ТВЕРДОТЕЛЬНАЯ РЕВОЛЮЦИЯ В ТЕЛЕВИДЕНИИ

Телевизионные системы на основе приборов с зарядовой связью, систем на кристалле и видеосистем на кристалле

Под редакцией
доктора техн. наук А. А. Умбиталиева и
доктора техн. наук А. К. Цыцулина

«Радио и связь»

2006

ББК 32.94
Т 26
УДК 621.396 (024)

Твердотельная революция в телевидении: Телевизионные системы на основе приборов с зарядовой связью, систем на кристалле и видеосистем на кристалле/ В. В. Березин, А. А. Умбиталиев, Ш. С. Фахми, А. К. Цыцулин, Н. Н. Шипилов; Под ред. А. А. Умбиталиева и А. К. Цыцулина. – М.: Радио и связь, 2006. – 312 с., ил. – **ISBN 5-256-01814-0**

Книга является логическим продолжением выпущенной на 20 лет раньше книги «Твердотельное телевидение» и посвящена второму этапу твердотельной революции в научно-прикладном телевидении, обусловленному рождением новых приборов микроэлектронной техники – систем на кристалле и видеосистем на кристалле, которые благодаря единой технологии КМОП позволили объединить на одном кристалле фотоприёмник, устройства развертки, оцифровки и цифровой обработки видеосигналов. Книга отражает главные тенденции современной техники связи – тотальный переход к «рукотворным кристаллам», т. е. большим интегральным схемам и срастание телевизионной и вычислительной техники. Теоретические основы проектирования сформулированы с позиций системотехники, теории информации и теории оптимального приёма сигналов.

Для инженерно-технических работников, связанных с разработкой информационных систем передачи и обработки изображений; может быть полезна аспирантам и студентам старших курсов компьютерных и радиотехнических специальностей.

ББК 32.94

Рецензенты:
доктор технических наук, профессор **Б. Н. Формозов**
(БГТУ «Военмех» им. Д.Ф.Устинова),
доктор технических наук, профессор **В. Г. Немудров**
(НИИМА «Прогресс»)

ISBN 5-256-01814-0

© В. В. Березин, А. А. Умбиталиев, Ш. С. Фахми, А. К. Цыцулин, Н. Н. Шипилов, 2006

ПРЕДИСЛОВИЕ

Книга является логическим продолжением выпущенной на 20 лет раньше специалистами ВНИИ Телевидения книги «Твердотельное телевидение (телевизионные системы с переменными параметрами на ПЗС и микропроцессорах)» и посвящена второму этапу твердотельной революции в научно-прикладном телевидении. Первый этап характеризовался внедрением рожденных в начале 1970-х годов достижений твердотельной электроники – приборов с зарядовой связью и микропроцессоров. Новый этап твердотельной революции в телевидении обусловлен не столько совершенствованием этих приборов, сколько рождением и внедрением новых приборов микроэлектронной техники – «систем на кристалле» и «видеосистем на кристалле», которые благодаря единой КМОП технологии позволили объединить на одном кристалле фотоприёмник, устройства развёртки, оцифровки и цифровой обработки видеосигналов.

Основной характерной чертой современных тенденций в телевидении является срастание его с компьютерной техникой. Это срастание ранее инициировалось специалистами телевизионной техники и выразилось в рождении твердотельного телевидения как «связки» приборов с зарядовой связью и микропроцессоров. В настоящее время, в период второго этапа твердотельной революции в телевидении, прежние инициативы дополняются усилиями специалистов компьютерной техники, благодаря чему концепция «телекамера – это вынесенная наружу часть компьютера» перестала нуждаться в популяризации. Примером срастания телевизионной и компьютерной техники является данная книга. Так же как и авторы первой книги «Твердотельное телевидение», авторы данной книги тесно связаны с НИИ телевидения, являясь либо его сотрудниками, либо «смежниками». Совсем не случайно в состав авторов вошли специалисты не только по телевизионной, но и по вычислительной технике. Книгу в целом следует рассматривать как важный и своевременный сигнал о необходимости обращения пристального внимания специалистов прикладной телевизионной техники к новому революционному изменению технологии. Причём технологию следует понимать двояко: новая технология создания СБИС дополняется технологией создания (системой автоматизированного проектирования, САПР) систем на них. Заслугой авторов, в частности, является инициатива включе-

ния проблематики «видеосистем на кристалле» в «Программу развития национальной технологической базы». Книга показывает, что состояние отечественной телекоммуникационной отрасли, в которую входят телевизионное вещание и прикладное телевидение, можно охарактеризовать как противоречивое, включающее наличие достаточного интеллектуального задела при необеспеченности отечественной микроэлектронной базой. Важным выводом из книги является необходимость сочетания усилий в области создания систем и в области совершенствования технологии создания их компонентов, которые только в единстве могут сократить отставание России от других государств, снизить объём импорта.

Предлагаемая вниманию читателей книга органически вписывается в русло проводимых Управлением радиоэлектронной промышленности и систем управления ФАП России мероприятий согласно «Стратегии развития электронной промышленности на ближайшие годы» и «Концепции структурных преобразований радиоэлектронной промышленности оборонно-промышленного комплекса России». Одной из ключевых задач указанной стратегии является обеспечение создания и массового производства перспективной электронной продукции, в том числе на основе широкого использования функциональной микроэлектроники. Массовая продукция типа аппаратуры цифрового вещательного и прикладного телевидения должна обеспечить рынок для современного микроэлектронного производства. Уже создаются научные, проектные, технические и технологические центры отрасли, включая «дизайн-центры» по производству СБИС и элементов функциональной микроэлектроники «система на кристалле», и можно прогнозировать появление отечественных разработок «видеосистем на кристалле».

Надеюсь, что знакомство читателей, в первую очередь сегодняшних и будущих производителей радиоэлектронных систем, с достижениями и возможностями нового поколения твердотельной микроэлектроники в прикладном телевидении станет новым импульсом для разработки оригинальных систем и концепций для обеспечения экономической, информационной и военной безопасности страны.

Начальник Управления
радиоэлектронной промышленности и систем управления
Федерального агентства по промышленности,
доктор технических наук *Ю. И. Борисов*

ОТ АВТОРОВ

Данное издание по сути служит продолжением книги «Твердотельное телевидение» [1] и отражает ряд важных изменений в телевидении, в первую очередь прикладном, и опирается в основном на работы авторов. Связь с книгой [1] подчёркивается не только тематикой и акцентом на единение телевизионной и компьютерной техники, но и структурой книги, в которой главы и даже ряд параграфов имеют близкие названия. Конечно, полного совпадения названий не может быть, так как прошедшие со времени издания книги [1] 20 лет ознаменованы не просто эволюционным развитием техники, а привели ко второму этапу твердотельной революции в телевидении. Главным содержанием нового этапа стало появление систем на кристалле и видеосистем на кристалле как проявление резкого усиления срастания телевизионной и компьютерной техники в единое информационное ядро систем связи.

Авторы преследовали цель показать основные тенденции развития прикладного телевидения, обусловленные взаимовлиянием теории и практики, науки и технологии. При этом они исходили из того, что прикладное твердотельное телевидение – уже достаточно зрелая область телекоммуникационной техники, и теоретическим аспектам должно уделяться внимание не меньшее, чем технологическим. Здесь авторы следовали концепции С. Лема, который в книге «Сумма технологии» писал: «Сущностью "эмпирической эпохи" в технологии было не столько отсутствие теории, сколько её вторичность. Сначала возникла паровая машина, а потом термодинамика; сначала самолёт, а потом теория полёта; сначала строили мосты, а потом научились их рассчитывать. Я бы рискнул утверждать даже, что технологическая эмпирика стремится развиваться до тех пор, пока это вообще возможно. Эдисон пытался изобрести что-то вроде "атомного двигателя", но из этого ничего не вышло и не могло выйти: ибо методом "проб и ошибок" можно построить динамо-машину, но атомный реактор – никогда. <…> Следует обратить внимание на различный подход к модели со стороны ученого и со стороны технолога. Технолог, получив возможность "синтеза живого организма" – если такова была его цель, – удовлетворится "конечным продуктом". Учёный – по крайней мере, учёный в классическом понимании – стремится детально изучить "теорию синтеза организмов". Учёный жаждет алгоритма, технолог же скорее походит на садовника, который, сажая дерево и срывая яблоки, не заботится о том, "как яблоня это сделала". Учёный считает

такой узкоутилитарный, прагматический подход прегрешением против канонов полного познания. Нам кажется, что в будущем обе эти позиции изменятся». Указанная С. Лемом вторичность теории в книге [1] проявилась в том, что по инициативе издательства «Радио и связь» раздел, посвящённый теории проектирования, был перемещён с первого места на третье. Повтор такой структуры (без повтора содержания) лишь подчеркивает преемственность данного издания по отношению к книге [1], хотя авторы высоко оценивают роль теории в создании систем прикладного телевидения и прилагают усилия к достижению гармонии в отношениях науки и технологии.

Авторы надеются, что данная работа (не являющаяся учебным пособием и не повторяющая многие известные положения) вносит вклад в решение научных проблем создания телевизионной аппаратуры и аппаратуры вычислительных средств, ориентированных на решение задач выработки решений в реальном времени, исчисляемом «по часам объекта». Объектом наблюдения может быть любой динамический объект – от космических объектов (далёких или близких, естественных или искусственных) до биологических объектов (от простейших до нарушителей границ). Особенностью решаемых задач является моделирование высокопроизводительных систем, предназначенных для выработки решений и отвечающих жёстким ограничениям. Не менее важно требование повышения быстродействия, что на системном уровне означает повышение оперативности автоматизированных систем управления (АСУ), включающих в себя телевизионную и вычислительную подсистемы.

При проектировании систем авторы опираются на научную гипотезу о возможности разработки телевизионно-компьютерных средств на основе разработки математической модели аппаратуры. Конечно, по выражению Р. Беллмана «моделирование – это искусство прохождения между капканом переупрощения и болотом переусложнения». Авторы понимают моделирование не как упрощенчество, а как борьбу с фоновой и шумовой информацией, как выделение доминантного свойства системы.

Авторы выражают признательность рецензентам за полезные советы и сотрудникам ФГУП «НИИТ» и его смежникам, в первую очередь, ОАО ЦНИИ «Электрон», ФГУП НИИ «Пульсар», ФГУП НИИМА «Прогресс» и СПбГЭТУ «ЛЭТИ», чьи работы влились в общее русло твердотельной революции в телевидении. В ряде разделов отражены результаты, полученные в рамках научного проекта 05-01-08037 (2006 г.), выполненного при поддержке Российского фонда фундаментальных исследований.

ВВЕДЕНИЕ

В.1. Научно-прикладное телевидение как отрасль информатики

Под информатикой традиционно понимают компьютерную технику, подразделяемую на аппаратные и программные средства. При этом важнейшим аспектом создания обоих этих средств является их сложность. Самую большую сложность от компьютерных систем потребовало телевидение, так как обработка в реальном времени динамических изображений характеризуется предельными значениями быстродействия вычислений, скорости передачи информации по каналам связи и ёмкости памяти. Современная телевизионная техника всё более срастается с компьютерной техникой. Это явление затронуло и телевизионное вещание, и прикладное телевидение. Если на ранней стадии взаимодействия телевизионной техники с компьютерной преобладали методы, основанные на использовании компьютеров при проектировании систем и при обработке видеоинформации в сложных прикладных комплексах [1, 2], то в настоящее время произошел существенный поворот к срастанию этих областей техники. Такое срастание имеет множество предпосылок и проявлений. Полезно напомнить, что Н. Винер на раннем этапе развития телевидения полагал, что принцип телевизионной развёртки по своему значению в технике превзойдёт значение самого телевидения. Хотя прогноз Винера о скромной роли телевидения в жизни общества и не сбылся, важна и положительная часть его утверждения: в настоящее время ни один компьютерный монитор и практически ни один мобильный телефон не обходится без развёртки изображения.

Телевизионные системы – это информационные машины, олицетворяющие триединство информационной *теории*, *технологии* реализации метода согласования источника с каналом связи и собственно *техники* в ходе её функционирования (вместе с результатом – полученной *видеоинформацией*). Системы передачи и обработки информации, в первую очередь связи и её раздела телевидения, создаются при опоре на теорию кодирования для каналов и источников. Единый комплекс, включающий телевизионную систему, систему (сеть) связи и компьютер (компьютеры), создаётся на базе принципов системотехники (см., напр. [110]).

Слияние телевидения и компьютеров в единое информационное ядро является требованием времени для множества приложений – от телевизионных роботов до АСУ промышленных и оборонных объектов, в том числе АСУ их охраны. Срастание компьютерной техники, телевидения и техники связи не только востребовано, но и возможно благодаря *теоретическому и технологическому единству* изделий этих отраслей техники, обеспечиваемое микроэлектроникой. Возрастание сложности электронных компонентов привело к возникновению принципиально новых изделий микроэлектроники – систем на кристалле для обработки изображений и видеосистем на кристалле, объединяющих в себе КМОП[1] фотоприёмник и устройство цифровой обработки. Это, с точки зрения телевидения, обусловило второй этап твердотельной революции в телевидении, а с точки зрения информатики окончательно закрепило за телевидением статус её важнейшего раздела – видеоинформатики. В ходе этих методологических изменений концепция «телекамера – это вынесенная наружу часть компьютера», обсуждавшаяся на начальном этапе вливания телевидения в информатику, для нового поколения проектировщиков стала тривиальной.

Телевидение влилось в информатику не только как потребитель схемотехнических и программных продуктов, оно привнесло в неё серьёзные результаты теории и технологии связи. Конечно, этот подход известен давно, и Н. Винер писал: «исходя из моих общих идей, я рассматривал автоматические вычислительные машины также как одну из форм систем связи». Однако именно в телевидении, а не в компьютерной технике родились современные теоретико-информационные методы обработки многомерных зашумлённых сигналов, включившие в себя методы оптимального приёма сигналов, известные в радиолокации. Руководящим тезисом при этом было мнение основоположника компьютерной техники Дж. Фон Неймана, более полувека назад отметившего, что «вполне возможно, что мы уже недалеки от пределов сложности, которую можно достичь в искусственных автоматах, не продвигаясь особенно в глубь теории информации». Важно также, что один из важнейших методов построения высокопроизводительных компьютеров – распараллеливание вычислений – в форме метода параллельного (многоэлементного) накопления – родился в телевидении существенно раньше, чем

[1] КМОП – сокращение от комплементарная (с дополняющими типами переходов в полупроводнике) структура металл–окисел–полупроводник, в англоязычной литературе CMOS (complementary metal–oksid–semiconductor).

в вычислительной технике. Правда, в силу отсутствия в то время развитой информационной теории разработчики телевизионной техники говорили о накоплении энергии сигнала, а не об имевшем место в действительности накоплении информации.

Важным методологическим принципом в синтезе сложных, в том числе телевизионно-компьютерных, систем является применение единственного критерия качества, определяемого *целью* системы, и отнесение сколь угодно большого количества параметров и свойств системы к ограничениям, искомым параметрам или показателям помехоустойчивости. В системотехнике основным критерием качества работы системы обычно считается вероятность решения поставленной задачи, т. е. вероятность выполнения системой своей цели**.** Главным ограничением считается количество выделяемых ресурсов. Под ресурсами можно понимать актуальную для мобильных телевизионных систем массу; в иных случаях это может быть энергопотребление, габариты и стоимость (или совокупность этих показателей). Формальная запись задачи синтеза имеет вид *уравнения системы*, позволяющего найти набор искомых параметров как аргумент максимума вероятности решения системой целевой задачи [111]. Решение уравнения системы может быть получено как аналитически, так и с помощью различных итерационных процедур, известных в исследовании операций как методы динамического программирования. Методика аналитического решения уравнений этого типа освоена, в частности при вычислении эпсилон-энтропии случайных процессов. Основным свойством решения уравнения системы является то, что в оптимальной системе отводимые ресурсы целесообразно распределить между всеми её составляющими так, чтобы существовало определённое равновесие между компонентами системы.

Идея равновесия лежит и в основе информационного проектирования систем связи реального времени, ядром которого является *уравнение связи*, формализующее *принцип минимума информации при заданном её качестве* [3]. Решение уравнения связи преследует цель достижения равновесия в скорости создания информации источником сигнала и скорости передачи информации через канал связи. Под каналом связи в последующем изложении понимается широкий круг устройств связи телевизионных источников (обычно – телекамер) с компьютерным получателем вне зависимости от локализации вычислительной подсистемы – будь то на том же кристалле, что и массив фоточувствительных элементов, будь то в виде удалённого компьютера.

Инженерные задачи адаптации телекамер и классификации

несемантических образов (поставленные в книге [1]) уже во многом решены и научно-прикладное телевидение уже стало важным разделом информатики. Сегодня для его развития актуальны создание программируемых адаптивных однокристальных кодеров и декодеров, вплоть до размещения на одном кристалле фотоприёмной матрицы, кодера источника (устройства сжатия) и либо решающего устройства, либо (при удалённом пользователе) – кодера канала.

В.2. Модели систем и сигналов в прикладном телевидении

Для моделирования телевизионной системы необходим учёт двух указанных в её названии признаков: то, что это цельная система, и то, что эта система (машина) информационная. Опора на системный подход ведёт к методологии познания частей на основании целого и целостности, в отличие от классического подхода, ориентирующегося на познание целого через части. Не случайно *целостность и цель* – однокоренные слова, поэтому проектирование сложной компьютерно-телевизионной системы необходимо начинать с формулировки её цели, целевой функции. Такой подход требует учёта роли телевизионной системы в метасистеме связи, включающей получателя (в общем случае смысловой информации). Вне зависимости от типа получателя (человек или автомат) его, строго говоря, нельзя рассматривать в духе традиционной модели *системы* связи, использованной К. Шенноном для анализа *канала* связи. Термин «получатель» с позиции целостной системы становится некорректным, это уже не пассивный элемент, а *экспериментатор*. Он не просто наблюдает, а осуществляет *связь с явлениями Природы*, задаёт ей вопросы, на которые и получает ответ. С точки зрения философии, научно-прикладное телевидение является инструментом познания Природы. Концепция отражения человеческим сознанием окружающей природы явно принижает творческую роль исследователя и экспериментатора. Создателям систем прикладного телевидения ближе мнение Ф. Бэкона, утверждавшего, что «Природа не выражает себя свободно, но говорит, только будучи спрошенной». Эта роль создателей систем прикладного телевидения отражается включением в используемые при проектировании статистические методы наряду с теорией информации и теорией решений ещё и теории планирования эксперимента.

В отличие от классической модели системы связи по Шеннону, в прикладном телевидении не человек-телеграфист, а *телекамера* – *источник информации.* Из физических процессов (из того, что называют стимулом или «сюжетом») телекамера, созданная экспериментатором и направляемая на сюжет рукой экспериментатора, формирует информацию (сигналы), передающую ответы Природы

на заданные экспериментатором вопросы. Но для того, чтобы задавать вопросы, человек должен иметь цель и располагать некоторой априорной информацией. Эта предварительная связь наблюдателя с объектом отражается с помощью обобщённой модели связи человека с природным объектом, можно сказать – метамодели связи [3], которую здесь будем трактовать несколько иначе, используя новый (апостериорный) образ, замещающий априорный образ после постановки вопроса и получения на него доступного ответа [4] (рис. В.1).

Схема рис. В.1 показывает, что информация циркулирует по замкнутой треугольной цепи: объект–сигнал–субъект. Процесс передачи сигнальной информации неотделим от процесса создания образа объекта в сознании человека. Отсутствовавшая ранее связь

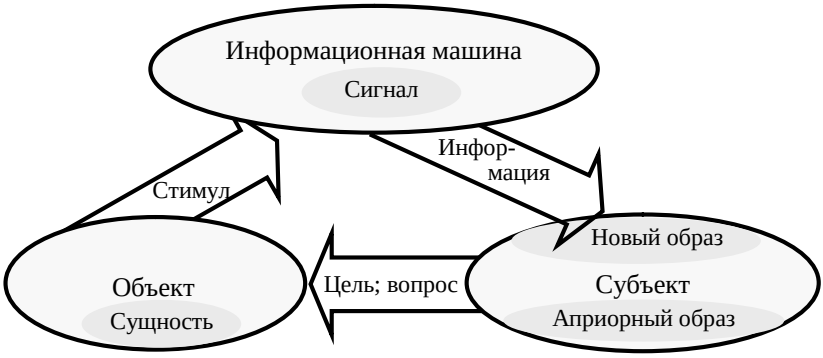


Рис.В.1. Обобщённая модель системы связи

получателя с объектом – это не исследованный ещё К. Шенноном канал с переспросом, а информационный процесс формирования цели связи с учётом априорной информации об объекте, имеющейся у получателя. Априорная информация об объекте тоже является образом, вне понятия которого нет информации. Конечно, априорный образ часто сильно отличается от апостериорного образа, формируемого в результате связи, и апостериорная информация больше априорной.

Вместе с тем при решении частных задач связи пользуются разомкнутой информационной моделью системы. Традиционно она состоит из источника и получателя, объединённых каналом связи, через который в условиях помех передаётся информация о сигнале. А. Н. Колмогоров отмечал, что информация не является скалярной величиной, хотя на заре становления теории информации Хартли счёл возможным считать информацию скалярной и постулировал исключение из рассмотрения смысловых («психо-

логических») факторов. Учёт лишь доминатного сигнала, т. е. подход к информации как к скалярной величине позволил описать её количественно и разработать теорию, основы которой созданы К. Шенноном, А. Н. Колмогоровым и Н. Винером. Теория информации на начальном этапе сводилась к статистической теории каналов связи, опирающейся на введённую Шенноном меру взаимной информации случайных процессов, заданных своими плотностями вероятностей.

Огромное разнообразие сюжетов прикладного телевидения может быть ранжировано по количеству априорной информации. Например, задача разведки с летательного аппарата при необходимости динамического совмещения изображений соседних телевизионных кадров характеризуется очень большой априорной неопределённостью (априори могут считаться приближённо известными лишь автокорреляционная функция изображения, моделируемого случайным полем, и модуль максимального смещения изображения относительно предыдущего кадра). Меньшей априорной неопределенностью, хотя и достаточно большой в силу наблюдения случайных полей, характеризуется задача классификации текстур, т. е. ситуация, когда решение принимается на основе измерения параметров автокорреляционной функции. Другая группа задач, принципиально отличающаяся существенно меньшей априорной неопределённостью сигналов – обнаружение и оценивание координат малоразмерных объектов, когда форма сигнала каждого отдельного объекта определяется функцией рассеяния точки объектива. Для этой группы задач характерны две градации априорной неопределённости: когда неизвестно количество объектов, неизвестны их интенсивности и координаты (как в задачах астроориентации), и когда известно количество объектов и их интенсивность, но неизвестны координаты, как в задаче оценки искажений волнового фронта методом Гартмана.

Информацию, которая участвует в создании полезных для наблюдателя (доминантных) образов, следует назвать полезной. Однако существуют бесполезные и даже ложные образы, а значит, существуют еще и бесполезная, и ложная информация. Среднее положение между полезной и шумовой (ложной) информацией в этой триаде занимает бесполезная или малополезная фоновая (контекстная) информация. Такая классификация, хотя и сохраняет важнейший инструмент – меру количества информации Шеннона, но делает важный шаг в переходе от скалярной информации к нескалярной.

Опора на обобщённую модель связи позволяет ставить задачи синтеза системы связи в целом, включая выбор и канала, и источника сигнала, т. е. включая и *передачу сигнала, и его рождение*. Модель прикладной телевизионной системы предполагает управление источником информации (телекамерой) получателем-экспериментатором и проектировщиком, а модель вещания – управление источником сторонним *манипулятором* (которых может быть много: идеологи, редакторы, операторы). Информационный облик прикладной телекамеры целиком в руках разработчика, а не регламентирован неким стандартом, как в вещании. В частности, малокадровое телевидение, ставшее незаменимым в космических приложениях – это не суррогат телевизионного вещания, а концепция, охватывающая все способы оптимального согласования источника с сюжетом и каналом.

Рассмотрение верхней ветви обобщённой модели связи рис. В.1 в приложении к телевидению даёт описание преобразований стимула в сигнал и передачи сигнала наблюдателю. В этой информационной модели телекамера выступает в роли кодера, воспринимающего искажённый сигнал (рис. В.2).

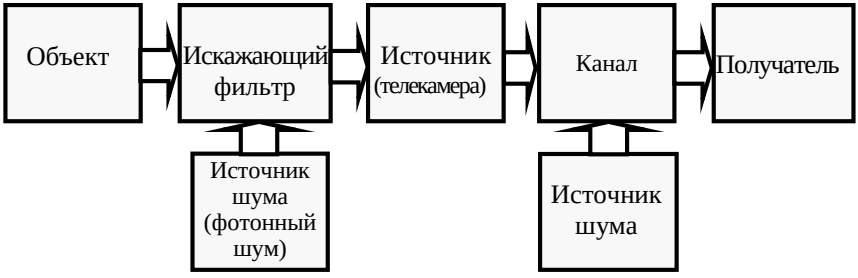


Рис. В.2. Обобщённая модель передачи информации, учитывающая два источника шумов – в оптическом тракте и в канале связи телевизионной системы с получателем

Обязательность учёта в информационной модели как минимум двух источников шума определяется тем, что в системе связи, передающей видеоинформацию, неизбежно присутствуют шумы двух типов – фотонные и тепловые. Хотя разработчики фотоприёмников и телекамер постоянно снижают уровень собственных шумов этих приборов, однако свести такие шумы к нулю можно лишь при температуре абсолютного нуля.

Важнейшим отличием указанной двухкаскадной схемы передачи информации является неизбежность передачи через канал связи не только полезной, но и *шумовой информации*, т. е.

информации о входном фотонном шуме. Проблема борьбы с шумовой информацией возникает и может быть исследована уже при учёте хотя бы одного источника шума (кроме шума канала связи). Наличие многочисленных источников шумов в реальной телевизионной камере уточняет картину, не меняя главных положений. Принципиальной является невозможность сведения влияния указанных источников шума к одной точке информационной модели.

Основной задачей, решаемой с помощью информационной модели рис. В.2, является анализ рождения сигналов и прохождения сигналов и шумов через различные преобразователи информации (кодеры), часть из которых может быть доступна для управления, а часть – не доступна для управления получателем. Важнейшими элементами в информационной модели являются [1 – 4]:

- равновесное согласование источника с каналом (формализуемое уравнением связи), состоящее в выравнивании скорости создания информации источником и скорости передачи информации по каналу;
- повышение качества кодирования информации, осуществляемое минимизацией информационного риска, состоящего из потери доминантной (полезной) информации и шумовой информации.

Информационная модель, даже оптимизированная с учётом указанных позиций, позволяет построить оптимальную систему передачи синтаксической информации. Проблемы же семантики в передаваемой видеоинформации пока плохо поддаются формализации.

В.3. Этапы твердотельной революции в телевидении

В истории телевидения следует выделять три главных периода, сменявших, конечно, друг друга не скачком, а постепенно, через определённые переходные процессы, но толчком к которым были революционные изобретения и преодоление крупных *технологических* барьеров. Поэтому иногда, при пристальном внимании к истории, такие переходные периоды можно выделять в самостоятельные разделы, обязательно помня о существовании глобальных периодов. При этом создание новой техники каждый раз являлось не только целью, но и средством достижения более высокой цели. Конечно, для изобретателя разработка прибора на конкретном отрезке его жизни может быть самоцелью, но, в конечном счете, цель любой человеческой деятельности – человек, или, иными словами, любое техническое достижение следует оценивать в первую очередь в социальном аспекте. Поэтому и наличие трёх главных периодов в истории телевидения следует рассматривать с социальных позиций с учётом технологических факторов. Указанные три периода опреде-

ляются ролью телевидения в жизни общества: первый период – до рождения вещательного телевидения, второй – от его рождения до достижения зрелости в региональном (национальном) масштабе, третий – от начала глобального охвата и завоевания ведущей роли в массовой информации, науке, образовании, контроле производственных процессов и безопасности. Конечно, как и любая другая наука, телевидение включает наряду с философскими проблемами технологические, решение которых и обеспечивает достижение поставленной цели. Первый период развития телевидения (до начала массового телевизионного вещания) характеризовался реализацией одного из важнейших принципов телевидения – развёртки – с помощью механического сканирования. На этом этапе были освоены механические системы с очень малым числом элементов разложения. Первые изображения, переданные по радиоканалу, имели разложение всего на несколько десятков строк. Конечно, при большом желании это можно назвать началом телевизионных передач, но правильнее считать этот период предысторией телевидения, так как победное шествие телевидения по планете могло начаться только после преодоления барьера качества, который соответствует чёткости примерно в 300 строк. Второй период развития телевидения – эпоха электронного телевидения, начавшаяся с изобретения и изготовления в США в 1933 г. В. К. Зворыкиным электронно-лучевой передающей телевизионной трубки (иконоскоп). В 1934 г. эту трубку воспроизвели в Ленинграде (Б. В. Круссер) и работы по телевидению приобрели новые цели, связанные с разворачиванием новой отрасли промышленности, которая должна была обеспечить массовое телевизионное вещание. В нашей стране системы электронного телевидения с умеренной чёткостью 180 строк впервые были реализованы в Ленинграде в 1934 г.

Основными достижениями периода электронно-лучевого телевидения можно считать создание сети телевизионных центров и рождение космического телевидения. Эти два достижения перечислены в порядке влияния на общество, но следует подчеркнуть, что рождение космического телевидения в нашей стране является серьезнейшим предметом нашей национальной гордости.

Принципиально новый период развития телевидения связан с рождением твердотельного телевидения, которое отражает главную тенденцию современной техники связи – тотальный переход к «рукотворным кристаллам», т. е. большим интегральным схемам.

В твердотельном телевидении на смену электронно-лучевым передающим трубкам пришли твердотельные фотоприёмники на основе приборов с зарядовой связью (изобретены в 1970 г.). Техно-

логия изготовления этих твердотельных приборов, разработанная в США (У. Бойл и Г. Смит) и быстро освоенная в нашей стране (Ф. П. Пресс) во многом схожа с технологией компьютерных узлов – процессоров и оперативной памяти. Это родство позволило создавать телекамеры на хорошо сопрягаемых узлах аналоговой и цифровой схемотехники. Но еще важнее стала методология, которая объединила телекамеру и компьютер в единое целое: аналогично тому, как физиологи осознали, что глаз – это вынесенная наружу часть мозга, так и разработчики систем прикладного телевидения осознали, что телекамера – это вынесенная наружу часть компьютера.

Говоря о твердотельной революции в телевидении необходимо помнить, что важным стимулом для неё было космическое телевидение, важной вехой в создании которого было подписание С. П. Королёвым 22.08.1956 г. (до запуска первого искусственного спутника Земли) технического задания ВНИИТу на разработку систем космического телевидения. Великий конструктор космических ракет не мыслил выход в космос без телевидения. Выданное им ТЗ имело две цели: создание телевизионной аппаратуры для беспилотного космического аппарата и для пилотируемого космического аппарата. С. П. Королёв сразу наметил два направления развития космонавтики – пилотируемую и беспилотную. Роль пилотируемой космонавтики широко признана, но роль беспилотной космонавтики недостаточно осознана. Именно беспилотная космонавтика была стимулом расширения функций телевидения в космосе и срастания её с компьютерной техникой. Наглядной иллюстрацией этого процесса является расширение сети спутников-ретрансляторов телевизионных сигналов. Кроме того, интенсивно внедряемое в вещании цифровое телевидение (цифровая передача и обработка, в том числе межкадровая) появилось на три десятилетия раньше именно в космических малокадровых системах [5].

Первый этап твердотельной революции в телевидении *связан с появлением приборов с зарядовой связью (ПЗС) и микропроцессоров* [1]. Фотоэлектрическое преобразование в полупроводнике (в твердом теле, кристалле) было освоено ещё в электронно-лучевом телевидении, и знаменитая книга «Телевидение» В. К. Зворыкина и Д. А. Мортона в ряду физических принципов телевидения подробно отражает значимость свойств полупроводников и их описаний с помощью зонной теории. Параллельное накопление зарядового изображения в слое полупроводника тоже давно было освоено в электронном телевидении (при считывании сигнала электронным пучком). Поэтому твердотельное телевидение – это, в первую

очередь, новый *технологический* метод реализации другого этапа рождения телевизионного сигнала – развёртки изображения. Поэтому ПЗС вполне могли бы называться приборами с зарядовой развёрткой.

Революционность появления твердотельного телевидения связана далеко не только с уменьшением габаритов благодаря исключению генератора пучка электронов и громоздкого устройства управления этим пучком для развёртки изображения. Огромное значение имеют стабильность параметров твердотельного фотоприёмника и жёсткость растра телекамеры, сделавшая эффективной межкадровую обработку сигналов изображения. Большое значение имеет и исключительная надёжность твердотельных фотоприёмников, и имеющиеся случаи выхода из строя телекамер ныне связаны в первую очередь с ненадёжностью дополнительных, обслуживающих элементов телекамер (интерфейсов питания и связи). Не случайно создатель первых отечественных матриц ПЗС Ф. П. Пресс на вопрос о надёжности его приборов отвечал со скромным достоинством: «Они вечны». И если учесть ещё и метрологические свойства твердотельных фотоприёмников (линейную свет–сигнальную характеристику и большой динамический диапазон) и простоту изменения параметров разложения изображения, то нельзя не признать твердотельной революции в телевидении.

Сейчас, по прошествии более трети века с начала первого этапа твердотельного телевидения, электронное телевидение почти полностью уступило место твердотельному. Остались лишь небольшие временные «островки» в области инфракрасной (ИК) техники и аппаратуры, устойчивой к интенсивной радиации. Наличие таких переходных периодов, когда используются и старые, и новые приборы, не противоречит утверждению о свершившейся твердотельной революции в телевидении. Так, в первые годы внедрения ПЗС со стороны приверженцев старой электронно-лучевой техники (можно назвать их подход тактическим, не учитывающим стратегии развития отрасли) к новой твердотельной технике было много нареканий из-за недостаточной чёткости, большого количества дефектных элементов, недостаточной чувствительности и заметного темнового тока, затрудняющего реализацию малокадрового режима. С годами все эти недостатки были устранены, а явные преимущества остались. И типичное для сторонников вакуумной техники (обеспечивавшей в начале твердотельной революции лучшее качество изображений) выражение «побалуются и бросят» кануло в Лету.

Наглядным примером такого переходного периода служит также борьба старых электронно-лучевых приёмных трубок (кинескопов)

с новыми жидкокристаллическими устройствами отображения видеоинформации. Никакие ухищрения в электронных трубках (вроде размещения электронной пушки параллельно экрану для уменьшения толщины монитора, рис. В.3) не спасут их, и они, в конце концов, полностью сойдут с исторической сцены. В переходной период от электронного к твердотельному телевидению разработчики еще постоянно оглядывались на электронно-лучевые приборы – не только для сравнения твердотельных фотоприёмников с телевизионными передающими трубками, но и для реализации отображения телевизионной информации на электронно-лучевых трубках.

Рис. В.3. Изображение кинескопа с электронной пушкой, почти параллельной экрану, на экране жидкокристаллического монитора

Вакуумные приборы не сошли сразу со сцены при изобретении твердотельных фотоприемников, т. к. долгое время оставались единственными приборами отображения телевизионных изображений. Создание приемной телевизионной трубки Б. Л. Розингом (1908 г.) и заметно опередило создание передающей телевизионной трубки его учеником В. К. Зворыкиным (1933 г.), и заметно пережило его. Сейчас, на втором этапе твердотельной революции, передающие электронно-лучевые трубки, верой и правдой прослужившие телевидению полвека, с иронией называют «фоточувствительными радиолампами» и «изображающими радиолампами», проводя прямую аналогию с произошедшим ранее вытеснением радиоламп транзисторами. «Соотношение сил» электронно-лучевых приборов и жидкокристаллических систем отображения информации в переходной период, на втором этапе твердотельной революции в телевидении, видно, в частности, из того, что в 2006 г. реклама компьютерных центров предлагает покупателям «более 15 моделей мониторов ЭЛТ и более 300 мониторов ЖК». Кроме того, наиболее массово выпускаемые приборы – цифровые фотоаппараты и мобильные телефоны строятся с использованием исключительно жидкокристаллических мониторов.

Приобретя благодаря твердотельной технологии такие важные свойства, как надёжность, компактность (см. рис. В.4), жёсткость растра, прикладное телевидение стало незаменимым в системах безопасности различных объектов, на транспортных средствах, в космосе, армии, авиации и на флоте.

Вместе с тем малые габариты не являются единственным фактором, определяющим победу твердотельного телевидения над электронно-лучевым. Миниатюрные видиконы не намного больше охлаждаемых матричных ПЗС, применяемых в астрономии (эти матрицы должны иметь достаточно большие площади и размеры элемента разложения для согласования с телескопами), рис. В.5. и В.6.

Рис. В.4. Видиконы (разработки ЦНИИ «Электрон») и матричные ПЗС сопоставимой площади изображения (разработки ЦНИИ «Электрон» и ЦНИИ «Пульсар»)

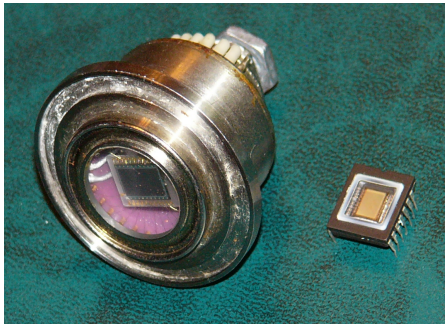


Рис. В.5. Охлаждаемая матрица ПЗС разработки ЦНИИ «Электрон» и неохлаждаемая матрицы ПЗС разработки фирмы Sony

Не менее важными, если не более важными, стали надёжность, лёгкость перестройки параметров и сопряжения с компьютерной техникой.

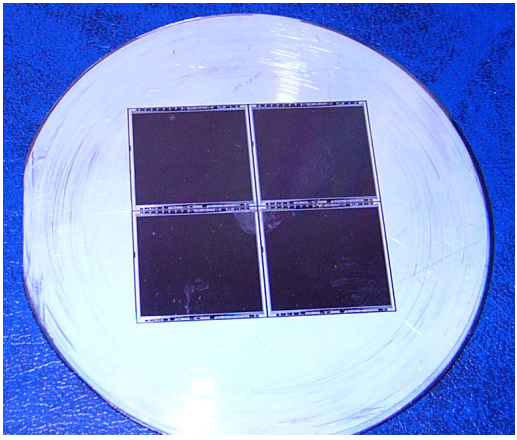


Рис. В.6. Кремниевая пластина диаметром 70 мм с астрономическими матрицами ПЗС класса 1,3 Мегапиксела. Разработки ЦНИИ «Электрон»

Практика применения прикладного телевидения стимулирует создание разнообразных фотоприёмников видимого и инфракрасного диапазонов, в том числе ПЗС – линейных и матричных. Если в системах безопасности крупных объектов обычно идут по пути создания весьма разветвлённых систем с большим количеством относительно простых телекамер, то во многих специальных системах идут по пути увеличения чёткости передаваемого изображения до уровня, превышающего даже чёткость телевидения унифицированной высокой чёткости. Основной тенденцией развития ПЗС для прикладных систем является *увеличение количества передаваемой информации* путём увеличения числа элементов и увеличения быстродействия. Массово выпускаются матрицы ПЗС с числом элементов 2 … 8 Мегапикселов (типовые значения для цифровых фотоаппаратов), созданы матрицы с десятками Мегапикселов. Производители линейных ПЗС уверенно освоили число элементов 5 … 12 тысяч элементов и выпускают множество их модификаций, которые характеризуются различными значениями размеров элемента (от 4×4 мкм до 20×20 мкм) и скоростью считывания от 2 до 50 МГц. Наблюдается устойчивая взаимосвязь между числом элементов и достигаемой кадровой (для линейных приборов – строчной) частотой, так как почти все матричные ПЗС имеют один (редко два или четыре) канала считывания сигнала изображения. Например, фирма Atmel выпустила линейные ПЗС с 4 выходами, что подняло эквивалентную частоту считывания до 160 МГц. Эта основная тенденция определяет развитие таких

отраслей, как создание цифровых фотоаппаратов (включая встраиваемые в мобильные телефоны), компьютерные сканеры и факсы, а также выпускаемые существенно меньшими тиражами космические и авиационные системы наблюдения.

То, что прикладное телевидение только став твердотельным по-настоящему внедрилось в практику, видно из такого примера. Оптико-электронные методы применялись в задаче измерения силы тяжести – гравиметрии – давно, но долго использовали одноэлементные фотоприёмники с механической развёрткой изображения специально формируемой световой метки, координаты которой определялись текущим значением силы тяжести. Никто не пытался применять электронно-лучевые трубки в оптико-электронных преобразователях гравиметров. Только тогда, когда прикладное телевидение готово было одновременно обеспечить и очень высокую надёжность, безоператорную работу в течение многих лет, компактность и высокую точность измерения параметров сигналов, оно стало востребованным. Поэтому оптико-электронная гравиметрия скачком перешла от механического телевидения к твердотельному, минуя фазу электронно-лучевого.

Завершение первого этапа твердотельной революции в телевидении, т. е. переходного этапа от электронно-лучевого телевидения к компьютерному и твердотельному телевидению совпало с осознанием человечеством наступления эры глобальной информатизации, когда телевидение и видеотехника с помощью космических средств и Интернет охватили всю Землю.

Второй этап твердотельной революции*, т. е. резкого изменения технологии*, прикладная телевизионная техника переживает в последние годы. В видимом и ближнем ИК диапазоне качественный скачок осуществлён благодаря переходу от главного типа фотоприёмников на ПЗС к *КМОП фотоприёмникам*. Принципиальное отличие этих приборов от ПЗС состоит в методе реализации *развёртки*. Координатная адресация к каждому элементу вместо протаскивания заряда каждого элемента сначала по всему столбцу, а потом по всей строке позволяет избежать искажений, связанных с неэффективностью переноса заряда. В ИК технике качественный скачок обусловлен как наращиванием площади и числа элементов в криогенно охлаждаемых фотоприёмных матрицах (в первую очередь ПЗС с барьерами Шоттки на основе PtSi, а также матрицах с координатной адресацией на основе InSb и $CdHg_xTe_{1-x}$), так и освоением *неохлаждаемых матричных* приборов с чувствительными элементами на основе микроболометров

и *КМОП устройствами считывания* видеоинформации. Если первый этап твердотельной революции связан с рождением матричных приборов с зарядовой связью и микропроцессоров, то нынешний этап характеризуется следующими чертами:

• совершенствование технологии изготовления матричных ПЗС и КМОП фотоприёмников позволило массово выпускать приборы высокой чёткости, существенно превышающей требования действующего телевизионного вещания – с числом элементов в единицы и *десятки мегапикселов* [16–21];

• новая технология создания *КМОП фотоприёмников* открыла перспективу заметного повышения качества формируемых телевизионных сигналов благодаря возможности *многоканального считывания* (в пределе – с применением отдельного АЦП для каждого элемента разложения) и введения *усиления до воздействия шумов считывания*. Развитие этих преимуществ КМОП фотоприёмников в ближайшем будущем должно вытеснить из практики формирования телевизионных сигналов не только вакуумные приборы усиления слабых оптических сигналов (электронно-оптические преобразователи, ЭОП), по и ПЗС [22, 23].

• совершенствование технологии изготовления микроэлектронных устройств обработки сигналов привело к синтезу разнородных вычислительных устройств – микропроцессоров, запоминающих устройств и программируемой логики в пределах одной микросхемы, которые получили название *системы на кристалле* [6–15];

• применение КМОП технологии к реализации матричных фотоприёмников породило новую ветвь твердотельных устройств, объединяющих в себе фотоприёмник, устройства развертки, квантования и обработки изображения, которые названы *видеосистемами на кристалле* [21 – 40];

• стремительный рост объёмов продаж твердотельных телекамер для охранной индустрии и цифровых фотоаппаратов и мобильных телефонов со встроенными цифровыми фотоаппаратами сделал телевидение *массовым* не только в смысле приёма, но и *передачи изображений* [5, 41, 167];

Совокупность этих черт позволяет говорить о *революционном*, а не эволюционном характере развития телевизионной техники. Качественный скачок связан с освоением технологии ультрафиолетовой литографии, давшей субмикронную точность изготовления приборов: ведущими в микроэлектронике странами

(в том числе Тайванем), уже освоена точность 90 нм, ведутся работы по запуску технологии с точностью 65 нм и 45 нм.

После двух десятилетий работы с матрицами вещательного формата (около 0,4 мегапиксела) телевидение быстро преодолело мегапиксельный барьер. Сейчас в цифровых фотоаппаратах, объём продаж которых за последние два года удвоился, массово применяются пяти- и десятимегапиксельные матрицы; в 2005 г. появились матрицы с 22 Мегапикселами, в 2006 г. – фотоаппараты с матрицей 39 и даже с 160 Мегапикселов. Такие большие массивы фотоприёмных матриц (при 16 разрядном АЦП 160 Мегапикселов – это уже 2,5 Гбит/кадр) ещё более заостряют вопрос об эффективном кодировании (сжатии) сигналов изображений при их передаче. Для решения этой проблемы широким фронтом ведутся работы по созданию различных однокристальных кодеров для цифрового телевидения, телевидения высокой чёткости и прикладных систем.

Срастание компьютерной техники, телевидения и техники связи, характерное для второго этапа твердотельной революции в телевидении, усиливается благодаря *технологическому единству* изделий этих отраслей техники, обеспечиваемое микроэлектроникой. Это срастание ранее инициировалось только специалистами телевизионной техники и выразилось в рождении твердотельного телевидения как «связки» приборов с зарядовой связью и микропроцессоров. В настоящее время набрало силу встречное движение к телевидению со стороны специалистов компьютерной техники (раньше телевизионщиков наладивших прямые связи с производителями заказных СБИС), ставшее одной из причин второго этапа твердотельной революции в телевидении. Так как современный этап создания телевизионно-компьютерных систем характеризуется единством технологии фотоприёмников и вычислительной подсистемы, полезно сравнить темпы роста числа транзисторов в микропроцессорах и фотоприёмных матрицах. Такое сравнение весьма полезно и наглядно, в частности, и потому, что приборы с зарядовой связью и микропроцессоры были изобретены одновременно в 1971 году. Рост числа транзисторов в микропроцессорах хорошо описывается законом Мура [155]: при линейной шкале времени и логарифмической шкале числа транзисторов эта зависимость представляет собой прямую, соответствующую удвоению числа транзисторов в СБИС примерно за 1,8 года.

Рост числа элементов в ПЗС имеет несколько участков (рис. В. 7):

• первый, когда благодаря регулярной структуре фотоприёмника темпы роста числа элементов фотоприёмника заметно опережали темпы роста числа транзисторов в процессорах;

• второй, когда разработчики фотоприёмных матриц сосредоточились на качественных показателях (количество дефектов, темновой ток, эффективность переноса, шум считывания и т. п.), в результате чего темпы роста числа элементов резко снизились, и со временем число транзисторов в процессорах превзошло число элементов разложения в ПЗС;

• третий участок характеризуется восстановлением темпов роста числа элементов разложения в фотоприёмниках. Различие в абсолютном значении числа транзисторов в процессорах и числа

Рис. В.7. Рост числа элементов фотоприёмников и процессоров от момента изобретения до настоящего времени

элементов в фотоприёмных матрицах будет сохраняться в силу принципиального различия между элементами, работающими с двоичными сигналами и с квазинепрерывными (число накапливаемых электронов в элементе матрицы исчисляется десятками и сотнями тысяч);

- четвёртый участок характеризуется уменьшением скорости роста числа элементов и числа транзисторов, т. е. «насыщением» этих зависимостей. Этот эффект для цифровых систем обусловлен уменьшением числа электронов на бит информации (на сегодня это сотни электронов на бит, это число убывает пропорционально квадрату линейного размера элемента (транзистора), т. е. точности технологии).

Второй излом в темпах роста числа элементов в фотоприёмниках наглядно показывает наличие второго этапа твердотельной революции в телевидении. С этим же моментом времени – 2000 г. – совпадает появление на рынке мобильных телефонов со встроенными фотокамерами, которые быстро стали самым популярным типом цифровых камер.

Причинами такого революционного изменения скорости роста элементов в СБИС фотоприёма и обработки информации являются, с одной стороны, постоянно возрастающие потребности в получении и передаче видеоинформации. С другой, технологической стороны, этот излом обусловлен освоением двух направлений: субмикронной технологии (благодаря ультрафиолетовой литографии) и кремниевым пластинам большого размера. Если на начальном этапе полупроводниковой техники в 1960-х г.г. для изготовления транзисторов кристаллы вырезались из цилиндров диаметром 30 мм, а в ходе первого этапа твердотельной революции в телевидении в 1980-х г.г. типичны были пластины диаметром 70 мм (см. рис В.6.), то на втором этапе твердотельной революции в телевидении типичны пластины диаметром 200 мм (рис. В.8, [51]), и начинается освоение пластин диаметром 300 мм.

Скорость роста технологических возможностей микроэлектроники впечатляет: на заре твердотельной революции в телевидении для оценки экономической эффективности разрабатываемых микросхем была введена специальная размерность *цент/бит*. В настоящее время эта размерность перестала соответствовать уровню технологии – стоимость одного бита снизилась в *миллион* раз. Особенно это заметно для СБИС памяти и фотоприёмников – благодаря регулярной структуре

они обладают *минимальными затратами* на каждый переданный бит информации.

Вместе с тем *нельзя беспредельно увеличивать точность технологии*, и уже близок момент насыщения кривой роста числа элементов в кристалле (и вычислителей, и фотоприёмников) и *прекращения эры действия закона Мура*. Этому насыщению есть

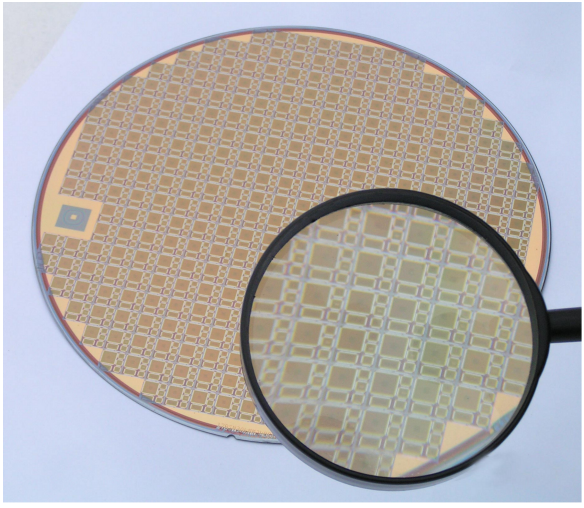


Рис. В.8. Кремниевая пластина диаметром 200 мм с комплектом БИС БПФ разработки «ОТИК дизайн»

несколько причин – физических, информационных (необходимый минимум числа электронов на бит) и экономических [5, 6].

КМОП фотоприёмники по сравнению с ПЗС дешевле и позволяют интегрировать на одном кристалле с ним ряд других элементов, в первую очередь – АЦП и интерфейсы каналов связи (рис. В.9). Важнейшим фактором является размещение в каждом элементе КМОП фотоприёмника нескольких транзисторов (при технологии с точностью 0,1 мкм – до 40, а при технологии 45 нм – до 256 [6, 26]).

Рост точности технологии обуславливает продвижение в КМОП матрицах от пассивного элемента к активному (с аналоговым усилением) и далее – к элементу с цифровой обработкой. Это открывает широкие возможности размещения устройств обработки сигналов в каждом элементе, вплоть до реализации в каждом элементе АЦП с плавающей запятой. Использование докоммутационного (предшествующего развёртке) усиления в КМОП фотоприёмниках должно завершить твердотельную

революцию не только в обычных телекамерах, но и в приборах ночного видения, пока использующих электронно-оптические (вакуумные) преобразователи. Такие важные преимущества

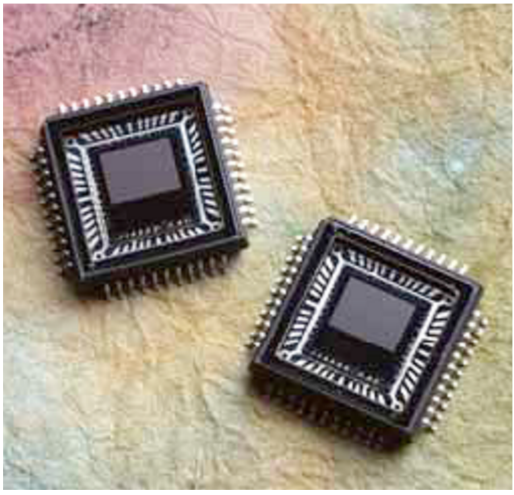


Рис. В.9.
КМОП матричные фотоприёмники даже на взгляд отличаются от ПЗС тем, что не вся площадь кристалла занята матрицей – часть площади расходуется на управление развёрткой, на множество АЦП и интерфейсы каналов связи

привели к тому, что по данным исследовательской фирмы In-Stat/MDR начала 2005 года уже в 2005 г. объём продаж КМОП фотоприёмников должен впервые превзойти объём продаж ПЗС фотоприёмников, а в ближайшие три года темпы роста отгрузок КМОП фотоприёмников будут в 7 раз превосходить аналогичный показатель для ПЗС; в целом же рынок твердотельных матричных фотоприёмников будет увеличиваться в среднем на 30% ежегодно. Важные технолого-экономические преимущества КМОП фотоприёмников обусловили пристальное внимание к ним и уже обоснованы перспективы заметного повышения их чувствительности вплоть до достижения характеристик не худших, чем у ПЗС. Идеи о потенциальных преимуществах КМОП фотоприёмников высказывались относительно давно, и ещё в 1990 г. была опубликована работа [28], в которой ПЗС были названы динозаврами. Ультрафиолетовая технология уже сделала КМОП фотоприёмники конкурентоспособными, а в будущем эти представители нового этапа твердотельной революции должны окончательно вытеснить ПЗС [29].

Единство КМОП технологии изготовления фотоприёмников и компьютерных СБИС обусловило возможность интеграции их в единый прибор, причём целью такого объединения является не только уменьшение массогабаритных показателей и энергопотребления (острая потребность в уменьшении которых стоит, например, при создании телевизионных систем для малогабаритных беспилотных летательных аппаратов), но и перенос центра сложности обработки видеоинформации и принятия решений из внешнего вычислителя непосредственно в телекамеру. Это в корне меняет роль телекамеры в системе – из относительно простого формирователя изображения она превращается в интеллектуальный источник видеоинформации и решений, становится системообразующим устройством.

Телекамера и сейчас уже методологически неотделима от компьютера, но в ближайшем будущем телекамера прикладной телевизионной системы будет составлять единое целое с компьютером или процессором цифровой обработки. Причём это единство не только технологическое, но и методологическое и конструктивное: размещение на одном кристалле делает реализуемой многоканальную связь телекамеры с вычислителем.

В результате такого срастания фотоприёмника с устройством обработки на втором этапе твердотельной революции в телевидении остановился рост доли фоточувствительного слоя в общем объёме телекамеры (рис. В. 10).

Эта *относительная* характеристика телекамеры, не менее чем их абсолютные конструктивные или эксплуатационные показатели, показывает наличие важнейших этапов в телевидении. Использованная характеристика относительного объёма фотослоя наглядно отражает и революционные переходы от одного этапа к другому, и эволюционные изменения внутри этапа.

Наличие второго этапа твердотельной революции характеризуется не только возрастанием относительной доли фотослоя благодаря размещению на том же кристалле устройств управления развёрткой (что было в силу технологических причин недоступно в ПЗС). Достаточно неожиданным (если учитывать только историю телевидения) оказалось уменьшение этой доли из-за занятия части площади кристалла устройством обработки изображения.

Вместе с тем зрительный анализатор человека тоже состоит не только из сетчатки глаза. Как отмечал при анализе структуры ЭВМ ещё Дж. Фон Нейман, «примерно пятую часть головного мозга составляет его зрительный аппарат, который, насколько нам

известно, не занят ничем другим, кроме принятия решений относительно зрительных аналогий».

Указанная пятая часть мозга примерно на четыре порядка больше по объёму, чем сетчатка глаза. Это даёт представление о соотношении требуемой сложности телекамеры и компьютера, принимающего семантические решения. Но, кроме того, и объясняет изменение знака наклона кривой эволюции телекамер на рис. В.10, приходящееся на переход от первого ко второму этапу твердотельной революции в телевидении или на переход к видеосистемам на кристалле.

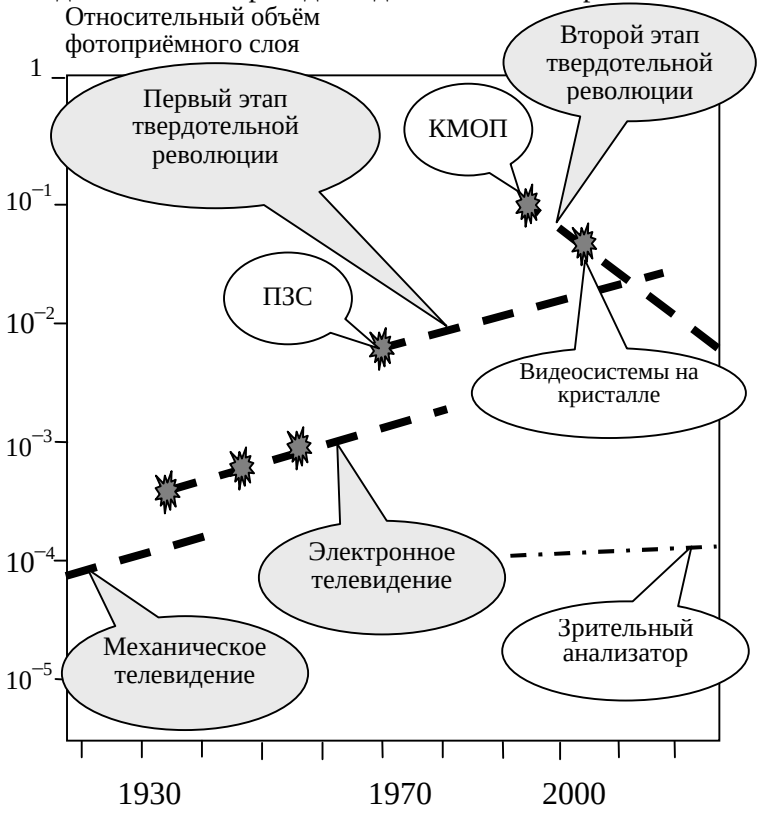


Рис. В.10. Эволюция относительного объёма фотоприёмного слоя в телекамере на различных этапах развития телевидения (для сравнения приведён относительный объём сетчатки глаза; на графике этапа электронного телевидения отмечены различные типы передающих трубок: иконоскоп, суперортикон, видикон)

Рождение видеосистем на кристалле не только прекратило дискуссию на тему, является ли телекамера вынесенной наружу частью компьютера (аналогично тому, что глаз является вынесенной наружу частью мозга), но и обратило взор разработчиков на необходимость распределения площади кристалла между фотоприёмником и вычислителем. В прикладном телевидении давно осознано, что телевизионную систему, состоящую из оптической и фотоприёмной подсистем необходимо проектировать как единое целое [1]. Однако именно второй этап твердотельной революции радикально усилил этот тезис: телевизионная система, объединяющая в себе оптическую, фотоприёмную и вычислительную подсистемы должна проектироваться как единое целое. Развитие видеосистем на кристалле превращает однокристальную камеру в центральный (в пределе – единственный) узел системы формирования и обработки изображений, для которой становится актуальной задача распределения ресурсов между фотоприемником и подсистемой обработки для наилучшего выполнения системой стоящей перед ней цели. Твердотельное телевидение отвергает принцип декомпозиции, согласно которому можно анализировать (синтезировать) систему по частям.

Поэтому проектирование видеосистем на кристалле (и многих других сложных функциональных блоков, т. е. систем на кристалле) должно предусматривать обязательное участие разработчика систем, в которой сложные функциональные микроэлектронные блоки будут выполнять основные функции. Современная технология позволяет перенести создание системы на уровень формирования структур на пластине, т. е. технология микроэлектроники превращается из технологии создания компонентов в технологию выпуска систем [8–16].

Изменение технологии изготовления СБИС для прикладных телевизионных систем, в состав которых вошли системы на кристалле и видеосистемы на кристалле, изменили методы теоретического синтеза и проектирования систем и устройств. Если на начальном этапе развития телевидения согласно методологии создателя электронного телевидения В. К. Зворыкина в проектировании преобладали методы физики, то новый этап порождает новый метод проектирования, который характеризуется высокой скоростью создания систем, резким возрастанием роли системного подхода и значимости программно-алгоритмического обеспечения. Появившиеся в результате технологического прорыва и интенсивно

развивающиеся видеосистемы на кристалле не только оправдали надежды разработчиков в части достигаемых возможностей, но и породили ряд проблем, среди которых – повышение сложности не только самого изделия микроэлектронной техники, не только сложности его изготовления, но и сложности его применения. Поэтому проектировщики систем прикладного телевидение во след освоению микропроцессоров и программируемых логических матриц должны освоить широкий набор средств САПР для эффективного использования систем на кристалле и видеосистем на кристалле.

Изложенное находится в русле подготовленного Управлением радиоэлектронной промышленности и систем управления ФАП России документа «Стратегия развития электронной промышленности на ближайшие годы» [8, 16]. Необходимо «только» проведение отечественной радиоэлектронной промышленностью в рамках программы «Национальная технологическая база» в 2007–2011 гг. работ не только по ставшим уже традиционными вычислительным *системам на кристалле*, но и комплекса работ, включающих решение задач микроэлектронной технологии, прикладной телевизионной техники и разработки средств автоматизированного проектирования по проблематике *видеосистем на кристалле*.

На данном этапе развития наиболее перспективное направление разработок видеоинформатики – создание интеллектуальных телевизионно-компьютерных систем, которые позволят реализовать системы технического и искусственного зрения, сопоставимые по характеристикам с биологическим зрением [7]. Уже сейчас число пользователей Интернет в мире превысило миллиард, а число пользователей мобильных телефонов – два миллиарда. В ближайшем будущем связь, включая мобильную телевизионную связь, должна стать поистине *глобальной*. Изложенное обосновывает необходимость отнести создание устройств обработки сигналов с помощью систем на кристалле и прикладных телевизионных систем с использованием видеосистем на кристалле к технологии федерального уровня, способной обеспечить современной элементной базой научный, промышленный и оборонный комплексы России.

Глава 1

ТВЕРДОТЕЛЬНЫЕ ТЕЛЕВИЗИОННЫЕ КАМЕРЫ

Твердотельное телевидение порождает новые возможности создания и обработки видеоинформации и не должно сводиться только к рассуждениям об «элементной базе». Твердотельные матрицы не являются «твердотельными аналогами передающих телевизионных трубок», как порой можно слышать от ветеранов создания телевизионных систем. Вакуумная техника, пусть и малогабаритная, не даёт однокристальных систем и, освоив многоэлементное накопление, принципиально остановилась перед многоканальным адаптивным считыванием, являющимся крайне важным аспектом в единении телевизионной и компьютерной техники, широко применяющей метод распараллеливания вычислений (метод «множественного потока данных»). Конечно, так же, как и в электронном телевидении, в твердотельном телевидении существует несколько направлений развития телекамер, грубо разделяемых на два класса: массовый выпуск дешёвых телекамер преимущественно для охранных и производственных целей и создание прецизионных телевизионных систем для научных исследований (в первую очередь астрономических) и для обороны. Но в целом в период электронного телевидения, не смотря на выпуск тысяч «промышленных телевизионных установок», надежды на рождение «телевизионной автоматики» были не реальны, по существу массовым прикладное телевидение стало лишь с появлением твердотельных телекамер.

С момента рождения твердотельного телевидения на протяжении почти двух десятилетий центральным моментом разработки фотоприёмных матриц и телекамер на них была борьба с помехами, борьба за чувствительность и чёткость изображений. Не случайно существенный объём главы про телекамеры книги [1] был посвящён борьбе с детерминированными помехами матриц ПЗС, и лишь на четвёртом месте находились вопросы управления параметрами разложения изображения – регулировки чёткости и кадровой частоты для согласования телекамеры с освещённостью и динамикой сюжета.

В последние годы благодаря совершенствованию технологии изготовления телевизионных матриц, в первую очередь, благодаря освоению ультрафиолетовой литографии, достигнуто резкое увеличение чёткости (количества элементов в матрицах), снижение уровня собственных шумов и темнового сигнала, сокращение количества дефектов. Такое повышение качества видеосигнала выдвинуло перестройку параметров телекамер на более почётное место. Важно, что идея реализации перестройки параметров телекамер в период электронного телевидения (в 1970-е годы) была во ВНИИТе экспериментально подтверждена, но в силу громоздкости реализации развёртки с переменными параметрами и зависимости сигнала от скорости считывания практического применения не нашла. И только твердотельное телевидение открыло возможность оперативной регулировки и адаптации параметров разложения.

Твердотельное телевидение существенно шире его упрощённого понимания, и особенности реализации развёртки или особенности спектральных диапазонов не исключают систему из семейства твердотельных телекамер. Так, и инфракрасные космические системы наблюдения на матричных или линейных фотоприёмниках – это по сути телекамеры, и массово используемые телефаксы и компьютерные сканеры – это тоже малокадровые телекамеры на линейных твердотельных фотоприёмниках. Цифровые фотоаппараты, многие из которых имеют режимы съёмки видеоизображений, тоже становятся разновидностью телекамер с переменными параметрами – с обменом пространственной чёткости на временну́ю [41]. Любопытной особенностью твердотельной революции в фотографии является учёт традиций и инерционности пользователей, для которых не только фокусные расстояния объективов даются с указанием эквивалента для стандартного размера фотоплёнки, но и настройки чувствительности оперируют единицами чувствительности фотоплёнки.

Характерной особенностью начала второго этапа твердотельной революции в телевидении является одновременное присутствие на рынке телекамер на ПЗС и КМОП фотоприёмниках при тенденции вытеснения ПЗС [18–42, 53, 55–58]. Оба эти типа фотоприёмников имеют в основе кремниевую пластину, в силу чего они имеют близкие спектральные характеристики, оба под влиянием вычислительной техники приобретают форматы, совместимые с форматами, отображаемыми компьютерами. Во многих случаях, особенно в сфере применения массовых дешёвых телевизионных систем выбор между ними осуществляется не из технических, а из экономических соображений.

1.1. Телекамеры на ПЗС

За десятилетия победного шествия ПЗС выпущено огромное количество книг и статей по структуре, параметрам и возможностям телекамер (напр., [1, 2, 30, 43–52, 65, 74–85, 129–136]), поэтому изложение этого раздела – сугубо конспективное. Ещё раз подчеркнём, что в процессе рождения телевизионного сигнала крайне важны этапы фотоэлектрического преобразования, накопления зарядов, развёртка изображения и усиление электрических сигналов. И три этапа – фотоэлектрическое преобразование в слое полупроводника, накопление и усиление – освоены ещё на этапе электронно-лучевого телевидения. Более полувека назад В. К. Зворыкин прекрасно описывал с позиций зонной теории полупроводников (твёрдого тела) методы выбора материалов передающих телевизионных трубок, частности по параметрам красной границы фотоэффекта. Важные акценты В. К. Зворыкин делал и на параметрах *дискретизации сигналов* в телекамерах, подчёркивая, что *только учёт конечных значений разрешающей способности зрения по полю кадра и по времени делают электронное телевидение реализуемым*. Эффект накопления сигналов в полупроводниковой мишени электронно-лучевых трубок также был хорошо исследован. Даже понятие о паразитных сигналах и специфических искажениях передаваемых образов – не нововведение эпохи ПЗС, а описанные В. К. Зворыкиным особенности работы телекамер на электронно-лучевых трубках. Поэтому и не удивительно, что многие популярные публикации из разряда «ПЗС – это очень просто» сразу начинают рассказ о процессе *развёртки* изображения в матрицах ПЗС, проводя аналогии в переносе заряда с переносом жидкости с помощью последовательности «стаканов». Именно новый метод реализации развёртки породил использовавшийся на начальном этапе развития матричных фотоприёмников термин «самосканирующие приборы».

В ПЗС светочувствительным элементом (пикселом) может являться обедненная область МОП ёмкости, возникающая при подаче на фазный электрод обедняющего напряжения рис. 1.1 *а*, либо же обеднённая область обратносмещенного фотодиода, рис. 1.1 *б*. Светочувствительный элемент второго типа предпочтительнее, хотя бы с той точки зрения, что фотодиодная структура имеет существенно более высокий коэффициент сбора светового потока из-за отсутствия слоёв поликремния, поглощающего световой поток.

Чтобы реализовать в телекамере кодирование многомерного

сигнала (яркость – функция пространственных координат, времени и длины волны) для согласования с одномерным каналом, в котором сигнал является функцией только одной координаты – времени, в приборах с зарядовой связью накопленные заряды переносятся от каждого пиксела к соседнему вплоть до считывающего устройства (рис. 1.2).

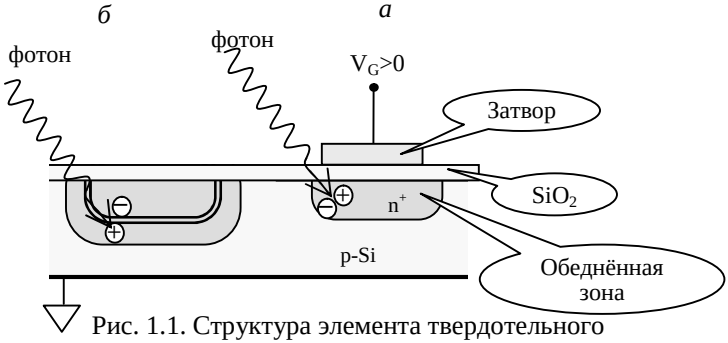


Рис. 1.1. Структура элемента твердотельного фотоприёмника

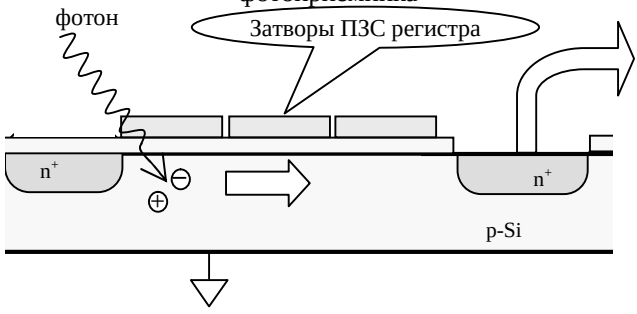


Рис. 1.2. Структура считывания сигнала в твердотельном фотоприёмнике

Для осуществления такого переноса под электродами регистров ПЗС формируются динамически изменяющиеся потенциальные ямы, под которыми концентрируются фотогенерированные электроны. Подробно фазы переноса зарядов показаны на рис. 1.3, где проиллюстрирована динамика изменения потенциальных ям под электродами регистра ПЗС. Структура ПЗС регистра состоит из слоя кремния *p*-типа (подложка), изолирующего слоя (диоксид кремния) и набора пластин – электродов. Один из электродов смещен более положительно, чем соседний, и именно под ним происходит накопление заряда.

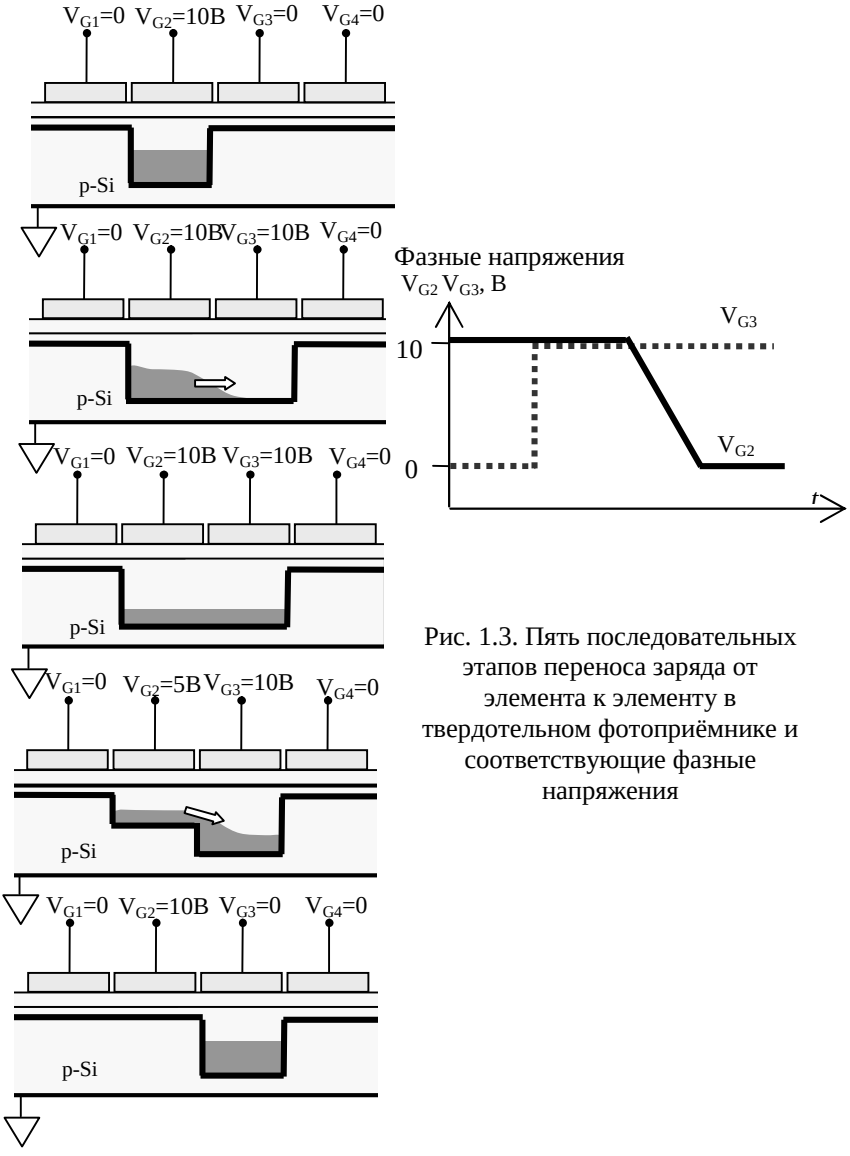


Рис. 1.3. Пять последовательных этапов переноса заряда от элемента к элементу в твердотельном фотоприёмнике и соответствующие фазные напряжения

Если подать небольшой положительный потенциал на чётные электроды регистра ПЗС, а на нечётные – нулевой потенциал относительно подложки, то под положительно смещенными электродами образуются потенциальные ямы. При поглощении кремнием фотона генерируется пара носителей заряда – электрон и дырка. Электростатическое поле в области элемента вытесняет «дырку» вглубь кремния, а электроны накапливаются в потенциальной яме под электродами, к которым подведен положительный потенциал. Здесь они могут храниться достаточно длительное время (в современных приборах без охлаждения до нескольких секунд), поскольку дырки в обедненной области отсутствуют и электроны не рекомбинируют. Заряд, накопленный под одним электродом, в любой момент может быть перенесен под соседний электрод, если его потенциал будет увеличен, в то время как потенциал первого электрода будет уменьшен (рис. 1.3). Если все электроды ПЗС находятся под нулевым потенциалом, то накопление генерированных светом электронов не происходит, хотя пары дырка – электрон под воздействием света рождаются.

Для создания *линейных* ПЗС, используемых при механическом сканировании по одной из координат – в факсах, сканерах, в космических системах наблюдения – такого регистра ПЗС оказывается достаточно, но для реализации полностью электронного сканирования (развёртки) непосредственно в твёрдом теле полупроводника требуется реализация *матрицы* с двумя ортогональными регистрами переноса зарядов.

Двухмерный массив (матрицу) элементов получают с помощью стоп-каналов, разделяющих электродную структуру ПЗС на столбцы. Стоп-каналы – это узкие области, которые формируются специальными технологическими приёмами в приповерхностной области и препятствуют растеканию заряда под соседние столбцы.

Большинство типов ПЗС матриц массового применения состоят из двух областей – области накопления и области хранения. Ниже рассмотрены особенности реализации развёртки изображения в матрицах ПЗС двух основных топологий [1, 43, 47, 52, 111, 131–134]:

- с кадровым переносом;
- со строчным переносом.

В англоязычной литературе их обозначают, соответственно, индексами: FT (Frame Transfer), IT (Interline Transfer). Исторически первыми были разработаны матричные ПЗС с кадровым переносом (рис. 1.4.).

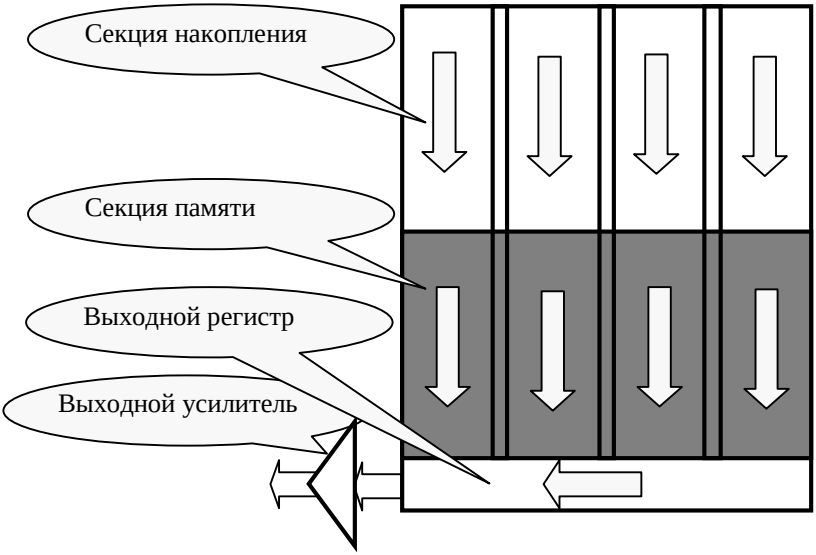


Рис. 1.4. Матричный ПЗС с кадровым переносом

Этот прибор содержит светочувствительную секцию накопления, состоящую из вертикальных ПЗС регистров, отделенных друг от друга областями стоп-каналов. Непосредственно к секции накопления примыкает секция памяти, имеющая ту же структуру и то же число элементов, что и секция накопления. В конце секции памяти расположен горизонтальный ПЗС регистр. Число элементарных ячеек этого регистра равно числу столбцов в секциях накопления и памяти. В конце горизонтального регистра расположено выходное устройство. Секция памяти, выходной горизонтальный регистр и выходное устройство экранируются от падающего света напылением алюминиевого покрытия на поверхность кристалла.

В течение времени накопления на один электрод секции накопления (в трёхфазных ПЗС для увеличения чувствительности возможно и на два электрода каждого пиксела) подаётся

положительный потенциал, образующий двухмерный массив обеднённых областей. Генерируемые падающим светом носители заряда собираются в близлежащих обеднённых областях, причём до переполнения потенциальной ямы их количество линейно зависит от экспозиции. По окончании времени накопления во время обратного хода по кадру на фазные электроды секции накопления и секции памяти подаются импульсы переноса. Накопленный двухмерный массив зарядовых пакетов за время переноса параллельно сдвигается из секции накопления в секцию памяти. Так как число ячеек в секции памяти равно числу ячеек в секции накопления, то каждый накопленный в секции накопления зарядовый пакет займёт соответствующее место в секции памяти. После того как все зарядовые пакеты перенесены в секцию памяти, процесс накопления зарядовых пакетов возобновляется.

Одновременно с этим в секции памяти во время обратного хода по строке зарядовые пакеты построчно переносятся в горизонтальный выходной регистр. Для этого на один из электродов горизонтального регистра подаётся положительный потенциал, соответствующий образованию под ним потенциальной ямы, а на электроды секции памяти подаются импульсы, соответствующие параллельному сдвигу массива зарядовых пакетов на один трёхфазный элемент по направлению к горизонтальному регистру. Таким образом, очередная строка зарядовых пакетов оказывается в горизонтальном регистре, а следующая за ней – располагается в той строке секции памяти, которая непосредственно прилегает к регистру.

Далее, во время прямого хода по строке, зарядовые пакеты считываются из горизонтального регистра через выходное устройство, для чего на электроды этого регистра подаются импульсы переноса. По окончании считывания строки зарядовых пакетов из секции памяти сдвигается следующая строка и процесс повторяется. После того, как будут считаны все строки из секции памяти, в неё возможен перенос массива зарядовых пакетов, накопленного в секции накопления. Для реализации чересстрочной развёртки накопление в различных полукадрах реализуется под различными фазными электродами секции накопления.

Матричные ПЗС с кадровым переносом при специально выполненном уменьшении толщины кристалла могут освещаться со стороны подложки. Это обеспечивает высокий квантовый выход – до 98%. Благодаря непосредственному облучению светом

поверхности кремния эти матрицы обладают высокой чувствительностью в ближней инфракрасной области спектра. Это обусловливает широкое применение телекамер на базе матриц с кадровым переносом в астрономии и в охранных системах для задач ночного наблюдения, в том числе с ИК подсветкой.

Современные матричные ПЗС с кадровым переносом последнего поколения, применяемые в астрономии, имеют, как правило, несколько раздельно управляемых по вертикали секций накопления заряда, несколько горизонтальных регистров переноса заряда и несколько выходных устройств. Возможность переноса заряда в обоих направлениях по вертикали в секциях накопления и по горизонтали в регистре переноса позволяют повысить суммарную скорость считывания кадра и делают эти приборы более предпочтительными для задач, требующих одновременного выполнения условий высокой чувствительности, высокого быстродействия и высокого динамического диапазона. Примером такого прибора может служить выполненный по технологии Full Frame матричный ПЗС производства фирмы Atmel AT71201M. Фоточувствительная матрица прибора представляет собой квадрат со стороной 45 мм (диагональ 2½") и количеством элементов 4096×4096. Размер элемента 11×11 мкм. Структура ПЗС AT71201M и основные режимы считывания показаны на рис. 1.5. Данный прибор имеет четыре раздельно управляемые секции накопления, два горизонтальных регистра и четыре выходных устройства. При этом горизонтальные регистры также разделены на независимо управляемые области переноса. В данном приборе имеется возможность одновременного переноса части зарядовых пакетов в сторону одного выходного устройства, а оставшейся – в противоположную сторону. Спектральная чувствительность прибора лежит в пределах длин волн 400…1100 нм с максимумом на длине волны 770 нм. Максимальная частота переноса в горизонтальном регистре составляет 40 МГц с эффективностью 0.99998 на один перенос. Средний шум считывания составляет 25 электронов, динамический диапазон 5000, ёмкость потенциальной ямы фоточувствительного элемента порядка 120000 электронов. Средний темновой ток фоточувствительных элементов составляет 4 пА/см^2, а для горизонтального регистра 10 пА/см^2 при температуре 25°С. Это эквивалентно темновому току порядка 30 и 100 электронов на элемент разложения в секунду соответственно. При охлаждении прибора до температуры –40°С следует ожидать средний темновой ток меньше одного электрона на элемент

разложения за кадр при максимальной кадровой частоте, составляющей 7.4 кадра/с. Матричные ПЗС такого класса позволяют обеспечивать соотношение времени накопления и времени считывания, необходимое для исключения эффекта смаза для меньших времен накопления, а потому могут использоваться в быстродействующих телевизионных системах различного назначения, в том числе и в астрономических системах реального времени.

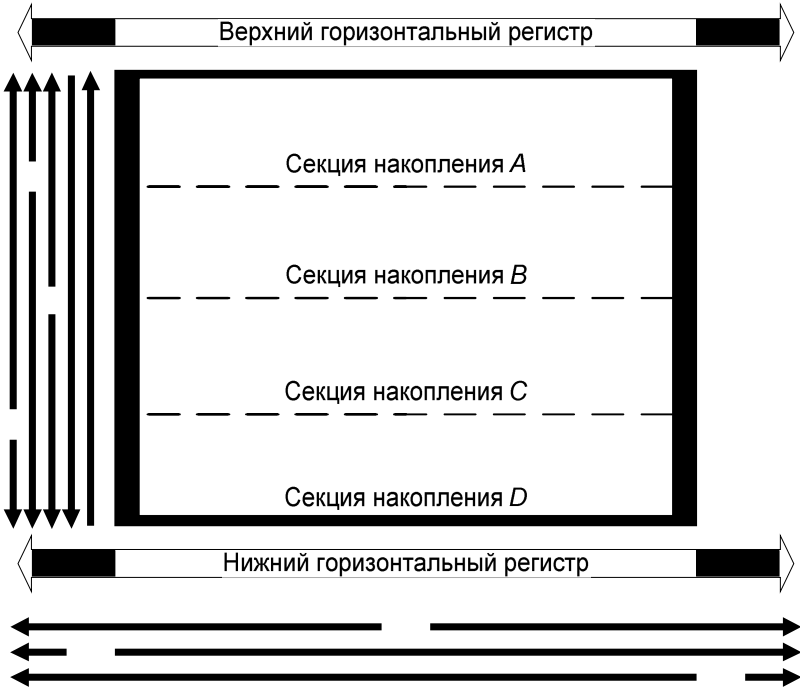


Рис. 1.5. Структура и режимы считывания матричного ПЗС AT71201M

Недостатком матриц ПЗС с кадровым переносом является вертикальный смаз от ярких деталей изображения. Во время переноса массива зарядовых пакетов из секции накопления в секцию памяти световой поток продолжает генерировать фотоэлектроны. Так как каждая потенциальная яма проходит весь столбец светочувствительной секции (частично до начала накопления, частично после), она неминуемо накопит некоторое количество зарядов от всех точек изображения, лежащих на данном

столбце. В реальных изображениях встречаются области, освещённость которых во много раз превышает средний уровень освещённости наблюдаемой сцены. Наличие этих областей приводит к образованию светлого вертикального столбца, являющегося продолжением яркой детали. Уровень вертикального смаза прямо пропорционален освещённости и площади яркой детали изображения, а также длительности переноса зарядов из секции накопления в секцию памяти и может превышать порог зрительного восприятия. Эффект вертикального смаза в этих матрицах может быть устранен только перекрытием светового потока на время переноса механическим затвором, который, в частности, применяется в цифровых фотоаппаратах при любых организациях фотоприёмных матриц. Дополнительным фактором, ограничивающим сферу применения ПЗС с кадровым переносом, является значительная площадь матрицы, обусловленная наличием секции памяти. Поскольку стоимость микросхем пропорциональна четвёртой степени диагонали кристалла, то цена этих ПЗС достаточно высока.

Более поздней разработкой матричных приборов является ПЗС со строчным переносом. Для накопления зарядовых пакетов в них используются столбцы обратносмещенных фотодиодов *p*-типа (Hole-accumulation diode, HAD). В непосредственной близости от каждого столбца фотодиодов находится нечувствительный к свету вертикальный ПЗС регистр, отделенный от фотодиодов фотозатвором. В первых матрицах ПЗС со сточным переносом роль фотозатвора выполнял отдельный поликремниевый электрод. В более поздних разработках его роль выполняет часть затвора вертикального ПЗС регистра, выступающая за край скрытого канала переноса зарядов. В конце вертикальных ПЗС регистров расположен горизонтальный ПЗС регистр с выходным устройством (рис. 1.6). Все регистры ПЗС – вертикальные и горизонтальный – экранированы от падающего света. Во время накопления зарядовых пакетов в фотодиодах на фотозатвор подается низкий потенциал, обеспечивающий потенциальный барьер между фотодиодами и вертикальным регистром. По окончании накопления на фотозатвор кратковременно подаётся положительный потенциал, разрешающий перенос зарядовых пакетов из фотодиодов в потенциальные ямы, образованные в вертикальных регистрах. Затем с фотозатвора снимается положительное

смещение, и накопление зарядовых пакетов в фотодиодах возобновляется.

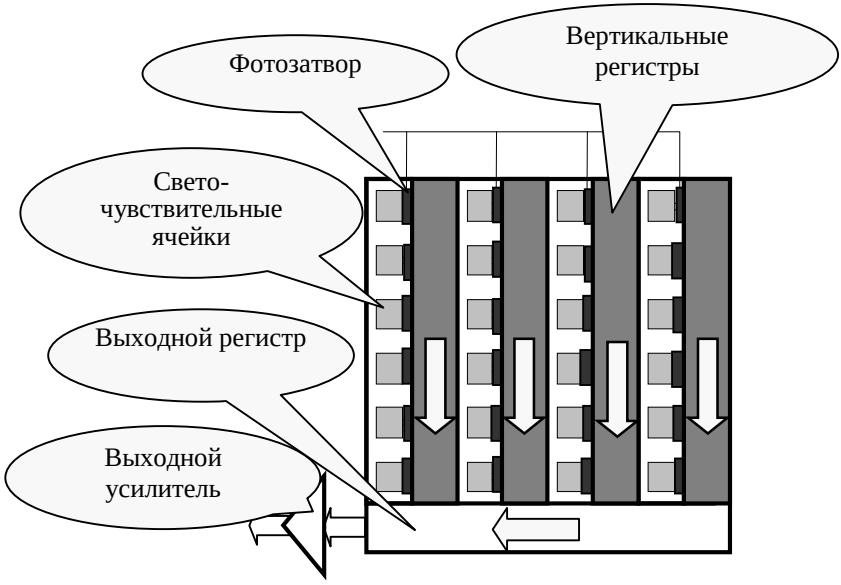


Рис. 1.6. Матричный ПЗС со строчным переносом

Зарядовые пакеты из вертикальных ПЗС регистров построчно переносятся в горизонтальный ПЗС регистр, из которого поэлементно считываются через выходное устройство. Перенос из светочувствительных фотодиодов в вертикальные регистры осуществляется во время обратного хода по кадру, а перенос зарядовых пакетов из вертикальных регистров в горизонтальный регистр – во время обратного хода по строке. После того как все строки зарядовых пакетов будут считаны, возможен перенос следующего двухмерного массива зарядовых пакетов из фотодиодов.

Достоинством матричных ПЗС со строчным переносом является малый уровень смаза, связанный с тем, что перенос всех зарядовых пакетов в защищённые от света вертикальные ПЗС регистры происходит в течение короткого промежутка времени. Основной недостаток матричных ПЗС со строчным переносом – неполное использование светового потока вследствие наличия нечувствительных вертикальных регистров. Для преодоления этого недостатка была реализована технология нанесения на поверхность

кристалла микролинз (рис. 1.7, а). При использовании этой технологии над каждым элементом матрицы ПЗС расположена микролинза, собирающая на элемент световой поток, падающий на не чувствительные к свету регистры переноса заряда. В развитие этого метода фирма Sony разработала технологию Hyper-HAD, позволяющую повысить чувствительность матриц ПЗС за счёт увеличения размера микролинз (рис. 1.7, б).

Примером может служить матричный ПЗС производства фирмы Sony ICX415AL с диагональю ½", количеством фоточувствительных элементов 782 × 582 и размером элемента 8,3 ×8,3 мкм. Максимальная частота переноса в горизонтальном регистре составляет 29,5 МГц с эффективностью 0.995 на один перенос, что позволяет достигать кадровой частоты 50 кадров/с. Спектральная чувствительность прибора лежит в пределах длин волн 400 … 1000 нм с максимумом на длине волны 500 нм. Ёмкость потенциальной ямы фоточувствительного элемента составляет порядка 40000 электронов. Средний темновой ток составляет 180 пА/см^2 при температуре 25°С.

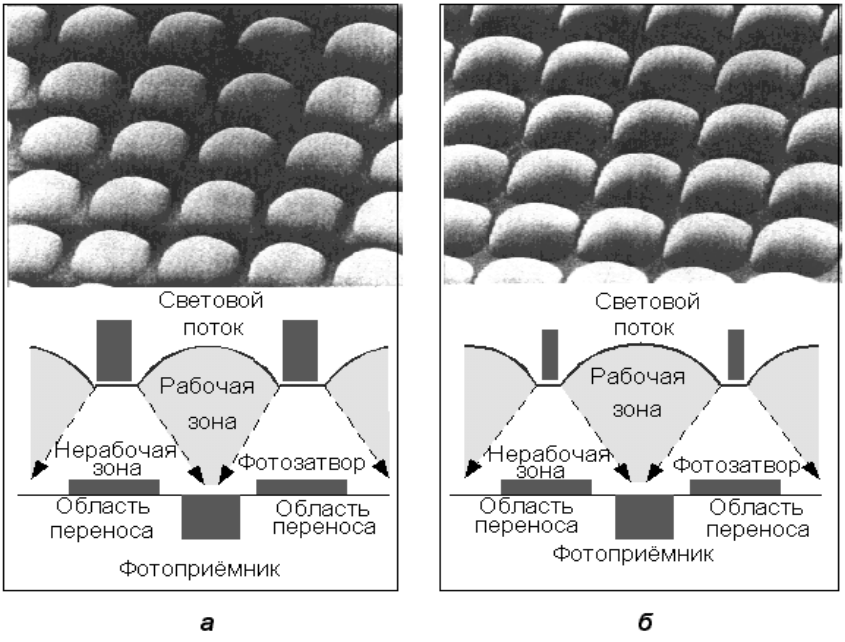


Рис. 1.7. Структура матриц ПЗС: *а* – стандартная, *б* – Hyper-HAD

Совокупность технологических приёмов, включающих использование микролинз, и, главное, повышение точности изготовления матриц ПЗС благодаря освоению ультрафиолетовой литографии, привели к тому, что по сравнению с начальным этапом твердотельной революции в телевидении существенно изменились шумовые характеристики фотоприёмников. Если на начальном этапе развития ПЗС преобладали различные источники шумов, связанные в первую очередь с несовершенством приборов (шумы темновых токов, шумы переноса), то в настоящее время матричные ПЗС вышли на уровень преобладания фотонного шума сюжета. Конечно, иногда встречаются экзотические сюжеты с контрастом, близким к 100%, но практически при всех реальных сюжетах, даже во многих астрономических приложениях, контраст нельзя считать стопроцентным. А это означает, что фотонный шум фона в современных твердотельных матричных фотоприёмниках, как правило, превышает собственные шумы матрицы.

В результате характеристика уровня шумов телекамеры в зависимости от освещённости сцены принимает вид, существенно отличный от типичной для электронно-лучевого телевидения, где чаще всего преобладали шумы видеоусилителя. И только в редких случаях малокадровых телевизионных систем удавалось увидеть разницу в уровне шума на тёмных и на светлых участках изображения.

Важнейшей характеристикой матричного ПЗС как регистрирующего устройства является чувствительность – минимальная освещённость на объекте, при которой достигается заданное качество изображения. Критерий качества при этом может быть различным и зависеть от задачи, поставленной перед телевизионной системой. Например, в задаче обнаружения сигналов – это может быть отношение сигнал/шум. Чувствительность матричного ПЗС характеризуется несколькими параметрами. Первый параметр – квантовая эффективность, определяющая интегральную чувствительность. В современных ПЗС при использовании микролинз удаётся достичь высокой квантовой эффективности около 60…70%, что обеспечивает при вещательных параметрах разложения пороговую освещённость на матрице, исчисляемую сотыми долями люкса. Этот параметр позволяет грубо оценить чувствительность телекамеры в целом, хотя и нуждается в уточнении в каждом конкретном случае исходя из цели системы, которая определяет важные *переменные* (регулируемые)

свойства телекамер – разрешающую способность по полю и по времени и *постоянные* – пороговое отношение сигнал/шум и коэффициент использования светового потока, определяемый соотношением спектральных характеристик излучения и фотоприёмной матрицы. Для фотоприёмников видимого диапазона спектра, в частности ПЗС, в паспортах на приборы указывают характеристику чувствительности, связывающую выходное напряжение фотоприёмника U с экспозицией, равной произведению освещённости E и времени накопления $T_н$. Обозначается она латинской буквой S (от англ. sensitivity – чувствительность) и измеряется в вольтах, делённых на люкс×секунда): $S = U/ET_н$. Эта характеристика фотоприёмника зависит от площади элемента разложения: чем больше площадь, тем больше сигнал при той же экспозиции.

Ядром вычисления любого из показателей чувствительности является учёт эффекта накопления, проявляющегося в том, что потенциальное отношение сигнал/шум определяется квадратным корнем из числа накопленных электронов, генерированных квантами света. Для характеристики систем, чувствительных в видимой области спектра, плотность потока фотонов характеризуют освещённостью, измеряемой в люксах. Для характеристики яркости звёзд пользуются понятием звёздной величины m, связанной с создаваемой ей освещённостью E логарифмической зависимостью, причём принято, что освещённость в 1 лк создаётся объектом, звёздная величина которого $m_0 = -13{,}89$: $m = -13{,}89 - 2{,}5\,lg\,E$.

Изменение яркости источника на одну звёздную величину означает изменение её в 2,5 раза, а на 5 звёздных величин – в 100 раз. При расчёте потенциальной чувствительности необходимо перейти от любого из возможных способов задания облучённости к плотности потока фотонов. Соотношения между различными способами определения плотности потока фотонов приведены в табл. 1.1, в которой облучённость пересчитана в количество фотонов при упрощающем предположении монохроматического излучения той же плотности энергии, что и в иных способах измерения. Данная таблица может быть продолжена и далее, так как астрономические системы могут осуществлять поштучный счёт фотонов при времени наблюдения, измеряемом часами, а внеатмосферные астрономические системы могут накапливать фотоны в течение нескольких суток.

Таблица 1.1

Освещённость E, лк	Звездная величина m	Облучённость в видимой области спектра, Вт/см^2	N фотонов, 1/(с·см^2) (λ = 0,5мкм)
$1{,}2\times10^5$ (Солнце)	–26,6	0,07	3×10^{17}
1	$m_0 = -13{,}89$	$4{,}7\times10^{-7}$	10^{12}
0,3 (Луна)	– 12,6	$1{,}5\times10^{-7}$	3×10^{11}
10^{-2}	– 8,9	$4{,}7\times10^{-9}$	10^{10}
10^{-4}	– 3,9	$4{,}7\times10^{-11}$	10^{8}
10^{-6}	1,1	$4{,}7\times10^{-13}$	10^{6}
10^{-8} (порог зрения)	6,1	$4{,}7\times10^{-15}$	10^{4}
10^{-10}	11,1	$4{,}7\times10^{-17}$	10^{2}
10^{-12}	16,1	$4{,}7\times10^{-19}$	1

При слабых потоках фотонов при оценке потенциальной чувствительности необходимо учитывать не только сигнальные фотоны, но и фоновые. Из-за рассеяния света в атмосферах планет, в том числе Земли, фон не является нулевым, небо даже ночью не абсолютно чёрное. Так, для наземных наблюдений в безлунную ночь яркость площадки неба площадью в одну квадратную секунду дуги примерно равна яркости звезды 22-й звёздной величины. Фон неба при выносе телевизионной астрономической аппаратуры за пределы атмосферы уменьшается незначительно, приблизительно всего вдвое, так как кроме рассеяния света в атмосфере существует зодиакальный свет, излучение звёзд Млечного пути и освещённых ими туманностей. В условиях наличия фона чувствительность телевизионной системы определяется флуктуациями потока фотонов в сигнале и в фоне. В силу независимости флуктуаций числа фотонов на полезном сигнале и на фоне отношение сигнал/шум Ψ определится из среднего накопленного числа сигнальных N_s и фоновых N_b фотонов:

$$\Psi^2 = N_s^2 \,/\,(N_s + N_b).$$

Фундаментальным фактором, ограничивающим чувствительность, является шум, имеющий по крайней мере два источника: шум квантовой структуры света и собственный шум телекамеры. При идеальной нешумящей телекамере её чувствительность зависит

только от фотонного шума – флуктуаций числа фотонов относительно среднего значения. Из-за независимости фотонов в потоке света количество фотонов в элементе разложения в конкретном кадре подчинено статистике Пуассона. В результате достижимое отношение сигнала к фотонному шуму пропорционально корню квадратному из числа фотонов, принятых телекамерой от объекта наблюдения. Здесь в первом приближении исключаются из рассмотрения зависимость энергии излучения, коэффициентов отражения и квантового выхода от длины волны. Сюжет в поле зрения ТВ камеры характеризуется коэффициентами отражения объекта $k_о$ и фона $k_ф$; способность фотоприёмника к накоплению определяется площадью элемента разложения Δ^2, квантовым выходом ŋ, усреднённым по длине волны, и временем накопления $T_н$. Для телекамер, используемых в составе систем наблюдения, важно задать пороговое отношение сигнал/шум $\Psi_{пор}$, обеспечивающее заданную достоверность принятия решения наблюдателем. Так как уровень фотонного шума зависит от накопленного сигнала, то в задаче обнаружения сигнала объекта порог принятия решения располагается не симметрично, а ближе к менее освещённой детали (объекту или фону). При определении пороговой освещённости это учитывается усреднением среднеквадратичных значений шума на сигнале и на фоне. В первом приближении при не слишком большой разнице между числом накопленных от объекта и от фона фотонов в качестве оценки дисперсии шума можно пользоваться полусуммой дисперсий шума по каждой гипотезе, что ведёт к учёту полусуммы коэффициентов отражения.

Учёт этих параметров позволяет найти выражение для предельной чувствительности, т. е. пороговой освещенности сцены для реальной ситуации обнаружения и опознавания малоконтрастных объектов:

$$E_{пор} = \frac{2\Psi_{пор}^2 (k_ф + k_о)}{(k_о - k_ф)^2 \left(\frac{D}{f}\right)^2 \tau N \eta \Delta^2 T_н},$$

где τ – коэффициент пропускания света объективом; D – эффективный диаметр входного зрачка объектива; f – фокусное расстояние объектива, $N \approx 2 \cdot 10^{12}$– потенциально доступное ПЗС количество фотонов на $1\,см^2$ в $1\,с$ при равномерном спектре и освещённости в видимом диапазоне 1 лк. При большой разнице

в коэффициентах отражения объекта и фона в числитель вместо коэффициента отражения фона $k_ф$ следует подставить полусумму коэффициентов отражения объекта и фона (это соответствует применению порога принятия решения о наличии объекта, обеспечивающего равенство ошибок из-за пропуска объекта и из-за ложной тревоги).

Почти неизменяемыми в процессе проектирования параметрами являются квантовая эффективность и пороговое отношение сигнал/шум. Квантовая эффективность большинства серийных ПЗС составляет 0,5…0,6 в диапазоне видимых длин волн. Усреднённая по всему спектральному диапазону 0,4…1,1 мкм квантовая эффективность кремниевых ПЗС меньше примерно вдвое. Площадь элемента разложения в матрицах ПЗС имеет тенденцию к уменьшению (наиболее распространены матрицы с площадью элемента 10…50 мкм^2, для мобильных телефонов освоен выпуск матриц с площадью элемента на порядок меньше – около 2×2 мкм), и только в астрономических системах, имеющих большие значения кружка рассеяния, используются матрицы с площадью элемента 200…400 мкм^2. Пороговое отношение сигнал/шум – это константа, зависящая от свойств сигнала (в первую очередь – от его площади) и от принимаемых наблюдателем вероятностях ошибки принятия решения (пропуска сигнала и ложной тревоги). Типичные значения порогового отношения сигнал/шум находятся в пределах 2…10.

Контраст объекта относительно фона обычно определяется как отношение разности уровней объекта и фона к их сумме. Контрастная чувствительность телевизионной системы, определяемая через минимальный контраст $K_{пор}$ обнаруживаемых деталей изображения, зависит от тех же параметров, что и чувствительность:

$$K_{пор}^2 = 2\Psi_{пор}^2 \Big/ \left[E(k_o + k_ф) \left(\frac{D}{f} \right)^2 \tau N \eta \Delta^2 T_н \right]$$

Это потенциальное значение порогового контраста достигается далеко не всегда. Часто оно ограничивается накопительной возможностью элемента фотоприёмника (его ёмкостью или предельным числом $N_э$ фотоэлектронов). Это означает, что фотоприёмники могут накапливать сигналы лишь за время, при котором сохраняется линейность преобразования свет–сигнал. Пороговый контраст телекамеры на ПЗС определяется простой формулой:

$$K_{пор} = \Psi_{пор} / \sqrt{N_э}.$$

Поэтому для повышения контрастной чувствительности следует применять фотоприёмники с большей ёмкостью, и, следовательно, с большей площадью элемента разложения. Обычно ёмкость элемента ПЗС составляет около 1000 электронов на квадратный микрометр. Практически достигаемое значение порогового контраста в телекамерах на матричных ПЗС составляет от 0,5% – для ПЗС с размером элемента 14×14 мкм, до 2% – для матриц размером элемента 3,3×3,3 мкм. Это означает, что контрастная чувствительность матриц малого размера становится хуже 1%, т. е. хуже контрастной чувствительности зрения. То есть квантовая структура света через посредство фотонного шума ставит предел уменьшения размеров элемента фотоприёмной матрицы, и матрицы с малым размером элемента порядка 1,5×1,5 мкм (пороговый контраст примерно 5%) станут самыми малогабаритными. Важно, что и волновая структура света из-за дифракции на входном зрачке объектива ограничивает размеры матрицы полем зрения около 1×1 мм, что при оцененном предельно малом размере элемента даёт чёткость примерно равную телевизионному стандарту.

Второй параметр, характеризующий чувствительность телекамеры – монохроматическая чувствительность, т. е. спектральная характеристика зависимости чувствительности от длины волны света. Практически все ПЗС (и КМОП [107]) имеют спектральную характеристику чувствительности в диапазоне от 0,3…0,4 мкм до 1,1 мкм с некоторыми вариациями в коротковолновой (синей и ультрафиолетовой) области и ближней ИК области спектра (рис. 1.8).

Доминирующая роль ПЗС видна даже из такого примера. Идеологи использования телевизионных систем для исследования Земли из космоса считают [129], что спектральный диапазон 0,4…1,0 мкм является *оптимальным* для решения многих практических задач, связанных с изучением растительности и почв. Это утверждение нуждается в комментариях:

- употребление термина «оптимальный» здесь некорректно, так как никакой математической задачи на поиск экстремума качества авторы не решали, они лишь считают, что этот диапазон соответствует некоторому (ими не формализованному) компромиссу;

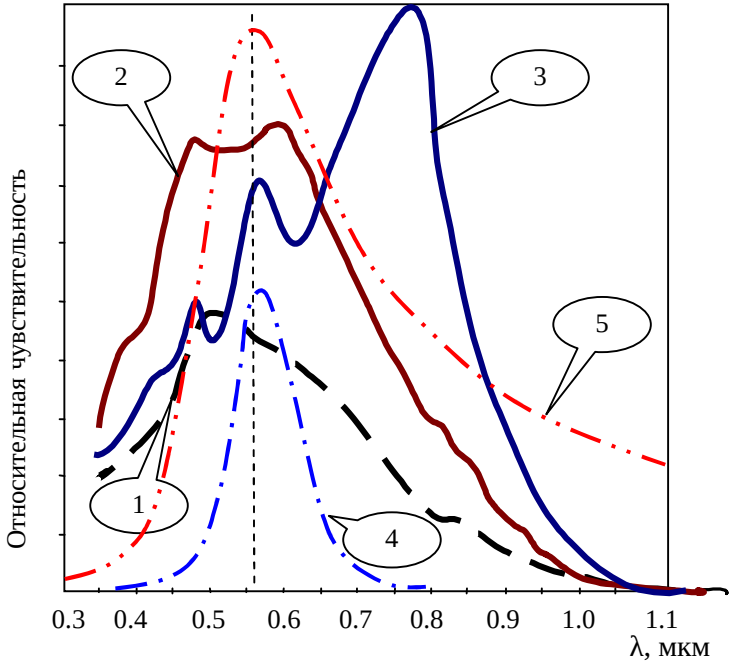


Рис. 1.8. Спектральные характеристики чувствительности ПЗС: *1* – со строчным переносом, *2* – Exview ПЗС, *3* – ПЗС с кадровым переносом; для сравнения приведены спектральная характеристика зрения *4* и спектральная характеристика излучения абсолютно чёрного тела с температурой 6000К, которая служит огибающей спектра излучения Солнца *5*

- указанный спектральный диапазон действительно охватывает участок спектра, содержащий максимум излучения Солнца (555 нм) и почти 60% энергии этого излучения (рис. 1.8);
- указанный спектральный диапазон полностью охватывает участок спектра, видимый зрительным анализатором человека (от 0,38 мкм до 0,78 мкм с максимумом на длине волны 555 нм), рис. 1.8
- указанный спектральный диапазон соответствует диапазону чувствительности кремниевых ПЗС (рис 1.8) – массово выпускаемых фотоприёмников с хорошими характеристиками квантовой эффективности, линейности, числа элементов разложения, ценой и т. д.

Если бы электронная промышленность выпускала дешёвые ПЗС из другого полупроводника (например, арсенида галлия) такого же высокого качества, но с достигаемой без охлаждения красной границей 1,7 или 2,2 мкм, то в статье [129] авторы с ещё большим удовольствием (процент использования энергии солнечного излучения в данном случае ещё больше) назначили бы на роль «оптимального» диапазона этот другой, технологически обеспеченный диапазон.

Для задач, требующих повышенного быстродействия, в астрономических системах применяются матричные ПЗС со строчным переносом заряда, в специальных случаях в считывающем регистре организуются несколько выходных устройств (более четырёх, показанных на рис. 1.5).

В настоящее время производство ПЗС различной организации освоило большое число производителей, каждый из которых выпускает множество типоразмеров матричных и линейных ПЗС. Выпускаются матрицы и большого размера для научных исследований (с диагональю, определяющей класс матрицы, 2/3 дюйма и более – уже до 2,5 дюймов) и малого размера для любительских видеокамер, охранных телевизионных систем и мобильных телефонов со встроенными фотоаппаратами и видеокамерами – класса от 1/3 до 1/10 дюйма. Например, в середине 2006 г. ведущая фирма по выпуску малоформатных ПЗС – фирма Sony – выпускала 8 типов цветных и чёрно-белых матриц класса 2/3 дюйма (диагональ 11 мм), 32 типа цветных и чёрно-белых матриц класса 1/2 дюйма (диагональ 8 мм), 21 тип цветных и чёрно-белых матриц класса 1/3 дюйма (диагональ 6 мм), 18 типов цветных и чёрно-белых матриц класса 1/4 дюйма (диагональ 4 мм), 2 типа цветных матриц класса 1/6 дюйма (диагональ 3 мм). Одна из фирм, ведущих по выпуску ПЗС научного назначения – фирма Kodak – выпускала 15 типов матриц с чересстрочной развёрткой (KAI) и 24 типа матриц с полной засветкой (KAF). Ведущими фирмами по выпуску ПЗС с высокой скоростью считывания (до 50 МГц), востребованных в научном и космическом приборостроении, являются фирмы Atmel (выпускающая линейные ПЗС) и Dalsa (выпускающая матричные и линейные ПЗС). В таблице 1.2 приведены сведения о некоторых ПЗС матрицах, дающих представление о разнообразии освоенных типов этих приборов.

Таблица 1.2

Наименование матрицы	Кол-во пикселов	Размер диагонали мм	Эффективные пикселы (H×V)	Частот кадров, Гц	Размер пиксела, мкм
Sony					
ICX419ALL	440K	8	752×582	30	8,4×8,4
ICX229AK	440K	4	752×582	50	4,2×4,2
ICX238AKE	380K	3	768×494	50	3,1×3,1
ICX205AK	1,4M	8	1392×1040	15	4×4
ICX252AK	3,2M	8	2078×1558	15	3×3
Texas Instrument					
TC253SPD-30	325K	6,08	656×496	60	7,4×7,4
TC285SPD-30	1M	11,2	1004×1002	20	8×8
Dalsa					
FTT1010M	1M	16,8	1024×1024	60	12×12
FTF2020M	4M	34,44	2048×2048	30	12×12
FTF3020C	6M	44,4	3072×2048	20	12×12
FTF4027M	11M	43,3	4008×2672	7,5	9×9
FTF4052M	22M	60,1	4008×5344	3,6	9×9
FTF5066C	33M	59,9	4992×6668	2	7,2×7,2
Kodak					
KAI-0340	VGA	5,9	640×480	110/210	9×9
KAI-1011	1M	12,9	1008×1018	15/30	9×9
KAI-2020	1,9M	14,8	1600×1200	15/30	7,4×7.4
KAI-4021	4,2M	21,4	2048×2048	8/15	7,4×7,4
KAI-16000	16M	43,4	4872×3248	1/3	7,4×7,4
KAF-1602	1,6M	16,6	1536×1024	2,2	9×9
KAF-16802	16,6M	51,9	4080×4080	0,5	9×9
KAF-39000	39M	61,3	7216×5412	0,9	6,8×6,8
Thomson					
TH 7895MVB	250K	13,8	512×512	20	19×19
THX 7899M	4M	40,6	2048×2048	5	14×14
ЦНИИ «Электрон»					
ФППЗ 32M	0.45M	20,7 ×15,7	768×580	33	27×27
ISD-075AP	1.6M	17.12×18.2	1225×1300	3	14×14
ФППЗ 12M1	1.2M	13.3×15	1024×1152	16	13×13

В результате совершенствования технологии матричных ПЗС они применяются в самых разнообразных приложениях – от простых задач наблюдения в стационарных условиях нормального освещения, типичного для офисных систем безопасности [46, 111], до астрономических наблюдений при большом времени накопления, обеспечиваемого глубоким охлаждением матриц ПЗС [43, 44, 47]. Если для достижения времени накопления, измеряемого часами и

Рис. 1.9. Астрономические телекамеры на охлаждаемых с помощью термоэлектрических холодильников матричных ПЗС, предназначенные для использования совместно с ЭОП в составе астрономического многоканального телевизионно-компьютерного комплекса) [123]; матрицы ПЗС класса ½" фирмы Sony

сутками, требуется охлаждение до криогенных температур, то для достижения времени накопления, исчисляемого минутами, достаточно охлаждения с помощью термоэлектрических холодильников, основанных на эффекте Пельтье (рис. В.5, рис. 1.9) [44, 47, 121]. В свете акцента в настоящем издании на

твердотельную революцию в телевидении, подчеркнём, что холодильник на эффекте Пельтье в отличие от холодильных машин (например, работающих по циклу Стирлинга), являются типичными твердотельными приборами. Важными направлениями работ в области создания телекамер на матричных ПЗС являются обеспечение наблюдений при малых контрастах изображения, как, например, под водой (рис. 1. 10) [17] и при большой динамике сюжета, имеющей место, в частности, при анализе волнового фронта, искажённого турбулентностью атмосферы (рис. 1.11) [124].

Рис. 1.10. Телекамера для морских применений; матрицы ПЗС класса ½" фирмы Sony

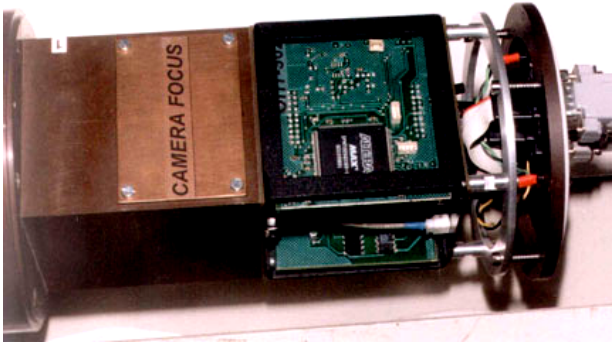


Рис. 1.11. Телекамера для астрономических применений в составе скоростного датчика волнового фронта (кадровая частота до 200Гц); матрицы ПЗС класса ½" фирмы Sony

Современные телекамеры рис. 1.9 – рис. 1.12, также как и массово выпускаемые бескорпусные телекамеры, составляющие основу относительно простых бытовых телекамеры (рис. 1.13), по конструкции резко отличаются от первых телекамер на ПЗС (рис. 1.14), выполненных в традиционном для видиконного камеростроения виде.

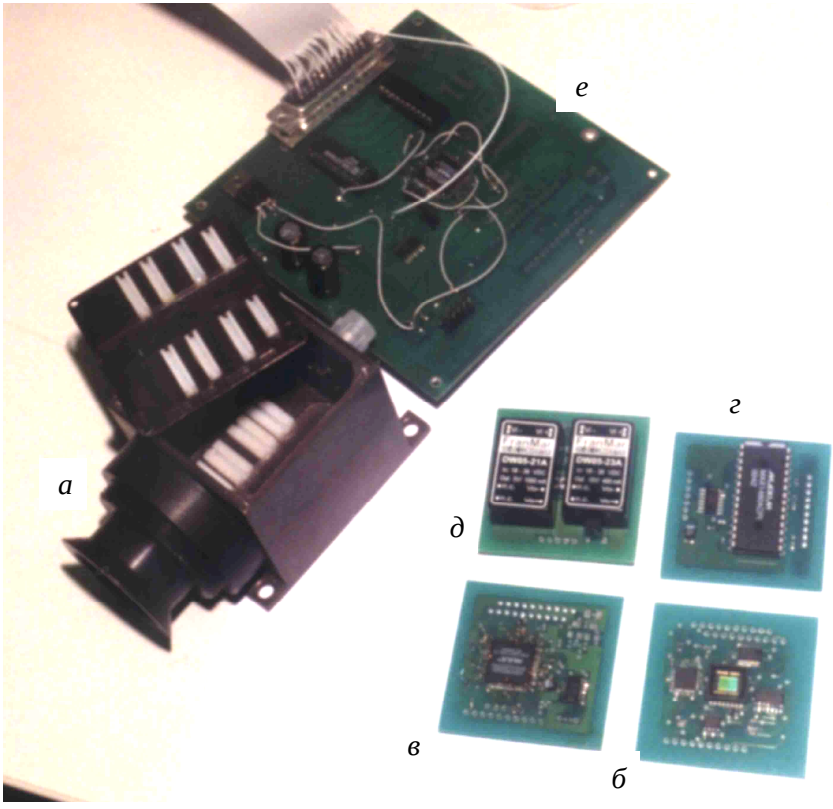


Рис. 1.12. Типовая бортовая телекамера
на матричном ПЗС класса 1/4" фирмы Sony;
корпус *а*, плата матрицы ПЗС и видеотракта *б*, плата синхронизации и адаптации времени накопления от 10^{-4} до 5 с *в*, плата интерфейса RS-485 *г*, плата формирователей питающих напряжений телекамеры из бортовой сети 27 В *д*, плата приёмника-декодера *е*

В результате развития телекамер на ПЗС на первом этапе твердотельной революции в телевидении они стали главным типом

Рис. 1.13. Одноплатная бескорпусная телекамера на матричном ПЗС класса 1/3" фирмы Sony

Рис. 1.14. Первые отечественные телекамеры на матрицах НПО «Пульсар» в музее телевидения ФГУП «НИИТ»

фотоприёмников видимого диапазона и составили основу рынка всех приложений телевидения – от охранных телекамер, цифровых фотоаппаратов, сканеров и факсов до космических телескопов, таких как телескоп имени Хаббла. Если изобретение фотографии лишь усилило проницающую способность телескопов, то твердотельные матричные фотоприёмники на основе ПЗС в астрометрии и линейные ПЗС в спектрометрии в сочетании с компьютерами «закрыли» этап фотографии в астрономии, оставив ей лишь короткий в историческом смысле эпизод перевода уже имеющихся фотографий (порядка миллиона) в электронную форму для хранения в памяти компьютеров. Характерно, что такое преобразование делается также с помощью сканеров на твердотельных телевизионных приборах – линейных ПЗС, являющихся по сути телекамерами, которые в соответствии с «золотым» правилом малокадрового телевидения при преобладании требования высокой разрешающей способности по полю зрения над требованием

высокой слитности изображения реализуются с использованием механического сканирования [4].

После того, как твердотельное телевидение в виде телекамер на матричных ПЗС впервые объединили в себе достоинства приборов с кадровым накоплением (видиконы), приборов с линейной свет–сигнальной характеристикой (диссекторы), и когда была решена проблема борьбы с неоднородностями темнового сигнала, могло показаться, что разработчики имеют в руках идеальный прибор. Однако в твердотельном телевидении имеется много нерешённых проблем.

Первая группа проблем связана с обеспечением работы в большом диапазоне освещённостей, когда при очень низкой освещённости наблюдению мешают не только собственные шумы телекамер, но даже и свечение транзисторов выходного узла матрицы ПЗС, а при очень большом перепаде освещённости яркие объекты вызывают искажения, связанные с растеканием зарядов. Борьба с растеканием избыточных зарядов (антиблуминг) имеет большие, но ограниченные возможности: удаётся погасить сигналы, в тысячу раз превышающие сигнал насыщения, но полностью устранить влияние пересветок не удаётся (рис.1.15).

Рис. 1. 15. Пример изображения с типичным искажением растекания зарядов из-за локальной пересветки

Вторая группа проблем ПЗС связана с важнейшей тенденцией слияния телевизионной и компьютерной техники, требующей реализации не только многоканального накопления, но и многоканального считывания сигнала, и не только в привычной

ранее аналоговой форме, но и цифровой – цифровые кодеры источника и устройства обработки для принятия решений, часто выполняемые с использованием распараллеливания вычислений, предъявляют соответствующие требования и к источнику видеоинформации. Однако технология ПЗС оказалась не совместимой с технологией, используемой при изготовлении аналого-цифровых преобразователей и других необходимых элементов цифрового кодирования сигналов. Совмещение технологий оказалось возможным только при переходе к КМОП технологии.

1.2. Телекамеры на КМОП

Важная тенденция срастания телевизионной техники с компьютерной техникой, инициируемая специалистами-компьютерщиками, проявляется, в частности, в том, что лидерами новой технологии являются фирмы, до последнего времени не представленные на рынке разработок в области фотоприёмных устройств – такие как Micron, Atmel, Motorola и др., но имеющие богатый опыт в области разработок СБИС различного назначения, в том числе микросхем памяти, имеющих сотни миллионов транзисторов на кристалле. Важно также то, что указанное срастание телевизионной техники с компьютерной во многом стимулировано развитием космической техники, и можно считать закономерным, что первые матричные КМОП фотоприёмники появились в лаборатории NASA (национального космического агентства США) [21] .

На первом этапе твердотельной революции в телевидении ПЗС казались чудом, перевернувшим всё телевидение и обеспечившим по настоящему массовое его распространение, так как при переходе от электронно-лучевых трубок к матрицам телекамеры не только резко уменьшились по габаритам, не только стали легко сопрягаемыми с компьютерами, но и обрели удивительную стабильность характеристик. Характерно, что первое десятилетие развития ПЗС сопровождалось неким скептицизмом специалистов телевизионной техники в отношении ПЗС из-за ещё не преодолённых тогда трудностей: достижения в матрицах малых темновых токов, устранения дефектных пикселов, достижения высокой квантовой эффективности и малого уровня шума считывания и эффективности переноса зарядов. На втором этапе твердо-тельной революции в телевидении картина психологического восприятия нового поколения техники во многом повторяется: предложения по

концентрации усилий на новом классе приборов – КМОП фотоприёмниках часто отвергаются под предлогом временного их отставания от ПЗС по чувствительности (точнее – по уровню шумов и паразитных сигналов). То, что это явление временное, специалистам микроэлектронной техники было ясно давно, и ПЗС предрекалось вымирание, как динозаврам [28]. В связи с этой противоречивой ситуацией необходимо ответить на вопрос: чем КМОП сенсоры отличаются от ПЗС, почему КМОП перспективнее, чем ПЗС и что мы имеем уже сегодня на фронте создания КМОП телекамер?

В видимом и ближнем ИК диапазоне осуществлён качественный скачок благодаря переходу от главного типа фотоприёмников на ПЗС к КМОП фотоприёмникам. Принципиальное отличие этих приборов от ПЗС состоит в **методе реализации развёртки**. Координатная адресация к каждому элементу вместо протаскивания заряда каждого элемента сначала по всему столбцу, а потом по всей строке позволяет избежать искажений, связанных с неэффективностью переноса заряда. Телекамеры нового поколения, основанные на КМОП фотоприёмниках, объединяют в себе фотоприёмник, устройства развертки, квантования и обработки изображения, и в силу своей структурной сложности переходят в класс видеосистемам на кристалле [25, 26, 31–40, 53].

Новая технология создания КМОП фотоприёмников открыла перспективу заметного повышения качества формируемой видеоинформации благодаря возможности многоканального считывания (в пределе – с применением отдельного АЦП для каждого элемента разложения [30–32, 40, 53]) и введения усиления до воздействия шумов считывания. Развитие этих преимуществ КМОП фотоприёмников в будущем может вытеснить из практики формирования телевизионных сигналов не только вакуумные приборы усиления слабых оптических сигналов (электронно-оптические преобразователи, ЭОП), по и ПЗС.

Цифровые телевизионные камеры, использующие фотоприёмники, изготовленные по технологии комплиментарных структур метал–оксид–полупроводник (КМОП, английский термин CMOS) получили название КМОП сенсоров или КМОП камер. В английской технической литературе наибольшее распространение получила аббревиатура CIS (CMOS Image Sensor). Использование технологии КМОП имеет целый ряд преимуществ по сравнению с давно развивающейся и широко применяемой технологией приборов с зарядовой связью (ПЗС), причём большая

часть преимуществ связана с так называемой координатной или XY-адресацией, т. е. возможностью произвольного доступа в произвольном порядке к любому пикселу изображения.

Твердотельные фотоприёмники с координатной адресацией начали своё развитие значительно раньше изобретения Бойлом и Смитом технологии ПЗС в 1969 г. В 60-е годы большие надежды связывались с фототранзисторными матрицами, имеющими в своём составе сдвиговые регистры, осуществляющие сканирование матрицы фотодиодов по горизонтали и вертикали [131–133]. Однако недостаточное развитие кремниевой технологии в те годы не позволило получить качественное изображение. Чувствительность фотоприёмников была низкой, а само изображение было поражено так называемым геометрическим шумом, проявляющимся в виде полосатой структуры на изображении. Данная структура была в первую очередь связана с неравномерностью чувствительности и темнового тока столбцов, строк и пикселов фотоприёмника.

Фотоприёмники на основе координатной адресации также пытались построить с использованием технологии приборов с зарядовой инжекцией (ПЗИ, Charge Injection Device, CID). Лидером в этой области на протяжении ряда лет была фирма Hitachi, однако низкая чувствительность, высокие значения флуктуационного и геометрического шума не позволяли этим преобразователям конкурировать с технологией ПЗС, быстро преодолевшей свои «детские» болезни.

В середине и, в особенности, в конце 90-х годов технология ПЗС была благодаря повышению точности изготовления (ультрафиолетовая литография) доведена практически до совершенства. Квантовая эффективность лучших фотоприёмников приблизилась к 100%, темновые заряды и собственные шумы упали до единиц электронов на пиксел даже в относительно недорогих, серийно выпускающихся приборах. Однако особенность многоуровневой технологии ПЗС не позволяла интегрировать на тот же кристалл, что и фотоприёмник, остальные узлы телевизионной камеры, такие как синхрогенератор, видеопроцессор, схемы управления, АЦП.

Технологией ПЗС владело ограниченное число фирм-производителей электронной техники, в первую очередь активно работающих на рынке видеоприложений. К ним можно отнести фирмы Sony, Sharp, Toshiba, Matsushita (Япония), LG, Samsung (Корея), Philips, Thomson, Texas Instruments, Fairchild, Dalsa, Kodak (Европа и Северная Америка). С другой стороны, в это же время

бурно развивалась технология производства микросхем статической и динамической памяти для персональных компьютеров на основе технологии КМОП, обычно имеющих организацию с произвольным координатным доступом. В этом направлении работали многие технологические гиганты электронной промышленности, такие как Hewlett-Paccard, Atmel, Micron (США) [18–24]. Степень интеграции этих, относительно просто организованных интегральных схем превысила миллиард транзисторов на кристалле [6].

Следующим шагом в развитии КМОП сенсоров была идея об объединении светочувствительной секции, представляющей собой матрицу фотодиодов и набор считывающих шин и управляющих регистров, аналогичных применяемым в микросхемах памяти. Однако ёмкость шин считывания оставалась высокой, что не позволяло говорить о высоких чувствительности и отношении сигнал/шум. Прорыв в этой области произошёл в 1993 г. в связи с изобретением Эриком Р. Фоссамом технологии Активный пиксел (APS, Active Pixel Sensor) [28]. Введение усилительного каскада в каждый светочувствительный элемент позволило избавиться от проблем, связанных с ёмкостью считывающих шин и снизить геометрический шум до приемлемых значений.

Дальнейшее развитие КМОП сенсоров было связано с увеличением числа транзисторов в пикселе, интеграции на тот же кристалл аналоговых и цифровых схем управления и обработки видеосигнала, повышением чувствительности и отношения сигнал/шум, увеличения числа элементов разложения и т. п. Всё это стало возможным благодаря тому, что технология КМОП позволяет достаточно просто размещать на том же кристалле, что и фотоприёмник, такие сложные узлы, как АЦП. Развитие КМОП сенсоров за последние 15 лет характеризуется несколькими этапами, определяемыми изменением их структуры [53].

Первое поколение КМОП сенсоров носит название Сенсор с пассивным пикселом (Passive Pixel Sensor, PPS). Их особенностью является наибольшая связь с породившими их микросхемами динамической или статической памяти. Данный вид сенсора, помимо матрицы фотодиодов содержит: регистр, осуществляющий выбор строки; ключевые полевые транзисторы, осуществляющие коммутацию фотодиодов выбранной строки к вертикальным видеошинам; аналоговый мультиплексор, осуществляющий подключение видеосигнала с выбранного фотодиода к выходному усилителю (рис. 1.16).

Высокая ёмкость вертикальной видеошины, суммирующаяся с ёмкостью горизонтальной шины, не позволяет получить приемлемый уровень видеосигнала при ограниченном объёме зарядового пакета и, как следствие, данный вид сенсора обеспечивает низкое отношение сигнал/шум.

Для снижения влияния емкостей в сенсор с пассивным пикселом были введены усилители видеосигнала на каждый столбец (рис. 1.17).

Такое решение позволило устранить влияние на амплитуду видеосигнала ёмкости горизонтальной видеошины. Влияние ёмкости вертикальной шины на видеосигнал было преодолено в КМОП сенсоре с активным пикселом (Active Pixel Sensor, APS, рис. 1.18).

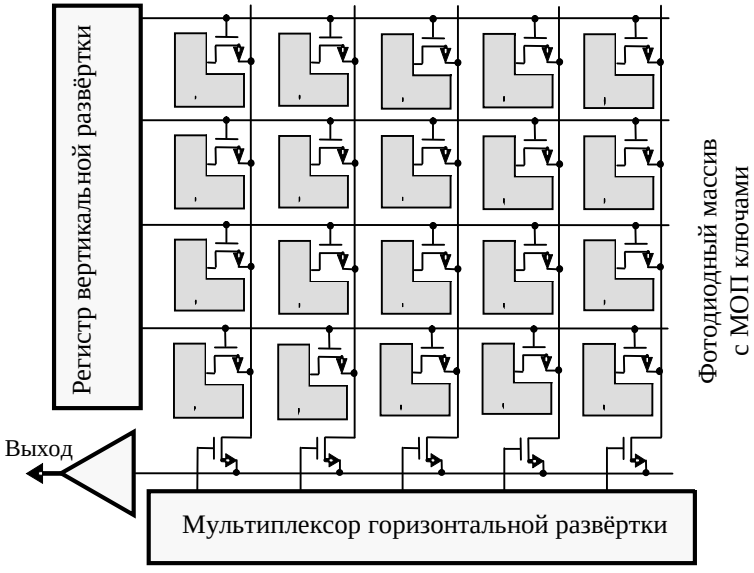


Рис. 1.16. КМОП сенсор с пассивным пикселом

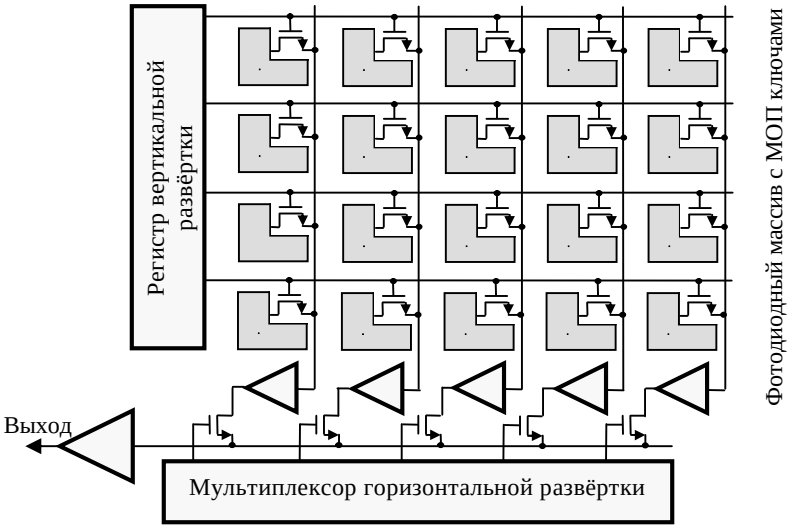


Рис. 1.17. КМОП сенсор с пассивным пикселом
и активным столбцом

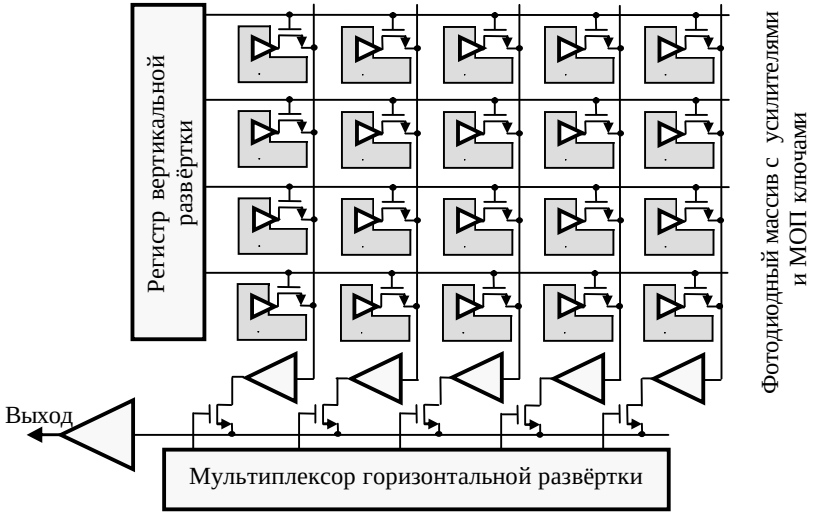


Рис. 1.18. КМОП сенсор с активным пикселом
и активным столбцом

Революционность этой технологии связана с тем, что усилитель (в простейшем случае однотранзисторный истоковый повторитель) размещается непосредственно в каждом пикселе фотоприемника. Тем самым ёмкость считывания и ёмкость вертикальной видеошины оказываются разделёнными этим усилителем. При этом ёмкость считывания может иметь тот же порядок, что и в ПЗС – единицы фемтофарад [29]. С учётом усилителя коэффициент преобразования заряда в напряжение оказывается не ниже, а даже выше, чем в фотоприёмниках на ПЗС. В современных матричных ПЗС ёмкость плавающей диффузионной области, в которой осуществляется считывание заряда, составляет величину порядка 0,01 пФ, что соответствует коэффициенту преобразования в 15 … 25 мкВ/е. В КМОП фотоприёмнике ёмкость считывания обычно того же порядка, однако применение активного усилителя (докоммутационного усиления) позволяет добиться эквивалентного ***коэффициента преобразования на порядок большего, чем у ПЗС*** – до 250 мкВ/е, что положительно сказывается на шумовых характеристиках фотоприёмника. Это в корне решает проблему чувствительности и отношения сигнал/шум. Развитие технологии Активного Пиксела в дальнейшем шло по пути улучшения качественных характеристик фотодиодов и транзисторов, а также по пути увеличения транзисторов в каждом пикселе.

Другое важное направление развития КМОП сенсоров было связано с тем, что как уже говорилось выше, КМОП технология позволяет интегрировать различные узлы телевизионной камеры на тот же кристалл и даёт разработчику возможность использования очень большого числа полевых транзисторов для реализации тех или иных узлов.

Следующим шагом на этом пути была реализация аналого-цифрового преобразователя на том же кристалле, рис. 1.19.

Одной из особенностей КМОП сенсоров является возможность реализации высокой частоты опроса, превышающей частоту считывания зарядовых пакетов из ПЗС. В связи с этим требования по быстродействию к АЦП, установленном в КМОП сенсоре, оказываются весьма жесткими.

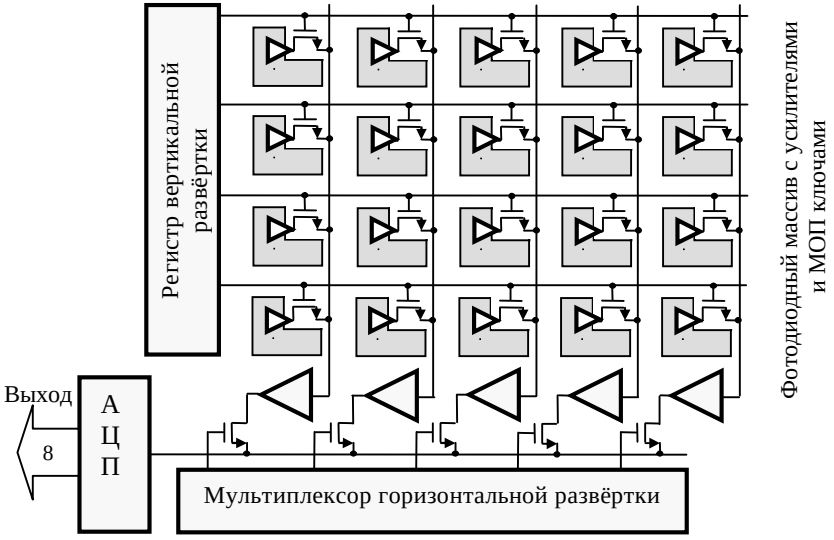


Рис. 1.19. КМОП сенсор с активным пикселом, активным столбцом и одним АЦП

Рис. 1.20. КМОП сенсор с активным пикселом и АЦП на каждый столбец

Высокая частота считывания приводит к широкой полосе частот видеосигнала и худшим шумовым характеристикам, а большая частота преобразования АЦП – к большому выделению мощности.

Для преодоления этого недостатка в КМОП сенсорах АЦП стали располагать из расчета один АЦП на каждый столбец и коммутацию осуществлять в цифровом виде (рис. 1.20). Это позволило снизить частоту преобразования каждого из АЦП в число раз, равное числу столбцов в фоточувствительной секции, а также в соответствующее число раз уменьшить полосу частот видеосигнала.

Дальнейшее развитие пошло по пути движения к КМОП сенсору с цифровым пикселом (Digital Pixel Sensor, DPS), рис. 1.21 [40].

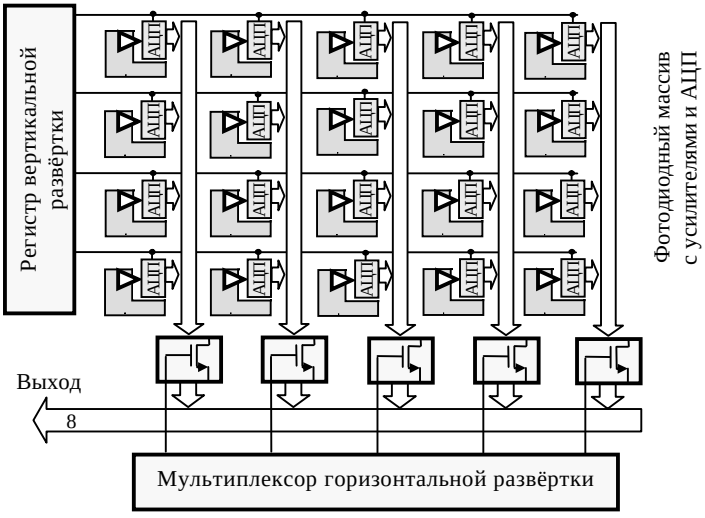


Рис. 1.21. КМОП сенсор с активным цифровым пикселом

Особенностью этих преобразователей является интеграция в каждый пиксел не только активного усилителя, но и непосредственно АЦП. При этом частота преобразования АЦП становится равной кадровой частоте фотоприёмника, а полоса частот видеосигнала уменьшается до минимальных значений. В настоящее время известны КМОП сенсоры с цифровым пикселом размером 9,8×9,8 мкм (при точности изготовления

0,18 мкм) содержащие в каждом пикселе АЦП, состоящий из 37 транзисторов [26, 40].

Как и в матричных ПЗС, светочувствительным пикселом в КМОП сенсоре может являться обеднённая область МОП-ёмкости, возникающая при подаче на фазный электрод обедняющего напряжения рис. 1.1 *а*, либо же обеднённая область обратносмещённого фотодиода, рис. 1.1 *б*. Светочувствительный элемент второго типа предпочтительнее, хотя бы с той точки зрения, что фотодиодная структура имеет существенно более высокий коэффициент сбора светового потока из-за отсутствия слоёв поликремния, поглощающего световой поток. Простейшей является структура пассивного пиксела, рис. 1.22 [18].

Помимо фотодиода она содержит ключевой транзистор VT, который по сигналу RS (Row Select, Выбор столбца), подключается к вертикальной видеошине. Таким образом, к этому пикселу подходит две шины – видеошина и шина, по которой подается сигнал считывания RS. Особенностью такого типа пиксела является высокий коэффициент использования площади, т. к. площадь транзистора много меньше площади фотодиода, и высокий уровень шумов, связанный с тем, что фотодиод подключён непосредственно к вертикальной видеошине, имеющей большую ёмкость.

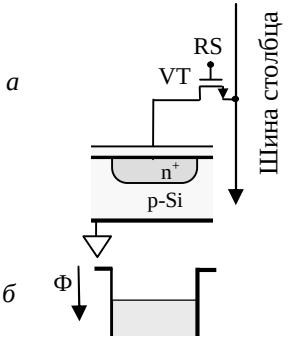


Рис. 1.22. Структура пассивного пиксела *а* и соответствующая диаграмма потенциала *б*

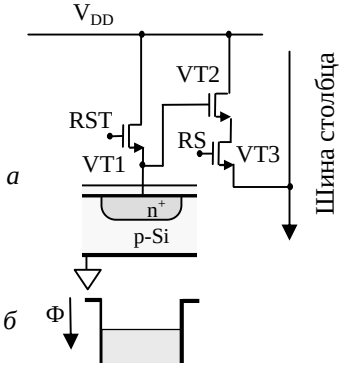


Рис. 1.23. Структура активного пиксела *а* и соответствующая диаграмма потенциала *б*

Структура активного пиксела содержит уже, как минимум, 3 полевых транзистора (рис. 1.23) и четыре шины – вертикальную видеошину, шины по которым подаются сигналы RS и RST, а также шину питания V_{DD}.

Транзистор VT3, управляемый сигналом RS, выполняет ту же функцию, что и в пассивном пикселе. Транзистор VT1, управляемый сигналом RST (Reset, сброс) служит для сброса накопленного заряда, а гальванометрический транзистор VT2 выполняет роль истокового повторителя, осуществляющего усиление видеосигнала по току. Нагрузкой данного транзистора является источник тока, расположенный на конце столбца за пределами фоточувствительной секции. Коэффициент использования площади у такого Пиксела значительно ниже за счет большего числа транзисторов, однако шумовые характеристики значительно лучше, чем в пассивном пикселе.

Дальнейшим развитием активного пиксела было введение в него плавающей диффузионной области, рис. 1.24. В данном виде пиксела накопленный заряд переносится из фотодиода в плаваю-

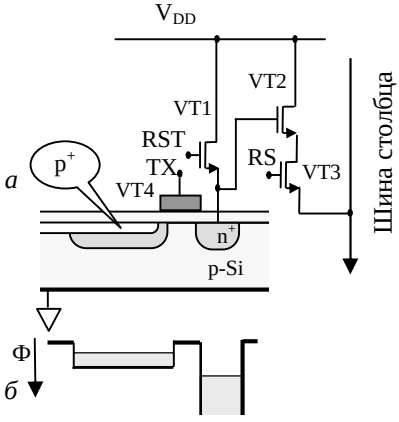


Рис. 1.24. Структура активного пиксела с дополнительным затвором *а* и соответствующая диаграмма потенциала *б*

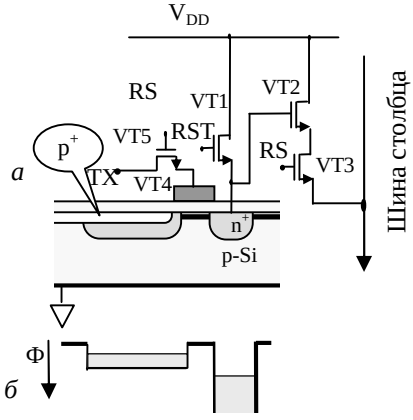


Рис. 1.25. Структура активного пиксела с дополнительным затвором и усилителем в цепи затвора *а* и соответствующая диаграмма потенциала *б*

щую диффузионную область подачей положительного смещения на затвор переноса TX [18, 53].

Возможность реализации плавающей диффузионной области с очень малой ёмкостью позволяет увеличить коэффициент преобразования заряда в напряжение и разделить процессы накопления и считывания по пространству. С учётом затвора переноса такой пиксел содержит четыре транзистора и пять шин. По сравнению с предыдущим пикселом в данном случае возникает необходимость в шине управления затвором переноса. Такой пиксел имеет малый коэффициент использования полезной площади, однако обеспечивает малый уровень флуктуационных и геометрических шумов.

С целью уменьшения числа шин управления и увеличения коэффициента использования полезной площади сигналы RS и TX можно объединить, рис. 1.25. Для структур пиксела с плавающей диффузионной областью формируемый видеосигнал имеет трёхуровневую форму, характерную для фотоприёмников на основе ПЗС. Наличие попиксельного сброса приводит к появлению шума установки узла детектирования заряда, так называемого *KTC*-шума. Как следствие, для обработки такого видеосигнала в структуру КМОП сенсора необходимо интегрировать схему двойной коррелированной выборки (ДКВ), т. е. узел, характерный практически для всех линейных и матричных ПЗС. Особый интерес представляют пикселы с логарифмической характеристикой накопления, рис. 1.26. При указанном на рис. 1.26 включении полевого транзистора соотношение между током и напряжением на затворе носит логарифмический характер. Указанное свойство позволяет получить фотоприёмник с расширенным динамическим диапазоном по свету и со свет–сигнальной характеристикой, аналогичной таковой у человека, которая согласно закону Вебера–

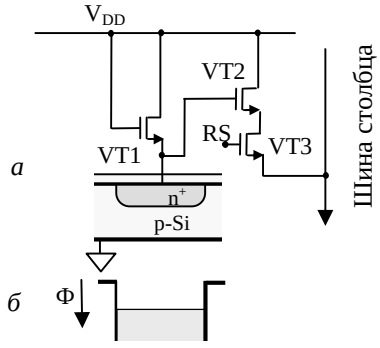


Рис. 1.26. Структура активного пиксела с логарифмической шкалой *а*, и соответствующая диаграмма потенциала *б*

Фехнера имеет логарифмический характер и учитывается в телевидении с самого его зарождения.

Важность принципа накопления в телевидении неоспорима, именно с его помощью осуществляется групповой счёт фотонов (см. раздел 3.4.). КМОП телекамеры в отличие от фотоприёмников предыдущего поколения на ПЗС, могут обеспечить сверхширокий линейный динамический диапазон.

Технической основой увеличения динамического диапазона является усложнение структуры пиксела путём введения элементов аналоговой или цифровой памяти. На рис. 1.27*а* приведена принципиальная схема пиксела, обеспечивающего линейный динамический диапазон по входной освещенности 132 дБ – *более миллиона раз* [35], а на рис. 1.27*б* его топология. Пиксел состоит из фотодиода VD1, первичной накопительной емкости C1, компаратора DA1, дополнительных емкостей хранения C1 и C2 и логики считывания.

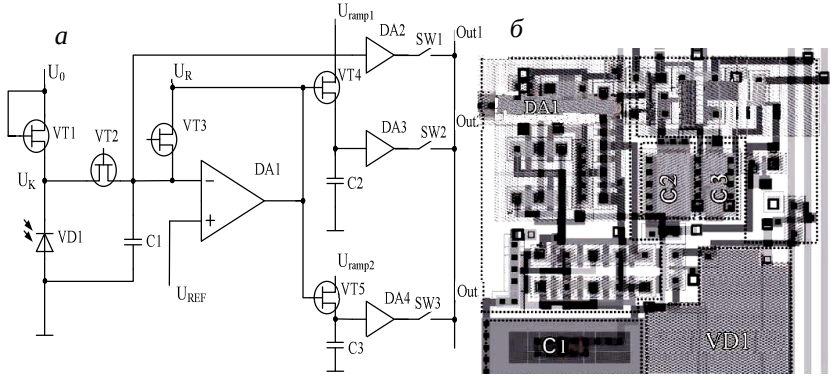


Рис. 1.27. Принципиальная схема *а* и топология *б* пиксела с расширенным динамическим диапазоном

В настоящее время производство КМОП сенсоров освоило большое число производителей. В таблице приведены сведения о целом ряде сенсоров с большим числом пикселов, ориентированных в первую очередь на применение в цифровой фотографии.

Таблица 1.3

Наименование сенсора	Кол-во пикселов	Тип, размер	Эффективные пикселы (H×V)	Частота кадров, Гц	Размер пиксела, мкм
Sony					
IMX006	1.33M	1/3.2	1296×985	15	3.45×3.45
IMX011	2.13M	1/3	1656×1249	15	2.925×2.925
IMX007AQ	12.4M	1.8	4288×2848	5	5.49×5.49
Sharp					
PJ24P3BA0ET	3.27M	1/3.2	2096×1560	3	2.2×2.2
PJ23R3BA0ET	4.06M	1/2.5	2332×1740		2.5×2.5
RJ23T3BA0ET	6.20M	1/2.5	2872×2160		2.05×2.05
RJ23U3BA0ET	7.40M	1/2.5	3096×2328		1.9×1.9
PJ21V3BA0ET	8.50M	1/1.8	3320×2496	5	2.2×2.2
PJ21W3BA0ET	10.6M	1/1.7	3696×2776		2.05×2.05
Panasonic					
MN39960	7.38M	4/3	3136×2352	3	5.6×5.6
Samsung					
S5K4AAF	1.31M	1/4	1280×1024		2.8×2.8
S5K3C1FX	3.15M	1/3	2048×1536		2.25×2.25
S5K2E1FX	5.16M	1/2.5	2616×1960	10	2.2×2.2
Micron					
MT9M112	1.31M	1/4	1280×1024	15	2.8×2.8
MT9M413	1.31M	20mm	1280×1024	500	12×12
MT9T012	3.17M	1/3,2	2056×1544	15	2.2×2.2
MT9E001	7.99M	1/2.5	3264×2448		1.75×1.75
Philips					
FXA1012	2.09M	2/3	1616×1296	5	5.1×5.1
Canon					
EOS 10D(704F)	6.29M	22.7x15.1	3072×2048	3	7.4×7.4
EOS 350D	8.00M	22.2x14.8	3456×2304	3	6.4×6.4
EOS 5D(706P)	12.8M	35mm	4368×2912	3	8.2×8.2
EOS 1DsII	16.6M	35mm	4992×3328	4	7.2×7.2
Kodak					
KAS-3100	3.15M	1/2.7	2048×1536	12	2.7×2.7
KAS-5000	5.04M	1/1.8	2592×1944	6	2.7×2.7

Nikon					
D2H	4.02M	23.7x15.5	2464×1632	8	9.6×9.6
Atmel					
AT71200M	8.05M	35mm	3500×2300	2.8	10×10
Foveon					
FO18-50-F19	1.50M	1/1.8	1420×1060	7	5×5
F7X3-C9110	3.40M	20.7x13.8	2268×1512	4.4	9.12×9.12
Fx17-78-F13	4.82M	20.7x13.8	2652×1768	5	7.8×7.8
Cypress					
CYIHDSC9000AA	9.03M	1.8	3710×2434	5	6.4×6.4

Типичная структура современного КМОП сенсора (рис. 1.28), соответствующая организации прибора, показанной на рис. 1.19, включает схемы двойной коррелированной выборки (ДКВ).

Вариант конкретной реализации КМОП сенсора (рис. 1.29) соответствует организации прибора, показанной на рис. 1.20 и отличается наличием не только многих каналов считывания с фотоприёмного массива, но и множеством каналов связи сенсора с внешним компьютером.

Фотодиодная матрица с усилителями и КМОП ключами управляется регистром вертикальной развёртки. Видеосигнал со всех пикселов строки поступает на столбцовые усилители и при необходимости обрабатывается схемой ДКВ. Далее сигнал при помощи аналогового мультиплексора горизонтальной развёртки подается на видеопроцессор и встроенный АЦП. Оцифрованный видеосигнал обрабатывается цифровым процессором, в котором могут быть осуществлены такие преобразования видеосигнала, как гамма-коррекция и апертурная коррекция, баланс чёрного и белого, формирование цветоделённых или яркостного и цветоразностных сигналов и т. п. В общем случае в КМОП сенсоре (как и во многих других системах на кристалле) могут присутствовать выходные ЦАП, формирующие аналоговый выходной сигнал или же кодер системы NTSC или PAL. Большинство сенсоров формируют выходной видеосигнал в цифровой форме. Внешнее управление сенсором осуществляется обычно при помощи какого-либо стандартного интерфейса – I^2C, SPI и т. д. При помощи данного интерфейса можно задать такие параметры, как размер окна считывания, коэффициент усиления видеопроцессора, время

накопления, значение гамма-характеристики и т. п. Управление всеми узлами сенсора осуществляется от блока управления.

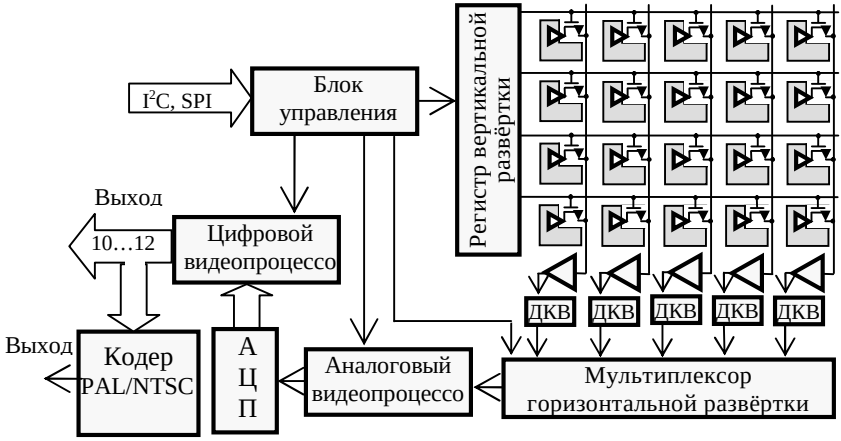


Рис. 1.28. Типовая структура КМОП сенсора с активным пикселом, активным столбцом, аналоговым и цифровым процессором

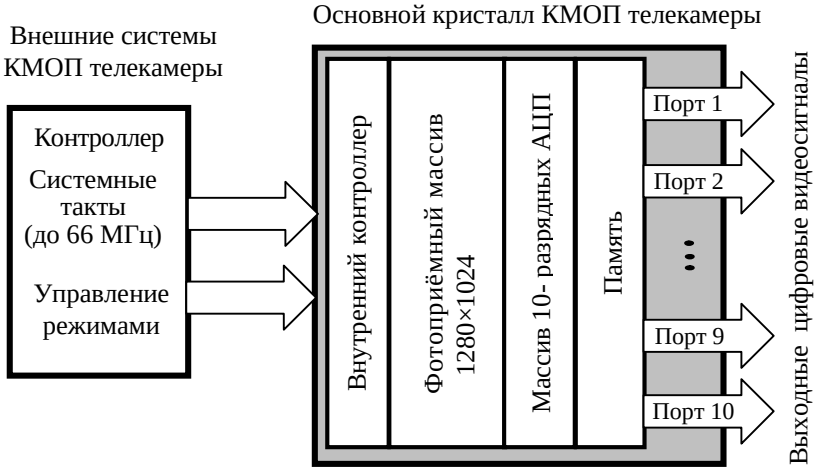


Рис. 1. 29. Структура КМОП сенсора MT9M413 фирмы Micron с активным пикселом и АЦП на каждый столбец, размещаемой в 280-выводном керамическом корпусе

В состав КМОП сенсора могут входить оперативные и постоянные запоминающие устройства, необходимые, например,

для коррекции дефектных пикселов. Таким образом, развитие технологии КМОП сенсоров реально привело к созданию однокристальной цифровой телевизионной камеры, реализующей функции цифровой и аналоговой обработки. Эта технология в англоязычной литературе получила название camera-on-chip [25, 26]. Расширение функциональных возможностей КМОП телекамер, ставшее первой причиной их конкурентоспособности по отношению к ПЗС, обострило внимание к их характеристикам, среди которых главными являются *разрешающая способность*, *чувствительность* (ограниченная не только шумами, но и *ложными сигналами*), *динамический диапазон* и *эксплуатационные характеристики* (масса, габариты, энергопотребление и т. п.) – все те характеристики телекамер, о значении учёта которых писал ещё создатель электронного телевидения В. К. Зворыкин.

Разрешающая способность фотоприёмника в первую очередь определяется числом светочувствительных элементов и их геометрическим размером. В отличие от технологии ПЗС, КМОП технология позволяет достаточно простыми средствами организовать фотоприёмник, содержащий очень большое число пикселов – десятки миллионов, и имеющий очень большую площадь кристалла, например, совпадающую с размером кадра стандартной фотоплёнки. Это объясняется тем, что при увеличении размера кристалла в ПЗС вероятность захватить фатальный дефект быстро возрастает. В КМОП сенсоре такой дефект вызовет поражение одного единственного пиксела, значение видеосигнала которого может быть интерполировано по соседним элементам, в том числе непосредственно в самом сенсоре. В матричных ПЗС такой дефект очень часто приводит к неработоспособности всего преобразователя в целом. Тем самым, выход годных приборов, выполненных по КМОП технологии, особенно при большой площади кристалла, существенно выше, нежели для ПЗС технологии.

Чувствительность фотоприёмника определяется коэффициентом сбора светового потока, квантовой эффективностью и собственными шумами. Наличие большого числа транзисторов в каждом пикселе приводит к снижению коэффициента использования светового потока в КМОП сенсорах, однако применение хорошо отработанной в матричных ПЗС со строчным переносом технологии микролинз, расположенных над фотодиодом (on-chip-microlenses) позволяет добиться очень высоких значений коэффициента сбора, приближающихся в перспективе к 100%.

Использование оптимизированных структур современных фотодиодов в качестве светочувствительного элемента, максимизирует квантовую эффективность, а наличие встроенного в пиксел активного усилителя теоретически позволяет достичь меньшего уровня шумов, чем в ПЗС структурах. Также благотворным фактором, влияющим на снижение собственного шума является уменьшение полосы частот обработки видеосигнал при применении параллельной обработки видеосигналов всех пикселов строки или же всего фотоприёмника. Уже в существующих КМОП сенсорах с активным пикселом коэффициент преобразования заряда в напряжение может составлять сотни мкВ/электрон, тогда как в лучших фотоприемниках на ПЗС – десятки мкВ/электрон [53].

Однако недостаточная отработанность технологии, проявляющаяся в том числе в трудности создания мегапиксельного массива малошумящих транзисторов, приводит к тому, что шумы КМОП фотоприёмников пока превышают шумы ПЗС преобразователей и их чувствительность при равенстве таких факторов, как время накопления и размер светочувствительного элемента, пока уступает чувствительности матричных ПЗС.

Спектральная чувствительность в целом идентична характеристикам матричных ПЗС и, в первую очередь, определяется типом и особенностями примененного кремниевого светочувствительного элемента – МОП ёмкости или фотодиода. В то же время известны КМОП сенсоры, в которых схемы считывания выполнены по кремниевой технологии, а в качестве фоточувствительных элементов использованы специальные слои. Такая структура позволяет создавать ИК КМОП фотоприёмники.

Ложные сигналы, возникающие в процессе фотоэлектрического преобразования и считывания видеосигнала, в КМОП сенсорах связаны с темновым током и разбросом чувствительности. С точки зрения темнового тока фотодиод КМОП сенсора может быть оптимизирован так же, как и в матричных ПЗС со строчным переносом. Однако наличие схем усиления и обработки, в том числе расположенных непосредственно в пикселе, приводит к тому, что тепло выделяется непосредственно в том месте, где накапливаются фотоэлектроны, что вызывает рост темнового тока. Одной из основных проблем современных КМОП сенсоров является наличие так называемого геометрического шума (Fixed Pattern Noise) – детерминированной компоненты, вызванной

неравномерностью видеосигнал соседних столбцов и строк. Геометрический шум проявляется обычно в виде вертикальной структуры, особенно проявляющейся на чёрном при малой освещённости. В настоящее время фактором, ограничивающим пороговую чувствительность КМОП фотоприёмников в первую очередь является именно геометрический, а не флуктуационный шум. Данное ограничение носит технологический характер и будет преодолено по мере совершенствования технологии и увеличении вычислительной мощности цифрового процессора, расположенного на кристалле. В то же время в КМОП сенсорах полностью отсутствуют артефакты, вызванные неэффективностью переноса и так называемый вертикальный смаз (вертикальный столб на изображении от ярких деталей), присущие матричным ПЗС. Это объясняется тем, что заряд считывается в том же месте, где и формируется и механизмы образования смаза и неэффективности отсутствуют. Проблемы блуминга, возникающего при пересветках, решаются теми же способами, которые отработаны в матричных ПЗС. Наблюдению объектов при малых освещенностях препятствует свечение полевых транзисторов. В матричных ПЗС свечение обычно связано с транзисторами выходного узла и проявляется в виде фона, плавно уменьшающегося от левого верхнего угла изображения [53, 54]. В КМОП сенсорах транзисторы расположены непосредственно в пикселе и их свечение может привести к образованию паразитного заряда, вносящего к тому же свои дробовые шумы.

Динамический диапазон КМОП сенсоров определяется максимальным зарядовым пакетом, накопленным в светочувствительном элементы и шумами считывания или же геометрическим шумом. С этой точки зрения динамический диапазон КМОП сенсора соизмерим с динамическим диапазоном матричного ПЗС с аналогичным размером пиксела. Однако, возможность варьирования характеристик светочувствительного элемента и схем считывания позволяет построить КМОП сенсор с логарифмической зависимостью выходного сигнала от освещенности и довести рабочий динамический диапазон до 132 дБ [35]. Известным методом расширения динамического диапазона в матричных ПЗС является суммирование зарядовых пакетов соседних пикселов по горизонтали и вертикали непосредственно на преобразователе до момента считывания [1]. В КМОП сенсорах с координатной адресацией такое суммирование (binning) затруднительно, однако

известны сенсоры, в которых может быть организовано суммирование произвольного числа зарядовых пакетов соседних элементов [31].

В современных матричных ПЗС зарядовый пакет при своём движении к выходному устройству может проходить макрорасстояния, составляющие десятки миллиметров. При этом скорость переноса обычно ограничена частотой тактирования фазных электродов выходного регистра и редко превышает 40 МГц. Дальнейшее увеличение частоты переноса приводит к повышению неэффективности, увеличению потребляемой мощности, снижению управляющей способности и т. п. Такое ограничение частоты считывания приводит к соответствующим ограничениям достижимой кадровой частоты и требует для её увеличения методов параллельного считывания и последующей сшивки изображения. В КМОП сенсорах частоты опроса пикселов в 50 МГц являются типичными, и в лучших образцах данная частота превышает 100 МГц. С учётом того, что распараллеливание выходов в КМОП сенсорах реализуется даже проще, чем в матричных ПЗС, достижение высоких кадровых частот и, соответственно, лучших динамических характеристик для КМОП сенсоров является более простой задачей. В настоящее время на рынке быстродействующих камер представлен целый ряд фирм, реализующих кадровые частоты выше 1 кГц [55, 56].

Эксплуатационные характеристики КМОП сенсоров существенно лучше, чем у матричных ПЗС. Выход годных при использовании технологии КМОП существенно выше, чем при использовании ПЗС технологии. Этим обусловлена малая стоимость КМОП сенсоров по сравнению с матричными ПЗС, особенно при большой площади кристалла. В отличие от матричных ПЗС, которым требуется несколько, причем достаточно высоких, питающих напряжений, для КМОП требуется одно низковольтное напряжение питания. Обычно это +5 или +3,3 вольта. Потребляемая мощность также оказывается во много раз меньше, чем для камеры на матричном ПЗС. Габаритные размеры и масса для КМОП сенсоров также во много раз меньше, чем для ПЗС камеры, что объясняется интеграцией всех основных узлов камеры на одном кристалле. Для камеры на матричном ПЗС требуется по крайней мере две–три дополнительных микросхемы плюс большое число дополнительных элементов. Интеграция всех функций на одном кристалле благоприятно сказывается и на надёжности

камеры. Высокие значения эксплуатационных параметров и характеристик позволяют активно использовать КМОП сенсоры в мобильных телефонах, где габаритные размеры, масса и потребляемая мощность весьма критичны.

КМОП сенсоры успешно завоевывают такую перспективную область телевизионных систем наблюдения, как сетевые телекамеры. Цифровая по своей сути структура КМОП сенсора хорошо согласуется с IP-технологиями, и в настоящее время целый ряд фирм выпускает камеры на основе КМОП технологии для построения видеосетей [48, 57, 58, 84].

Новые функциональные возможности КМОП сенсоров связаны с возможностью *адаптации* КМОП телекамер к изменениям не только уровня освещённости, но других свойств сюжета – динамике наблюдаемых объектов и их числу. Так, одной из самых востребованных функций является задание размера и положение фрагмента изображения, с которого производится считывание видеосигнала (Region of interest) [31, 32, 106]. При этом реализуется обмен поля зрения и числа элементов фотоприёмника на частоту кадров, т. е. при уменьшении размеров фрагмента в соответствующее число раз возрастает кадровая частота. Многие современные КМОП сенсоры позволяют задать размеры и положение фрагмента с точностью до четырёх пикселов. Гипотетически нет никаких ограничений на задание нескольких фрагментов с соответствующей кадровой частотой для каждого. Координатное считывание позволяет комбинировать режимы чересстрочного/прогрессивного считывания. Так, в видеокамере, при видеозаписи КМОП сенсор может работать в чересстрочном режиме со стандартным разложением, а при считывании разложение может изменяться на прогрессивное.

Новая технология формирования цветного изображения Foveon, основана на различной глубине проникновения видимого излучения в глубину кремния. Так, фотоны синего излучения поглощаются в поверхностном слое, зеленого излучения – глубже, а красного – в глубине кремния [134]. В КМОП сенсоре существует возможность расположить три фотодиода друг над другом и, тем самым, осуществить цветоделение. Каждый из фотодиодов поглощает свою часть спектрального диапазона. При этом достигается целый ряд преимуществ по сравнению с известными методами цветоделения, такими как использование цветоделительных призм или кодирующих светофильтров. Технология

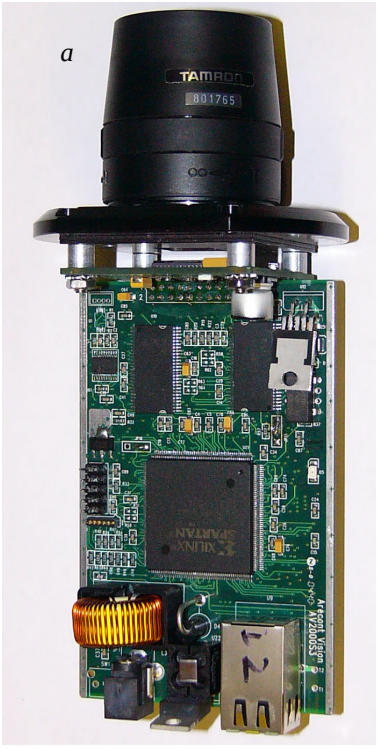
Foveon обеспечивает 100%-е использование светового потока, т. к. каждый из фотонов видимого диапазона поглощается в одном из фотодиодов, а также полную разрешающую способность в красном, зелёном и синем участках спектра [135].

В КМОП сенсорах, также как и в матричных ПЗС может быть реализован режим электронного затвора для улучшения качества передачи динамических изображений. Однако в большинстве КМОП сенсоров реализуется так называемый бегущий электронный затвор (Roller shutter). Его особенностью является то, что считывание в КМОП сенсоре осуществляется построчно. При этом время накопление в различных строках хоть и одинаково по длительности, но смещено друг относительно друга. Это приводит к тому, что движущиеся объекты хоть и не испытывают динамического смаза, однако искажаются геометрии-чески. Так, если круглый объект движется по изображению сверху вниз, по направлению считывания строк, то его изображение превратиться в эллипс, вытянутый по вертикали. Вертикальные линии, двигающиеся в горизонтальном направлении, оказываются наклонными и т. п. Наличие таких искажений особенно заметно при уменьшенном времени накопления и является еще одно проблемой применения КМОП сенсоров, ожидающей своего решения [53].

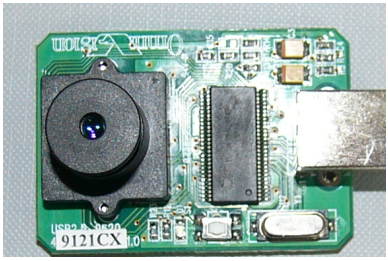
Одной из новых особенностей видеосистем на кристалле является возможность реализации разнообразных алгоритмов считывания. Именно такое разнообразие возводит этап развёртки изображений в видеосистемах на кристалле в ранг микропрограммирования эксперимента. Реализация всевозможных видов развёртки открывает путь существенной экономии ёмкости памяти оперативных запоминающих устройств, необходимых при фильтрации изображений с помощью так называемых «многоэлементных апертур». Кроме того, сложная форма развёртки может существенно упростить операцию увеличения площади зоны накопления, необходимую для обмена разрешающей способности системы на чувствительность. В частности, для дополнительной пространственной обработки, например, усреднения нескольких элементов изображения или медианной фильтрации, при традиционной прогрессивной развёртке требуются дополнительные элементы аналоговой или цифровой памяти. Так называемая развёртка Z-сканирования позволяет проводить пространственное усреднение групп элементов 2×2, 4×4, 8×8 без применения дополнительных запоминающих устройств [31].

В настоящее время рынок телевизионных камер на КМОП сенсорах стремительного догоняет и обгоняет рынок камер на матричных ПЗС – мобильные телефоны, цифровые фотоаппараты, видеокамеры, системы видеонаблюдения, быстродействующие телевизионные камеры – вот неполный перечень областей, где КМОП сенсоры находят всё большее применение. На рис. 1.30 *а* приведена фотография телевизионной камеры фирмы ArecontVision (США), выполненной на 2-мегапиксельном сенсоре фирмы Micron и предназначенной для работы в IP сетях. На рис. 30 *б* представлен внешний вид отладочного модуля OV-9121-ECX производства фирмы «Omni Vision»на базе 1,3-мегапиксельного КМОП сенсора с интерфейсом USB2.0. На рис. 1 30 *в* приведён внешний вид телевизионной камеры, разработанной на кафедре ТВ СПбГЭТУ «ЛЭТИ» на основе 3-мегапиксельного КМОП сенсора фирмы Micron и имеющей интерфейс USB2.0 и программное обеспечение для ввода изображений в компьютер.

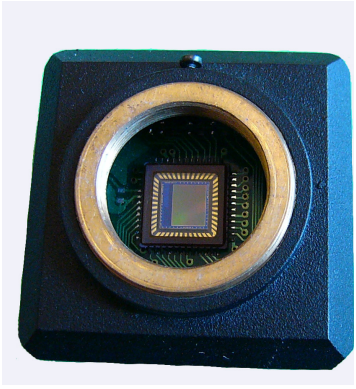
Компания Sharp сообщила о разработке самого маленького в мире модуля CMOS, предназначенного для использования в сотовых телефонах и других портативных цифровых устройствах. Модуль под названием LZ0P396D имеет размер 1/11 дюйма. Разрешение в этом устройстве составило 110 тыс. пикселей. Толщина модуля – всего 2,43 мм. Как говорят в компании, помимо съемки фотографий данная модель может записывать видео с частотой 30 кадров в секунду. Важную тенденцию в развитии КМОП телекамер отражает прибор ADCC-3960 – 1,3-мегапиксельный КМОП датчик изображения фирмы Avago Technologies (Agilent Technologies). В данном датчике реализована технология Agilent, которая позволяет получать чёткие изображения с хорошей цветопередачей практически при всех условиях освещения. Данный датчик изображения может использоваться в вычислительной, медицинской технике и индустриальных применениях. Новая усовершенствованная архитектура (рис. 1.31) и получение пикселя основаны на технологии, которая обеспечивает десятикратное уменьшение шумов, что позволяет получать более качественное цветное изображения, понизить темновой ток, повысить чувствительность в синей области.



а



б



в

г

Рис. 1.30. Телекамеры на КМОП фотоприёмниках с внешними устройствами различной сложности от включающей сложные цветовые коррекции (*а*) до простой (*г*, справа), содержащей во внешнем устройстве лишь кварцевый генератор (*г*, слева от КМОП телекамеры)

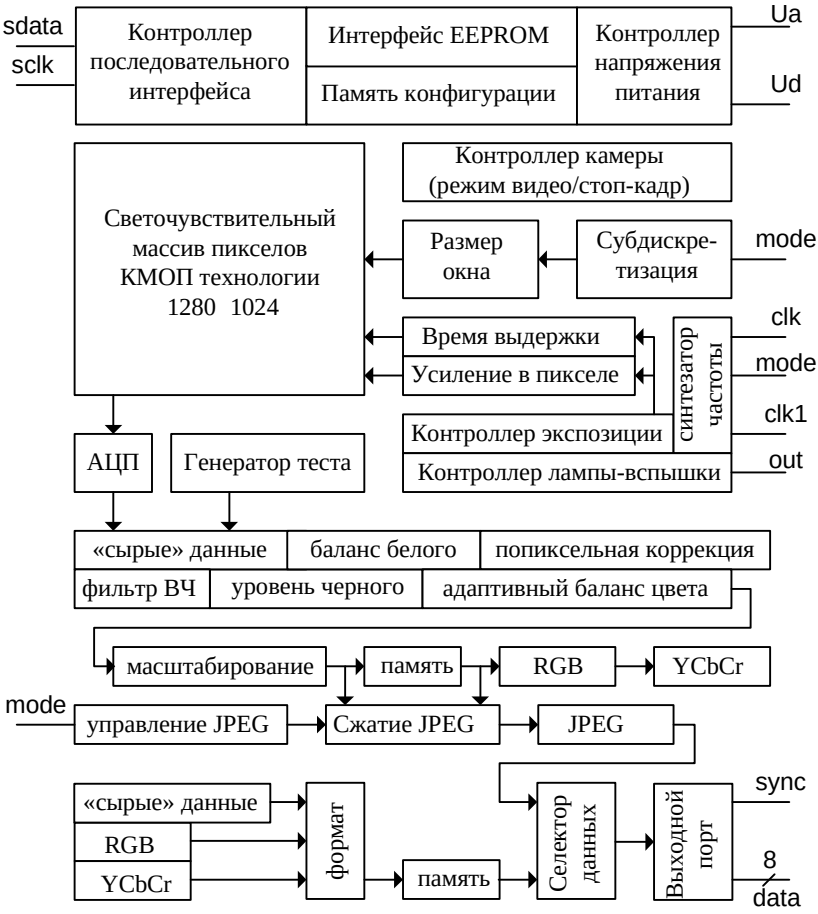


Рис. 1.31. Структура видеосистемы на кристалле ADCC-3960 фирмы Avago Technologies, построенной на основе 1,3-мегапиксельного КМОП датчика изображения

ADCC-3960 – полностью интегрированная система на кристалле, которая включает КМОП датчик с процессором для обработки изображения и сжатия по протоколу JPEG, 10-разрядный АЦП. Данный КМОП датчик изображения использует 1:3 оптический формат, способен обрабатывать 15 кадров в секунду при полном SXGA разрешении и 30 кадров в секунду при VGA

разрешении, имеет размер элемента 3,3×3,3 мкм^2 и содержит 1280 горизонтальных и 1024 вертикальных активных элементов, имеет программируемую синхронизацию с фотовспышкой на базе светодиодов и ксеноновых ламп, автоматическую коррекцию плохих элементов и четырехкратное цифровое увеличение.

Перспективы телекамер на КМОП фотоприёмниках связаны с тем, что именно этот тип приборов будет наиболее ярко олицетворять объединение техники связи и рукотворных кристаллов. Перспективы телекамер на КМОП фотоприёмниках объясняются их двумя принципиальными особенностями реализации кодирования источника и кодирования канала. Во-первых, это реализация непосредственно в кристалле докоммутационного усиления изображений, ранее осуществлявшегося только с помощью электронно-оптического преобразования в вакуумных приборах с использованием высоких напряжений. Но свойства, влияющие на надёжность прибора – необходимость обеспечения вакуума и поддержки работы прибора киловольтами – не главные недостатки электронно-оптических преобразователей (ЭОП, в англоязычной литературе называемых усилителями изображений, Image Intensifier, I I).

Гораздо важнее их информационные характеристики, принципиально отличающиеся от характеристик твердотельных фотоприёмников: потенциальная чувствительность и разрешающая способность. Потенциальная чувствительность телекамер на КМОП фотоприёмниках выше, чем у ПЗС или видиконов, не только благодаря докоммутационному усилению, но и потому, что в КМОП удаётся реализовать многоканальное считывание с увеличением числа каналов от одного до числа столбцов и даже до числа элементов в матрице. Это позволяет резко сократить полосу частот в каждом парциальном канале считывания и уменьшить влияние собственных шумов прибора.

Низкая квантовая эффективность электронно-оптических преобразователей по сравнению с твердотельными фотоприёмниками связана с различием используемого вида фотоэлектрического эффекта: в твердотельных приборах используется внутренний фотоэффект, в ЭОП – внешний фотоэффект. Разрешающая способность электронно-оптических преобразователей в лучшем случае достигает 30 линий на миллиметр, что в несколько раз уступает возможностям твердотельных матриц.

Во-вторых, КМОП телекамеры быстро эволюционируют в сторону видеосистем на кристалле, в которых осуществляется различная обработка видеоинформации – и кодирование источника, и кодирование канала – непосредственно на том же кристалле, на котором расположен фотоприёмный массив элементов. Новая топология прибора наглядно иллюстрирует новые проблемы синтеза видеосистем на кристалле, проявляющиеся в задаче распределения площади кристалла между фотоприёмником и устройством обработки сигналов и в задаче проектирования множества параллельных быстродействующих каналов связи фотоприёмного массива с вычислителями (последние для увеличения производительности давно строятся с использованием метода распараллеливания вычислений). Уже существуют видеосистемы на кристалле, осуществляющие преобразование потока фотонов в поток координат объектов в задаче измерения динамических искажений волнового фронта (предметно ориентированный кодер источника). На пороге появление видеосистем на кристалле, объединяющих фотоприёмную матрицу большого формата (телевидение высокой чёткости) и кодер источника (т. е. устройство сжатия изображения) в одном из перспективных стандартов. Сейчас наиболее популярны MPEG-2 и MPEG-4, но будут и более эффективные. Так же существуют видеосистемы на кристалле, сочетающие в себе первый этап кодирования источника (накопление и усиление сигналов) с кодированием канала, ориентированным на компьютерные сети – USB, Ethernet (пока эти элементы в КМОП телекамерах выполняются на отдельных БИС, рис. 1. 26). Поэтому можно прогнозировать эволюцию КМОП телекамер к новому типу СБИС – видеосистемам на кристалле высокой системной интеграции, включающей в себя и фотоприёмный массив, и все виды кодирования – кодирование источника и кодирование канала, обеспечивая равновесное согласование источника с каналом, когда выравниваются скорость создания информации источником и скорость её передачи через канал связи, и реализуется принцип минимума информации при назначенном качестве информации.

Важной задачей при проектировании твердотельных телекамер является борьба с темновыми сигналами, обусловленными тепловой генерацией зарядов. Темновой сигнал фотоприёмника снижает качество формируемого изображения по трём причинам:

- уменьшается диапазон воспроизводимых сигналов;

- возрастает собственный шум телекамеры за счёт шума темнового сигнала, среднеквадратичное значение которого равно корню квадратному из накопленного темнового заряда;
- неравномерность темнового заряда по полю является коррелированной от кадра к кадру помехой и делает необходимой периодическую калибровку аппаратуры с целью запоминания и вычитания из информационного кадра оценки темного кадра.

Заряд Q_{Ti}, накопленный в конкретном i-м элементе накопления из-за влияния тепла, а не света, зависит от плотности темнового тока I_{oi} в этом элементе, времени накопления $T_н$ и температуры T [44, 45]:

$$Q_{mi} = I_{oi} T_н T^{3/2} exp[-E(T)/2kT] \ ,$$

где $E(T)$ – ширина запрещённой зоны кремния (зависящая от температуры); k – постоянная Больцмана.

В практике применения кремниевых фотоприёмников на ПЗС эта формула трансформировалась в эмпирическое правило: *темновой ток при охлаждении убывает вдвое на каждые 7*^{o}.

Так как темновой заряд сильно зависит от температуры, естественно стремление к её снижению. Этот аргумент часто выдвигают в качестве ограничения к применению КМОП фотоприёмников, так как размещение дополнительных элементов на том же кристалле, что и фотоприёмник, повышает тепловыделение на кристалле, и, следовательно, его температуру. Вместе с тем практика применения твердотельных телекамер показывает, что в обоих типах телекамер – и с ПЗС матрицами, и с КМОП матрицами – основное тепловыделение обусловлено внешними цепями – стабилизаторами напряжений, генераторами синхросигналов, процессорами управления телекамерой. При этом тепловыделение самих матриц соизмеримо, а тепловыделение обслуживающих цепей в телекамерах на матрицах ПЗС оказывается даже больше, чем в телекамерах на матрицах КМОП (рис. 1.32). В свете характерных тенденций современной телевизионной техники следует отметить, что данные изображения получены с помощью твердотельной телекамеры, построенной на ИК микроболометрической матрице с КМОП мультиплексором. Схемотехнические методы компенсации неравномерных темновых сигналов хорошо освоены в телевизионных системах измерения координат точечных объектов [1, 2].

Поэтому на определённом этапе развития КМОП предельная чувствительность камер будет достигаться применением внешних устройств компенсации темновых сигналов; на последующих этапах можно ожидать повышения однородности массива фоточувствительных элементов, как это было достигнуто в ПЗС.

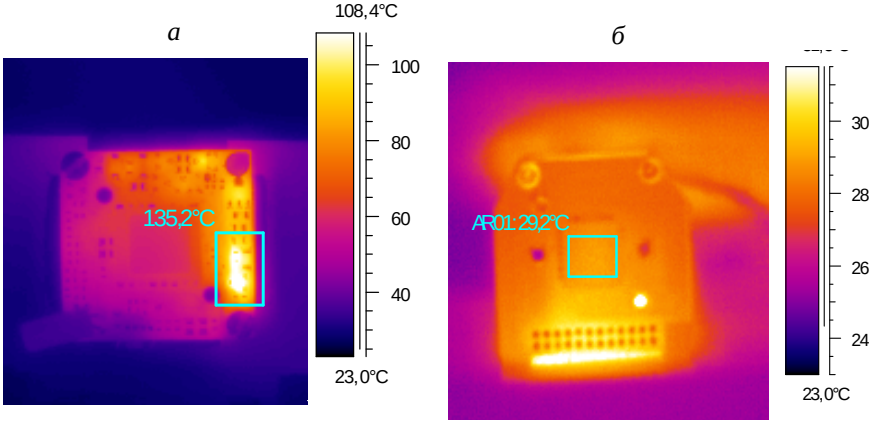


Рис. 1. 32. Распределение температур на матричной плате в телекамере на ПЗС фирмы Sony *а* и на матричной плате КМОП телекамеры на основе КМОП сенсора фирмы Omni Vision *б*

Итак, КМОП телекамеры обладают потенциальными характеристиками чувствительности, разрешающей способности, динамического диапазона, быстродействия, надёжности и компактности, позволяющими им, как высшей форме рукотворных кристаллов, стать главными видеоинформационными приборами. Они имеют целый ряд качественно новых функциональных возможностей, не реализуемых не только в электронно-лучевых приборах, но и в матричных ПЗС. КМОП телекамеры, эволюционируя к видеосистемам на кристалле, определяют наиболее перспективное направление видеоинформатики – создание интеллектуальных телевизионно-компьютерных систем, которые позволят реализовать системы технического и искусственного зрения реального времени, сопоставимые по характеристикам с биологическим зрением и существенно превосходящие его по быстродействию.

1.3. Адаптация в твердотельных телекамерах

Традиционно основной задачей адаптации в электронном телевидении было обеспечение максимального качества видеоинформации в условиях как можно большей априорной неопределённости в уровне освещения сюжета, или, что аналогично, достижение как можно большего динамического диапазона сигналов. При этом разработчики телевизионной техники в качестве реализованного в природе метода адаптации всегда «имели под рукой» зрительный анализатор человека, диапазон чувствительности которого при изменении освещённости от дневной к ночной составляет не менее восьми порядков (от 10^5 до 10^{-3}… 10^{-4} лк). Самым первым шагом на пути реализации принципа адаптации в телевидении была реализация устройств автоматической регулировки усиления видеотракта (АРУ). Борьба за расширение априорного интервала неопределённости в уровне освещённости сцены в электронно-лучевом телевидении породила разнообразные методы, связанные с управлением от видеосигнала напряжениями на различных электродах трубки (в видиконных системах – в первую очередь на сигнальной пластине). Однако и эти методы по сути своей относились к автоматической регулировке усиления и не давали радикального увеличения диапазона априорной неопределённости в освещённости сюжета. Попытки реализации адаптации времени и площади зоны накопления сигнала в электронно-лучевом телевидении предпринимались, в том числе и во ВНИИ телевидения, однако к практическому применению они оказались из-за громоздкости непригодными.

Адаптация телевизионных камер к уровню освещённости сюжета путём обмена разрешающей способности по полю и по времени на чувствительность стала реальностью только в твердотельном телевидении [1, 2]. Группирование зарядовых пакетов, позволяющее увеличивать отношение сигнал/шум при потере разрешающей способности, реализует принцип минимума информации при заданном её качестве (см. гл. 3) – количество передаваемых отсчётов и требуемая пропускная способность канала связи сокращается до тех пор, пока получатель остаётся удовлетворён качеством получаемой видеоинформации.

Одной из первых камеру с увеличенным временем накопления сигнала выпустила фирма Matsushita Electric (торговая марка

Panasonic). Режим получил название «Electronic sensitivity enhancer» и обеспечивал увеличение времени накопления от 1 до 32 телевизионных полей, т. е. с 1/50 до 0,64 с, что привело к улучшению чувствительности до 32 раз (сигнал возрастает в 32 раза, отношение сигнал/шум в $\sqrt{32} \approx 5{,}6$ раза). Телекамеры с режимом «Electronic sensitivity enhancer» выпускают практически все ведущие производители телевизионного оборудования. В таких камерах при использовании матриц ПЗС с технологией EXview HAD и асферических объективов достигается чувствительность до 0,0002 лк при отношении сигнал/шум 20 дБ. Вторым вариантом адаптивного накопления сигнала является суммирование зарядов с соседних элементов матрицы ПЗС [1, 2]. Изменяя режимы развёртки изображения, можно обеспечить сложение зарядов соседних элементов на затворе выходного транзистора и соседних строк на электродах выходного регистра ПЗС. В этом случае также происходит сложение сигнала до воздействия шума, а группирование сигналов соседних десяти элементов приводит к десятикратному увеличению сигнала. В то же время рост площади эффективной апертуры приводит к снижению разрешающей способности ТВ-камеры.

Третьим вариантом адаптивного накопления сигналов является суммирование сигналов всех каналов в цветных телекамерах при снижении освещённости сцены ниже определённого порога. Этим достигается примерно четырёхкратное повышение чувствительности ценой потери информации о цвете объектов. Выигрыш в чувствительности цветных ТВ камер при переходе в ночной чёрно-белый режим может составлять около десяти раз при исключении из оптической схемы отрезающего ИК фильтра. Однако цветопередача в дневных условиях при этом ухудшается.

Оригинальный метод адаптации реализован фирмой Matsushita Electric в так называемых камерах Super Dynamic. В них используется специальная матрица ПЗС с *двумя ячейками памяти на каждый накопительный фотодиод*. Одна из ячеек памяти используется для хранения и считывания сигналов, сформированных при большом времени накопления, другая – при малом времени накопления. Удвоение частоты считывания вертикального и горизонтального регистров обеспечивает стандартную длитель-

ность видеосигнала одного поля при наличии в нём «короткого» и «длинного» компонентов, чередующихся через строку.

Демультиплексирование и цифровая обработка этих сигналов обеспечивают преобразование характеристики свет–сигнал камеры к логарифмическому виду. В результате мгновенный динамический диапазон ТВ-камер Super Dynamic первого поколения превышает аналогичный параметр обычных ТВ камер в 20…40 раз. В камерах Super Dynamic второго поколения этот выигрыш увеличен ещё в два раза. В КМОП телекамерах динамический диапазон за счёт использования этого метода (применения нескольких различных ячеек памяти в каждом пикселе) уже превышает миллион (!) [35].

Идея мультиплексирования накопленного за различное время зарядового изображения может решаться не только методом Super Dynamic с наделением каждого пиксела двумя ячейками аналоговой памяти. Так как этот метод в ПЗС со строчным переносом обычно ведёт к уменьшению динамического диапазона каждой из этих ячеек памяти, то радикальное увеличение динамического диапазона (расширение априорного интервала освещённостей сцены) может быть реализовано и в матрицах ПЗС с кадровым переносом. Для этой цели в них, так же, как и в Super Dynamic, организуются два кадра с различной экспозицией и требуется два массива аналоговой памяти, но благодаря вынесению памяти из секции накопления удаётся расширить динамический диапазон без уменьшения ёмкости ячеек хранения зарядового изображения [138]. При не слишком большой подвижности наблюдаемой сцены могут быть организованы не только пары кадров с чередующимся временем накопления, но и группы кадров по три, для чего может быть использована матрица, показанная на рис. 1.5, с тем лишь отличием, что роль секции накопления будет выполнять лишь одна из четырёх секций, а три другие должны быть защищены от света.

Особое место в ряду адаптивных систем занимают системы адаптации параметров разложения (автоматического изменения числа группируемых зарядовых пакетов) при адаптации телевизионной системы к этапам наблюдения, т. е. при переходе от обнаружения объектов к оцениванию их координат [2, 69–72] (см. главу 2).

Новые функциональные возможности адаптации связаны с возможностью автоматического управления КМОП телекамерами к изменениям не только уровня освещённости, но и других свойств

сюжета – динамике наблюдаемых объектов и их числу. Так, одной из самых востребованных функций является задание размера и положение фрагмента изображения, с которого производится считывание видеосигнала (Region of interest). При этом реализуется обмен поля зрения и числа элементов фотоприемника на частоту кадров, т. е. при уменьшении размеров фрагмента в соответствующее число раз возрастает кадровая частота. Для наилучшего выполнения целевой функции КМОП телекамерой для выделения глобального движения используются не все пикселы, а только т. н. «умные» пикселы, располагающиеся на периферии фоточувствительной матрицы. Эти сложные пикселы отличаются от остальных элементов изображения повышенной сложностью и включают в свой состав 20 дополнительных транзисторов для выполнения вычислительных функций. Такое усложнение структуры пиксела приводит к уменьшению коэффициента заполнения и, как следствие, ухудшению фотоэлектрических характеристик такого пиксела, поэтому в центральной области изображения применение таких пикселов нецелесообразно. В итоге переменная по полю архитектура позволяет на 99% сократить быстродействие и/или сложность вычислителя путём переноса этапа предобработки изображения непосредственно в фотоприёмник. Такая структура фотоприёмника с ещё большим основанием может быть названа *foveal system*, так как обеспечивает свойства телевизионной системы, ещё более близкие к свойствам зрительного анализатора, у которого уменьшение разрешения на периферии зрения сопровождается повышенной чувствительностью этой периферии к движению объектов. Это делает КМОП телекамеру аналогично зрительному анализатору *адаптивной не только к сюжету, но и к этапам наблюдения:* на первом этапе наблюдения максимизируется чувствительность системы, на втором – точность оценки параметров обнаруженных объектов.

Рассмотренные методы согласования параметров телекамеры с наблюдаемой сценой не исчерпывают функциональных возможностей современных формирователей видеосигнала. При этом при использовании матриц ПЗС включение в состав телекамеры микропроцессора позволяет реализовать ряд новых функций: электронное увеличение, детектирование движения в выбираемых оператором зонах, инверсию уровня белого и т. д. [79]. Адаптация считывания в телекамерах на ПЗС применяется не только при формировании

сигнала изображения, но и при прецизионной оценке темнового сигнала в астрономических малокадровых телекамерах [80]. Все указанные методы адаптации параметров телекамер при использовании КМОП матриц либо уже реализованы, либо принципиально могут быть реализованы без использования внешнего микропроцессора, целиком на основе вычислительных ресурсов однокристальной телекамеры.

Несомненно, что по мере развития прикладного телевидения и увеличения сложности решаемых с его помощью задач, количество перестраиваемых параметров и диапазоны регулировок в твердотельных телекамерах будут увеличиваться.

1.4. Твердотельные телекамеры инфракрасного диапазона

Прикладная телевизионная техника переживает новый этап революционного изменения технологии создания телевизионных систем не только видимого диапазона длин волн, но и инфракрасного диапазона. Матрицы ИК диапазона используют те же принципы развёртки изображений, что и матрицы видимого диапазона – зарядовую связь или координатную адресацию с помощью КМОП мультиплексоров. Основные отличия твердотельных матричных фотоприёмников ИК диапазона от матриц видимого диапазона приходятся на первый этап рождения сигнала – фотоэлектрический эффект, для обнаружения которого на фоне теплового шума обычно требуется охлаждение фотоприёмника.

Выделяют четыре поколения тепловизионных приборов [139–142, 145–148]. Нулевое поколение основано на применении одноэлементных приемников излучения и двумерной развертки изображения с помощью сканирующей оптико-механической системы (применяемое в геостационарных метеорологических системах). Первое поколение основано на применении одномерных линеек фотоприёмников и одномерной оптико-механической развёртки изображения, второе поколение – на применении матриц фотоприёмников в виде 2…6 линеек с ВЗН (временная задержка и накопление) и одномерной оптико-механической развёртки изображения.

Новое третье поколение основано на применении двумерных многоэлементных матриц фотоприемников (FPA – Focal Plane

Array) с параллельным накоплением сигнала и «самосканированием» без использования оптико-механических систем развертки изображения, которые разработчиками механических сканеров были не корректно названы «смотрящими» [142, 144] и которые во многих приложениях вытеснили механически сканирующие системы (конечно, тоже «смотрящие»). В 90-е годы XX века появились тепловизионные приборы на QWIP матрицах или QWIP детекторах (QWIP – Quantum Well Infrared Photodetector) с высокой технологичностью, воспроизводимостью, однородностью параметров по элементам с форматом от 256×256 до 640×512. QWIP матрицы обладают способностью управления спектральной чувствительностью и возможностью перейти в будущем от гибридных структур фокальных матриц к монолитным [141]. Современные криогенно охлаждаемые матричные приёмники могут обладать ёмкостью элемента до 10^6…10^7 электронов, квантовым выходом 80%...99,9% и собственным шумом считывания 1…20 электронов на элемент [145].

Качественный скачок в характеристиках ИК телекамер обусловлен как наращиванием числа элементов в криогенно охлаждаемых фотоприёмных матрицах (в первую очередь ПЗС с барьерами Шоттки и матрицах с координатной адресацией), так и освоением *неохлаждаемых* матричных приборов с чувствительными элементами на основе *микроболометров* и КМОП устройствами считывания видеоинформации. В качестве материалов для *пироэлектрических* приемников излучения используются цирконаты свинца, ниобаты и титанаты бария–стронция, сополимеры виниленфторида. В качестве материалов для микроболометров используются модификации окислов ванадия V_xO_y, поликристаллический и аморфный кремний. Современные ИК матрицы *высокой квантовой эффективности* могут быть выполнены на основе различных материалов – халькогенидов свинца (PbS, PbSe), соединения кадмий–ртуть–теллур – HgCdTe (КРТ), антимонида индия (InSb), силицида платины (PtSi), примесных кремния (Si:x) и германия (Ge:x), многослойных структур с квантовыми ямами на базе GaAs/AlGaAs – QWIP детекторов [86–93, 139–145].

При выборе проектировщиком типа фотоприёмника для конкретной задачи важнейшей характеристикой является зависимость его чувствительности от длины волны излучения. Так как чувствительность телевизионной системы зависит от параметров

объектива, времени накопления и размеров элемента разложения (см. раздел 1.1), то создатели ИК фотоприёмников часто пользуются характеристикой «удельной» чувствительности D^* [см×Гц$^{1/2}$×Вт$^{-1}$] (от англ. detectivity – обнаружительная способность), зависящей от длины волны излучения (рис. 1.33). Эта характеристика зависит от длины волны излучения λ, квантового выхода фотоприёмника на выбранной длине волны и интенсивности потока фотонов фона N_B: $D^*(\lambda) = \frac{\lambda}{2hc}\sqrt{\frac{\eta(\lambda)}{N_B}}$. Обнаружительная способность фотоприёмника D^* не зависит от площади элемента разложения и является характеристикой материала, из которого изготовлена светочувствительная поверхность фотоприёмника в конкретных условиях наличия внешнего фона с заданной интенсивностью.

Синтез фотоприёмника с требуемой спектральной характеристикой основан на создании полупроводниковых материалов с шириной запретной зоны, соответствующей энергии квантов принимаемого электромагнитного излучения. Для инфракрасных систем пороговый контраст может быть вычислен, но он уже не является достаточно наглядной характеристикой чувствительности. Гораздо лучше её отражает пороговая разность температур. Основной исходной посылкой для её вычисления является всё то же ограничение чувствительности флуктуациями фотонного потока, распределённого по закону Пуассона.

При выборе спектральной характеристики материала фотоприёмника (рис. 1.33) необходимо учитывать:

- наличие отдельных «окон прозрачности» атмосферы;
- наличие максимума приращения количества излучаемых квантов, который приходится не на максимум спектральной плотности (длина волны которого λ_{max} вычисляется по формуле Вина), а на более короткую длину волны $\lambda_{max\delta T} \approx 0{,}8\lambda_{max}$.

Чувствительность системы ИК диапазона чаще всего характеризуют разностью температур, эквивалентной шуму или пороговой разностью температур. Эти характеристики отличаются множителем – пороговым отношением сигнал/шум, которое, как и при определении чувствительности систем видимого диапазона, в разных задачах может иметь несколько различные значения в пределах 2…10.

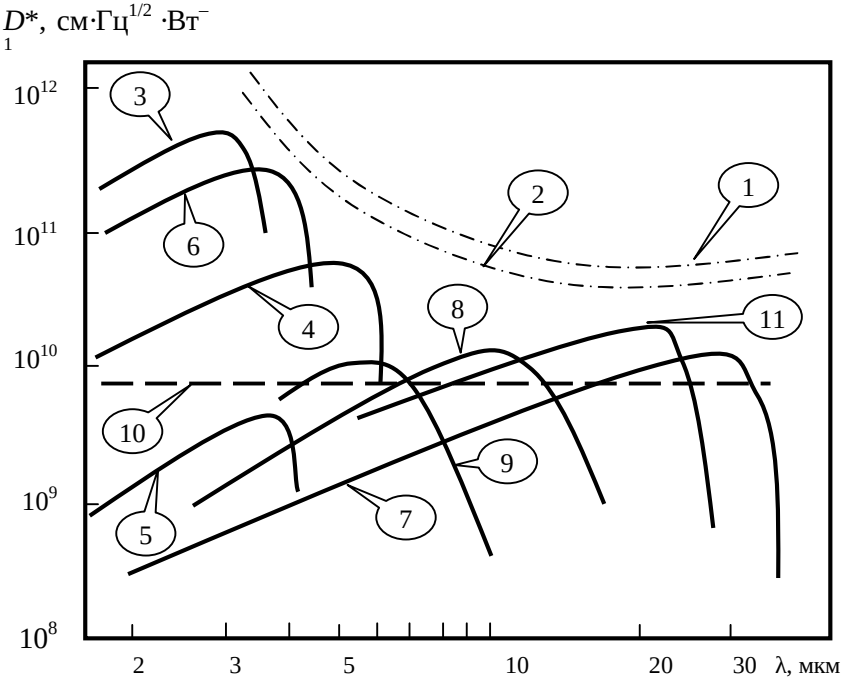


Рис. 1.33. Спектральные характеристики чувствительности различных материалов, применяемых для наблюдения в ИК-области спектра: *1* – предельное значение для фотовольтаического детектора с полем зрения 2π при ограничении чувствительности фотонным шумом фона с температурой 300 К; *2* – то же для фотосопротивлений; *3* – InAs при 77 К; *4* – InSb при 77 К; *5* – InAs при 300 К; *6* – PbS при 77 К; *7* – Ge:Zn при 4,2 К; *8* – Ge:Hg при 30 К; *9* – Ge:Au при 77 К; *10* – термисторный болометр при 300 К; 11 – Si:B при 10 К

Как и пороговая освещённость, пороговая разность температур зависит от порогового отношения сигнал/шум $\Psi_{пор}$, относительного отверстия объектива *D/f*, квантового выхода фотоприёмника η и накопительных свойств – размера элемента Δ и времени накопления $T_н$. Дополнительно учитываются коэффициент серости излучателя $k_с$ и коэффициент использования спектрального диапазона μ. Пороговая разность температур составляет

$$\delta T_{\text{пор}} = \frac{2\Psi_{\text{пор}}}{3\left(\frac{D}{f}\right)\Delta\sqrt{\sigma' T \tau k_{\text{с}} \mu T}}.$$

Подробно расчёт обнаружительной способности системы рассмотрен в работах [44, 45, 49]. Пороговая разность температур определяется пользователем системы и включает назначаемое им пороговое отношение сигнал/шум. Так как изготовители ИК матриц рассчитывают на различных пользователей, то они пользуются другой характеристикой чувствительности системы – *эквивалентной шуму разностью температур* (Noise Equivalent Temperature Difference). Этот показатель, соответственно, в несколько раз меньше и измеряется уже в милликельвинах (см. табл. 1.4). Для современных матричных ИК телевизионных систем эквивалентная шуму разность температур при наблюдении крупной детали составляет 15…100 мК. Вместе с тем, так же как и при наблюдении в видимом диапазоне, контрастная чувствительность при увеличении пространственной частоты ухудшается (рис. 1.34).

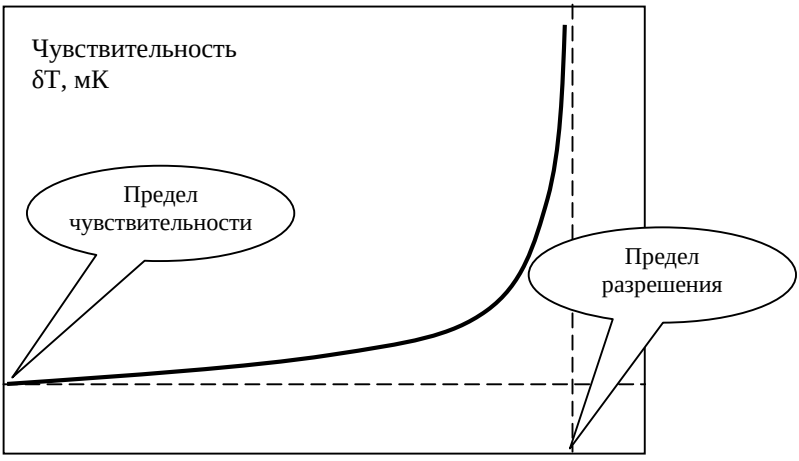


Рис. 1. 34. Зависимость эквивалентной шуму разницы температур от пространственной частоты

Как видно из рис. 1.33, квантово-предельная чувствительность ИК систем достигается при криогенном охлаждении фотоприёмника, а работающие при нормальной температуре 300K болометры уступают им по чувствительности. Поэтому во многих, в первую очередь оборонных, приложениях предпочтение отдаётся приборам с использованием ИК матриц на базе КРТ. Самая высокая чувствительность и возможность их работы в области спектра 1…20 мкм является важным преимуществом.

Однако отсутствие газовой холодильной машины, снижающей ресурс системы, обуславливает перспективу массового применения именно микроболометрических матриц. Основными преимуществами этих приборов являются: отсутствие оптико-механической развертки изображения и, соответственно, малые масса, габариты и энергопотребление, отсутствие вибраций.

Принцип работы микроболометра заключается в изменении сопротивления материала при поглощении ИК излучения, а устройство матрицы определяется набором таких микроболометров, опрашиваемых КМОП мультиплексорами аналогично КМОП приборам видимого диапазона. Типичным представителем этих приборов является длинноволновый неохлаждаемый микроболометрический детектор UL01021E французской фирмы Sofradir. Он имеет структуру, показанную на рис. 1.35 и следующие характеристики:

- число элементов разложения 320×240;
- светочувствительная площадь 14,4×10,8 мм^2;
- шаг между элементами 45 мкм;
- филл-фактор более 80%;
- постоянная времени (по уровню –3 дБ) примерно 4 мс;
- частота считывания до 5 МГц (кадровая частота 65 Гц).

Хотя в инфракрасной технике электронно-лучевые приборы ещё не совсем сошли со сцены (ещё применяются пировидиконы и видиконы с большой площадью мишени из InSb), однако эти «светочувствительные радиолампы» методично вытесняются матричными фотоприёмниками. В области создания ИК матриц интенсивно работает множество фирм, среди которых Raytheon IR Operation, Boeing, Lockheed-Martin, FLIR Systems, Hughes, Sofradir, ЦНИИ «Электрон», НПО «Орион», создающие для унификации собственных разработок системы модульного типа. В мире

20 фирм в семи странах выпускают более 50-ти типов ИК матриц, примеры продукции которых представлены в табл. 1.4 [142, 144].

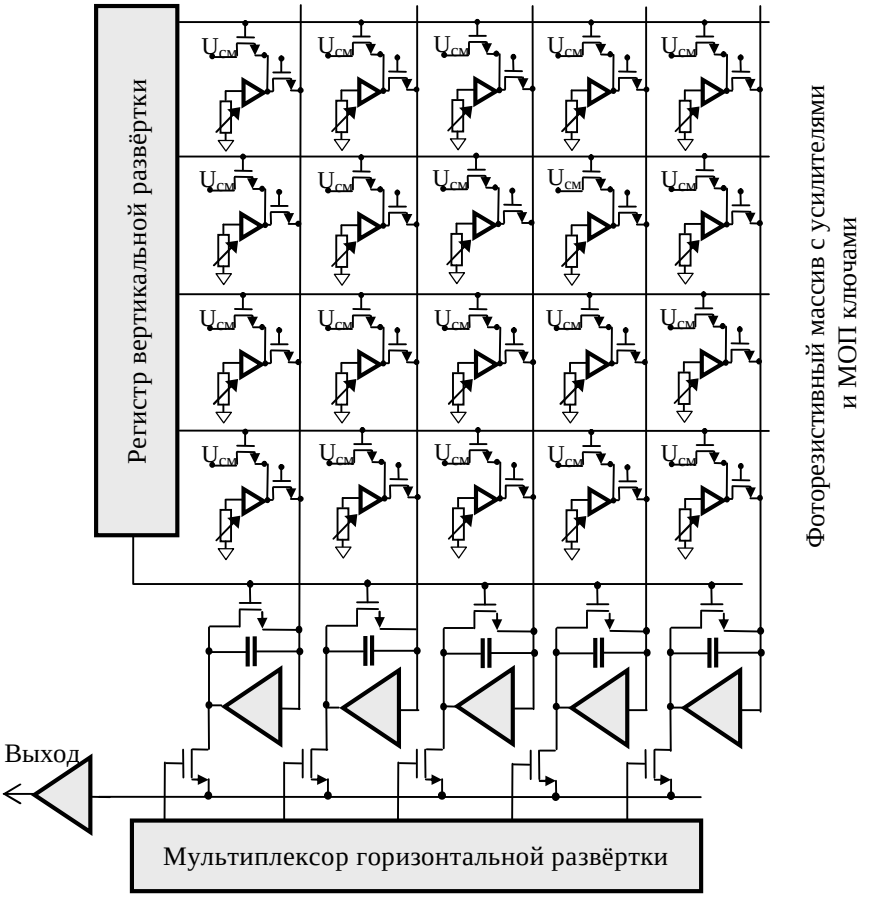


Рис. 1.35. Микроболометричекая матрица для дальнего ИК диапазона с КМОП мультиплексорами

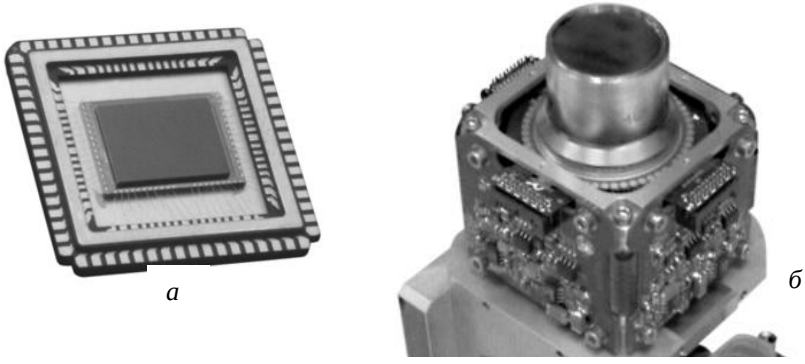
Таблица 1.4.

Тип матрицы	Диапазон мкм	Рабочая температура К	Эффективные пикселы (H×V)	Чувствительность мК	Размер пиксела мкм
Raytheon, США					
QWIP	8…12	70	256×256	15	28×28

QWIP	8…12	70	640×486	25	24×24
Микро-болометр	8…14	300	320×236	100	50×50
LIR, Франция					
КРТ	3…5	77	640×480	20	23×23
Rockwell, США					
КРТ	8…12	77	256×256		40×40
Hughes, США					
КРТ	3…4,5	300	256×256	50	40×40
КРТ	3…4,5	300	256×256	65	30×30
InSb	0,5…5,4	80	256×256		30×30
InSb	0,5…5,4	80	640×480		20×20
InSb	0,5…5,4	30…50	1024×1024		27×27
InSb	0,5…5,4	30…50	2048×2048		27×27
Mitsubishi, Япония					
Ge:Si/Si (барьер Шоттки)	8…12	43	512×512	80	34×34
ЦНИИ “Электрон”, Россия					
PbS	1,5…4	80	128×128	20	60×60
PtSi (барьер Шоттки)	0,5…5	80	340×280	30	45×45
НПО “Орион”, Россия					
InSb	0,5…5,4	80	256×256	50	50×50
КРТ	8…12	80	256×256	50	50×50
LIR, Франция					
КРТ	3…5	77	640×480	14	23×23
Микро-болометр	8…14	300	256×64	50	50×50
Lockheed, США					
Микро-болометр	8…14	300	320×240	50	48×48
Микро-болометр	8…14	300	327×245	100	25×25
MES, Великобритания					
Пиро-электрик	8…12	300	384×288	130	40×40

Рынок этих систем жестко контролируется. Для экспорта тепловизоров с матрицей 640×480 пикселей из США требуется разрешение Президента страны. Это товар стратегического значения. В России предприятия «Циклон» и «Орион» ведут разработки микроболометрических матриц [144]. До настоящего времени подавляющее большинство тепловизионных систем, использующихся в России и за рубежом, были на основе криоохлаждаемых преобразователей, что резко ограничивало их широкое применение.

Внешний вид типичной ИК матрицы (рис. 1.36*а* [144]) мало отличается от внешнего вида КМОП матрицы видимого диапазона, показанного на рис. В.9. Основное отличие, как показывает сравнение табл. 1.3 и табл. 1.4, состоит в том, что ИК матрицы с тем же числом элементов, что и матрицы видимого диапазона, имеют существенно большие размеры. В целом же ИК телекамеры даже по внешнему виду (рис. 1.36*б* [144]) заметно отличаются от телекамер видимого диапазона в силу специфических требований по прочностным и температурным характеристикам. Одно из отличий состоит в требовании к спектральным характеристикам объективов и входного окна (обычно выполняемым из германия, обеспечивающего в силу большей диэлектрической проницаемости и, соответственно, коэффициента преломления, более высокое качество изображения и «отсечку» фона, т. е. коротковолнового диапазона длин волн до 2 мкм).



а *б*

Рис. 1.36. ИК матрица *а* и матричный блок с ней с электронными платами и германиевым входным окном *б*

ИК телекамера, также как и телекамеры видимого диапазона, содержат объектив и несущую конструкцию, однако отличаются от них конструкцией приёмного модуля (рис. 1.37 [144]), содержащей

ИК матрицу, охладитель, работающий по циклу Сплит-Стирлинга и электронный блок, включающий буферный усилитель и аналого-цифровой преобразователь.

Рис. 1.37. Приёмный криогенно охлаждаемый модуль ИК твердотельной телекамеры

Важными отличительными чертами ИК матриц и телекамер на них по сравнению с телекамерами видимого диапазона всегда будет больший на порядок размер элемента разложения. Это свойство имеет причинами волновую природу излучения (размер функции рассеяния точки объективом пропорционален длине волны) и квантовую структуру излучения (для борьбы с фотонным шумом интенсивного внешнего и приборного фона требуется большая накопительная способность элемента). Кроме того, в силу более сложной технологии изготовления ИК матрицы отличаются от матриц видимого диапазона существенно большим количеством дефектных элементов и разбросом чувствительности элементов по полю. Неоднородности сигналов элементарных фотоприёмников матрицы предварительно корректируются в аналоговой форме, преобразуются в цифровую и корректируются с использованием данных, полученных в процессе калибровки. В результате получаются тепловые изображения, в каждом конкретном сеансе калибруемые по соответствующей шкале температур (см. рис 1.32, рис. 1.38).

Существенно более высокая стоимость ИК матриц обостряет внимание к повышению разрешающей способности телекамер на их основе какими-либо иными методами кроме повышения числа элементов разложения. Особенно остро проблема нехватки чёткости стоит в системах различения удалённых объектов, когда дискретная структура растра ИК матрицы, усугублённая зазорами между элементами, приводит к существенным искажениям формы сигнала, в том числе проявляющимся в неинвариантности формы отклика к форме оптического воздействия. Эта проблематика является нетрадиционной для классической радиотехники, так как в отличие от

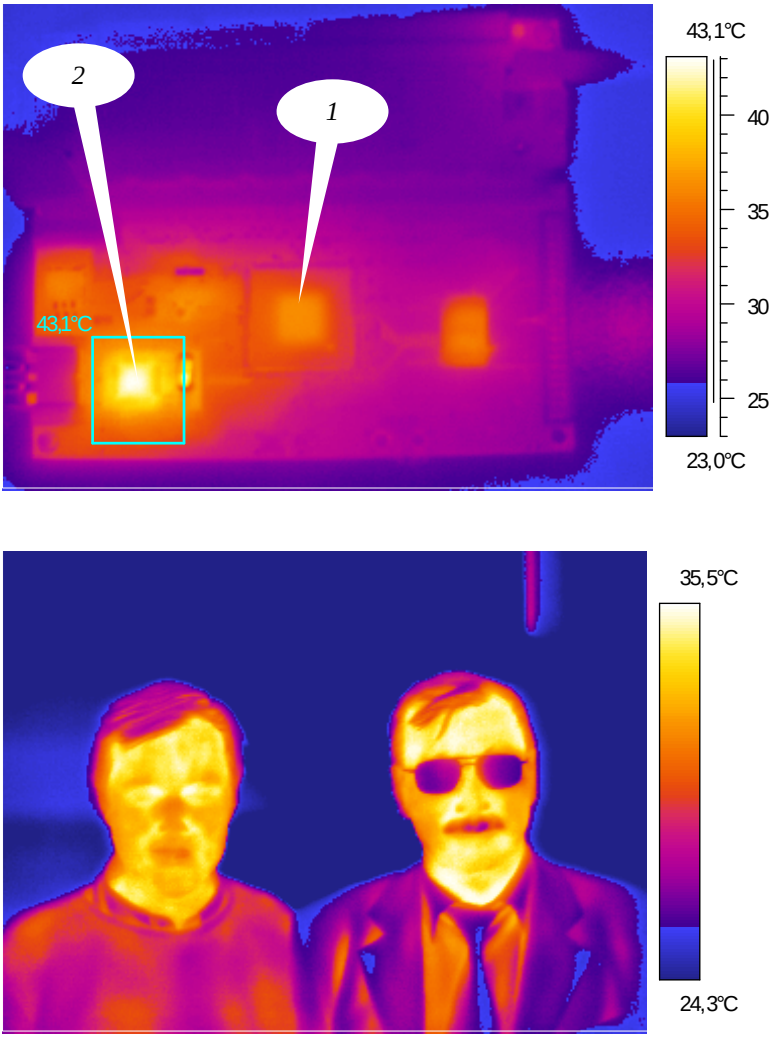


Рис. 1.38. примеры изображений, формируемых ИК микроболометрической телекамерой: плата обработки видеоинформации *1* с системой на кристалле и интерфейсом Ethernet *2* (см. раздел 4.2) *а* и групповой портрет *б*

систем с непрерывным входом и непрерывным выходом и систем с дискретным входом и дискретным выходом матричные фотоприёмники при непрерывном входном воздействии формируют дискретный отклик. Как следствие, даже при линейности свет–

сигнальной характеристики, аналого-дискретные преобразователи уникальны: их импульсные характеристики не связаны с их частотными характеристиками преобразованием Фурье. Для оценки таких искажений введена мера неинвариантности системы и предложен метод взаимообмена разрешающей способности по времени на разрешающую способность по полю [2]. Этот метод состоит из формирования нового кадра с увеличенным количеством отсчётов, полученных по совокупности нескольких кадров при смещении фотоприёмной матрицы относительно визирной оси телевизионной системы. Такое смещение может создаваться с помощью специальных пьезоэлектрических двигателей, а в некоторых системах, например, космических, может быть следствием естественного дрейфа визирной оси системы. В силу интегрирования света внутри элемента такой новый кадр не обеспечивает полной разрешающей способности, которую обеспечивает матрица с аналогичным количеством действительных, а не виртуальных элементов. Например, при сдвиге матрицы в фокальной плоскости на полэлемента по каждому из направлений и формировании синтезированного кадра по двум таким «полукадрам», частота Найквиста действительно удваивается, однако глубина модуляции на частотах, превышающих частоту Найквиста для исходного кадра в таком синтезированном кадре существенно меньше [2]. Для построения синтезированного кадра по нескольким смещённым на долю элемента изображениям требуется пропорциональное увеличение ёмкости оперативной памяти. Значение количества дополнительных позиций должно получаться не эвристически, а аналитически, как решение задачи на минимум информационного риска (см. раздел 3.1), составленного из потери полезной информации из-за дискретности растра (уменьшающейся с увеличением позиций смещаемой матрицы), и сложности устройства восстановления сигнала, увеличивающейся с ростом этих позиций. Возможны другие варианты информационного риска, например, учитывающий не просто сложность устройства обработки, а его загрузку обработкой присутствующих в изображении шумов (в первую очередь – шумов фона). Количество дополнительных смещённых кадров заранее не известно и определяется с учётом априорной информации о сюжетах и весовых коэффициентов для потери полезной информации и для вычислительных затрат. Дополнительные позиции матрицы, при возможности управления ими, вводятся не в произвольные точки растра, а в минимизирующие ошибку

восстановления изображения и максимизирующие эффективность использования вводимого кадрового запоминающего устройства – в точки, равноудалённые от центров ближайших элементов. Для восстановления сигнала в окрестности центра элемента при наличии четырёх дополнительных отсчётов требуется формирование и запоминание всего одного дополнительного кадра, сформированного со сдвигом относительного первого на пол-элемента по вертикали и пол-элемента по горизонтали. Из этого одного дополнительного кадра при интерполяции сигнала в окрестностях центра каждого элемента потребуется обратиться к четырём элементам разложения, окружающим элемент исходного кадра .

Матричные ИК приборы во многих приложениях вытесняют сканирующие приборы не только в силу эксплуатационных преимуществ. Массово используемый ИК диапазон не простирается далее 14 мкм, а для этих длин волн при типовой температуре фотоприёмника менее 300К шум квантовой структуры внешнего фона существенно больше тепловых шумов приёмника. Это означает, что накопление сигналов во времени при прочих равных условиях должно привести к увеличению отношения сигнал/шум пропорционально корню квадратному из отношения времени накопления. Это значит, что матричные ИК фотоприёмники, имеющие более 10^4 элементов, потенциально должны обеспечивать отношение сигнал/шум на два порядка больше, чем сканирующие одноэлементные и на порядок больше, чем сканирующие строчные фотоприёмники. Поэтому выбор типа развёртки, так же как и в телевизионных системах видимого диапазона, осуществляется согласно «золотому» правилу малокадрового телевидения: если нужны широкий угол зрения и большая разрешающая способность на местности, то следует применить сканирующую систему (строчную или одноэлементную), если нужна высокая слитность воспроизведения движения объектов, то следует применить систему с кадровым накоплением. А достигаемый с помощью сканирующих одноэлементных систем выигрыш в амплитуде отклика от точечной цели примерно в два раза (нет потери на не инвариантность к сдвигу) не может компенсировать проигрыша из-за уменьшения времени накопления. Это означает, что матричные фотоприёмники наилучшим образом осуществляют согласование источника сигнала с получателем, считая фотоны группами (см. раздел 3.3). Следует говорить не о борьбе с «недостатками» матричного фотоприёма, а о минимизации информационного риска (см. разд. 3.1.) в ходе извлечения информации о полезном сигнале из входного потока фотонов системой на матричных фотоприёмниках.

Глава 2

ДИСКРЕТНО-АНАЛОГОВАЯ ОБРАБОТКА СИГНАЛА ИЗОБРАЖЕНИЯ

Важнейшим элементом концепции построения твердотельных телевизионных систем с использованием видеосистем на кристалле (ВСнК), кроме упомянутого во введении *распределения площади* между вычислителем и фотоприёмником, является *распределение обязанностей* между ними по ряду процедур кодирования изображений. В задачу проектировщика входит разгрузка вычислителя от некоторых рутинных задач, которые быстрее и точнее могут быть выполнены фотоприёмником. И хотя в силу влияния шумов связь может принести лишь конечное количество информации и в принципе является цифровой, ряд операций при выполнении в аналоговом виде могут обеспечить и более высокую помехоустойчивость, и существенно сократить потребные *время* и *площадь* кристалла. Важно, что при построении ВСнК задачи распределения площади и обязанностей между вычислителем и фотоприёмником взаимосвязаны. Такая связь обусловлена тем, что операция запоминания значения сигнала одного элемента разложения в аналоговом виде требует площади, равной площади накопительного элемента (порядка 1 мкм^2 на тысячу электронов), а для этой же операции в цифровом виде потребуется несколько десятков транзисторов не считая АЦП.

Ряд приёмов кодирования изображений в ходе дискретно-аналоговой обработки сигналов, посвящённых достижению потенциальной помехоустойчивости в задаче совместного обнаружения–оценивания сигналов точечных объектов, опирается на ту же концепцию уменьшения влияния шума считывания, что и при реализации обмена разрешающей способности на чувствительность [1] (см. разделы 1.3, 2.1). Различие состоит в том, что *группирование зарядовых пакетов, реализующее принцип минимума информации при заданном её качестве* (см. гл. 3), осуществляется не одинаково по всему полю, а в двух вариантах неоднородности. В системах с большой априорной неопределённостью о координатах и интенсивности сигналов, например, в астродатчиках, группирование сигналов соседних элементов производится адаптивно, в зависимости от интенсивности сигнала точечного объекта. В системах

наведения на единственную точечную цель переменная по полю чёткость характеризуется наличием зоны полной чёткости в районе визирной оси системы, окружённой полем с сокращённой чёткостью. Разделение изображения на зоны с различной разрешающей способностью по полю и/или по времени целесообразно не только при наблюдении точечных объектов, но и как процедура, предшествующая семантическому анализу изображений. Выделение *доминантных* зон (или, что то же самое, отнесение прочих зон к фону или контексту) позволяет получить ряд важных преимуществ. Можно отметить повышение чувствительности в зонах с пониженным разрешением, но важнейшим полезным эффектом от переменной по полю чёткости является повышение кадровой частоты за счёт уменьшения числа считываемых с матрицы строк. Повышение кадровой частоты твердотельной телевизионной системы может быть реализовано как для всей матрицы в случае ПЗС, так и для её отдельных участков в случае КМОП.

Общим в согласовании источника оптического сигнала с вычислителем координат при наблюдении точечных и сложных объектов является согласование размеров импульсной характеристики объектива («кружка рассеяния») с размером элемента разложения в матрице. Такое согласование должно учитывать цель системы (критерий качества) – обнаружение объекта или измерение его координат, а также контраст объекта относительно фона и уровень собственных шумов телекамеры.

С методологической точки зрения в группе приёмов кодирования изображений дискретно-аналоговыми методами выделяется метод скоростного формирования сигналов проекций накопленного изображения [2] без считывания сигнала изображения через единственное устройство, наличие которого в матрицах предполагает использование единственного канала связи. Считывание накопленного при этом методе сигнала производится с суммированием зарядов каждого столбца непосредственно в матрице, т. е. реализуется принципиально нереализуемая в электронно-лучевых трубках *параллельная обработка* с числом каналов, равным числу столбцов в фотоприёмной матрице, обеспечивающая выигрыш в быстродействии на три порядка. Телекамеры становятся источником с несколькими выходами («множественный поток данных») и почти одновременно могут формировать и сигналы проекций, и обычный сигнал изображения.

2.1. Обнаружение и оценивание координат объектов

Обнаружение сигналов точечных объектов. В простой постановке задачи обнаружения оптического сигнала точечного объекта (звезды) $\alpha S(x - x_o)$ считается известной форма сигнала $S(x)$ и неизвестными амплитуда сигнала α и координата x_o. При проверке гипотезы о наличии сигнала минимум риска достигается оптимизацией значения порога и максимизацией отношения сигнал/шум, обеспечивающей максимум правдоподобия при аддитивном гауссовском шуме. Для упрощения синтеза аналого-дискретной обработки сигнала в фотоприёмнике твердотельной телекамеры обычно выбор порога считают известной задачей: при равенстве цен ошибок и вероятностей гипотез порог обеспечивает равенство вероятностей ошибок первого и второго рода. При этом максимум правдоподобия достигается при максимуме отношения сигнал/шум Ψ. Для его вычисления учитывают различие дисперсий шума на уровне фона и на сигнале, считая шумы сигнала, фона и телекамеры независимыми. Тогда зависимость отношения сигнал/шум от числа сигнальных N_s фотоэлектронов, числа фоновых N_b фотоэлектронов и дисперсии D шума считывания составляет:

$$\Psi^2 = \frac{N_s^2}{N_s + N_b + D}. \qquad (2.1)$$

Форма полезного сигнала обычно не известна абсолютно точно; бывают известны диапазоны изменения её основных характеристик. Обычно можно считать, что априорная неопределённость не функциональная, а параметрическая. В этом случае вариации формы полезного сигнала могут быть сведены к вариации некоторых параметров: амплитуды, радиуса (зависящего от качества фокусировки объектива), эксцентриситета (зависящего от наличия комы объектива) и т. п. Возможны несколько путей решения задачи совместного обнаружения и оценивания в условиях параметрической априорной неопределенности. Один из путей – нахождение оптимальных процедур для среднего значения параметров с последующей оценкой устойчивости показателей качества к отклонению параметров. Второй путь – применение алгоритмов, не являющихся оптимальными для принятой модели типичного сигнала, но обладающих большой устойчивостью к изменению параметров сигнала [4, 65, 68]. При обоих подходах форму полезного сигнала либо варьируют (напр., [60]), либо просто

принимают гауссовской с радиусом r и неизвестными амплитудой α и координатой x_o.

$$s(x) = \alpha S(x - x_o) = \frac{\alpha}{\sqrt{2\pi}} \exp\left[-\left(\frac{x - x_o}{2r} \right)^2 \right]. \qquad (2.2)$$

Усреднённое по фазе полезного сигнала относительно дискретной структуры твердотельного фотоприёмника отношение сигнал/шум вычисляется с учётом шумов фона и шумов считывания, дисперсии которых не зависит от номера отсчета:

$$\Psi^2 = \frac{\frac{1}{\Delta}\left[\int_0^{\Delta} dx_o \int_0^{\Delta} \alpha S(x - x_o)\, dx \right]^2}{\frac{1}{\Delta} \int_0^{\Delta} dx_o \int_0^{\Delta} \alpha S(x - x_o)\, dx + B\Delta + D}. \qquad (2.3)$$

При оценке амплитуды сигнала и проверке гипотезы о его наличии обычно рассматривается отсчёт, содержащий максимум сигнала, при оценке координаты сигнала – группа отсчётов, содержащая максимальный сигнальный отсчет. При обнаружении сигнала полагается, что решение о его наличии принимается путем сравнения амплитуд сигналов без дополнительной фильтрации. Это означает, что согласование интервала дискретизации с полезным сигналом звезды (по радиотехнической терминологии – с импульсной характеристикой объектива, по оптической – с функцией рассеяния точки, по телевизионной – с «кружком рассеяния» объектива) является исчерпывающей процедурой оптимизации. Поэтому в первом приближении можно положить $x_o = (i + ½)\Delta$ и, введя обозначение $\Phi(x)$ для табулированного интеграла вероятности $\Phi(x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$ привести формулу (2.3) к виду:

$$\Psi^2 \approx \frac{\alpha^2 \Phi^2\left(\frac{\Delta}{2r}\right)}{\alpha \Phi\left(\frac{\Delta}{2r}\right) + B\Delta + N_c}. \qquad (2.4)$$

Анализ показывает, что максимум помехоустойчивости обнаружения зависит от интенсивности α сигнала звезды и интенсивности B фона (рис. 2.1). Отношение α/B, которое может трактоваться как потенциальный контраст звезды относительно

фона, определяет не только достижимое отношение сигнал/шум, но и оптимальный размер элемента разложения.

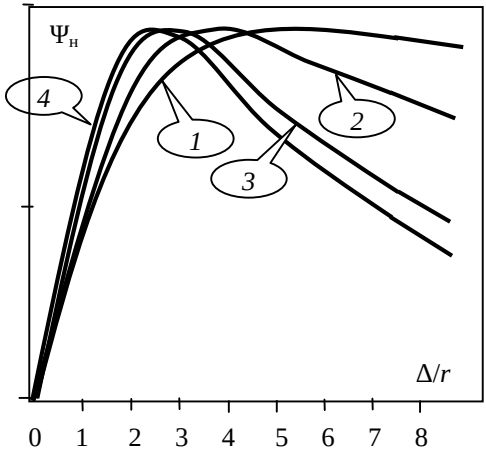


Рис. 2.1. Зависимость нормированного отношения сигнал/шум $\Psi_н^2 = \Psi^2 / \Psi_{max}^2$ от соотношения размеров элемента разложения и функции рассеяния точки объективом Δ/r при постоянном уровне полезного сигнала α=100 и шума считывания D=3 и различных значениях фона
1) B=1,
2) B=10;
3) B=100,
4) B=1000.

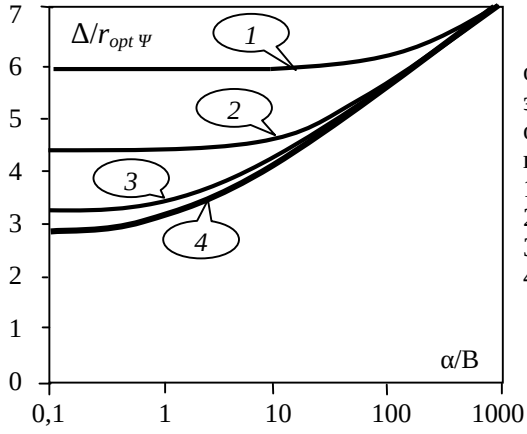


Рис. 2.2. Зависимость оптимального размера элемента разложения при обнаружении от потенциального контраста :
1) B=10, D=1000;
2) B=10, D=100;
3) B=10, D=10;
4) B=10, D=1;

Зависимость оптимального размера элемента от потенциального контраста (рис. 2.2) может быть выражена приближенной формулой:

$$\Delta_{o\,\psi} \approx 1{,}45\, r \max\left\{\left[1{,}5 + \log\frac{\alpha}{B}\right];\ \max\left[1{,}9;\ \left(2 + \log\frac{D}{B}\right)\right]\right\}.$$

В широком диапазоне изменения соотношения шумов фона и считывания оптимальный по критерию отношения сигнал/шум нормированный на радиус сигнала звезды размер элемента разложения изменяется в пределах от 3 до 7 (рис. 2.2). Формула (2.5) позволяет осуществить предварительную оценку помехоустойчивости обнаружения и найти ориентировочное значение оптимального соотношения размеров зоны накопления и кружка рассеяния объектива не только в хорошо исследованном случае преобладания шумов фона, но и в условиях действия трёх источников шума, т. е. в широком диапазоне изменения контраста звезды относительно фона.

Оценивание координат сигналов точечных объектов. Построение оптимальной телевизионной системы для второй задачи астрономических телевизионных систем – оценивания координат (параметров смещения) изображения – распадается на две *взаимосвязанные* задачи:

- вычисление оптимального интервала дискретизации Δ непрерывного оптического сигнала*;*
- выбор оптимального (приемлемого по вычислительным затратам) алгоритма вычисления оценки.

Связь этих задач проявляется в том, что *значение оптимального интервала дискретизации зависит от используемого алгоритма*. Это подчёркивает единство телевизионной системы, в которой три важнейших звена – *объектив, фотоприёмник и устройство обработки не могут просто сопрягаться попарно, а должны оптимизироваться совместно.*

Оптимизация измерения координат звезд требует применения другого критерия помехоустойчивости. Наиболее употребительным на этом этапе является минимум среднеквадратической ошибки (СКО) измерения координаты x_0. В отличие от задачи обнаружения, в которой обработка для принятия решения сводится к сравнению с порогом и не подвержена существенным вариациям, при оценке координаты возможны различные алгоритмы оценки. В случае присутствия единственного источника аддитивного шума оптимальный алгоритм оценки координаты включает операцию согласованной фильтрации и нахождения нуля производной. При этом для исключения грубых ошибок прибегают к временной селекции основного максимума сигнала.

Вариации алгоритмов не только связаны с различиями формы полезного сигнала, но также стимулированы различным

количеством априорной информации о форме полезного сигнала и различными отношениями проектировщиков системы к сложности вычислителя, которую надо затратить на получение некоторой точности измерений (см. раздел 3.1).

При измерении координат точечных объектов, например, звёзд, форма сигнала определяется формой импульсной характеристики объектива, которая и определяет оптимальные характеристики вычислителя координат. Оптимальная по критерию минимума СКО процедура измерения координат сигнала для непрерывного аргумента при аддитивном шуме состоит в свёртке принятого наблюдения $r(x, y)$ с импульсной характеристикой согласованного фильтра $h_o(x, y) = s(-x, -y)$ и нахождении аргумента максимума полученного сигнала свёртки. Потенциальная точность (дисперсия среднеквадратичной ошибки) измерения координаты определяется с учётом энергетического отношения сигнал/шум и среднеквадратичной полосы частот видеоимпульса β^2. Минимальный риск R_E формирования неправильной оценки, или, в данном случае, дисперсия ошибки вычисляется по формуле Вудворда [66]:

$$R_E = M\left[(\hat{x}_o - x_o)^2 + (\hat{y}_o - y_o)^2\right] \geq \varepsilon_W = \frac{1}{\beta^2 \Psi^2}. \quad (2.5)$$

При использовании иных алгоритмов оценки координаты (условно оптимальных или квазиоптимальных) ошибка измерения $\varepsilon > \varepsilon_W$. Так как полезный сигнал наблюдается на априорном поле $0 \leq x \leq X$, $0 \leq y \leq Y$, то количество получаемой в результате измерения информации определяется логарифмом отношения априорного и апостериорного интервалов (множитель $2\pi e$ отражает различие энтропии априорного равномерного и апостериорного гауссовского распределений) [66]:

$$I = \frac{1}{2} \log \frac{X^2 Y^2}{\varepsilon} \leq \frac{1}{2} \log \frac{X^2 Y^2}{(2\pi e)^2 \varepsilon_W}.$$

Увеличение ошибки оценивания координаты ε по сравнению с потенциальной ошибкой ε_W приводит к уменьшению количества полезной информации на величину [4]

$$\Delta I = \frac{1}{2} \log \frac{\varepsilon}{\varepsilon_W}.$$

Важной причиной уменьшения количества информации по сравнению с потенциальным значением является нереализуемость

непрерывного двумерного фильтра, так как непрерывность несовместна с накоплением и считыванием многомерного телевизионного сигнала, вызывающего дискретизацию аргумента. Иными словами, для реализации непрерывного согласованного фильтра потребовалась бы не только бесконечная пропускная способность канала связи, но и бесконечная вычислительная сложность алгоритма оценки координат. При наличии бесконечно большого количества априорной информации и отсутствии ограничения сложности для аддитивного шума потенциальная точность обеспечивается применением согласованного фильтра и устройства вычисления аргумента максимума отклика на его выходе. При дискретном аргументе согласованному фильтру соответствует вычисление координат с интерполяцией наблюдений методом наименьших квадратов. Примерами неполной априорной информации о форме сигнала могут быть различные по полю и во времени аберрации оптической системы при измерении координат звезд астродатчиком или искажения, вызванные турбулентной атмосферой в гартмановском анализаторе волнового фронта. При переходе к субоптимальным процедурам учёт отклонения формы импульсной характеристики фильтра от среднего значения формы сигнала является средством уменьшения сложности алгоритма оценки координат. На практике использует несколько основных алгоритмов, обладающих низкой сложностью благодаря абстрагированию от конкретной формы полезного сигнала.

В плане решения задачи о согласовании размера элемента разложения с полезным сигналом важно подчеркнуть, что оптимальный размер элемента разложения зависит не только от критерия качества, но и от выбранного алгоритма вычисления оценки координаты сигнала [4, 69]. При этом сложность алгоритма может изменяться двояко: параметрически и не параметрически. В первом случае сложность алгоритма при выбранном типе алгоритма, например, состоящем в вычислении аргумента максимума свертки входного сигнала с импульсной характеристикой квазисогласованного фильтра, определяется числом членов импульсной характеристики фильтра. Во втором случае переход от одного типа алгоритма к другому осуществляется дискретно.

Оценка влияния изменения типа алгоритма на достижимую точность иллюстрируется рис. 2.3, на котором приведены расчётные зависимости СКО от соотношения размеров полезного сигнала и интервала дискретизации для гауссовской модели функции рассеяния точки объектива с эффективным радиусом r для разных

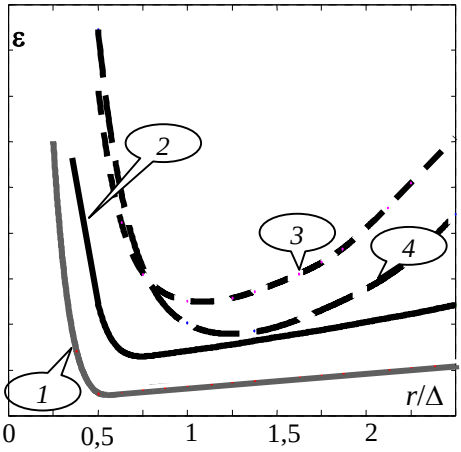


Рис. 2.3. Зависимость относительной ошибки ε измерения координат сигнала гауссовской формы при аддитивном шуме для различных алгоритмов:
- по методу наименьших квадратов *1*,
- по центру тяжести *2*,
- по нулю производной *3*
- по фронтам *4*

алгоритмов оценивания координат. Анализ показывает, что различные алгоритмы обладают не просто различной точностью, но и минимум СКО для них достигается при различном размере элемента разложения.

Более того, можно показать [128], что, например, вычисление координат минимума бимодального сигнала (формируемого некоторыми типами оптических систем из сигнала звезды) для достижения минимума ошибки должно производиться по различным алгоритмам в зависимости от соотношения амплитуд сигналов в группе отсчетов, примыкающих к минимальному. Это означает, что оптимальный алгоритм должен быть адаптивен к форме сигнала.

Поэтому в сложной телевизионно-компьютерной системе все её подсистемы – оптическая подсистема, фотоприёмник, квантователь сигнала и вычислитель оценок параметров сигнала – должны быть адаптивными, и в силу взаимного влияния синтезироваться совместно. Взаимосвязь точности и сложности вычислений формализуется с помощью понятия эпсилон-сложности [126], которая, как и сложность передачи сигналов, т. е. эпсилон-энтропия, зависит от логарифма допустимой ошибки ε:

$$S \sim log \frac{1}{\varepsilon}.$$

Учёт наличия максимальной ошибки, равной дисперсии равномерного распределения в пределах поля зрения и минимальной ошибки ε_W, обусловленной наличием шума в сигнале

изображения, приводит к тому, что сложность алгоритма определяется не самой ошибкой измерения, а разностью между достигаемой и минимальной ошибками.

$$S \approx log \frac{X^2 Y^2}{\left(2\pi e\right)^2 \left(\varepsilon - \varepsilon_W\right)}.$$

Так же, как и эпсилон-энтропия, эпсилон-сложность равна нулю при допустимой ошибке, превышающей априорный интервал существования полезного сигнала, и стремится к бесконечности при приближении ошибки вычислений к потенциальной границе ошибки ε_W. Точную границу сложности для рассматриваемой задачи вычислить трудно, однако можно найти некоторые оценки на основе анализа свойств конкретных алгоритмов исходя из необходимого количества сложений $N_\oplus$ и умножений $N_\otimes$.

Соотношение сложностей операций умножения и сложения при различных аппаратных реализациях различны, но в первом приближении целесообразно допустить оценку сложности операций вычисления суммы по разрядности m чисел, а оценку сложности умножения – по квадрату этой разрядности [127]:

$$S \approx N_\oplus m^2 + 2N_\otimes m.$$

Характер этой зависимости, потеря полезной информации и совокупный информационный риск (равный сумме потери полезной информации и взятой с выбираемым проектировщиком весовым коэффициентом сложности) показаны на рис. 2.4. В этой же плоскости точность–сложность представлены зоны, соответствующие нескольким типовым алгоритмам вычисления оценки координат, характеризующие возможные вариации взаимообмена точности и сложности при изменениях соотношения размера сигнала к размеру элемента разложения и разрядности квантователя. Анализ этих зависимостей показывает, что основные используемые алгоритмы можно разделить на три класса:

- высокоточные алгоритмы, обладающие высокой сложностью и реализуемые с большими задержками на вычисление,
- алгоритмы реального времени с обычно приемлемой точностью, отличающейся от потенциальной в 3…5 раз,
- весьма упрощенные алгоритмы, в которых взаимообмен между точностью и сложностью приводит к существенному отступлению от потенциальной точности.

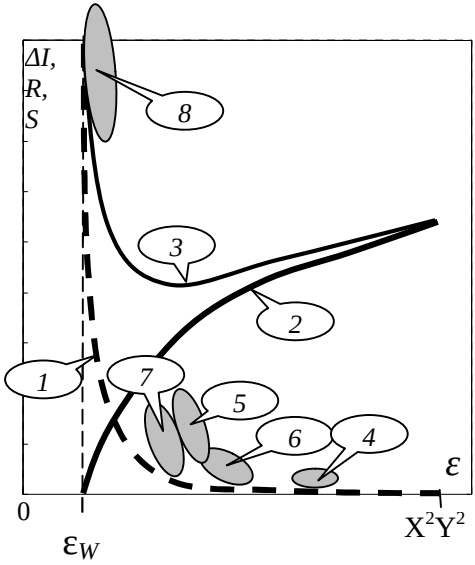


Рис. 2.4. Взаимосвязь сложности и точности измерения координат.

Нижние границы
- сложности алгоритма *1*,
- потери *ΔI* полезной информации *2*,
- информационного риска *R 3*;

Зоны соотношения точности и сложности *S* при различном числе уровней квантования для алгоритмов измерения координат по:
- номеру элемента разложения с максимальным сигналом *4*,
- нулю производной сигнала *5*,
- фронтам сигнала *6*,
- центру тяжести сигнала *7*
- методу наименьших квадратов *8*

Наилучший *компромисс между точностью и сложностью* обеспечивает алгоритм вычисления координат по центру тяжести. При этом раздельная оптимизация алгоритма вычисления и сопряжения фотоприёмника с объективом может привести к проигрышу по достижимой точности в два – три раза.

Различные упрощенные алгоритмы отличаются не только в плане соотношения точность–сложность, но и устойчивостью к отклонению формы сигнала, которая может возникать, в частности, из-за различных по полю и во времени аберраций оптической системы.

Алгоритма оценки координаты сигнала по центру тяжести полезного сигнала (метод центроида) [60, 68– 70] не только обеспечивает наилучший компромисс между точностью и сложностью, но и обладает высокой устойчивостью к недостатку априорной информации о форме сигнала. Поэтому взаимообмен точности и сложности при фиксированном алгоритме и изменении его параметров рассмотрим на примере алгоритма оценки координаты сигнала по методу центроида. Если абстрагироваться от сложности реализации элементов разложения очень малой площади и роста дисперсии шума считывания при увеличении числа элементов в матричном фотоприемнике, можно рассмотреть модель обработки сигнала, включающую реализацию накопления полезного сигнала в два этапа: на фотоприёмнике и в согласованном (квазисогласованном) фильтре.

Такое комбинированное накопление перед операцией измерения с точки зрения достижимой точности при сделанной оговорке даст более точную оценку координат, но ценой существенного увеличения сложности вычислителя.

При этом выигрыш в дисперсии ошибки будет асимптотически стремиться к трехкратному [124] при стремлении сложности алгоритма S (пропорциональной числу членов импульсной характеристики фильтра n) к бесконечности (рис.2.5).

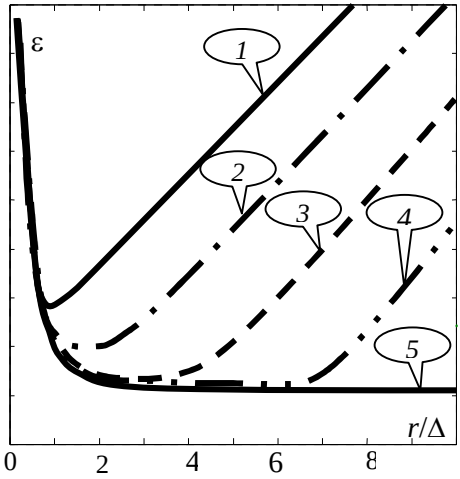


Рис. 2.5. Зависимость относительной ошибки измерения координат сигнала методом центра тяжести при оптимальной дискретизации без фильтрации *1* и при избыточно подробной дискретизации с фильтрацией при числе n членов дискретной двумерной импульсной характеристики:

- n=4 *2,*
- n=16 *3,*
- n=64 *4*
- $n \to \infty$ *5*

Анализ показывает, что эффективность использования сложности системы фотоприёмник–вычислитель (включающий двумерный фильтр) по мере повышения подробности дискретизации оптического сигнала фотоприёмником заметно уменьшается. Этот эффект обуславливает стремление проектировщиков реализовывать телевизионно-компьютерные системы измерения координат объектов так, чтобы по возможности избежать необходимой процедуры промежуточной фильтрации сигнала до вычисления координат.

Взаимовлияние характеристик подсистем телевизионно-компьютерной системы особенно наглядно проявляется в задаче распределения площади $S_к$ кристалла между фотоприёмником и устройством цифровой обработки сигналов в рамках единой видеосистемы на кристалле. Выпускаемые в настоящее время видеосистемы на кристалле (КМОП матрицы) характеризуются распределением площади кристалла $S_к$ так, что управление

системой, системы синхронизации, обработки и канального кодирования видеоинформации в совокупности занимают площадь $S_в$, на порядок меньшую, чем площадь $S_ф$ фотоприёмного массива элементов. Видеосистемы на кристалле, в которых площадь кристалла, занимаемая вычислительной подсистемой, больше, чем площадь фотоприёмника, пока характеризуются весьма скромным числом элементов разложения. Такое положение можно считать временным, так как оно противоречит основным тенденциям развития телевизионно-компьютерных систем, в которых сложность компьютерной обработки изображений постоянно возрастает. Конечно, проблема оптимального распределения площади кристалла между фотоприёмником и процессором в общем случае аналитически трудно разрешима, и подходы к ней можно искать в области методов динамического программирования. Но можно выявить основные тенденции на примере некоторых относительно простых моделей изображений, например, для измерения характеристик волнового фронта с помощью динамического теста Гартмана. В этой задаче обычно минимизируют ошибки, обусловленные дискретизацией волнового фронта на парциальные пучки и дискретизацией совокупного потока фотонов на телевизионные кадры при ограничении плотности потока фотонов [33, 121, 124].

Особенностью динамического теста Гартмана является необходимость измерения координат совокупности световых пятен в реальном времени с большой скоростью: от нескольких десятков до нескольких сотен пар координат при кадровой частоте, существенно превышающей вещательную – от 100 Гц до 1000 Гц. Здесь кроме шумовой ошибки ε_W и ошибки дискретизации $\varepsilon_д$ дополнительно учтём ошибки вычислений, вызванные ограничением вычислительной сложности, определяющей ошибку вычислений $\varepsilon_в$.

Если считать, что число парциальных световых пучков (число пятен гартманограммы) и кадровая частота выбраны оптимально [4, 121], то при известных размерах элементов разложения в фотоприёмнике Δ и элементарной ячейки в вычислителе $\Delta_в$ можно охарактеризовать зависимость количества информации, получаемой в результате измерения координаты одного пятна в одном кадре, от соотношения площадей фотоприемника $S_ф$ и вычислителя $S_в = S_к - S_ф$ формулой:

$$I = \tfrac{1}{2} log \frac{S_ф^2}{\varepsilon_W + \varepsilon \ + \varepsilon} \approx \tfrac{1}{2} log \frac{S_ф^2}{\varepsilon_W + \dfrac{\Delta^4}{S_ф^2} + S^2 \ 2^{-\frac{S_к - S_ф}{\Delta_в^2}}}.$$

Зависимость извлекаемой с помощью видеосистемы на кристалле информации от распределения площади кристалла между фотоприёмником и вычислителем показана на рис. 2.6.

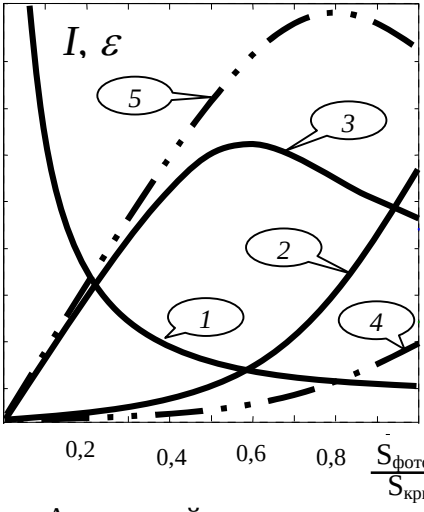


Рис. 2.6. Зависимость от относительной площади фотоприёмника:

- ошибки измерения координат $\varepsilon_д$, обусловленной дискретностью растра фотоприемника *1*,
- ошибки вычислений $\varepsilon_в$ *2* и количества информации *I 3* при равенстве площадей фотоприёмной ячейки и вычислительной ячейки,
- ошибки вычислений $\varepsilon_в$ *4* и количества информации *I 5* при площади вычислительной ячейки, на порядок меньшей, чем площадь накопительной ячейки

Анализ этой зависимости показывает, что оптимальное распределение площадей фотоприёмной и вычислительной подсистем во многом определяется соотношением площадей элемента разложения в фотоприёмнике и элементарной ячейки вычислителя. Так как размер накопительной ячейки фотоприёмника выбирается исходя из требуемой большой накопительной способности, то он оказывается существенно больше, чем размер ячейки вычислителя.

Имеющееся различие в указанных размерах в ходе развития микроэлектроники постоянно увеличивается – по мере освоения всё более высокой точности изготовления видеосистем на кристалле количество транзисторов на единицу площади постоянно возрастает, а количество элементов разложения остается примерно постоянным. Поэтому для имеющегося в настоящее время типовых размеров видеосистем на кристалле и размеров накопительных и вычислительных ячеек можно считать, что в реальном времени

вычисления должны вестись с помощью относительно простого алгоритма, оценивающего центры тяжести сигнальных пятен.

Можно прогнозировать, что дальнейшее развитие видеосистем на кристалле позволит со временем перейти даже в скоростных системах к высокоточным алгоритмам, основанным на методе наименьших квадратов и учитывающим зависимость формы сигнала (оптических аберраций) от координат сигнала в кадре, в настоящее время применяемым лишь в медленных телевизионно-компьютерных системах, например в системах создания каталогов звёздных полей.

Проведённый анализ системы измерения координат объектов показывает методологическое единство проектирования телевизионно-компьютерной системы как неразрывного процесса, не допускающего декомпозицию системы на подсистемы, в ходе которого три важнейших звена – объектив, фотоприёмник и устройство компьютерной обработки не могут просто сопрягаться попарно, а должны оптимизироваться совместно.

Компромиссный с позиций точности и сложности алгоритм центра центроида ниже проанализируем для одномерного дискретного сигнала, когда оценка координаты вычисляется по формуле, следующей из определения центра тяжести [1, 60, 65, 159]:

$$\hat{x} = \sum_i i\Delta U_i \Big/ \sum_i U_i \ . \qquad (2.6)$$

Метод центра тяжести обосновывают трактовкой отклика оптической системы на излучение точечного источника как плотности вероятности его координат. При этом считается, что «процедура получения оценок координат по центру тяжести не зависит ни от вида, ни от масштаба реакции оптической системы на излучение точечного источника» [60]. Для алгоритма центроида (2.6) вычисления по каждой оси ведут по трем отсчётам, т. е. оптимальный по критерию минимума СКО размер элемента разложения считают близким к удвоенному радиусу r в формуле (2.2) [1], [65]. Этот результат, полученный в предположении о преобладании собственных шумов матрицы ПЗС, нуждается в уточнении, так как требует учёта фотонного шума сигнала и различной зависимости шума внешнего фона и шума считывания от размера элемента разложения.

Различие в вычислении отношения сигнал/шум и СКО в основном состоит в том, что при вычислении координат

решающую роль играет шум не отсчёте с максимум сигнала звезды, а в соседних с ним. Это обусловлено тем, что алгоритм (2.8) может быть преобразован к виду, содержащему номер i_m отсчёта с максимальным значением сигнала звезды и дробную часть, вычисляемую через значения сигнала в соседних отсчётах с номерами i_{m-1} и i_{m+1}:

$$\hat{x} = \Delta\left(i_m + \frac{\sum_{j=1}^{\infty} j(U_{i_m - j} - U_{i_m + j})}{\sum_{j=-\infty}^{\infty} U_{i_m + j}} \right).$$

В широко распространенном варианте использования алгоритма при дискретизации сигнала звезды по каждой оси на три элемента разложения такая оценка определяется сигналами в двух соседних с максимальным отсчетах с номерами i_{m-1} и i_{m+1}:

$$\hat{x} = \Delta\left(i_m + \frac{U_{i_m-1} - U_{i_m+1}}{U_{i_m-1} + U_{i_m} + U_{i_m+1}} \right). \qquad (2.7)$$

Ошибка измерения координат при использовании алгоритма центроида имеет несколько компонент, обусловленных шумом. Известно [66], что «шум вызывает два явления: одно – обычное уменьшение точности, другое – неоднозначность. Оценка может быть отброшена от своего правильного положения настолько, что она будет казаться соответствующей совершенно ложному значению координаты». Неоднозначность из-за ложных тревог вне полезного сигнала устраняется с помощью введения порога обнаружения γ. Вместе с тем специфика обработки дискретных отсчётов сигнала при оценке координат в отличие от непрерывного варианта задачи не позволяет полностью абстрагироваться от неоднозначности результатов вычислений. Даже при достаточно высоком пороге обнаружения и при исключении ложных тревог вне сигнала при слишком подробной дискретизации полезного сигнала возникает неоднозначность из-за неправильного назначения отсчета на роль соответствующего истинной координате сигнала. Поэтому полная ошибка измерения координат ε будет содержать как минимум две компоненты: $\varepsilon_н$, определяемую «обычным уменьшением точности», и которая может быть названа нормальной ошибкой, и $\varepsilon_а$, определяемую неоднозначностью внутри самого сигнала, которая может быть названа аномальной ошибкой: $\varepsilon = \varepsilon_н + \varepsilon_а$. Нормальная

ошибка родственна пропуску полезного сигнала при обнаружении. Аномальная ошибка, родственная ложной тревоге при обнаружении сигнала, определяется вероятностью неправильного определения отсчёта, сигнал в котором максимален. Нормальная ошибка в общем случае разложения сигнала звезды на большое число элементов вычисляется по формуле

$$\varepsilon_н = 2\Delta^2 \left(\sum_{j=1}^{\infty} j\sigma^2_{i_m - j} \right) / \left(\sum_{j=-\infty}^{\infty} U_{i_m - j} \right)^2 \qquad (2.8)$$

В частном случае использования алгоритма (2.7) для вычисления нормальной ошибки, воспользовавшись независимостью шумов сигнала, фона и считывания, и тем, что дисперсия суммы независимых случайных величин равна сумме их дисперсий, из формулы получим:

$$\varepsilon_н \approx \Delta^2 \frac{2(\sigma_s^2 + \sigma_b^2 + \sigma_c^2)}{\left(\sum U_i\right)^2} = \Delta^2 \frac{\left\{\alpha\left[\Phi\left(\frac{3\Delta}{2}\right) - \Phi\left(\frac{\Delta}{2}\right)\right] + 2(B\Delta + D)\right\}}{\alpha^2 \Phi^2\left(\frac{3\Delta}{2}\right)}. \qquad (2.9)$$

Анализ формулы (2.8) показывает, что нормальная ошибка при увеличении размера элемента монотонно возрастает.

Аномальная ошибка определяется вероятностями $p(U_{i_m+j} > U_{i_m})$ превышения смесью сигнала и шума в соседних элементах над содержащим истинный максимум сигнала звезды:

$$\varepsilon_а = \sum_{j=-\infty}^{\infty} (j\Delta)^2 p(U_{i_m+j} > U_{i_m}) = 2\sum_{j=1}^{\infty} (j\Delta)^2 p(U_{i_m+j} > U_{i_m}). \qquad (2.10)$$

Можно показать [69], что для гауссовских случайных величин ς и ν, с помощью которых моделируются сигналы соседних элементов разложения в фотоприёмнике, искомая вероятность составляет

$$p(\varsigma > \nu) = \frac{1}{2}\left[1 - \Phi\left(\frac{\bar{x}_\nu - \bar{x}_\varsigma}{\sqrt{\sigma_\nu^2 + \sigma_\varsigma^2}}\right)\right]. \qquad (2.11)$$

С учётом (2.2) можно найти амплитуды сигнала для центрального и соседнего с ним отсчётов. С учётом зависимости шума фона от размера элемента аналогично использованному в формуле (2.4) для отношения сигнал/шум можно найти дисперсии шума в этих элементах разложения. Тогда аномальная ошибка составит:

$$\varepsilon_a \approx \Delta^2\left[1-\Phi\left(\frac{\frac{\alpha}{2}\left\{3\Phi\left(\frac{\Delta}{2}\right)-\Phi\left(\frac{3\Delta}{2}\right)\right\}}{\sqrt{\frac{\alpha}{2}\left[\Phi\left(\frac{\Delta}{2}\right)+\Phi\left(\frac{3\Delta}{2}\right)\right]+2(B\Delta+N)}}\right)\right]. \qquad (2.12)$$

Анализ формулы (2.11) показывает, что аномальная ошибка при увеличении размера элемента в отличие от нормальной ошибки монотонно убывает. Эти две противоположные тенденции и определяют наличие оптимального для измерения координат размера элемента разложения. Расчет полной ошибки измерения координат в общем случае производится с использованием формул (2.8) для нормальной ошибки и (2.10) для аномальной ошибки.

При расчёте ошибок для малых значений аргумента число слагаемых тем больше, чем меньше размер элемента разложения. В пределе, при стремлении размера элемента разложения к нулю, суммы в формулах (2.8) и (2.10) имеют бесконечное число слагаемых. Например, при $\Delta/r = ½$ в формуле суммирование ведется для $j \le 4$, при $\Delta/r = 1$ для $j \le 2$, для больших значений аргумента $j = 1$.

Расчет показывает, что в широком диапазоне изменения соотношения интенсивности сигнала и шумов фона и считывания имеется оптимальное значение размера элемента $\Delta/r_{opt\,\varepsilon} \approx 1{,}5$ (рис. 2.7).

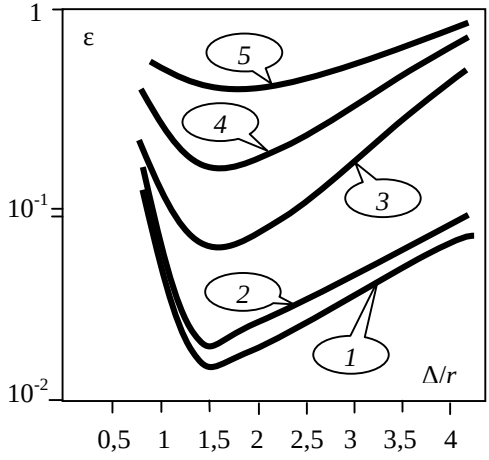


Рис. 2.7. Зависимость среднеквадратической ошибки измерения координат звезды от соотношения размеров элемента разложения и функции рассеяния точки объективом при различных соотношениях шума фона и собственного шума телекамеры
1) α=100, B=5, D=5;
2) α=100, B=5, D=10;
3) α=100, B=50, D=5;
4) α=100, B=100, D=5;
5) α=10, B=1, D=1.

Это означает, что расчёт, учитывающий три источника шума (полезного сигнала, фона и считывания) и влияние двух типов ошибок при использовании алгоритма измерения центра тяжести подтвердил правило дискретизации сигнала звезды на три элемента

по каждому направлению для достижения минимума ошибки измерения координат. Вместе с тем соотношение между радиусом сигнала и размером элемента уточнено. При этом, как и при доминировании шумов телекамеры, точность измерения координат растёт с увеличением отношения сигнал/шум. При соответствующей калибровке системы (напр., [65]) для коррекции неоднородности чувствительности элементов матрицы ошибка измерения координат при типовой ёмкости элемента разложения порядка 10^5 электронов может достигать нескольких сотых долей элемента разложения.

Сопоставление зависимостей отношения сигнал/шум и СКО оценки координат от соотношения размеров элемента разложения и функции рассеяния точки объективом показывает, что их экстремумы достигаются при существенно различных значениях аргумента. Самым важным является то, что интервал дискретизации, оптимальный для оценивания координаты, практически постоянен $\mathbf{\Delta/r_{opt\,\varepsilon} \approx const}$, а оптимальный для обнаружения – зависит от контраста и соотношения шумов фона и камеры $\mathbf{\Delta/r_{opt\,\psi} = var}$.

При этом, чем меньше дисперсия собственного шума телекамеры по сравнению с дисперсией шума фона, тем сильнее выражены указанные экстремумы. Важными количественными соотношениями являются возможные выигрыши в помехоустойчивости при правильном выборе критерия качества. Так, при типичном для астродатчиков соотношении сигнала и шумов значение СКО может быть в 3 раза меньшим, чем при размере элемента разложения, оптимальном по критерию максимума отношения сигнал/шум. Аналогично, значение отношения сигнал/шум может быть в 1,8 раза больше, чем при размере элемента разложения, оптимальном по критерию минимума СКО. Важным конструктивным результатом такого сопоставления является тезис: *изменение критерия помехоустойчивости при переходе от обнаружения к оцениванию координат ведёт к адаптации размера элемента к интенсивности полезного сигнала*. Соотношение размера элемента при измерении координат и обнаружении позволяет строить систему так, чтобы при измерении координат размер элемента составлял 1,5 радиуса функции рассеяния точки, а при обнаружении увеличивался в 2–3 раза [67].

Такое различие оптимальных параметров системы при её различных целях может быть использовано двояко: созданием параллельной системы и реализацией адаптивного алгоритма её работы.

Параллельное считывание с организацией каналов различной чёткости. В системах различного назначения могут реализовываться различные компромиссы между чёткостью и быстродействием. Более того, максимальная помехоустойчивость определяется, в частности, количеством группируемых зарядовых пакетов, число которых зависит от формы и размеров выделяемого объекта. Когда в одном и том же кадре могут быть объекты различных классов, то целесообразно вести обработку изображений по раздельным файлам, сформированным из одного и того же накопленного зарядового изображения при разных количествах группируемых элементов. Согласно методу «пирамиды» [125] возможно формирование многих параллельных каналов считывания с различной чёткостью. Совместное обнаружение–оценивание точечных объектов в телевизионных камерах на матричных фотоприёмниках в режиме реального времени возможно путём последовательного не разрушающего заряды считывания сигналов с полной и сокращённой чёткостью [64, 72]. В этом режиме матричный ПЗС (включая устройство считывания зарядов) исполняет роль общей ветви системы обработки для совместного обнаружения–оценивания, а разветвление на самостоятельные алгоритмы осуществляется с помощью демультиплексора, включённого непосредственно после специальной схемы выборки. Специфика схемы выборки состоит в том, что при суммировании зарядовых пакетов по строке в выходном устройстве с плавающей диффузионной областью формируется видеосигнал, представляющий собой дискретно увеличивающуюся сумму видеосигналов соседних элементов разложения ПЗС. В этом случае для формирования каналов с различным пространственным разрешением и чувствительностью достаточно с помощью устройств выборки–хранения формировать разности между соответствующими значениями видеосигнала, а для формирования сигнала с максимальным отношением сигнал/шум – между последним значением, равным сумме всех видеосигналов и уровнем чёрного. Данный алгоритм обработки видеосигнала [72] эквивалентен схеме двойной коррелированной выборки, что позволяет устранить шум установки потенциала узла детекти-рования заряда. Количество каналов на выходе демультиплексора определяется набором требуемых размеров зон накопления и зависит от соотношения шумов считывания и фотонных шумов фона. Данный метод обладает избыточностью и поэтому полезен при большой априорной неопределённости сигналов.

Адаптивное считывание с переменным размером зоны накопления. Важным конструктивным результатом проведённого параметрического синтеза является обоснование адаптации размера элемента к задаче обнаружения или оценивания и к интенсивности полезного сигнала. Соотношение оптимального размера элемента и радиуса кружка рассеяния объектива при измерении координат и обнаружении позволяет строить систему так, чтобы при измерении координат размер элемента составлял 1,5 радиуса кружка рассеяния, а при обнаружении увеличивался в 2…3 раза.

Согласование зоны накопления с кружком рассеяния объектива может производиться группированием зарядовых пакетов до передачи через выходное устройство, т. е. методом переменной по полю чёткости ([2, 71], см. раздел 2.2), уменьшающим влияние шумов канала считывания зарядового изображения. Переменная чёткость, реализуемая непосредственно в фотоприёмной матрице, принципиально отличается от таковой, организованной путём компьютерной обработки считанного обычным способом изображения. Указанное отличие состоит в том, что шум считывания в обычной системе добавляется к каждому элементу разложения, а в системе с переменной чёткостью он добавляется один раз на группу элементов.

Вместе с тем при создании твердотельных систем реального времени, например, астродатчиков или систем контроля космического пространства, заранее разделить кадр на фрагменты с полной и сокращённой чёткостью, как это описано в [2, 71], удается редко. Адаптация дискретизации сигналов точечных объектов к их интенсивности позволяет достичь потенциальной помехоустойчивости и для обнаружения и для оценивания координат. При интенсивных сигналах звезды дискретизация должна быть подробной (предельной для матрицы), относительный размер элемента $\Delta \approx 1{,}5r$. При малой интенсивности сигнала должно осуществляться увеличение размера зоны накопления. Такая адаптация имеет аналогию в виде АРУ «по точечному объекту» и требует включения в состав телекамеры порогового устройства, которое обычно используется в вычислителе оценок координат точечных объектов. Значение порога определяется через функцию $\Phi^{-1}(p_{лт})$ (обратную интегралу вероятности) от заданной вероятности ложной тревоги $p_{лт}$. Превышение порога сигналом конкретного элемента переводит считывающее устройство в режим сброса после каждого элемента, а при сигналах, меньших порога, считывание производится с группированием зарядовых пакетов

[69, 70]. Такое группирование сопровождается управлением нормировкой сигнала в видеотракте, учитывающей размер зоны накопления. При сложении двух зарядовых пакетов устанавливается коэффициент передачи K = 1/2, а при сложении трёх зарядовых пакетов K = 1/3. Реализация алгоритма работы адаптивной телевизионной системы (рис. 2.8) позволяет при сохранении требований к уровню ложной тревоги повысить чувствительность к слабым источникам. Возможный выигрыш в чувствительности определяется соотношением значения порогов на различных шагах адаптации. С учётом переменного значения нормирующего коэффициента передачи порог на *k*-м шаге определяется СКО шума вне полезного сигнала:

$$\gamma_k = \frac{\Phi^{-1}(p_{лт})}{k\Delta}\sqrt{kB\Delta + D}. \qquad (2.13)$$

В зависимости от соотношения уровня шума фона и шума считывания соотношение порогов, а, следовательно, и выигрыш в чувствительности адаптивной системы по сравнению с неадаптивной, изменяется. При практической реализации описанной адаптивной системы уровень шума считывания измеряется по защищенным от воздействия оптического излучения элементам матрицы ПЗС, а уровень шума фона (так же как и собственно уровень фона) измеряется в оценивающе-вычитающем устройстве системы. В зависимости от уровня фона расчётный выигрыш в чувствительности, учитывающий и изменение порога обнаружения (уровня шума), и изменение уровня слабых сигналов при группировании зарядовых пакетов, составляет для группирования двух зарядовых пакетов от 1,1 до 1,6, а при группировании трёх зарядовых пакетов – от 1 до 1,8 [69].

Экспериментальная проверка метода адаптации производилась на специально разработанной телекамере [121] на охлаждаемом с помощью двухкаскадного термоэлектрического холодильника матричном ПЗС с прогрессивной разверткой и строчным переносом ICX415AL фирмы Sony. Для повышения точности быстрого измерения среднего значения темнового сигнала матриц применен метод сложения непосредственно в матрице зарядовых пакетов закрытых от света элементов считывающего регистра.

Формат матриц ПЗС 1/2″ со светочувствительной поверхностью 4,8×6,4 мм; переключаемое время накопления:

- режим 1: кадровая частота 25 Гц (8 бит/элемент);

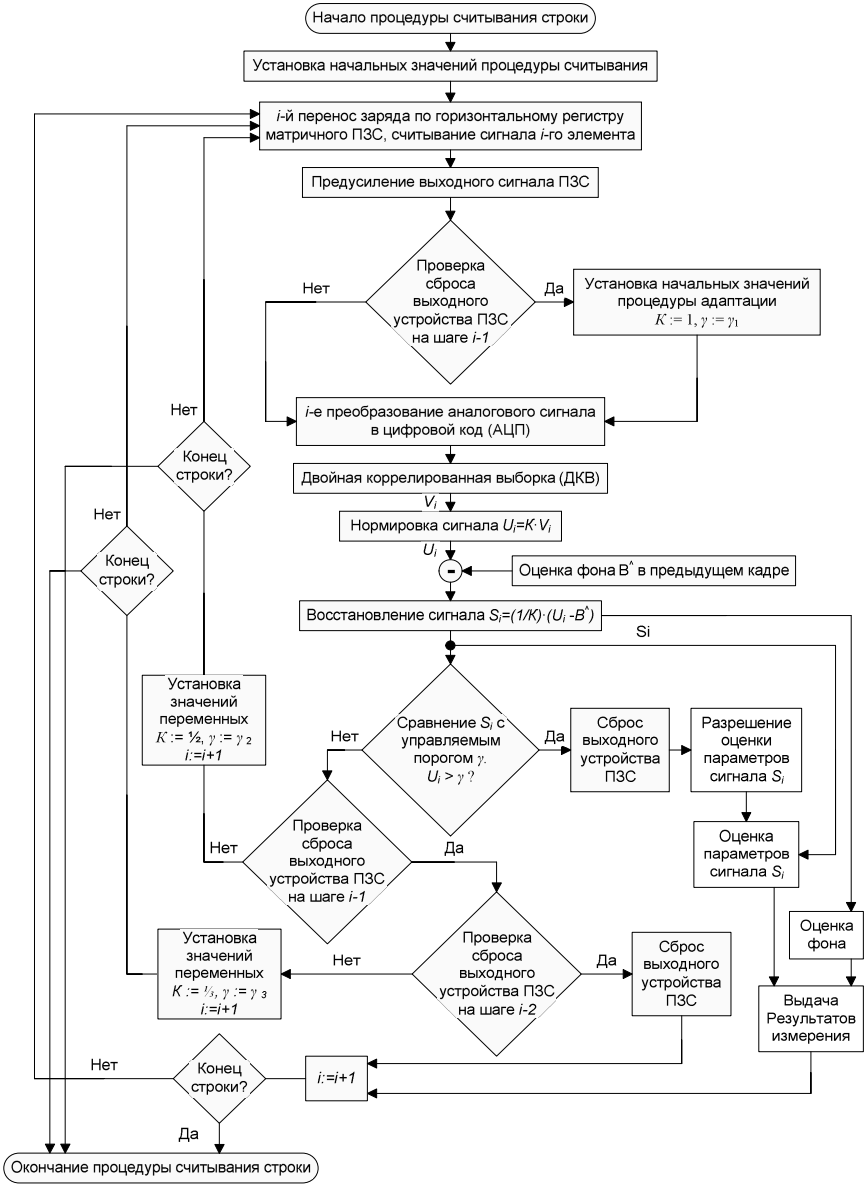
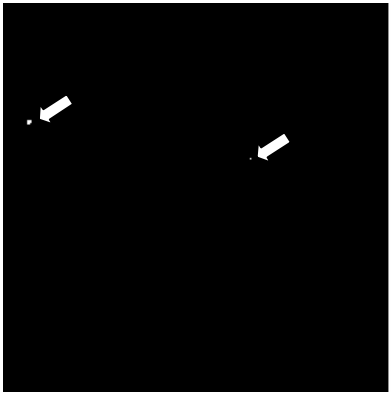
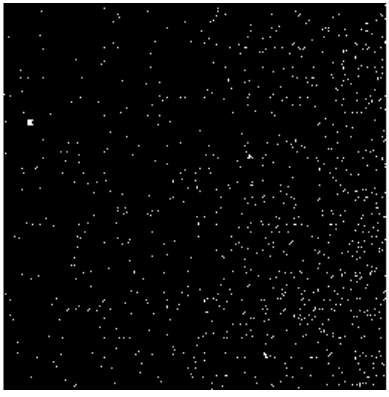


Рис. 2.8. Алгоритм адаптации размера элемента разложения матричного ПЗС к значению накопленного заряда

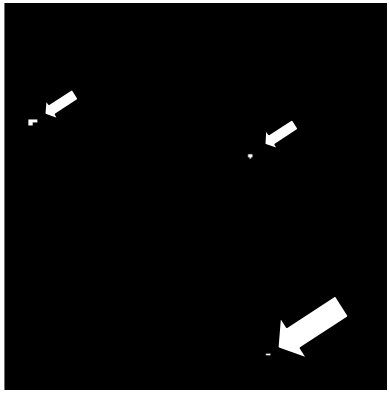
• режим 2: время считывания 0,2 с, 12 бит/элемент; время накопления 1, 3, 8, 16, 32, 64, 128, 180 *с*, устанавливается программно. Результаты эксперимента [70] подтвердили расчёты и показали, что метод адаптивной дискретизации сигналов точечных объектов реализуем и даёт выигрыш в помехоустойчивости (рис. 2.11). Показанные на рис. 2.10 пороги *1* и *2* обеспечивают при различных режимах дискретизации одинаковое значение ложной тревоги, меньшее 1 за кадр.



а



б



в

Рис. 2.9. Иллюстрация эффективности адаптивного метода дискретизации сигналов точечных объектов:

• при подробной дискретизации и заданной вероятности ложной тревоги в элементе разложения $p_{л} \le 10^{-6}$ обнаруживается только два объекта из трёх имеющихся, *а*;

• при той же дискретизации для обнаружения третьего объекта необходимо понизить порог до уровня, приводящего к большой вероятности ложной тревоги, *б*;

• при адаптивной дискретизации два интенсивных сигнала дискретизируются подробно, но дополнительно обнаружен грубо дискретизированный слабый сигнал третьего объекта, *в*

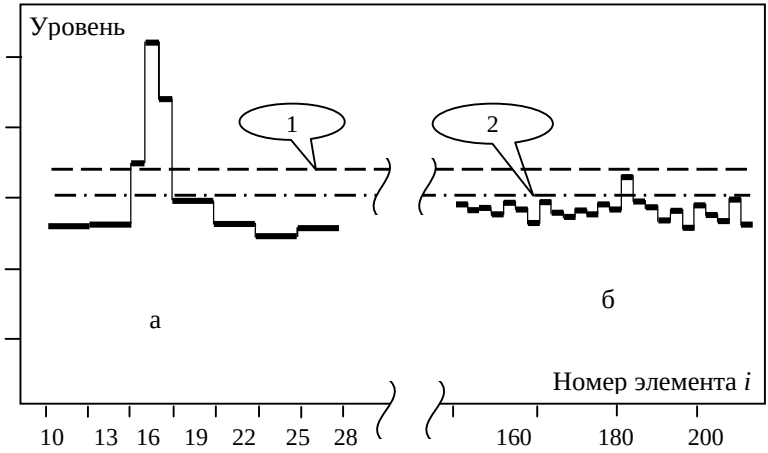


Рис. 2.10. Осциллограммы сигнала при адаптивной дискретизации. Интервал дискретизации для слабых сигналов, не превышающих порог *1*, в три раза больше, чем для интенсивных сигналов *а*; слабый сигнал, не превышающий порог *1* при подробной дискретизации, превышает порог *2* и обнаруживается при адаптивной дискретизации, *б*

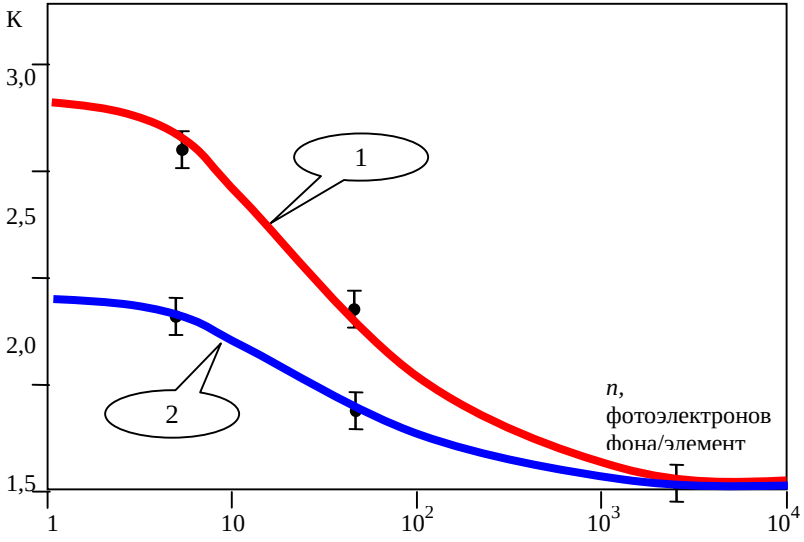


Рис. 2.11. Экспериментальная зависимость выигрыша в чувствительности адаптивной астрономической телекамеры от интенсивности внешнего фона при охлаждении матрицы ПЗС *1* и без охлаждения *2*

Выигрыш в чувствительности, как показывают расчёт и эксперимент при различных условиях наблюдения (при изменениях интенсивности внешнего фона и температуры матрицы ПЗС), уменьшается при увеличении как собственных шумов телекамеры, так и шумов внешнего фона (рис. 2.11).

Полученный в эксперименте максимальный выигрыш в помехоустойчивости в 2,6 раза вместо найденной оценки в 1,8 раза обусловлен примерно в полтора раза большим значением размера кружка рассеяния объектива телекамеры по сравнению с оптимальным, обеспечивающим минимум среднеквадратической ошибки измерения координат. При использовании разработанного метода адаптации для сигналов, уровень которых ниже значения порога *1*, достигаемая точность измерения координат в силу более грубой дискретизации, конечно, хуже (уже не «субпиксельная»), чем для сигналов большой интенсивности.

Вместе с тем достигнутый выигрыш в чувствительности обеспечит наблюдение дополнительного числа слабых объектов в поле зрения при том же времени накопления.

В системах астроориентации требования к точности измерения координат звёзд экспоненциально уменьшаются при увеличении количества наблюдаемых звёзд. Поэтому измерение координат слабых звёзд даже с пониженной точностью (до одного элемента разложения) позволяет увеличить достоверность идентификации наблюдаемого участка звёздного неба.

Значимость, казалось бы, незначительного выигрыша в энергетике сигналов определяется существенным приростом информации при классификации созвездий в системах астроориентации, что иллюстрирует роль информационной парадигмы при проектировании твердотельных телевизионных систем обнаружения–оценивания сигналов.

Измерение координат динамических объектов. В соответствии с имеющейся априорной информацией о характере движения объекта разделяют две модели его движения: по случайной траектории и по траектории, описываемой аналитической функцией с неизвестными параметрами. К первой модели относятся задачи, типичным примером которых является измерение параметров движения летательных аппаратов, включая искусственные спутники Земли.

Вторая модель характерна для анализа отклонений визирной оси системы от наблюдаемого объекта после вычитания постоянной скорости его движения. Математические аспекты обработки таких случайных элементов траекторий стали основой теории оценивания случайных процессов Колмогорова–Винера. На практике это соответствует слежению за маневрирующими целями [62, 63] и измерению координат световых пятен в системах анализа искажений волнового фронта, возникающих в силу турбулентности атмосферы [124].

Первичная, можно сказать, основная дискретно-аналоговая обработка видеоинформации в твердотельных телекамерах производится непосредственно в фотоприёмной матрице. Пространственно-временна́я дискретизация изображения фотоприёмником – главнейший инструмент для достижения потенциального качества передачи телевизионных сигналов и измерения их параметров. В отличие от вещания, в научно-прикладном телевидении вследствие меньшей априорной неопределённости передаваемых сигналов такая оптимизация получает аналитическое решение. Чаще всего встречается задача оптимизации либо пространственной, либо временно́й дискретизации. Но в общем случае разделить эти процессы нельзя. Наиболее наглядно совместная оптимизация пространственно-временно́го сигнала осуществляется в телевизионном анализаторе волнового фронта в силу достаточно простой модели сигналов. При анализе волнового фронта методом Гартмана (или его модификациями, например, методом Гартмана–Шака) с применением телевизионных систем пространственно-временная дискретизация волнового фронта производится в два этапа. Пространственная дискретизация на парциальные световые пучки осуществляется в оптическом тракте, временна́я дискретизация на кадровые временные интервалы – функция телекамеры. Искажения волнового фронта оцениваются по динамике изменения взаимного расположения пятен гартманограммы (рис. 2.12) [4, 124].

Обычно оптимизацию пространственной и временно́й дискретизации осуществляют раздельно: число парциальных пучков N (микролинз в маске Гартмана–Шака) выбирают исходя из соотношения радиуса телескопа R и радиуса корреляции волнового фронта r_F. При этом временну́ю дискретизацию считают уже выбранной так, что она, опираясь на временной интервал корреляции τ, обеспечивает заданное быстродействие системы.

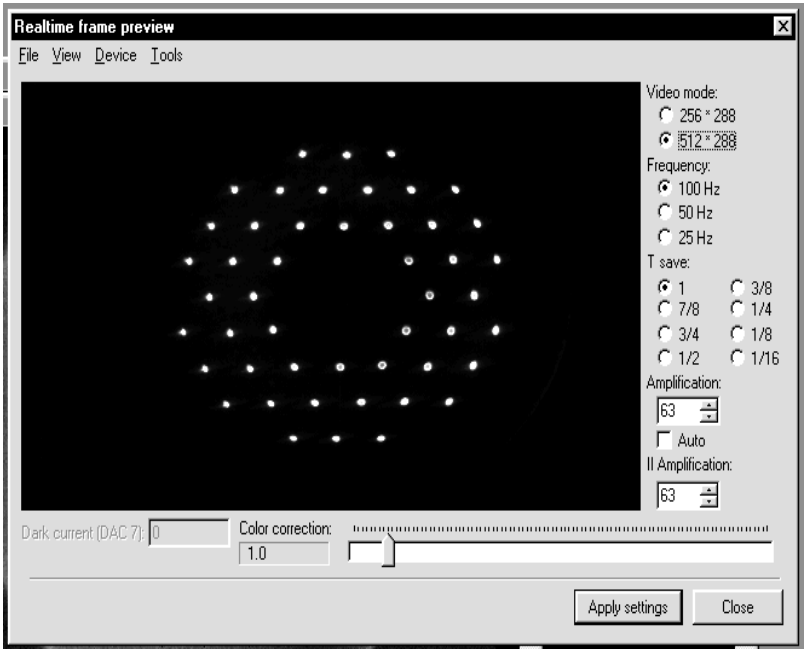


Рис. 2. 12. Изображение гартманограммы и поля управления параметрами твердотельной телекамеры

Гипотеза о возможности рассмотрения волнового фронта как случайного поля с разделимыми временно́й и пространственной корреляционными функциями позволяет осуществить параметрический синтез анализатора волнового фронта новым методом. Он приводит к совместной оптимизации пространственно-временной дискретизации волнового фронта по критерию минимума среднеквадратичной ошибки оценивания пространственно-временно́го случайного поля при ограничении точности фотонными шумами светового сигнала. Такой подход позволяет выделить три компоненты ошибок измерения волнового фронта (координат пятен гартманограммы):

- линейной пространственной ошибки, уменьшающейся пропорционально числу N парциальных пучков,
- линейной временной ошибки, уменьшающейся пропорционально кадровой частоте $f_к = 1/T_к$,

• шумовой ошибки, увеличивающейся пропорционально кадровой частоте и числу N парциальных пучков растра Гартмана (числа микролинз маски Гартмана–Шака).

Результатом синтеза в этом случае являются формулы для оптимального числа N_{opt} парциальных пучков и оптимального времени кадра $T_{ко}$. Коэффициенты пропорциональности в данных формулах различны при необходимости учёта задержки на считывание изображения в телекамере (экстраполяции волнового фронта в системах адаптивной оптики) и в случае исследования волнового фронта без учёта задержек на считывание. В обоих случаях имеют место функциональные зависимости [4]:

$$T_{ко} \sim \tau^{3/4} (R/r_F)^{1/2}; \qquad N_{opt} \sim (R/r_F)\, \tau^{1/2}$$

Важным аспектом этих формул является зависимость оптимальной пространственной дискретизации не только от пространственных параметров волнового фронта, но и от временных, а также зависимость оптимальной кадровой частоты не только от временных параметров волнового фронта, но и от пространственных.

Одна из основных характеристик пространствено-временно́й дискретизации случайного поля – отношение значения *смаза* m_v изображения пятна гартманограммы к его эффективному радиусу r. На этапе становления телевидения считалось, что пространственно-временна́я дискретизация телевизионного изображения должна приводить к значению смаза порядка элемента разрешения. Встречались утверждения о равенстве смаза размеру точечного объекта, половине элемента разрешения, и т. п. Но все эти оценки *допустимого* смаза опирались на концепцию *деградации* разрешающей способности, а не на принцип оптимальности. Приведённые соотношения позволяют получить иной вывод [4]: при наблюдении случайно движущегося сигнального пятна размером r *оптимальная* по критерию минимума среднеквадратической ошибки измерения координат дискретизация даёт существенно меньше значение смаза. Оно обратно пропорционально отношению сигнал/шум Ψ в сигнальном пятне: $m_v = \frac{r}{\Psi}$. Это означает, что при *оптимальной* дискретизации смаз случайно движущегося пятна настолько мал, что амплитуда межкадровой разности равна среднеквадратическому уровню шума, а движение пятна по фотоприёмнику характеризуется *слитностью траектории*. Этот важный результат, состоящий в обосновании

оптимальной пространственно-временно́й дискретизации волнового фронта, может быть распространён и на другие классы подвижных изображений и систем, в частности на селекцию, и траекторные измерения и сопровождение движущихся целей.

Измерение координат специально формируемых сигналов. В современном приборостроении существует множество задач, которые решаются телевизионными методами, и сфера применения телевидения постоянно расширяется. Многие приборы строятся для измерения неэнергетических параметров оптических сигналов, таких как координаты звезд или специально формируемых реперов (меток). Такие приборы разделяются на два класса, отличающихся способом формирования сигналов: либо на основе естественных объектов, например, звёзд на небе, либо рукотворных, как светодиод в оптико-электронном гравиметре. Ниже рассматривается измерение координат специально формируемой метки как типичная задача измерительных систем прикладного телевидения на примере оптико-электронного преобразователя гравиметра – прибора измерения силы тяжести, используемого в специальных навигационных комплексах (рис. 2.13). В этих приборах координата светового сигнала (метки) изменяется под воздействием ускорений, обусловленных движением прибора в поле тяготения Земли. То, что прикладное телевидение только став твердотельным по-настоящему внедрилось в практику, видно из такого примера. Оптико-электронные методы применялись в задаче измерения силы тяжести – гравиметрии – давно, но долго использовали одноэлементные фотоприёмники с механической развёрткой изображения специально формируемой световой метки, координаты которой определялись текущим значением силы тяжести. Никто не пытался применять электронно-лучевые трубки в оптико-электронных преобразователях гравиметров. Только тогда, когда прикладное телевидение готово было одновременно обеспечить и очень высокую надёжность, безоператорную работу в течение многих лет, компактность и высокую точность измерения параметров сигналов, оно стало востребованным. Поэтому оптико-электронная гравиметрия скачком перешла от механического телевидения к твердотельному, минуя фазу электронно-лучевого.

Потенциальная точность измерения координат как естественных, так и специально формируемых сигналов, наблюдаемых на фоне аддитивного шума, определятся по формуле Вудворда и

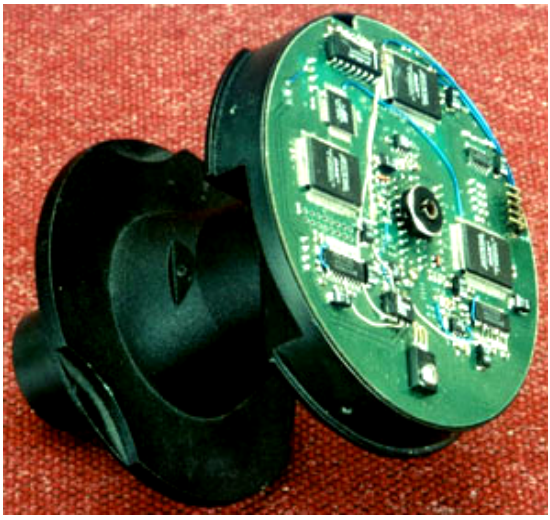


Рис. 2.13. Оптико-электронный преобразователь гравиметрического датчика на двух линейных ПЗС фирмы Sony разработки ЦНИИ «Электроприбор»

ограничивается значениями фотонного шума и энергией сигнала. Различие в приборах указанных классов состоит в том, что для первого класса энергия и форма сигнала не управляются проектировщиком, а в приборах второго класса открываются возможности для увеличения точности, так как проектировщик волен выбирать форму сигнала, задаваясь длиной волны излучения, пространственным распределением яркости в метке и зависимостью яркости от времени.

Синтез формы оптического сигнала в приборах второго класса для увеличения точности измерения его координат учитывает специфику многоэлементных твердотельных фотоприёмников (ограничения размеров фотоприёмника, размеров его элемента и динамического диапазона), и использует следующие методы:

- проектирование прибора ведётся с позиций системного подхода, который диктует *совместное рассмотрение оптической системы, фотоприёмника и вычислителя как единого целого*. Совместная вариация параметров всех подсистем позволяет извлекать максимум информации из каждого сеанса наблюдения: например, в гравиметре это достигается вариацией свойств упругой системы фиксации пробной массы, фокусного расстояния коллимирующего объектива и длиной строки фотоприёмника;

- в качестве чувствительного элемента системы используются линейные, а не матричные фотоприёмники, так как в линейных приборах достигается *максимальная длина строки*, которая играет роль априорного интервала неизвестной координаты световой метки. При этом выбирается специальный тип линейного фотоприёмника, имеющего *малый размер вдоль строки* (согласованный с размером функции рассеяния точки объективом) и *большой размер элемента в поперечном направлении*, для увеличения отношения сигнала метки к шуму. Такие линейные приборы применяются не для формирования изображений, а там, где полезный сигнал, как и в гравиметре, является функцией одной координаты – например, в спектроскопии;
- численное значение отношения размеров элемента фотоприёмника вдоль строки и функции рассеяния точки объективом находится для конкретных, обычно квазиоптимальных *быстрых алгоритмов* вычисления оценок координат, чаще всего – по центру тяжести метки.
- источник излучения выбирается согласованным с фотоприёмником по длине волны излучения так, чтобы обеспечить минимальную ошибку измерений при ограничениях на спектральную характеристику фотоприёмника и излучаемую мощность.

Благодаря применению этих методов был создан гравиметр, позволяющий оценить координату прибора в вертикальной плоскости с точностью в несколько сантиметров [128]. Однако имеется практическая потребность повышения этой точности. Ниже рассматривается наиболее перспективный метод повышения точности измерения координат оптической метки – переход от одноэлементной метки к многоэлементной.

Следствием ограниченного размера элемента фотоприёмника является предел уменьшения размеров функции рассеяния точки, что ограничивает среднеквадратическую полосу частот в формуле Вудворта. Учёт ограничения размера элемента фотоприёмника ставит предел для уменьшения размеров функции рассеяния точки, что даёт естественную границу для среднеквадратической полосы частот сигнала β^2 в формуле Вудворта. Ограничение на динамический диапазон сигналов (связанный с ёмкостью

потенциальной ямы и площадью элемента; удельная ёмкость составляет примерно 1000 электронов на мкм^2) также даёт предел увеличению отношения сигнал/шум. В этой ситуации задачу синтеза формы сигнала необходимо расширить, включив в допустимое множество сигналы сложной формы, широко применяемые в связи и в радиолокации (шумоподобные сигналы, коды Баркера, сигналы с линейно изменяющейся внутри импульса частотой, ЛЧМ, см. раздел 4.5). Такой подход, по-прежнему опирающийся на формулу Вудворта (отношение сигнал/шум Ψ^2 в ней – энергетическое, введенное Нортом, и достигаемое на выходе согласованного фильтра), приводит в рамках указанных ограничений к использованию многоштриховой метки с одинаковыми амплитудами, число штрихов в которой определяется по максимуму получаемой информации.

Оптимальное соотношение шага дискретизации изображения метки фотоприёмником и периода метки находится, так же как и при измерении координат точечного объекта, с учетом нормальной и аномальной ошибок. Аналитическое решение указанной задачи синтеза даёт два результата в виде длины многоштриховой метки, в шесть раз меньшей длины строки (априорного интервала неопределенности положения метки, рис. 2.14) и периода метки, примерно в 6 раз большем интервала дискретизации, рис. 2.15.

Такая метка с числом n штрихов потенциально должна дать выигрыш в точности измерения координат в $\sqrt{n}$ раз по сравнению с использованием одного штриха и формулу Вудворда для дисперсии ε ошибки измерения координат можно представить в виде $\varepsilon=1/(n\,\Psi^2\,\beta^2)$.

Специфика приложения этого результата к практике состоит в отказе от использования фильтра, согласованного со сложным сигналом многоштриховой метки, и, также как и при одноштриховой метке, использовании быстрых простых алгоритмов. Важным результатом синтеза сигнала сложной формы в варианте квазиоптимальных алгоритмов является выигрыш в точности, превышающий $\sqrt{n}$ раз, и достигающий $\sqrt{(2n-1)}$ раз в силу возможности измерения координат не только максимумов, но и

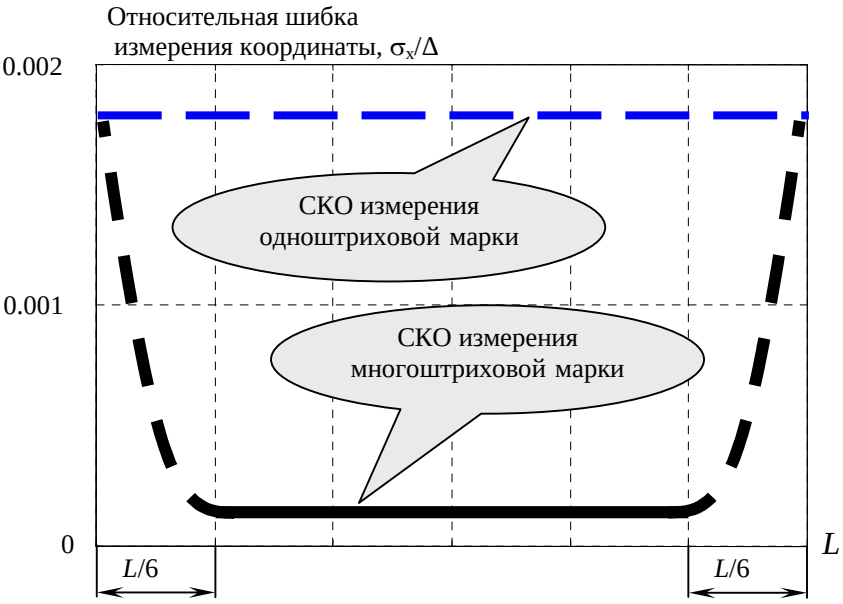


Рис. 2.14. Зависимость ошибки измерения координат многоштриховой метки от ее положения на линейном фотоприёмнике, обусловленная уменьшением числа штрихов на краях поля зрения

минимумов сигнала («провалов» между импульсами). Такое увеличение точности не является чудом, а лишь почти восстанавливает потерю точности, свойственную простым быстрым алгоритмам.

Дополнительный выигрыш в точности при измерении координат многоштриховой оптической метки возникает в силу большей точности измерения координат минимумов, чем координат максимумов из-за меньшего уровня фотонного шума в используемых отсчетах сигнала, так как дисперсия фотонного шума равна среднему значению числа фотоэлектронов в каждом отсчёте.

Совокупность рассмотренных методов обеспечивает создание прецизионных приборов с относительной погрешностью единичного измерения координат световой метки меньшей, чем в системах с одноштриховой меткой, на два порядка.

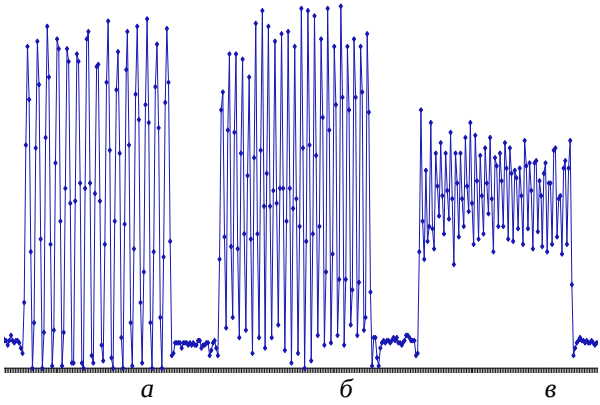


Рис. 2.15. Осциллограммы многоштриховой метки при оптимальном соотношении периода метки и шага дискретизации *а*, при несколько меньшем периоде, когда количество информации о координате метки уменьшается из-за корреляции ошибок измерения координат максимумов и минимумов *б*, и при существенно меньшем периоде метки, когда количество информации о координате метки уменьшается из-за уменьшения глубины модуляции сигнала метки *в*

В результате понятие субпиксельной точности измерения координат (десятые и сотые доли элемента), получившее широкое распространение, в первую очередь в приложении к астродатчикам, приобретает новое звучание. Твердотельные телевизионные системы при необходимых калибровках (учёте разброса чувствительности элементов) позволяет выйти на такой высокий уровень точности, что её ограничивает какой угодно другой компонент, только не телевизионно-компьютерная система.

2.2. Переменная чёткость в твердотельных телекамерах

Разработчики прикладных телевизионных систем при решении многих вопросов сверяют свои решения со свойствами зрительного анализатора человека. Так как чёткость зрения убывает к периферии и максимальна только в небольшой центральной зоне (называемой *fovea*), то были предприняты многочисленные попытки создания прикладных телевизионных систем с переменной по полю чёткостью. В период электронного телевидения такой режим удавалось практически реализовать только при группировании сигналов соседних элементов разложения во внешнем по

отношению к телекамере устройстве цифровой обработки изображений.

В твердотельном телевидении эта идея не одинаковой по всему полю зрения чёткости стала реализуемой, а её практическая потребность обосновывалась наличием априорной информации о неоднородной смысловой структуре изображения, желанием повышения помехоустойчивости и быстродействия телевизионной системы. В частности, чёткость может быть сокращена во всех участках кадра, кроме «зон особого внимания». Наличие таких зон характерно для телевизионных охранных систем, в которых пользователь настраивает систему на максимальную чувствительность (в том числе к движению объектов) в некотором множестве зон, но не во всём кадре.

Переменная по полю чёткость актуальна не только в охранных системах, но и в системах наведения на движущиеся удалённые объекты, используемых в космонавтике и оборонных системах. Для обеспечения малого текущего промаха при наведении на подвижный объект необходимо быстро отслеживать малые флуктуации его траектории. Однако при значительных ускорениях наблюдаемых объектов текущий промах характеризуется двумя значениями: средней ошибкой в режиме сопровождения движущегося объекта и вероятностью срыва слежения. В идеале обе эти характеристики должны быть как можно меньше. Малое среднее значение промаха позволяет достичь цели системы при меньших энергетических затратах. Малая вероятность срыва сопровождения позволяет повысить вероятность выполнения системой своей цели. Для уменьшения среднего значения промаха требуются высокая угловая разрешающая способность и высокая кадровая частота, а для уменьшения вероятности срыва сопровождения – высокая кадровая частота и большое поле зрения. Одновременно достичь в системах наведения высокую угловую разрешающую способность, большое поле зрение и высокую кадровую частоту затруднительно, а, главное, и не нужно. Кроме того, следует считаться с тем, что чем больше считывается элементов с матричного фотоприёмника, тем шире полоса частот и выше уровень шума, увеличивающий и средний промах, и вероятность срыва слежения.

Решение проблемы наведения визирной оси телевизионной системы на подвижный объект достигается использованием концепции адаптации системы к этапам наведения, которое

разделяется на грубое наведение (в широком поле зрения) и точное наведение (в узком поле зрения). Ключевыми пунктами в такой адаптации является автоматическое изменение структуры считываемого сигнала в зависимости от величины текущего промаха. При малом промахе сигнал должен формироваться таким образом, чтобы минимизировать шумовую ошибку измерения координат объекта. При этом можно пожертвовать качеством изображения вне сопровождаемого объекта. При большом промахе, возникающем при манёвре объекта, нужно максимизировать отношение сигнала объекта к шуму, а разрешающей способностью, то есть точностью измерения его координат, можно пренебречь. В результате адаптивная телевизионная система наведения может обладать наилучшими показателями качества и по точности измерения координат объекта в режиме сопровождения, и по вероятности срыва сопровождения при его манёвре. Концепция адаптации параметров телевизионной системы к этапам наведения реализуется в матричных фотоприёмниках с зарядовой связью и с координатной адресацией с помощью различных алгоритмов, так как принципиальное отличие в способах развёртки изображения в них ведёт к различию в возможных методах обработки сигнала непосредственно в матрице.

Сложность достижения малого значения текущего промаха при использовании матричных ПЗС состоит в том, что естественная задержка на накопление фотонов (равная в среднем половине времени кадра $T_К$) усугубляется дополнительной задержкой на считывание накопленного зарядового изображения объекта.

В результате в матрице ПЗС суммарная задержка (при равенстве времени накопления и времени считывания) зависит от координаты элемента разложения в кадре и изменяется от $0{,}5T_К$ для самого близкого к считывающему узлу элемента до $1{,}5T_К$ для самого удалённого от считывающего узла элемента. Из-за этой задержки в системах наведения с телекамерами на матричных ПЗС текущий промах оказывается в среднем в два раза большим, чем в телекамерах на КМОП [83].

Реализация переменной чёткости в телекамерах на КМОП по сравнению с телекамерами на ПЗС имеет две важные особенности. Во-первых, в КМОП однокристальную телекамеру заложен режим управления скоростью считывания в выбранном фрагменте, благодаря которому в нём может быть существенно повышена кадровая

частота по сравнению с остальным полем зрения. Во-вторых, в КМОП-матрицах возможна реализация всевозможных видов развёртки, открывающих путь существенного упрощения операции увеличения площади зоны накопления, необходимой для обмена разрешающей способности системы на чувствительность.

В частности, для дополнительной пространственной обработки, например, усреднения нескольких элементов изображения, при традиционной прогрессивной развёртке требуются дополнительные элементы аналоговой или цифровой памяти. Так называемая развёртка Z-сканирования (рис 2.16, [31]) позволяет проводить пространственное усреднение групп элементов 2×2, 4×4, 8×8 без применения дополнительных запоминающих устройств.

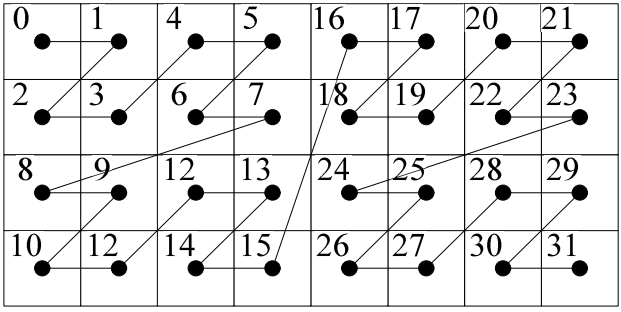
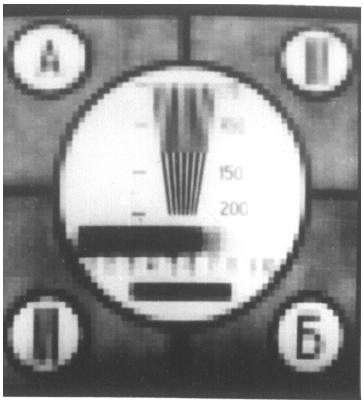


Рис. 2.16 Траектория Z-развёртки

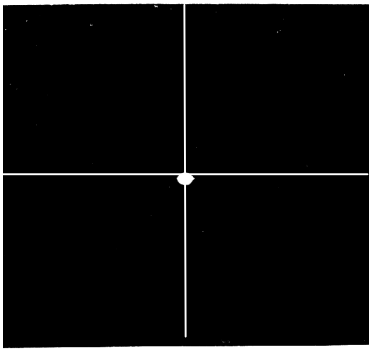
Z-развёртка в КМОП-телекамерах делает достаточно просто реализуемые *множественные выходы* телекамеры, отличающиеся различной (от «страницы к странице») чёткостью, потребной при анализе сложных сюжетов [72].

Выделение доминантной зоны высокой чёткости в изображении, формируемом твердотельной телекамерой, может быть дополнено уменьшением скорости считывания выделенного фрагмента для уменьшения шума считывания [2].

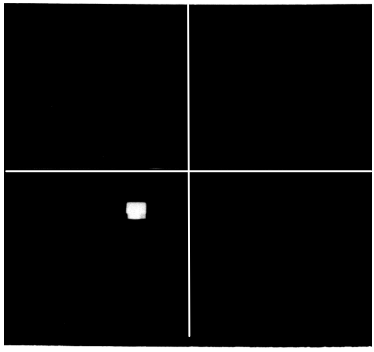
Первая в мире экспериментальная аппаратура с переменной по полю чёткостью, реализованной непосредственно в матрице ПЗС, была создана во ВНИИ телевидения [71, 2] и использовалась в ходе исследований по созданию космических телевизионных систем наведения (рис. 2. 17). В ней в соответствии с различием оптималь-



а



б



в

Рис. 2.17. Иллюстрация одного из методов адаптации к этапам наведения – метода переменной по полю чёткости. Изображение испытательной таблицы *а*, полученное при полной чёткости в центре поля зрения и сокращённой в 4 раза чёткости по вертикали и горизонтали; изображение удалённой цели в центре поля зрения *б* и вне центральной зоны *в*

ных зон накопления при обнаружении и измерении координат (см. 2.1) в периферийной зоне кадра чёткость сокращалась в 4 раза.

За рубежом этот результат был повторен позже [81], и получил там официальное наименование *foveal vision systems* т. е. систем со структурой сетчатки глаза. К сожалению, российские специалисты не всегда знают об отечественном приоритете в данной области твердотельного телевидения [82].

2.3. Вычисление проекций изображения

Проекции многомерных функций на какую либо ось – это результат их интегрирования вдоль другой ортогональной ей оси. Восстановление изображений по его проекциям – важная задача, впервые математически решенная Радоном, и применяемая

в медицине, геологии и неразрушающем контроле. Известно, что для абсолютно точного восстановления многомерной функции требуется бесконечно много её проекций. В практических системах в силу принципа достаточности точности, опирающегося на понятие шумовой информации, ограничиваются конечным числом проекций, число которых в зависимости от отношения сигнал/шум и методов вычислений имеет порядок десятков и сотен [160]. Вместе с тем, в практике прикладного телевидения имеется (в смысле требуемого числа проекций) вырожденная задача, когда не требуется восстановление самого изображения, а требуется лишь быстрая оценка его параметра. Типичным представителем такой задачи является оценка смещения изображения.

Многочисленные эксперименты по совмещению изображений принимали структуру телевизионной системы заданной и не выходили из концепции *раздельной* работы телекамеры и вычислителя измерения смещения изображений. Пересмотр этого ранее незыблемого принципа стал возможен в результате твердотельной революции в телевидении, когда проекции изображения стало возможным формировать не во внешнем компьютере, а непосредственно в фотоприёмной матрице.

Впервые в мире формирование проекций изображений непосредственно в матрице ПЗС осуществлено во ВНИИТе [2, 67]. Эта операция предназначалась для измерения координат специального репера, материализующего визирную ось системы и импульсно проецируемого на матрицу в перерывах между накоплением сигнала наблюдаемого сюжета. За счёт быстрого считывания проекций достигалось получение координат визирной оси за время, пренебрежимо малое по сравнению с временем кадра. Для этого сразу после краткой экспозиции световых пятен фазный электрод считывающего регистра отключался от генератора фазных напряжений и подключался к усилителю, на выходе которого за время одной строки формировался сигнал вертикальной проекции. После переноса всего массива элементов в нижний регистр производилось обратное переключение и считывание из регистра сигнала горизонтальной проекции (рис. 2.18).

Методология видеосистем на кристалле не просто использует наработки первого этапа твердотельной революции, полученные с использованием матриц ПЗС. Глубинное срастание фотоприёмника с вычислителем заставляет в системах реального времени пересмотреть взгляд на свойства, функции и параметры и фотоприёмника, и вычислителя. Так, в задаче совмещения изображений, актуальной

при наблюдениях с дистанционно пилотируемых летательных аппаратов, открывается путь к теснейшему переплетению фотоприёмной матрицы и вычислителя. Фотоприемник в видеосистеме на кристалле начинает выполнять ранее не свойственные ему функции вычислений проекций, а вычислитель переходит от оценок смещения самого изображения к оценкам смещения его проекций. Важно, что совместная работа фотоприёмной матрицы и компьютера в ходе вычисления межкадрового смещения изображений может дать дополнительный, «незапланированный» выигрыш в быстродействии формирования требуемой оценки смещения.

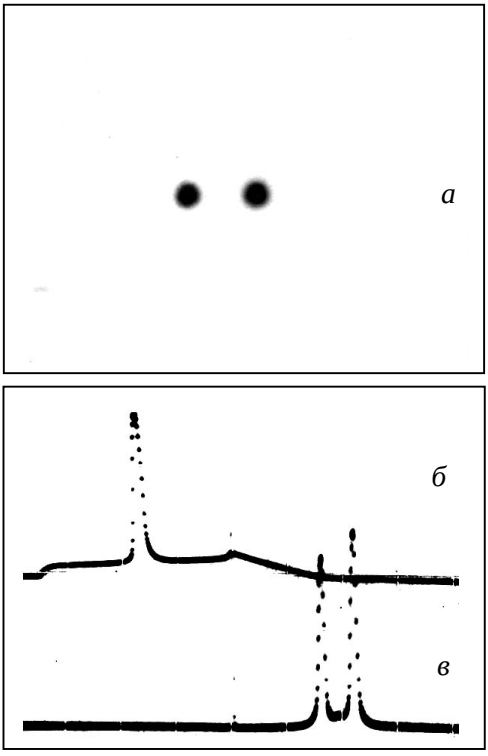


Рис. 2.18. Формирование проекций изображения светового репера в матрице ПЗС.
Изображение, формируемое матрицей на основном выходе в стандартном режиме *а*,
Сигнал проекции изображения на вертикальную ось, считываемый с фазного электрода считывающего регистра *б*,
Сигнал проекции на горизонтальную ось, считываемый через стандартный выходной каскад во время, предшествующее накоплению сигнала полного изображения *в*

Точность измерения смещения по полному изображению выше, чем по его проекциям. Вместе с тем проекции изображения однозначно связаны с ним в силу теоремы о проекционном срезе [160], согласно которой спектр Фурье проекции, полученной под углом θ, является сечением двумерного преобразования Фурье исходного изображения вдоль линии, проходящей через начало координат

плоскости частот (ω_1, ω_2) и составляющей угол θ с осью ω_1. В результате для большинства реальных изображений достигается субпиксельная точность оценки смещения по их проекциям.

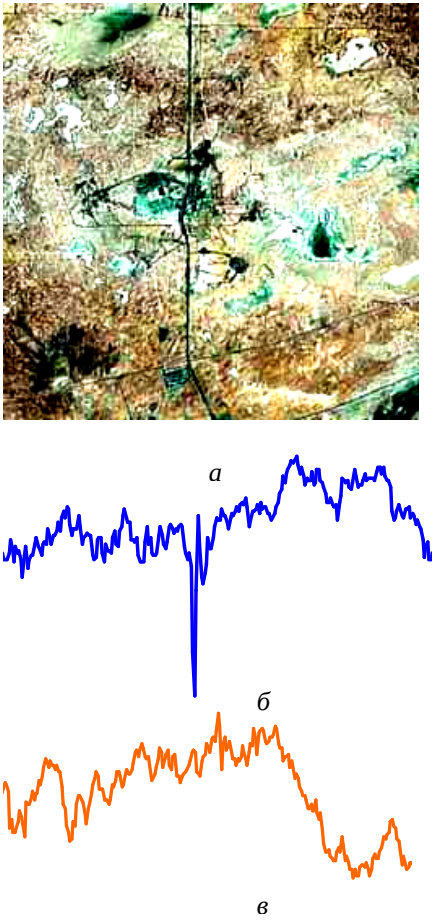


Рис. 2.19. Изображение местности *а* и сигналы его проекций на горизонтальную *б* и вертикальную *в* оси. Анализ сигналов проекций, так же как и анализ полного изображения, показывает возможность дополнительного сокращения вычислительных затрат в тех случаях, когда сигнал содержит локальные зоны с большим значением производной (как на сигнале *б* проекции изображения местности на горизонтальную ось)

Достигаемая ошибка может быть на два порядка меньше размера элемента благодаря тому, что накопление сигнала осуществляется не раздельно для каждого элемента, а по всей строке или по всему столбцу. Поэтому отношение сигнал/шум в сигналах проекций имеет порядок нескольких сотен при практически том же интервале корреляции, что и исходное изображение (рис. 2.19).

Особенностью структурной схемы, реализующей распределение функций вычисления смещения между вычислителем и фотоприёмником [61], становится использование демультиплексора, разделяющего видеосигнал и сигналы проекций, и замена запоминающего устройства емкостью в один кадр запоминающим устройством емкостью в две строки (для запоминания двух проекций сигнала, рис. 2.20*а*). Особенностью функционирования системы становится использование двух интервалов времени накопления сигнала и двух интервалов его считывания различными способами за время формирования одного кадра изображения (рис. 2.20*б*).

Требуемое время накопления сигнала проекций (на рис. 2.20*б* «накопление 1») на два порядка меньше времени накопления сигнала изображения (на рис. 2.20*б* «накопление 2») по двум причинам: достаточное отношение сигнал/шум может быть достигнуто благодаря интегрированию зарядов с совокупности элементов, равной числу строк или числу элементов в строке; кроме того, при неоправданном увеличении времени накопления при формировании проекций может не хватить накопительной способности считывающего регистра ПЗС. При не слишком малой освещенности, когда время накопления изображения может быть существенно меньше времени считывания кадра, моменты времени накопления сигнала для формирования проекций и для накопления сигнала изображения почти совпадают.

Анализ временных диаграмм показывает, что оптическое изображение воспринимается матричным фотоприёмником дважды с небольшим зазором на время считывания проекций (время двух строк). Благодаря этому важной особенностью реализации измерения смещения методом проекций, получаемых непосред-ственно в матрице ПЗС, в отличие от традиционной обработки самого изображения, является получение оценок смещения кадра не просто без задержек, но и даже намного раньше, чем сам этот кадр будет считан.

Сочетание новейших достижений в технологии, воплощаемых в видеосистемах на кристалле, с нетрадиционными методами алгоритмической обработки, охватывающими и фотоприемник, и вычислительную подсистему, и позволят решить проблему стабилизации изображений при наблюдении с динамичных объектов.

В свете построения телевизионных систем с применением нового класса приборов – видеосистем на кристалле, сочетающих в себе и фотоприёмную матрицу, и устройство цифровой обработки сигналов, важным становится ещё один аспект.

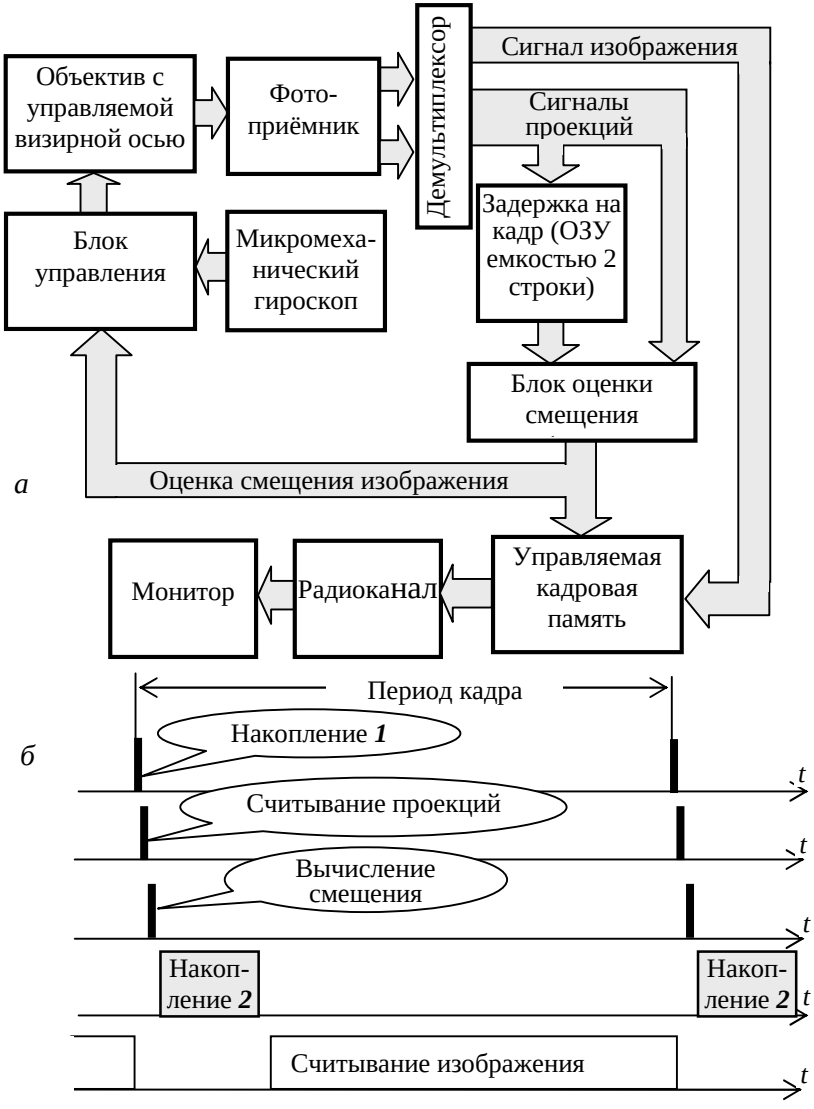


Рис. 2.20. Структурная схема *а* и временные диаграммы *б* системы, измеряющей смещение изображений до считывания самого изображения для его стабилизации при наблюдениях с малогабаритных летательных аппаратов

Так же, как и в задаче измерения координат пятен гартманограммы (рис. 2.6), в любой видеосистеме на кристалле важно корректно распределить полную площадь кристалла между фотоприёмником и вычислителем. Поэтому возложение рутинных функций вычисления проекций изображения на фотоприёмную матрицу и освобождение от неё вычислителя сокращает потребную площадь вычислителя, увеличивает и площадь фотоприёмника и количество информации об изображении и его проекциях.

В ряде приложений, например при классификации случайных полей (текстур) решения принимаются не исходя из их смыслового образа, а по превышению порога некоторых статистик. В случае, когда в качестве таких статистик выступает асимметрия автокорреляционной функции или спектра Фурье, видеосистема на кристалле вновь позволяет эффективно распределить функции вычисления между фотопрёмным массивом и собственно вычислителем. Особенностью такого построения видеоинформационной системы является нетрадиционный способ формирования проекций. Так как в этой задаче (в отличие от задачи оценки смещения) уже недостаточно двух проекций, то для увеличения их числа до требуемого значения могут быть сформированы *виртуальные* оси. В простейшем случае формирования проекций на оси, наклонённые к границам кадра под углом $\theta = 45^{o}$, текущий сигнал проекции перед суммированием со следующей строкой смещается на один элемент разложения. Введение других правил смещения текущего значения проекции позволяет создать большой набор виртуальных осей, потребный для классификации изображений по их спектрам.

С точки зрения компьютерной системотехники методы измерения проекций (как на действительные, так и на виртуальные оси) позволяют не только на два порядка повысить скорость формирования параметров световых сигналов, но и переводят твердотельную телекамеру в новый класс приборов с множественным потоком данных, способствующим извлечению полезной (доминантной в данном эксперименте) видеоинформации с минимальными вычислительными затратами.

Глава 3

ПРОЕКТИРОВАНИЕ СИСТЕМ НА ОСНОВЕ ТЕОРИИ ИНФОРМАЦИИ

Теория телевидения – это раздел теории связи, такой же, как теория радиолокации, отличающийся от неё в основном диапазоном длин волн, служащих для передачи информации. Теория информации, срастаясь с теорией вычислительной техники (различными разделами дискретной математики – теорией алгоритмов, теорией чисел и т. п.), должна определять стратегию при выборе структуры и параметров информационных машин и определять границы возможного.

Формальной логической основой теории информации является в основном аксиоматика теории вероятностей, дополненная новыми понятиями, введение которых имеет характер аксиом. А. Н. Колмогоров отмечал, что «*в качестве итога исследований аксиоматического направления можно считать выясненным, что никакой другой столь же естественной скалярной сводной характеристики информации, содержащейся в одном случайном объекте ξ относительно другого случайного объекта η, кроме $I(\xi,\eta)$, существовать не может*» [150]. Здесь указывается на главенство понятия *взаимной информации* (а не энтропии). Важнейшие понятия теории информации, наполненные конкретным смыслом, – лишь экстремальные значения функционала Шеннона для взаимной информации – энтропия H источника и пропускная способность C канала связи. Не случайно А. Н. Колмогоров писал: «*я настаиваю на той идее, что основным понятием, допускающим обобщение на совершенно произвольные непрерывные сообщения и сигналы, является не непосредственно понятие энтропии, а понятие количества информации $I(\xi,\eta)$*». Эти понятия стали опорными в формулировке приводимых ниже первых трёх аксиом теории информации, позволивших доказать множество теорем и построить теорию передачи «чистой» информации. Однако часто информация предстаёт в сложной смеси сигналов и помех. В частности, известно, что главной проблемой семантической обработки информации является выделение нужных фактов и образов из смеси фактов, не относящихся к доминантным в данной постановке задачи. Даже известны мнения о том, что груда фактов (для телевидения это «груда фотонов»), не обработанная аналитиком, не является информацией. Поэтому обойтись *одним*

понятием взаимной информации не удаётся. Следует обратить внимание на мысль А. Н. Колмогорова, отмеченную им восклицательным знаком: «*Однако так как "информация" по своей природе не обязана быть (и в действительности не является!) скалярной величиной, то никакие аксиоматические исследования указанного направления не могут ответить на вопрос о том, сколь полно характеризует величина $I(\xi,\eta)$ интересующую нас информацию. Некоторый ответ на этот вопрос … дают теоремы, обосновывающие общие принципы теории передачи информации, состоящие в том, что при $H < C$ в принципе возможно производить передачу без систематически возрастающего запаздывания, если подлежащая передаче информация … надлежащим образом кодируется*».

Эти два тезиса составляют конструктивную («положительную») часть аксиоматической основы теории информации, исследующей передачу по каналам с шумом сигналов, наблюдаемых в смеси с другими сигналами и шумами. «Отрицательной» частью основы теории передачи информации о полезном сигнале, наблюдаемом в сложной смеси сигналов и помех, является несогласие с тезисом (оформленным в виде теорем скалярной информации) о том, что «кодирующая система может уменьшить, а в лучшем случае сохранить имеющуюся информацию». Последний тезис является следствием игнорирования идеи А. Н. Колмогорова о том, что информация не является скалярной величиной, и причиной многих неудач в применении теории информации к практике телевидения. Сумма этих (частично интуитивных) положений позволяет без претензии на полноту системы аксиом теории нескалярной информации дополнить её четырьмя новыми аксиомами и распространяет применение теории информации на передачу многокомпонентных сигналов.

1. Существует неотрицательная характеристика информационной связи двух случайных объектов, называемая *взаимной информацией*, которой разумно приписать логарифмическую связь с числом и вероятностной структурой возможных значений случайных объектов.

2. Существует *минимум* взаимной информации, который характеризует скорость создания информации источником, называемый *энтропией* источника.

3. Существует *максимум* взаимной информации, который характеризует скорость передачи информации через канал связи, называемый *пропускной способностью* канала связи.

4. Существует *взаимная информация* между выходным сигналом и *любой из компонент* входного воздействия системы связи – *доминантным* сигналом, *фоновыми* сигналами и *шумами*.

5. Кодирование и декодирование может и имеет целью *повысить качество* информации.

6. Качество переданной информации определяется соотношением взаимных информаций между выходным сигналом системы связи и *различными компонентами* входного воздействия.

7. Качество (эффективность) кодирования информации определяется соотношением потери информации о *доминантной компоненте* входного воздействия и *торможения* информации об остальных компонентах воздействия.

Теория, как и практика, не стоит на месте. В результате развития в теории информации давно выделились несколько направлений в соответствии с тремя подходами к теории информации – вероятностным, комбинаторным и алгоритмическим [150], в каждом из которых получено большое количество результатов (см., напр., [3, 126, 151]). Практически самостоятельно развиваются её различные ветви, такие как семиотика [149]. Однако «достаточно полной, законченной, устраивающей всех строгой теории информации не придумано» [123]. Поэтому в задачи теории информации современного этапа входит интеграция её различных ветвей, должных вооружить проектировщиков систем единой методологией синтеза информационных машин с учётом вероятностных свойств обрабатываемых сигналов (образов) и алгоритмического подхода к оптимизации их структуры при учёте доступной сложности. Такая теория информации в отличие от уже развитых, должна иметь целью передачу не только сигналов и информации, но и образов [3].

Период электронно-лучевого телевидения в основном совпадает по времени с многочисленными, часто мало конструктивными, попытками приложения теории Шеннона к кодированию видеоинформации. Твердотельная революция в телевидении стимулировала развитие теории, учитывающей, в отличие от теории Шеннона, требование реального времени и наличие входного фотонного шума в кодируемом сигнале. Важным шагом на этом пути, соответствующем потребности объединения различных ветвей теории связи, является информационная концепция [3], основу которой составило уравнение связи. Оно, с сохранением шенноновской меры информации, позволило выйти за рамки блокового кодирования Шеннона и стало ядром теории

телевидения как связи в реальном времени. Новая концепция утверждает, что вся связь, являясь фрагментом процесса познания мира, осуществляется исходя из принципа минимума информации при назначаемом качестве получаемой информации (или качестве кодирования). Этот принцип является развитием понятия об эпсилон-энтропии источника информации [150], но оперирующим уже не только энергетическими, но и информационными показателями качества. Теперь обоснована эквивалентность понятий эпсилон-энтропии источника и эпсилон-пропускной способности канала [3] и доказано, что существует не один метод кодирования, а разные методы кодирования, которые удовлетворяют требованию информационного равновесия, понимаемого как равенство скорости создания информации, которую мы хотим передать, и скорости передачи информации через канал. Уравнение связи позволяет учесть влияние не только шума канала, но и входного шума. Такой учёт является следствием приведённых аксиом теории информации и осуществляется благодаря введению понятий не только *полезной информации о доминантном сигнале*, но и *шумовой информации* – передаваемой получателю информации о входном шуме [1–4]. Целью информационного проектирования становится не просто достижение равновесия, а такого, когда удовлетворяются требования критериев максимума качества информации или максимума качества кодирования информации. Так же, как минимум среднего риска в теории решений приводит к критерию максимума правдоподобия, максимум качества кодирования приводит к критерию минимума информационного риска, состоящего из взвешенной суммы потери полезной информации и переданной шумовой информации [4].

Резюмируя свои работы по математической теории связи, Шеннон писал, что «Основные идеи теории связи не новы. Однако теперь значительно глубже понимается действие шума в канале и важность статистических свойств сообщений, предназначенных к передаче». Подобно этому, уравнение связи сделало понятнее роль равновесия в системе связи, а критерий минимума информационного риска позволил лучше понять влияние нескольких источников шума, отвергнув сведение всех источников шумов в одну точку.

Теория информации, включающая концепцию связи [1–4], призвана обосновать такое построение системы, при котором проектировщик уверен в том, что для имеющегося набора {цель + априорная информация + набор ограничений} синтезированная система оптимальна.

3.1. Применение математической теории связи в телевидении

Статистическая теория передачи информации формализуется при задании свойств случайного сигнала и канала связи. Случайный сигнал X задаётся плотностью вероятности $q(x)$, а сигнал на выходе канала Z имеет плотность вероятности $p(z)$. Переход от передаваемого сигнала к получаемому осуществляется за счёт переходной характеристики канала $p(z|x)$. Связь (мера сходства) между передаваемым и получаемым сигналами может быть охарактеризована совместной плотностью вероятности $p(x, z)$. Для характеристики системы связи уже недостаточно вычисления ошибок, а требуется мера, характеризующая точность и сложность передачи. Совокупность плотностей вероятности объединена К. Шенноном в меру количества взаимной информации между случайными величинами $I(X, Z)$, или, что то же самое, между их плотностями вероятности $I(q, p)$ [174]:

$$I(X,Z)=\iint p(x,z)\log\frac{p(x,z)}{q(x)q(z)}dxdz=\iint q(x)p(z|x)\log\frac{q(x)p(z|x)}{q(z)}dxdz=I(q,p).$$

Синтез системы передачи информации состоит в выборе функций распределения источника сигнала и получаемого кода. Такой выбор имеет двойственный характер: можно искать наилучшее распределение $q(x)$ источника при заданном канале $p^*(z)$ и можно искать наилучшее распределение кода $p(z)$ при заданном источнике $q^*(x)$. Совместная оптимизация и источника, и канала позволяет найти оптимальную пару функций распределения сигнала q^* и переходную вероятность канала p^*, т. е. синтезировать уравновешенную систему, описываемую уравнением связи [3]. Такое согласование обеспечивает равновесие в скорости рождения сигнала и его передачи и обобщает неравенство Шеннона, связывающее скорости создания информации источником (энтропию) и передачи информации по каналу связи (пропускную способность).

Все системы передачи информации синтезируются при наличии некоторой совокупности ограничений $\{G_j\}$, при которых должно достигаться равновесие между скоростью создания информации источником и скоростью её передачи по каналу связи. Поэтому уравнение связи записывается при совокупности ограничений. Ограничения – это не критерии, хотя их иногда и путают, относя к «критериям второго рода» [108, 110]. Такая инверсия критериев и ограничений выводит проектирование из рамок теории связи и ведёт к мнениям о том, что «критерий максимума экономической эффективности целесообразно выбрать

основным, а остальные критерии взять в качестве ограничивающих» [108]. Применение новых технологий позволит более эффективно решать задачи кодирования, так как численное значение ограничений на ёмкость памяти и быстродействие по мере совершенствования технологии всё более «играет в пользу» разработчика систем. Это открывает путь к реализации всё более сложных алгоритмов, но всё равно оставляет сложность системы в ранге ограничений. Ограничения в рамках теории информации могут быть весьма разнообразными:

- на качество информации, характеризуемое отношением полезной и шумовой информации в полученном сигнале;
- на качество кодирования информации, характеризуемое информационным риском, т. е. суммой потери полезной информации и переданной шумовой информации;
- на доступную мощность в канале;
- на доступную площадь усиления кодирующего фильтра;
- на сложность цифрового кодера;
- на допустимую задержку передаваемого сигнала.

В качестве примера такого ограничения К. Шеннон использовал ограничение на мощность сигнала в канале. Проектировщик должен учитывать большую совокупность ограничений, часть которых $\{E_i\}$ может относиться к источнику, а часть $\{E_n\}$ – к каналу. Поэтому и в уравнение связи кроме количества взаимной информации входят указанные ограничения [3]:

$$\max_{\{E_i\};\, q = \text{var}} I(q, p^*) = \min_{\{E_n\};\, p = \text{var}} I(q^*, p) = I(q^*, p^*). \qquad (3.1)$$

Существует ряд решений уравнения связи, соответствующих оптимальным системам. Среди них и двоичный симметричный канал, и непрерывный гауссовский канал, на примере которого наиболее наглядна роль качества информации и методы её повышения. То, что и дискретный, и непрерывный каналы связи могут быть оптимальными, оставляет проектировщикам свободу выбора типа системы. При любом типе канала его важнейшей характеристикой является пропускная способность. Количественно она выражается с помощью функционала Шеннона как наибольшая средняя взаимная информация, максимизированная по всем распределениям входного сигнала при каком-либо ограничении.

Кодирование сигналов (в том числе изображений) с предварительным преобразованием в спектральную область и обработкой вычисленных спектральных коэффициентов разложения в ряд по

различным системам базисных функций ведёт своё происхождение от задачи об эпсилон-энтропии [150]. Наиболее общие теоретические исследования опираются на разложение Карунена–Лоэва. Теория передачи сигналов с погрешностью предписывает метод кодирования, основанный на усечении спектра сигнала при равномерном распределении среднеквадратической ошибки (СКО) передачи по передаваемым спектральным компонентам сигнала. Впоследствии этот метод кодирования широко использовался как в теоретических исследованиях, так и при построении практических алгоритмов сжатия сигналов, включая изображения [112–119]. Для канала с гауссовским сигналом со спектром $\{\lambda_k\}$, кодером $\{G_k\}$, и шумом с дисперсией σ_2^2 предельная пропускная способность при ограничении мощности в канале определяется через спектральное отношение сигнал/шум $\Psi_k^2 = \lambda_k G_k^2 / \sigma_2^2$:

$$C = \max I(X, Z) = \frac{1}{2} \sum_{k=1}^{\infty} \log(\Psi_k^2 + 1). \qquad (3.2)$$

Можно оценить пропускную способность телевизионной системы в предположении, что скорость передачи ограничивает не радиоканал, а телекамера, в первую очередь – фотоприёмник. Для оценки такой пропускной способности можно воспользоваться трактовкой телекамеры как кодера в схеме на рис. В.2, в которой для борьбы с шумом канала используется накопление за время кадра на площади элемента разложения и усиление в *a* раз. При большом числе фотоэлектронов в каждом элементе разложения функцию распределения шума в нём можно считать гауссовской и воспользоваться выражением для пропускной способности гауссовской системы. В первом приближении можно считать, что полезный сигнал равен произведению контраста *K* сигнала в каждом элементе относительно соседнего элемента на среднее число *N* фотонов, накапливаемых в одном элементе разложения за время кадра. При независимости передаваемых кадров пропускную способность телевизионной системы с квадратным растром $Z \times Z$ элементов разложения можно вычислить по формуле [3]:

$$C = \frac{Z^2}{2T_к} \log (K^2 N^2 + N) \frac{a^2}{\sigma_2^2}.$$

Приведённое значение пропускной способности учитывает энергию полезного сигнала K^2N^2, дисперсию фотонного шума $\sigma_1^2 = N$ и приведённое значение дисперсии шума канала σ_2^2 / a^2, а

также показывает, что пропускная способность телевизионной системы растёт по мере увеличения не только разрешающей способности и числа принимаемых в элементе фотонов, но и по мере увеличения докоммутационного усиления.

Линейное кодирование зашумлённого случайного сигнала. В модели передачи непрерывного сигнала по гауссовскому каналу для отражения влияния входного (фотонного) шума, как и на рис. В.2, должны быть учтены два источника шума (рис. 3.1). В схеме линейного кодирования сигнал с дисперсиями спектральных компонент $\{\lambda_k\}$ неконтролируемо искажается фильтром с частотной характеристикой $\{B_k\}$ и учтён канал связи, который содержит кодер $\{R_k\}$ и декодер$\{Q_k\}$. Применение уравнения связи к этой схеме позволяет найти и параметры оптимальных кодера и декодера, и достижимые характеристики системы.

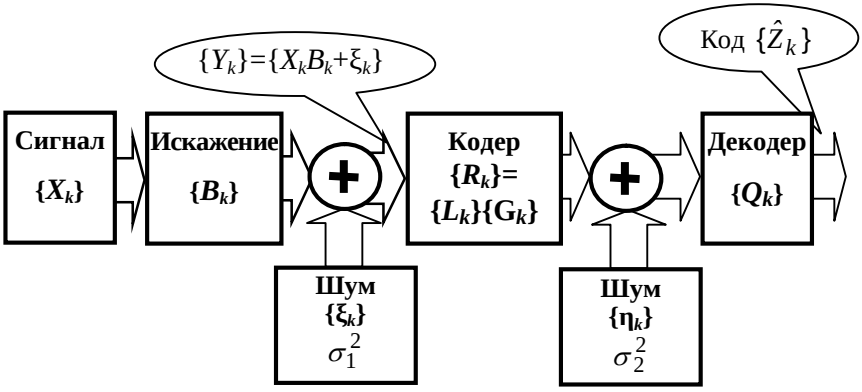


Рис. 3.1. Модель линейного кодирования зашумлённого сигнала

Оптимальный декодер $\{Q_k\}$ является фильтром Винера, учитывающим преобразование сигнала двумя другими фильтрами (искажающим и кодирующим). Оптимальный кодер $\{R_k\}$ состоит из двух звеньев: первое является кодером источника, реализуемым как фильтр Винера $\{L_k\}$, компенсирующим влияние искажающего фильтра формированием оптимальной оценки сигнала (условного математического ожидания) с дисперсиями $\{\hat{\lambda}_k\}$, второе – кодером канала, реализуемым как конечномерный фильтр $\{G_k\}$. Оптимальный линейный фильтр Винера со спектральной

характеристикой $L_k = \frac{\lambda_k B_k}{\lambda_k B_k^2 + \sigma_1^2} = \frac{1}{B_k} \frac{\Psi_k^2}{\Psi_k^2 + 1}$ имеет две характерные области: те спектральные компоненты, где отношение сигнала к шуму велико, усиливаются для компенсации влияния искажающего фильтра, а те, которые слабее шума, дополнительно подавляются. В телевидении фильтр Винера применяется давно, он входит в состав усилительного тракта телекамер и служит для коррекции искажений, вносимых в процессе считывания зарядового изображения паразитными емкостями фотоприёмника и входных каскадов усилителей, и называется апертурным корректором. Минимальная СКО, составляет

$$\varepsilon_B = \sum_{k=1}^{\infty} \frac{\lambda_k \sigma_1^2}{\lambda_k B_k^2 + \sigma_1^2}. \qquad (3.3)$$

Минимальная СКО (3.3) требует использования *всех* спектральных компонент сигнала вплоть до бесконечности, даже тех, где отношение сигнал/шум $\Psi_k{}^2 = \lambda_k B_k^2 / \sigma_1^2 << 1$. Иными словами, *минимальная* достаточная статистика, формируемая фильтром Винера, содержит *бесконечное количество кодовой* информации (взаимной информации между выходным кодом и смесью доминантного сигнала и шума) $I(Y, Z)=\infty$ при *конечном количестве полезной* информации, т. е. взаимной информации $I(X, \hat{X}) = I(X, Y) < \infty$ между полезным сигналом X и его оценкой $\hat{X}$, вычисляемой по формуле (3.2).

Усечение числа коэффициентов бесконечного ряда кодирующим фильтром осуществляется именно с целью приведения в соответствие конечного количества полезной информации с конечным числом передаваемых спектральных компонент. Усечение осуществляется везде, где усиленные кодером в a раз спектральные коэффициенты оценки полезного сигнала $\{\hat{X}_k\}$ «тонут» в шуме канала, т. е. спектральное отношение сигнал/шум $\hat{\Psi}^2 = \hat{\lambda} a^2 / \sigma_2^2$ становится меньше 1.

Информационное равновесие в схеме рис. 3.1. согласно уравнению (3.1) достигается не при произвольных ограничениях, а при их взаимосвязанной совокупности. Так, при ограничении качества рождаемой источником информации заданной СКО,

равенство скорости создания информации и скорости её передачи достигается при ограничении площади усиления кодирующего фильтра. Использование СКО в качестве показателя качества информации и другого ограничения на кодер, например, рассмотренного Шенноном ограничения на мощность в канале, не позволяет достичь равновесия, требуемого уравнением (3.1).

Доступное усиление кодирующего фильтра и граничная полоса частот связаны с ограничением на его площадь усиления (сумму квадратов коэффициентов):

$$\sum_{k=1}^{\infty} G_k^2 \le E \quad . \tag{3.4}$$

Частотные характеристики кодирующего и декодирующего фильтров, обеспечивающих информационное равновесие в системе (равенство ε-энтропии источника и пропускной способности канала при минимуме СКО), отличаются только постоянными коэффициентами усиления [2 – 4]:

$$G_{k\ \mathrm{opt}}^2 = a^2 \frac{\hat{\Psi}^2 - 1}{\hat{\Psi}^2}; \qquad Q_{k\ \mathrm{opt}}^2 = \frac{1}{a^2} \frac{\hat{\Psi}^2 - 1}{\hat{\Psi}^2}.$$

Для монотонно убывающих спектров сигнала (например, описываемыми полиномами Баттерворта) спектральные характеристики кодера и декодера также монотонно убывают. Это отличает их от оптимальных линейных фильтров, обеспечивающих минимум СКО при ограничении мощности в канале [152]. Такая монотонно убывающая спектральная характеристика оптимального кодера весьма близка по форме к реальной частотно-контрастной характеристике твердотельных фотоприёмников [2].

В оптимальной линейной системе передачи непрерывных сигналов количество передаваемой полезной информации, вычисляемой с использованием функционала Шеннона в уравнении связи, составляет[2–4]

$$I(X,Z) = \frac{1}{2}\sum_{k=1}^{N} \log \frac{a^2 \hat{\lambda}_k}{\sigma_2^2} = \frac{1}{2}\sum_{k=1}^{N} \log \hat{\Psi}_k^2 = \sum_{k=1}^{N} \log \hat{\Psi}_k . \tag{3.5}$$

Формула (3.5) отличается от знаменитой формулы (3.2) Шеннона для пропускной способности. Хотя количественное различие этих двух формул для пропускной способности и не велико, но по смыслу – существенно: кодирование информации в реальном времени ценой допущения небольших потерь полезной информации в ходе рождения сигналов позволяет достичь лишь

немногим меньшей пропускной способности, но позволяет отказаться от блокового кодирования и использования бесконечного спектра передаваемого сигнала.

Так как информация не является скалярной величиной, то ни качество переданной информации, ни качество кодирования не могут быть охарактеризованы только пропускной способностью, да и любым другим *одним* числом. Через канал связи вынужденно передаётся ещё и входной шум $\{\xi_k\}$, в результате чего выходной код содержит шум $\{\omega_k\}$, имеющий две составляющие, определяемые входным шумом и шумом канала $\{\eta_k\}$: $\omega_k = G_k Q_k \xi_k + Q_k \eta_k$. Передаваемая через канал кодовая информация $I(Y,Z)$ состоит из *полезной (доминантной) информации* и *шумовой информации*: $I(Y, Z) = I(X, Z) + I(\xi, \omega)$. Уровень шума канала по разному влияет на полезную и на шумовую информацию. При увеличении шума канала и полезная информация и шумовая информация убывают, но с разной скоростью. При $\varepsilon < 2\varepsilon_B$ это различие формализуется соотношением, связывающим потерю полезной информации $\Delta I = I(X, Y) - I(X,Z)$ и шумовую информацию $I_ш = I(\xi, \omega)$ [2–4]:

$$\Delta I \times 2^{2I_ш} \approx 1. \qquad (3.6)$$

Если в системе применяются произвольные фильтры, а не оптимальные, то соотношение (3.6) приобретает вид неравенства, схожего по структуре с соотношением неопределённостей Гейзенберга. Потеря информации при оценке формы гауссовского случайного сигнала определяется логарифмом отношения дисперсий фактической и потенциальной ошибок: $\Delta I = \frac{1}{2} log \frac{\varepsilon}{\varepsilon_B}$.

Учёт не только полезной информации, но и шумовой, позволяет оптимизировать систему передачи как по согласованию источника с каналом, так и по согласованию передаваемого кода с требованиями получателя по *качеству кодирования информации*. Информационный риск R_I служит той же идее оценки достижимого качества кодирования (передачи информации с погрешностью), что и ε-энтропия, однако в отличие от неё учитывает не только качество информации, но и качество кодирования, определяемое соотношением *потери* полезной информации и *переданной шумовой информации*. *Информационный риск* вычисляется с учётом цен потерь разного рода (которые по смыслу аналогичны ценам в критерии Байеса проверки гипотез) [4]:

$$R_I = c_0 \Delta I + c_1 I_ш . \qquad (3.7)$$

В простейшем случае равенства цен потери полезной информации и загрузки канала и получателя шумовой информацией *минимум информационного риска* становится аналогом критерия идеального наблюдателя в теории решений и достигается при равенстве СКО из-за первого и второго источников шума в схеме рис. 3.1:

$$R_I = \Delta I + I_{\text{ш}} . \qquad (3.8)$$

Конструкция критерия минимума информационного риска отражает стремление проектировщика найти компромисс между потерей полезной информации и загрузкой канала и получателя шумовой (ложной) информацией. На семантическом уровне этот компромисс означает поиск компромисса между истиной и ложью. То, что такой компромисс достигается не при слишком строгих правилах, знал, например, Рабиндранат Тагор, утверждавший: «*Закройте дверь перед всеми ошибками, и Истина не сможет войти!*». В рассматриваемом случае линейного кодирования это означает предостережение от попытки назначить чрезмерно высокую цену шумовой информации. Такая тенденция характерна для оценки качества информации по информационному к. п. д. η, равному отношению полезной информации и шумовой информации: $\eta = I(X, Z)/I_{\text{ш}}$. Однако этот показатель достигает максимума при слишком большой ошибке передачи (рис. 3.2). Поэтому перспективнее применение критерия минимума информационного риска (3.7), характеризующего *качество кодирования*, т. е. то, насколько рационально распорядился кодер доступной ему *смесью полезной и шумовой информации*. Информационный риск не просто отражает приведённое мнение писателя, но и *дополняет* его двойственной формулировкой: «*Распахните дверь для Истины, и в неё хлынет Ложь*».

Выбор полной ошибки передачи по характеру оптимизационной задачи аналогичен выбору порога в задаче проверки гипотез: в обеих задачах ищется компромисс между потерей информации и шумовой информацией. Вариант (3.7) с весовыми коэффициентами родственен критерию Байеса, вариант без них (3.8) – критерию идеального наблюдателя.

В результате такого компромисса (3.8) коэффициент шума системы связи становится равным двум (полная ошибка передачи равна удвоенной ошибке, обеспечиваемой фильтром Винера), количество передаваемой шумовой информации в каждом отсчёте уменьшается от бесконечности до полбита, а потеря полезной информации остаётся на невысоком уровне полбита на символ.

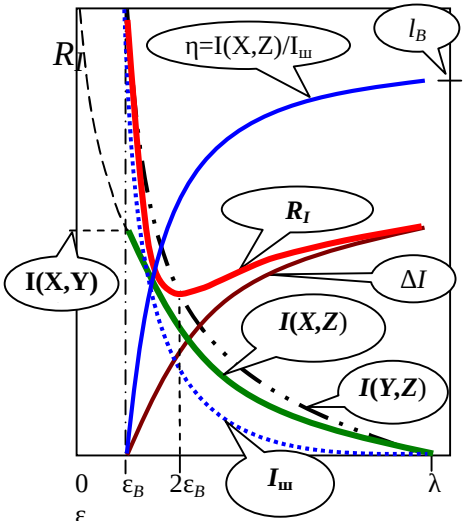


Рис. 3.2. Зависимость информационного риска R_I и его составляющих – потери полезной информации ΔI и шумовой информации $I_ш$ от среднеквадратичной ошибки ε; минимум информационного риска при равенстве цен потери полезной информации и шумовой информации достигается при полной ошибке $\varepsilon=2\varepsilon_B$, т. е. при коэффициенте шума, равном двум

Обобщения понятия информационного риска на случай сложной входной смеси сигналов $\{X_i\}$, состоящей из доминантной компоненты и K помеховых (фоновых и/или шумовых, как, например, при селекции движущихся целей на сложном фоне) компонент, приводит к критерию

$$R_c = c_0 \Delta I_0 + \sum_{i=1}^{K} c_i I(Z, X_i).$$

В уравновешенной с точки зрения минимума информационного риска системе скорость передачи всей кодовой информации равна потенциальной информации о полезном (доминантном) сигнале. Обе лишь незначительно отличаются от скорости передачи информации о полезном сигнале (3.5):

$$I(Y, Z) = I(X, Y) = \frac{1}{2} \sum_{k=1}^{N} log\ (\hat{\Psi}_k^2 + 1).$$

При отсутствии входного шума ограничение (3.4) может назначаться произвольно – при любом его конечном значении происходит существенная потеря полезной информации. Однако при наличии входного шума ограничение площади усиления (3.4) не может назначаться произвольно. Наличие минимального значения информационного риска означает, что равновесие в скорости создания и передачи информации достигается при условии: *оптимальный линейный кодер усекает те и только те*

спектральные компоненты, где отношение сигнал/шум, определяемое совокупным влиянием входного шума и шума канала, становится меньше порогового значения. В практике телевидения взаимообмен между потерей полезной информации и шумовой информацией наиболее нагляден при выборе коэффициента докоммутационного усиления: шум видеотракта нужно уменьшать до тех пор, пока он не сравняется с фотонным шумом.

Кодирование при ограничении количества информации о кодере. Сложность систем изучается, прежде всего, для создания методов её снижения [110] до уровня, минимально достаточного для обеспечения заданного качества информации. Взаимосвязь точности (СКО ε) и сложности S, изучаемая теорией сложности как разделом теории алгоритмов (эпсилон-сложность [126]) и теорией информации (ε-энтропия [150]), во многих случаях имеет простую формализацию: ошибка убывает экспоненциально при росте сложности при обработке незашумлённых сигналов. При *передаче* через канал с гауссовским шумом зашумлённых гауссовских сигналов соотношение между точностью (потерей полезной информации) и сложностью передачи (кодовой ε-энтропией) формализуется соотношением неопределённости, похожим на соотношение (3.6), отличаясь от него тем, что произведение входящих величин не является константой, а зависит от входного отношения сигнал/шум:

$$\Delta I \times 2^{2I(Y,Z)} \approx \Psi^2 + 1. \qquad (3.9)$$

Так же как из соотношения (3.6) был сконструирован критерий минимума информационного риска (3.7), из соотношения (3.9) может быть сконструирован критерий вида:

$$R_{I(2)} = c_0 \Delta I + c_1 I(Y,Z). \qquad (3.10)$$

При *обработке* зашумлённых сигналов при росте доступной сложности *S* экспоненциально убывает дополнительная (алгоритмическая) ошибка, добавляющаяся к минимально возможной шумовой ошибке (для линейной фильтрации эта ошибка фильтра Винера (3.3)):

$$S \sim log \frac{1}{\varepsilon - \varepsilon_B}.$$

Эта зависимость приводит к тому, что на практике (см. раздел 4.4) используют критерий качества кодирования, родственный критерию (3.10), связывающий потерю полезной информации с вычислительной сложностью системы кодирования:

$$R_{I(3)} = c_0 \Delta I + c_1 S \,. \qquad (3.11)$$

Ограничение сложности является фундаментальным и важным при любом подходе к теории связи. Например, в семиотике отмечается, что «в зависимости от кодируемого объекта выбирается соответствующая форма знака; выбор велик, и он определяется различными слагаемыми ситуации общения, намерениями кодирующего и физическими возможностями, находящимися в данный момент в его распоряжении» [149].

Одной из двух важнейших задач кодирования видеоинформации наряду с кодированием канала является кодирование источника, часто осуществляемое с использованием предварительного перехода от временного представления сигнала к спектральному представлению. Такое кодирование источника называют кодированием с преобразованием. В рамках кодирования с преобразованием можно выделить две характерные задачи. Первая – для статистически заданного сигнала при заданной системе базисных функций (например, разложение Карунена–Лоэва) и при заданной сложности хранения (передачи) нужно найти распределение сложности (ёмкости памяти) по спектральным компонентам. Эта задача отличается от вычисления ε-энтропии тем, что выходной (хранимый) сигнал имеет дискретное распределение, а не непрерывное. Вторая задача – для статистически заданного сигнала при заданной сложности хранения (передачи) нужно определить систему базисных функций и распределение сложности (ёмкости памяти) между числом базисных функций и их разрядностью.

Первая задача формализуется как поиск совокупности *целочисленных* значений $\{m_k\}$ разрядов хранимых спектральных компонент при ограничении их суммы заданным целым числом M (совокупной ёмкости памяти): $\sum_{k=1}^{\infty} m_k \le M$. В этой задаче[2], так же как и при вычислении ε-энтропии, учитывается две компоненты ошибки: в полосе пропускания и вне её. Ошибка за счёт усечения спектра после N-го спектрального отсчёта вычисляется точно так же, как и в классической задаче об ε-энтропии в виде суммы $\sum_{k=N+1}^{\infty} \lambda_k$. Ошибка в полосе пропускания в этой задаче из-за

[2] Решение этой задачи получено совместно с В. В. Варюхиным.

требования целочисленности ёмкости памяти каждого разряда уже не имеет равномерного распределения по спектру. В результате при заданном значении M совокупной ёмкости памяти суммарная ошибка становится большей, чем вычисляемая по формуле для «второй» ε-энтропии без требования целочисленности значений ёмкости памяти спектральных компонент. Аналитическое решение даёт оценки для числа разрядов m_k и числа N хранимых спектральных компонент $m_k = ent\left(\frac{1}{2} log_2 \frac{\lambda_k}{\lambda_N}\right)$, подтверждаемые модельным экспериментом средствами табличного процессора Excel для сигналов со спектрами Баттерворта (рис. 3.3).

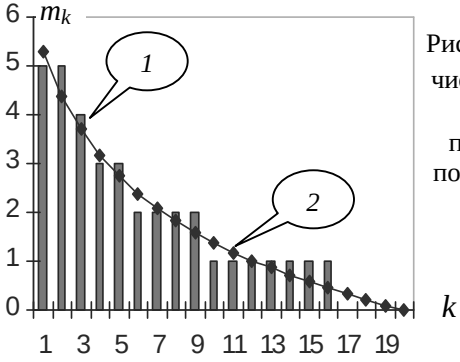


Рис. 3.3. Распределение по спектру числа разрядов m_k *1* для сигнала со спектром *2*, описываемым полиномом Баттерворта первого порядка с эквивалентной полосой частот K

$$\lambda_k = \frac{K}{K^2 + k^2}$$

Вторая задача имеет следующие особенности. В силу сложности формирования базисных функций разложения Карунена–Лоэва и отсутствия для него методов сокращения вычислительных затрат (как в быстром преобразовании Фурье) практики используют более простые системы базисных функций. При этом для различных систем базисных функций (оптимальных для различных статистических свойств сигналов: косинусного, наклонного, Хаара, Адамара и др.[165]) путём моделирования показана разная скорость сходимости ряда, иллюстрируемая разной скоростью уменьшения ошибки передачи при увеличении числа передаваемых спектральных компонент (членов ряда) [118, 119]. При этом разложение Карунена–Лоэва характеризуется максимальной скоростью сходимости ряда, меньшей скоростью характеризуется косинусное преобразование, самой малой – разложение Уолша–Адамара. Традиционный подход к определению числа базисных функций при кодировании опирается на определение ошибки, вызванной лишь усечением спектра

полезного сигнала, когда считается, что в полосе пропускания ошибка кодирования равна нулю. Такой подход к кодированию лишь скрыто подразумевает ограничение сложности кодера, выделяя лишь один из показателей – число базисных функций без учёта другого показателя – разрядности базисной функции. Следствием такого подхода к кодированию с переходом к спектральному представлению сигналов стал ряд стандартизованных алгоритмов кодирования изображений типа JPEG, MPEG [107, 117, 171], в которых используется дискретное косинусное преобразование, число разрядов базисных функций просто считается достаточным, выбираемым исходя из монотонного убывания ошибки при увеличении разрядности отсчётов базисных функций. Однако при синтезе кодеров, особенно однокристальных (выполняемых в виде одной СБИС), следует учитывать полный объём памяти, отводимой на значения базисных функций, определяемый произведением числа функций на их разрядность. Поэтому разрядность базисных функций и их число не фиксированы, а находятся по максимуму качества кодирования при заданном ограничении сложности кодера.

Моделирование с использованием типовых сигналов изображений показывает [153], что доступную сложность S следует распределять между числом базисных функций N и разрядностью m их отсчетов таким образом, чтобы обеспечить минимум СКО передачи сигналов и/или максимума пропускной способности (рис. 3.4.). При этом в отличие от первой задачи число разрядов базисных функций для различных спектральных компонент не зависит от их номера.

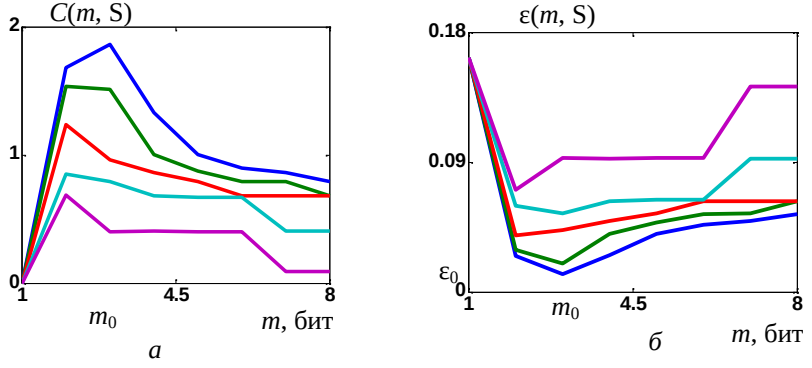


Рис. 3.4. Зависимость пропускной способности *а* и среднеквадратической ошибки передачи *б* от разрядности представления базисных функций для сигнала при различных значениях ограничения сложности кодирования S

Значения оптимальной разрядности m_0 для одного и того же сигнала при одном и том же значении ограничения сложности S при оптимизации по разным критериям (минимума СКО ε(S) или максимума пропускной способности C(S)) близки, но не совпадают. Это отличает рассмотренную модель кодирования с ограничением сложности от идеального линейного кодирования, обеспечивающего строгое совпадение параметров фильтра, оптимизируемого по критериям и минимума СКО, и максимума пропускной способности.

Переход от ограничения числа передаваемых спектральных компонент сигнала к ограничению сложности кодера при кодировании с преобразованиями меняет приоритет при выборе базисных функций. В первом случае, когда сложность базисных функций не учитывается, выбираются те разложения, которые имеют максимальную скорость сходимости ряда (разложение Карунена–Лоэва), во втором случае лучшее качество кодирования обеспечивают разложения с базисными функциями, требующими меньшего количества информации для их хранения (например, Уолша–Адамара).

Так же как и в линейном кодировании, где ограничение на площадь усиления должно выбираться исходя из соотношения полезной и шумовой информации в сигнале, ограничение сложности кодера не может назначаться произвольно, а должно назначаться исходя из правила обнуления спектра передаваемого сигнала всюду, где спектральное отношение сигнал/шум ниже порога. Общая тенденция кодирования зашумлённых сигналов состоит в том, что чем больше отношение сигнал/шум, тем больше требуемая оптимальная сложность кодера, обеспечивающего максимальное качество кодирования.

3.2. Квантование сигналов в условиях фотонного шума

Обработка видеоинформации в современных компьютерах осуществляются в цифровой форме, а главным элементом кодирования накопленного фотоприёмником изображения для его согласования с цифровым каналом связи является квантователь, или аналого-цифровой преобразователь (АЦП). Теоретическую основу квантования составляет теория ε-энтропии, основа которой создана А. Н. Колмогоровым и К. Шенноном. А. Н. Колмогоров, в частности, ввёл две формы ε-энтропии, связанных со способом ограничения ошибки – через единичную вероятность P непревыше-ния расстоянием ρ(ξ, ξ') между исходной величиной ξ

и её аппроксимацией ξ' заданного значения ε, либо через математическое ожидание M дисперсии этого расстояния: «*Кажется естественным потребовать, чтобы* P{ρ(ξ, ξ')≤ε}=1, *или чтобы* Mρ²(ξ, ξ')≤ε². *Эти два вида «ε-энтропии» распределения* P_ξ *будем обозначать* $H_\varepsilon^0(\xi)$, $H_\varepsilon(\xi)$» [150].

Точного значения ε-энтропии первого вида $H_\varepsilon^0(\xi)$ для общего случая А. Н. Колмогоров не привёл (есть лишь оценки сверху, ограниченные системой неравенств [150]), и, хотя она как бы специально сконструирована для анализа аналого-цифрового преобразования, все обращаются к ε-энтропии второго вида, не смотря на то, что она эффективно применима в первую очередь к линейному кодированию в гауссовских каналах. На этом пути применения «второй» ε-энтропии $H_\varepsilon(\xi)$ к квантованию найдено, в частности, что при квантовании гауссовского сигнала при надлежащем выборе порогов скорость кода оказывается лишь *незначительно больше* ε-энтропии (примерно 0,3 бит/символ). При этом известные исследователи [176] даже удивляются тому, как немного можно достичь с помощью сложных систем кодирования *независимых гауссовских величин* (устранить указанную «избыточность» 0,3 бит/символ), если квантование уже оптимизировано по значениям порогов. Удивляться здесь нечему: квантование – единственная процедура перевода непрерывной величины в дискретную, и если бы велось сравнение с ε-энтропией первого вида $H_\varepsilon^0(\xi)$, то обнаружилось бы абсолютное совпадение достижимой скорости кода с ε-энтропией.

Кроме технологических аспектов и реализации сжатия непрерывных сигналов квантование является важнейшей операцией по выделению доминантной информации из шумовой информации. Эту функцию квантования рассмотрим на характерном для телевидения примере борьбы с фотонным шумом.

Важной спецификой видеоинформационных систем является мультипликативность всегда присутствующей помехи – фотонного шума сигнала. Так как излучение фотонов происходит случайным образом, то плотность вероятности накопления конкретного числа *n* фотов при среднем значении сигнала в элементе $\mathcal{N}$ фотонов подчинена закону Пуассона. Дисперсия σ^2 флуктуаций реально принимаемого числа *n* фотонов, равна среднему значению оптического сигнала $\mathcal{N}$:

$$\sigma^2 = M\{[\mathcal{N} - n]^2\} = \mathcal{N}. \qquad (3.12)$$

Известным решением задачи передачи информации по дискретному каналу при такой специфической мультипликативной помехе, искажающей входной сигнал, является квантование с переменным шагом. Компромиссное значение шага квантования оказывается с точностью до некоторого коэффициента, вычисляемого из энергетических соображений, пропорционально квадратному корню от уровня сигнала [120]. К определению коэффициента пропорциональности можно подойти с более обоснованных информационных позиций информационного риска (3.6). Соотношение шага квантования δ и СКО шума $\sigma = \sqrt{\varepsilon}$ при конкретной амплитуде видеосигнала определяется выбранными проектировщиком системы весами потери полезной информации и загрузки канала связи шумовой информацией: Этот подход в рамках энергетических понятий известен как метод создания «уравновешенной» системы [113]. Более общий взгляд с учётом весовых коэффициентов показывает, что при $c_0 > c_1$ цена потери полезной информации больше, поэтому число разрядов квантователя увеличивают. При $c_0 < c_1$ имеется обратная ситуация, т. е число разрядов уменьшают. В результате такой оптимизации квантование с переменным шагом, возрастающим пропорционально квадратному корню от уровня сигнала, при некотором соотношении переменного шага квантования и СКО фотонного шума [120] $\delta / \sigma = \sqrt{6/\pi e}$ (т. е. $\delta \approx 0.84\sigma$) обеспечивается глобальный минимум информационного риска, составляющий 1 бит/символ. В этом случае даже при переменном значении уровня фотонного шума обеспечивается постоянство коэффициента шума кодера (квантователя) во всем диапазоне изменения амплитуд сигналов. Однако компьютерная обработка сигналов, представленных с помощью переменного шага, резко повышает сложность алгоритма вычисления различных параметров сигнала (не говоря о трудностях реализации такого квантователя в условиях температурного дрейфа пороговых уровней).

На практике эту трудность преодолевают просто – игнорируют этот теоретический результат и используют квантование с постоянным шагом (оптимальное при аддитивном шуме) и для случая мультипликативного фотонного шума. В качестве некоторого приближения к оптимальному неравномерному квантованию можно рассматривать нелинейное преобразование до АЦП (гамма-коррекцию) с обратным преобразованием в приёмнике [59].

Число m разрядов АЦП, равное двоичному логарифму числа уровней квантования, определяется из доступного отношения

сигнал/шум, которое зависит от ёмкости $N_э$ элемента разложения фотоприёмника (измеряемой в числе электронов). Можно показать [120], что оптимальное значение числа разрядов m_o при *равномерном* квантовании составляет:

$$m_o \approx ½log(2N_э log N_э). \qquad (3.13)$$

При квантовании зашумленных сигналов всегда нужно знать соотношение шага квантования и СКО шума. Однако в рассматриваемой задаче квантования с постоянным шагом при мультипликативном шуме такое соотношение является не числом, а функцией от амплитуды сигнала. Вместе с тем можно попытаться охарактеризовать фотонный мультипликативный шум некоторым средним значением по всему диапазону амплитуд сигналов. Если считать, что сигнал имеет равномерную плотность вероятности, то среднее значение СКО фотонного шума, равного корню квадратному из амплитуды сигнала, составит $\sigma_{ср} = \sqrt{N_э / 2}$. Тогда из формулы (3.13) с учётом (3.12) можно получить оптимальное для равномерного квантования отношение шага квантования δ к усредненному по динамическому диапазону сигнала значению СКО фотонного шума [120]:

$$\frac{\delta}{\sigma_{ср}} = \frac{1}{\sqrt{log N_э}}. \qquad (3.14)$$

Значение минимального информационного риска для равномерного квантования вычисляется подстановкой оптимального числа разрядов (3.13) в формулы для шумовой информации и потери полезной информации:

$$R_o = \Delta I_o + I_{шо}; \qquad \Delta I_o = ¼; \qquad I_{шо} = ½log(2log N_э). \qquad (3.10)$$

Для реальных фотоприёмников на ПЗС накопительная способность элемента разложения $N_э$ находится в диапазоне $10^4 \ldots 10^6$. и минимальный информационный риск при равномерном квантовании в 2…3 раза больше, чем достижимый при неравномерном квантовании, шаг которого согласован с мультипликативным фотонным шумом. В частности, при типовом значении $N_э \approx 6\times10^4$ минимальный риск составляет 2,5 бита на элемент разложения. Рост числа уровней квантования обусловлен ростом динамического диапазона и при постоянном шаге квантования соответствует пренебрежимо малому весовому коэффициенту при шумовой информации по сравнению с весовым коэффициентом при потере полезной информации. Для типового значения накопительной способности элемента разложения $N_э = 64\times10^3$

(16 двоичных разрядов, 2 байта/пиксел) это даёт квантование с теми же 16 двоичными разрядами, или 2 байта/пиксел. Равенство цен потери полезной информации и шумовой информации в критерии минимума информационного риска ведёт к тому, что для типового значения ёмкости потенциальной ямы $N_э = 64\times10^3$ оптимальное число разрядов квантователя с равномерным шагом составляет 10,5 (1,25 байта/пиксел). Применяемое в массово выпускаемой телевизионной аппаратуре квантование на 8 разрядов (1 байт/пиксел) соответствует соотношению весовых коэффициентов, равное 6 ($c_0 \approx c_1/6$), т. е. фактически считается, что потеря полезной информации в шесть раз менее опасна, чем шумовая информация. Такой подход проявляется в том, что при достаточно большой освещённости сюжета шумы на изображении практически не заметны, но на некоторых изображениях проявляются «артефакты», например, в виде ложных контуров на полях с медленно изменяю-щейся яркостью.

Для телекамер на матрицах ПЗС с относительно большой емкостью элемента разложения порядка 2×10^5 ($\log_2 N_э \approx 17{,}5$) оптимальным является квантование на 11,5 разрядов (почти 1,5 байта/пиксел). Вместе с тем в таких астрономических телекамерах цена потери полезной информации очень высока ($c_0 > c_1$), и поэтому применяется большее число разрядов квантователя. Например, в малокадровых астрономических телекамерах производства «Электрон-оптроник», применяется равномерное квантование на 14 или 16 разрядов. Эти значения отражают противоположную тенденцию – в 2 раза большей цены потери полезной информации по сравнению с шумовой информацией (при $m = 14$ шумовая информация составляет примерно 5 бит/элемент, потеря полезной информации 10^{-2} бит/элемент) и в 4 раза (при $m = 16$ шумовая информация составляет 6,75 бит/элемент, потеря полезной информации 10^{-3} бит/элемент), чем цена шумовой информации (рис. 3.5). Повышение числа разрядов до 18 в этом примере означает *поштучный* (а не групповой) счёт числа электронов, накопленных в каждом элементе разложения, и среднюю загрузку шумовой информацией канала связи и вычислителя, составляющую (при равномерной плотности вероятности амплитуд сигналов) половину разрядов квантователя.

Анализ показывает, что при равномерном квантовании из-за чрезмерной загрузки канала связи и компьютера шумовой информацией оптимальный с точки зрения информационного риска шаг квантования имеет достаточно большое значение. Это

означает, что метод равномерного квантования оптических

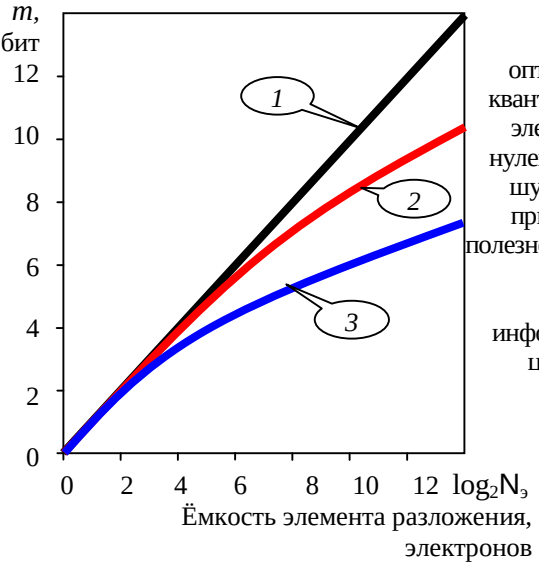


Рис. 3.5. Зависимость оптимальной разрядности квантователя *m* от емкости $N_э$ элемента разложения при нулевой цене загрузки канала шумовой информацией *1*, при равенстве цен потери полезной информации и шумовой информации *2* и при цене шумовой информации в 6 раз большей цены потери полезной информации *3*

сигналов при преобладающем влиянии фотонного мультипликативного шума вступает в противоречие со стремлением проектировщиков одновременно уменьшать ошибки квантования и повышать эффективность использования каналов связи и компьютерной обработки в реальном времени. Поэтому актуален поиск иного локального минимума информационного риска, который обеспечивал бы большую эффективность передачи, чем при равномерном квантовании, но обладал бы его преимуществами с точки зрения сложности компьютерной обработки сигналов.

Разрешение указанного противоречия между стремлением к оптимизации канала, связывающего телекамеру и компьютер, и ограничением сложности последнего лежит на пути применения кусочно-линейной аппроксимации зависимости шага квантования от уровня сигнала [120]. Принципиальным отличием равномерного квантования от оптимального неравномерного квантования является недостаток плотности уровней квантования в области малых сигналов и её избыток в области больших сигналов. В связи с этим становится целесообразным разбиение всей шкалы амплитуд сигналов на участки с тем, чтобы частично устранить этот недостаток. При этом ключом к поиску наилучшей передаточной характеристики кодека связи телекамеры с компьютером

является обеспечение кратности шагов квантования на различных участках. Анализ информационного риска для кусочно-равномерного двухшкального квантования (рис. 3.6) показал, что выигрыш по сравнению с равномерным квантованием имеет место в большинстве практических случаев и делает двухшкальный метод перспективным в измерительных телевизионных системах.

Двухшкальная система может быть реализована различными способами, например, с применением аналоговых или цифровых пороговых устройств; с применением двух АЦП или одного АЦП, преобразующего сигнал с удвоенной частотой элементов. Из соображений температурной и временной стабильности сквозной характеристики кодера предпочтительнее использование одного АЦП, работающего с удвоенной частотой элементов.

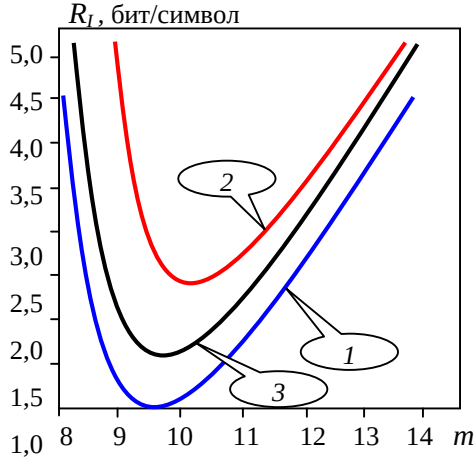


Рис. 3.6. Зависимость информационного риска от числа разрядов при неравномерном квантовании с непрерывно изменяющимся шагом (1), равномерном квантовании (2), и двухучастковом квантовании с шагом квантования первой четверти диапазона в четыре раза меньшим, чем в остальной части диапазона амплитуд сигналов (3) для ёмкости элемента разложения $N_э \approx 6{\cdot}10^4$ электронов

Таким образом, применение критерия минимума информационного риска для формализации равновесного информационного согласования телекамеры с процессором позволяет обосновать разбиение шкалы квантования на участки с разным шагом.

3.3. Адаптивный групповой счет фотонов

Две «вечные» заботы проектировщиков телевизионных систем – борьба за качество информации и качество её кодирования могут быть наглядно продемонстрированы на примере оптимизации совместной процедуры дискретизации и квантования зашумлённых оптических сигналов. При этом борьба за качество информации

предполагает повышение отношения полезной информации к шумовой информации, а качество кодирования оценивается соотношением потери полезной информации и переданной шумовой информации. Положение о зависимости методов кодирования от намерений получателя по дальнейшей обработке принятых сигналов, известно давно. В связи с развитием видеоинформатики благодаря увеличению роли компьютерной обработки передаваемых (и, особенно, принимаемых) сигналов многие решенные ранее теоретические задачи передачи информации переосмысливаются на основе учёта ограничений на сложность и быстродействие, свойственных вычислительным системам принятия решений. При таком согласовании особенно актуален учёт ограничения на сложность вычислителя в системах реального времени, например, в бортовых системах астро-ориентации [122].

Любая телевизионная камера выполняет функцию не просто приёмника света, а функцию приёмопередатчика, цель которого не только извлечь информацию из потока фотонов, но и закодировать её для передачи по видеоканалу с ограниченной пропускной способностью для достижения наименьшей ошибки. Телевизионный счётчик фотонов выдает в видеоканал видеоинформацию, сильно зашумлённую фотонным шумом. Передача этой видеоинформации требует большой пропускной способности канала, значительная часть которой будет загружена избыточной шумовой информацией. Такой счётчик собирает весь поступающий световой поток, но фотоны в телевизионном элементе не накапливает, т. к. на каждый элемент приходится один фотон. При этом возникает гигантский разрыв между кодовой (Терабайты) и полезной (Килобайты) информацией. По этой причине приходится отказываться от режима счёта отдельных фотонов (хотя и стремиться к использованию фотоприёмников с максимальной квантовой эффективностью). Известно, что хотя адаптирующаяся в темноте сетчатка глаза и способна к счёту отдельных фотонов, способность эта не используется при наблюдениях днём.

Понятие счета фотонов был введено А. Роузом для обозначения принципа и средства достижения максимальной чувствительности оптико-электронных систем, и, как он надеялся, максимальной информативности изображений. Адаптация счета фотонов в телевидении традиционно реализуется управлением докоммутационным усилением, которое в электронном телевидении обычно выбиралось на таком уровне, когда заметность шума телекамеры и фотонного шума была примерно одинаковой. Лишь в астрономических системах в режимах наблюдения при слабых

фонах усиление до считывания сигнала выбиралось на уровне, обеспечивающем счёт отдельных фотонов, то есть на уровне, когда шум телекамеры становился пренебрежимо малым по сравнению с фотонным шумом. Современные телекамеры с такой адаптацией строятся на основе сочетания электронно-оптических преобразователей (ЭОП) и матричных ПЗС [47]. Такие телекамеры обладают очень широким динамическим диапазоном, охватывающим и очень слабые световые потоки, которые анализируются путем счёта (измерения координат) отдельных фотонов, и очень интенсивные световые потоки, которые анализируются методом группового счёта фотонов. К сожалению, такие гибридные приборы в силу нежёсткого растра ЭОП не реализуют потенциальной точности измерения координат, и их применение в астрономии ограничено – например, в системах анализа динамических искажений волнового фронта с помощью метода Гартмана. В данной работе делается акцент на другой адаптации, т. е. не на автоматическом переходе от счёта отдельных фотонов к групповому счёту, а на автоматическом управлении характеристиками камеры в рамках группового счёта фотонов. Групповой счёт фотонов, теоретические основы которого изложены в работе [73], при использовании реальных фотоприемников, в том числе матричных ПЗС преобразуется в счёт фотоэлектронов, особенностью которого, в первую очередь, является значение квантового выхода, меньшее 100%. Вместе с тем, термин счёт фотонов обычно распространяют и на счёт фотоэлектронов. Основными характеристиками группового счёта фотонов в матричном фотоприёмнике являются размер пространственной зоны, в которой производится группирование фотоэлектронов и размер группы фотоэлектронов, соответствующей одному уровню квантования амплитуды сигнала в АЦП. Хотя в астрономической телевизионной аппаратуре обычно используют фиксированные значения размера зоны накопления и шага квантования, с помощью которых реализуется обычный (не адаптивный) групповой счёт фотонов, оба эти параметра должны быть адаптивны к уровню сигнала: и интервал дискретизации (раздел 2.1, [69, 70]), и шаг квантования (раздел 2.3, [120]).

Групповой счёт фотонов достаточно просто формализуется в абстрактном случае отсутствия априорной информации о сюжете, когда известно только то, что фотонный поток дискретен и распределён по закону Пуассона [73]. Учёт конкретной функции

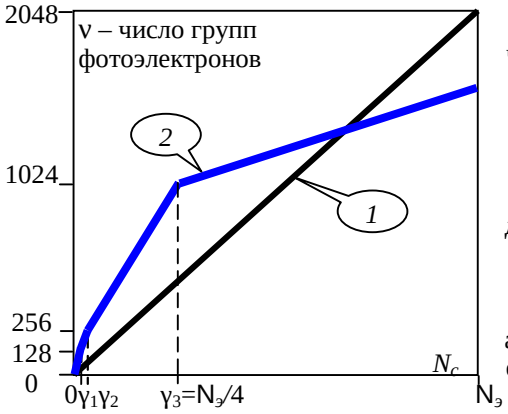


Рис. 3.7. Зависимость числа ν групп фотоэлектронов в отсчете с максимальным откликом от интенсивности сигнала N_c (полного числа фотоэлектронов от точечного объекта) при равномерной дискретизации и равномерном квантовании *1* и при адаптивном групповом счете фотоэлектронов *2* при адаптивном счете фотонов для емкости элемента $N_э = 6{,}4\times10^4$

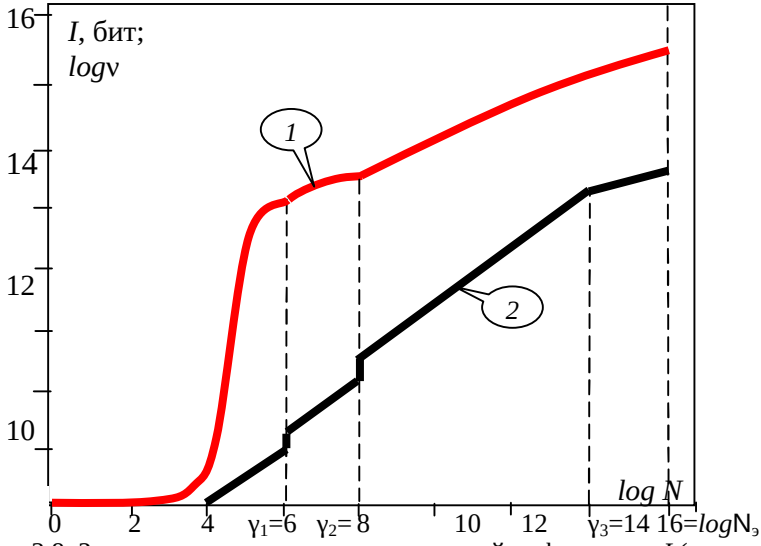


Рис. 3.8. Зависимость количества извлекаемой информации *I* (по одной из координат) при обнаружении–оценивании координаты точечного объекта для различных вариантов накопления *1*, число ν групп фотоэлектронов от звезды *2* от количества *N* сигнальных фотоэлектронов для матрицы с числом элементов на строке 10^3 и накопительной способностью элемента $N_э \approx 64\times10^3$

распределения сигнала приводит к возможности оптимизации квантования, т. е. обоснования размеров групп фотонов.

Наиболее наглядно групповой счёт фотонов формализуется в астрономических приложениях телевизионной техники – задаче

обнаружения и оценивания координат звёзд (точечных объектов) как единство дискретизации и квантования. Для этой задачи и дискретизация входного сигнала (раздел 2.1, рис. 2.8) и их квантование (раздел 3.3, рис. 3.4) при оптимальной процедуре приёма фотонного потока и передаче цифрового кода становятся адаптивными. При этом при низкой плотности потока фотонов наблюдаемой звезды размер зоны накопления больше, чем при высокой плотности, а шаг квантования – наоборот увеличивается при увеличении с ростом плотности потока фотонов (рис. 3.7, 3.8).

Для характеристики группового счёта фотонов полезно ввести численный показатель числа ν групп фотоэлектронов в отсчете с максимальным откликом от интенсивности сигнала N_c (полного числа фотоэлектронов от точечного объекта). В практически реализуемой системе с дискретно изменяющимися размерами зоны накопления и шага квантования число групп фотоэлектронов (так же как и количество информации) при увеличении числа принятых фотонов растёт не линейно. При малой плотности потока фотонов число групп в отсчёте с максимум сигнала растёт быстро, а затем всё медленнее (рис. 3.7).

Как показано в разделе 2.1, система, осуществляющая адаптивный счёт фотонов от звезды, имеет два порога γ_1 и γ_2, при превышении сигналом которых уменьшается размер зоны накопления (уменьшается число складываемых зарядовых пакетов соседних пикселов). Превышение сигналом третьего порогового уровня $\gamma_3 = N_э/4$ (четверти ёмкости элемента разложения) ведёт, как обосновано в разделе 3.3, к переходу к шагу квантования, в 4 раза бóльшему, чем при слабых сигналах. В результате, благодаря адаптации, групповой счёт фотонов обеспечивает согласование скорости создания информации источником и пропускной способности канала связи телекамеры с вычислителем координат, т. е. позволяет извлечь из принимаемого потока информацию о наблюдаемой звезде с минимальными потерями при минимальной загрузке канала шумовой информацией.

3.4. Кодирование источника: сжатие видеоинформации

Целью кодирования источника (сжатия сигналов) является экономия пропускной способности каналов связи и вычислительных ресурсов компьютерных устройств обработки. Потребность в сжатии видеоинформации была осознана давно – достаточно привести два примера. Во-первых, сжатие информации (в соответствии с принятой тогда терминологией – сокращение полосы

частот) составляло цель малокадрового метода передачи телевизионных сигналов, разработанного С. И. Катаевым в 1930-х гг. Во-вторых, намечая стратегию развития телевидения, В. К. Зворыкин в 1954 г. писал: «*Могут быть разработаны методы для передачи сигналов изображения, соответствующих различию между соседними кадрами. Так как это различие обычно содержится лишь в небольшой части общего поля изображения, число элементов изображения, которое должно быть передано для обеспечения необходимой чёткости, может быть таким путём значительно уменьшено. Таким образом, передача сигналов различия изображений может быть использована для значительного уменьшения необходимой полосы видеочастот*». Эта идея была главной при разработке методов кодирования движущихся изображений – MPEG (см. раздел 4.4).

Проблема сжатия изображений, которой десятилетиями занимается большое количество организаций и специалистов, имеет множество решений в связи с разнообразием *критериев качества*, *априорной информации* о сигналах (статистики наблюдаемых сюжетов) и имеющихся *ограничений*. Значение сжатия видеосигнала в телевидении определяется постоянной нехваткой пропускной способности каналов связи. Надежды на реализацию сжатия сигналов связаны с тем, что реальные изображения имеют явно выраженные коррелированные области, т. е. формируемый телекамерой сигнал обладает некоторой *избыточностью*. Иными словами, предпосылкой к возможности сжатия изображений является наличие в любом телевизионном кадре не только полезной, но другой информации – бесполезной фоновой и вредной шумовой.

По месту в системе связи устройство сжатия относится к части кодирующего устройства – кодеру источника. Это устройство дополняет работу *главного устройства сжатия – телекамеры*, которая сжимает поток фотонов до потока элементов разложения. Эффективность сжатия характеризуют *коэффициентом сжатия*, равным отношению скоростей кодов на входе и на выходе устройства сжатия. Коэффициент сжатия сигнала в телекамере отражает групповой счёт фотонов и имеет существенно большее значение, чем коэффициент сжатия последующего кодера. Грубо коэффициент сжатия информации телекамерой может быть оценен по ёмкости элемента разложения твердотельного фотоприёмника, что в среднем составляет десятки и сотни тысяч раз.

Сжатие сигналов подразделяется на два класса – без потерь информации и с потерями информации или, иначе, – на обратимое

и на необратимое кодирование. Получателем видеоинформации в прикладном телевидении является не «телезритель», а эксперт, для которого потеря полезной информации весьма ощутима. Эту ситуацию можно обозначить так: экспериментатор назначает в критерии информационного риска весовые коэффициенты явно в пользу полезной информации, для него некоторое количество шумовой информации не так уж и страшно. Вероятные ошибки при сжатии с потерями трудно предсказуемы, что породило скептическое отношение к сжатию видеосигналов в космическом телевидении, поскольку исходной уникальной информации легко нанести непоправимый вред. Теоретическую основу всех способов сжатия изображений составляют:

- основная теорема Шеннона для канала с шумом, согласно которой можно безошибочно передавать информацию до тех пор, пока энтропия сообщения не превысит пропускной способности канала связи. Это обусловливает постоянство ошибки передачи при увеличении коэффициента сжатия до некоторого критического значения, за которым ошибка монотонно возрастает;
- теория передачи информации с погрешностью (ε-энтропия), устанавливающая зависимость между задаваемой допустимой ошибкой передачи и требуемой скоростью передачи информации;
- метод разложения сигналов по собственным функциям (разложение Карунена–Лоэва), позволяющий преобразовать изображение в совокупность независимых случайных величин и применить теорию передачи с погрешностью для оценки к случайным процессам и полям.
- методы кодирования при ограничении сложности кодирующего устройства, опирающиеся на различные подходы к достижению эпсилон-сложности путём сокращения вычислительной сложности алгоритмов кодирования (рекурсивные функции, модулярные вычислительные структуры, пирамидальные методы и т. д.).

Для источника сигнала, являющегося гауссовской случайной величиной (так называемый «источник без памяти»), квантование является весьма эффективным способом сжатия. Поэтому в случае коррелированного гауссовского источника обычно сначала преобразуют передаваемую последовательность кадров в последовательность некоррелированных символов («выбеливают» сигнал), а затем квантуют. Эта общая идея, заключающаяся в разложении сложного источника на независимые подысточники, кодируемые отдельно друг от друга, служит основным подходом к проектированию практических систем сжатия (раздел 4.4). В результате такого кодирования передаются не все спектральные компоненты

представления исходного изображения, а лишь те, в которых полезная информация превышает шумовую информацию.

То, насколько близко удаётся приблизиться к ε-энтропии в сочетании с многочисленными практическими методами кодирования означает, что на практике чаще всего решают не столько проблему сжатия, сколько проблему преодоления *априорной неопределённости* о статистике передаваемых нестационарных сигналов в совокупности с нечётко сформулированными требованиями к *сложности* кодера. Это часто переводит решение проблемы сжатия из ранга теоретических задач в ранг изобрета-тельских задач (см. гл. 4). Решая эти задачи, проектировщики проявляют своё искусство, используя теоретические основы в качестве некоторого эталона, который не даёт конкретного алгоритма, но с которым можно сравниваться в уровне качества кодирования.

3.5. Модулярные вычислительные структуры обработки видеоинформации

Характер задач, решаемых в реальном времени вычислительными средствами систем на кристалле, предъявляет всё более жёсткие требования к их производительности, причём требования, как правило, недостижимые в рамках классической машинной арифметики и архитектуры фон Неймана. В связи с этим, в последние годы резко активизировались научно-прикладные исследования в области создания высокоскоростных *средств модулярной обработки информации* на основе матричных и параллельно-конвейерных структур СБИС.

Как известно [137, 161, 162], *модулярная машинная арифметика* базируется на применении систем счисления в остаточных классах, относящихся к числу непозиционных.

Системой счисления в остаточных классах (СОК) принято называть [137] систему, в которой целое положительное число A представляется в виде кортежа $(\alpha_1, \alpha_2, \cdots, \alpha_n) \Leftrightarrow A$ наименьших положительных остатков (*вычетов*) от деления A на выбранные *модули* $p_1, p_2, \ldots, p_n$ (основания системы [137]). Замечательное свойство таких систем заключается в том, что любая *модулярная операция* вида $A \circ B = C \ \left(\mathrm{mod}\left(p_1 \cdot p_2 \cdots p_n\right)\right)$ сводится к вычислению элементов результирующего кортежа $(\xi_1, \xi_2, \cdots, \xi_n) \Leftrightarrow C$ путём выполнения аналогичной операции над одноименными

компонентами кортежей $(\alpha_1, \alpha_2, \cdots, \alpha_n) \Leftrightarrow A$ и $(\beta_1, \beta_2, \cdots, \beta_n) \Leftrightarrow B$:

$$\xi_i = \alpha_i \circ \beta_i \pmod{p_i},\ i = \overline{1,n}. \tag{3.16}$$

Не будучи строго формальным, приведённое определение позволяет, тем не менее, понять и оценить фундаментальные преимущества СОК перед позиционными системами счисления.

Основными достоинствами СОК являются:

- глубокий естественный параллелизм вычислений, порождаемый взаимной независимостью вычетов, которые могут обрабатываться при выполнении модулярных операций *параллельными каналами*;
- практически недостижимая в позиционных системах счисления *возможность однотактной табличной реализации* операций при одинаковой длительности тактов как сложения/вычитания, так и умножения чисел;
- «приспособленность» СОК к синхронной конвейерной обработке данных [177];
- отсутствие погрешностей вычислений, характерных для «позиционных» процессоров и обусловленных неравномерностью распределения чисел, задаваемых в форме с плавающей точкой;
- простота представления и рекордно высокая скорость обработки в СОК комплексных чисел, кватернионов, полиномов и других нетривиальных математических объектов [178–180];
- возможность обнаружения и парирования отказов в реконфигурируемых *модулярных вычислительных структурах* (МВС) за счет введения в СОК избыточных модулей [181–183];
- высокая регулярность таблично-матричных МВС, и как следствие, простота топологического проектирования СБИС на их основе.

Вместе с тем, *системам остаточных классов присущи недостатки, часть из которых носит проблемный характер*. Среди них:

- отсутствие эффективных методов деления, поразрядного сравнения, масштабирования и округления чисел, представленных в СОК;
- квадратичная зависимость

$$V_i = O\left(p_i^2\right) \tag{3.17}$$

объема V_i таблиц, применяемых для формирования результирующих вычетов (3.16), от значений используемых модулей p_i;

- трудоёмкость перевода чисел из СОК в позиционные системы счисления и, как следствие, сложность получения так называемых *позиционных характеристик* чисел по их модулярным

кодам (определения знака результата, обнаружения его аддитивного или мультипликативного переполнения, получения абсолютных значений величин и т. п.). Достаточно отметить, что существующие на сегодняшний день методы определения позиционных характеристик имеют временну́ю сложность не ниже

$$O\left(n \log_2 n\right). \tag{3.18}$$

В силу указанных обстоятельств, проводимые более полувека научно-технические изыскания в области модулярной обработки информации фактически всегда были подчинены одной и той же стратегической цели – компенсации недостатков СОК и как можно более полному использованию их преимуществ с учетом достижений микроэлектроники [184–186].

С целью более глубокого анализа отмеченных выше проблем и изложения новых методов их решения введём ряд формальных теоретико-числовых определений.

СОК как алгебраическая структура. Пусть $p_1, p_2, \ldots, p_n$ – попарно взаимно простые натуральные числа и $\prod_{i=1}^{n} p_i = \mathcal{P}$. Согласно китайской теореме об остатках в конечном кольце $\mathbb{Z}_{\mathcal{P}}$ существует единственное решение системы сравнений

$$\begin{cases} x \equiv \alpha_1 \pmod{p_1}, \\ x \equiv \alpha_2 \pmod{p_2}, \\ \quad \ldots\ldots \\ x \equiv \alpha_n \pmod{p_n}. \end{cases}$$

Тогда СОК с заданным *набором модулей* $p_1, p_2, \ldots, p_n$ можно определить как *биекцию*

$$\mathbb{S}\left(\mathcal{P}, n\right) = \mathbb{S}\left(p_1 \cdot p_2 \cdots p_n\right): \mathbb{Z}_{\mathcal{P}} \leftrightarrow \overset{n}{\underset{i=1}{\times}} \mathbb{Z}_{p_i}, \tag{3.19}$$

где $\times$ – символ декартова произведения множеств.

Первичным (определяющим) параметром любой СОК $\mathbb{S}\left(p_1 \cdot p_2 \cdots p_n\right)$ является набор её модулей $\left\{p_1, p_2, \cdots, p_n\right\}$, а вторичными (производными) – порядок $\mathcal{P} = \prod_{i=1}^{n} p_i$ и размерность n. *Порядок* $\mathcal{P}$ системы задает интервал чисел $\left[0, \mathcal{P}\right)$, на котором замыкаются модулярные вычисления. *Размерность* n – количество модулей СОК.

Сложность и сравнимость СОК. *Мерой сложности* системы $\mathbb{S}\left(p_1 \cdot p_2 \cdots p_n\right)$ будем считать сумму квадратов всех её модулей:

$$W\left(\mathbb{S}\right) = \sum_{i=1}^{n} p_i^2 \,. \qquad (3.20)$$

Пусть заданы две системы счисления в остаточных классах $\mathbb{S}\left(p_1 \cdot p_2 \cdots p_n\right)$ и $\mathbb{S}'\left(p'_1 \cdot p'_2 \cdots p'_{n'}\right)$, модули каждой из которых упорядочены по убыванию: $p_1 > p_2 > ... > p_n$; $p'_1 > p'_2 > ... > p'_{n'}$.

Системы $\mathbb{S}\left(p_1 \cdot p_2 \cdots p_n\right)$ и $\mathbb{S}'\left(p'_1 \cdot p'_2 \cdots p'_{n'}\right)$ *линейно сравнимы,* если при $n' \le n$ для всех $i = \overline{1,n'}$ справедливо неравенство $p'_i \le p_i$:

$$\mathbb{S}(p_1 \cdot p_2 \cdots p_n) \succ \mathbb{S}(p'_1 \cdot p'_2 \cdots p'_{n'}) \Leftrightarrow (n \ge n') \wedge \left(\forall i \in \{1,2,...,n'\}, p_i \ge p'_i\right). \qquad (3.21)$$

Очевидно, что отношение «$\succ$» линейной сравнимости СОК несимметрично. Будем называть систему $\mathbb{S}(\mathcal{P},n)$ *доминирующей* над $\mathbb{S}'(\mathcal{P}',n')$, если $\mathbb{S}(\mathcal{P},n) \succ \mathbb{S}'(\mathcal{P}',n')$.

Аксиома 3.1. Из двух линейно сравнимых СОК большей сложностью обладает доминирующая система:

$$\mathbb{S}(\mathcal{P},n) \succ \mathbb{S}'(\mathcal{P}',n') \Rightarrow W\left[\mathbb{S}(\mathcal{P},n)\right] > W\left[\mathbb{S}'(\mathcal{P}',n')\right]. \qquad (3.22)$$

Равносильные СОК. Две системы счисления в остаточных классах $\mathbb{S}\left(\mathcal{P}_1,n_1\right)$ и $\mathbb{S}\left(\mathcal{P}_2,n_2\right)$ *равносильны,* если одинаковы их порядки:

$$\mathcal{P}_1 = \mathcal{P}_2 = \mathcal{P} \Rightarrow \mathbb{S}\left(\mathcal{P},n_1\right) \cong \mathbb{S}\left(\mathcal{P},n_2\right). \qquad (3.23)$$

Пример 3.1. Равносильными СОК являются системы порядка $\mathcal{P} = 210$ с наборами модулей: {2,3,5,7}, {6,5,7}, {3,10,7}, {3,5,14}, {2,15,7}, {2,5,21}, {2,3,35}, {30,7}, {6,35}, {5,42}, {3,70},{10,21}, {14,15}, {2,105}, {210}.

Систему счисления в остаточных классах с единственным модулем $p_1 = \mathcal{P}$ назовем *вырожденной СОК.*

Первообразной СОК $\hat{\mathbb{S}}\left(\mathcal{P},n^*\right)$ порядка $\mathcal{P}$ назовем систему, модули которой $p_1^{e_1}, p_2^{e_2}, \cdots, p_{n^*}^{e_{n^*}}$ являются элементами канонического (Гауссова) разложения числа $\mathcal{P}$ на сомножители: $\mathcal{P} = p_1^{e_1} \bullet p_2^{e_2} \bullet \cdots \bullet p_{n^*}^{e_{n^*}}$, где $p_1, p_2, ..., p_{n^*}$ – простые, $e_1, e_2, ..., e_{n^*}$ – натуральные числа.

Нетрудно заметить, что среди равносильных СОК порядка $\mathcal{P}$ максимальной размерностью n^* обладает первообразная система, а минимальной размерностью – вырожденная СОК с модулем $p_1 = \mathcal{P}$.

Аксиома 3.2. Количество попарно различных равносильных СОК порядка P равно числу разбиений на множестве модулей первообразной системы $\hat{\mathbb{S}}\left(\mathcal{P}, n^*\right)$, т. е. числу Белла B_{n^*} [187].

Аксиома 3.3. Ни одна из равносильных СОК порядка $\mathcal{P}$ не содержит модулей, меньших, чем модули первообразной системы $\hat{\mathbb{S}}\left(\mathcal{P}, n^*\right)$.

Так, в примере 3.1 первообразная система $\hat{\mathbb{S}}\left(2\cdot 3\cdot 5\cdot 7\right) = \hat{\mathbb{S}}\left(210, 4\right)$ имеет размерность $n^* = 4$ и набор минимальных по величине модулей $\{2,3,5,7\}$. Количество равносильных систем порядка $\mathcal{P} = 210$ равно $B_4 = 15$ (одна размерностью 4, шесть размерностью 3, семь размерностью 2 и одна вырожденная с $n = 1$).

Аксиома 3.4. Вырожденная система $\mathbb{S}\left(p,1\right)$ является доминирующей по отношению ко всем равносильным ей системам и, следовательно, обладает по сравнению с ними наибольшей сложностью.

Лемма 3.1. Среди всех равносильных систем порядка $\mathcal{P}$ минимальной сложностью характеризуется первообразная система $\hat{\mathbb{S}}\left(\mathcal{P}, n^*\right)$.

Доказательство. Справедливость леммы вытекает из соотношения $\forall\left(a_i \geq 2\right) \in \mathbb{N},\ \prod_{i=1}^{m} a_i^2 > \sum_{i=1}^{m} a_i^2$, означающего, что переход к любой равносильной системе группировкой модулей первообразной системы сопровождается ростом меры сложности.

Анализ выражений (3.16) – (3.22) и леммы 3.1 показывает, что свойство (3.23) равносильности СОК можно рассматривать как один из наиболее простых и естественных источников оптимизации МВС путём преобразования первоначально заданной (не обязательно первообразной или вырожденной) системы (3.19) в равносильную ей систему, удовлетворяющую определенным ограничениям и являющуюся оптимальной (либо близкой к ней) по выбранному критерию. Возможные варианты постановки оптимизационных задач:

1. *Задача глобальной минимизации объёма таблиц МВС* сводится к разложению на сомножители модулей первоначально заданной СОК $\mathbb{S}\left(\mathcal{P}, n\right)$, т. е. к построению равносильной ей первообразной системы $\hat{\mathbb{S}}\left(\mathcal{P}, n^*\right)$.

2. *Задача локальной минимизация объема таблиц МВС при заданном числе каналов* n°. Данная задача сводится к поиску «наилучшего» n° – блочного разбиения $Q=\left\{B_1,B_1,...,B_{n^{\circ}}\right\}$ на множестве $\left\{p_1^{e_1},p_2^{e_2},\cdots,p_{n^*}^{e_{n^*}}\right\}$ модулей первообразной системы $\widehat{\mathbb{S}}\left(\mathcal{P},n^*\right)$. При этом в качестве претендентов на включение в Q рассматриваются блоки, произведение элементов каждого из которых находится в окрестностях точки $\sqrt[n^{\circ}]{\mathcal{P}}$.

Пример 3.2. Все возможные решения сформулированных выше задач применительно к семейству равносильных систем из примера 3.1 приведены в табл. 3.1.

Таблица 3.1.

Набор модулей СОК	2, 3, 5, 7	6, 5, 7	3, 10, 7	3, 5, 14	2, 15, 7	2, 5, 21	2, 3, 35	30, 7	6, 35	5, 42	3, 70	10, 21	14, 15	2, 105	210
Мера сложности СОК	87	110	158	230	278	470	1238	949	1261	1789	4909	541	421	11029	44100
Число каналов МВС	4	3	3	3	3	3	3	2	2	2	2	2	2	2	1

В примере 3.2 минимальным результирующим объемом таблиц, используемых для формирования вычетов (3.16), характеризуется МВС, которая соответствует первообразной СОК с набором модулей $\{2,3,5,7\}$. Для нее $W\left[\mathbb{S}\left(2\cdot3\cdot5\cdot7\right)\right]=87$. Она же отличается максимальной временной сложностью вычисления позиционных характеристик $O\left(4\log_2 4\right)=8$. В то же время, система с набором оснований $\{14,15\}$ имеет показатели сложности $W\left[\mathbb{S}\left(14\cdot15\right)\right]=421$ и $O\left(2\log_2 2\right)=2$. Иными словами, переход от первообразной СОК $\widehat{\mathbb{S}}\left(2\cdot3\cdot5\cdot7\right)$ к равносильной ей системе $\mathbb{S}\left(14\cdot15\right)$ обеспечивает четырехкратное повышение скорости формирования позиционных характеристик модулярных кодов при примерно пятикратном росте аппаратной сложности МВС.

Рассмотренный пример иллюстрирует основное противоречие, требующее своего разрешения в ходе системного проектирования МВС с табличной реализацией модулярных операций*: при*

сокращении размерности и сохранении порядка применяемых СОК рост объема таблиц опережает повышение скорости вычислений. Очевидно, что *радикальное решение указанной проблемы лежит на пути исключения квадратичной зависимости (3.17) и замены её зависимостью, близкой к линейной.* Исследования авторов настоящей монографии показали, что такая замена осуществима при переходе от классических СОК к *многоступенчатым.*

Многоступенчатые СОК. Первоначальная идея построения многоступенчатых систем счисления в остаточных классах (МСОК) принадлежит, видимо, авторам работы [137]. Позднее она была повторена в [162, 188, 189] со следующим комментарием: «При использовании многоступенчатой СОК обработка данных будет вестись только в высшей ступени, а результат вычисления должен преобразовываться в форму, наиболее предпочтительную для выхода. Результаты вычислений всегда будут правильными, если не было выхода из диапазона представления чисел, как в старшей, так и в нижней ступени».

Однако, к сожалению (и это будет показано далее), «обработка данных только в высшей ступени» МСОК приводит к получению *преимущественно неправильных результатов*, если не предпринимать необходимых мер по предотвращению «выхода из диапазона представления чисел».

Соответствующая процедура введения в допустимые диапазоны всех вычетов, формируемых не только в старшей, но и в других ступенях МСОК, именуется далее *редуцированием вычетов*.

Пусть задана некоторая *главная* СОК $\mathbb{S}(\mathcal{P},n)=\mathbb{S}(p_1\cdot p_2\cdots p_n)$. Тогда *многоступенчатой СОК*, содержащей в качестве главной систему $\mathbb{S}(p_1\cdot p_2\cdots p_n)$, будем называть систему $\overline{\overline{\mathbb{S}}}(\mathcal{P},n,m)$ с числом ступеней $m\ge 2$, в которой вычеты хотя бы по одному из модулей $p_i\in\{p_1,p_2,...,p_n\}$ представляются (*отображаются*) в другой (*возможно также многоступенчатой!*) системе $\mathbb{S}(\mathcal{P}',r)=\mathbb{S}(p_1'\cdot p_2'\cdots p_r')$ с модулями $p_1',p_2',\cdots,p_r'$, *меньшими чем* p_i, размерностью $r\ge 2$ и порядком $\mathcal{P}'=p_1'\cdot p_2'\cdots p_r'$.

Наиболее наглядной формой представления МСОК является *структурное дерево*, пример которого показан на рис. 3.9.

Корень структурного дерева МСОК отождествляется с порядком $\mathcal{P}$ главной системы, узлы *k*-го уровня (на рис. 3.9 изображены «жирными» окружностями) – с транзитными модулями

k-й ступени, а вершины *k*-го уровня – с рабочими 5 модулями *k*-й ступени. При этом *транзитным* считается модуль некоторой системы *k*-й ступени, вычеты по которому представляются многокомпонентными *субкортежами* с помощью одной из *отображающих* систем (*k* + 1)-й ступени. Вычеты по *рабочим* модулям однокомпонентны.

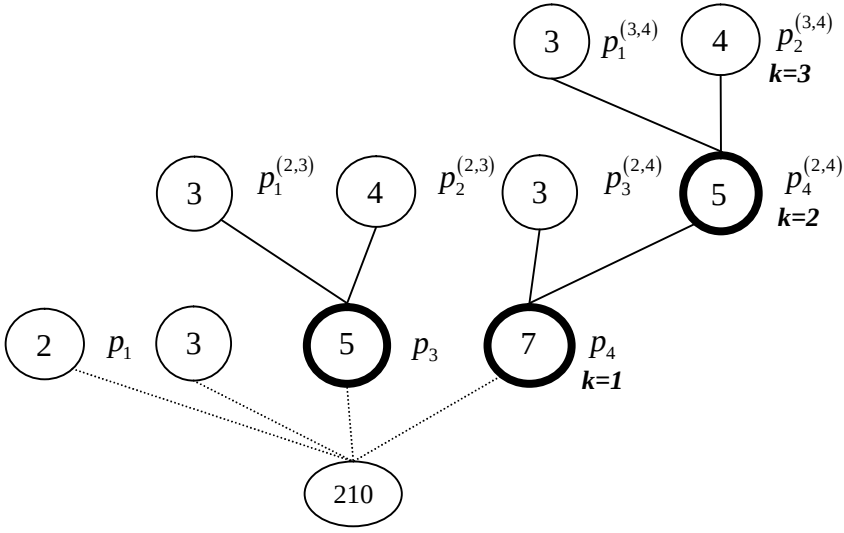


Рис. 3.9. Пример структурного дерева трёхступенчатой СОК

Более компактным представлением МСОК может служить префиксно-скобочная форма. Так, например, МСОК, структурное дерево которой показано на рис. 3.9, имеет префиксно-скобочную форму вида: (2, 3, ^{5}(3, 4), ^{7}(3, ^{5}(3, 4))). Здесь надстрочными префиксами перед скобками обозначены транзитные модули, а числами в строке – рабочие модули МСОК. Уровень вложенности соответствующих скобок определяет номер ступени МСОК.

В общем случае, многоступенчатая СОК $\overline{\overline{\mathbb{S}}}(\mathcal{P},n,m)$ с числом ступеней ≥ 2 описывается алгебраической структурой

$$\overline{\overline{\mathbb{S}}}(\mathcal{P},n,m)=\left\{\mathbb{S}(p_1\cdot p_2\cdots p_n),\ \bigcup_{k=2}^{m}\mathfrak{M}^{(k)}\right\},$$

где $\mathbb{S}(p_1\cdot p_2\cdots p_n)$ – СОК первой ступени (*главная система*);

$\bigcup_{k=2}^{m}\mathfrak{M}^{(k)}$ – множество семейств отображающих систем;

$\mathfrak{M}^{(k)}=\left\{\mathbb{S}\left(\mathcal{P}^{(k,j_{k-1})},n^{(k,j_{k-1})}\right)\middle|\ k=\overline{2,m};\ j_{k-1}=\overline{n'_{k-1}+1,n''_{k-1}}\right\}$ – семейство отображающих систем *k*-й ступени МСОК;

n'_{k-1} – количество рабочих модулей в (*k-1)*-й ступени МСОК;

n''_{k-1} – суммарное количество транзитных модулей (*k-1)*-й ступени, равное числу отображающих систем в *k*-й ступени МСОК;

$\mathcal{P}^{(k,j_{k-1})}$ и $n^{(k,j_{k-1})}\geq 2$ – соответственно порядок и размерность отображающей системы $\mathbb{S}\left(\mathcal{P}^{(k,j_{k-1})},n^{(k,j_{k-1})}\right)$, используемой для представления вычетов по транзитному модулю $p_{j_{k-1}}$.

Если в некоторой МСОК каждый транзитный модуль $p_{j_{k-1}}$, порядок $\mathcal{P}^{(k,j_{k-1})}$ сопоставленной ему отображающей системы $\mathbb{S}\left(\mathcal{P}^{(k,j_{k-1})},n^{(k,j_{k-1})}\right)$ и порядок $\mathcal{P}$ главной системы $\mathbb{S}\left(p_1\cdot p_2\cdots p_n\right)$ связаны соотношением

$$2\left(p_{j_{k-1}}-1\right)\leq\mathcal{P}^{(k,j_{k-1})}<\left(p_{j_{k-1}}-1\right)^2, \qquad (3.24)$$

то такая система называется *многоступенчатой системой аддитивного типа*. Если же для МСОК связь между указанными параметрами имеет вид

$$\left(p_{j_{k-1}}-1\right)^2\leq\mathcal{P}^{(k,j_{k-1})}<\mathcal{P}, \qquad (3.25)$$

то система называется *многоступенчатой системой мультипликативного типа*.

Нетрудно заметить, что на рис. 3.9 представлена МСОК аддитивного типа. Действительно, для неё справедливо:

$2\left(p_3-1\right)<p_1^{(2,3)}\cdot p_2^{(2,3)}<\left(p_3-1\right)^2$, т. е. 8 < 12 < 16;

$2\left(p_4-1\right)<p_3^{(2,4)}\cdot p_4^{(2,3)}<\left(p_4-1\right)^2$, т. е. 12 < 15 < 36;

$2\left(p_4^{(2,4)}-1\right)<p_1^{(3,4)}\cdot p_2^{(3,4)}<\left(p_4^{(2,4)}-1\right)^2$, т. е. 8 < 12 < 16.

Для упрощения дальнейших рассуждений воспользуемся следующими обозначениями:

q – некоторый транзитный модуль МСОК;

${}^{q}\mathbb{S}\left(u_1\cdot u_2\cdots u_z\right)$ – отображающая система, сопоставленная транзитному модулю q, т. е. используемая для представления и покомпонентной обработки вычетов по модулю q;

$U = \prod_{h=1}^{z} u_h$ – порядок отображающей системы ${}^{q}\mathbb{S}(u_1 \cdot u_2 \cdots u_z) = {}^{q}\mathbb{S}(U, z)$;

$\mathrm{Red}(x, q) = |x|_q = x(\mathrm{mod}\ q)$ – операция (функция) редуцирования (взятия, свертывания) числа x по модулю q.

Если транзитному модулю q сопоставлена отображающая система ${}^{q}\mathbb{S}(U, z)$, такая что результат произвольной конечной цепочки модулярных операций над вычетами по модулю q может быть получен выполнением той же цепочки операций в системе ${}^{q}\mathbb{S}(U, z)$ и редуцированием *только последнего из формируемых в ней результатов,* то ${}^{q}\mathbb{S}(U, z)$ будем называть отображающей системой *с возможностью финального редуцирования результатов (ФР-системой).*

Лемма 3.2. Отображающая система ${}^{q}\mathbb{S}(U, z)$ обладает возможностью финального редуцирования результатов тогда и только тогда, когда *q* делит *U*.

Доказательство. Если q делит U, то $U = tq$, где $t \geq 2$ – натуральное число, и $\mathbb{Z}_q$ является фактор-кольцом конечного кольца $\mathbb{Z}_U$. Тогда, согласно теореме о гомоморфизмах [190], группы по сложению $\mathbb{Z}_U^{+}$ и по умножению $\mathbb{Z}_U^{\times}$ гомоморфны соответственно группам $\mathbb{Z}_q^{+}$ и $\mathbb{Z}_q^{\times}$. Следовательно, для любого $x \in \mathbb{Z}_U$ функция редуцирования $\mathrm{Red}(x, q)$ суть гомоморфное отображение $\mathrm{Red}(x, q): \mathbb{Z}_U \to \mathbb{Z}_q$ или, с учетом определения (3.19),

$$\mathrm{Red}(x, q): {}^{q}\mathbb{S}(u_1 \cdot u_2 \cdots u_z) \to \mathbb{S}(q). \qquad (3.26)$$

С другой стороны, если q не делит U, то $\mathbb{Z}_q$ не может быть фактор-кольцом кольца $\mathbb{Z}_U$ и гомоморфного отображения (3.26) не существует. Лемма доказана.

Пример 3.3. Пусть $q = 6$, ${}^{q}\mathbb{S}(u_1 \cdot u_2 \cdots u_z) = {}^{6}\mathbb{S}(3 \cdot 4)$ и необходимо вычислить выражение $y_1 = |3 \cdot 3 + 4 \cdot 2|_6$ над вычетами по модулю $q = 6$. Для вычисления данного выражения непосредственно в системе $\mathbb{S}(6)$ можно использовать цепочку модулярных операций

вида: $|3\cdot 3|_6 = 3 \mapsto |4\cdot 2|_6 = 2 \mapsto |3+2|_6 = 5.$

Аналогичная цепочка операций в отображающей системе ${}^{6}\mathbb{S}(3\cdot 4)$ будет иметь вид:

$$|(0,3)\cdot(0,3)|_{12} = (0,1) \mapsto |(1,0)\cdot(2,2)|_{12} = (2,0) \mapsto |(0,1)+(2,0)|_{12} = (2,1).$$

Полученный в отображающей системе конечный результат – двухкомпонентный субкортеж (2,1) – есть не что иное, как представление *приведенного* (наименьшего положительного) вычета 5, редуцирование которого не требует каких-либо действий, кроме проверки его на принадлежность интервалу $[0,6)$:

$$\mathrm{Red}((2,1),6) = (2,1).$$

Заметим, что промежуточные результаты $(0,1) = 9_{(10)}$ и $(2,0) = 8_{(10)}$ не являются приведенными по модулю 6.

Пример 3.4. Для исходных условий примера 3.3 рассмотрим варианты вычисления выражения $y_2 = |3\cdot 3 + 4\cdot 2 - 2|_6$:

$$|3\cdot 3|_6 = 3 \mapsto |4\cdot 2|_6 = 2 \mapsto |3+2|_6 = 5 \mapsto |5-2|_6 = |5+4|_6 = 3;$$

$$|(0,3)\cdot(0,3)|_{12} = (0,1) \mapsto |(1,0)\cdot(2,2)|_{12} = (2,0) \mapsto |(0,1)+(2,0)|_{12} =$$

$$= (2,1) \mapsto |(2,1)+(1,0)|_{12} = (0,1).$$

В данном случае, поскольку $U/q = 12/6 = 2$, редуцирование конечного результата $(0,1) = 9_{(10)}$ сводится к однократному вычитанию из него модуля 6:

$$\mathrm{Red}((0,1),6) = |(0,1)+(0,2)|_{12} = (0,3).$$

Пример 3.5. Пусть в МСОК мультипликативного типа $q = 11$, ${}^{q}\mathbb{S}(u_1\cdot u_2\cdots u_z) = {}^{11}\mathbb{S}(105,3) = {}^{11}\mathbb{S}(3\cdot 5\cdot 7)$ и необходимо вычислить выражение $y_3 = |9\cdot 10 + 7\cdot 3|_{11}$ над вычетами по модулю $q = 11$. В данном случае $q = 11$ не является делителем $U = 105$. Цепочки модулярных операций в вырожденной системе $\mathbb{S}(11)$ и отображающей системе ${}^{11}\mathbb{S}(3\cdot 5\cdot 7)$ будут иметь вид:

$$|9\cdot 10|_{11} = 2 \mapsto |7\cdot 3|_{11} = 10 \mapsto |2+10|_{11} = 1;$$

$$|(0,4,2)\cdot(1,0,3)|_{105} = (0,0,6) \mapsto |(1,2,0)\cdot(0,3,3)|_{105} = (0,1,0) \mapsto$$

$$\mapsto |(0,0,6)+(0,1,0)|_{105} = (0,1,6).$$

В подтверждение леммы 3.1 полученный в отображающей системе конечный результат $(0,1,6)=6_{(10)}$ приведен по модулю q = 11, *но ошибочен*, так как истинное значение y_3 равно 1. Поскольку ${}^{11}\mathbb{S}(3\cdot5\cdot7)$ не является ФР-системой, восстановление истинного результата из полученного субкортежа $(0,1,6)=6_{(10)}$ невозможно.

Лемма 3.3. Для транзитного модуля q, равного ненулевой степени простого числа, не существует отображающей системы с возможностью финального редуцирования результатов.

Доказательство. По определению многоступенчатой СОК любая ее отображающая система ${}^{q}\mathbb{S}(u_1\cdot u_2\cdots u_z)={}^{q}\mathbb{S}(U,z)$, сопоставленная транзитному модулю q, имеет попарно взаимно простые модули, меньшие чем q ($u_1<q,\ u_2<q,...,u_z<q$). Однако, если $q=p^e$, где p – простое, а e – натуральное числа, то порядок $U=u_1\cdot u_2\cdots u_z$ отображающей системы делится на q без остатка только в том случае, когда q либо является одним из модулей u_h, либо входит в разложение одного из модулей u_h этой системы, т.е. тогда и только тогда, когда $u_h\geq q$, что противоречит определению МСОК. Лемма доказана.

Многоступенчатую СОК, в которой сложность любой отображающей системы ${}^{q}\mathbb{S}(U,z)$ не превышает сложности ни одной из равносильных систем порядка q, назовем *ранжированной МСОК.*

Вполне очевидно, что *применение неранжированных многоступенчатых СОК лишено практического смысла.* Действительно, если существует некоторая система $\mathbb{S}(q,k)\cong\mathbb{S}(q,1)$, такая что $W\left[\mathbb{S}(q,k)\right]\leq W\left[{}^{q}\mathbb{S}(U,z)\right]$, то переход от вырожденной системы $\mathbb{S}(q,1)$ к равносильной ей $\mathbb{S}(q,k)$ предпочтительней, чем использование отображающей системы ${}^{q}\mathbb{S}(U,z)$, поскольку в этом случае модулярные вычисления в равносильной системе $\mathbb{S}(q,k)$ требуют тех же или даже меньших аппаратных затратах и при этом никак не связаны с проблемой редуцирования результатов.

Иллюстрацией сказанному может служить приведенный выше пример 3.3. В этом примере взамен отображающей системы

${}^{6}\mathbb{S}(3\cdot 4)$ с мерой сложности $W\left[{}^{6}\mathbb{S}(3\cdot 4)\right]=9+16=25$ целесообразнее использовать первообразную систему $\widehat{\mathbb{S}}(6,2)=\mathbb{S}(2\cdot 3)$, для которой $W\left[\mathbb{S}(2\cdot 3)\right]=4+9=13$ и выражению $y_1=\left|3\cdot 3+4\cdot 2\right|_6$ соответствует цепочка операций

$$\left|(1,0)\cdot(1,0)\right|_6=(1,0)\mapsto\left|(0,1)\cdot(0,2)\right|_6=(0,2)\mapsto\left|(1,0)+(0,2)\right|_6=(1,2).$$

Теорема 3.1. Для реализации в ранжированной МСОК безошибочных модулярных вычислений необходимо редуцирование всех результатов в каждой из отображающих систем.

Доказательство. Согласно лемме 3.2 промежуточные результаты можно не редуцировать, если они получены в отображающей ФР-системе ${}^{q}\mathbb{S}(U,z)$, для которой $U=tq$, где $2\le t<q$ – натуральное число. Одновременно, из леммы 3.3 следует, что, если ${}^{q}\mathbb{S}(U,z)$ – ФР-система, то q – составное число.

Пусть $q=p_1^{e_1}\cdot p_2^{e_2}\cdots p_h^{e_h}$ – *расширенное каноническое разложение* числа q, содержащее степени полного ряда простых чисел, начиная от p_1, ближайшего к q снизу, и заканчивая $p_h=2$. В таком разложении отдельные показатели степеней могут быть нулями. Пусть, далее, аналогичное разложение для натурального числа $t<q$ имеет вид: $t=p_1^{g_1}\cdot p_2^{g_2}\cdots p_h^{g_h}$.

Тогда $U=tq=p_1^{g_1+e_1}\cdot p_2^{g_2+e_2}\cdots p_h^{g_h+e_h}$, причем

$$\left(\forall i\in\{1,2,...,h\}\colon p_i^{g_i+e_i}\ge p_i^{e_i}\right)\wedge\left(\exists j\in\{1,2,...,h\}\colon p_i^{g_j+e_j}>p_i^{e_j}\right)\quad(3.27)$$

С учетом этого, сложность первообразных систем $\widehat{\mathbb{S}}(q,h-b')$ и ${}^{q}\widehat{\mathbb{S}}(U,h-b'')$ будет определяться выражениями

$$W\left[\widehat{\mathbb{S}}(U,h-b')\right]=\sum_{i=1}^{h}\left(p_i^{e_i}\right)^2-b',\qquad(3.28)$$

$$W\left[{}^{q}\widehat{\mathbb{S}}(U,h-b'')\right]=\sum_{i=1}^{h}\left(p_i^{g_i+e_i}\right)^2-b'',\qquad(3.29)$$

где b' и b'' – количество единичных сомножителей в расширенном каноническом разложении соответственно числа q и числа $U=tq$.

Поскольку $U=tq$, то справедливо неравенство $b'\ge b''$, в связи с чем из выражений (3.27) – (3.29) вытекает соотношение

$$W\left[\widehat{\mathbb{S}}\left(q,h-b'\right)\right]<W\left[{}^{q}\widehat{\mathbb{S}}\left(U,h-b''\right)\right]. \qquad (3.30)$$

Согласно лемме 3.1 среди всех равносильных систем порядка U минимальной сложностью обладает первообразная ${}^{q}\widehat{\mathbb{S}}\left(U,h-b''\right)$, но в соответствии с неравенством (3.30) даже она сложнее первообразной системы $\widehat{\mathbb{S}}\left(q,h-b'\right)$, что противоречит определению ранжированной МСОК.

Таким образом, если МСОК содержит хотя бы одну отображающую ФР-систему, то такая МСОК не относится к классу ранжированных. Справедливо и обратное – ранжированная МСОК не может содержать ни одной отображающей ФР-системы, а посему в ней *для исключения возможных ошибок должны подвергаться редуцированию не только конечные, но и все промежуточные результаты вычислений в каждой из отображающих систем*. Теорема доказана.

Редуцирование результатов в МСОК. Операция $\mathrm{Red}\left(x,q\right)=\left|x\right|_{q}=x\left(\mathrm{mod}\ q\right)$ сводится к нахождению *остатка от деления* x на q: $\mathrm{Red}\left(x,q\right)=x-\left[x/q\right]\cdot q$. Сложность её реализации применительно к многоступенчатым СОК заключается в том, что соответствующие *действия должны выполняться в отображающих системах.*

При модулярном сложении чисел сумма двух вычетов α и β по транзитному модулю q находится в интервале $0\le\alpha+\beta\le 2\left(q-1\right)$ и редуцирование результата возможно в соответствии с выражением

$$\mathrm{Red}\left(\alpha+\beta,q\right)=\begin{cases}\alpha+\beta, & \alpha+\beta<q\\ \alpha+\beta-q, & \alpha+\beta\ge q\end{cases} \qquad (3.31)$$

В этом случае основная трудность заключается в сравнении чисел $\left(\alpha+\beta\right)$ и q, каждое из которых представлено своим z-компонентным субкортежем вычетов в отображающей системе ${}^{q}\mathbb{S}\left(u_1\cdot u_2\cdots u_z\right)$. Известные многотактные алгоритмы [191] сравнения чисел в СОК характеризуются временно́й сложностью $O\left(z\right)$ и поэтому непригодны для применения в быстродействующих МВС. Гораздо более эффективным является ассоциативно-табличный метод редуцирования, варианты реализации которого представлены на рис. 3.10 , 3.11.

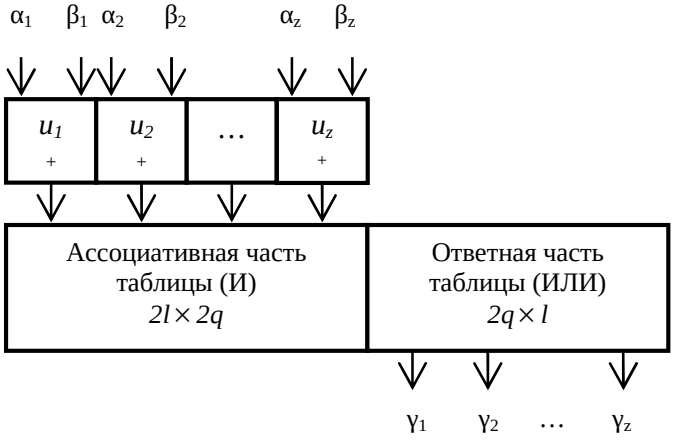


Рис. 3.10. Редуцирование суммы вычетов полной подстановкой

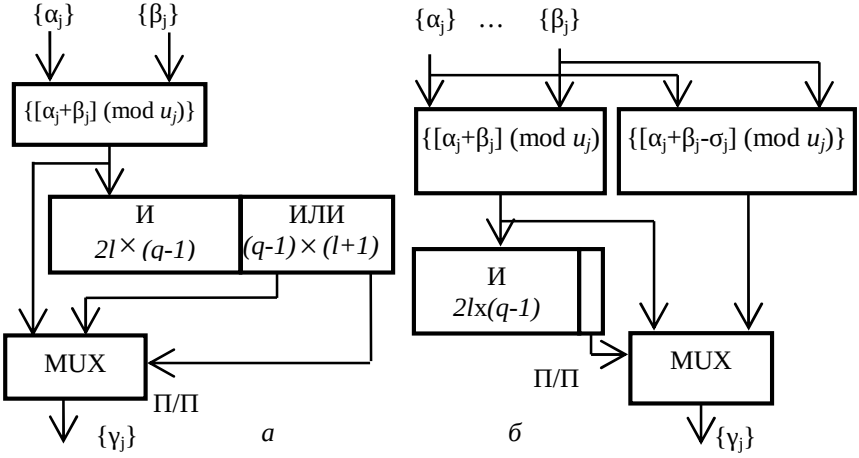


Рис. 3.11. Редуцирование суммы вычетов частичной подстановкой с применением одной *а* и двух *б* групп сумматоров

В схеме на рис. 3.10 с помощью ассоциативной таблицы производится замена *каждого* из $(2q-1)$ l-битовых субкортежей, среди которых $q-2$ соответствуют нередуцированным, а q – редуцированным значениям суммы вычетов $\alpha+\beta$. В качестве подстановочной таблицы может использоваться ассоциативное

запоминающее устройство (АЗУ) или программируемая логическая матрица (ПЛМ) [192].

В схеме на рис. 3.11а с помощью ассоциативной таблицы осуществляется обнаружение и преобразование только тех субкортежей, которым соответствуют значения суммы вычетов $q \le \alpha + \beta \le 2(q-1)$. В случае $0 \le \alpha + \beta \le q-1$ сформированный субкортеж по сигналу П/П («Попадание/Промах») снимается непосредственно с выходов группы табличных сумматоров.

В схеме на рис. 3.11б ассоциативная таблица используется лишь для проверки принадлежности суммы вычетов $\alpha + \beta$ диапазону $[q, 2(q-1)]$.

С помощью альтернативной группы сумматоров вычисляется величина $\alpha + \beta - q$. Редуцированный результат снимается с выходов основной или альтернативной группы сумматоров в зависимости от значения сигнала П/П.

При модулярном умножении чисел произведение $\alpha \cdot \beta$ двух вычетов по транзитному модулю q находится в диапазоне $[0, (q-1)^2]$. В связи с этим, редуцирование результатов-произведений в отображающих системах так же сложно, как и деление чисел, а применение ассоциативно-табличных методов не дает положительного эффекта. Так, в схеме, аналогичной приведённой на рис. 3.10, требуется z табличных умножителей общим объёмом $W' = \sum_{j=1}^{z} u_z^2$ ячеек и ассоциативная редуцирующая таблица с числом ячеек (строк), равным $W'' = (q-1)^2 - q$. При этом, в силу квадратичной зависимости W' от u_j и W'' от q, как правило, выполняется соотношение $W' + W'' \ge q^2$.

В качестве иллюстрации сказанному обратимся к примеру 3.5, где $q = 11$, ${}^{q}\mathbb{S}(u_1 \cdot u_2 \cdots u_z) = {}^{11}\mathbb{S}(3 \cdot 5 \cdot 7)$. Для него: $W' = 3^2 + 5^2 + 7^2 = 83$; $W'' = 10^2 - 11 = 89$; $W' + W'' = 172$. В то же время, таблица умножения вычетов непосредственно по модулю $q = 11$ содержит всего лишь 121 строку.

Таким образом, применение многоступенчатых СОК мультипликативного типа, удовлетворяющих условию (3.25), не снимает проблемы квадратичной зависимости объема таблиц от значений используемых модулей.

С учётом данного вывода, будем рассматривать далее только аддитивные МСОК. При этом будем понимать под *машинной арифметикой в системе остаточных классов*, модулярную арифметику, использующую представление чисел дополнительными кодами (в искусственной форме со смещением нуля). В работе [137] показаны необходимые и достаточные условия функциональной полноты такой арифметики относительно модулярных операций алгебраического сложения, вычитания и умножения чисел:

1*) наличие в системе $\mathbb{S}(p_1 \cdot p_2 \cdots p_n)$ модуля $p_1 = 2$;

2*) возможность выполнения поразрядных операций $\gamma_i = |\alpha_i + \beta_i|_{p_i}$ и $\lambda_i = |\alpha_i \cdot \beta_i|_{p_i}$, где $\alpha_i, \beta_i \in \{0,1,...,p_i - 1\}$, $i = \overline{2,n}$;

3*) формирование вычета γ_1 суммы и вычета λ_1 произведения дополнительных кодов чисел по булевым формулам $\gamma_1 = \alpha_1 \,\&\, \beta_1 \vee \overline{\alpha}_1 \,\&\, \overline{\beta}_1$ и $\lambda_1 = \overline{\alpha}_1 \,\&\, \overline{\beta}_i$.

При выполнении данных условий алгебраическое вычитание чисел сводится к изменению знака вычитаемого на противоположный и последующему сложению кодов. Для изменения знака числа X на противоположный используется равенство $(-X)_{\text{дк}} = X_{\text{дк}} \cdot (-1)_{\text{дк}}$.

Теорема 3.2. Табличная машинная арифметика в ранжированной МСОК аддитивного типа обладает функциональной полнотой относительно модулярных операций алгебраического сложения, вычитания и умножения, если модули главной системы $\mathbb{S}(p_1 \cdot p_2 \cdots p_n)$ – простые положительные числа и $p_1 = 2$.

Доказательство. Модуль $p_1 = 2$ не может быть транзитным, поэтому условие 3*) выполнимо, независимо от особенностей МСОК.

Пусть простое положительное число q>2 – один из модулей главной системы. S($p_1 \cdot p_2 \cdots p_n$). Тогда [190, 193] существует примитивный элемент $g \in \breve{\mathbb{Z}}_q^{\times}$, порождающий циклическую мультипликативную группу $\breve{\mathbb{Z}}_q^{\times}$, такую что

$$\forall \alpha, \beta \in \breve{\mathbb{Z}}_q^{\times}: \left(\alpha = \left|g^{ind(\alpha)}\right|_q, \beta = \left|g^{ind(\beta)}\right|_q\right) \wedge \left(ind(\alpha), ind(\beta) \in \mathbb{Z}_{q-1}^{+}\right) \wedge \left(\alpha \neq \beta \Rightarrow ind(\alpha) \neq ind(\beta)\right)$$

При этом группы $\breve{\mathbb{Z}}_q^{\times} = \{1,2,...,q-1\}$ и $\mathbb{Z}_{q-1}^{+} = \{0,1,...,q-2\}$

изоморфны, а их взаимно однозначное отображение имеет вид: $\left|ind\left(\alpha\right)+ind\left(\beta\right)\right|_{q-1}=ind\left(\left|\alpha\cdot\beta\right|_q\right)$ или

$$\left|\alpha\cdot\beta\right|_q=\left|g^{\left|ind(\alpha)+ind(\beta)\right|_{q-1}}\right|_q. \qquad (3.32)$$

Предположим далее, что q – транзитный модуль, которому сопоставлена отображающая система ${}^q\mathbb{S}\left(U,z\right)$ второй ступени МСОК. Тогда, в силу определений (3.24) и (3.31), сумма $a+b$ любых двух целых $a,b\in\left\{0,1,...,q-1\right\}$ имеет единственное представление в отображающей системе ${}^q\mathbb{S}\left(U,z\right)$, при условии, что промежуточные результаты во всех отображающих системах третьей и более высоких ступеней МСОК редуцированы.

Если для представления неопределенного значения $ind\left(0\right)$ в отображающей системе ${}^q\mathbb{S}\left(U,z\right)$ использовать уникальный субкортеж, например, $\overrightarrow{\left(ind\left(0\right)\right)}=\left(u_1,u_2,...,u_z\right)$, и полагать $ind\left(0\right)+ind\left(x\right)=ind\left(0\right)$, то из равенства (3.32) и единственности представления величин $\left(a+b\right)\in\left[0,2q-2\right)$ z-компонентными субкортежами $\overrightarrow{\left(a+b\right)}=\left(\delta_1,\delta_2,...,\delta_z\right)$ вытекает однозначность следующих табличных преобразований:

$$\overrightarrow{\left(\alpha+\beta\right)}\xrightarrow{f_1(\alpha,\beta,q)}\overrightarrow{\left|\alpha+\beta\right|_q}=\vec{\gamma};$$

$$\overrightarrow{\left(\alpha+\beta\right)}\xrightarrow{f_2(\alpha,\beta,q)}\overrightarrow{ind\left(\left|\alpha+\beta\right|_q\right)}=\overrightarrow{ind\left(\gamma\right)};$$

$$\overrightarrow{\left(ind\left(\alpha\right)+ind\left(\beta\right)\right)}\xrightarrow{f_3(\alpha,\beta,q)}\overrightarrow{\left|\alpha\cdot\beta\right|_q}=\vec{\lambda};$$

$$\overrightarrow{\left(ind\left(\alpha\right)+ind\left(\beta\right)\right)}\xrightarrow{f_4(\alpha,\beta,q)}\overrightarrow{ind\left(\left|\alpha\cdot\beta\right|_q\right)}=\overrightarrow{ind\left(\lambda\right)};$$

$$\overrightarrow{\left(\alpha+\beta\right)}\xrightarrow{f_5(\alpha,\beta,q)}\vec{\tilde{\gamma}};$$

$$\overrightarrow{\left(\alpha+\beta\right)}\xrightarrow{f_6(\alpha,\beta,q)}\overrightarrow{ind\left(\tilde{\gamma}\right)};$$

$$\overrightarrow{\left(ind\left(\alpha\right)+ind\left(\beta\right)\right)}\xrightarrow{f_7(\alpha,\beta,q)}\vec{\tilde{\lambda}};$$

$$\overrightarrow{\left(ind\left(\alpha\right)+ind\left(\beta\right)\right)}\xrightarrow{f_8(\alpha,\beta,q)}\overrightarrow{ind\left(\tilde{\lambda}\right)}.$$

При этом поразрядной операции сложения вычетов в главной СОК (см. условие 2*)) соответствует эквивалентное преобразо-

вание f_1 в МСОК, а операции умножения вычетов – суперпозиция преобразований f_2 и f_3:

$$\left|\alpha_i+\beta_i\right|_{p_i} \Leftrightarrow f_1(\alpha_i,\beta_i,p_i)=\mathrm{Red}(\alpha_i+\beta_i,p_i);$$

$$\left|\alpha_i\cdot\beta_i\right|_{p_i} \Leftrightarrow f_3\left(f_2(\alpha_i,0,p_i),f_2(0,\beta_i,p_i),p_i\right).$$

Таким образом, функциональная полнота модулярной машинной арифметики в МСОК аддитивного типа обеспечивается сложением вычетов по простым модулям или индексов таких вычетов с применением табличных преобразований f_1, f_2, f_3. Теорема доказана.

Заметим, что функционально полный набор табличных преобразований $\{f_1, f_2, f_3\}$ минимален, а все множество преобразований $\{f_1, f_2, ..., f_8\}$ избыточно с точки зрения функциональной полноты. *Выбор конкретного набора табличных преобразований является одной из главных оптимизационных задач системного проектирования многоступенчатых МВС*, от решения которой существенно зависят: а) объем памяти, необходимой для хранения исходных данных и промежуточных результатов модулярных вычислений; б) аппаратная сложность каналов МВС и, в первую очередь, занимаемая таблицами площадь на кристалле СБИС; в) производительность МВС. К сожалению, ввиду значительного объема материала, обзор методов решения указанной задачи остается за рамками настоящей монографии.

Пример 3.6. Пусть задана многоступенчатая аддитивная СОК вида $\left(2,7,{}^{11}(3,7)\right)$ с главной системой $\mathbb{S}(2\cdot 7\cdot 11)$ и представлением чисел дополнительными кодами: $\forall X\in[-76,+76]$: $\mathrm{X}_{\text{дк}}=X+7\cdot 11=X+77$. Код каждого числа X хранится в памяти МВС в виде пары кортежей: кортежа вычетов $\overrightarrow{\chi_i}$ и кортежа индексов $\overrightarrow{ind(\chi_i)}$. Для обработки вычетов по модулю 11 и их индексов используется канал $\left\langle \mathrm{m}^{11}(3,7)\right\rangle$, общий вид которого представлен на рис. 3.10. Функционирование сумматоров отображающей системы ${}^{11}\mathbb{S}(3\cdot 7)$ описывается двумя таблицами сложения (табл. 3.2, 3.3).

Индексы вычетов по модулю 11, соответствующие примитивному элементу $g=2$, представлены в табл. 3.4. Преобразования $f_1, f_2, ..., f_8$ осуществляются согласно табл. 3.5.

Таблица 3.2

a / b	**0**	**1**	**2**	**3**
0	0	1	2	3
1	1	2	0	3
2	2	0	1	3
3	3	3	3	3

Таблица 3.3

a / b	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
0	0	1	2	3	4	5	6	7
1	1	2	3	4	5	6	0	7
2	2	3	4	5	6	0	1	7
3	3	4	5	6	0	1	2	7
4	4	5	6	0	1	2	3	7
5	5	6	0	1	2	3	4	7
6	6	0	1	2	3	4	5	7
7	7	7	7	7	7	7	7	7

Таблица 3.4

$\chi_i = \left\|2^{ind(\chi_i)}\right\|_{11}$	0	1	2	3	4	5	6	7	8	9	10
$ind(\chi_i)$	$\wr$	0	1	8	2	4	9	7	3	6	5

Неопределённое значение индекса $ind(0) = \wr$ задается в отображающей системе ${}^{11}\mathbb{S}(3\cdot 7)$ субкортежем вида $(3,7)$.

Рассмотрим порядок работы канала $\left\langle \mathrm{m}^{11}(3,7)\right\rangle$ на примере вычисления выражения $Y = A-(B+C\cdot D)\cdot E+B$ при $A=72$; $B=-12$; $C=3$; $D=4$; $E=2$.

Указанным выше исходным данным соответствуют пары кортежей:

$$A_{\text{дк}} = 149_{(10)} \Leftrightarrow \overrightarrow{\chi_i} = (1;2,(0,6)); \ \overrightarrow{ind(\chi_i)} = (2,(0,2));$$

$$B_{\text{дк}} = 65_{(10)} \Leftrightarrow \overrightarrow{\chi_i} = (1;2,(1,3)); \ \overrightarrow{ind(\chi_i)} = (2,(2,5));$$

$$C_{\text{дк}} = 80_{(10)} \Leftrightarrow \overrightarrow{\chi_i} = (0;3,(0,3)); \ \overrightarrow{ind(\chi_i)} = (1,(2,1));$$

$$D_{\text{дк}} = 81_{(10)} \Leftrightarrow \overrightarrow{\chi_i} = (1;4,(1,4)); \ \overrightarrow{ind(\chi_i)} = (4,(2,2));$$

$$E_{\text{дк}} = 79_{(10)} \Leftrightarrow \overrightarrow{\chi_i} = (1;2,(2,2)); \ \overrightarrow{ind(\chi_i)} = (2,(1,1)).$$

Таблица 3.5

Номер строки	Сумма вычетов или индексов	Субкортеж суммы	f_1	f_2	f_5	f_6	f_3	f_4	f_7	f_8
1	0	**0,0**	0,0	3,7	0,0	3,7	1,1	0,0	1,3	2,5
2	1	**1,1**	1,1	0,0	1,3	2,5	2,2	1,1	0,2	0,6
3	2	**2,2**	2,2	1,1	0,2	0,6	1,4	2,2	1,0	1,0
4	3	**0,3**	0,3	2,1	2,1	0,3	2,1	0,3	0,3	2,1
5	4	**1,4**	1,4	2,2	1,0	1,0	2,5	1,4	0,6	0,2
6	5	**2,5**	2,5	1,4	0,6	0,2	1,3	2,5	1,1	0,0
7	6	**0,6**	0,6	0,2	2,5	1,4	0,2	0,6	2,2	1,1
8	7	**1,0**	1,0	1,0	1,4	2,2	1,0	1,0	1,4	2,2
9	8	**2,1**	2,1	0,3	0,3	2,1	0,3	2,1	2,1	0,3
10	9	**0,2**	0,2	0,6	2,2	1,1	**0,6**	**0,2**	**2,5**	**1,4**
11	10	**1,3**	**1,3**	**2,5**	**1,1**	**0,0**	1,1	0,0	1,3	2,5
12	11	**2,4**	0,0	3,7	0,0	3,7	2,2	1,1	0,2	0,6
13	12	**0,5**	1,1	0,0	1,3	2,5	1,4	2,2	1,0	1,0
14	13	**1,6**	2,2	1,1	0,2	0,6	2,1	0,3	0,3	2,1
15	14	**2,0**	0,3	2,1	2,1	0,3	2,5	1,4	0,6	0,2
16	15	**0,1**	1,4	2,2	1,0	1,0	1,3	2,5	1,1	0,0
17	16	**1,2**	2,5	1,4	0,6	0,2	0,2	0,6	2,2	1,1
18	17	**2,3**	0,6	0,2	2,5	1,4	1,0	1,0	1,4	2,2
19	18	**0,4**	1,0	1,0	1,4	2,2	0,3	2,1	2,1	0,3
20	19	**1,5**	2,1	0,3	0,3	2,1				
21	20	**2,6**	0,2	0,6	2,2	1,1				
22	≀	**3,7**					0,0	3,7	0,0	3,7

При вычислении $Y = 72 - (-12 + 3 \cdot 4) \cdot 2 - 12 = 60_{(10)}$ в канале $\left\langle \mathrm{m}^{11}(3,7) \right\rangle$ будет выполнена следующая цепочка действий:

$$C \cdot D \Rightarrow (2,1) + (2,2) = (1,3) \xrightarrow{f_3} (1,1);$$

$$-(B + C \cdot D) \Rightarrow (1,3) + (1,1) = (2,4) \xrightarrow{f_6} (3,7);$$

$$(-(B + C \cdot D)) \cdot E \Rightarrow (3,7) + (1,1) = (3,7) \xrightarrow{f_3} (0,0);$$

$$A + (-(B + C \cdot D)) \cdot E \Rightarrow (0,6) + (0,0) = (0,6) \xrightarrow{f_1} (0,6);$$

$$\left(A+\left(-(B+C\cdot D)\right)\cdot E\right)+B \Rightarrow (0,6)+(1,3)=(1,2)\left\langle \begin{array}{l} \xrightarrow{f_1}(2,5); \\ \xrightarrow{f_2}(1,4). \end{array} \right.$$

Полученные субкортежи (2,5) и (1,4) суть представление вычета и индекса вычета по модулю 11 искомого результата $Y_{\text{дк}} = 60+77 = 137_{(10)} \Leftrightarrow \overrightarrow{\chi_i} = \left(1;2,(2,5)\right);\ \overrightarrow{ind\left(\chi_i\right)} = \left(2,(1,4)\right).$

Для вычисления данного результата потребовалось 5 тактов.

Расширенный набор преобразований, приведенный в табл. 3.5, обеспечивает максимальную скорость вычислений за счет однотактной реализации выражений $A+B$, $-\left(A+B\right)$, $A\cdot B$, $-\left(A\cdot B\right)$ и $-\left(X\right)=-\left(X+0\right)=X\cdot(-1)$, а также за счёт одновременного формирования и хранения в памяти вычетов и их индексов.

В случае применения минимального набора преобразований f_1, f_2, f_3 и хранения в памяти МВС только вычетов, структура канала упрощается. Ёмкость ответной части ассоциативной таблицы (см. рис. 3.10) сокращается примерно с $2q\times 8l$ бит до $2q\times 3l$ бит, т.е. более чем в 2,5 раза. Объём памяти, необходимой для хранения операндов и результатов, уменьшается примерно вдвое. Вместе с тем, производительность МВС падает, поскольку операция умножения становится трехтактной (два такта на преобразование вычетов в индексы и один такт на сложение индексов с преобразованием полученной суммы в произведение). Кроме того, становится невозможной непосредственная реализация выражений $-\left(A+B\right)$, $-\left(A\cdot B\right)$ и $-\left(X\right)=-\left(X+0\right)$, а для замены знака числа X требуется умножение $X\cdot(-1)$ длительностью не менее двух тактов. Очевидно, что *изменение производительности МВС будет существенным образом зависеть от частотного спектра выполняемых операций, т.е. от характера реализуемых вычислений.*

Для рассмотренного примера 3.6 минимальная сложность канала $\left\langle \mathrm{m}^{11}(3,7)\right\rangle$ будет определяться суммарной ёмкостью двух ПЗУ и одной ПЛМ: табличный сумматор по модулю 3 – $V_{\mathrm{m}3}^{\text{ПЗУ}} = 2^4\times 2 = 32$ бит; табличный сумматор по модулю 7 – $V_{\mathrm{m}7}^{\text{ПЗУ}} = 2^6\times 3 = 192$ бит; ассоциативная таблица, осуществляющая преобразования f_1, f_2, f_3, – $V_f^{ПЛМ} = 10\times 22 + 22\times 15 = 550$ бит; итого – 774 бит. При прямой табличной реализации операций сложения и умножения по модулю

11 требуются два ПЗУ общим объёмом $V_{\text{m11}}^{\text{ПЗУ}} = \left(2^8 \times 4\right) \times 2 = 2048$ бит или одна ПЛМ ёмкостью $V_{\text{m11}}^{ПЛМ} = 16 \times 121 + 121 \times 8 = 2804$ бит.

Как видно из примера 3.6, даже для небольшого модуля *q = 11* применение многоступенчатой СОК даёт примерно трёхкратное сокращение аппаратных затрат в канале табличной МВС.

Для получения более общих оценок рассмотрим схему двухступенчатого канала $\left\langle \text{m}^{\,q}(15,14,...) \right\rangle$, представленную на рис. 3.12.

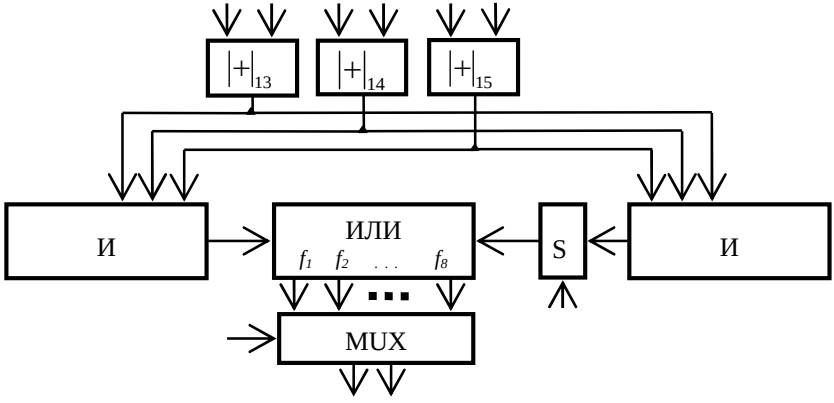


Рис. 3.12. Канал $\left\langle \text{m}^{\,q}(15,14,...) \right\rangle$ двухступенчатой МВС с реструктуризированной ассоциативной таблицей

Её особенность состоит в использовании реструктури-зированной таблицы преобразований. Ассоциативная часть таблицы (матрица И-термов) разделена на две половины, подключенные к ответной части (матрице ИЛИ-термов) способом «Монтажное ИЛИ». Каждая из трёх матриц содержит ровно *q* строк. Общая ёмкость матриц составляет $V_{\text{mq}}^{ПЛМ} = \left(2l \times q\right) + \left(q \times k_f l\right) + \left(2l \times q\right) = lq\left(4 + k_f\right)$, где l – суммарная длина (разрядность) кортежей, формируемых в отображающей системе ${}^{q}\mathbb{S}\left(u_1 \cdot u_2 \cdots\right) = {}^{q}\mathbb{S}\left(15 \cdot 14 \cdots\right)$, а $3 \le k_f \le 8$ – количество используемых преобразований из множества $\left\{f_1, f_2, ..., f_8\right\}$. Одна из матриц И-термов подключается к матрице ИЛИ-термов через сдвигатель S. Такая организации ассоциативной таблицы использует повторяемость её частей относительно строки

с номером q при сложении и строки с номером $q-1$ при умножении чисел (см. табл. 3.5). При этом сложность матрицы ИЛИ-термов сокращается вдвое.

Каждый табличный сумматор $|+|_{u_j}$ при $u_j \in \{13,14,15\}$ может быть выполнен в виде типового блока ПЗУ ёмкостью 256 четырехразрядных слов (1 Кбит). Ассоциативная таблица преобразований может быть построена на основе программируемых матриц типа «И–НЕ», «ИЛИ–НЕ» или реконфигурируемых модулей памяти [194].

Параметры двух ранжированных МСОК аддитивного типа, использующих в старшей ступени модули $u_j \in \{13,14,15\}$, представлены в табл. 3.6.

Таблица 3.6

Отображающая СОК / Параметры МСОК	${}^{q}\mathbb{S}(15\cdot 14)$	${}^{q}\mathbb{S}(15\cdot 14\cdot 13)$
Мера сложности отображающей СОК	$15^2+14^2=421$	$15^2+14^2+13^2=590$
Наименьший простой транзитный модуль главной системы, удовлетворяющий условию ранжированности МСОК	$q_{min}=23$	$q_{min}=29$
Порядок отображающей системы	$\mathcal{P}=210$	$\mathcal{P}=2730$
Наибольший простой транзитный модуль главной системы, удовлетворяющий условию аддитивности МСОК	$q_{max}=103$	$q_{max}=1361$
Тип двухступенчатого канала МВС	$\left\langle \mathrm{m}^{\,q}(15,14)\right\rangle$	$\left\langle \mathrm{m}^{\,q}(15,14,13)\right\rangle$
Емкость ассоциативной таблицы в канале [бит]	$8q\left(4+k_f\right)$	$12q\left(4+k_f\right)$
Суммарная сложность канала	$8q\left(4+k_f\right)+2\cdot 2^{10}$	$12q\left(4+k_f\right)=3\cdot 2^{10}$

На рис. 3.13, 3.14 показаны графики, характеризующие *отношение* аппаратной сложности одноступенчатых каналов $\langle \mathrm{m}\, q\rangle$ к сложности замещающих их двухступенчатых каналов, представленных в табл.3.6.

Расчёт сложности $V_{\langle \mathrm{m}\, q\rangle}$ каждого одноступенчатого канала проводился для двух вариантов его табличной реализации: $V_{\langle \mathrm{m}\, q\rangle}^{\text{ПЗУ}} = 2^{2r} \times 2r$ и $V_{\langle \mathrm{m}\, q\rangle}^{\text{ПЛМ}} = 4r \times q^2 + q^2 \times 2r = q^2 \cdot 6r$, где $r = \left\rceil \log_2 (q-1) \right\lfloor$ – разрядность вычетов по модулю q. Сложность двухступенчатых каналов оценивалась по формулам, приведённым в нижней строке табл. 3.6.

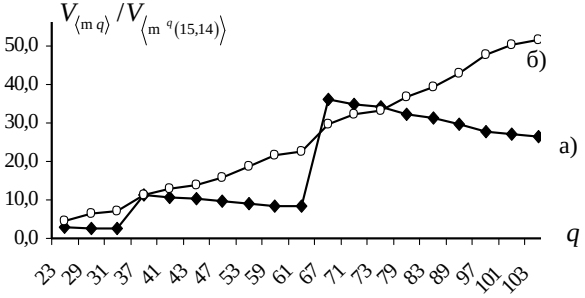


Рис. 3.13. Отношение аппаратной сложности таблично реализованных каналов $\langle \mathrm{m}\, q\rangle$ и двухступенчатых каналов $\langle \mathrm{m}^{\,q}(15,14)\rangle$:
а) канал $\langle \mathrm{m}\, q\rangle$ на ПЗУ; б) канал $\langle \mathrm{m}\, q\rangle$ на ПЛМ

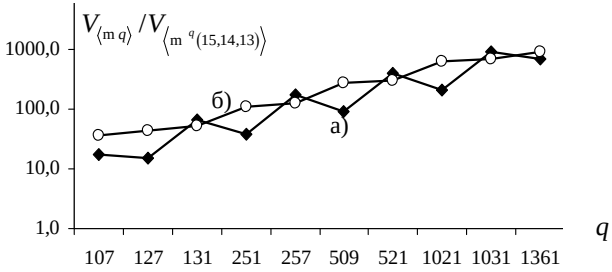


Рис. 3.14. Отношение аппаратной сложности таблично реализованных каналов $\langle \mathrm{m}\, q\rangle$ и двухступенчатых каналов $\langle \mathrm{m}^{\,q}(15,14,13)\rangle$:
а) канал $\langle \mathrm{m}\, q\rangle$ на ПЗУ; б) канал $\langle \mathrm{m}\, q\rangle$ на ПЛМ

При этом, исходя из условия равноценной производительности одно- и двухступенчатых каналов, использовался параметр $k_f = 4$ (полагалось, что в каждом из двухступенчатых каналов применяется набор преобразований $\{f_1, f_2, f_3, f_4\}$).

Полученные результаты показывают, что *переход от одноступенчатых СОК к многоступенчатым позволяет избавиться от «тирании больших таблиц», обеспечивая сокращение объема последних на один–три порядка*. Замечательное свойство аддитивных МСОК – линейная зависимость аппаратных затрат в каналах МВС от модулей главной системы – делает возможным:

- радикальное расширение диапазона представления данных без потери производительности МВС (в качестве модулей главной системы можно использовать простые числа от 3 до 509 и более);
- многократное сокращение временных затрат на вычисление позиционных характеристик чисел и их перевод в двоичную систему счисления при сохранении заданного диапазона представления данных. (Например, использование в качестве модулей главной системы девяти простых чисел от 271 до 317 вместо шестнадцати чисел от 5 до 61 сохраняет порядок диапазона представления данных $\approx 10^{22}$ и сокращает оценку (3.18) примерно вдвое. При этом суммарная сложность табличной МВС снижается более чем в 10 раз!).

Уникальным достоинством аддитивных МСОК является также возможность существенного повышения производительности МВС за счёт глубокой конвейеризации модулярных вычислений на уровне микроопераций (элементарных преобразований в каждой отдельно взятой ступени).

Таким образом, аддитивные многоступенчатые системы счисления в остаточных классах открывают новую перспективу в развитии модулярной арифметики, находящей всё более широкое применение в цифровой обработке сигналов и изображений, в том числе с использованием систем на кристалле.

Глава 4

ОБРАБОТКА ТЕЛЕВИЗИОННЫХ СИГНАЛОВ В СИСТЕМАХ НА КРИСТАЛЛЕ И ВИДЕОСИСТЕМАХ НА КРИСТАЛЛЕ

Обработка видеоинформации с целью кодирования и классификации изображений всегда была и будет актуальной в телевидении. Непреходящая актуальность проблемы обработки обусловлена тем, что большая априорная неопределённость часто побуждает проектировщика решить ту или иную задачу методом *перебора* вариантов. А для изображений с реальным числом элементов – это *трансвычислительная* задача, сложность которой определяется числом различимых градаций в степени числа элементов разложения, что далеко выходит за пределы возможностей каких-либо мыслимых компьютеров. Поэтому практической реализации поддаются лишь те задачи, сложность которых с ростом числа элементов разложения растёт полиномиально, а не экспоненциально. Не случайно Р. Беллман подчеркивал, что «сегодня, как никогда раньше в истории науки, характер теоретической постановки стремится идти рука об руку с вычислительной осуществимостью».

Изобретение множества методов обработки сигналов в совокупности с развитием технологии устройств обработки позволяют осуществлять кодирование изображений в реальном времени при весьма больших потоках информации. Вычислительная техника, во многом стимулируемая телевидением, наращивает своё быстродействие, которое требует пропускной способности, производительности и ёмкости памяти на несколько порядков больших, чем иные источники сигналов. В первую очередь именно телевидение заставило создателей компьютеров выдвинуть эпохальные требования к скорости передачи и вычислений – «даёшь Терагерц, Терабит, Терафлоп!».

Одним из ключевых инструментов повышения производительности устройств обработки видеоинформации является распараллеливание обработки. Телевидение освоило параллельное накопление ещё на этапе электронно-лучевой развёртки, а параллельную обработку компьютеры освоили несколькими десятилетиями позже [126]. Совокупность параллельного накопления, параллельного считывания и параллельных вычислений реализована только на втором этапе твердотельной революции в телевидении с появлением систем на кристалле и видеосистем на кристалле.

Развитие технологии СБИС направлено на рост быстродействия, интеграции и на возможность объединения на одном кристалле разнородных архитектур и устройств (например, программируемой логики, процессорного ядра, цифроаналоговых и аналого-цифровых преобразователей). Если ранее местом соединений, задаваемых разработчиком, была печатная плата, то сейчас усилия технологов направлены на реализацию всего многообразия возможных соединений в одном кристалле, с учетом ограничений на задержку сигнала, функциональную сложность СБИС и ресурс соединений. Подобные СБИС, объединяющие функционально разнородные структуры (например, процессор, память, программируемая логика и даже фотоприёмный массив), уже существуют и получили название системы на кристалле (СнК, SoC – System-on-Chip). Такая архитектура имеет аналогию в виде *разделения скоростной функции распознавания изображений и медленной функции логического мышления* между правым и левым полушариями человеческого мозга. Подобное решение позволяет значительно увеличить сложность обработки информации при сохранении скорости её обработки, так как конструктивно и процессор, и аппаратное ядро объединены на одном кристалле и есть возможность сочетать в рамках одной СБИС программную и аппаратную реализации. Ранее реализация подобного многообразия функций была доступна только в рамках одной печатной платы или нескольких, с соответствующим увеличением габаритов, ростом энергопотребления и уменьшением быстродействия, т. е. со снижением основных функциональных характеристик изделия.

Для широкого внедрения систем на кристалле в практику отечественной высокотехнологичной промышленностью взят курс на ускоренное освоение микроэлектронных технологий нового поколения и создание межотраслевой вертикально интегрированной системы сквозного проектирования СБИС и аппаратуры [8, 16]. Для решения этой основополагающей задачи развития электроники в России будет создано и модернизировано более 15 ключевых производств ЭКБ, реализована сеть из 40…50 центров проектирования СБИС типа «система на кристалле» и создан межотраслевой центр фотошаблонов. Производственную базу микроэлектроники (кремниевые фабрики) предполагается создавать как имеющие национальное значение инвестиционные проекты. Ожидается, что в 2011 году технологический уровень изделий в серийном производстве будет соответствовать 0,13 – 0,09 мкм и резко уменьшится доля импортной ЭКБ, что значительно укрепит обороноспособность и безопасность России [8].

4.1. Обработка видеоинформации в эпоху систем на кристалле

Базовой единицей, способной обрабатывать информацию, является логический элемент. Для уменьшения массы и габаритов, повышения надежности и быстродействия логические элементы или вентили технологически размещаются на одном кристалле, образуя интегральную микросхему (ИМС). Для реализации разнообразных алгоритмов обработки информации требуются различные классы ИМС:

- микросхемы малой и средней степеней интеграции, реализующие стандартные неизменяемые функции (логические элементы, счётчики, дешифраторы, регистры и т. д.). Логика обработки информации на таких ИМС задается разработчиком путем внешних электрических соединений (так называемая жесткая логика). Этот путь построения цифровых устройств – интеграция на печатной плате – является экстенсивным, так как при сложных алгоритмах обработки сигналов происходит увеличение количества микросхем, рост массогабаритных показателей и энергопотребления, снижение надежности из-за увеличения количества соединений. В настоящее время существующие аппаратные решения, использующие ИМС класса «жесткой логики», могут быть перенесены на более компактные конструкции программируемой логики;
- микросхемы высокого уровня интеграции (микропроцессоры, цифровые сигнальные процессоры, микроконтроллеры). Логика функционирования подобных устройств основана на последовательном выполнении заданной разработчиком программы обработки (так называемая микропроцессорная реализация). В настоящее время существует большое разнообразие процессоров различных архитектур, диапазонов тактовых частот (до единиц Гигагерц), с сокращенным (RISC) и со сложным (CISC) наборами команд, с большим разнообразием по потребляемой мощности, по габаритам, по доступному периферийному оборудованию и по назначению. Непременной составной частью таких систем являются запоминающие устройства команд и данных, предназначенные для организации и описания целевой функции устройства;
- микросхемы высокого уровня интеграции программируемой логики (программируемая матричная логика, перепрограммируемые вентильные матрицы, FPGA, FLEX). Функционирование устройств на таких микросхемах основано на задаваемой разработчиком внутренней конфигурации связей внутренних стандартных логических элементов («гибкая логика»). В настоящее время число программируемых вентилей на кристаллах достигает единиц миллионов, а

время задержки – до единиц наносекунд. Данная технология построения устройств обработки имеет много общего с «жесткой логикой», но здесь достигается интеграция на уровне кристалла СБИС, а конфигурация связей между стандартными макроячейками задается программированием соединений внутри единого кристалла. Поэтому существует ограничение на сложность конфигурации («разводки») кристалла;

- полностью заказные ИМС, алгоритм функционирования которых задается на этапе изготовления, а значит отсутствует возможность последующей модификации. Данные ИМС имеют наивысший коэффициент полезного использования площади кристалла, как следствие – наивысшую интеграцию и наибольшее быстродействие; однако длительный цикл проектирования и изготовления делает их практически недоступными для использования в уникальных системах, но эффективными для массовой аппаратуры.

Устройства на «жесткой» и на «гибкой» логике вследствие возможности выполнения *параллельной* обработки имеют потенциально большее системное быстродействие, чем устройства на «микропроцессорной реализации». Микропроцессорные устройства могут реализовать более сложный алгоритм обработки информации, чем тот, который доступен аппаратной логике. Такой взаимообмен системного быстродействия и сложности обработки должен учитываться при синтезе алгоритмов обработки видеоинформации.

Различные классы ИМС предполагают и различные технологии проектирования устройств на их основе с различным уровнем автоматизации проектирования. Так, проектирование устройств на основе «жесткой» логики требует знаний номенклатуры и функциональных характеристик стандартных ИМС, формализованных методик синтеза и эвристических подходов к реализации нетиповых устройств при невысоком уровне автоматизации разработки. При проектировании «микропроцессорной реализации» аппаратная часть конфигурируется из типовых узлов (процессор, память, периферия), а основной акцент переносится на программное обеспечение, которое может происходить как на низком уровне – ассемблер, так и на разнообразных языках высокого уровня при среднем уровне автоматизации. Проектирование устройств на основе «гибкой» логики требует наиболее высокого уровня автоматизации проектирования с обязательным присутствием функционального и временно́го моделирования в различных пакетах САПР. В этом случае часто невозможна компиляция проекта и конфигурация кристалла без соответствующего программного обеспечения, т. е. рост сложности ИМС влечет за собой рост сложности средств

проектирования и отладки.

Высокая сложность СБИС приводит к увеличению сроков разработки, удорожает её, требует значительных затрат высококвалифицированного инженерного труда и применения всё более сложных аппаратных и программных средств проектирования. Исходя из этого, возникла необходимость применения идеи «детского конструктора», когда роль отдельных кубиков выполняют уже отлаженные, опробованные и хорошо зарекомендовавшие себя «сложные функциональные блоки» – СФ-блоки. По определению СФ-блок (IP-блок, IP – Intellectual Property – интеллектуальная собственность) представляет собой модель функционально законченного устройства, описанную соответствующим образом и предназначенную для использования в качестве функционального модуля (ядро процессора, блок памяти, интерфейс и т. п.) при разработке СБИС класса система на кристалле.

Разработчику системы на кристалле уже не нужно проектировать СБИС «с нуля», а можно воспользоваться отлаженным набором СФ-блоков. Чем больше библиотека СФ-блоков будет покрывать поле задач, тем быстрее и с меньшими затратами будет выполнена разработка СБИС [6, 12]. Например, в состав САПР фирмы Cadence входит более 3000 разнообразных модулей включая коммуникационные системы сотовой связи WCDMA, cdma2000, GSM/GPRS/HSCSD/EDGE, беспроводные сети стандартов IEEE802.11, Bluetooth, многообразные кодеры–декодеры цифрового телевидения и др. [27, 157, 158, 164, 165].

В основе методологии проектирования лежит принцип повторного использования сложных функциональных блоков (IP-блоков, СФ-блоков), разрабатываемых целенаправленно. По аналогии с системой на плате, где в качестве компонентов выступают готовые ИМС, система на кристалле конструируется из повторно используемых блоков. IP-блоки могут быть двух типов:

- soft-IP, описанные на уровне регистровых передач (RTL-уровне);
- hard-IP, описанные на топологическом уровне.

Сегодня для обозначения повторно используемых блоков (reuse) наиболее часто применяют термин IP-блок, т. е. блок, представляющий собой объект интеллектуальной собственности. Однако в ходу и другие термины: СФ-блок, macro, core (обычно обозначает блоки типа центральный процессор CPU или процессор цифровой обработки сигналов DSP). СФ-блоки можно многократно использовать при проектировании различных схем обработки

сигнала. Разработчик аппаратуры, имея системные модели IP-блоков, задействует их при проектировании своего устройства, сокращая сроки разработки и увеличивая надежность и безошибочность проекта.

Система на кристалле (СнК, SoC – System on Chip) – это функционально полный программно-аппаратный комплекс, который включает в себя собственно СБИС и программное обеспечение (ПО). СБИС выполняет заложенную аппаратную функцию, а ПО, под управлением которого происходит обработка данных, разрабатывается и вводится в программирующее устройство. Высокие затраты и сопутствующие риски при создании СнК полагают, что проекты СБИС должны планироваться и создаваться как платформы, рассчитанные на различные приложения и дочерние продукты. Смена ПО при одинаковой структуре СБИС даёт возможность реализовать различные системы обработки сигналов.

Систему на кристалле определяют [12] как сложную интегральную схему, объединяющую в одном чипе все основные функциональные элементы полного конечного продукта. Соответствующая терминология стала оформляться несколько лет назад, когда уровень развития технологии позволил реализовывать на кристалле не только отдельные устройства, но и функционально законченные системы. В целом понятие система на кристалле предполагает любую более или менее сложную и функционально завершенную систему, содержащую, как правило, микропроцессорное ядро с разнообразными сложными периферийными устройствами обработки и реализуемую на отдельном кристалле.

В качестве иллюстрации на рис. 4.1*а* приведена топология системы на кристалле декодера MPEG-2, представляющего собой декодер источника (разработка ФГУП «НИИМА «Прогресс» в кооперации с ФГУП «НИИТ» и ГУП «Элвис»). Модули процессора, памяти, синтезатора частоты и аппаратного декодера MPEG-2 – суть IP-блоки, разработанные дизайн-центрами и интегрированные в общую систему на кристалле. На рис. 4.1*б*. – фото партии СБИС класса система на кристалле декодера MPEG-2 и тестовой платы, выполняющей ту же функцию.

Подобная методология и соотношение площадей кристалла между готовыми IP-блоками и вновь разработанными характерна и для другой важной составляющей телевизионной системы – декодером канала [165]. «System» в понятии система на кристалле гораздо важнее чем «Chip»: то, что основные функциональные блоки *проектируются* как часть интегрированного целого, важнее того, что они все будут физически размещены на одной подложке[12].

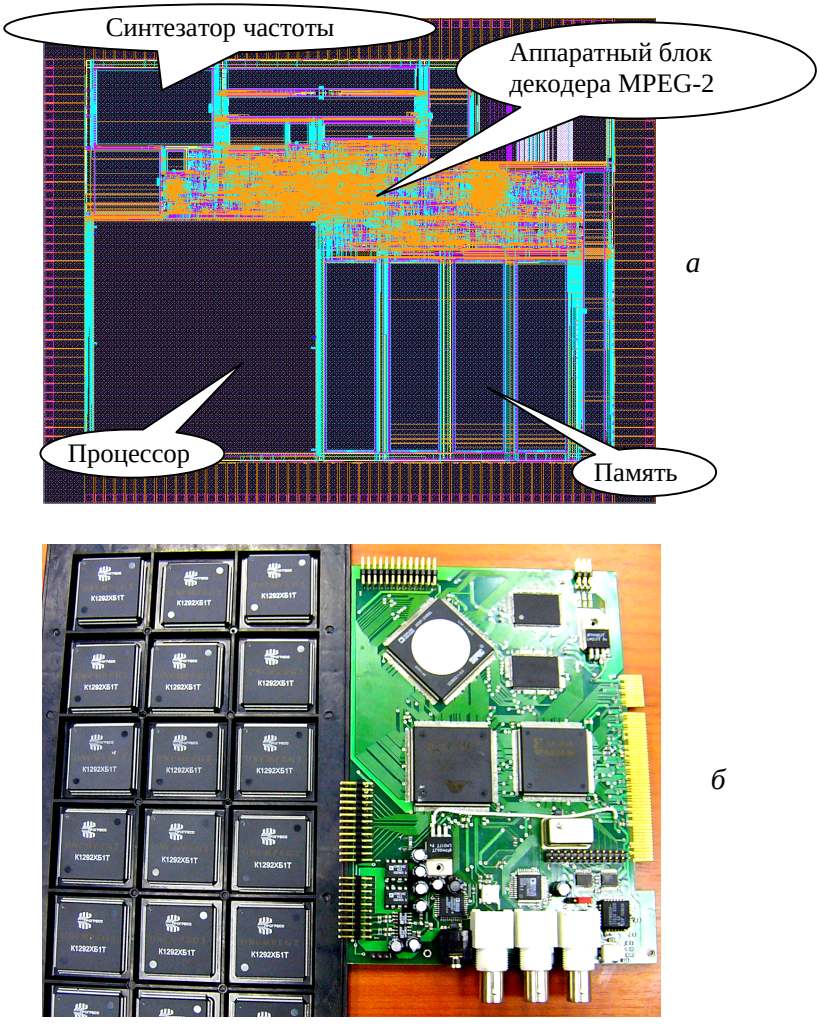


Рис. 4.1 Система на кристалле: декодер MPEG-2; топология *а*
и партия СБИС с платой прототипа *б*

Решения типа система в корпусе(SiP – System in package) могут оказаться более дешевыми, чем единственная подложка, но при этом можно по прежнему считать, что SiP проектируется как интегрированная система, с различными компромиссами [12].

Взаимное влияние телевизионной техники и микроэлектроники проявляется не только в том, что телевизионная техника использует СБИС, но и сами СБИС создаются с применением

телевизионно-компьютерных систем (ТВКС). Наиболее активно телевизионная техника применяется на этапах контроля качества СБИС (рис. 4.2) и печатных узлов, построенных на их основе (рис. 4.3 и 4.4.) (по материалам ЗАО Предприятие ОСТЕК).

Наблюдается устойчивый рост доли *смешанных цифро-аналоговых* систем в общем объеме выпуска систем на кристалле [9], по данным компании Cadence Design Systems в 2006 г. их доля достигнет70% от дискретно-аналоговых систем [158]. Поэтому такая смешанная обработка видеоинформации, отраженная в гл. 2, является актуальной на современном этапе развития телевизионно-компьютерной техники. Появились и новые виды систем на кристалле [5, 8, 37], например сочетающие формирование и обработку изображений, поэтому в маршрут проектирования должны быть включены этапы по совместной верификации (тестированию) цифровой, аналоговой и видео составляющих проекта.

Современный уровень полупроводниковой технологии, с проектными нормами менее 350 нм, позволяющий размещать на кристалле сотни миллионов транзисторов (с возможностью реализовывать на кристалле одновременно процессоры, память, цифровую логику, аналоговые узлы, интерфейсы и т. д.) [6, 8, 12], *определил революционные изменения в микроэлектронике и в целом радиоэлектронной индустрии*.

Одна из проблем при проектировании СнК состоит в создании IP-блоков. Практика показывает, что стоимость reuse-блока в среднем в 10 раз превышает стоимость аналогичного однократно используемого блока, а для процессоров эта величина на порядок выше. По этой причине reuse-блоки предназначены для общих логически формализуемых задач, например для кодирования изображения (MPEG, wavelet). СФ-блоки являются рыночным товаром, в западных странах объем продаж IP-блоков оценивается на уровне свыше 10 миллиардов долларов в год. Например, широко распространенные процессорные ядра ARM7, ARM9 стоят порядка сотен тысяч долларов для конкретного применения.

Принципиальная особенность СнК - это наличие программируемых блоков, поэтому СнК - это не просто интегральная схема, а комплекс, в состав которого входит как аппаратная часть (собственно кристалл), так и программная – встраиваемое ПО.

Поскольку система на кристалле – изделие не только сложное, но и при небольших объемах выпуска – дорогое, то ошибки проектирования должны быть исключены. Поэтому маршрут проектирования должен включать операции по совместной

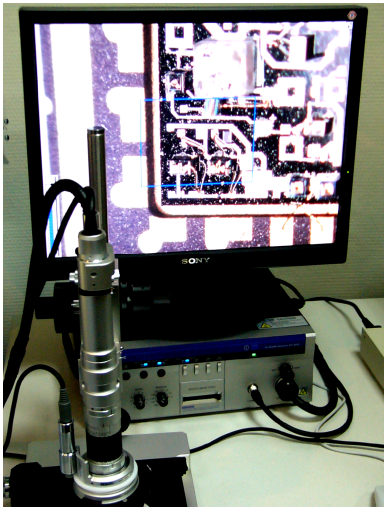


Рис. 4.2. Телевизионно-компью-терная система контроля СБИС

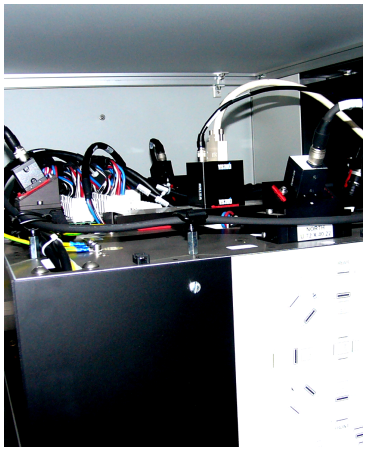


Рис. 4.3. Многоракурсная телевизионно-компьютерная система контроля печатных узлов (вид изнутри)

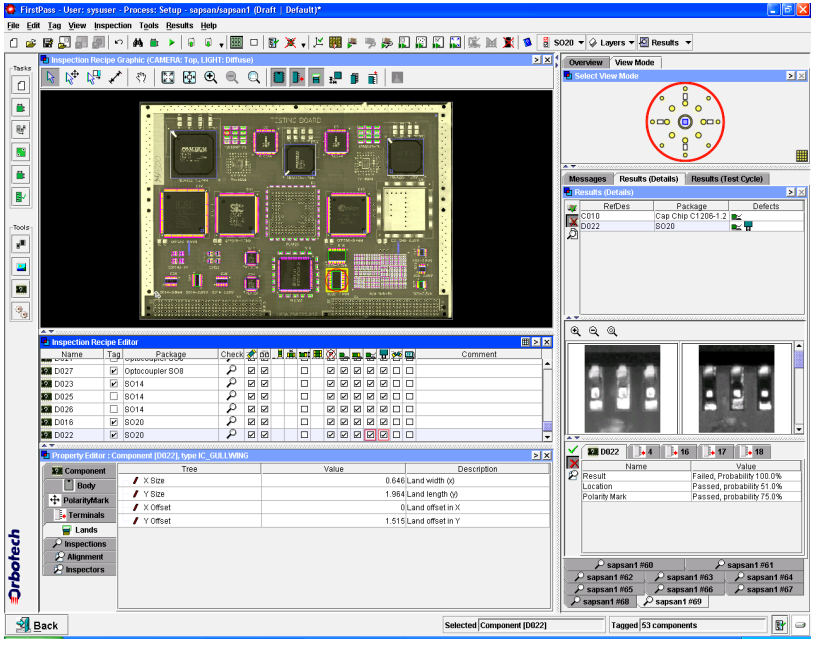


Рис. 4.4. Вид экрана монитора телевизионно-компьютерной системы контроля качества печатных плат.

верификации и отладке программной и аппаратной частей проекта.

При разработке микроэлектронных изделий необходимо учитывать следующие факторы:

- в условиях рынка прибыль от изделий электроники сильно зависит от временных затрат на проектирование;
- увеличение степени интеграции делает задачу верификации микросхем качественно более сложной;
- технические параметры интегральных микросхем, такие, как производительность, площадь кристалла и потребляемая мощность, являются ключевыми в продвижении товара на рынок;
- команды разработчиков СБИС обладают различными знаниями и опытом в области проектирования и часто при выполнении проектов разнесены территориально.

Как и на более ранних этапах развития твердотельного телевидения [1, 2], одной из целей обработки видеоинформации в эпоху систем на кристалле является повышение их быстродействия. Под быстродействием телевизионно-компьютерной системы понимается количество принимаемых решений за единицу времени. С позиций системного подхода, при анализе и синтезе телевизионно-компьютерной системы как *единого целого*, разработчика, интересует совокупное быстродействие фотоприёмника и сочлененного с ним процессора. Быстродействие телевизионно-компьютерных систем можно повышать широким набором способов и алгоритмов, разделяемых на два класса. Первый основан на обработке без потери полезной информации. Второй повышает производительность за счет компромиссных потерь информации, он основан на невозможности абсолютно точных измерений и допустимости увеличения существующих потерь информации [4].

Существует экстенсивный метод повышения быстродействия – наращивание частоты процессора обработки, однако он имеет физические пределы и ведёт к росту рассеиваемой на кристалле мощности. Больший интерес представляют следующие направления повышения быстродействия:

- распараллеливание вычислений;
- сокращение вычислительных затрат путем разработки быстрых алгоритмов, учитывающих специфику сигнала;
- структурный синтез телевизионно-компьютерных систем для повышения быстродействия.

Основные вычислительные операции при решении типовых задач обработки изображений могут быть сведены к набору операций над матрицами. По вычислительным затратам эти алгоритмы достаточно сложны и требуют примерно N^3 или N^4 операций умножения для каждого набора из N входных отсчетов [116]. Если исключить прямое повышение частоты процессора, то увеличение на несколько порядков производительности для обработки в реальном времени должно осуществляться эффективным использованием параллелизма при вычислениях.

Параллельные компьютеры предполагают концентрацию вычислительных ресурсов, ресурсов памяти и высокоскоростных каналов ввода-вывода для решения вычислительных задач, с использованием принципов параллельной обработки. Самым непосредственным способом реализации параллельной обработки сигналов является простое соединение набора процессоров к общей шине, однако такой путь повышает производительность только до тех пор, пока не наступают ограничения, связанные с проблемами соединения с общей шиной. Для мультипроцессорной обработки с общей шиной существует конфликт Минского между N процессорами с коллективным распределением ресурсов, соединенными с общей шиной, ограничивающий производительность значением log_2N. Организация параллельных вычислительных систем на основе суперкомпьютеров для тех же алгоритмов позволила достичь производительности, соответствующей закону Амдала [116, 163], т. е. N/log_2N.

Наивысшей производительностью в задачах обработки изображений обладают систолические архитектуры, совмещающие достоинства конвейерных и матричных вычислителей. Они обеспечивают наиболее перспективную совокупность характеристик для обработки сигналов в реальном времени: параллелизм на уровне модулей с пропускной способностью, прямо пропорциональной числу вычислительных ячеек, простое управление, синхронный поток данных, локальные связи, достаточную гибкость для выполнения матричных операций. Для систолической обработки для широкого класса алгоритмов существует линейная зависимость между количеством процессоров N и достигаемой производительностью (рис. 4.5).

Решаемые телевизионно-компьютерной системой задачи могут быть разбиты на два класса: вычислительно ограниченные и ограниченные по вводу–выводу.

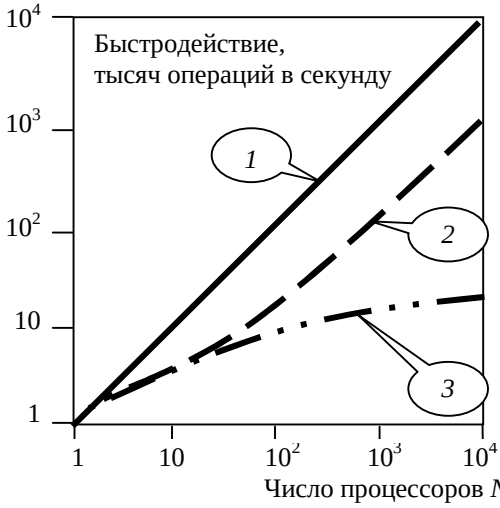


Рис. 4.5. Зависимость увеличения быстродействия, обеспечиваемого одновременно работающими N процессорами.
1 – 100% производительность (систолические решетки);
2 – закон Амдала;
3 – предположения Минского (ограничения, связанные с шинным конфликтами в многопроцессорной системе)

К первому классу относятся алгоритмы, число вычислительных операций в которых превосходит полное число элементов ввода–вывода. Примером этого класса задач является алгоритм перемножения матриц, т. к. в нём каждый элемент одной матрицы умножается на все элементы строки дугой матрицы. Ограниченными по вводу–выводу являются задачи, в которых число операций меньше полного числа элементов, предназначенных для ввода–вывода. Примером такой задачи является сложение матриц, т. к. полное число операций не превышает числа элементов в обеих матрицах, а на каждую операцию сложения нужно проделать минимум три обращения к памяти: чтение одного элемента, чтение другого, сохранение суммы.

Увеличить производительность процессора обработки для задач ограниченных по вводу–выводу (кроме увеличения пропускной способности), можно при использовании систолической архитектуры. Такая система состоит из множества связанных между собой ячеек, каждая из которых выполняет простую операцию. Информация в систолической системе передаётся между ячейками в конвейерном режиме [116]. Такая организация находит применение в телевизионно-компьютерных системах для задач, ограниченных по вводу-выводу, и где многочисленные операции выполняются над каждым элементом данных одним и тем же повторяющимся способом. При обработке видеоинформации широко применяются линейные методы,

которые обеспечивают возможность создания регулярных структур процессоров. Можно унифицировать процессоры обработки изображений на следующие группы:

- процессоры, предназначенные для операций непосредственно над видеосигналом. К ним относятся фильтры, корреляторы, измерители координат объектов. Главные черты такой обработки описаны в гл. 2 и могут быть эффективно реализованы в СнК;
- процессоры для операций над корреляционными матрицами: перемножения, сложения и обращения. Эти операции могут быть реализованы на реконфигурируемой СнК, описанной в разделе 4.2;
- процессоры сжатия изображений и канального кодирования, к которым предъявляются дополнительные требования поддержки разнообразных алгоритмов и стандартов сжатия. Эти вопросы рассмотрены в разделах 3.3 и 4.3.

Применение систем на кристалле, благодаря сочетанию параллельной обработки информации и сложных логических алгоритмов, позволяет достичь большего быстродействия, чем достижимое при использовании микропроцессоров [1, 36]. Если первый этап твердотельной революции в телевидении в области технологии обработки был основан на отходе от «жесткой» логики и переходу к микропроцессорам, то сегодня является очевидной необходимость сочетания программируемой логики и микропроцессоров – что и реализуется в системах на кристалле, в частности реконфигурируемых, описанных в следующем разделе.

4.2. Особенности проектирования телевизионных систем с использованием систем на кристалле

Подход к проектированию изделий микроэлектроники, сложившийся в XX веке, уже не оправдывает себя. Исходя из производительности разработчика порядка 100 вентилей в день и сложности микросхем порядка 10^7 вентилей, для разработки такой СБИС потребуется 500 человеко-лет, что, безусловно, неприемлемо. Кроме того, при производстве изделий микроэлектроники постоянно растёт доля затрат на разработку программного обеспечения, поэтому при раздельной разработке программной и аппаратной составляющих резко возрастает вероятность выявления ошибок на этапах тестирования или эксплуатации, что также приводит к дополнительным затратам. Выход состоит в изменении методологии проектирования СБИС, и наиболее перспективным направлением представляется методология проектирования «система на кристалле».

Ведущие полупроводниковые и приборостроительные компании, активно внедряющие системы на кристалле, признают, что необходима новая методология проектирования подобных систем и новый алгоритм взаимодействия разработчиков систем и аппаратуры, собственно разработчиков систем на кристалле, а также производителей микросхем. Для достижения такой цели организована специальная международная ассоциация Virtual Socket Interface Alliance, задача которой – формировать методологию проектирования систем на кристалле, стандарты на описания сложных функциональных блоков и библиотек виртуальных элементов, маршруты проектирования и верификации таких устройств [36].

Примером успешных полупроводниковых компаний являются производители ПЛИС Xilinx и Altera, которые относятся к так называемым «fables» компаниям, т. е. не имеющим собственных фабрик. Основной вид деятельности таких предприятий – высокоуровневая разработка структур СБИС и производство средств САПР [12]. Среди государств лидером в области только производящих, так называемым «pure-play» компаний является Тайвань.

Реконфигурируемые системы на кристалле для цифровой обработки. Непрерывное усложнение методов цифровой обработки требует не только экстенсивного роста показателей элементной базы (быстродействия, количества эквивалентных вентилей), но и новых архитектурных решений системного уровня интеграции [24]. Особо остро стоит задача интеграции процессорного ядра и массива программируемой логики, что продиктовано необходимостью сочетания в реальных изделиях аппаратной и программной обработки. Существует два вида методов проектирования систем на кристалле – алгоритмически ориентированные методы и методы «платформенного» проектирования.

В технических решениях производителей изделий типа «реконфигурируемая система на кристалле» (РСнК, – Configurable System On Chip) в основном присутствует два подхода:

- реализация синтезируемого (как правило, специализированного) процессорного ядра в виде загружаемой в ПЛИС структуры;
- реализация жесткого (как правило, стандартного) аппаратного процессорного ядра на одном кристалле с логической матрицей.

Первый подход к построению систем на кристалле применен компанией Xilinx (8/32-битные ядра PikoBlaze, MicroBlaze), компанией Altera (16/32-битное синтезируемое ядро Nios) и др. Достоинства такого подхода – возможность использования

стандартных микросхем программируемой логики (ПЛИС) соответствующих производителей и возможность конфигурации архитектуры процессора под конкретную специфику задачи. Недостатков у такого подхода несколько больше:

- уменьшение ресурсов выбранной ПЛИС пропорционально размеру и разрядности синтезируемого ядра;
- аппаратная привязка синтезируемого процессорного ядра к конкретному семейству ПЛИС;
- ограниченное число производителей средств разработки программного обеспечения (ассемблер, компилятор Си и т. д.) для нестандартного ядра;
- реализация периферийных стандартных блоков (таймеры, контроллеры DMA, контроллер прерываний, интерфейс с внешней памятью и т. д.) за счет ресурсов ПЛИС;
- повышенное энергопотребление, вызванное формированием структуры процессора из стандартных ячеек, неоптимизированных по скоростным режимам работы.

Второй подход – реализация на одном кристалле аппаратного процессорного ядра совместно с программируемой логикой – представлен следующими производителями: Altera (система Excalibur, 32-битный процессор ARM), Atmel (система FPSLIC, 8-битный контроллер AVR), Triscend (система E05, 8-битный контроллер 8051, система A7, 32-битный процессор ARM). Такое решение технологически более трудоемко, так как связано с совмещением различных технологий изготовления приборов: микропроцессорной и FPGA. Недостаток такого подхода – жесткая архитектура процессора, которая ограничивает гибкость его применения в специфичных задачах. Достоинства данного решения являются антиподами недостатков первого и состоят в следующем:

- ресурсы программируемой логики кристалла не используются для реализации процессорного ядра;
- средства разработки программного обеспечения для стандартных процессоров предлагаются несколькими независимыми производителями.

Архитектура реконфигурируемой системы на кристалле A7 фирмы Triscend (рис. 4.6) имеет вид ИМС, содержащей как традиционные для микропроцессоров узлы, так и матрицу программируемой логики, выполненной по технологии FPGA, а также средства их взаимодействия [36, 95]. Интеграция этих технологий часто осуществляется лишь на печатной плате, что приводит к необходимости проектирования сопряжения процесс-

сора и программируемой логики, к росту энергопотребления интерфейсных схем, а в итоге увеличивает время проектирования, потребляемую мощность и массогабаритные показатели изделия.

Центральный процессор включает в себя 32-разрядное RISC-ядро ARM7TDMI, которое имеет общее адресное пространство кода и данных объемом 4 Гбайт и 32 регистра общего назначения. Большинство команд выполняется за один такт, т. е. для максимальной тактовой частоты менее чем за 20 нс. Для ускорения умножения используется аппаратный умножитель с поддержкой функции умножения с накоплением, актуальной для задач цифровой обработки сигналов. Кроме RISC-инструкций в систему команд входят CISC-подобные инструкции, предназначенные для загрузки или сохранения группы регистров, трехадресные инструкции, для которых два регистра источника и регистр результата можно задавать независимо, и т. д., что позволяет достичь наивысшей производительности.

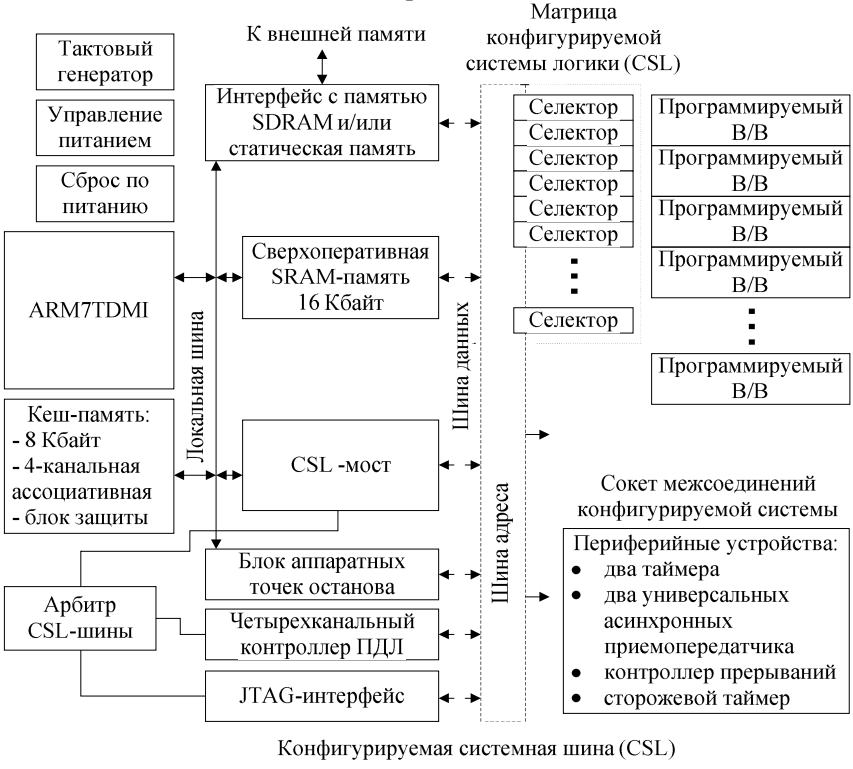


Рис. 4.6. Архитектура реконфигурируемой системы на кристалле

В настоящее время процессоры с архитектурой ARM широко распространены и являются лидерами по соотношениям «производительность/потребляемая мощность» и «цена/качество», поэтому выбор ядра ARM7TDMI для построения первой в мире 32-разрядной реконфигурируемой системы вполне оправдан.

Ядро ARM7TDMI, в зависимости от задачи, может работать в двух состояниях – ARM или THUMB. В состоянии ARM ядро исполняет 32-битные инструкции, выровненные на границе 4-байтного слова, и имеет наиболее широкий набор команд. В состоянии THUMB ядро оперирует с усечённым набором инструкций, при этом команды выровнены на границе полуслова. Состояние выбирается программно, исходя из разрядности интерфейса подсистемы памяти, ограничений на размер кода и на скорость выполнения приложения, и может изменяться в процессе работы пользовательского приложения, обеспечивая максимальную гибкость в применениях (рис. 4.7).

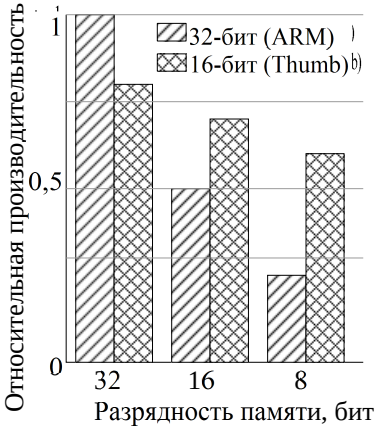


Рис. 4.7. Производительность ядра

Состояние Thumb во многих случаях позволяет сократить объём кода на 30…50 %, а производительность при использовании внешней 8 или 16 разрядной памяти может оказаться выше, чем в состоянии ARM, за счёт меньшего количества обращений к памяти. Ядро ARM7TDMI имеет 7 режимов работы, названия и назначение которых приведены в табл. 4.1. Режим работы может быть изменён программно или автоматически при возникновении внешнего прерывания или исключения, например при выборке из памяти несуществующей инструкции. В зависимости от режима работы и состояния ядра изменяется видимость и набор доступных регистров, допустимость выполнения некоторых инструкций. Так, при настройке механизма кеширования и защиты памяти, встроенного в систему на кристалле Triscend A7, выполняются команды сопроцессора, что предполагает нахождение ядра в режиме Supervisor.

Таблица 4.1. Режимы работы ядра ARM7TDMI

Режим	Назначение
User	Пользовательский режим, нормальное состояние выполнения программ
FIQ	Поддержка скоростных процессов
IRQ	Поддержка прерываний обычного приоритета
Supervisor	Защищённый режим
Abort	Возникновение аномальной ситуации при выборке команды или данных
System	Системный режим
Undefined	Выполнение неопределённой инструкции

Программная поддержка такой работы заключается, прежде всего, в распределении памяти, задании отдельных областей стека и точек входа исключительных ситуаций для различных режимов. Обычно это выполняется в C-startup-коде непосредственно после сброса процессора.

Система прерываний реализует 7 различных видов прерываний (табл. 4.2), часть из источников прерываний отвечает за исключительные ситуации, связанные с выполнением программы, а другая часть обслуживает запросы от внешнего аппаратного окружения. Выполнение кода исключительной ситуации начинается с фиксированных предопределенных адресов в начальной области памяти.

Благодаря встроенному контроллеру прерываний все многообразие внешних запросов прерываний с помощью схем приоритета и мультиплексирования сводится к двум стандартным ARM прерываниям:

- быстрое прерывание (FIQ) с временем реакции в наихудшем случае 28 циклов процессора (420 нс при тактовой частоте 60 МГц) для обработки критичных ко времени процессов;
- обычное прерывание (IRQ) для стандартной реакции на внешнее воздействие. Запросы от источников прерываний вызываются высоким активным уровнем, приоритет источников прерываний задается программно. Для организации гибкой логики обслуживания прерываний контроллер включает в свой состав 10 регистров.

Таблица 4.2. Типы векторов прерываний ядра ARM7TDMI

Тип исключения	Приоритет	Причина или источник	Состояние ядра	Адрес вектора
Быстрое прерывание	3	Запрос от контроллера на быстрое прерывание	FIQ	0x1C
Обычное прерывание	4	Запрос от контроллера на обычное прерывание	IRQ	0x18
Резерв	-	-	-	0x14
Ошибка данных	2	Ошибка устройства защиты памяти	Abort	0x10
Ошибка выборки	5	Выполнение ядром несуществующей инструкции	Abort	0x0C
Программное прерывание	6	Выполнение ядром инструкции SWI	Supervisor	0x08
Неопределённая инструкция	6	Выполнение ядром внешней инструкции без реакции сопроцессора	Undefined	0x04
Сброс	1	Любой источник системного сброса. Программный сброс м. б. произведен путем перехода на адрес 0x0	Supervisor	0x00

Управление исключительными ситуациями производится через таблицу векторов прерываний, расположенную в нижней части карты памяти начиная с адреса 0x0. Для каждого из семи видов прерываний в этой таблице зарезервировано 8 байт памяти. В таблице векторов содержатся команды перехода следующих видов:

- выполняющиеся за 1 такт системной частоты (*B label*) и позволяющие передавать управление относительно текущего положения счетчика команд вперед или назад на 32 Мбайт;
- выполняющиеся за 2 такта системной частоты команды загрузки счетчика команд (*ldr pc, adr*) и позволяющие осуществлять переход в полном пространстве 4 Гбайт.

Цель применения таких команд – передача управления непосредственно на подпрограмму обработки исключительной ситуации. Системы на кристалле A7 содержат встроенную систему тактирования. Программно или при первоначальной конфигурации можно выбрать следующие источники для цепей тактирования процессора и массива программируемой логики:

- внутренний *RC*-генератор с низкой стабильностью и диапазоном частот 15…30 МГц;

- вход внешней синхронизации;
- внутренний синтезатор частоты (Clock Synthesizer), работающий от внешнего часового кварцевого резонатора 32 768 кГц, обеспечивающий выходную частоту в пределах от 1 до 60 МГц.

Во время работы возможна смена источников тактирования, изменение частоты синтезатора и прекращение генерации для уменьшения энергопотребления устройства.

Подсистема памяти в составе Triscend A7 является важной составляющей кристалла, она оптимизирована на достижение максимальной производительности, простоты подключения различных типов запоминающих устройств и включает в себя следующие блоки:

- внутреннее быстрое ОЗУ объемом 16 Кбайт (StratchPad SRAM);
- внутреннюю кеш-память команд и данных объемом 8 Кбайт (Cache);
- контроллер интерфейса с внешней системой памяти различного типа с суммарным объемом до 4 Гбайт (Memory Interface Unit).

Кроме того, к подсистеме памяти можно отнести встроенные буфера ускорения обмена с внешней динамической и статической памятью, аппарат переназначения адресов, защиты и приоритетов областей памяти («ремапирование»), механизм конвейерной предвыборки команд и данных (рекомендуемый при частоте более 40 МГц). Кратко рассмотрим составляющие подсистемы памяти.

Операции записи–чтения во внутреннее ОЗУ выполняются за один такт, причем можно одновременно задействовать старшие и младшие 8 Кбайт, например, ядро по локальной шине записывает данные в младшую половину ОЗУ, а в это же время по каналу DMA производится чтение ячейки из старшей половины. Это делает внутреннее ОЗУ удобным для использования в качестве «моста» между ядром ARM7TDMI и массивом программируемой логики. Внутренняя память является наиболее быстрой, поэтому при распределении адресного пространства в неё целесообразно помещать наиболее критичный по времени исполнения код, например обработчики прерывания. Компиляторы языка Си пакетов MULTI2000 и ADS обладают возможностью гибко настраивать карту памяти.

Внутренняя кеш-память «прозрачна» для выполняемой программы, считываемые данные и инструкции могут сохраняться в ней, и для их использования в дальнейшем не потребуется длинный цикл выборки обращения к внешней шине, все транзакции будут выполняться с локальной шиной, что значительно повысит производительность. Исследования [102]

механизма программного цикла ввода потока видеоданных показали, что при включении механизма кеш-памяти выборка команд из внешней динамической памяти происходит лишь до момента первого прохождения тела цикла, в дальнейшем остаются только импульсы чтения данных. Имеется механизм назначения областей памяти, для которой включен аппарат кеширования. Естественно, назначать кеширование целесообразно для относительно медленных устройств памяти, например FLASH или SDRAM, кешировать область внутреннего ОЗУ необязательно.

Контроллер интерфейса с внешней системой памяти ориентирован на поддержку статической и динамической памяти различной разрядности и информационной ёмкости, а также обеспечивает настройку временных параметров выборки, параметров регенерации, режимов энергопотребления [103]. Возможно подключение стандартных модулей DIMM объёмом до 256 Мбайт, а при организации страничной адресации возможен и больший объём внешнего динамического ОЗУ [121].

Важной особенностью контроллера является автоматическое проведение нескольких транзакций записи–чтения в зависимости от разрядности интерфейса внешнего ЗУ. Так, операция записи 4-байтного числа во внешнюю SDRAM-память с шириной данных шины 1 байт ведёт к «прозрачному» для разработчика выполнению четырёх байтовых циклов записи. Наличие встроенного программируемого контроллера позволяет упростить подключение разнообразных видов памяти и отказаться от организации поддержки динамического ОЗУ на программируемой логике.

Массив конфигурируемой логики предназначен для создания оригинальных пользовательских устройств обработки на одном кристалле с процессором и содержит в зависимости от типа микросхемы до 3200 логических ячеек. Каждая логическая ячейка содержит 4-входовую таблицу перекодировки, схему расширителя ускоренного переноса, триггер с входами тактирования разрешения и сброса [36]. Ячейки конфигурируются при включении питания из внешней FLASH-памяти или через интерфейс JTAG от персонального компьютера в режиме отладки. Они могут выполнять комбинационную или последовательную обработку сигналов, могут быть сконфигурированы в модули памяти, в том числе и двухпортовой. В матрице конфигурируемой логики предусмотрены 6 глобальных сигналов синхронизации с повышенной нагрузочной способностью, длинные и короткие связи для облегчения внутрикристальной разводки.

Процесс проектирования аппаратной части на матрице конфигурируемой логики имеет много общего с подходами для микросхем FPGA [98, 99, 105, 194]. Его особенностью является описание и моделирование взаимодействия аппаратной и программной частей проекта через механизм адресных селекторов и системную шину.

Непосредственно к массиву программируемой логики подключаются порты ввода-вывода, каждый из которых имеет в своем составе 4 триггера для синхронизации ввода-вывода, а также содержит дополнительные узлы, которые обеспечивают:

- управление временем нарастания выходного сигнала;
- управление выходным током вывода;
- поддержку выходных уровней для внешнего окружения с питанием 2.5 или 3.3 В;
- совместимость с шиной PCI 3.3 В;
- управление входным гистерезисом.

Порт ввода-вывода помимо традиционной фиксации входа подтягивающим резистором к потенциалу земли или питания может обеспечить интересный режим Bus Follower – отслеживания и запоминания последнего активного состояния вывода, что может быть использовано, в частности, для снижения динамического энергопотребления. Максимальное количество доступных линий ввода-вывода зависит от типа кристалла и выбранного варианта интерфейса с внешней памятью и достигает 252.

Стандартные периферийные устройства являются традиционными для микропроцессорных систем и включают в себя:

- два 16-битных таймера, задающих начальное состояние и работающих в режиме счёта внутренних тактовых импульсов;
- два последовательных интерфейса UART, совместимых с 16C450, с поддержкой сигналов квитирования и встроенной памятью типа FIFO объёмом 16 байт;
- 16-входовой контроллер прерываний (описан ранее);
- сторожевой таймер (32-bit Watchdog Timer) с возможностью формирования импульса сброса;
- схему сброса (Power-On Reset) с возможностью раздельного сброса только процессорного ядра, только массива конфигурируемой логики или кристалла в целом;
- схему управления энергопотреблением (Power Control).

Использование стандартных периферийных устройств кристалла Triscend A7 значительно облегчает процесс проектирования и высвобождает ресурсы матрицы программируемой логики для реализации только «нетиповых» функций.

Как известно [12, 27], механическое объединение разнородных структур без средств их взаимодействия ещё не образует системы. В качестве такого объединяющего звена служит высокопроизводительная конфигурируемая системная шина (CSI Bus) и механизм адресных селекторов. Системная шина (рис. 4.8) имеет разрядность 128 бит, является строго синхронной (все процессы фиксируются по переднему фронту тактовой частоты ядра процессора) и на ней представлены текущие адрес, данные и управляющие сигналы взаимодействия процессора, контроллеров DMA, конфигурируемой логики и арбитра шины. Шина поддерживает 8-, 16- и 32-битные транзакции с минимальной длительностью в один такт процессора.

Арбитр шины разделяет запросы на ресурс шины от ядра процессора, четырех контроллеров DMA и схемы JTAG-отладки реального времени. Максимальная пропускная способность конфигурируемой шины достигает 240 Мбайт/с. Важным моментом для разработчиков является возможность компьютерного моделирования процессов взаимодействия аппаратного и программного обеспечения при обмене по системной шине. При этом существующая в среде разработки VHDL-модель шины позволяет ускорить совместную отладку различных частей проекта [102].

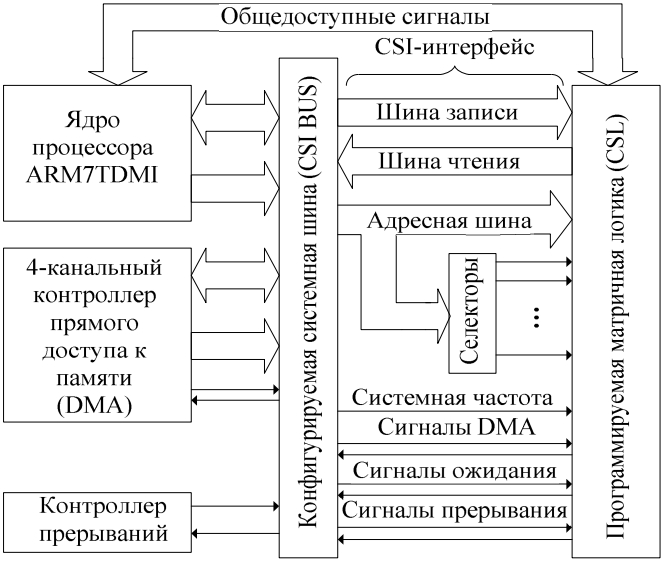


Рис. 4.8. Конфигурируемая системная шина и ее окружение

Адресные селекторы предназначены для обеспечения доступа к массиву программируемой логики со стороны процессора или контроллера DMA без использования ресурсов собственно встроенной матрицы FPGA. Селектор (рис. 4.8) представляет собой гибкую аппаратную схему сравнения адреса в диапазоне 4 Гбайт с числом, характеризующим этот ресурс (Match Value) с учётом выравнивания (Byte Enable [3:0]) и направления доступа (чтение или запись). После настройки селектора становится известным его адрес, который используется при разработке математического обеспечения и характеризует программно-доступный ресурс.

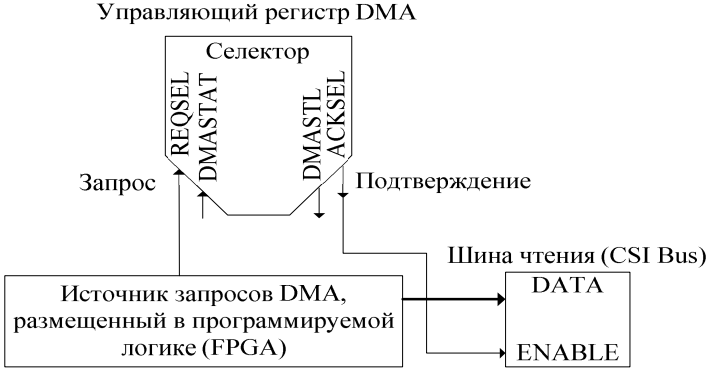


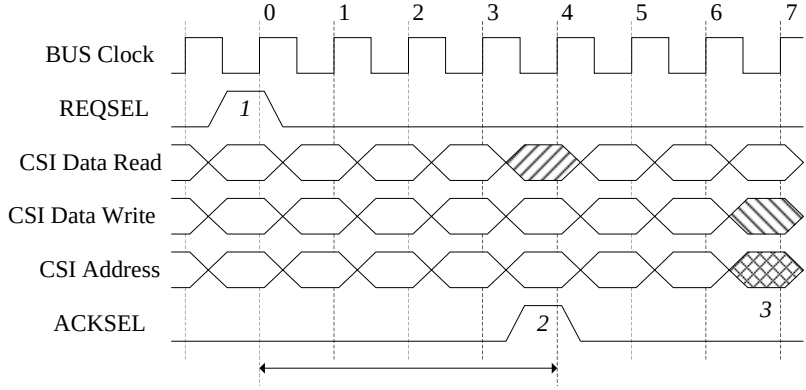
Рис. 4.8. Структура селектора DMA при записи в память

Выходными сигналами селектора являются сигнал разрешения установки данных на конфигурируемую шину со стороны программируемой логики (сигнал чтения) и сигнал стробирования для ввода данных с шины в массив FPGA (сигнал записи). Обычно такая транзакция проходит за один период тактовой частоты, но если устройство работает асинхронно, то используется механизм ожидания готовности устройства, что весьма удобно для подключения приборов с невысоким быстродействием.

Селектор DMA, предназначенный для переноса данных из (или в) массива конфигурируемой логики, дополнительно имеет вход требования цикла передачи (REQSEL) данных и выход предоставления такой операции (ACKSEL). На рис. 4.9 приведена временная диаграмма взаимодействия селектора DMA, системной шины и программируемой логики при записи информации в память.

Количество селекторов определяется типом кристалла и позволяет, например, легко разместить в системе на кристалле 50 программно-доступных для чтения и записи 32-битных регист-

ров, не используя аппаратные ресурсы собственно матрицы логики. Таким образом, доступные механизмы межблочного взаимодействия, в частности адресные селекторы, действительно интегрируют ресурсы и превращают разрозненные элементы в элегантную систему.



Минимальная задержка между сигналами REQSEL и ACKSEL 4 такта

- считываемые данные из CSL по каналу DMA

- данные, записываемые в память

- адрес памяти, формируемый контроллером DMA

Рис. 4.9. Временная диаграмма записи в память

Еще одним общесистемным узлом является четырехканальный контроллер прямого доступа в память. Каждый канал контроллера – это самостоятельное активное устройство на шине, передающее блочную информацию без участия процессора в режимах: «память–память», «память–массив программируемой логики», «массив программируемой логики–память». Если достаточно тривиальный механизм пересылок «память–память» значительно увеличивает системное быстродействие, инициируется ядром и похож на подобные механизмы в других процессорах, то два других режима являются более интересными и новыми для практики.

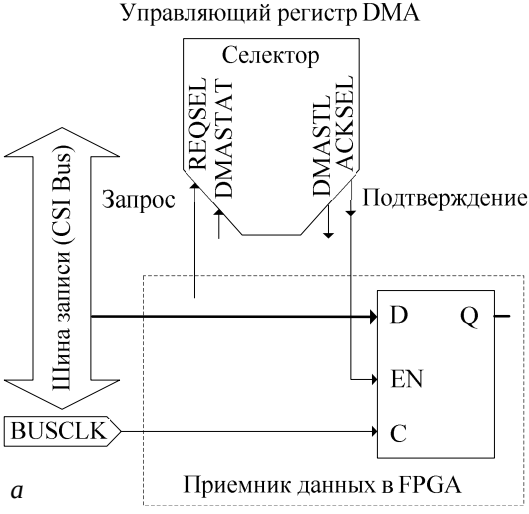
На рис. 4.10 приведена схема взаимодействия программируемой логики селектора DMA и системной шины при выполнении передачи «массив программируемой логики–память» и соответствующая временная диаграмма работы.

Аппаратная логика в соответствии со своим алгоритмом функционирования в произвольный момент (1) формирует сигнал REQSEL, который информирует выбранный канал DMA о готовности данных для передачи в память по текущему для

выбранного канала адресу записи. Спустя не менее четырех периодов тактовой частоты арбитр шины предоставляет временной интервал контроллеру DMA для чтения данных (2), при этом вырабатывается сигнал ACKSEL, который разрешает массиву FPGA выставить данные на системную шину. Эти данные фиксируются контроллером DMA и в момент времени (3) передаются, например, в контроллер внешней памяти для записи в динамическое ОЗУ.

В случае отличающейся разрядности источника данных и разрядности SDRAM соответствующие дополнительные транзакции добавляются автоматически контроллером внешней памяти доступно для разработчика системы. Контроллер DMA позволяет накапливать последовательные запросы REQSEL от матрицы логики и в дальнейшем также последовательно переносить данные в память. Количество запросов, полученных контроллером DMA, но реально не переданных, содержится в 10-битном регистре и программно доступно для анализа.

В случае DMA-пересылки «память–массив программируемой логики» (рис. 4.10) инициатором обмена опять является аппаратное устройство, размещенное в FPGA, которое в момент времени (1) формирует сигнал REQSEL, поступающий в соответствующий канал контроллера DMA. В момент времени (2) информация считывается из внешней памяти, а в момент времени (3) считанные данные записываются, например в регистр, размещенный в программируемой логике.



а

б

Рис. 4.10. Структура *а* и временная диаграмма *б* взаимодействия селектора DMA, программируемой логики и системной шины при чтении из памяти

Диаграммы на рис. 4.9, 4.10 – идеализированные, в зависимости от загрузки шины задержки могут быть больше, но, не смотря на это, поток информации достигает своего приёмника за счёт высокой тактовой частоты и механизмов буферизации. Для повышения достоверности передачи существует аппаратный узел подсчёта контрольной суммы при DMA-передаче, механизм обнаружения и генерации повторных запросов. За одну пересылку контроллер DMA способен перенести между источником и приёмником информации до 64 Кбайт.

Наиболее простой режим работы контроллера DMA заключается в периодической инициализации процессором механизма DMA (direct mode). При больших объемах пересылаемой информации, например при вводе видеоданных, такой постоянный контроль над пересылками может быть затруднительным. В этом случае разработчику предоставляется еще один режим работы DMA-контроллера descriptor mode, при котором в любой доступной памяти создается последовательность команд инициализации DMA-посылок, фактически вторая (третья, …, пятая) параллельно выполняющаяся программа. Контроллер шины самостоятельно, без участия ядра процессора, по завершении текущей DMA-пересылки выбирает из памяти следующее двойное слово, определяющее

новые параметры инициализации DMA-посылки, и вновь инициализирует массовую передачу данных. Такое решение позволяет значительно разгрузить процессор от рутинной работы по пересылке информации, сведя её к периодическому мониторингу.

Реконфигурируемые системы на кристалле нашли практическое применение в формировании и обработке изображений. Так, блок (рис. 4.11), входящий в состав трёхканальной телевизионно-компьютерной система анализа волнового фронта [80, 121], включающий СБИС реконфигурируемой СнК, обеспечивает формирование изображения, его ввод, накопление, предварительную обработку, управление и канальное кодирование для передачи по сети Ethernet. Применение СнК совместно с СБИС аппаратного стека протоколов TCP/IP позволило достичь рекордных скоростей передачи по сетевому интерфейсу Ethernet 10/100Mb для встраиваемых приложений [104]. Возможности логического управления СБИС СнК позволяют реализовать метод высокоточного измерения среднего сигнала темнового тока ПЗС в астрономических телевизионных камерах [80] и другие методы обработки видеоинформации, рассмотренные в главе 2.

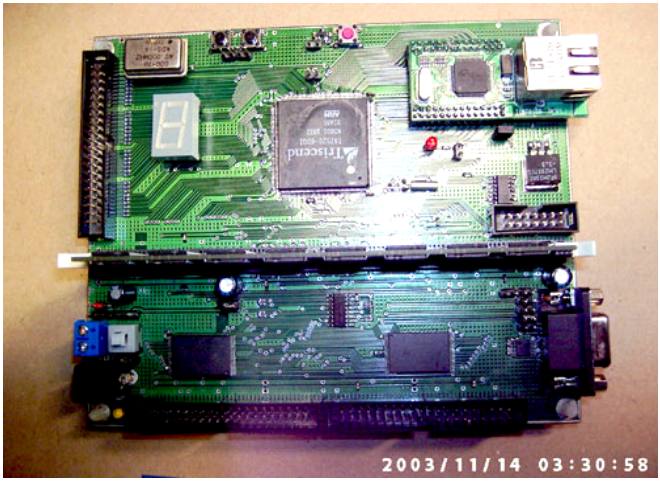


Рис. 4.11. Устройство ввода и обработки видеоинформации на основе технологии система на кристалле

Таким образом, в реконфигурируемых системах на кристалле обеспечивается глубокая интеграция микроконтроллерного ядра и разнообразных «жёстких» встроенных периферийных устройств

через системную шину с массивом программируемой логики. Для разработчика это означает гибкую возможность реализации различных аппаратно-программных проектов в едином окружении, а при смене структуры аппаратной обработки новая реализация размещается в том же кристалле без изменения печатной платы. Появление устройств класса реконфигурируемая СнК составляет новый этап интеграции в электронике. Дальнейшее развитие этого направления позволит решать более сложные задачи, а соответствующие изделия станут более распространенной элементной базой, чем ранее стала микропроцессорная технология.

Видеосистемы на кристалле для формирования телевизионного изображения. Появление СнК характерно не только для вычислительной техники. Телевизионная техника на рубеже веков переживает аналогичный этап революционного изменения технологии. До него формирование и обработка изображения разделялись, в том числе и по технологии реализации. Современный этап характеризуется единством технологии большинства подсистем. Она имеет целью не только уменьшение массы, габаритов и энергопотребления, но и перенос центра сложности обработки видеоинформации и принятия решений из внешнего вычислителя непосредственно в телекамеру. Это в корне меняет роль телекамеры в системе – из простого формирователя изображения она превращается в интеллектуальный источник видеоинформации и решений. Телекамера и сейчас уже методологически неотделима от компьютера, а скоро она будет составлять единое технологическое и методологическое целое с компьютером или процессором цифровой обработки [37, 38, 39, 109]. Движение к такому будущему идёт с двух сторон. Со стороны средств цифровой обработки информации появилось новое направление – система на кристалле, сочетающее различные парадигмы такой обработки. Со стороны средств формирования изображения происходит объединение на одном кристалле датчика изображения и аналогового видеопроцессора. Это привело к появлению новой элементной базы – однокристальных цифровых телекамер или видеосистем на кристалле (ВСнК, VSoC Vision System on Chip,). Такие камеры выпускаются многими производителями, число которых превышает 50 и непрерывно растёт.

Наиболее приспособлена к такой интеграции оказалась КМОП технология. Приведённые в разделе 1.2 варианты структур элемента КМОП матрицы охватывают широкий диапазон сложности – от простейших до использующих сотни транзисторов обеспечивающих

реализацию в каждом элементе многоразрядного АЦП [26, 37, 53]. Использование докоммутационного усиления в КМОП фотоприёмниках завершит твердотельную революцию не только в обычных телекамерах, но и в приборах ночного видения, пока использующих электронно-оптические (вакуумные) преобразователи. Есть предпосылки создания однокристальных цифровых камер с устройствами аналоговой, цифровой и нейроподобной обработки изображения [26, 38, 110, 188]. Перспективное направление разработок – создание интеллектуальных камер, которые позволят реализовать системы технического и искусственного зрения, сопоставимые по характеристикам с биологическим зрением [7].

Основные процессы формирования и преобразования изображений показаны на рис. 4.12. Структуры конкретных видеосистем на кристалле объединяют указанные этапы преобразований в различные конфигурации

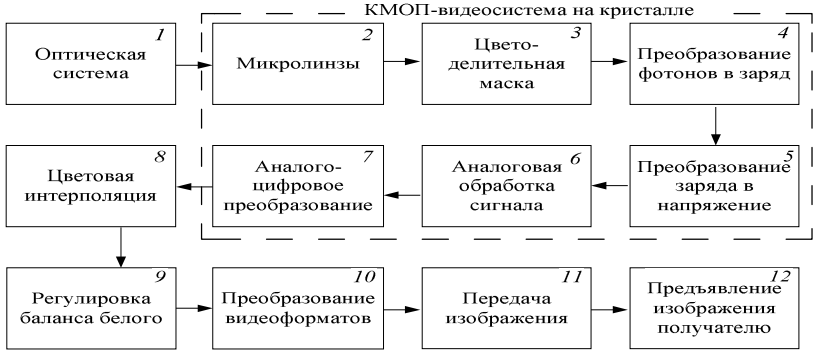


Рис. 4.12. Видеосистема на кристалле в типовой телевизионной системе

Функционирование телевизионной аппаратуры, например параметры накопления, развёртки, усиления и кодирования задаётся программным путем с помощью конфигурационных регистров.

Рассмотрим характеристики и структурные схемы ряда однокристальных камер VGA-разрешения, обеспечивающих полную обработку видеосигнала с цифровым выходом. Основное внимание уделяется не фотоэлектрическим характеристикам преобразователя свет–сигнал, входящего в состав ВСнК, рассмотренным в гл. 1, а особенностям аппаратно-программного окружения, которые превращают эти приборы в новый класс устройств формирования и обработки изображений. Первый представитель ВСнК – VS6552, используемый в мобильных приложениях (рис. 4.13). Она включает в включает в свой состав КМОП матрицу, формирователь адреса строк

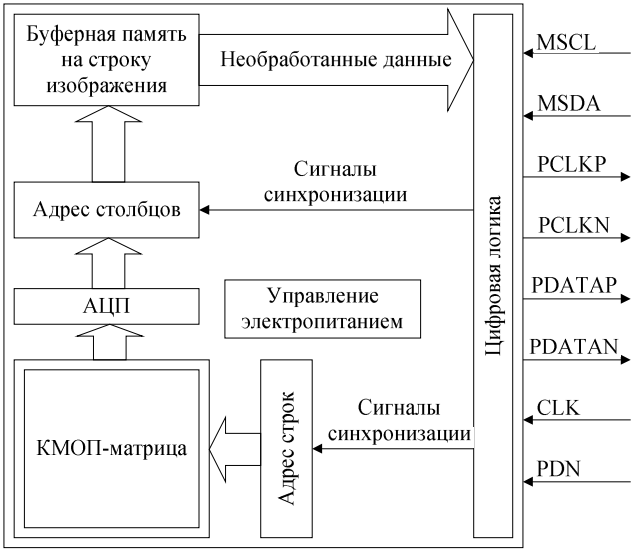


Рис. 4.13. Видеосистема на кристалле VS6552 фирмы STMicroelectronics

и столбцов, аналого-цифровой преобразователь АЦП, буферную память на строку изображения, схемы интерфейса и управления питанием. Данная видеосистема на кристалле характеризуется простотой реализации, её отличительными особенностями по сравнению с аналогами являются:

- совмещение в едином модуле объектива и формирователя изображения;
- применение АЦП с последовательным выходом и передачей видеоинформации по дифференциальному высокоскоростному интерфейсу LVDS.

Видеодатчик VS6552 по интерфейсу LVDS подключается к специализированному процессору STV0974, который обеспечивает обработку информации о цвете согласно кодирующему свето-фильтру RGB, управление экспозицией и организацию вывода видеоинформации в форматах YCbCr, RGB или JPEG.

На отечественном рынке известны однокристальные камеры производства OmniVision Technologies [25, 36], архитектуру которых рассмотрим на примере OV7640 (рис. 4.14). Из внешних компонентов для подключения микросхемы требуются только фильтрующие конденсаторы. В состав OV7640/OV7141 входят:

- матрица изображения (полное разрешение 682×482 элемента изображения);

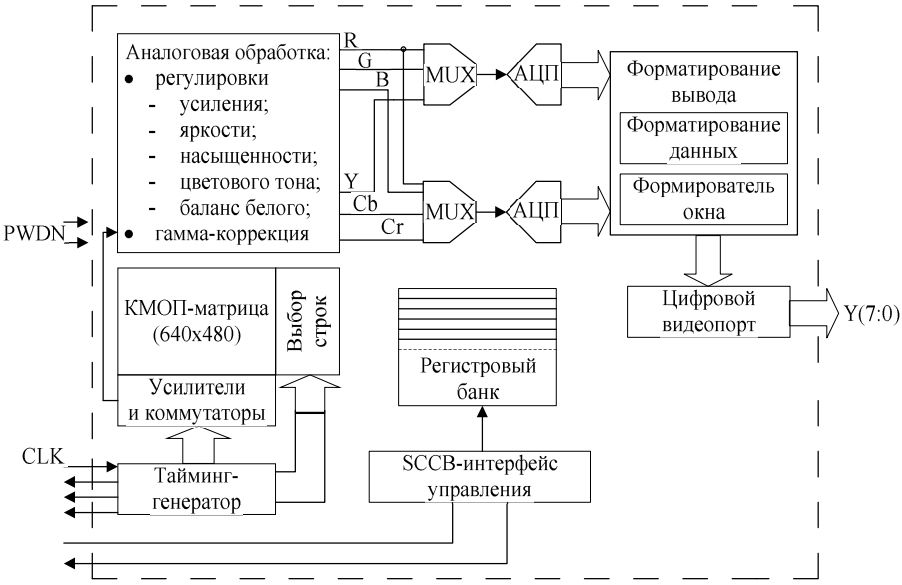


Рис. 4.14. Видеосистема на кристалле OV7640 фирмы OmniVision

- схемы аналоговой обработки;
- схема тактового генератора, которая осуществляет формирование кадра изображения в формате VGA или QVGA, управление временем экспозиции, формированием внешних синхронизирующих импульсов. Для своей работы схема тактового генератора использует внешнюю синхронизацию частотой до 27 МГц;
- блок форматирования с цифровым видеопортом, который обеспечивает постобработку цифрового видеосигнала, например вывод определенной зоны светочувствительной матрицы (windowing) и трансляцию выходного двоичного кода;
- трёхпроводной интерфейс управления SCCB и набор регистров, которые обеспечивают цифровую конфигурацию однокристальной видеокамеры.

Общее число байт информации, которой управляется микросхема V7640/OV7141, превышает 80. Кроме стандартных настроек с помощью интерфейса управления можно получать статистику о среднем значении сигнала изображения в каналах яркости и цветности, изменять направление развёртки и её вид (чересстрочная или прогрессивная), управлять энергопотреблением микросхемы. Это позволяет организовывать гибкие перестраиваемые системы получения изображения в зависимости от текущей

цели аппаратуры [96, 97]. Такая ВСнК была использована в аппаратуре мобильной телевизионной связи, описанной в разделе 4.5.

ВСнК фирмы Kodak (рис. 4.15) позиционируется как однокристальная камера с большим динамическим диапазоном – до 110 дБ. Полный формат светочувствительной матрицы составляет 648×448 элементов, по 4 пиксела с каждой стороны матрицы используются для формирования оптического «чёрного». В начале каждого кадра производится очистка накопленных зарядов, после окончания времени накопления зарядовое изображение с помощью двойной коррелированной выборки преобразуется в электрический сигнал изображения. Далее аналоговый сигнал подаётся на два каскада усиления: общий и избирательный по цвету. Коэффициент усиления каждого каскада регулируется цифровым кодом в пределах 0…15 дБ. Усиленный сигнал каждой цветовой составляющей через аналоговый мультиплексор поступает на АЦП. Цифровой 12-разрядный сигнал изображения проходит блок коррекции дефектных пикселов, схемы регулировки уровня «чёрного» и поступает на выход кристалла в программируемом формате 8, 10, 12 бит.

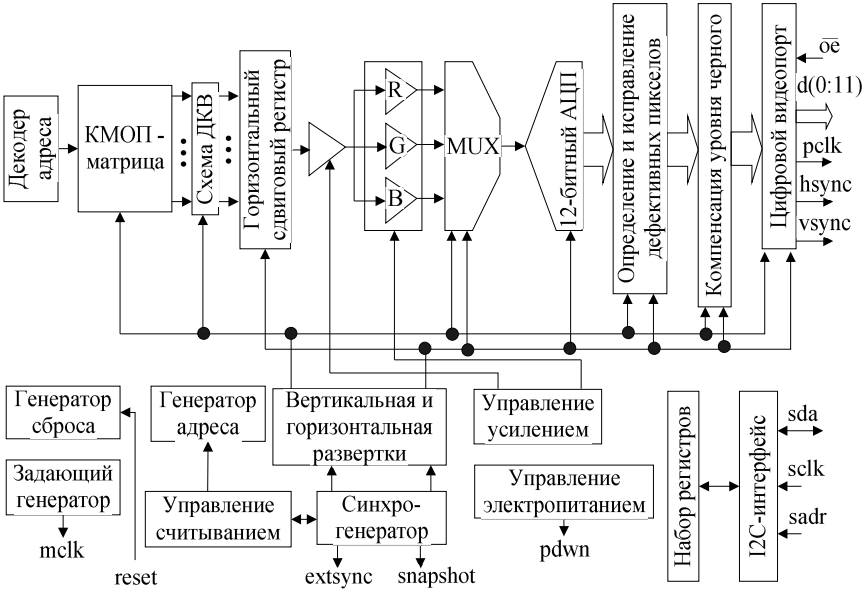


Рис. 4.15. Видеосистема на кристалле КАС-9628 фирмы Kodak

Блок коррекции может программно отключаться, а логика его функционирования включает в свой состав 4 программируемых регистра пороговых уровней. Коррекция уровня «черного» основана на вычитании из сигнала изображения оценки его уровня. Отдельный программно доступный регистр определяет смещение уровня при аналого-цифровом преобразовании. Блок синхронизации тактируется от внешнего сигнала частотой до 48 МГц и содержит схемы вертикальной и горизонтальной развёртки. Программируемые регистры определяют окно вывода информации, которое может располагаться в произвольном месте светочувствительной матрицы. Программным путем определяются несколько режимов субдискретизации, позволяющих прореживать пикселы при формировании цифрового изображения, и режимы прогрессивного или чересстрочного разложения. Результирующая частота кадров определяется входной тактовой частотой, размерами выводимого окна, видом субдискретизации и может регулироваться в широких пределах. Высокая кадровая частота (до 1 кГц и выше) весьма актуальна при классификации динамических изображений, например в адаптивной оптике. Один и тот же прибор ВСнК может использоваться и в сверхбыстродействующих, и в малокадровых телевизионных системах. Для реализации малокадрового режима характерно использование программирования синхронизации для покадровой съёмки. Логика работы такого режима определяется входным сигналом «Запрос кадра», программируемой задержкой на несколько кадров и программируемым временем накопления. Выходной сигнал «Захват кадра» имеет высокий активный уровень во время формирования кадра изображения, что может быть использовано для организации дополнительной подсветки или механического затвора.

Для достижения широкого динамического диапазона освещённостей может использоваться программно-управляемая кусочно-линейная аппроксимация свет–сигнальной характеристики, с помощью которой можно реализовать квазиоптимальное квантование для мультипликативного фотонного шума (раздел 3.2). Важным узлом ВСнК является набор регистров конфигурации и интерфейс управления. Общее количество программно доступных байт в системе более 40, они доступны по двухпроводному интерфейсу I^2C.

ВСнК фирмы Micron MT9V112 (рис. 4.16) ориентирована на применение в составе мобильного оборудования, и большинство функций обработки сигнала в ней выполняются автоматически

(без настройки). К ним относятся компенсация уровня «чёрного», автоматический баланс «белого», подавление эффекта мелькания, вызванного люминесцентными лампами, коррекция дефектных пикселов, цветовая обработка. ВСнК MT9V112 содержит 10-разрядный АЦП для первичного квантования и внутренней обработки сигнала. На выходы подаётся 8 разрядный код. Синхронизируется MT9V112 от источника тактовой частоты 27 МГц, блок развёртки обеспечивает прогрессивное разложение с частотой кадров 15 Гц при формате 640×480 элементов, а при частоте кадров до 90 Гц реализуется формат 320×240 элементов.

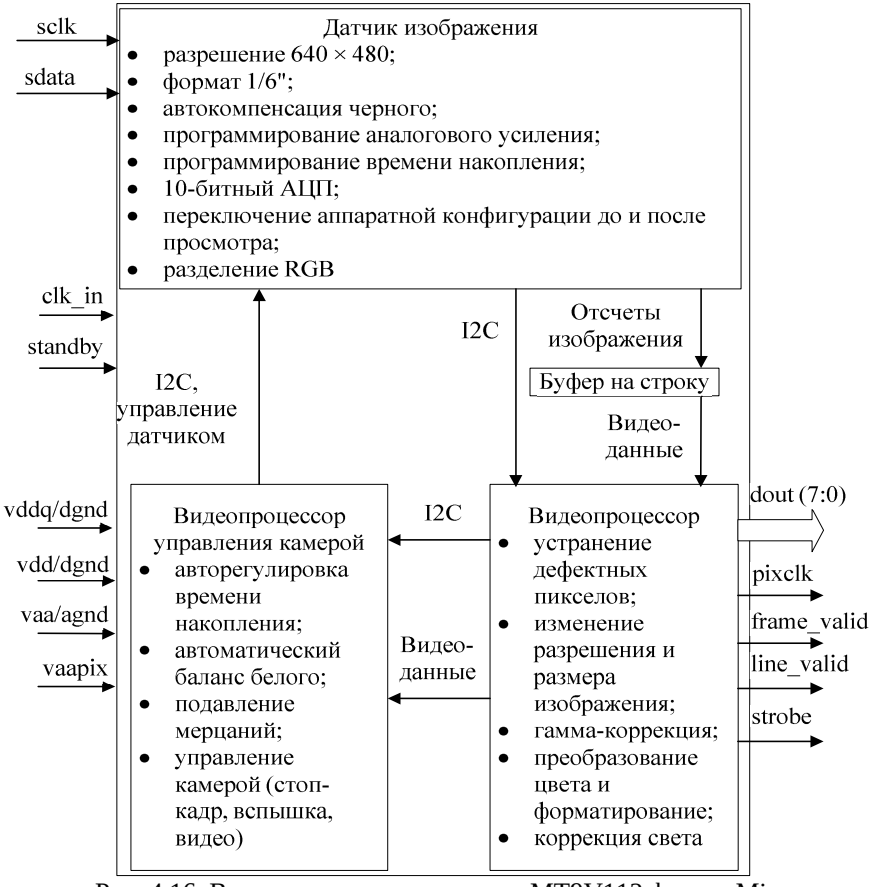


Рис. 4.16. Видеосистема на кристалле MT9V112 фирмы Micron

К конфигурируемым через интерфейс I^2C относятся такие параметры, как усиление, опорное напряжение АЦП, настройка электронного затвора. Подобно микросхеме КАС-9628 в MT9V112 имеется возможность организации покадровой съёмки с одновременной синхронизацией осветителя, например лампы-вспышки.

К видеосистемам на кристалле повышенного разрешения (SVGA) относится микросхема ZDM33220 (рис. 4.17).

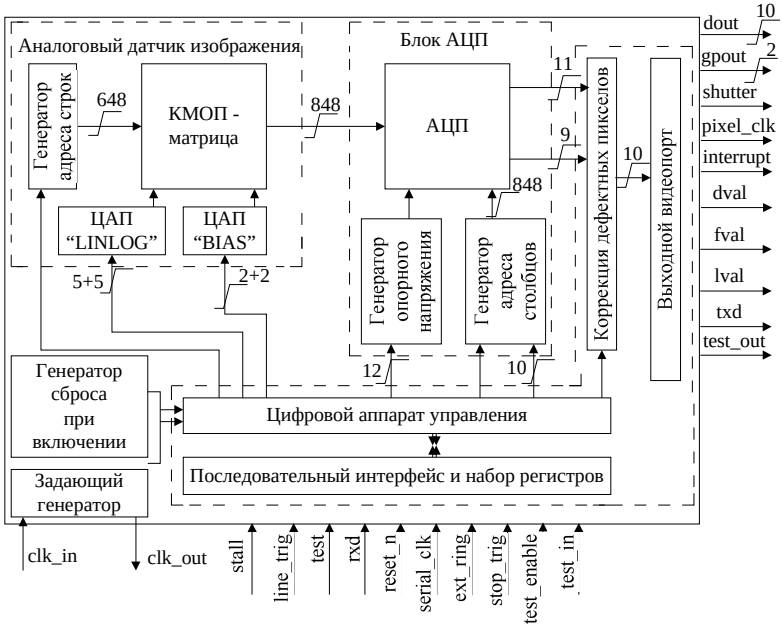


Рис. 4.17. Видеосистема на кристалле ZDM33220 фирмы ZMD AG

От описанных ранее она отличается большим размером элемента изображения и оптическим форматом ⅔ дюйма. Структурная схема ZDM33220 включает КМОП датчик изображения, программируемые схемы развёртки (генераторы адреса строк и столбцов), цифро-аналоговые преобразователи (ЦАП) коррекции свет–сигнальной характеристики для увеличения динамического диапазона (LinLog), ЦАП установки уровня «черного» (Bias), генератор опорного напряжения, 12-битный АЦП видеосигнала, схему коррекции дефектных элементов, выходной регистр со схемой синхронизации, схему последовательного интерфейса и регистровый банк, программируемый цифровой автомат управления, схему тактового питания и схему управления энергопитанием. Данная ВСнК имеет

расширенные варианты программной конфигурации. Важной функцией этой ВСнК является наличие до трёх программно-управляемых окон изображения. При этом возможен обмен пространственного разрешения на временно́е: частота считывания такой зоны определяется частотой опроса элемента и размером зоны и может достигать нескольких килогерц.

В системах на матричных фотоприемниках на основе КМОП технологии из-за асинхронности накопления возникают искажения изображений движущихся объектов. Такая асинхронность накопления при наблюдении малоразмерных подвижных объектов перестаёт быть недостатком, так как уменьшает временные задержки при считывании сигналов и промах при слежении за ними [4, 83].

Революционность перехода к видеосистемам на кристалле проявляется даже в том, что изменяется подход к *определению чувствительности*. Для ВСнК MT9M413 изготовители декларируют чувствительность 1600 *бит/(люкс×с)* вместо использования традиционной размерности *Вольт/(люкс×с)* [24, 38, 39].

Важнейшим свойством видеосистем на кристалле явилась возможность реализации *множественных каналов связи* между фотоприёмной матрицей и вычислителем, так как наращивание количества каналов связи при раздельном исполнении фотоприёмной матрицы и устройства обработки неизбежно будет ограничиваться технологически допустимым количеством выводов микросхем и ростом задержек. Если в ПЗС число таких каналов в уникальных приборах доводили до 16, то в ВСнК это число легко может составить число столбцов. Например, в приборе MT9M413 фирмы Micron (см. раздел 1.2) на каждый столбец фотоприёмной КМОП матрицы, приходится свой 10-битный АЦП. Между фотоприёмной матрицей прибора MT9M413 и внешним вычислителем организовано 10 каналов связи, работающих со скоростью до 66 МГц.

Такая структура видеосистемы на кристалле несмотря на кажущееся обилие АЦП резко снижает требования к подобным преобразователям по быстродействию. При этом суммарная производительность достигает 660 МГц. Каждый АЦП имеет программную настройку и калибровку, кадровая частота при сканировании всей матрицы «активных пикселов» (1280×1024) достигает 500 Гц, а в режиме частичного сканирования – 4000 Гц.

Видеосистемы на кристалле сочетают в рамках одного кристалла фотоприёмную матрицу и процессор обработки. В связи с этим актуально распределение общих ресурсов (площади, выводов кристалла, потребляемой мощности) между этими разно-

родными элементами, каждый из которых тем лучше выполняет свои функции, чем большая сложность ему доступна. Вопросы распределения таких общих ресурсов являются актуальным предметом системного уровня проектирования СБИС (см. раздел 2.1).

Появление нового класса видеосистем на кристалле ведёт к изменению метода проектирования телевизионных устройств. Если на начальном этапе развития телевидения согласно методологии создателя электронного телевидения В. К. Зворыкина в проектировании преобладали методы физики, то новый этап порождает новый метод проектирования, который характеризуется высокой скоростью создания систем благодаря использованию средств САПР, резким возрастанием роли системного подхода и значимости программно-алгоритмического обеспечения.

4.3. САПР телевизионно-компьютерных систем с использованием систем на кристалле и видеосистем на кристалле

Современные методики проектирования устройств обработки сигналов на базе сложных программируемых микросхем базируются на широком применении САПР на большинстве этапов реализации алгоритмов обработки сигналов [10, 14, 156, 163].

На сегодня реально не существует ни одной интегрированной системы САПР, способной эффективно решать задачи проектирования во всем диапазоне возможных способов реализации. Производители интегрированных САПР поддерживают весь маршрут проектирования (от спецификации до физической реализации), но только для определенного класса ИМС. Все интегрированные системы опираются на определённый способ физической реализации и выстроены по вертикальному принципу. Интегрированные маршруты проектирования состоят из отдельных этапов, каждый из которых имеет свой пользовательский интерфейс, систему команд и базу данных. Основные этапы проектирования (спецификация проекта, моделирование, верификация, логический синтез, физическая реализация) определены достаточно хорошо, что позволяет без потери эффективности и дополнительных затрат на интеграцию комбинировать средства проектирования различных производителей, выбирая те, которые наилучшим образом отвечают решаемым задачам.

Для обеспечения унификации маршрутов проектирования необходимо отойти от сегментирования средств разработки по способам физической реализации или производителям САПР и перейти к их делению по классам задач, решаемых на отдельных этапах маршрута проектирования [156, 163]. В этом обобщённом случае выделяются следующие уровни проектирования:

- системный (иначе архитектурный) уровень, где производится определение и спецификация основных функций проектируемого устройства обработки сигналов, создание исполняемой системной модели;
- алгоритмический уровень, заключающийся в описании и анализе алгоритмов, протоколов, сценариев системы и общих функциональных характеристик;
- функциональный уровень (иначе уровень регистровых передач), связанный с разработкой спецификаций функциональных блоков на языках высокого уровня описания аппаратуры – Verilog или VHDL, и их верификацией;
- логический уровень (иначе вентильный уровень), полагающий переход в элементный базис производителя БИС средствами логического синтеза;
- схемотехнический уровень, на котором выполняется реализация проекта на физическом уровне с детальным размещением на конкретном кристалле, трассировкой, генерацией тестовых структур и верификацией топологии.

Модель представления проекта системы на кристалле содержит три области описания (поведенческая, структурная и физическая), представляющие различные аспекты единого процесса проектирования [163].

Поведенческая область определяет поведение проектируемой СнК или её функциональные возможности. Структурная область определяет логическую структуру в виде комплексного набора взаимосвязанных элементов и является интегрирующим фактором между поведенческим и физическим описанием. Физическая область определяет реализацию абстрактных составляющих структурной области виде конкретных физических компонент (табл. 4.3).

Системный и алгоритмический уровень проектирования представляет интерес для разработчика аппаратуры. Например, рассмотренный в разделе 2.1 вопрос о распределении общей площади видеосистемы на кристалле между фотоприёмником и вычислителем для выполнения целевой функции устройства решается в ходе системного проектирования (физическая область). Схемотехнический уровень, включая топологию, является сферой интересов проектировщика БИС, а функциональный и логический уровни – областью совместных интересов сторон. Такое представление преломляется в создании специализированных дизайн-центров системного и кристального уровней, вооруженных системами САПР, ориентированных на изготовление БИС.

Таблица 4.3.

Уровни проектирования	Области представления проекта		
	Поведенческая	Структурная	Физическая
Системный	Описание функционирования системы на языке внешнего проектирования	Крупные структурные элементы (процессоры, память, контроллеры)	Топология кристалла, ограничения на мощность, площадь количество выводов и т. д.
Алгоритмический	Описание проекта на алгоритмических языках	Аппаратные модули, соответствующие процессам преобразования	Группировка функционально подобных операций, макроблочная декомпозиция
Функциональный	Регистровые передачи на уровне абстракций	Компоненты (АЛУ, регистры) необходимые для организации регистровых передач	Детализация топологии кристалла
Логический	Логические выражения, таблицы истинности	Триггеры, вентили, ключи	Элементарные логические ячейки
Схемотехнический	Дифференциальные уравнения отдельных компонентов	Элементарные компоненты (транзисторы, резисторы и т. д.)	Геометрические образы слоев кристалла, фотошаблоны

До недавнего времени задачи системного уровня проектирования решались без специальных программных средств. Переход к методологии проектирования систем на кристалле требует автоматизации такого процесса ввиду следующих причин:

- сложность систем на кристалле не позволяет проверить их работоспособность, используя только аналитические методы;
- при спиралевидной модели проектирования (рис. 4.18) возникают рекурсивные возвраты с нижних уровней проектирования к верхним, что предполагает использование унифицированных программных средств поддержки разработки.

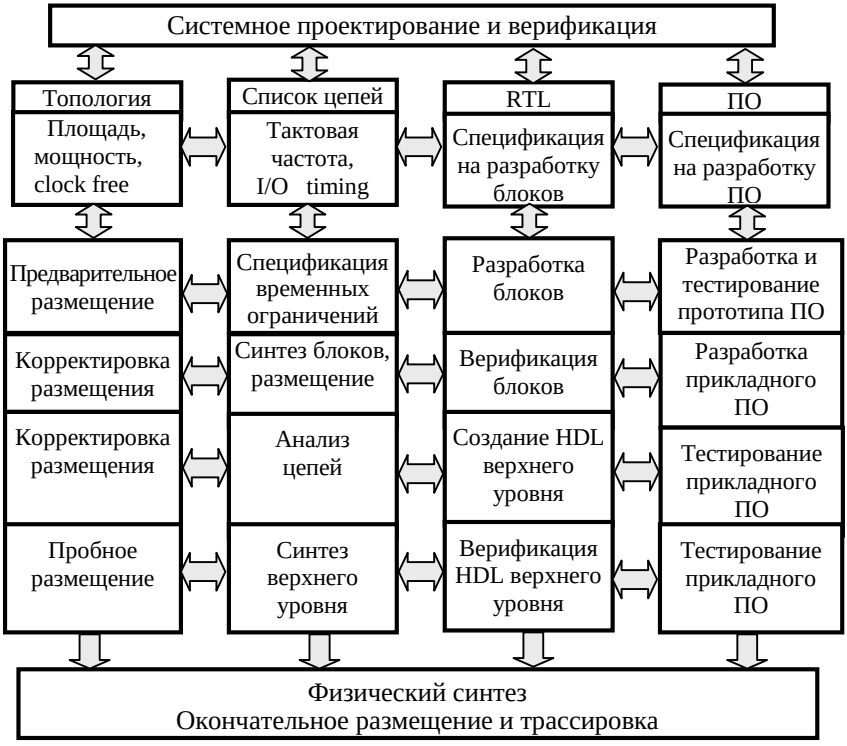


Рис. 4.18. Спиралевидная модель процесса проектирования с использованием технологии «система на кристалле»

Начальный этап разработки включает *рекурсивные* итерации, верификацию и уточнение набора спецификаций до такой степени детализации, чтобы на их основе можно было создавать описание процесса обработки сигналов на уровне регистровых передач. Быстрая разработка чётких, полных и последовательных спецификаций – отдельная сложная, трудоёмкая и ответственная задача. Спецификации определяют поведение системы и подразделяются на функциональную, определяющую интерфейс системы, и архитектурную, описывающую состав и взаимодействие между частями блока. Спецификации разрабатываются как для программной, так и аппаратной частей проекта.

Для программной части необходимо описать выполняемые функции, временные параметры, быстродействие, интерфейс к аппаратной части, структуру и ядро. Для аппаратной части спецификация должна включать внешний интерфейс к другим

блокам (выводы, шины, протоколы), временные параметры быстродействия, особенности физической реализации (площадь кристалла, энергопотребление).

На этапе концептуального проектирования системы, выбора её параметров и алгоритма функционирования наиболее подходящими для анализа протекающих процессов являются пакеты MatCAD, MatLab, LabView. Подобные САПР позволяют имитировать работу системы и обладают высокоуровневыми средствами анализа процессов, происходящих в модели. На этом этапе можно выделить три уровня детализации:

- общую концепцию построения системы и операционной среды окружения, определение статических и динамических сценариев, планирование целевых задач;
- архитектуру системы с моделированием и анализом производительности систем, сетевых архитектур и протоколов, пропускной способности каналов;
- микроархитектуру системы с моделированием и анализом алгоритмов, протоколов, конкретизацией методов обработки и программно-аппаратной декомпозицией.

Первые два уровня относятся к задаче концептуального построения и анализа системы и к ним применимы базовые понятия системотехники о внешнем и внутреннем проектировании. На концептуальном уровне при проектировании ВСнК необходимо применение математической теории связи с учетом качества информации (см. гл. 3). Третий уровень (микроархитектуры) непосредственно связан с последующим этапом функционального проектирования. Процесс проектирования на концептуальном уровне (рис. 4.19.) начинается с формулировки задач и требований к СнК.

На этапе функционального моделирования определяется правильность логического и схемотехнического построения системы без учёта реальных временных задержек распространения сигнала в кристалле.

На фазе временного анализа, после выполнения процедур размещения и разводки, САПР позволяют извлечь информацию о временных задержках и при помощи внешних программных средств промоделировать работу устройства [94].

На этапах функционального и временного моделирования целесообразно использование пакетов САПР моделирования фирм Aldec, Mentor Graphics, Cadence. На этапе физической верификации и натурных экспериментов работоспособность устройства или отдельных его частей можно проверить в реальном времени, загрузив выходной конфигурационный файл проекта на целевую плату прямо из инструментального компьютера.

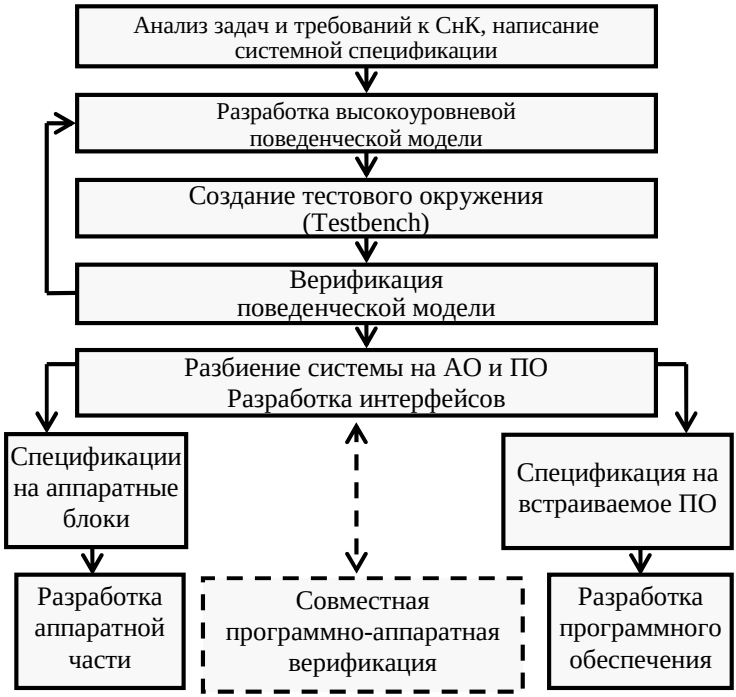


Рис. 4.19. Маршрут концептуального проектирования системы на кристалле

Современные требования к процессу разработки цифровых устройств заключаются в *возрастании сложности алгоритмов* обработки при ограничениях на время проектирования и выпуска продукции. Для традиционного подхода к разработке сложных аппаратно-программных изделий характерно последовательное выполнение во времени типовых проектных процедур. Такой характер проектирования не только существенно задерживает начало этапа разработки программной части системы, но и значительно увеличивает риск итерационного возврата к повторному выполнению аппаратной части. Недостатки традиционного подхода нельзя устранить только сокращением времени выполнения отдельных этапов, необходимо изменение методологии проектирования. Такая методология называется сопряженным проектированием [12] и заключается в совмещении во времени этапов моделирования, разработки и отладки аппаратной и программной составляющих изделия. Реконфигурируемые СнК в совокупности со средствами автоматизации уже в настоящее время

позволяют практически реализовать все этапы сопряженного проектирования [96, 97] (рис. 4.20). Использование такого подхода при фиксированной сложности изделия приводит к значительному сокращению времени проектирования и минимизирует риск проектных ошибок. Кроме того, как показал реальный опыт сопряженного проектирования, значительный положительный эффект даёт повторное использование ранее разработанных и верифицированных элементов проекта.

Сопряженное проектирование – процесс параллельного проектирования аппаратных и программных средств, при котором оценивается целесообразность выбора аппаратной или программной реализации определенного фрагмента проекта. Этот процесс также позволяет проектировщикам увидеть, как система могла бы работать с данным разделением аппаратных средств ЭВМ и программного обеспечения.

Сопряженная верификация – это проверка возможности совместного функционирования ПО и аппаратных средств ЭВМ [94, 100–102]. С другой стороны, во время этого процесса анализируется, будут ли одинаково решаться специфические задачи проекта при аппаратной или программной реализации.

Сопряженное моделирование – это совместное моделирование работы аппаратных и программных средств, при которой аппаратные средства устройств обработки описаны, например, в форме программных моделей. Возможны две трактовки понятия сопряжённого проектирования: либо это весь процесс проектирования, а совместная верификация и симуляция – его отдельные подэтапы, либо это начальный этап проектирования, результатом которого является декомпозиция проблемы [12, 98], а совместная верификация и симуляция – это последующие и самостоятельные этапы проектирования.

В качестве примеров маршрутов проектирования и САПР рассмотрим средства компаний Triscend (ныне Zylogic) и Cadence.

Для сопровождения проектирования устройств обработки сигналов на основе технологии РСнК производитель таких микросхем (ZyLogic) предлагает программную среду под названием «FastChip». Этот программный пакет обеспечивает управление всем процессом реализации устройств на основе технологии РСнК, имеет характерные черты других подобных САПР, а потому его возможности представляют интерес для анализа и исследования. На современном этапе наблюдается тенденция специализации средств автоматизированного проектирования

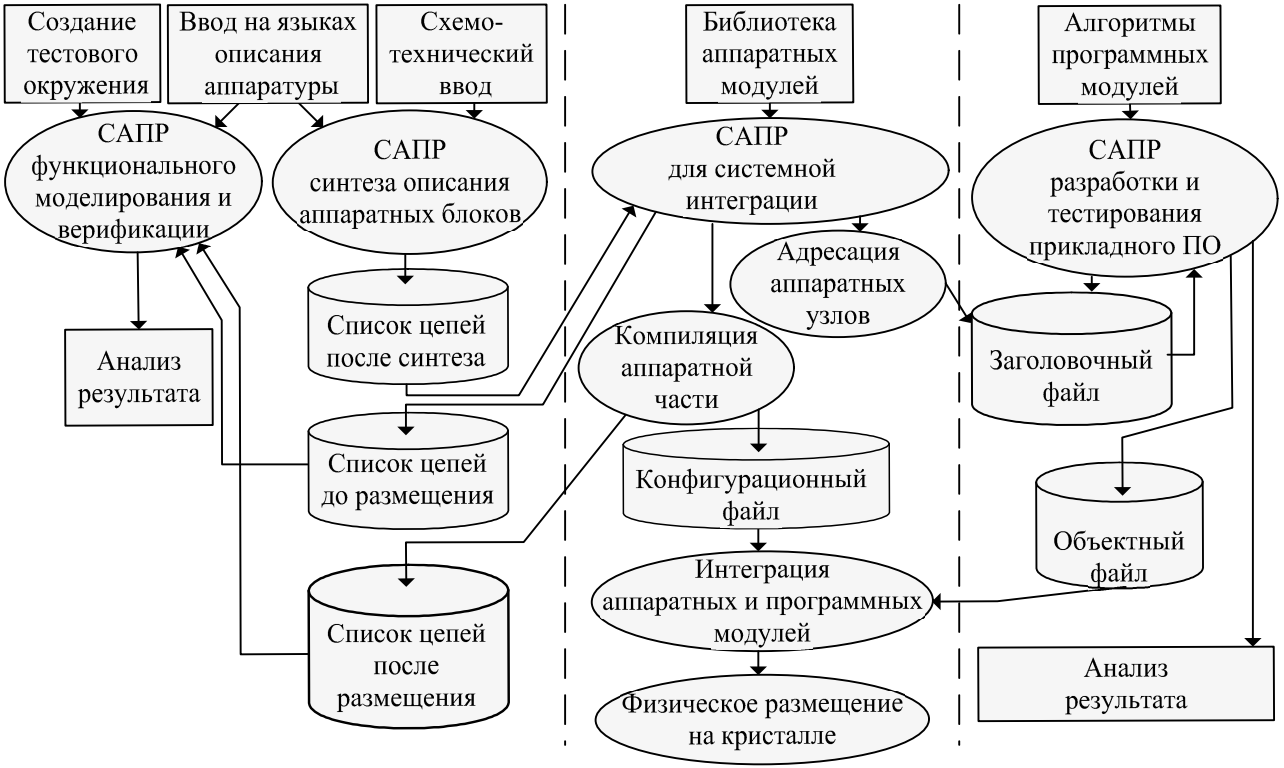


Рис. 4.20. Методология сопряженного проектирования применительно к реконфигурируемым системам на кристалле

электронных изделий, что является важнейшим условием эффективности разработки и ускорения выпуска продукции. Такой пример сочетания специализации, функциональности и открытости современных САПР демонстрирует указанная среда разработки.

Следует отметить, что идеология разработки устройств на основе РСнК отличается от проектирования устройств в многокристальном исполнении. Это связано с тем, что РСнК представляют собой приборы системного уровня интеграции, т. е. сочетают на одном кристалле как разнородные аппаратные структуры (процессор, программируемая логика, оперативная память), так и архитектурные элементы их взаимодействия (адресные селекторы, конфигурируемая системная шина). Разработка цифровых устройств на основе РСнК характеризуется интеграцией с помощью САПР двух потоков проектирования:

- реализации аппаратной составляющей системы на основе конфигурируемой системной логики CSL (ячейки матрицы FPGA и средства их межсоединений);
- разработки программной составляющей системы для встроенного микропроцессорного ядра, используя ресурсы как встроенных периферийных элементов, так и возможности интерфейса с конфигурируемой логикой.

Основные характеристики пакета FastChip Version 2.6 определяют методологию сопряженного проектирования. Пакет обеспечивает все этапы проектирования цифровых устройств и помогает достичь максимальной эффективности и производительности. САПР FastChip позволяет конфигурировать структурные составляющие всех частей СнК: элементы конфигурируемой системной логики, устройства интерфейса с внешней памятью, стандартные периферийные блоки (таймеры, порты ввода–вывода, контроллер прерываний), элементы синхронизации и межсистемного взаимодействия.

В качестве основных показателей САПР выделим следующие:

- поддерживаемые системой проектирования семейства ИМС;
- характеристика способов ввода аппаратной конфигурации;
- возможности компиляции аппаратной части проекта;
- средства разработки программной части проекта;
- средства верификации проекта;
- средства системной интеграции и использования САПР сторонних производителей.

Среда FastChip предоставляет разработчику как полноэкранный графический интерфейс, так и возможность работы в режиме командной строки. САПР обеспечивает средства взаимодействия с пакетами проектирования сторонних производителей разработки как для разработки аппаратной конфигурации для массива программируемой логики, так и программной составляющей для микропроцессорного ядра.

Микросхемы класса РСнК реализуют законченную микроэлектронную систему на одном кристалле, на котором размещаются стандартное микропроцессорное ядро, периферийные устройства и программируемая логика. В зависимости от разрядности используемого процессорного ядра существуют 8- и 32-битные семейства, различающиеся типом процессора, составом и характеристиками встроенных периферийных узлов.

Семейство 8-разрядных устройств РСНК Triscend E5 является исторически первым вариантом класса РСнК. Оно основано на широко известном высокопроизводительном ядре «turbo» 8032, совместимом по данным и коду с устройствами архитектуры MSC-51. Производительность ядра составляет до 10 MIPS при тактовой частоте 40 МГц. В состав 8-разрядных семейств входят следующие микросхемы: TE502S08, TE502S16, TE502S32, TE502S40, TE502S64. На кристалле, кроме обычной для архитектуры MSC-51 периферии, размещены следующие аппаратные узлы:

- дополнительный регистр-указатель данных DPTR;
- программируемый сторожевой таймер;
- блок сопряжения с внешней памятью, позволяющий без вспомогательных средств подключать устройства памяти как с параллельным, так и с последовательным интерфейсом;
- до 64 Кбайт внутренней оперативной памяти данных и программ;
- двухканальный контроллер прямого доступа в память;
- JTAG-интерфейс стандарта IEEE 1149.1.

Следующим шагом в развитии РСнК стало появление 32-разрядных семейств, выполненных на основе процессорного ядра ARM. Первое из этих семейств – A7S, описанное в разделе 4.2, его развитием и расширением стало A7V.

Методология проектирования устройств цифровой обработки на основе РСнК предполагает широкое применение средств автоматизации, поэтому актуальным является анализ

функциональности соответствующих САПР. Для разработки аппаратной конфигурации САПР FastChip обеспечивает возможность интерактивного выбора составных фрагментов схемы и средства задания их межсоединений. Применение САПР предполагает эффективные, наглядные и контролируемые средства описания проекта, при этом наиболее распространенными являются графический и текстовый способы.

Собственные средства описания проекта FastChip основаны на графическом представлении с использованием допустимых для САПР библиотечных элементов и пользовательского базиса. Собственные средства ввода аппаратной конфигурации в массив встроенной конфигурируемой логики – библиотечные модули, на основе которых разработчик может определять схемотехнику проекта. Набор предлагаемых средой FastChip предопределённых элементов достаточно широк (около 150), постоянно расширяется и может быть разделен на следующие группы:

- типовые параметризируемые элементы вычислительной техники (логические элементы, мультиплексоры, счетчики и т. д.), конкретные параметры которых (разрядность, коэффициент пересчета, начальное состояние) могут назначаться разработчиком;
- специализированные элементы взаимодействия микропроцессорной части РСнК и матрицы программируемой логики (адресные селекторы, элементы системной шины, схемы синхронизации);
- специфичные для устройств РСнК составляющие (буферные элементы, порты ввода–вывода);
- типовые узлы вычислительных систем (контроллеры интерфейсов, графических индикаторов и т. д.), часть параметров которых варьируется проектировщиком.

Библиотечные модули от простых (логические вентили, сумматоры, умножители, компараторы, счетчики, элементы памяти) до сложных (блок шифрования стандарта DES) представляют собой готовые к использованию элементы принципиальной схемы (IP-блоки). Такие параметризируемые конструкции в терминологии САПР FastChip называются «Soft Modules», что отражает возможность их лёгкой модификации. Функции модуля выполняются аппаратно с помощью внутрикристальной программируемой логики и библиотечные модули не содержат никакой информации о своей компоновке внутри FPGA, следовательно, их быстродействие зависит от конкретного варианта размещения, трассировки и загруженности кристалла.

После определения необходимых библиотечных модулей разработчик размещает выбранные блоки в рабочей области (проектном окне) экрана САПР и при необходимости определяет их символические имена. Далее проектировщик настраивает изменяемые параметры выбранных модулей, например разрядность счётчика, начальное состояние регистров и т. д., установкой конкретных числовых значений в полях конфигурации модуля. Среда FastChip оперативно отражает оценку затрат на реализацию (число требуемых логических ячеек, портов ввода–вывода, адресных селекторов) в зависимости от сложности проекта.

Электрические связи между модулями задаются вводом существующего или созданием нового имени цепи, т. е. элементы между собой соединяются путём сопоставления имен. Такие имена устанавливаются для портов (внешних выводов) модуля, тем самым определяются требуемые электрические связи между компонентами схемы. Для разработчика доступен полный список цепей, который состоит из предопределенных, т. е. имеющихся на начальном этапе проектирования сигналов, и новых, установленных пользователем имён цепей. К предопределённым цепям относятся сигналы системной частоты, сброса, встроенных периферийных устройств, например запросы DMA, прерывания и т. д. Такие сигналы заводятся в логическую матрицу и могут участвовать в функционировании аппаратной части сигналов. Электрические связи между модулями могут быть отображены в окне проектирования с указанием имен элемента и цепи.

Ещё раз отметим, что встроенный способ ввода описания проекта в среде FastChip – графический, в базисе допустимых для САПР библиотечных модулей. Однако это не является существенным ограничением, поскольку открытая архитектура пакета FastChip легко интегрируется со средствами проектирования программируемой логики сторонних производителей. Достоинство графического способа описания – наглядность и хорошее восприятие структуры проекта, что связано с образным мышлением разработчика и привычным представлением электрических схем [156].

Информация о структуре логических элементов FPGA Triscend включена в базу данных ряда сторонних САПР синтеза программируемой логики. Это позволяет задать алгоритм функционирования и описать в текстовом или графическом виде целиком проект или его фрагмент во внешней САПР, синтезировать логическую структуру и далее путём переноса (импорта) создать собственный модуль в среде FastChip.

Созданный модуль проектировщика включается при этом в состав библиотеки пакета и становится доступным для использования.

Описание проекта в среде FastChip, содержащее Soft Modules, также может быть экспортировано во внешние среды синтеза, моделирования и схемотехнического ввода, что позволяет эффективно использовать технологические средства проектирования и отладки. Обмен информацией между САПР осуществляется через файлы стандарта EDIF 2.0.0. Таким образом, среда FastChip практически не ограничивает свободы разработчика при конфигурировании аппаратных ресурсов кристалла.

Аппаратная часть может быть реализована следующими способами:

- из библиотечных элементов, предварительно разработанных и заданных в среде FastChip;
- с помощью схемных или текстовых решений на основе инструментальных средств синтеза, представленных внешними системами проектирования;
- комбинацией этих двух способов.

Проектирование на основе СнК охватывает три типа ресурсов:

1. Встроенные аппаратные блоки, такие, как центральный процессор, системная память, контроллер прямого доступа к памяти и другие, недоступные для изменения разработчиком.

2. Реализованные полностью на FPGA аппаратные модули (счетчики, регистры, арифметические узлы, логические вентили), логика работы и электрические связи которых могут быть изменены проектировщиком для выполнения требуемых функций.

3. Интерфейсные модули связи между FPGA и внутренней системной шиной (адресные селекторы, программно-доступные ресурсы), предназначенные для интеграции перечисленных блоков.

Средства синтеза аппаратной части проекта предназначены для описания конфигурации последних двух типов ресурсов. Собственные средства описания проекта в среде FastChip представляют собой графический ввод в базисе предопределённых модулей. Такой способ задания функционирования в ряде случаев ограничивает класс решаемых задач или неудобен для разработчика, сковывает его образы заранее заданными конструкциями. В этом случае проектировщик для ввода проекта и задания алгоритма работы может использовать САПР, приведённые в табл. 4.4. Для текстового описания проекта допустимы синтезируемые конструкции языков описания аппаратуры, например VHDL. Используемый синтезатор должен поддерживать кристаллы Triscend, т. е. необходимо определение технологического базиса:

схемотехники макроячейки FPGA и портов ввода–вывода, низкоуровневых логических, арифметических, последовательных примитивов, реализуемых на таких структурах. Библиотека Triscend в этом случае предоставляет низкоуровневые примитивы, содержащиеся в подключаемых к внешней САПР файлах.

Табл. 4.4.

Тип САПР	Выполняемые функции	Ограничения
Synplify/ Synplify	Текстовый ввод описания проекта на языках описания аппаратуры; использование библиотек и примитивов; графическое представление проекта на уровне регистровых передач и на технологическом уровне; синтез проекта; экспорт синтезированного проекта в формате EDIF в САПР FastChip	Версия пакета Synplify более 6.1.3
Active-HDL	Ввод проекта на языке описания аппаратуры; описание функционирования при помощи редактора машины состояний и редактора блок-диаграмм; функциональное моделирование; подготовка описания проекта для синтеза	Версия САПР более 5.1. Требуется внешний САПР для синтеза проекта
OrCAD Express	Ввод описания проекта с помощью схемного редактора; использование графических примитивов и макроэлементов Triscend; синтез проекта; экспорт синтезированного проекта в формате EDIF в САПР FastChip	Версия САПР более 9.0. Отсутствует ввод текстового описания проекта
Protel	Графическое описание проекта с помощью схемного редактора; использование графических примитивов Triscend; экспорт синтезированного проекта в формате EDIF в САПР FastChip	Версия Protel 99 SE и выше. Отсутствует ввод текстового описания проекта
View Draw	Ввод описания проекта с помощью схемного редактора; использование графических примитивов и макроэлементов; синтез проекта; экспорт синтезированного проекта в формате EDIF в САПР FastChip	Отсутствует ввод текстового описания проекта
FPGA Express	Текстовый ввод описания проекта на языках описания аппаратуры; использование библиотек и примитивов Triscend; синтез проекта; экспорт синтезированного проекта в формате EDIF в САПР FastChip	Версия САПР 3.4 и выше

Разработчик на языке описания аппаратуры может использовать поведенческое или структурное задание алгоритма функцио-

нирования. Средства компиляции служат для преобразования ранее введённого разработчиком описания аппаратной части в выходной файл, содержащий информацию о настройке логических ячеек и их связей. Другими словами, целью компиляции является создание файла, отражающего размещение описанного ранее проекта в конкретном кристалле с трассировкой всех его электрических связей. Процесс компиляции (bind) состоит из нескольких этапов:

1. Проверка ресурсных ограничений (check constraints – IO, Device, Address) с целью нахождения грубых ошибок проектирования, например соединения выходных цепей, и определение критических системных факторов.

2. Распределение памяти программно-доступных элементов в FPGA (address allocation), т. е. назначение реальных физических адресов программно-доступным компонентам внутри массива программируемой логики.

3. Проекция логических модулей проекта на макроячейки PCHK (mapping), т. е. привязка абстрактных логических элементов к реальным физическим ресурсам.

4. Размещение выбранных макроячеек внутри микросхемы (placement), т. е. определение наилучшего места их размещения внутри кристалла в соответствии с данными о межсоединениях компонентов и требуемом быстродействии.

5. Разводка электрических связей внутри кристалла (routing) между размещёнными компонентами, т. е. выбор реальных физических путей прохождения сигналов, исходя из заданных временных требований и возможностей трассировки, определяемых доступными линиями связи и программируемыми матрицами соединений.

6. Временной анализ (timing analysis) полученной конфигурации FPGA и новая итерация (переход на п. 3) при невыполнении заданных пользователем временных ограничений (timing constraints) на задержки распространения сигнала в критичных цепях схемы.

7. Формирование выходного файла-образа (bit stream generation) с расширением csl, содержащего данные об используемом кристалле, алгоритме функционирования логических ячеек, портов ввода–вывода и структуре межсоединений.

Процесс компиляции является наиболее требовательным к аппаратным ресурсам инструментального компьютера, в частности, весьма желательно наличие не менее 512 Мбайт ОЗУ, иначе возрастает время компиляции и увеличивается нагрузка на жёсткий

диск. Большие временные затраты объясняются тем, что в случае компиляции САПР решает сложные многоитерационные задачи при наличии ограничений, имеющие много общего с процессами размещения элементов и разводки печатных плат. Применение технологии СнК, в какой-то мере уменьшает сложность разводки собственно печатной платы, но при этом возрастает машинная трудоемкость внутрикристалльной разводки. Архитектура трассировочных ресурсов и алгоритмы размещения и разводки нацелены на минимизацию задержки на наиболее длинных связях, и, тем самым, на создание конфигурации с наилучшей системной производительностью. Кроме загрузочного файла в процессе компиляции создается файл отчёта в формате html, содержащий, в частности, следующую информацию о проекте:

- примененные разработчиком библиотечные аппаратные модули и параметры их настройки;
- использованные ресурсы программируемой логики (CSL Resourse Utilization), в т. ч. со статистикой по отдельным показателям (таблицы перекодировки, триггеры, адресные селекторы и т. д.);
- использованные блоки ввода–вывода и их характеристики (номер вывода, направление передачи, и т. д.);
- межмодульные связи внутри проекта со статистикой соединений (цепи, имеющие несколько источников, ненагруженные цепи);
- распределение памяти для программно доступных элементов внутри программируемой логики, необходимое для связи аппаратной и программной составляющих проекта;
- расчётные данные (без реальной загрузки в FPGA) статического временного анализа электрической схемы, реализованной в результате компиляции.

Алгоритмы компиляции САПР FastChip оптимизированы для автоматического прохождения большинства проектов. При необходимости пользователь может управлять процессами размещения и разводки путём задания временных ограничений на быстродействие проекта, буферизацией цепей тактирования, заданием относительного размещения макроячеек внутри FPGA. Среда позволяет варьировать параметрами компиляции и выбрать наиболее подходящий ей режим прохождения, исходя из качества разводки и затрат времени.

Поскольку в микросхемы класса РСнК входит ядро на основе универсального микропроцессора, идеология разработки програм-

много обеспечения имеет много общего с традиционным процессом программирования микроконтроллеров. Отличительным моментом является процесс связывания программного и аппаратного обеспечения на этапе системной интеграции.

Программная часть разрабатывается с помощью внешних относительно среды FastChip компиляторов. Для 8-разрядного семейства РСнК могут быть применены средства проектирования для широко распространенного в промышленности процессорного ядра MSC-51, в частности Franklin Software, IAR, Tasking. Разработчик программного обеспечения может использовать как язык ассемблера, так и Си. Наиболее полную поддержку процесса проектирования и отладки обеспечивают инструментальные средства фирм Keil Software и Arhimedes Software.

Входящий в состав РСнК Triscend A7 процессор ARM7TDMI представляет собой один из вариантов широко распространенного ядра ARM и поддерживается, например, следующими компиляторами и системами разработки:

- ARM Development Suite (ADS) фирмы ARM;
- CodeWarrior фирмы Metrowerks;
- Diab Compiler Suite фирмы Wind River;
- компилятор и отладчик на основе лицензии GNU;
- MULTI 2000 фирмы Green Hills Software;
- CrossWork фирмы Rowley Associates;
- EWARM фирмы IAR Systems.

Для первых четырёх средств разработки возможна пошаговая отладка приложений в реальном времени с использованием специализированных отладочных возможностей интерфейса JTAG.

Для поддержки проектирования программной части производитель выпускает пакеты примеров и драйверов нижнего уровня для типовых применений – Triscend SDK. Такой набор позволяет освоить в сжатые сроки типовое аппаратно-программное обеспечение и перейти собственно на кодирование специфики требуемой задачи. Кроме того, для семейства TE5 на этапе системной интеграции могут создаваться инициализирующие процедуры периферийных узлов.

В результате процесса разработки программного обеспечения создаётся объектный код для последующего объединения его с аппаратной конфигурацией и загрузки в реальный кристалл для выполнения [101]. Идеология построения пакета FastChip наглядно демонстрирует тенденцию совместных разработок САПР, которая

заключается в кооперации усилий различных производителей. Так как информационный поток проектирования устройств на основе РСнК многоплановый, то пакеты САПР должны занимать в нём соответствующее место исходя из своей функциональности. Среда FastChip в процессе маршрута проектирования обеспечивает, прежде всего, специфичные для РСнК этапы (компиляция, загрузка и внутрисхемная отладка) и системную интеграцию, остальные же функции выполняются сторонними САПР.

К этапу системной интеграции для САПР FastChip относится процесс генерации (Generate) заголовочного файла (обычно с расширением .h) для программной части, содержащего символьные и двоичные данные о распределении памяти программно-доступных элементов проекта. Данный файл содержит назначение логических адресов программно доступных областей кристалла соответствующим символьным константам (именам). В дальнейшем этот файл участвует в кодировании программной части встроенного процессора (архитектуры MSC-51 или ARM в зависимости от семейства), при этом обеспечивается доступ из программы к аппаратным ресурсам, размещённым в программируемой логике. Процесс генерации является важнейшим для проектирования связи аппаратного и программного обеспечения конфигурируемой системы на кристалле. Сформированный заголовочный файл используется вне среды САПР FastChip компиляторами с языка Си или ассемблера сторонних производителей.

Процесс настройки встроенных аппаратных ресурсов также относится к системной интеграции. Такая настройка зависит от типа кристалла и может включать в себя конфигурацию интерфейса с внешней памятью, способов формирования системной частоты, режимов работы встроенных периферийных узлов. Результаты такой настройки отражаются в содержимом конфигурационного файла, заголовочного файла и в возможном создании инициализирующих подпрограмм для периферийных узлов, что характерно для проектировании устройств семейства РСнК TE5.

Далее при помощи редактора доступные элементы ввода–вывода размещаются на конкретных контактах кристалла, т. е. организуется связь ранее спроектированной аппаратной части и внешнего окружения. Также в редакторе определяются режимы работы портов ввода-вывода, которые имеют настраиваемую схемотехнику для ряда конкретных условий функционирования (управление временем нарастания импульса, выходным током

вывода, входным гистерезисом, совместимостью с шиной PCI 3.3 В и т. д.). Таким этапом системной интеграции организуется внешний электрический интерфейс разрабатываемой схемы.

Возможность интеграции с САПР сторонних производителей характеризует степень «открытости» среды разработки. Проектирование устройств на основе технологии РСнК предполагает несколько этапов: концептуальный синтез алгоритмов, моделирование, разработку аппаратных средств, написание программного обеспечения, системную интеграцию, совместную отладку, тестирование. При реализации указанных функций в рамках одной САПР потребуется большое количество разнородных методов проектирования, что существенно увеличит размеры, сложность освоения и время выхода на потребителя такой универсальной среды проектирования.

Другой подход, который успешно демонстрирует среда FastChip – качественное обеспечение собственных, внутренних функций системы, описанных ранее, и целенаправленная поддержка потока проектирования как программного, так и аппаратного обеспечения с помощью внешних САПР. Такой подход оправдан как с позиций специализации (качества легче достигнуть при отказе от универсальности), так и с позиций выбора и состязательности – пользователь не замыкается на закрытом продукте, а может выбирать требуемую САПР из ряда предлагаемых. Возможности и показатели качества САПР синтеза и моделирования, предоставляемые различными компаниями, как правило, существенно отличаются, однако в настоящее время существует тенденция возрастания конкурентных преимуществ средств проектирования компаний, специализирующихся на выпуске именно САПР, а не компаний, производящих элементную базу.

Средства проектирования сторонних производителей при разработке устройств на основе технологии РСнК работают в дружественном окружении среды FastChip на следующих этапах:

- синтеза аппаратной части проекта для выполнения алгоритмов обработки сигналов на встроенном FPGA;
- разработки программного кода для внутреннего процессора и его автономная отладка при помощи эмулятора;
- моделирования аппаратной части проекта на всех фазах его прохождения;
- совместной верификации аппаратного и программного обеспечения.

Принципиальная возможность интеграции САПР различных производителей обеспечивается единым форматом представления данных описания электронных устройств (Electronic Design Interchange Format – EDIF). Этот формат поддерживается всеми производителями САПР в рамках требований к программам автоматизации проектирования электронных устройств (Electronic Design Automation – EDA). Важнейшую роль в интеграции играет техническая политика фирмы Triscend, в том числе и поддержка совместного проектирования и отладки аппаратно-программного обеспечения. Так, для осуществления синтеза аппаратной части проекта в САПР сторонних производителей специалисты фирмы Triscend подключают и тестируют в подобных средах синтеза собственные библиотеки, структуры и описания низкоуровневых технологических примитивов, а также организуют сквозной поток проектирования и его сопровождение. Поддержка разработки и отладки программной части требует создания специфичного для различных компиляторов программного окружения, а также специализированных модулей для пошаговой отладки с использованием JTAG-интерфейса.

Подобный цикл работ по интеграции со сторонними САПР проводится и для процесса моделирования аппаратной части проекта. Для выполнения временного моделирования с учетом расчетных данных о быстродействии среда FastChip имеет возможность экспорта списка соединений, конфигурации и информации о задержках во внешнюю САПР для тестирования.

Верификация (тестирование) разрабатываемого устройства является важнейшей частью процесса проектирования. Как правило, эта процедура проводится несколькими итерациями, причём каждая фаза процесса проектирования предполагает целесообразные методы и средства верификации [102]. На рис. 4.20 отражены основные этапы описанного процесса проектирования для устройств Triscend A7.

Разноплановость целей, которые ставятся при тестировании устройств обработки сигналов, затрудняет создание универсальных отладочных средств, поэтому на каждом этапе целесообразно применять специализированные приёмы. Собственные средства верификации проекта в САПР FastChip выполняют следующие функции:

- проверку на этапе компиляции на ресурсные, адресные ограничения и грубые ошибки проектирования;

- статический временной анализ задержек распространения сигналов внутри синтезированной на FPGA электрической схемы;
- аппаратную отладку программного обеспечения и анализ конфигурируемой логики в реальном времени с помощью механизма JTAG-сканирования.

Наиболее важным для практики является поддержка средой проектирования физического контроля, когда разработчику предоставляется возможность в интерактивном режиме или в командной строке считывать реальные состояния электрических цепей, анализировать содержимое ОЗУ и программно-доступных ресурсов, реализованных в конфигурируемой логике, управлять состоянием процессора и задавать точки останова со сложными условиями вызова. Таким образом, САПР FastChip эффективно решает вопросы отладки, причем наибольшее внимание уделено физической верификации. Средства проектирования компании Cadencex [158] охватывают почти все уровни проектирования сложных систем на кристалле, приведенные в табл. 4.3.

Средства системного проектирования, включая Coware Signal Processing Worksystem (SPW2000), Coware Hardware Design System (HDS2000) и соответствующие библиотеки позволяют определить функциональные параметры СнК, провести декомпозицию на программные и аппаратные модули, обеспечивают моделирование на высоком уровне абстракции. Программный модуль i-Architect предназначен для поддержки работы с библиотеками IP-блоков и поиска IP-блоков по заданным характеристикам в сети Internet.

Моделирование в среде SPW2000 может осуществляться разнообразными способами, в частности поддерживаются как традиционные языки программирования и моделирования (C/C++, MatLab), так и языки описания аппаратуры (SystemC, Verilog, VHDL). При моделировании можно использовать источники реальных сигналов, например при проверке IP-блока MPEG2 кодера подавать на вход системы моделирования файл цифровых данных канала связи. Удобный графический редактор BDE (Block Diagram Editor) представляет проектируемое устройство подлежащее анализу как иерархическую блок-диаграмму, которая содержит как общее описание системы, так и отдельные блоки. Отдельная поведенческая модель и спецификация блока доступна разработчику в случае, если IP-компонент присутствует в библиотеке модулей. В состав библиотек (Communication Library for SPW, Multimedia Design Kit for SPW) IP-блоков содержится более 3000 разных модулей: коммуникационные системы сотовой связи WCDMA, cdma2000, GSM/GPRS/HSCSD/EDGE, беспроводные

сети стандартов IEEE802.11, Bluetooth, многообразные кодеры–декодеры цифрового телевидения и т. д. [156, 163].

При моделировании на платформе SPD2000 используется представление чисел с плавающей запятой. Для реализации в составе ИМС необходимо преобразование в алгоритмы с фиксированной запятой, которое выполняет пакет HDS2000. С его помощью из структурного описания системной модели формируется её представление на языках описания аппаратуры Verilog/VHDL. Большим подспорьем для разработчика, применяющего САПР системного уровня компании Cadence, является возможность описания проекта на смешанных уровнях: системном и уровне регистровых передач. Это позволяет интегрировать сложные изделия путём повторного использования ранее разработанных блоков. Узлы смешанной аналого-цифровой обработки, радиочастотные модули и печатные платы изделий электронной техники разрабатываются с использование пакетов программ Cadence: Shematic Editor, Spectre, SpectreRF.

Программная среда Cadence Incisive Unified Simulator предназначена для программно-аппаратного проектирования, отладки, верификации и генерации тестов на этапе от системного уровня до функционального. С помощью программной среды Cadence RTL Compiler производится переход с функционального уровня проектирования на логический. С применением описанных САПР можно быстро получить ответ на вопрос о реализуемости в системе на кристалле алгоритма обработки, оценки площади кристалла, энергопотребления, количества вентилей и т. д.

Планирование кристалла, размещение IP-блоков, интерактивную и автоматическую трассировку всего кристалла в целом выполняют средства САПР кристального уровня: Layout-XL, Assura, Chip Assembly кристалла. Все продукты компании Cadence интегрированы в сквозной маршрут, верифицированный и поддерживаемый технологическими библиотеками IP-блоков ведущих производителей. Применение средств САПР одинаково важно как на системном уровне проектирования, так и на последующих уровнях: функциональном, логическом, топологическом. При этом средства САПР должны обеспечивать совместимые входные и выходные форматы данных для сквозного сопряженного проектирования. Соответствующие дизайн-центры должны быть оснащены средствами САПР для прохождения маршрута проектирования. Обобщенный маршрут проектирования СБИС система на кристалле (рис. 4.21), применяемый в ФГУП «НИИМА

«Прогресс», использует средства САПР Cadence и Synopsys на всех этапах проектирования. Высокие единовременные затраты, в том числе на средства САПР и сопутствующие риски означают, что проекты систем на кристалле должны планироваться и создаваться как платформы, рассчитанные на различные приложения и дочерние продукты [12]. В целом задача современного проектирования устройств обработки сигналов на основе СБИС СнК – выполнять каждый этап работы быстрее, с более надежными и достоверными результатами.

Технология СнК является прекрасной технической основой сопряженного проектирования, поскольку такие системы в значительных пределах конструктивно-инвариантны к аппаратным и программным решениям. Технологической основой сопряженного проектирования являются САПР, интегрирующие в себе средства проектирования и отладки аппаратно-программных систем.

Хотя для массовой обработки малоподвижных изображений (например, спутниковых) еще производится крупные системы цифровой обработки, но с появлением технологии СнК перспективными является создание специализированных систем цифровой обработки изображений (рис. 4.22) обработки изображений.

Использование предварительно разработанных IP-блоков для проектирования СнК стало необходимым для быстрого выхода на рынок. В связи с этим и решаются следующие основные задачи автоматизации проектирования систем обработки изображений на основе этих блоков:

- разработка архитектуры системы;
- интеграция IP-блоков и новых блоков в единую СнК;
- верификация на всех этапах проектирования.

Под СнК обработки изображений подразумевается универсальная САПР СнК, которая включает набор подсистем кодирования и декодирования изображений, где каждая предназначена для обработки определенного класса изображений. Методы кодирования и декодирования изображений часто базируются на спектральном преобразовании видеосигнала (см. раздел 3.1). С другой стороны, появление технологии СнК с большими возможностями на одном кристалле как программной, так и аппаратной реализации привело к развитию прямых методов представления и обработки изображений, использованию синтаксических описаний, где основная проблема заключается в преобразовании структуры потока видеоданных либо для хранения (сжатия, кодирования), либо для быстрых преобразований и обработки.

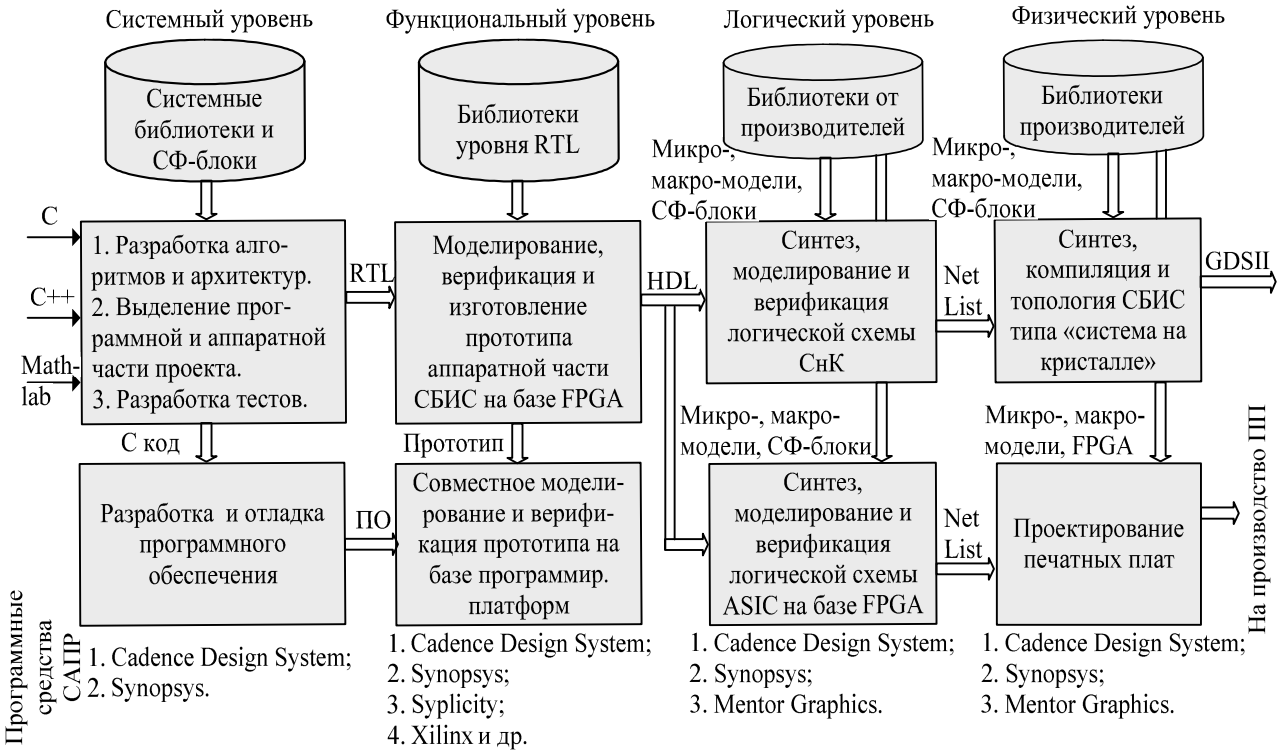


Рис. 4.21. Обобщенный маршрут автоматизированного проектирования СБИС типа «системам на кристалле»

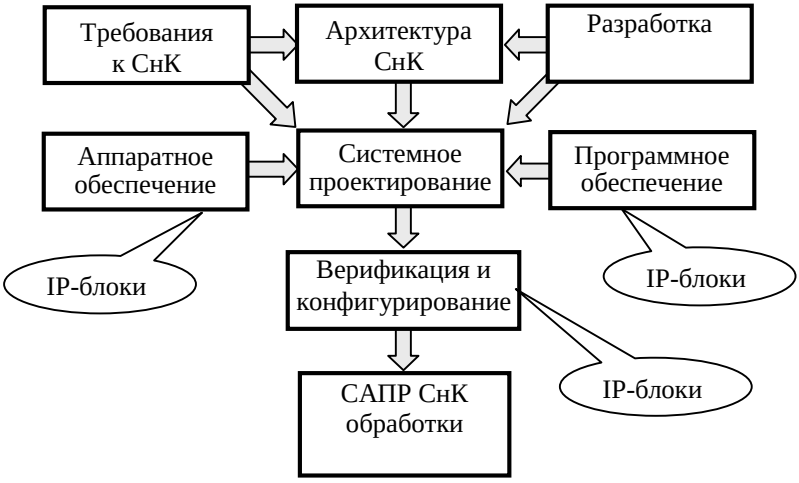


Рис. 4.22 Этапы проектирования СнК обработки изображений

Такая двойственность привела к появлению систем, распараллеливающих либо вычислительные операции, либо доступ к памяти. Естественным при решении задач сжатия изображений, рассмотренных в разделе 4.4, является применение методов, либо максимально адаптированных к исходному кадру, либо в максимальной степени независимых от класса изображений.

4.4. Сжатие сигналов изображений в системах на кристалле

Теоретические аспекты проблемы сжатия (кодирования источника) телевизионных изображений освещены в гл. 3, в настоящем разделе рассматриваются практические вопросы кодирования изображений. Как справедливо считается [171], нет смысла говорить о том, какой алгоритм сжатия изображения лучше, если не обозначены класс изображения, класс приложений и критерии сравнения. Это означает, что разработчик телевизионной аппаратуры в своей практике явно или неявно использует критерий информационного риска (см. раздел 3.1.) Поиск компромисса между коэффициентом сжатия и сложностью аппаратно-программного обеспечения характерен для всех практических алгоритмов сжатия. Как правило, с увеличением сложности аппаратно-программных ресурсов и времени кодирования увеличивается коэффициент сжатия.

Сжатие изображений в системах мобильной телевизионной связи. Системы наблюдения могут использовать два варианта кодирования: с потерями и без потерь. Алгоритмы сжатия без потерь используются в системах, ориентированных на формирование базы данных в форме видеоматериалов, которая может быть представлена в качестве доказательства в судебных инстанциях, или предназначенной для последующей обработки и анализа с целью выявления закономерностей (научные цели, космические исследования, мониторинг окружающей среды, медицинская диагностика, промышленная дефектоскопия и т. п.). Все остальные системы используют алгоритмы сжатия с потерями, которые существенно более эффективны с точки зрения компактности формируемого кода, чем алгоритмы без потерь.

Всё многообразие методов сжатия изображений можно классифицировать на две большие группы: без использования и с использованием межкадровой корреляции, то есть межкадрового предсказания. В первом случае каждый кадр изображения кодируется сам по себе, независимо от других кадров. К таким методам относятся стандартизованные форматы JPEG, JPEG2000, другие алгоритмы вейвлетного и фрактального кодирования.

Другая, значительно более многообразная, группа методов кодирования изображений предусматривает использование временного предсказания. В простейшем случае, кодированию подвергается межкадровая разность (например, в алгоритме Дельта-Вейвлет). В более сложных алгоритмах (стандарты MPEG-1, MPEG-2, MPEG-4-video (одна из реализаций которого широко известна как DivX), MPEG-4-AVC (международный стандарт ISO/IEC 14496-10), Motion-Vawelett, Windows Media (Microsoft Corp.), One2 и др. используется компенсация движения.

Преимущества методов внутрикадрового предсказания :

1. Относительно низкие затраты памяти и вычислительной мощности на техническую реализацию.
2. Малая зависимость степени сжатия (при сохранении качества изображения) от наличия движущихся объектов и частоты смены кадров.
3. Возможность начала декодирования с произвольного кадра, что важно при возможности «обрыва» канала связи или разрушения цифрового потока по иным причинам.

Основным недостатком этих методов является сравнительно невысокий коэффициент сжатия, который для большинства реальных изображений и методов составляет 10…30 раз при визуальном сохранении хорошего качества изображения.

В отличие от этого, методы с межкадровым предсказанием, и, в частности, с компенсацией движения, позволяют достичь значительно более высоких степеней сжатия, достигающих 30…300 и выше при сохранении хорошего качества изображения.

Методам с временным предсказанием присущи недостатки:

1. Значительные затраты памяти для хранения предсказателей и вычислительной мощности для компенсации движения. Требуемая для кодирования вычислительная мощность может на порядки превышать эту величину для методов без предсказания. Многие методы пока не имеют микроэлектронной реализации и могут работать в реальном времени только на мощных вычислительных платформах, например, многопроцессорных ПК.

2. Зависимость степени сжатия (при сохранении качества изображения) от наличия движущихся объектов и вида движения, а также частоты смены кадров. При уменьшении частоты смены кадров временное предсказание становится малоэффективным и степень сжатия становится соизмеримой с методами без предсказания.

3. Возможность начала декодирования имеется только с т. н. опорных кадров (не использующих межкадрового предсказания), которые периодически, как правило, один раз в 0.5…10 с, вводятся в состав кодированной последовательности. Это увеличивает время возобновления декодирования при разрушении цифрового потока.

4. Высокая эффективность методов достигается лишь при хорошем отношении сигнал/шум исходного видеоматериала. При зашумлении входного изображения (например, вследствие недостаточной освещенности объекта) их эффективность существенно снижается, поскольку некоррелированный шум не поддается предсказанию.

При построении систем наблюдения, использующих сжатие видеоинформации с потерями, предъявляется ряд специфических требований к алгоритму кодирования:

1. Алгоритм должен быть максимально эффективен, т. е. обеспечивать максимальную степень сжатия при визуально приемлемом качестве изображения или при формировании декодированного изображения, укладывающегося в заданный допуск объективно измеряемых параметров (разрешающей способности, уровня шумов, искажений воспроизведения полутонов, крутизны переходной характеристики и т. п.).

2. Использование межкадрового предсказания существенно увеличивает эффективность алгоритма, однако в ряде случаев приходится отказываться от использования межкадровой декоре-

ляции и применять алгоритмы, основанные на внутрикадровом кодировании. Наиболее часто межкадровые декоррелирующие алгоритмы неприемлемы в системах, в которых передача сжатого цифрового потока сопряжена с возможным отбрасыванием отдельных кадров из-за высокого уровня помех. Поскольку отбрасываемые кадры могут быть использованы (а заранее это неизвестно) в качестве опорных, то в результате удаления всего лишь одного кадра будет разрушаться или искажаться целая группа кадров, предсказанных из удалённого. Аналогично, принятые с искажениями опорные кадры распространят эти искажения на всю группу (т. н. трек ошибки предсказания). Алгоритмы с межкадровым предсказанием становятся малоэффективными при низкой частоте обновления кадров – когда корреляция между отдельными изображениями становится пренебрежимо малой. При воспроизведении архивированных видеоматериалов часто должен быть обеспечен быстрый доступ к любому из кадров. Такое требование обычно присутствует для систем автоматической архивации с последующим просмотром записанной информации при возникновении чрезвычайных событий с целью оценки предыстории. В любом случае, если присутствует требование исключения межкадрового предсказания, то оно автоматически исключает применение алгоритмов типа MJPEG, MPEG, Motion Wavelet, Delta-Wavelet.

3. Алгоритм не должен быть излишне сложен, поскольку при аппаратном исполнении кодека увеличит его стоимость, габаритно-массовые характеристики, потребляемую мощность, надежность. При программном исполнении для кодирования–декодирования в реальном времени может потребовать для его реализации многопроцессорной системы.

4. С целью оперативности получения информации задержка кодирования–декодирования не должна превышать заданного значения. Это требование исключает применение в системе межкадрового предсказания, поскольку при таком кодировании имеется сильная неравномерность объема сжатых кадров – разностные изображения, предсказанные из одного кадра (P-frame), сжимаются в 2…3 раза более эффективно, чем не предсказанные (I-frame), а кадры основанные на двух ранее переданных изображениях (B-frame) – 4…9 раз. Из-за неравномерности цифрового потока приходится его пропускать через накопительный буфер, размер которого в среднем складывается из размера I-кадра и суммарного размера B-кадров, передаваемых между двумя P-кадрами. В системах, использующих стандарт MPEG-2 с

постоянным размером буфера, задержка от момента кодирования до воспроизведения может составлять 0,5…3 *с* и зависит от скорости сжатого цифрового потока (для больших скоростей задержка меньше).

5. Алгоритм должен быть адаптивен к классу наблюдаемых изображений. Формально это означает, что для конкретного класса изображений может быть разработан максимально эффективный кодер. Однако при предъявлении такому кодеру иного (выпадающего из указанного класса) изображения эффективность может оказаться хуже, чем у универсального алгоритма (т. е. рассчитанного на «почти любые» реальные сцены). Очевидно, что разработанное в рамках подобного требования устройство будет иметь узкий рынок своего применения. Поэтому в настоящее время в системах наблюдения используют универсальные алгоритмы.

6. Желательно использовать стандартный или широко распространенный алгоритм сжатия, поскольку это позволит использовать дешёвые аппаратные (готовые ИМС) или программные средства, не требующие дополнительных затрат на разработку, ремонт (переустановку системы), перепроектирование с целью введения новых функциональных возможностей. Имеется один недостаток в этом требовании – в силу постоянного совершенствования алгоритмов некоторые из них устаревают, а поддерживающие их аппаратные средства снимаются с производства.

7. Допустимость некоторой потери информации без существенного ухудшения качества восстановленного изображения (высокая помехоустойчивость алгоритма).

В настоящее время в системах наблюдения используются два вида кодирования сигналов изображения с целью сжатия: кодирование на основе формата JPEG и кодирование на основе вейвлет-преобразования.

Формат JPEG стандартизован (ISO/IEC 10918) на основе рекомендации ITU-T T.81 в 1992 г. В настоящее время существует множество систем наблюдения, основанных на этом формате. Однако, начиная с 1999 г. новых микросхем, поддерживающих этот стандарт не производится. Это вызвано несколькими причинами:

- в сравнении с новым стандартом JPEG2000, формат JPEG примерно в 2…3 раза менее эффективен;
- микросхемы имеют относительно узкий рынок сбыта (в сравнении с вещательным телевидением), поэтому контингент фирм, выпускающих эту продукцию, ограничен;

- старые ИМС в основном удовлетворяют производителей ранее разработанных и поставляемых систем наблюдения;
- с развитием компьютерных технологий алгоритм может быть реализован другими способами. Кодирующая часть, имеющая жесткие требования по размерам – на цифровых сигнальных процессорах (DSP), декодирующая (в силу выполнения многих функций – отображения, архивирования) – на обычных ЭВМ.

Поэтому, разработанные в прошлом тысячелетии микросхемы, как правило, выполненные по технологии 0,35 мкм:

- имеют высокое энергопотребление;
- требуют внешней обвязки (памяти, интерфейсов преобразования сигналов) и рассчитаны на определенные компоненты (не поддерживают SDRAM, DDR-RAM и т. п.);
- выполнены в больших по размерам корпусах;
- могут быть в любой момент сняты с производства.

Достоинствами данного алгоритма является относительная простота реализации и, в силу широкой распространенности, возможность замены используемого оборудования на близкое по своим характеристикам.

Стандартизованный в 2000 г. (ISO/IEC 15444) стандарт JPEG2000, основан на вейвлет-преобразовании, позволяющем сжать изображение до 40…60 раз. Множество программ, «поддерживающих» этот формат кодирования зачастую оказываются несовместимыми между собой (файл, закодированный одной программой, не отображается, или отображается некорректно другой программой). Многие зарубежные фирмы по разработке программ (например, LuraWave) до сих пор продолжают настаивать на том, что «их вейлет-преобразование» лучше (файлы с расширением LWF), чем JPEG2000, но для обеспечения совместимости вынуждены вводить поддержку JPEG2000.

Существует множество других алгоритмов сжатия, уже стандартизованных или находящихся на этапе стандартизации, которые не могут быть использованы из-за несоответствия специфике систем наблюдения. В частности, они эффективны, если используют межкадровое предсказание, но ни один из них (кроме MPEG-2) не реализован в однокристальном кодере, и все не могут кодировать чёрно-белое изображение. Некоторые из них представлены в таблице 4.5.

Таблица 4.5

Стандарт, алгоритм, кодек	Достоинства	Недостатки
SMPTE 421M (VC-1, WindowsMedia 9)	Сжатие – до 90 раз, декодер реального времени на IBM PC, есть однокристальный декодер, есть SDK	Сложность собственной реализации алгоритма
TrueMotion VP6	Сжатие – до 90 раз, декодер реального времени на IBM PC,	Стандартизация не предусматривается, нет открытых кодов
RealVideo 10	Сжатие – до 90 раз, декодер реального времени на IBM PC,	Очень неустойчивый декодер, стандартизация не предусматривается, нет открытых кодов
ISO/IEC 14496-10 (H.264/AVC)	Сжатие – до 80 раз, , есть однокристальный декодер	Встроенная помехозащита
ISO/IEC 14496-2 (MPEG-4 в т.ч. DivX)	Сжатие – до 60 раз, декодер реального времени на IBM PC	Устаревший и распространенный только в мультимедиа формат с не очень хорошим качеством
ISO/IEC 13818-2 (MPEG-2)	Сжатие – до 40 раз, декодер реального времени на IBM PC, есть однокристальный кодер и декодер	Устаревший формат, предназначенный для DVD и вещания.

Вейвлет-технология кодирования представляет собой математическое преобразование изображений, поддерживающее обработку полного изображения, а не только его отдельных областей, это ядро, на котором базируется стандарт JPEG 2000. Еще в 1994 г. эксперты по обработке видеосигналов из Analog Devices обратили внимание на огромный потенциал вейвлет-технологии в области видеокомпрессии и видеозаписи. Специалистами фирмы Analog Devices разработаны алгоритмы обработки видеосигнала, которые сделали возможной высокоэффективную вейвлет-обработку с низкими требованиями к памяти и малым энергопотреблением. Вейвлет-технология SURFTM (Spatial Ultraefficient Recursive Filtering) компании Analog Devices включает эти запатентованные методы, теперь реализованные в микросхеме ADV202. Это первая система на кристалле в электронной промышленности, поддерживающая новый стандарт сжатия изображений JPEG 2000. Нацеливаясь на один из самых

быстрорастущих сегментов рынка бытовой электроники – сегмент цифровых фотокамер – СБИС обеспечивает аппаратное ускорение, требуемое для стандарта JPEG 2000, малое энергопотребление и низкую стоимость. Базируясь на вейвлет-технологии, которая известна под названием SURFTM, СБИС позволяет производителям цифровых фотокамер донести все преимущества стандарта JPEG 2000 до конечного пользователя:

- превосходное качество изображения;
- масштабируемость при сжатии, что позволяет пользователю извлекать изображения с различным качеством и разрешением;
- контроль над качеством изображения и использованием памяти;
- устойчивость к ошибкам при передаче изображений в зашумленной среде.

JPEG 2000 коренным образом меняет способы управления изображением, получаемым с цифровых фотокамер и вебкамер систем безопасности, а также способы воспроизведения изображений, передаваемых по каналам Интернет или сотовому телефону. С выпуском нового чипа JPEG 2000 для цифровых камер фирма Analog Devices оказалась на переднем крае поддержки стандарта JPEG 2000, в конечном счете, отвечая запросам потребителей. Такое решение является примером и для отечественных приборостроительных компаний, так как микросхемы кодеров изображения имеют широкий спрос в промышленности. Первый стандарт сжатия изображений JPEG был создан в 1987 г., в попытке уместить изображение на дискете. В новом «цифровом мире» требования потребителя к записи, передаче и получению сложных изображений посредством цифровых фотокамер, Интернет-каналов, сотовых телефонов и принтеров/копировальных устройств намного превышают возможности исходного стандарта JPEG. Новый стандарт JPEG 2000, созданный лидерами индустрии электроники, может использоваться для записи и передачи сложных фотоизображений по проводным и беспроводным сетям [170].

Мировой рынок (в 2000 г. – $4,3 млрд., в 2005 г. – $8 млрд.) заполнен нетребовательными к энергопотреблению и недорогими цифровыми фотокамерами. Стандарт сжатия изображений JPEG 2000 идеально подходит для цифровых фотокамер и открывает для пользователя массу возможностей.

Так, режим постоянного качества в цифровой фотокамере с ADV-JP2000 (Constant Quality Mode) позволяет создавать очень маленькие файлы. Например, в случае простого изображения файл

будет просто крошечным. А для сложного изображения алгоритм выберет наилучшее возможное качество, не превышая запрограммированный максимальный размер файла. Фотокамера на основе JPEG 2000 сможет сохранить, в среднем, на 50% больше данных, чем фотокамера на основе JPEG.

Режим точного размера файла (Exact File Size Mode) – размер создаваемых файлов в точности равен запрограммированному значению, что позволяет пользователю получить конкретное число кадров наилучшего возможного качества.

Режим выбора пользователем размера файла хранимого изображения (User-Selectable File Size of Stored Images) – позволяет пользователю понизить качество некоторых изображений и, таким образом сэкономив память, записать большее число кадров. В JPEG 2000 это преобразование происходит без перекодировки изображения, т. е. в реальном времени и при минимальном потреблении энергии. Это главное преимущество перед JPEG, которому для выполнения такой же задачи потребуется несколько секунд и гораздо больше энергии.

Предварительный просмотр изображений в различном масштабе (Preview of Images in Multiple Sizes) позволяет пользователям просмотреть файл, настроив размер изображения, и это не требует дополнительных вычислений при масштабировании или отдельного хранения каждого изображения. А вот цифровым фотокамерам, которые используют JPEG, потребуется сохранить каждый просмотр в отдельном файле, что занимает много флэш-памяти и замедляет процесс предварительного просмотра.

Расширенная поддержка файловых форматов (Extensive File Format Support) – JPEG 2000 обеспечивает усовершенствованный формат файлов, который поддерживает все аудио- и видео-стандарты, используемые сегодня в цифровых фотокамерах. Таким образом, это обеспечивает общий стандарт использования мульти-медийных файлов с цифровыми фотокамерами.

Микросхема выполнена по КМОП технологии с проектными нормами 0,18 мкм, потребляет 100 мВт в активном режиме и менее 100 мкА в режиме ожидания, имеет номинальное напряжение от 1,5 В до 1,8 В и 3,3-вольтовый I/O и представляет собой миниатюрную систему на кристалле: корпус LFBGA с 48 выводами. ADV-JP2000 способна сжать пять 3-мегапиксельных изображений за 1 секунду. Кроме того, ADV-JP2000 обеспечивает сжатие без потерь до 10 бит на компоненту и сжатие с потерями до 14 бит на компоненту. Это даёт существенные преимущества в

эффективности в сравнении с JPEG-стандартом, который давал только 8 бит на компоненту и не мог обеспечить одновременное (в одном формате) сжатие без потерь и с потерями.

«На базе нашей вейвлет-технологии SURFTM ADV-JPG 2000 реализует всю мощь видеокомпрессии JPEG 2000 для цифровых изображений», – сказал Роджер К. Смит, менеджер Analog Devices, Inc. «Пользователи не только получат более полный контроль над качеством и количеством изображений, сохраняемых в цифровых фотокамерах, но и смогут передавать сложные изображения по Интернет-каналам. Пользователь на приёмном конце, исходя из доступной полосы пропускания, сможет сам выбрать подходящий размер и качество изображения».

Триангуляционные методы. Одним из перспективных направлений в области кодирования и декодирования статических изображений умеренной детальности являются триангуляционные методы, применение которых известно из задач распознавания образов и из задач восстановления двумерных изображений по нерегулярно расположенным опорным точкам. Классификация триангуляционных алгоритмов кодирования и декодирования изображений (рис. 4.23) тесно связана с блочным методом проектирования СнК, когда проектируется комбинация новых IP-блоков и IP-блоков многократного использования из библиотек (от разных поставщиков и от собственных проектных групп). В основу систематизации (рис. 4.23) алгоритмов сжатия и восстановления изображения положены следующие понятия:

- число полигонов
- форма полигонов;
- взаимосвязанность сжатия и восстановления;
- расположения опорных точек (ОТ) в пределах полигона;
- критерии поиска.

Многочисленные стандарты и отсутствие универсального инструмента применения этих блоков делает эту задачу нетривиальной. При проектировании алгоритмов, основанных на ПРМ, сложность аппаратной реализации возрастает по мере повышения степени сжатия, однако быстрое развитие технологии СнК и повышение плотности размещения вентилей на кристалле делает некоторое возрастание сложности преодолимой проблемой.

Основные преимущества триангуляции следующие:

- адаптивность, кодер подстраивается под данные – там, где опорные точки (ОТ) разрежены, треугольники – крупнее, а там, где есть сгущение – мельче (рис. 4.24);

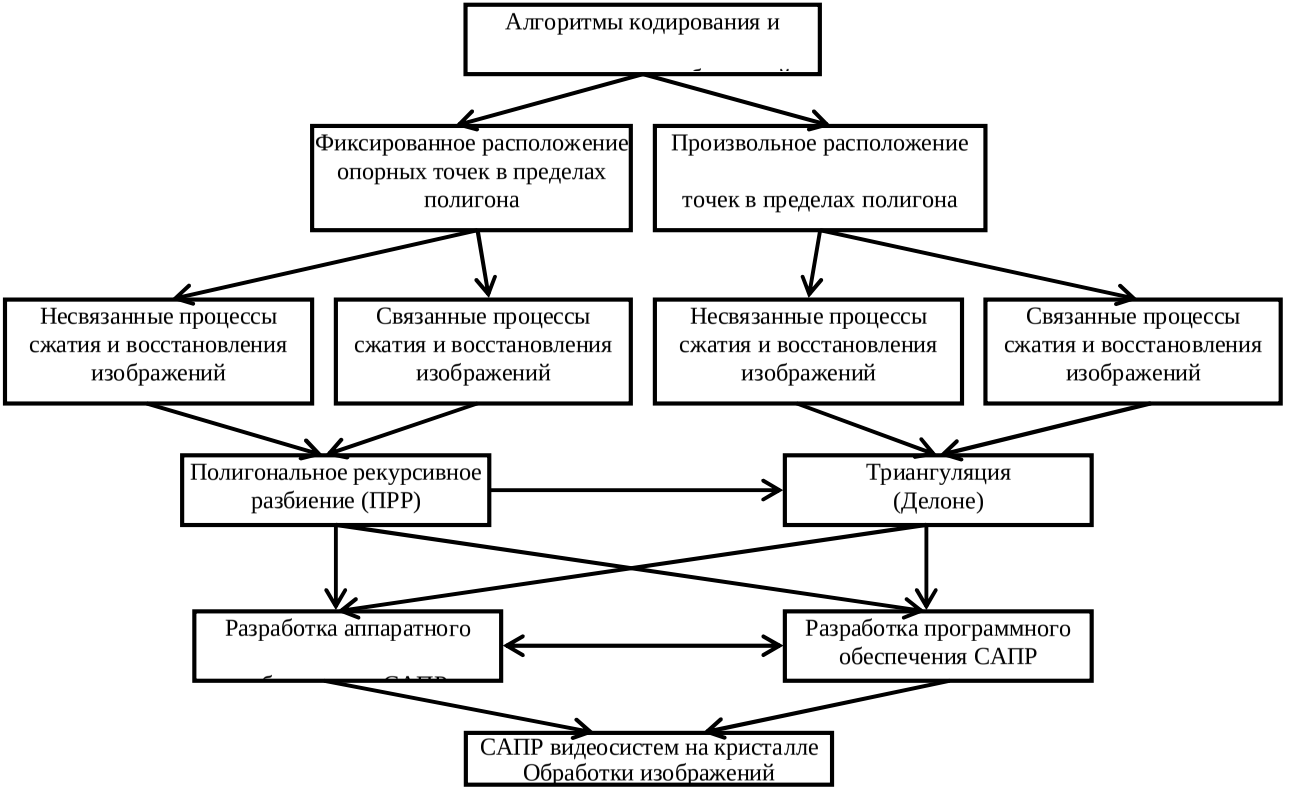


Рис. 4.23. Классификация триангуляционных алгоритмов кодирования и декодирования

• число треугольников не превышает удвоенного числа опорных точек;

• треугольная сетка требует существенно меньших вычислительных затрат, чем прямоугольная при той же точности восстановления.

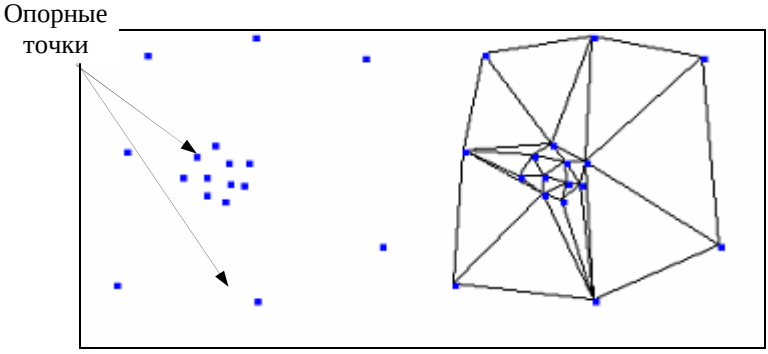


Рис. 4.24. Триангуляция опорных точек изображения

Известен (но мало изучен) метод триангуляции Делоне [172, 173], эффективный по достижению высокого коэффициента сжатия, т. е. приближения к энтропии изображений умеренной детальности.

Высокая точность аппроксимации обеспечивается высокой сложностью реализации кодера. На прямоугольной сетке можно за несколько шагов локализовать область с перепадом яркости, но при триангуляции поиск треугольника и соединение опорных точек непересекающимися отрезками затруднены. Для сокращения вычислительных затрат (при статических изображениях умеренной детальности) целесообразен вариант триангуляционного метода с регулярной сеткой и рекурсивной реализацией метода полихотомии, получивший название **пирамидально-рекурсивного метода** (**ПРМ**). Его суть заключается в том, что исходное изображение описывается упорядоченной последовательностью фрагментов различных уровней (L), расположенных обычно друг над другом (отсюда и название пирамида изображения (рис. 4.25)).

При таком представлении изображение разбивается на равные полигоны, возможно не пересекающиеся. Затем процедура повторяется для каждого блока до тех пор, пока его размер не станет равным элементу исходного изображения. Каждому блоку приписывается значение, называемое яркостью или цветом и являющееся

его обобщенной характеристикой [125, 173]. Разбиение изображений, не являющихся тривиальными, представляет собой одну из самых сложных задач обработки изображений. Конечный успех анализа изображений во многом определяется точностью разбиения.

Пирамидально-рекурсивные методы разбиения основываются на одном из двух базовых свойств сигнала яркости: разрывности и однородности. В первом случае подход состоит в разбиении изображений на основании резких изменений сигнала, таких как перепады яркости на изображении. Вторая категория методов использует разбиение изображения на области, однородные в смысле заранее выбранных критериев, например пороговая обработка, выращивание областей, слияние и разбиение областей.

Рис. 4.25. Пирамидальная структура процесса кодирования ПРМ

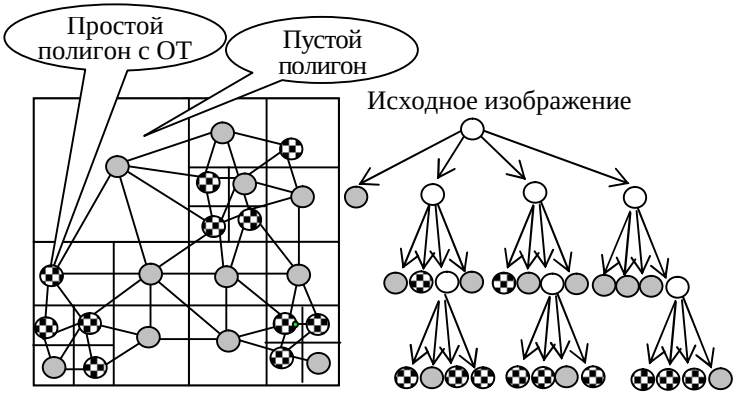
Регулярность пирамидальных структур позволяет создавать эффективные видеосистемы обработки изображений из-за возможности распараллеливания алгоритмов. Развитие ПРМ сдерживалось несколькими факторами:

- не разработан адекватный аппарат оперирования с данными, представленными в таком виде;
- нерешенными остаются многие вопросы реализации пирамидальных представлений на существующих ЭВМ как параллельного, так и последовательно типа;
- отсутствие чёткого определения *полезной информации* (опорные точки на исходном изображении) и критерия разбиения.

Развитие технологии систем на кристалле позволило перейти от теоретического моделирования к реализации быстрых ПРМ алгоритмов обработки видеоинформации.

Алгоритмы кодирования и декодирования изображений. Обработка проводится на изображении размером $Z \times Z$ точек (полигон нулевого уровня) с яркостным разрешением 2^m градаций. Ёмкость памяти для хранения цифрового изображения $C = Z^2 C_k$, где C_k скорость входного кода в битах на элемент.

При кодировании (рис. 4.26*а*) происходит последовательное разбиение составных полигонов до тех пор, пока все полигоны не станут простыми.



б Восстановление *а* Сжатие

Рис. 4.26. Пирамидальный процесс сжатия и восстановления изображения

Простой полигон – это полигон, где содержатся основные данные об объекте (либо опорная точку, либо пустой полигон, не содержащий полезной информации) и который не требует дальнейшей детализации. Составной полигон требует дальнейшей детализации. Алгоритм кодирования (рис. 4.27) формирует сжатое описание изображения, содержащее:

- информацию об опорных точках;
- информация о пустых полигонах;
- информация о структуре данных (квадродерево исходного изображения).

- При декодировании для каждой опорной точки осуществляется поиск соседних точек, до тех пор, пока не будут найдены все точки, затем они объединяются в треугольники и закрашиваются (рис. 4.28). Основное внимание при разработке алгоритмов уделяется обеспечению регулярности при поиске объектов на исходных изображениях (т. е. опорных точек).

Начало

Анализ полигона

Если имеется объект, то разбиение и сохранение новых полигонов.

Иначе кодирование опорных точек.

Конец

Рис. 4.27. Структура алгоритма кодирования ПРМ

Начало

Прием сжатого описания изображения

для всех ОТ поиск соседних точек

Триангуляция ОТ

Закраска треугольников

Рис. 4.28. Структура алгоритма декодирования ПРМ

Оценка эффективности кодирования и декодирования ПРМ. Предельный коэффициент сжатия без потери полезной информации определяется отношением исходной скорости кода к энтропии изображения. Введём следующие обозначения: J – число опорных точек (ОТ; M – максимальное число разбиений; d – число полигонов, получаемых при разбиении исходного полигона, т. е. основание полихотомии[3]; Z^2 – число пикселей исходного изображения; C_J – количество информации о местоположении опорных точек в кадре; m – разрядность яркости опорных точек, p_j – вероятность опорных точек (т. е. усреднённое отношение J/Z^2).

На поле изображения, содержащем Z^2 элементов, J опорных точек могут распределяться $C_{Z^2}^{J} = \frac{Z^2!}{J!(Z^2 - J)!}$ способами.

При равновероятном распределения опорных точек по полю все комбинации независимы, и энтропия изображения вычисляется по формуле Шеннона [174]:

[3] В экспериментах исследовано разбиение априорной области на различное число частей: d =2, … 6, т. е. дихотомия (деление на две части), трихотомия (деление на три части), тетрахотомия (деление на четыре части) и т. д.

$$H(p_j) = \log C_N^J = -p_j \log p_j - (1-p_j)\log(1-p_j).$$

Знание энтропии $H(p_j)$ опорных точек ничего не говорит о способе кодирования, при котором достигается равновесие в системе связи (равенство скорости создания информации и скорости передачи). Непосредственный способ кодирования информации об опорных точках потребует пропускной способности $C_н = J\,2\log Z$, которая сильно зависит от числа Z^2 элементов в кадре и существенно превышает энтропию $H(p_j)$. Поэтому при большом числе J опорных точек можно кодировать изображения следующим образом: на каждый элемент изображения отводится один бит, означающий наличие или отсутствие опорной точки, для чего потребуется $C_1 = Z^2$ бит. При $J > Z^2/2$ следует кодировать не расположение опорных точек, а номера элементов, не являющихся опорными точками. Это справедливо для любых способов кодирования, т. к. энтропия $H(p_j)$ симметрична относительно $p_j = ½$; далее ограничимся диапазоном $0 \le p_j \le ½$. Таким образом, под непосредственным способом кодирования подразумевается комбинация указанных способов, и требуемая пропускная способность в битах на кадр составит:

$$C_н = \begin{cases} p_j\, 2\log Z, & 0 < p_j \le \dfrac{1}{2\log Z}, \\ 1, & \dfrac{1}{2\log Z} < p_j \le 0{,}5. \end{cases}$$

Достижимый коэффициент сжатия $K(p_j)$ определяется отношением скоростей кода на входе и выходе кодера [175]:

$$K(p_j) = \frac{p_j m + C_н}{p_j m + C(p_j)}, \text{ где}$$

$$C(p_j) = \begin{cases} \log d(p_j M - \dfrac{1}{d^M}\sum LZ_L + E(E+1),\ d = 2 \\ \log d(p_j M - \dfrac{1}{d^M}\sum_{L=1}^{M} LZ_L + p_j \dfrac{2dE+E+1}{d-1},\ d > 2 \end{cases},$$

L – уровень разбиения, E – отношение пустых к заполненным полигонам, d – число новых полигонов после разбиения, Z_L – число заполненных (т. е. покрывающих ОТ) полигонов уровня L.

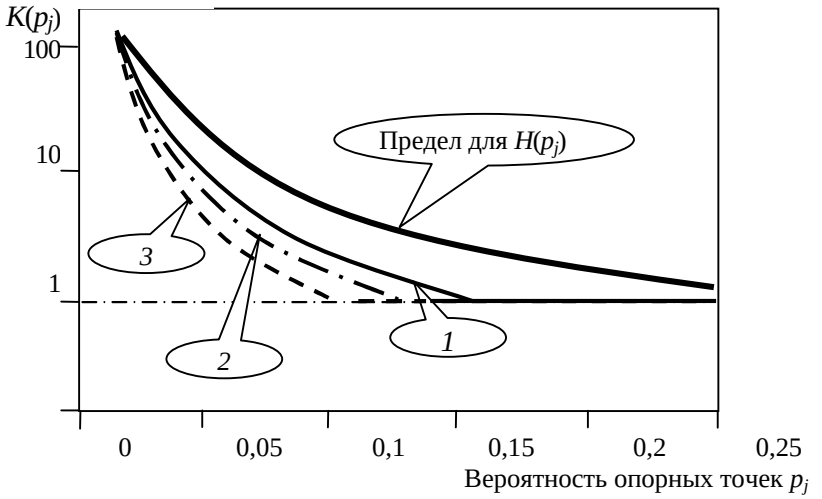


Рис. 4.29. Зависимость коэффициента сжатия от числа опорных точек при различных видах разбиения d=3 *1*, d=4 *2*, d=6 и d=2 *3*

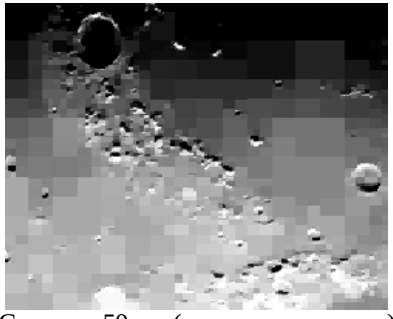
Достижимый коэффициент сжатия определяется энтропией изображения, которая зависит от числа опорных точек (рис. 4.29). Анализ результатов показывает, что различные основания полихотомии приводят к различному приближению к энтропии изображения. Вместе с тем требуемая пропускная способность канала связи $C(p_j)$ слабо зависит от числа разбиений M и при числе элементов в кадре $Z^2 > d^M > 10^3$ этой зависимостью можно пренебречь (ошибка на уровне десятых долей процентов). Так как оптимальное основание кода составляет иррациональное число e, то на практике наиболее близкий к энтропии источника результат достигается при трихотомии (рис. 4.28). Расчёт показывает, что при $d = 3$ или $d = 4$ можно достичь скорости кода, близкой к энтропии совокупности опорных точек $H(p_j)$.

Результаты моделирования. ПРМ алгоритмы кодирования и декодирования реализованы на языке высокого уровня C++Builder, используемом на начальной стадии системного проектирования СБИС класса СнК. Апробация ПРМ проведена при использовании разбиения на каждом шаге кодирования априорного интервала на 4 части ($d = 4$) на изображениях различных классов, в том числе – «лунная поверхность» (рис. 4.30) и «портрет» (рис. 4.31) при переменном коэффициенте сжатия исходного изображения размером 320×320 элементов и 8 разрядами сигнала яркости.

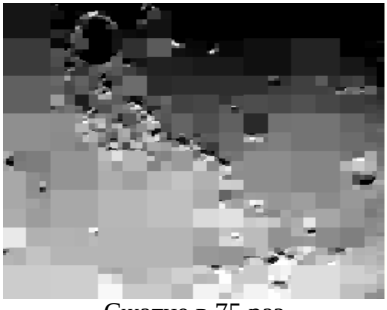
Исходное изображение

Сжатие в 10 раз (отлично)

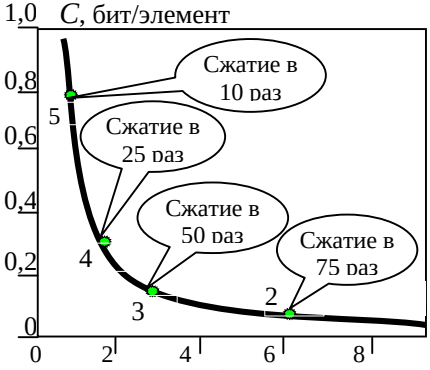
Сжатие в 25 раз (хорошо)



Сжатие в 50 раз (удовлетворительно)



Сжатие в 75 раз
(неудовлетворительно)



Зависимость требуемой пропускной способности от допустимой ошибки

Рис. 4.30 Взаимосвязь требуемой пропускной способности и допустимой среднеквадратической ошибки ($\sigma^2=\varepsilon$) для алгоритма ПРМ при d=4

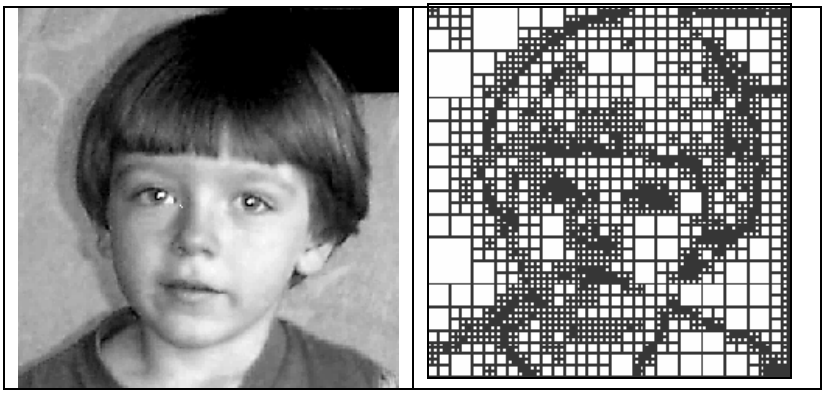


Рис. 4.31. Иллюстрация работы алгоритма ПРМ при разбиении на каждом уровне кодирования априорного интервала на 4 части ($d = 4$)

Приведённые, с целью достижения высокой степени сжатия при приемлемом качестве, анализ и эксперименты определяют основные характеристики для аппаратной реализации ПРМ алгоритмов, полагаемых в основу оценки требуемой производительности СБИС кодера и декодера на базе технологии СнК:

- скорость кода в битах/кадр и количество опорных точек;
- оптимальное сочетание числа полигонов, получаемых при разбиении исходного полигона d, числа уровней разбиения L, дисперсии ошибки ε и требуемой пропускной способности C.

Сравнение ПРМ с другими методами кодирования (рис. 4.32) показывает, что при заданной скорости кода C различные алгоритмы (ПРМ, Wavelet, JPEG) обеспечивают различное качество изображений. При этом качество изображений в смысле СКО при пирамидально-рекурсивном методе выше, чем при Wavelet и JPEG. В данном эксперименте скорость кода приблизительно равна энтропии совокупности опорных точек $C \approx H(p_j) \approx 3$ Кбайт/кадр, коэффициент сжатия $K \approx 25$.

Синтез алгоритмов сжатия видеоинформации и определение параметров устройств для реализации этих алгоритмов во многом опирается на модифицированный критерий минимума информационного риска $R_{I(3)}$ (3.11). На современном этапе твердотельной революции на кристалле размещается всё большее количество вентилей, и, как следствие, проектировщики назначают всё меньшую цену C_1 сложности W по сравнению с ценой C_0 потерь ΔI полезной информации.

Зависимость требуемой пропускной способности от допустимой СКО при различных методах кодирования

Wavelet

JPEG

ПРМ

Рис. 4.33. Сравнение алгоритмов сжатия по достижимой ошибке при коэффициенте сжатия 25

Поэтому повышенная сложность пирамидально-рекурсивного метода кодирования источника по сравнению с рассмотренными методами не критична. При этом лучшее приближение к энтропии источника делает пирамидально-рекурсивные методы перспективными для применения к кодированию статических изображений умеренной детальности, в том числе при аппаратной реализации метода с помощью систем на кристалле.

4.5. Системы на кристалле в мобильной телевизионной связи

Любая связь обеспечивает передачу конечного количества информации, и поэтому имеет тенденцию к реализации в цифровом виде. Общие тенденции сращивания телевизионной и компьютерной техники усиливаются единством микроэлектронной технологии и проявляются в том, что СнК активно внедряются в технику модуляции видеосигналов [166], и в том, что распространенные связные интерфейсы GSM/GPRS (пакетная передача данных), Wi-Fi (связь в нелицензируемом диапазоне радиочастот 2.4 ГГц.) становятся инструментарием прикладного телевидения[167–170].

Традиционным для Интернет является понятие – «последняя миля», которое отражает важность доставки информации конечному пользователю. Вместе с тем интеграция телевизионной камеры и связных интерфейсов позволяет решить проблему «*первой мили*» – передачи видеоизображения с мобильной точки его рождения. Перспективность этого направления можно проиллюстрировать следующими примерами.

Компания Nikon, выпустила новую цифровую фотокамеру *P1*, имеющую встроенную технологию Wi-Fi для беспроводной передачи изображений на Wi-Fi совместимые фотокамеры, компьютеры и принтеры. Поддержка беспроводных сетей (IEEE 802.11b/g) обеспечивает возможности беспроводной съёмки, передачи изображений и прямой печати фотографий. Функция беспроводной съёмки и передачи изображений реализует мгновенную «живую» передачу изображений сразу, как только они были сняты; беспроводная передача обеспечивает передачу изображений «по требованию» с фотокамеры на компьютер; беспроводная печать обеспечивает прямую печать фотографий на любом совместимом принтере при помощи дополнительного беспроводного адаптера печати [169].

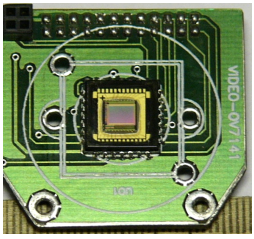
Компания CANON выпустила фотоаппарат Digital IXUS Wireless – беспроводная компактная цифровая фотокамера. Стиль и производительность IXUS теперь соответствуют новым требованиям беспроводных технологий – фотокамера, поддерживающая стандарт Wi-Fi (IEEE 802.11b), предоставляет новый уровень гибкости и комфорта при цифровой фотосъёмке. Digital IXUS WIRELESS имеет несколько инновационных функций, среди которых беспроводная передача изображений и беспроводное управление функциями съёмки с персонального компьютера. Фотоаппарат обеспечивает свободное беспроводное соединение для печати и передачи изображений из любой комнаты дома. Фотокамеру можно быстро и легко подключить к нескольким

устройствам (до 8 штук), в том числе точкам беспроводного доступа, что позволит выполнить интеграцию с домашними беспроводными сетями. В автоматическом режиме камера передаёт на персональный компьютер каждый отснятый кадр, а ПО, входящее в комплект поставки, обеспечивает отображение изображений на дисплее компьютера в режиме реального времени для просмотра фотографий. Дистанционное управление съёмкой: реализует беспроводное управление функциями фотосъёмки аппарата с помощью персонального компьютера. Во избежание перехвата для беспроводной передачи данных в Digital IXUS WIRELESS используется защищённое соединение. Для обеспечения дополнительной защиты данных, наряду с отраслевым стандартом WEP, камера также поддерживает WPA-PSK кодирование [167].

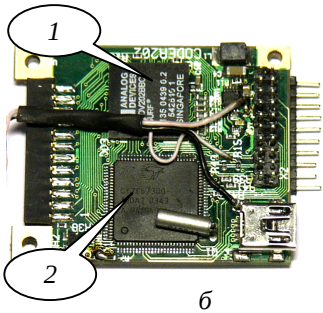
Эти примеры показывают актуальность разработок малогабаритных цифровых видеокамер, способных работать в нелицензированном диапазоне радиочастот 2.4 ГГц. Этот диапазон, именуемый ISM (Industrial, Scientific, Medicine – промышленный, научный и медицинский), используется во многих странах для безлицензионного доступа к среде передачи. Телевизионный комплекс, включающий абонентский терминал и видеокамеру с радиоинтерфейсом, обеспечивает совместное решение задач, связанных с управлением видеокамерой с удаленного терминала по радиоканалу, увеличением качества и объёма передаваемой по радиоканалу видеоинформации, определением достижимого качества кодирования и оперативности работы в широком диапазоне условий эксплуатации. Новизна комплекса состоит в апробации методов построения системы связи абонентского терминала с видеокамерой, функционирующего при участии одного оператора терминала. Видеокамера с радиоинтерфейсом включает в свой состав следующие функциональные узлы: фотопреобразователь – видеосистему на кристалле (рис. 4.34*а*), кодер источника формата JPEG 2000 *1*, с центральным процессором *2* (рис. 4.34*б*), кодер канала – приёмопередатчик радиоинтерфейса 2.4 ГГц (рис. 4.34*в*).

Фотоприёмник на основе КМОП технологии класса «видеосистема на кристалле» выполнен на приборе типа OV7141 компании Omnivision. (см. разделы 1.1, 4.2). Применение такого однокристального решения позволяет сосредоточить основные усилия на разработке программного обеспечения управления режимами формирования изображения.

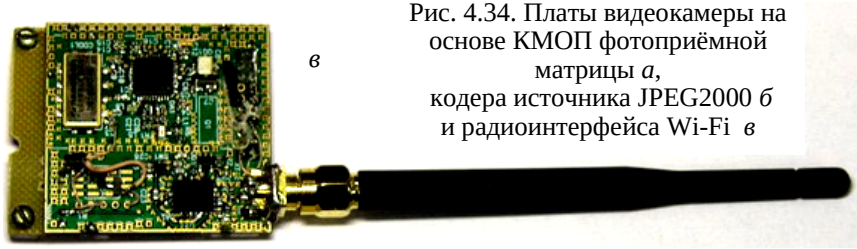
Кодер цифрового видеосигнала выполнен с применением СБИС ADV202 фирмы Analog Devices.



а



б



в

Рис. 4.34. Платы видеокамеры на основе КМОП фотоприёмной матрицы *а*, кодера источника JPEG2000 *б* и радиоинтерфейса Wi-Fi *в*

На конференции, посвящённой телевидению высокой чёткости DisplaySearch HDTV в августе 2005 г. компания Analog Devices продемонстрировала беспроводную передачу сжатого в реальном масштабе времени видеоизображения высокой чёткости для бытового применения. Применение микросхемы ADV202, единственного на настоящий момент времени устройства, реализующего стандарт JPEG 2000 высокой чёткости в реальном масштабе времени, резко сокращает видимые искажения изображения, возникающие в беспроводных каналах передачи, создавая хорошее качество изображения даже в зашумлённых условиях [169**]**.

В то время, как другие кодирующие технологии оказались дорогими, и не способными на практике обеспечить в реальном масштабе времени компрессию изображения высокой чёткости, стандарт JPEG2000 позволяет передавать высококачественные изображения высокой чёткости с малыми задержками и предназначен для использования в двух основных вариантах работы – кодирования и декодирования цифрового потока в формате JPEG2000. В варианте кодера, кодек осуществляет:

- по включении питания загрузку микропрограммного кода из энергонезависимой памяти в управляющий коммуникационный микропроцессор CY7C67300, фирмы Cypress;

• формирование опорных тактовых сигналов для работы видеодатчика и ядра микросхемы сжатия ADV202;

• получение команд управления и данных от платы приемопередатчика, их обработку в коммуникационном микропроцессоре и трансляцию команд на видеодатчик;

• приём цифрового видеосигнала от видеодатчика (8-битная шина D[7:0], сопровождаемая тактовым сигналом PCLK, сигналами синхронизации строк HREF и кадров VSYNC);

• кодирование сигнала изображения в формате JPEG 2000 и выдачу сжатого потока на микропроцессор. Инициатором выдачи является кодер ADV202, который выставляет сигнал прерывания, в ответ на который микропроцессор выполняет процедуру чтения некоторой (заранее определённой) порции данных (от 64 до 512 32-разрядных слов). Чтение производится по 16-разрядной шине HDATA с установкой необходимого адреса ADDR[3:0] и сопровождением стробирующими сигналами;

• обработку компрессированного цифрового потока и принятие решения о его выдаче в канал связи (на плату приемопередатчика);

• чтение состояния регистров видеодатчика по команде от абонента и передачу в канал связи прочитанных значений.

Структурная схема беспроводной видеокамера и взаимодействие однокристальной СБИС ADV202 с контроллером показана на рис. 4.35.

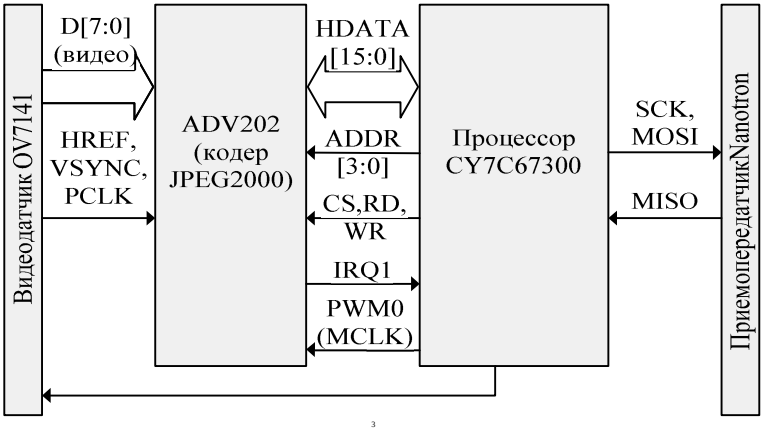


Рис. 4.35. Структурная схема беспроводной видеокамеры

Радиочастотный приемопередатчик системы видеонаблюдения. Беспроводные технологии передачи данных в настоящее время интенсивно развиваются. Для диапазона ISM от

2.4 до 2.5 ГГц, существует несколько вариантов организации радиоканала: Wi-Fi, Wireless USB, WiMAX, ZigBee(IEEE 802.15.4) и др. При выборе подходящей технологии передачи видеоинформации для мобильной видеокамеры необходимо учитывать: скорость передачи, радиус действия, энергопотребление, помехоустойчивость и сложность программного обеспечения. Результатом проведенного сравнительного анализа беспроводных технологий передачи стал выбор приёмопередатчиков фирмы Nanotron для организации двунаправленного обмена по радиоканалу между видеокамерой и приёмным терминалом.

Приёмопередатчики компании Nanotron диапазона 2.4 ГГц предназначены для беспроводной передачи данных, где скорости ZigBee и Bluetooth становится не достаточно, а устройства Wi-Fi и WiMAX не могут применяться в малогабаритных системах из-за высокой сложности программного обеспечения и повышенного энергопотребления.

Линейно-частотная модуляция (ЛЧМ, Chirp Spread Spectrum), используемая приёмопередатчиками Nanotron, представляет собой один из методов расширения спектра и позволяет повысить помехоустойчивость за счет того, что мощность сигнала «размывается» по спектру, и при воздействии помех фиксированной частоты теряется только часть передаваемого сигнала, так что двоичная информация может быть затем восстановлена в приёмнике. (В радиолокации ЛЧМ применяется для повышения точности измерения временного положения отражённых от целей радиоимпульсов). Формирование передаваемого ЛЧМ-сигнала и обработка принимаемого осуществляются с помощью дисперсионной линии задержки, выполненной на базе фильтра выполненного по технологии поверхностных акустических волн. ЛЧМ импульсы, используемые приёмопередатчиками Nanotron для передачи двоичных данных, имеют фиксированную длительность (1 мкс) и линейно нарастающую или спадающую частоту несущей. Полоса частот канала при этом составляет 64 МГц и значительно превышает ширину частотных каналов таких технологий как ZigBee и Bluetooth, которые также используются для беспроводной передачи в данном диапазоне. Это даёт возможность приёмопередатчикам Nanotron работать на более высоких скоростях и с более высокой степенью надёжности передавать информацию в условиях сложной помеховой обстановки (однако такая ширина спектра не позволяет использовать больше двух сетей в пределах радиовидимости).

По сравнению с технологией Wi-Fi, которая также имеет широкую полосу канала, приёмопередатчики Nanotron имеют лучшее соотношение показателей «дальность передачи – скорость передачи – энергопотребление» благодаря тому, что первичная обработка ЛЧМ импульса выполняется аналоговым способом. Поэтому приёмопередатчики Nanotron могут использоваться в носимых устройствах, работающих от аккумуляторов, аналогично мобильному телефону.

Основными особенностями приёмопередатчиков NanoNET TRX являются скорость передачи до 2 Мб/с, радиус действия до 900 метров на открытом пространстве и встроенный MAC-контроллер с поддержкой различных методов доступа к среде передачи. Мощность передатчика может программно изменяться в пределах от 1 мкВт до 6.3 мВт, что позволяет обменивать энергопотребление на дальность связи.

В настоящее время формируются дополнительные требования к стандарту IEEE 802.15.4. Этот стандарт будет иметь обозначение IEEE 802.15.4.a и его основным отличием будет введение жестких требований к физическому каналу передачи данных. Технология CSS компании Nanotron является одной из двух технологий, рассматриваемых в качестве решения для реализации альтернативного физического канала.

Радиочастотный приёмопередатчик видеокамеры предназначен для организации двунаправленного обмена по радиоканалу с платой приёмопередатчика абонентского терминала.

Основные характеристики платы приёмопередатчика:

- используемая элементная база – трансивер nanoNet TRX NA2TR1, выполненный по технологии система на кристалле;
- допустимое напряжение питания – 2.4 … 3.6 В;
- интерфейс обмена с ведущим контроллером – последовательный, SPI;
- максимальная скорость передачи данных – 2 Мбит/с;
- полоса частот приемопередающего тракта 64 МГц;
- чувствительность приемника при вероятности ошибки (BER) 10^{-3} и скорости передачи 1 Мбит/с – 92 дБм;
- потребляемый ток при передаче (усиление 11 дБм) – 86 мА;
- потребляемый ток в режиме приема – 40 мА;
- потребляемый ток в дежурном режиме – 1 мкА;
- температурный диапазон работы от –40 до +85 °С.

Система на кристалле nanoNet TRX NA2TR1 (рис. 4.36) включает:

- полностью интегрированный аналоговый приемник (от выхода антенны до цифровой демодуляции);
- полностью интегрированный передатчик (от цифрового входа до высокочастотного усилителя мощности, который может быть подключен ко входу антенны);
- программируемый цифровой контроллер, взаимодействующий с внешним микроконтроллером через последовательный интерфейс SPI. Этот контроллер включает контроллер обработки в основной полосе частот (baseband-контроллер) и контроллер доступа к среде (MAC-контроллер – Media Access Control). Baseband-контроллер выполняет обработку данных (разбиение данных на кадры, коррекцию ошибок, шифрование–дешифрование данных и т. д.), MAC-контроллер реализует алгоритмы доступа к среде CSMA/CA, TDMA и гибридные алгоритмы доступа к среде передачи (радиоканалу). Используя ЛЧМ модуляцию трансивер оперирует импульсами ЛЧМ длительностью символа $T_{symbol} = 1$ мкс при эффективной полосе частот 64 МГц. Трансивер, используя такую модуляцию, позволяет выбрать программно одну из трёх физических скоростей передачи: 500 Кбит/с, 1000 Кбит/с и 2000 Кбит/с.

Рис. 4.36. Схема состава приёмопередатчика Nanotron

Внешний микропроцессор взаимодействует с системой на кристалле через буфер типа FIFO (First In, First Out). Трансивер оперирует с двумя буферами FIFO, каждый из них имеет объём 1024 бит приемного буфера и 1024 бит передающего буфера. Буфер FIFO предназначен для временного сохранения принятого пакета или загрузки готового для передачи пакета. Эти блоки памяти могут быть программно сконфигурированы для доступа независимо один от другого, т. е. в приёмный буфер поступают данные от высокочастотной части почти одновременно с передачей данных из

передающего буфера и обменом информацией с указанными блоками памяти по интерфейсу SPI.

Дополнительный программируемый блок содержит часы реального времени, схему управления питанием, схему измерения напряжения батареи и несколько подстроечных и калибровочных функций для аналоговой части системы на кристалле. Все функции этого блока управляются программно. Индикатор уровня принимаемого сигнала (RSSI) может быть считан внешним микроконтроллером. Значение уровня RSSI является необходимым при организации доступа к среде передачи CSMA/CS для получения информации о занятости радиоканала.

Обработка принимаемого сигнала в системе на кристалле nanoNet TRX NA2TR1 включает в свой состав:

- скремблирование – вычисление CRC кода и его подстановка при передаче и приёме и проверка такого кода при приёме пакета. (типы CRC, поддерживаемые трансивером: ISO/IEC3309, CCITT X.25, X.75, ETS 300 125 / IEC 60870-5-1 / CCITT-32);
- шифрование–дешифрование с длиной ключа 128 бит;
- канальное кодирование методом Хаффмана для коррекции ошибок.

Скремблирование, шифрование–дешифрование, избыточное кодирование и тип применяемого проверочного кода CRC выбираются программно. В качестве усилителя мощности для платы приёмопередатчика использована микросхема SE2526A фирмы SiGe, предназначенная для увеличения радиуса действия радиоканала. Она обеспечивает усиление сигнала по мощности до 20 дБ. Схема её включения в состав платы приёмопередатчика полагает программно-управляемое переключение мощности.

Построенная таким образом малогабаритная система мобильной видеосвязи решает поставленную проблему «*первой мили*» – мобильной связи с источником видеоинформации благодаря адаптации коэффициента сжатия и мощности передатчика к свойствам канала связи.

Подобные мобильные телевизионные системы позволяют сменить парадигму телевидения, которое ранее было исключительно централизованным [114]. Новая услуга – «видео по запросу» как оперативная обратная связь между телецентром и зрителем является первым шагом в *децентрализации* и *глобализации* передачи изображений, а мобильная (пока малокадровая) передача телевизионных сигналов между большим числом пользователей в сети является решающим шагом.

ЗАКЛЮЧЕНИЕ

В силу молодости самой твердотельной революции в телевидении термин «Твердотельное телевидение» 20 лет назад был новым, но ныне он устоялся, и новый этап твердотельной революции важен не только сам по себе, но ещё и служит подтверждением важности её первого этапа. То, что, при написании книги [1] твердотельное телевидение только набирало силу, нашло отражение, в частности, в том, что её авторы, энтузиасты твердотельного телевидения, в заключении к книге делали уступки традиции и прогнозировали: «Конечно, сохранится и значительный объём «старой» аппаратуры типа традиционных видиконных ПТВУ с постоянными параметрами разложения для визуального контроля». Теперь уже почти никто не вспоминает о видиконных телекамерах при проектировании новых систем. Время написания настоящей книги – тоже переходной период от первого ко второму этапу твердотельной революции в телевидении. Поэтому её авторы могли бы также написать о сохранении значительного объёма «старой» аппара-туры, понимая под ней телекамеры на ПЗС. Однако история подсказывает, что тактические методы, отражающие консерва-тизм в технике, не противоречат стратегическому подходу, ориентированному на существенную новизну. Поэтому, предсказывая окончательную победу КМОП фотоприёмникам и видеосистемам на кристалле, можно говорить лишь о времен-ном сохранении «старой» аппаратуры. Технология КМОП оказалась прорывной и одной из важнейших конверсионных технологий. Её дешевизна по сравнению с ПЗС, совместимость с технологией устройств обработки видеосигналов, лучшая радиационная стойкость создают предпосылки для необратимости второго этапа твердотельной революции. КМОП мультиплексоры становятся главными «ответственными» за реализацию развёртки не только в приборах видимого диапазона, но и в ИК матрицах.

Важнейшим аспектом твердотельной революции в телевидении является органическое вливание техники телевидения в единый поток систем связи. Уже сейчас число пользователей Интернет в мире превысило миллиард, а число пользователей мобильных телефонов – два миллиарда. В ближайшем будущем связь, включая основанную на технологии интерактивного цифрового телевидения мобильную телевизионную связь, должна стать поистине *глобальной.*

Сегодня никакая система связи уже не может создаваться в отрыве от *рукотворных кристаллов*. Единство всей *техники* связи обеспечивается едиными *теорией* проектирования на основе теории информации и микроэлектронно-компьютерной *технологией*.

Результатом твердотельной революции в прикладном телевидении стало его массовое внедрение во многие отрасли хозяйства, науки и обороны. Твердотельное телевидение используется всюду: в системах безопасности и в АСУ промышленных и оборонных объектов, в космосе и под водой, причём не только в очевидных областях наблюдения изобра-жений, но и в неожиданных областях приборостроения, например, в гравиметрии, позволяющей осуществлять навигацию самолетам и подводным лодкам. Твердотельная революция, неизбежно приведшая к *слиянию телевидения и вычислительной техники*, в частности, привела к существенному изменению технологии астрономических наблюдений. Если изобретение фотографии лишь усилило проницающую способность телескопов, то современные твердотельные матричные фото-приёмники в сочетании с компьютерами «закрыли» этап фотографии в астрономии, оставив ей лишь короткий в истори-ческом смысле эпизод перевода уже имеющихся фотографий (порядка миллиона) в электронную форму для хранения в памяти компьютеров. Характерно, что такое преобразование делается также с помощью сканеров на твердотельных телевизионных приборах – линейных ПЗС.

Прикладные телевизионно-компьютерные системы по мере развития становятся всё сложнее и всё точнее. Ещё Дж. Фон Нейман предупреждал, что «учёт ошибок становится тем важнее, чем больше усложняется система». Поэтому высокую сложность и высокую точность можно реализовать, только вооружившись новейшими достижениями теории информации и теории алгоритмов. В этом свете авторы считают свою книгу, посвящённую самой информационноёмкой отрасли связи, вкладом в осознание обществом смены парадигмы: как бы ни была ценна энергия, наступила эра информации и информационная парадигма вытесняет энергетическую. Не случайно говорят о глобальной информатизации, и никто не говорит о глобальной «энергетизации». Новая *информационная парадигма* отражается и в учёте не только технологических аспектов создания телевизионных систем, но и в научном подходе, неизбежно оперирующем взаимосвязью *точности, сложности и априорной*

неопределённости при рождении и передаче видеоинформации. Этот подход привёл к опоре на понятие равновесия при информационном проектировании систем связи реального времени, формализующего *принцип минимума информации при заданном её качестве* посредством *уравнения связи* [3]. Решение уравнения связи преследует цель достижения равновесия в скорости создания информации источником сигнала и скорости передачи информации через канал связи. Решение этих задач открывает путь к решению ключевых проблем связи, которая должна в обозримом будущем перестать служить только прикладным целям, а стать *фундаментальной наукой*, которая оформится как раздел теории познания. При этом телевидение, являясь лидером в передаче информации, должно в свои теоретические основы включить не только сигнальную и информационную концепции связи, но и концепцию рождения и передачи образов.

Твердотельная революция в телевидении началась тогда, когда уровень точности изготовления СБИС составлял 3…5 мкм, её второй этап, связанный с появлением видеосистем на кристалле, начался при освоении ультрафиолетовой литографии и проектных норм точности 0,35 мкм [6]. И хотя сегодня ведущие производители СБИС владеют технологиями уровня 0,065 мкм, но и строительство фабрик с проектными нормами 0,18 …0,13 мкм (в сочетании с производством пластин кремния диаметром 200 мм), предусмотренное Федеральной программой перевооружения радиоэлектронной отрасли, выводит отечественную микроэлектронику на уровень, *минимально достаточный* для полностью самостоятельного выпуска видеосистем на кристалле класса «активный пиксел». Для создания видеосистем на кристалле класса «цифровой пиксел» требуются более высокие проектные нормы, и (будем надеяться – временная) кооперация отечественных системных и кристальных центров проектирования СБИС с зарубежными изготовителями кристаллов. Управление радиоэлектроники и систем управления РФ в рамках программы «Национальная технологическая база» на 2007–2011 вкладывает средства и создание сети дизайн-центров, и в перевооружение собственного микроэлектронного производства [8]. В результате, можно надеяться, российские предприятия перестанут «со стороны наблюдать за происходящими в мире фотоэлектроники процессами, не имея возможности принять в них какое-то участие, так как для выпуска современных КМОП датчиков необходимо существенное переоснащение производства» [42].

Появление систем на кристалле и видеосистем на кристалле не только расширило сферу применения прикладного телевидения (например, благодаря возможности создания телевизионной системы весом несколько граммов, получили распространение беспилотные летательные аппараты, стартующие с ладони), но и *сменили метод проектирования* прикладных систем, ещё более срастив его с проектированием главных систем информатики – компьютерами. Если ранее местом соединений, задаваемых разработчиком, была печатная плата, то современные технологии направлены на реализацию максимально возможного, в пределе – всего многообразия соединений внутри одного кристалла. Это заставляет заново оценивать и пересматривать, дополнять сложившиеся методы проектирования. Такие требования резко повышают значимость современных САПР и обязательно должны быть поддержаны на этапе подготовки специалистов, что и делается сейчас в рамках Программы подготовки специалистов для высокотехнологичных отраслей промышленности. В определённом смысле и данную книгу можно рассматривать как скромный вклад в подготовку кадров для отечественной промышленности, занятой разработкой систем прикладного телевидения. Книга должна влиться в общий поток отечественных книг по радиоэлектронике, отражающий возрождение интереса к этой отрасли. Весьма показательно, что после периода спада в книгоиздательском деле «за два года выпущено 80 книг по различным вопросам науки и техники, в том числе и электроники; в этом году должно выйти ещё порядка 80 новых книг» [16].

Наука и технология постоянно развиваются, и у «Твердотельного телевидения» должны будут появляться продолжения – как у Дюма: после «20 лет спустя» было еще «10 лет спустя». Так же, как на первом этапе твердотельной революции нельзя было угадать свойства техники её второго этапа, сейчас трудно предсказать методы совершенствования прикладной телевизионной техники будущего. Можно, конечно, надеяться, что общее направление «нанотехнологий» породит и «наноинформатику». И, помня, «что знание относительно, и только вера абсолютна» и «что лучше один раз увидеть, чем сто раз услышать», можно с уверенностью сказать, что никто никогда не откажется развивать прикладное телевидение и видеоинформатику – главный парк информационных «зрячих» машин, многократно усиливающих мощь человеческого познания.

Литература

1. Твердотельное телевидение: Телевизионные системы с переменными параметрами на ПЗС и микропроцессорах /Л. И. Хромов, Н. В. Лебедев, А. К. Цыцулин, А. Н. Куликов, М.: Радио и связь, 1986. – 184 с.
2. Хромов Л. И., Цыцулин А. К., Куликов А. Н. Видеоинформатика: передача и компьютерная обработка видеоинформации. М.: Радио и связь, 1991. – 192 с.
3. Хромов Л. И. Теория информации и теория познания. СПб, РФО, 2006. – 200 с.
4. Цыцулин А. К. Телевидение и космос. СПб, СПбГЭТУ «ЛЭТИ», 2003. – 228 с.
5. Умбиталиев А. А., Цыцулин А. К. Основные направления развития телевизионной техники. Вопросы радиоэлектроники, сер. Техника телевидения, 2006, №1, с. 3–12.
6. Авдонин Б. Н., Мартынов В. В. Электроника. Вчера… Сегодня. Завтра?: Очерки по истории, технологии, экономике. М.: ИКП «Дека», 2005 – 600 с.
7. Алфёров Ж. И. Перспективы электроники в России. Гетеро-структурная электроника и акустоэлектроника //Электроника: наука, технология, бизнес. 2004, №6, с. 90–93.
8. Борисов Ю. И. Отечественная электронная промышленность и компонентная база. Перспективы развития. //Электроника: Наука, технология, бизнес. 2006, №2, с. 6–9.
9. Борисов Ю. И. Первая отечественная система на кристалле с быстро-действующими ЦАП/АЦП 600 Мвыборок/с по двум квадратурным каналам // Электроника: наука, технология, бизнес. 2004, № 2, с. 36–42.
10. Борисов Ю. И., Шабанов Б. М. Одно из направлений развития САПР для создания сложных технических систем // Информационные технологии, 2003, № 10, с. 2–17.
11. Борисов Ю. И. Новый шаг отечественной микроэлектроники// Наука и технологии в промышленности. 2003, №12, с. 3 – 11.
12. Немудров В. Г., Мартин Г. Системы на кристалле. Проблемы проектирования и развития. М.: Техносфера, 2004. – 216 с.
13. Немудров В. Г. «Прогресс» в проектировании отечественных СБИС. Электроника: наука, технология, бизнес. 2002, №6, с. 5–9.
14. Бухтев А., Немудров В. Системы на кристалле. Новые тенденции /Электроника: наука, технология, бизнес. 2004, №3, с. 52–56.
15. Кривченко И. Системы на кристалле: общее представление и тенденции развития //Компоненты и технологии. 2001, № 6, с. 49-54.
16. Концепция и стратегия развития электроники России. // Электроника: наука, технология, бизнес. 2006, № 3, с. 4–17.
17. www. niitv.ru
18. www. Dalsa. Com
19. www. Toshiba. com
20. www. Atmel. com

21. www. Hewlett-Paccard. com
22. www. Sony. com
23. www. Omni Vision. com
24. **www.micron.com/products/imaging/products/mt9m413.html**.
25. Ракович Н. Н. CameraChip: полная видеосистема на кристалле //Компоненты и технологии. 2004, №1, с. 128 – 130.
26. Стемпковский А., Шилин В. КМОП-фотодиодные СБИС. Перспективная элементная база однокристалльных систем приема и обработки информации// Электроника: наука, технология, бизнес. 2003, №2, с. 14–20.
27. Березин В. В. Видеосистемы на кристалле – новый этап интеграции в телевизионной технике // Изв. вузов России. Радиоэлектроника. 2005, № 3, с. 31–41.
28. E. R. Fossum. «Active pixel sensors – Are CCD's dinosaurs?» Charge-Coupled Devices and Optical Sensors III, Proc. SPIE, vol. 1900, pp. 2–14, 1993.
29. Kozlovski L.J. , Stanley D., Luo J., Tomasini A. et al. Theoretical basis and experimental confirmation: why a CMOS imager is superior to a CCD. Proceedings of the SPIE, vol. 3698, FL, April 1999, pp. 388–396.
30. Кузнецов Ю. А., Скрылёв А. С. Шилин В. А. Разработка основ проектирования и субмикронной технологии изготовления ФПЗС нового поколения//НТК «Пути развития телевизионных фотоэлектронных приборов и устройств на их основе», СПб, 2001, с. 89–90.
31. Morton (Z) Scan Based Real-Time Variable Resolution CMOS Image Sensor // E. Artyomov, Y. Rivenson, G. Levi, O. Yadid-Pecht / IEEE transactions on circuits and systems for video technology, V. 15, № 7, pp. 947–952.
32. Active-Pixel Image Sensors With Programmable Resolution" (NPO-19510), NASA Tech Briefs, Vol. 20, No. 5, (May 1996), p. 26.
33. Nirmaier T, Droste D., Bille J. Hartmann-Shack sensor ASIC's for real-time adaptive optics in biomedical physics. IIIS Congress on systemics, cybernetics and informatics, v. XIII, Orlando, 2002, p. 1–5.
34. E. Artyomov, O. Yadid-Pecht. Adaptive multiple resolution CMOS active pixel sensor, IEEE ISCAS, Vancouver, BC, Canada, pp. III-786-789, May 2004.
35. Novel CMOS image sensor with a 132-dB dynamic range / D. Stoppa, A. Simoni, L. Gonzo M. Gottardi // IEEE journal of solid-state circuits, vol. 37, N 12, dec 2002, pp. 1846–1852
36. Березин В. В., Фахми Ш. С. Проектирование устройств обработки сигналов на основе технологии «система на кристалле». СПб.: СпбГЭТУ «ЛЭТИ», 2005. – 148 с.
37. Березин В. В. Видеосистемы на кристалле: новые системные возможности// Вопросы радиоэлектроники, сер. Техника телевидения, 2006, вып. 1, с. 63–71.
38. Березин В. В., Умбиталиев А. А., Цыцулин А. К. Структурный синтез «видеосистем на кристалле» повышенного быстродействия// Электронные

средства и системы управления: Докл. междунар. науч.-практ. конф. Томск, изд. Инст. оптики атмосферы СО РАН, 2005, ч.1, с.242 – 245 .
39. Березин В. В., Умбиталиев А. А., Цыцулин А .К. Анализ архитектур устройств класса «видеосистемы на кристалле»// Проблемные вопросы сбора, обработки, передачи и защиты информации в сложных радиотехнических системах. Материалы 7-й межв. НТК СПб, г. Пушкин, 2005 г. стр. 122–125.
40. Передовая технология компании Pixim//«CCTV Фокус», №4, 2003, с. 12–15.
41. Смирнов В. Д. Цифровые фотоаппараты. СПб, Петербургский институт печати, 2002. – 192 с.
42. Минкин В. А. , Татаурщиков С. С., Штамм А. И. Ближайшие перспективы развития твердотельных датчиков изображения//XII НТК «Пути развития телевизионных фотоэлектрических приборов и устройств на их основе», СПб, 2001, с. 67–68.
43. Пресс Ф. П. Фоточувствительные приборы с зарядовой связью. М.: Радио и связь, 1991. – 264с.
44. Формозов Б. Н. Введение в криогенную микроэлектронику. СПб.: Наука, 2001. – 328 с.
45. Формозов Б. Н. Аэрокосмические фотоприёмные устройства в видимом и инфракрасном диапазонах. СПб.: СПбГУАП, 2002. – 121 с.
46. Телевизионные камеры «ЭВС»//CCTV'2005. Каталог продукции, с. 49.
47. Отечественные УФ и ИК ФПЗС и цифровые камеры на их основе. /Вишневский Г. И., Выдревич М. Г., Нестеров В. М., Ривкинд В. Л.// Электроника: наука, технология, бизнес. 2003 N8 стр. 18 – 22.
48. Малков А. В. IP-камеры для безопасности крупных городов//CCTV'2005. Каталог продукции, с. 31, 32.
49. Арутюнов В. А., Иванов В. Г., Каменев А. А., Прокофьев А. Е. Методика оценки потенциальных характеристик обнаружения малоразмерных аэрокосмических целей многоспектральной оптико-электронной аппаратурой на матричных фотоприемниках// Вопросы радиоэлектроники, сер. Техника телевидения, 2006, вып. 2, с. 47–69.
50. Кузнецов Ю. А., Шилин В. А. Микросхемотехника БИС на приборах с зарядовой связью. М.: Радио и связь, 1988. – 160 с.
51. Зайцев В. В. Центр проектирования СБИС «ОТИК-Дизайн»: опыт, возможности, кооперация// Практические аспекты разработки отечественных СБИС класса «Система на кристалле», Геленджик, 2006.
52. Носов Ю. Р., Шилин В. А. Полупроводниковые приборы с зарядовой связью. М.: Сов.Радио, 1976. – 144 с.
53. Манцветов А. А., Цыцулин А. К. Телекамеры на КПОП фотоприём-никах. Вопросы радиоэлектроники, сер. Техника телевидения, 2006, вып. 2, с. 70–89.
54. Куликов А. Н. Телевизионные наблюдения в сложных условиях//Специальная техника, 2000, № 4.
55. Сверхскоростные КМОП-телекамеры MotionPro HS//Security news, 10.07.2006.
56. www.fastcam.com.

57. www.axis.com**.**
58. www.arecontvision.com**.**
59. Скляр Б. Цифровая связь. Теоретические основы и практическое применение. М.: «Вильямс», 2003. – 1104 с.
60. Захаренков В. Ф., Раковский Ю. Н., Шибаев С. Н. Об ошибках определения координат точечного источника методом центроида//Оптический журнал, 2001, №8, с. 61–66.
61. Березин В. В., Гатаулин В. М., Цыцулин А. К. Стабилизация изображения при наблюдении с беспилотных летательных аппаратов // Изв. вузов. Приборостроение, 2005, № 4, с. 65–72.
62. D. Navarro, G. Cathebras and F. Gensolen A block matching approach for movement estimation in a CMOS retina: principle and results // Procceeding of the European Solid-State Circuits Conference ESSCIRC-03, 2003, pp.615–618.
63. T. Komuro, I. Ishi, M. Ishikawa and A. Yoshida. High Speed Target tracking vision chip // Computer Arkhitectures for Machine Percetion, 2000. Proceedings. Fifth IEEE International Work, pp.49–56.
64. J. Schreiter et al. Design of Parallel Preprocessing Image Sensors // 3 Conf. on Knowledge Based Intelligent Information Engineering Systems KES'99, Adelaide, Australia, Aug. 1999.
65. Звёздные координаторы систем ориентации космических аппаратов/Аванесов Г. А., Воронков С. В., Форш А. А., Куделин М. И. // Известия вузов. Приборостроение, 2003, вып. 4, с. 66–69.
66. Вудворт Ф. М. Теория вероятностей и теория информации с применением в радиолокации. М.: Сов. Радио, 1955.– 128 с.
67. Дополнительные каналы считывания в матричных ПЗС/ А. Н. Куликов, С. А. Иванов, А. К. Цыцулин и др.// Техника средств связи, сер. Техника телевидения, 1990, вып. 2, с. 34–39.
68. Березин В. В. Обработка сигнала матричного ФППЗ при определении координат малоразмерного объекта//Радиоэлектроника и связь. 1992, №2–3, с. 71–76
69. Березин В. Б., Березин В. В., Соколов А. В., Цыцулин А. К. Адаптация считывания в астрономической телекамере на матричном ПЗС // Известия вузов. Сер. Радиоэлектроника, 2004, вып. 4, с. 36–45.
70. Березин В. Б., Березин В. В., Переспелов А. В, Цыцулин А. К. Адаптация дискретизации в астрономической системе на матричном ПЗС// Вопросы радиоэлектроники, сер. Техника телевидения, 2006, вып. 1, с. 72–78.
71. Переменная чёткость в твердотельных телевизионных системах /С. А. Иванов, А. Н. Куликов, Д. А. Довжиков, и др.//Техника средств связи, сер. Техника телевидения, 1989, вып. 7, с. 27–33.
72. Манцветов А. А., Березин В. Б., Цыцулин А. К. Совместное обнаружение и оценивание параметров сигналов телекамерами на ПЗС// Телевидение. Передача и обработка изображений. 3 международная конференция, 3–6 июня 2003, СПб, с. 80, 81.
73. Ноценко В. С., Хромов Л. И. Принцип счета фотонов в телевидении// Техника средств связи, сер. Техника телевидения, 1991 вып 1, с. 3–7.

74. Смелков В. М., Иванов С. А. Камеры с предельной чувствительностью на ПЗС. Техника средств связи, сер. Техника телевидения. 1985, вып. 2, с. 26–32.
75. CCD Imaging Cameras Provide High Precision // Europhotonics, October/November 2003, p. 40.
76. Никулин О. Ю., Неизвестный С. И. Приборы с зарядовой связью – основа современной телевизионной техники. Основные характерис-тики ПЗС// Специальная техника. 1999, №5, с.17–29.
77. Вахромеева О. С., Манцветов А. А., Шиманская К. А. Характеристики чувствительности телевизионных камер на матричных приборах с зарядовой связью// Известия ВУЗов России. Радиоэлектроника. 2004, вып. 4, с. 25–35.
78. Quinn Gary. High-speed Imaging Requires Progressive-Scan Technology. Europhotonics, April/May 2004 p.p. 26–27.
79. Вето А. В., Жильцов В. С., Скакун А. В. Цифровые ТВ камеры для наблюдения и охраны// Системы безопасности, связи и телекоммуникаций. – 1999, № 25, с. 78 – 80.
80. Метод высокоточного измерения среднего сигнала темнового тока ПЗС в астрономических телевизионных камерах/ Березин В. Б., Березин В. В. и др. // Известия СПбГЭТУ «ЛЭТИ», сер. Радиоэлектро-ника и телекоммуникации. 2004, вып. 1, с. 39–42.
81. Шэндл Д. Датчики изображения со структурой сетчатки для систем технического зрения// Электроника, 1993, №17, с. 7, 8.
82. Пахомов А. Н. Состояние и перспективы систем активного зрения// Зарубежная радиоэлектроника, 1999, №2, с. 57–65.
83. Выбор типа фотоприёмника в телевизионной системе измерения пространственно-временных координат объектов/ А. А. Манцветов, А. К. Цыцулин, С. Н. Коноплёв и др.//Изв. СПбГЭТУ «ЛЭТИ», 2002, вып. 3, с. 24–27.
84. Комаров В. В. Возможности российских цифровых ПЗС камер серийного производства для астрономических приложений/ XIX Международная НТК по фото-электронике и приборам ночного видения, 23–26 мая 2006, М., с. 173.
85. Телевизионные датчики изображения в радиационных полях / В. М. Коняев, С. С. Красовский, И. Н. Суриков, В. И. Флеров. Рига: Зинатне, 1989. –157 с.
86. История, современное состояние и перспективы развития полупроводниковых матричных ИК-фотодиодных структур и технологий / В. М. Акимов, Е. А. Климанов, В. П. Лисейкин и др. // Прикладная физика, 1999, № 1, с. 3–15.
87. Тарасов В. В., Якушенков Ю. Г. Инфракрасные системы «смотрящего типа» М.: Логос, 2004. – 444 с.
88. Модуль цифровой обработки ИК-изображений с матричных фотоприёмных устройств/Ю. Борисов, А. Грошев, В. Михайлов и др.//Компоненты и технологии. 2002, №2, с. 29–30.

89. Цыцулин А. К., Ресовский В. А., Березин В. В. Космос и малокадровое телевидение//Информация и космос. 2005, №2, с. 80–86
90. Овсянников В. А. Филиппов В. Л. Критерий целесообразности использования в несканирующих тепловизионных приборах микросканирования/ XIX Международная НТК по фотоэлектронике и приборам ночного видения, 23–26 мая 2006, М., с. 83.
91. Сравнение матричных ФПУ на основе InSb с различными типами охлаждаемых мультиплексоров/ Чишко В. Ф., Касаткин И. Л., Лопухин А. А., Голубков А. В.// XIX Международная НТК по фото-электронике и приборам ночного видения, 23–26 мая 2006, М., с. 118.
92. Особенности работы и характеристики МФПУ формата 128х128 и 256×256 на основе InSb в режиме «мгновенной фотографии» и конвейерного считывания/ Чишко В.Ф., Касаткин И.Л., Лопухин А.А., Голубков А.В.//XIX Международная НТК по фотоэлектронике и приборам ночного видения, 23–26 мая 2006, М., с. 118.
93. Унифицированный тепловизионный прибор на основе микроболометрической матрицы/Тареев А. М., Бондаренко М. М., Чеча Э. П. и др.// XIX Международная НТК по фото-электронике и приборам ночного видения, 23-26 мая 2006, М., с. 157.
94. Программирование на C++. Учеб пособие/А. Аверкин, А. Бобровский, В. Веснич и др.; Под ред. А. Хомоненко. М.: Корона принт, 2003. – 512 с.
95. Березин В. В. Золотухо Р. Н. 32-разрядная реконфигурируемая система на кристалле фирмы Triscend. Компоненты и технологии. 2003, №2, с. 14 – 20.
96. Березин В. В. Методология автоматизированного проектирования с применением технологии «система на кристалле» (аппаратная часть) // Промышленные контроллеры и АСУ. 2004, № 11, с. 37–41.
97. Березин В. В. Методология автоматизированного проектиро-вания с применением технологии «система на кристалле» (программная часть и системная интеграция) // Промышленные контроллеры и АСУ. 2004, № 12, с. 38–40.
98. Грушвицкий Р. И., Мурсаев А. Х., Угрюмов Е. П. Проектирование систем на микросхемах программируемой логики. СПб.: БХИ-Петербург, 2002. 608 с.
99. Стешенко В. Б. EDA. Практика автоматизированного проектирования радиоэлектронных устройств. М.: Издатель Молгачева С. В., Издательство «Нолидж», 2002. – 768 с.
100. Евтушенко Н., Немудров В., Сырцов И. Методология проектирования систем на кристалле. Основные принципы, методы, программные сред-ства // Электроника: наука, технология, бизнес. 2003, № 6. с. 7–11.
101. Алексеев С. Ю., Березин В. В., Золотухо Р. Н. Разработка программного обеспечения реконфигурируемых систем на кристалле семейства A7 фирмы Triscend // Компоненты и технологии. 2004, № 2, с. 138–142.

102. Березин В. В., Золотухо Р. Н., Фахми Ш. С. Отладка аппаратно-программного обеспечения реконфигурируемых систем на кристалле // Компоненты и технологии. 2003, № 7, с. 118–122.
103. Березин В. В., Фахми Ш. С. Использование возможностей контроллера динамической памяти как составной части «системы на кристалле» // Неразрушающий контроль и диагностика окружающей среды, материалов и промышленных изделий. СПб., 2004, с. 213–225. (Межвуз. сб. СЗТУ. Вып. 9.)
104. Березин В. В. Аппаратный стек протоколов TCP/IP в сетях Ethernet 10/100 Mb // Телекоммуникации. 2004, №4, с. 18–21.
105. Березин В. В., Лячек Ю. Т., Фахми Ш. С. Автоматизация проектирования электронных устройств: Учеб. пособие. СПб.: СПбГЭТУ «ЛЭТИ», 2003. – 80 с.
106. Nirmayer T., Droste D., Bille J/ Hartmann-Shack sensor ASIC's for real-time adaptive optics in biomedical physics/ IIIS Congress on systemics, cybernetics and informatics, Vol. XIII, Orlando, 2002.
107. Цифровое преобразование изображений: Учебн. пособие для вузов / Быков Р. Е., Фрайер Р., Иванов К. В., Манцветов А. А.; под ред. Быкова Р. Е.. – М.: Горячая линия–Телеком, 2003. – 228 с.
108. Кириллов В. И. Высокоэффективные системы информационного обмена для пространственно разнесённых телевизионных комплексов. Минск, Вышэйша школа, 1989. – 212 с.
109. Иванов В. Г., Березин В. В., Цыцулин А. К. Проблемы прикладного телевидения// Телевидение. Передача и обработка изображений. 3 Межд. конф., 3–6 июня 2003, СПб, с. 13 – 17.
110. Клир Дж. Системология. М.: Радио и связь, 1990. – 544 с.
111. Никитин В. В., Цыцулин А. К. Телевидение в системах физической защиты. СПб, СПбГЭТУ «ЛЭТИ», 2001. – 135 с.
112. Цифровая обработка телевизионных и компьютерных изображений: Под ред. Зубарева Ю. Б. и Дворковича В. П., 2-е изд., М.: НАТ, 1997.
113. Биллингсли Ф. Влияние шума аппаратуры цифровой обработки изображений // Обработка изображений и цифровая фильтрация/ Под ред. Т. Хуанга. М.: Мир, 1979. – 320 с.
114. Цыцулин А. К., Березин В. В. Малокадровое телевидение: от рождения до наших дней//Факультету радиотехники и телекоммуникаций 60 лет/ Под ред. Н. В. Лысенко. СПб.: СПбГЭТУ «ЛЭТИ», 2005. С. 178–191.
115. Jain A.K. Fundamentals of image processing, Englewood Cliffs, New Jersey, Prentice Hall,1989, pp. 150–153.
116. СБИС для распознавания образов и обработки изображений: Пер. с англ./ Под ред. К.Фу. М.: Мир, 1988. – 248 с.

117. Зубарев Ю. Б., Кривошеев М. И., Красносельский И. Н. Цифровое телевизионное вещание. – М.: НИИ Радио, 2001. – 568 с.
118. Прэтт У. Цифровая обработка изображений.– М.: Мир, 1982, кн. 1. – 312 с.
119. Ахмед Н., Рао К. Р. Ортогональные преобразования при обработке цифровых сигналов. – М.: Связь, 1980. – 248 с.
120. Реализация аналогово-цифрового преобразования с переменным шагом квантования./ Березин В.Б., Наумов Е. В., Березин В.В. и др.//Известия высших учебных заведений, серия Радиоэлектроника, 2004, №3, с.57–65.
121. Трёхканальная телевизионно-компьютерная система анализа волнового фронта/Войцехович В. В., Цыцулин А. К., Березин В. В., и др. // Известия ВУЗов. Приборостроение. 2005, №9, с. 29–34.
122. Телевизионная съёмка кометы Галлея / Аванесов Г. А., Зиман Я. Л., Тарнопольский В. И. и др. М.: Наука, 1989. – 295 с.
123. Шамис А. Л. Поведение, восприятие, мышление: проблемы создания искусственного интеллекта. М.: Едиториал УРСС, 2005. – 224 с.
124. Быстродействующая телевизионная измерительная система для оценки искажений волнового фронта методом Гартмана / Войцехович В. В., Анкудинов С. Н., Манцветов А. А. и др. // Оптический журнал, 2000, № 2, с. 113–119.
125. Форсайт Д. А., Понс Ж. Компьютерное зрение. Современный подход. М.: Вильямс, 2004. – 928 с.
126. Трауб Дж., Васильковский Г., Вожьняковский Х. Информация, неопределённость, сложность. М.: Мир, 1988. – 184 с.
127. Рабинер Л., Гоулд Б. Теория и применение цифровой обработки сигналов. М., Мир, 1978. – 848 с.
128. Соколов А. В., Манцветов. А. А., Цыцулин А. К., Умников Д. В. Телевизионное измерение координат специально формируемых оптических сигналов// Вопросы радиоэлектроники, сер. Техника телевидения, 2006, вып. 2, с. 90 – 95.
129. Кондратьев К. Я., Федченко П. П. Особенности методики спектрометрических измерений природных объектов // Исследование Земли из Космоса, 1993, № 1, с. 15–24.
130. Здор С. Е., Колинько В. И., Суриков И. Н. Тенденции развития современных телевизионных систем наблюдения космических объектов//XIII НТК «Современные проблемы развития телевизи-онных фотоэлектронных приборов», СПб, 2006, с. 43, 44.
131. Достижения в технике передачи и воспроизведения изображений. Т. 3/ Под ред. Б. Кейзана. М. Мир, 1980. – 312 с.
132. Ермаков О. Н. Прикладная оптоэлектроника. М. Техносфера, 2004. – 416 с.

133. Котов Б. А. Матричные преобразователи оптической информации в видимой области спектра//Фотоприёмники и фотопребразователи. Сб. научн. трудов под ред. Ж. И. Алфёрова и Ю. В. Шмарцева. Л.: Наука, 1986, с. 131–147.
134. Полупроводниковые формирователи сигналов изображений/ Под ред. П. Йесперса, Ф. Ван де Виле и М. Уайта. М.: Мир, 1979. – 574 с.
135. www.foveon.com.
136. Эклз М., Сим Э., Тритон К. Детекторы слабого излучения в астрономии. М., Мир, 1986. – 200 с.
137. Акушский И. Я., Юдицкий Д. И. Машинная арифметика в остаточных классах. – М.: Сов. Радио, 1968. – 440 с.
138. Смелков В. М. Метод зарядовой фильтрации изображений охраняемых объектов в обнаружителе движения на ПЗС для уменьшения ложных тревог//Специальная техника, 2002, №4, с.12–16.
139. Кощавцев Н. Ф., Федотова С. Ф. Состояние и перспективы развития техники ночного видения//Прикладная физика, 1999, вып. 2, с. 141–145.
140. Ерофейчев В. Г. Инфракрасные фокальные матрицы //Оптичес-кий журнал, 1995, № 2, с. 12–20.
141. Ушакова М. Б. Тепловизоры на основе неохлаждаемых микроболо-метрических матриц: современное состояние зарубежного рынка и перспективы развития. ОНТИ ГУП “НПО “Орион”, М., 2001.
142. . Pengelley R., Hewish M. In the heat of the night. Jane’s International Review. 2001, Vol. 34, No. 10, pp. 49 – 57.
143. Сверхбольшие интегральные схемы и современная обработка сигналов/ Под ред. С. Гуна, Х. Уайтхауса, Т. Кайлата. М., Радио и связь, 1989. – 472 с.
144. Волков. В. Г., Ковалёв А. В., Федчишин В. Г. Тепловизионные приборы нового поколения.//Специальная техника, 2001, №6, с. 16 – 21.
145. Тарасов В. В., Якушенков Ю. Г. Инфракрасные системы «смотрящего типа». М.: Логос, 2004, – 444 с.
146. Wijewarnasuriya P.S., at al., HgCdTe detectors and FPAs for remote sensing application/ Proceedings of the SPIE, vol. 3698, FL, April 1999, pp. 538–544.
147. Пономаренко В. П., Филачёв А. М. Инфракрасная техника и электронная оптика. Становление научных направлений (1946 – 2006). М.: Физматкнига, 2006. – 334 с.
148. Филачёв А. М., Таубкин И. И., Тришенков М. А. Твердотельная фотоэлектроника. Физические основы. М.: Физматкнига, 2005. – 384 с.
149. Соломоник А. Позитивная семиотика. Мн.: МЕТ, 2004. – 191 с.
150. Колмогоров А. Н. Теория информации и теория алгоритмов. М.: Наука, 1987. – 304 с.

151. Успенский В. А., Семёнов А. Л. Теория алгоритмов: основные открытия и приложения. М.: Наука, 1987. – 288 с.
152. Зубакин И. А., Цыцулин А. К. Влияние вида ограничения на линейное кодирование// Вопросы радиоэлектроники, сер. Техника телевидения, 2006, вып. 1, с. 79–87.
153. Зубакин И. А., Цыцулин А. К. Моделирование влияния ограничения сложности кодера на качество кодирования изображений с преобразованием// Вопросы радиоэлектроники, сер. Техника телевидения, 2006, вып. 2, с. 32–40.
154. 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. – 240 с.
155. Дьяконов В. П. Закон Мура и компьютерная математика// Exponenta Pro. Математика в приложениях. 2003, №1, с. 82–86.
156. Фахми Ш. С. Автоматизация проектирования БИС на базе «система на кристалле». СПб.: СПбГЭТУ «ЛЭТИ», 2006. – 80 с.
157. Томас Ф., Иванов А. САПР микроэлектроники. Этапы большого пути. // Электроника: наука, технология, бизнес. 2006, № 3, с. 82–85.
158. Борович. А. Технологии глобального синтеза компании Cadance// Электроника: наука, технология, бизнес. 2006, № 5, с. 102–105.
159. Казанцев Г. Д., Курячий М. И., Пустынский И. Н. Измерительное телевидение. М., Высшая школа, 1994. – 288 с.
160. Даджион д., Мерсеро Р. Цифровая обработка многомерных сигналов. М.: Мир, 1988. – 275с.
161. Амербаев В. М. Теоретические основы машинной арифметики. – Алма-Ата: Наука, 1976. – 320 с.
162. Нейрокомпьютеры в остаточных классах. Кн. 11 / Червяков Н.И., Сахнюк П.А., Шапошников А.В., Макоха А.Н.: Учебное пособие для вузов (Научная серия «Нейрокомпьтеры и их применение», ред. А.И. Галушкин). – М.: Радиотехника, 2003. – 272 с.
163. Микропроцессорные системы: Учебн. пособие для вузов / Е. К. Александров, Р. И. Грушвицкий, М. С. Куприянов и др. Под общ. ред. Д. В. Пузанкова. – СПб.: Политехника, 2002. – 935 с.
164. Казеннов Г. Г. Основы проектирования интегральных схем и систем/ М.: БИНОМ. Лаборатория знаний, 2005.
165. Бухтев А. Среда проектирования компании Cadence/ CHIP NEWS, 2003, №4.
166. Бауэр Г. Э., Клюшкин В. И., Александров А. В. Разработка специализированной СБИС типа «система на кристалле» COFDM демодулятора телевизионного сигнала наземного цифрового телевидения//НТК «Микроэлектронные системы», М., 12 окт. 2006.

167. http://www.it-life.ru/news/2005/09/05/2118
168. http://www.sonyhouse.info/tovar-5069569275-92
169. www.electronicstalk.com
170. http://www.cctvfocus.ru
171. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео/Ватолин Д., Ратушняк А., Смирнов М., Юкин В. – М.: ДИАЛОГ-МИФИ, 2002. – 384 с.
172. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2005.
173. Александров В. В., Горский Н. Д. Представление и обработка изображений: Рекурсивный подход. Л.: Наука, 1985.
174. Шеннон К. Работы по теории информации и кибернетике. М.: 1963. – 832 с.
175. Фахми Ш. С. Аналитическая модель оценки эффективности кодирования информации по опорным точкам// Изв. СПбГЭТУ «ЛЭТИ» 2006, вып.2, с. 71–76.
176. Витерби А. Д., Омура Дж. К. Принципы цифровой связи и кодирования. М., Радио и связь, 1982. – 536 с.
177. Коляда А. А., Пак И. Т. Модулярные структуры конвейерной обработки цифровой информации. – Минск: Университетское, 1992. – 256 с.
178. Амербаев В. М., Пак И. Т. Параллельные вычисления в комплексной плоскости. – Алма-Ата: Наука, 1984. – 182 с.
179. Коляда А. А., Ревинский В. В., Селянинов М. Ю., Чернявский А. Ф. Применение минимально избыточного модулярного кодирования для быстрого умножения комплексных чисел в системах цифровой обработки сигналов // Весці Акадэміі навук Беларусі. Сер. фіз.-тэхн. Навук. – 1996. – № 1.
180. Инютин С. А. Компьютерная модулярная алгебра квадратичного диапазона и область её приложения // Вестник Тюменского госуниверситета. – Тюмень, 2001, № 2. – с. 141–148.
181. Торгашов В. А. Система остаточных классов и надежность ЦВМ. М.: Сов. Радио, 1973. – 118 с.
182. Архитектура адаптивной параллельно-конвейерной нейронной сети для коррекции ошибок в модулярных нейрокомпьютерных системах / Червяков Н. И., Галкина В. А, Стрекалов Ю. А., Лавриненко С.В. // Нейрокомпьютеры: разработка, применение. – 2003, № 6, с. 47–60.
183. Построение систем повышенной надёжности на основе аппарата модулярной арифметики с применением современных методов и средств проектирования. А. Л. Стемпковский, А. И. Корнилов,

М. Ю. Семенов, О. В. Ласточкин, В. С. Калашников // Международная НТК «50 лет модулярной арифметике». – 2005, 15 с., www.computer-museum.ru.
184. Модулярные вычислительные структуры: вчера, сегодня, завтра. А. А. Коляда, А. Ф. Чернявский // Международная НТК «50 лет модулярной арифметике». – 2005. – 12 с., www.computer-museum.ru.
185. Халиков Э. М., Шаменков Н. А., Шипилов Н. Н. О применении арифметики остаточных классов для аппаратной компрессии/декомпрессии телевизионных изображений//4-я международная НТК «Телевидение: передача и обработка изображений», СПб, 2005, с. 48–49.
186. Алдохина В. Н., Малтыков Д. В., Неверов В. В., Савченко О. В., Умбиталиев А. А., Халиков Э. М., Шипилов Н. Н. Способы построения высокоскоростных модулярных видеокодеков// 7-я межведомственная НТК «Проблемные вопросы сбора, обработки, передачи и защиты информации в сложных радиотехнических системах». – Пушкин, 2005, с. 30–36.
187. Липский В. Комбинаторика для программистов. – М.: Мир, 1988. – 200 с.
188. Червяков Н. И., Сахнюк П. А., Шапошников А. В. Пути эффективного использования иерархических структур фрагментов нейронных сетей на ПЛИС Xilinx при аппаратной реализации цифровых фильтров с параллельной обработкой данных // Нейрокомпьютеры: разработка, применение. – 2001, № 10, с. 26–33.
189. Червяков Н. И., Сахнюк П. А., Шапошников А. В., Макоха А. Н. Архитектура нейрокомпьютеров, функционирующих в системе остаточных классов // Нейрокомпьютеры: разработка, применение. – 2003, № 5, с. 17–28.
190. Головина Л. И. Линейная алгебра и некоторые ее приложения. – М.: Наука, 1979. – 392 с.
191. Полисский Ю. Д. Сравнение чисел в системе остаточных классов. – 2005. – 6 с., www.computer-museum.ru.
192. Кузелин М. О., Кнышев Д. А., Зотов В. Ю. Современные семейства ПЛИС фирмы Xilinx. Справочное пособие. – М.: Горячая линия – Телеком, 2004. – 440 с.
193. Постников М. М. Теория Галуа. – М.: Факториал Пресс, 2003. – 304 с.
194. Стешенко В. Б. ПЛИС фирмы «Altera»: элементная база, система проектирования и языки описания аппаратуры. – М.: Издательский дом «Додэка-XXI», 2002. – 576 с.

СОДЕРЖАНИЕ

Предисловие…….. 3
От авторов…….. 5
Введение…… 7
В.1. Научно-прикладное телевидение как отрасль информатики 7
В.2. Модели систем и сигналов в прикладном телевидении….. 10
В.3. Этапы твердотельной революции в телевидении…….. 14

Глава 1. **Твердотельные телевизионные камеры**…… 32

1.1. Телекамеры на ПЗС…… 34
1.2. Телекамеры на КМОП…… 59
1.3. Адаптация в твердотельных телекамерах…… 88
1.4. Твердотельные телекамеры ИК диапазона……... 92

Глава 2. **Дискретно-аналоговая обработка сигнала изображения** 105

2.1. Обнаружение и оценивание координат объектов…….. 107
2.2. Переменная четкость в твердотельных телекамерах…….. 139
2.3. Вычисление проекций изображений……... 143

Глава 3. **Проектирование телевизионных систем на основе теории информации**…….. 150

3.1. Применение математической теории связи в телевидении 154
3.2. Квантование сигналов в условиях фотонного шума…….. 167
3.3. Адаптивный групповой счет фотонов …… 173
3.4. Кодирование источника: сжатие видеоинформации ……. 177
3.5. Модулярные вычислительные структуры обработки видеоинформации……..……..180

Глава 4. **Обработка телевизионных сигналов в системах на кристалле и видеосистемах на кристалле** 206

4.1. Обработка видеоинформации в эпоху систем на кристалле 208
4.2. Особенности проектирования телевизионных систем с использованием систем на кристалле…… 218
4.3. САПР телевизионно-компьютерных систем с использованием систем на кристалле и видеосистем на кристалле……... 243
4.4. Сжатие сигналов изображений в системах на кристалле.. 267
4.5. Системы на кристалле в мобильной телевизионной связи 287

Заключение…… 295
Литература…….…… 299

Научное издание

БЕРЕЗИН ВИКТОР ВЛАДИМИРОВИЧ
УМБИТАЛИЕВ АЛЕКСАНДР АХАТОВИЧ
ФАХМИ ШАКИБ СУБХИЕВИЧ
ЦЫЦУЛИН АЛЕКСАНДР КОНСТАНТИНОВИЧ
ШИПИЛОВ НИКОЛАЙ НИКОЛАЕВИЧ

ТВЕРДОТЕЛЬНАЯ РЕВОЛЮЦИЯ В ТЕЛЕВИДЕНИИ. ТЕЛЕВИЗИОННЫЕ СИСТЕМЫ НА ОСНОВЕ ПРИБОРОВ С ЗАРЯДОВОЙ СВЯЗЬЮ, СИСТЕМ НА КРИСТАЛЛЕ И ВИДЕОСИСТЕМ НА КРИСТАЛЛЕ

ИБ № 3202

Подписано в печать 30.11.06 Формат 84×108/32 Бумага тип. №2 Печать высокая.
Усл. печ. л. 19,0. Тираж 600 экз. Заказ № _____
Издательство «Радио и связь». 101000, Москва, Почтамт, а/я 693
Издательская лицензия №010164 от 29.01.97

Санкт-Петербургская типография «Триумф». Санкт-Петербург, ул. Чугунная, 14.