это обстоятельство существенно облегчает доказательство идеальности кода (2.39).

Пропускная способность гауссовского одномерного канала (дискретного по времени) определяется формулой Шеннона:

$$C=\frac{1}{2}\log\left(1+\frac{E_s}{\sigma^2_2}\right)=\frac{1}{2}\log\left(1+\frac{g^2\hat{\lambda}}{\sigma^2_2}\right). \qquad (2.40)$$

Докажем, что кодовая ε-энтропия равна пропускной способности. Для этого преобразуем первое уравнение из системы (2.37) к виду:

$$\hat{\lambda}/\varepsilon_V=1+\gamma(\gamma\hat{\lambda}+\varepsilon_V)/\varepsilon_V.$$

Подставив в это равенство значение коэффициента γ из выражения (2.39) и перейдя от ошибки ε_V к дисперсии шума $\sigma^2{}_2$ канала, получим следующее равенство:

$$\hat{\lambda}/\varepsilon_V=1+g^2\hat{\lambda}/\sigma^2{}_2.$$

С учетом полученного равенства преобразуем формулу кодовой ε-энтропии (2.29) к виду (2.40):

$$H_Y(\varepsilon_V)=\frac{1}{2}\log(1+g^2\hat{\lambda}/\sigma^2{}_2) \qquad (2.41)$$

при $\varepsilon_V<\lambda$.

Из совпадения (2.40) и (2.41) следует, что выполняется условие Шеннона для идеального кода

$$H_Y(\varepsilon^*_V)=C. \qquad (2.42)$$

С помощью функции, обратной ε-энтропии (2.29), найдем явное решение уравнения:

$$\frac{\varepsilon^*_V}{\hat{\lambda}}=2^{-2C}. \qquad (2.43)$$

Формула (2.43) дает оптимальную минимальную ошибку передачи оценки $\hat{X}$ как экспоненциальную функцию пропускной спо-

собности канала. Эта ошибка может быть нулевой только при бесконечной пропускной способности:

$$\lim_{C\to\infty} \varepsilon^*_V = 0.$$

Полная ошибка ε^* при идеальном кодировании складывается из ошибки фильтра B и ошибки передаточного фильтра:

$$\varepsilon^* = \varepsilon_B + \varepsilon^*_V.$$

Решение уравнения (2.42) графически показано на рис. 2.1 точкой пересечения верхней кривой с прямой, параллельной оси абсцисс. Уравнение идеального линейного кодирования имеет решение, определяющее минимальное значение второго члена в сумме (2.26):

$$\varepsilon^*_V = \min_{\{Z\}} d(\hat{X}, Z) = d(\hat{X}, \hat{Z}) = \sigma_2^2/a^2.$$

С такой минимальной среднеквадратической ошибкой ε^*_V идеальный код $\hat{Z}$ воспроизводит оценку $\hat{X}$ для получателя информации.

Обратим внимание на то, что найденный идеальный код (2.39) является оптимальным решением трех вариационных задач: минимизации среднеквадратической ошибки восстановления полезного сигнала, вычисления кодовой ε-энтропии источника и вычисления пропускной способности:

$$\min d(X, Z) = d(X, \hat{X}) + d(\hat{X}, \hat{Z}) = \varepsilon_B + \sigma^2{}_2/a^2;$$
$$\min I(Y, Z) = I(\hat{X}, \hat{Z}) = H_Y(\varepsilon^*_V);$$
$$\max I(Y, Z) = I(\hat{X}, \hat{Z}) = C.$$

Иначе говоря, указанные три вариационные задачи (одна из теории фильтрации и две из теории информации) имеют совпадающие решения в одномерном гауссовском случае. В этом причина того, что мы весьма просто нашли идеальный код.

Перейдем теперь к описанию свойств идеального кодирования с учетом влияния шумовой информации. Их можно описать с помощью следующих информационных равенств:

$$I(\hat{X}, \hat{Z}) = I(X, \hat{Z}) + I(\xi, \hat{\mathcal{U}});$$
$$I(X, \hat{X}) = I(X, \hat{Z}) + \Delta I. \qquad (2.44)$$

Первое равенство (2.44) показывает, что сигнальная и шумовая информации создают единый поток кодовой информации $I(\hat{X}, \hat{Z})$, на передачу которой полностью расходуется пропускная способность канала в соответствии с уравнением Шеннона (2.42). Поэтому пропускную способность (в двоичных единицах) можно рассматривать как величину, состоящую из сигнальной части C_X и шумовой $C_ш$:

$$C = C_X + C_ш, \qquad (2.45)$$

где $C_X = I(X, \hat{Z})$; $C_ш = I(\xi, \hat{\mathcal{U}})$.

Представление о двух разных слагаемых пропускной способности раньше отсутствовало в теории информации, так как модель системы связи Шеннона на рис. В.1 не включала шума источника. Такое представление (2.45) не нарушает утверждения о том, что приспособленность канала к передаче информации характеризуется единственным числом C, а лишь учитывает специфику информации, поступающей от зашумленного источника.

Сравним равенство (2.26), обеспечивающее минимум среднеквадратической ошибки, с равенством (2.44), описывающим информационное свойство идеального кода:

$$d(X, \hat{Z}) = d(X, \hat{X}) + d(\hat{X}, Z) = \varepsilon_B + \varepsilon^*_V;$$
$$I(\hat{X}, \hat{Z}) = I(X, \hat{X}) + [-\Delta I + I(\xi, \hat{\mathcal{U}})]. \qquad (2.46)$$

Сопоставим друг другу первые и вторые члены выражений (2.46). Первые члены $d(X, \hat{X})$ и $I(X, \hat{X})$ отражают потенциальные возможности в получении информации о полезном сигнале X, содержащейся в оценке $\hat{X}$. Однако они не могут быть доставлены получателю из-за необходимости передачи оценки $\hat{X}$ по зашумленному каналу. В результате возникают ошибки передачи, описываемые вторыми членами выражений (2.46).

В первое равенство (2.46) входит «энергетическая» ошибка передачи оценки $\hat{X}$ — среднеквадратическая ошибка ε^*_V, которая добавляется к ошибке ε_B одномерного аналога фильтра Винера. Второе равенство (2.46) содержит две информационные ошибки, которые являются потерей информации ΔI и шумовой информацией. Полная информационная оценка ошибки передачи складывается из потери сигнальной информации и из шумовой информации:

$$\Delta I + I(\xi, \hat{\mathcal{U}}).$$

Потеря сигнальной информации согласно формуле (2.34) родственна «энергетической» ошибке:

$$\Delta I = \frac{1}{2} \log\left(1 + \frac{\varepsilon^*_V}{\varepsilon_B}\right) = \frac{1}{2} \log\left(1 + \frac{\sigma_2^2}{a^2\,\varepsilon_B}\right).$$

Такая потеря монотонно падает с уменьшением «энергетической» ошибки передачи:

$$\lim_{\varepsilon^*_V \to 0} \Delta I = 0.$$

Шумовая информация в соответствии с выражением (2.45) равна числу двоичных единиц $C_ш$ пропускной способности, расходуемой на передачу шума источника по каналу. Шумовая информация является не просто бесполезной, но и вредной для получателя, принимающего по информации то или иное решение. Поэтому возникает естественное беспокойство о соотношении сигнальной C_X и шумовой $C_ш$ составляющих пропускной способности. Обратимся к характеристикам на рис. 2.1 и решим вариаци-

онную задачу по нахождению минимума полной информационной ошибки при вариации энергетической ошибки:

$$\min_{\{\varepsilon^*_V\}} [\Delta I + I(\xi, \hat{\mathcal{U}})].$$

Используя формулы (2.34) и (2.35), легко найти, что искомый минимум достигается, когда ошибка передачи оценки X равна ошибке фильтра Винера для сигнала Y:

$$\varepsilon^*_V = \varepsilon_B \text{ или } \varepsilon^* = 2\varepsilon_B. \tag{2.47}$$

Подставив найденное значение ошибки (2.47) в (2.34) и (2.35), найдем, что минимальная сумма потерь информации и шумовой информации составит $(1+l_B)/2 \approx 1$ бит. Для передачи сигнала X с ошибкой (2.47) необходима пропускная способность, соответствующая уравнению (2.42):

$$C_B = H_Y(\varepsilon^*_V = \varepsilon_B) = \frac{1}{2}\log\frac{\hat{\lambda}}{\varepsilon_B} \leqslant \frac{1}{2}\left(\frac{\hat{\lambda}}{\varepsilon_B} - 1\right)\log e.$$

Найденное условие (2.47) позволяет разбить диапазон задания кодовой и сигнальной ε-энтропий, представленных кривыми на рис. 2.1, на два участка, разделяемых точкой $2\varepsilon_B$ на оси абсцисс:

$$\varepsilon_B < \varepsilon < 2\varepsilon_B;\ 2\varepsilon_B \leqslant \varepsilon < \lambda.$$

Оценим на этих участках отношение $C_X/C_ш$, которое характеризует расхождение двух кривых на рис. 2.1. С помощью (2.31) и (2.35) легко показать, что это отношение на верхней границе равно энергетическому отношению сигнал-шум:

$$\lim_{\varepsilon \to \lambda} I(X, \hat{Z})/I(\xi, \hat{\mathcal{U}}) = \lambda/\sigma^2_1.$$

Так как обычно энергетическое отношение сигнал-шум велико, это означает, что две кривые на рис. 2.1 близки. Расхождение кривых несколько увеличивается при переходе к точке $\varepsilon = 2\varepsilon_B$. Можно дать следующую приближенную оценку:

$$\frac{I(X, \hat{Z})}{I(\xi, \hat{\mathcal{U}})} \approx \frac{1}{2}\frac{\lambda}{\sigma^2_2} \qquad \text{при } \varepsilon = 2\varepsilon_B.$$

Полученные оценки показывают, что на участке $2\varepsilon_B \leqslant \varepsilon \leqslant \lambda$ отношение $C_X/C_ш$ достаточно велико, а кривые на рис. 2.1 расходятся слабо. Это означает, что на этом участке на передачу шумовой информации пропускная способность канала практически не расходуется.

Совсем иная картина наблюдается на узком участке, примыкающем к значению ε_B ошибки одномерного фильтра Винера для зашумленного сигнала Y. Достижение чрезмерной точности в передаче непрерывного сигнала X, т. е. получение слишком малых потери информации и величины ошибки ($\varepsilon < 2\varepsilon_B$), наталкивается на принципиальную трудность — интенсивный рост расхода пропускной способности на передачу шумовой информации. Кривые на рис. 2.1 иллюстрируют интенсивный рост шумовой информации (см. расхождение кривых) на узком участке диапазона значений ошибок в окрестности ($\varepsilon_B < \varepsilon < 2\varepsilon_B$). Для этого участка при $\lambda \gg \sigma^2_1$ формулы (2.34) и (2.35) дают *информационное соотношение неопределенности:*

$$\Delta I \cdot 2^{2I(\xi, \hat{\mathcal{U}})} \approx 1;\ \text{в частном случае}$$

$$\lim_{\lambda/\sigma^2_1 \to \infty} \lim_{\varepsilon \to \varepsilon_B} \Delta I \cdot 2^{2I(\xi, \hat{\mathcal{U}})} = 1/\ln 4.$$

Передача сигнала X с ошибкой ε_B или передача оценки $\hat{X}$ с нулевой ошибкой невозможна, так как потребовала бы бесконечной пропускной способности из-за шумовой информации:

$$\lim_{\varepsilon \to \varepsilon_B} \hat{Z} = \hat{X};\ \lim_{\varepsilon \to \varepsilon_B} C = \infty;\ \lim_{\varepsilon \to \varepsilon_B} I(\xi, \hat{\mathcal{U}}) = \infty.$$

Нарисованная выше картина влияния шумовой информации на пропускную способность канала в двухкаскадной одномерной схеме на рис. 1.2 при идеальном кодировании была получена нами для гауссовского случая. Но в основных своих чертах она может быть распространена и на негауссовский дискретный случай, когда используется квадратический критерий. Обратимся к кривым на рис. 2.2. Дискретный зашумленный источник имеет энтропию полезного сигнала $H(X)$. Кодовая ε-энтропия при нижней границе ошибки ε'_B конечна и равна энтропии дискретного сигнала Y:

$$H_Y(0) = H(Y).$$

Уравнение идеального кодирования для негауссовского случая с учетом шумовой информации приобретает следующий вид:

$$H_Y(\varepsilon^*_V) = \frac{n_2}{n_1}(C_X + C_ш),$$

где $C = C_X + C_ш$.

Решение уравнения графически показано на рис. 2.2 пересечением кривой кодовой ε-энтропии и прямой с ординатой $\frac{n_2}{n_1}C$. Идеальное кодирование позволяет передать дискретную оценку $\hat{X}$ с нулевой ошибкой $\varepsilon^*_V = 0$, если выполняется условие

$$H(Y) = C_0 n_2/n_1.$$

Канал с пропускной способностью C_0 может быть загружен в основном передачей шумовой информации. Отношение $C_ш/C_0$ при этом будет равно (см. рис. 2.2)

$$\frac{C_ш}{C_0} = \frac{H(Y) - [H(X) - \Delta I]}{H(Y)}.$$

Не располагая явными формулами для ε-энтропий, трудно дать аналитическую оценку отношения $C_ш/C_0$. Но очевидно, что это

отношение зависит от отношения энтропии чистого полезного сигнала $H(X)$ и энтропии оптимальной оценки зашумленного сигнала $H(\hat{X})$.

При рассмотрении кодирования дискретных сигналов Шеннон ввел понятие об относительной энтропии дискретного источника, определив ее как отношение энтропии источника к максимальному значению, которого могла бы достичь энтропия при тех же символах. Говоря об относительной энтропии Шеннон отмечает: «...она является максимальным сжатием, которое можно осуществить кодированием при помощи того же самого алфавита. Единица минус относительная энтропия есть избыточность» [3, с. 267].

Распространим понятие о коэффициенте сжатия или избыточности на дискретные зашумленные источники. Для этого обратимся к графикам на рис. 2.2, иллюстрирующим «раздвоение» ε-энтропии источника, который создает дискретный зашумленный сигнал Y. Кодовая ε-энтропия достигает конечного наибольшего значения, равного энтропии $H(Y)$ при ошибке ε'_B:

$$0 \leqslant H_Y(\varepsilon - \varepsilon'_B) \leqslant H_Y(0) = H(Y).$$

Максимальное значение, которое могла бы достичь энтропия $H(Y)$ при тех же символах в отсутствии между ними статистических связей, обозначим $H_{\text{макс}}(Y)$. Кодовая ε-энтропия источника при идеальном кодировании равна $H_Y(\varepsilon^* - \varepsilon'_B)$. Следуя Шеннону, определим коэффициент максимального сжатия кодовой информации при идеальном коде как отношение

$$A = H_{\text{макс}}(Y) / H_Y(\varepsilon^* - \varepsilon'_B). \qquad (2.48)$$

Вместо коэффициента сжатия можно использовать избыточность:

$$1 - 1/A.$$

Распространение определения коэффициента сжатия (2.48) с дискретного на непрерывный источник, ε-энтропия которого дана на рис. 2.1, встречает затруднение. Оно вызвано неограниченным ростом кодовой информации при ошибке $\varepsilon = \varepsilon_B$. Однако обычно проектировщик, стремясь к экономии пропускной способности, допускает не слишком малую ошибку $\varepsilon > 2\varepsilon_B$, при которой отношение $C_X/C_{\text{ш}}$ близко к отношению сигнал-шум $\beta^2\lambda/\sigma^2_1$. Тогда кодовую ε-энтропию $I(Y, \hat{Z})$ можно считать конечной. Ограничим кодовую информацию, присущую не идеальному линейному коду, некоторым максимальным значением

$$I(Y, Z) \leqslant I_{\text{макс}}(Y, Z).$$

При сделанных допущениях определение коэффициента сжатия (2.48) можно распространить и на идеальное линейное кодирование:

$$A = I_{\text{макс}}(Y, Z) / I(Y, \hat{Z}).$$

В данном разделе мы относительно просто для одномерного гауссовского случая нашли решения двух задач, поставленных в 1.4: учли влияние шума источника и отыскали идеальный линейный код. На первый взгляд кажется, что эти результаты легко, с помощью разложения Карунена — Лоэва (1.1), распространить и на многомерный случай, когда в схеме на рис. 1.2 действуют гауссовские процессы. Однако, как мы увидим ниже, здесь возникают трудности, делающие такое обобщение проблемным.

2.4. СИГНАЛЬНАЯ И ШУМОВАЯ ИНФОРМАЦИЯ ГАУССОВСКИХ ПРОЦЕССОВ

Обобщим результаты, изложенные в разделах 2.2 и 2.3, применительно к одномерному варианту схемы на рис. 1.2, на случай, когда в схеме сигналами являются не случайные величины X, Y, Z, а гауссовские случайные процессы $X(t)$, $Y(t)$, $Z(t)$ непрерывного времени t. Методика обобщения использует статистическую независимость коэффициентов разложения Карунена — Лоэва. Это позволяет источник гауссовского процесса заменить «произведением» множества одномерных источников. Гауссовский канал, непрерывный по времени, на рис. 1.2 «распараллеливается» в множество одномерных каналов. И тем не менее многомерная картина лишь в общих чертах повторяет одномерную. В многомерном случае возникают новые трудности, вызванные необходимостью учета распределения ошибки по многим степеням свободы.

Воспользуемся разложением Карунена — Лоэва (1.1) и спектральным представлением фильтров (1.3). Источник в схеме на рис. 1.2 порождает зашумленный сигнал

$$Y(t) = \sum_{k=1}^{\infty} Y_k \varphi_k(t) = \sum^{\infty} \beta_k X_k \varphi_k(t) + \sum_{k=1}^{\infty} \xi_k \varphi_k(t).$$

Коэффициенты ряда X_k независимы в совокупности с нулевыми математическими ожиданиями и дисперсиями λ_k. Коэффициенты разложения белого шума также независимы с дисперсиями σ^2_1 и, следовательно, коэффициенты Y_k также независимы и равны $Y_k = \beta_k X_k + \xi_k$. Код $Z(t)$ ищется в виде

$$Z(t) = \sum_{k=1}^{\infty} Z_k \varphi_k(t).$$

Средняя взаимная информация между процессами $Y(t)$ и $Z(t)$, которую назовем кодовой, в силу независимости коэффициентов равна сумме взаимных информаций между коэффициентами

$$I[Y(t),\ Z(t)] = \sum_{k=1}^{\infty} I(Y_k Z_k); \qquad (2.49)$$

$$I(Y_k, Z_k) = \frac{1}{2} \log \left[1 + \frac{r_k^2 \left(\beta_k^2 \lambda_k + \sigma_1^2 \right)}{\sigma_2^2} \right],$$

где β_k, r_k — коэффициенты фильтров (1.3).

Выходной сигнал в схеме на рис. 1.2 содержит как полезную составляющую, так и шумовую. Шум на выходе равен

$$\mathcal{U}(t) = Z(t) \text{ при } X(t) = 0.$$

В разделе 2.2 была введена и вычислена шумовая информация для каждого коэффициента в отдельности (2.19). В силу независимости коэффициентов X_k, Y_k, Z_k шумовая информация для непрерывных процессов будет равна сумме шумовых информаций для каждого коэффициента:

$$I[\xi(t), \mathcal{U}(t)] = \sum_{k=1}^{\infty} I(\xi_k, \mathcal{U}_k);$$

$$I(\xi_k, \mathcal{U}_k) = \frac{1}{2} \log\left(1 + \frac{r_k^2 \sigma_1^2}{\sigma_2^2}\right). \tag{2.50}$$

Сигнальная информация будет равна

$$I[X(t), Z(t)] = \sum_{k=1}^{\infty} I(X_k, Z_k); \tag{2.51}$$

$$I(X_k, Z_k) = \frac{1}{2} \log\left(1 + \frac{r_k^2 \lambda_k \beta_k^2}{r_k^2 \sigma_1^2 + \sigma_2^2}\right).$$

Аналогично равенству (2.20) *кодовая информация равна сумме сигнальной и шумовой:*

$$I[Y(t), Z(t)] = I[X(t), Z(t)] + I[\xi(t), \mathcal{U}(t)]. \tag{2.52}$$

Расстояние между процессами $X(t)$, $Y(t)$ и $Z(t)$ с учетом (2.25) равно:

$$d[X(t), Z(t)] = \sum_{k=1}^{\infty} [d(X_k, \hat{X}_k) + d(\hat{X}_k, Z_k)];$$

$$d[Y(t), Z(t)] = \sum_{k=1}^{\infty} d(Y_k, Z_k).$$

Для каждой тройки коэффициентов X_k, Y_k, Z_k можно найти, пользуясь методом, изложенным в 2.2, кодовую ε-энтропию и сигнальную ε-энтропию. Для экономной передачи процесса $X(t)$ через зашумленный канал необходимо кодирование источника, дающее минимально возможное количество информации между процессами $Y(t)$ и $Z(t)$ и обеспечивающее восстановление сигнала $X(t)$ с ошибкой ε. При этом ошибка восстановления каждого коэффициента равна ε_k, а суммарная ошибка равна ε:

$$\varepsilon = \sum_{k=1}^{\infty} \varepsilon_k.$$

Необходимо найти оптимальный код $\hat{Z}(t)$, который обеспечивает минимум кодовой информации

$$I[Y(t), \hat{Z}(t)] = \min_{\{\varepsilon_k\}} \sum_{k=1}^{\infty} I(Y_k, Z_k) \tag{2.53}$$

при ограничении

$$\sum_{k=1}^{\infty} \varepsilon_k = \varepsilon.$$

Отличие от одномерного случая заключается не только в нахождении кодовой ε-энтропии как суммы элементарных ε_k-энтропий, но и в нахождении распределения ошибки ε_k по коэффициентам спектра. Поэтому минимум в выражении (2.53) находится путем вариации ошибок ε_k при условии постоянства их суммы ε. Эту задачу удобно решить с помощью теоремы Куна — Таккера, согласно которой выполняется равенство

$$\frac{\partial}{\partial \varepsilon_k} I[Y(t), Z(t)] = -\frac{1}{\mu},$$

где μ — множитель Лагранжа.

Множитель Лагранжа μ находится из условия ограничения среднеквадратической ошибки:

$$\sum_{k=1}^{N} (\mu + \varepsilon_{Bk}) + \sum_{k=N+1}^{\infty} \hat{\lambda}_k = \varepsilon,$$

или

$$\varepsilon_B + N\mu + \sum_{k=N+1}^{\infty} \hat{\lambda}_k = \varepsilon,$$

где $\varepsilon_B = \sum_{k=1}^{\infty} \varepsilon_{Bk} = \sum_{k=1}^{\infty} \frac{\lambda_k \sigma_1^2}{\beta_k^2 \lambda_k + \sigma_1^2}$ — ошибка фильтра Винера;

$\hat{\lambda}_k = \lambda_k - \varepsilon_{Bk}$.

Минимум в выражении (2.35) достигается при следующем распределении ошибок:

$$\varepsilon_k - \varepsilon_{Bk} = \begin{cases} \mu & \text{при } \varepsilon_k - \varepsilon_{Bk} < \hat{\lambda}_k;\ k < N, \\ \hat{\lambda}_k & \text{при } \varepsilon_k - \varepsilon_{Bk} \geqslant \hat{\lambda}_k;\ k \geqslant N. \end{cases}$$

Суммарная ошибка при оптимальном распределении ее спектральных составляющих превышает ошибку фильтра Винера на величину

$$\varepsilon_V = \varepsilon - \varepsilon_B = \sum_{k=1}^{\infty} \varepsilon_{Vk} = N\mu + \sum_{k=N+1}^{\infty} \hat{\lambda}_k, \tag{2.54}$$

где $\varepsilon_{Vk} = \varepsilon_k - \varepsilon_{Bk}$.

Подчеркнем, что первый член ошибки (2.54) имеет выбеленный спектр по N степеням свободы. Второй член дает ошибку усечения разложения Карунена — Лоэва оценки $\hat{X}(t)$. Подставив найденный оптимальный спектр ошибки в формулу многомерной кодовой ε-энтропии, получим обобщенную формулу Колмогорова:

$$I[Y(t), \hat{Z}(t)] = H_Y(\varepsilon_V) = \sum_{k=1}^{\infty} H(\varepsilon_{Vk}) = \frac{1}{2} \sum_{k=1}^{N} \log \frac{\hat{\lambda}_k}{\mu}. \tag{2.55}$$

Обобщенная формула Колмогорова (2.55) сводит кодовую ε-энтропию источника гауссовского процесса к множеству одномерных источников с элементарными одномерными ε_{Vk}-энтропиями. Это множество источников всегда является конечномерным, так как усекаются элементарные источники, для которых ошибка восстановления превышает собственную дисперсию:

$$H_Y(\varepsilon_{Vk})=0 \quad \text{при } \varepsilon_{Vk}\geqslant\hat{\lambda}_k;\ k\geqslant N.$$

Число степеней свободы N находится из уравнения:

$$\varepsilon_{Bk}+\mu=\lambda_k \quad \text{при } k=N,$$

$$\text{или } \varepsilon_{Vk}=\hat{\lambda}_k \quad \text{при } k=N.$$

Обобщенная формула Колмогорова отличается от формулы, приведенной в работе [4, с. 48] заменой дисперсий λ_k незашумленного сигнала $X(t)$ на дисперсии $\hat{\lambda}_k$ оценки $\hat{X}(t)$. Необходимость такой замены при вычислении в гауссовском случае ε-энтропии зашумленного источника была отмечена в статье Р. Л. Добрушина и Б. С. Цыбакова [19].

Аналогичным способом обобщим на многомерный случай формулу для вычисления сигнальной ε-энтропии (2.31). Для этого решим вариационную задачу:

$$I[X(t),\ Z^*(t)]=\min_{\sum_{k=1}^{\infty}\varepsilon_k=\varepsilon}\sum_{k=1}^{\infty}I(X_k,\ Z_k). \tag{2.56}$$

Вариационную задачу (2.56) решим также с помощью теоремы Куна — Таккера, согласно которой должно выполняться равенство:

$$\frac{\partial}{\partial\varepsilon_k}I[X(t),\ Z(t)]=-\frac{1}{\Theta_4},$$

где Θ_4 — множитель Лагранжа.

Минимум (2.56) будет достигаться при следующем спектральном распределении ошибки [22]:

$$\varepsilon_k=\begin{cases}\varepsilon_{Bk} & \text{при } \varepsilon_{Bk}\geqslant\Theta_4,\\ \Theta_4 & \text{при } \varepsilon_{Bk}<\Theta_4,\\ \lambda_k & \text{при } \Theta_4>\lambda_k.\end{cases}$$

Множитель Лагранжа находится из уравнения

$$\sum_{k=1}^{N_1}\varepsilon_{Bk}+N_2\Theta_4+\sum_{k=N_2+1}^{\infty}\lambda_k=\varepsilon.$$

Числа N_1 и N_2 находятся из соотношений:

$$\varepsilon_{Bk}=\Theta_4 \quad \text{при } k=N_1;$$

$$\Theta_4=\lambda_k \quad \text{при } k=N_2.$$

Подставив значения для ε_k в выражение (2.56), получим формулу

$$I[X(t),Z^*(t)]=\frac{1}{2}\sum_{k=1}^{N_1}\log\frac{\lambda_k}{\varepsilon_{Bk}}+\frac{1}{2}\sum_{k=N_1+1}^{N_2}\log\frac{\lambda_k}{\Theta_4}. \tag{2.57}$$

Формула сигнальной ε-энтропии (2.57) не является обобщением прежней формулы (2.31), когда коды Z^* и $\hat{Z}$ совпали в одномерном случае. Отличие многомерной сигнальной ε-энтропии от одномерной заключается в том, что минимумы (2.53) и (2.56) достигаются при разных кодах $\hat{Z}(t)$ и $Z^*(t)$. Специфика формулы (2.57) выражается также и в том, что реализация первых ее членов N_1 требует передачи коэффициентов X_k с ошибкой фильтра Винера:

$$I(X_k,\ \hat{X}_k)=\frac{1}{2}\log\frac{\lambda_k}{\varepsilon_{Bk}};\quad k\in[1,\ N_1].$$

Как было показано в 2.3, такая передача требует бесконечной пропускной способности канала и поэтому ε-энтропия (2.57) гауссовского сигнального процесса неосуществима. Минимум (2.56) не достигается при оптимальном коде $\hat{Z}(t)$, так как приходится тратить большее число двоичных единиц для получения заданного значения ошибки ε:

$$I[X(t),\ \hat{Z}(t)]>I[X(t),\ Z^*(t)].$$

Вместо сигнальной ε-энтропии в многомерном случае будем пользоваться понятием о сигнальной информации $I[X(t),\ \hat{Z}(t)]$, содержащейся в оптимальном коде. Именно эта информация в сумме с шумовой информацией, как следует из (2.52), создает ε-энтропию при оптимальном коде $\hat{Z}(t)$:

$$I[Y(t),\ \hat{Z}(t)]=I[X(t),\ \hat{Z}(t)]+I[\xi(t),\ \hat{\mathcal{U}}(t)],$$

где $\hat{\mathcal{U}}(t)=\hat{Z}(t)$ при $X(t)=0$; $\hat{\mathcal{U}}(t)$ — выходной шум при оптимальном коде.

В 2.1 мы рассмотрели винеровские ошибки, возникающие в однокаскадной схеме на рис. 1.3 и в двухкаскадной схеме на рис. 1.2. Было сформулировано правило покаскадного суммирования винеровских ошибок. Хорошо известно, что для схемы на рис. 1.3 с фильтром $G=1$ еще Винером была дана формула вычисления взаимной информации между гауссовским процессом $Y(t)$ и сигналом $X(t)$ [50]:

$$I[X(t),\ Y(t)]=\frac{1}{2}\sum_{k=1}^{\infty}\left(1+\frac{\lambda_k}{\sigma_1^2}\right).$$

Формула Винера наглядно проиллюстрировала зависимость сигнальной информации от спектрального отношения сигнал-шум λ_k/σ^2_1. Нам же желательно выявить влияние винеровской ошибки ε_B, с которой условное математическое ожидание $\hat{X}(t)$ в схеме на

рис. 1.3 оценивает сигнал $X(t)$. Для этого достаточно переписать формулу Винера, вычленив в ней винеровскую ошибку:

$$I[X(t),\ Y(t)] = I[X(t),\ \hat{X}(t)] = \frac{1}{2}\sum_{k=1}^{\infty}\log\frac{\lambda_k}{\varepsilon_{Bk}}.$$

Оптимальная оценка $\hat{X}(t)$ содержит ту же полезную информацию, что и смесь $Y(t)$, но в более «сжатом» виде, так как дает меньшую ошибку воспроизведения входного сигнала $X(t)$:

$$\mathbf{M}[X(t)-\hat{X}(t)]^2 \leqslant \mathbf{M}[X(t)-Y(t)]^2.$$

Очевидно, что вышеприведенная формула для $I[X(t),\ Y(t)]$ будет справедлива не только для частного случая фильтра $G=1$, но и при любом заданном передаточном фильтре G. Более того, правило покаскадного сложения винеровских ошибок позволяет распространить эту формулу и на многокаскадные схемы (см. рис. 1.2). Подставив винеровскую ошибку (2.11) в формулу (2.51), получим

$$I[X(t),\ \tilde{Y}(t)] = I[X(t),\ \tilde{X}(t)] = \frac{1}{2}\sum_{k=1}^{\infty}\log\frac{\lambda_k}{\varepsilon_{Bk}+d(\hat{X}_k,\ \tilde{X}_k)}.$$

Видно, что формулы для сигнальной информации в схемах на рис. 1.3, 1.2 аналогичны: в них входят спектры сигнала λ_k и винеровских ошибок. Очевидно, что увеличение каскадов, осуществляющих необратимое линейное преобразование, ведет к неустранимой потере сигнальной информации. Потеря сигнальной информации в двух каскадах схемы на рис. 1.2:

$$\Delta I = I[X(t),\ \hat{X}(t)] - I[X(t),\ \tilde{X}(t)] = \frac{1}{2}\sum_{k=1}^{\infty}\log\left[1+\frac{d(\hat{X}_k,\ \tilde{X}_k)}{d(X_k,\ \hat{X}_k)}\right].$$

Полученная формула показывает, что потеря информации определяется отношением спектров винеровской ошибки для двух каскадов к ошибке ε_{Bk} для первого каскада. Очевидно, что формула справедлива и для случая многих каскадов. Возрастание потери сигнальной информации с ростом числа каскадов отражает сложение винеровских ошибок каскадов. Однако покаскадное увеличение потери сигнальной информации подчинено более сложному правилу (см. 3.8).

С помощью оптимизации кодирующего фильтра R во втором каскаде нельзя уменьшить потерю информации или винеровскую ошибку, присущую первому каскаду. Можно лишь уменьшить потерю сигнальной информации во втором каскаде или уменьшить его винеровскую ошибку до значения (2.16). При этом потеря сигнальной информации в двух каскадах снизится до величины

$$\Delta I_{\text{мин}} = I[X(t),\ \hat{X}(t)] - I[X(t),\ \hat{Z}(t)] = \frac{1}{2}\sum_{k=1}^{\infty}\log\left[1+\frac{d(\hat{X_k},\ \hat{Z_k})}{B}\right].$$

В 2.4 было показано, что зашумленный источник создает два потока информации: сигнальную информацию и шумовую. Шумовая информация определяется по формуле (2.50). Преобразуем эту формулу, применив понятие о спектральном коэффициенте шума к линейному кодеру R с каналом:

$$b_k = \frac{r_k^2\sigma_1^2+\sigma_2^2}{r_k^2\sigma_1^2}.$$

Подставив это выражение в формулу (2.50), найдем зависимость шумовой информации от спектрального коэффициента шума:

$$I[\xi(t),\ \mathcal{U}(t)] = \frac{1}{2}\sum_{k=1}^{\infty}\log\frac{b_k}{b_k-1}.$$

При снижении спектрального коэффициента шума до единицы ($b_k \to 1$) шумовая информация стремится к бесконечности. Такая картина справедлива не только в гауссовском случае, но и для способов дискретной передачи с квадратическим критерием точности передачи, для которых справедлив рис. 2.2. Высказывания по этому вопросу, изложенные в 2.3, полностью распространимы и на многомерный случай.

2.5. ТЕОРЕМА ОБ ИДЕАЛЬНОМ ЛИНЕЙНОМ КОДИРОВАНИИ

В 2.3 мы уже отмечали, что совпадение решений трех вариационных задач в одномерном гауссовском случае позволило весьма просто доказать, что оптимальный код $\hat{Z}$ (2.39) является идеальным по Шеннону кодом. Распространение этого утверждения на многомерный случай сталкивается с серьезными трудностями, которые превратили этот вопрос в дискуссионный.

Главным был вопрос о том, может ли оптимальное линейное кодирование быть абсолютно оптимальным, т. е. удовлетворять условию идеального кодирования (1.17). Шеннон, приведя доказательство теоремы об идеальном кодировании, подчеркнул двойственность источника и канала [3, с. 607]. Она проявляется, с одной стороны, в том, что вычисление пропускной способности C канала соответствует нахождению источника, согласованного с каналом. С другой стороны, вычисление ε-энтропии $H(\varepsilon)$ источника соответствует нахождению канала, согласованного с источником и заданным уровнем искажения. Развивая эту мысль, Р. Галлагер ввел так называемый «тест-канал» (термин автора), с помощью которого вычисляется ε-энтропия источника, порождающего гауссовский процесс [18]. Естественно, что в вычислениях шум источника, как не входящий в классическую схему Шеннона (см. рис. В.1), не учитывался. Структура тест-канала Галлагера пояснена в книге схемой, о которой он пишет: «В отличие от дискретного по времени гауссовского источника тест-канал для источника, порождающего гауссовский случайный процесс, в общем случае не используется источником со скоростью, равной

пропускной способности (при этом принимаются соответствующие мощности ограничения на входе канала)» [18]. Процитированная фраза показывает, что Галлагер сравнивал ε-энтропию гауссовского источника с пропускной способностью (1.18), вычисленной Шенноном при ограничении «мощности» или «энергии» E_s. Как показано в 1.2, при ограничении «энергии» E_s сигнала оптимальным линейным кодером служит фильтр Костаса (1.12). В работе [16] с учетом [18] для оптимального линейного кодирования Костаса (с учетом шума источника) показано, что ε-энтропия источника меньше пропускной способности, т. е. уравнение (1.17) не выполняется. В создавшемся положении, казалось бы, осталось лишь одно — признать недосягаемость идеального по Шеннону кода для всех классов кодирования с помощью линейных фильтров в гауссовском случае. Это мнение стало повсеместным и подкреплялось высказыванием о том, что линейное кодирование всегда хуже нелинейного кодирования Шеннона.

Иной подход к проблеме поиска идеального линейного кодирования был дан в [22—24]. Он базировался на утверждении о существовании идеального линейного кода в многомерном случае в виде следующей теоремы.

Теорема. В гауссовском случае линейное кодирование, осуществляемое кодирующим фильтром, состоящим из передаточного фильтра с фильтром Винера, и декодирующим фильтром Винера, обеспечивает передачу по каналу с заданной пропускной способностью C информации с минимальной ошибкой, являющейся решением уравнения Шеннона для идеального кодирования при $n_1=n_2$.

Доказательство. Теорема утверждает не только существование идеального по Шеннону линейного кода при $n_1=n_2$, но и дает конструктивный способ его осуществления с помощью линейных фильтров, правда, только в частном гауссовском случае. Напомним уравнение идеального кодирования Шеннона (1.17) при одинаковой длине кодов источника и канала, т. е. при $n_1=n_2$:

$$H(\varepsilon^*_V)=C;\ \varepsilon^*_V=H^{-1}(C).$$

Нам надо доказать, что оптимальное линейное кодирование, описанное в 2.1, является абсолютно оптимальным, т. е. удовлетворяет уравнению (1.17). В этой связи можно говорить не только о двойственности источника и канала, которую отмечал Шеннон [3, с. 607], но и «тройственности» задач нахождения ε-энтропии, пропускной способности и минимума ошибки (максимума точности передачи). Суть доказательства теоремы будет заключаться в установлении факта

$$\hat{Z}_k=\gamma_k\hat{X}_k+W_k, \qquad (2.58)$$

где γ_k — коэффициент линейного фильтра; W_k — коэффициент гауссовского процесса $W(t)$ с дисперсией σ^2_{Wk}.

Соотношение (2.27) в применении к гауссовским процессам $\hat{X}(t)$ и $\hat{Z}(t)$ принимает вид

$$\hat{X}_k=\hat{Z}_k+V_k.$$

Гауссовский процесс $V(t)$, не зависящий от процесса $\hat{X}(t)$, выражает ошибку передачи оценки $\hat{X}(t)$ кодом $\hat{Z}(t)$:

$$\varepsilon_V=\mathbf{M}[V^2(t)]=\mathbf{M}[\hat{X}(t)-\hat{Z}(t)]^2=\sum_{k=1}^{\infty}\mathbf{M}(V^2_k).$$

Применив систему уравнений (2.37) к каждой степени свободы, найдем параметры кода (2.58):

$$\gamma_k=1-\frac{\varepsilon_{Vk}}{\hat{\lambda}_k};\quad \sigma_{Wk}=\gamma\varepsilon_{Vk}.$$

Приведем код (2.58) к виду, аналогичному преобразованию (2.38):

$$\hat{Z}_k=\frac{\sqrt{\gamma_k}}{a}(a\sqrt{\gamma_k}\,\hat{X}_k+\nu_k);\quad k\in[1,\ N], \qquad (2.59)$$

где a — коэффициент усиления; $\nu_k=\dfrac{aW_k}{\sqrt{\gamma_k}}$ — шум с дисперсией

$$\sigma_2^2=\mathbf{M}(\nu_k^2)=a^2\sigma^2_{Wk}/\gamma_k=a^2\mu.$$

В преобразование (2.59) вошли коэффициенты гауссовского процесса, которые имеют одинаковую дисперсию и потому относятся к белому шуму. Для лучшей расшифровки конструкции кода (2.59) перепишем его в виде:

$$\hat{Z}_k=q_{k\,\text{опт}}(g_{k\,\text{опт}}\hat{X}_k+\nu_k), \qquad (2.60)$$

где $q_{k\,\text{опт}}=\sqrt{\gamma_k}/a$; $g_{k\,\text{опт}}=a\sqrt{\gamma_k}$.
Выражение (2.60) показывает, что конструкцией линейного кода служит схема на рис. 1.2, в которой фильтр L является фильтром Винера, а оптимальные фильтры G и Q отличаются лишь на постоянную a^2.

Теперь перейдем к доказательству того, что код (2.60) является идеальным линейным кодом, т. е. обеспечивает выполнение уравнения Шеннона (1.17). Для этого сначала заметим, что с помощью кода (2.60) реализуется ε-энтропия (2.55), когда ошибка передачи ε_V имеет спектр (2.54):

$$\varepsilon_{Vk}=\begin{cases}\mu, & k<N;\\ \hat{\lambda}_k, & k\geqslant N.\end{cases}$$

Запишем уравнение Шеннона (1.17), использовав формулу Колмогорова (2.55):

$$\frac{1}{2}\sum_{k=1}^{N}\log\frac{\hat{\lambda}_k}{\mu^*}=C. \qquad (2.61)$$

В уравнении (2.61) величина μ^* выражает выбеленную часть спектра ($k<N$) нижней границы ошибки ε^*_V при заданной про-

пускной способности C гауссовского канала, представленного сумматором с белым шумом на рис. 1.2. В отличие от Р. Галлагера мы рассматриваем схему на рис. 1.2 не как тест-канал (искусственное сооружение, дающее легко запоминаемую картину вычисления ε-энтропии) [18], а как схему с настоящим, реальным непрерывным каналом, в котором выполняется уравнение (2.61).

Для нахождения решения уравнения (2.61) вычислим пропускную способность гауссовского канала. При этом мы будем строго следовать определению Шеннона, однако с тем уточнением, которое Шеннон внес в работе 1959 г. [3, с. 608]. Кодовая информация для схемы на рис. 1.2 равна

$$I(Y(t),\ Z(t)] = \frac{1}{2}\sum_{k=1}^{\infty}\log\left(1+\frac{g_k^2\hat{\lambda}_k}{\sigma_2^2}\right). \tag{2.62}$$

Решим вариационную задачу на отыскание максимума кодовой информации (2.62), достигаемого с помощью вариации передающего фильтра G при ограничении его «энергии» E_g:

$$C(E_g) = \max_{\{g_k^2\}} I[Y(t),\ Z(t)]. \tag{2.63}$$

Построим функцию Лагранжа:

$$\Phi(g_k^2) = \frac{1}{2}\sum_{k=1}^{\infty}\log\left(1+\frac{g_k^2\hat{\lambda}_k}{\sigma_2^2}\right) - \frac{1}{a_1^2}\sum_{k=1}^{\infty}g_k^2,$$

где $a^2{}_1$ — множитель Лагранжа.
Уравнение Эйлера имеет вид:

$$\partial\Phi(g^2{}_k)/\partial g^2{}_k = 0. \tag{2.64}$$

Решив уравнение Эйлера (2.64), определим оптимальные значения варьируемых коэффициентов:

$$\tilde{g}_k^2 = \begin{cases} a_1^2\dfrac{\Psi_k^2-1}{\Psi_k^2}, & k<N; \\ 0, & k\geqslant N, \end{cases} \tag{2.65}$$

где $\Psi^2{}_k = \dfrac{a_1^2\hat{\lambda}_k}{\sigma_2^2}$ — спектральное отношение сигнал-шум в канале.

Получим искомую формулу вычисления пропускной способности гауссовского канала при ограничении E_g:

$$C(E_g) = \sum_{k=1}^{\infty}C_{gk} = \frac{1}{2}\sum_{k=1}^{N}\log\left(1+\frac{\tilde{g}_k^2\hat{\lambda}_k}{\sigma_2^2}\right) = \frac{1}{2}\sum_{k=1}^{N}\log\Psi_k^2. \tag{2.66}$$

Формула (2.66) представляет пропускную способность гауссовского канала при ограничении E_g как сумму «элементарных» пропускных способностей множества параллельных одномерных каналов. При этом происходит усечение тех одномерных каналов, которым не выделена «энергия» усиления E_g передаточного фильтра:

$$C_{gk}=0 \text{ при } g_{k\,\text{опт}}=0,\quad k\geqslant N.$$

Подставим выражение (2.66) в уравнение (2.61):

$$\frac{1}{2}\sum_{k=1}^{N}\log\frac{\hat{\lambda}_k}{\mu^*} = \frac{1}{2}\sum_{k=1}^{N}\log\Psi_k^2. \tag{2.67}$$

Решение уравнения идеального линейного кодирования (2.67) дает нижнюю границу ошибки передачи:

$$\varepsilon_V^* = \sum_{k=1}^{\infty}\varepsilon_{Vk}^*;\quad \varepsilon_{Vk}^* = \begin{cases}\mu=\sigma_2^2/a^2, & k<N; \\ \hat{\lambda}_k, & k\geqslant N.\end{cases} \tag{2.68}$$

Сравнив ошибки (2.15), (2.54) и (2.68), полученные при решении трех вариационных задач, обнаружим, что эти ошибки совпадают, если выполняется равенство:

$$\mu^* = \sigma^2{}_2/a^2{}_1 = \sigma^2{}_2/a^2. \tag{2.69}$$

Равенство (2.69) достигается с помощью простого выравнивания коэффициентов усиления ($a=a_1$) фильтра G. Видно, что в этом случае совпадают спектральные характеристики передающих фильтров (2.16) и (2.65), которые дают одинаковые минимальные значения ошибок передачи:

$$\tilde{g}^2{}_k = g^2{}_{k\,\text{опт}};\quad \varepsilon^*{}_V = \varepsilon_g.$$

Оба фильтра (2.16) и (2.65) определяются по спектральному отношению сигнал-шум $\Psi^2{}_k$ в канале и имеют ограниченный запас «энергии» усиления. Декодирующий фильтр при оптимальном коде относится к типу фильтров Винера и в нашем случае совпадает с точностью до коэффициента усиления с оптимальным передаточным фильтром.

Таким образом, мы нашли решение (2.68) уравнения Шеннона (2.67) и доказали, что решения трех вариационных задач совпали. «Тройственность», которую мы доказали, можно формализовать с помощью выражений

$$H(\varepsilon^*{}_V) = H(\varepsilon_g) = I[\hat{X}(t),\ \hat{Z}(t)];$$
$$C(E_g) = I[\hat{X}(t),\ \hat{Z}(t)];$$
$$\varepsilon_g = H^{-1}(C) = d[\hat{X}(t),\ \hat{Z}(t)]. \tag{2.70}$$

Идеальный код (2.60) $\hat{Z}(t)$ реализуется с помощью простых линейных фильтров и обеспечивает выполнение доказываемой теоремы. Указанные линейные фильтры позволяют достичь равенства ε-энтропии источника и пропускной способности гауссовского канала с достижением минимальной ошибки ε_g, являющейся решением уравнения Шеннона (1.17) и одновременно решением вариационной задачи в теории фильтрации при ограничении «энергии»

усиления E_g. Полная минимальная ошибка для двухкаскадной схемы на рис. 1.2 равна сумме двух винеровских ошибок

$$\varepsilon^* = \mathbf{M}[X(t) - \hat{Z}(t)] = \varepsilon_B + \varepsilon_g = \varepsilon_B + \sum_{k=1}^{\infty} \hat{\lambda}_k 2^{-2C_{gk}}. \qquad (2.71)$$

Если бы в гауссовском случае при одинаковых длинах кодов источника и канала ($n_1 = n_2$) применить нелинейные кодирующий и декодирующий фильтры по способу кодирования Шеннона, то ошибка передачи стала бы больше ε_g.

В общем (негауссовском) случае это утверждение неверно. Кроме того, не следует забывать, что способ кодирования Шеннона предполагает изменение отношения n_1/n_2. Однако для частного случая, оговоренного в теореме, кодирование с помощью линейных фильтров будет лучше, чем с помощью нелинейных. Тем самым мы доказали, что рассматриваемое линейное кодирование является не просто оптимальным, но и абсолютно оптимальным в гауссовском случае при $n_1 = n_2$.

Обратим внимание на важный момент в доказательстве теоремы. Уравнение идеального линейного кодирования (2.67) с учетом определения (2.66) может рассматриваться как простое распространение уравнения Шеннона с одномерного случая на N-мерный, когда уравнение (2.43) применяется для каждой из N степеней свободы. В этом аспекте доказательство теоремы, казалось бы, носит тривиальный характер. Но вряд ли формулу (2.66), с помощью которой получено уравнение (2.67), можно назвать тривиальной. Наоборот, отличие формулы (2.66) от самой известной формулы Шеннона (1.18) содержит элемент неожиданности и может вызвать недоумение. Но это своеобразие формулы (2.66) вызвано спецификой определения пропускной способности непрерывного (в нашем случае гауссовского) канала. Шеннон отмечал, что пропускная способность непрерывного канала, строго говоря, бесконечна [3, с. 408]. Только условная пропускная способность может иметь конечное значение. Знаменитая формула Шеннона (1.18) справедлива при условии ограничения «энергии» сигнала E_s на выходе кодера и числа N_s степеней свободы. Однако теория Шеннона не регламентирует выбор типа ограничения [3, с. 308]. Мы исходим из того, что пропускная способность канала, определяемая с учетом ограничения на кодер, по существу, служит характеристикой не чисто канала, а канала совместно с кодером. Именно ограничение, присущее данному классу кодеров, обеспечивает конечную максимальную скорость передачи информации по каналу. Шенноновское ограничение кодера (E_s, N_s) относилось к классу нелинейных блочных кодеров. Нужно ли было сохранять шенноновское ограничение при поиске идеального кода в новом классе кодеров — линейных фильтров? Костас первым встал на путь применения энергетического ограничения Шеннона к классу линейных кодеров (фильтров) [10], но этот путь не привел к идеальному линейному кодированию [16—18]. Мы же учли, что все

линейные кодирующие фильтры обладают ограниченной «энергией» усиления E_g, и с помощью этого ограничения получили новую формулу (2.66) для вычисления пропускной способности. Из-за наличия двух типов ограничения кодеров и двух формул (1.18) и (2.66) встает вопрос о сравнении между собой трех способов кодирования — идеального блочного, способа Костаса и идеального линейного. Эти способы характеризуются соотношениями:

$$H(\varepsilon^*_V) = C(E_s, N_s),$$
$$H(\varepsilon_s) < C(E_s, N_s),$$
$$H(\varepsilon_g) = C(E_g).$$

Пропускная способность (2.66), хотя и меньше предельного значения (1.18), т. е. $C(E_g) < C(E_s)$, но может быть всегда сделана численно равной шенноновской пропускной способности

$$C(E_s, N_s) = C_P;\ C(E_g) = C_P.$$

При заданном значении пропускной способности C_p способ Костаса дает большую ошибку, чем два других:

$$\varepsilon^*_V(C_P) = \varepsilon_g(C_P) < \varepsilon_s(C_P).$$

Такое сравнение разных способов стало возможным благодаря использованию уравнения Шеннона при разных типах ограничения кодера.

2.6. СВОЙСТВА ИДЕАЛЬНОГО ЛИНЕЙНОГО КОДИРОВАНИЯ

Простота математического описания безблочного линейного кодирования позволяет наглядно представить процессы преобразования информации при идеальном кодировании в гауссовском случае. Источник информации выдает смесь $Y(t)$ сигнала $X(t)$ и шума $\xi(t)$, которая содержит полезную сигнальную информацию $I[X(t), Y(t)]$ на фоне огромного количества шумовой информации. Кодовая информация (2.52), передаваемая при неидеальном кодере, превышает количественно кодовую ε-энтропию (2.55):

$$I[Y(t), Z(t)] = I[X(t), \hat{Z}(t)] + I_s + I[\xi(t), \mathcal{U}(t)], \qquad (2.72)$$

где $I_s = I[X(t), Z(t)] - I[X(t), \hat{Z}(t)]$.

В кодовой информации (2.72) *избыточными являются два члена: избыточная сигнальная информация* I_s *и шумовая информация* (2.50). Количество кодовой информации, поступающей в канал, зависит от того, насколько кодеру удастся устранить избыточную информацию. Идеальный кодер сожмет в A раз информацию (2.72) до ε-энтропии (2.55), устранив полностью сигнальную избыточность и уменьшив шумовую избыточность ценой некоторой потери ΔI полезной информации:

$$H_Y(\varepsilon_g) = I[Y(t), \hat{Z}(t)] = I[X(t), \hat{Z}(t)] + I[\xi(t), \hat{\mathcal{U}}(t)];$$
$$I[X(t), \hat{Z}(t)] = I[X(t), Y(t)] - \Delta I. \qquad (2.73)$$

Коэффициент сжатия A определит устранимую избыточность по Шеннону (2.48):

$$1-\frac{1}{A}=\frac{I_s+I[\xi(t),\ \mathcal{U}(t)]-I[\xi(t),\ \hat{\mathcal{U}}(t)]}{I[Y(t),\ Z(t)]}. \qquad (2.74)$$

Если добиться равенства пропускной способности канала и кодовой информации (2.72), то это приведет к большой ошибке, т. е. низкому качеству передачи. Сжатие кодовой информации с помощью идеального кодера до величины (2.73) преследует цель снижения ошибки передачи до нижней границы, определяемой уравнением Шеннона. Так как «сжатая» кодовая информация (2.73) полностью заполняет канал с пропускной способностью (2.66), последнюю можно, как в одномерном случае (2.45), считать состоящей из сигнальной C_X и шумовой $C_ш$ составляющих:

$$C(E_g)=C_X+C_ш;\ C_X=I[X(t),\ \hat{Z}(t)],\ C_ш=I[\xi(t),\ \hat{\mathcal{U}}(t)]. \qquad (2.75)$$

Избыточность (2.74) показывает, что существование шумовой информации представляет новую проблему, более сложную, чем выявленная Шенноном проблема статистической избыточности сообщений от незашумленных источников. Сигнальная избыточность полностью устранима идеальным кодированием, тогда как шумовая избыточная информация уменьшается лишь до неустранимой величины. Для оценки степени бедствия, вытекающего из факта существования неустранимой шумовой избыточности, удалось установить ее зависимость от потери полезной сигнальной информации. Найденное для идеального линейного кодирования обменное соотношение между неустранимой частью шумовой информации и потерей полезной информации описано с помощью простых формул в одномерном случае (см. рис. 2.1). Эта закономерность может быть обобщена и на многомерный случай. Найденная закономерность дает следующую картину. Если получателю информации не нужна слишком малая ошибка передачи ($\varepsilon_g > \varepsilon_B$) и, следовательно, слишком малая потеря полезной информации, то идеальный кодер оставляет небольшое количество шумовой информации, занимающей небольшую часть пропускной способности ($C_ш \ll C_X$). Но если получателю необходима очень большая точность передачи, т. е. ошибка и потеря информации слишком малы ($\varepsilon_g < \varepsilon_B$), то не только надо увеличивать пропускную способность (2.75), но и приходится расходовать ее очень неэффективно из-за передачи шумовой информации ($C_ш \gg C_X$). Происходит это потому, что шум источника создает информацию, которую идеальный кодер не может подавить и которая существенно превышает полезную:

$$I[\xi(t), \hat{\mathcal{U}}(t)] \gg I[X(t), Y(t)].$$

В этом случае при принятии решений получатель вынужден перерабатывать огромное количество шумовой информации, превышающее количество полезной для получателя информации. Пал-

лиативным способом решения сформулированной проблемы является поиск разумного компромисса между потерей полезной информации и шумовой информацией путем выбора разумно достаточной точности идеального кодирования. Радикальное решение указанной проблемы требует выхода за рамки идеального кодирования, основанного на достаточно простых формальных правилах. Видимо, необходимо семантическое кодирование, гипотеза о котором включает меру с двойным логарифмом [2, с. 139]:

$$I_{сем}=\log I[Y(t),\ \hat{Z}(t)]=\log\log N_C.$$

Только семантическое кодирование способно будет снизить остаточную шумовую информацию до достаточно малой величины δ:

$$I_{сем}=\log I[Y(t),\ \hat{Z}(t)]=\log I[X(t),\ \hat{Z}(t)]+\delta.$$

Сказанное выше подтверждает, что идеальное линейное кодирование дает новые нетривиальные результаты для теории. Но оно имеет также важное значение и для практики. На практике система связи должна доставлять информацию получателю не только с требуемой точностью, но и своевременно. Задержка в доставке информации может сделать ее ненужной даже при сколь угодно большой точности воспроизведения. Поэтому в технике связи важна так называемая передача в реальном времени. Шеннон отмечал, что способы кодирования без задержек более важны для приложений, чем блочные способы, но они хуже разработаны теоретически [3, с. 462]. Если стремиться к универсальности способа и предъявить к нему требование нулевой ошибки, как это было свойственно Шеннону, то трудно было рассчитывать на создание теории кодирования без задержек. Однако инженеру важны и частные, не слишком универсальные способы кодирования. Примером такого частного решения и служит идеальное линейное кодирование. Отсутствие у него блочных задержек, достаточно больших при коде Шеннона, весьма привлекательно для инженеров.

Простота конструкции идеального линейного кодирования поясняется рис. 2.3 и 2.4. Мы строили идеальный код, используя разложение Карунена — Лоэва, что отражено на рисунках в виде блока преобразования координат. Затем идет идеальный кодирующий фильтр со спектральной характеристикой (2.65) и импульсной характеристикой, описываемой с помощью оператора Гильберта — Шмидта. Импульсная характеристика определена на конечном временно́м (или пространственно-временном) интервале и

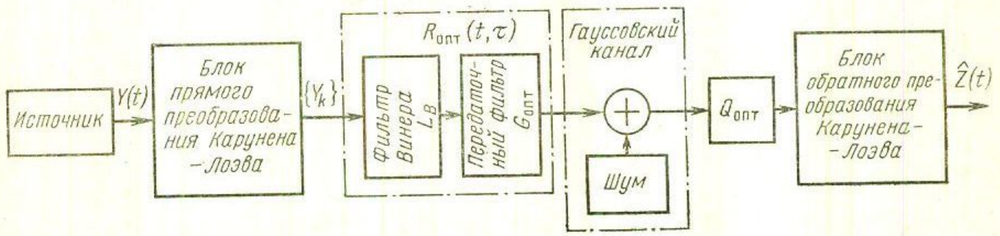


Рис. 2.3. Структурная схема кодирования с преобразованием Карунена — Лоэва

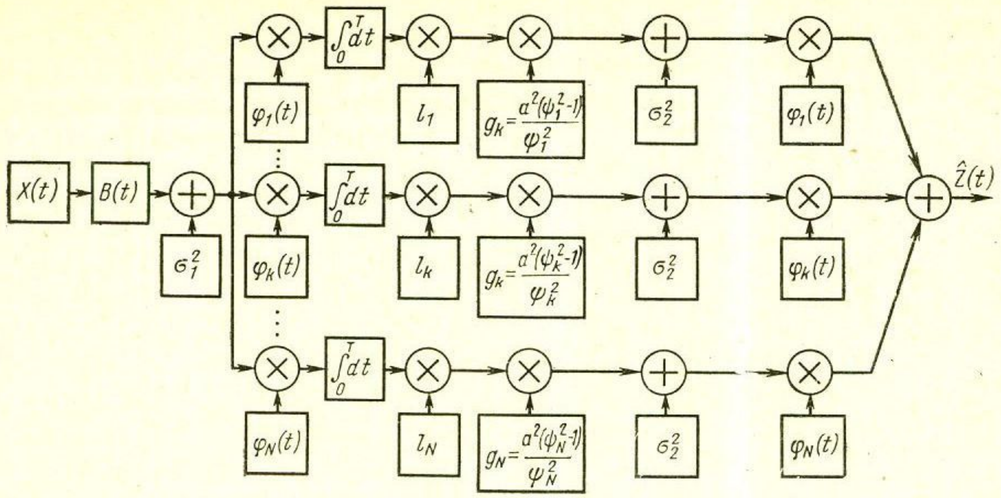


Рис. 2.4. Структурная схема параллельной системы линейного кодирования с преобразованием Карунена — Лоэва

имеет N степеней свободы. Кодирующий фильтр усекает ряд Карунена — Лоэва по пороговому отношению сигнал-шум и усиливает надпороговые спектральные коэффициенты сигнала ($\Psi^2_k > 1$) перед каналом с белым шумом. На рис. 2.4 гауссовский канал представлен как N параллельных «элементарных» каналов с пропускными способностями C_{gk}. На приемной стороне сигнал декодируют с помощью фильтра Винера и выполняют обратное преобразование. Получателю информации доступен исходный полезный сигнал $X(t)$ в виде функции

$$Z(t) = \sum_{k=1}^{N} Z_k \varphi_k(t).$$

Сигнал, по которому получатель принимает решение, представляет условное математическое ожидание исходной функции:

$$\hat{Z}(t) = \mathbf{M}[X(t) \mid Z(t)].$$

Эта наилучшая оценка полезного сигнала достигается с помощью идеального кодирующего фильтра и декодирующего фильтра Винера.

Схемы на рис. 2.3, 2.4 позволяют сравнить идеальное линейное кодирование и так называемое «кодирование на основе преобразования» или «кодирование с преобразованием» [60, с. 698; 63]. Очевидно сходство этих двух методов. Отсутствие квантования спектральных коэффициентов в схеме на рис. 2.3 не является отличительным признаком, так как может быть введено на тех же основаниях, что и во втором методе. Отмеченное сходство показывает, что метод «кодирование с преобразованием» представляет собой один из вариантов реализации идеального линейного кодирования. Математическое описание свойств идеального линейного кодирования позволяет инженерам сравнивать данную реализа-

цию метода с идеалом и компромиссно выбирать параметры субоптимального варианта.

Выводы по главе. Один из основных выводов утверждает, что оптимальное линейное кодирование при некоторых условиях является абсолютно оптимальным, т. е. обеспечивает минимальную среднеквадратическую ошибку, совпадающую с решением уравнения идеального кодирования Шеннона (при $n_1 = n_2$). Этот вывод составил содержание теоремы об идеальном линейном кодировании. Сама постановка вопроса об идеальном по Шеннону линейном кодировании содержала элемент неожиданности. Согласно теореме Шеннона достижение минимальной ошибки требует в общем случае очень сложного правила кодирования информации (нелинейный блоковый код) и приводит к большому запаздыванию ее выдачи на выходе канала. В отличие от этого линейное кодирование реализуется с помощью простых линейных фильтров и не требует временной задержки на формирование блоков Шеннона. Поэтому казалось вполне естественным, что линейное кодирование *во всех случаях*, в том числе при $n_1 = n_2$, должно давать ошибку, большую, чем при нелинейном блоковом кодировании. В этом плане оценивалась лишь степень приближения ошибки оптимального линейного кодирования к минимальной ошибке из уравнения Шеннона [16, 17]. Большая эффективность сложного способа кодирования Шеннона, проявляющаяся в устранении ошибки, признавалась вне конкуренции. На этом фоне и была неожиданной попытка найти хотя бы некоторые частные случаи, когда линейное кодирование обладало бы той же эффективностью при устранении ошибки, что и идеальный код Шеннона.

Эта попытка привела к доказательству теоремы об идеальном безблочном линейном кодировании, которое выполнено для гауссовского случая. Гауссовское распределение случайных процессов — обязательное условие для доказательства теоремы. Другим условием является переход от ограничения на «энергию» сигнала, использованного в работах Шеннона и Костаса, к ограничению «энергии» фильтра. Хотя важность выбора ограничения давно известна из вариационного исчисления, заранее было не ясно, что без смены ограничения не удастся доказать теорему. Вопрос смены ограничения был труден и в методологическом плане, ведь надо было отойти от ограничения, примененного Шенноном и Костасом и признанного наиболее физичным. Здесь помогло понимание того факта, что энергетический спектр информационных сигналов — их «энергия» — измеряется не в джоулях, а является безразмерной величиной, точно так же, как спектральные характеристики и «энергия» фильтров. Безразмерные энергии сигнала и фильтра имеют одинаковую физичность или одинаковую нефизичность. К этому можно добавить, что ограничение «энергии» фильтра имеет в радиотехнике достаточно давнее происхождение, известно как «широкополосность» ламп или «площадь усиления».

Теорема об идеальном линейном кодировании относится не к типу теорем о существовании, а имеет конструктивный характер.

Она дала простую конструкцию идеального линейного кодера в виде двузвенного фильтра, составленного из фильтров Винера для борьбы с входным шумом и передаточного фильтра, который в силу нового ограничения отличается от фильтра Костаса. Оптимальный передающий фильтр определяется коэффициентом усиления и спектральным отношением сигнал-шум. Его конечномерная спектральная характеристика может быть аппроксимирована прямоугольной функцией со срезом спектрального отношения сигнал-шум по пороговому значению. Принципиальная новизна заключается здесь в том, что теперь мы видим, кто и в каких пределах усекает неограниченный (нефинитный) спектр разложения Карунена — Лоэва, которое описывает зашумленный сигнал источника информации. Так как передаточный фильтр усекает по порогу число степеней свободы, то отпадает неприятная потребность в моделировании входных сигналов, заданных на конечном отрезке T, сингулярными функциями с ограниченным (финитным) спектром. Идея А. Н. Колмогорова о вырождении бесконечномерных процессов в конечномерные, высказанная им при выводе формулы для ε-энтропии, получает здесь свою дальнейшую конкретизацию. Приобретают новую трактовку и старейшие эвристические представления в радиотехнике об идеальном неискажающем фильтре с прямоугольной спектральной характеристикой.

Из теоремы следует, что сложнейшее нелинейное кодирование Шеннона при равной длине блоков источника и канала ($n_2=n_1$) дает в гауссовском случае ошибку, большую, чем простое идеальное линейное кодирование без блоковой задержки. Этот вывод должен несколько развеять «страх» инженеров, которые давно и успешно применяют на практике понятие о линейных фильтрах, перед математической сложностью теории Шеннона. Конечно, надо помнить, что в общем негауссовском случае и для произвольных каналов нет универсального способа кодирования, лучшего нелинейного блокового кодирования Шеннона.

Не менее важным выводом, относящимся не только к гауссовскому случаю, является обнаружение эффекта расщепления ε-энтропии зашумленного источника и введение понятия о шумовой информации для объяснения этого эффекта. Шумовая информация определена как средняя взаимная информация между шумом на выходе канала и шумом источника. В гауссовском случае шумовую информацию можно вычислить по простой формуле через спектральные коэффициенты шума кодера с каналом. Шумовая информация в сумме с полезной информацией о входном сигнале равна кодовой информации, поступающей в канал. Элемент неожиданности проявился здесь в двустороннем вредном влиянии шума источника. С одной стороны, шум источника уменьшает полезную сигнальную информацию, содержащуюся на входе кодера. Это проявление шума учитывалось и ранее. Но оказалось, что шум источника рождает еще собственную шумовую информацию, которая в сумме с полезной сигнальной информацией загружает канал. Важное значение шумовой информации заключается в том,

что она является лимитирующим фактором для передачи конечной полезной сигнальной информации с помощью конечной пропускной способности.

Между потерей полезной информации и шумовой информацией возможен обмен согласно информационному соотношению неопределенности (2.4). Конечность пропускной способности канала создает проблему выбора между стремлением устранить потерю полезной информации и желанием не передавать по каналу шумовую информацию. Наличие в идеальном кодере передаточного фильтра позволяет решить указанную проблему путем компромисса. Это означает, что полная адекватность переданной по каналу информации объекту наблюдения невозможна.

При сохранении квадратического критерия качества можно расширить рамки теории безблокового ($n_1=n_2$) кодирования и на негауссовские сигналы, правда, лишившись простых спектральных формул. Безблоковое кодирование ближе к практически реализуемым сегодня способам кодирования, чем блоковое.

Изложенная в этой главе теория безблокового кодирования может быть распространена и на кодирование при идеальной передаче по двоичному симметричному каналу равномерно распределенной случайной величины. Ключевым моментом было вычисление ε-энтропии источника случайной величины при квадратическом критерии точности с последующей проверкой уравнения (1.17) [30]. При ее нахождении случайная величина представляется числовым рядом, который усекается кодером, осуществляющим квантование с предусилением. При таком информационном подходе необходимость квантования для равномерного закона, как и дискретизации для гауссовского, возникает естественным путем как средство реализации идеальной передачи (1.17), а не как способ приближенного вычисления.

Глава 3.

КОДИРОВАНИЕ ИЗОБРАЖЕНИЙ В РЕАЛЬНОМ ВРЕМЕНИ

Обычно рассматривают кодирование после телекамеры, считая телекамеру источником информации. Однако послекамерному кодированию предшествует камерное кодирование — преобразование света от наблюдаемого сюжета (источника информации) в цифровой электрический сигнал, осуществляемое в реальном времени, т. е. с задержкой на один телевизионный кадр. Кодирование в телекамере осуществляется в два этапа: кодирование в ПЗС, включая видеоусилитель, и кодирование в аналого-цифровом преобразователе (АЦП) с предварительной дискретно-аналоговой обработкой. Кодирование в ПЗС включает адаптивное управление

процессами пространственно-временного накопления и считывания. Здесь мы покажем, как управлять временем и зоной накопления (суммирования) с помощью специального сигнала, формируемого до считывания и даже до завершения процесса накопления. Получение такого сигнала осуществляется практически без временной задержки, а процесс накопления зарядового (скрытого) изображения не подвергается искажению или зашумлению. Осуществляя быстрое неразрушающее измерение параметров зарядового изображения, можно сформировать специальный опережающий сигнал для управления ходом процесса считывания. Сигналы, управляющие считыванием, использованы и для управления параметрами дискретно-аналоговой обработки электрического сигнала перед АЦП.

3.1. ПРИМЕНЕНИЕ ТЕОРИИ ИНФОРМАЦИИ В ТЕЛЕВИДЕНИИ

Применение теории информации в телевидении должно начинаться с выяснения того, насколько хорошо математическая модель отражает специфику телевизионных систем. Дополнение модели на рис. В.1 шумом источника и учет его влияния в модели на рис. В.3 (гл. 2) позволил сблизить модель с практической системой, так как в телевидении все источники информации создают зашумленный сигнал. Однако остаются следующие трудности моделирования:

статистическое описание полезного сигнала источника;

наличие в телевизионной системе, в отличие от модели на рис. В.3, не одного канала, а нескольких последовательно соединенных каналов, например двух разнотипных каналов, как в модели на рис. В.2;

формализация критерия качества телевизионной системы с учетом ее назначения.

Проблема выбора математического критерия качества хотя и относится к наиболее дискуссионным в телевидении, но, по существу, сужается до вопроса о применимости квадратичного критерия точности (среднеквадратической ошибки). Рассматривая этот вопрос в аспекте применения теории информации в телевидении, необходимо помнить, что только в гауссовском случае теория дает информационные формулы в явном виде, что наглядно было показано в гл. 2. Отказ от применения квадратического критерия означал бы невозможность применения этих информационных формул в телевидении. Нельзя было бы применить формулу Колмогорова для вычисления ε-энтропии источника, что сделало бы менее плодотворным применение понятия ε-энтропии и, следовательно, теоремы Шеннона 1959 г. Пришлось бы отказаться от применения фильтра Винера в телевидении, не говоря уже об идеальном линейном кодировании, изложенном в гл. 2.

Назначением рассматриваемого нами класса телевизионных систем является проверка гипотез, или классификация изображений, осуществляемая в реальном времени с помощью видеокомьютера (получателя информации). Решение за время телевизионного кадра должен выдать видеокомпьютер, а не зрительный анализатор человека. Поэтому для класса систем «от света до решения» нет необходимости дискутировать о том, насколько правильно среднеквадратическая ошибка отражает субъективное качество восприятия изображений человеком. Принятие модели на рис. В.2 с квадратическим критерием точности открыло путь к содержательному применению теории информации в телевидении.

Вопрос о том, что считать источником информации, а что каналом при моделировании многокаскадной телевизионной системы, не столь очевиден, как кажется с первого взгляда. Трудность вызвана тем, что в теории информации модель системы (рис. В.1) содержит только один канал, математически описываемый как сумматор сигнала с шумом, а в телевизионной системе имеется много источников шума и, соответственно, можно выделить несколько сумматоров сигнала с шумом. Это ведет к модели системы с несколькими каналами связи, которые могут относиться и к дискретным и к непрерывным каналам. Поэтому телевизионная система относится к типу смешанных систем [3, с. 246], наиболее трудному для теоретического анализа. На рис. В.2 была дана модель смешанной системы с двумя каналами, которая применима к телевизионной системе, включающей преобразование «от света до решения».

Модель на рис. В.2 отражает синтез телекамеры и видеокомпьютера для выработки решений по видеоинформации, содержащейся во входных оптических изображениях, доставляемых электромагнитным полем от сцены к телекамере. Инженерам хорошо известно, что именно телекамера вносит основной вклад в ошибку воспроизведения полезного сигнала и необратимую потерю информации, и поэтому изъятие телекамеры из кодирования недопустимо. Однако недостаточно провозгласить, что телекамера является кодером. Необходимо математически описать процесс кодирования в телекамере. Камерное кодирование относится к смешанному типу кодирования, так как сначала преобразует путем параллельного накопления заряда в фотоприемнике на ПЗС аналоговое (на макроскопическом уровне) оптическое изображение в дискретно-аналоговое изображение, а затем квантует и преобразует его в цифровую форму с помощью АЦП. Камерное кодирование представлено на рис. В.2 двумя типами каналов и кодеков. Только применение квадратического критерия точности позволяет получить приемлемое для инженеров математическое описание камерного кодирования. Так, линейное кодирование, изложенное в гл. 2, достаточно хорошо моделирует аналоговую часть камерного кодирования.

Если каналы и кодеки в схеме на рис. В.2 описываются в рамках теории информации, то видеокомпьютер относится к но-

вой ветви этой теории — теории алгоритмов [4]. В гл. 5 будут описаны телевизионные роботы, для функционирования которых достаточно классификации изображений по формализованным правилам, вытекающим из теории решений [12], не обращаясь к сложной проблеме опознавания семантических образов. Наиболее популярной является многоальтернативная задача проверки M гипотез [8]. Пусть видеокомпьютер телевизионного робота способен за время кадра $T_к$ принять n_0 решений, каждое из которых соответствует выбору из M гипотез. Скорость принятия решений будет равна $n_0/T_к$ решений в секунду. Если сопоставить решению $\log M$ дв. ед., то решающая (а не пропускная) способность видеокомпьютера в двоичных единицах в секунду будет равна:

$$V_{реш} = (n_0 \log M)/T_к.$$

Цифровой канал в системе на рис. В.2 имеет пропускную способность в двоичных единицах в секунду C, которая определяет скорость поступления цифровой видеоинформации в компьютер. Хотя решающая способность компьютера и пропускная способность его канала ввода выражаются в одних и тех же единицах, они характеризуют разные узлы системы. Характеристику $V_{реш}$ нельзя трактовать как «пропускную способность системы в целом», так как она характеризует получателя информации, а не канал.

Полезный сигнал о наблюдаемом сюжете совместно с шумом доставляет телекамере электромагнитное поле с длиной волны из оптического диапазона. Источник оптического поля можно считать дискретным (фотонным) на микроскопическом уровне и непрерывным на макроскопическом уровне.

На рис. 2.2 были приведены характеристики кодовой и сигнальной ε-энтропии дискретного источника при использовании квадратического критерия точности. Эти характеристики и понятие о шумовой информации могут быть отнесены и к пуассоновскому источнику. Если бы телекамера выполняла функции идеального блокового кодера Шеннона, то сигнал от пуассоновского источника можно было бы экономно передать по дискретному каналу с пропускной способностью $C_X + C_ш$ при минимальной среднеквадратической ошибке ε^* (рис. 2.2):

$$H_Y(\varepsilon^* - \varepsilon'_B) = (C_X + C_ш) n_2/n_1, \tag{3.1}$$

где n_1, n_2 — длины блоков кодеров источника и канала. Максимальное значение энтропии в двоичных единицах фотонного источника соответствует сумме $N_ф$ всех фотонов, поступающих в телекамеру за время кадра $T_к$, и приближенно равно [2]:

$$H_{макс}[Y(t)] = N_ф. \tag{3.2}$$

Информационный поток, оцениваемый формулой (3.2), является статистически избыточным. Полное сокращение избыточности достигается с помощью идеального кодирования. Коэффициент

максимального сжатия (2.48), который можно достичь с помощью кодирования Шеннона, равен:

$$A_м = N_ф/H_Y(\varepsilon^* - \varepsilon'_B) = N_ф/(C_X + C_ш) n_2/n_1. \tag{3.3}$$

Столь большой коэффициент сжатия достигается ценой не только увеличения сложности телекамеры до уровня идеального кодера Шеннона, но и значительного увеличения задержки во времени, необходимой для формирования достаточно длинных блоков, и, следовательно, ценой временной задержки принятия решения. Если же телекамера является лишь счетчиком фотонов, но не выполняет функции идеального нелинейного кодера Шеннона, то канал, связывающий телекамеру с компьютером, должен будет передавать кодовую информацию в двоичных единицах за секунду со средней скоростью $N_ф/T_к$. Эта скорость передачи чрезмерно велика при освещенности, с которой обычно работают в телевидении. Только в астрономии при наблюдении весьма слабых по силе света объектов телевизионный счетчик фотонов может иметь достаточно малую скорость $N_ф/T_к$, которая соизмерима с быстродействием ЭВМ. Скорость передачи будет расти пропорционально увеличению силы света и быстро превзойдет возможности любого практического канала. Возникает противоречие в телевидении между стремлением к индивидуальному счету фотонов и скоростью передачи информации по каналу. Указанное противоречие разрешается в телекамере с помощью принципа многоэлементного накопления зарядов фотоэлектронов [2], который можно назвать и принципом группового счета фотонов. В соответствии с этим принципом ПЗС с небольшой долей идеализации можно представить как мозаику из $N_э$ жестко расположенных элементов — счетчиков фотоэлектронов, каждый из которых выдает число просуммированных (накопленных) им за время кадра $T_к$ фотоэлектронов [27]. Фотоэлектроны внутри каждой группы, накопленной элементом ПЗС, не различаются. Зарядовое изображение в секции накопления ПЗС представляет таблицу из $N_э$ чисел, оценивающих число фотоэлектронов в каждой группе. Поэтому ПЗС можно назвать групповым счетчиком фотонов. Принцип многоэлементного накопления фотоэлектронов всегда рассматривался в телевидении как основное средство достижения наибольшего отношения полезного сигнала к фотонному шуму. Расчет энергетического отношения сигнал-шум при групповом счете фотонов был дан еще в 40-х годах А. Роузом [26]. Мы же подчеркнем значение *принципа группового счета фотонов как средства сжатия кодовой информации* [2]. Грубая оценка кодовой информации в двоичных единицах, содержащейся в зарядовом изображении, равна

$$I_э = \frac{N_э}{2} \log \frac{N_ф}{N_э}. \tag{3.4}$$

Для считывания накопленной информации за время кадра необходим канал считывания (видеоканал) со скоростью передачи $I_э/T_к$. Принцип накопления обеспечивает первичное сжатие кодо-

вой информации с меньшим коэффициентом, чем коэффициент (3.3):

$$A_э = \frac{N_ф}{I_э} = \frac{N_ф}{N_э \log \frac{N_ф}{N_э}} > 1.$$

Принцип многоэлементного накопления фотоэлектронов и принцип идеального кодирования Шеннона возникли в разных областях знаний. Первый возник в практике и был впервые реализован в телевидении в начале 30-х годов в передающей трубке — иконоскопе [36]. Второй принцип появился в теории информации и воплотился в основной теореме Шеннона [3]. Мы дали принципу многоэлементного счета фотонов трактовку как одного из способов кодирования — безблочного, работающего в реальном времени.

Пуассоновский случай получил распространение в телевидении не только потому, что хорошо отражает фотонную природу фотоэффекта, но и потому, что он дал простую формулу расчета энергетического отношения сигнал-шум, примененную А. Роузом для оценки предельной чувствительности [26]. Однако информационные характеристики в пуассоновском случае не имеют простых формул. Формулы (3.2) и (3.4) были даны лишь как приближенные без оценки ошибки приближения. Как видно из гл. 2, только гауссовский случай дает простые информационные формулы, что и позволило развить теорию линейного кодирования.

Моделирование с переходом от пуассоновской модели источника к гауссовской открывает путь применения информационно-спектральной теории линейного кодирования (гл. 2) для описания первого этапа кодирования в телекамере, осуществляемого в аналоговой форме. На втором этапе кодирования дискретно-аналоговый телевизионный сигнал, пройдя предварительную обработку, кодируется поэлементно с помощью АЦП и в двоичной форме присутствует на выходе телекамеры. Преобразование зарядового изображения в двоичный сигнал на выходе телекамеры сопровождается необратимой потерей информации. Поэтому кодовая информация на выходе телекамеры будет меньше значения (3.4)

$$I_0 = N_э \log \mathcal{N} < N_э \log \sqrt{\frac{N_ф}{N_э}},$$

где $\log \mathcal{N}$ — число разрядов АЦП.

Камерное кодирование можно характеризовать коэффициентом сжатия, значение которого в первом приближении равно

$$A_и = N_ф / I_0 = N_ф / N_э \log \mathcal{N}. \qquad (3.5)$$

Коэффициент сжатия (3.5) существенно больше единицы, так как обычно $N_э \ll N_ф$. Камерное кодирование является безблоковым, поэтому остается возможность дополнительного сокращения избыточности с помощью цифрового послекамерного кодирования (см. 4.2).

Обращаясь к истории применения теории информации в телевидении и радиолокации, следует отметить ее малую практическую значимость. В телевидении это подтверждается тем, что все современные учебники по-прежнему излагают теоретические основы курса, опираясь не на теорию информации, а на физиологию зрения. Теоретические основы радиолокации излагаются с применением теории решений, но без применения теории информации.

Первопричина такого странного положения заключается в сложности теории Шеннона и разобщенности между теорией информации и практикой построения информационных систем. Наличие такой разобщенности и необходимость ее преодоления является известным фактом. Преодолеть эту разобщенность, т. е. перекинуть мост между математиками, развивающими теорию информации, и инженерами входило в задачу известной книги Р. Галлагера [18]. Однако вышедшие после нее книги по цифровой обработке изображений, как в США, так и у нас [55, 56, 59, 60, 63] по-прежнему не базируются на обобщенной теореме Шеннона [3, с. 587], т. е. практически не применяют теорию информации. Для преодоления разрыва теории и практики весьма перспективным было обращение к кодированию с помощью линейных фильтров, хорошо освоенных инженерами. Кроме того, такой способ кодирования не требует формирования сложного блокового кода. Центральным моментом здесь был вопрос о существовании идеального линейного кодирования. И хотя линейному кодированию было посвящено много работ, по-видимому, теорема об идеальном кодировании (гл. 2) ранее не появлялась в книгах. Легко заметить, что если способ идеального линейного кодирования (рис. 2.4) дополнить квантованием, то он аналогичен способу кодирования на основе преобразования Карунена — Лоэва [59, т. 2, с. 701]. Передаточный фильтр в схеме на рис. 2.4 выполняет функцию селектора спектральных отсчетов, превышающих пороговое отношение сигнал-шум. У. Прэтт отмечал, что способ кодирования с преобразованием основан на идее усечения концентрированного спектра Фурье [59, т. 2, с. 699]. Однако идея усечения спектра Фурье была реализована в линейном кодировании Костаса [10, 14] и в тест-канале, формирующем ε-энтропию [4, 18]. Оставалось неясным, каковы обоснования усечения спектра при кодировании в рамках теории Шеннона (вне Шеннона теория кодирования пока неизвестна). Поэтому теорема об идеальном линейном кодировании обосновывает теоретически идею усечения спектра Фурье (вырождения степеней свободы) при кодировании с преобразованием и тем самым перекидывает мост от теории к практике.

Коэффициент сжатия (3.3) кодовой информации распределяется между камерным кодированием (3.5) и послекамерным (см. 4.2). В первом приближении можно считать, что полный коэффициент сжатия (3.3) равен произведению двух коэффициентов, характеризующих камерное и послекамерное кодирование:

$$A_м = A_и A_ш,$$

где $A_ш$ — коэффициент сжатия послекамерного кодирования.

Из формул (3.3) и (3.5) найдем грубую оценку коэффициента сжатия послекамерного кодирования:

$$A_ш = \frac{N_э \log \mathcal{N}}{H(\varepsilon^* - \varepsilon'_B)}. \qquad (3.6)$$

Формула (3.6) показывает, что для нахождения коэффициента сжатия с помощью цифрового послекамерного кодирования надо использовать квадратический критерий точности и вычислить кодовую энтропию источника. Так как послекамерному цифровому кодированию уделяется достаточно много внимания, мы не будем его касаться, а изложим не изученную проблему кодирования в телекамере на ПЗС.

3.2. ОПРЕДЕЛЕНИЕ КАДРОВОЙ ЧАСТОТЫ ТЕЛЕКАМЕР

Основным инструментом сжатия сигнала в процессе аналогового кодирования является пространственно-временное накопление, реализуемое в фотоприемнике. Так как после аналогового кодирования в видеосигнал добавляются шумы канала считывания, необратимо искажающие сигнал, то выбор параметров накопления становится первоочередной задачей проектирования телекамер.

Робототехнические телекамеры проектируются для решения конкретных задач с помощью видеокомпьютеров. Замена получателя-человека автоматическим получателем позволила определять параметры разложения изображений в телекамерах по характеристикам сюжета, а не характеристикам зрительного анализатора. Поэтому у робототехнических телекамер кадровая частота может быть как очень низкой (до 10^{-1} Гц), так и сверхвысокой — до 2000 Гц. В 3.3—3.5 будут описаны способы регулировки кадровой частоты по яркостным характеристикам сюжета. Здесь мы рассмотрим выбор кадровой частоты по динамике сюжета, которая может определяться собственным движением объекта, движением телекамеры, установленной на манипуляторе робота или, как это имеет место при астрономических наблюдениях, изменениями формы и положения изображения из-за турбулентности атмосферы [86, 103]. Для астрономических систем, в которых используется адаптивная оптика для компенсации турбулентности атмосферы, частота кадров телевизионных систем должна иметь порядок 200 Гц.

Частота кадров выбирается при наблюдении подвижных изображений исходя из компромисса между необходимостью увеличения времени накопления сигнала для его выделения из шума и необходимостью уменьшения времени воздействия искажающего фактора — скоростного смаза. Принятию решений по подвижным изображениям предшествуют два этапа кодирования — аналоговый в телекамере и цифровой в видеопроцессоре. На видеопроцессор возлагается роль декодера, на телекамеру — роль кодера изображений. Выбор частоты кадров определяет требования по быстродействию как к телекамере, так и к видеопроцессору, что и определяет важную роль правильного выбора кадровой частоты. При вычислении кадровой частоты мы опираемся на модели сигналов и их преобразований в процессе камерного кодирования — накопления и дискретизации. Важной чертой камерного

кодирования является реальность времени, т. е. недопустимость существенных задержек. Важно, что увеличение задержек в процессе накопления увеличивает ошибки экстраполяции положения объектов. Поэтому частота кадров при необходимости экстраполяции оказывается выше, чем при оценивании без упреждения. Так, если объект движется по случайной траектории, то при экстраполяции его положения на время одного кадра (в течение которого зарядовое изображение считывается из секции хранения) для достижения минимально возможной ошибки потребуется почти вдвое более высокая кадровая частота.

Динамика сюжета может быть учтена с помощью линейных искажений, вносимых фильтром B в схеме на рис. 1.2. Такая модель позволяет применить результаты гл. 2 к определению частоты кадров в телекамере при рассмотрении сигнала изображения в виде гауссовского поля. Этот сигнал разлагается в ряд по прямоугольным функциям φ_k с шириной Δ вида [2, 58—60]

$$\varphi_k = \mathrm{rect}[(t - k\Delta)/\Delta].$$

Разложение по прямоугольным функциям при покадровой дискретизации вместо оптимального разложения по базису Карунена — Лоэва определяется упомянутой выше необходимостью уменьшения ошибок экстраполяции и сложности видеокомпьютера при восстановлении пространственно-временного процесса. Переход от оптимального базиса Карунена — Лоэва к разложению по прямоугольным функциям ведет к замене формы импульсной характеристики, эквивалентная ширина которой отражает способность фильтра накапливать сигнал для борьбы с шумом. Для временного аргумента она соответствует времени кадра $T_к$, для пространственного — определяется отношением T/N интервала задания сигнала T к числу степеней свободы N, определяемому из (2.74). Для сигнала со спектром Баттерворта при достаточно большом произведении длительности на эквивалентную полосу частот импульсная характеристика идеального линейного кодера имеет вид, близкий к $\{(\sin tN/T)/(tN/T),\ t \in [0, T]\}$, приближением к которому является прямоугольная импульсная характеристика ПЗС с почти такой же эквивалентной шириной (см. 3.8). Здесь же воспользуемся предположением о независимости кадров и рассмотрим разложение сигнала по времени по прямоугольным функциям, а по пространственным координатам — по базису Карунена — Лоэва. Это позволит применить многомерные информационные формулы из гл. 2. Переход к разложению по прямоугольным функциям несколько снизит количество взаимной информации $I[X(t), Z(t)]$, не изменив качественной картины — наличия оптимальной частоты кадров, выбираемой по малой величине смаза изображения.

Для борьбы со смазом время накопления можно регулировать и при постоянной частоте кадров («электронный затвор»). Этот способ применяют для автоматической регулировки чувствительности и изучения быстропротекающих процессов. При необходи-

мости принятия решения в реальном времени (см. 4.2, 5.4) для уменьшения задержки управления предпочтительнее регулировка и частоты кадров.

Оптимальная частота кадров или обратная ей величина время кадра могут быть определены по динамике сюжета для двух моделей изображения: случайно движущегося сигнала известной формы и регулярно движущегося случайного сигнала. В первом случае время кадра определяется необходимым превышением межкадрового приращения положения объекта над ошибками измерения [2]:

$$\mathbf{M}\{[X(t)-X(t-\Delta_0)]^2\}\approx 2{,}4\varepsilon_{\text{т мин}}.$$

При линейной экстраполяции положения объекта оптимальное время кадра определится отношением сигнал-шум $\Psi^2=\lambda_0/\sigma^2$ и интервалом корреляции траектории его движения [2]: $\Delta_0\approx 2\tau/\Psi$.

Для нахождения оптимальной частоты кадров при второй модели будем считать, что дисперсия P полезного сигнала — гауссовского поля (и дисперсии его спектральных отсчетов λ_k разложения в ряд Карунена — Лоэва) пропорциональна квадрату средней плотности $\bar{N}_\Phi$ потока фотонов от изображения и времени накопления, которое примем равным времени кадра $T_к$ [1]:

$$P=\bar{N}^2_\Phi T^2_к. \tag{3.7}$$

Шумы источников в схеме рис. 1.2 считаем не коррелированными, не зависящими от сигнала. В общем случае шум канала ξ_k можно уменьшать пропорционально уменьшению кадровой частоты многократным обращением к каналу (многократным, неразрушающим заряды считыванием сигнала каждого элемента матрицы ПЗС). Ниже будем полагать, что шум канала не зависит от времени накопления, а дисперсия входного шума пропорциональна первой степени времени накопления и квадрату плотности потока фотонов фонового изображения (не несущего полезной информации и подавляемого идеальным линейным кодером):

$$\sigma^2{}_1=\bar{N}^2_\Phi T_к. \tag{3.8}$$

Искажения подвижных изображений, вызываемые скоростным смазом, обычно моделируются сверткой изображения с импульсной характеристикой линейного фильтра в схеме рис. 1.2. При движении изображения относительно накопителя со скоростью V возникает его смещение на длину $VT_к$ в направлении движения изображения. В простейшем случае движения вдоль одной из границ поля зрения (вдоль строки или столбца фотоприемника) искажения скоростного смаза описываются наиболее просто и матрица искажений преобразуется в прямоугольную импульсную характеристику

$$h(t)=\frac{1}{VT_к}\operatorname{rect}\frac{t-VT_к}{VT_к}. \tag{3.9}$$

Воспользуемся для упрощения предположением о независимости телевизионных кадров, равномерности движения изображения

и экспоненциальности корреляционной функции изображения $K(t)=K(0)\times\exp|t/\tau|$. Здесь τ — эквивалентная ширина корреляционной функции, которой соответствует энергетический спектр, описываемый полиномом Баттерворта первого порядка с эквивалентной шириной спектра пространственных частот F [8, 14]:

$$\Lambda_k=F/(F^2+k^2).$$

Эта простая модель соответствует случайному процессу с наиболее медленным убыванием спектра при его заданной эквивалентной ширине.

Для применения формулы (2.61) идеального линейного кодека следует найти преобразование импульсной характеристики (3.9) по базису $\{\varphi_k\}$ Карунена — Лоэва для полезного сигнала со спектром Карунена — Лоэва, описываемым полиномом Баттерворта первого порядка

$$\lambda_k=\Lambda_k P=\Lambda_k\bar{N}^2_\Phi T^2_к. \tag{3.10}$$

Собственные функции $\{\varphi_k\}$ для этого сигнала являются косинусоидами и синусоидами, частоты которых не находятся в гармоническом соответствии. Вместе с тем при больших номерах базисных функций сигнала со спектром (3.10) имеет место их асимптотическое приближение к базису Фурье с косинусоидами и синусоидами, находящимися в гармоническом соотношении [8]. Корреляционная матрица такого сигнала симметрична относительно главной диагонали и является Теплицевой матрицей [31]. Матрица коэффициентов дискретного преобразования Фурье Теплицевой матрицы при увеличении ее размеров асимптотически стремится к диагональной матрице. Это позволяет дискретное преобразование Фурье такого сигнала приближенно считать преобразованием Карунена — Лоэва. Иными словами, при большом числе элементов разложения $N_э$ собственные функции $\{\varphi_k\}$ сигнала со спектром (3.10) асимптотически приближаются к базису дискретного преобразования Фурье. Такое асимптотическое поведение базисных функций, типичное для стационарных процессов, позволяет приближенно считать спектр β_k разложения импульсной характеристики (3.9) равным его фурье-преобразованию:

$$\beta_k=\frac{\sin\pi k VT_k}{\pi k VT_к}. \tag{3.11}$$

Подстановка характеристик сигнала, шума и искажения (3.7)—(3.11) в формулу (2.61) для идеального кодера показывает, что он вырожден, а при тех номерах k, где входное отношение сигнал-шум велико, имеет место подъем частотной характеристики, компенсирующий искажение, вносимое фильтром $\{\beta_k\}$ (рис. 3.1).

Особенностью наблюдения подвижных изображений является зависимость как отношения полезного сигнала к шуму, так и искажений скоростного смаза от времени накопления. Подстановка этих зависимостей сигнала, шума и искажения в формулу (2.51) для взаимной информации между полезным оптическим сигналом

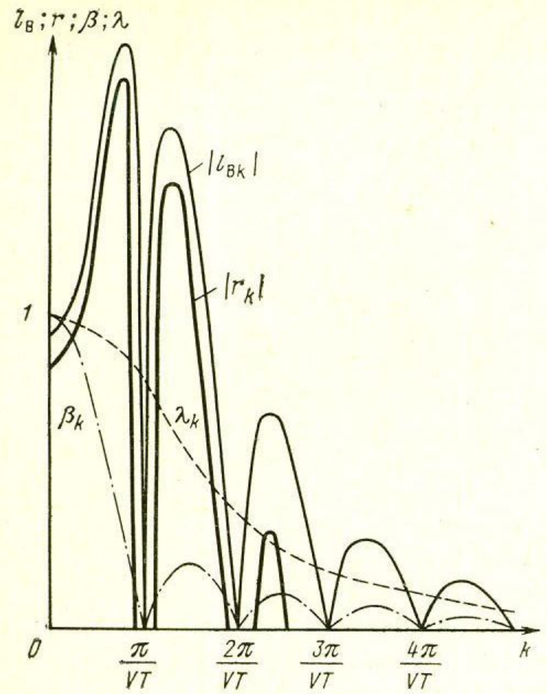


Рис. 3.1

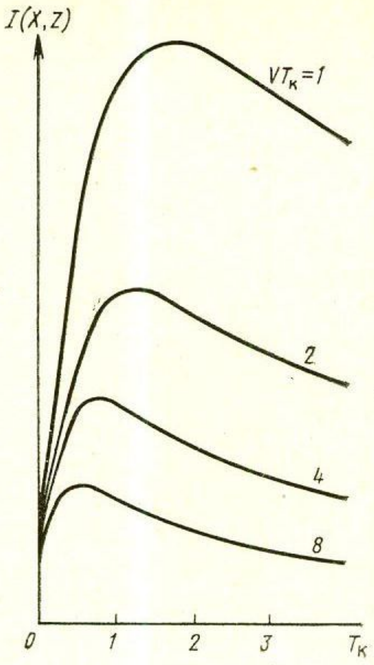


Рис. 3.2

Рис. 3.1. Спектральная характеристика идеального линейного кодера $r_{k\,опт}$ и винеровского фильтра $l_{в\,k}$ для сигнала со спектром λ_k Баттерворта первого порядка при линейном скоростном смазе изображения β_k

Рис. 3.2. Зависимость количества информации от времени кадра

и кодом выявляет зависимость передаваемого в кадре количества информации от времени накопления

$$I[X(t),\ Z(t)] = \frac{1}{2}\sum_{k=I}^{N_э}\log\left[1+\frac{\bar{N}_ф^2\,\Lambda_k\sin^2(\pi k VT_к)\,r_{k\,опт}^2}{\pi^2 k^2 V^2 T_к^2\left(r_{k\,опт}^2\,\bar{N}_ф^2\,T_к+\sigma_2^2\right)}\right], \quad (3.12)$$

где $r_{k\,опт}$ определяется формулой (2.61).

При построении видеоинформационных систем решается традиционная задача минимизации количества информации $I[Y(t), Z(t)]$, требуемого для передачи полезного сигнала с заданной ошибкой. При этом допустимая ошибка не может назначаться произвольно, а зависит от накопленного за кадр количества информации. Информация о полезном сигнале накапливается в течение времени накопления не монотонно. Вначале преобладает рост накопления энергии полезного сигнала, затем все сильнее сказывается рост искажения скоростного смаза. Расчет взаимной информации по формуле (3.12) подтверждает наличие максимума при некотором значении времени кадра, зависящем от скорости движения изображения (рис. 3.2).

Известно, что оптимальное время кадра определяется компромиссом между увеличением отношения сигнал-шум и увеличением смаза [1]. Такой компромисс имеет место и при максимизации количества информации о полезном сигнале в кадре. Определение оптимального времени кадра на основе формулы (3.12) производится обычным путем нахождения производной и приравнивания ее нулю:

$$dI[X(t),\ Z(t)]/dT_k = 0.$$

Аналитическое решение такого уравнения легко получается при аппроксимациях спектра сигнала λ_k и искажения β_k в виде прямоугольных функций. Это грубое приближение дает простую формулу для оптимального времени кадра, которое совпадает с обеспечивающим заданную разрешающую способность [1]:

$$T_к = 1/2FV. \quad (3.13)$$

Для нахождения оптимального времени кадра при более точных аппроксимациях по формуле (3.12) требуется численное интегрирование. Для спектра сигнала (3.10) оптимальное время кадра примерно равно [32]:

$$T_к \approx 1/2F^{1/3}V^{2/3}. \quad (3.14)$$

Количество информации в кадре (3.12) и оптимальное время кадра (3.14) с увеличением скорости сдвига изображения уменьшаются (рис. 3.2). Это не мешает при весьма слабых шумах на входе и в канале восстановлению смазанных изображений с помощью соответствующего декодера даже при большом значении смаза $VT_к \gg 1$ [54—56, 63]. Однако при оптимальном времени кадра, обеспечивающем малый смаз, количество информации в кадре будет больше, а ошибка восстановления меньше, чем при декодировании сильно смазанных изображений. Из-за влияния источников шума смаз обратим лишь частично. Поэтому, хотя искажающий фильтр в схеме рис. 1.2 и не варьируется, уменьшение его влияния достигается слежением за изображением. Таким образом, максимум информации о подвижном объекте достигается при совместной оптимизации кодера и времени кадра, а минимум ошибки передачи требует кроме того оптимального декодера. Оптимальное время кадра (3.13) и (3.14) соответствует смазу $VT_к$ вполэлемента разложения. Поэтому идеальный кодер подвижных изображений имеет спектральную характеристику, внешне отличающуюся

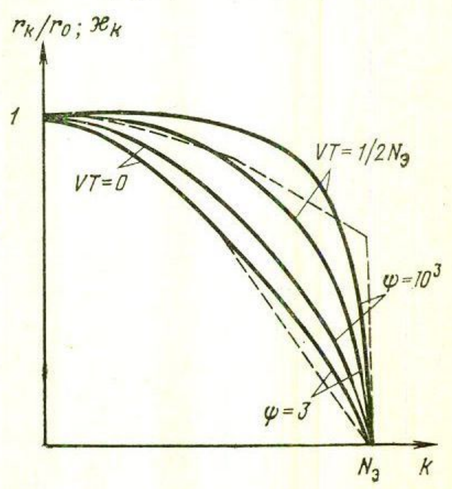


Рис. 3.3. Спектральная характеристика разложения сигнала по прямоугольным функциям $\varkappa_k$ (границы области неинвариантны к сдвигу, штриховая линия) и идеального линейного кодера r_k при оптимальном времени кадра

от показанной на рис. 3.1 (построенной для большой величины смаза, для которой характерны несколько участков спектра с ненулевыми коэффициентами). Для малой величины смаза вполэлемента разложения спектральная характеристика оптимального линейного кодера имеет лишь одну зону ненулевых коэффициентов. При этом спектральная характеристика (передаточная функция) разложения по прямоугольным функциям (при оптимальном выборе размеров элементов разложения Δ и времени кадра $T_к$) в среднем лишь на 8% отличается от спектральной характеристики оптимального кодера (рис. 3.3). Такое жесткое условие малой величины скоростного смаза вынуждает изыскивать возможности для повышения кадровой частоты телевизионных камер на ПЗС [41].

3.3. КОДИРОВАНИЕ В ПРОЦЕССЕ НАКОПЛЕНИЯ ФОТОНОВ

В последнее десятилетие значительные усилия были затрачены на совершенствование ПЗС. В результате из фотоприемников, которые, как считалось раньше, «вряд ли станут приборами массового производства» [42], ПЗС превратились в «один из главных фотоприемников излучения видимого диапазона» [35]. В лучших ПЗС шумы канала считывания снижены до нескольких электронов на зарядовый пакет, что обеспечивает динамический диапазон более 200 тыс. раз. Число элементов выпускаемых матриц достигает 2048×2048 и не является предельным. Темновые токи снижены до нескольких десятков пикоампер на квадратный сантиметр, что позволяет отказаться от неизбежного ранее охлаждения при временах накопления до 10 с [39]. Быстродействующие ПЗС работают на тактовых частотах в десятки мегагерц, достаточных для использования их в вещательных телевизионных системах высокой четкости [35]. Отмеченные свойства в сочетании со стабильностью растра, прочностью, надежностью, а также малыми габаритными размерами и энергопотреблением, позволяют с успехом использовать ПЗС практически во всех областях научного, промышленного и вещательного телевидения.

К настоящему времени накопилось большое число публикаций, посвященных разработке, измерению параметров и практическому применению ПЗС в различных телевизионных системах. Часто ПЗС рассматривают как твердотельный аналог (заменитель) передающей электронно-лучевой трубки, отличающийся от нее только улучшенными массогабаритными характеристиками и технологией изготовления.

С нашей точки зрения главное достоинство ПЗС в их новых функциональных возможностях. Метрологические возможности ПЗС, обусловленные жестким растром, линейностью световой характеристики и низким шумом канала считывания были изложены в [2]. Здесь мы рассмотрим новые функциональные возможности адаптивного камерного кодирования. Камерное кодирование в реальном времени охватывает преобразования «от света до цифры»

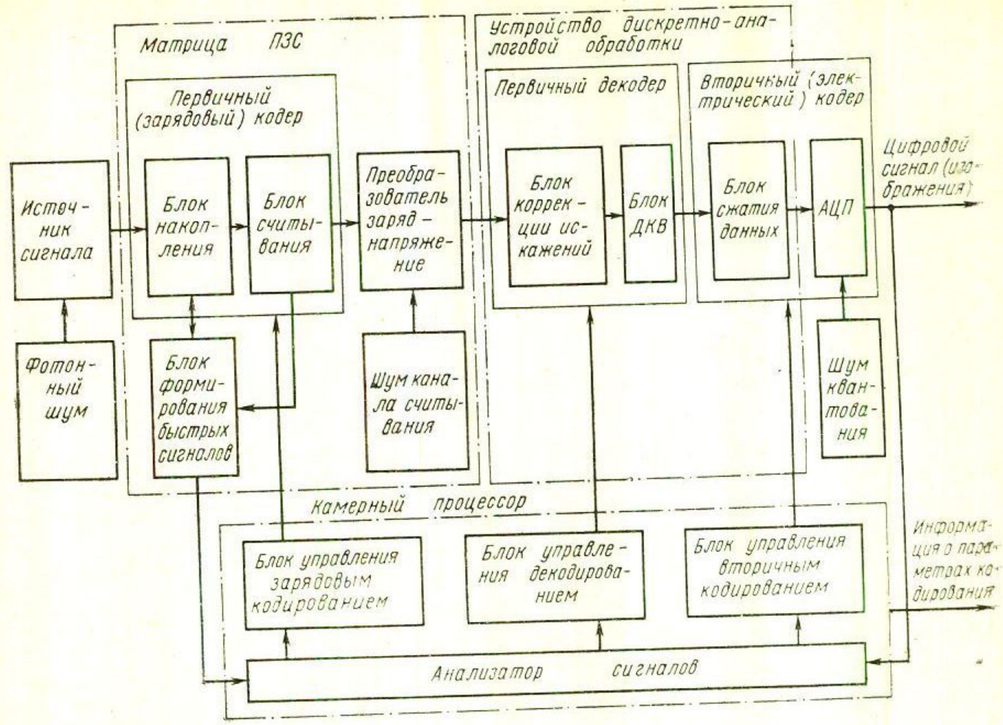


Рис. 3.4. Структурная схема камерного кодирования

в соответствии со схемой (рис. 3.4). Разделение процесса кодирования на два этапа обусловлено наличием двух шумящих каналов связи в ПЗС-камере: аналогового, характеризующегося шумом канала считывания фотоприемника, и цифрового. Матрица ПЗС реализует первичное (зарядовое) кодирование. Оно необходимо для адаптивного согласования динамического диапазона и чувствительности ПЗС с изменяющейся плотностью потока фотонов входной смеси сигнала и шума. Первичный кодер содержит блоки накопления и считывания. Он позволяет оперативно перестраивать время накопления заряда, частоты считывания кадров, строк, элементов и апертуру (площадь) считывания элементов. Кодирование зарядового изображения обеспечивает минимальные потери информации при прохождении сигнала через первый шумящий канал телевизионной камеры. Снижение потерь сопровождается появлением дополнительной информации о переменных параметрах кодирования, которую необходимо передавать далее синхронно с сигнальной информацией. Количество дополнительной информации характеризует отличие адаптивного режима ПЗС от стандартного режима телевизионной развертки и растет с увеличением степеней свободы управления фотоприемником. Нужно отметить, что характерной чертой безразверточных параллельных фотоприемников является значительное преобладание этой информации (временной и геометрической), которая становится в них основной, над обычной амплитудной информацией. Рассмотренное первичное

адаптивное кодирование позволяет передать большее количество информации о сигнале ценой отхода от режима стандартной телевизионной развертки.

Вторичное (электрическое) кодирование направлено на борьбу с шумами квантования. Оно необходимо потому, что динамический диапазон АЦП, как правило, на несколько порядков у́же динамического диапазона фотоприемной матрицы ПЗС. Вторичное кодирование реализуется в устройстве дискретно-аналоговой обработки сигнала (см. рис. 3.4) и обеспечивает сжатие динамического диапазона сигнала до АЦП. Возможности этого этапа кодирования были подробно рассмотрены в [2] и обусловлены наличием интенсивных фоновых составляющих в сигнале изображения. Эта помеха может оперативно оцениваться и вычитаться из суммарного сигнала в реальном времени.

Этапу вторичного кодирования предшествует декодирование сигнала после первого канала связи. Получателю нужна информация лишь о полезном объекте. Поэтому декодер телекамеры должен восстанавливать не все изображение, подавляя фоновую помеху. Декодирующий фильтр не будет обратный кодирующему. Первичный декодер содержит схему двойной коррелированной выборки (ДКВ) и блок коррекции передаточной функции ПЗС. Декодер компенсирует искажения, вносимые ПЗС, и обеспечивает линейность обмена напряжения выходного сигнала на время накопления или площадь элемента накопления в режиме адаптивного кодирования. Схема вторичного декодера, расположенного после цифрового канала связи, определяется типом устройства-получателя информации. При наблюдении на экране видеоконтрольного устройства с выхода декодера должен поступать стандартный телевизионный сигнал. При подключении автоматического исполнительного устройства роль декодера выполняет классификатор изображений (см. 4.2).

В камерном кодировании (см. рис. 3.4) определяющую роль играет первичный кодер, поскольку он расположен ближе ко входу камеры. Важным достоинством ПЗС-камеры является высокая эффективность этого кодера. Она обусловлена двумя факторами:

широкими диапазонами перестройки параметров матрицы ПЗС, позволяющими адаптироваться к оптическому сигналу, изменяющемуся в диапазоне изменения освещенностей до 10^{12} раз [2];

возможностью скоростного управления кодированием по дополнительным не разрушающим заряд сигналам, снимаемый с ПЗС.

Первичное адаптивное кодирование в ПЗС-матрице можно разделить на два взаимосвязанных этапа: кодирование при накоплении и кодирование при считывании. Кодирование в процессе накопления осуществляется одинаково для всех элементов изображения и является главным средством борьбы с шумом фотоэлектронов. Основным параметром этого кодирования является частота кадров, а вспомогательным — время накопления в режиме «электронного затвора» [2]. Рассмотрим адаптивные режимы ПЗС, реализующие кодирование при накоплении.

Режим переменной частоты кадров является важнейшим режимом при реализации адаптивного кодирования [38]. Изменяя частоту кадров, можно пропорционально обменивать чувствительность ПЗС-камеры на быстродействие.

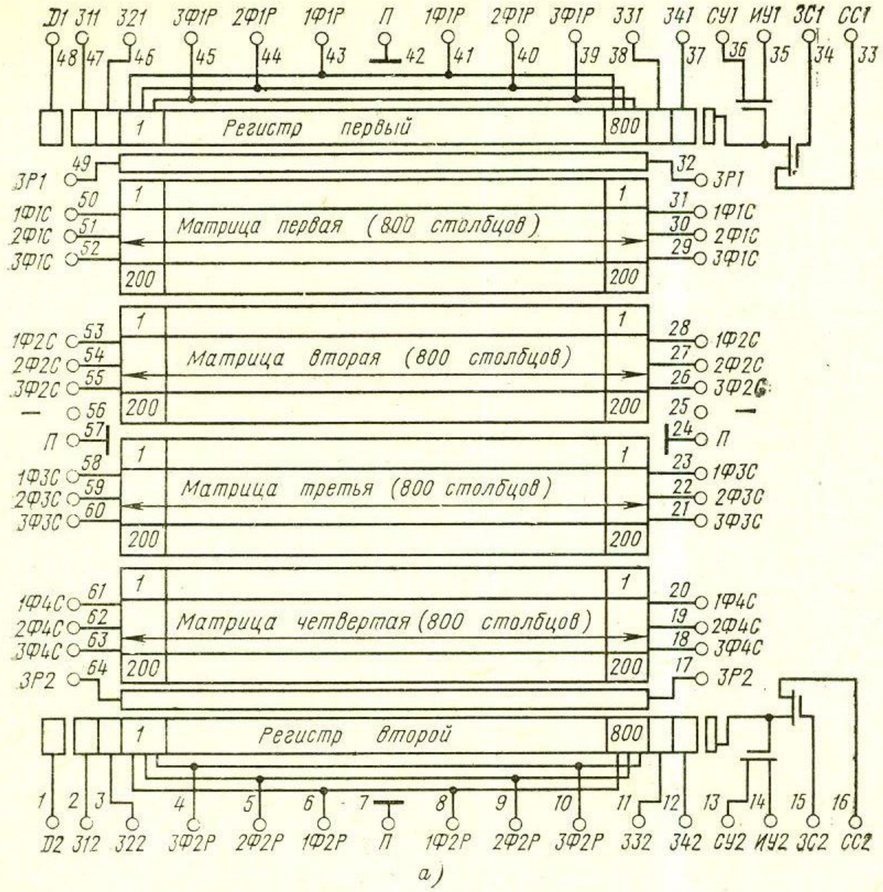
Регулирование чувствительности необходимо из-за нехватки динамического диапазона фотоприемной матрицы ПЗС, который значительно у́же диапазона входных освещенностей. Обычно чувствительность камеры регулируют с помощью светофильтров, которые уменьшают одновременно с числом фотонов фона и число фотонов полезного сигнала. С появлением ПЗС стали применять электронный аналог светофильтра, т. е. уменьшать время накопления при постоянной частоте кадров. Но при этом также теряются фотоны, так как накопление идет с промежутками. Эти потери приводят к снижению отношения сигнал-шум. В ПЗС-камерах имеется возможность регулировать чувствительность и не теряя фотонов. Для этого при увеличении освещенности фона нужно повышать частоту кадров. Конечно, число фотонов в каждом кадре также уменьшается. Но если из электрического сигнала каждого кадра, применяя известные методы обработки, вычесть фон, а затем кадры сложить, то число фотонов сигнала останется прежним, а отношение сигнал-фон улучшится [38]. Поэтому регулирование частоты кадров является важнейшей составной частью адаптации фотоприемника. Практически адаптация частоты кадров стала возможной лишь при появлении ПЗС, тактовые частоты которых могут легко перестраиваться в широком диапазоне.

Наименьшая частота кадров ограничена темновым током ПЗС. В широком диапазоне температур плотность темнового тока ПЗС снижается примерно в 2 раза при уменьшении температуры на каждые 8 °С [34]. Учитывая эту зависимость и задавая максимально допустимое отношение $b_т$ уровня темнового тока к уровню насыщения, наименьшую частоту кадров можно найти по формуле [40]

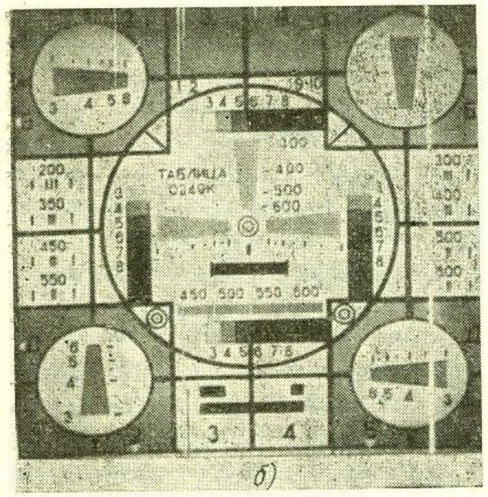
$$F_{к\ мин} = S_э J_0 \exp\{0{,}0866\,(T-293)\}/10^6 b_т q_e Q_н,$$

где $S_э$ — суммарная площадь элемента в секциях накопления и хранения заряда, см^2; J_0 — начальная плотность темнового тока при температуре +20 °С, нА/см^2; T — температура в градусах Кельвина; $Q_н$ — управляющая способность (емкость потенциальной ямы) элемента секции накопления ПЗС, в электронах; $q_e = 1{,}6 \cdot 10^{-16}$ К — заряд электрона.

Например, для матрицы ПЗС 1200ЦМ7 со средним значением плотности темнового тока 0,1 нА/см^2 при комнатной температуре и при $b_т = 0{,}03$ наименьшая частота кадров составит 0,3 Гц. При этом за время кадра темновой ток будет заполнять не более 3% полной емкости потенциальных ям. Охлаждение ПЗС позволяет значительно снизить наименьшую частоту кадров. Так, включение встроенного в матрицу с числом элементов 800×800 [45] термохолодильника обеспечивает время кадра до 10 мин при том же относительном значении уровня темнового заряда. Использование



а)



б)

Рис. 3.5. Структурная схема матрицы ПЗС ФПЗС14 с числом элементов 800×800 (*а*) и изображение (*б*), формируемое в режиме засветки всей площади матрицы ПЗС

криогенного охлаждения позволяет увеличить время кадра на несколько порядков. Нужно отметить, что в ПЗС есть возможность понизить наименьшую кадровую частоту и без использования охлаждения, учтя, что плотность темнового тока уменьшается при снижении напряжений на электроде, реализующем накопление заряда [2]. Для этого либо накапливают заряд при подаче на электрод ПЗС пилообразного напряжения, либо вводят автоматическое включение «микрорежима» при малых уровнях освещенности [43]. С использованием этих режимов даже без охлаждения большеформатной матрицы ПЗС (рис. 3.5,*а*) можно получить качественные изображения на частоте кадров 0,1 Гц (рис. 3.5,*б*).

Наибольшая частота кадров ограничена быстродействием ПЗС. Имеются данные о превосходстве ПЗС с объемным каналом над ПЗС с поверхностным каналом [2, 35]. Однако в значительной мере быстродействие серийных приборов определяется конструктивным исполнением электродной структуры, подводящих шин и топологией выходных устройств. Кроме того, существенную роль играет форма управляющих импульсов на высоких частотах. В [41] показано, что эффективность переноса заряда отечественных *p*-канальных матриц на высоких частотах значительно улучшается, если в конце вершины управляющего импульса сформирован дополнительный положительный импульс с амплитудой 1...3 В. Этот импульс предотвращает обратный перенос заряда при его быстром выталкивании из потенциальной ямы каждого элемента. Для *n*-канальных приборов этот импульс меняет полярность.

Стремление к низкому уровню шума канала считывания серийных ПЗС заставляет уменьшать емкость, а следовательно, и площадь затвора выходного транзистора устройства считывания. В результате снижается крутизна МОП-транзистора, что уменьшает его коэффициент передачи на высоких частотах. При стандартном включении выходного транзистора в режиме истокового повторителя коэффициент передачи по напряжению на частоте 20 МГц обычно не превышает 0,2. Применение следящей обратной связи (рис. 3.6,*а*) обеспечивает единичный коэффициент передачи до частоты 20 МГц. В результате выходной каскад не лимитирует быстродействие ПЗС. С выхода матрицы ПЗС снимается выходной сигнал в смеси с превышающей его в несколько раз синхронной тактовой помехой. Возможность подавления тактовой помехи на высоких частотах режекторным фильтром ограничена (из-за одновременного подавления высокочастотных компонентов полезного сигнала). Вместо него следует применить мостовую схему выборки-хранения (СВХ) на рис. 3.6,*б*, подавляющую тактовую помеху более чем на порядок [41]. Настройка схемы по максимуму отношения сигнал-помеха производится резистором *R6*, определяющим зарядный ток во время выборки. Приведенные на рис. 3.6 схемы работоспособны на тактовых частотах регистра ПЗС до 30 МГц. Использование СВЧ-транзисторов позволит дополнительно увеличить максимальную тактовую частоту в ПЗС-камере.

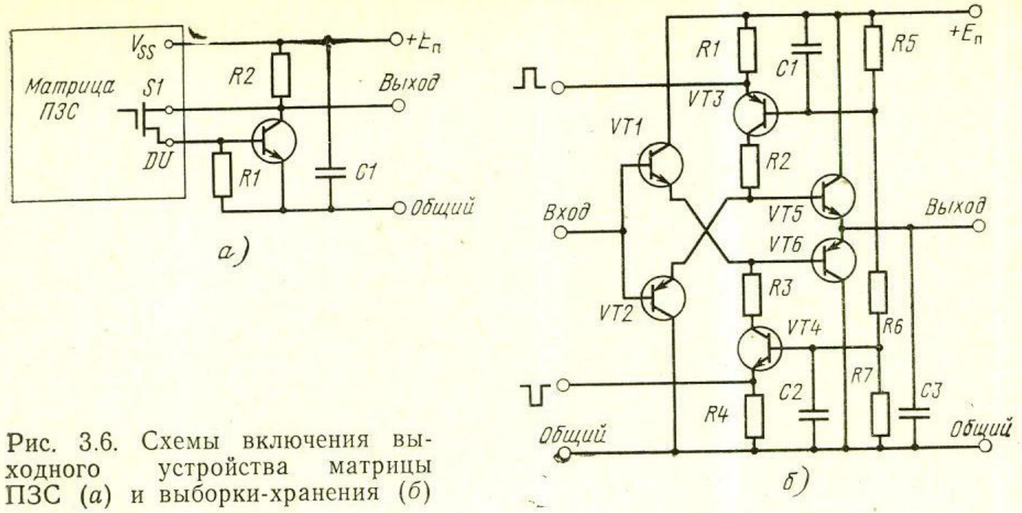


Рис. 3.6. Схемы включения выходного устройства матрицы ПЗС (*а*) и выборки-хранения (*б*)

Эксперимент по исследованию высокочастотных свойств ПЗС показал, что быстродействие определяет в первую очередь не канал переноса заряда, а конструкция и электрические параметры проводников, подводящих управляющие импульсы: алюминиевых шин на кристалле и внешних соединительных проводников. Из-за несимметричного расположения подводящих шин и их высокого сопротивления в матрице 800×800 элементов при использовании стандартных управляющих импульсов эффективность переноса заряда в выходном регистре становится неприемлемо плохой уже на тактовой частоте $F_т=5$ МГц (рис. 3.7). Использование асимметричной расфазировки и указанных выше дополнительных положительных выбросов на управляющих импульсах расширяет диапазон рабочих частот до 15 МГц. В матрице ПЗС К1200ЦМ1 удается обеспечить хорошую эффективность переноса заряда до 20 МГц (рис. 3.8,*б*). Матрица работоспособна и на частоте 30 МГц (рис. 3.8,*в*), однако четкость на строке ухудшается в 3 раза. Но и в этом случае оказывается лимитирующим не быстродействие поверхностного канала переноса заряда, а сопротивление R подводящих напряжение к электродам регистра и подложке алюминиевых шин [41]. При одинаковой конструкции матрицы с объемным каналом имеют более высокое быстродействие и из-за того, что ем-

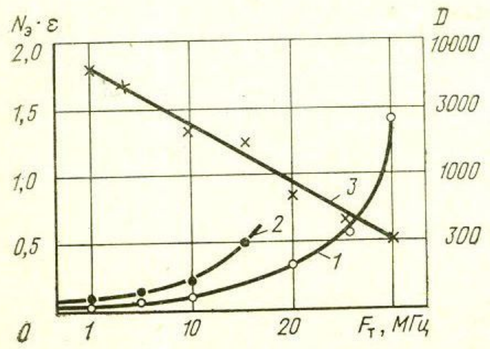


Рис. 3.7. Зависимость неэффективности переноса заряда $N_э$ ε в выходном регистре матрицы ПЗС К1200ЦМ1 (*1*), ФПЗС14 (*2*) и динамического диапазона D матрицы К1200ЦМ1 (*3*) от тактовой частоты регистра $F_т$

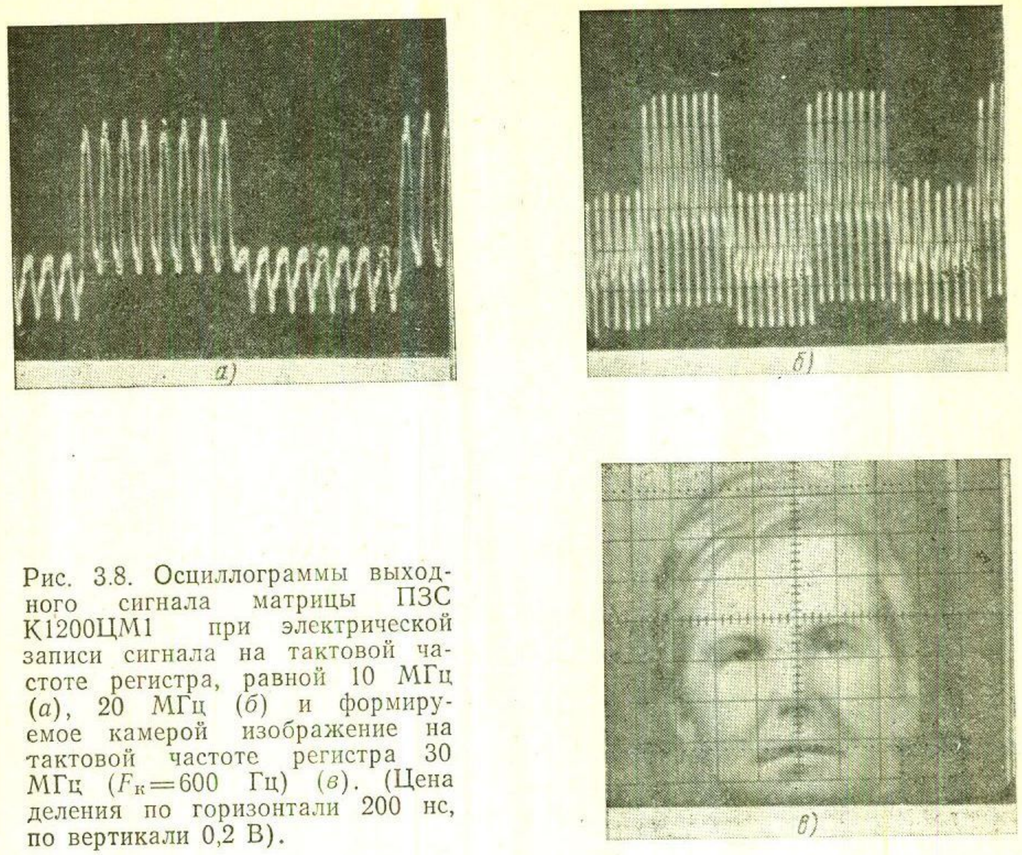


Рис. 3.8. Осциллограммы выходного сигнала матрицы ПЗС К1200ЦМ1 при электрической записи сигнала на тактовой частоте регистра, равной 10 МГц (*а*), 20 МГц (*б*) и формируемое камерой изображение на тактовой частоте регистра 30 МГц ($F_к=600$ Гц) (*в*). (Цена деления по горизонтали 200 нс, по вертикали 0,2 В).

кость C электродной структуры объемного канала оказывается меньшей, в результате электрический RC-фильтр, искажающий подводимые к ней импульсы, имеет меньшую постоянную времени.

Измерив максимальную частоту выходных регистров $f_{р\,макс}$ для матрицы формата $N_{э1}\times N_{э2}$, имеющей $m_{вых}$ параллельно работающих выходов, максимальную частоту кадров можно определить

$$F_{к\,макс}=f_{р\,макс}b_а m_{вых}/N_{э1}N_{э2},$$

где $b_а$ — коэффициент укрупнения апертуры по кадру, равный числу строк, суммируемых в выходных регистрах ПЗС.

Например, для матриц К1200ЦМ1 и К1200ЦМ7 максимальные частоты кадров в режиме одноэлементной апертуры составляют соответственно 600 и 200 Гц, а при восьмикратном укрупнении апертуры по кадру достигают 4,8 и 1,6 кГц соответственно.

Режим переменного времени накопления используется при постоянной частоте кадров как средство защиты от пересветки. В ряде случаев такой режим нужен как средство подавления фона, когда полезный сигнал импульсный и заранее известно время его действия. Диапазон изменения (уменьшения) времени накопления ограничивают два фактора. На низких частотах кадров основным фактором является неидеальная оптическая изоляция секций на-

копления и хранения. За время отвода избыточного заряда из секции накопления в стоковую область, небольшая часть света проникает в секцию хранения, вызывая образование помехи. В стандартных ПЗС уровень помехи становится недопустимо большим при отношениях времени накопления ко времени кадра, меньших 1/500. Подведение светозащитного экрана непосредственно к поверхности секции памяти позволяет работать при отношениях времени накопления ко времени кадра 1/10000 [40].

На высоких частотах кадров главным ограничением становится смаз изображения, возникающий из-за накопления заряда от изображения во время переноса заряда из секции накопления в секцию хранения. Количество паразитного заряда от смаза обратно пропорционально частоте переноса в параллельных секциях, которая ограничена значениями 1—2 МГц в зависимости от типа матрицы ПЗС. Поэтому минимальное время накопления в матрицах К1200ЦМ1 и К1200ЦМ7 должно быть не менее 100 и 200 мкс соответственно. Режим переменного времени накопления используется в адаптивных камерах как вспомогательный, когда уже достигнута максимальная частота кадров, но имеется возможность пересветки фотоприемника.

3.4. ДОПОЛНИТЕЛЬНЫЕ КАНАЛЫ СЧИТЫВАНИЯ В ПЗС-ФОТОПРИЕМНИКЕ

Любая секция или регистр матрицы ПЗС состоит из близко расположенных конденсаторов, объединенных двумя видами связи: зарядовой и электрической. Двойная связь придает массиву конденсаторов особые свойства:

1) прикладывая к обкладкам конденсаторов (электродам ПЗС) определенную последовательность управляющих импульсов, можно перемещать вдоль поверхности локализованные зарядовые пакеты;

2) измеряя напряжение или ток в электрической цепи конденсаторов (в цепи фазного электрода ПЗС), можно определять находящийся в потенциальных ямах заряд и его приращения.

Второе свойство ПЗС, в отличие от первого, оказалось к настоящему времени недостаточно изученным. Тем не менее оно открывает новые возможности адаптивного кодирования зарядового изображения. Любой фазный электрод можно включить в режим считывания сигнала, т. е. организовать в матрице ПЗС дополнительные каналы считывания сигналов. Они дадут дополнительную информацию, причем без разрушения и искажения зарядовых пакетов, продвигающихся по основному каналу считывания [109].

Недостаточное внимание к дополнительным каналам считывания в ПЗС объясняется широко распространенным мнением об их малой информативности. Считается, что с электрода секции или регистра ПЗС можно снять лишь сигналы, характеризующие интегральный уровень заряда. Однако двойная связь ПЗС-конденсаторов позволяет из сигнала фазного электрода извлечь значительно большее количество информации, организовав режим амплитудной развертки [46]. В этом режиме, подробно рассматриваемом в 3.5, из сигнала фазного электрода формируется сигнал распределения заряда по уровням. Такой сигнал несет количество информации, достаточное для управления камерным кодированием.

Преимущество дополнительной информации заключается в том, что она извлекается с минимальной задержкой и является опережающей по отношению к основной информации, формируемой методом телевизионной развертки. Уменьшение времени задержки объясняется тем, что дополнительные каналы расположены ближе ко входу фотоприемника. Преимущество дополнительной информации реализуется наиболее полно, когда она используется для управления процессами накопления и считывания сигнала в основном канале ПЗС. При этом практически полностью устраняются динамические ошибки регулирования, которые возникают при обычном управлении с использованием обратной связи по основному каналу.

Рассмотрим подробнее возможности дополнительных каналов считывания, которые могут быть организованы в универсальных матрицах ПЗС, таких как К1200ЦМ1, К1200ЦМ7 и др. Каналы считывания разделяются на две группы: последовательные, разрушающие заряд каналы *2* и *5*, и параллельные, не разрушающие заряда каналы *1*, *3*, *4*, *6* (рис. 3.9). В последних имеется возможность многократного считывания информации без ее разрушения. Наибольшей информативностью среди этих каналов обладает канал *3*, практическая реализация которого изложена в 3.5 [46]. Выходом этого канала является один из фазных электродов секции накопления матрицы ПЗС. По сигналу этого канала можно контролировать процесс накопления практически без задержки, следить за динамикой изображения «внутри» телевизионного кадра. Нужно отметить, что динамическая внутрикадровая информация полностью теряется после завершения накопления кадра и отсутствует в других каналах. Поэтому канал считывания *3* (рис. 3.9) незаменим и может эффективно применяться для скоростного управления другими каналами считывания. Анализируя непрерывно изменяющееся распределение накапливаемого заряда на выходе этого канала (см. 3.5), можно обнаружить моменты резких изменений освещенности изображения и, быстро управляя переносом заряда, исключить пересветку ПЗС.

Не разрушающие заряд параллельные каналы считывания *4* и *6* используются для скоростной, опережающей основной сигнал, автоматической настройки декодера и декоррелятора (см. рис. 3.9) основного канала. С выходов этих каналов поступает информация о распределении амплитуд сигнала всего кадра и очередной строки, которые только через некоторое время появятся на выходе основного канала *5*. По этим сигналам можно заранее настроить видеотракт основного канала. При этом полностью исключаются динамические ошибки таких инерционных узлов, как

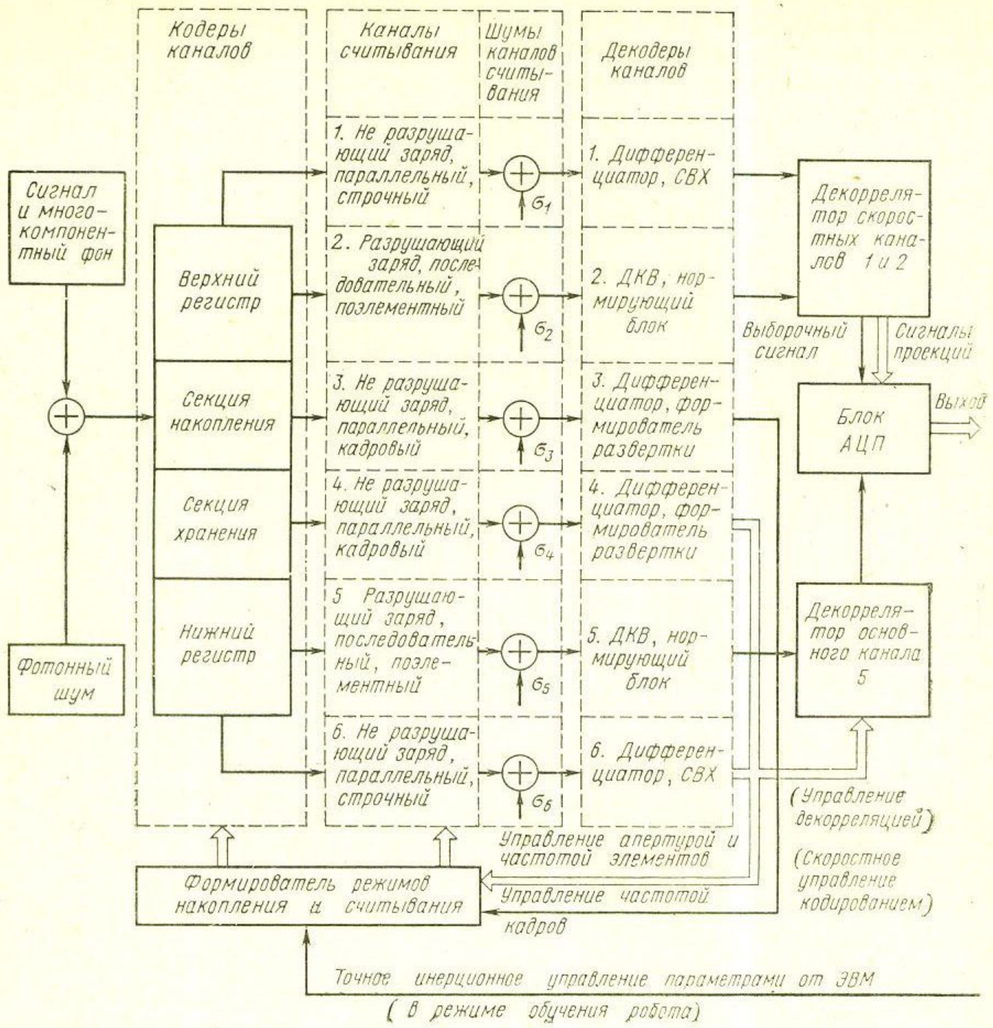


Рис. 3.9. Структурная схема ПЗС-камеры, реализующей адаптивное многоканальное считывание сигналов, $\sigma_1 \ldots \sigma_6$ — среднеквадратические значения шумов считывания (ДКВ — двойная коррелированная выборка)

автоматические регуляторы усиления и размаха (АРУ и АРР) [2]. Заметный выигрыш может быть получен при использовании информации в каналах *4* и *6* для управления режимами переменной скорости считывания или переменной площади считывающей апертуры в канале *5* (см. 3.6). Обычно адаптация режима считывания производится по результату анализа сигнала в том же канале, режим которого адаптируется. При этом возникают два вида ошибок: динамическая, из-за задержки управления на время кадра, и линейная. Используя опережающую информацию с выходов каналов *4* и *6*, можно уменьшать ошибки регулирования в канале *5*.

Каналы считывания *3* и *4*, подключенные к электродам секций накопления и хранения матрицы ПЗС, передают лишь амплитудную информацию. Поэтому они не могут использоваться для быстрого измерения координат объектов. В стандартных камерах для этого используют канал считывания *5*. Но скорость измерений, определяемая частотой кадров в таком режиме, не велика (см. 3.3) и в ряде случаев не достаточна для контроля за быстро протекающими процессами. Значительно повысить скорость измерения координат объектов можно, используя дополнительные каналы считывания *1* и *2*. Рассмотрим механизм формирования сигналов, который можно назвать режимом считывания проекций изображения. Для его реализации используются два дополнительных канала: электрод одной из фаз верхнего регистра (канал *1*) и выходное устройство этого регистра (канал *2*, рис. 3.9). Режим считывания проекций осуществляется в два этапа. После обязательного накопления сигнала производят быстрый перенос всего зарядового изображения в верхний регистр, при одновременном измерении тока в цепи фазного электрода регистра. Как только зарядовые пакеты от изображения объекта достигнут верхнего регистра, в цепи его электрода появится ток. Измерив число тактов вертикального переноса от его начала до момента, когда ток примет наибольшее значение, получим координату *Y* объекта. Отметим, что на этом этапе измерения зарядовые пакеты не разрушаются, а только суммируются по вертикали, оставаясь в верхнем регистре. На втором этапе регистр включают в режим переноса строки и аналогично измеряют координату *X* объекта, при этом происходит разрушение зарядовых пакетов. Сформированные так сигналы на выходах дополнительных каналов считывания *1* и *2* представляют из себя проекции изображения по осям *X* и *Y*.

Режим формирования проекций сигнала был реализован в адаптивной телевизионной камере на матрице ПЗС К1200ЦМ7, в которой помимо формирования сигнала изображения в канале *5*, производилось скоростное измерение двух вспыхивающих на время 30 нс горизонтально расположенных световых пятен (рис.

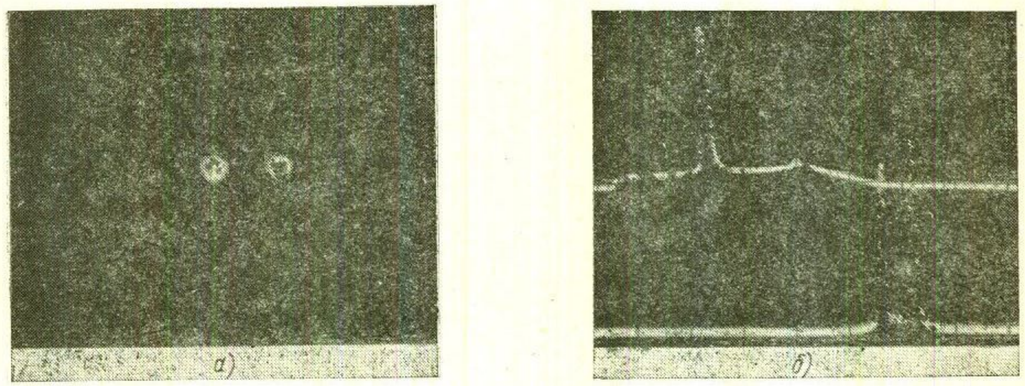


Рис. 3.10. Изображение на экране видеоконтрольного устройства (ВКУ) световых пятен (*а*) и осциллограммы проекций этого изображения на горизонтальную и вертикальную оси (*б*)

3.10,*а*). Для этого сразу после появления пятен включались два дополнительных канала считывания *1* и *2*. Осциллограммы сигналов на выходах этих каналов (рис. 3.10,*б*) свидетельствуют о достаточной точности и чувствительности скоростного измерения координат, производимого за тысячные доли времени вещательного кадра [109].

Рассмотренные режимы работы дополнительных каналов считывания в матричных ПЗС являются лишь небольшой частью тех возможностей, которые могут быть реализованы в адаптивной телевизионной камере. Для повышения быстродействия и информативности считывания необходимо создание специальных многоучастковых матриц ПЗС, снабженных сервисными секциями и регистрами, позволяющими реализовать различные режимы накопления и считывания для различных участков изображения.

3.5. БЫСТРОДЕЙСТВУЮЩИЕ УПРАВЛЯЮЩИЕ СИГНАЛЫ

В структуре всех многоэлементных фотоприемников, в том числе и ПЗС, заложен принцип последовательной телевизионной развертки. Он позволяет упростить конструкцию и управление фотоприемником за счет съема сигналов с сотен тысяч фоточувствительных элементов через один выход и управления всем массивом фотоприемника с помощью всего лишь двух генераторов разверток.

Такие упрощения уменьшают потенциальные возможности многоэлементных фотоприемников. Ограниченный динамический диапазон массива жестко связанных, работающих в едином режиме фоточувствительных элементов, приводит к снижению точности измерения телевизионной камерой параметров сложного изображения. Лишь некоторые элементы оказываются в условиях, благоприятных для точного измерения. Остальные либо насыщаются на участках изображения с большой освещенностью, либо работают с низким отношением сигнал-шум. В результате средняя по полю точность измерения уровня освещенности оказывается значительно ниже той, которая могла бы быть достигнута при согласовании параметров каждого фоточувствительного элемента с характеристиками соответствующих участков оптического изображения. Еще больший проигрыш возникает при использовании телевизионной камеры, реализующей последовательную развертку для анализа динамики изображения, когда производится измерение приращений освещенности.

Перспективным вариантом представляется безразверточная параллельная система фоточувствительных элементов, снабженная устройством независимого параллельного управления и устройством параллельной обработки сигналов, выполненная на одном кристалле. Фоточувствительные элементы такой системы должны обладать возможностями независимого непрерывного неразрушающего считывания накопленного заряда, независимого управления (сброс накопленного заряда) и возможностью объединения до

считывания в группы. Идеальным «рабочим материалом» для создания таких фотоприемников являются ПЗС, приборы с зарядовой инжекцией (ПЗИ), а также аналоговые и цифровые СБИС на базе КМОП-транзисторов. На современном уровне технологии могут быть созданы малоформатные параллельные фотоприемники. Имеется перспективный вариант последовательно-параллельных адаптивных телевизионных систем с внешним накопителем, обладающих характеристиками, приближающимися к параллельным системам [38]. В структуре последовательно-параллельных систем можно выделить три важных момента:

1) распараллеливание фоточувствительного массива ПЗС на крупные участки с независимым управлением и считыванием сигнала;

2) первичная быстродействующая адаптация частоты кадров каждого участка фотоприемника, частоты и площади апертуры считываемых элементов;

3) вторичная адаптация частоты кадров каждого элемента изображения во внешнем цифровом накопителе.

В настоящее время уже созданы матрицы ПЗС с многоучастковым накоплением и считыванием [2]. Вопросы обработки сигнала изображения во внешнем цифровом накопителе также достаточно освещены. Однако потенциальные возможности скоростной адаптации частоты кадров не реализованы. Существующие методы предполагают управление частотой кадров по информации, заключенной в обычном электрическом сигнале изображения. В этом случае полная задержка для i-го элемента j-й строки

$$\tau_{\text{с}} = T_{\text{к}}\left(1+\frac{j}{N_{\text{э}2}}+\frac{i}{N_{\text{э}}}\right)+T_{\text{р}}, \qquad (3.15)$$

где $N_{\text{э}}=N_{\text{э}1}N_{\text{э}2}$ — число элементов фотоприемника; $N_{\text{э}1}$ — число элементов в строке; $N_{\text{э}2}$ — число строк; $T_{\text{к}}$ — время кадра (накопления); $T_{\text{р}}$ — время принятия решения (вычисления) в камерном процессоре.

Эта задержка складывается из времени накопления, времени считывания заряда и времени принятия решения. При этом количество информации в сигнале изображения

$$I[X(t),\ Z(t)] \leqslant N_{\text{э}} \log_2 \mathcal{N}, \qquad (3.16)$$

где $\mathcal{N}$ — число разрешаемых по амплитуде уровней.

Видно, что при выделении управляющих сигналов из сигнала изображения количество информации (3.16) достигается путем значительной задержки (3.15), которая может превышать время двух телевизионных кадров. Эта задержка приводит к значительным ошибкам при наблюдении в условиях быстро изменяющейся освещенности.

Уменьшить время задержки можно изменением механизма считывания сигнала в ПЗС, используя способ формирования сигнала управления чувствительностью телевизионной камеры на ПЗС [46]. Такое неразрушающее заряд считывание сигнала может про-

изводиться многократно в течение времени кадра и не влияет на накапливаемые в потенциальных ямах ПЗС зарядовые пакеты. Задержка при считывании управляющего сигнала $\mathscr{Z}(t)$

$$\tau_у = (N_{эз}+2)/f_т, \tag{3.17}$$

где $N_{эз}$ — число отсчетов разрешаемых уровней заряда в измерении; $f_т$ — тактовая частота отсчетов.

Дополнительная задержка на два такта представляет задержку вычисления, которая возникает за счет двойного дифференцирования, которое, как будет показано ниже, необходимо для выделения информации о распределении заряда. Количество информации в сигнале управления $\mathscr{Z}(t)$, считанном таким образом, меньше содержащегося в сигнале изображения (3.16) и составляет

$$I[X(t),\ \mathscr{Z}(t)] \leqslant \mathscr{N} \log N_э. \tag{3.18}$$

Реально из-за больших емкостей фазных электродов секции накопления шум канала считывания при этом способе больше, чем при считывании сигнала изображения через выходное устройство ПЗС. Поэтому практически количество информации оказывается меньше определяемого (3.18):

$$I[X(t),\ \mathscr{Z}(t)] \leqslant \mathscr{N}' \log N_э, \tag{3.19}$$

где $\mathscr{N}'$ — число разрешенных уровней по амплитуде после двойного дифференцирования.

Знак равенства в (3.16), (3.18) и (3.19) достигается при равномерном распределении яркости.

Несмотря на меньшее количество информации при неразрушающем считывании, решающим обстоятельством в вопросе о целесообразности использования такого сигнала является то, что он может быть получен за значительно меньший интервал времени — реально до 1000 раз быстрее периода кадровой частоты вещательного телевидения.

Поэтому при адаптации по этим сигналам, несмотря на несколько большую статическую ошибку управления, будет достигнута минимальная динамическая ошибка, что крайне важно при наблюдении в условиях быстро изменяющейся освещенности.

Рассмотрим подробно механизм реализации не разрушающего считывания сигнала с электродов секции накопления ПЗС [46]. Пусть накопление заряда в секции накопления происходит под двумя фазными электродами (рис. 3.11). При неравномерной засветке секции накопления в каждом элементе накапливается разное количество заряда. Будем плавно увеличивать потенциал правого на рисунке электрода подачей напряжения U_2 (рис. 3.12,*а*). При этом в каждом элементе глубина правых потенциальных ям уменьшается, а заряд перетекает из правых потенциальных ям в левые. Во внешней цепи левого электрода возникает ток, равный сумме токов в каждом элементе. Сначала этот ток максимальный, так как заряд есть в каждом элементе. По мере возрастания потенциала наступит момент, когда кончится заряд в элементе с

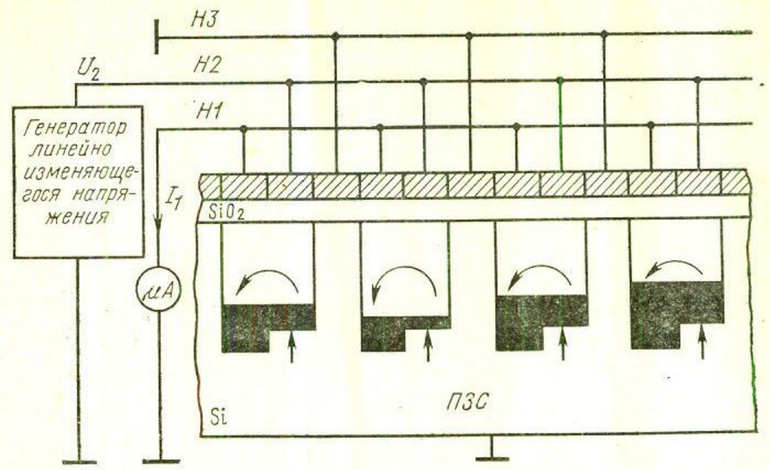


Рис. 3.11. Формирование быстрого управляющего сигнала баз разрушения накапливаемого заряда путем его параллельного переноса

наименьшим количеством заряда. При этом суммарный ток уменьшается (рис. 3.12,*б*). Затем кончится заряд в следующем элементе и ток снова уменьшится. Так будет продолжаться, пока не кончится заряд в элементе с наибольшим зарядом. При этом ток станет равным нулю, а весь заряд секции накопления окажется в левых потенциальных ямах (см. рис. 3.11). В сформированном так сигнале заключена информация о распределении заряда. Чтобы ее выделить, нужно продифференцировать сигнал и поменять знак. В результате получается гистограмма $N_i(q_i)$ (число элементов ПЗС с зарядом q_i) распределения заряда по всей поверхности секции накопления матрицы ПЗС (рис. 3.12,*в*). Таким образом, рассмотренный режим — это не разрушающая и параллельная операция над массивом зарядовых пакетов. Поэтому его

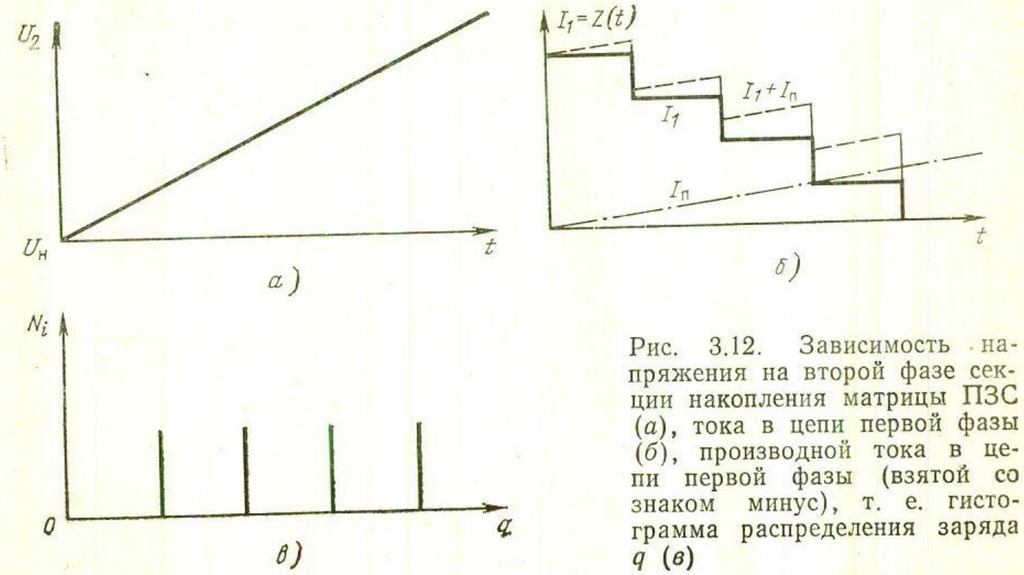


Рис. 3.12. Зависимость напряжения на второй фазе секции накопления матрицы ПЗС (*а*), тока в цепи первой фазы (*б*), производной тока в цепи первой фазы (взятой со знаком минус), т. е. гистограмма распределения заряда q (*в*)

можно проводить многократно и быстро, за тысячную долю времени кадра. Точность считывания ограничена коммутационной помехой. Из-за емкостных связей ПЗС-структуры напряжение развертки накладывается на полезный сигнал (рис. 3.13). Поскольку сигнал $\mathcal{Z}(t)=I_1$ и помеха $I_п$ независимы, можно определить приращение напряжений на измерительном электроде за счет действия каждого тока в отдельности:

$$\Delta U_{1\,\text{внеш}} = \Delta U_2 \frac{C_{12}}{C_{12}+C_{13}+\dfrac{C_{ок1}C_{об1}}{C_{ок1}+C_{об1}}} = b_1 \Delta U_1;$$

$$\Delta U_{1\,\text{внутр}} = \frac{1}{C_{об1}+C_{ок1}+C_{13}+C_{12}} \sum_{i=1,\,j=1}^{MN} Q_{ij} = b_2 \sum_{i=1,\,j=1}^{MN} Q_{ij};$$

$$\Delta U_1 = \Delta U_{1\,\text{внутр}} + \Delta U_{1\,\text{внеш}},$$

где $\Delta U_{1\text{внутр}}$ — приращение напряжения на электроде первой фазы от тока внутренней цепи, вызванного движением зарядов $\sum_{i=1,\,j=1}^{M,N} Q_{ij}$ в потенциальных ямах; $\Delta U_{1\text{внеш}}$ — приращение напряжения на электроде первой фазы от перезаряда внешней цепи; ΔU_2 — скачок напряжения на электроде второй фазы (развертывающее напряжение).

Суммарный заряд в матрице

$$\sum_{i=1,\,j=1}^{MN} Q_{ij} = (\Delta U_1 - k_1 \Delta U_2)/k_2.$$

Это выражение определяет операцию компенсации сигнала на электроде второй фазы, необходимую для выделения полезного сигнала. Определить значения коэффициентов b_1 и b_2 можно после измерения емкостей $C_{ок}$, $C_{об}$, C_{12}. Для матрицы ПЗС К1200ЦМ1 $C_{об}=(200\ldots400)$ пФ, $C_{12}=C_{23}=C_{13}=600$ пФ, $C_{ок}=400$ пФ. Тогда $b_1=0{,}3$, $b_2=0{,}1\ldots0{,}2$.

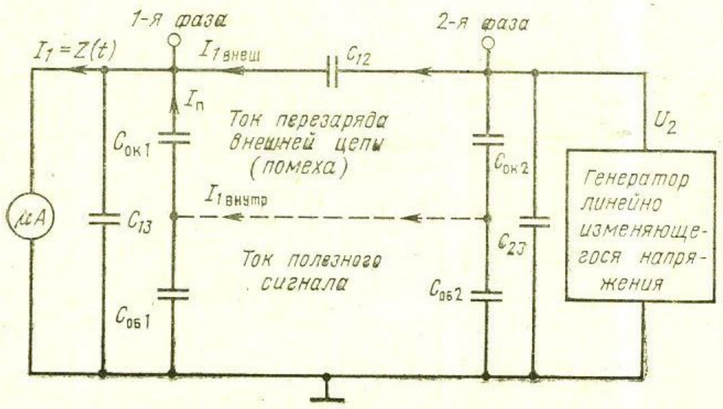


Рис. 3.13. Эквивалентная схема ПЗС-структуры трехфазной секции, поясняющая механизм формирования сигнала и помехи

Способ бескоординатной развертки был проверен в телевизионной камере на матрице ПЗС К1200ЦМ1 на частоте кадров 1 Гц. Плотность темнового тока составляла 70 пА/см² (интегральный уровень темнового сигнала $q_т=0{,}8\%$ от насыщения). На первом этапе эксперимента были сняты семейства характеристик при равномерной засветке секции накопления матрицы ПЗС (рис. 3.14). Степень размытости кривых зависит от неравномерности чувствительности матрицы ПЗС по полю и от того, насколько равномерным было распределение освещенности при засветке. Измеренное разрешение по горизонтали (рис. 3.14) определяется числом дискретов измерительного сигнала. В эксперименте число отсчетов было выбрано 200, при этом разрешение составило 0,5% по шкале заряда относительно уровня насыщения. Разрешение по вертикали определяется шумом канала считывания с электрода матрицы ПЗС. Динамический диапазон сигнала в эксперименте составил 400. *Период измерения управляющего сигнала $\mathcal{Z}(t)$ был равен 20 мкс, что в 1000 раз меньше периода вещательного кадра.* Таким образом, количество информации, полученное при измерении, составило $I[X(t), \mathcal{Z}(t)]=200\log_2 400\approx2000$ бит/измерение. При времени измерения $T_и=20$ мкс скорость кода $I[X(t), \mathcal{Z}(t)]T_и=10^8$ бит/с. Сравнивая это значение со скоростью создания информации по сигналу изображения для матрицы К1200ЦМ1, работающей на частоте кадров телевизионного вещания (частота элементов 5 МГц), получим $I[X(t), \mathcal{Z}(t)]T_к=5\cdot10^7$ бит/с. Видно, что скорости создания информации по основному выходу камеры (сигналу изображения) и по выходу с электродов секции накопления примерно равны. В последнем случае возможен обмен количества информации (точности измерения гистограммы яркости) на скорость измерения.

На втором этапе эксперимента была проверена работа неразрушающего считывания при неравномерной засветке секции накопления. На нее был спроецирован семиградационный клин (рис.

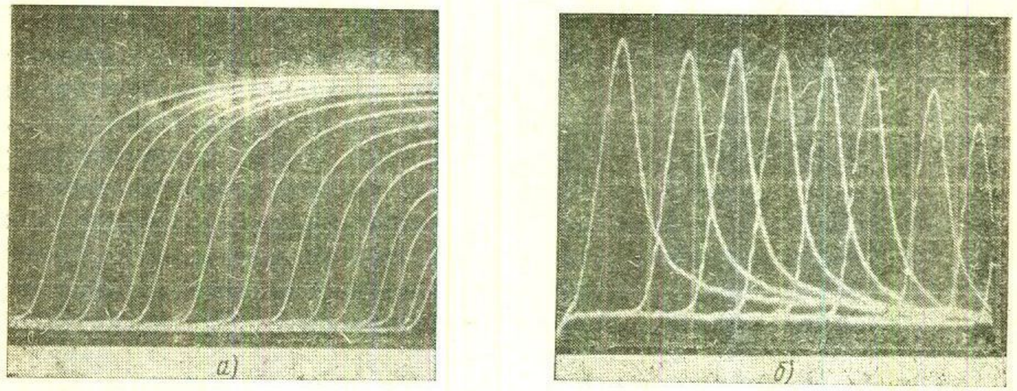


Рис. 3.14. Осциллограммы семейства характеристик тока в цепи электрода первой фазы (*а*) и его производной (*б*) от времени (заряда) при различной интенсивности равномерной засветки секции накопления

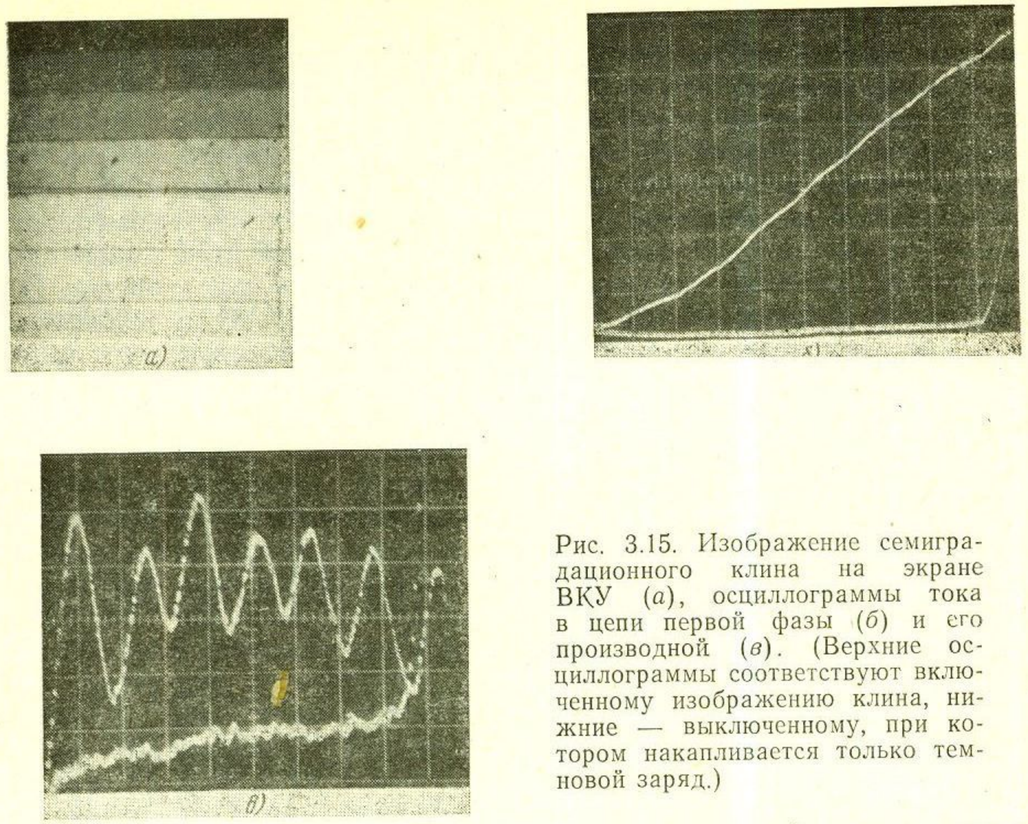


Рис. 3.15. Изображение семиградационного клина на экране ВКУ (*а*), осциллограммы тока в цепи первой фазы (*б*) и его производной (*в*). (Верхние осциллограммы соответствуют включенному изображению клина, нижние — выключенному, при котором накапливается только темновой заряд.)

3.15,*а*,*б*). При этом на осциллограмме производной сигнала тока электрода секции накопления (рис. 3.15,*в*) отчетливо различалось 7 максимумов, соответствующих семи градациям сигнала. Таким образом, точность неразрушающего считывания вполне достаточна для выделения скоростных сигналов управления частотой кадров адаптивной камеры.

3.6. АДАПТИВНОЕ КОДИРОВАНИЕ В ПРОЦЕССЕ СЧИТЫВАНИЯ

Характеристики и структуру телевизионных систем, особенно робототехнических, принято сравнивать со свойствами зрения человека [26]. Нередко обсуждается вопрос о целесообразности разделения телевизионного кадра на участки аналогично свойству человеческого зрения, характеризующегося различной четкостью на периферии и в центре (в фовеальной области). Это должно дать существенное сокращение полосы частот, т. е. сжатие сигнала за счет уменьшения четкости при воспроизведении общей обстановки с сохранением четкости при передаче полезных объектов [59].

Практическое создание систем с переменной четкостью в электронно-лучевом телевидении наталкивалось на непреодолимые трудности и относилось к послекамерному кодированию. Лишь с освоением твердотельного телевидения [2] за счет использования

возможностей ПЗС переменная четкость в телевизионных камерах стала реальностью [105]. Важно, что она реализуется в процессе камерного кодирования и относится к адаптивному процессу считывания зарядового изображения.

Выбор скорости считывания и четкости в различных участках растра основывается на теории идеального линейного кодирования (см. гл. 2). Оптимальные локальные четкость и скорость считывания определяются по входному спектральному отношению сигнал-шум с учетом допустимой среднеквадратической ошибки, назначаемой автоматическим получателем информации. В этом случае разные фрагменты изображения считываются с разными скоростями и апертурами. Рассмотрим подробнее режимы адаптивного кодирования в процессе считывания.

Режим переменной скорости считывания позволяет перераспределением частот считывания в пределах одного кадра улучшить качество считывания некоторого фрагмента изображения ценой ухудшения качества считывания остального массива элементов. Существуют различные варианты этого режима. В [36] элементы с малым количеством заряда считываются медленно, т. е. с пониженным шумом считывания, а элементы с большим количеством заряда — быстро. Другим вариантом является режим, когда медленно считывается лишь небольшой фрагмент, охватывающий интересующую область изображения, а остальной массив считывается быстро. В ряде случаев остальной массив может считываться с максимальной скоростью, а соответствующий сигнал не передаваться в канал связи.

Возможность адаптивного считывания обусловлена тем, что во многих случаях изображение полезного объекта занимает лишь небольшую (менее 1%) часть всей площади изображения. Поэтому после определения координат интересующего участка изображения целесообразно организовать такой режим считывания, который бы обеспечил наиболее благоприятные условия для считывания заряда с этого участка. Остальной массив изображения может быть считан с худшим качеством, а в ряде случаев может считываться на максимальной скорости и инжектироваться в стоковую область. Выигрыш в качестве считывания достигается за счет зависимости шума считывания от тактовой частоты (см. рис. 3.7). Например, замедление скорости считывания в небольшом фрагменте с 20 до 2 МГц обеспечивает выигрыш в отношении сигнал-шум для слабых сигналов в 4 ... 5 раз.

Для реализации режима переменной частоты считывания необходимо весь массив элементов, за исключением выбранного фрагмента, считывать с частотой несколько превышающей тактовую частоту регистра в обычном режиме $f_{т1}$. Задав коэффициент ускорения считывания $b_{сч}$, равный отношению новой более высокой частоты к частоте $f_{т1}$, можно определить частоту медленного считывания $f_{т2}$ фрагмента изображения, в котором находится объект:

$$f_{т2}=f_{т1}b_{сч}n_{э1}n_{э2}/[N_{э1}N_{э2}(b_{сч}-1)+n_{э1}n_{э2}],$$

где $N_{э2}$ — число строк матрицы ПЗС; $N_{э1}$ — число элементов матрицы ПЗС; $n_{э2}$ — число строк во фрагменте медленного считывания; $n_{э1}$ — число элементов фрагмента.

Размер фрагмента медленного считывания зависит от размера интересующего объекта, скорости его движения, его сложности и динамики мешающего фона. При заданном отношении частот $f_{т1}/f_{т2}$ размер фрагмента можно определить $n_{э1}n_{э2}=N_{э1}N_{э2}(b_{сч}-1)/[b_{сч}\times\times(f_{т1}/f_{т2}-1)]$. Например, при отношении частот $f_{т1}/f_{т2}=10$, $b_{сч}=1{,}1$ в матрице ПЗС 1200ЦМ7 с числом элементов 288×360 получаем число элементов фрагмента $n_{э1}\times n_{э2}=1036$, что при квадратной форме примерно соответствует стробу 32×32 элемента.

Рассмотренный вариант адаптивного считывания эффективен при относительно небольших размерах наблюдаемого объекта. При наблюдении точечных источников алгоритм может быть видоизменен. Главная трудность состоит в том, что до выделения объекта не ясно, какие элементы изображения считывать с минимальным шумом. Решение качественно может состоять в следующем. До выделения всю отведенную площадь фрагмента медленного считывания разбиваем на более мелкие и независимые друг от друга зоны. Например, фрагмент 32×32 элемента разбиваем на 64 зоны форматом 4×4 элемента каждая. Этими зонами будем пользоваться для проверки ложных тревог, возникающих в текущем кадре, «набрасывая» их на «подозрительные» элементы в следующем кадре (рис. 3.16,*а*). Благодаря значительно меньшему шуму считывания в этих зонах резко уменьшится время проверки ложных тревог. Те зоны, которые освободились после проверки очередного подозрительного элемента, используют в следующих кадрах для проверки новых появившихся ложных тревог. Такие проверки элементов могут проводиться до тех пор, пока не будет обнаружен полезный объект. После этого все зоны объединяются в один фрагмент, охватывая обнаруженный объект для более точного измерения его параметров (рис. 3.16,*б*). В некото-

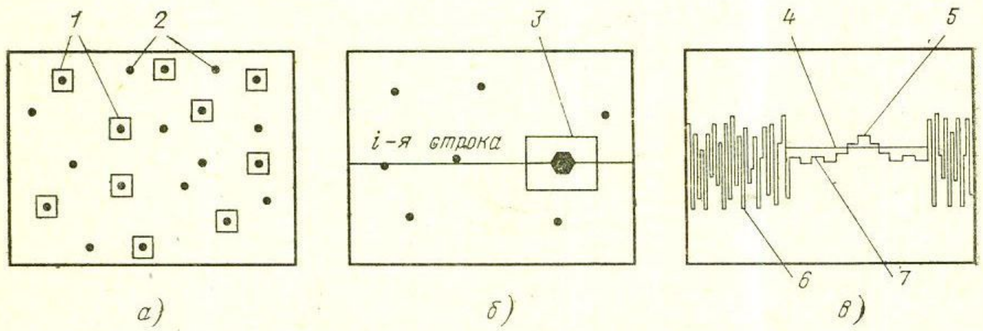


Рис. 3.16. Схематическое изображение, формируемое адаптивной ПЗС-камерой в режиме поиска объекта (*а*), в режиме измерения параметров объекта (*б*) и осциллограмма *i*-й строки, проходящей через объект (*в*); *1* — фрагменты медленного считывания, охватывающие ложные тревоги предыдущего кадра; *2* — ложные тревоги текущего кадра; *3* — объединенный фрагмент медленного считывания, охватывающий объект; *4* — порог; *5* — объект; *6* — шум быстрого считывания; *7* — шум медленного считывания

рых случаях алгоритм адаптации считывания может быть найден методами «задачи о разладке» в рамках статистического последовательного анализа аналогично, как он был определен для режима обнаружения с переменной частотой кадров [44].

Важным способом сжатия сигнала (сокращения полосы частот), обеспечивающим компромиссное распределение коэффициента сжатия между телекамерой и видеокомпьютером (см. 3.1 и 4.2), является **режим переменной по полю четкости.** В этом режиме полезный объект передается с полной четкостью, а окружающий его фон — с уменьшенной. При этом фон приобретает роль полезного, хотя и вспомогательного сигнала — контекста. Подход к распределению четкости в системах с получателем-человеком и автоматическим получателем различен. В первом случае снижают четкость при передаче подвижных объектов [59]. При определении характеристик движения объекта с помощью видеокомпьютера телевизионного робота следует поступать наоборот: сокращать четкость вне зоны полезного объекта [105]. При больших изменениях положения полезного объекта зона полной четкости должна быть подвижной.

Укрупнение апертуры в ПЗС легко получить при сложении зарядов нескольких элементов в потенциальной яме выходного устройства и сложении зарядов нескольких строк в потенциальных ямах выходного регистра [2, 40]. Возможны различные варианты этого режима. В простейшем случае апертуру укрупняют одновременно во всем массиве элементов ПЗС. Возможны режимы переменной апертуры по полю изображения. Во многих случаях полезный сигнал занимает незначительную часть всей площади изображения. Тогда целесообразно в небольшом фрагменте (например, 50×50 элементов), охватывающем интересующую область изображения, считывать сигнал с полной четкостью, а весь остальной массив считывать с укрупненной апертурой. Такой режим [105] позволяет на порядок и более сократить полосу частот сигнала изображения либо соответственно увеличить максимальную рабочую кадровую частоту. Диапазон укрупнения апертуры зависит от управляющей способности $Q_н$ секции накопления. При максимальных уровнях сигнала возможно лишь девятикратное увеличение апертуры: 3×3 в матрицах К1200ЦМ1 и К1200ЦМ7, так как управляющие способности $Q_в$ выходного устройства и $Q_р$ элемента нижнего регистра, а также элементов нижнего регистра и секции накопления $Q_н$ относятся как 3 : 1 [40]. При уменьшении заряда диапазон укрупнения апертуры увеличивается [40]:

$$D_а=(Q_в/Q_р)\,(Q_р/Q_н)/[(Q_э+q_т)/Q_н],$$

где $q_т$ — темновой ток, накопленный за время кадра.

Ранее экспериментально было реализовано 64-кратное укрупнение апертуры по всему полю [2]. При адаптивном кодировании интересующий фрагмент изображения считывают с низкой скоростью в режиме одноэлементной апертуры, а остальной массив

изображения считывают в режиме укрупненной апертуры (например, 4×4 элемента изображения). В этом режиме достигнуто сужение полосы частот сигнала изображения в 8 раз. Особенно проста реализация адаптивного считывания, когда частоты следования укрупненных элементов и элементов внутри фрагмента одинаковы и равны $f_{т3}$. При этом в камере можно использовать один видеотракт с переключением усиления внутри и вне фрагмента обратнопропорционально площади апертуры. Задав коэффициенты укрупнения элементов по кадру $b_{э2}$ и по строке $b_{э1}$, можно определить новую пониженную частоту считывания элементов $f_{т3}$.

$$f_{т3}=f_{т1}\{n_{э1}n_{э2}b_{э2}b_{э1}+(N_{э1}-n_{э2})\times$$

$$\times[N_{э2}-n_{э2}(b_{э2}-1)]\}/N_{э1}N_{э2}b_{э1}b_{э2}.$$

Например, при реализации в матрице ПЗС 1200ЦМ1 фрагмента размером $n_{э1}\times n_{э2}=40\times 40$ элементов и квадратной апертуре $b_{э1}=b_{э2}=4$ вне фрагмента достигается примерно десятикратное снижение тактовой частоты транзистора сброса матрицы ПЗС, а соответственно и полосы частот сигнала изображения. Достоинством режима является уменьшение числа выходных элементов изображения, для запоминания и обработки которых потребуется



Рис. 3.17. Изображения, формируемые адаптивной камерой на матрице ПЗС К1200ЦМ7 в режиме переменной площади считывающей апертуры на частоте кадров 25 Гц при коэффициенте укрупнения апертуры $g_M=g_N=4$

ОЗУ меньшей емкости. Полное число выходных элементов составит

$$N'_э=n_{э1}n_{э2}+[(N_{э2}-n_{э2})(N_{э1}-n_{э1})]/b_{э1}b_{э2}+$$

$$+[n_{э2}(N_{э1}-n_{э1})]/b_{э1}.$$

Видно, что режим адаптивного считывания при переменной площади считывающей апертуры позволяет не только улучшить пороговое отношение сигнал-шум внутри фрагмента, но и значительно сократить требуемую пропускную способность канала связи, снизить требования к быстродействию видеопроцессора.

Целесообразность переменной четкости связана с заметным превышением оптимального интервала дискретизации в задаче обнаружения по сравнению с задачей оценивания координат. Для типичной в астрономии модели сигнала — гауссовского видеоимпульса [87, 97, 107] — такое превышение более чем двукратное. Это ведет к концепции построения систем, адаптивных не только к изменениям характеристик сюжета, но и к этапам выделения полезной информации о наблюдаемых объектах.

При практической реализации адаптивного считывания необходимо точно синхронизировать обычные и укрупненные элементы, в противном случае возможно появление переходных границ (окантовки) фрагмента. Рассмотренный режим был использован в телекамере на матрице ПЗС 1200ЦМ7 с кадровой частотой 25 Гц (рис. 3.17). Размер фрагмента был выбран 50×100 элементов, частота элементов выходного сигнала составляла 350 кГц. Измеренная чувствительность камеры в 3 раза выше, чем при включении матрицы 1200ЦМ7 в стандартный режим при той же частоте кадров.

3.7. СПЕЦИФИКА ДИСКРЕТНО-АНАЛОГОВОЙ ОБРАБОТКИ СИГНАЛА

В телевизионной камере сигнал изображения проходит ряд последовательно включенных оптико-электронных каскадов. Каждый из них в соответствии с правилом покаскадного суммирования ошибок (2.7) искажает и зашумляет сигнал. Ошибки каскадов различны и связаны с их динамическими диапазонами. Качество телекамеры определяется каскадами с минимальным динамическим диапазоном. В ПЗС-камере три таких каскада: потенциальные ямы матрицы ПЗС; выходное устройство матрицы ПЗС; АЦП.

Адаптивное кодирование в ПЗС-камере позволяет сдвигать динамические диапазоны каскадов как отдельно (относительно друг друга), так и совместно (относительно шкалы входных освещенностей, рис. 3.18). Например, при перестройке частоты кадров и площади считывающей апертуры сдвигаются динамические диапазоны всех каскадов одновременно. В процессе адаптации частоты считывания и параметров видеотракта сдвигаются только ди-

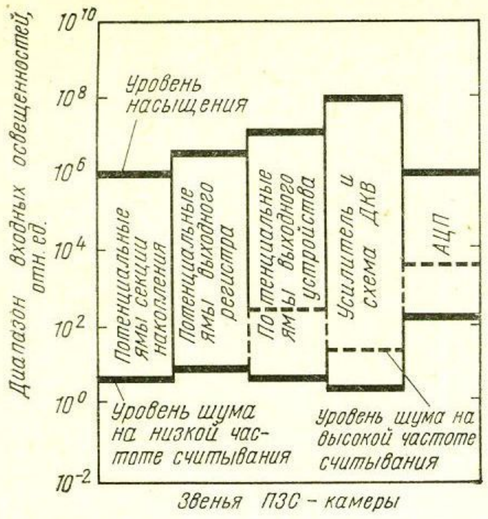


Рис. 3.18. Динамические диапазоны звеньев ПЗС-камеры

намические диапазоны выходного устройства матрицы ПЗС и АЦП соответственно. Необходимость адаптивного кодирования обусловлена сужением динамических диапазонов. Одновременный сдвиг диапазонов нужен потому, что любой из них уже диапазона входных освещенностей (см. рис. 3.18). Раздельное автоматическое смещение диапазонов требуется из-за их последовательного сужения в реальной телевизионной камере. На ее выходе расположен каскад с минимальным динамическим диапазоном — АЦП, открывающее цифровой канал связи. Кодером этого канала является адаптивный видеотракт, реализующий дискретно-аналоговую обработку сигнала изображения [2].

Главной особенностью такого видеотракта является наличие нескольких параллельных каналов. Часть из них обусловлена многоучастковостью матрицы ПЗС [2]. Участки матрицы ПЗС адаптируются независимо в соответствии с параметрами попадающих на них фрагментов изображения, а сигналы изображения считываются через раздельные выходы. Другая часть каналов видеотракта служит для съема и обработки скоростных управляющих адаптацией сигналов. Каждый участок матрицы ПЗС обслуживают два канала, один из которых предназначен для съема сигнала изображения $\mathcal{Z}(t)$, а другой — для сигналов управления $\mathcal{Z}(t)$. В адаптивной камере обычно используют две скорости считывания сигнала изображения — высокую (более 5 МГц) и низкую (менее 500 кГц). Большое различие частот приводит к разбиению сигнального канала на два канала. Низкочастотный канал должен быть малошумящим и точным. Для этого его выполняют с использованием операционных усилителей и КМОП-аналоговых коммутаторов. В высокочастотном канале для достижения быстродействия приходится жертвовать высокой точностью и широким динамическим диапазоном. Общим элементом каналов является буферный каскад и выравниватель усиления и смещения. Буферный каскад (см. рис. 3.6,*а*) служит для согласования видеотракта с выходным МОП-транзистором матрицы ПЗС. Выравнивание коэффициента передачи по напряжению камеры необходимо в режиме переменной площади апертуры. Для этого служит усилитель с управляемыми коэффициентом усиления по напряжению и напряжением смещения. Коэффициент усиления устанавливается обратно пропорционально площади апертуры поступающего на вход элемента. В этом случае при равномерной засветке

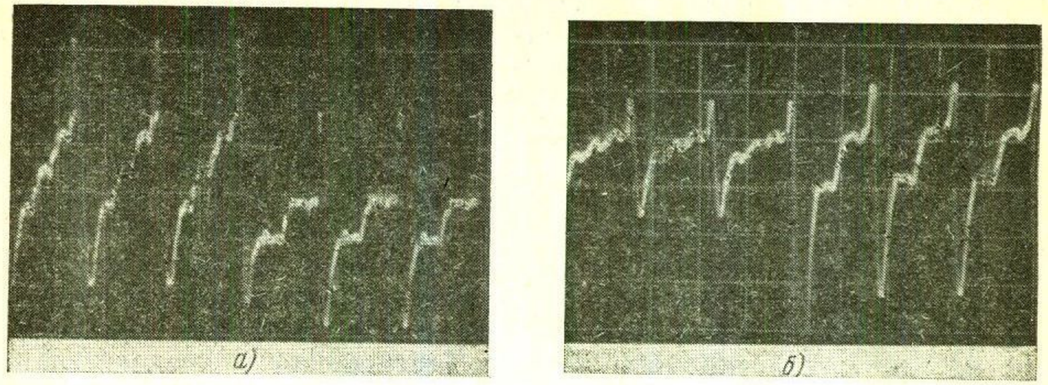


Рис. 3.19. Осциллограммы участка строки на выходе матрицы ПЗС (*а*), на выходе коммутируемого усилителя (*б*), иллюстрирующие переход от укрупненной апертуры к полной четкости изображения

напряжение сигнала элемента на выходе нормализатора не будет зависеть от площади апертуры элемента (рис. 3.19). При укрупнении апертуры из-за суммирования темнового заряда, а также из-за нелинейности выходного устройства может происходить изменение смещения напряжения сигнала. Для его компенсации выравниватель содержит переключаемые цепи смещения. Двойная коррелированная выборка предназначена для компенсации шума установки транзистора сброса выходного устройства, а также для компенсации температурного дрейфа. Особенностью ДКВ адаптивной камеры является большее, чем обычно требуется, отношение времени разряда ко времени заряда конденсаторов схем фиксации и выборки. Это отношение возрастает пропорционально максимальному размеру считывающей апертуры по строке.

Устройство дискретно-аналоговой обработки адаптивной камеры предназначено для сжатия динамического диапазона сигнала перед квантованием [2]. Если в нем реализуется пространственная фильтрация, то импульсная характеристика фильтра должна перестраиваться в соответствии с площадью апертуры поступающих на вход элементов.

Одновременно с сигналом изображения на выход адаптивной камеры должны поступать сигналы о параметрах телевизионного кадра: времени кадра, времени накопления, координаты и размеры участков с различной апертурой и различной скоростью считывания. Без этих данных невозможна ни дальнейшая цифровая обработка изображения, ни его визуализация, для которой необходим адаптивный преобразователь параметров разложения, который бы обеспечивал точное преобразование изменяющегося входного сигнала изображения в стандартный сигнал телевизионного вещания. Такое преобразование возможно без потерь, когда частота кадров и формат изображения не превышают вещательных. В противном случае необходима предварительная запись на специальный видеомагнитофон с последующим замедленным просмотром, либо запись с пропуском ряда кадров.

Канал управления видеотракта адаптивной камеры может формировать несколько управляющих сигналов. Адаптация частоты кадров и времени накопления осуществляется по сигналу, снимаемому с электрода секции накопления в процессе накопления заряда. Этот же сигнал в конце интервала накопления используется для адаптации параметров видеотракта. Регулирование площади считывающей апертуры и частоты считывания производится по сигналам, снимаемым с электрода выходного регистра и с выходного затвора. При съеме сигналов с электродов секции и регистра методом неразрушающего считывания (3.5) на входе видеотракта устанавливается коммутатор, подключающий к электроду поочередно генератор управляющих напряжений и входной каскад. Этот каскад выполняется в виде дифференциального преобразователя ток — напряжение для одновременного осуществления первого дифференцирования зарядового сигнала и компенсации развертывающего напряжения. Затем следует дифференциатор, на выходе которого формируется распределение заряда по уровням, которые сравниваются с несколькими порогами в блоке компараторов. Сигналы управления могут формироваться и из обычного сигнала изображения. Однако при этом значительно возрастет задержка (3.15) в цепи управления до длительности двух кадров, и наблюдение быстроизменяющихся по освещенности процессов станет невозможно. Использование сигнала управления, опережающего входной сигнал, обеспечивает минимальные динамические ошибки регулирования даже в случаях резких изменений параметров изображения в соседних кадрах.

3.8. ИНФОРМАЦИОННЫЕ ХАРАКТЕРИСТИКИ ТЕЛЕВИЗИОННЫХ КАМЕР

Выше были рассмотрены практические вопросы двухэтапного камерного кодирования по схеме на рис. 3.4. Здесь мы остановимся на информационных характеристиках телекамер, входящих в состав информационно-решающих систем, т. е. осуществляющих первый шаг кодирования информации на пути принятия решения о наблюдаемом сюжете. Такие системы описываются в рамках двух теорий — теории решений [12, 8] и теории информации [3, 4]. При этом проверка гипотез получателем, основывающаяся на теории решений, называется также классификацией или многоальтернативным обнаружением. Важно, что при построении декодеров в теории информации используют понятие теории решений (например, максимума правдоподобия). Теория решений, однако, не использует понятий теории информации, так как не использует ограничений типа пропускной способности канала связи. Поэтому эти две теории не подменяют, а дополняют друг друга. Такое дополнение мы рассмотрим в 4.2, здесь ограничимся информационными оценками. Спецификой работы информационно-решающих систем является сложность сюжета, содержащего кроме полезных классифицируемых объектов интенсивный мешающий фон и шум.

Подавление фона и классификация входят в функции видеокомпьютера (см. соответственно 4.1 и 4.2), телекамера же осуществляет первичное сжатие информации, содержащейся в световом потоке. Мешающий фон, обычно превышающий полезный сигнал, моделируют гауссовским полем, что позволяет получить оценки эффективности работы телекамеры и процессора видеокомпьютера.

Информационно-решающая система в целом осуществляет сжатие информации от энтропии потока фотонов H_Y (равной числу фотонов $N_ф$) до решающей способности $V_{реш}$ (см. 3.1). При этом число гипотез, различаемых получателем, которое определяется его решающей способностью $V_{реш}$, на много порядков меньше числа гипотез канала $M_к = 2^C$.

Коэффициент сжатия A распределяется между телекамерой и процессором с учетом ε-энтропии $H_Y(\varepsilon)$ в битах в секунду на выходе телекамеры:

$$A = A_к A_п; \quad A_к = N_ф / H_Y(\varepsilon); \quad A_п = H_Y(\varepsilon)/V_{реш}.$$

Увеличение коэффициента сжатия изображения в телекамере, зависящего от назначаемой ошибки ε, ведет как к упрощению процессора, так и к возрастанию ошибок решения. В результате автоматический получатель как бы сам выдвигает требования к информационным характеристикам телекамеры для достижения собственной максимальной эффективности. Это отражается и в связи чувствительности телекамеры, определяемой отношением пропущенного количества информации к поступающему [26], с пропускной способностью канала связи. Такая взаимосвязь получателя и системы передачи объединяет две части информационно-решающей системы.

Максимальная решающая способность при заданной суммарной сложности телекамеры и процессора достигается при распределении коэффициентов сжатия информации между ними поровну (см. 4.2): $A_к = A_п$. Тогда для ε-энтропии получаем:

$$H_Y(\varepsilon) = \sqrt{N_ф V_{реш}}.$$

Выходной сигнал телекамеры на ПЗС (с учетом декоррелятора и квантователя) при накопительной способности Q электронов на потенциальную яму, числе элементов в кадре (матрице) $N_э$, числе решений за кадр $V_{реш} T_к$ будет иметь энтропию в битах за секунду:

$$H_Y(\varepsilon) = \sqrt{V_{реш} T_к Q N_э / T_к}.$$

В пересчете на один элемент это составит:

$$H_Y(\varepsilon) T_к / N_э = \sqrt{V_{реш} T_к Q / N_э}.$$

Для системы коррекции волнового фронта [86, 103] необходима решающая способность более чем 10^4 решений в секунду. При частоте кадров около 200 Гц, матрице ПЗС с числом элементов 10^6 и емкостью потенциальной ямы 10^5 фотоэлектронов от камеры и видеокомпьютера требуется сжатие информации примерно на 4

порядка, что соответствует скорости кода на выходе телекамеры 2 ... 4 бит/элемент. Важно, что это значение близко к 2, т. е. к числу уровней квантования оптимальной по минимуму среднеквадратической ошибки оценки $\hat{Z}$ гауссовского сигнала X, при котором достигает максимума информационный КПД — отношение $I(X, Z)/I(Y, Z)$ [2].

Распределение сжатия между телекамерой и процессором так, что каждый из них сжимает поток информации на четыре порядка, наглядно показывает их взаимное влияние, позволяющее говорить об информационно-решающей системе как о самостоятельной системной единице.

Требуемые информационные характеристики телекамеры определяются с позиций системного подхода, структура же ее связана с реализуемым в реальном времени кодированием, которое в телекамере разделяется на аналоговое и цифровое. В отличие от идеального блокового кодирования Шеннона, совмещающего процессы дискретизации и квантования, в телекамерах они разнесены и осуществляются соответственно в ПЗС и АЦП. Такое вынужденное разделение увеличивает ошибку воспроизведения и выдвигает вопрос о распределении ошибок по каскадам.

Критерий минимума суммарной информационной ошибки. Распределение ошибок по каскадам должно осуществляться исходя из компромисса между потерей полезной информации и возникновением шумовой информации. Возможности взаимообмена между ними проиллюстрируем характеристикой, приведенной на рис. 3.20 [104]. Она отражает зависимость полезной информации $I(X, Z)$, нормированной на ее потенциальное значение $I(X, Y)$, от доли шумовой информации в кодовой ε-энтропии. Здесь мы для приближенных оценок воспользуемся одномерными формулами из гл. 2 для взаимных информаций. Многомерные точные формулы трудно непосредственно применить к оценке характеристик системы, использующей разложение по базису, отличному от базиса Карунена — Лоэва. При этом *качественная картина бесконечного роста шумовой информации при стремлении ошибки к лимитируемой входным шумом при переходе от одномерных к многомерным формулам сохраняется.* Параметром характеристики рис. 3.20 является входное отношение сигнал-шум. При заданном входном отношении сигнал-шум перемещение вдоль кривой осуществляется изменением назначаемой ошибки, т. е. изменением усиления, обеспечивающего задаваемое превышение входного зашумленного сигнала над шумом канала. Это эквивалентно вариации шага квантования при постоянном уровне входного сигнала. Важно, что увеличение отношения сигнал-шум увеличивает количество полезной информации и уменьшает количество шумовой информации $I(\xi, \mathcal{U}) \triangleq I_{ш}$. На рис. 3.20 штриховая линия разделяет все возможные соотношения полной ошибки и ошибки, обусловленной входным шумом, на две области: правее этой линии шумовая информация превышает потерю полезной, левее — наоборот. Этой линии соответствует полная ошибка, вдвое большая ε_B. При достаточно большом входном отношении сигнал-шум и достаточно большой назначаемой ошибке шумовая информация составляет весьма малую часть кодовой ε-энтропии, которая в пределе равна

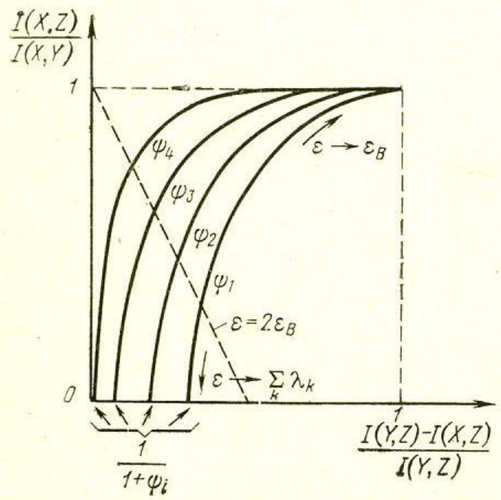


Рис. 3.20. Нормированная полезная информация как функция загрузки канала шумовой информацией

$$\lim_{\varepsilon \to \lambda} I(\xi, \hat{\mathcal{U}})/I(Y, \hat{Z}) = 1/(1+\psi^2).$$

Наглядность взаимообмена потери полезной информации на увеличение шумовой информации $I_ш$ по рис. 3.20 дополняется аналитическим выражением для информационного соотношения неопределенности (см. 2.3 и 4.4), справедливого для малых ошибок ε: $\Delta I \cdot 2^{2I_ш} \approx 1$.

Необходимость учета как потери полезной, так и шумовой информации, загружающей без пользы канал связи и расходующей производительность видеокомпьютера, позволяет сформулировать *критерий минимума суммарной информационной ошибки*:

$$W = \min_{\varepsilon} \{c_1 \Delta I_{мин} + c_2 I_ш\}. \qquad (3.20)$$

Весовые коэффициенты c_1 и c_2 вводятся для отражения «цены» информационных ошибок того или иного вида аналогично ценам в критерии Байеса для проверки статистических гипотез [8, 14]. В частном случае равенства цен ошибок $c_1 = c_2$ критерий (3.20) принимает вид

$$W = \min_{\varepsilon} \{\Delta I_{мин} + I_ш\}. \qquad (3.21)$$

По существу, критерий (3.21) использовался при нахождении компромиссной ошибки (2.47), которая вдвое превышает винеровскую ошибку ε_B. Вариацией весовых коэффициентов в (3.20) можно изменять соотношение ошибок первого и последующих каскадов. Важным свойством критерия минимума суммарной информационной ошибки (3.20) и (3.21) является возможность обоснованного распределения среднеквадратических ошибок ε_Σ по каскадам и при произвольном числе каскадов J (при фиксированной среднеквадратической ошибке $\varepsilon_1 = \varepsilon_B$ первого каскада с неварьируемым искажающим фильтром B на рис. 1.2). Используя критерий (3.21), получаем оптимальное распределение суммарной среднеквадратической ошибки ε_Σ, приводящее к *равномерному распределению потери полезной информации по каскадам*:

$$\Delta I_{j\,мин} = 1/2 \text{ бит при } I_ш \approx 1/2 \text{ бит}. \qquad (3.22)$$

При этом среднеквадратическая ошибка ε_Σ в многокаскадной системе распределяется по каскадам не равномерно, а так, что ошибка каждого каскада равна сумме ошибок предшествующих каскадов:

$$\varepsilon_j = \varepsilon_1 2^{(j-2)}, \quad j \geqslant 2. \tag{3.23}$$

Это соответствует минимальной суммарной информационной ошибке

$$W = \sum_{j=1}^{J} [\Delta I_{j\,\text{мин}} + I_\text{ш}] \approx \frac{1}{2} \log \frac{\varepsilon_\Sigma^2}{\varepsilon_1 \varepsilon_J} \approx J/2; \quad J \geqslant 2.$$

В частном случае двух каскадов, подробно рассмотренном в гл. 2, $W \approx 1$ бит (см. 2.3). Этот случай при достаточно больших отношениях сигнал-шум характеризуется *равенством пропускных способностей каналов: первичного с неварьируемым фильтром B и вторичного с кодером R*:

$$C_2 = C_1 - \Delta I_\text{мин} + I_\text{ш},$$

$$C_1 = \frac{1}{2} \log (\psi^2 + 1); \quad C_2 = \frac{1}{2} \log \psi^2.$$

Критерий минимума суммарной информационной ошибки поясняет качественную картину работы передаточного фильтра (2.16) и (2.65), заключающуюся в убывании спектра шумовой информации и усечении тех степеней свободы, в которых полезная информация не превышает полбита [см. (2.65) и (3.22)].

Оценим степень приближения информационных характеристик телекамеры к идеальным, используя в качестве эталона для аналогового этапа идеальное линейное кодирование, критерий минимума суммарной информационной ошибки и простые модели многомерных сигналов изображений и их параметрически оптимальное преобразование, содержащее этапы аналогового и аналого-цифрового кодирования.

Аналоговое кодирование уступает идеальному прежде всего из-за использования вместо орт Карунена — Лоэва упрощенной системы базисных прямоугольных функций. Такие функции хорошо аппроксимируют импульсную характеристику матричного ПЗС как по пространственным координатам, так и по времени [2].

Для выявления специфики преобразования сигналов при разложении по прямоугольным функциям мы воспользуемся различными моделями сигналов. Здесь рассмотрим гауссовские сигналы, а в 4.3 — «телеграфный» сигнал [14], так как для него наиболее наглядна неинвариантность к сдвигу. Так же как в 3.2, рассмотрим простую модель гауссовского процесса, являющегося функцией двух пространственных и временной координат. Для упрощения будем считать корреляционную функцию аксиально симметричной, т. е. ее можно разложить на независимые компоненты по всем переменным; это позволит рассмотреть разложение сигнала как функции одной переменной.

Разложение по прямоугольным базисным функциям позволяет достичь быстрой переработки информации, так как задержка на формирование проекций сигнала на перекрывающиеся орты Карунена — Лоэва отсутствует. Такое приобретение оплачивается ценой ухудшения других информационных характеристик. Во-первых, теряется независимость отсчетов, во-вторых, повышается минимально достижимая ошибка, которая обязательно превысит лимитируемую входным шумом.

Для уменьшения ошибки представления в этом разложении требуется большее число отсчетов, чем при разложении Карунена — Лоэва. Для гауссмарковского процесса такое превышение составляет примерно 1,6 раза [14]. Вместе с тем, как показано в 3.2 (см. рис. 3.3), разложение гауссмарковского сигнала по прямоугольным функциям обеспечивает хорошее приближение передаточной функции системы к спектральной характеристике идеального кодера, состоящего из винеровского и передаточного фильтров (гл. 2). Это при линейной интерполяции обеспечивает превышение ошибки $\varepsilon_{r\,\text{мин}}$ над лимитируемой входным шумом: $\varepsilon_{r\,\text{мин}} \approx 1{,}2\varepsilon_B$ (рис. 3.21). Увеличение ошибки из-за неидеальности фильтрации в фотоприемнике соответствует потере полезной информации на аналоговом этапе кодирования в телекамере примерно 0,5 бит/элемент. К этой потере добавляется потеря, обусловленная шумом канала считывания, достаточно малая при больших сигналах, и достигающая 1 бит при сигналах, близких к пороговым.

Потеря независимости отсчетов в телекамере проявляется в их корреляции. Для гауссмарковских сигналов коэффициент корреляции двух отсчетов определяется расстоянием между ними $K(n\Delta) = \rho^n$, где ρ — коэффициент корреляции соседних элементов. Важно, что при оптимальной дискретизации, дающей минимум среднеквадратической ошибки, такая корреляция существенна. Она растет пропорционально увеличению входного отношения сигнал-шум, так как оптимальный интервал дискретизации при этом уменьшается [2]: $\Delta_0 \approx 2\tau/\psi$. Наличие такой корреляции приводит к существенной неэффективности использования канала считывания в ПЗС по сравнению с идеальным по Шенонну.

Кодирование в телекамере характеризуется указанной бесполезной загрузкой канала, потерей информации и передачей шумо-

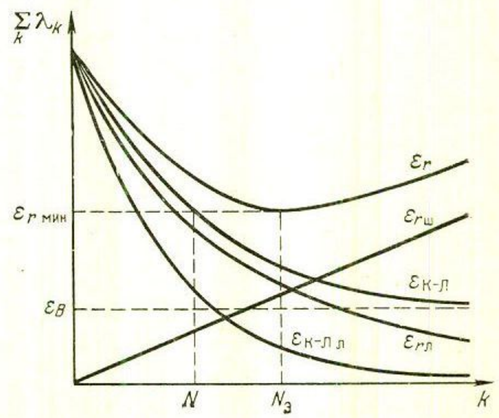


Рис. 3.21. Зависимость ошибок от числа степеней свободы при разложении Карунена — Лоэва: линейной $\varepsilon_{\text{К—Л л}}$ и полной $\varepsilon_{\text{К—Л}}$ и разложении по прямоугольным функциям: линейной $\varepsilon_{r\text{л}}$, шумовой $\varepsilon_{r\text{ш}}$, полной ε_r

вой информации. Степень отклонения информационных характеристик канала считывания в ПЗС при гауссовской модели сигнала может быть оценена сверху с учетом максимально достижимого отношения сигнал-шум. Оно определяется при лимитирующем влиянии входного шума емкостью потенциальной ямы элемента ПЗС. Оценки характеристик кодирования в ПЗС-камерах будем определять для ориентировочных значений емкости потенциальной ямы (управляющей способности) $Q_н = 10^6$ электронов и среднеквадратического значения шума канала считывания $\sigma_2 = 10^2$ электронов. Максимальное отношение сигнал-шум в этом случае определится входным шумом с дисперсией $\sigma^2{}_1 = Q : \psi = Q/\sqrt{Q + \sigma^2{}_2} \approx \approx 10^3$. Межэлементная корреляция приводит к тому, что ε-энтропия, составляющая [59] $\frac{1}{2}\log\lambda(1-\rho^2)/\varepsilon$ уменьшается по сравнению с пропускной способностью в битах на $\frac{1}{2}\log(1-\rho^2)$. Коэффициент корреляции $\rho = K(\Delta_0)/K(0) = \exp(-\Delta_0/\tau) \approx 1-2/\psi$. Поэтому пропускная способность канала считывания превышает ε-энтропию источника более, чем на 4 бит/элемент.

Вместе с полезной информацией и не рационально загружающей канал бесполезной, но в принципе устранимой информацией, через канал передается и шумовая информация. Воспользовавшись с указанными оговорками одномерными формулами для взаимных информаций, шумовую информацию $I_ш = \frac{1}{2}\log\left(1 + \frac{g^2\sigma_1^2}{\sigma_2^2}\right)$ можно оценить для максимальных сигналов в 1,5 бит/элемент. При этом каждый элемент несет о входном полезном сигнале около 0,5 бит информации [2]. В результате пропускная способность канала считывания, определяемая соотношением емкости потенциальной ямы и шума считывания $C = \frac{1}{2}\log\left(1 + \frac{Q_н}{\sigma_2^2}\right)$ распределяется между полезной, шумовой и бесполезной устранимой информацией в соотношении 0,5 : 1,5 : 4,5. То, что на полезную информацию расходуется менее 10% пропускной способности канала считывания, обусловлено ограниченностью необходимых операций кодирования до считывания — усиления и декорреляции.

Эти операции кодирования — декорреляция и усиление — осуществляются в телекамере перед квантованием в АЦП. Это позволяет улучшить информационные характеристики телекамеры в целом. Декорреляция обычно осуществляется методами дифференциальной импульсно-кодовой модуляции [58—61], обычно с помощью линейной фильтрации.

Цифровое кодирование является многоуровневым принятием решения и опирается на процедуру квантования, которая в соответствии с рис. В.2 дополняет накопление и дискретизацию по схеме рис. 2.4 в их борьбе с шумовой информацией и вносит вклад в сжатие сигнала. Квантование гауссовского сигнала даже

при поэлементном безблоковом кодировании ($n_1 = n_2 = 1$) позволяет достичь хорошего приближения к кодовой ε-энтропии непрерывного сигнала. Превышение скорости кода над ε-энтропией (устранимая ценой задержек бесполезная информация) составляет около 0,3 бит/элемент [25, 111].

Выбор числа разрядов (уровней) квантования сигнала $\hat{Z}$ в АЦП производится по требуемой для надежной работы автоматического получателя среднеквадратической ошибке. Диапазон изменения назначаемых получателем ошибок распределяется на три участка в соответствии с различными взглядами на роль потери информации и шумовую информацию. Первый участок соответствует *предельно малой потере полезной информации* из-за квантования. Этот предел связан с ограничениями на энергию $E_к$, отводимую на усиление кода одного непрерывного отсчета (элемента). Эта энергия распределяется по разрядам кода для минимизации ошибки [112] аналогично распределению энергии усиления E_g, рассмотренного в гл. 2. В этом случае при квантовании усечение числа разрядов разложения континуального числа будет определяться шумом цифрового канала так же, как шум непрерывного канала усекает число членов разложения Карунена — Лоэва непрерывной функции при дискретизации. При импульсно-кодовой модуляции и передаче разрядов кода ортогональными сигналами будет иметь место оптимальное число m_0 разрядов, обеспечивающее минимум суммарной ошибки ε, состоящей из ошибки квантования (кодирования) $\varepsilon_к$ и ошибки $\varepsilon_д$ передачи по каналу (декодирования).

Ошибка кодирования убывает с ростом числа разрядов, так как энтропийная мощность ошибки квантования при оптимальном выборе порогов определяется дисперсией сигнала λ и числом разрядов [61]: $\varepsilon_к = \lambda/(2^{2m}-1)$. Ошибка декодирования при ограничении $E_к$ возрастает с увеличением числа разрядов, так как определяется отношением сигнал-шум в канале и для гауссовского шума составит $\varepsilon_д = 2\sum_{i=1}^{m_0} 2^i\Phi(-\psi_i/2)$, где Φ — интеграл вероятности.

Оптимальное распределение энергии по степеням свободы (сигналам разрядов в канале) для передачи гауссовского сигнала по дискретному каналу с аддитивным белым гауссовским шумом со спектральной плотностью мощности $\mathcal{N}_0/2$ определяется зависимостью $\psi_i \approx 3{,}3\sqrt{m_0 - i}$. Такая зависимость аналогична передаточному фильтру (2.2) и определяет оптимальное конечное число разрядов m_0, зависящее от доступного отношения суммарного сигнала элемента к шуму в канале:

$$m_0 \approx 0{,}6\,(2E_к/\mathcal{N}_0)^{1/3}.$$

Хотя из-за ограничения $E_к$ число разрядов m_0 может быть и не слишком большим, но методологически этот подход пренебрегает шумовой информацией, которая в соответствии с информационным соотношением неопределенности (см. 2.4) возрастает при уменьшении потери полезной информации.

Учет шумовой информации может вызвать желание проектировщика полностью ее подавить. Для этого в дополнение к накоплению и дискретизации следует произвести квантование с шагом, намного превышающим ошибку ε_B в сигнале, связанную с входным шумом (второй участок диапазона ошибок). Например, если шаг квантования на порядок превышает среднеквадратическое значение этой ошибки, то «вероятность присвоения двоичного числа неправильному уровню» [55] будет порядка 10^{-20}. Это соответствует *практически полному подавлению шумовой информации*, однако приводит к заметному уменьшению полезной информации.

Третий, наиболее интересный участок диапазона ошибок определяется компромиссом между потерей полезной информации и передачей шумовой (см. рис. 3.20). Этот компромисс в *соответствии с критерием минимума суммарной информационной ошибки предписывает выбор мощности ошибок квантования по формуле (3.23), равной сумме мощностей ошибок оптического канала и канала считывания зарядового изображения.* Это дает опирающееся на решение вариационной задачи теоретическое обоснование эмпирическому правилу приближенного (с коэффициентами 2 ... 3,5) равенства среднеквадратических значений шума в сигнале и ошибок квантования [55]. Развитый нами подход, опирающийся на учет шумовой информации, так же как указанное эмпирическое правило, отражает нецелесообразность неограниченной загрузки каналов бесполезной информацией. При выборе разрядности следует учитывать, что шумы квантования надо вычислять, исходя из энтропийной мощности, которая учитывает зависимость ошибок квантования от полезного сигнала [61]. Компромиссный выбор энтропийной мощности ошибки квантования, равной ошибке оценивания полезного сигнала на выходе аналогового кодера схемы рис. 3.4, приводит к тому, что при квантовании шумовая информация уменьшается, а полезная почти сохраняется. Для принятых моделей сигнала и характеристик ПЗС указанное квантование снижает шумовую информацию с 1,5 до 0,5 бит/элемент. При этом *максимум информационного КПД достигается при двух разрядах квантования* и составляет примерно 30% [2]. Распределение пропускной способности цифрового нешумящего канала между полезной информацией, шумовой неустранимой информацией и бесполезной, но устранимой за счет задержек на формирование блоков, осуществляется при таком квантовании примерно поровну.

Приведенные оценки эффективности кодирования зашумленного гауссовского поля при покадровой и поэлементной дискретизации, декорреляции и поэлементном квантовании позволяют сделать вывод: телекамера, осуществляющая аналоговое и цифровое кодирование в соответствии с рис. 3.4, из-за неидеальности базисных функций и влияния шума канала считывания может передавать около трети потенциальной полезной информации, требуя для этого из-за влияния входного шума, ошибок квантования и скоростного безблокового кодирования пропускную способность, примерно втрое большую кодовой ε-энтропии (полезной информации при равенстве ошибок, вносимых аналоговыми и цифровым каскадами). Избыточность цифрового кода на выходе телекамеры, обусловленная входным шумом, неустранима. Вторая компонента избыточности устранима лишь ценой увеличения задержек.

Выводы по главе. Введенная в гл. 2 многокаскадная модель передачи информации и сформулированное правило суммирования винеровских покаскадных ошибок позволило в данной главе проанализировать передающую часть телевизионной системы как многокаскадную и сформулировать критерий минимума информационной ошибки, состоящей из потери полезной информации и шумовой информации. Этот критерий позволяет обоснованно распределить среднеквадратическую ошибку по каскадам и со своих позиций поясняет потребность изменения спектральных коэффициентов и вырождения числа степеней свободы передаточного фильтра.

Оконечным звеном видеоинформационной системы является автоматический получатель с ограниченной решающей способностью, предъявляющий требования ко всем каскадам в виде среднеквадратической ошибки. Скоростное безблоковое кодирование в телекамере осуществляет поэтапное сжатие видеоинформации на четыре порядка (одинаково с видеокомпьютером) путем накопления, декорреляции и квантования от потока фотонов до кода, скорость которого по возможности приближается к ε-энтропии. Накопление сигнала в ПЗС определяется пространственной и временной дискретизацией, выбираемыми по сюжету с учетом влияния шумов. При его переменных свойствах телекамеры на ПЗС могут адаптироваться изменением параметров считывания зарядового изображения, например при разделимости сюжета на фон и объект вводить переменную по полю четкость. Возможности адаптации ПЗС-камер расширяются благодаря организации нескольких дополнительных скоростных каналов считывания. С их выходов снимаются сигналы, управляющие аналоговым кодированием от накопления в ПЗС до декорреляции. Для рациональной загрузки канала связи камеры с ЭВМ квантователь в соответствии с критерием минимума суммарной информационной ошибки обеспечивает компромисс между потерей полезной информации и передачей шумовой. В результате совместных усилий аналогового и цифрового кодеров, телекамера из-за упрощенных базисных функций и шума ПЗС может передавать треть полезной информации. При этом из-за входного шума, ошибок квантования и неприменимости блокового кодирования потребуется пропускная способность втрое большая ε-энтропии для указанной компромиссной ошибки.

Глава 4.

КОМПЬЮТЕРНАЯ ПЕРЕРАБОТКА ИНФОРМАЦИИ С ПРИНЯТИЕМ РЕШЕНИЯ В РЕАЛЬНОМ ВРЕМЕНИ

В видеоинформационных системах «от света до решения» граница между кодеком и получателем однозначно не проводится. Поэтому компьютерная переработка информации составляет специфический этап, который объединяет в себе элементы кодирования и принятия решения. Принятие решения в реальном времени диктует необходимость преобразования видеокомпьютером цифрового сигнала телекамеры в решение с частотой кадров. Поэтому алгоритмы переработки информации в быстродействующих системах телевизионных роботов не могут быть слишком сложными. Ниже излагается подход к выбору алгоритмов с позиций теории передачи информации и с использованием понятий теории решений, в основном отношения правдоподобия.

Система управления телевизионным роботом (см. гл. 5) связана с телекамерой через видеокомпьютер (рис. 4.1). Особенностью систем реального времени является переход от ЭВМ общего назначения, реализующих последовательные алгоритмы Тьюринга и фон Неймана, к многопроцессорным компьютерам, реализующим параллельную переработку информации. При этом видеокомпьютер является специфическим узлом системы связи, в котором, благодаря алгоритмической переработке информации с принятием решения, пропускная способность канала вывода информации существенно меньше, чем у канала ввода. Процесс переработки видеоинфомации в решение реализует последовательность: информация — оценка — решение, и обычно распадается на три взаимосвязанных этапа. На первом предварительном этапе осуществляют компенсацию искажающих изображение факторов (см. 4.1). Важнейшим среди них является влияние фонового изображения. При переработке информации на этом этапе приходится также осуществлять компенсацию влияния различных деформаций изображений и компенсацию разброса чувствительности по полю. На втором этапе производится анализ изображе-

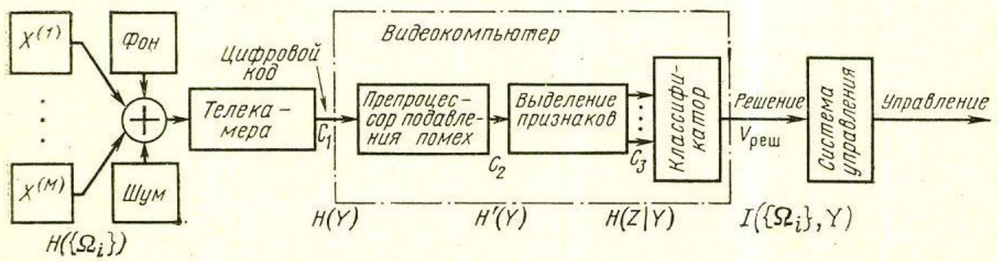


Рис. 4.1. Модель переработки информации видеокомпьютером

ний для их классификации (см. 4.2). Здесь выявляется взаимодополнение понятий теории информации и теории решений (см. 3.8) при определении ограничений, возникающих в задаче классификации изображений. На третьем этапе формируется вектор параметров объектов (см. 4.4), используемый автоматическим получателем (см. гл. 5) для формирования поведения. Взаимосвязь этих этапов обусловлена, в частности, тем, что вектор параметров объектов используется и для управления исполнительным механизмом и для принятия решения. Все три этапа переработки информации характеризуются методологическим единством на основе теории связи, в первую очередь ее основных разделов — теории передачи информации и теории обнаружения и оценивания.

4.1. ПРЕДВАРИТЕЛЬНАЯ ПЕРЕРАБОТКА ВИДЕОИНФОРМАЦИИ

Предварительная переработка видеоинформации преследует цель подготовки изображения для осуществления процедуры классификации. В ее функции входит компенсация фонового изображения и влияния искажающих факторов, например смаза изображения. Обычно предварительная переработка видеоинформации реализуется как фильтрация изображений.

Фильтрация для выделения сигнала из фона основывается на тех различиях пространственно-временных характеристик полезного сигнала и фона, которые не были полностью использованы при выборе параметров разложения изображения при камерном кодировании. Часто возникает ситуация, когда фон по полю и во времени неоднороден и фотоприемник дискретизирует оптимальным образом сигналы, наиболее быстро изменяющиеся. Дополнительная фильтрация позволяет повысить помехоустойчивость выделения сигнала, обеспечивая оптимальную (обычно в смысле минимума среднеквадратической ошибки) оценку фонового изображения и его вычитание.

Искажения изображений вносятся различными внешними и внутренними эффектами. К внутренним относятся специфические свойства фотоприемников: для ПЗС, например, — это неэффективность переноса заряда, неравномерность темнового тока. К внешним искажающим эффектам относятся различные постоянные и переменные искажения оптических сигналов на пути от источника до фотоприемника. При этом постоянными во времени, но переменными по полю искажениями являются аберрации оптической системы. Значительные временные искажения вносятся турбулентностью атмосферы. Промежуточное положение по скорости изменения во времени занимают искажения, вносимые подвижностью изображений. В ряде приложений, в частности в робототехнике, движение изображений имеет как равномерную, так и случайные компоненты. Все эти искажения ведут к потере информации и могут быть скомпенсированы лишь частично (см.

3.2). Несмотря на разнообразие статистических свойств внешних искажений их объединяет необходимость устранения ошибок путем фильтрации изображений в соответствии со схемой рис. 1.2. Минимизация ошибок должна осуществляться оптимизацией характеристик кодирующего устройства, в роли которого выступает телекамера (см. 3.3).

Реальные телекамеры всегда формируют интересующие наблюдателя сигналы не на идеально ровном фоне, а в смеси с изображениями посторонних объектов и паразитными сигналами фотоприемников. Поэтому различные задачи видеоинформационных систем — выделение сигнала из помех, классификация изображений, оценка их параметров — начинаются с предварительной фильтрации изображений. Во многих приложениях мешающий фон можно моделировать случайным сигналом. В этих распространенных на практике случаях система передачи изображений рис. 4.1 моделируется с учетом двух источников сигнала — полезного и фонового. При этом считается, что площадь и интенсивность полезного сигнала меньше, чем фонового, который и определяет суммарную плотность вероятности изображения. Задача фильтрации полезного сигнала из его смеси с фоном и шумом сводится к выделению сигнала из смеси двух по-разному коррелированных случайных сигналов. При этом отсчеты шума, как правило, не коррелированы, а фоновый сигнал коррелирован и по полю и по времени. Известные процедуры выделения сигнала из «окрашенного» шума опираются на формирование отношения правдоподобия и имеют двухкаскадную каноническую реализацию, состоящую из фильтра, согласованного с полезным сигналом, и оценивающе-вычитающего устройства [8].

Предварительная переработка информации необходима для получения наиболее точной оценки изображения. Информационные характеристики изображений на этом этапе переработки играют двойную роль. Во-первых, потенциальное количество информации $I(X(t), Y(t))$ определяет достижимую ошибку оценивания фонового изображения, моделируемого случайным процессом [8]

$$\varepsilon_в = 2\sigma^2_1 I(X(t), Y(t)).$$

Во-вторых, ошибка оценивания фонового изображения выступает в роли лимитирующего фактора в процессе принятия решения. Если в отсутствии фонового изображения вероятности ошибочных решений потенциально ограничены энергетическим отношением сигнал-шум $\psi = E/\sigma^2_1$, то при наличии фона, который необходимо компенсировать, ошибки в решении определяются энергетическим отношением сигнала к ошибке оценивания E/ε. Естественно, что чем хуже оценка (чем сильнее ε превышает $\varepsilon_в$), тем больше вероятность ошибочного решения, принимаемого классификатором. В автоматических видеоинформационных системах именно этот аспект и вызывает пристальное внимание к информационным характеристикам телекамер (см. 3.8).

Фильтрация сигналов изображений может быть линейной и нелинейной. Отличия между ними состоят в следующем. Во-первых, нелинейная фильтрация обычно сложнее линейной, т. е. требует больших вычислительных затрат. Во-вторых, нелинейные фильтры часто специфичны и при отклонении характеристик реальных сигналов от принятой модели резко снижают помехоустойчивость системы. Вместе с тем, в ряде приложений, характеризуемых стабильностью формы сигналов, применение нелинейной фильтрации может дать хорошие результаты [55]. В качестве примера приведем результаты нелинейной фильтрации сигналов плоских неоднородных фигур при выделении их из неравномерного фона (рис. 4.2). Вместе с тем линейные фильтры получили существенно большее применение. Это связано с тремя обстоятельствами [3, с. 706]. Во-первых, линейные фильтры абсолютно оптимальны при фильтрации гауссовских сигналов. Во-вторых, линейные фильтры просты, их легко синтезировать и существует развитая теория их построения, чего нельзя сказать о нелинейных фильтрах. В-третьих, можно применять линейную теорию просто из-за отсутствия другой теории, так как неполное решение лучше, чем отсутствие решения вообще.

Проиллюстрировано применение линейной фильтрации может быть на задаче *контрастирования* полезного сигнала при его выделении из фона. Повышение контраста сигнала, тривиальное при равномерном фоне, становится сложной задачей при пространственно неоднородном подвижном фоне. При оценке качества конт-

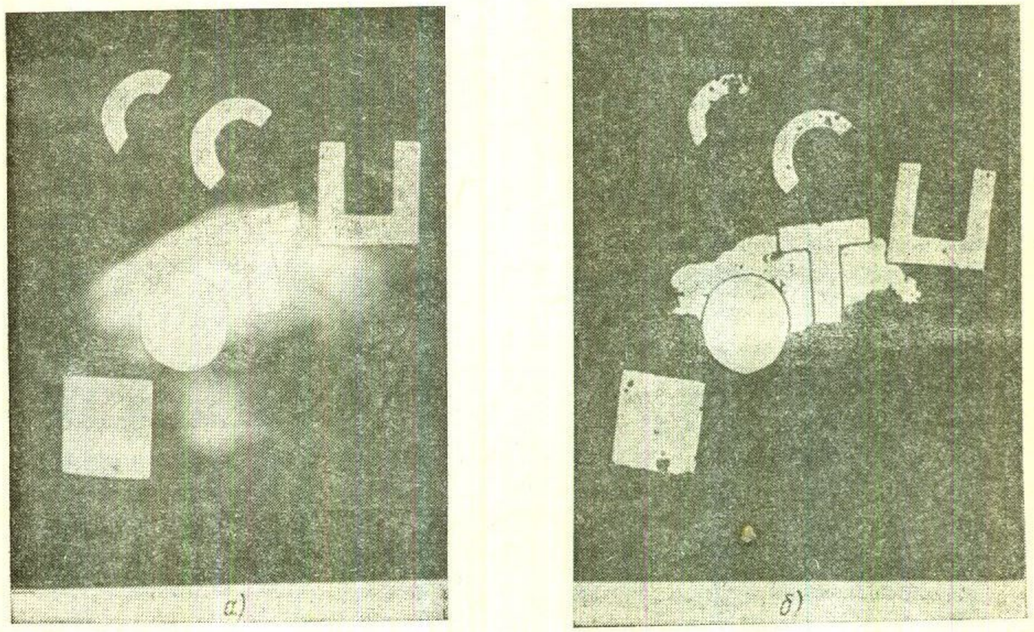


Рис. 4.2. Изображение группы объектов на неравномерном фоне до (*а*) и после (*б*) нелинейной фильтрации

растирования сигнала следует расширить понятие контраста, учтя неравномерность фона, которая определит *локальный контраст*. Потенциальное качество подавления фона определяется шумом фона (имеющим место и при равномерном фоне) и энергией на степень свободы переменной по полю случайной компоненты фона. Пусть изображение содержит кроме интересующего наблюдателя *подвижного* малоконтрастного объекта сложный подвижный фон с гауссовской функцией распределения. Тогда сигнал объекта будет контрастирован наилучшим образом, если из сигнала $\mathbf{Z}_n$ текущего кадра будет вычтена оптимальная оценка фонового сигнала (условное математическое ожидание), построенная с помощью пространственно-временного винеровского фильтра:

$$\mathbf{Z}^*_n = \mathbf{Z}_n - \mathbf{M}(Z_n | \mathbf{Z}_{n-1}, \mathbf{Z}_{n-2}, \dots).$$

Будем считать, что наблюдению доступны лишь два кадра — текущий $\mathbf{Z}_n$ и предыдущий $\mathbf{Z}_{n-1}$. Тогда оптимальная линейная оценка первого кадра по второму будет определяться функцией регрессии сигнала $\mathbf{Z}_{n-1}$ на сигнал $\mathbf{Z}_n$ и ее можно будет вычислить через автоковариационную $\mathbf{\Sigma}_{22}$ и взаимноковариационную $\mathbf{\Sigma}_{12}$ матрицы первого и второго кадров изображения [81]:

$$\mathbf{M}(\mathbf{Z}_n | \mathbf{Z}_{n-1}) = \mathbf{Z}_{n-1}\mathbf{\Sigma}_{12}\mathbf{\Sigma}^{-1}{}_{22}. \qquad (4.1)$$

Структурная схема видеопроцессора (рис. 4.3), соответствующая этому алгоритму, может быть реализована по параллельно-конвейерному методу. Особенностью корреляторов в этом вычислителе является возможность использования небольших размеров по полю измеряемых корреляционных матриц. Необходимо лишь, чтобы этот размер был не менее смещения изображения от кадра к кадру. При этом, если такое смещение может достигать заметной величины, следует произвести предварительно измерение смещения и компенсацию целочисленного сдвига изображения. Возможность сокращения вычислительных затрат при

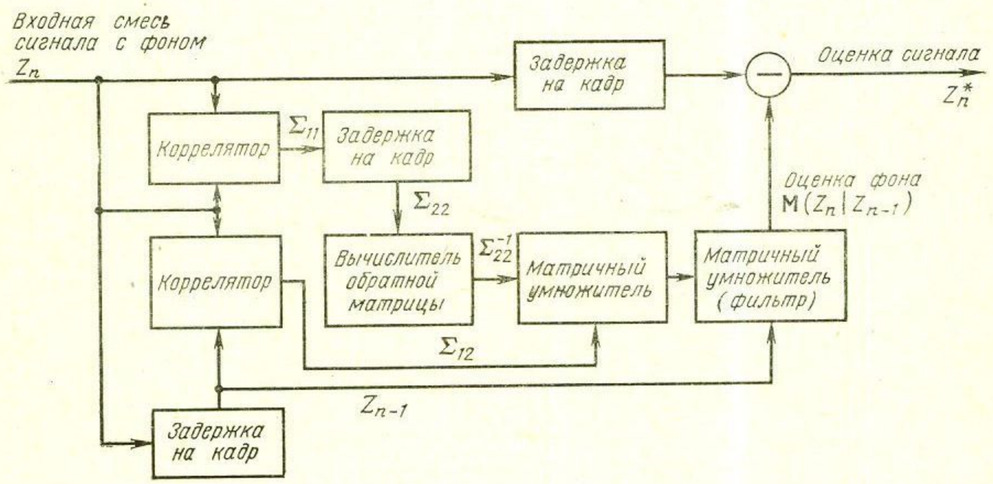


Рис. 4.3. Структурная схема видеопроцессора, реализующего линейный алгоритм (4.1) подавления фона

этом базируется на применении (на первом этапе компенсации движения) методом полихотомии (см. 4.4 и 4.5). Эффективность линейного оценивания подвижных изображений иллюстрируется рис. 4.4, показывающим, что алгоритм (4.1) позволяет получить оценку с ошибкой, тонущей в шуме, даже для изображения, отличающегося от гауссовского.

Выделению полезного сигнала может препятствовать множество фоновых изображений. При этом мешающие объекты группируются по видимой скорости перемещения. Для компенсации такого многокомпонентного фона потребуется многокадровая память, число кадров в которой на единицу превышает число фоновых компонентов с различными скоростями движения. Важной особенностью алгоритмов многокадровой переработки телевизионных изображений является возможность работы в реальном времени, т. е. принятия решения сразу же по окончании текущего кадра. Задержка на несколько кадров, пропорциональная числу кадровых ОЗУ, означает лишь увеличение времени «выхода на рабочий режим»; в установившемся режиме дополнительные задержки, кратные времени кадра, не возникают. Остаются лишь задержки, возникающие в ходе вычислений, которые обычно стремятся выполнить за время кадрового «гасящего сигнала». Этот термин, возникший при становлении электронно-лучевого вещательного телевидения, пока используется и в видеоинформационных системах, хотя смысл его в твердотельных телевизионных системах с автоматическим получателем изменился. Для ПЗС, например, это время переноса зарядового изображения из секции накопления в секцию хранения (см. гл. 3).

Важно, что линейная фильтрация может быть легко распространена и на случай большего числа доступных анализу кадров. Полное решение одновременного анализа всех кадров [80] достаточно громоздко. Вместе с тем, в практике видеоинформационных систем весьма распространена ситуация, когда наблюдению полезного объекта мешает фон, содержащий две компоненты —

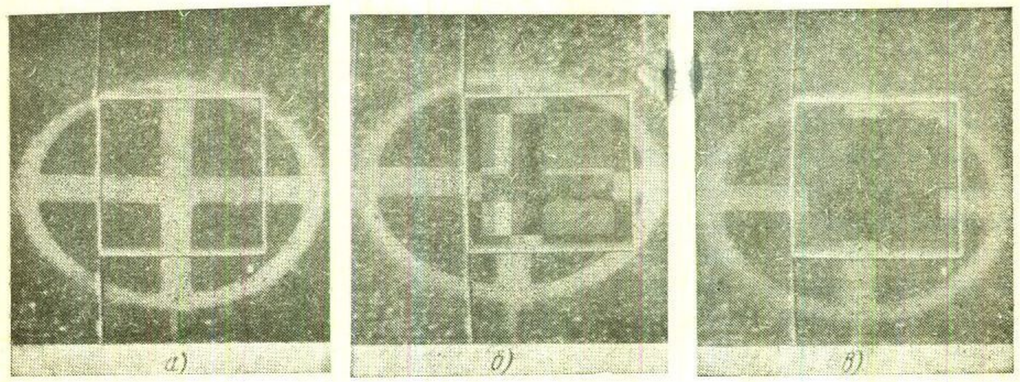


Рис. 4.4. Оценка подвижного изображения: *а*) исходное; *б*) межкадровая разность; *в*) вычитание оценки по линейному алгоритму (4.1)

подвижную и неподвижную. Подвижная компонента обусловлена движением телекамеры (например, в мобильных роботах, см. гл. 5), неподвижная — темновым сигналом фотоприемника. Для компенсации такого двухкомпонентного фона необходимо и достаточно проанализировать три кадра изображения. При этом возможны варианты линейных алгоритмов выделения полезных сигналов, различающихся расположением оцениваемого кадра в группе анализируемых. Для минимизации ошибки оценивания следует формировать оценку среднего из трех кадров $\mathbf{M}(\mathbf{Z}_{n-1}|\mathbf{Z}_n, \mathbf{Z}_{n-2})$. Уменьшение ошибки при этом достигается за счет увеличения задержки на время кадра. Если в системе это не допустимо, то формируют оценку $\mathbf{M}(\mathbf{Z}_n|\mathbf{Z}_{n-1}, \mathbf{Z}_{n-2})$ последнего кадра по предыдущим. В обоих случаях предварительно измеряются взаимокорреляционные матрицы Σ_{ij} трех кадров $ij=(1, 2, 3)$. Алгоритм работы принимает в этом случае следующий вид [81]:

$$\mathbf{M}(\mathbf{Z}_n|\mathbf{Z}_{n-1}, \mathbf{Z}_{n-2}) = \mathbf{\Sigma}_{13}\mathbf{\Sigma}_{33}^{-1}\mathbf{Z}_{n-2} + (\mathbf{\Sigma}_{12} - \mathbf{\Sigma}_{13}\mathbf{\Sigma}_{33}^{-1}\mathbf{\Sigma}_{32})\times$$
$$\times(\mathbf{\Sigma}_{22} - \mathbf{\Sigma}_{23}\mathbf{\Sigma}_{33}^{-1}\mathbf{\Sigma}_{32})^{-1}(\mathbf{Z}_{n-1} - \mathbf{\Sigma}_{23}\mathbf{\Sigma}_{33}^{-1}\mathbf{Z}_{n-2}). \quad (4.2)$$

Так же как и алгоритм (4.1), алгоритм (4.2) является функцией регрессии сигналов $\mathbf{Z}_{n-1}$ и $\mathbf{Z}_{n-2}$ на $\mathbf{Z}_n$. Отличительной особенностью алгоритма (4.2) является его адаптивность к изменению структуры фона. Несмотря на то, что алгоритм (4.2) предназначен для компенсации двухкомпонентного фона, он является оптимальным в классе фильтров с двумя кадровыми ОЗУ и при наличии лишь одной компоненты фона может быть полезен для компенсации однокомпонентного фона в условиях грубой пространственной дискретизации, вызывающей неинвариантность сигнала фона к его сдвигу (см. раздел 4.3).

Отметим, что при линейной фильтрации, особенно при межкадровой, использующей кадровые ОЗУ с большой емкостью памяти, встает задача экономии этой памяти (см. раздел 4.5). Так как емкость памяти в определенном смысле эквивалентна пропускной способности канала связи [4], то решением этой задачи является полученный в гл. 2 передаточный фильтр, который в рассматриваемом случае обеспечит минимум объема памяти ОЗУ при заданной ошибке и предварительном представлении изображения его спектральными коэффициентами.

Применение критерия минимума среднеквадратической ошибки и гауссовской модели обусловливает линейность обработки изображений. Она, в свою очередь, обеспечивает возможность создания регулярных структур процессоров параллельно-конвейерного типа или систолических. Это является важной предпосылкой для реализации быстродействующих процессоров обработки изображений в виде СБИС [53]. Реализация набора СБИС для обработки сигналов позволит на основе транспьютерной методологии в большой степени унифицировать видеопроцессоры предварительной обработки изображений. Можно выделить три группы

функциональных узлов, для которых возможна разработка заказных СБИС:

1. СБИС, предназначенные для операций непосредственно над видеосигналом. К ним относятся устройства фильтрации, корреляторы, измерители координат объектов. Отличительной чертой этих устройств является применение алгоритма свертки или подобных ему алгоритмов.

2. СБИС для операций над корреляционными матрицами: перемножения, сложения, приведения к треугольному виду и обращения. Все эти операции реализуются на реконфигурируемом массиве процессоров, т. е. процессорной матрице с перестраиваемыми связями.

3. Элемент систолического обрабатывающего массива для выполнения арифметических операций. Такая СБИС должна включать в себя двунаправленные порты ввода-вывода, стек, умножитель, делитель, сумматор и схему извлечения квадратного корня.

В результате предварительной переработки информации в реальном времени происходит не только компенсация искажений сигнала, но и некоторое сокращение требуемой пропускной способности для передачи изображения. Это означает, что в схеме рис. 4.1 пропускная способность C_2 несколько меньше пропускной способности устройства ввода. Основное сокращение пропускной способности устройства вывода видеокомпьютера, определяемой его решающей способностью $V_{\text{реш}}$ (см. 3.1), достигается за счет принятия решения в классификаторе.

4.2. КЛАССИФИКАЦИЯ ИЗОБРАЖЕНИЙ

Классификация изображений, или распознавание образов — это центральная проблема видеоинформатики. Она захватывает интересы как родственных видеоинформатике теоретических дисциплин кибернетики и искусственного интеллекта, так и их практического приложения — робототехники. Проблема классификации изображений, будучи междисциплинарным направлением, очень сложная. Мало того, что вывод о большой трудности задачи распознавания образов считают основным итогом ее изучения, но эта высокая сложность усугубляется из-за влияния внешнего фона (см. раздел 4.1) и неинвариантности отклика дискретного фотоприемника к сдвигу, повороту и масштабированию изображения (см. раздел 3.6). Отличительными чертами классификации изображений в реальном времени являются обработка подвижных изображений, позволяющая различать их по характеру движения, и специфические ограничения на вычислительную сложность и быстродействие процессора.

Проблему классификации изображений можно формулировать как в статистической [8, 31, 72—74, 76], так и в нестатистической [52, 73, 74] постановке. Обе эти постановки допускают классификацию изображений на основе анализа как пространственных

признаков, так и спектральных характеристик излучения (так называемые спектрозональные видеоинформационные системы [71]). При построении классификаторов изображений, применимых в телевизионных роботах, используют статистические и лингвистические методы распознавания изображений на основе анализа их пространственных признаков. Важной чертой статистической постановки является то, что она позволяет не только синтезировать структуру классификатора, но и оптимизировать его параметры и определить помехоустойчивость, т. е. достоверность классификации. Здесь мы не преследуем цель давать обзор состояния вопроса классификации изображений, за которым читателю следует обратиться к работам [73, 74], а лишь воспользуемся статистическим подходом в целях выявления потенциальных *возможностей* и принципиальных *ограничений*, присущих классификации изображений. Основу статистического подхода к классификации образов составляет *байесовская теория принятия решений* [8, 76]. Этот подход предполагает наличие у классификатора определенного объема априорной информации о свойствах различаемых изображений и помех. Получение априорной информации классификатором называют его обучением (с учителем или без него). В типичной, естественно сильно упрощенной, задаче классификации изображений считается заданным ансамбль из M классов изображений Ω_i, $i=1, \ldots, M$. Предъявляемые классификатору сигналы изображений X^i обладают энергией $E_i=\sum_k (X^i{}_k)^2$ (k — номера элементов изображения). Изображения различных классов моделируются сигналами либо с известной формой (и неизвестными мешающими параметрами), либо случайными (текстурами). Мешающими факторами в обоих случаях является белый шум с дисперсией σ^2 и неизвестность параметров сигналов — сдвига, поворота, масштаба — в первом случае. Байесовская теория решений позволяет найти минимум среднего риска, т. е. потерь от ошибочных решений. В частном случае, когда потери считаются не зависящими от номеров классов, минимизация среднего риска обеспечивается за счет минимизации вероятности ошибок классификации. При этом задача различения M гипотез $H(\Omega_i) \triangleq H_i$ при обеих моделях сигналов сводится к формированию $M(M-1)/2$ отношений правдоподобия и сравнению их с порогом γ_{ij}. Обозримое аналитическое решение этой задачи достигается лишь при гауссовском распределении сигналов $\mathbf{Z}(\Omega_i)$, т. е. если она сводится к общей гауссовской задаче [8, т. 1] для сигналов известной формы, либо при различении гауссовских случайных полей [8, т. 3]. Классификация в обоих случаях опирается на вычисление достаточных статистик, использующих наблюдения полезного сигнала $X^i(t)$ в смеси с шумом $\xi(t)$

$$Y(t)=X^i(t)+\xi(t), \quad i=1, \ldots, M.$$

Задача различения сигналов известной формы решена Хелстромом [1, 14]. Ее решение основано на использовании фильтра,

согласованного с разностным сигналом по сравниваемым гипотезам. В обобщенном случае сигналов известной формы, когда изображения могут иметь произвольные сдвиг, поворот и масштаб, такое решение практически не применимо, и используют переход к пространству признаков, инвариантных к этим изменениям формы сигнала. В задаче различения случайных сигналов, как будет показано ниже, решение опирается также на разностную характеристику сигналов по соответствующим гипотезам. Только это уже не разностный сигнал, а разность энергетических спектров, которая определяет структуру оптимального линейного различающего фильтра. В обоих случаях при классификации осуществляется переход к пространству признаков, в котором каждый класс изображений характеризуется вектором признаков $\mathbf{Z}=\sum Z_k$, а каждая компонента Z_k которого задается не каким-либо одним числом, а целой областью значений. Размер этих областей монотонно убывает с ростом отношения сигнал-шум $\psi_i=E_i/\sigma^2$.

Априорная информация, необходимая для вычисления достаточных статистик в пространстве признаков при сигналах известной формы, состоит из математических ожиданий $\mathcal{M}_i$ и ковариационных матриц рассеяния Σ^i, а при случайных сигналах — из ковариационных матриц K_i или их спектров $\sum_k \lambda^i{}_k$ разложения по ортогональным базисным функциям $\{\varphi_k\}$, обеспечивающим независимость спектральных компонент $\lambda^i{}_k$, Z_k, ξ_k при истинности любой гипотезы H_i.

Исходная форма записи критерия отношения правдоподобия оперирует вероятностями данного наблюдения $\mathbf{Z}$ при соответствующих гипотезах:

$$\Lambda_{ij}(\mathbf{Z})=\frac{p(\mathbf{Z}|H_i)}{p(\mathbf{Z}|H_j)} \underset{H_j}{\overset{H_i}{\gtrless}} 1; \quad \ln \Lambda_{ij}(\mathbf{Z}) \underset{H_j}{\overset{H_i}{\gtrless}} 0. \tag{4.3}$$

В общем случае классификации сигналов известной формы достаточные статистики вычисляются в виде квадратичных форм. Отношение правдоподобия для классов Ω_i и Ω_j имеет вид [8, т. 1, 72—74, 76]:

$$l_{ij}(\mathbf{Z})=(\mathbf{Z}-\mathcal{M}_i)^{\mathrm{T}}\Sigma^{-1}{}_i(\mathbf{Z}-\mathcal{M}_i)-(\mathbf{Z}-\mathcal{M}_j)^{\mathrm{T}}\Sigma^{-1}{}_j(\mathbf{Z}-\mathcal{M}_j) \underset{H_j}{\overset{H_i}{\gtrless}} \gamma_{ij}. \tag{4.4}$$

Порог γ_{ij} определяется ковариационными матрицами, математическими ожиданиями и априорными вероятностями $p(\Omega_i)$ и $p(\Omega_j)$ по сравниваемым гипотезам. В простом случае равных ковариационных матриц достаточные статистики становятся линейными и отношение правдоподобия принимает вид

$$l_{ij}(\mathbf{Z})=\mathbf{Z}^{\mathrm{T}}\Sigma^{-1}(\mathcal{M}_i-\mathcal{M}_j) \underset{H_j}{\overset{H_i}{\gtrless}} \gamma_{ij}. \tag{4.5}$$

Геометрически выбор гипотез по (4.4) соответствует выбору $H(\Omega_i)$ того класса, до которого данный вектор $\mathbf{Z}$ имеет наименьшее махаланобисово расстояние, выбор по (4.5) осуществляется

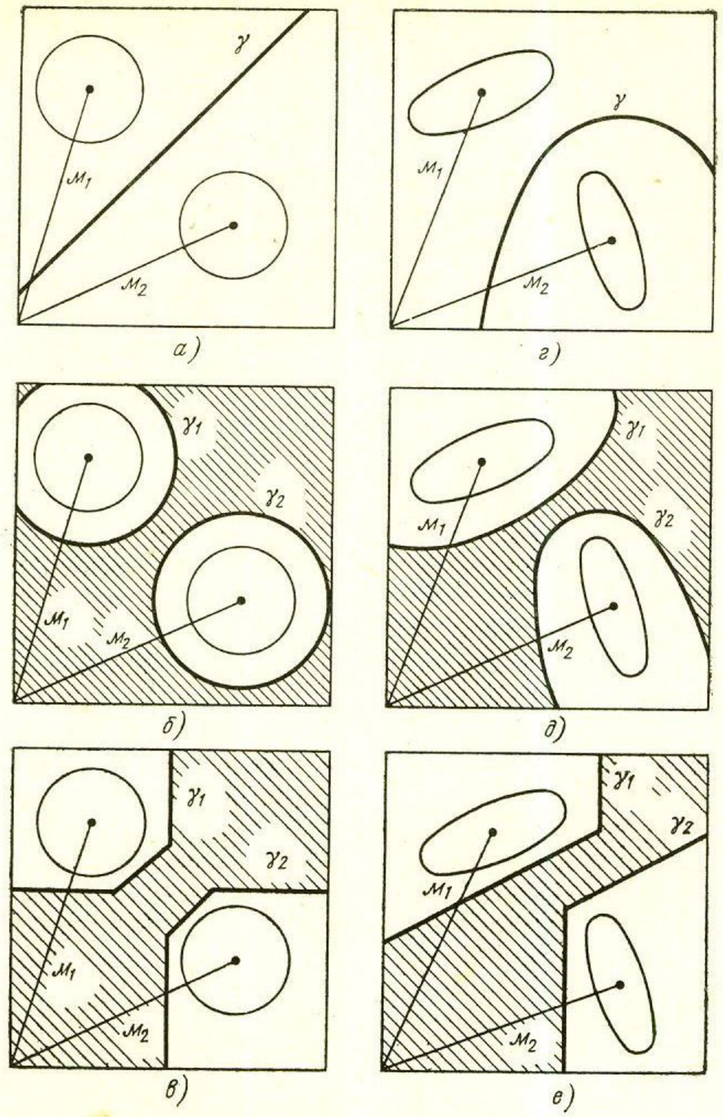


Рис. 4.5. Пороги классификации γ_i при равных (*а*—*в*) и неравных (*г*—*е*) ковариационных матрицах (заштрихованы области отказа от решения)

по минимуму евклидова расстояния. При равных ковариационных матрицах оптимальная различающая функция линейна (рис. 4.5,*а*), при неравных—нелинейна (рис. 4.5,*г*). Нелинейность оптимальной различающей функции появляется и при введении отказа от решения — отнесения к фону, т. е. к $(M+1)$-му классу объектов, имеющих признаки, сильно отличающиеся от математических ожиданий для заданных классов (рис. 4.5,*б*, *д*). Упрощение вычислений в этом случае достигается за счет использования нескольких линейных различающих функций (рис. 4.5,*в*, *е*).

Классификация случайных сигналов также осуществляется с помощью критерия отношения правдоподобия (4.3), которое (при

нулевых математических ожиданиях всех различаемых процессов и наблюдений на фоне белого шума с дисперсиями спектральных отсчетов σ^2) выражается через спектральные коэффициенты $\lambda^i{}_k$ различаемых стационарных сигналов:

$$\sum_k \frac{\lambda_k^i - \lambda_k^j}{\left(\lambda_k^i + \sigma^2\right)\left(\lambda_k^j + \sigma^2\right)} \cdot Z_k^2 \underset{H_j}{\overset{H_i}{\gtrless}} \sum_k \ln \frac{\lambda_k^i + \sigma^2}{\lambda_k^j + \sigma^2} \triangleq \gamma_{ij}. \tag{4.6}$$

Согласно методике построения отношения правдоподобия для случайных процессов [8, т. 3], суммирование в (4.6) ведется сначала по конечному числу членов. Затем, устремляя его к бесконечности, можно перейти к эквивалентным формам записи отношения правдоподобия через частотную характеристику $\mathcal{H}^{ij}{}_k$ или импульсную характеристику $h_{ij}(t,\ u)$ оптимального фильтра:

$$\int\int Z(t)\, h_{ij}(t,\ u)\, Z(u)\, dt\, du \underset{H_j}{\overset{H_i}{\gtrless}} \gamma_{ij}; \tag{4.7}$$

$$\mathbf{Z}^{\mathrm{T}} \mathcal{H}_{ij} \mathbf{Z} \underset{H_j}{\overset{H_i}{\gtrless}} \gamma_{ij}.$$

Хотя достаточная статистика, входящая в отношение правдоподобия (4.6), вычисляется с помощью нелинейного преобразования (используется операция возведения в квадрат), фильтр $\mathcal{H}^{ij}$ относится к классу оптимальных линейных фильтров, который назовем различающим. Его спектральная характеристика может быть записана через спектры различаемых сигналов $\lambda^i{}_k$ и $\lambda^j{}_k$ в нескольких удобных либо для вычислений, либо для понимания видах, например, включающих частотные характеристики винеровских фильтров $l^i{}_{Bk}$ и $l^j{}_{Bk}$ оценки сигналов по соответствующим гипотезам:

$$\mathcal{H}_k^{ij} = \frac{\lambda_k^i - \lambda_k^j}{\left(\lambda_k^i + \sigma^2\right)\left(\lambda_k^j + \sigma^2\right)} = \frac{1}{\lambda_k^j + \sigma^2} - \frac{1}{\lambda_k^i + \sigma^2} =$$

$$= l_{Bk}^i\, l_{Bk}^j \left(\frac{1}{\lambda_k^j} - \frac{1}{\lambda_k^i}\right). \tag{4.8}$$

Порог γ_{ij} в критерии отношения правдоподобия (4.6), (4.7) определяется разностью априорных количеств информации по соответствующим гипотезам:

$$\gamma_{ij} = \frac{1}{2} \sum_k \ln \frac{\lambda_k^i + \sigma^2}{\lambda_k^j + \sigma^2} = \frac{1}{2} \sum_k \ln \left(1 + \frac{\lambda_k^i}{\sigma^2}\right) -$$

$$- \frac{1}{2} \sum_k \ln \left(1 + \frac{\lambda_k^j}{\sigma^2}\right) = I(Y(t),\ X^i(t)) - I(Y(t),\ X^j(t)). \tag{4.9}$$

Критерий отношения правдоподобия (4.6), (4.7) предписывает оптимальную процедуру переработки информации в ходе принятия решения и может быть реализован в виде ряда канонических

схем, наиболее известной из которых является структура «фильтр—квадратор—интегратор». В частном случае обнаружения случайного процесса, когда сигнал по одной из гипотез (H_j) равен нулю, критерий отношения правдоподобия (4.6), (4.7), различающий фильтр (4.8) и порог (4.9) переходят к более простой форме [8, т. 3].

При классификации случайных сигналов видеокомпьютер может моделироваться схемой, состоящей из $M(M-1)/2$ каналов, в каждом из которых осуществляется линейное кодирование с передачей сигналов через канал связи с ограниченной пропускной способностью при наличии ограничения на кодирующее устройство G^{ij} (рис. 4.6). Кодек в каждом канале этой структуры аналогичен кодеку, детально рассмотренному в гл. 2. Он состоит из линейных оптимальных по критерию минимума среднеквадратической ошибки различающего L^{ij}, передаточного G^{ij} и винеровского декодирующего Q^{ij} фильтров. Отличие от кодека (2.71) состоит в том, что линейный входной фильтр L^{ij} оценивает не просто входной сигнал, а в соответствии с (4.8) выделяет различие между сигналами по соответствующим гипотезам. При том же ограничении суммы квадратов коэффициентов кодера оптимальный передаточный фильтр в схеме рис. 4.6 определится также спектральным отношением сигнал-шум $(\psi^{ij}{}_k)^2 = a^2\,\hat{\lambda}^{ij}{}_k/\sigma^2{}_2$ на выходе оптимального различающего фильтра L^{ij}:

$$(g_k^{ij})^2 = \frac{(\psi_k^{ij})^2 - 1}{(\psi_k^{ij})^2}.$$

Спектр оцениваемого различия в сигнале $\hat{\lambda}^{ij}{}_k$ определяется усреднением спектров сигналов $\lambda^i{}_k$ и $\lambda^j{}_k$, прошедших оптимальный различающий фильтр $l^{ij}{}_k$. При равных вероятностях гипотез H_i и H_j спектр

$$\hat{\lambda}^{ij}{}_k = [(\lambda^i{}_k + \lambda^j{}_k)/2 + \sigma^2{}_1]\,(l^{ij}{}_k)^2.$$

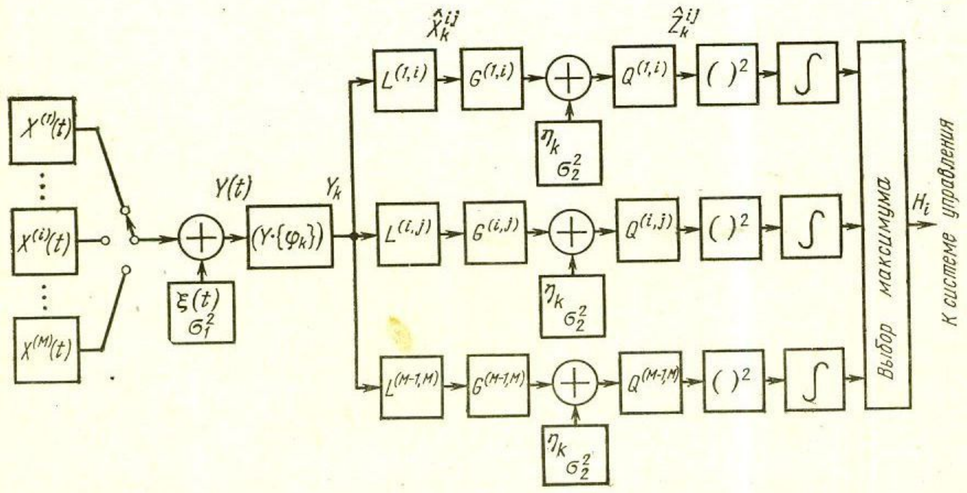


Рис. 4.6. Оптимальный многоканальный приемник для различения гауссовских сигналов

Подчеркнем, что оптимальный передаточный фильтр G^{ij} имеет конечное число N ненулевых коэффициентов ($N = N' + [N''' - N'']$ на рис. 4.7). Оптимальное число N коэффициентов (признаков) находится исходя из ограниченной суммы их квадратов, которая в существенной мере определяет сложность видеокомпьютера. Это означает, что число признаков при различении гауссовских случайных сигналов усекается решением вариационной задачи при наличии ограничения (см. гл. 2), а не относится к ситуации, когда «в некоторых случаях (малая ЭВМ, большой объем данных) необходимо по причинам чисто вычислительного характера на этапе обучения уменьшить число признаков» [74]. Аналогичный подход к вырождению числа ранжированных по информативности независимых признаков применим и к сигналам известной формы.

Для выявления потенциальных возможностей классификации изображений ниже рассмотрим типичную для телевизионных роботов ситуацию классификации изображений сигналов известной формы. Пусть наблюдается изображение, состоящее из $N_э$ элементов разложения твердотельного фотоприемника. При этом на растре порознь или одновременно могут присутствовать изображения X^i, принадлежащие к одному из M классов Ω_i. Каждый сигнал i-го класса может иметь произвольные в рамках заданного растра с $N_э$ элементами координаты, поворот и масштаб. Вероятностная структура пространства признаков соответствует общей

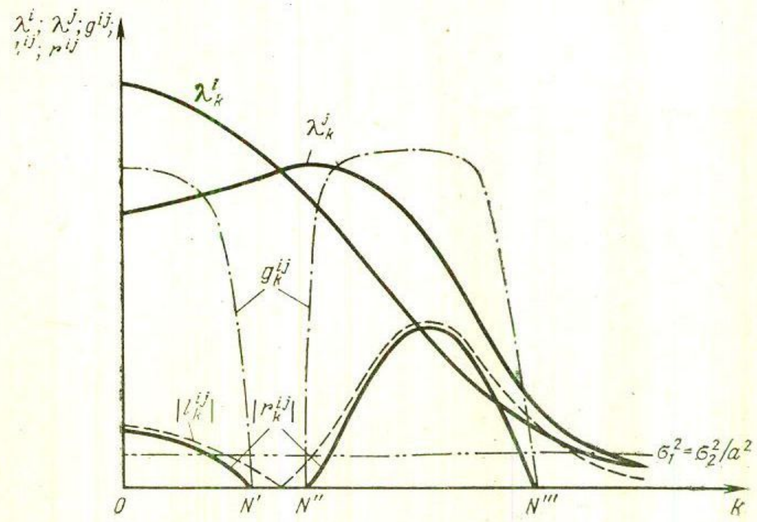


Рис. 4.7. Спектральные характеристики различаемых гауссовских сигналов λ^i, λ^j и оптимальных фильтров: передаточного g^{ij}, различающего l^{ij} и их произведения r^{ij}

гауссовской задаче и классификация осуществляется на основе отношения правдоподобия (4.4).

В этих условиях достоверность классификации определяется выбором признаков образов и числом этих признаков. Подчеркнем, что выбор признаков образов в большинстве случаев является неформализуемой процедурой, более того, именно проблема выбора признаков является «сердцевиной всех затруднений распознавания образов» [74].

Выбор признаков переходит из ранга неформализуемого поиска в ранг задач статистического синтеза лишь при моделях изображений в виде случайных полей. В этом случае возможен формализованный выбор признаков образов — в их роли выступают проекции на базисные функции разложения изображений в ортогональные ряды.

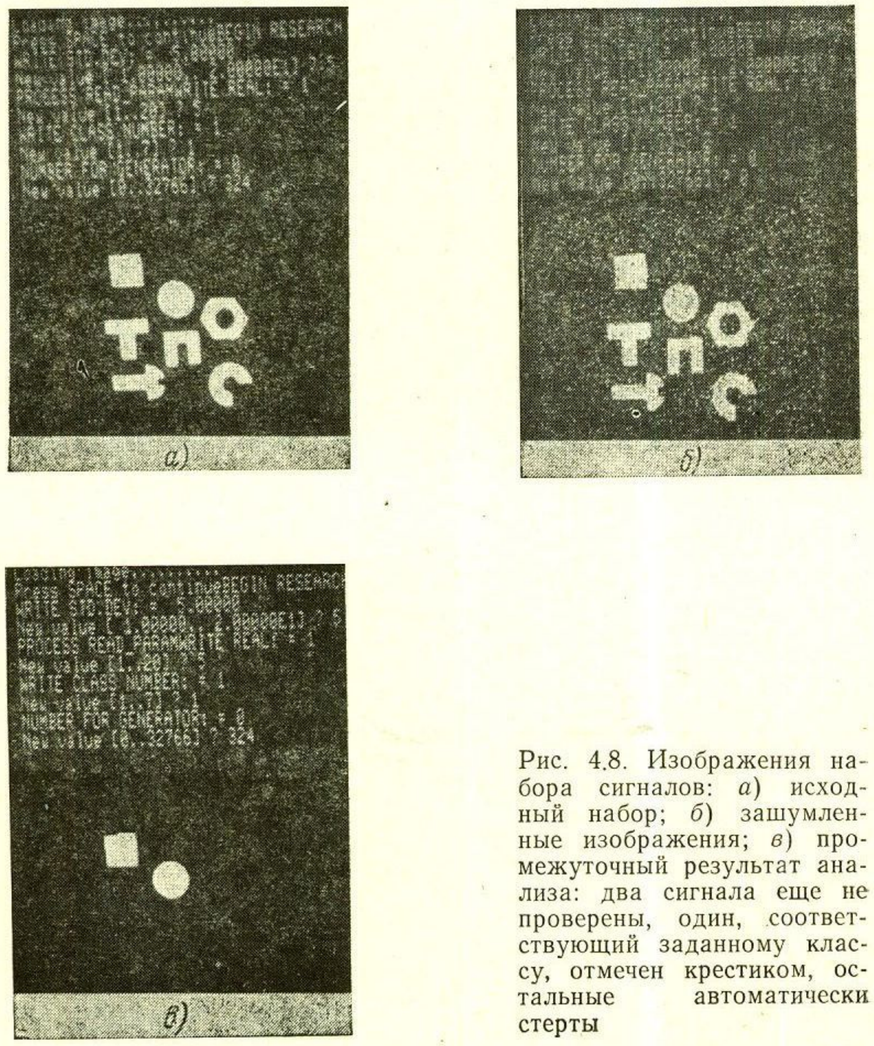


Рис. 4.8. Изображения набора сигналов: *а*) исходный набор; *б*) зашумленные изображения; *в*) промежуточный результат анализа: два сигнала еще не проверены, один, соответствующий заданному классу, отмечен крестиком, остальные автоматически стерты

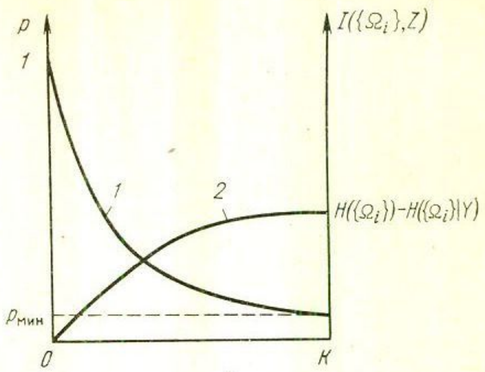


Рис. 4.9. Зависимость ошибки (*1*) и полезной информации (*2*) от числа признаков

Основное требование к признакам при различении изображений известной формы состоит в их *инвариантности* к изменениям образов — сдвигу, повороту, масштабу. Для простых бинарно-квантованных изображений (рис. 4.8,*а*) это могут быть функции их геометрических параметров — отношения квадрата периметра к площади или момента инерции к площади и тому подобные нормированные характеристики.

Признаком объектов может быть их ориентация относительно сторон растра. Для ее оценки используют построение по методу наименьших квадратов прямой, проходящей через центр тяжести сигнала объекта, и максимально удаленные друг от друга точки его контура [89]. Для повышения скорости вычисления ориентации объекта следует применять алгоритмы, допускающие *распараллеливание* вычислений. Например, ориентация объекта может быть оценена по параллельно измеряемым его сечениям, проведенными под разными углами через центр тяжести, или по проекциям на различные направления. При гауссовском распределении сигналов ошибки классификации вычисляются по энергетическому отношению сигнал-шум $\psi_{ij}=E_{ij}/\sigma^2$. Здесь через E_{ij} обозначена энергия разностного сигнала для образов классов Ω_i и Ω_j при их наилучшем совмещении путем сдвига и поворота. При этом минимальная средняя ошибка классификации зависит от числа используемых признаков и монотонно убывает от единицы до

$$p_{\text{мин}} = \frac{2}{M(M-1)} \sum_{i,\ j;\ j<i}^{M(M-1)/2} p_{ij\ \text{мин}} =$$

$$= \frac{4}{M(M-1)} \sum_{i=1}^{M} \sum_{j=1,\ j<i}^{(M-1)/2} \Phi(-\psi_{ij}/2),$$

где Φ — интеграл вероятности (рис. 4.9). То, что ошибка классификации с ростом числа признаков убывает не до нуля, вызывает желание разработчика усечь число используемых признаков. Сделано это может быть с учетом рассмотрения информационных характеристик сюжета и видеокомпьютера.

Предъявляемое классификатору изображение характеризуется энтропией ансамбля $H(\{\Omega_i\})$, определяемой вероятностями появления образов по формуле Шеннона:

$$H(\{\Omega_i\}) = -\sum_{i=1}^{M} p(\Omega_i) \log p(\Omega_i). \qquad (4.10)$$

В простейшем случае равенства вероятностей появления образов

$p(\Omega_i)=\text{const}$ энтропия ансамбля изображений определяется по формуле Хартли:

$$H(\{\Omega_i\})=\log M. \tag{4.11}$$

Само изображение характеризуется кодовой энтропией независимо от данной задачи классификации. Вычисление кодовой энтропии в гауссовском случае рассмотрено в гл. 2. В случае наблюдения бинарных изображений она определяется по формуле Шеннона, в которую входит вероятность принятия элементом изображения «единичного» значения, а суммирование проводится по всем $N_э$ элементам растра. Растр, на котором могут наблюдаться произвольные бинарные изображения, характеризуется энтропией

$$H(\mathbf{Y})=-N_э\sum_{i=1}^{2} p_i\log p_i=N_э. \tag{4.12}$$

Если объект (объекты) имеет размер, определяемый числом элементов $N_{об}$, то кодовая энтропия изображения не превысит

$$H'(\mathbf{Y})=N_{об}\log\frac{N_э}{N_{об}}. \tag{4.13}$$

Это значение кодовой энтропии достигается лишь при статистической независимости отсчетов, составляющих изображение образа, которой обладает белый шум. Более точная оценка кодовой энтропии бинарного изображения вычисляется по формуле (4.13) при условии, что изображение подвергнуто идеальному кодированию по Шеннону. Можно приближенно считать, что это предельное число необходимых отсчетов оценивается числом контурных элементов $N_к$. Это является определенным объяснением целесообразности выделения контуров на изображении (декорреляции) при обработке сигналов в процессоре классификатора изображений [20, 91]. Вектор признаков характеризуется условной энтропией

$$H(\mathbf{Z})=\sum_{j=1}^{K}\sum_{i=1}^{M}\log\frac{\mathcal{M}_{ij}}{\varepsilon_{ij}}, \tag{4.14}$$

где $\mathcal{M}_{ij}$ — математическое ожидание i-го признака для j-го класса; ε_{ij} — ошибка измерения этого признака. После принятия решения в пользу какой-либо гипотезы остаточная неопределенность, т. е. условная энтропия, определяется вероятностями ошибочных решений

$$H(\{\Omega_i\}|\mathbf{Y})=-\sum_{i=1}^{M}\sum_{j=1,\ j<1}^{(M-1)/2} p_{ij}\log p_{ij}.$$

Количество информации, извлекаемой из изображения, определяется разностью безусловной энтропии ансамбля изображений и условной энтропии:

$$I(\{\Omega_i\},\ \mathbf{Z})=H(\{\Omega_i\})-H(\{\Omega_i\}|\mathbf{Y}). \tag{4.15}$$

Изменяя число K используемых признаков, можно с различной

степенью приблизиться к минимальной ошибке и количеству информации (4.15) (см. рис. 4.9). Во всех практических случаях безусловная энтропия ансамбля (4.11) существенно меньше не только безусловной кодовой энтропии (4.12) и (.413), но и условной энтропии (4.14). При любой величине ошибки классификации энтропия вектора признаков $H(\mathbf{Z}|\mathbf{Y})$ больше апостериорной энтропии решения. При этом чем ближе ошибка классификации к минимально возможной, тем сильнее это превышение (рис. 4.10).

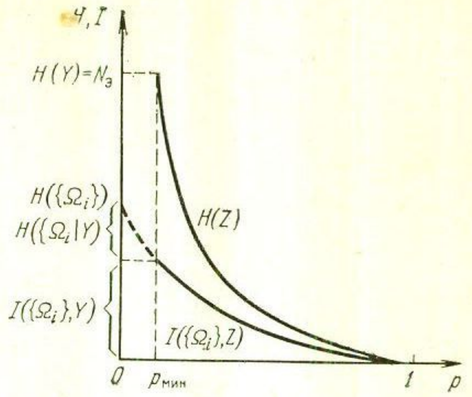


Рис. 4.10. Зависимость энтропии вектора признаков $H(Z)$ и полезной информации $I(\{\Omega_i\}, Z)$ от вероятности ошибки p

Последовательная переработка информации по схеме рис. 4.1 позволяет переходить от сигналов с большой энтропией к сигналам с меньшей энтропией. Это означает, что существует система неравенств

$$H(\mathbf{Y})>H'(\mathbf{Y})>H(\mathbf{Z}|\mathbf{Y})>H(\{\Omega_i\})>$$
$$>I(\{\Omega_i\}|\mathbf{Y})>I(\{\Omega_i\}|\mathbf{Z}). \tag{4.16}$$

Неравенства (4.16) являются важными ограничениями в классификации изображений. Они наглядно показывают, например, что не следует пытаться различать слишком большое число классов изображений, не связанное с числом $N_к$ контурных элементов сигналов объектов. Из неравенств (4.16) и формул (4.11) и (4.13) следует, что ни при какой системе признаков число надежно различаемых классов не превысит величины $M=2^{N_{об}}$. Это эквивалентно неравенству, связывающему количество информации (4.15) с числом контурных элементов образа [108]:

$$I(\{\Omega_i\}|\mathbf{Z})<N_к. \tag{4.17}$$

Отметим, что уменьшение масштаба сигнала объекта относительно растра с заданным числом элементов $N_э$ вызывает увеличение ошибок классификации в силу неинвариантности отклика дискретного фотоприемника к сдвигу изображения (см. раздел 4.3).

Выражение (4.17) является информационным ограничением возможностей классификации изображений, в определенной мере объясняющим эмпирическое правило Джонсона [110], находившего число элементов дискретного растра, необходимое для различения изображений. Информационное ограничение (4.17) позволяет также находить *компромисс в распределении сжатия изображений при накоплении и дискретизации в твердотельном фотоприемнике и при компьютерной переработке* (см. раздел 3.8). Чрезмерное сжатие сигнала в фотоприемнике сокращает возможности классификации изображений из-за возникающих ошибок неинвариантности к сдвигу (см. раздел 4.3), приводящих к сниже-

нию информативности системы признаков (4.14). Чрезмерное же увеличение числа элементов изображения при ограниченной решающей способности (сложности) видеокомпьютера ведет к уменьшению числа вычисляемых признаков и, следовательно, к увеличению ошибок классификации. Оптимальная четкость связана с отношением сложности видеокомпьютера к числу признаков, достаточному для практически достоверной классификации.

При постоянстве суммарного коэффициента сжатия информационно-решающей системы, равного отношению энтропии принимаемого потока фотонов к решающей способности системы, желательно распределять коэффициенты сжатия информации в телекамере и видеокомпьютере поровну (см. разделы 3.1 и 3.8). Целью такого согласованного распределения сжатия, особенно актуального в телевизионных роботах (см. разделы 5.3 и 5.4), является сокращение четкости на периферии поля зрения ПЗС-телекамер, рассмотренное в разделе 3.6.

Зависимость вероятности ошибки от характеристик сигнала и видеопроцессора проиллюстрируем на примере классификации простых бинарных равновероятных изображений, представленных на рис. 4.8,*а*. Для них вероятность минимально возможных ошибок будет определяться симметричной матрицей Δ_{ij}, элементами которой являются коэффициенты, показывающие, во сколько раз энергия разностного сигнала E_{ij} превосходит энергию сигнала изображения E_i:

$$E_{ij}=\Delta_{ij}E_i,$$

$$\Delta_{ij}=\begin{vmatrix} 0 & 1 & 0{,}5 & 1{,}1 & 0{,}4 & 0{,}7 & 0{,}7 \\ 1 & 0 & 0{,}8 & 0{,}6 & 1{,}1 & 0{,}5 & 0{,}6 \\ 0{,}5 & 0{,}8 & 0 & 0{,}8 & 1{,}1 & 0{,}6 & 0{,}5 \\ 1{,}1 & 0{,}6 & 0{,}8 & 0 & 1 & 0{,}7 & 0{,}6 \\ 0{,}4 & 1{,}1 & 1{,}1 & 1 & 0 & 0{,}9 & 0{,}9 \\ 0{,}7 & 0{,}5 & 0{,}6 & 0{,}7 & 0{,}9 & 0 & 0{,}2 \\ 0{,}7 & 0{,}6 & 0{,}5 & 0{,}6 & 0{,}9 & 0{,}2 & 0 \end{vmatrix}.$$

В качестве признаков использовались отношение квадрата периметра к площади, отношение максимальной и минимальной хорд, проходящих через центр тяжести фигуры, отношение квадрата расстояния от центра тяжести до ближайшей точки фигуры к площади и т. п. Для этих изображений и признаков распределения сигналов близки к гауссовским, а ковариационные матрицы — к симметричным. Отклонения от симметрии матриц незначительны и

Таблица 4.1

Признак	Класс 1		2		3		4		5		6	
	$\mathcal{M}$	σ^2	$\mathcal{M}$	σ^2	$\mathcal{M}$	σ^2	$\mathcal{M}$	σ^2	$\mathcal{M}$	σ^2	$\mathcal{M}$	σ^2
1	30	1,06	30	1,08	35	1,4	26	0,7	12,4	0,9	15	1
2	1,07	0,03	1,65	0,05	1,04	0,04	1,09	0,04	1,07	0,1	1,07	0,11
3	13	0,3	12,6	0,4	14,4	0,4	15	0,5	13	1	17	1
4	2	0,08	2	0,06	2,45	0,09	2,18	0,08	1,3	0,08	1,8	0,11
5	0,1	0,03	0		0,1	0,04	0		0		0	

распределение сигналов можно характеризовать их средним значением и дисперсией, приведенными в табл. 4.1.

При наличии входного шума (см. рис. 4.8,*б*) дисперсии признаков увеличиваются, однако при не слишком малых отношениях сигнал-шум целесообразно применять простой линейный классификатор (4.4) в предположении о равенстве ковариационных матриц Σ_i [76]. В ходе эксперимента линейный классификатор последовательно анализировал все предъявленные изображения, отмечая центр тяжести фигуры, относимой им к заданному классу (рис. 4.8,*в*). Естественно, что признаки не эквивалентны по вкладу в устранение априорной энтропии и могут быть ранжированы. Это видно по различиям в отношениях $\mathcal{M}_i/\sigma^2_i$ в табл. 4.1. Поэтому при каждом числе K признаков имеется разброс вероятности ошибки. Этот разброс увеличивается с возрастанием уровня шума на изображении и с уменьшением числа признаков (рис. 4.11). В ходе разработки классификатора стремятся использовать независимые признаки, добавление которых уменьшает ошибку. Однако синтез независимых признаков не формализован и добавление признаков может и не уменьшить ошибку. Более того, на практике часто наблюдается, что добавление новых признаков ухудшает, а не улучшает качество работы [76]. Для иных групп классов изображений и признаков, чем использованных на рис. 4.11, можно проиллюстрировать указанную аномалию зависимостью ошибки классификации от числа признаков (рис. 4.12).

Вместе с тем изобретение новых и отбор старых признаков, исключающих возрастание ошибки при добавлении новых признаков, позволяют экспериментально подтвердить зависимости на рис. 4.9, полученные из теоретических соображений. Экспериментальные результаты показывают, что при малом числе классов $M=7$ малое число используемых признаков $K=3$—4 позволяет получать почти потенциальное количество информации $H(\{\Omega_i\})$ (рис. 4.13).

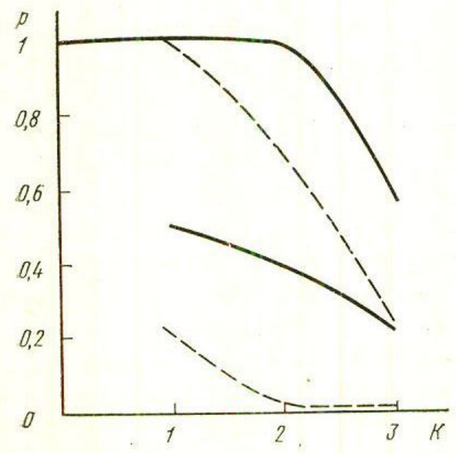


Рис. 4.11. Экспериментальные зависимости ошибки классификации от числа признаков для энергетического отношения сигнал-шум $\psi=20$ (сплошные линии) и $\psi=30$ (штриховые линии). Даны границы разброса при вариации признаков

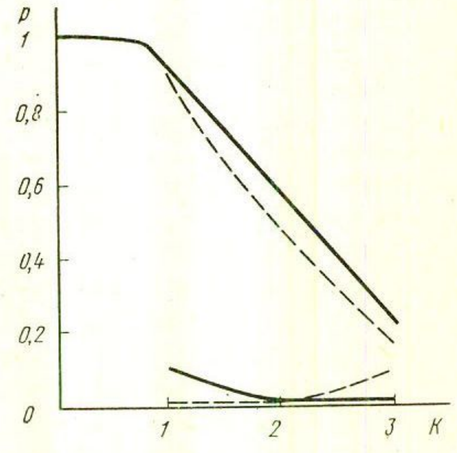


Рис. 4.12. Экспериментальные зависимости ошибки классификации от числа признаков при их независимости (сплошные линии) и наличии зависимости (штриховые линии). Даны границы разброса при вариации признаков

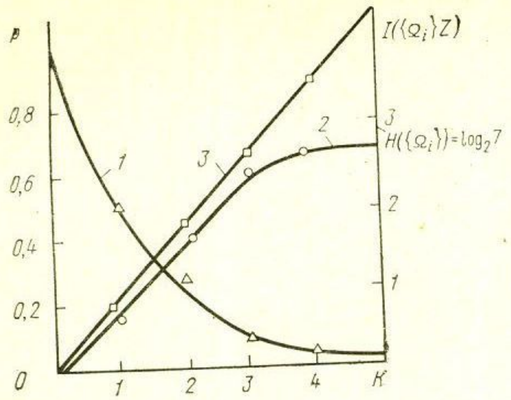
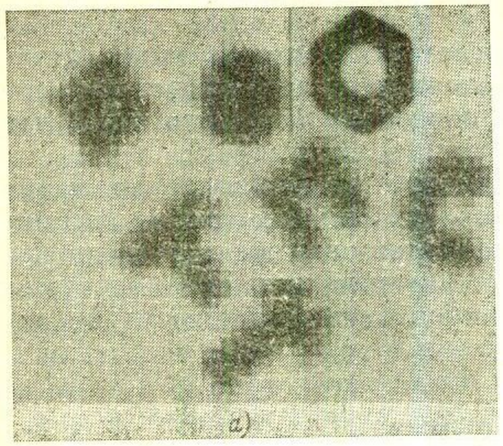


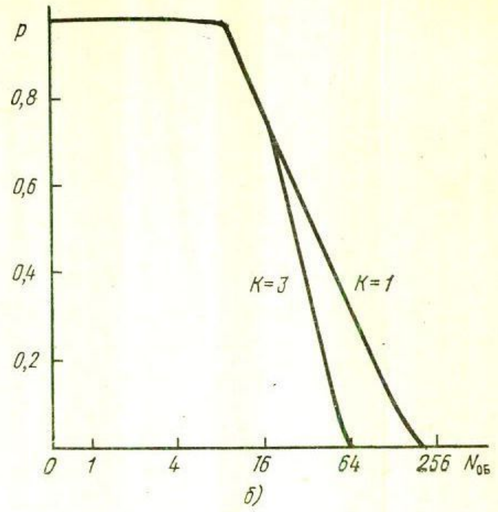
Рис. 4.13. Экспериментальные зависимости усредненной ошибки p (*1*), полезной информации $I(\{\Omega_i\}, Z)$ (*2*) и вычислительной сложности $C_в$ (*3*) классификатора от числа признаков

При достаточно большом числе элементов $N_{об}$ объекта на растре, существенно превышающем число гипотез, и низком уровне шума можно достичь безошибочной классификации. Однако $N_{об}$, т. е. масштаб изображения объекта, может уменьшиться при движении из-за удаления или попадания в зону пониженной четкости адаптивной ПЗС-камеры (рис. 4.14,*а*, см. 3.3). В этом случае ошибки классификации возрастают (рис. 4.14,*б*) и количество извлекаемой из изображения информации $I(\{\Omega_i\}, Z)$ снижается (рис. 4.14,*в*), в любых случаях оставаясь меньше величины, определяемой неравенством (4.17) [108].

В силу неравенств (4.16) от совокупности измеренных признаков требуется, чтобы она обладала энтропией не меньшей, чем принимаемое решение. Это ограничение не слишком жесткое и при небольшом числе классов позволяет ограничиться сравнительно малым числом признаков.

Дальнейшее увеличение числа K признаков несколько уменьшает ошибку классификации, которая ни при каком числе признаков не становится меньше минимального значения $p_{мин}>0$. Вместе с этим, все более медленным, продвижением к достоверной классификации приходится все больше усложнять видеокомпьютер. При этом практически линейно с ростом числа признаков увеличиваются как алгоритмическая, так и вычислительная сложности. Алгоритмическая сложность определяется минимальной длиной программы, по которой ведется классификация, а вычислительная определяется произведением числа шагов выполнения программы и емкости памяти, необходимой для вычислений [75]. Сложность видеокомпьютера более полно отражает вычислительная сложность: например, декоррелирующие фильтры с межэлементным и межкадровым вычитанием имеют одинаковую алгоритмическую сложность, но отличающуюся в число элементов разложения $N_э$ раз вычислительную сложность. Экспериментальная зависимость вычислительной сложности $C_в$ от числа признаков приведена на рис. 4.13 [108]. Эта ситуация убывания эффективности использования вычислительных ресурсов с приближением к потенциальной ошибке характерна для всех информационных систем — как каналов связи (см. гл. 2), так и устройств оценки параметров (см. раздел 4.3). При наличии ограничения какого-либо вида такое снижение эффективности вызывает необходимость вырождения, т. е. усечения числа членов ряда. При ограничениях на кодер усекается его спектральная характеристика (передаточный фильтр, фильтр Костаса, см. гл. 2), при ограничении сложности классификатора усекается число вычисляемых признаков.





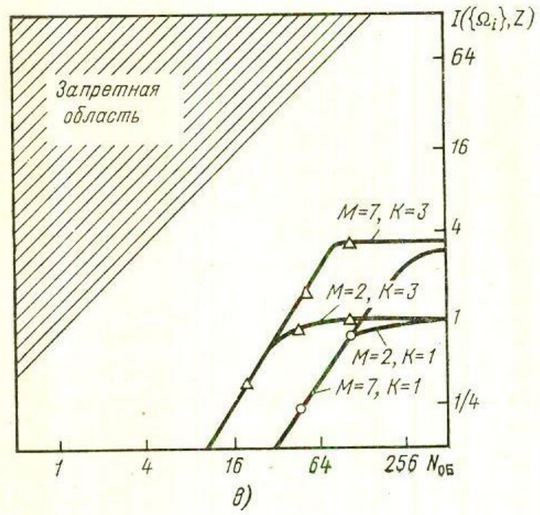


Рис. 4.14. Изображение классифицируемых образов при различном числе элементов разложения на образ (*а*), зависимость ошибки классификации (*б*) и количества информации (*б*) от числа элементов разложения на образ

Итак, идеальная структура различения гауссовских сигналов использует совокупность оптимальных линейных различающих и передаточных фильтров. Для обеих рассмотренных моделей сигналов получение полезной информации при классификации ограничивается следующими факторами:

энергией разностного сигнала для классифицируемых изображений известной формы при их наилучшем совмещении и разностью энергетических спектров для случайных сигналов;

неустранимыми искажениями, вносимыми шумом и ошибками компенсации фона;

энтропией классифицируемых изображений, определяемой, в первую очередь, числом контурных элементов образа;

вычислительной сложностью классификатора, его емкостью памяти и быстродействием.

При линейном различении гауссовских сигналов вычислительная сложность связана с суммой квадратов коэффициентов передаточного фильтра, оптимальным образом усекающего число членов ряда (признаков).

Два последних фактора определяют информационные характеристики сигнала и классификатора. Их роль подробно рассмотрена ниже — значение и пути экономии сложности видеопроцессора в разделе 4.5, а проявление ограниченности числа элементов изображения объектов в виде ошибок неинвариантности к сдвигу — в следующем разделе.

4.3. НЕИНВАРИАНТНОСТЬ К СДВИГУ

В процессе переработки светового сигнала в решение для достижения минимума ошибок классификации необходимо достичь минимума среднеквадратической ошибки при оценке сигналов. Наблюдение подвижных изображений предусматривает анализ группы кадров, в которых сигналы изображений могут быть искажены по-разному для его различных положений. Это обуславливает потребность использования для оценки качества работы видеопроцессора не только среднеквадратической ошибки оценки сигнала в данном кадре, сопоставляющей вход системы с ее выходом, но и среднеквадратической оценки разброса сигналов на выходе при движении объектов, т. е. сопоставляющей *выход системы с ее выходом* при вариации положения сигнала на входе.

Практическая переработка информации при камерном кодировании ведется с усечением числа членов ряда, в который разлагается полезный сигнал. Важно, что при таком усечении теряется обратимость процедуры разложения в ряд и, как следствие, теряется инвариантность сдвигу изображения (рис. 4.15). Характер проявления неинвариантности к сдвигу различен при использовании различных сигналов и рядов. Так, если использовать в качестве сигнала гармоническое колебание (типичный тестовый сигнал при измерениях передаточных характеристик), характеризуемое ограниченным спектром, и проектировать его на конечное число неперекрывающихся базисных функций и в соответствии с разложением по Котельникову восстановить непрерывный сигнал с



Рис. 4.15. Изображение на дискретном растре телекамеры на ПЗС тестового сигнала (слова BASIC) при различной фазировке относительно растра при интерполяции нулевого порядка

помощью интерполяции функциями $(\sin x)/x$, то неинвариантность к сдвигу проявится в виде искажений сигнала, сконцентрированных на краях интервала [58]. Пренебрежение краевыми эффектами в сочетании с использованием сигналов (при измерениях) с ограниченным спектром и интерполирующих функций, каждая из которых имеет ненулевое значение почти на всем изображении, позволяет в некоторых видеоинформационных системах не учитывать ошибки неинвариантности к сдвигу. Но это не означает, что и сама проблема борьбы с ошибками неинвариантности к сдвигу отсутствует.

Практика видеоинформатики показывает, что все реальные сигналы далеки от гармонических и имеют неограниченный спектр, а восстановление изображения по отсчетам с помощью интерполирующих функций $(\sin x)/x$, каждая из которых задана на всем изображении, требует огромных *вычислительных затрат* и не может быть осуществлена в телевидении в *реальном масштабе времени*. Борьба за *экономичность* и *быстродействие* привели к использованию в твердотельном телевидении простых интерполирующих функций, например треугольного вида.

Важность задачи оценивания привела к широкому распространению критериев качества оценивания, основанных на разности сигнала и его оценки, особенно среднеквадратического критерия. При этом под ошибкой ε понимается усредненная за интервал T мощность разности выходного сигнала оператора $L\{\cdot\}$ и оцениваемого сигнала $X(t)$:

$$\varepsilon = \frac{1}{T}\int_0^T [X(t) - L\{X(t)\}]^2\, dt. \qquad (4.18)$$

С помощью этого критерия минимума среднеквадратической ошибки (4.18) решаются многие задачи оценивания случайных процессов и полей в оптике, радиолокации и телевидении. В частности, результатом такого статистического синтеза являются линейные фильтры Винера, Калмана, Костаса, а также линейный кодек, синтезированный в гл. 2.

Сопоставление выхода оператора $L\{\cdot\}$ и его входа для непрерывных операторов дает исчерпывающую характеристику преобразования сигнала $X(t)$ в сигнал $Z(t)=L\{X(t)\}$. Такие непрерывные операторы являются инвариантными к сдвигу, т. е. отвечают условию [14]:

$$\text{если } Z(t)=L\{X(t)\}, \text{ то } L\{X(t-\Theta)\}=Z(t-\Theta). \qquad (4.19)$$

Импульсная характеристика $h(t, \Theta)$ инвариантных к сдвигу операторов зависит только от разности аргументов $t-\Theta$. Это условие является необходимым и достаточным, т. е. инвариантный во времени оператор формирует выходной сигнал $Z(t)$, являющийся сверткой входного сигнала с импульсной характеристикой $h(t-\Theta)$. Для инвариантных к сдвигу операторов импульсную характеристику часто обозначают как $h(t)$, подразумевая, что сигнал подается в момент $\Theta=0$.

Подчеркнем, что и дискретные операторы при дискретном воздействии являются инвариантными к сдвигу; для них справедливо условие (4.19) с той лишь оговоркой, что аргументы t и Θ принимают только целочисленные значения. Именно в этом смысле дискретная сумма свертки дискретных последовательностей эквивалентна непрерывному интегралу свертки непрерывных функций.

Для оценки близости дискретной системы к непрерывной может быть полезна мера неинвариантности к сдвигу — среднеквадратическая ошибка, определяемая разностью сигналов на выходе оператора при вариации времени воздействия одного и того же сигнала на входе. При этом возможно определение двух видов ошибок неинвариантности к сдвигу [77] — дифференциальной, зависящей от сдвига Θ.

$$\nu(\Theta) = \frac{1}{T}\int_0^T [L\{X(t)\} - \delta(t+\Theta)\, L\{X(t-\Theta)\}]^2\, dt \qquad (4.20)$$

и интегральной

$$\nu = \frac{1}{T^2}\int_0^T\int [L\{X(t)\} - \delta(t+\Theta)\, L\{X(t-\Theta)\}]^2\, dt d\Theta, \qquad (4.21)$$

где $\delta(t+\Theta)$ — оператор сдвига на интервал Θ.

Естественно, из условия (4.19) следует, что для операторов, инвариантных к сдвигу, эти ошибки равны нулю. Для неинвариантных к сдвигу операторов существует широкий диапазон соотношений среднеквадратической ошибки неинвариантности ν (4.21) и среднеквадратической ошибки оценивания ε (4.18). При рассмотрении дискретизации сигналов часто считают эквивалентную ширину импульсной характеристики дискретизатора равной шагу дискретизации. Вместе с тем возможны и иные соотношения между ними. В частности, при реализации чересстрочного разложения изображения в матричных ПЗС с кадровым переносом эквивалентный шаг дискретизации оказывается меньшим размера накопительной ячейки, в матрицах ПЗС со строчно-кадровым переносом зона накопления вдвое меньше элемента дискретизации изображения. Экспериментально установлено также [67], что выборка на элементе дискретизации «мгновенного» или максимального значения приводит к существенно большей ошибке неинвариантности к сдвигу при наблюдении объектов, размеры которых соизмеримы с размером элемента разложения. Оценить диапазон изменения ошибки неинвариантности к сдвигу (4.21) можно, рассмотрев усечение до конечного числа членов разложения случайного сигнала $X(t)$ с ограниченной энергией по ортонормированному базису $\{\varphi_i(t)\}$:

$$L\{X(t)\} = \sum_{i=1}^{Q} \varphi_i(t)\, X_i = \sum_{i=1}^{Q} \varphi_i(t) \int_{-\infty}^{\infty} \varphi_i(t)\, X(t)\, dt.$$

Среднеквадратическая ошибка (4.18) представления сигнала $X(t)$ оператором $L\{X(t)\}$ в этом случае является энергией усеченного «хвоста» спектра сигнала и составит

$$\varepsilon = \sum_{i=Q+1}^{\infty} X_i.$$

Введя в подынтегральное выражение (4.21) дважды с различными знаками $X(t)$ и перегруппировав члены, получим:

$$\nu = \frac{1}{T^2}\int_0^T\int [(X(t) - L\{X(t)\}) - (X(t) - \delta(t+\Theta)\, L\{X(t-\Theta)\})]^2\, dt d\Theta.$$

Введем коэффициент корреляции ошибки (4.18) ε:

$$\rho_\varepsilon = \frac{\frac{1}{T^2}\int_0^T\int [X(t) - L\{X(t)\}]\,[X(t) - \delta(t+\Theta)\, L\{X(t-\Theta)\}]\, dt d\Theta}{\frac{1}{T}\int_0^T [X(t) - L\{X(t)\}]^2\, dt}.$$

Тогда среднеквадратическая ошибка (4.21) неинвариантности к сдвигу будет удовлетворять условию

$$\nu = 2\varepsilon(1-\rho_\varepsilon),$$

ограничивающему диапазон ее изменения и связывающему ее с среднеквадратической ошибкой (4.18) усечения энергетического спектра сигнала.

В качестве примера вычисления ошибок неинвариантности к сдвигу рассмотрим твердотельную телевизионную камеру, осуществляющую накопление сигнала изображения на массиве $Q \times Q$ квадратных элементов с ребром Δ. Она может быть описана оператором, одномерный вариант которого имеет вид [2]:

$$L\{X(t)\} = \sum_{i=1}^{Q} \varphi(t - i\Delta) \int_{-\infty}^{\infty} \frac{1}{\Delta}\, \mathrm{rect}\, \frac{t - i\Delta}{\Delta}\, X(t)\, dt, \qquad (4.22)$$

где φ — интерполирующая функция.

В качестве входного воздействия выберем наиболее быстро меняющийся при фиксированных дисперсии λ и интервале корреляции τ сигнал с экспоненциальной корреляционной функцией («телеграфный» сигнал):

$$K(t) = \lambda \exp(-|t|/\tau). \qquad (4.23)$$

Для этого сигнала при достаточно больших отношениях сигнал-шум оптимальная интерполирующая функция аппроксимируется линейной зависимостью [77]:

$$\varphi(t) = \max\{(1-|t|/\Delta),\ 0\}. \qquad (4.24)$$

Подстановка выражений (4.22)—(4.24) в (4.20) позволяет найти дифференциальную ошибку неинвариантности к сдвигу:

$$\nu(\Theta) = \frac{\lambda\Delta}{3\tau} \min\left\{\left[\frac{\Theta}{\Delta} - \mathrm{ent}\,\frac{\Theta}{\Delta}\right], \left[1 - \frac{\Theta}{\Delta} + \mathrm{ent}\,\frac{\Theta}{\Delta}\right]\right\}. \qquad (4.25)$$

Эта формула отражает зависимость ошибки неинвариантности к сдвигу не только от интервалов дискретизации и корреляции, но и от сдвига Θ. Ошибки неинвариантности к сдвигу сигнала возникают не только на краях интервала задания сигнала. Формула (4.25) показывает существование и *периодическую зависимость ошибки неинвариантности к сдвигу на всем интервале задания сигнала.* Заметим, что формула (4.25) иллюстрирует инвариантность оператора дискретизации (4.22) к целочисленному сдвигу (рис. 4.16) аналогично свойствам модуляции [3, с. 294].

Периодичность среднеквадратической ошибки оценивания (4.18) и среднеквадратической ошибки неинвариантности к сдвигу (4.20) наглядно проявляется в системах телевизионной астрономии при условии соизмеримости размеров импульсной характеристики оптической системы (аппаратной функции или «кружка рассеяния» объектива) и элементарной накопительной ячейки. Известно, что сигнал на выходе дискретного фотоприемника может меняться в зависимости от его фазы относительно дискретной структуры фотоприемника от

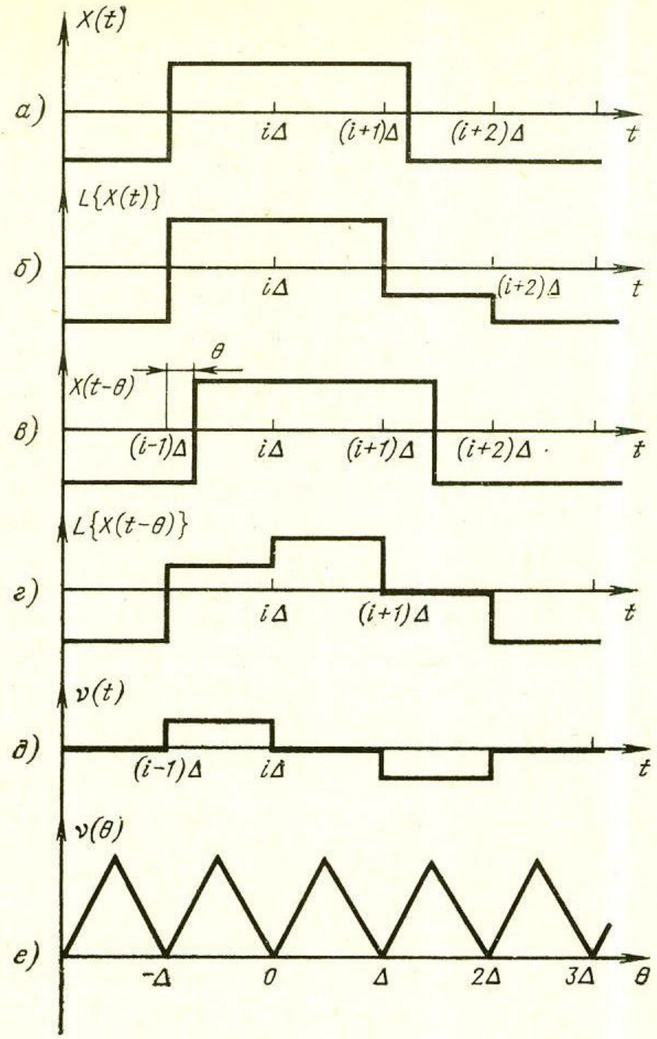


Рис. 4.16. Определение ошибки неинвариантности к сдвигу при ступенчатой интерполяции сигнала: *а*) исходный телевизионный сигнал; *б*) сигнал на выходе; *в*) смещенный исходный сигнал; *г*) сигнал на выходе при смещенном входном сигнале; *д*) разностный сигнал; *е*) зависимость разностного сигнала от смещения

номинального значения до уменьшенного вдвое при попадании на границу раздела двух элементов фотоприемной матрицы и вчетверо при попадании на границу раздела четырех элементов матрицы.

Подставив выражения (4.22)—(4.24) в формулу (4.21), найдем интегральную ошибку неинвариантности к сдвигу:

$$\nu = \frac{\lambda \Delta}{12\tau}. \qquad (4.26)$$

Для данного примера интегральная ошибка оказывается численно равной ошибке (4.18) при подстановке в нее выражений (4.22)—(4.24). Это позволяет считать, что для оператора дискретизации (4.22) линейная ошибка (4.18) явля-

ется оценкой сверху для ошибки (4.21). Хотя дисперсия ошибки неинвариантности к сдвигу (4.26) определяется лишь видом оператора и спектральными свойствами сигнала и не зависит от функции распределения сигнала, функция распределения ошибки неинвариантности к сдвигу определяется функцией распределения фильтруемого сигнала. Так, в рассмотренном примере оценивания смещающегося «телеграфного» сигнала, имеющего функцию распределения по амплитуде u $p_X(u) = \frac{1}{2}\delta(u - |\sqrt{\lambda}|)$, функция распределения ошибки неинвариантности к сдвигу имеет вид

$$p_\nu(u) = \frac{\tau - \Delta}{\tau}\delta(u) + \frac{\Delta}{\tau}\operatorname{rect}\frac{|u|}{\sqrt{\lambda}}.$$

В случае гауссовского оцениваемого сигнала ошибки неинвариантности, так же как и ошибка (4.18), гауссовские.

В отсутствии шума сдвиг непрерывного сигнала на выходе непрерывной системы может быть точно измерен. При этом взаимная информация между двумя выходными реализациями одного и того же непрерывного сигнала инвариантной к сдвигу системы (хотя и искажающей входной сигнал $X(t)$) равна бесконечности: $I\{L\{X(t)\},\ \delta(t+\Theta)L\{X(t-\Theta)\}\} = \infty$. В этих же идеализированных условиях отсутствия шума взаимная информация между двумя выходными реализациями одного и того же непрерывного сигнала при неинвариантности к сдвигу конечна. Для рассмотренного примера дискретного фотоприемника $Q \times Q$ элементов эта взаимная информация составит

$$I\{L\{X(t)\},\ \delta(t+\Theta)L\{X(t-\Theta)\}\} = \frac{Q^2}{2}\log\frac{\lambda(1-\rho^2)^2}{\nu},$$

где $\rho = \exp(-\Delta/\tau)$ — коэффициент корреляции соседних отсчетов входного незашумленного изображения.

Так как в реальных системах сигналы всегда наблюдаются в смеси с шумом (см. гл. 1, 2), то исходя из формулы (4.25) можно выбрать интервал дискретизации так, чтобы ошибка неинвариантности к сдвигу маскировалась шумом. Если же дискретизация «грубая», т. е. интервал Δ выбран не в результате оптимизации, а исходя из наличия имеющихся твердотельных фотоприемников и требуемых углов зрения, то ошибки неинвариантности к сдвигу становятся значительными. В этих условиях бывает целесообразно использовать специальные меры для уменьшения ошибок неинвариантности к сдвигу. К таким мерам относится обмен временной и пространственной четкости видеоинформационной системы.

При различении подвижных изображений часто встречается ситуация, когда оценивание пространственно-временного процесса ведется в различных условиях дискретизации по полю и во времени. Наиболее часто не хватает пространственной четкости (разрешающей способности). В твердотельных телевизионных камерах, обладающих высокой светочувствительностью и соизмеримым с кружком рассеяния объектива размером накопительного элемента (элемента разложения), четкость ограничивается не шумом, а пространственной дискретизацией. При этом высшая воспроизводимая пространственная частота определяется интервалом дискретизации изображения по соответствующему аргументу (по вертикали или по горизонтали): $F_в = 1/2\Delta_в$; $F_г = 1/2\Delta_г$ (рис. 4.17,*а*). Вместе с тем известно, что операция сканирования изображения с помощью

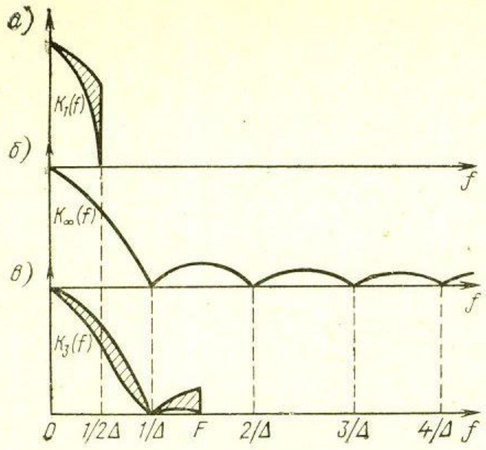


Рис. 4.17. Передаточные функции прямоугольного накопительного «окна» без сканирования (*а*), при непрерывном сканировании (*б*), при дискретном сканировании при $m=3$ (*в*). (Заштрихована область неинвариантности к сдвигу). [78]

некоторого «окна», в роли которого в нашем случае выступает накопительный элемент твердотельного фотоприемника (ячейка ПЗС) с линейным размером Δ, дает возможность получения отклика на любых частотах F [14]. Интенсивность спектра сигнала, полученного путем сканирования, хотя и убывает с ростом частоты, но не ограничена величиной $F<\infty$ (рис. 4.17,*б*). Такая картина получается в предположении, что наблюдатель может различать весь континуум значений отклика, аппаратной функции «окна» и, что особенно важно в нашем рассмотрении, значений сдвига накапливающего «окна» при сканировании.

Если сканирование накопительного элемента осуществлять дискретно, то выявятся два существенных момента. Во-первых, для формирования полного отклика потребуется уже не континуальное число «элементарных» откликов, а их счетное множество. Если считать область задания исследуемого сигнала конечной, но и число m дискретных положений «окна» будет конечным. Для многоэлементных фотоприемников такое сканирование достаточно осуществлять на интервале элемента разложения (размера «окна»). Во-вторых, форма частотной характеристики с прозрачностью «окна» $f(t)$ и последовательным смещением его на m положений внутри него будет определяться преобразованием Фурье от $f(t)$ вплоть до верхней частоты $F=m/2\Delta$ (рис. 4.17,*в*). Последнее положение позволяет применить метод сканирования фотоприемника для повышения пространственного разрешения [82].

Для рассмотренного примера «телеграфного» сигнала и прямоугольного накопительного «окна» ошибка неинвариантности к сдвигу (4.21) при росте коэффициента перекрытия m убывает от значения (4.26), асимптотически стремясь к нулю. Среднеквадратическая ошибка (4.18) при этом убывает от значения (4.26) до половины этого значения [77]:

$$\nu=\frac{\lambda\Delta}{12\tau m},\quad \varepsilon=\frac{\lambda\Delta}{12\tau}\left(\frac{m+1}{2m}\right). \tag{4.27}$$

Именно благодаря уменьшению среднеквадратической ошибки (4.27) с увеличением коэффициента перекрытия накопительных элементов метод сдвига фотоприемника применяется для повышения пространственной четкости [82].

Вследствие (4.27) *с помощью введенной меры ошибки неинвариантности к сдвигу* (4.21) *часто можно оценить среднеквадратическую ошибку* (4.18) *оценивания оптического сигнала электрическим кодом на выходе телекамеры на ПЗС несмотря на различие физических носителей сигнала* $X(t)$ *и кода* $Z(t)$ [118].

При различении подвижных изображений использование перекрывающихся накопительных элементов матричного фотоприемника позволяет повысить пространственную четкость за счет использования информации о сигнале объекта из группы кадров, в которых положение изображения объекта относительно дискретного поля элементов различно. Для повышения пространственной четкости в m раз потребуется увеличение сложности устройства переработки информации и емкости его памяти на m кадровых ОЗУ. Очевидно, что при этом происходит обмен четкости по времени на четкость по полю. Вместе с тем эффективность использования ОЗУ ухудшается с ростом числа запоминаемых кадров, так как в соответствии с формулой (4.27) ошибки убывают все медленнее.

Можно считать, что информационный КПД использования памяти определяется отношением количества информации о световом изображении объекта в синтезированном по сигналу из m ОЗУ кадре (коде $Z(t)$) к емкости памяти ОЗУ $I_{\text{ОЗУ}}=mI_1$ (I_1 — емкость памяти ОЗУ на кадр). Для гауссовского сигнала эффективность использования ОЗУ составит

$$\eta=I\{X(t),\ Z(t)\}/I_{\text{ОЗУ}}=(\log\lambda/\varepsilon)/m.$$

Определение требуемого числа m кадровых ОЗУ для повышения пространственной четкости может производиться на основе требуемой ошибки оценивания сигнала объекта или ошибки неинвариантности к сдвигу. Назначаемая ошибка определяет ε-энтропию изображения и, следовательно, емкость памяти ОЗУ.

В упрощенном варианте устройств переработки видеоинформации, когда все m телевизионных кадров участвуют в синтезируемом изображении с одинаковыми весами, оптимальное число объединяемых кадров зависит от изменчивости формы сигнала во времени. В этом случае определяющую роль будет играть соотношение точности пространственной и временной дискретизации изображения фотоприемников. Если сигнал, имеющий интервал временной корреляции τ_t и пространственной корреляции τ_s, дискретизирован с интервалами Δ_t и Δ_s, то оптимальное число дополнительных кадровых ОЗУ, обеспечивающее минимум среднеквадратической ошибки (4.18) и (4.21), составит [80].

$$m\approx\sqrt{\frac{\tau_t\,\Delta_s}{\Delta_t\,\tau_s}}.$$

Эффективность использования памяти в таком упрощенном варианте ниже, чем при идеальном кодировании, использующем разложение Карунена — Лоэва. Кроме того, не могут быть достигнуты потенциальные значения среднеквадратической ошибки (4.18) и (4.21). Вместе с тем отход от оптимальности позволяет сократить алгоритмические затраты как по объему программы, так и по количеству вычислений за время элемента разложения (см. 4.5).

Наиболее эффективно использование перекрытия накопительных элементов, т. е. дополнительных кадровых ОЗУ в условиях грубой пространственной дискретизации, свойственной телевизионным роботам (см. гл. 5).

Уменьшение скреднеквадратической ошибки неинвариантности к сдвигу (4.21) достигается при любом способе увеличения отношения размера накопительного элемента к эквивалентному шагу дискретизации. Так, сдвиг фазы накопительных элементов от строки к строке, т. е. формирование «шахматной» структуры растра, уменьшает максимальное значение ошибки неинвариантно-

сти к сдвигу: малоразмерная деталь может в худшем случае уменьшиться уже не в четыре, а только в три раза. Важно, что введение перекрытия накопительных элементов уменьшает как среднеквадратическую ошибку оценивания (4.18) (что проявляется в подавлении ложных сигналов на изображении — муара), так и среднеквадратическую ошибку (4.21) неинвариантностик сдвигу — вариации муарового паразитного сигнала при изменениях фазировки изображения относительно дискретного растра твердотельного фотоприемника.

Введенные меры для дифференциальной (4.20) и интегральной (4.21) ошибок неинвариантности к сдвигу позволяют оценить близость непрерывно-дискретной системы (непрерывный вход, дискретный выход) к непрерывной. Строго говоря, из-за наличия ненулевой ошибки неинвариантности к сдвигу непрерывно-дискретная система не может быть эквивалентной непрерывной. Вместе с тем эквивалентность этих систем может иметь не абсолютный смысл, а с точностью до ошибок в работе системы переработки информации, вызываемых неизбежными шумами — как на входе системы, так и в канале связи. Шум проявляет себя трояко: он ограничивает, во-первых, точность измерения формы сигналов, во-вторых, точность измерения их сдвига друг относительно друга, в-третьих, загружает шумовой информацией канал связи (см. гл. 2). Два первых проявления влияния шума обуславливают допустимость не слишком больших (маскируемых шумом) ошибок неинвариантности к сдвигу. Третье проявление делает нецелесообразным чрезмерное подавление ошибок неинвариантности к сдвигу, достигаемое ценой снижения эффективности использования устройств памяти и каналов связи.

Таким образом, шум определяет предел необходимой степени инвариантности непрерывно-дискретной системы к сдвигу. Система, в которой ошибки неинвариантности к сдвигу маскируются шумом, может считаться практически эквивалентной непрерывной системе.

4.4. ИЗМЕРЕНИЕ ПАРАМЕТРОВ СИГНАЛОВ ИЗОБРАЖЕНИЙ

Во многих видеоинформационнах системах продуктом переработки информации — конечным или промежуточным — является измеренный параметр сигнала изображения. Согласно принятым определениям измерение — это процесс получения опытным путем численного отношения между измеряемой физической случайной величиной и некоторым ее значением, принятым за единицу сравнения [85].

Результат такого измерения называют оценкой и для нахождения способов оценивания и теоретических границ качества процедуры оценивания пользуются хорошо разработанным математическим аппаратом теории оценивания [8, 14, 84].

С точки зрения видеоинформатики в измерении параметров сигналов изображений существенны следующие вопросы: как сформировать изображение, чтобы получить максимум нужной информации; как переработать изображение в идеальную оценку измеряемой величины и в приемлемую, но полученную с разумными вычислительными затратами; как представить результат измерений для лучшей передачи получателю через канал связи; каково соотношение количества доставленной получателю полез-

ной информации к содержавшейся в исходном оптическом изображении и к объему кода, потребовавшегося для передачи. Качество оценки измеряемого параметра — вопрос комплексный, включающий в себя количественные ответы на все поставленные вопросы. Теория оценок позволяет дать ответ на вопрос о том, как, располагая априорными данными о функции распределения измеряемой случайной величины (т. е. ее моделью) и некоторыми критериями качества оценки, получить: а) идеальный алгоритм переработки информации, содержащейся в наблюдениях и б) границу достижимой точности измерения. Теория информации и теория алгоритмов позволяют дать ответ на вопросы о способе оценки при сокращении вычислительных затрат и оценке эффективности передачи извлеченной информации получателю через канал связи.

К оценкам предъявляется ряд требований, нацеленных как на извлечение полезной информации, так и на ее передачу. Так, оценка должна быть минимальной достаточной статистикой — это условие обеспечивает минимальную энтропию сообщения и, следовательно, требуемую пропускную способность для его передачи. С другой стороны, идеальная оценка должна быть несмещенной, состоятельной и эффективной. Несмещенная оценка имеет математическое ожидание, равное измеряемой случайной величине. Состоятельная оценка в отсутствии помех (неограниченно возрастающем отношении сигнал-шум) совпадает с истинным значением измеренного параметра. Эффективная оценка — это такая несмещенная и состоятельная оценка, которая характеризуется наименьшей возможной для имеющегося набора априорной информации дисперсией. Необходимым и достаточным условием эффективности оценки является достижение ее дисперсией потенциально малого значения, определяемого количеством информации по Фишеру об оцениваемом параметре. Эта величина опирается на функцию правдоподобия, определяемую как плотность вероятности $p(Y|X)$ принятого наблюдения $Y=X+\xi$ при истинном значении X измеряемого параметра. Неравенство, утверждающее, что никакая оценка $\tilde{X}$ не может иметь дисперсии, меньшей, чем величина, обратная количеству информации по Фишеру, называют неравенством Рао — Крамера [8, т. 1]:

$$\mathbf{M}\{(X-\tilde{X})^2\} \geqslant 1/\mathbf{M}\left\{\left[\frac{\partial \ln p(Y|X)}{\partial X}\right]^2\right\} = -1/\mathbf{M}\left\{\frac{\partial^2 \ln p(Y|X)}{\partial X^2}\right\}.$$

Эффективная оценка $\tilde{X}$, независимо от способа получения, должна удовлетворять этому условию со знаком равенства.

Примером границы Рао — Крамера может быть граница, связывающая дисперсию ошибки σ^2_τ идеальной оценки временного положения τ видеоимпульса известной формы с энергетическим отношением сигнал-шум

$$\psi^2 = \frac{E}{\sigma_1^2}, \quad \left(E = \int_{-\infty}^{\infty} |X(t)|^2 dt\right)$$

и среднеквадратической полосой частот $\bar{\omega}^2$ спектра Фурье $S(\omega)$ сигнала $X(t)$

$$\bar{\omega}^2 = \int_{-\infty}^{\infty} \omega^2 |S(\omega)|^2 d\omega \Big/ \left[\int_{-\infty}^{\infty} |S(\omega)|^2 d\omega \right].$$

Измерение временного положения импульсов, широко применяемое в импульсной радиолокации, актуально во многих видеоинформационных системах. При этом следует иметь в виду, что измерение координат объектов в плоскости изображения — это процедура двукратного измерения величины, аналогичной временному положению видеоимпульса.

Для этих условий граница точности определяется формулой Вудворда [9]:

$$\varepsilon_W = \mathbf{M}\{(\tau - \hat{\tau})^2\} = 1/\psi^2 \bar{\omega}^2. \qquad (4.28)$$

Воспользовавшись равенством Парсеваля

$$\int_{-\infty}^{\infty} |X(t)|^2 dt = \int_{-\infty}^{\infty} |S(\omega)|^2 d\omega.$$

формулу Вудворда можно привести к виду

$$\varepsilon_W = \mathbf{M}\{(\tau - \hat{\tau})^2\} = \sigma^2_1 \Big/ \int_{-\infty}^{\infty} |S(\omega)|^2 \omega^2 d\omega. \qquad (4.29)$$

Эта запись формулы Вудворда наглядно показывает, что вклад в точность измерения растет с ростом частоты спектральных компонент полезного сигнала. Обратно пропорциональная зависимость ошибки оценивания от отношения сигнал-шум понятна, роль же среднеквадратической полосы частот достаточно специфична. Она сказывается не только в уменьшении ошибки измерения временного положения видеоимпульса при уменьшении его длительности. Оказывается, что увеличения точности измерения следует ожидать и при сохранении длительности видеоимпульса при изменении его формы. Так, для видеоимпульса с гауссовской формой и эквивалентной длительностью r

$$X(t-\tau) = \exp\{-(t-\tau)^2/2r^2\}.$$

минимальная ошибка измерения временного положения

$$\mathbf{M}\{(\tau - \hat{\tau})^2\} = 2r^2/\psi^2. \qquad (4.30)$$

Гауссовская модель видеоимпульса широко применяется в радиолокации [9] и телевизионной астрономии [87]. В робототехнике сигналы обрабатываемых изделий часто моделируют прямоугольными видеоимпульсами [88—90].

Для видеоимпульса прямоугольной формы, независимо от его длительности, интеграл в (4.29) расходится, что означает бесконечную точность измерения, несмотря на влияние шума. Наличие такого порадоксального решения связано с тем, что формула

Вудворда (4.28) является асимптотической и справедлива при достаточно больших отношениях сигнал-шум. Существование предела применимости формулы Вудворда видно и из связи среднеквадратической полосы частот $\bar{\omega}^2$ с производной $X'(t)$ сигнала $X(t)$ по времени

$$\bar{\omega}^2 = \int_{-\infty}^{\infty} [X'(t)]^2 dt \Big/ \left[\int_{-\infty}^{\infty} |X(t)|^2 dt \right].$$

Временной расчет среднеквадратической частоты сигнала позволяет сразу ввести запрет на применение рассмотренного подхода к сигналам, моделируемым функциями с разрывами непрерывности. Вместе с тем, такие модели хотя и приводят иногда к упрощению описания, но для видеоинформационных систем являются чрезмерной идеализацией. Даже если в предметной области наблюдаемый объект и имеет разрывы непрерывности, то уже на входе фотоприемника они сглажены оптической системой. Это сглаживание осуществляется как за счет считающихся принципиально устранимыми аббераций, так и за счет неустранимой дифракции света на зрачке объектива.

В информационных системах, преобразующих оптический сигнал в электрический код с помощью матричных фотоприемников, для достижения потенциальной точности измерения координат следует согласовывать размер наблюдаемого объекта с размером элемента разложения. При этом радиус r импульсной характеристики объектива должен быть несколько менее элемента разложения изображения в фотоприемнике [2]. Это положение справедливо для объектов различных размеров: как определяемых импульсной характеристикой объектива, так и существенно больших. В этом смысле формула (4.30) является универсальной.

Вместе с тем неравенство Рао — Крамера и формулы, получаемые из его частных случаев, определяют лишь потенциальную границу точности, к которой можно лишь в той или иной мере приближаться. Для этого приближения потребуется не только указанный подбор соотношения размеров элемента разложения и импульсной характеристики объектива, но и наращивание сложности алгоритма вычисления оценки. Покажем это на простом примере вычисления оценки координат $\hat{s}_ц$, $\hat{u}_ц$ центра тяжести сигнала $X(s, u)$ изображения. Оценка координат центра тяжести осуществляется по формулам

$$\hat{s}_ц = \frac{\int_{-\infty}^{\infty}\int X(s, u)\, s\, ds\, du}{\int_{-\infty}^{\infty}\int X(s, u)\, ds\, du}, \quad \hat{u}_ц = \frac{\int_{-\infty}^{\infty}\int X(s, u)\, u\, ds\, du}{\int_{-\infty}^{\infty}\int X(s, u)\, ds\, du}. \qquad (4.31)$$

Для сигналов с круговой симметрией оценки координат (4.31) являются эффективными и должны обеспечивать потенциальную точность измерения, определяемую формулой (4.30). Для дости-

жения этой границы потребовалась бы бесконечная точность представления данных в формуле (4.31). Однако некоторая приемлемая точность измерения координат может быть достигнута и при конечной точности представления данных или, что то же самое, при упрощенных алгоритмах, использующих усеченные данные об изображении. При этом существует обмен между точностью оценки $\tilde{s}_{\text{ц}}$, $\tilde{u}_{\text{ц}}$ и сложностью алгоритма, эквивалентный обмену точности на пропускную способность, рассмотренной в гл. 2. При этом кодовая ε-энтропия может рассматриваться как оценка сложности ансамбля сообщений. Тогда график зависимости сложности от допустимой ошибки оценивания качественно совпадает с зависимостью ε-энтропии сигнала от задаваемой среднеквадратической ошибки на рис. 2.1. Отличие заключается в способах вычисления особых точек графика — минимальной и максимальной ошибках. Это, соответственно, ошибка, вычисляемая по формуле Вудворда (4.20) и ошибка, определяемая величиной априорного интервала S. Эта максимальная ошибка $\varepsilon = S^2/12$ соответствует самому простому алгоритму — отсутствию вычислений и принятию в качестве оценки середины априорного интервала независимо от наблюдений. Минимальная же ошибка ε_W, являясь непрерывной величиной, имеет бесконечную энтропию, которая может служить оценкой сложности алгоритма вычисления оценки. Важным практическим случаем вычисления оценки координат центра тяжести с сокращенными вычислительными затратами является алгоритм вычисления по бинарно-квантованному дискретному сигналу X_{ij} твердотельного фотоприемника, равному в i-м элементе j-й строки либо 0, либо 1. В этом простом случае оценки координат вычисляют по формулам:

$$\tilde{s}_{\text{ц}} = \frac{\sum_i \sum_j X_{ij}\, s}{\sum_i \sum_j X_{ij}}, \quad \tilde{u}_{\text{ц}} = \frac{\sum_i \sum_j X_{ij}\, u}{\sum_i \sum_j X_{ij}}. \qquad (4.32$$

Широкое использование упрощенного алгоритма (4.32) на практике, особенно в робототехнике, обусловлено тем, что он дает хороший компромисс между точностью и сложностью. Это происходит потому, что при достаточно больших отношениях сигнал-шум бинарное квантование, не подчеркивая влияния шума, лишь незначительно уменьшает энергию полезного сигнала. Так, для гауссовского яркостного сигнала дисперсия ошибки измерения центра тяжести $\{\tilde{s}_{\text{ц}}, \tilde{u}_{\text{ц}}\}$ при бинарном квантовании всего лишь менее чем вдвое превышает потенциальную границу, почти достижимую, например, при 10-разрядном кодировании сигнала и его переработке в соответствии с формулой (4.31). Сокращение же скорости кода на выходе видеопроцессора и сложности алгоритма вычисления оценки в 10 раз явно заманчиво. Это и обеспечивает простой алгоритм (4.32), характеризующийся максимальной эффективностью использования алгоритмических затрат.

Обменное соотношение между точностью и сложностью имеет универсальный характер. Для гауссовских сигналов при не слишком малых отношениях сигнал-шум и ошибках $\varepsilon < 2\varepsilon_B$ имеет место *соотношение неопределенности*, согласно которому кодовая информация $I(Y, Z)$, характеризующая сложность, связана с потерей полезной информации ΔI, характеризующей точность:

$$\Delta I_{\text{мин}} \cdot 2^{2I(Y,\,Z)} \geqslant 2^{2I(X,\,Y)}/\ln 4 = (\psi^2+1)/\ln 4. \qquad (4.33)$$

Обмен точности на сложность сопровождается обменом потери полезной информации на шумовую информацию в соответствии с *информационным соотношением неопределенности* (см. разделы 2.4 и 3.8):

$$\Delta I_{\text{мин}} \cdot 2^{2I_{\text{ш}}} \geqslant 1/\ln 4. \qquad (4.34)$$

При произвольных способах измерения (кодирования) в (4.33) и (4.34) выполняется неравенство, равенство же достигается при идеальном кодировании (см. разделы 2.3—2.5). Поэтому соотношения неопределенности (4.33) и (4.34) позволяют сравнивать различные эвристические алгоритмы оценки параметров, дающие различные приближения к потенциальной ошибке, определяемой неравенством Рао — Крамера. Критерием качества алгоритма должно служить отклонение фактического значения произведения информационных характеристик от задаваемого соотношениями неопределенности (4.33) и (4.34).

Обменное соотношение между потерей полезной информации и шумовой информацией, выражаемое информационным соотношением неопределенности (4.34), связано с критерием минимума суммарной информационной ошибки (3.20). Аналогично обменное соотношение между потерей полезной информации и кодовой информацией (сложностью) (4.33) порождает критерий минимума суммарных информационных потерь:

$$L = \min_{\varepsilon} \{c_1 \Delta I_{\text{мин}} + c_2 I(Y, Z)\}. \qquad (4.35)$$

В отличие от критерия (3.20), позволяющего найти оптимум при произвольном соотношении c_1 и c_2, минимум в (4.35) существует при $c_1/c_2 > 1$, т. е. при большей стоимости полезной информации по сравнению со сложностью.

Неоправданный рост сложности видеокомпьютера, загружаемого переработкой шумовой информации и проверкой ложных гипотез, ведет к *концепции достаточности точности* при переработке изображений в реальном времени.

Измерение параметров изображений хотя и является часто промежуточной операцией, но используется многократно. В качестве примера укажем вычисление признаков при классификации (см. раздел 4.2) и измерение смещения изображений, необходимое во многих приложениях — в многоракурсных и многоспектральных системах, а также в качестве предварительной процедуры при матричной фильтрации подвижных изображений (см. раздел 4.1).

Алгоритм работы измерителей вектора смещения изображений строится на основе нахождения координат максимума взаимокорреляционной функции двух изображений — корреляционно-экстремального метода. На первом этапе этого метода одно изображение сравнивается с другим, различно смещенным относительно первого. В классическом алгоритме усредняются произведения сигналов, в его модификациях — более простые функции (см. раздел 4.5).

Общей чертой всех измерителей параметров изображений является поиск компромисса между сложностью и точностью. При этом следует выделять два аспекта. Во-первых, эффективность использования сложности алгоритма всегда убывает по мере приближения к потенциальной ошибке, лимитируемой шумом. Во-вторых, следует выделять два типа видеоинформационных систем по степени целесообразности создания сложных алгоритмов, обеспечивающих близкую к потенциальной точность оценивания. К первому типу относятся системы со сложными и дорогостоящими механическими или оптическими подсистемами. Так, в астрономии, в том числе внеатмосферной [66, 97, 107], затраты на видеокомпьютер все равно уступают стоимости телескопа и оправдывают себя бесценным новым знанием. Ко второму типу относятся системы с механической подсистемой средней сложности — роботы, используемые в промышленности и рассматриваемые в следующей главе. В них, в соответствии с правилом суммирования винеровских ошибок (гл. 2) и критерием минимума суммарных информационных ошибок (см. 3.8), информационно-решающая подсистема должна обеспечивать точность, соизмеримую с точностью исполнительного механизма.

4.5. СОКРАЩЕНИЕ ВЫЧИСЛИТЕЛЬНЫХ ЗАТРАТ

Производительность компьютеров связана с экономией вычислительных затрат. Общий характер принципа экономии подразумевает широкий набор способов и алгоритмов, разделяемый на два класса. Первый экономит ресурсы без потерь полезной информации; он базируется на более полном использовании априорной информации. Второй экономит ресурсы за счет компромиссных потерь информации; он основан на невозможности абсолютно точных измерений и допустимости незначительного увеличения существующих потерь информации.

К неразрушающим полезную информацию методам относится обратимое кодирование, позволяющее в принципе сколь угодно приближаться к идеальному по Шеннону. Реальные каналы связи его не реализуют из-за возникающей сложности кодера: принцип экономии информации охватывает и эффективность использования канала и сложность кодера. Отражает это и критерий (3.20) минимума суммарных информационных ошибок и связанный с ним критерий (4.35). Шеннон отмечал [3, с. 453], что кодово-импульсная модуляция по достижимой скорости передачи на единицу полосы частот лишь незначительно уступает идеальному (неограниченно сложному) кодированию.

При наличии дискретного по времени канала связи с белым шумом с дисперсией σ^2 и передаче случайной величины с дисперсией λ пропускная способность C канала достигается при гауссовском распределении сигнала, которая превышает скорость передачи двоичного симметричного сигнала C' (при тех же дисперсиях сигнала и шума). При росте отношения сигнал-шум в канале пропускная способность C растет пропорционально его логарифму, скорость же передачи C' ограничена значением в один бит на символ (рис. 4.18). Отношение пропускных способностей C'/C имеет меньший единицы (примерно 0,8) максимум при отношении сигнал-шум около двух. Низкое значение скорости передачи информации при этом максимуме (полбита на символ) вынуждает повышать отношение сигнал-шум в канале, хотя при этом эффективность его использования уменьшается. *Чрезмерная надежность приобретается лишь при чрезмерной расточительности ресурсов.*

Экономия информации связана с понятием *минимальной достаточной статистики*, которая определяется по возможности сжатия кода без потерь информации. Достаточность имеет конкретный смысл, например, достаточность для оценивания параметра или проверки гипотез. Поэтому в сообщении может содержаться много минимальных достаточных статистик, например: при наблюдении последовательности отсчетов, состоящих из суммы неизвестного сигнала и гауссовского шума — сумма отсчетов; при наблюдении гауссовского случайного процесса на фоне гауссовского шума — свертка исходного сигнала с импульсной характеристикой винеровского фильтра.

Представление наблюдений в виде совокупности отсчетов *разложения в ряд* также не разрушает информации. При этом условием возможности точного обратного преобразования является полнота ортонормальной системы базисных функций [14]. Используют разложение в обобщенный ряд Фурье по гармоническим функциям, функциям Уолша, Эрмита, Чебышева и др. Для случайных процессов используют разложение в ряд Карунена — Лоэва. Его коэффициенты статистически независимы, что позволяет находить количество информации суммированием его для компонентов разложения. *Ряд Карунена — Ло-*

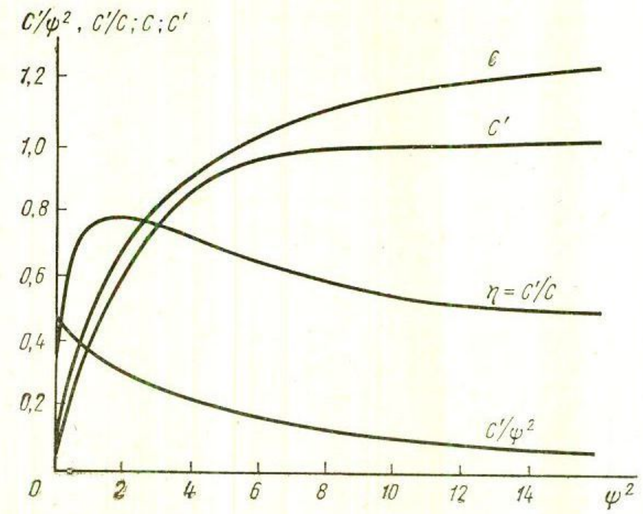


Рис. 4.18. Зависимость пропускной способности гауссовского и двоичного симметричного каналов от отношения сигнал-шум в канале

эва имеет предельно высокую скорость сходимости суммы ограниченного числа членов ряда к исходному сигналу. Для дискретного сигнала с N отсчетами разложение Карунена — Лоэва требуется около $2N^2$ умножений и сложений. Его используют при малом числе отсчетов и как эталон при оценке качества систем, так как не существует быстрого алгоритма для сокращения вычислений, как в быстром преобразовании Фурье, позволяющем сократить число вычислений с N^2 до $N \log N$ [31, 51, 62].

Структура вычислений, состоящая из умножений и сложений, имеет место и при вычислении взаимно корреляционных функций и свертки, для которых используются быстрые алгоритмы [51, 62, 66], позволяющие также сократить число вычислений до $N \log N$ (при одинаковой длине N сигнала и импульсной характеристики). Применение китайской теоремы об остатках [51, 57] позволяет преобразовать одномерную циклическую свертку в многомерную. Для некоторых длин последовательностей, в частности, соответствующих числам Мерсена (2^n-1) и числам Ферма (2^n+1), сокращение вычислительных затрат может быть даже более эффективным, чем в быстром преобразовании Фурье благодаря использованию лишь операций сдвига и сложения, т. е. предельного упрощения самой трудоемкой операции — умножения.

При измерениях параметров сигналов и поиске объектов используется принцип последовательного разделения априорного интервала на несколько частей, уменьшающихся с каждым шагом, метод *полихотомии*[1]. Число частей, на которые делится на каждом этапе априорный интервал S, определяется основанием a полихотомии. Это основание определяет число тактов m, за которое последовательное разбиение приводит к достигаемому апостериорному интервалу неопределенности Δs:

$$m = \mathrm{ent}(\log_a S/\Delta s + 1).$$

При основании $a=2$ имеем широко распространенный на практике метод *дихотомии* — рассечения априорного интервала на каждом этапе пополам. Аналитическое определение оптимального по минимуму суммарных затрат, состоящих из числа компараторов и числа тактов измерения, основания полихотомии аналогично определению оптимального основания кода [65]. Оптимальное значение основания полихотомии, обеспечивающее минимум информационных затрат, равно е. Различие в требуемых информационных затратах показано ниже:

a	2	е	3	4	8	16	32	64
$R/R_{\text{мин}}$	1,06	1	1,006	1,06	1,42	2,13	3,43	5,67

Затраты, наиболее близкие к соответствующим идеальному нереализуемому основанию, имеет полихотомия с основанием 3, т. е. трихотомия; она используется при измерении взаимно корреляционной функции при совмещении изображений [70].

Борьба за быстродействие квантователя, т. е. уменьшение числа тактов измерения, приводит к попыткам использовать основание полихотомии больше двух. Это означает переход к параллельно-последовательным квантователям, а в пределе — к параллельным, реализующим максимальное быстродействие ценой увеличения числа компараторов, т. е. за счет нерационального использования информационных ресурсов.

[1] От греческого поли — много, хотос — сечение.

Обработка информации в реальном времени ведется для уменьшения задержек (см. 5.4) и для экономии информации. Идеальная структура, использующая трехмерный ряд Карунена — Лоэва, требует большого объема памяти для хранения значений входного процесса и базисных функций. Поэтому фильтрацию сигналов осуществляют сверткой с импульсной характеристикой фильтра. Экономия памяти достигается за счет использования априорной информации — модели сигнала в виде марковского процесса, обусловившей переход к структурам с обратными связями — *рекурсивным фильтрам* (фильтры Калмана — Бьюси, которые применимы, к задачам, связанным с нестационарными сигналами и нелинейными системами). В них возможность экономии памяти определяется логарифмическим ростом памяти при росте числа слагаемых, а не линейным, как у трансверсальных фильтров.

Экономить память видеопроцессора удается и за счет организации *структуры фильтра*. В частности, в рекурсивных фильтрах могут запоминаться различные сигналы. Например, с точки зрения достигаемой ошибки оценивания сигнала фильтром с разностным уравнением $y_n = x_n - x_{n-1} + (1-\alpha) y_{n-1}$ варианты структуры, показанные на рис. 4.19, эквивалентны, хотя требуемая память для этих структур различна.

Приближенные вычисления, систематизированные А. Н. Крыловым, опираются на концепцию достаточности точности: раз результат, получаемый из неточных данных неточен и при точных вычислениях, то нет надобности в абсолютно точных вычислениях, нужна лишь уверенность, что суммарная погрешность не превышает заданных пределов. Эти пределы находятся по критерию минимума суммарной информационной ошибки (3.20) и (4.35).

В теории информации *приближенная передача сигналов* (т. е. с заданной погрешностью) опирается на понятие ε-энтропии. Идеальное кодирование, в частности идеальное линейное кодирование, рассмотренное в гл. 2, характеризуется усечением бесконечного числа членов ряда, представляющего процесс.

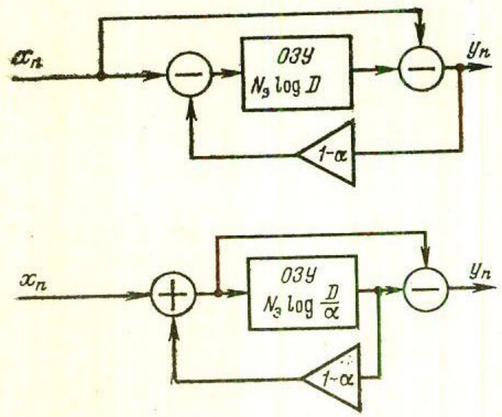


Рис. 4.19. Возможные структуры режекторных рекурсивных фильтров первого порядка

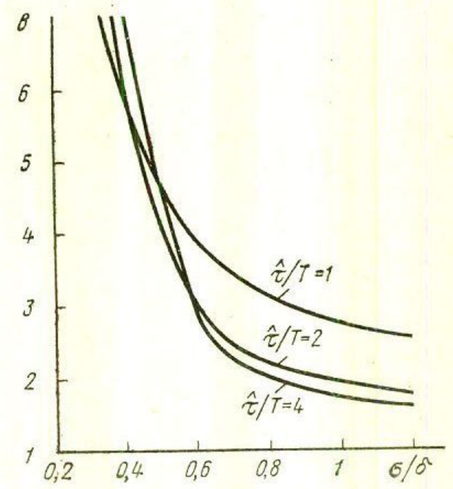


Рис. 4.20. Зависимость коэффициента шума цифрового рекурсивного фильтра от нормированного на шум σ шага квантования δ при различных постоянных времени фильтра τ/T

Примерами усечения спектральных коэффициентов частотных фильтров являются передаточный фильтр (гл. 2) и фильтр Костаса (гл. 1).

Аналогично производится усечение импульсной характеристики. Эквивалентная длительность импульсной характеристики винеровского фильтра обратно пропорциональна отношению сигнал-шум, а его частотная характеристика — прямо пропорциональна. Поэтому для экономии памяти сигналы с большим отношением сигнал-шум надо фильтровать во временной области (с усечением импульсной характеристики), а сигналы с малым отношением сигнал-шум — в частотной (с усечением частотной характеристики).

Квантование случайных величин разрушает полезную информацию, однако является необходимой процедурой для перехода от бесконечной энтропии кода $\hat{z}$ (гл. 2) к конечной энтропии цифрового кода и для разрушения шумовой информации (см. 3.8). Разрядность АЦП выбирается по ошибкам оценивания сигнала, т. е. по максимальному отношению сигнал-шум: $m = \log_2 \psi_{макс}$. *В видеокомпьютере при реализации фильтрации сигналов разрядность должна быть больше* и пропорциональна нормированной на время такта T (элемента, кадра) постоянной времени фильтра $\hat{\tau}$ [2]:

$$m = \log_2(\psi_{макс}\hat{\tau}/T).$$

Выбор мелкой разрядной сетки в АЦП и видеопроцессоре позволяет избежать ухудшения коэффициента шума цифровых рекурсивных фильтров, который, например для режекторного фильтра первого порядка (рис. 4.20), лишь при достаточно малых шагах квантования δ приближается к значениям для аналоговых прототипов. Экономия памяти рекурсивных фильтров при межкадровой обработке изображений достигается *управляемым округлением* [2], применяемым для предотвращения накопления ошибок округления, т. е. обеспечения абсолютной устойчивости фильтров.

Сокращение вычислительных затрат достигается при *усечении объемов выборки* сигнала при расчете его характеристик. Часто сигналы считают стационарными, поэтому прореживание наблюдений при вычислении функций распределения или корреляционных матриц позволяет на один-два порядка сократить вычислительные затраты. Возникающие ошибки экспериментально обнаруживаются лишь при нарушении стационарности.

Отступление от точных формул для сокращения информационных затрат используется при вычислении меры сходства изображений для их совмещения. Оптимальным является вычисление аргумента максимума взаимно корреляционной функции, обеспечивающее потенциальную точность измерения сдвига изображений при высокой сложности. *Упрощение алгоритма вычисления смещения получают нахождением модуля разности, или даже просто площади изображения, на которой разность данного сигнала и эталонного превышает заданный порог.* Это увеличивает ошибку измерения смещения, но часто оставляет на уровне долей элемента разложения. Любое упрощение алгоритма достигается ценой увеличения ошибки и критерием его качества становится приближение к предельному соотношению потери полезной информации и шумовой информации или сложности (см. 4.4).

4.6. ОБРАБОТКА ВИДЕОИНФОРМАЦИИ В ДИАЛОГЕ ЧЕЛОВЕКА С КОМПЬЮТЕРОМ

Задача обработки видеоинформации возникает во многих приложениях, где окончательное решение может принять только человек, компьютер же призван лишь облегчить его работу. Такая ситуация типична для исследования природных ресурсов [106] и астрономии [87, 107]. Существует много систем обработки изображений в интерактивном режиме работы компьютера, например HT680X (Венгрия), AGA820 (Швеция), RADIANS820 и PERICOLOR (Франция), СКИФ (СССР), MAGISCAN (Великобритания), MIDAS (ФРГ). Из-за разнообразия систем внедрение вычислительной техники в научные исследования пока процесс абсолютно хаотичный, приводящий к тому, что необходимая информация оказывается распыленной, а бóльшая часть — и вовсе недоступной [115].

Вместе с тем использование компьютеров переводит все исследования на новый качественный уровень. На современном этапе высококвалифицированные специалисты найдут в вычислительных машинах помощников, которые резко расширят человеческие возможности в познании [115].

Работа компьютера в диалоге с человеком включает не только анализ изображений, но и их синтез [116], который, в силу необходимости автоматизации конструирования и возможностей машинной графики, предшествовал появлению автоматического анализа, т. е. компьютерного зрения.

Системы обработки изображений создаются на базе универсального компьютера с использованием внешних устройств: кадрового ОЗУ, адресного процессора, управляющего доступом к фрагментам обрабатываемого изображения, АЦП, ЦАП, видеомонитора, графического дисплея. Например, система обработки и анализа изображений MIDAS, разработанная для Южной Европейской обсерватории и являющаяся основным инструментом обработки астрономических изображений, обеспечивает возможность использования стандартных алгоритмов анализа и большую легкость и гибкость интерактивного диалога при создании собственных алгоритмов. В этой системе используется суперминикомпьютер серии VAX с операционной системой VMS, транслятором Фортран, ОЗУ не менее 4 Мбайт и накопителем на магнитных дисках емкостью 150 Мбайт. Мощное программное обеспечение системы MIDAS имеет большой набор стандартных операций над изображениями, включающий: сокращение размерности, выделение признаков (выбор наиболее полезных переменных); визуализацию многомерных данных, индентификацию групп объектов или посторонних объектов; идентификацию спектрограмм; кластерный и дискриминантный анализ в различных модификациях; свертку с многоэлементными апертурами, позволяющими накапливать сигналы для борьбы с шумами и подавлять мешающие фоны; алгоритмы для борьбы с шумом, распределенным в

соответствии с законами Гаусса и Пуассона; определение моментов распределения вплоть до 4-го; построение гистограмм и другие операции, ориентированные на анализ астрономических изображений, получаемых с помощью телекамер на ПЗС.

Столь широкие возможности системы вызваны применением подхода, отличного от развиваемого в технике связи и принятого при реализации телевизионных роботов: в диалоговых системах человека с компьютером *человек является самым высокоинтеллектуальным, но зато и самым медленным звеном.* Поэтому в этих интерактивных системах придерживаются принципа сокращения суммарного времени решения задачи, для чего сосредотачивают внимание на повышении скорости обслуживания человека компьютером, оставляя практически без внимания вопросы экономии вычислительных ресурсов. Конечно, многие методы сокращения вычислительных затрат, рассмотренные в 4.5, используются при разработке аппаратной части и программного обеспечения диалоговой системы обработки изображений. Вместе с тем, диалоговым системам свойственны специфические особенности, не характерные для информационно-решающих телевизионных систем. В частности, использование при программировании языков высокого уровня преследует цель сокращения времени программирования, тогда как вопросы *избыточности* не являются определяющими. Например, Фортран обладает примерно 50%-ной избыточностью. Кроме того, различие в «стоимости» аппаратурных затрат и времени совместной работы человека и компьютера приводит к тому, что разрабатываемые для этих целей алгоритмы и программы оцениваются исходя из их алгоритмической сложности, опирающейся в основном на число шагов вычислений, а не из вычислительной сложности, учитывающей и требуемую память, т. е. аппаратурные затраты. Эта ситуация типична для использования универсальных компьютеров в диалоговом режиме [33]: «Программирование и решение задач на машинах является одним из видов обработки информации, поэтому возможно хотя бы качественное использование соображений теории информации, разработанной Шенноном, несмотря на то, что получение точных количественных оценок будет в данном случае весьма затруднительным. Применительно к нашему случаю основной принцип Шеннона гласит следующее. Для однородных групп часто встречающихся задач целесообразно иметь высокостандартные и тщательно отработанные программы».

В связи со сменой критерия качества системы обычно даже не ставится вопрос об экономичном кодировании в смысле оценки степени приближения метода кодирования к идеальному по Шеннону [3, с. 597], которое определяется уравнениями (1.14) и (1.15), а просто изыскиваются различные практические способы сжатия сигналов [117]. В связи с этим уместно напомнить мнение Дж. фон Неймана [49]: «Независимо от того, как вы будете работать на вычислительной машине, вы просто не в состоянии будете использовать ее возможности на 100%». Возможно, что качество работы компьютера в диалоговом режиме сможет быть оценено по количеству новых знаний — новой *семантической* информации, получаемой человеком в таком диалоге.

Выводы по главе. Компьютерная переработка информации содержит элементы кодирования и принятия решений. Видеокомпьютер является специфическим узлом системы связи, в котором алгоритмическая переработка изображений существенно уменьшает требуемую пропускную способность устройства вывода по сравнению с устройством ввода информации. Для этого видеокомпьютер последовательно осуществляет все этапы борьбы с фоном и шумом, выделения объектов, измерения их инвариантных к сдвигу, повороту и масштабу признаков, параметров положения и движения и классификации. При классификации подвижных изображений фильтрация изображений является пространственно-временной. Число одновременно анализируемых кадров определяется динамикой сюжета и числом различно движущихся его компонент. Возможности классификации ограничены влиянием шума, сложного фона, энергии разностного сигнала классифицируемых изображений при известной форме или разностью энергетических спектров при случайных сигналах.

Идеальный классификатор гауссовских сигналов содержит набор каналов, включающих оптимальные линейные фильтры — различающий, настроенный на разность сигналов, и передаточный (гл. 2). Важную роль играют информационные характеристики сигналов и видеопроцессоров: ограничение числа элементов, занимаемых изображением объекта, ограничивает энтропию изображения и проявляется в возникновении ошибок неинвариантности к сдвигу; вычислительная сложность классификатора, его объем памяти и быстродействие ставят предел числу различаемых гипотез и достоверности классификации изображений в реальном времени. Информационно-решающая система обладает внутренним единством из-за требований к точности передачи изображений, выдвигаемых автоматическим получателем. Его ограниченная сложность является причиной сокращения четкости ПЗС-телекамер на периферии поля зрения. Стремление повысить сложность решаемых задач при ограниченной сложности видеокомпьютера вызывает необходимость экономии объема вычислений на всех этапах переработки видеоинформации.

Глава 5.

ТЕЛЕВИЗИОННЫЕ РОБОТЫ

Робототехника является частью информатики — фундаментальной естественной науки, изучающей процессы передачи и переработки информации. В задачу информатики входит разработ-

ка триединой структуры *модель—алгоритм—программа* [7]. Телевизионные роботы олицетворяют раздел информатики — видеоинформатику, т. е. науку о передаче и переработке изображений.

Специфика видеоинформатики, в том числе телевизионных роботов, проявляется во всех трех звеньях указанной триады. Модель должна отражать свойства не только наблюдаемого изображения, но и особенности преобразования многомерных оптических сигналов (как функций пространственных координат, времени и длины волны) в электрический сигнал (код меньшей размерности). Важнейшей спецификой моделей видеоинформационных систем является необходимость учета воздействия на непрерывный оптический сигнал изображения источников шумов — как входного, так и шума канала (см. гл. 1, 2). Алгоритмы переработки видеоинформации последовательно решают большой комплекс взаимосвязанных, подчиненных единой цели задач оценивания, принятия решения и формирования управления, обеспечивающего оптимальное поведение робота. Важной спецификой алгоритмического обеспечения телевизионного робота является его ориентирование на максимальное быстродействие исполнения программы. Для программирования телевизионных роботов используются многочисленные языки как широко распространенные (Бейсик, Паскаль, Фортран, Си), используемые для реализации алгоритмов оценивания и принятия решения, так и специально разработанные (Юнимайт, ЯРС, Хелп и др.) — для реализации алгоритмов управления роботами. Важнейшей спецификой третьего звена видеоинформационной триады является создание программного обеспечения роботов на уровне заданий, а не команд.

5.1. ОБЛАСТИ ПРИМЕНЕНИЯ И КЛАССИФИКАЦИЯ

Существует много определений роботов. Так, считают, что это машины, в определенной степени копирующие поведение людей, или механические руки, управляемые ЭВМ [94]. Роботом называют также автомат с внешней обратной связью, синтезированный с общей схемой системы управления [99]. Эта схема содержит рецепторы, память, преобразователь информации и эффекторы (исполнительные устройства). При этом роботы не во всех случаях отличают от традиционных автоматов. Вместе с тем констатируется [7], что роботы являются принципиально новым классом систем, являющихся симбиозом механических, электрических и электронных узлов, отличающихся многоцелевым назначением и способностью самостоятельно или по команде изменять свое поведение в условиях изменяющейся среды. В этих, в общем верных, определениях не подчеркивается роль переработки информации, т. е. принадлежности роботов к информатике. Поэтому приведенные определения роботов целесообразно дополнить следующим: робот — *это сложная автономная искусственная система, перерабатывающая на основе заложенной в ней априорной информации сигналы из внешней среды и на основе извлеченной из них информации последовательно формирующая решения, управление и целенаправленное поведение.* Отметим, что в понятии телевизионного робота ключевую роль играет концепция обратной связи. Управляющий орган, включающий в себя телевизионную информационно-решающую и управляющую системы, передает исполнительному органу команды, осуществляя прямую связь. Этот же управляющий орган воспринимает, интерпретирует визуальную информацию, по которой контролирует исполнение команд, осуществляя обратную связь. Обратная связь, обеспечивающая точное сложное движение манипулятора робота или всего мобильного робота, должна включать измерения положения, скорости и ускорения исполнительного органа.

Публикации по «зрению роботов» идут большим потоком. Выпускаются книги (например, [88—94]), статьи, проводятся конференции и выставки, создаются новые научные советы академий наук, общества, журналы. Так, журнал Computer grafic, vision and image processing ежегодно публикует обзоры по обработке изображений, число ссылок в которых порой превышает 1000. Особое место в ряду популярных изданий по робототехнике занимает хорошо иллюстрированная книга [94]. Эта книга, а также изданная в США энциклопедия робототехники, показывают зрелость робототехники. Вместе с тем обилие публикаций по «зрению роботов» свидетельствует не только о достижениях, но и трудностях в решении важной научно-технической задачи создания надежных, экономичных, быстродействующих телевизионных роботов.

Широкий фронт работ по телевизионным роботам обусловлен многочисленностью областей их применения. Существует несколько сотен моделей промышленных роботов, созданы мобильные роботы для внутрицеховых и складских транспортных операций, сельского хозяйства, космических экспериментов. Экономически оправдано применение телевизионных роботов в производствах с быстрой сменой номенклатуры изделий и операций. Широкое применение телевизионные роботы получили, в частности, в микроэлектронике из-за непродолжительности спроса на ее изделия. Применение адаптивных телевизионных роботов позволяет сохранить большую часть оборудования (до 70%) при смене объекта производства, в то время как без них удается при аналогичной смене сохранить лишь до 20% оборудования [92].

Целесообразность применения телевизионных роботов может определяться неэкономическими факторами — когда операции не могут выполняться человеком в силу вредности среды или удаленности объекта, например при обслуживании атомных энергетических установок и в космических экспериментах.

Значимость телевизионных роботов характеризует также большая доля (до 80%) технологической информации в робототехнических производствах, получаемая с помощью оптических систем, а также большой объем продажи телевизионных систем для роботов, достигающий одной трети всего рынка зрительных систем. *Развитие телевизионных роботов — яркий пример положительной обратной связи между наукой и производством: потребности производства стимулируют развитие телевизионных роботов, они же с ростом технологических возможностей становятся все совершеннее.* В результате мировой парк роботов удваивается каждые два года. На современном этапе развития робототехники созданы многие сотни моделей роботов, существенная доля которых

выпускается серийно. Напомним, что макет первой «механической руки», очувствленной и управляемой ЭВМ, был разработан и собран в 1961 г. в Массачусетском технологическом институте Г. Эрнстом под руководством К. Шеннона и М. Минского. Сейчас выпускается много типов роботов, которые классифицируются по следующим признакам:

по степени автономности — полуавтоматические (с участием человека-оператора) и полностью автоматические, в которых человек участвует лишь на этапах создания и обучения робота;

по функциональному назначению — транспортные (мобильные) и манипуляционные;

по виду обрабатываемой информации — фотометрические и дальнометрические;

по способу освещения предметной области — с рассеянным и направленным (равномерным и структурированным) освещением;

по способу измерения дальности — стереоскопические (две или более телевизионных камер) и локационные (сканирующие и проектирующие систему реперных отметок, триангуляционные и измеряющие время);

по сложности системы управления — простые (1—2 степени свободы) и сложные;

по сложности классификатора изображений — простые (двумерные изображения, равномерный фон), повышенной сложности (двумерные изображения, сложный подвижный фон) и высокой сложности (трехмерные изображения, сложный трехмерный подвижный фон);

по типу фотоприемника — одноэлементные, строчные (линейные) и двумерные (твердотельные и вакуумные).

В процессе развития робототехники сформировалась определенная иерархия уровня интеллектуальности роботов, которую часто называют их поколением. Понятие поколения в этом применении весьма условно, так как широкий парк действующих ныне роботов состоит из представителей всех трех поколений.

К *первому поколению* относят роботы с программным управлением, в которых телевидение не используется. Они работают по жесткой программе, которая записывается в ЗУ при наладке.

Ко *второму поколению* относят роботы с адаптивным управлением. Для очувствления робота используются телевизионные камеры. Эти роботы приспосабливаются к заранее неизвестным и изменяющимся условиям работы. В них обязательно используется обратная связь по состояниям робота и внешней среды, визуальная информация используется для управления в реальном масштабе времени.

К *третьему поколению* относят интеллектуальные роботы. В них широко используются методы адаптации и искусственного интеллекта. Они характеризуются самонастройкой, самопрограммированием, самообучением и самоорганизацией целенаправленного поведения.

5.2. СТРУКТУРА, ХАРАКТЕРИСТИКИ И ОСОБЕННОСТИ РОБОТОТЕХНИЧЕСКИХ ТЕЛЕВИЗИОННЫХ СИСТЕМ

Рассмотрим кратко структуру телевизионных роботов. Автоматические роботы — это единый комплекс (рис. 5.1), состоящий

из четырех подсистем: телевизионной *1* (информационной), вычислительной *2* (решающей), управляющей *3* и исполнительной *4*, реализующей поведение робота. Принципиально важным с точки зрения информатики является объединение этих подсистем в единый организм с помощью цепи каналов связи (электрических *5*, механического *6* и оптического *7*).

Главным стратегическим направлением развития телевизионных роботов является переход к твердотельным телевизионным информационно-решающим системам на базе ПЗС и микропроцессоров. Сложность телевизионного робота характеризуется не только сложностью его структуры, но и сложностью программного обеспечения — самой дорогостоящей части микропроцессорных систем.

Информационно-решающая система, являющаяся предметом рассмотрения данной книги, состоит из телевизионной камеры и классификатора изображений, вырабатывающего то или иное решение. Эти подсистемы подразделяют на две группы в соответствии со способом установки телевизионной камеры: на неподвижном основании либо подвижном манипуляторе робота.

Управляющая подсистема объединяет вычислитель, преобразующий выработанные классификатором решения в управляющие команды, и исполнительный механизм. Высокая эффективность систем управления достигается применением специализированных языков программирования.

Создание сложной структуры телевизионных роботов преследует цели достижения необходимых качественных характеристик. При проектировании добиваются высокого уровня качества функционирования, используя критерии максимума достоверности, минимума временных или энергетических затрат — скорости при-

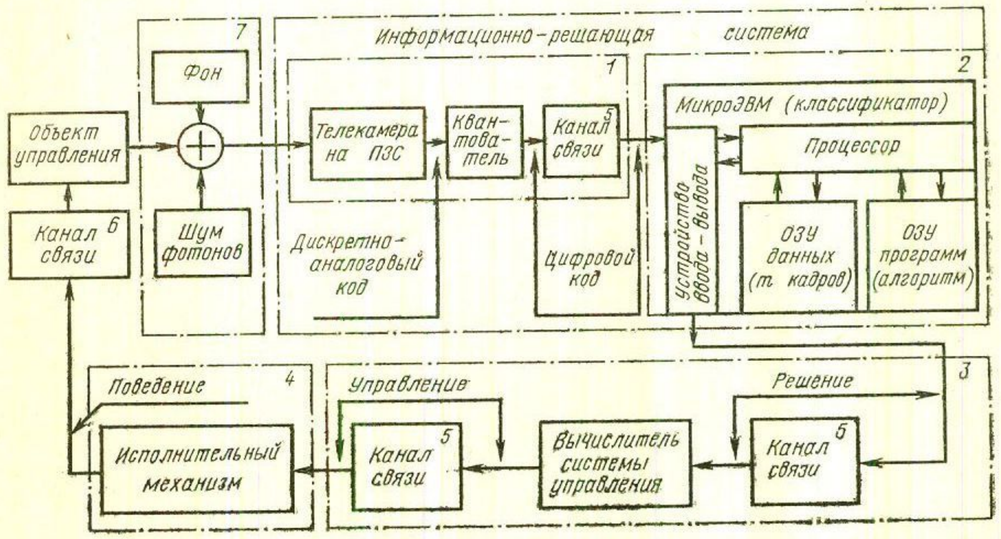


Рис. 5.1. Структурная схема переработки информации в телевизионном роботе

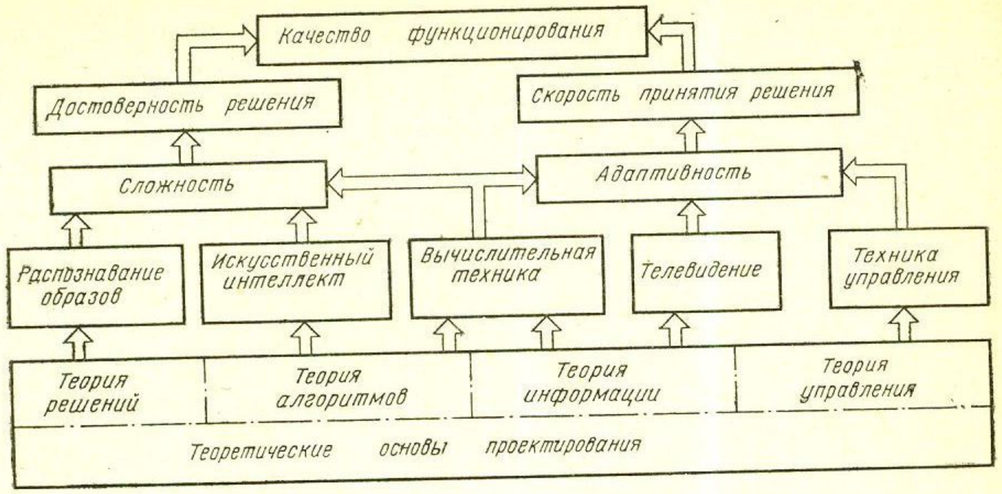


Рис. 5.2. Взаимосвязь проблем телевизионных роботов и путей их решения

нятия решения (рис. 5.2). Все звенья телевизионного робота должны быть согласованы по точности. В первую очередь это относится к точностям управляющей и информационно-решающей систем. При чрезмерно высокой точности информационно-решающей системы по сравнению с управляющей имеет место нерациональное использование пропускной способности (а также быстродействия, сложности), т. е. загрузка каналов связи (и памяти) информацией, которая для грубой механической системы является шумовой.

Возможности информационно-решающих систем телевизионных роботов определяются точностью и экономичностью преобразования оптического сигнала телекамерой в код и затем классификатором кода в решение. Эти качественные характеристики достигаются за счет адаптации и быстродействия информационно-решающей системы. Часто эти характеристики определяются не принципиальными возможностями электронной и вычислительной техники, а экономическими показателями всего телевизионного робота. Так, уникальные системы для автоматических астрономических наблюдений могут содержать в своем составе телескопы и большие универсальные ЭВМ, обрабатывающие изображения сверхвысокой четкости. Роботы в промышленном производстве, напротив, должны строиться при учете ряда ограничений. Ограничение решающей способности классификатора, выполненного на микропроцессоре, при стремлении к увеличению быстродействия приводит к упрощению алгоритмов принятия решений и сокращению четкости изображения (см. раздел 4.5). Это достигается снижением по сравнению с прочими телевизионными системами четкости и числа воспринимаемых градаций яркости. В ряде случаев четкость снижают всего до 1000 элементов (32××32), а сигнал квантуют всего на два уровня.

Информационно-решающие системы телевизионных роботов во многом имеют свою специфику, выделяющую их из общего класса видеоинформационных систем. Отметим важнейшие особенности робототехнических телевизионных систем.

1. *Подвижность изображения* объекта управления, имеющая место при установке телевизионной камеры как на неподвижном, так и на подвижном основании. В последнем случае сложность принятия решения еще более возрастает из-за того, что фон, окружающий объект управления, вследствие движения телевизионной камеры становится в координатах телевизионного растра подвижным. Робот с подвижной телевизионной камерой лучше использует четкость фотоприемника и точнее измеряет положение объекта управления в пространстве, снижает требования к калибровке системы позиционирования. Более того, движение телевизионной камеры в сочетании с использованием кадровых ОЗУ позволяет пропорционально их числу увеличить эффективную четкость телевизионной системы и снизить до необходимого предела вариации формы сигнала объекта управления при перемещении его изображения по фотоприемной матрице.

2. *Быстродействие*, обусловленное необходимостью работы в реальном масштабе времени. Это реальное время может составлять единицы и доли секунды. Именно по быстродействию телевизионный робот должен превзойти человека-оператора — это одна из основных целей робототехники.

3. *Компактность* и *экономичность*. Робот оборудуется телевизионной камерой и классификатором со стоимостью (сложностью), сравнимой со стоимостью (сложностью) манипулятора. Если стоимость классификатора слишком велика, то задачу целесообразно решать в полуавтоматическом режиме.

4. *Достоверность принятия решения.* Фиксированность алгоритма принятия решения телевизионным роботом при наличии внешних помех и необходимость быстрого принятия решения оставляет возможность ошибочного решения, поведение робота на основании которого может быть катастрофическим.

5. *Помехоустойчивость.* Достоверность принятия решения во многом зависит от помехоустойчивости, т. е. качества решения, принятого в условии наличия помех различных типов. Важнейшие из них: флуктуационные помехи — шумы фотоприемников и усилителей, фотонные шумы (играющие решающую роль при низких освещенностях); внешний фон (посторонние предметы), выделение полезной информации при наличии которого особенно затруднено при подвижности изображения фона (подвижности телевизионной камеры); помехи преобразования оптического изображения в цифровой код (дискретизация и квантование), приводящие к непостоянству формы сигнала объекта управления при его перемещении относительно телевизионной камеры (сдвиге, повороте, масштабировании).

6. *Измерение дальности* до объекта управления. Трехмерность пространства действия робота вносит существенное отличие в ме-

тоды классификации. Распространены стереоскопические телевизионные системы (содержащие до пяти телевизионных камер) и системы, сочетающие телевидение с лазерной дальномерной локацией. Именно в разработке алгоритмов распознавания трехмерных объектов велика роль методов искусственного интеллекта.

7. *Адаптивность.* Телевизионный робот должен быстро и качественно выполнять свои функции в условиях изменяющихся освещения и динамики объекта управления. Адаптивными делают все звенья робота: систему управления, классификатор и телевизионную камеру.

8. *Метрологичность.* Высокая достоверность решения телевизионного робота в условиях помех может быть достигнута лишь при высоком качестве преобразования оптического изображения в код. Такое качество обеспечивается линейностью преобразования света в электрический сигнал, прецизионностью воспроизведения геометрических характеристик изображения объекта управления и безынерционностью телевизионной камеры. Этими качествами в наивысшей степени обладают твердотельные телевизионные камеры на матричных ПЗС.

5.3. СЛОЖНОСТЬ И БЫСТРОДЕЙСТВИЕ

Различают два типа сложности — сложность системы и сложность проектирования системы [113]. Для оценки сложности используют понятия алгоритмической, вычислительной и информационной сложности. Последняя представляет собой развитие понятия ε-сложности, имеющей общие корни с понятием ε-энтропии [114]. Сложность создания системы имеет практический оттенок и определяется текущим уровнем развития науки и техники. Сложность создания телевизионных роботов связана не с телевизионной камерой и даже не с архитектурой ЭВМ, а с синтезом алгоритмов принятия решения и управления (реализуемых аппаратно или программно) [88].

Множество телевизионных роботов разделяется по типу обрабатываемых изображений. Большинство телевизионных роботов ограничивается анализом черно-белых (нецветных) изображений, число градаций яркости часто равно 2, а решаемые задачи не слишком сложные. Сложность решаемых задач определяется сложностью:

объекта и его движения в предметной плоскости;

фона и его движения в предметной плоскости;

формализации цели поведения робота;

манипулирования (числом степеней свободы механизма).

Сложность алгоритма удобно оценивать средним числом $\mathcal{D}$ операций на каждый из кодовых импульсов (знаков) телевизионного сигнала, которая зависит от сложности условий опознавания и качества примененного алгоритма. Его сложность может изменяться в широких пределах от единиц до тысяч операций на телевизионный элемент. Положения теории информации распространяются на устройства обработки сигналов: пропускная способность или емкость канала идентична емкости памяти ОЗУ — важнейшего элемента классификатора. Емкость памяти классификатора I, хранящая m кадров изображения, связана со скоростью кода R на выходе телевизионной камеры и быстродействием V (операций в секунду) классификатора и временем $T_р$ принятия решения в нем [95]:

$$I = mRT_к, \quad V = R\mathcal{D}/T_р.$$

Подвижность объекта управления усложняет телевизионный робот. Движение объекта вызывает смаз изображения, т. е. потерю информации [32]. В условиях наличия шумов смазанное изображение не может быть восстановлено с малой ошибкой, характерной для неподвижных изображений; потеря информации из-за смаза обратима лишь частично (см. раздел 3.2). Поэтому высокоточные телевизионные роботы компенсируют движение объекта управления дополнительным движением телевизионной камеры, т. е. слежением за объектом [83, 87, 88]. Стремление к увеличению точности работы телевизионного робота при неточном априорном значении скорости движения изображения вызывает необходимость адаптации к движению во всех звеньях: в телевизионной камере, кодере изображений, классификаторе и устройстве управления. Это достигается усложнением всех подсистем телевизионного робота. Наличие в схеме телевизионного робота (рис. 5.3) не менее трех контуров обратной связи наглядно показывает уровень его сложности.

Высокая сложность телевизионных роботов определяет необходимость использования:

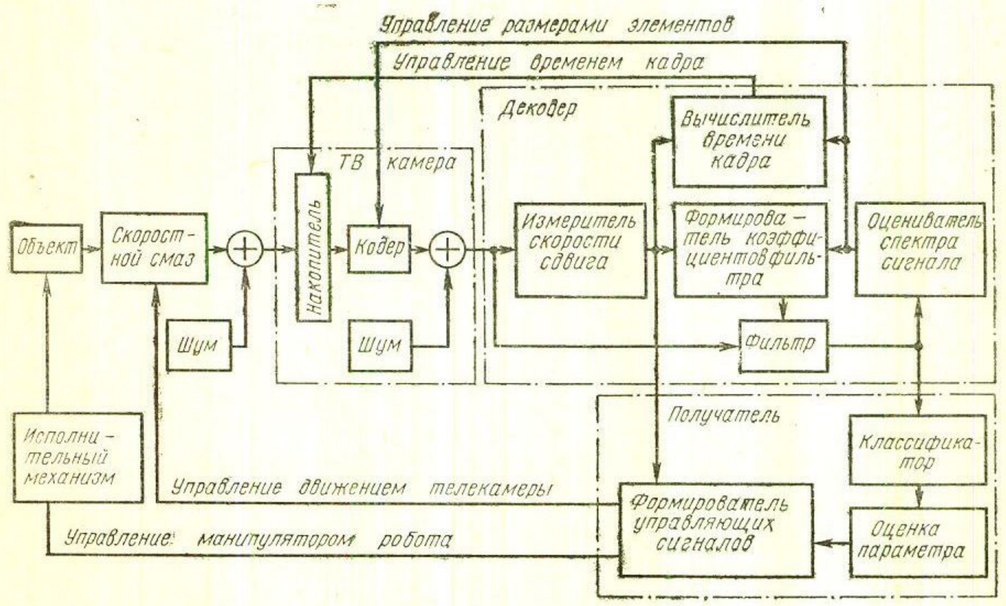


Рис. 5.3. Структурная схема универсального робота с информационно-решающей системой, обрабатывающей подвижные изображения

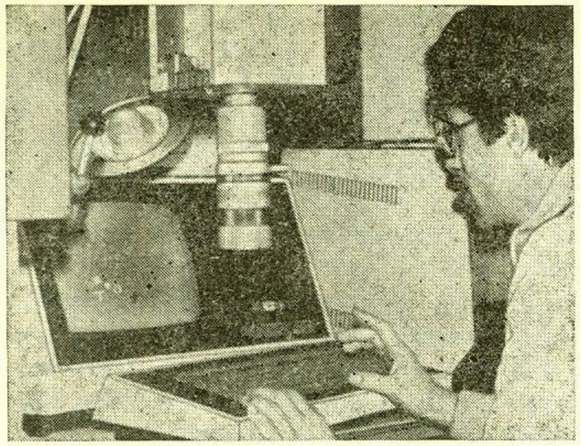


Рис. 5.4. Обучение телевизионного робота классификации изображений; на дисплее набор сигналов рис. 4.8

теории проектирования (см. рис. 5.3):

моделирования на универсальных ЭВМ сложных алгоритмов обработки изображения и поведения роботов;

макетирования плохо формализуемых особенностей;

метода проб и ошибок при наладке (обучении) телевизионного робота (рис. 5.4).

Классификатор — самая сложная подсистема телевизионного робота. В ней хранится вся информация о внешнем мире, которой располагает робот (заложенная при конструировании и обучении). Объем этой информации, хранимой в оперативных и перепрограммируемых постоянных запоминающих устройствах, часто измеряется мегабитами.

В классификаторах выполняются различные операции:

измерение параметров движения изображения для компенсации смаза (обычно корреляционно-экстремальными методами);

компенсация смаза («восстановление» изображений);

выделение контуров (для сокращения скорости кода);

принятие решений, которые делятся на три типа [96] — идентифицирующие, классифицирующие и анализирующие.

Так как минимизация ошибок классификации требует минимизации ошибок оценивания параметров изображения, то при определении сложности видеокомпьютера телевизионного робота следует воспользоваться вычислительным и информационным соотношениями неопределенности (4.33) и (4.34), которые служат основой *концепции достаточности точности*. Она утверждает необходимость отказа от стремления к очень высокой точности из-за неоправданных роста сложности и снижения быстродействия робота.

Использование роботов во многих приложениях оправдано только при их высоком быстродействии, которое должно превзойти «быстродействие» человека. Поэтому при синтезе алгоритмов работы телевизионных роботов часто минимизируют временные затраты. Борьба за быстродействие робота ведется четырьмя путями:

оптимизацией траектории движения робота (манипулятора);

разработкой экономичных алгоритмов переработки информации, характеризуемых минимально возможным числом шагов;

созданием специализированных видеопроцессоров: матричных, параллельных, систолических, магистральных, ассоциативных и ориентированных на умножение и обращение матриц;

построением видеопроцессоров в виде СБИС.

Причиной снижения скорости принятия решения может служить сложность решения задачи позиционирования, содержащей этапы:

определения позиции исполнительного органа (схвата);

определения поля зрения телевизионной камеры по данным дальномера;

помещения роботом детали в поле зрения дальномера. (Такая задача часто решается методом последовательных приближений. Время, затрачиваемое на различные этапы алгоритма, зависит от сложности изображения и объема программы.)

Повышение быстродействия телевизионного робота может быть достигнуто за счет специальных методов считывания сигналов с матричных ПЗС. К ним относятся метод процессора проекций (см. раздел 3.4) и метод переменной четкости (см. раздел 3.6). Каждый из них позволяет на порядок и более повысить частоту измерения параметров изображений без ухудшения точности.

При наблюдении быстропротекающих процессов может сложиться ситуация, когда четкости по полю изображения достаточно и ошибки классификации из-за неинвариантности к сдвигу изображения малы, но покадровая дискретизация оказывается слишком грубой даже на предельных кадровых частотах ПЗС-камер. Это ведет к большой ошибке неинвариантности к сдвигу сигнала во времени — смазу изображения. Для борьбы с такой ошибкой кроме использования фильтрации изображений (см. 4.2) применим способ, схожий с использованием многокадрового ОЗУ: для формирования точной оценки процесса могут быть использованы многоматричные ТВ камеры со сдвинутой по фазе кадровой синхронизацией матриц ПЗС.

В настоящее время существуют телевизионные роботы, проверяющие до 600 деталей в минуту, используя до 40 признаков для каждой детали; телевизионные роботы, распознающие свыше 520 изображений в минуту; робототехнические камеры для наблюдения быстродвижущихся объектов с частотой кадра от 60 до 2000 Гц [98].

К вопросу быстродействия роботов тесно примыкает необходимость экстраполяции сигналов, обеспечивающей функционирование робота в реальном времени.

5.4. ЭКСТРАПОЛЯЦИЯ ПОДВИЖНЫХ ИЗОБРАЖЕНИЙ

Телевизионный робот создается для преобразования наблюдения в конкретное поведение. Поэтому при всей его возможной интеллектуальности он не строит поведение по принципу «поживем — увидим». Именно скоростная переработка информации роботом позволяет ему превзойти человека в исполнении строго

формализованных операций. Эта картина делает понятной необходимость функционирования робота в реальном времени. Термин «реальное время» появился в противовес медленной последовательности переработок видеоинформации на универсальных ЭВМ. Сейчас часто этот термин понимают как ситуацию, при которой задачи решаются за время одного или, в крайнем случае, нескольких кадров. Робототехника вносит свое уточнение в понятие реального времени переработки информации: в сбалансированной вычислительно-механической системе ЭВМ должна перерабатывать информацию за время, не превышающее требуемого на элементарный акт поведения механической системы. При этом видеокомпьютер должен экстраполировать измеренные характеристики объекта на время такого действия робота.

Экстраполяция, т. е. прогноз положения наблюдаемого объекта, широко используется во многих видеоинформационных системах, в том числе и в радиолокации. При этом следует различать два аспекта экстраполяции в процессе:

накопления информации о наблюдаемом изображении для повышения достоверности принятия решения;

формирования управления для повышения точности функционирования исполнительных механизмов.

Типичным представителем первого типа экстраполяторов является траекторный накопитель информации, используемый в радиолокаторах при обнаружении и сопровождении отметок [68, 72].

Традиционно траекторные накопители осуществляют предсказание положения объекта (и его проверку) в следующем цикле обзора (в телевидении — в следующем кадре), опираясь на измеренные положения объекта в предыдущих циклах обзора и на модель движения объекта. Эта модель может быть случайной или полиномиальной. Последняя типична для обзорных систем; в точных системах, где детерминированное движение компенсируют, применяют случайную модель движения. При случайной модели движения для получения наилучшего прогноза используют оценку, полученную с помощью фильтрации, оптимальной по критерию минимума среднеквадратической ошибки (фильтры Винера и Калмана). Экстраполяторы первого типа характеризуются равноотстоящими узлами и в детерминированной модели требуемая оценка может быть вычислена по формулам Ньютона. Эти формулы позволяют построить полином степени, на единицу меньшей числа узлов интерполяции. При длительном наблюдении и простых моделях движения, когда число узлов заметно больше требуемой степени интерполирующего полинома, применяют метод наименьших квадратов, позволяющий сгладить возмущающее влияние шума на измерение координаты. В простейшем случае равномерного движения объекта с неизвестной заранее скоростью по прямолинейной траектории прогнозируемая оценка координат s_{n+1} определяется ее оценкой в текущем s_n и предыдущем s_{n-1} кадрах:

$$s_{n+1}=s_n+(s_n-s_{n-1}).$$

Специфика экстраполятора второго типа в системах управления состоит в том, что время, необходимое для осуществления поведения, не связано жестко с временем кадра. Здесь подразумевается, что поведение телевизионного робота связано с управлением детерминированной функцией. Поэтому расчетные формулы для экстраполированного положения объекта, принципиально не меняясь, включают в себя кроме времени кадра $T_к$ еще и время исполнения $T_и$. В это время входят все процедуры переработки информации, следующие за накоплением ее в фотоприемнике: для ПЗС с кадровым переносом это время считывания зарядов, равное времени накопления, время передачи сигналов, время принятия решения, время выработки управляющих сигналов и время исполнения рабочим органом предписанного ему поведения. Это приводит к использованию в экстраполяторах второго типа интерполяционных формул Лагранжа для неравноотстоящих узлов. Эта формула, так же как и формулы для равноотстоящих узлов, характеризуется степенью полинома на единицу меньшей числа узлов интерполяции. Так же как и в первом случае при длительном наблюдении возможно сглаживание траектории применением метода наименьших квадратов. При линейной модели движения объекта в поле зрения телевизионного робота экстраполированное на требуемый момент положение объекта составит

$$s_{n+1}=s_n+(s_n-s_{n-1})(T_к+T_и)/T_к.$$

В случае неравномерного движения (по модулю скорости и направлению) расчетные формулы усложняются пропорционально степени аппроксимирующего полинома из-за увеличения числа узлов интерполяции (наблюдаемых кадров), однако множитель $(T_к+T_и)/T_к$ сохраняется.

В телевизионных роботах, наблюдающих неравномерно движущиеся объекты и использующих фотоприемники с кадровым накоплением повысить точность экстраполяции его координат можно, измерив не только его межкадровое смещение, но и величину *скоростного смаза* изображения. Такая экстраполяция основана на разложении функции, описывающей координаты объекта, в ряд Тейлора. Так же как и при разложении сигнала в ряд Карунена — Лоэва (см. гл. 2) наличие шумов ведет к усечению числа членов ряда, в данном случае — учитываемых производных координат объекта. По одному смазанному изображению объекта надежно могут быть измерены его координаты и скорость. Более высокие производные координат объекта не выделяются из шумов при правильном выборе частоты кадров, согласно которому скоростной смаз должен быть малым (см. раздел 3.2). Подчеркнем, что *измерение скорости объекта по величине смаза позволяет получить оценку, не зависимую от оценок координат объекта в соседних кадрах.* Такое использование для экстраполяции координат и скорости объекта, измеренных по одному кадру, аналогично способу кодирования, использующему «значение функции и ее производной в точках отсчета, взятых через одну»

[3, с. 437]. В результате такой экстраполяции уменьшается время между получением оценки координат и прогнозируемым моментом: при экстраполяции по двум кадрам его среднее значение τ равно $1{,}5T_к$, при экстраполяции по одному кадру $\tau \approx T_к$. В результате измерение скоростного смаза позволяет при случайном движении объекта в 1,5 раза повысить точность совмещения исполнительного механизма робота с подвижным объектом.

Так как увеличение времени переработки информации ведет к возрастанию ошибок предсказания положения объекта, то представляется возможным выявить связь между требуемым быстродействием видеокомпьютера V и точностными характеристиками робота.

Будем считать, что основную долю времени исполнения составляет время формирования решения и управления $T_р$. Модель движения объекта в поле зрения примем случайной, описываемой стационарным случайным процессом $X(t)$ с гауссовской функцией распределения вероятностей и известными дисперсией λ и корреляционной функций $K(t)$. Координаты предмета измеряются с ошибкой, вызванной влиянием внешнего шума с известной дисперсией σ^2_1.

Такая модель экстраполяции позволяет для наблюдения дискретного (одномерного) процесса воспользоваться результатами для одномерного случая (см. раздел 2.3), в том числе формулами (2.29) и (2.31) для кодовой и сигнальной ε-энтропий с той лишь оговоркой, что входящая в них минимальная ошибка ε_B при экстраполяции станет больше, чем при оценивании. Не вникая в сложности, связанные с экстраполяцией случайного процесса с бесконечным прошлым и использованием разложения Вольда [101], рассмотрим упрощенный случай *предсказания случайного процесса по его одному наблюдению.* Известно [4, 11], что для двух случайных величин взаимная информация определяется коэффициентом корреляции ρ между ними:

$$I[X,\ Y] = -\frac{1}{2}\log(1-\rho^2). \tag{5.1}$$

В рассматриваемом случае коэффициент корреляции равен значению корреляционной функции $\rho = K(T_р)$. Взаимная информация двух значений гауссовского случайного процесса, разделенных временным интервалом $T_р$, при наблюдении зашумленного процесса $Y(t) = X(t-T_р) + \eta(t)$ также составит $-\frac{1}{2}\log(1-\rho^2)$, однако коэффициент корреляции меняется из-за влияния шума. В результате взаимная информация между значением случайного процесса и его предсказанным значением при упрощенной оценке определится одномерной формулой (см., например, [85]):

$$I[X(t),\ X(t-T_р)+\eta(t)] = \frac{1}{2}\log\frac{\lambda+\sigma_1^2}{\lambda(1-\rho^2)+\sigma_1^2} =$$

$$= \frac{1}{2}\log\frac{\lambda}{\varepsilon_B[1+\psi^2(1-\rho^2)]}. \tag{5.2}$$

Здесь $\psi^2 = \lambda/\sigma^2_1$ — отношение сигнал-шум. В этом разделе, как и в 2.3, λ — дисперсия одномерных случайных величин, а не отсчетов Карунена — Лоэва. Очевидно, что при отсутствии шума формула (5.2) совпадает с (5.1). Важно, что при уменьшении интервала времени $T_р$, разделяющего момент измерения и прогнозируемый момент, происходит увеличение коэффициента корреляции ρ. В пределе, при $T_р = 0$ коэффициент корреляции ρ равен 1. Этот предельный случай переводит формулу (5.2) в другой важный вариант — формулу Винера для взаимной информации случайного сигнала и его смеси с шумом [50]:

$$I[X(t),\ X(t)+\eta(t)] = \frac{1}{2}\log\frac{\lambda+\sigma_1^2}{\sigma_1^2} = \frac{1}{2}\log\frac{\lambda}{\varepsilon_B}.$$

Влияние ошибок, возникающих при переработке информации в видеопроцессоре (из-за ограниченной точности вычислений), может быть учтено аналогично влиянию ошибок канала связи. В этом случае влияние всех источников ошибок вычислений сводится к влиянию одного источника шума с дисперсией σ^2_1 и порождающего энергетический порог передачи информации $\mu = \varepsilon^* v$ [см. (2.46)]. Сигнальная ε-энтропия из (2.33) преобразуется к виду

$$I[X(t),\ Z(t)] = \frac{1}{2}\log\frac{\lambda}{\varepsilon_B[1+\psi^2(1-\rho^2)]+\mu}. \tag{5.3}$$

Для получения явной зависимости быстродействия видеокомпьютера от точности передачи предположим, что случайный процесс $X(t)$, описывающий траекторию движения, имеет экспоненциальную корреляционную функцию с интервалом корреляции τ. Тогда при не слишком больших интервалах предсказания $T_р$ коэффициент корреляции

$$\rho \approx 1 - T_р/\tau. \tag{5.4}$$

Время принятия решения видеокомпьютером связано с его быстродействием V (операций в секунду), сложностью алгоритма $\mathcal{D}$ (операций/бит), ε-энтропией кода $Z(t)$, несущего информацию об измеренных параметрах объекта, и размерностью $N_э$ входного сигнала (числом элементов разложения в телекамере):

$$T_р = N_э \mathcal{D} I(Y(t),\ Z(t))/V.$$

Если приближенно считать, что W компонент вектора движения объекта (смещения и скорости по различным направлениям) независимы и имеют одинаковую дисперсию, то кодовая ε-энтропия в W раз превысит ее для каждой компоненты, определяемой из (2.29)

$$I((Y(t),\ Z(t)) = \sum_{i=1}^{W} R_i = W I_i[Y(t),\ Z(t)] = \frac{W}{2}\log\frac{\hat{\lambda}}{\mu}. \tag{5.5}$$

Здесь $\hat{\lambda}$, как и в гл. 1, 2, — дисперсия оптимальной оценки сигнала. Время принятия решения тогда составит

$$T_{\text{р}}=\frac{N_{\text{э}}\mathcal{D}W}{2V}\log\frac{\hat{\lambda}}{\mu}. \tag{5.6}$$

Подстановка коэффициента корреляции (5.4) и времени принятия решения (5.6) в формулу (5.3) для взаимной информации между истинным и предсказанным значениями координат объекта дает выражение для сигнальной ε-энтропии

$$I[X(t),\ Z(t)]=\frac{1}{2}\log\frac{\lambda}{\varepsilon_B\left[1+\dfrac{\psi^2 N_{\text{э}}\mathcal{D}W}{V}\log\dfrac{\hat{\lambda}}{\mu}\right]+\mu}. \tag{5.7}$$

Анализ формулы (5.7) показывает, что *большие ошибки экстраполяции координат объекта имеют место как при слишком грубой передаче, так и при слишком точной.* При слишком грубой передаче ошибка экстраполяции растет в результате неточности знания исходного положения предмета, при слишком точной она растет из-за необходимости увеличения времени экстраполяции, так как растет время переработки информации. Существование этих двух противоположных тенденций позволяет найти оптимальную точность передачи сигналов. Дифференцируя выражение (5.7) по величине порога и приравнивая производную нулю, находим

$$\mu=\frac{\hat{\lambda}\mathcal{D}N_{\text{э}}W}{\tau V}. \tag{5.8}$$

В соответствии с *критерием минимума суммарной информационной ошибки* (см. раздел 3.8) будем считать, что потери информации (2.34) равны шумовой информации (2.35). Для такой информационно сбалансированной переработки информации порог μ равен винеровской ошибке оценивания отсчета процесса $\mu=\varepsilon_B$. Так как винеровская ошибка

$$\varepsilon_B=\frac{\lambda\sigma_1^2}{\lambda+\sigma_1^2}=\frac{\hat{\lambda}}{\psi^2}, \tag{5.9}$$

то кодовая ε-энтропия предсказываемой координаты и при идеальном кодировании пропускная способность для каждой измеряемой компоненты движения составит [см. (2.67) и (2.70)]:

$$I_i[Y(t),\ Z(t)]=\frac{1}{2}\log\frac{\hat{\lambda}_i}{\mu}=\frac{1}{2}\log\frac{\hat{\lambda}_i}{\varepsilon_{Bi}}=\log\psi_i. \tag{5.10}$$

Рассмотрение такой информационно сбалансированной системы, в которой порог μ равен винеровской ошибке, позволяет преобразовать формулу (5.8) к виду, определяющему требуемое быстродействие видеокомпьютера:

$$V=\psi^2\mathcal{D}N_{\text{э}}W/\tau. \tag{5.11}$$

При таком быстродействии обеспечивается время принятия решения, которое может быть получено подстановкой выражений (5.10) и (5.11) в формулу (5.6):

$$T_{\text{р}}=\frac{\tau}{\psi^2}\log\psi. \tag{5.12}$$

Это время переработки информации видеокомпьютером при экстраполяции положения объекта является оптимальным. При меньшем времени потребуется неразумно большее быстродействие видеокомпьютера, при большем времени будут возникать существенные ошибки экстраполяции положения объектов и точность телевизионного робота будет низкой по вине его информационно-решающей системы. Отметим, что $T_{\text{р}}$ из (5.12) является неким приближением к эквивалентной ширине T_B импульсной характеристики фильтра Винера для рассмотренного сигнала с экспоненциальной корреляционной функцией:

$$T_B=\tau/\sqrt{\psi^2+1}\approx\tau/\psi.$$

Грубо говоря, формулы (5.11) и (5.12) для быстродействия видеокомпьютера и времени принятия решения показывают, что весь информационный контур обратной связи телевизионного робота, включая объединенные каналами связи его информационно-решающую и управляющую подсистемы, должен иметь малую инерционность (постоянную времени). Она в соответствии с правилом суммирования винеровских ошибок и критерием минимума суммарных информационных ошибок должна распределяться поровну между этими подсистемами. Требуемое быстродействие растет вместе с потенциально достижимой точностью измерения положения объектов.

5.5. ПЕРСПЕКТИВЫ РАЗВИТИЯ

Для правильного понимания перспектив развития телевизионных роботов важно сравнить их возможности с возможностями человека. Следует признать, что простейшие функции выполняются рукотворными системами лучше, чем человеком. Так, эффективность преобразования электромагнитного поля в код в телевизионных камерах может быть сделана значительно лучшей, чем у органов зрения человека (в силу больших диапазонов длин волн и времени накопления изображений). Однако *истинные трудности робототехники не в сборе информации, а в ее интерпретации.* Не переоценивая возможностей робота, следует помнить, что главной задачей кибернетики считается создание искусственного интеллекта, элементы которого являются отличительной чертой телевизионных роботов третьего поколения. Поэтому неизбежно сравнение робота с человеком не только по скорости принятия решений, но и по структуре и методам принятия решений.

Сравнение мозга человека с ЭВМ весьма распространено (см., например, [49]). Так, структура видеокомпьютера робота анало-

гично мозгу может содержать искусственную нейронную сеть, построенную на ЗУ с произвольной выборкой. Телевизионный робот (манипуляционного типа) в целом называют часто системой «глаз—рука». Вместе с тем этот упрощенный двухзвенный образ призван отобразить систему, состоящую из четырех подсистем — телевизионной, классификатора, управляющей подсистемы и исполнительного механизма. Поэтому развернутый антропоморфный образ телевизионного робота, вероятно, должен выглядеть как «глаз — головной мозг — спинной мозг — рука». В связи с сопоставлением мозга человека и робота ставится вопрос о целесообразности придания роботу не только левого, но и правого полушария [100]. При этом отправным моментом рассуждения считается различие функций этих полушарий человеческого мозга. Распространена версия, что это различие состоит в том, что левое полушарие рационально-логическое, а правое реализует мышление на уровне чувственных образов. Именно такая трактовка функций правого полушария мозга позволяет не только ставить вопрос о его необходимости роботу, но и выдвигать как центральную проблему, ждущую своего разрешения в теории искусственного интеллекта, проблему выявления метапроцедур правого полушария и увязки их действия с метапроцедурами левого полушария.

Вместе с тем в настоящее время укрепляется уверенность в том, что правое полушарие человеческого мозга помимо метапроцедур, к ярким представителям которых относится интуиция, осуществляет и ряд более простых алгоритмов, среди которых выделим на наш взгляд важнейшие для приложения к робототехнике.

Во-первых, что обязанность именно правого полушария бороться с самыми различными помехами — «зрительным шумом», т. е. фоновой помехой (см. 4.1). «Настоящий» шум — фотонный — подавляется, насколько возможно, непосредственно в фотоприемнике; для биологической системы — это сетчатка глаза. Во-вторых, это способность к экстраполяции (прогнозу) на основании зрительных образов. Экстраполяция изменений сигнала телевизионным роботом — это необходимая процедура обеспечения надежности его работы. Попытка отказа от экстраполяции изменения фона во времени осложняет, а в некоторых случаях и вовсе исключает возможность работы в реальном масштабе времени — в классификаторе изображений возникают либо неприемлемые ошибки, либо неприемлемые задержки. В частности, экстраполяция фона на один или несколько телевизионных кадров вперед характерна для алгоритмов выделения подвижных объектов на подвижном фоне. Для простых движений фона (линейного смещения) и при малой внутренней динамике — отмеченные две функции правого полушария мозга (борьба с фоновой помехой, использующей экстраполяцию) уже выполняются классификатором изображений [77]. Отметим, что эти две элементарные — для мозга — процедуры требуют существенных алгоритмических

и аппаратурных затрат. Изложенное позволяет сказать, что на вопрос: «Нужно ли роботу правое полушарие?» — следует дать утвердительный ответ. При этом можно считать, что простейшие функции правого полушария уже формализуются и моделируются.

Классификаторы изображений, строящиеся по теории несемантической информации и решений, не могут ставить задачу опознавания семантических образов, которую зрительный анализатор решает в повседневной жизни. Но *по несемантическим показателям функционирования телевизионные роботы могут и должны превзойти человека.* К таким показателям относят *скорость принятия решений* и точность оценки пространственно-временных координат при приеме электромагнитного поля во всем диапазоне длин волн.

Благодаря усилиям разработчиков интеллектуальные возможности роботов расширяются, но не настолько, чтобы он стал умнее «творца». Робот создается для конкретных условий, для того же, чтобы понять неожиданное, нужен человек. По образному выражению Н. Винера «окончательной моделью кошки может быть только кошка». Сколь ни совершенен робот, человек увидит больше.

В области программно-алгоритмического обеспечения телевизионных роботов усилия будут направлены на создание программных средств для объединения телевизионных и решающих систем и языков для параллельной переработки информации [98]. В будущем половина роботов будет иметь возможность автономного программирования, причем возможно что уже к 2000 году будет широко использоваться стандартный язык аналитического программирования роботов, свыше половины которых будут телевизионными [98], т. е. перестанут работать «на ощупь» и обретут зрение.

ЗАКЛЮЧЕНИЕ

Читатель, ознакомившись с данной книгой и с книгами [1, 2], легко обнаружит, что все они едины в своем отражении этапов развития отечественного прикладного телевидения: малокадровое телевидение, твердотельное телевидение и видеоинформатика. Малокадровое телевидение явилось первым серьезным шагом в отходе от «оков» стандарта вещательного телевидения в целях реализации в инженерной практике сужения полосы видеочастот. Это позволило не только увеличить дальность передачи в телевидении, но и открыло первые возможности применения полупроводниковых аналого-цифровых преобразователей и обработки сигналов изображений от натурных сцен в реальном масштабе времени с помощью ЭВМ [1].

Твердотельное телевидение явилось следующим этапом в углублении процесса компьютеризации, охватившего телевидение, открыв возможность создания на базе единой кремниевой технологии нового класса систем «от света до принятия решения».

Впервые в истории телевидения в этих системах не надо было восстанавливать изображение для восприятия глазом, так как компьютер, а не человек принимал решение по информации в самом сигнале. Это сразу устранило дискуссию о применимости квадратичного критерия точности в проектировании систем. «Телевидение без телевизора» — так можно было бы назвать этот новый класс систем. Поэтому понятно желание авторов книги [2] рассматривать твердотельное телевидение как видеоинформатику. Но, если книга [2] заканчивалась словами о видеоинформатике, то данная книга вся подчинена этой теме.

Слово «информатика», как раньше кибернетика, имеет слишком широкий смысл и потому лишена математического единства. Видеоинформатика рассмотрена в книге как более узкая дисциплина, интересы которой представлены телевизионными роботами. Но и в таком аспекте видеоинформатика охватывает телевизионную и вычислительную технику в контуре управления. Поэтому достижение методологического единства является необходимым условием развития видеоинформатики. Оно возможно на основе творческого применения теории информации и ее новой ветви — теории алгоритмов [3, 4]. К сожалению, телевидение, впрочем как и оптика и радиолокация, не располагало опытом успешного приложения теории информации, видимо, как из-за разобщенности между математиками и инженерами, так и в силу застандартизованности вещательного телевидения.

Что изменилось к лучшему в этом отношении? Прежде всего, доказана теорема об идеальном линейном кодировании (гл. 2)*. Согласно теореме способ оптимального линейного кодирования с передаточным фильтром (а не фильтром Костаса) обеспечивает минимальную ошибку, являющуюся решением уравнения идеального кодирования Шеннона в частном гауссовском случае при единичном отношении длин блоков (т. е. $n_1=n_2=1$). Это означает, что такой способ кодирования является в гауссовском случае не просто оптимальным, но и абсолютно оптимальным, т. е. является лучшим среди линейных и нелинейных способов кодирования при условии выделения на один кодовый символ источника одного кодового символа канала ($n_1=n_2$). Теорема дает весьма простую конструкцию идеального кода Шеннона с использованием линейных фильтров, хотя только для частного гауссовского сигнала. Это устраняет, пусть хоть и в частном случае, страх инженеров перед чрезмерной сложностью и неконструктивностью «математических» кодера и декодера в теореме Шеннона. Оптимальный передаточный фильтр, составляющий основу идеального линейного кодирования, раскрывает конструктивно смысл концепции Колмогорова о конечном числе степеней свободы для случайных функций, изложенную при выводе формулы для ε-энтропии [4]. Тем самым решается проблема оптимальной по квадратичному критерию точности дискретизации с учетом входного шума и шума канала непрерывных сообщений, типичным примером которых служит макроскопически непрерывные оптические изображения.

* Успех поиска идеального линейного кодирования не в последнюю очередь объясняется тем, что коллектив авторов статей [22—24] состоял из представителей отраслевой науки и математиков: специалиста по математической статистике А. Б. Ковригина и его ученика А. В. Мартынихина.

Другой важной проблемой является повышение точности оценки зашумленных непрерывных сообщений (изображений) ценой экономного увеличения пропускной способности канала. Новый аспект этой проблемы вызван обнаружением эффекта расщепления ε-энтропии зашумленного источника и объяснением его с помощью понятия о шумовой информации, которое определено как средняя взаимная информация между шумом источника и шумом на выходе канала. Ранее Р. Л. Добрушин и Б. С. Цыбаков показали, что зашумленный источник может характеризоваться в гауссовском случае одной ε-энтропией (кодовой), вычисляемой по формуле Колмогорова путем замены полезного сигнала на условное математическое ожидание [19]. Однако влияние шума источника не ограничивается такой заменой. Было обнаружено, что ε-энтропия (кодовая) зашумленного источника является суммой полезной и шумовой информации, в чем и заключается эффект расщепления ε-энтропии.

Соотношение между полезной и шумовой информацией в кодовой ε-энтропии зависит от задаваемой ошибки передачи по каналу или от задаваемой потери полезной информации. Передача по каналу, которой сопутствует большое количество шумовой информации, вредна не только из-за неэффективности расхода пропускной способности, но и тем, что может вызвать перегрузку видеокомпьютера и отказ от принятия решения. Наличие шумовой информации делает невозможной полную адекватность переданной по каналу информации объекту наблюдения. В рамках принципа идеального линейного (безблокового) кодирования возможен обмен между потерей полезной информации и шумовой информацией по закону, отражаемому информационным соотношением неопределенности. Этим соотношением необходимо пользоваться при выборе между стремлением уменьшить потерю полезной информации и желанием не загружать канал и получателя шумовой информацией. Нахождение такого компромисса обеспечивает предложенный критерий минимума суммарной информационной ошибки, позволяющий в многокаскадной системе передачи информации обоснованно распределить среднеквадратическую ошибку по каскадам. При равных ценах потери полезной информации и шумовой информации для всех каналов межкаскадных связей оптимальное с точки зрения этого критерия распределение ошибки ведет к равномерному по каскадам распределению потери полезной информации.

Теория линейного кодирования с учетом шумовой информации применена к первичному многокаскадному преобразованию света в электрический ток. Хотя в телевидении давно применяют

линейные фильтры для моделирования процессов оптоэлектронного преобразования, но при этом считают фильтры заданными, что отражает положение в вещательных камерах на видиконах. Появление ПЗС открыло возможность автоматического управления процессами накопления и считывания, т. е. изменения фильтров, моделирующих эти процессы (см. гл. 3). В этом случае линейное многокаскадное кодирование корректно описывает преобразование света в сигнал. Такой подход вооружает проектировщиков телекамер на ПЗС теорией линейного кодирования. Открывается возможность инженерного применения простых информационных формул и теоремы об идеальном линейном кодировании.

Оптимизация кодирования в телевизионной камере затруднена разнообразием и непредсказуемостью наблюдаемых оптических сигналов. Преодоление этой трудности возможно только при быстрой адаптации параметров кодирования в соответствии с изменениями оптического сигнала. Компьютеризация твердотельного телевидения выдвинула два основных требования к адаптивному кодированию телевизионных камер — выделение быстрых управляющих параметрами кодирования сигналов и достижение широких диапазонов перестройки параметров кодирования.

В настоящее время этим требованиям способны удовлетворить твердотельные телевизионные камеры на ПЗС-фотоприемниках. Они позволяют организовать дополнительные параллельные каналы неразрушающего измерения параметров зарядового изображения. Параллельное считывание и формирование управляющих кодированием сигналов осуществимо практически без задержек. Управление параметрами ПЗС-камеры достигается в диапазонах более широких, чем у других типов камер.

Кодирование в процессе накопления осуществляется во всех элементах изображения одинаково и является главным средством борьбы с шумом фотоэлектронов. Кодирование зарядового изображения в процессе считывания является средством борьбы с шумом считывания ПЗС. Ошибки передачи по каналу считывания в соответствии с теоремой Шеннона определяются пропускной способностью канала считывания и соотношением числа букв в канале к числу букв в источнике. Это ведет к существованию двух инженерных путей уменьшения ошибок передачи накопленного зарядового изображения. Во-первых, снижение шума считывания возможно при одновременном уменьшении тактовой частоты и фильтрации нижних частот выходного сигнала. Во-вторых, возможно многократное обращение к каналу при считывании одного и того же элемента. Этот метод может быть реализован как в известной структуре распределенного усилителя зарядов, так и на обычном устройстве с плавающим затвором.

Переработка информации в видеоинформатике ведется в несколько этапов. Важнейшие среди них — камерное кодирование и классификация изображений. Эти два взаимосвязанных этапа обеспечивают радикальное сжатие, осуществляя преобразование

изображений от света до решения, т. е. от числа фотонов до числа решений. В информационно-решающих системах это сжатие распределяется поровну между телекамерой и видеокомпьютером. При общей решающей способности более 10^4 решений в секунду (до ста за кадр) сжатие достигает в каждом из этих блоков 10^4. Отклонение от такого баланса ведет либо к увеличению ошибок классификации при более сильном сжатии в телекамере (грубой дискретизации), либо к неоправданному росту быстродействия и сложности видеокомпьютера. Важным способом сжатия изображений в телекамере является кодирование в процессе считывания зарядов с переменной четкостью для полезного объекта и фона (контекста).

Для компьютера сжатие обеспечивается подавлением фона, переходом в пространство признаков объектов и принятием решений. Если для камерного кодирования характерны адаптация накопления и считывания, то для компьютерной переработки изображений (см. гл. 4) характерны методы сокращения вычислений, линейные и нелинейные методы разделения сигналов и фонов и классификация изображений. Сокращение вычислений опирается как на теорию информации и алгоритмов, так и на эвристические методы приближенных вычислений. Высокая эффективность линейных методов при переработке изображений позволяет развивать унификацию видеопроцессоров на базе транспьютерной методологии. Линейные методы позволяют дополнить камерное кодирование в тех важных случаях, когда необходима коррекция внешних искажений изображения, например при его смазе из-за движения объектов или турбулентности атмосферы. Классификация или распознавание изображений выдвигает требования инвариантности решения к масштабированию, повороту и сдвигу изображений.

Важным практическим приложением видеоинформатики являются телевизионные роботы (см. гл. 5). Главной особенностью их функционирования по сравнению с другими видеоинформационными системами является формирование поведения, т. е. реакции на изображения и их изменение в реальном масштабе времени. Существенная трудность в построении классификаторов изображений телевизионных роботов обусловлена наличием сложного пространственно неоднородного фона и изменениями сигнала фона объектов управления (масштаб, ракурс, сдвиг) при движении робота или его манипулятора с закрепленной на нем телевизионной камерой.

Телевизионным роботам доступна классификация лишь простых изображений, но в этой задаче они могут достичь решающей способности более 10^4 решений за секунду, что превышает возможности человека на три порядка. Создание таких быстродействующих роботов иллюстрирует широкие возможности видеоинформатики.

СПИСОК ЛИТЕРАТУРЫ

1. **Брацлавец П. Ф., Росселевич И. А., Хромов Л. И.** Космическое телевидение. — М.: Связь, 1973. — 248 с.
2. **Твердотельное** телевидение/Л. И. Хромов, Н. В. Лебедев, А. К. Цыцулин, А. Н. Куликов. — М.: Радио и связь, 1986. — 184 с.
3. **Шеннон К.** Работы по теории информации и кибернетике/Пер. с англ. под ред. Р. Л. Добрушина, О. Б. Лупанова. — М.: ИЛ, 1963. — 830 с.
4. **Колмогоров А. Н.** Теория информации и теория алгоритмов. — М.: Наука, 1987. — 304 с.
5. **Кибернетика.** Дела практические: Сб. статей. — М.: Наука, 1984. — 190 с.
6. **Кибернетика.** Становление информатики: Сб. статей. — М.: Наука, 1986. — 192 с.
7. **Информатика** и научно-технический прогресс: Сб. статей. — М.: Наука, 1987. — 190 с.
8. **Ван Трис Г.** Теория обнаружения, оценок и модуляции: В 3 т./Пер. с англ. под ред. В. И. Тихонова и В. Т. Горяинова. — М.: Сов. радио, 1972—1977. — Т. 1—3.
9. **Вудворд Ф. М.** Теория вероятностей и теория информации с применениями в радиолокации/Пер. с англ. под ред. Г. С. Горелика. — М.: Сов. радио, 1955. — 128 с.
10. **Costas J. P.** Coding with linear systems//Proc. IRE. — 1952. — Vol. 40. — P. 1101—1103.
11. **Колесник В. Д., Полтырев Г. Ш.** Курс теории информации. — М.: Наука, 1982. — 416 с.
12. **Вальд А.** Статистические решающие функции//Позиционные игры: Сб. статей/ Пер. с англ. под ред. Н. И. Воробьева и И. Н. Врублевской. — М.: Наука, 1963. — С. 300—522.
13. **Хартли Р.** Передача информации//Теория информации и ее приложения/ Пер. с англ. под ред. А. А. Харкевича. — М.: Физматгиз, 1959. — С. 5—35.
14. **Френкс Л.** Теория сигналов/Пер. с англ. под ред. Д. Е. Вакмана. — М.: Сов. радио, 1974. — 344 с.
15. **Мартынихин А. В., Хромов Л. И.** Оптимальные фильтры с конечным числом степеней свободы в прикладном телевидении//ТСС ТТ[1]. — 1982. — Вып. 1. — С. 3—10.
16. **Цыбаков Б. С.** Теория передачи непрерывных сообщений по непрерывным каналам: Дис. ... докт. техн. наук. — М., 1970. — 399 с.
17. **Маригодов В. К.** Помехоустойчивая обработка информации: методы оптимального линейного предискажения и коррекции. — М.: Наука, 1983. — 202 с.
18. **Галлагер Р.** Теория информации и надежная связь/Пер. с англ. под ред. М. С. Пинскера, Б. С. Цыбакова. — М.: Сов. радио, 1974. — 719 с.
19. **Добрушин Р. Л., Цыбаков Б. С.** Передача информации с дополнительным шумом//Проблемы передачи информации. — 1963. — Вып. 14. — С. 21—42.
20. **Эксперименты** с машинным зрением/В. П. Андреев, Д. А. Белов, Г. Г. Вайнштейн, Е. А. Москвина. — М.: Наука, 1987. — 128 с.
21. **Хромов Л. И., Литвинчук Л. А.** Оценка пропускной способности телевизионного канала с учетом линейных искажений и двух источников шумов// ТСС ТТ. — 1981. — Вып. 1. — С. 3—10.
22. **Теория** кодирования непрерывного сигнала и ее применение в прикладном телевидении/Л. И. Хромов, А. Б. Ковригин, А. В. Мартынихин, А. К. Цыцулин//ТСС ТТ. — 1985. — Вып. 1. — С. 3—11.
23. **Применение** теоремы Шеннона в прикладном телевидении/А. Б. Ковригин, Л. И. Хромов, А. В. Мартынихин, А. К. Цыцулин//ТСС ТТ. — 1986. — **Вып. 2. — С. 30—39.**
24. **Идеальное** линейное кодирование телевизионных сигналов/Л. И. Хромов, А. Б. Ковригин, А. В. Мартынихин, А. К. Цыцулин//ТСС ТТ. — 1987. — Вып. 4. — С. 3—10.
25. **Витерби А. Д., Омура Дж. К.** Принципы цифровой связи и кодирования/ Пер. с англ. под ред. К. Ш. Зигангирова. — М.: Радио и связь, 1982. — 536 с.
26. **Роуз А.** Зрение человека и электронное зрение/Пер. с англ. под ред. В. С. Вавилова. — М.: Мир, 1977. — 216 с.
27. **Бринкен Н. О., Хромов Л. И.** Влияние параметрической неопределенности фона на структуру оптимальной телевизионной камеры, использующей идеальный многоэлементный фотоприемник//ТСС ТТ. — 1979. — Вып. 1. — С. 3—9.
28. **Левит А. Б.** Введение в общую теорию телевидения. — М.: Сов. радио, 1967. — 116 с.
29. **Лебедев Д. С., Цуккерман И. И.** Телевидение и теория информации. — Л.: Энергия, Ленингр. отд-ние, 1965. — 219 с.
30. **Ковригин А. Б., Мартынихин А. В., Хромов Л. И.** Теория кодирования в реальном масштабе времени//СС ТТ. — 1990. — Вып. 7. — С. 3—12.
31. **Ахмед Н., Рао К. Р.** Ортогональные преобразования при обработке цифровых сигналов/Пер. с англ. под ред. И. Б. Фоменко. — М.: Связь, 1980. — 248 с.
32. **Цыцулин А. К.** Построение телевизионных систем для наблюдения подвижных изображений//ТСС ТТ. — 1988. — Вып. 4. — С. 3—11.
33. **Ляпунов А. А.** Проблемы теоретической и прикладной кибернетики. — М.: Наука, 1980. — 336 с.
34. **Секен К., Томпсет М.** Приборы с переносом заряда/Пер. с англ. под ред. В. В. Поспелова, Р. А. Суриса. — М.: Мир, 1978. — 327 с.
35. **Yates G. J., Myers D. R.** Characterization of Elektrooptic Anomales Associated with Transient of Fast Readout Charge-coupled Devices//Optical Engineering. — 1987. — Vol. 26(8). — P. 747—756.
36. **Sadowski.** Signal Processing for Time Delay and Integrated Charge-coupled Device (TDI—CCD) in the Panoramic Scan Mode//Proc. SPIE. — 1981. — Vol. 282. — P. 115—128.
37. **А. с.** 1327323 СССР, МКИ[3]H04N 5/235. Адаптивная телевизионная камера на матрице приборов с зарядовой связью/А. Н. Куликов, А. В. Мартынихин, Л. И. Хромов, А. К. Цыцулин. — Опубл. 1987, Бюл. № 28.
38. **Зайцев В. П., Крылков В. Ф., Куликов А. Н.** Телевизионная система с перестраиваемой частотой кадров//ТСС ТТ. — 1985. — Вып. 5. — С. 11—18.
39. **Телекамеры** на матрице ПЗС с расширенным динамическим диапазоном/ В. А. Зеленов, А. Н., Куликов, А. Н. Марков, С. Н. Нарышкин//ТСС ТТ.— 1986. — Вып. 4. — С. 31—36.
40. **Куликов А. Н.** Оценка диапазонов перестройки параметров разложения в телекамере на матрице ПЗС//ТСС ТТ. — 1986. — Вып. 4. — С. 47—54.
41. **Быстродействующая** телекамера на матричном ПЗС с поверхностным каналом/М. В. Клебанов, А. Н. Куликов, А. Н. Марков, А. К. Цыцулин.//ТСС ТТ. — 1986. — Вып. 5. — С. 12—18.
42. **Пресс Ф. П.** Формирователи видеосигнала на приборах с зарядовой связью. — М.: Радио и связь, 1981. — 136 с.
43. **Довжиков Д. А., Куликов А. Н., Филиппов В. Н.** Оптимизация режима работы матричного ПЗС при низких условиях освещенности//ТСС ТТ. — 1983. — Вып. 5. — С. 38—45.
44. **Ковригин А. Б., Куликов А. Н., Мартынихин А. В.** Адаптация времени кадра в телевизионных камерах на ПЗС//ТСС ТТ. — 1986. — Вып. 2. — С. 17—22.

[1] Здесь и далее ТСС ТТ — Техника средств связи. Сер. Техника телевидения.

45. **Нестационарные** помехи в телевизионных камерах на ПЗС/Д. А. Довжиков, М. В. Клебанов, А. Н. Куликов, А. Н. Марков//ТСС ТТ. — 1987. — Вып. 4. — С. 23—28.
46. **А. с.** 1417210 СССР, МКИ³H04*N* 5/228. Способ формирования сигнала управления чувствительностью телевизионной камеры на матрице ПЗС/А. Н. Куликов, Л. И. Хромов. — Опубл. 1988, Бюл. № 12.
47. **Думаревский Ю. Д., Ковтонюк Н. Ф., Савин А. И.** Преобразования изображений в структурах полупроводник-диэлектрик. — М.: Наука, 1987. — 176 с.
48. **Титце У., Шенк К.** Полупроводниковая схемотехника/Пер. с англ. под ред. А. Г. Алексенко. — М.: Мир, 1982. — 512 с.
49. **Дж. фон Нейман.** Теория самовоспроизводящихся автоматов/Пер. с англ. под ред. В. И. Варшавского. — М.: Мир, 1971. — 382 с.
50. **Винер Н.** Кибернетика или управление и связь в животном и машине/Пер. с англ. под ред. Г. Н. Поварова. — М.: Наука, 1983. — 343 с.
51. **Маккеллан Дж. Х., Рейдер Ч. М.** Применение теории чисел в цифровой обработке сигналов/Пер. с англ. под ред. Ю. И. Манина. — М.: Радио и связь, 1983. — 264 с.
52. **Ульман Ш.** Принципы восприятия подвижных объектов/Пер. с англ. под ред. В. Л. Стефанюка. — М.: Радио и связь, 1983. — 167 с.
53. **СБИС** для распознавания образов и обработки изображений: Сб. статей/Под ред. К. Фу; Пер. с англ. под ред. В. А. Абрамова. — М.: Мир, 1988. — 247 с.
54. **Василенко Г. И., Тараторин А. М.** Восстановление изображений. — М.: Радио и связь, 1986. — 302 с.
55. **Обработка** изображений и цифровая фильтрация: Сб. статей/Под ред. Т. Хуанга; Пер. с англ. под ред. Е. З. Сороки, В. А. Хлебородова. — М.: Мир, 1979. — 318 с.
56. **Применение** цифровой обработки сигналов: Сб. статей/Под ред. Э. Оппенгейма; Пер. с англ. под ред. А. М. Рязанцева. — М.: Мир, 1980. — 552 с.
57. **Грудзинский М. А., Захаров Ю. И., Титков Б. В.** Полиномиальный вариант китайской теоремы об остатках для вычисления быстрой дискретной циклической свертки//ТСС ТТ. — 1987. — Вып. 7. — С. 44—50.
58. **Ярославский Л. П.** Цифровая обработка сигналов в оптике и голографии: Введение в цифровую оптику. — М.: Радио и связь, 1987. — 296 с.
59. **Прэтт У.** Цифровая обработка изображений: В 2 т./Пер. с англ. под ред. Д. С. Лебедева. — М.: Мир, 1982. — 2 т.
60. **Методы** передачи изображений. Сокращение избыточности: Сб. статей/Под ред. Прэтта У. К.; пер. с англ. под ред. Л. С. Виленчика. — М.: Радио и связь, 1983. — 264 с.
61. **Величкин А. И.** Передача аналоговых сообщений по цифровым каналам связи. — М.: Радио и связь, 1983. — 240 с.
62. **Быстрые** алгоритмы в цифровой обработке изображений: Сб. статей/Под ред. Т. С. Хуанга; пер. с англ. под ред. Л. П. Ярославского. — М.: Радио и связь, 1984. — 221 с.
63. **Цифровое** кодирование телевизионных изображений/И. И. Цуккерман, Б. М. Кац, Д. С. Лебедев и др.; под ред. И. И. Цуккермана. — М.: Радио и связь, 1981. — 239 с.
64. **М. И. Кривошеев, Л. С. Виленчик, И. Н. Красносельский.** Цифровое телевидение/Под ред. М. И. Кривошеева. — М.: Связь, 1980. — 263 с.
65. **Пенин П. И.** Системы передачи цифровой информации. — М.: Сов. радио, 1976. — 366 с.
66. **Селиванов А. С.** Космос — Миру. Телевизионные системы исследования планет. — М.: Знание, 1990. — 64 с.
67. **Голушко М. Н., Кулешов В. Н., Цыцулин А. К.** Блок цифровой межкадровой обработки для малокадрового телевидения/ТСС ТТ. — 1979. — Вып. 4. — С. 40—47.
68. **Кузьмин С. З.** Основы проектирования систем цифровой обработки радиолокационной информации. — М.: Радио и связь, 1986. — 352 с.
69. **Зворыкин В. К., Мортон Д. А.** Телевидение/Пер. с англ. под ред. С. И. Катаева. — М.: ИЛ, 1956. — 780 с.

70. **А. с.** 974602 СССР, МКИ³H04*N* 7/02. Устройство совмещения телевизионных изображений/А. К. Цыцулин, А. М. Захаров, М. Н. Голушко и др. — Опубл. 1982, Бюл. № 42.
71. **Зубарев Ю. Б., Сагдулаев Ю. С.** Спектрозональная селекция оптических изображений. — Ташкент: Фан, 1987. — 108 с.
72. **Бакут П. А., Жулина Ю. В., Иванчук Н. А.** Обнаружение движущихся объектов/Под ред. П. А. Бакута. — М.: Сов. радио, 1980. — 288 с.
73. **Горелик Д. Л., Гуревич И. Б., Скрипкин В. А.** Современное состояние проблемы распознавания. — М.: Радио и связь, 1985. — 160 с.
74. **Распознавание** образов: состояние и перспективы/К. Верхаген, Р. Дейн, Ф. Грун и др.; пер. с англ. под ред. И. Е. Гуревича. — М.: Радио и связь, 1985. — 104 с.
75. **Успенский В. А., Семенов А. Л.** Теория алгоритмов: основные открытия и приложения. — М.: Наука, 1987. — 288 с.
76. **Дуда Р., Харт П.** Распознавание образов и анализ сцен/Пер. с англ. под ред. В. Л. Стефанюка. — М.: Мир, 1976. — 512 с.
77. **Цыцулин А. К., Голушко М. Н., Мартынихин А. К.** Различение изображений в условиях неинвариантности к сдвигу//ТСС ТТ. — 1987. — Вып. 1. — С. 24—32.
78. **Хромов Л. И., Цыцулин А. К.** Анализ процесса дискретизации сигналов в некоторых системах прикладного телевидения//ТСС ТТ. — 1980. — Вып. 1. — С. 18—26.
79. **Интерполяция** дискретного видеосигнала телевизионной камерой на матричном фотоприемнике/Н. О. Бринкен, М. Н. Голушко, В. Н. Макаров, А. К. Цыцулин//ТСС ТТ. — 1980. — Вып. 1. — С. 27—34.
80. **Голушко М. Н., Макаров В. Н.** Специфика применения теории решений в научно-прикладном телевидении//ТСС ТТ. — 1985. — Вып. 4. — С. 18—25.
81. **Андерсон Т.** Введение в многомерный статистический анализ/Пер. с англ. под ред. Б. В. Гнеденко. — М.: Физматгиз, 1963.
82. **Harada H. a. o.** A Swing CCD Image Sensor//J. Inst. Television Engenering of Japan. — 1983. — Vol. 37, N 10. — P. 826—832.
83. **Сизов В. П.** Синтез оптимальных цифровых телевизионных следящих систем//ТСС ТТ. — 1978. — Вып. 1. — С. 13—22.
84. **Тихонов В. И.** Оптимальный прием сигналов. — М.: Радио и связь, 1983. — 320 с.
85. **Кавалеров Г. И., Мандельштам С. М.** Введение в информационную теорию измерений. — М.: Энергия, 1974. — 375 с.
86. **Компьютерная** оптика: Сб. статей/ Под ред. Е. П. Велихова и А. М. Прохорова. — Вып. 1. — М.: ИППИ — ИОФАН, 1987. — 192 с.
87. **Телевизионная** астрономия/А. Н. Абраменко, Е. С. Агеев, В. Ф. Анисимов и др.; Под ред. В. Б. Никонова. — М.: Наука, 1984. — 272 с.
88. **Техническое** зрение роботов: Сб. статей/Под ред. А. Пью; пер. с англ. под ред. Г. П. Катыса. — М.: Машиностроение, 1987. — 320 с.
89. **Микропроцессорные** системы управления в робототехнике: Сб. статей/Под ред. И. М. Макарова, Д. Е. Охоцимского, Е. П. Попова. — М.: Наука, 1984. — 176 с.
90. **Системы** очувствления и адаптивные промышленные роботы: Сб. статей/ Под ред. Е. П. Попова, В. В. Клюева. — М.: Машиностроение, 1985. — 255 с.
91. **Белянин П. Н.** Робототехнические системы для машиностроения. — М.: Машиностроение, 1986. — 254 с.
92. **Жаботинский Ю. Д., Исаев Ю. В.** Адаптивные промышленные роботы и их применение в микроэлектронике. — М.: Радио и связь, 1985. — 105 с.
93. **Куафе Ф.** Взаимодействие робота с внешней средой: Пер. с фр. — М.: Мир, 1985. — 287 с.
94. **Не** счесть у робота профессий/Под ред. П. Марша; пер. с англ. под ред. В. С. Гурфинкеля. — М.: Мир, 1987. — 183 с.
95. **Проблемы создания телевизионных роботов**/М. А. Грудзинский, В. Ф. Крылков, Л. И. Хромов, А. К. Цыцулин//ТСС ТТ. — 1987. — Вып. 2. — С. 3—10.
96. **Тимофеев А. В.** Управление роботами. — Л.: ЛГУ, 1986. — 240 с.

97. **Орбитальная** астрофизическая обсерватория «Астрон». Итоги 5 лет работы/А. А. Боярчук, Е. Д. Гернет, Р. Е. Гершберг и др.//Космические исследования. — 1988. — Т. 26, № 6. — С. 917—933.
98. **Porter A. L., Rossini F. A.** Robotics in the Year 2000//Robotic Today. — 1987. — Vol. 9, N 3. — P. 27—28.
99. **Криницкий Н. А.** Алгоритмы и роботы. — М.: Радио и связь, 1983. — 168 с.
100. **Поспелов Д. А.** Фантазия или наука: на пути к искусственному интеллекту. — М.: Наука, 1982. — 224 с.
101. **Горбунов А. К., Пинскер М. С.** Эпсилон-энтропия с задержкой при малой среднеквадратической ошибке воспроизведения//Проблемы передачи информации. — 1987. — Т. 23, № 2. — С. 3—8.
102. **Теория** моделей в процессах управления/Б. Н. Петров, Г. М. Уланов, И. И. Гольденберг, С. В. Ульянов — М.: Наука, 1987. — 223 с.
103. **Атмосферная** нестабильность и адаптивный телескоп: Сб. статей./Под ред. А. Х. Курмаева, В. С. Шевченко. — Л.: Наука, Ленингр. отд-ние, 1988. — 145 с.
104. **Хромов Л. И., Цыцулин А. К.** Теория информации и телевизионные роботы//ТСС ТТ. — 1989. — Вып. 4. — С. 25—32.
105. **Переменная** четкость в твердотельном телевидении/С. А. Иванов, А. Н. Куликов, А. К. Цыцулин и др.//ТСС ТТ. — 1989. — Вып. 7. — С. 34—40.
106. **Ханцеверов Ф. В., Остроухов В. В.** Моделирование космических систем изучения природных ресурсов Земли. — М.: Машиностроение, 1989. — 264 с.
107. **Телевизионная** съемка кометы Галлея/Г. А. Аванесов, Я. Л. Зиман, В. И. Тарнопольский и др.; Под ред. Р. З. Сагдеева. — М.: Наука, 1989.— 296 с.
108. **Александров А. М., Цыцулин А. К.** Ограничения в задаче классификации бинарных изображений на дискретном растре.//ТСС ТТ. — 1990.—Вып. 7.— С. 73—81.
109. **Дополнительные** каналы считывания в матричных ПЗС/А. Н. Куликов, С. А. Иванов, А. К. Цыцулин и др.//ТСС ТТ. — 1990. — Вып. 2. — С. 34—39.
110. **Ллойд Л.** Системы тепловидения/Пер. с англ. под ред. А. И. Горячева.— М.: Мир, 1980. — 416 с.
111. **Кошелев В. Н.** Квантование с минимальной энтропией//Проблемы передачи информации. — 1963. — Вып. 4. — С. 151—156.
112. **Фалькович С. Е.** Оценка параметров сигналов. — М.: Сов. радио, 1970. — 334 с.
113. **Солодовников В. В., Тумаркин В. И.** Теория сложности, системный синтез и автоматизация проектирования//Информатика, управление, вычислительная техника. — 1987. — Вып. 1. — С. 82—133.
114. **Трауб Дж., Васильковский Г., Вожьняковский Х.** Информация, неопределенность, сложность/Пер. с англ. под ред. А. Г. Сухарева. — М.: Мир, 1988. — 184 с.
115. **Вильсон К. Дж.** Задачи науки и техники в свете нового Японского вызова//ТИИЭР. — 1984. — Т. 72, № 1. — С. 162—179.
116. **Павлидис Т.** Алгоритмы машинной графики и обработки изображений/Пер. с англ. под ред. И. Б. Гуревича. — М.: Радио и связь, 1986. — 292 с.
117. **Шарыпин Е. В.** Подход к организации сжатия изображений в персональных системах обработки изображений//Системы и средства информатики.— 1989. — Вып. 1. — С. 93—97.
118. **Иванов С. А., Цыцулин А. К.** Измерение среднеквадратической ошибки оценивания оптических сигналов телекамерой на ПЗС//ТСС ТТ. — 1990. — Вып. 7. — С. 56—63.

СОДЕРЖАНИЕ

Предисловие 3
От авторов 4
Введение 5
В.1. Компьютеризация и видеоинформатика 5
В.2. Некоторые направления развития видеоинформатики . . . 6

Глава 1. Элементы теории фильтрации и теории информации . . . 12

1.1. Спектральный метод описания сигналов и фильтров . . . 12
1.2. Вариационная задача в теории линейной фильтрации . . . 17
1.3. Принцип блокового кодирования 20
1.4. Некоторые задачи развития теории информации 24

Глава 2. Кодирование в реальном времени при квадратическом критерии точности 27

2.1. Линейное кодирование при ограничении на фильтр 28
2.2. Раздвоение ε-энтропии из-за шумовой информации (одномерный случай) 37
2.3. Идеальный код с учетом шумовой информации (одномерный случай) 45
2.4. Сигнальная и шумовая информация гауссовских процессов . . 53
2.5. Теорема об идеальном линейном кодировании 59
2.6. Свойства идеального линейного кодирования 65

Глава 3. Кодирование изображений в реальном времени 71

3.1. Применение теории информации в телевидении 72
3.2. Определение кадровой частоты телекамер 78
3.3. Кодирование в процессе накопления фотонов 84
3.4. Дополнительные каналы считывания в ПЗС-фотоприемнике . . 92
3.5. Быстродействующие управляющие сигналы 96
3.6. Адаптивное кодирование в процессе считывания 102
3.7. Специфика дискретно-аналоговой обработки сигнала . . . 107
3.8. Информационные характеристики телевизионных камер . . . 110

Глава 4. Компьютерная переработка информации с принятием решения в реальном времени 120

4.1. Предварительная переработка видеоинформации 121
4.2. Классификация изображений 127
4.3. Неинвариантность к сдвигу 142
4.4. Измерение параметров сигналов изображений 150
4.5. Сокращение вычислительных затрат 156
4.6. Обработка видеоинформации в диалоге человека с компьютером 161

Глава 5. Телевизионные роботы 163

5.1. Области применения и классификация 164
5.2. Структура, характеристики и особенности робототехнических телевизионных систем 166
5.3. Сложность и быстродействие 170
5.4. Экстраполяция подвижных изображений 173
5.5. Перспективы развития 179
Заключение 181
Список литературы 186

Научное издание

ХРОМОВ ЛЕОНИД ИОСИФОВИЧ, ЦЫЦУЛИН АЛЕКСАНДР КОНСТАНТИНОВИЧ
КУЛИКОВ АЛЕКСАНДР НИКОЛАЕВИЧ

ВИДЕОИНФОРМАТИКА. ПЕРЕДАЧА И КОМПЬЮТЕРНАЯ ОБРАБОТКА ВИДЕОИНФОРМАЦИИ

Заведующая редакцией Г. И. Козырева. Редактор В. И. Ченцова.
Обложка художника А. С. Дзуцева. Художественный редактор Н. С. Шеин.
Техн. редактор З. Н. Ратникова. Корректор Т. В. Дземидович

ИБ № 2391

Сдано в набор 10.05.90 Подписано в печать 18.10.90
Формат 60×90 1/16 Бумага типогр. № 2 Гарнитура литературная Печать высокая
Усл. печ. л. 12,0 Усл. кр-отт. 12,25 Уч.-изд. л. 13,46 Тираж 5000 экз.
Изд. № 22753 Заказ № 51 Цена 3 р. 50 к.
Издательство «Радио и связь». 101000 Москва, Почтамт, а/я 693

Типография изд-ва «Радио и связь». 101000 Москва, ул. Кирова, д. 40