

Министерство образования Республики Беларусь
Учреждение образования
«Белорусский государственный университет
информатики и радиоэлектроники»

Факультет телекоммуникаций

Кафедра систем телекоммуникаций

А. П. Ткаченко, А. Л. Хоминич

ЦИФРОВОЕ ТЕЛЕВИДЕНИЕ

В двух частях

Часть 1

КОДИРОВАНИЕ ИСТОЧНИКА СООБЩЕНИЙ В СИСТЕМАХ ЦИФРОВОГО ТЕЛЕВИЗИОННОГО ВЕЩАНИЯ

*Рекомендовано УМО по образованию
в области информатики и радиоэлектроники
в качестве учебно-методического пособия
для направлений специальности 1-45 01 01-01 «Инфокоммуникационные
технологии (системы телекоммуникаций)»,
1-45 01 01-04 «Инфокоммуникационные технологии
(цифровое теле- и радиовещание)»*

Минск БГУИР 2015

УДК 621.397.132(075.8)

ББК 32.943я73

Т48

Рецензенты:

кафедра последиplomного образования учреждения образования

«Высший государственный колледж связи»

(протокол №2 от 03.10.2014);

заведующий лабораторией компьютерной графики

государственного научного учреждения

«Объединенный институт проблем информатики

Национальной академии наук Беларуси»,

кандидат технических наук, доцент В. В. Ткаченко

Ткаченко, А. П.

Т48

Цифровое телевидение. В 2 ч. Ч. 1 : Кодирование источника сообщений в системах цифрового телевизионного вещания : учеб.-метод пособие / А. П. Ткаченко, А. Л. Хоминич. – Минск : БГУИР, 2015. – 162 с. : ил.

ISBN 978-985-488-681-7 (ч. 1).

Посвящено анализу методов кодирования источника сообщений в системах цифрового телевизионного вещания (ЦТВ).

Рассматриваются и математически обосновываются параметры компонентных сигналов ЦТВ стандартной и высокой четкости. Приводится детальная классификация систем ЦТВ, анализируются их преимущества, поясняются принципы формирования программных и транспортного потоков.

Дается классификация основных методов сокращения избыточности изображений и их математическое обоснование. Приведены пути повышения эффективности кодирования и рекомендации по их практическому использованию.

Предназначено студентам, изучающим учебные дисциплины «Телевизионные системы» и «Цифровое телевидение».

УДК 621.397.132(075.8)

ББК 32.943я73

ISBN 978-985-488-681-7 (ч. 1)

ISBN 978-985-488-680-0

© Ткаченко А. П., Хоминич А. Л., 2015

© УО «Белорусский государственный университет информатики и радиоэлектроники», 2015

СОДЕРЖАНИЕ

СПИСОК СОКРАЩЕНИЙ	4
ВВЕДЕНИЕ	9
1 ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ЦИФРОВОГО ПРЕДСТАВЛЕНИЯ СИГНАЛОВ ТЕЛЕВИЗИОННЫХ ИЗОБРАЖЕНИЙ.....	13
1.1 Обобщенная структурная схема цифровой системы передачи телевизионных изображений	13
1.2 Классификация систем цифрового телевизионного вещания	21
1.3 Преимущества систем ЦТВ	45
1.4 Теоретическое обоснование параметров исходных телевизионных сигналов для их цифрового представления	48
1.5 Цифровое представление компонентных сигналов яркости и цветоразностных.....	53
1.6 Параметры цифровых компонентных сигналов.....	59
2 МЕТОДЫ СОКРАЩЕНИЯ ИЗБЫТОЧНОСТИ ИЗОБРАЖЕНИЙ.....	73
2.1 Классификация методов сокращения избыточности.....	73
2.1.1 Обоснование необходимости компрессии изображений	73
2.1.2 Избыточность изображений	75
2.1.3 Классификация методов сокращения избыточности изображений.....	80
2.1.4 Критерии оценки качества отображения	84
2.2 Статистическое кодирование	86
2.2.1 Оценка избыточности изображений	86
2.2.2 Код Хаффмана	89
2.2.3 Арифметическое кодирование	94
2.2.4 Кодирование длин серий	96
2.2.5 LZW-кодирование	97
2.3 Кодирование с предсказанием	100
2.3.1 Классификация методов кодирования с предсказанием	100
2.3.2 Общие принципы кодирования с предсказанием	101
2.3.3 Предсказание с компенсацией движения	117
2.4 Кодирование с преобразованием	123
2.4.1 Общие принципы кодирования с преобразованием	123
2.4.2 Математические основы ортогональных преобразований	128
2.4.3 Дискретное преобразование Фурье	131
2.4.4 Дискретное-косинусное преобразование.....	134
2.4.5 Преобразование Уолша – Адамара.....	139
2.4.6 Вейвлетное преобразование	143
2.5 Гибридное кодирование.....	145
ЗАКЛЮЧЕНИЕ.....	152
ПРИЛОЖЕНИЕ А	153
ПРИЛОЖЕНИЕ Б	155
ЛИТЕРАТУРА	157

СПИСОК СОКРАЩЕНИЙ

АБГШ	–	аддитивный белый гауссовский шум
АМ	–	амплитудная модуляция
АСБ	–	аппаратно-студийный блок
АЧХ	–	амплитудно-частотная характеристика
АЦП	–	аналого-цифровое преобразование, аналого-цифровой преобразователь
ВОЛС	–	волоконно-оптическая линия связи
ГВЗ	–	групповое время задержки
ДИКМ	–	дифференциальная импульсно-кодовая модуляция
ДКП	–	дискретно-косинусное преобразование
ДФФ	–	дискретное преобразование Фурье
ДПУА	–	дискретное преобразование Уолша – Адамара
ДМВ	–	дециметровые волны
ИКМ	–	импульсно-кодовая модуляция
КАМ	–	квадратурная амплитудная модуляция
КАС	–	конец активной строки (в цифровом сигнале)
КМОП	–	комплементарная структура металл–оксид–полупроводник
КЦТВ	–	кабельное цифровое телевизионное вещание
МВ	–	метровые волны
МСЭ	–	Международный союз электросвязи
МСЭ-Р	–	сектор радиосвязи МСЭ
МСЭ-Т	–	сектор стандартизации электросвязи МСЭ
МЦТВ	–	мобильное цифровое телевизионное вещание
НАС	–	начало активной строки (в цифровом сигнале)
НТВ	–	непосредственное ТВ вещание (спутниковое)
НЦТВ	–	наземное цифровое телевизионное вещание
ОЭП	–	оптико-электронный преобразователь
ПВФ	–	пространственно-временная функция
ПЗС	–	прибор с зарядовой связью
ПолРК	–	поляризационное разделение каналов
ПрРК	–	пространственное разделение каналов
РАВИС	–	аудиовизуальная информационная система реального времени
РТПС	–	радиотелевизионная передающая станция
РЧР	–	радиочастотный ресурс
СВЧ	–	сверхвысокие частоты
СКТ	–	система кабельного телевидения
СП	–	сингулярное (S-) преобразование
СЦ	–	сигнал цветности
СЦТВ	–	спутниковое цифровое телевизионное вещание
СЯ	–	сигнал яркости
ТВ	–	телевидение, телевизионный
ТВК	–	телевизионный канал

ТВЧ	– телевидение высокой четкости
ТПК	– телевидение повышенного качества (повышенной четкости)
ТПР	– телевидение пониженного разрешения (пониженной четкости)
ТСЧ	– телевидение стандартной четкости
ТСВЧ	– телевидение сверхвысокой четкости
ТУВЧ	– телевидение ультравысокой четкости
ФНЧ	– фильтр нижних частот
ЦА	– центральная аппаратная (телецентра)
ЦАП	– цифроаналоговое преобразование, цифроаналоговый преобразователь
ЦРС	– цветоразностный сигнал
ЦСП	– цифровая система передачи
ЦТВ	– цифровое телевидение, цифровое ТВ-вещание (компонентное)
ЦТП	– цифровая телевизионная программа
ЧРК	– частотное разделение каналов
ЭОП	– электронно-оптический преобразователь
AES	– Audio Engineering Society – Общество инженеров звукотехники, США
ATSC	– Advanced Television Systems Committee – Комитет улучшенных (перспективных) телевизионных систем, по имени которого названа принятая в США система цифрового телевидения
AVC	– Advanced Video Coding – расширенное (улучшенное) видеокодирование
BCH	– Bose-Choudhuri-Nocquenghem – Боуза-Чоудхури-Хоквингема (помехоустойчивый код)
BER	– Bit Error Rate – вероятность ошибки на бит
CCD	– Charge Coupled Device – см. ПЗС
CIF	– Common Image Format – единый формат изображения (международный формат изображения с разрешением 1920×1080 пикселей) – Common Interchange Format – единый формат обмена (формат изображения с разрешением 352×288 пикселей и частотой кадров до 30 Гц, используемый в видео-конференц-связи)
СММВ	– China Multimedia Mobile Broadcasting – система мобильного мультимедийного вещания КНР
СМОС	– Complementary Metal-Oxide Semiconductor – см. КМОП
COFDM	– Coded Orthogonal Frequency Division Multiplexing – ортогональное частотное разделение каналов (модулированных несущих) с кодированием
CWDM	– Coarse Wavelength Division Multiplexing – разреженное мультиплексирование с разделением по длинам волн
DMB	– Digital Multimedia Broadcasting – цифровое мультимедийное вещание, система цифрового телевизионного вещания

- DOCSIS – Data Over Cable Service Interface Specifications – стандарт (интерфейс) передачи данных по кабелю (обратный канал в СКТ)
- DVB-T, C, H, S, SH – Digital Video Broadcasting – цифровое телевизионное вещание, Terrestrial – наземное (НЦТВ), Cable – кабельное (КЦТВ), Handheld – мобильное (с приемом на портативные устройства) (МЦТВ), Satellite – спутниковое (СЦТВ), Satellite Handheld – спутниковое мобильное
- DVB-T2, C2, S2, – цифровое телевизионное вещание – НЦТВ, КЦТВ и СЦТВ второго поколения
- DVB-H2 – МЦТВ второго поколения, обозначаемое также DVB-NGH – Next Generation Handheld – новое поколение МЦТВ
- DVB-3DTV – DVB-3Dimensional Television – DVB стандарт цифрового трехмерного телевидения
- DWDM – Dense Wavelength Division Multiplexing – плотное мультиплексирование с разделением по длинам волн
- EAV – End of Active Video – см. КАС
- EBU – European Broadcasting Union – Европейский вещательный союз
- EDTV – Extended Definition Television – см. ТПК
- ETSI – European Telecommunications Standards Institute – Европейский институт телекоммуникационных стандартов
- FEC – Forward Error Correction – прямое исправление ошибок
- FFT – Fast Fourier Transform – быстрое преобразование Фурье
- FTTx – Fiber To The x – волокно до «х», общее обозначение абонентского участка сети, в которой от узла связи волоконно-оптический кабель доходит до определенного места – точки X, а далее, до абонента – медный кабель (в FTTH – волокно); FTTB – FTT Building – волокно до здания; FTTC – FTT Curb – волокно до микрорайона, квартала или группы домов; FTTH – FTT Home – волокно до жилища (квартиры или отдельного коттеджа); FTTN – FTT Node – волокно до сетевого узла
- HDTV – High Definition Television – см. ТВЧ
- HD-SDI – High Definition Serial Digital Interface – последовательный цифровой интерфейс для передачи сигналов телевидения высокой четкости, имеет также разновидности 3G-, 6G- и 12G-SDI, отличающиеся допустимой скоростью передачи данных
- HEVC – High Efficiency Video Coding – высокоэффективное видеокодирование
- IEC – International Electrotechnical Commission – Международная электротехническая комиссия, МЭК
- IPTV – Internet Protocol for Television – телевидение по протоколу Интернет: услуги предоставляются по выделенной и управляемой оператором сети с гарантированным качеством обслуживания QoS

- ISDB – Integrated Services Digital Broadcasting – цифровое вещание с интеграцией служб, принятая в Японии система ЦТВ
- ISO – International Standard Organization – Международная организация по стандартизации, ИСО
- ITU-R,T – International Telecommunication Union – Международный союз электросвязи (МСЭ, R – сектор радиосвязи, T – сектор стандартизации электросвязи)
- JPEG – Joint Picture Expert Group – объединенная группа экспертов по изображениям, наименование группы стандартов компрессии статических изображений
- LDPC – Low Density Parity Check – низкая плотность проверок на четность (помехоустойчивый код)
- LDTV – Low Definition Television – см. ТПР
- LMDS – Local Multipoint Distribution System – локальная многоточечная распределительная система (в диапазоне 22...29 ГГц)
- MAC – Multiplexed Analog Components – мультиплексирование (временное) аналоговых компонент, система телевидения повышенного качества и HD-MAC – высокой четкости
- MediaFlo – Media Forward Link Only – только прямое медиа-соединение, система мобильного вещания США
- MFN – Multi Frequency Network – многочастотная сеть
- MMDS – Microwave Multipoint Distribution System – многоточечная СВЧ распределительная система (в диапазоне ниже 10 ГГц, обычно 2,5...2,7 ГГц)
- MPEG – Motion Pictures Expert Group – группа экспертов по движущимся изображениям, наименование группы стандартов компрессии динамических изображений
- Multicast – многоадресная передача: «точка – многоточка»
- MVDS – Multipoint Video Distribution Systems – многоточечная видеораспределительная система (в диапазоне 10 ГГц и выше)
- NTSC – National Television Systems Committee – национальный комитет ТВ систем, США, по имени которого названа система цветного телевидения (аналоговая композитная)
- OTT – Over the Top Internet Television – телевидение (видеоконтент) через Интернет: услуги предоставляются по открытой, неуправляемой сети Интернет
- PAL – Phase Alternation Line – «изменение фазы от строки к строке», наименование системы цветного телевидения (аналоговой композитной)
- PON – Passive Optical Network – пассивная оптическая сеть
- PSF – Progressive Segment Frame – сегментный кадр при прогрессивной развертке
- QAM – Quadrature Amplitude Modulation – см. КАМ

- QEF – Quasi Error Free – практически без ошибок (в ЦТВ квазибезошибочный прием обеспечивается при одной ошибке в час, что соответствует вероятности ошибок 10^{-11})
- QoS – Quality of Service – качество обслуживания
- QPSK – Quadrature Phase-Shift Keying – квадратурная фазовая модуляция
- RAVIS – Real-time Audio-Visual Information System – см. РАВИС
- SAV – Start of Active Video – см. НАС
- SDI – Serial Digital Interface – последовательный цифровой интерфейс для передачи сигналов телевидения стандартной четкости
- SDTV – Standard Definition Television – см. ТСЧ
- SECAM – Sequentiel Couleur avec Memoire – «последовательная передача цветов с запоминанием», наименование системы цветного телевидения (аналоговой композитной)
- SFN – Single Frequency Network – одночастотная сеть
- SMPTE – Society of Motion Pictures and Television Engineers – Общество инженеров кино и телевидения, США
- SPI – Synchronous Parallel Interface – синхронный параллельный интерфейс
- TCP/IP – Transport Control Protocol / Internet Protocol – протокол управления передачей / протокол Интернета
- TRS – Timing Reference Signals – опорный сигнал синхронизации
- UHDTV – Ultra High Definition Television – см. ТСВЧ, ТУВЧ
- Unicast – одноадресная передача: «точка – точка»
- VC – Video Coding – видеокодирование, общее обозначение семейства стандартов компрессии изображений (VC-1...VC-3), принятых SMPTE
- VoD – Video on Demand – видео по запросу, требованию
- Wi-Fi – Wireless Fidelity – беспроводные Hi-Fi (High-Fidelity – высокая верность, точность) устройства, системы
- WiMAX – Wireless MAN (Metropolitan Area Network) Exchange – беспроводная городская (мегаполиса) сеть обмена сообщениями
- xDSL – (x – префикс) Digital Subscriber Line – семейство стандартов цифровой абонентской линии: более 20 асимметричных и симметричных стандартов – от ADSL до VDSL

ВВЕДЕНИЕ

Современное развитие электронных средств передачи информации характеризуется глубоким проникновением цифровых методов в процессы формирования, обработки и передачи (записи) сигналов с одной стороны, а с другой – их интеграцией и конвергенцией с компьютерными технологиями. Это позволяет многие операции над цифровыми сигналами осуществлять программно.

Теле- и радиовещание не являются исключением – во всем мире происходит процесс перехода к цифровым методам, который в одних городах и странах уже завершен, в других идет бурная цифровизация с конечной целью – завершить её в период до 2015 г. По-видимому, в этот срок не все уложатся, но переход неизбежен, поскольку цифровое телевизионное (ТВ) вещание (ЦТВ) по сравнению с аналоговым имеет существенные преимущества, подробно анализируемые в данном учебно-методическом пособии.

Главное преимущество заключается в том, что усилиями ученых и инженеров многих стран удалось не только сохранить радиочастотный ресурс (РЧР) – уже сложившееся распределение радиочастотных ТВ-каналов, но и «вписать» в полосу частот одного канала несколько цифровых ТВ-программ (ЦТП) стандартной и даже высокой четкости [1 – 5].

Совсем недавно концепция «6+7+8 = XXI век», предложенная М. И. Кривошеевым (с 1970 по 2000 год Председатель исследовательской ТВ-комиссии сектора радиосвязи Международного союза электросвязи – МСЭ-Р), казалась фантазией и даже утопией. Действительно, для передачи цифровых ТВ-сигналов требуются каналы с пропускной способностью примерно от 270 Мбит/с до 3 Гбит/с. Существующие же полосы частот наземных каналов вещательного и кабельного телевидения 6, 7 или 8 МГц (в зависимости от принятого в стране стандарта) на такие скорости не рассчитаны. Но, благодаря разработке высокоэффективных алгоритмов сжатия (сокращения избыточности), методов помехоустойчивого кодирования, использованию многопозиционных видов цифровой модуляции, успехам в микроэлектронике и программном обеспечении, сложная задача была решена.

Теория преобразования компонентных ТВ сигналов (яркости и цветоразностных) в цифровой вид, а также сокращения в них избыточности и рассматривается в пособии. Эта первая операция получила название *кодирования источника*. Вместе с еще двумя – кодированием канала (помехоустойчивым кодированием) и модуляцией они составляют процесс формирования радиосигнала ЦТВ. Последние будут рассматриваться во второй части учебно-методического пособия.

Необходимо отметить, что вещание по аналоговым композитным системам цветного телевидения NTSC, PAL и SECAM сыграло большую роль в накоплении теоретического и практического опыта работы со сложными

сигналами и системами. Однако совместная передача сигнала яркости и сигнала цветности в общей полосе частот путем частотного уплотнения их дискретных спектров (так обеспечивается совместимость с черно-белым телевидением) явилась причиной многих видов искажений. Последние отсутствуют в цифровых ТВ системах, использующих временное уплотнение цифровых компонентных сигналов яркости и двух цветоразностных. Вместе с ними уплотняются цифровые стереофонические сигналы звукового сопровождения, звукового вещания и дополнительные данные.

Европейские страны и большинство других стран для ЦТВ приняли семейство стандартов DVB (Digital Video Broadcasting – дословно – цифровое видеовещание), которое только по виду среды распространения сигналов подразделяется на наземное (Terrestrial), спутниковое (Satellite) и кабельное (Cable) – DVB-T, DVB-S и DVB-C (НЦТВ, СЦТВ и КЦТВ).

В наиболее промышленно развитых европейских странах внедрение DVB начиналось с 1994 – 1995 гг. по системе телевидения стандартной четкости ТСЧ (SDTV – Standard Definition Television), условно обозначаемой 625/50/2:1/4:3, поскольку телевидение высокой четкости – ТВЧ (HDTV – High Definition Television) считалось роскошью. Только США начали ЦТВ по американской системе ATSC (Advanced Television Systems Committee – Комитет улучшенных (перспективных) телевизионных систем) и по ТСЧ, и по ТВЧ (рассмотрено в п. 1.2) и в середине 2009 г. прекратили вещание по аналоговой системе NTSC.

По прошествии нескольких лет после начала тестового вещания стала вырисовываться существенная и отличительная особенность ЦТВ по сравнению с аналоговыми системами, с момента принятия которых прошло более 60 (NTSC) и 45 (PAL, SECAM) лет. За столь продолжительный период совершенствовалась элементная база, внедрялись новые методы обработки сигналов, но параметры и характеристики аналоговых систем, утвержденные соответствующими стандартами, практически не изменялись [6 – 10].

Напротив, цифровые системы передачи и записи изображений все время развиваются и совершенствуются, расширяются их функциональные возможности. Приведем только основные этапы развития систем DVB по состоянию на начало 2015 г. Появление источников сообщений – ТВ передающих камер с отношением сигнал/шум порядка 60 дБ для ТВЧ и более 65 дБ – для ТСЧ, явилось причиной использования 12- и 16-разрядного аналого-цифрового преобразования (АЦП) сигналов основных цветов при их формировании и постепенного перехода от 8-разрядного представления к 10- и даже 12-разрядному при передаче компонентных (яркостного и цветоразностных) сигналов.

Необходимость передачи возрастающих объемов информации с высоким качеством в ограниченной полосе стандартного ТВ канала была побудительным мотивом для разработки методов сокращения избыточности с большей степенью компрессии (рассмотрено в разд. 2) – от MPEG-1 до MPEG-4

AVC/H.264 и H.265, а также способов помехоустойчивого кодирования канала и высокоэффективных видов цифровой модуляции [11 – 25, 47, 48].

Результаты не замедлили сказаться – разработаны и внедряются системы ЦТВ второго поколения DVB-T2, S2 и C2, появились системы мобильного приёма на портативные устройства с небольшими экранами – DVB-H и DVB-SH, созданы системы интерактивного ТВ, «видео по запросу» (Video on Demand), Интернет-вещание, значительно возросло количество программ, передаваемых по стандарту высокой четкости HDTV [26 – 28, 52, 63]. Для привлечения абонентов российская компания «Континент ТВ» предлагает всем, кто подключится к базовому пакету «Классический» на оборудовании Continent, смотреть бесплатно в течение 2012 г. 10 разнообразных по тематике каналов.

Следует отметить, что и в нашей стране эффективно внедряются наземное, кабельное, спутниковое, мобильное ЦТВ и вещание по протоколу Интернет – IPTV (Internet Protocol Television). Только по линии РУП «Белорусский радиотелевизионный передающий центр» уже установлено более 90 цифровых телевизионных передатчиков, покрывающих территорию, на которой проживает более 98 % населения. Каждый из передатчиков транслирует 8 ЦТП (цифровых ТВ программ) по стандарту ТСЧ. Активно внедряется ТВ вещание на коммерческой основе по системе DVB-T2, количество введенных в эксплуатацию передатчиков приближается к сотне.

Национальный оператор электросвязи РУП «Белтелеком» в 2008 г. ввел в коммерческую эксплуатацию в г. Минске интерактивную систему ТВ-видео по запросу (IPTV) по цифровой абонентской линии (xDSL), получившую в 2009 г. брэнд ZALA. С 15 ноября 2010 г. начато вещание пакета программ и по стандарту ТВЧ, а с мая 2011 г. услуга IPTV предоставляется региональными дата-центрами в областях в режиме unicast («точка – точка»), т. е. по индивидуальному запросу пользователей в удобное для них время. При этом трансляция ТВ каналов в режиме реального времени multicast («точка – многоточка») по-прежнему осуществляется с головного центра обработки данных. В целом по индексу информационно-коммуникационных технологий, принятому МСЭ, Республика Беларусь в 2014 г. впервые вошла в число 40 ведущих стран, опережая Россию, Украину, Казахстан, Польшу и др.

Казалось бы, достигнут предел мечтаний: ТВЧ на большом и плоском экране с возможностью индивидуального запроса программ в любое время, что может быть лучше? – только трехмерное изображение – 3D. Пока это стереоскопическое телевидение (не в полном смысле объемное), но вышедшее из стен научно-исследовательских лабораторий. Тестовое вещание программ стереоскопического телевидения ведется со спутников, например с Eutelsat EUROBIRD 9A (позиция 9° в.д.) или Astra (позиция 19,2° в.д.). ТВ программы в формате 3D HDTV также доступны у некоторых провайдеров потокового телевизионного вещания по технологиям OTT и IPTV, многие кинотеатры Европы перешли от тестовых к реально прямым демонстрациям различных событий.

Международная организация DVB Project, курирующая развитие и стандартизацию ЦТВ, в 2011 г. приняла окончательную версию спецификаций стандарта DVB-3DTV.

На международной выставке «Связь–Экспоком '2011» в Москве было продемонстрировано несколько новых оригинальных решений в области ЦТВ. Ученые из МНИТИ и томской компании Triaxes Vision разработали совместимую с имеющимися у населения аналоговыми и цифровыми телевизорами российскую систему цифрового трехмерного телевидения 2D/3D DVB-T. Просмотр программ трехмерного ТВ на обычных ТВ приемниках в виде плоского изображения обеспечивается тем, что стереоскопический ТВ сигнал 3D содержит две компоненты: сигнал изображения 2D и дополнительный сигнал глубины Depth. Такое решение позволяет передавать по сетям ЦТВ программы стереоскопического ТВ телевидения 2D/2D DVB-T и в то же время их принимать и отображать – как на обычных, так и на стереоскопических телевизорах 3D со встроенным декодером сигнала трехмерного ТВ (формат 2D+Depth).

В нашей стране с 1 июня 2011 г. начато тестовое стереоскопическое цветное 3D ТВ вещание, которое доступно подписчикам HD-пакета ZALA. Для их просмотра необходимо иметь телевизор с функцией 3D и активные 3D-очки.

Поскольку обычная киноплёнка имеет разрешение на порядок больше, чем ТСЧ, повышение качества изображения не ограничивается внедрением ТВЧ – уже созданы системы ТВ ультравысокой четкости (UHDTV) для видеоинформационных систем и цифрового кинематографа [19, 52, 54].

Это далеко не полный перечень новых технологий в ЦТВ. Они отражены в огромном количестве рекомендаций Международного союза электросвязи, стандартах Международной организации стандартизации (ISO), Международной электротехнической комиссии (IEC), технических отчетах и стандартах европейского института стандартов по телекоммуникациям (ETSI), рекомендациях Европейского союза вещания (EBU) и других организаций. В этих документах преимущественно утверждаются требования к параметрам и характеристикам, алгоритмам обработки сигналов и т. п., т. е. даются ответы на вопросы «Как должно быть?», «Каким требованиям должна отвечать та или иная система, устройство?», но в меньшей степени на вопрос «Почему так, а не иначе?».

Учитывая ограниченность объема любого учебного издания и данного учебно-методического пособия в том числе, в нем в доступной для студентов форме, но не в ущерб строгости изложения, рассматриваются на физическом и математическом уровнях вопросы кодирования источника сообщений в системах ЦТВ.

Во второй части учебно-методического пособия предполагается проанализировать канальное (помехоустойчивое) кодирование и модуляцию, алгоритмы компрессии изображений MPEG, а также оконечные устройства ТВ систем – преобразователи свет–сигнал и сигнал–свет.

1 ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ЦИФРОВОГО ПРЕДСТАВЛЕНИЯ СИГНАЛОВ ТЕЛЕВИЗИОННЫХ ИЗОБРАЖЕНИЙ

1.1 Обобщённая структурная схема цифровой системы передачи телевизионных изображений

Обобщённая модель цифровой системы передачи (ЦСП) информации включает три фундаментальных процесса: кодирование – декодирование (на передающей и приёмной сторонах соответственно) источника, кодирование – декодирование канала и модуляцию – демодуляцию.

Структурная схема модели ЦСП, приведённая на рисунке 1.1, очень слабо зависит от типа передаваемой информации. Последний определяет конкретные параметры и характеристики системы. Правда, это замечание относится к ЦСП, в которых передаётся только один вид информации, например, аудио или видео. Все современные ЦСП являются многоканальными и предназначены для одновременной передачи многих видов информации. Поэтому схема на рисунке 1.1 является обобщающей [1 – 5, 12].

На передающей стороне все виды обработки сообщений в блоках 2 – 6 служат цели преобразования их в такие сигналы, которые наиболее подходят для передачи по каналу конкретного типа – по открытому пространству (радио- или атмосферному оптическому каналу), либо по направляющей среде (кабелю металлическому или световодному, т. е. волоконно-оптическому). В источнике сообщений 1 информация представляется в виде сообщения, которое в блоке 3 преобразуется в электрический сигнал, называемый часто первичным. При передаче речи, музыки (т. е. аудиоинформации) таким преобразователем выступает микрофон. При передаче изображений – ТВ передающая камера. В обоих случаях первичные сигналы являются аналоговыми.

При передаче звуковых сообщений одномерная функция звукового давления $D(t)$ преобразуется в одномерный электрический сигнал $U(t)$ ¹. При передаче изображений источником сообщений являются объекты, оптическое изображение которых нужно ещё построить. Эту задачу выполняет оптическая система 2, которая на светочувствительной поверхности передающей трубки 3 (в общем случае выполняющей роль оптико-электронного преобразователя – ОЭП) строит плоское оптическое изображение $L_{\text{ОИ}}(x, y, t)$. В цветном телевидении отражённый от объекта световой поток сначала расщепляется светоделительной оптикой на три цветоделённых потока $L_{\text{ОИ.Р}}(x, y, t)$, $L_{\text{ОИ.Г}}(x, y, t)$, $L_{\text{ОИ.В}}(x, y, t)$, которые и поступают на три передающие трубки.

Таким образом, оптическое изображение является многомерной пространственно-временной функцией (ПВФ), которая в ТВ камере преобразует-

¹ Далее функция времени $U(t)$ будет записываться без аргумента t для упрощения записи и только там, где это принципиально – $U(t)$.

ся в одномерный ТВ сигнал: сначала оптическое изображение преобразуется в электронное (рельеф зарядов, фотопроводимостей или потенциальных ям в зависимости от типа ОЭП: электровакуумные передающие трубки, ПЗС- или КМОП – матрицы – CCD, CMOS), а затем с помощью электронной развёртки ПВФ преобразуется в сугубо временную и формируется видеосигнал $U(t)$. В этом и состоит принципиальное различие между процессами преобразования в электрический сигнал изображений и любой другой информацией [3, 5, 9].



Рисунок 1.1 – Обобщённая структурная схема цифровой системы передачи ТВ изображений

На приёмной стороне в блоках 8 – 11 производятся обратные операции для восстановления информации в исходном виде с минимально возможными искажениями. Принятый сигнал $\tilde{U}(t)$ на входе электронно-оптического преобразователя – ЭОП 11 всегда будет отличаться от переданного. Объясняется это неизбежными искажениями информации ввиду неидеальности процессов её прямого и обратного преобразования, отличием характеристик тракта от идеальных, а также действием внутренних и внешних помех. В каждом конкретном случае в зависимости от вида передаваемой информации нормируются допустимые искажения сигнала или вероятность ошибок. Случайный характер сообщений, сигналов, а также помех обусловил важнейшее значение теории вероятностей в теории связи и вещания. Вероятностные свойства сигналов и сообщений, а также среды, в которой передается сигнал, позволяют определять количество передаваемой информации и ее потери.

Описанием сигнала может быть некоторая математическая функция времени $x(t)$, отображающая реальный физический процесс: напряжение, ток, напряженность поля или другая характеристика. Однако исчерпывающее описание сигнала не всегда требуется. Для решения ряда задач достаточно более общего описания в виде нескольких параметров, характеризующих основные свойства сигнала, подобно тому, как это делается в системах транспортирования. Так, с точки зрения перевозки важно указать габаритные размеры и массу предметов (грузов), другие их свойства (например цвет) являются несущественными.

Сигнал также является объектом транспортирования (передачи), а техника телекоммуникаций и вещания – по существу техникой транспортирова-

ния (передачи) сигналов по каналам связи и вещания (каналам передачи). В учебнике по теории электрической связи [50] к параметрам сигнала, которые являются основными с точки зрения его передачи, отнесены: длительность сигнала T_C , его динамический диапазон D_C и ширина спектра F_C . Кроме того, введена более общая и наглядная характеристика – объём V_C передаваемого сигнала:

$$V_C = T_C F_C D_C, \quad (1.1)$$

где D_C – динамический диапазон – отношение максимальной мощности сигнала к наименьшей мощности, при которых обеспечивается требуемое качество передачи (в ЦСП – допустимая вероятность ошибок);

T_C – длительность сигнала, определяющая интервал времени, в пределах которого сигнал существует;

F_C – ширина спектра, занимаемая сигналом.

Динамический диапазон выражается обычно в децибелах. D_C речи диктора, например, равен 25 – 30 дБ, небольшого вокального ансамбля 45 – 65 дБ, симфонического оркестра 70 – 95 дБ. Во избежание перегрузок канала в радиовещании D_C часто сокращают до 45 – 50 дБ. При передаче ТВ изображений D_C изменяется в широких пределах – от 30 до 60 дБ.

Наконец, ширина спектра сигнала F_C даёт представление о скорости изменения сигнала внутри интервала его существования. В технике связи спектр сигнала сознательно сокращают. Это обусловлено тем, что аппаратура и линии связи имеют ограниченную полосу пропускаемых частот, а также необходимостью передачи в смежных полосах частот других сигналов. Сокращение спектра осуществляется исходя из допустимых искажений сигнала. Например, в телефонии требуется, чтобы речь была разборчивой и корреспонденты могли узнать друг друга по голосу. Для этого достаточно передать речевой сигнал в полосе от 0,3 до 3,4 кГц. Аналогично необходимая ширина спектра ТВ сигнала определяется требуемой чёткостью изображения и для принятого у нас стандарта D/K (625 строк, 50 Гц – частота полей) равна 6 МГц [6].

Поскольку первичные сигналы при передаче аудио- и видеоинформации являются аналоговыми, то кодирование источника предусматривает, во-первых, аналогово-цифровое преобразование (АЦП) и, во-вторых, собственно сжатие данных (уменьшение избыточности или компрессию), позволяющее существенно сократить скорость цифрового потока.

Задача кодера источника 4 состоит в том, чтобы с учетом статистических (вероятностных) и структурных свойств ТВ-изображений и сигналов, а также психофизиологических характеристик получателя сообщений – зрительной системы человека, – наиболее компактно представить объем первичного сигнала.

Цель кодера – выражение цифрового сигнала минимальным количеством двоичных символов в единицу времени с тем, чтобы сузить полосу

пропускания канала (либо в заданной полосе передать большее количество информации), сократить время передачи или уменьшить среднюю мощность сигнала в канале.

Скорость цифрового потока $V_{\text{инф}}$ на выходе АЦП пропорциональна частоте дискретизации $f_{\text{д}}$ и разрядности двоичного кодирования m

$$V_{\text{инф}} = f_{\text{д}} \log_2 N = f_{\text{д}} m, \text{ бит/с}, \quad (1.2)$$

где N – количество уровней квантования, $N = 2^m$.

Скорость $V_{\text{инф}}$ называют также информационной скоростью передачи, поскольку она характеризует количество информации. Современные методы устранения пространственно-временной (статистической), структурной и психофизиологической избыточности, присущей ТВ изображениям, позволяют скорость $V_{\text{инф}}$ уменьшить в $k_{\text{СЖ}}$ раз без заметного снижения качества воспроизводимого изображения: $V_{\text{инф.СЖ}} = V_{\text{инф}} / k_{\text{СЖ}}$ (сжатие в десятки, а иногда и в сотни раз).

Кодирование канала (в блоке 5) применяется для исправления ошибок, возникающих при приёме цифрового сигнала из-за действий различных помех и искажений. В общем случае кодирование канала может быть реализовано путём перезапросов (повторов) искажённых пакетов информации или путём коррекции ошибок при использовании специальных корректирующих кодов. В трактах вещания применяется только прямое исправление ошибок – FEC (Forward Error Correction). В любом случае кодирование канала увеличивает объём передаваемых данных – общую скорость цифрового потока $V_{\text{КК}}$ (при этом информационная её часть остаётся неизменной):

$$V_{\text{КК}} = V_{\text{инф.СЖ}} / R_{\text{К}}, \text{ бит/с}, \quad (1.3)$$

где $R_{\text{К}}$ – относительная скорость канального (помехоустойчивого) кодирования, $R_{\text{К}} < 1$.

Под $R_{\text{К}}$ понимают отношение количества бит на входе кодера к количеству бит на его выходе. Например, при $R_{\text{К}} = 1/2$ на каждый информационный бит приходится один проверочный. Исправляющая способность такого кодирования максимальна, но скорость цифрового потока при этом увеличивается в два раза. Обычно в системах ЦТВ используется каскадное кодирование, содержащее две ступени: первая называется внешней (по отношению к каналу), а вторая – внутренней. В системах ЦТВ DVB первого поколения применяется кодирование Рида – Соломона ($R_{\text{РС}}$) и сверточное ($R_{\text{СВ}}$), тогда относительная скорость канального кодирования в (1.3) определяется произведением $R_{\text{К}} = R_{\text{РС}} R_{\text{СВ}}$.

Модуляция используется для преобразования сигналов, представленных со скоростью $V_{\text{КК}}$, в радиосигналы (модулированные колебания) заданной полосы частот, что обеспечивает их передачу по конкретному физическому каналу. Дополнительным свойством сложных (многопозиционных) видов модуляции является более плотная упаковка данных в частотной обла-

сти, когда на единицу полосы пропускания приходится больше передаваемой информации.

В результате модуляции высокочастотное несущее колебание дискретно изменяется по фазе или по амплитуде и фазе одновременно, принимая M их значений. При этом длительность символа модулированного сигнала в $\log_2 M$ раз больше, чем у двоичного сигнала на входе, а символьная скорость во столько же раз меньше битовой на входе модулятора.

С другой стороны, символьная скорость связана с полосой частот канала F_K соотношением

$$V_{\text{СИМВ}} = F_K / b_p, \text{ симв/с}, \quad (1.4)$$

где b_p – коэффициент расширения полосы.

При использовании одночастотных методов передачи (одна несущая на канал с одним из видов многопозиционной модуляции – QPSK, QAM, PSK, APSK или др.) численное значение коэффициента расширения выбирается в пределах 1,1 – 1,4.

При использовании многочастотного метода передачи с COFDM (Coded Orthogonal Frequency Division Multiplexing – ортогональное частотное разделение каналов с кодированием) $b_p = 1,185$ (при DVB-T с полосой радиоканала 8 МГц). Суть этого метода более полно отражает расширенный перевод – ортогональное частотное мультиплексирование (большого числа модулированных несущих) с кодированием. Необходимость в COFDM будет пояснена ниже.

Обычно задана ширина полосы канала F_K , поэтому её не расширяют, но символьную скорость уменьшают в b_p раз. Тогда можно записать

$$V_{\text{СИМВ}} = B_{\text{КК}} / \log_2 M = F_K / b_p, \text{ симв/с}, \quad (1.5)$$

и, учитывая выражения (1.3)...(1.5), получить выражение для определения максимально допустимой скорости цифрового потока на выходе кодера источника, которую можно «вписать» в канал с шириной полосы F_K

$$V_{\text{ИНФ.СЖ}} = B_{\text{КК}} R_K = V_{\text{СИМВ}} R_K \log_2 M = F_K R_K \log_2 M / b_p, \text{ бит/с}. \quad (1.6)$$

От выбора методов выполнения рассмотренных трёх фундаментальных процессов и их параметров зависит эффективность ЦСП, т. е. степень использования пропускной способности канала, которая определяется известной теоремой Шеннона.

Формула (1.7), представленная ниже, имеет исключительно важное значение в теории передачи информации и цифровой связи. Она устанавливает взаимосвязь максимальной скорости передачи информации C_K в полосе канала F_K с отношением средних мощностей сигнала и шума $P_C/P_{\text{Ш}}$:

$$C_K = F_K \log_2(1 + P_C / P_{\text{Ш}}), \text{ бит/с}. \quad (1.7)$$

Выражение (1.7) определяет пропускную способность частотно-ограниченного непрерывного канала с аддитивным белым гауссовым шумом (АБГШ) при ограничении средней мощности передаваемого сигнала значением P_C . Как следует из выражения (1.7), пропускная способность в канале с шумом имеет конечное значение только при ограничении мощности передатчика. В канале без шума или канале с шумами, но без ограничения мощности передатчика, отношение сигнал/шум и соответственно пропускная способность стремятся к бесконечности.

При проектировании и анализе ЦСП наибольший интерес представляет пропускная способность, отнесённая к единице полосы частот

$$C_K / F_K = \log_2(1 + P_C / P_{\text{ш}}), \text{ бит}/(\text{с} \cdot \text{Гц}). \quad (1.8)$$

Формула (1.8) определяет максимальную удельную скорость передачи и применяется для оценки спектральной эффективности ЦСП (таблица 1.1).

Таблица 1.1 – Результаты расчёта удельной скорости передачи

$P_C / P_{\text{ш}}, \text{ дБ}$	-10	-5	0	5	10	15	20	25	30
$C_K / F_K, \text{ бит}/(\text{с} \cdot \text{Гц})$	0,14	0,4	1	2	3,47	5,04	6,68	8,33	10

По данным таблицы 1.1 построен график (рисунок 1.2) зависимости удельной скорости с АБГШ от отношения сигнал/шум. Причём вид модуляции выбирается таким, чтобы обеспечить передачу $n = 2F_K T_K$ символов в полосе F_K за время T_K . Фактически он определяет идеальную верхнюю границу, к которой стремятся приблизиться, оптимизируя те или иные параметры ЦСП.

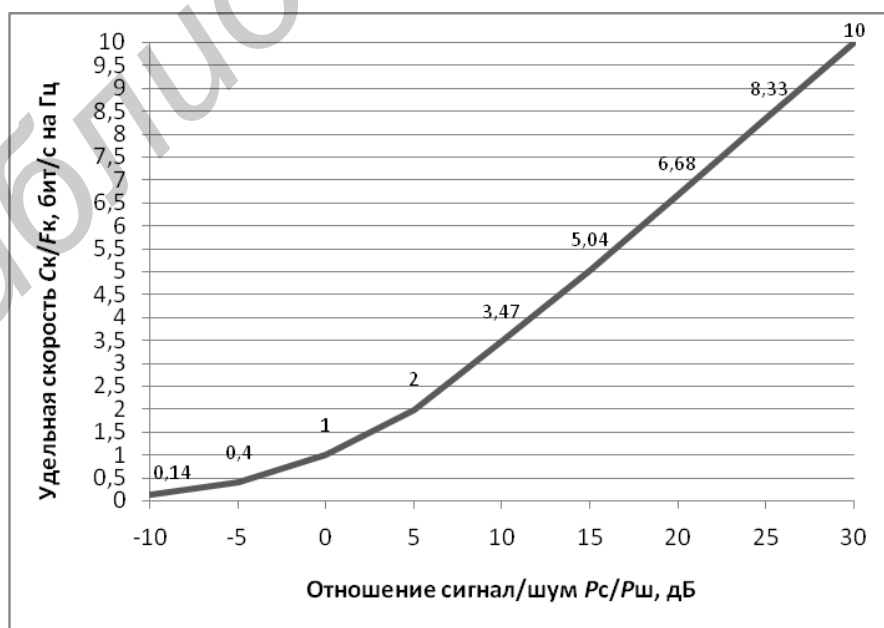


Рисунок 1.2 – Зависимость удельной скорости от отношения $P_C/P_{\text{ш}}$

По аналогии с объёмом сигнала V_C (1.1) вводят понятие объёма (ёмкости) канала V_K

$$V_K = F_K T_K D_K. \quad (1.9)$$

Очевидно, что канал не искажает сигнал, если выполняется условие

$$V_C < V_K. \quad (1.10)$$

В простейшем случае сигнал согласуют с каналом по всем трём параметрам, т. е. добиваются выполнения условий

$$F_C < F_K, T_C < T_K, D_C < D_K. \quad (1.11)$$

Тогда объём сигнала полностью «вписывается» в объём канала. Однако неравенство (1.10) может выполняться и тогда, когда одно или два из неравенств (1.11) не выполнены. Это означает, что можно производить обмен длительности T_C на ширину спектра F_C или ширины спектра F_C на динамический диапазон D_C и т. д.

Например, записанный в течение $T_{C, \text{зап}}$ сигнал с шириной спектра $F_C = 3,1$ кГц можно передать по каналу, полоса пропускания которого не 4 кГц, а только $F_K = 400$ Гц (не нарушая условия $V_C < V_K$). Для этого воспроизводить и передавать сигнал необходимо за время, примерно в 10 раз большее того, за которое он был записан (т. е. $T_{\text{воспр}} = 10T_{C, \text{зап}} = T_K$). При этом все частоты исходного сигнала уменьшаются в 10 раз – происходит обменная операция полосы на время передачи. На приёмной стороне сигнал сначала записывается в память и после воспроизведения со скоростью, в 10 раз большей, восстанавливается исходный сигнал с параметрами F_C и T_C .

Преобразуем выражение (1.6) и по аналогии с (1.8) запишем выражение для удельной скорости цифрового потока

$$V_{\text{инф}} / F_K = R_K (\log_2 M) / b_p. \quad (1.12)$$

Критериями оценки систем ЦТВ являются спектральная $\gamma_{C, \text{Э}}$ (1.12) и энергетическая $\gamma_{\text{Э, Э}}$ эффективности, определяющие количество информации, которое можно передать в заданной полосе F_K , и отношение средних мощностей сигнала и шума $\text{ОСШ} = P_C / P_{\text{ш}}$, обеспечивающие квазибезошибочный приём данных. В отличие от аналоговых систем в цифровых под $\gamma_{\text{Э, Э}}$ чаще используют нормированное ОСШ – как отношение, приходящееся на бит энергии E_b к спектральной плотности мощности шума N_0 . При этом E_b можно представить как мощность сигнала, умноженную на время передачи бита T_b , равную также тактовому интервалу T_T , а N_0 – как мощность шума $P_{\text{ш}}$, деленную на ширину полосы F_K . Поскольку время передачи бита и скорость их передачи $V_{\text{инф. СЖ}}$ (или тактовая частота) взаимно обратны, T_b можно заменить на $1 / V_{\text{инф. СЖ}}$. Тогда $E_b / N_0 = (P_C / P_{\text{ш}}) \times (F_K / V_{\text{инф. СЖ}})$ [52]. Подставляя $P_C / P_{\text{ш}}$ в ранее полученное равенство, находим:

$$R_K(\log_2 M)/b_P = \log_2(1 + (E_b/N_0) \times (B_{\text{инф.сж}}/F_K)), \text{ бит/с Гц.} \quad (1.13)$$

Анализ выражения (1.13) показывает, что в канале с фиксированной полосой частот F_K повысить пропускную способность можно только увеличением числа значащих позиций сигнала M . Но такой сигнал более чувствителен к искажениям вследствие неидеальности характеристик канала и менее помехозащищён. Поэтому потребуется либо увеличение отношения несущая/шум – $P_C/P_{\text{ш}}$, что не всегда возможно, либо применение более мощных кодов, исправляющих ошибки, что также имеет свои ограничения, т. к. возникает необходимость уменьшения информационной ёмкости сигнала, см. (1.6).

Системы ЦТВ второго поколения DVB-T2, S2 и C2 являются лучшим подтверждением сформулированных выводов. В них применяются: 1) методы помехоустойчивого кодирования с большей исправляющей способностью, что обеспечивает требуемую вероятность ошибок при меньшем отношении $P_C/P_{\text{ш}}$; 2) виды цифровой модуляции с большей спектральной эффективностью. Так, например, в системе кабельного телевидения (КЦТВ) второго поколения DVB-C2, с учётом меньшей подверженности линий связи (коаксиальных и особенно волоконно-оптических) действию внешних помех, число позиций M квадратурной амплитудной модуляции увеличено до 4096 (рисунок 1.3). При этом уровень сигнала в линии должен быть таким же, как и уровень несущей изображения $p_{\text{н.и.}}$ в аналоговой системе, но скорость передаваемого цифрового потока может составлять 81 Мбит/с [51]. В этой же работе приводятся результаты расчета с 65536-QAM, при которой уровень сигнала необходимо увеличить на 12 дБ, но ничего не сообщается о том, как при этом обеспечиваются требования по нелинейным искажениям.

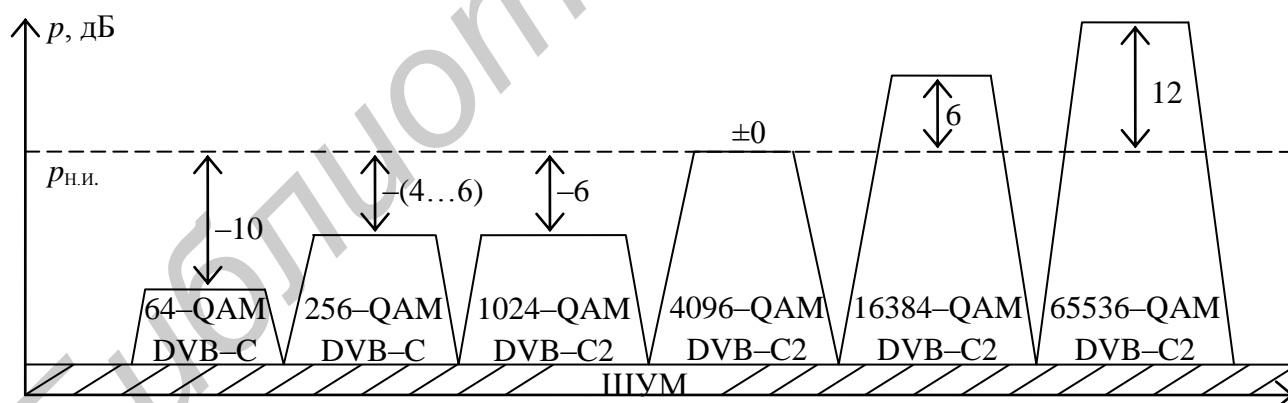


Рисунок 1.3 – Сравнение уровней сигналов в системах КЦТВ при разных значениях M при QAM-модуляции с уровнем несущей изображения в аналоговой системе

В современных системах ЦТВ с помощью самых передовых методов обработки и передачи сигналов (сжатие информации, каскадное помехоустойчивое кодирование, полосноберегающая модуляция и др.) достигается хорошее приближение к границе Шеннона (разница составляет менее 1 дБ).

1.2 Классификация систем цифрового телевизионного вещания

Современное состояние в области ЦТВ характеризуется большим разнообразием систем ЦТВ, применяемых способов сжатия (значительного уменьшения избыточности), формирования программных и транспортных потоков, методов помехоустойчивого кодирования и высокоэффективных многопозиционных видов цифровой модуляции.

Не претендуя на исчерпывающую классификацию, системы ЦТВ можно разделить на 17 классификационных признаков (рисунок 1.4). Первый признак (А) делит системы по виду передаваемых изображений на монохромные (черно-белые), цветные, стереоскопические и объемные. Поскольку ТВ вещание с электронной разверткой начиналось (в 1930-е годы) с передачи черно-белых изображений, то и системы А1 включены в классификацию. Это сделано еще и потому, что в ЦТВ черно-белые изображения передаются при трансляции снятых на черно-белую пленку кинофильмов. В последующие годы были приняты стандарты с различным числом строк в кадре – 405 (Великобритания), 525 (США) и 819 (Франция). Стандарт на 625 строк впервые был принят в СССР в 1945 г., и он стал наиболее распространенным. При внедрении цветного телевидения по системам NTSC (1953 г.) и PAL, SECAM (1967 г.) были оставлены только два значения числа строк разложения изображения – 525 (страны Северной Америки, Япония и Корея) и 625 (остальные) при чересстрочной развертке с частотой полей (кадров) 60 (30) и 50 (25) Гц соответственно. В системе NTSC по сравнению с черно-белым ТВ-значения частот полей (кадров) уменьшены в 1,001 раза – 59,94 (29,97) Гц.

Длительное время все операции по формированию, обработке, записи, передаче и приему сигналов осуществлялись в аналоговом виде. Даже когда в телефонии стали применяться цифровые методы передачи с импульсно-кодовой модуляцией, ТВ-вещание оставалось аналоговым еще долгие годы. Элементная база того времени, да и весь уровень развития техники не позволяли осуществить столь резкий скачок (почти в 4220 раз) – от 64 кбит/с (скорость в основном цифровом канале) к 270 Мбит/с и более.

Появление микросхем, пусть и малой степени интеграции, открыло путь к проведению научно-исследовательских и опытно-конструкторских работ (НИР и ОКР) по созданию комплекса цифрового оборудования аппаратно-студийного блока (АСБ) и центральной аппаратной (ЦА) телецентра. Усилиями нескольких коллективов ученых и инженеров предприятий, заводов и вузов, но в основном Всесоюзного научно-исследовательского института телевидения (г. Ленинград) в течение 1975 – 1982 гг. были выполнены ряд НИР и ОКР, которые завершились разработкой и настройкой первых в мире цифровых АСБ и ЦА, а также комплекса измерительных приборов. Сигнал яркости дискретизировался с частотой 12,5 МГц, а цветоразностные сигналы – поочередно через строку с частотой 3,125 МГц. Использовалось 8- и 7-разрядное двоичное кодирование. Скорость цифрового потока составляла 121,875 Мбит/с.

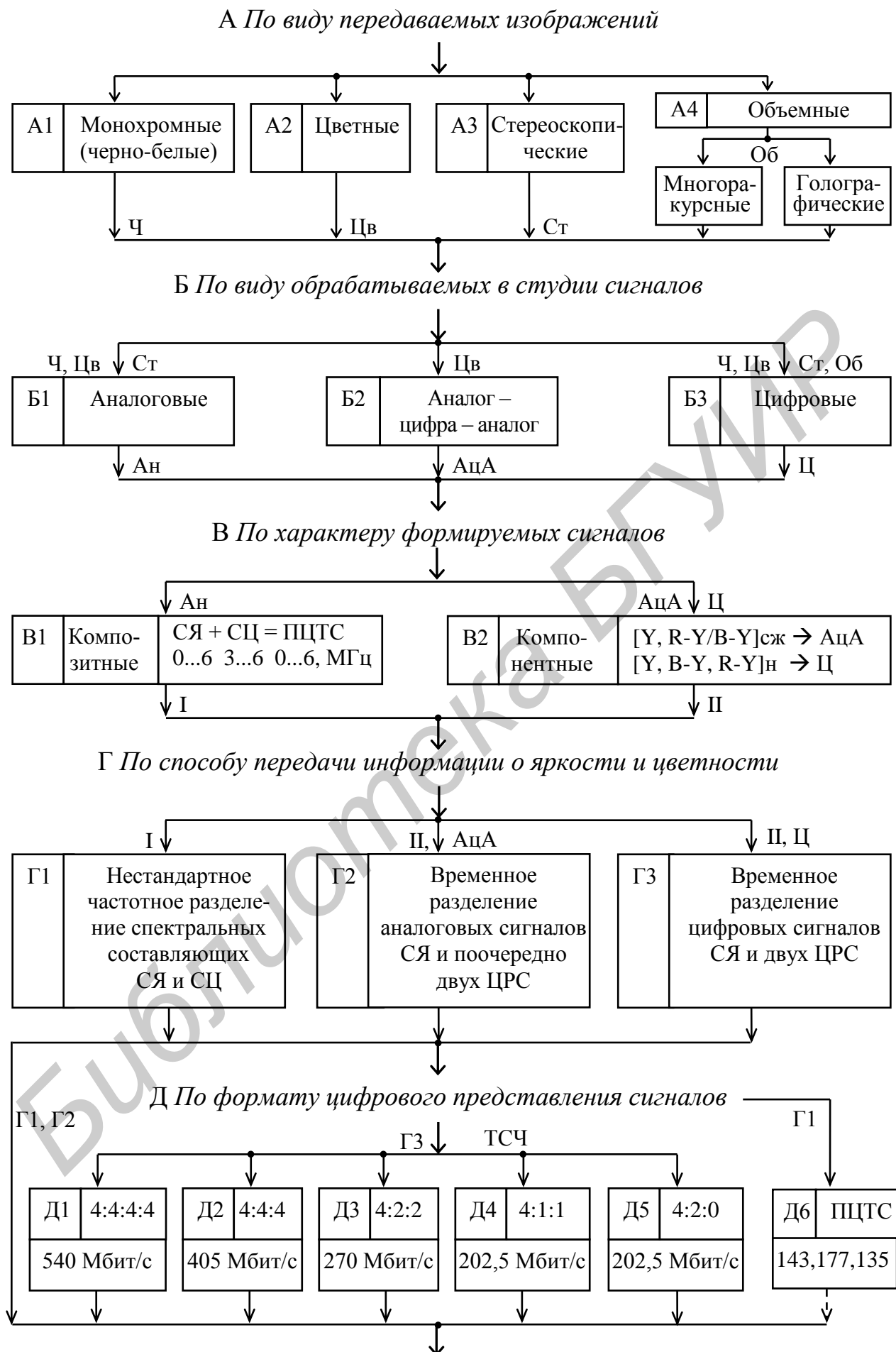


Рисунок 1.4 – Классификация систем цифрового телевизионного вещания

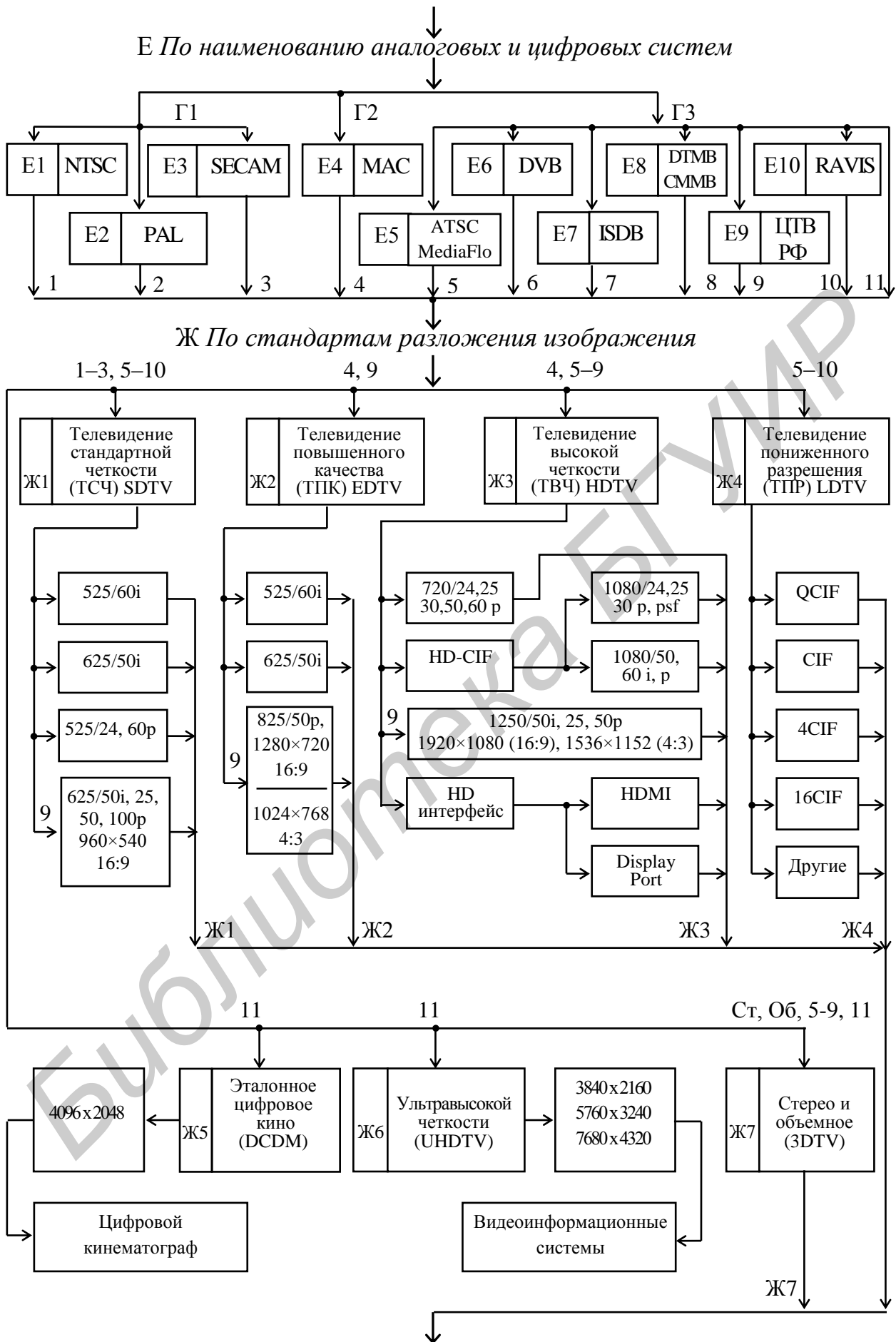


Рисунок 1.4, лист 2

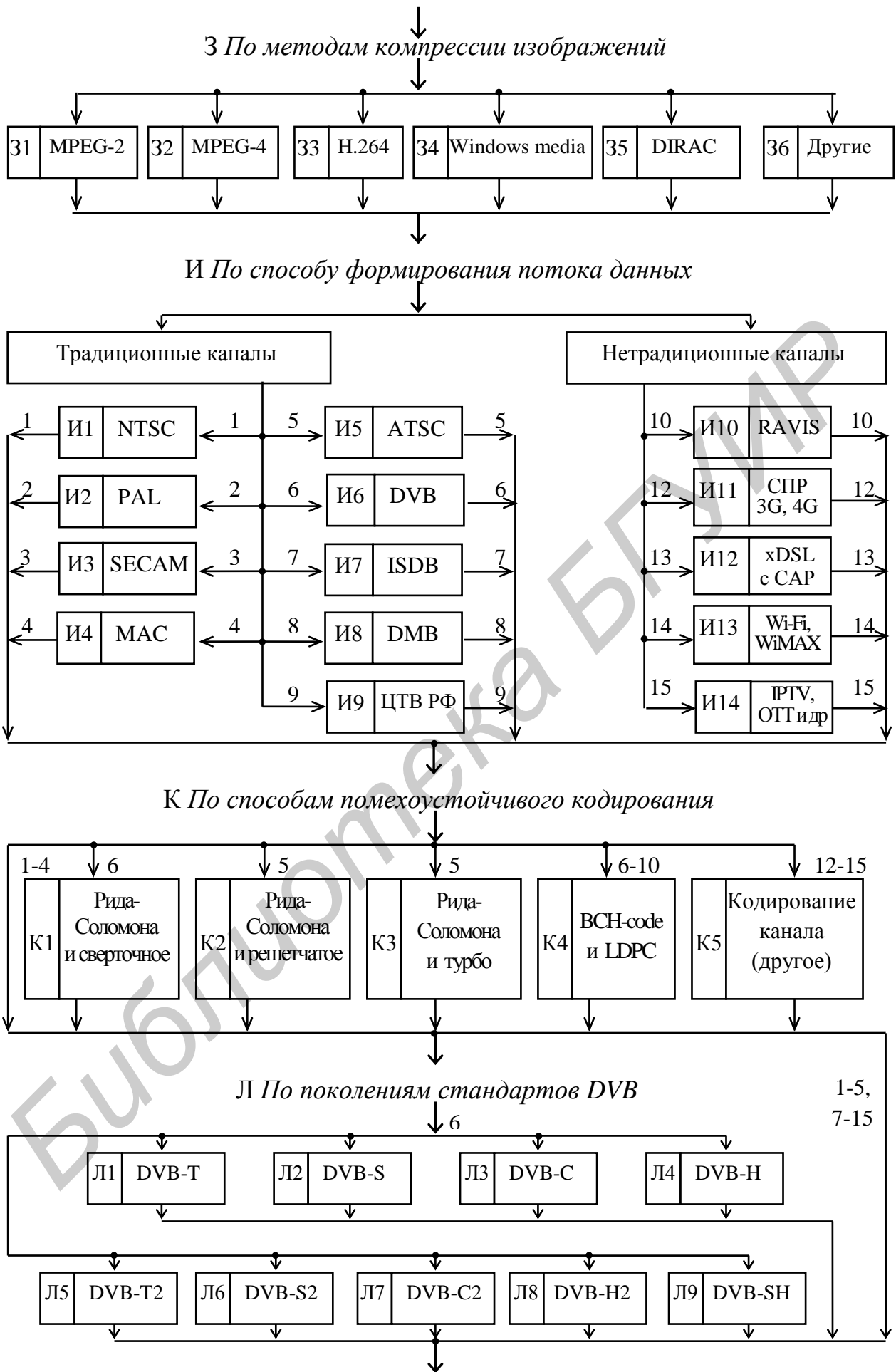


Рисунок 1.4, лист 3

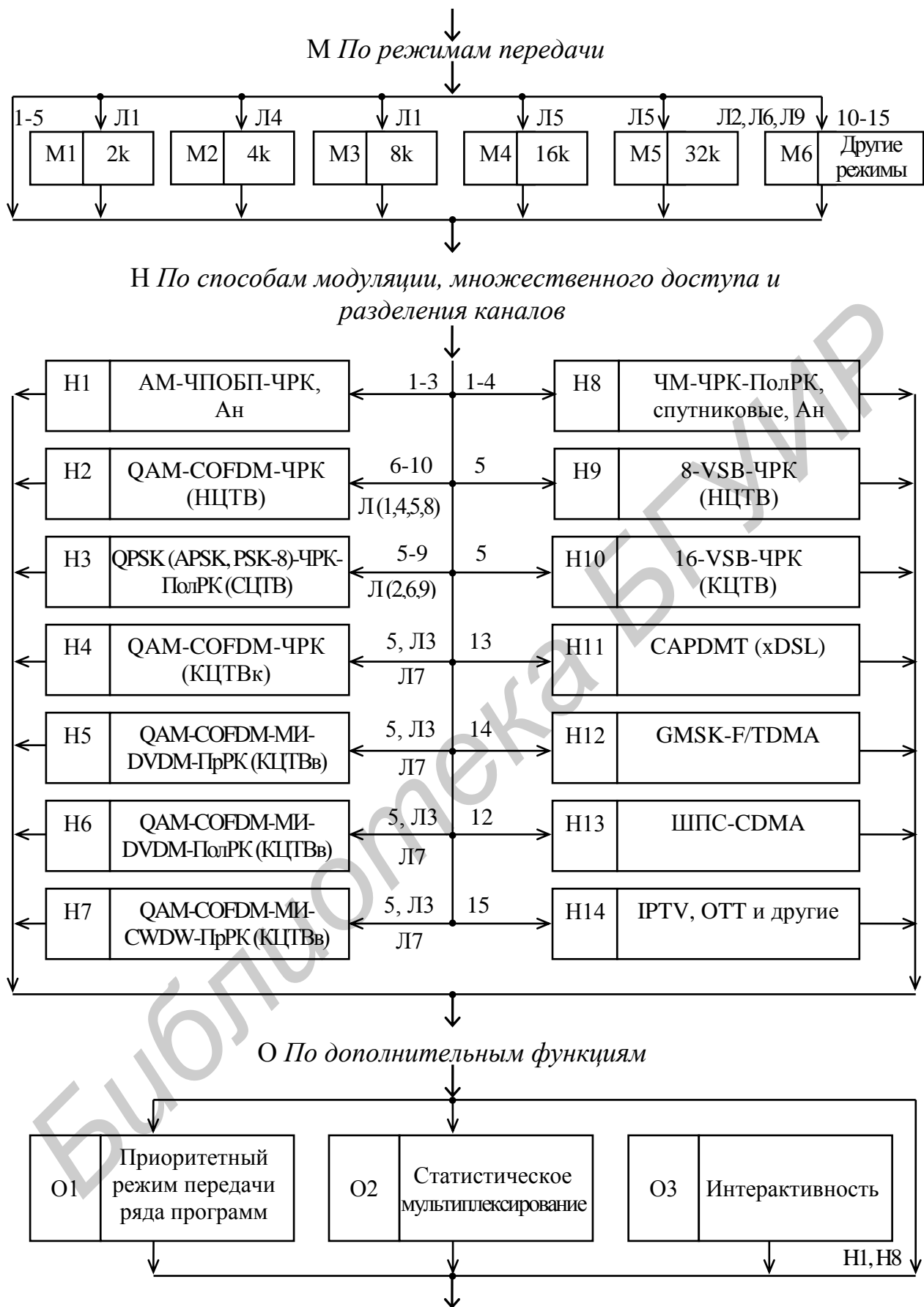


Рисунок 1.4, лист 4

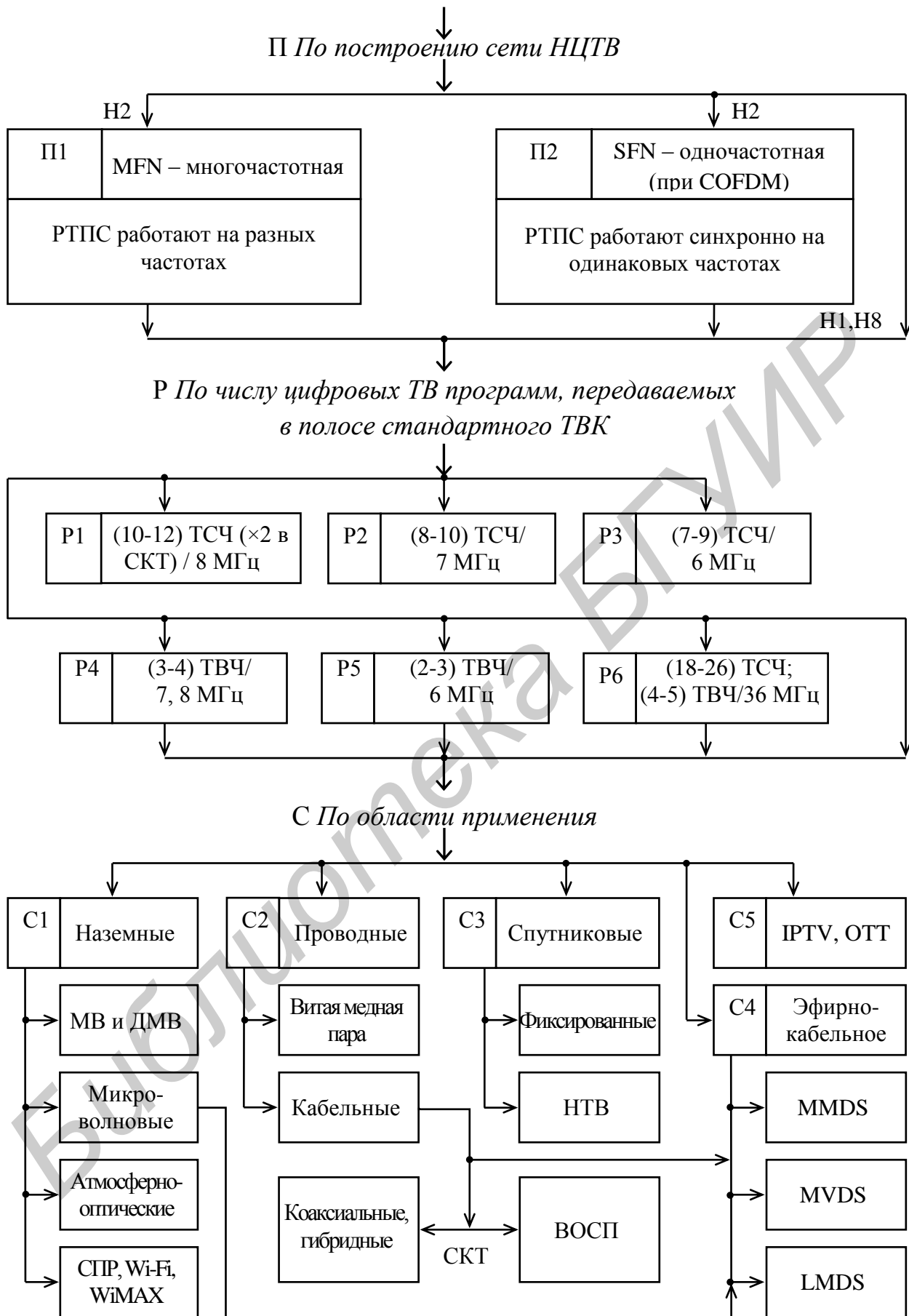


Рисунок 1.4, лист 5

Советские специалисты принимали активное участие в работе Международного консультативного комитета по радио (МККР, сейчас это сектор радиосвязи МСЭ-Р) и знали, что готовятся рекомендации по цифровому телевидению с параметрами: $f_{д.у} = 13,5$ МГц, $f_{д.црс} = 6,75$ МГц при 8-ми разрядном кодировании. ВНИИТ в 1982 – 1984 гг. по программе «Студия» разработал и изготовил цифровое оборудование для АСБ и ЦА с такими параметрами почти на год раньше, чем МККР принял соответствующую рекомендацию.

В мире в то время была разработана лишь одна экспериментальная цифровая студия французской фирмы «Томсон» (г. Ренн, 1985 г.), но по объему функций она уступала АСБ и не имела ЦА. Обе разработки – и советская, и французская – по уже упоминавшимся причинам были громоздки и много потребляли энергии. Об объеме оборудования на существующей тогда отечественной элементной базе можно только догадываться, если вспомнить, что первый видеомагнитофон для записи сигналов цветного ТВ на широкую магнитную ленту, изготовленный в США фирмой «Амрех», весил 1,5 тонны. Отсутствие же эффективных алгоритмов сжатия и помехоустойчивого кодирования не позволяли начать опытное вещание.

В 1990-е годы в СССР произошли известные события, и работы по ЦТВ на некоторое время были практически свернуты. Наверстывать упущенное всегда сложно, тем более в такой стремительно развивающейся отрасли как информационные технологии.

Разделение систем на стереоскопические (А3) и объемные (А4) сделано нами специально, чтобы подчеркнуть различное их толкование, несмотря на то, что «stereos» в переводе с греческого обозначает «твердый, объемный, телесный, пространственный». Иными словами стерео – это и есть объемное. В чем же тогда их различие в нашем контексте?

Мы видим объемное изображение объектов (предметов), во-первых, благодаря наличию двух органов зрения, разнесенных на определенное расстояние – базис и, во-вторых, – возможности изменения точки наблюдения и, следовательно, видения другого ракурса. Очевидно, технически проще всего реализовать первую функцию: съёмку вести двухобъективным (с призмёнными насадками) киносъёмочным аппаратом так, чтобы на обычной киноплёнке фиксировались два расположенных друг под другом левое и правое изображения (стереопара). Поэтому плёнка при съёмке должна продвигаться сразу на два кадра. Когда появилась широкая (70 мм) киноплёнка, в СССР впервые в мире учеными и инженерами научно-исследовательского кинофото института (НИКФИ) был разработан стереоскопический аппарат, обеспечивающий съёмку на один кадр плёнки двух изображений – левого и правого. Они получили широкое распространение в кинотеатрах прежде всего Москвы и Ленинграда. Еще больший успех ожидал их в США, когда американская техническая делегация ознакомилась с разработками НИКФИ.

Аналогично и в ТВ, только уже на уровне сигналов: требуется передавать два сигнала, соответствующие стереопаре. Следовательно, в стереоцвет-

ном ТВ передают полный цветовой ТВ-сигнал $U_{\text{ПЦТС.л}}(t)$ левого изображения и $U_{\text{ПЦТС.п}}(t)$ правого одновременно (требуется занятие двух ТВК), либо поочередно по кадрам (полям). Аналоговые системы стереоцветного ТВ (они названы МРТИ-1, МРТИ-2, автор – А. П. Ткаченко, и ЛЭИС – кафедра телевидения Ленинградского электротехнического института связи им. проф. М. А. Бонч – Бруевича, которую возглавлял тогда известный ученый Павел Васильевич Шмаков, один из родоначальников многих новых проектов по развитию ТВ, автор первой в мире монографии по стереотелевидению черно-белому и цветному (1966 г.), под его редакцией были изданы несколько учебников «Телевидение») подробно описаны в [10]. Опытные передачи черно-белых и цветных стереоизображений впервые в СССР и одни из первых в мире как раз и велись из небольшой студии ЛЭИСа (сейчас СПбГТУ-им. проф. М. А. Бонч – Бруевича).

Вторую и первую функции зрения обеспечивают объемные системы ТВ – многоракурсные (имеются сообщения о видеосъемке с помощью 32-х ТВ-камер, расположенных вокруг объекта съемки) и голографические (разработано несколько вариантов голографических экранов (дисплеев). В настоящее время стереоскопические и объемные системы ТВ реализуются в цифровом виде по стандарту DVB – 3DTV, см. также Введение и [54, 64].

По виду обрабатываемых в студии сигналов (признак Б) системы ТВ разделены на аналоговые, аналог – цифра – аналог и цифровые. В такой последовательности они исторически и развивались вместе с непрерывным повышением требований к качеству изображения, обеспечиваемому цветными ТВ-системами, которые к 1988 г. находились в эксплуатации уже долгие годы – NTSC (35 лет), PAL и SECAM (21 год). На этапе перехода от ТСЧ к ТВЧ было несколько новых систем, получивших название «телевидение повышенного качества» – ТПК (EDTV – Enhanced Definition Television). Наиболее интересными были системы MAC (Multiplexed Analog Components – мультиплексирование (по времени) аналоговых компонентных сигналов). ТПК обозначает системы, основанные на стандартах разложения изображения 525/625, но обеспечивающие повышенную разрешающую способность, обычный (4:3) или расширенный (16:9) формат кадра при неизменных или новых стандартах излучения радиотелевизионного сигнала.

Системы MAC использовались для передачи сначала программ ТСЧ по аналоговым каналам спутникового ТВ с частотной модуляцией, а затем и ТВЧ. Для того времени это было незаурядным решением – уплотнить во времени аналоговые компонентные сигналы СЯ и ЦРС с расширенными полосами частот 5,6 МГц для СЯ (по сравнению с 5,0 МГц для большинства стран Европы) и 2,8 МГц для ЦРС (что более, чем в два раза шире, чем в композитных системах NTSC, PAL и SECAM (1,3 МГц)). Причем в спутниковом канале ширина полосы частот не изменялась. Сигналы СЯ и ЦРС преобразовывали в цифровой вид и сжимали во времени в 1,5 и 3 раза соответственно, т. е. записывали в память и считывали быстрее. Затем их преобразовывали в аналоговый вид и получали СЯ и ЦРС с увеличенной во столько же

раз шириной полосы – до 8,4 МГц (происходит обмен полосы на время передачи, о котором шла речь на с. 20 этого пособия). Подаваемый на частотный модулятор земной станции сигнал представлял собой уплотненные во времени в течение строки сигнала: синхронизации и звукового сопровождения в цифровом виде (в интервале СГИ), один из ЦРС (например «красный») в течение $\sim 17,3$ мкс ($T_{\text{стр.акт}}/3 = 52 \cdot 106 / 3 = 17,3$) мкс) и сигнал яркости в течение $\sim 34,6$ мкс. В другой строке вместо «красного» передавался «синий» ЦРС. Поскольку спектр реальных сигналов, и МАС сигнала в том числе, спадает с ростом частоты, и в нем нет модулированных поднесущих цветности и звукового сопровождения, то не было необходимости увеличивать ширину полосы ЧМ тракта.

По характеру формируемых сигналов (признак В) ТВ системы могут быть композитными (В1), к которым относятся NTSC, PAL и SECAM (полосы частот указаны для стандартов D/K и L), и компонентными (В2), в которых ЦРС обозначены «сж» – сжатые во времени, «н» – нормированные.

По способу передачи информации о яркости и цветности (признак Г) системы делятся на три типа: к Г1 – отнесены системы PAL, SECAM и NTSC, к Г2 – системы МАС, Г3 – цифровые системы.

По формату цифрового представления сигналов ТСЧ системы разделены на 6 типов, для каждого из них указаны соотношения между частотами дискретизации СЯ и двух ЦРС: 1 – 3,375 МГц, 2 – 6,75 МГц и 4 – 13,5 МГц при формате кадра 4:3. Приведены также скорости мультиплексированного несжатого цифрового потока.

В Д6 приведены скорости для преобразованных в цифровой вид полных цветовых ТВ сигналов NTSC, PAL и SECAM. Этот вариант применялся в первые годы внедрения ЦТВ.

Признак Е делит аналоговые и цифровые системы по их наименованию: Е1 – Е4 NTSC, PAL, SECAM, МАС; Е5 – американская (США) ATSC; Е6 – европейская DVB, принятая в большинстве стран мира; Е7 – японская ISDB; Е8 – система DMB, имеющая три модификации – китайскую DTMB (и CMMB – эфирно-спутниковую) южно-корейскую и европейскую; Е9 – системы DVB по принятым в Российской Федерации (РФ) стандартам (ЦТВ РФ); Е10 – RAVIS, российская система цифрового наземного мультимедийного и звукового вещания РАВИС.

Принятая в США – Е5 – система ЦТВ ATSC (Advanced Television Systems Committee – Комитет перспективных (улучшенных) систем телевидения) получила название по имени разработавшего ее комитета по аналогии с ранее принятой аналоговой системой NTSC (National Television System Committee – национальный комитет телевизионных систем). При этом для мобильного ЦТВ принята система MediaFlo [1, 2, 19, 22, 24, 25].

Системы DVB (Digital Video Broadcasting) – Е6 – различаются по области применения: DVB-T (Terrestrial) – наземного цифрового телевизионного вещания (НЦТВ) первого поколения и DVB-T2 – второго поколения; DVB-S (Satellite) – спутникового цифрового телевизионного вещания (СЦТВ) перво-

го поколения и DVB-S2 – второго поколения; DVB-C (Cable) – кабельного цифрового телевизионного вещания (КЦТВ) первого поколения и DVB-C2 – второго поколения; DVB-H (Handheld) – мобильного цифрового телевизионного вещания (МЦТВ) и DVB-SH (Satellite Handheld), в которых прием сигналов осуществляется на мобильные терминалы [1, 2, 12, 22...28].

Семейство стандартов DVB перечисленными не ограничивается, имеется еще более 30 стандартов ETSI (European Telecommunication Standard Institute – Европейский институт стандартов по телекоммуникациям), среди которых: DVB-MC – на микроволновые многоточечные распределительные системы MMDS в диапазоне ниже 10 ГГц; DVB-MS – на многоточечные видеораспределительные системы MVDS в диапазоне выше 10 ГГц; DVB-TXT, DVB-SI – на передачу сигналов телетекста и сервисной информации; DVB-DSNG – системы спутникового сбора новостей; DVB-RCC, DVB-RCP, DVB-RCD, DVB-RCG и DVB-RCL – интерактивные каналы: для кабельной распределительной сети, по коммутируемой телефонной сети общего пользования или ISDN, через усовершенствованную беспроводную систему связи DECT, через систему GSM и для локальной распределительной системы LMDS соответственно; DVB-CS – на спутниковые распределительные ТВ-системы с коллективной антенной SMATV; DVB-DATA – передачи данных по любым сетям DVB и др.

Система ISDB (Integrated Services Digital Broadcasting – цифровое вещание с интеграцией служб) – E7 – принята в Японии и имеет много общего с европейской DVB [1, 2, 22, 24, 25].

Системы DMB имеют одинаковое обозначение (аббревиатуру), но различное наименование в зависимости от разработчика.

Стандарты РФ – E9 – на цифровые системы ТПЧ и ТВЧ [55 – 59] разработаны с учетом соответствующих рекомендаций ИТУ-R, европейских стандартов на системы DVB и Американского национального института по стандартизации и Общества инженеров кино и телевидения SMPTE (США). Особенностью цифровой системы ТВЧ (для которой принято расстояние наблюдения, равное $3h$, h – высота экрана) является возможность использования двух стандартов разложения изображения [55]. Первый содержит 1250 полных строк в кадре: $1250/25$, $50p$, $1250/50i$ с форматом кадра 16:9 (4:3), количеством пикселей в активной части кадров $1080 (1125) \times 1920$ и форматом представления 4:2:2 (при $25p$, $50i$), и 4:2:0 (при $50p$).

Частота дискретизации сигналов R , G , B , Y при частоте кадров 50 Гц (формат 4:2:0) равна $f_{дY} = 148,5$ МГц, а ЦРС – $f_{дЦРС} = 74,25$ МГц. Для формата 4:2:2 ($25p$, $50i$) соответствующие частоты равны 74,25 МГц и 37,125 МГц. Полное число отсчетов в строке для обоих форматов равно 2376 для сигналов R , G , B , Y и по 1188 – для каждого ЦРС.

Воспроизведение широкоформатного изображения высокой четкости должно осуществляться в формате кадра 16:9 с числом элементов изображения 1920×1080 . При этом сделано существенное дополнение: для цифрового ТВЧ предпочтительным является воспроизведение изображения высокой

четкости унифицированным растром с форматом отображения 15:9 и числом отображаемых элементов 1920×1152 . В результате этого внизу растра появится поле 1920×72 для отображения дополнительной информации, например, речевого или иноязычного сопровождения в цифро-буквенном виде или полного отображения изображения с форматом кадра 4:3 и числом отображаемых элементов 1536×1152 .

Второй стандарт разложения изображения в кадре содержит 1125 полных строк: $1125/25$, $50p$ и $1125/50i$ с форматом кадра 16:9, числом активных элементов изображения в кадре 1920×1080 . Частоты дискретизации при прогрессивной развертке ($50p$) совпадают с разложением на 1250 строк, а при развертках с $25p$ и $50i$ – с аналогичными для 1250 строк. Полное число отсчетов в строке при указанных параметрах разложения составляет: для сигналов R , G , B , Y – 2640 и по 1320 – для каждого ЦРС. Этих данных достаточно для нахождения скорости уплотненных во времени сигналов при формировании последовательного цифрового потока, что будет показано в п. 1.4.

В РФ разработан также государственный стандарт [58] на систему ЦТВ повышенной четкости (качества) – ТПЧ с построчным разложением на 825 строк, частотой кадров 50 Гц, форматом передаваемого изображения 16:9 (4:3), при активном числе строк в кадре 720 (768) и расстоянии наблюдения $l_H = 4h$. Частоты дискретизации приняты равными 74,25 МГц для СЯ и 37,125 МГц для ЦРС, при полном числе отсчетов в строке, равном 1800 (для R , G , B , Y) и 900 (для ЦРС). Здесь тоже рекомендуется воспроизводить изображение ТПЧ растром с форматом 15:9 и числом отображаемых элементов 1280×768 . На поле 1280×48 внизу растра можно отображать ту же информацию, что и в стандарте на 1250 строк, или полное изображение с форматом кадра 4:3 и числом отображаемых элементов изображения 1024×768 (4:3).

Следующим стандартом РФ на ЦТВ [59] является широкоформатное ТВ стандартной четкости – ТСЧ с форматом изображения 16:9 и разложением изображения на 960×540 активных элементов в кадре при полном числе строк 625 и активном – $540/576$. Расстояние наблюдения принято равным $l_H = 6h$. Предусмотрены параметры разложения $625/50i$, 25 , 50 , $100p$ с частотами дискретизации СЯ (ЦРС) – 18,5625 (9,28125) МГц – для $50i$ и $25p$; 37,125 (18,5625) МГц и 74,25 (37,125) МГц – для $50p$ и $100p$ соответственно. Полное число отсчетов в строке для сигналов R , G , B , Y (ЦРС) составляет 1188 (594).

Воспроизведение изображения должно осуществляться в формате 16:9 с числом элементов 960×540 предпочтительно на экранах с форматом 15:9 и числом элементов 960×576 . Возникающее поле 960×36 внизу растра используется как и в предыдущих стандартах или для полного отображения изображения с форматом 4:3 и числом элементов 768×576 .

Получилось столь большое разнообразие форматов передаваемых ТВ-изображений (4:3, 16:9) обычных, широкоформатных и широкоэкранных кинофильмов (от 1,5:1 до 2,75:1), форматов отображения и устройств воспроизведения (4:3, 5:4, 15:9, 16:9, 16:10), при котором трудно осуществлять их со-

гласование. По-видимому, оно должно производиться в телевизоре (дисплее) в ручном режиме либо автоматически.

Система E10 – RAVIS (Real-time Audiovisual Information System) – аудиовизуальная информационная система реального времени, разработана в РФ [60, 61] на основании патента [62]. Система предназначена для организации цифрового наземного звукового и мультимедийного радиовещания в I и II полосах ОВЧ – диапазона частот. В настоящее время под аналоговое ЧМ – вещание отведено две полосы в этом диапазоне: 65,8 – 74 МГц и 87,5 – 108 МГц (при этом участок 92 – 100 МГц отведен пятому ТВК). При стереофоническом вещании на частотный модулятор подается комплексный стереосигнал, сформированный в первой полосе по методу принятой еще в СССР полярной модуляции, а во второй – по международной системе с полярной модуляцией. Система РАВИС с существенно расширенными функциями должна заменить аналоговые системы с ЧМ при сохранении частотных распределений и предоставлении новых мультимедийных услуг: по передаче в одном стандартном радиоканале ЧМ – вещания более 10 звуковых стереопрограмм CD – качества или нескольких программ многоканального звука, возможности трансляции видеопрограмм для мобильного потребителя и реализации ТВ вещания в малонаселенных пунктах при резком удешевлении системы вещания.

Для видеосжатия используется оригинальная реализация кодера H.264/AVX, основанная на российских патентах «ноу-хау». В системе реализован высокоэффективный кодер звука HE-AAC, а также кодирование многоканального звука MPEG Surround, более помехоустойчивое кодирование (как в DVB-T2) и технология OFDM. Сравнение характеристик систем DRM+, IBOC-FM (HD Radio, Digital System C) и РАВИС показывает, что последняя значительно превосходит известные [61].

По стандартам разложения изображения (признак Ж) системы ЦТВ можно разделить на семь типов, из которых Ж1÷Ж4 относятся к ТВ, а Ж5 и Ж6 – к цифровому кинематографу и видеоинформационным системам. В отдельную группу Ж7 выделены системы стереоскопического 3D телевидения, которые могут иметь различное разрешение. Телевидение стандартной чёткости – ТСЧ обычно обозначают четырьмя цифрами 625 / 50 / 2:1 / 4:3 и 525 / 60 / 2:1 / 4:3, которые определяют количество строк разложения, частоту полей, кратность чересстрочной (interlaced – *i*) развёртки и формат кадра – отношение ширины к высоте передаваемого ТВ изображения. При переходе на цифровое ТВ системы ТСЧ дополнились возможностью применения прогрессивной (progressive – *p*), т. е. построчной развёртки и формата кадра 16:9.

Здесь указаны только четыре параметра, характеризующие особенности разложения изображения в процессе его развёртки. Действующие сейчас 10 стандартов на системы аналогового ТВ вещания (*B, D, G, H, I, K, K1, L, M, N*), принятые ещё в эпоху чёрно-белого ТВ, описываются ещё рядом параметров, которые пронумеруем с учётом четырёх предыдущих: 5) диапазон волн (метровые, дециметровые); 6) ширина полосы частот ТВ сигнала; 7) вид

модуляции несущих изображения и звука; 8) разность частот между несущими; 9) степень подавления нижней боковой полосы частот; 10) ширина полосы радиоканала вещательного телевидения; 11) уровни, форма и длительность синхронизирующих и гасящих импульсов.

С принятием аналоговых систем цветного телевидения (NTSC-1953 г., PAL и SECAM – 1967 г.) все перечисленные параметры остались в силе т. к. эти системы были разработаны с условием обеспечения совместимости с системами чёрно-белого ТВ. Появились новые параметры, характеризующие: 12) способ формирования сигналов основных цветов U'_R, U'_G, U'_B ; 13) способ формирования сигнала яркости U'_Y и первичных ЦРС U'_{R-Y} и U'_{B-Y} ; 14) способ формирования преобразованных ЦРС: I и Q (NTSC), V и U (PAL), D_R и D_B (SECAM); 15) способ формирования сигналов цветности U_{CI} и цветовой синхронизации U_{CIC} ; 16) способ обеспечения перемежения спектральных составляющих сигналов U'_Y и U_{CI} ; 17) метод формирования полного цветового ТВ сигнала; 18) метод формирования радиосигнала вещательного ТВ, равного сумме радиосигналов изображения и звукового сопровождения [3, 4, 6, 8, 9, 10].

При этом значения высшей частоты спектра ТВ сигнала F_B и ширина полосы частот радиоканала ΔF_{PK} (параметры 6 и 10), принятые в чёрно-белом ТВ, сохранились в цветном и в зависимости от стандарта имеют следующие значения ($F_B / \Delta F_{PK}$):

- 6/8 МГц (стандарты $D, K, K1, L$);
- 5,5/8 МГц (стандарт I);
- 5/7(8) МГц (стандарты B, G, H);
- 4,2/6 МГц (стандарты M, N).

Стандарт M (525 строк) принят в США, Японии и некоторых других странах; в большинстве других государств используются остальные девять стандартов (625 строк).

Каждая из трёх систем NTSC, PAL и SECAM может применяться с любым из 10 стандартов, образуя 30 комбинаций. На практике их число почти в два раза меньше: например, нет стран, в которых используется система SECAM с развёрткой на 525 строк (M) и 625 строк (I, N), не применяется система PAL со стандартом L , а система NTSC применяется только по стандарту M .

Из приведённых 18 параметров, характеристик и способов формирования и обработки сигналов в цифровых системах ТСЧ сохранены следующие: 1...6, 9 (только в ATSC), 10, 12 и 13, поз. 7, 11, 14 и 18 выполняются принципиально по-другому, а поз. 8, 15, 16, 17 не используются.

В системе ATSC сохранены частоты полей $f_{\text{пол}} = 60/1,001 = 59,94$ Гц и кадров $f_k = 30/1,001 = f_{\text{пол}}/2 = 29,97$ Гц при чересстрочной развёртке, как было в NTSC. Необходимость изменения частот полей и кадров по сравнению с 60 и 30 Гц, принятыми в чёрно-белом телевидении, обоснована в [8]. Дополнительно в ATSC приняты ещё два значения частот кадров 30 и 60 Гц при

прогрессивной развёртке. Поэтому когда указывают частоты полей и кадров 30 или 60 Гц, под ними нужно понимать 29,97 и 59,94 Гц.

Кроме того, в стандарте с 525-ю строками (США и ряд других стран) предусмотрен режим работы передающих ТВ камер с прогрессивной развёрткой с частотой 24 Гц, совпадающей с частотой съёмки кинофильмов. Необходимость в этом обусловлена тем, что при демонстрации кинофильмов, снятых с частотой 24 кадра в секунду, по ТВ с частотой полукадров (полей) 60 Гц и частотой кадров 30 Гц возникают сильно заметные искажения изображения (особенно быстро движущихся объектов) и звукового сопровождения (высота звука увеличивается в 30/24 раза). Для их уменьшения усложняют телекинопроектор, в котором из каждых четырех кинокадров формируют 5 ТВ кадров (10 полукадров): когда передается первый (т. е. нечетный) кинокадр – образуют один ТВ кадр (2 поля), на втором (четном) кинокадре – образуют 1,5 ТВ кадра (3 поля = 2 + 1 повторенное) и так далее поочередно – 2 поля, 3 поля. Поскольку 75 % вечернего и ночного времени в ТВ вещании тратится на показ кинофильмов (24 к/с), то при работе телекинопроектора с прогрессивной развёрткой и частотой кадров 24 Гц искажения не возникают. В ТВ приемниках для устранения яркостных мельканий необходимо преобразование частоты кадров 24 Гц в 48 или 72 Гц. В простейшем варианте на статических изображениях в модуле преобразования вида развёртки запоминают каждый кадр и воспроизводят дважды или трижды, на подвижных изображениях алгоритм обработки значительно усложняется.

Оценка разрешающей способности аналоговых ТВ систем осуществляется по специальным испытательным таблицам, штриховым мирам и по измерительным сигналам. При этом за счет развёртки изображение в вакуумной передающей трубке дискретизируется только по вертикали на $z_{\text{АКТ}}$ активных строк и воспроизводится таким же в кинескопе. Поэтому размер минимальной детали определяется диаметром электронного луча.

Высшая частота спектра ТВ сигнала при известных параметрах электронной развёртки определяется по выражению [3, 4, 8, 10]

$$F_E = p \cdot \frac{kz f_K}{2} \cdot \frac{1 - \beta}{1 - \alpha}, \quad (1.14)$$

где p – коэффициент Келла;

$k = b/h$ – формат передаваемого ТВ изображения;

f_K – частота кадров;

α и β – относительные потери времени на строчный и кадровый гасящие импульсы.

Подставляя в (1.14) численные значения, принятые в межгосударственном стандарте на систему вещательного ТВ [6]: $k = 4/3$, $z = 625$, $f_K = 25$ Гц, $\alpha = 0,18$, $\beta = 0,08$, получаем $F_B = 7,3$ МГц, а с учетом коэффициента Келла $p = 0,82$ – $F_B = 6$ МГц. Если тракт передачи видеосигнала аппроксимировать идеальным ФНЧ с частотой среза, равной 6 МГц, длительность им-

пульса на выходе фильтра не может быть уже $1/2F_B = 1/(2 \cdot 6 \cdot 10^6) = 83$ нс по уровню 0,5 (или фронта между уровнями 0,1...0,9), либо 166 нс – по нулевому уровню. Значение 83 нс является теоретическим пределом минимального размера одной детали – элемента изображения (ЭИ). Практически длительность ЭИ получается в 1,5 – 2 раза большей по ряду причин: идеальный ФНЧ нереализуем и в этом нет необходимости, т. к. нужно будет корректировать характеристику группового времени запаздывания, а в тракте передачи цифровых двоичных сигналов склон АЧХ специально делается пологим для уменьшения межсимвольных искажений.

Современные ТВ передающие камеры строятся на ПЗС- или КМОП-матрицах, у которых количество светочувствительных элементов дискретно как по горизонтали (строкам), так и по вертикали. Их общее количество ($N = N_{Э.Х.} \cdot N_{Э.У.}$) и определяет число ЭИ в кадре. В этом случае выражение для расчета F_B имеет вид [29]

$$F_B = p \cdot N_{Э.Х.} \cdot N_{Э.У.} \cdot f_k \cdot (1-\beta) \cdot (1-\alpha) / 2. \quad (1.15)$$

Если выполняется стандарт разложения по вертикали $N_{Э.У.} = z_{АКТ} = z \cdot (1-\beta)$ и горизонтали $N_{Э.Х.} = N_{Э.СТР} = k \cdot z_{АКТ} = k \cdot z \cdot (1-\beta)$, выражение (1.15) принимает вид (1.14).

При преобразовании ТВ сигнала в цифровой вид с частотой дискретизации $f_D = 13,5$ МГц формируются отсчеты или выборки, следующие с периодом $T_D = 74,074$ нс, создающие ортогональную пространственную структуру отсчетов (дискретную по двум направлениям), которые условно можно считать элементами изображения.

После преобразования на приемной стороне цифрового сигнала в аналоговую форму он подается на цветной кинескоп (люминофорный слой которого дискретен – мозаичный или штриховой), либо на матричные дисплеи (ЖК, плазменная или светодиодная панель, проектор), размер ЭИ на которых определяется одной физической ячейкой панели.

Оценка разрешающей способности ТВ системы с учетом изложенных факторов числом воспроизводимых ЭИ затруднительна. Поэтому за единицу измерения разрешающей способности принята «телевизионная линия» (ТВЛ). Под разрешающей способностью (раздельно по горизонтали и по вертикали), измеренной в ТВЛ, понимают максимальное суммарное количество светлых и темных линий, которые может воспроизвести устройство или система в целом, если испытательный сигнал прошел весь тракт. Здесь имеется некоторая аналогия с измерением разрешающей способности оптических приборов – числом периодов оптической волны, приходящимся на миллиметр или градус.

В компьютерной технике разрешающую способность дисплеев оценивают в пикселах (pixel = picture element – элемент изображения). Под ним понимают минимальный адресуемый элемент двумерного растрового изображения, для которого независимым образом можно задать яркость, цветовой

тон и насыщенность. Следовательно, пикселом называются точки, на которые делится экран дисплея, при графическом режиме работы или точки, из которых формируется изображение на струйных или лазерных принтерах [13].

Разрешающую способность устройств воспроизведения изображения в телевидении можно оценивать в пикселах, но ТВ системы в целом измеряют в ТВЛ.

Телевидение повышенного качества – Ж2 – ТПК (EDTV – Extended Definition Television) по межгосударственному стандарту определяется как система ТВ с неизменным стандартом разложения изображения на телецентре (т. е. таком же, как и в ТСЧ), но обеспечивающая лучшее качество воспроизводимого изображения на приемной стороне. Примером такой системы ТПК являлась уже описанная ранее система МАС. Для приема таких сигналов на стандартный цветной телевизор необходимо было кроме спутниковой приемной установки иметь декодер МАС. В связи с успехами ЦТВ спутниковые системы МАС сейчас не используются. В то же время в РФ принят стандарт на систему ЦТВ повышенной четкости с полным числом строк 825 и построчной разверткой с частотой кадров 50 Гц.

К ТВЧ – Ж3 – на рисунке 1.4 отнесены только системы с числом строк разложения 720, 1125 и 1250 при различных видах развертки – чересстрочной (*i* – interlaced) или прогрессивной (*p*) с различными значениями частот полей и кадров. Но в США по системе HDTV ATSC ТВ вещание начато по 18 разновидностям – сочетаниям количества строк разложения – 525, 720 и 1125 (1080 активных), типа развертки – *i* и *p* и числа полей и кадров 24, 30 и 60*p*, 24*psf*, 60*i*. При этом телевизоры должны принимать любой из сигналов и автоматически перестраиваться на нужный вид развертки. По-видимому, стандарт 525 даже с 60*p* нельзя относить к ТВЧ.

Усилиями 11-й Исследовательской Комиссии МСЭ удалось стандартизировать цифровой сигнал не только стандартного, но и высокого качества. Принят единый стандарт ТВЧ для производства и международного обмена программами с 1125 строками разложения изображения и разными значениями частот полей и кадров (таблица 1.2). В соответствии с рекомендацией ITU-R BT.709-5 сохраняется единый формат изображения (Common Image Format) HD-CIF в виде матрицы с 1080 активными строками в кадре (т. е. отсчетами по вертикали) и с 1920 квадратными элементами изображения (пикселями) в активной части строки (т. е. по горизонтали). При этом значения единых параметров не зависят от частоты кадров, значения которой могут быть равными 60, 50, 30, 25 и 24 Гц.

Необходимо отметить, что задолго до начала цифрового ТВ вещания в Японии было разработано все необходимое оборудование для системы ТВЧ на 1125 строк и велось спутниковое вещание с использованием аппаратуры сжатия MUSE. В Европе разрабатывалась система HD-MAC. Трансляция олимпийских игр из Сеула велась по системе ТВЧ MUSE, а олимпийских игр из Франции и Испании в 1992 г. – по системе HD-MAC.

В стандарт включены новые значения частоты кадров при прогрессивной развертке, в том числе 24 Гц, используемой в кинематографии. Это обеспечивает тождественность электронной и киноверсии фильма, что упрощает международный обмен программами в виде кинофильмов, которые можно редактировать в электронном виде и затем передавать по цифровым каналам связи.

Для формирования сигналов изображения определены прогрессивный (*p*) и чересстрочный (*i*) способы развертки. Для транспортировки сигналов, сформированных первым способом, можно применять канал передачи прогрессивных кадров, либо сегментированных (*psf*).

Таблица 1.2 – Значения параметров для HD-CIF ТВЧ на 1125 строк

Вид развертки	60 <i>p</i>	30 <i>p</i>	30 <i>psf</i>	60 <i>i</i>	50 <i>p</i>	25 <i>p</i>	25 <i>psf</i>	50 <i>i</i>	24 <i>p</i>	24 <i>psf</i>
Частота полей, кадров, сегментированных кадров	60	30	60		50	25	50		24	48
Кратность развертки	1:1	1:1		2:1		1:1		2:1		1:1
Номинальная полоса аналоговых сигналов, МГц	60	30		60				30		
Частота дискретизации, МГц: R, G, B, Y R _R , R _B	148,5 74,25	74,25 37,125		148,5 74,25				74,25 37,125		

Термин «сегментированный кадр» означает, что ТВ сигнал сформирован прогрессивным способом, а передается в виде двух сегментов, один из которых содержит нечетные строки, а другой – четные. Для распределения сигналов, сформированных вторым способом, необходимо применять канал передачи чересстрочных полей.

Все четыре семейства систем ЦТВ DVB, ATSC, ISDB и DMB обеспечивают передачу ТВ сигналов как стандартной SDTV, так и высокой четкости HDTV. Однако, только в США переход в 1998 г. на цифровое ТВ вещание начался сразу на HDTV. В Европе и в других регионах переход от аналога к цифре начался с SDTV, поскольку HDTV считалось роскошью.

В европейских странах интерес к ТВ вещанию по стандарту HD-DVB проявился примерно с 2005 г. В настоящее время количество ТВ программ, передаваемых в цифровом виде по спутниковым каналам (СЦТВ), значительно превышает число аналоговых (см. тж. Введение).

В классификационный признак Ж вошли и цифровые интерфейсы высокого разрешения – HD, которые пришли на смену аналоговым, что значи-

тельно упрощает транспортировку данных от видеоконтроллера до матричных дисплеев на ЖК, OLED и плазменных панелях. Интерфейс HDMI (High Definition Multimedia Interface – мультимедийный интерфейс высокого разрешения) был разработан международной ассоциацией стандартизации видеоэлектроники VESA (Video Electronics Standards Association) в конце 2002 г. как единый мультимедийный стандарт для бытовой и компьютерной электроники. При его разработке за основу был взят интерфейс DVI (Digital Video Interface). В 2006 г. ассоциация VESA утвердила спецификацию интерфейса нового поколения, названного DisplayPort версии 1.1. Стандарт DisplayPort заявлен как не подлежащий лицензированию и не облагаемый какими-либо выплатами, универсальный цифровой интерфейс для передачи аудио- и видеосигнала. Его внедрение позволяет производителям отказаться от интерфейсов VGA, LVDS, DVI и HDMI (таблица 1.3).

Унификация внутреннего и внешнего интерфейсов, как и отсутствие лицензионных выплат, позволяет производителям электроники снизить стоимость конечной продукции.

В DisplayPort предусмотрена поддержка стандарта HDCP защиты данных при широкополосной передаче от незаконного копирования, применяемого также в HDMI и получившего широкое распространение. В отличие от HDMI интерфейс DisplayPort обеспечивает передачу высококачественного звука без потерь и поддерживает современные звуковые форматы: Dolby Audio, Dolby True HD и Dolby Digital Plus.

В этот же классификационный признак для полноты картины включены и системы с более высоким разрешением – Ж5 и Ж6, чем HDTV – цифрового кинематографа (DCDM) с разрешением 4096×2048 пикселей и ультравысокой четкости (UHDTV) с разрешением от 3840×2160 до 7680×4320 пикселей, которые Международным союзом электросвязи рекомендованы и для видеоинформационных систем с большими экранами и соотношением сторон 16:9 – LSDI (Large Screen Digital Imagery). В настоящее время разработано 17 рекомендаций МСЭ по данной тематике. Некоторые параметры этих систем приведены в таблице 1.4 [52 – 54].

Следует отметить огромный вклад в развитие ЦТВ в мире профессора М. И. Кривошеева, который в течение 30 лет (1970–2000 гг.) был Председателем ТВ исследовательской комиссии (11-й ИК) МККР МСЭ (с 1993 г. – сектор радиосвязи МСЭ-Р). Его концепция «6 + 7 + 8 = XXI век» подчинила исследователей и разработчиков различных систем ЦТВ как ТСЧ, так и ТВЧ одной цели – обеспечить передачу нескольких цифровых ТВ программ ТСЧ или ТВЧ в стандартных полосах частот [1, 53].

Первая рекомендация по стандарту ТВЧ была принята в 1989 г. и по инициативе М. И. Кривошеева были продолжены исследования по повышению четкости ТВ изображения. Ее результатом стал принятый в 1999 г. единый мировой стандарт цифрового ТВЧ с разрешением 1920×1080 пикселей – нулевой уровень в иерархии будущих цифровых ТВ систем ультравы-

сокой четкости UHD TV. Они ориентированы на создание международных видеоинформационных систем (ВИС) и цифровой кинематограф.

Таблица 1.3 – Основные характеристики дисплейных интерфейсов

Параметр	Типы интерфейсов и значения параметров			
	DisplayPort	LVDS	DVI	HDMI
Поддержка аналогового интерфейса	Отсутствует	Отсутствует	Версия DVI-I	Отсутствует
Количество каналов передачи данных и синхронизирующих	1, 2 или 4 канала, синхронизация в основном канале	8 каналов, 2 синхроканала	3 – 6 каналов (два типа разъемов), 1 синхроканал	3 (тип А) и 6 (тип В) каналов, 1 синхроканал
Скорость передачи данных на канал	1,6 или 2,7 Гбит/с (версия 1.1), до 8,2 Гбит/с (версия 1.3),	945 Мбит/с	Макс. 1,65 Гбит/с	Макс. 1,65 Гбит/с
Суммарная пропускная способность, Гбит/с	1,6 – 25,6	7,56	4,95 и 9,9	4,95 и 9,9
Поддержка передачи аудиоданных	Полная	Отсутствует	Отсутствует	Полная
Дополнительный канал	1 Мбит/с (версия 1.1), 720 Мбит/с (версия 1.3)	Отсутствует	DDC	100 Мбит/с (начиная с версии 1.4)
Канальное кодирование	ANSI 8 В/10 В	Отсутствует	TMDS	TMDS
Защита данных	DPCP; HDCP	Отсутствует	Отсутствует	HDCP
Протокол передачи данных	Микропакетный режим	Последовательный поток данных	Последовательный поток данных	Последовательный поток данных
Использование в качестве внутреннего интерфейса	Присутствует	Только как внутренний интерфейс	Отсутствует	Отсутствует

Уже предложены новые форматы изображения UHD TV: первый уровень – 3840×2160 , второй 5760×3240 , третий – 7680×4320 ; и многоканальное звуковое сопровождение 22.2. Это не просто плод теоретических исследований: японская компания NHK реализовала систему UHD TV с разрешением 7680×4320 .

Таблица 1.4 – Основные характеристики большеэкранный системы отображения информации

Параметр	Значения параметров	
	Формат 3840×2160	Формат 7680×4320
Формат кадра	16 : 9	
Число отсчетов в активной части строки	3840	7680
Число активных строк в кадре	2160	4320
Структура дискретизации	Ортогональная	
Направление развертки	Слева направо, сверху вниз	
Формат пиксела	1 : 1 (квадратные пиксели)	
Соотношение частот дискретизации сигналов яркости и цветоразностных	4:2:0, 4:2:2, 4:4:4	
Частота кадров, Гц	24*, 25, 30*, 50, 60*	
Разложение изображения	Прогрессивное	
Число бит/пиксел	10, 12	
* Для систем 24, 30 и 60 Гц возможны также значения этих частот, деленные на 1,001.		

Эти системы ориентированы на демонстрацию разнообразной визуальной информации, в том числе и специально созданной телевизионной, на больших плоских экранах, установленных в многолюдных местах: на площадях, стадионах, улицах, вокзалах и т. д. ВИС заменяют во многом ставшие привычными неподвижные панно, стенды с афишами и рекламой. Подавать сигналы на такие экраны можно по волоконно-оптическим линиям связи, так как скорость цифрового потока для третьего уровня UHD TV составляет 16,6 Гбит/с в режиме 4:2:2 и при частоте кадров 25 Гц.

Председателем международной группы по проблемам мировой стандартизации цифровых мультимедийных ВИС избран в мае 2009 г. проф. М. И. Кривошеев, который с 2000 г. является Почетным председателем Исследовательской Комиссии по вещанию МСЭ-Р.

В 2005 г. был принят первый стандарт на системы цифровой кинематографии DCI (Digital Cinema System Specification), а в январе 2008 г. были официально приняты новые стандарты SMPTE (Society of Motion Picture and Television Engineers – общество инженеров по подвижным изображениям и телевидению США) для цифрового кинематографа D-Cinema Distribution Master (DCDM), D-Cinema Packaging MXF JPEG 2000 и др. В настоящее время цифровые стандарты кинопроизводства находят все большее применение в мировой практике.

В феврале 2011 г. ФГУП «Космическая связь» и компания «СинеЛаб Дата Дливверн» объявили о запуске проекта доставки цифрового кино через спутник непосредственно в кинотеатры на всей территории России и СНГ (за исключением удаленных регионов Дальнего Востока). По данным Digital Cinema.ru к 1 декабря 2011 г. около 660 кинотеатров (1332 зала) оснащены цифровыми проекторами. Приемная сеть, развернутая компанией «СинеЛаб», не имеет аналогов в России и к началу 2012 г. включает в себя 160 станций. До конца года их планируется установить во всех цифровых кинотеатрах России.

В состав систем ТВ пониженного (низкого) разрешения – ТПР (LDTV – Low Division Television) вошли системы, которые обеспечивают более низкое разрешение по сравнению с ТСЧ (SDTV). В соответствии с рекомендацией ITU-T BT.601 стандарту SDTV соответствует разрешение 720×576 пикселей. В LDTV под CIF (Common Interface Format) – общий формат интерфейса (сжатых видеоданных) по рекомендации ITU-T H.261 понимается цветное изображение с разрешением $1/2$ от стандартного, т. е. 352×288 пикселей. CIF расшифровывают еще и как Common (Source) Intermedia Format. Тогда формат QCIF (Quarter) дает разрешение, равное $1/4$ от стандартного, т. е. 176×144 (таблица 1.5).

В соответствии с Рекомендацией ITU-T H.263 форматы цифровых ТВ сигналов, передаваемых по компьютерным сетям, расширены, и они дополнились форматами SQCIF (Sub-Quarter CIF), 4CIF и 16CIF. Два последних не вписываются в принятую нами классификацию – системы LDTV.

Приведенные в таблице 1.5 значения максимальной скорости передачи двоичных символов соответствуют частоте кадров 30 (29,97) Гц, которая в зависимости от передаваемого сюжета может быть уменьшена до 5...15 Гц, т. е. в 2 – 6 раз. Во столько же раз уменьшится и скорость передачи в соответствии с выражениями (1.2) и (1.15).

Необходимость таких форматов LDTV объясняется нецелесообразностью передачи изображений со стандартным разрешением при воспроизведении на экранах дисплеев очень малых размеров, например как у сотовых телефонов. Следует, однако, отметить, что появление гибких дисплеев с размером экрана, как у малогабаритного телевизора, вынуждает и в системах мобильного цифрового телевизионного вещания (МЦТВ) передавать изображения со стандартным разрешением и даже более высоким.

Таблица 1.5 – Параметры цифровых ТВ сигналов по Рекомендации ITU-T H.263

Формат	Сигнал яркости	Цветоразностные сигналы	Максимальная скорость передачи, Мбит/с
	Число пикселей		
SQCIF	128 × 96	64 × 48	2,9
QCIF	176 × 144	88 × 72	8,7
CIF	352 × 288	176 × 144	34,8
4CIF	704 × 576	352 × 288	139
16CIF	1408 × 1152	704 × 576	557

По методам компрессии изображений (признак 3) системы разделены на 6 видов (см. раздел 2).

Следующий классификационный признак И (на рисунке 1.4) – по способу формирования потока данных и радиосигнала – делит системы на два типа: использующие традиционные (И1 – И9) и нетрадиционные (И10 – И14) (т. е. альтернативные) каналы ТВ вещания. К первому типу отнесены системы ЦТВ DVB, ATSC, ISDB и DMB. Во второй тип включены системы подвижной радиосвязи (СПР) третьего и четвертого поколения 3G, 4G; семейство систем цифровой абонентской линии – xDSL (x – prefix, Digital Subscriber Line), количество которых превышает 30, если учитывать все разновидности. Сюда также вошли беспроводные широкополосные системы доступа Wi-Fi (Wireless Fidelity – беспроводная точность) и WiMAX (World Wide Interoperability for Microwave Access – корпоративный доступ в микроволновом диапазоне) и семейство стандартов IPTV, под которыми понимаются системы цифрового ТВ вещания по протоколу Интернет (TCP/IP, IPTV, Web-TV, ONLINE-TV и Mobile-TV) [52, 63, 67].

Конкуренцию системе МЦТВ DVB-H могут составить СПР 3G и 4G, WiMAX, частично Wi-Fi и IPTV при организации беспроводного доступа в Интернет. Если передавать сигнал в стандарте ТВЧ, то его воспроизведение на миниатюрном экране является достаточно сложной задачей прежде всего потому, что невозможно, например, на дисплее сотового телефона отобразить целиком все ТВ изображение высокого разрешения. Его нужно масштабировать с учетом многих факторов, чтобы сделать просмотр максимально комфортным, но не в ущерб информационной наполненности кадра. Именно в этом направлении сейчас ведется интенсивная работа в шестой исследовательской комиссии (ИК №6) МСЭ-Р, отвечающей за вещание, по созданию устройства так называемого интеллектуального масштабирования. Здесь речь

идет о том, что нельзя просто вырезать из ТВЧ-картинки какую-то фиксированную часть и отображать ее на дисплее миниатюрного абонентского терминала. В каждый момент времени размер выделяемой части относительно целого изображения может меняться. В ИК №6 МСЭ-Р как раз и создают стандартное устройство, которое автоматически выполняло бы такое интеллектуальное масштабирование. Его создание позволит существенно повысить эффективность перефильтрации контента для МЦТВ и снизить расходы, связанные с этим. При этом перефильтрация можно будет осуществлять на любой стадии, будь то прямая трансляция или монтаж различных версий записанной программы.

Во всех системах ЦТВ для обеспечения квазибезошибочного приема применяется каскадное помехоустойчивое кодирование, состоящее из двух ступеней. По мере развития систем ЦТВ совершенствовались и типы кодов для обнаружения и исправления ошибок при декодировании: если в первом поколении систем ЦТВ применялось кодирование Рида – Соломона (как внешнее) и сверточное (как внутреннее) с декодированием по алгоритму Витерби, то в системах второго поколения применяются методы кодирования, позволяющие увеличить радиус зоны обслуживания.

Так, в варианте разработанной в КНР системы DMB-T в качестве внешнего кода выбран код Рида – Соломона, а в качестве внутреннего может использоваться параллельно-каскадированный (т. е. турбо), решетчатый (trellis) или систематический свёрточный код. В системе НЦТВ DVB-T2 используется BCH-CODE (Боуза – Чоудхури – Хоквигвема) в качестве внешнего и LDPC (Low Density Parity-Check Code – код с низкой плотностью проверок на чётность) в качестве внутреннего. Это сочетание по сравнению с DVB-T обеспечивает выигрыш в помехозащищённости примерно на 3 дБ при прочих равных условиях.

Способы помехоустойчивого кодирования будут проанализированы в Части 2 учебного пособия.

В признак Л включены только основные стандарты DVB, хотя ETSI разработан уже более 100 стандартов по ЦТВ.

По режимам передачи – признак М – системы разделены на 6 видов от $2k$ до $32k$ в зависимости от размера массива быстрого преобразования Фурье (Fast Fourier Transform), которое применяется для формирования большого числа поднесущих при COFDM [1, 2, 25].

По способам модуляции, множественного доступа и деления каналов (признак Н на рисунке 1.4) системы ЦТВ разделены на 14 основных видов. Здесь приняты следующие сокращения: QAM-COFDM-ЧПК в системах НЦТВ: QAM (Quadrature Amplitude Modulation – квадратурная амплитудная модуляция), COFDM (Coded Orthogonal Frequency Division Multiplex – ортогональное частотное уплотнение (большого числа модулированных несущих) с кодированием), ЧПК – частотное деление каналов; аналогично для системы ATSI – 8/16 VSB (Vestigial Side Band – метод модуляции с частичным подавлением одной боковой полосы). В СЦТВ применяются QPSK (APSK,

PSK-8) – ЧРК-ПолПК: QPSK (Quadrature Phase Shift Keying – квадратурная относительная фазовая модуляция); APSK (Amplitude Phase Shift Keying – амплитудно-фазовая модуляция); ПолПК – поляризационное разделение каналов. Системы КЦТВ_к по коаксиальному кабелю, а КЦТВ_в – по волоконно-оптическому; МИ – модуляция по интенсивности и DVDM (Dense Wavelength Division Multiplexing – плотное спектральное уплотнение по длинам волн) в волоконно-оптических системах передачи; GMSK-F / TDMA (Gaussian Minimum Shift Keying – Frequenz / Time Division Multiple Access) – гауссовская модуляция с минимальным частотным сдвигом с частотным / временным множественным доступом; ШПС-CDMA – широкополосная система с кодовым разделением и множественным доступом. Сюда же вошли и различные системы вещания через интернет (например OTT) и по протоколу интернет.

Среди дополнительных функций – признак О на рисунке 1.4 – выделены только три, описанные в [25].

По построению сетей (признак П) системы делятся на многочастотные (MFN) и одночастотные синхронные (SFN).

В наземной системе ЦТВ необходимо обеспечить уверенный прием на стационарную наружную антенну (фиксированный прием), переносную антенну внутри помещения и на портативную антенну в движущемся транспорте (мобильный прием). При этом во втором и особенно в третьем варианте качество принимаемых радиосигналов сильно зависит от возможных многочисленных отраженных сигналов. Вследствие их интерференции с прямым лучом возникают селективные замирания. Если задержка одного из лучей равна или больше половины длительности радиосимвола, резко увеличивается количество цифровых ошибок и прием становится невозможным.

Специально для борьбы с помехами при многолучевом приеме был разработан метод передачи COFDM-символов, о котором шла речь в п. 1.1. Сущность его заключается в следующем: высокоскоростной последовательный цифровой поток после помехоустойчивого кодирования (в кодере канала) поступает на COFDM-модулятор. В нем этот поток разделяется на большое число низкоскоростных потоков, которые методом QPSK или QAM модулируют свои несущие (активное число которых может изменяться от 1512 – в режиме $2k$ до 24176 – в режиме $32k$).

Благодаря большому числу несущих длительность символа в каждом из параллельных потоков оказывается в тысячи раз больше, чем в исходном последовательном потоке. Частотный разнос между несущими и, соответственно, длительность символа выбираются такими, что спектры модулированных несущих оказываются ортогональными: на центральной частоте каждой из несущих огибающие спектров всех остальных модулированных несущих проходят через ноль. Большая длина символа обеспечивает хорошую защиту от межсимвольных искажений, обусловленных интерференцией, так как отраженные сигналы чаще поражают не весь символ, а только его часть. Последнее объясняется тем, что длительность символа для каждой несущей

намного больше типичной задержки отраженных сигналов при многолучевом распространении.

Вместо подверженного селективному замиранию широкополосного канала с одной несущей в системе с COFDM передается большое число независимо подвергаемых замиранию узкополосных каналов. При этом небольшие группы несущих могут, вследствие замирания, быть полностью подавлены без потерь принимаемой информации.

Для увеличения степени защиты от многолучевого приема часть длительности символа отводится на защитный интервал, в течение которого передается как бы возвращенная назад во времени копия последующей части символа – в этот промежуток времени оценка значения символа в декодере не производится.

Система ЦТВ с COFDM и защитным интервалом имеет еще одно уникальное свойство – возможность организации синхронной сети вещания (SFN), когда большое число ТВ передатчиков работают строго на одной частоте и синхронно передают одинаковые символы COFDM. Если расстояние между передатчиками выбрано таким образом, что разница в запаздывании сигналов от соседних передатчиков не превышает длительности защитного интервала, приемники в SFN сети не испытывают помех от наложения сигнала.

По числу цифровых ТВ программ, передаваемых в полосе стандартного ТВ канала (признак Р на рисунке 1.4) системы разделены на 6 видов в зависимости от ширины полосы радиоканала в НЦТВ и КЦТВ, а также СЦТВ. Последним семнадцатым признаком (С) является область применения, которая понятна из рисунка 1.4 и не требует пояснений.

1.3 Преимущества систем ЦТВ

В научно-технической литературе обычно приводят 3 – 4 признака, которые характеризуют преимущества ЦТВ. Ниже приведены 15 признаков-преимуществ систем ЦТВ, разработанных на основе опыта проведенных научно-исследовательских работ на кафедре СТК БГУИР в области цифрового теле- и радиовещания.

1 Сохранение существующего плана распределения радиочастотного спектра (РЧС) в наземном, кабельном и спутниковом телевидении. Занимаемые полосы частот в ТВ вещании обсуждались в подразделе 1.2.

2 Передача в полосе частот одного радиоканала (в котором в аналоговой системе передается одна программа телевидения вместе с монофоническим звуковым сопровождением) до 5 – 6 цифровых программ со стереофоническим звуковым сопровождением при стандарте компрессии MPEG-2 и до 10 – 12 – при MPEG-4 вместе с несколькими цифровыми программами стереофонического звукового вещания и дополнительными данными. Концепция «6 + 7 + 8 = XXI век» (см. подраздел 1.2) не только для

ТВЧ, но и для ТСЧ казалась вначале фантазией, поскольку не было уверенности в практическом достижении таких больших коэффициентов сжатия, которые вместе с другими мерами позволили бы добиться столь высокой спектральной эффективности (см. также классификационный признак Р на рисунке 1.4).

Достижения поразительные. Чтобы в этом убедиться, рассчитаем скорость суммарного цифрового потока уплотнённых во времени сигналов яркости (СЯ) и двух цветоразностных (ЦРС) для одной и десяти программ ТСЧ (при частоте кадров $f_k = 25$ Гц) и одной программы ТВЧ (при $f_k = 25$ и 50 Гц). Режим 4:2:2 для ТСЧ обозначает, что $f_{д.У} = 13,5$ МГц, а $f_{д.ЦРС} = 6,75$ МГц. Тогда по выражению (1.2) при 10-разрядном кодировании $m = 10$ находим для ТСЧ

$$\begin{cases} B_{\Sigma,1} = f_{д.У}m + 2f_{д.ЦРС}m = 13,5 \cdot 10^6 \cdot 10 + 2 \cdot 6,75 \cdot 10^6 \cdot 10 = 270 \text{ Мбит/с;} \\ B_{\Sigma,10} = 2,7 \text{ Гбит/с.} \end{cases}$$

В ТВЧ (1920×1080) при $f_k = 25$ Гц: $f_{д.У} = 74,25$ МГц, $f_{д.ЦРС} = 37,125$ МГц; при $f_k = 50$ Гц значения частот дискретизации удваиваются. Аналогично по выражению (1.2) вычисляем

$$\begin{cases} B_{\Sigma,1} = 74,25 \cdot 10^6 \cdot 10 + 2 \cdot 37,125 \cdot 10^6 \cdot 10 = 1,485 \text{ Гбит/с;} \\ B_{\Sigma,1(50\text{Гц})} = 2 \cdot B_{\Sigma,1} = 2,97 \text{ Гбит/с.} \end{cases}$$

Теперь можно оценить, насколько трудной стояла задача – «вписать» сигналы с такими скоростями в полосу частот 8, 7 и даже 6 МГц. Кстати, необходимая инфраструктура для внутростудийного обмена такими потоками без сжатия уже создана [23].

3 Значительная экономия РЧС, которая образует так называемый «цифровой дивиденд». Это колоссальное преимущество, так как проблема электромагнитной совместимости радиоэлектронных средств является самой острой для всех стран. Например, в США только от сдачи в аренду высвобождающегося РЧС страна получит 25 млрд. долларов прибыли. В европейских странах благодаря использованию этого ресурса получают до 3,5 % ВВП (внутреннего валового продукта). В РФ этот показатель к концу 2006 г. составил 1,5 %. Прибыль отрасли подвижной связи (системы 3G и 4G, которые являются конкурентами мобильному ТВ вещанию) в 15 странах ЕС составила 105,6 млрд евро, создано 2,8 млн новых рабочих мест.

4 Одновременная передача с цифровыми ТВ программами нескольких программ цифрового радиовещания и/или дополнительной информации без расширения полосы частот $\Delta F_{РКВТ}$ радиоканала вещательного телевидения.

5 Повышенная помехоустойчивость систем передачи, записи и хранения сигналов цифровых ТВ программ.

6 Экономия энергопотребления ТВ передатчиками. Пусть $P_{\text{изл}} = 10$ кВт, если КПД равен 20 %, то потребляемая аналоговым передатчиком мощность составляет 50 кВт/час. В НЦТВ обычно $P_{\text{изл}} = 1$ кВт (реже $P_{\text{изл}} = 5$ кВт). Таким образом, с учетом передачи в одном радиоканале вещательного ТВ нескольких цифровых программ, на передачу одной цифровой ТВ-программы расходуется (0,5 – 1) кВт/ч, т. е. экономия энергопотребления составляет 50 – 100 раз.

7 Возможность внедрения ТВ-систем с новыми стандартами разложения изображения – ТВЧ (HDTV) и ТУВЧ (UHDTV).

8 Возможность создания интерактивных ТВ систем, в том числе «видео по запросу» (VoD) [1, 2, 24, 25], см. тж. Введение.

9 Применение приоритетного (иерархического) режима передачи, при котором одна или несколько программ из передаваемого на каком-то ТВК пакета, например, для мобильного приема, передаются с большей помехозащищенностью [25].

10 Возможность статистического мультиплексирования при передаче более 8 – 9 программ.

11 Возможность внедрения одночастотной синхронной сети с целью экономии РЧС (см. признак П и описание SFN-сети).

12 Создание мобильных ТВ систем наземного (DVB-H) и спутникового (DVB-SH) вещания.

13 Возможность распределения ТВ программ по любым цифровым системам передачи, обладающим необходимой пропускной способностью – ISDN, системам беспроводного широкополосного доступа (БШД) Wi-Fi и WiMAX, сети Интернет (IPTV), сетям подвижной радиосвязи третьего и четвертого поколений, микроволновым распределительным системам MMDS, LMDS и MVDS.

14 Расширение режимов уверенного приема – на стационарные, переносные и мобильные ТВ-приемники.

15 Возможность превращения ТВ приемника в многофункциональный информационный терминал.

В заключение следует отметить, что переход на цифровое вещание является доминирующим фактором возможности сближения (конвергенции) и интеграции различного рода информационных платформ (ТфОП, ПД, ИС-NGN, Интернет, БШД, сети WLAN и др.) в единое инфокоммуникационное пространство. Каждая из них предлагает различные услуги, а также, способы доставки информации. Но потребителя, как правило, не интересует техническая сторона способов доставки этой информации. Здесь важен конечный результат – предоставление услуг требуемого качества.

1.4 Теоретическое обоснование параметров исходных телевизионных сигналов для их цифрового представления

Источниками телевизионных сигналов при передаче из студий программных телецентров являются ТВ передающие камеры. В них формируются три аналоговых сигнала основных цветов U_R^* , U_G^* , U_B^* , хотя уже выпускаются камеры с выходами указанных (или других) сигналов в цифровом виде.

Будем исходить из того, что первичная обработка аналоговых сигналов осуществляется в аппаратной. Как известно [3, 4, 8, 9], спектральные характеристики чувствительности $S_R(\lambda)$, $S_G(\lambda)$, $S_B(\lambda)$ трех датчиков (например ПЗС-матриц), включая светоделительную оптику камеры, должны соответствовать кривым смещения $\bar{r}_\Pi(\lambda)$, $\bar{g}_\Pi(\lambda)$, $\bar{b}_\Pi(\lambda)$ в системе основных цветов приемника. Для этого в камере требуется восемь датчиков. Чтобы ее не усложнять в характеристиках $S(\lambda)$, реализуют только основные положительные участки, а нереализуемые побочные положительные и отрицательные частично компенсируются в цветокорректирующей матрице, в которой формируются сигналы U_R , U_G , U_B . Последние (после гамма-коррекции с показателем γ) обеспечивают наилучшее воспроизведение эталонных тестовых изображений.

Затем в кодирующей матрице с целью обеспечения совместимости с системой чёрно-белого телевидения образуются три сигнала U_Y' , U_{R-Y}' , U_{B-Y}' . Один из них – сигнал яркости (СЯ) U_Y' передаёт информацию только о яркости. Он формируется из гамма-корректированных сигналов основных цветов U_R^γ , U_G^γ , U_B^γ по принятому для ТСЧ выражению

$$U_Y' = 0,299U_R^\gamma + 0,587U_G^\gamma + 0,114U_B^\gamma, \quad (1.16)$$

а два других – цветоразностных сигнала (ЦРС) U_{R-Y}' и U_{B-Y}' передают информацию о цветности (цветовом тоне и насыщенности):

$$U_R^\gamma - U_Y' = U_{R-Y}' = 0,701U_R^\gamma - 0,587U_G^\gamma - 0,114U_B^\gamma, \quad (1.17)$$

$$U_B^\gamma - U_Y' = U_{B-Y}' = -0,299U_R^\gamma - 0,587U_G^\gamma + 0,886U_B^\gamma. \quad (1.18)$$

Каждый из них на выходе кодирующей матрицы занимает ширину полосы частот 0...6 МГц в соответствии с межгосударственным стандартом (РФ и РБ) на систему ТВ вещания [6]². По этому же стандарту и рекомендациям МСЭ на цифровое ТСЧ и ТВЧ-значение $\gamma = 0,45$.

Формирование СЯ U_Y' по выражению (1.16) указывает на его отличие от $U_Y^\gamma = (0,299U_R + 0,587U_G + 0,114U_B)^\gamma$. Очевидно, что $U_Y' \neq U_Y^\gamma$. Это одна из

² С целью упрощения записи в дальнейшем показатель гамма-коррекции в сигналах основных цветов будет обозначаться U' , и только там где это принципиально важно – U_Y^γ .

причин того, что принцип постоянной яркости выполняется не полностью [8]. Представление информации о цвете с помощью (СЯ) и двух ЦРС оказалось настолько продуктивным и оптимальным решением, что оно используется не только в аналоговых системах цветного ТВ вещания (NTSC, PAL, SECAM), но и в цифровых системах как ТСЧ, так и ТВЧ.

Во-первых, разрешающая способность зрения человека к цветным деталям в 3 – 4 раза меньше (зависит от цвета), чем к чёрно-белым. Поэтому полосу частот ЦРС можно во столько же раз сократить (до 1,3 МГц в NTSC, PAL, SECAM).

Во-вторых, зрение более чувствительно к искажению яркости, менее – к искажению цветового тона и ещё менее – насыщенности. Поэтому информация о яркости передаётся только одним сигналом – U'_Y , в отличие от варианта передачи сигналов основных цветов U'_R , U'_G и U'_B , каждый из которых передаёт информацию о яркости своего цвета. Вариант передачи сигналов U'_R , U'_G и U'_B также уступает выбранному.

В-третьих, три ЦРС по сравнению с сигналами основных цветов являются линейно зависимыми друг от друга. Поэтому третий ЦРС U'_{B-Y} не передаётся и формируется непосредственно в приёмнике (телевизоре) из двух принятых:

$$U'_{G-Y} = -0,509U'_{R-Y} - 0,194U'_{B-Y}. \quad (1.19)$$

Преимущество ЦРС состоит и в том, что при чёрно-белых сюжетах они обращаются в ноль.

Полученные сигналы в телевизоре подвергаются обратному матрицированию – каждый ЦРС суммируется с СЯ. Образуются три сигнала основных цветов $U'_{R.кин}$, $U'_{G.кин}$ и $U'_{B.кин}$, пригодные для подачи на кинескоп или другое воспроизводящее устройство.

Полученные сигналы основных цветов не полностью идентичны тем, которые были на выходе ТВ передающей камеры (см. $U(t)$ на рисунке 1.1). Объясняется это не только тем, что сигналы при передаче искажаются под действием помех и неидеальности характеристик. Здесь речь идет о так называемых системных ограничениях, которые сознательно закладываются в систему при ее проектировании и разработке. Одним из них является сокращение полосы частот ЦРС с 6 МГц до 1,3 МГц. Поэтому сигналы $U'_{R.кин}$, $U'_{G.кин}$ и $U'_{B.кин}$ получают следующими:

$$\left. \begin{aligned} U'_{R.кин} &= (U_R^{\gamma} - U_Y^{\gamma})_{НЧ} + U_Y^{\gamma} = (U_R^{\gamma} - U_Y^{\gamma})_{НЧ} + (U_Y^{\gamma} - U_Y^{\gamma})_{ВЧ} = U_R^{\gamma} - U_Y^{\gamma} + U_Y^{\gamma}, \\ U'_{G.кин} &= U_G^{\gamma} - U_Y^{\gamma} + U_Y^{\gamma}, \\ U'_{B.кин} &= U_B^{\gamma} - U_Y^{\gamma} + U_Y^{\gamma}, \end{aligned} \right\} \quad (1.20)$$

где индексами НЧ и ВЧ обозначены низкочастотные (0...1,3 МГц) и высокочастотные (1,3...6 МГц) компоненты сигналов.

Следовательно, в кодирующем устройстве сигналы основных цветов занимают полосу частот от 0 до 6 МГц, т. е. и крупные, и мелкие детали представлены в цвете, а на приемной стороне мелкие детали, соответствующие как следует из выражения (1.20) ВЧ компонентам 1, 3...6 МГц, воспроизводятся черно-белыми.

В цифровом телевидении также учитываются особенности цветового зрения для сокращения полосы частот ЦРС, но, как будет показано в п. 1.5, здесь предусмотрены четыре режима (или формата).

Таким образом, СЯ и ЦРС (называемые еще компонентными сигналами) согласованы со свойствами зрения, с простотой формирования сигналов основных цветов в телевизоре и поэтому их можно считать исходными для формирования канального сигнала с целью его передачи по открытому пространству (радиоканалу) или по какой-нибудь направляющей среде (металлическому или волоконно-оптическому кабелю).

Системы цифрового ТВ вещания независимо от вида канала являются *компонентными* и обеспечивают одновременную передачу нескольких цифровых ТВ программ и дополнительных данных путем временного уплотнения (ВРК) компрессированных цифровых сигналов СЯ и ЦРС, а также стереофонического звукового сопровождения и звукового вещания.

В аналоговых системах ТВ вещания принципиально нет другой возможности передавать ЦРС кроме как в пределах полосы частот СЯ (примерно от 3 до 6 МГц для стандартов D/K и L). Для этого путем модуляции цветовой поднесущей одновременно двумя ЦРС (NTSC и PAL) или поочередно через строку одним (SECAM) формируют сигнал цветности (СЦ) и суммируют его с СЯ так, чтобы их дискретные спектры перемежались (нестандартное частотное разделение каналов). Образующийся полный цветовой ТВ сигнал $U_{\text{ПЦС}} = U_Y' + U_{\text{СЦ}} + U_{\text{СЦС}}$ (СЦС – сигнал цветовой синхронизации) называется *композитным* (composite), как и сами системы цветного телевидения.

В этом и состоит одно из многих различий *компонентных цифровых и композитных аналоговых* систем ТВ вещания. Практически все виды искажений, присущие совместной передаче в общей полосе частот СЯ и СЦ и проанализированные в [8, 9, 14, 15], отсутствуют в цифровых системах ТВ вещания.

В цифровом телевидении возможны и другие варианты обработки сигналов, например в цифровой вид преобразуют аналоговые сигналы U_R^* , U_G^* и U_B^* , а все остальные операции осуществляют над цифровыми сигналами [21, 24]. Качество воспроизводимого изображения при этом самое высокое, но полосы частот трех сигналов должны быть одинаковыми и их нельзя сокращать. А это требует большей скорости цифрового потока (по сравнению с вариантом U_Y' , U_{R-Y}' , U_{B-Y}') и, как следствие, усложняется обработка сигналов и увеличивается объем памяти для хранения информации. Этот вариант рекомендуется для международного обмена ТВ программами.

Относительные значения аналоговых сигналов основных цветов, СЯ и ЦРС при передаче вертикальных цветных полос с номенклатурой 100/0/100/0, рассчитанных по выражениям (1.16)...(1.18), приведены в таблице 1.6, а временные диаграммы – на рисунке 1.5 (см. рисунки А.1 и А.2 (Приложение А)). Из нее следует, что сигналы U'_R , U'_G , U'_B и U'_Y (без синхросигналов) принимают значения от 0 до 1. Во всех странах, использующих стандарт разложения на 625 строк, принято, что указанные сигналы на нагрузке 75 Ом создают напряжение 700 мВ, а вместе с синхросигналом (в одном из сигналов, например U'_Y) – 1 В (1000 мВ). Здесь учтено отношение уровней синхросигнала и видеосигнала, равное 3:7 [3, 4, 6, 9, 10].

В США (и других странах, использующих стандарт на 525 строк) отношение этих уровней составляет 4:10. Это различие стало причиной того, что при том же размахе полного сигнала в 1 В, расстояние между уровнями гашения и белого составляет не 700 мВ, а $1000 \times 10/14 = 714,285714$ мВ. Это число не только трудно для запоминания, но и неудобно для практического применения. Поэтому институт радиоинженеров США (IRE) ввел специальную единицу измерения, которую так и назвали – единица IRE.

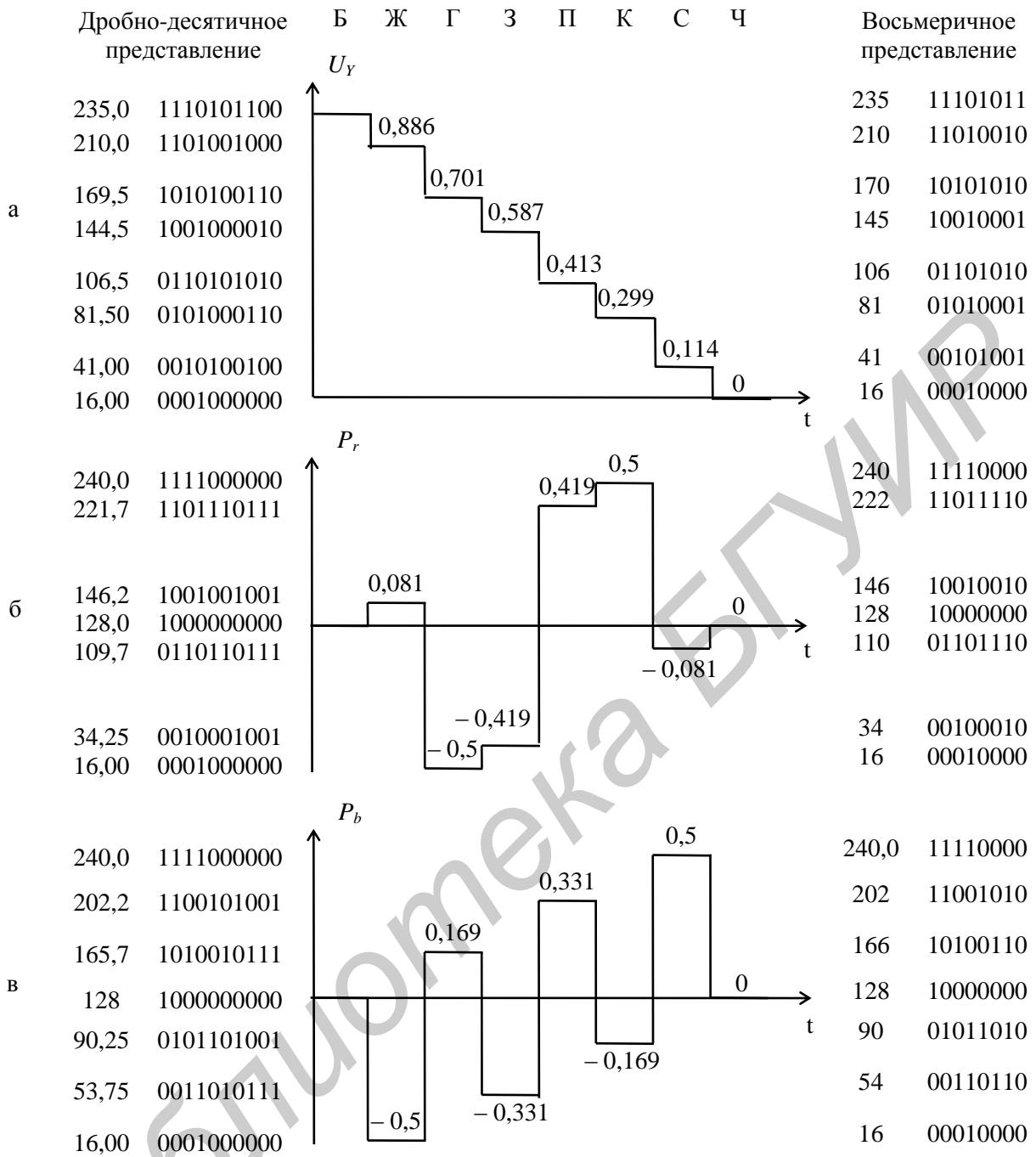
Одна единица IRE равна 7,14 мВ, разница уровней белого и гашения составляет 100 IRE, а размах полного сигнала равен 140 IRE. Защитный интервал между уровнями черного и гашения составляет 7,5 IRE (53,6 мВ).

Из таблицы 1.6 также следует, что относительные значения U'_{R-Y} находятся в пределах от минус 0,701 до +0,701, а U'_{B-Y} – от минус 0,886 до +0,886. Поскольку сигналы U'_{R-Y} и U'_{B-Y} являются двуполярными, их значения необходимо нормализовать так, чтобы размах каждого равнялся размаху СЯ (1,0 отн. ед., т. е. 700 мВ). Для этого они должны изменяться от минус 0,5 (350 мВ) до +0,5 (+350 мВ), т. е.

$$P_r = |U'_{R-Y}|_{\text{МАКС}} / k_r = 0,5; \quad P_b = |U'_{B-Y}|_{\text{МАКС}} / k_b = 0,5. \quad (1.21)$$

Таблица 1.6 – Относительные значения сигналов основных цветов, СЯ и нормированных ЦРС

Цвет	U'_R	U'_G	U'_B	U'_Y	U'_{R-Y}	U'_{B-Y}	P_r	P_b
Белый	1,0	1,0	1,0	1,0	0	0	0	0
Желтый	1,0	1,0	0	0,886	0,114	-0,886	0,081	-0,5
Голубой	0	1,0	1,0	0,701	-0,701	0,229	-0,5	0,168
Зеленый	0	1,0	0	0,587	-0,587	-0,587	-0,418	-0,33
Пурпурный	1,0	0	1,0	0,413	0,587	0,587	0,418	0,33
Красный	1,0	0	0	0,229	0,701	-0,229	0,5	-0,168
Синий	0	0	1,0	0,114	-0,114	0,886	-0,081	0,5
Черный	0	0	0	0	0	0	0	0



а – СЯ; б – нормированные ЦРС P_r ; в – P_b

Рисунок 1.5 – Временные диаграммы

Учитывая (1.21) и данные таблицы 1.6, определим численные значения коэффициентов k_r и k_b :

$$\begin{aligned}
 k_r &= \left| U'_{R-Y} \right|_{\text{МАКС}} / 0,5 = 0,701 / 0,5 = 1,402, \\
 k_b &= \left| U'_{B-Y} \right|_{\text{МАКС}} / 0,5 = 0,886 / 0,5 = 1,772.
 \end{aligned}
 \tag{1.22}$$

Учитывая (1.17, 1.18) и (1.22), находим формулы, устанавливающие взаимосвязь между сигналами основных цветов U'_R , U'_G , и U'_B и нормированными ЦРС P_r и P_b :

$$P_r = U'_{R-Y} / k_r = (0,701U'_R - 0,587U'_G - 0,114U'_B) / 1,402 = 0,5U'_R - 0,419U'_G - 0,081U'_B, \quad (1.23)$$

$$P_b = U'_{B-Y} / k_b = (-0,299U'_R - 0,587U'_G + 0,114U'_B) / 1,772 = -0,169U'_R - 0,331U'_G + 0,5U'_B.$$

Аналогично (1.20) в телевизоре из СЯ U'_Y и ЦРС P_r и P_b образуют сигналы основных цветов (для подачи на кинескоп или другое воспроизводящее устройство) по выражениям:

$$U'_R = U'_Y + 1,402P_r, \quad U'_G = U'_Y - 0,713P_r - 0,344P_b, \quad U'_B = U'_Y + 1,772P_b - 0,344P_b. \quad (1.24)$$

Необходимо отметить, что в цифровом ТВ исходные аналоговые ЦРС U'_{R-Y} , U'_{B-Y} , а следовательно, и P_r , P_b в зависимости от формата их цифрового представления (см. подраздел 1.5) могут иметь ширину полосы, равную 5,75; 2,375 или 1,375 МГц.

1.5 Цифровое представление компонентных и цветоразностных сигналов яркости

В цифровом ТВ вещании исходными сигналами для цифрового представления могут быть сигналы основных цветов U^*_R , U^*_G , U^*_B или сигнал яркости U'_Y и два цветоразностных сигнала U'_{R-Y} и U'_{B-Y} . Рассмотрим процесс формирования цифровых сигналов из аналоговых СЯ и ЦРС по обобщенной структурной схеме (рисунок 1.6). Для того, чтобы эти сигналы имели одинаковые размахи в относительных единицах, их уровни нормализуются в масштабных усилителях А1.1...А1.3. В А1.2 и А1.3 по выражениям (1.23) формируются ЦРС P_r и P_b . Уровни всех сигналов были приведены в таблице 1.6, а их форма – на рисунке 1.5.

Как было обосновано в подразделе 1.4, в системах ЦТВ, как и в аналоговых системах цветного телевидения, учитываются свойства зрения для сокращения полосы частот ЦРС. Но если в NTSC, PAL и SECAM она сокращается до 1,3 МГц (более чем в 4 раза по сравнению с полосой частот СЯ 0..6 МГц для стандартов D, K, K1, L), в ЦТВ предусмотрено четыре варианта (режима или формата), которые условно обозначают 4:4:4, 4:2:2, 4:2:0, 4:1:1 для ТСЧ, в ТВЧ используются только первый и второй.

С одной стороны, эти режимы указывают на соотношение частот дискретизации СЯ и ЦРС. При формате ТВ изображения 4:3 частота дискретизации СЯ в соответствии с Рекомендацией МСЭ-Р ВТ.601-6 [30] принята равной 13,5 МГц для всех стандартов разложения изображений на 625 и 525 строк, описанных в подразделе 1.2.

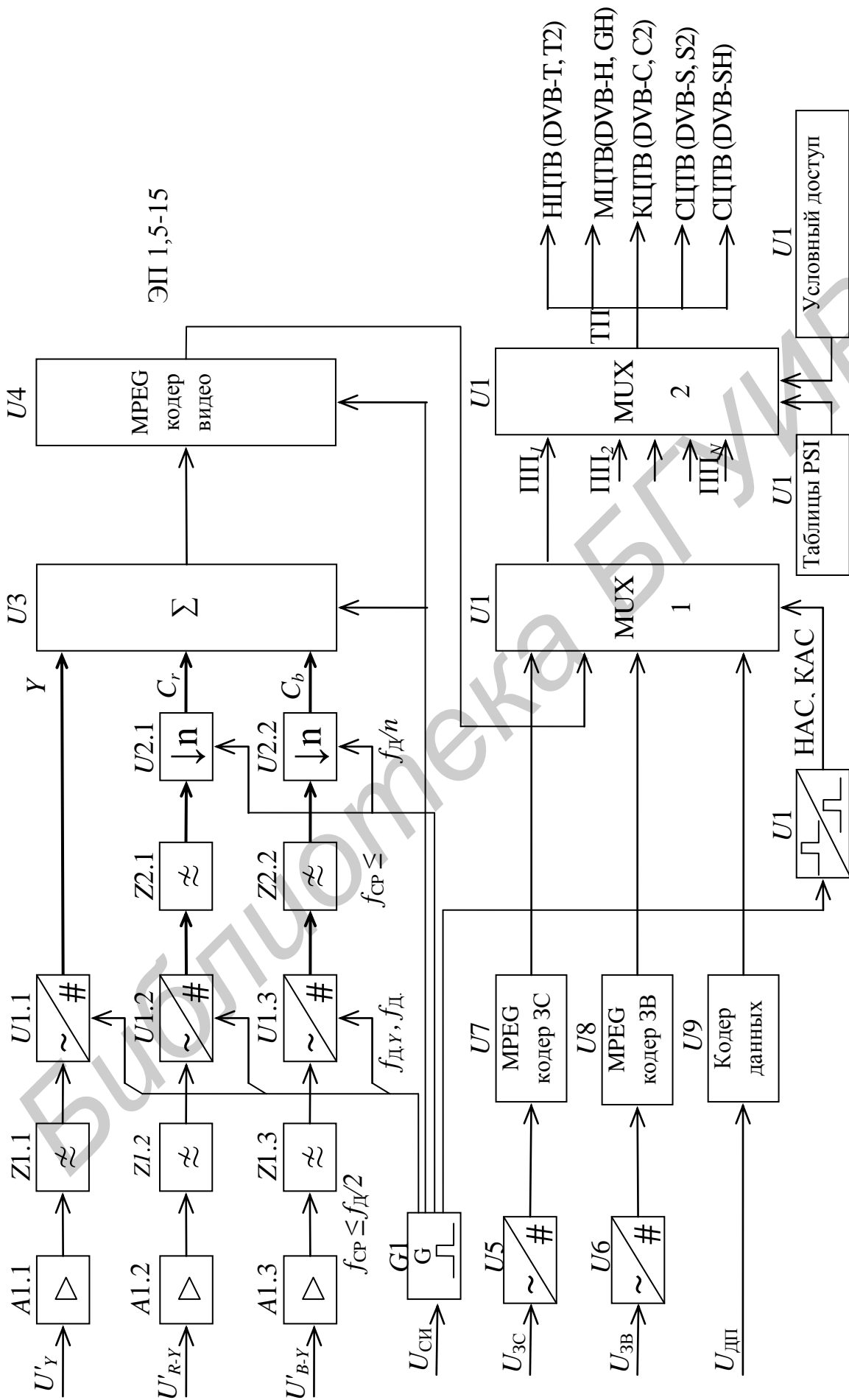


Рисунок 1.6 – Обобщенная структурная схема формирования цифровых ТВ-сигналов

Тогда в четвертом режиме $f_{д.у} = 13,5$ МГц, а $f_{д.црс} = f_{д.у}/4 = 3,375$ МГц, во втором и третьем $f_{д.црс} = f_{д.у}/2 = 6,75$ МГц, а в первом $f_{д}$ одинакова для трех сигналов и равна 13,5 МГц. С другой – условные обозначения указывают на соотношение значений полос частот аналоговых сигналов СЯ и ЦРС (таблица 1.7).

Как следует из таблицы 1.7 и рисунка 1.7, в формате 4:2:2 происходит снижение цветовой чёткости по горизонтали в 2 раза по сравнению с эталонным форматом 4:4:4, а в 4:1:1 – в 4 раза. Учитывая свойства зрения, для большинства сюжетов это не очень заметно.

Таблица 1.7 – Основные параметры сигналов ЦТВ при различных форматах

Форматы	4:4:4	4:2:2	4:2:0	4:1:1
$\Delta F_{у}$, МГц	5,75	5,75	5,75	5,75
$\Delta F_{црс}$, МГц	5,75	2,75(3)	2,75(3)	1,375(1,5)
$f_{д.у}$, МГц	13,5	13,5	13,5	13,5
$f_{д.црс}$, МГц	13,5	6,75	6,75	3,375
B_{Σ} , Мбит/с, $m = 10$	405	270	202,5	202,5

В формате 4:2:0 отсчёты ЦРС прорежены в 2 раза не только по горизонтали, но и по вертикали. Этот формат предпочтительнее для систем с 625 строками и полосой частот СЯ 5,0; 5,5 или 6 МГц, а 4:1:1 – для систем с 525 строками и $\Delta F_{у} = 4,2$ МГц [5].

Структура расположения отсчётов в пространстве кадра (рисунок 1.7) наглядно поясняет принятое обозначение форматов.

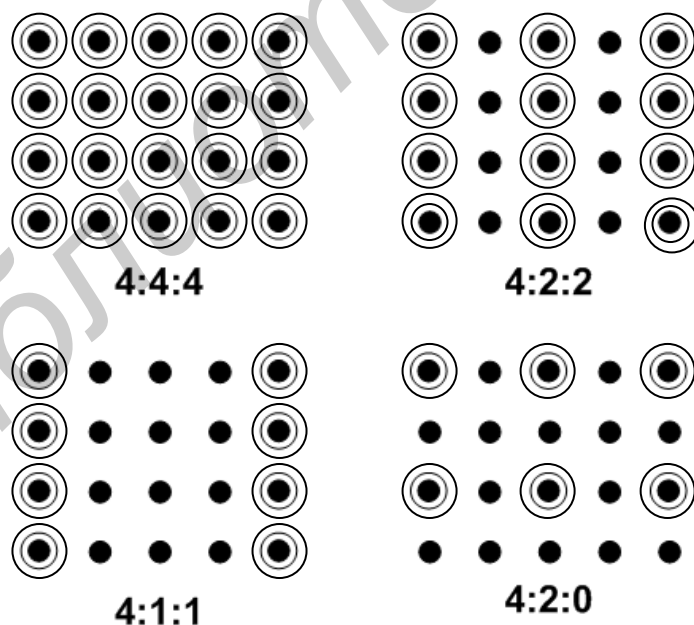
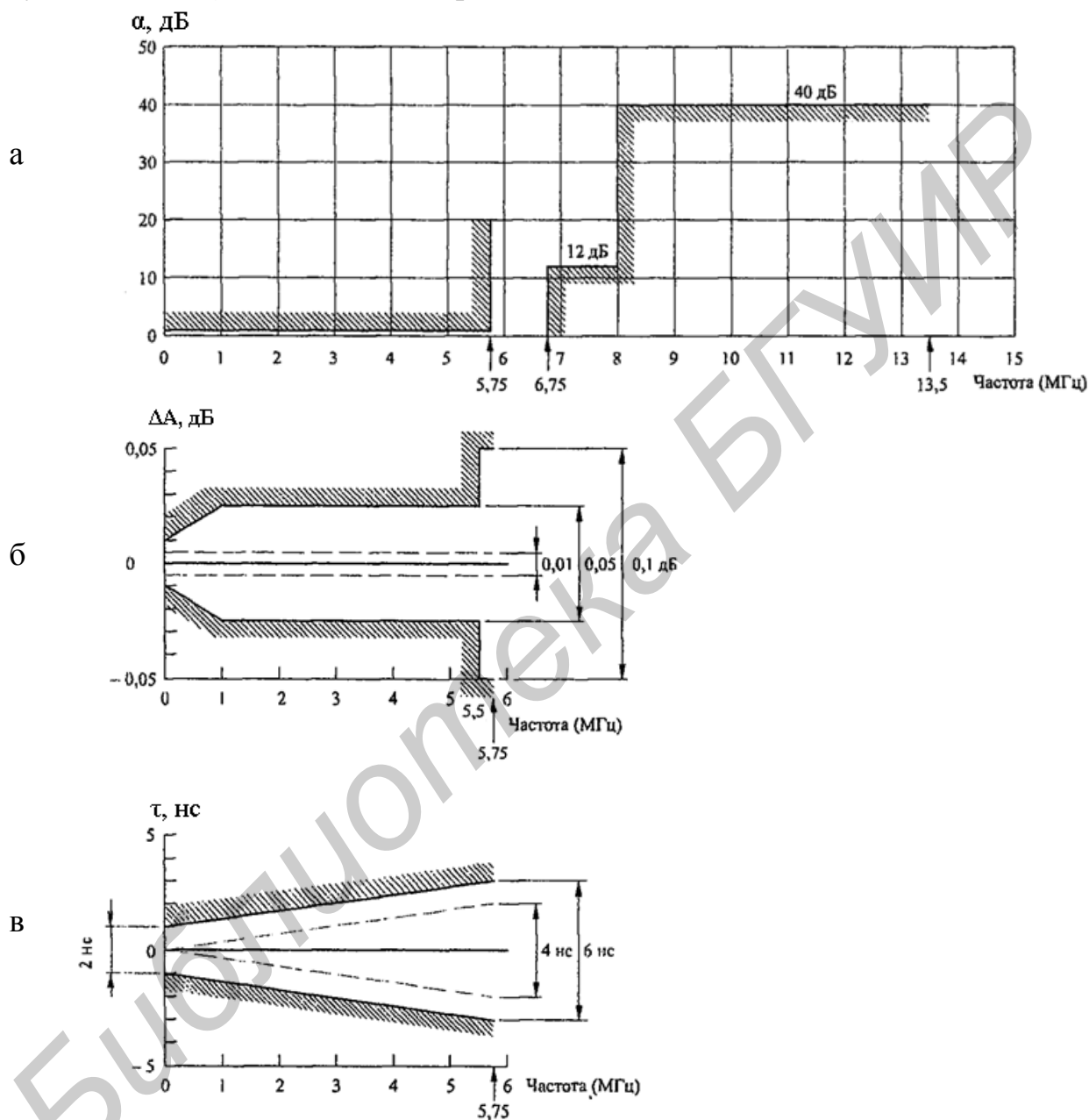


Рисунок 1.7 – Пространственная структура расположения отсчётов – СЯ «●» и двух ЦРС «○»

В учебном пособии авторов [5] были детально проанализированы процессы, связанные с аналого-цифровым преобразованием ТВ сигнала: выбор

численного значения частоты дискретизации, количества уровней квантования, защищённости от шумов квантования, требования к отношению сигнал/шум квантования в зависимости от отношения сигнал/шум датчика ТВ сигнала. Поэтому здесь эти вопросы не рассматриваются.

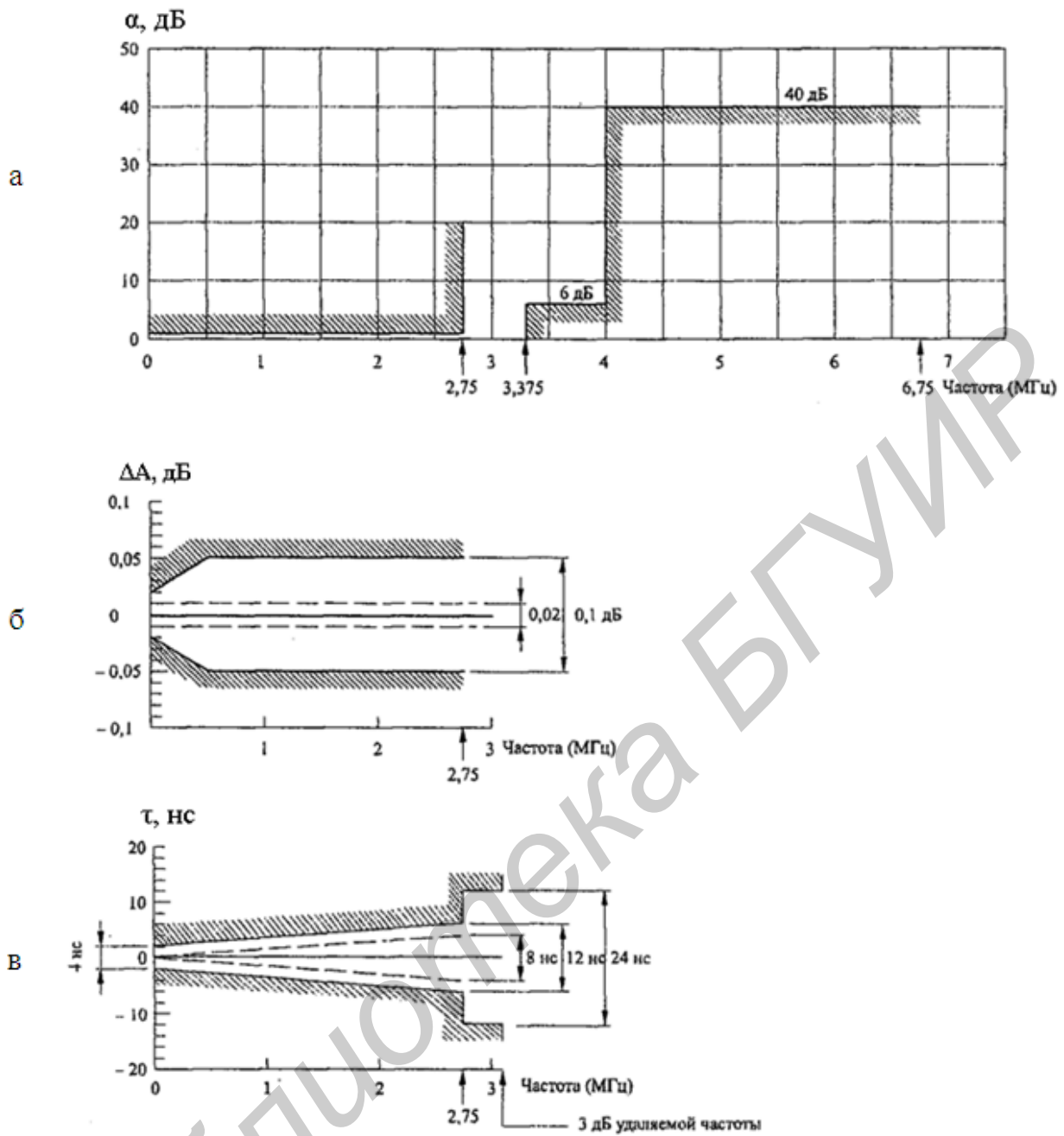
С помощью ФНЧ Z1.1...Z1.3 сигналы ограничиваются на полосе (рисунки 1.8 и 1.9) и подаются на три АЦП U1.1...U1.3.



Примечание – Наименьшая частота в (б) и (в) равна 1 кГц (вместо 0 МГц)

а – характеристика затухания; б – характеристика неравномерности АЧХ в полосе пропускания; в – ГВЗ для сигналов U_R^y , U_G^y , U_B^y или Y , U_{Cr} , U_{Cb} в режиме 4:4:4

Рисунок 1.8 – Требования к характеристикам пред- и постфильтров нижних частот



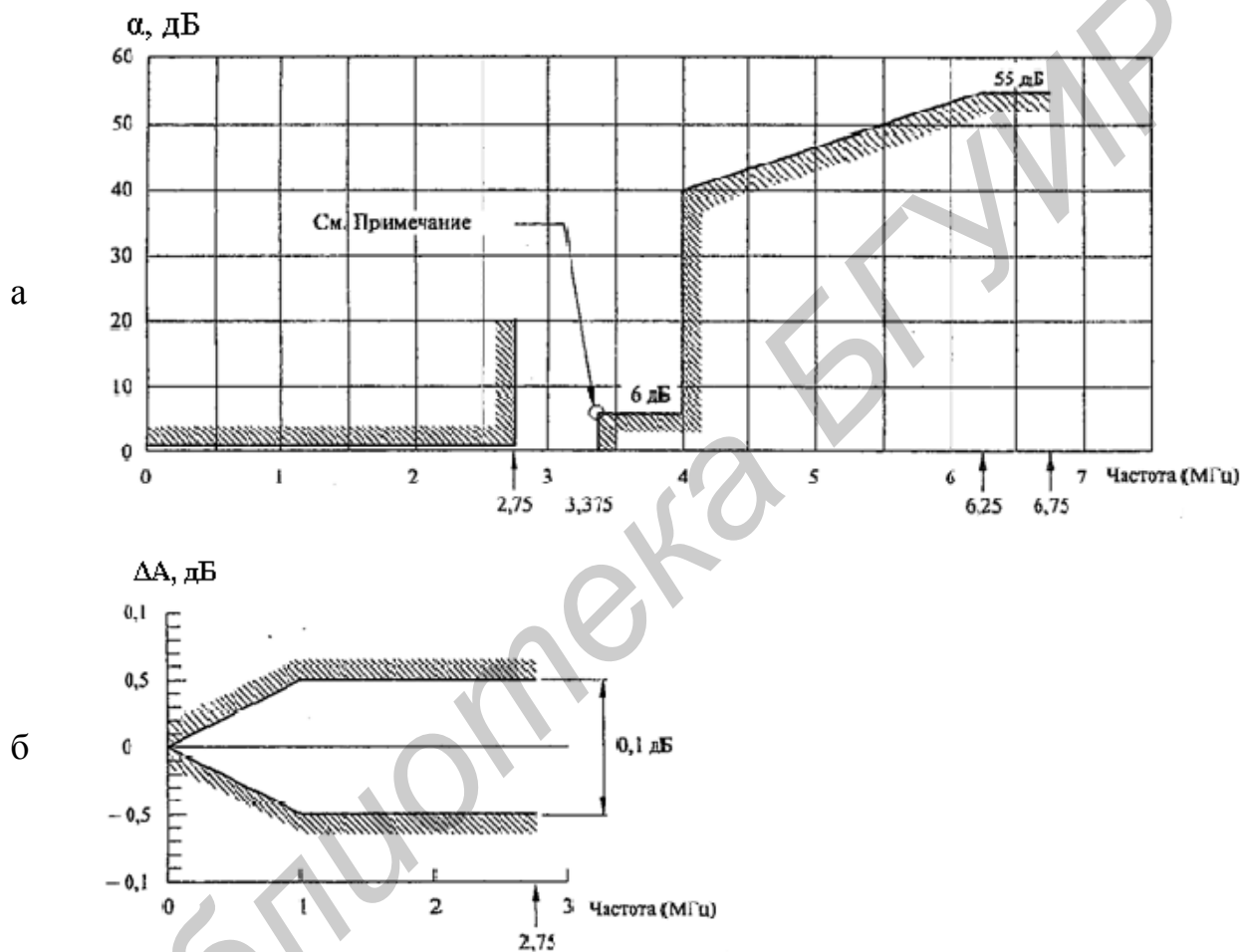
Примечание – Наименьшая частота в (б) и (в) равна 1 кГц (вместо 0 МГц)

а – характеристика затухания; б – характеристика неравномерности АЧХ в полосе пропускания; в – ГВЗ для цветоразностных сигналов в режиме 4:2:2 и 4:2:0

Рисунок 1.9 – Требования к характеристикам пред-и постфильтров нижних частот

Предфильтры устанавливаются на входе АЦП, а постфильтры – на выходе ЦАП. С целью упрощения ФНЧ спад его АЧХ (рисунок 1.8, а) начинается на частоте 5,75 МГц (в отличие от 6 МГц в аналоговой системе). Допуски на неравномерность АЧХ и характеристики группового времени задержки (ГВЗ) в полосе пропускания приведены на рисунке 1.8, б; в (с учетом фазового корректора, включенного на выходе ФНЧ).

Три АЦП могут работать с одинаковыми параметрами (частотой дискретизации f_d и разностью двоичного кодирования m). Поскольку полоса частот ЦРС в 2 (форматы 4:2:2 и 4:2:0) или 4 (формат 4:1:1) раза меньше полосы частот СЯ, для них выполняется операция прореживания (децимация, субдискретизация, т. е. передискретизация с уменьшением числа отсчетов в единицу времени) в 2 (или 4) раза соответственно в блоках $U2.1$ и $U2.2$. Назначение цифровых фильтров $Z2.1$ и $Z2.2$ (рисунок 1.10) аналогично $Z1.2$ и $Z1.3$. Прореживание сокращает суммарную скорость B_{Σ} (см. таблицу 1.7) потока данных в 1,5 – 2 раза соответственно по сравнению с форматом 4:4:4.



Примечание – В цифровом фильтре АЧХ (в линейном масштабе) должна иметь наклон, симметричный относительно точки половинной амплитуды, как показано на рисунке 1.10

а – характеристика затухания; б – характеристика неравномерности АЧХ в полосе пропускания; в – характеристики цифрового фильтра для преобразователя частоты дискретизации цветоразностных сигналов из 4:4:4 в 4:2:2

Рисунок 1.10 – Требования к характеристикам

1.6 Параметры цифровых компонентных сигналов

Компонентные цифровые сигналы (т. е. после АЦП) принято обозначать Y , C_r , C_b . В рекомендации МСЭ-Р [30] для записи кодовых слов используется двоичная, десятичная (d) и шестнадцатеричная (h) формы записи. Первоначально в ЦТВ применялось квантование на 256 ($M = 0, 1, \dots, 255$) уровней и восьмиразрядное кодирование ($m = 8$). Потом в связи с выпуском ТВ камер с отношением сигнал/шум 60 – 65 дБ возникла необходимость перехода на $M = 1024$ и соответственно на 10-разрядное кодирование. Чтобы избежать возможной путаницы между 8-разрядным и 10-разрядным представлением в Рекомендации МСЭ-Р [30] уровни квантования приведены в дробно-десятичном представлении (см. рисунок 1.5) и соответствуют 8- и 10-разрядному кодированию. При этом восемь старших разрядов (до десятичной точки) записываются как целая часть числа (при $m = 8$), а два дополнительных бита (цифры после десятичной точки), если они есть, записываются как его дробная часть, т. е. являются двумя младшими разрядами с шагом 0,25. Если дробная часть отсутствует, принимается, что её двоичное значение равно нулю.

$$\text{Например, число } 165,75d = \underbrace{10100101,11}_m=8,2; \quad \underbrace{\hspace{10em}}_m=10$$

10100101,00 при $m = 8$ есть 165 уровень,

10100101,00 при $m = 10 \rightarrow 165 \times 4 = 660$ уровень,

10100101,11 при $m = 8 \rightarrow 165,75$ уровень, что указывает на то, что при $m = 8$ этот уровень ближе к 166-му, таким он и будет в 8-разрядном АЦП.

10100101,11 при $m = 10 \rightarrow 165,75 \times 4 = 663$ уровень.

В рекомендации МСЭ-Р [30] уровни $0,00d$ и $255,75d$ зарезервированы для передачи данных синхронизации, уровни от $1,00d$ до $254,75d$ используются для видеосигнала. В соответствии с этой рекомендацией СЯ должен занимать только 220 8-разрядных ($m = 8$) или 877 10-разрядных ($m = 10$) уровней, а уровню чёрного соответствовать значение $16,00d$. При этом десятичные значения квантованного сигнала яркости Y вычисляются по выражению

$$Y = \text{int}\{(219U'_Y + 16)D\}/D, \quad (1.25)$$

где D принимает значения 1 или 4 при $m = 8$ и $m = 10$ соответственно.

Оператор $\text{int}()$ в выражении (1.23) присваивает значение 0 дробным частям в диапазоне от 0 до 0,4999 и +1 для дробных частей в диапазоне от 0,5 до 0,999, т. е. он округляет в большую сторону дробные части, превышающие 0,5.

Учитывая, что ЦРС должны занимать 255 8-разрядных или 897 10-разрядных уровней и что нулевым должен быть уровень $128,00d$

(поскольку ЦРС двуполярны, см. рисунок 1.5), десятичные значения квантованных ЦРС C_R и C_B вычисляются по выражениям

$$C_R = \text{int}\{(224U'_{CR} + 128) \times D\}/D, \quad (1.26)$$

$$C_B = \text{int}\{(224U'_{CB} + 128) \times D\}/D. \quad (1.27)$$

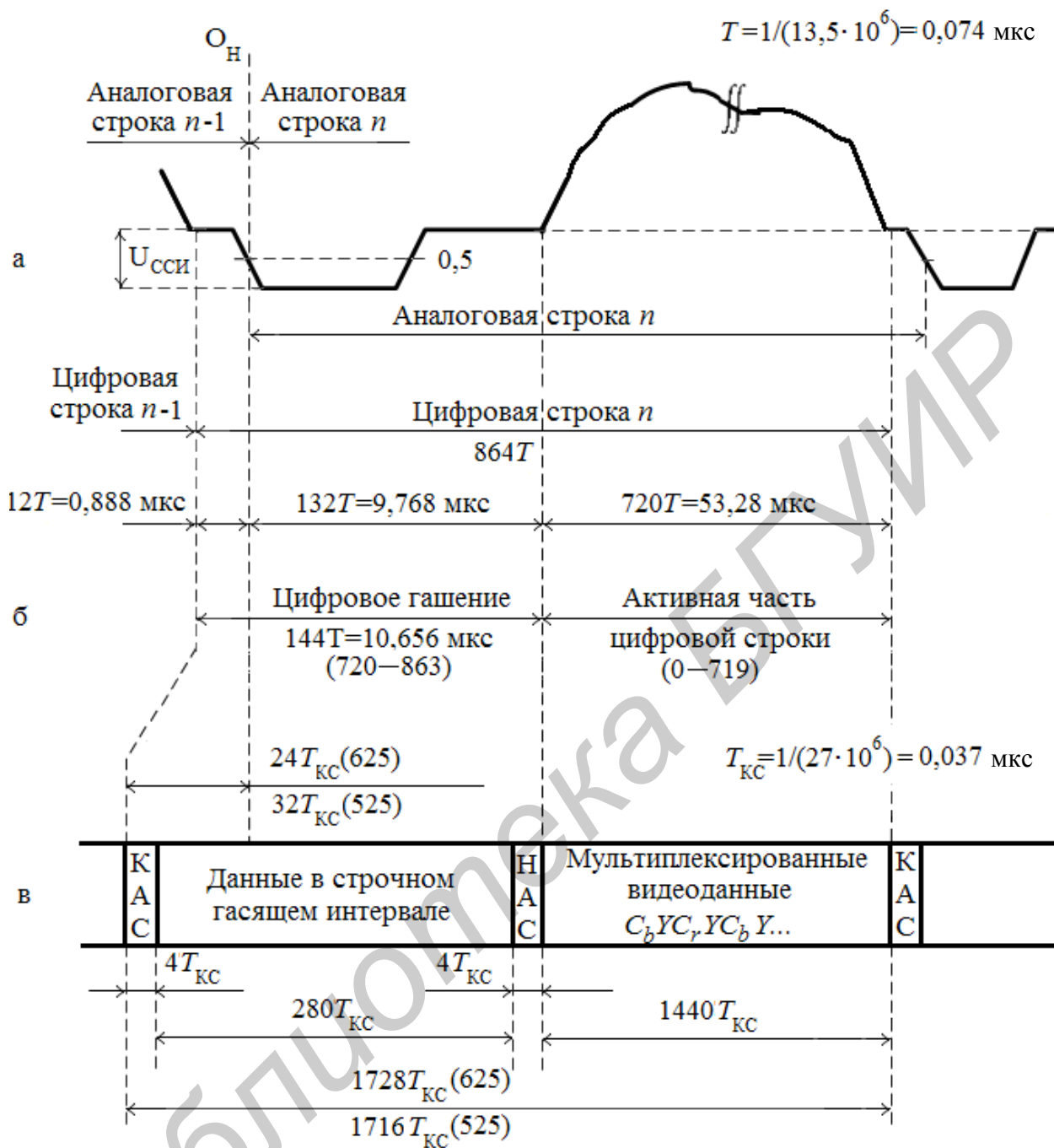
Чтобы от уровней в дробно-десятичном представлении (на рисунке 1.5 справа) перейти к разрешенным при $m = 8$, необходимо дробные значения, равные и превышающие 0,5, округлить до целого значения, а меньшие 0,5 – до нуля. Для перехода к разрешенным уровням при $m = 10$ необходимо дробно-десятичные значения уровней умножить на 4.

Значения основных параметров кодирования СЯ и ЦРС для ТСЧ в форматах 4:4:4 и 4:2:2 приведены в таблицах 1.8 и 1.9 соответственно. Значения основных параметров цифрового представления компонентных сигналов ТВЧ при различных видах развёртки приведены в таблице 1.10.

Временные соотношения между аналоговой и цифровой строками приведены на рисунке 1.11. В режиме 4:2:2 (рисунок 1.11) данные передаются в мультиплексированном виде со скоростью 270 Мбит/с: десятиразрядные кодовые слова следуют с периодом $T_{KC} = 1/(27 \cdot 10^6) = 0,037 \dots$ мкс, т. е. в интервале $T_{дУ}$ нужно представить четыре кодовых слова. Частота дискретизации является 864-й (или 858-й) гармоникой частоты строк для стандартов 625/50 (или 525/59,94), поэтому в полной цифровой строке будет 864 (858) отсчётов сигнала яркости (таблицы 1.8 и 1.9). При этом в активной части строки обоих стандартов используется одинаковое число отсчётов 720 (0 – 719).

Активная часть цифровой строки образует блок из 1440 кодовых слов $(720Y + 360C_b + 360C_r)$ мультиплексированных видеоданных в следующей последовательности: $C_b, Y, C_r, Y \dots$. Группа C_b, Y, C_r представляет собой значения (в двоичном коде) сигнала яркости и цветоразностных сигналов одного пикселя (элемента изображения); для следующего за ним элемента передается только значение сигнала яркости (без ЦРС).

В течение цифровой строки передаётся два опорных синхросигнала TRS (Timing Reference Signals) (рисунки 1.11 и 1.12: один – перед началом каждой активной строки – НАС (Start of Active Video, SAV), т. е. в конце строчного гасящего интервала, а другой – после окончания каждой активной строки – конец активной строки – КАС (End of Active Video, EAV). Каждый из синхросигналов НАС и КАС состоит из последовательности четырёх 10-разрядных ($m = 10$) кодовых слов (КС) длительностью $4T_{KC}$ (рисунок 1.13). Первое состоит из 10 двоичных единиц, что соответствует десятичному числу 1023 (при $m = 8$ будет 255).



а – аналоговая строка; б – цифровая строка; в – мультимплексированные данные (в режиме 4:2:2)

Рисунок 1.11 – Временные соотношения для сигналов ТСЧ

Таблица 1.8 – Значения основных параметров кодирования для формата 4:4:4

Параметры	Системы с 625/50i	Системы с 525/60i
1 Кодированные сигналы: Y, C_r, C_b или R, G, B	Получены из гамма-корректированных сигналов: U'_Y, U'_{R-Y}, U'_{B-Y} или U'_R, U'_G, U'_B	
2 Число отсчетов в полной строке для каждого сигнала	864	858
3 Структура дискретизации	Ортогональная для строк, полей и кадров. Три структуры дискретизации должны быть совмещены друг с другом и со структурой дискретизации сигнала яркости	
4 Частота дискретизации для каждого сигнала	13,5 МГц	
5 Вид кодирования	Равномерная ИКМ, 8 или 10 бит на отсчет	
6 Число отсчетов в активной части строки	По 720 для каждого сигнала	
7 Связь между аналоговой и цифровой синхронизацией строк	12 периодов синхронизации	16 периодов синхронизации
8 Соответствие между уровнями видеосигнала и квантования: – шкала; – R, G, B или Y ; – каждого ЦРС	<p>От $0,00d$ до $255,75d$ 220 (8-разрядных) или 877 (10-разрядных) уровней квантования с уровнем черного, соответствующим уровню $16,00d$ и пиком уровня белого, соответствующего уровню $235,00d$. Уровень сигнала может иногда превышать уровень $235,00d$ или быть ниже уровня $16,00d$.</p> <p>225 (8-разрядных) или 897 (10-разрядных) уровней квантования; в центральной части шкалы квантования нулевому сигналу соответствует уровень $128,00d$. Уровень сигнала может иногда превышать уровень $240,00d$ или быть ниже уровня $16,00d$</p>	
9 Используемые кодовые слова	Кодовые слова, соответствующие уровням квантования $0,00d$ и $255,75d$, используются только для синхронизации. Уровни с $1,00d$ по $254,75d$ используются для видеосигнала. Если 8-разрядные слова интерпретируются в 10-разрядные, к ним нужно добавить два младших, равных нулю, разряда	

Таблица 1.9 – Значения основных параметров кодирования для формата 4:2:2

Параметры	Системы с 625/50i	Системы с 525/60i
1 Кодированные сигналы: Y, C_r, C_b	Получены из гамма-корректированных сигналов: U'_Y, U'_{R-Y}, U'_{B-Y} .	
2 Число отсчетов в полной строке: – сигнала яркости, Y ; – каждого ЦРС (C_r, C_b)	864 423	858 429
3 Структура дискретизации	Ортогональная для строк, полей и кадров. Отсчеты C_r, C_b пространственно совмещены с нечетными отсчетами (1-й, 3-й, 5-й и т. д.) сигнала Y в каждой строке	
4 Частота дискретизации: – сигнала яркости; – каждого ЦРС	13,5 МГц 6,75 МГц	
5 Вид кодирования	Равномерная ИКМ, 8 или 10 разрядов на отсчет для сигнала яркости и каждого ЦРС	
6 Число отсчетов в активной строке: – сигнала яркости; – каждого ЦРС	720 360	
7 Связь между аналоговой и цифровой синхронизацией строк от конца активной цифровой строки до момента O_H	12 периодов синхронизации	16 периодов синхронизации
8 Соответствие между уровнями видеосигнала и квантования: – шкала; – сигнал яркости; – каждый ЦРС	От $0,00d$ до $255,75d$ 220 (8-разрядных) или 877 (10-разрядных) уровней квантования с уровнем черного, соответствующим уровню $16,00d$ и пиком уровня белого соответствующего уровню $235,00d$. Уровень сигнала может иногда превышать уровень $235,00d$ или быть ниже уровня $16,00d$. 225 (8-разрядных) или 897 (10-разрядных) уровней квантования; в центральной части шкалы квантования нулевому значению ЦРС соответствует уровень $128,00d$. Уровень сигнала может иногда превышать уровень $240,00d$ или быть ниже уровня $16,00d$	
9 Используемые кодовые слова	Кодовые слова, соответствующие уровням квантования $0,00d$ и $255,75d$ используются только для синхронизации. Уровни с $1,00d$ по $254,75d$ используются для видеосигнала. Если 8-разрядные слова интерпретируются в 10-разрядные, к ним нужно добавить два младших, равных нулю, разряда	

Таблица 1.10 – Значения основных параметров цифрового представления сигналов ТВЧ

Параметр	Значения, обозначающие системы									
	60p	30p	30psf	60i	50p	25p	25psf	50i	24p	24psf
1 Порядок представления отсчетов в системе с раз- верткой	Слева направо, сверху вниз В чересстрочных системах и системах с сегментным кадром первая активная строка поля 1 расположена сверху изображения									
2 Полное число строк	1125									
3 Частота полей/ кад- ров/сегментных кадров, Гц	60 (60/1,001)	30 (30/1,001)	60 (60/1,001)	50	25	50	24 (24/1,001)	48 (48/1,001)		
4 Коэффициент чересстрочности	1:1	2:1	1:1	1:1	2:1	2:1	1:1	1:1		
5 Частота кадров, Гц	60 (60/1,001)	30 (30/1,001)	50	25	25	24 (24/1,001)	27000 (27000/1,001)			
6 Частота строк ¹ , Гц	67500 (67500/1,001)	33750 (33750/1,001)	56250	28125						
7 Число отсчетов на пол- ную строку: – R, G, B, Y; – C _r , C _b	2200 1100	2640 1320								
8 Номинальная полоса аналоговых сигналов ² , МГц	60	30	60	30						
9 Частота дискретизации – R, G, B, Y, МГц	148,5 (148,5/1,001)	74,25 (74,25/1,001)	148,5	74,25						
10 Частота дискретизации ³ – C _r , C _b , МГц	74,25 (74,25/1,001)	37,125 (37,125/1,001)	74,25	37,125						

¹ допуск на частоты ± 0,001%;

² полоса частот для всех компонентов;

³ частота дискретизации сигналов C_B, C_R равна половине частоты дискретизации сигнала яркости

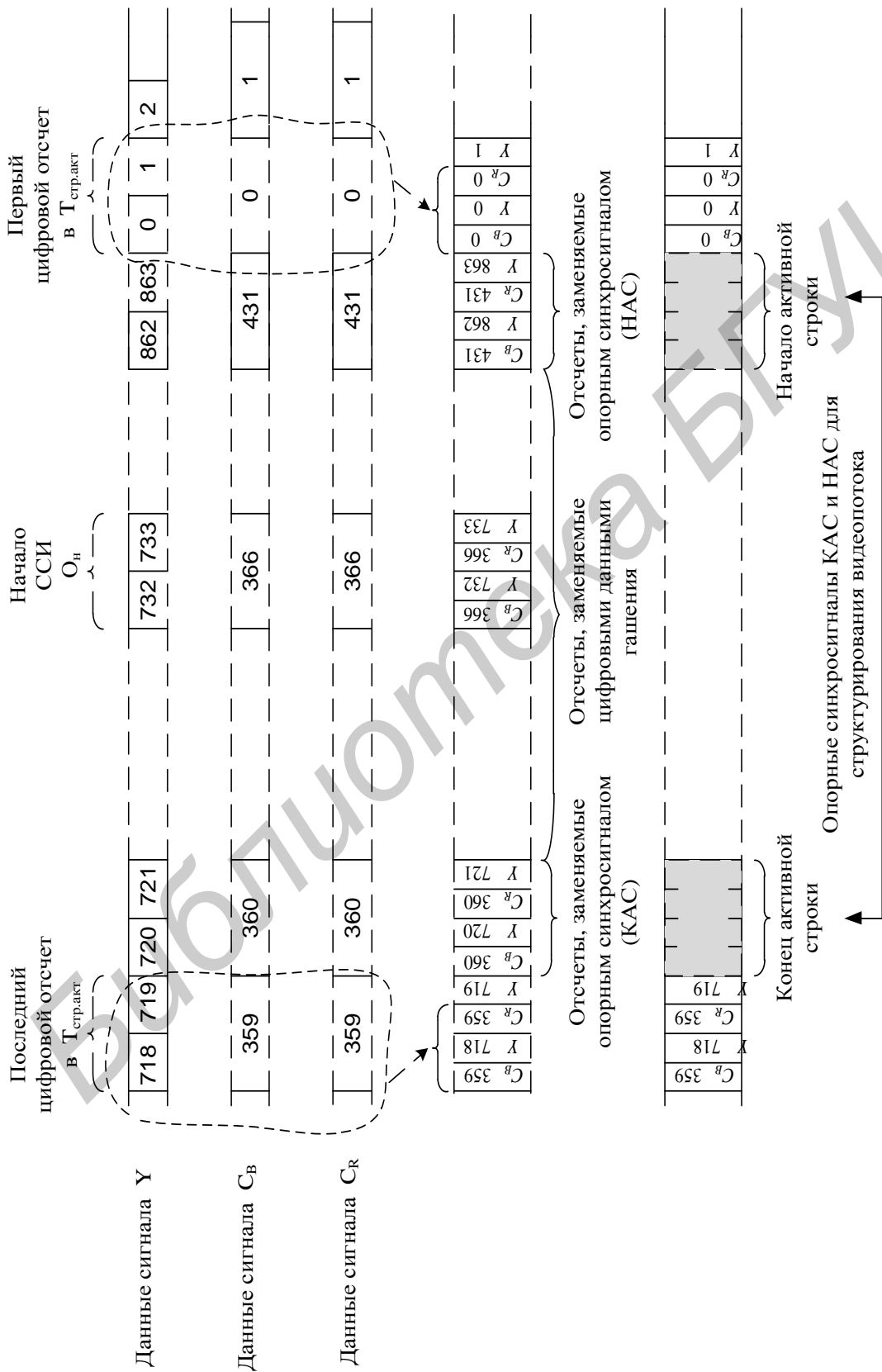


Рисунок 1.12 – Состав сигналов Y , S_v , S_r и мультиплексированные данные в интервале СГИ

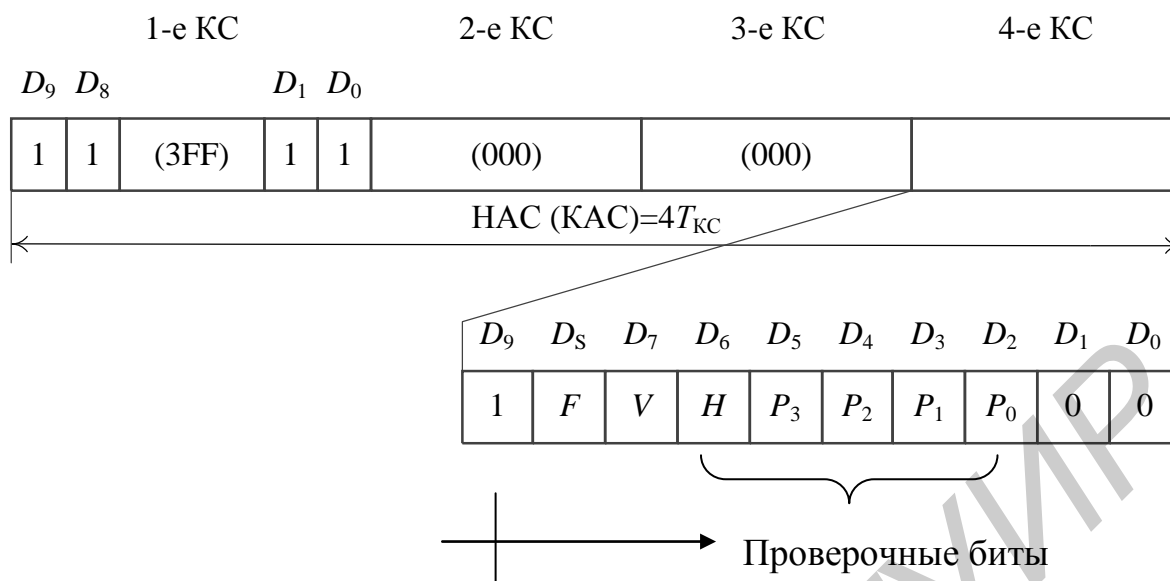


Рисунок 1.13 – Состав и значения опорных синхросигналов НАС и КАС

Следующие два КС равны нулю. Назначение двоичных разрядов последнего (четвертого) КС поясняется на рисунке 1.13 справа. Сигнал F идентифицирует номер поля. Его значение $F = 0$ соответствует первому полю (строка номер 1 по стандарту [6]), а $F = 1$ (строка 313) – второму полю. Сигнал V определяет начало и конец цифрового интервала гашения поля. В первом поле его значение $V = 1$ соответствует началу (строка 624), а $V = 0$ (строка 23) – концу интервала гашения; во втором поле $V = 1$ (строка 311) и $V = 0$ (строка 336). Сигнал H принимает значения $H = 0$ для НАС и $H = 1$ для КАС. В соответствии с [30] номер цифровой строки изменяется до момента 0_H (см. рисунки 1.11, 1.12).

Для обеспечения совместимости с существующими 8-разрядными интерфейсами, значения разрядов D_0 и D_1 не определены. Значения проверочных битов $P_0... P_3$, зависящие от значения F , V и H , приведены в таблице 1.11. В последних двух столбцах этой таблицы даны десятичная и шестнадцатеричная записи числа, передаваемого в четвертом КС синхросигнала, которое может иметь восемь допустимых значений. Проверочные биты позволяют исправить в приемнике однобитовые ошибки и обнаруживать двубитовые.

В интервалах от КАС до НАС видеоинформация не передается. Эти интервалы заполняются либо последовательностью уровней гашения; 128,00, 16,00, 128,00, 16,00 и т. д., либо какими-то дополнительными данными произвольного типа (например, преобразованными в цифровую форму сигналами звукового сопровождения).

Таблица 1.11 – Значения проверочных битов

D9	F	V	H	P3	P2	P1	P0	Шестнадцатеричная запись	Десятичная запись
1	0	0	0	0	0	0	0	80	128
1	0	0	1	1	1	0	1	9D	157
1	0	1	0	1	0	1	1	AB	171
1	0	1	1	0	1	1	0	B6	182
1	1	0	0	0	1	1	1	C7	199
1	1	0	1	1	0	1	0	DA	218
1	1	1	0	1	1	0	0	EC	236
1	1	1	1	0	0	0	1	F1	241

Сформированные компонентные цифровые сигналы яркости Y и цветоразностные C_R и C_B подаются на блок $U3$ предварительной обработки (БПО на рисунке 1.6), в котором осуществляются операции по временному уплотнению сигналов в форматах 4:2:2, 4:2:0 или 4:1:1.

В качестве примера рассмотрим процесс уплотнения в последовательный поток данных цифровых компонентных сигналов формата 4:2:2 (рисунок 1.14). На рисунке приведены несколько кодовых слов СЯ: Y_0, Y_1, Y_2, Y_3 , ЦРС C_{B0}, C_{B1} и ЦРС C_{R0}, C_{R1} , отдельно поступающих на блок БПО (рисунок 1.14, а, б, в). Длительность кодовых слов (КС) определяется частотой дискретизации и для формата 4:2:2 равна 74,(074) нс (Y) и 148,(148) нс (C_R, C_B). При их временном уплотнении кодовые слова должны иметь длительность 37,(037) нс (рисунок 1.14, г). Порядок следования КС определен рекомендацией МСЭ-Р [30]: C_{B0}, Y_0, C_{R0}, Y_1 и т. д. Следовательно, КС Y_0 сигнала яркости в БПО сжимается по времени в 2 раза и задерживается на 37 нс. Кодовое слово C_{B0} сжимается в 4 раза, во столько же раз сжимается кодовое слово C_{R0} и задерживается на 74,(074) нс. Длительность тактового интервала в уплотненном потоке равна 3,7(037) нс, что соответствует скорости 270 Мбит/с.

Аналогичным способом производится мультиплексирование и для формата 4:2:0 (рисунок 1.15), в котором последовательный цифровой поток SDI образуется в n -й строке путем временного уплотнения цифровых сигналов Y и C_B , а в $n+1$ -й – Y и C_R .

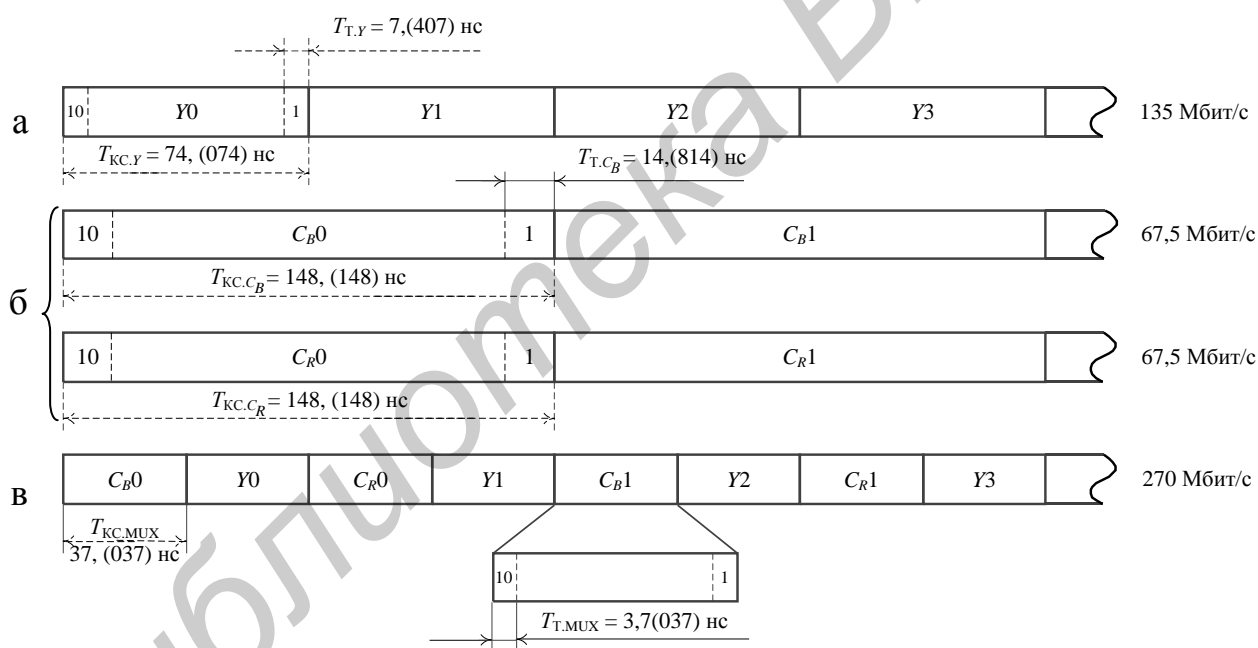
После MPEG кодирования в блоке $U4$ образуется элементарный цифровой поток со скоростью от 1,5 до 15 Мбит/с (в зависимости от конкретного назначения системы ЦТВ), который подается на первый мультиплексор $U5$. На него поступают также цифровые стереофонические сигналы звукового сопровождения, звукового вещания, дополнительных данных, а также синхросигналы НАС и КАС.

На выходе мультиплексора $U5$ образуется программный поток, который уплотняется с другими программными потоками во втором мультиплек-

соре $U12$, на который подаются данные таблиц сервисной информации (PSI) и условного доступа. Скорость транспортного потока (ТП) варьируется от 30 до 60 Мбит/с в зависимости от назначения системы (НЦТВ, МЦТВ, КЦТВ, СЦТВ) и допустимого числа уплотняемых программ. Полученный транспортный поток в соответствии со схемой на рисунке 1.1 подается на кодер канала, а затем на модулятор соответствующей системы ЦТВ.

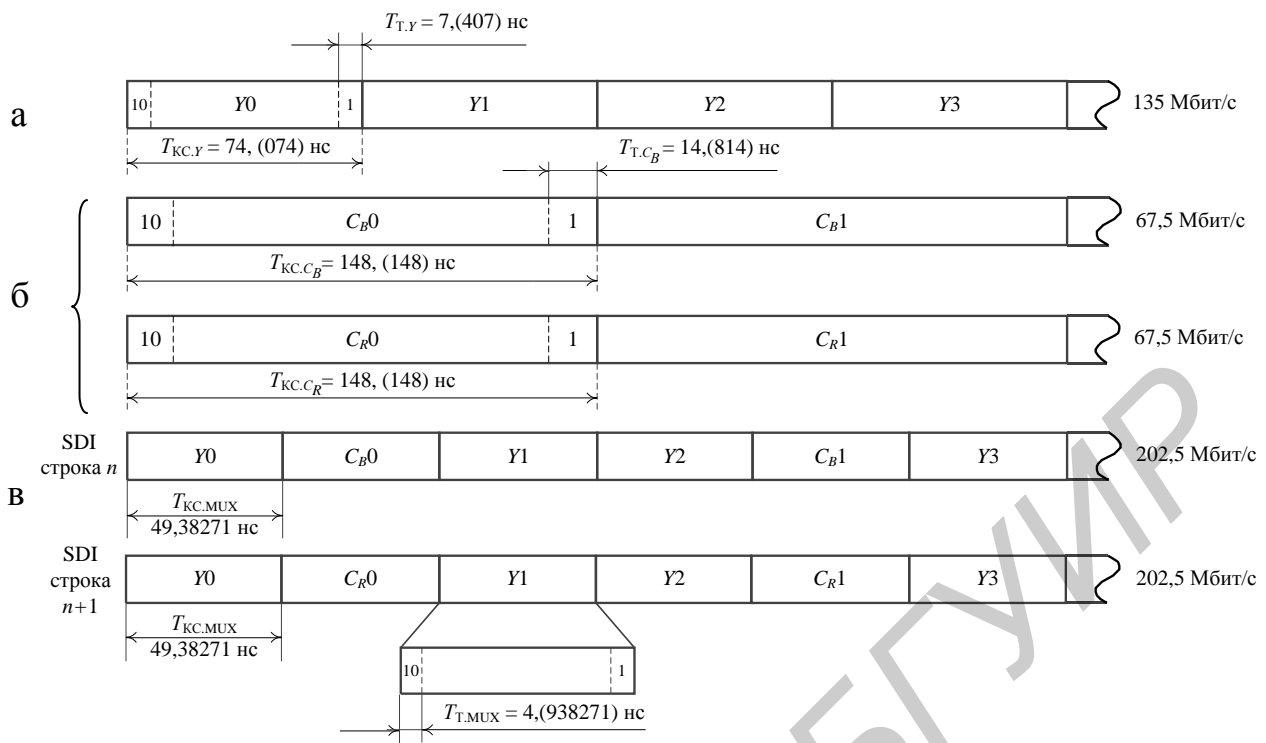
В ТВЧ аналоговый ТВ сигнал приведен на рисунке 1.16, а. Отличительной особенностью этого сигнала является форма синхросигнала в виде трехуровневого биполярного импульса. Ниже показано расположение цифровых строк: сигнала Y (рисунок 1.16, б); точно таким же образом могут располагаться три параллельных потока цифровых R, G, B сигналов. Уплотненные во времени цифровые ЦРС C_B и C_R приведены на рисунке 1.16, в.

По аналогии с мультиплексированием цифровых сигналов ТСЧ осуществляется временное уплотнение цифровых сигналов ТВЧ. В качестве примера на рисунке 1.17 приведено временное уплотнение компонентных цифровых сигналов в формате 4:2:2.



а – сигнал яркости Y ; б – цифровые ЦРС C_B, C_R ; в – последовательный поток данных

Рисунок 1.14 – Временное уплотнение компонентных цифровых сигналов в формате 4:2:2 (ТСЧ)



а – сигнал яркости Y ; б – цифровые ЦРС C_B, C_R ; в – SDI-поток

Рисунок 1.15 – Временное уплотнение компонентных цифровых сигналов в формате 4:2:0 (ТСЧ)

В системах ТВЧ в соответствии с Рекомендацией ITU-R BT.1361 [31] преобразование сигналов основных цветов U'_R, U'_G, U'_B в U'_Y, U'_{R-Y}, U'_{B-Y} выполняется по следующим формулам:

$$\begin{aligned}
 U'_Y &= 0,2126U'_R + 0,7152U'_G + 0,0722U'_B; \\
 U'_{R-Y} &= 0,7874U'_R - 0,7152U'_G - 0,0722U'_B; \\
 U'_{B-Y} &= -0,2126U'_R - 0,7152U'_G + 0,9278U'_B.
 \end{aligned}
 \tag{1.28}$$

Относительные значения аналоговых сигналов основных цветов, СЯ и ЦРС при передаче вертикальных цветных полос с номенклатурой 100/0/100/0, рассчитанных по выражению (1.21), приведены в таблице 1.7. Из нее следует, что сигналы U'_R, U'_G, U'_B и U'_Y (без синхроимпульсов) принимают значения от 0 до 1. Как было отмечено выше, указанные сигналы на нагрузке 75 Ом создают напряжение 700 мВ, а вместе с синхросигналом (в одном из сигналов, например U'_Y) – 1 В (1000 мВ).

Из таблицы 1.12 следует, что относительные значения U'_{R-Y} находятся в пределах от минус 0,7874 до +0,7874, а U'_{B-Y} – от минус 0,9278 до +0,9278. Поскольку сигналы U'_{R-Y} и U'_{B-Y} являются двуполярными, их значения необ-

ходимо нормализовать так, чтобы размах каждого равнялся размаху СЯ (1,0 отн. ед., т. е. 700 мВ).

Таблица 1.12 – Относительные значения сигналов основных цветов, СЯ и ЦРС

Состояние	U'_R	U'_G	U'_B	U'_Y	U'_{R-Y}	U'_{B-Y}	P_r	P_b
Белый	1,0	1,0	1,0	1,000	0,000	0,0000	0,0000	0,0000
Черный	0,0	0,0	0,0	0,000	0,000	0,0000	0,0000	0,0000
Красный	1,0	0,0	0,0	0,2126	0,7874	-0,2126	0,5000	0,1145
Зеленый	0,0	1,0	0,0	0,7152	-0,7152	-0,7152	-0,4541	-0,3854
Синий	0,0	0,0	1,0	0,0722	-0,0722	0,9278	-0,0458	-0,5000
Голубой	0,0	1,0	1,0	0,7874	-0,7874	0,2126	-0,5000	-0,1145
Пурпурный	1,0	0,0	1,0	0,2848	0,7152	0,7152	0,4541	0,3854
Желтый	1,0	1,0	0,0	0,9278	0,0722	-0,9278	0,0458	0,5000

Для этого они должны изменяться от минус 0,5 (350мВ) до +0,5 (+350 мВ), т. е.

$$P_r = |U_{R-Y}|_{\max} / k_r = 0,5; P_b = |U_{B-Y}|_{\max} / k_b = 0,5. \quad (1.29)$$

Учитывая (1.29) и данные таблицы 1.12, определяем численные значения коэффициентов k_r и k_b :

$$k_r = |U'_{R-Y}|_{\max} / 0,5 = 0,7874 / 0,5 = 1,5748; \quad (1.30)$$

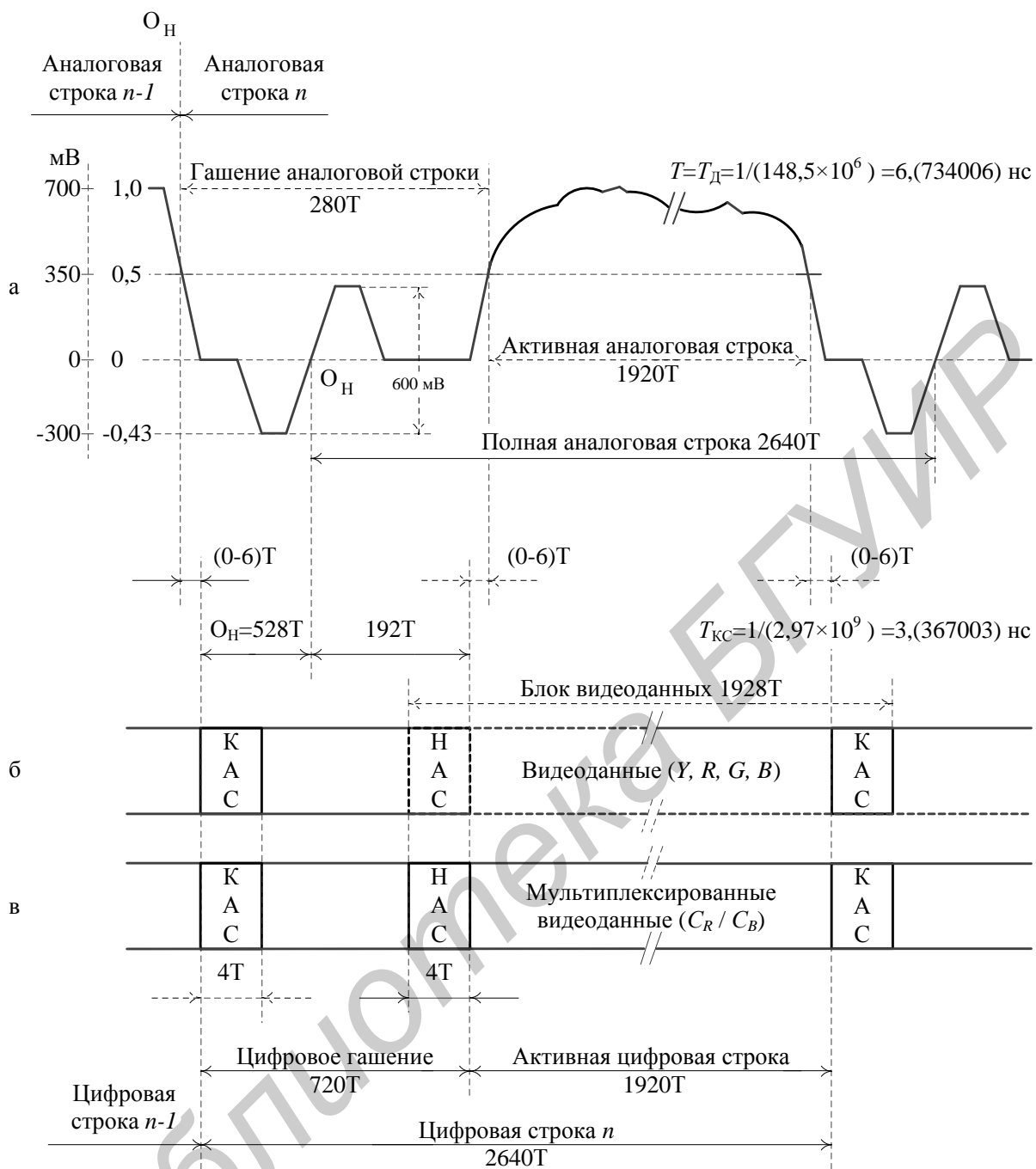
$$k_b = |U'_{B-Y}|_{\max} / 0,5 = 0,9278 / 0,5 = 1,8556.$$

Учитывая (1.28) и (1.30), находим выражения, устанавливающие взаимосвязь между сигналами основных цветов U'_R , U'_G и U'_B и нормированными ЦРС P_r и P_b :

$$P_b = \frac{U'_{B-Y}}{k_b} = \frac{-0,2126U'_R - 0,7152U'_G + 0,9278U'_B}{1,8556} = -0,1145U'_R - 0,3854U'_G + 0,5U'_B; \quad (1.31)$$

$$P_r = \frac{U'_{R-Y}}{k_r} = \frac{0,7874U'_R - 0,7152U'_G - 0,0722U'_B}{1,5748} = 0,5U'_R - 0,4541U'_G - 0,0458U'_B.$$

В ТВ приемниках сигналы основных цветов U^Y_R , U^Y_G и U^Y_B формируются из СЯ U'_Y и ЦРС P_r и P_b подобно выражению (1.24).



а – аналоговая строка; б, в – цифровые строки мультиплексированных данных (в режиме 4:2:2)

Рисунок 1.16 – Временные соотношения между аналоговой и цифровой строками в формате 4:2:2 (ТВЧ)

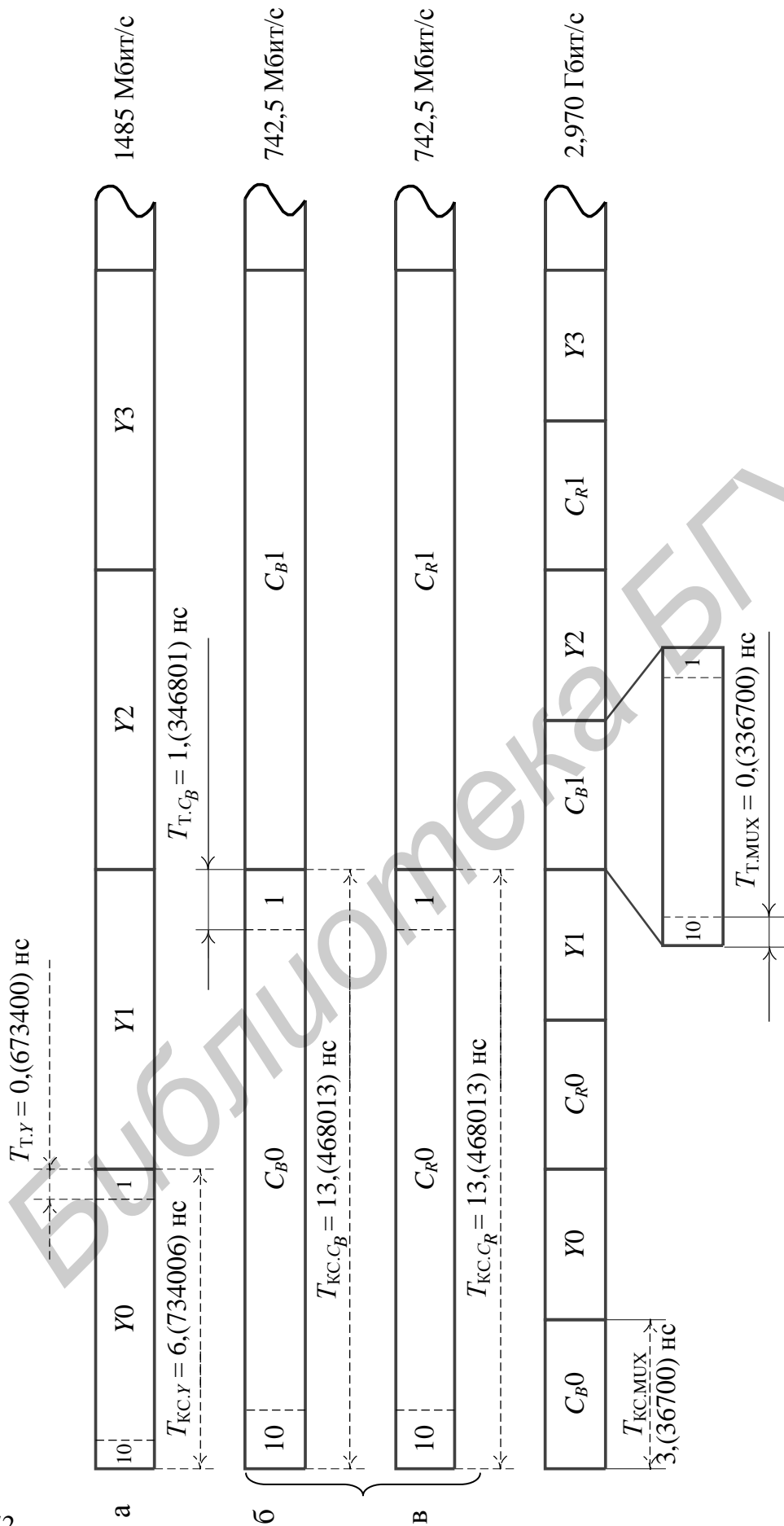


Рисунок 1.17 — Временное уплотнение компонентных цифровых сигналов в формате 4:2:2 (ТВЧ)

2 МЕТОДЫ СОКРАЩЕНИЯ ИЗБЫТОЧНОСТИ ИЗОБРАЖЕНИЙ

2.1 Классификация методов сокращения избыточности

2.1.1 Обоснование необходимости компрессии изображений

Как было показано в разделе 1, скорость потока данных, несущего информацию об изображении (цифрового видеосигнала), составляет, в зависимости от стандарта разложения и формата цифрового представления, от десятков Мбит/с до нескольких Гбит/с и более. В таблице 2.1 для сравнения приведены скорости потоков данных цифровых изображений типовых телевизионных систем – пониженного разрешения (ТПР, формат CIF), стандартной четкости с разложением на 625 строк (ТСЧ) и высокой четкости с разложением на 1125 (ТВЧ) и 720 строк (американский вариант ТВЧ). Также в ней приводятся объемы, занимаемые одним кадром, и длительности фрагмента ТВ программы, который можно записать на носитель емкостью 1 Гбайт. Во всех случаях брались данные, несущие информацию только о активной части изображения, без учета интервалов гашения (структурной избыточности). Примеры полных скоростей потоков данных можно увидеть в таблицах 1.5 и 1.7 предыдущего раздела.

Таблица 2.1 – Скорости потока данных цифровых изображений для некоторых типовых телевизионных систем

Параметры телевизионной системы			Скорость потока данных, Мбит/с	Объем одного кадра, МБайт	Длительность фрагмента объемом 1 Гбайт, с
Размер изображения, пикс.	Частота полей, Гц	Формат цифрового представления			
352×288 (CIF)	15i	4:2:2	24,3	0,193	346
		4:2:0	18,2	0,145	460
720×576 (ТСЧ)	25i	4:4:4	248,8	1,19	34,5
		4:2:2	165,9	0,79	51,8
		4:2:0	124,4	0,59	69,0
1920×1080 (ТВЧ)	25i	4:4:4	1244,2	5,98	6,7
		4:2:2	829,4	3,96	10,1
		4:2:0	622,1	2,97	13,4
	50p	4:4:4	2488,3	5,98	3,3
		4:2:2	1658,9	3,96	5,1
		4:2:0	1244,2	2,97	6,7
1280×720 (ТВЧ Ам.)	60p	4:4:4	1327,1	2,64	6,3
		4:2:2	884,7	1,76	9,5
		4:1:1	663,6	1,32	12,6

Далее в таблицах 2.2 – 2.4 приводятся данные о максимальной скорости передачи данных для цифровых интерфейсов, радиосистем и проводных (медных и оптических) систем передачи данных.

Таблица 2.2 – Скорости передачи данных некоторых цифровых интерфейсов

Интерфейс	SDI	IEEE 1394	USB	HDMI/DVI
Максимальная скорость передачи, Мбит/с	270 (SDI) 1485 (HD-SDI) 2970 (3G-SDI) 5940 (6G-SDI)	100...800 (в зав-ти от версии)	480 (ver. 2.0)	4,95 Гбит/с или 9,9 Гбит/с
Основное назначение	ТВ	Мульти-медиа	Мульти-медиа	ТВ, Мульти-медиа

Таблица 2.3 – Скорости передачи данных типовых радиосистем

Система/технология	DVB-T	DVB-T2	GSM/EGPRS (2,5G)	UMTS (3G)	HSDPA/HSUPA (3G+)	Wi-Fi	WiMAX
Максимальная скорость передачи, Мбит/с	31,7	50,3	0,4	3,84	21,6*	от 54** до 300	75**
Основное назначение	ЦТВ	ЦТВ	СПРС	СПРС	СПРС	СШБД	СШБД

* при ширине полосы частот радиоканала 5 МГц

** при ширине полосы частот радиоканала 20 МГц

Таблица 2.4 – Скорости передачи данных в проводных и оптических системах передачи

Система/технология	xDSL	Ethernet	SAN	STM-n
Максимальная скорость передачи, Мбит/с	51,84	10 (10 Base) ... 10 ⁴ (10G Base)	10 ⁴ (Fibre Channel)	n×155,52, n=1;4;16;64; 256 (перспект.)
Основное назначение	СД	ЛВС, СД	ЛВС	МСП

Сопоставление данных, приведенных в таблицах 2.1 и 2.2 – 2.4 показывает, что передача некомпьютеризированных данных возможна только посредством специализированных интерфейсов либо высокоскоростных систем передачи данных. Первые, однако, имеют ограничения по дальности передачи без регенерации: для металлических кабелей – от нескольких метров (USB, IEEE 1394) до 150 м (SDI), для оптических – до нескольких километров. Магистральные многоканальные системы передачи (МСП) STM-иерархии таких ограничений не имеют, но стоимость передачи данных с высокой скоростью в них также значительная, да и абонентские терминалы к ним напрямую не подключаются. Оборудование локальных вычислительных сетей (ЛВС) также теоретически позволяет передавать несжатые видеоданные в реальном времени, но при этом практически весь трафик сети будет создаваться всего двумя абонентами – отправителем и получателем видеоданных, множествен-

ный доступ к сети в это время навряд ли будет возможен. Поэтому на практике через ЛВС неkomпрессированные видеоданные передаются не в реальном времени, телецентры зачастую используют для этого специализированные SAN-сети на базе оптического интерфейса Fibre Channel.

Остальные рассматриваемые системы в большей либо меньшей степени пригодны для передачи только компрессированных видеоданных. При этом необходимо стремиться к результирующей скорости потока данных в расчете на одну программу в пределах 1–5 Мбит/с – для ТСЧ, 5–20 Мбит/с – для ТВЧ. Современные алгоритмы компрессии при таких условиях обеспечивают достаточно высокое качество изображений и звука, а телевизионное вещание становится доступным как при использовании специализированных систем цифрового телевизионного вещания (DVB, ATSC, ISDB и т. п.), причем в одном радиоканале возможна передача до десятка программ, так и для систем абонентского доступа (СД) технологий xDSL, Ethernet либо широкополосного беспроводного (СШБД). Системы подвижной радиосвязи (СПРС) имеют гораздо большие ограничения по скорости передачи данных, поэтому в них возможна передача изображений пониженного разрешения и частоты кадров, для которых возможно достижение скорости потока данных в 100–300 Кбит/с.

Компрессия изображений необходима и для их консервации (хранения). Несмотря на то, что емкости современных дисковых накопителей («винчестеров») превысили рубеж в 1 Тбайт/устройство, хранение несжатой видеоинформации нецелесообразно ни с технической, ни с экономической точки зрения.

Исключение составляют случаи, когда недопустимы даже малейшие потери качества – это студийная обработка при телепроизводстве и особенно – электронный кинематограф и системы цветной печати. В перечисленных системах используются специальные высокоскоростные интерфейсы (SDI, HD-SDI и др.), как электрические, так и на базе ВОЛС, и массивы дисковых накопителей (RAID-массивы), обеспечивающие передачу и хранение неkomпрессированных видеоданных либо видеофайлов. Для уменьшения объема последних могут использоваться только методы сжатия без потерь («архиваторы»).

2.1.2 Избыточность изображений

В любой цифровой системе, обеспечивающей передачу, хранение либо обработку изображений, два основных параметра аналого-цифрового преобразования – частота дискретизации (либо разрешающая способность) и разрядность двоичного кодирования – выбираются исходя из условия визуальной неотличимости при заданных условиях цифрового изображения от исходного аналогового. Выбор производится для наиболее сложного случая – максимального контраста изображения и минимального порогового

контраста, иные особенности зрительного восприятия, а также статистические свойства изображения при этом не учитываются. Поэтому сформированный таким образом цифровой сигнал содержит гораздо больший объем информации, чем требуется для обеспечения субъективно точной передачи изображений, т. е. обладает *избыточностью*.

Непосредственное вычисление избыточности изображений представляет собой крайне сложную инженерную задачу, особенно если речь идет о визуальной избыточности. Поэтому в качестве количественной меры эффективности компрессии изображений выступает коэффициент сжатия

$$K_{\text{КОМПР}} = \frac{V_{\text{ИСХ}}}{V_{\text{КОМПР}}}, \text{ или } K_{\text{КОМПР}} = \frac{B_{\text{ИСХ}}}{B_{\text{КОМПР}}}, \quad (2.1)$$

где $V_{\text{ИСХ}}$ – объем данных исходного изображения, байт;

$V_{\text{КОМПР}}$ – объем данных после компрессии, байт;

$B_{\text{ИСХ}}$ – скорость исходного потока данных, бит/с;

$B_{\text{КОМПР}}$ – скорость потока данных на выходе компрессора, бит/с.

Реже используют понятие относительной избыточности изображений [16]:

$$R_{\text{ИЗ}} = 1 - \frac{1}{K_{\text{КОМПР}}}. \quad (2.2)$$

Следует отметить, что при компрессии с потерями значение $R_{\text{ИЗ}}$ будет справедливым только при полной визуальной неотличимости исходного и восстановленного изображений. В некоторых случаях для обеспечения передачи изображений по каналам связи с ограниченной пропускной способностью коэффициент компрессии выбирается большим, что приводит к некоторым отличиям восстановленного изображения от исходного. При этом вычисленное по (2.2) значение относительной избыточности будет завышенным, так как в него войдут и невозполнимые потери изображения.

Количественную оценку эффективности использования битового потока данных для передачи информации об изображении дает эквивалентное удельное количество бит/отсчет, показывающее, сколько бит используется в среднем для передачи одного отсчета изображения:

$$B_{\text{ЭКВ}} = \frac{V_{\text{КОМПР}}}{X \cdot Y}, \quad (2.3)$$

где X и Y – размеры изображения по горизонтали и вертикали.

В дальнейшем будем различать три вида избыточности ТВ сигналов и изображений:

- структурную;
- статистическую;
- визуальную (психофизиологическую).

Структурная избыточность свойственна только ТВ сигналу и обусловлена наличием в нем кадровых (КГИ) и строчных (СГИ) гасящих импуль-

сов. При цифровой передаче изображений потребность в передаче КГИ и СГИ в явном виде отсутствует, поэтому объем передаваемой информации за счет этого можно сократить примерно на ΔV процентов:

$$\Delta V = \frac{N_{\text{СТР}}Z_{\text{К}} - N_{\text{СТР.А}}Z_{\text{К.А}}}{N_{\text{СТР}}Z_{\text{К}}} \cdot 100\%, \quad (2.4)$$

где $N_{\text{СТР}}$ – общее количество отсчетов в строке;

$N_{\text{СТР.А}}$ – количество активных отсчетов;

$Z_{\text{К}}$ – общее количество строк в кадре;

$Z_{\text{К.А}}$ – количество активных строк.

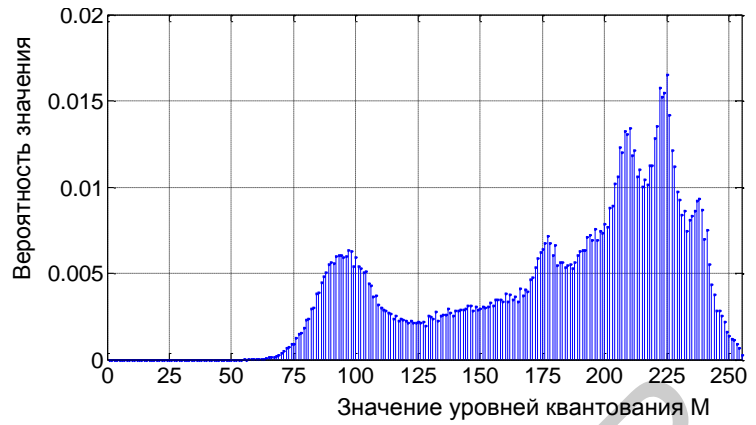
Для сигналов телевизионной системы стандартной четкости структурная избыточность составляет порядка 25 % (рисунок 2.1).



Рисунок 2.1 – Структурная избыточность изображения

Статистическая избыточность бывает двух типов – кодовая и межэлементная. Кодовая определяется различной вероятностью появления тех либо иных значений отсчетов изображения, и, соответственно, энтропия данного изображения не будет максимально возможной. В результате среднее число бит, необходимое для передачи каждого отсчета изображения, также будет меньше разрядности двоичного кодирования. Например, представленное на рисунке 2.2, а 8-битное изображение, характеризуемое плотностью распределения вероятностей значений отсчетов сигнала яркости, приведенной на рисунке 2.2, б, обладает энтропией $H \approx 7,45$, следовательно, его кодовая избыточность составляет $(8 - 7,45)/8 \approx 0,07 = 7\%$. Методы оценки кодовой избыточности и способы ее сокращения подробно рассматриваются в подразделе 2.2 настоящего пособия.

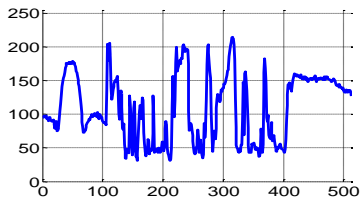
Межэлементная избыточность связана с корреляцией между отсчетами изображения в одной строке, в соседних строках и соседних кадрах. Так, большая часть изображения одного кадра, как правило, приходится на фон или задний план, где яркость постоянна или мало меняется в пространстве. Иллюстрация межстрочной корреляции приведена на рисунке 2.3.



а – исходное изображение, квантованное на 256 уровней (8 бит/отсчет)

б – плотность распределения вероятностей значений отсчетов изображения

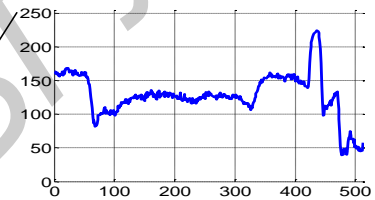
Рисунок 2.2 – Оценка кодовой избыточности изображений



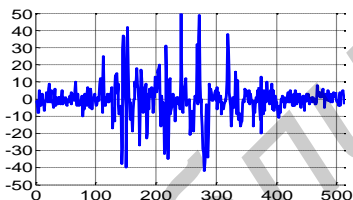
а – значения отсчетов изображения в строках 265, 266



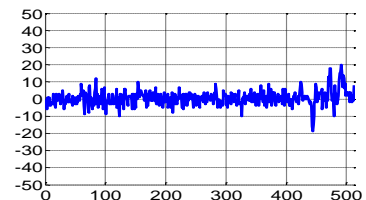
в – исходное изображение, размер – 512×512 элементов, разрядность квантования – 8 бит/отсчет



г – значения отсчетов изображения в строках 45, 46



б – разность между строками



д – разность между строками

Рисунок 2.3 – Корреляция между строками изображения

Видно, что на участках с сильно изменяющейся яркостью (средняя часть изображения на рисунке 2.3, в) соседние строки (рисунок 2.3, а) коррелированы достаточно слабо (рисунок 2.3, б). В то же время на участках с мало изменяющейся яркостью (верхняя часть изображения на рисунке 2.3, в) соседние строки (рисунок 2.3, г) коррелированы сильно, разность между ними минимальна (рисунок 2.3, д).

Комплексная оценка межэлементной избыточности отражена на рисунке 2.4. Изображения (рисунки 2.4, б и 2.4, в) представляют собой разность между исходным изображением (рисунок 2.4, а) и его копией, сдвину-

той на 1 элемент по горизонтали и вертикали соответственно. Поскольку разность между изображениями является двуполярной и достаточно малой, нулевой уровень (полное совпадение яркостей) смещен в область серого, а сама разность для наглядности представления увеличена в 5 раз. Анализ разностных изображений показывает достаточно высокую степень межэлементной корреляции как по горизонтали, так и по вертикали.



Рисунок 2.4 – Оценка пространственной межэлементной избыточности

Изображения в соседних кадрах также мало отличаются даже при наличии движущихся изображений (рисунок 2.5, а–в). Визуальная избыточность связана с учетом свойств зрения, позволяющим устранить часть информации об изображении, не существенно влияющей на его восприятие.

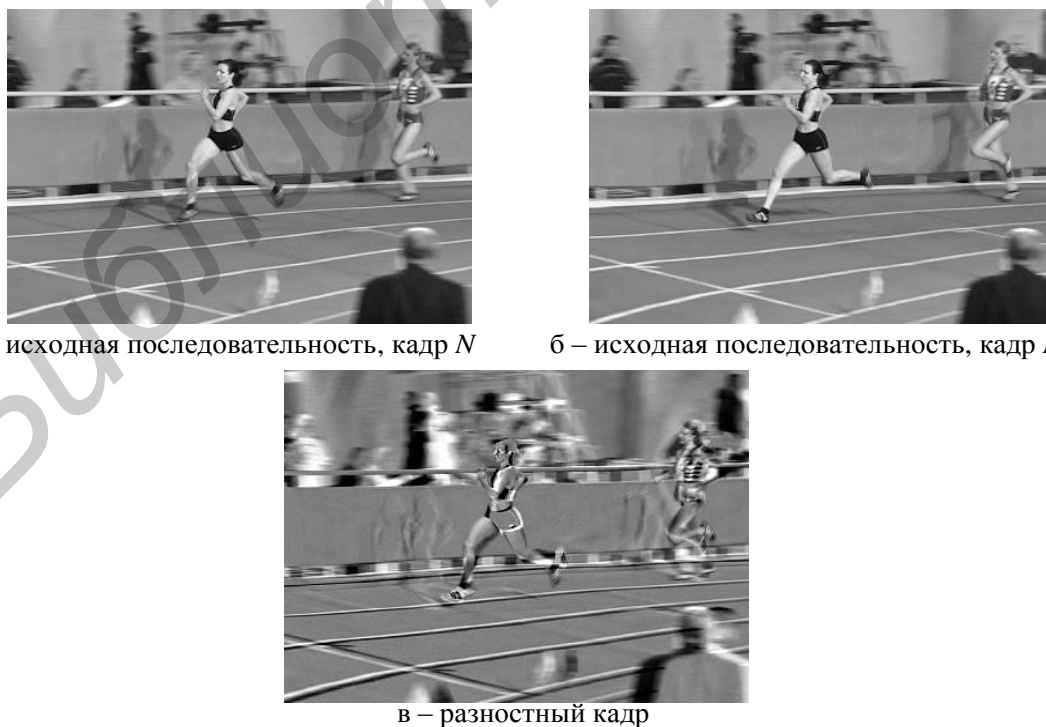


Рисунок 2.5 – Оценка временной межэлементной избыточности

Так, шумы квантования, различимые глазом на крупных деталях изображения в виде ложных контуров, малозаметны на мелких деталях и на резких перепадах яркости (рисунок 2.6). Цветовая разрешающая способность глаза в среднем в 3 – 4 раза меньше, чем яркостная [8], к сожалению, возможности черно-белой печати не позволяют отобразить это графически. Глаз также менее требователен к качеству отображения движущихся деталей.

Устранение визуальной избыточности наряду со статистической межэлементной является основным резервом при сжатии изображений и подробно рассматривается в последующих разделах настоящего пособия.

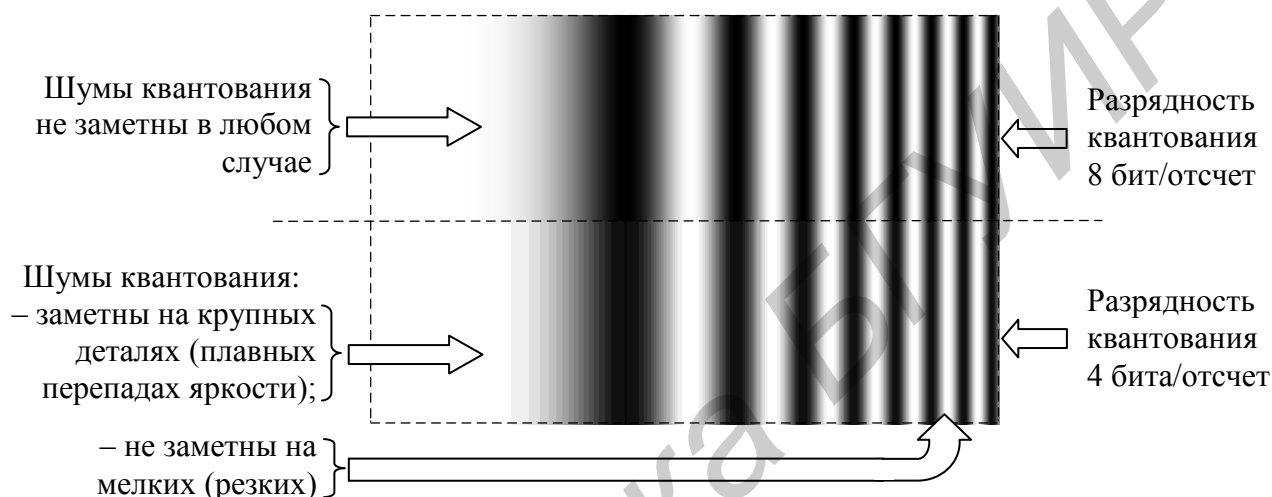


Рисунок 2.6 – Зависимость заметности шумов квантования от размеров элементов изображения

2.1.3 Классификация методов сокращения избыточности изображений

Современные методы компрессии изображений чрезвычайно разнообразны. Представленная в данном пункте их классификация (рисунок 2.7) не претендует на абсолютную полноту и включает в себя только наиболее характерные признаки, по которым можно разделить методы, применяющиеся для устранения избыточности статических и динамических изображений.

Прежде всего, в зависимости от точности восстановления изображения, компрессия изображений может быть осуществлена двумя способами – с потерями или без потерь.

Сжатие без потерь. Полученное после декомпрессии изображение (т. е. восстановленное) будет побитно совпадать с оригиналом, что дает возможность производить неограниченное количество циклов компрессии/декомпрессии без потери качества изображения. В процессе кодирования устраняется только статистическая избыточность. Поскольку коэффициент сжатия при этом редко превышает 50 %, непосредственно для кодирования отсчетов изображения данный метод не применяется.

В существующих системах компрессии изображений статистический, или, как его часто называют, энтропийный, кодер устанавливается на выходе видеокодера и обрабатывает уже закодированные данные – квантованные коэффициенты, вектора движения и т. п.

Сжатие с потерями. Реализующие это методы основаны на устранении психофизиологической избыточности и кодировании статистически зависимых символов. В этом случае восстановленное изображение побитно уже не совпадает с исходным. Точность восстановления при этом может устанавливаться в широких пределах – от полного визуального сходства до достаточно больших различий. Конкретные требования к качеству компрессированного изображения (степени совпадения с оригиналом) определяются назначением системы. Если приоритетным является максимально возможное качество изображения (примеры – студийная обработка изображений в телепроизводстве, профессиональная цифровая фотография), выбирается достаточно малый коэффициент сжатия; в случаях, когда определяющим фактором является пропускная способность канала связи (пример – видеоконференц-связь), допускают некоторые погрешности в точности передачи изображений, но коэффициент сжатия при этом выбирается максимально возможным.

Все современные стандартизованные методы компрессии изображений, применяемые в телевизионном и мультимедийном вещании, видеозаписи, цифровой фотографии – MPEG-1, 2, 4; JPEG/JPEG-2000, WMV, VC-1,2 и прочие – обеспечивают сжатие с потерями¹. Коэффициент сжатия может при этом варьироваться в широких пределах для обеспечения компромисса между качеством изображения и пропускной способностью канала связи. Общий недостаток методов компрессии с потерями – накопление ошибок при повторе циклов компрессии/декомпрессии.

Компрессия изображений обеспечивается за счет устранения их избыточности – визуальной и статистической. Последняя бывает двух типов – кодовая и межэлементная, методы их устранения тоже различны. И визуальная, и межэлементная избыточность делятся на два общих класса – временную (межкадровую) и пространственную (внутрикадровую). Структурная избыточность в данной классификации не рассматривается, т. к. присуща только телевизионным сигналам, а не изображениям в отдельности.

Временная избыточность устраняется кодированием с предсказанием, при котором обработке подвергаются не непосредственно отсчеты изображения, а разность между их фактическим и предсказанным значениями. Варианты реализации кодирования с преобразованием рассматриваются в последующих разделах. Непосредственное устранение только временной избыточности, без устранения пространственной, на практике не используется.

¹ Некоторые стандарты, в частности JPEG-2000, допускают и сжатие без потерь, однако, ввиду малого коэффициента компрессии при этом, указанный режим используется в достаточно редких случаях, например в перечисленных в пункте 2.1.1.

Для устранения пространственной избыточности могут использоваться кодирование с предсказанием, кодирование с преобразованием, гибридное кодирование, объединяющее в себе оба предыдущих метода, интерполяционное и фрактальное кодирование. Впрочем, фрактальное кодирование, перспективность которого достаточно широко обсуждалась в научно-технической литературе на протяжении последних десятилетий [65], в настоящее время используется крайне редко и в настоящем пособии далее не рассматривается.

В задачах компрессии без потерь для устранения пространственной избыточности может использоваться кодирование длин серий (КДС), наиболее подходящее для сжатия двухградационных изображений, его усовершенствованный вариант – кодирование битовых плоскостей, подходящее и для полутоновых изображений, а также словарные методы. Среди последних наибольшее распространение получили алгоритмы Лемпеля–Зива (LZ, LZ77, LZ78 и LZW – Лемпеля–Зива–Уэлча), широко использующиеся в компьютерной графике.

Кодеки динамических изображений (далее – видеокодеки) устраняют, как правило, и пространственную, и временную избыточность. Исключение составляют видеокодеки цифровых систем видеозаписи форматов DV, DVCAM, DVCPro и некоторые другие. Причина – необходимость обеспечения видеомонтажа с покадровой точностью без декомпрессии изображений. Методы компрессии статических изображений (группа JPEG), естественно, устраняют только пространственную избыточность.

Кодирование с преобразованием выполняется на базе дискретных двумерных ортогональных преобразований – Фурье (ДПФ), косинусного (ДКП), Уолша–Адамара (ДПУА) [7]. В отдельных случаях используются преобразования Хартли (ДПХ), S-преобразование (или сингулярное – СП) и др. Весьма эффективным и поэтому широко используемым в последнее время, является дискретное вэйвлетное (волновое) преобразование (ДВП).

Интерполяционные методы, включающие при кодировании усреднение отсчетов изображения с последующим их прореживанием, а при декодировании – экстраполяцию, тоже можно считать методами сжатия изображений с потерями. Они обеспечивают устранение визуальной избыточности, обусловленной, в частности, различной чувствительностью зрения к яркостным и цветовым переходам. Реализуются путем перехода от формата цифрового представления сигналов 4:4:4 к форматам с прореживанием цветоразностных сигналов (4:2:2, 4:1:1, 4:2:0 и др.), рассмотренным в предыдущем подразделе.

Кодовая избыточность устраняется путем использования кодов переменной длины (VLC-коды). В первых стандартах компрессии статических и динамических изображений (MPEG-1, 2; JPEG) использовались коды Хаффмана; в последующих (MPEG-4 AVC (H.264), H.265, JPEG-2000) применяются более эффективные коды Голомба и арифметическое кодирование.

В задачах компрессии изображений используются одномерная, двумерная (пространственная) и трехмерная (пространственно-временная) обра-

ботка данных. При одномерной обработке используются различные варианты ДИКМ и статистического кодирования, основанные на учете корреляции между соседними элементами изображения по горизонтали (по строке). При двумерной обработке учитывается корреляция между соседними элементами по обоим направлениям – горизонтали и вертикали. При трехмерной обработке учитывается также и корреляция между элементами с одинаковыми координатами соседних кадров (временная корреляция).

Во всех существующих стандартах компрессии статических и динамических изображений (JPEG, MPEG-1,2,4 и др.) используются комбинации из различных методов сжатия – основу внутрикадровой компрессии составляют методы кодирования с преобразованием на базе ДКП или ДВП, для межкадровой компрессии используется кодирование с предсказанием и компенсацией движения, сформированный поток данных подвергается в дополнение еще и кодированию статистическими кодами. В результате степень сжатия изображений может достигать десятков и даже сотен раз при достаточно хорошем качестве воспроизводимого изображения.

2.1.4 Критерии оценки качества отображения

Как было отмечено ранее, сокращение визуальной избыточности влечет потерю реальной количественной визуальной информации. Поскольку при этом может быть также утеряна и представляющая интерес информация, то весьма желательно иметь средства количественных оценок характера и величины потерь информации. В основу такого определения могут быть положены как объективные, так и субъективные критерии верности (точности) воспроизведения.

Если степень потери информации может быть выражена как функция исходного (входного) изображения и сжатого, а затем восстановленного (выходного) изображения, то такой подход называют объективным критерием верности воспроизведения. Наиболее простым здесь является критерий среднеквадратического отклонения (СКО) разности выходного и входного изображений. Пусть $f(x, y)$ означает входное изображение, а $f'(x, y)$ – восстановленное. Ошибка восстановления (разность между исходным и восстановленным изображениями) $\varepsilon(x, y)$ для элементов изображений $f(x, y)$ и $f'(x, y)$ определяется как

$$\varepsilon(x, y) = f'(x, y) - f(x, y), \quad (2.5)$$

а величина СКО (MSE – mean squared error) этих изображений равна

$$\|\varepsilon(x, y)\|^2 = \frac{1}{X \cdot Y} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} [f'(x, y) - f(x, y)]^2, \quad (2.6)$$

где $X \times Y$ – размеры изображения.

Величина корня из среднеквадратического отклонения (RMSE – root mean squared error) разности изображений $f'(x, y)$ и $f(x, y)$ будет равна:

$$\varepsilon = \sqrt{\frac{1}{X \cdot Y} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} [f'(x, y) - f(x, y)]^2}. \quad (2.7)$$

Другим близким объективным критерием верности воспроизведения является отношение сигнал/шум восстановленного изображения. Если с помощью простой перестановки членов в выражении (2.5) рассматривать изображение $f'(x, y)$ как сумму исходного изображения $f(x, y)$ и шума $\varepsilon(x, y)$, то отношение сигнал/шум, обозначаемое SNR , вычисляется по выражению:

$$SNR = \sqrt{\frac{\sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} [f'(x, y)]^2}{\sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} [f'(x, y) - f(x, y)]^2}}. \quad (2.8)$$

Недостатком использования критерия ОСШ для оценки качества восстановленных изображений является его зависимость от значений отсчетов изображения, т. е. от его средней яркости. В результате каждому испытательному изображению при одинаковых условиях будет соответствовать свое отношение сигнал/шум, что неудобно при оценке результатов. Поэтому, по аналогии с методом измерения ОСШ в системах аналогового ТВ, где в качестве значения сигнала берется разность между контрольными уровнями белого и черного, являющаяся постоянной, вычисляется пиковое отношение сигнал/шум – $PSNR$. Оно представляет собой отношение квадрата максимального количества уровней квантования, используемых в системе, к СКО изображений:

$$PSNR = \frac{(2^m - 1)^2}{\sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} [f'(x, y) - f(x, y)]^2}, \quad (2.9)$$

либо, при более удобной для выполнения сравнений логарифмической шкале,

$$PSNR = 10 \cdot \lg \frac{(2^m - 1)^2}{\sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} [f'(x, y) - f(x, y)]^2}, \text{ дБ}. \quad (2.10)$$

Критерий $PSNR$, несмотря на простоту вычислений, имеет и ряд недостатков. Для его вычисления необходимо иметь и исходную, и преобразованную последовательность, однако первая бывает не всегда доступна. Кроме того, не всегда легко проверить, имело ли само исходное изображение совершенное качество. Величина $PSNR$ также далеко не всегда коррелируется с субъективными оценками качества, например с рекомендациями ITU [33, 34].

Для изображений или последовательностей изображений большое $PSNR$ обычно указывает на высокое, а малое – на низкое качество преобразованных изображений по сравнению с исходными. Однако одинаковые величины $PSNR$ необязательно гарантируют одинаковое субъективное качество. Например, на рисунке 2.8 приведено исходное изображение и восстановленное после компрессии различными методами, характеризующиеся разным значением $PSNR$. Очевидно, что субъективная оценка качества всех трех изображений не столь однозначна. Следовательно, оценка качества изображений по критерию $PSNR$ является недостаточной и может использоваться только в качестве предварительной, а окончательное решение должно приниматься на основании субъективной экспертизы в соответствии с [33; 34].



Рисунок 2.8 – Неоднозначность соответствия объективных и субъективных критериев оценки качества изображений

Определение качества изображения с помощью субъективного оценивания часто является предпочтительным. Это может быть достигнуто путем показа «типичного» восстановленного изображения группе наблюдателей (экспертов) и усреднения их оценок. Оценивание может производиться как с использованием абсолютной шкалы оценок, так и путем попарного сравнения изображений $f'(x, y)$ и $f(x, y)$. Методам оценки качества изображений, помимо упомянутых рекомендаций ИТУ, посвящено большое количество специальной литературы, например, [35 – 37].

2.2 Статистическое кодирование

2.2.1 Оценка избыточности изображений

В предыдущих подразделах были рассмотрены основные пути сокращения объема данных, необходимых для представления изображений. Естественно, встает вопрос о минимальном количестве данных, необходимых для полного описания изображения без потери информации. Математически ре-

шает этот вопрос теория информации. Фундаментальная ее предпосылка заключается в том, что любой источник информации, в том числе и источник сигнала изображения, может быть описан как вероятностный процесс [38]. В соответствии с этим предположением случайное событие E , появляющееся с вероятностью $P(E)$, содержит

$$I(E) = \log \frac{1}{P(E)} = -\log P(E) \quad (2.11)$$

единиц информации. Основание логарифма в (2.11) определяет единицу измерения количества информации. Если основание равно 2, единица измерения называется бит.

Предположим, что источник информации генерирует случайную последовательность символов из конечного набора. Применительно к изображению это означает, что каждый его элемент (отсчет) может принимать одно из $M = 2^m$ значений, где m – разрядность двоичного кодирования.

Вероятность того, что элемент изображения принимает значение a_j , равна $P(a_j)$, причем $\sum_{j=1}^M P(a_j) = 1$.

Для описания совокупности вероятностей появления элементов изображения со значениями a_1, a_2, \dots, a_M используется вектор вероятностей $z = [P(a_1) P(a_2) \dots P(a_M)]$, показывающий, с какой вероятностью встречаются элементы со значением a_j . Графическое отображение функции $P(a_j)$ называется гистограммой вероятностей значений изображения.

В соответствии со сделанными предположениями и формулой (2.11) количество информации, содержащееся в элементе со значением a_j будет равно $I(a_j) = -\log P(a_j)$. Для изображения размером $X \times Y$ элементов согласно закону больших чисел значение a_j будут принимать $X \cdot Y \cdot P(a_j)$ элементов. Исходя из этого, среднее количество информации, содержащееся в изображении, будет равно

$$\begin{aligned} X \cdot Y [-P(a_1) \log P(a_1) - P(a_2) \log P(a_2) - \dots - P(a_M) \log P(a_M)] = \\ = -X \cdot Y \sum_{j=1}^M P(a_j) \log P(a_j), \end{aligned}$$

а среднее количество информации, приходящееся на один элемент изображения – энтропия (неопределенность) изображения $H(z)$ будет равна:

$$H(z) = -\sum_{j=1}^M P(a_j) \log P(a_j). \quad (2.12)$$

Если значения элементов изображения a_j равновероятны, энтропия максимальна и равна разрядности двоичного кодирования. Энтропию изображения, характеризующую только кодовую избыточность, без учета межэлементных связей, называют оценкой первого порядка. Она определяет

нижнюю границу для сжатия, которого можно достичь применением только кодов переменной длины. На рисунках Б.1 и Б.2 (Приложение Б) цветной вклейки приведены примеры испытательного и реального изображений, для сигналов которых построены гистограммы распределения вероятности значений и вычислена энтропия.

Все изображения являются цветными и все соответствующие им сигналы кодированы с разрядностью 8 бит/отсчет. Испытательное изображение синтезировано искусственным путём, его сигналы характеризуются наличием от 2 (сигналы основных цветов) до 8 (сигнал яркости) уровней, соответственно и их энтропии невелики – от 1 до 3 бит/элемент. Для сигналов реального изображения характерно гораздо более равномерное распределение вероятности мгновенных значений отсчётов сигналов, соответственно, и значения энтропии существенно больше – в данном случае от 5,34 до 7,59 бит/элемент. Это означает, что непосредственное использование кодов переменной длины для их компрессии даст максимальный выигрыш в $8/5,34 \dots 8/5,5 \approx 1,05 \dots 1,5$ раз (5...50 %).

Для оценки энтропии более высоких порядков вычисляются вероятности появления пар, а далее блоков элементов изображения с комбинациями значений из множества $\{a_1, a_2, \dots, a_M\}$. Поскольку количество возможных пар (либо более сложных комбинаций) растёт пропорционально степени количества элементов в комбинации, вычисление энтропий высоких порядков чрезвычайно затруднительно. Например, 8-битное изображение может содержать до $2^{8-2} = 2^{16}$ пар значений, $2^{8-3} = 2^{24}$ троек и т. д. Понятно, что размер изображения при этом должен быть достаточно большим, поскольку число возможных комбинаций зависит также и от количества его элементов.

С учетом того, что увеличение вычислительной сложности не идёт ни в какое сравнение с получаемым выигрышем в эффективности кодирования, обычно составляющим единицы, максимум – десятки процентов, на практике статистические методы кодирования реализуются без учета межэлементных связей и используются только для устранения кодовой избыточности. Для повышения эффективности их использования могут применяться различные методы снижения энтропии изображения – кодирование с предсказанием, кодирование с преобразованием, в том числе и с потерями, рассматриваемые в последующих разделах. Здесь же рассмотрим наиболее распространенные методы сокращения кодовой избыточности и некоторые алгоритмы сжатия изображений на их основе.

2.2.2 Код Хаффмана

Наиболее простым подходом к сжатию изображений без потерь является сокращение только кодовой избыточности. Как правило, кодовая избыточность присутствует при любом обычном двоичном кодировании значений элементов изображения. Как уже отмечалось выше, она может быть устранена кодированием уровней сигналов – яркостного, ЦРС, либо иных, несущих информацию об отсчетах изображения, при условии минимизации среднего числа бит, необходимого для представления значения одного элемента. Чтобы достичь этого, требуется разработка кода переменной длины (неравномерного кода), который наиболее вероятным уровням яркости присваивает самые короткие кодовые комбинации и наоборот.

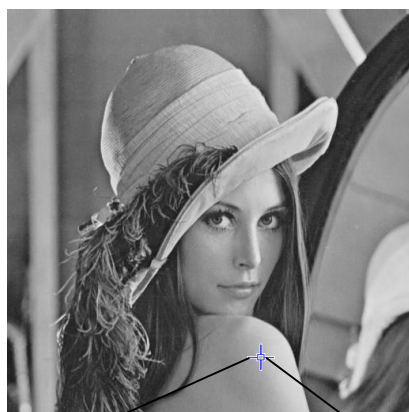
Самый известный метод сокращения кодовой избыточности был предложен Хаффманом¹ (Huffman) [39]. При независимом кодировании символов источника информации коды Хаффмана обеспечивают наименьшее число кодовых символов (битов) на символ источника.

Процедуру кодирования по Хаффману наиболее удобно рассматривать на базе простого примера. Поэтому возьмем в качестве источника фрагмент яркостной составляющей размером 4×4 реального изображения (рисунок 2.9, а). Гистограмма распределения вероятностей значений отсчетов для него приведена на рисунке 2.9, б.

Первым шагом в подходе Хаффмана является построение серии (множества уровней) редуцированных источников путем упорядочивания вероятностей значений рассматриваемых символов и «склеивания» символов с наименьшими вероятностями в один символ, который будет замещать их в редуцированном источнике следующего уровня. Этот процесс иллюстрируется на рисунке 2.10, а для двоичного кодирования (аналогично могут быть построены К-символьные коды Хаффмана). В двух левых колонках набор символов источника и их вероятности упорядочены сверху вниз в порядке убывания вероятностей.

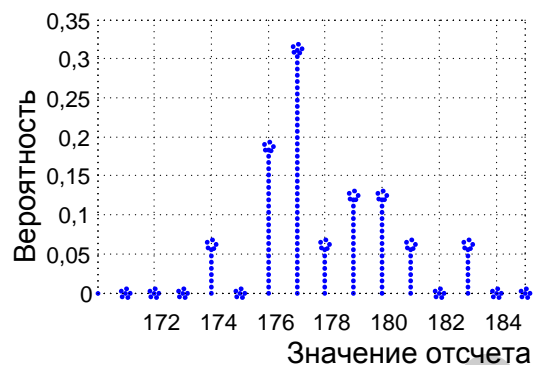
Для формирования первой редукции источника, символы с наименьшими вероятностями – в данном случае 0,0625 – объединяются в некоторый «составной символ» с суммарной вероятностью 0,1250. Этот составной символ и связанная с ним вероятность помещаются в список символов редуцированного источника, который опять упорядочивается в порядке убывания значений полученных вероятностей. Этот процесс повторяется до тех пор, пока не образуется модифицированный источник всего лишь с двумя оставшимися символами (крайняя правая колонка на рисунке).

¹ В русскоязычной литературе используется два варианта фамилии – Хаффман либо Хаффмен.



175	171	174	179	177	183	182	184
175	172	178	179	181	183	181	182
175	174	180	176	179	180	178	181
173	175	177	176	177	177	179	180
172	173	174	177	177	176	180	181
169	171	172	169	171	174	181	183
171	167	167	171	176	180	179	184

а



б

a_{ij}	0	1	2	3
0	178	179	181	183
1	180	176	179	180
2	177	176	177	177
3	174	177	177	176

в

а – фрагмент изображения; б – гистограмма распределения вероятностей значений; в – таблица значений отсчетов фрагмента изображения

Рисунок 2.9 – Пример выбора источника для кодирования по Хаффману

Исходный источник		Редуцированный источник					
Значение	Вероятность	1	2	3	4	5	6
177	0,3125	0,3125	0,3125	0,3125	0,3125	0,4375	0,5625
176	0,1875	0,1875	0,1875	0,2500	0,2500	0,3125	0,4375
179	0,1250	0,1250	0,1250	0,1875	0,2500	0,2500	
180	0,1250	0,1250	0,1250	0,1250	0,1875		
174	0,0625	0,1250	0,1250	0,1250			
178	0,0625	0,0625	0,1250				
181	0,0625	0,0625					
183	0,0625						

Рисунок 2.10 – Модификации источника по Хаффману

Второй шаг в процедуре кодирования по Хаффману состоит в кодировании каждого из модифицированных источников, начиная с источника с наименьшим числом символов (т. е. правого на рисунке 2.10), и возвращаясь обратно к исходному источнику (рисунок 2.11).

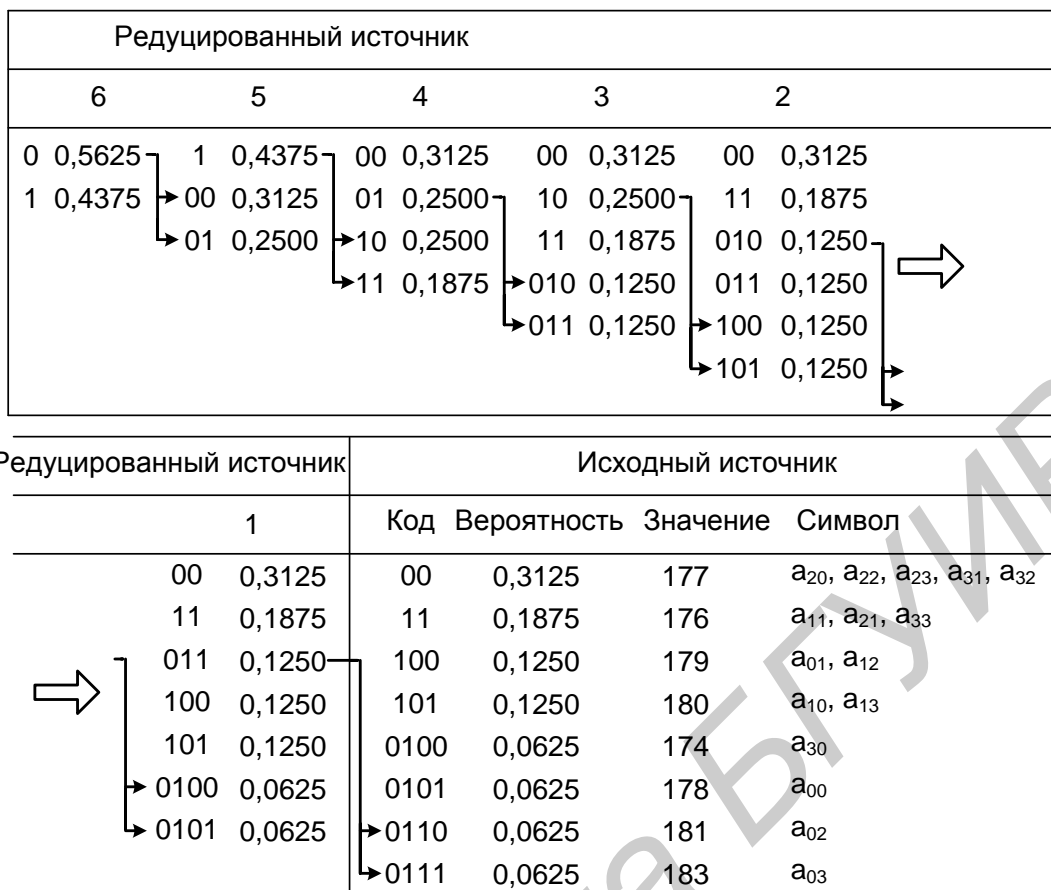


Рисунок 2.11 – Процедура построения кода Хаффмана

Для источника с двумя символами наименьшим двоичным кодом являются, конечно, символы 0 и 1. Как показано на рисунке 2.14, эти символы приписываются символам источника слева (выбор символов произволен — изменение 0 на 1 и наоборот даст абсолютно тот же результат). Поскольку символ текущего модифицированного источника, имеющий вероятность 0,5625, был получен объединением двух символов предыдущего модифицированного источника, то кодовый бит 0, выбранный для данного варианта, приписывается каждому из двух соответствующих символов предыдущего источника, затем эти коды произвольным образом дополняются символами 0 и 1, для отличия их друг от друга. Эта операция затем повторяется для модифицированных источников всех остальных уровней, пока не будет достигнут уровень исходного источника. Результирующий код приведен в колонке «Код» рисунка 2.11.

Таким образом, последовательность $a_{00} \dots a_{33}$ элементов рассматриваемого фрагмента изображения будет закодирована следующим образом:

Символ	a_{00}	a_{01}	a_{02}	a_{03}	a_{10}	a_{11}	a_{12}	a_{13}	a_{20}	a_{21}	a_{22}	a_{23}	a_{30}	a_{31}	a_{32}	a_{33}
Значение	178	179	181	183	180	176	179	180	177	176	177	177	174	177	177	176
Код	0101	100	0110	0111	101	11	100	101	00	11	00	00	0100	00	00	11

Процедура Хаффмана строит оптимальный код для набора символов и их вероятностей при условии, что символы кодируются по отдельности. После

того, как код построен, процесс кодирования/декодирования осуществляется простым табличным преобразованием. Код Хаффмана является мгновенным однозначно декодируемым блоковым кодом. Он называется блоковым кодом, поскольку каждый символ источника отображается в фиксированную последовательность кодовых символов. Он является мгновенным, потому что каждое кодовое слово в строке кодовых символов может быть декодировано независимо от последующих символов. Он является однозначно декодируемым, т. к. любая строка из кодовых символов может быть декодирована единственным образом. Таким образом, любая строка кодированных по Хаффману символов может декодироваться анализом отдельных символов в строке слева направо. Для двоичного кода, представленного на рисунке 2.11, анализ слева направо показывает, что в закодированной строке 01011000110011110111100101001100000100000011 первым правильным кодовым словом является 0101, которое есть код для символа a_{00} . Следующим правильным кодовым словом является 100, что соответствует символу a_{01} . Продолжая эти действия, получим декодированное сообщение в виде $a_{00}a_{01}a_{02}a_{03}a_{10}\dots a_{33}$.

Проведем оценку эффективности кодирования. Средняя длина полученного кода составит $L_{ср} = 44 \text{ бита} / 16 \text{ символов} = 2,75 \text{ бита/символ}$. Поскольку энтропия источника, вычисленная по (2.12) равна $2,727 \text{ бита/символ}$, то эффективность кода Хаффмана составит $2,727/2,75 = 0,99$. Безусловно, это весьма высокое значение. Однако если провести оценку степени сжатия рассматриваемого фрагмента изображения, то результаты не столь оптимистичны. Анализируемый фрагмент имел размеры $4 \times 4 = 16$ элементов, соответственно, для полного описания каждого из них достаточно кодовой комбинации длиной $\log_2 16 = 4$ бита. Следовательно, степень сжатия в данном случае составляет всего $4/2,75 = 1,45$ раза или 45 %.

Построение двоичного оптимального кода Хаффмана является нетривиальной задачей, когда нужно кодировать большое число символов. Для общего случая M исходных символов необходимо построить $M - 2$ редукций источника (см. рисунок 2.10) и выполнить $M - 2$ присвоения кода (см. рисунок 2.11). Так, построение оптимального кода Хаффмана для изображения с 256 уровнями сигнала яркости, требует 254 редукции источника и 254 присвоения кода. Ввиду вычислительной сложности этой задачи, иногда приходится жертвовать кодовой эффективностью для упрощения кодовой конструкции. Поэтому разработаны коды переменной длины, обеспечивающие компромисс между эффективностью и вычислительной сложностью [16].

Кодирование Хаффмана имеет также еще два недостатка, которые проявляются при его реализации на практике в видеокодеках. Во-первых, декодер должен использовать то же самое семейство кодов. Передача информации, содержащейся в таблице вероятностей, потребует дополнительных битов, что сокращает степень сжатия, особенно при кодировании коротких видеофрагментов. Во-вторых, таблицу вероятностей для длинной видеопоследовательности, необходимую для построения дерева Хаффмана (см. рисунки 2.9 и 2.10), можно определить только после просмотра всей видеопоследова-

тельности. Это может привести к недопустимой задержке процесса кодирования, передачи и воспроизведения видео на приемном конце. По этим причинам современные стандарты кодирования динамических и статических изображений определяют множества кодовых слов, исходя из распределения вероятностей для некоторого «обобщенного» видеоматериала. Для примера можно привести фрагмент таблицы значений кода переменной длины для представления разностей векторов движения в стандарте MPEG-4 – таблица 2.5.

Для пояснения отметим, что рассматриваемые разностные векторы движения кодируются в виде пары VLC-кодов (VLC – Variable Length Code): первый – для разности горизонтальной составляющей векторов движения соседних блоков (компонента x), второй – для разности горизонтальной составляющей (компонента y).

Таблица 2.5 – Коды переменной длины для разностей векторов движения в MPEG-4

Разность MVD, ед.	Код
0	1
+0,5	010
- 0,5	011
+1	0010
-1	0011
+ 1,5	00010
-1,5	00011
+2	0000110
-2	0000111
+2,5	00001010
-2,5	00001011
+3	00001000
-3	00001001
+3,5	00000110
-3,5	00000111
...	...

В таблице 2.5 показаны только первые 15 из 64 кодов. Остальные 49 (с длиной бит от 8 до 13) не приводятся. Отметим, что самые короткие коды отвечают малым разностям векторов движения (MVD – Movie Vector Displacement), например, вектор $MVD = 0$ представлен кодом в один бит – 1. Коды этой таблицы весьма близки к «истинным» кодам Хаффмана, так как каждому символу присвоен единственный код, часто встречающимся элементам приписаны более короткие коды и никакой код из таблицы не является префиксом другого кода этой же таблицы. Главное отличие этих кодов от «истинных» кодов Хаффмана заключается в том, что они построены на основе

предварительных вычислений с использованием обобщенных распределений, в данном случае – вероятностей значений векторов движения.

Наряду с кодами Хаффмана существует множество других семейств кодов VLC, которые могут быть полезными в приложениях кодирования видео. Серьезным недостатком кодов, построенных на основе схемы Хаффмана, является их сильная чувствительность к ошибкам. Искажение в одном бите последовательности этих кодов может привести к полной потере синхронизации при декодировании и к невозможности дальнейшего правильного декодирования последовательности.

Обратимые коды VLC (RVLC, Reversible Variable Length Codes), которые можно успешно декодировать в прямом и обратном направлении, способны исправить такую ошибку. Недостаток предварительно вычисленных кодов (см. табл. 2.5) состоит в том, что и кодер, и декодер должны хранить эти таблицы в памяти. Альтернативный подход заключается в использовании кодов, генерируемых автоматически («на лету»), если известен входной символ – например, экспоненциальных кодов Голомба (Exp-Golomb) [40] в стандарте MPEG-4.

2.2.3 Арифметическое кодирование

В отличие от рассмотренных ранее неравномерных кодов, *арифметическое кодирование* создает неблочные коды. В арифметическом кодировании не существует однозначного соответствия между символами источника и кодовыми словами. Вместо этого, вся последовательность символов источника (т. е. все сообщение) соотнесена с одним арифметическим кодовым словом. Само по себе кодовое слово задает интервал вещественных чисел между 0 и 1. С увеличением числа символов в сообщении, интервал, необходимый для их представления, уменьшается, а число единиц информации (например битов), требуемых для представления интервала, увеличивается. Каждый символ в сообщении уменьшает размер интервала в соответствии с вероятностью своего появления. Поскольку метод не требует, чтобы каждый исходный символ отображался в целое число кодовых слов (т. е. чтобы символы кодировались по одному), он достигает (в теории) границы, установленной теоремой кодирования без шума (первой теоремой Шеннона) [16, 38].

На рисунке 2.12 проиллюстрирован основной процесс арифметического кодирования. Здесь кодируется сообщение из пяти символов, порожденное четырехсимвольным источником: $a_1a_2a_3a_4$. В начале процесса кодирования предполагается, что сообщение занимает весь полуоткрытый интервал $[0, 1]$. Этот интервал изначально делится на четыре отрезка пропорционально вероятностям символов источника, которые приведены в таблице 2.6. Символу a_1 , например, соответствует подынтервал $[0, 0,2]$. Поскольку это первый символ кодируемого сообщения, то значит, интервал оставшейся части сообщения $\{a_2a_3a_4\}$ будет сужен до $[0, 0,2]$. На рисунке 2.12 полученный интервал растянут на полную вы-

ту, а приведенные значения на его концах соответствуют значениям суженного диапазона. Затем суженный диапазон также делится на отрезки, пропорционально вероятностям символов источника, и процесс повторяется со следующим символом сообщения. Таким образом, символ a_2 сузит подынтервал до $[0,04, 0,08]$, a_3 – до $[0,056, 0,072]$ и т. д. Последний символ сообщения, который должен быть зарезервирован для специального индикатора окончания сообщения, сужает диапазон до $[0,06752, 0,0688]$. Конечно, любое число в этом подынтервале – например, 0,068, может быть использовано для представления сообщения.

Сообщение из пяти символов, приведенное на рисунке 2.12, после арифметического кодирования требует для записи всего трех десятичных цифр. Это соответствует $3/5$ или 0,6 десятичных знаков на символ источника и весьма близко энтропии источника, которая, согласно (2.12), составляет 0,58 десятичных знаков (десятичных единиц) на символ. Принципиальное преимущество метода арифметического кодирования заключается в том, что передаваемое кодовое число (в нашем случае – это число 0,58), не ограничивается заданием целого числа битов для каждого отдельного символа кодируемой последовательности.

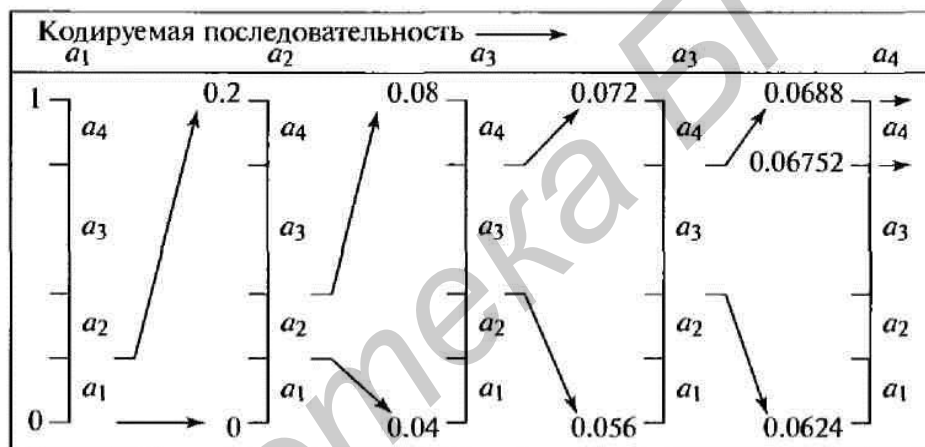


Рисунок 2.12 – Процедура арифметического кодирования

Таблица 2.6 – Пример арифметического кодирования

Символ	Вероятность	Исходный подынтервал
a_1	0,2	[0,0, 0,2)
a_2	0,2	[0,2, 0,4)
a_3	0,4	[0,4, 0,8)
a_4	0,2	[0,8, 1,0)

При увеличении длины кодируемой последовательности, результирующий арифметический код приближается к границе, устанавливаемой теоремой кодирования без шума. На практике два фактора мешают кодовым характеристикам приблизиться к данной границе вплотную: это необходимость включения некоторого символа окончания, позволяющего отделить одну кодовую последовательность от другой, и использование арифметики конечной точности. Для преодоления последней проблемы, при практической реализации арифметического кодирования применяются стратегии масштабирования и округления [41]. Согласно стратегии

масштабирования, каждый подынтервал перед разбиением его на отрезки, пропорциональные вероятностям символов, растягивается до диапазона $[0, 1)$. Стратегия округления гарантирует, что ограничения, связанные с конечной точностью вычислений, не препятствуют точному представлению кодовых подынтервалов.

Эффективность энтропийного кодирования зависит от точности моделей вероятностей символов и способом ее повышения является адаптация к параметрам кодируемой информации (контексту). Контекстное арифметическое кодирование (CAE, Context-based Arithmetic Encoding) основано на изучении локальных пространственных или временных вероятностных характеристик кодируемых последовательностей. Метод CAE используется в стандарте JBIG для сжатия бинарных изображений [43]. Он был также приспособлен для кодирования двоичных «шаблонов» в MPEG-4 Visual [42]. В стандарте MPEG-4 AVC / H.264 используется контекстно-адаптивное двоичное арифметическое кодирование (CABAC – Context-based Adaptive Binary Arithmetic Coding) [44, 45]. Оно достигает хороших результатов по сжатию данных, если выполняются следующие условия: вероятностная модель для каждого синтаксического элемента выбирается в соответствии с его контекстом; адаптивные оценки вероятностей основываются на локальных статистиках.

2.2.4 Кодирование длин серий

Альтернативой использованию кодов переменной длины является представление каждой строки изображения или битовой плоскости последовательностью длин, которая описывает протяженность серий соседних черных или белых пикселей. Этот метод, относящийся к *кодированию длин серий* (КДС), был разработан в 1950-х годах и вместе со своим двумерным расширением стал стандартным способом сжатия в факсимильном (ФАКС) кодировании. Основная идея состоит в том, что при сканировании строки слева направо обнаруживаются непрерывные серии из нулей или единиц, которые затем кодируются кодом их длины; кроме того, устанавливаются соглашения об определении значения каждой серии. Наиболее частыми способами задания значения серии являются следующие: задавать значение первой серии каждой строки, или постановить, что каждая строка начинается с белой серии, однако допустить, что ее длина может быть нулевой.

Хотя кодирование длин серий само по себе является весьма эффективным способом сжатия изображений, обычно можно дополнительно повысить степень сжатия путем неравномерного кодирования самих значений длин серий. К тому же, длины черных и белых серий могут кодироваться по отдельности, используя разные неравномерные коды, каждый из которых оптимизирован по своей статистике. Например, допуская, что символ a_j представляет черную серию длины j , можно оценить вероятность того, что символ a_j может быть порожден гипотетическим источником длин черных серий, путем деления числа черных серий длины j изображения на общее число черных серий.

Оценка энтропии этого источника длин черных серий, обозначаемая H_0 , получается подстановкой этих вероятностей в (2.12). Аналогичным образом можно подсчитать энтропию источника длин белых серий, обозначаемую H_1 . Приближенное значение общей энтропии изображения, кодированного длинами серий, составит

$$H_{дс} = \frac{H_0 + H_1}{L_0 + L_1}, \quad (2.13)$$

где L_0 и L_1 означают средние значения длин черных и белых серий.

Формула (2.13) дает оценку среднего числа битов на элемент изображения, требуемых для представления двоичного изображения, сжатого кодированием длин серий.

Основной недостаток кодирования длин серий заключается в том, что они подходят только для кодирования изображений с двумя градациями яркости. Для сжатия изображений с большим числом уровней яркости данный метод как минимум не эффективен. Поэтому в системах сжатия изображений подобные методы используются для кодирования квантованных коэффициентов, получаемых в результате кодирования с преобразованием. В этом случае действительно часто возникают длинные серии нулей, особенно при высоких степенях сжатия, поэтому кодирование длин серий целиком оправдано.

2.2.5 LZW-кодирование

Рассмотрев основные методы сокращения кодовой избыточности, перейдем к рассмотрению одного из нескольких методов сжатия без потерь, направленного на сокращение межэлементной избыточности изображения. Метод, называемый алгоритмом Лемпеля-Зива-Уэлча (Lempel-Ziv-Welch, LZW), отображает последовательности символов источника различной длины на равномерный код, не требуя априорного знания вероятностей появления кодируемых символов. Он интегрирован во многие широко используемые файловые форматы изображений, включая GIF (graphic interchange format), TIFF (tagged image file format), а также PDF (portable document format).

Концептуально LZW-кодирование базируется на том, что кодовая книга формируется в процессе кодирования последовательности символов источника (т. е. пикселей), и является индивидуальной для каждого сеанса. Каждое слово кодовой книги отображает серию из одного или нескольких символов источника. Формирование кода и его запись в кодовую последовательность происходят лишь тогда, когда при поступлении очередного символа источника образуется новая серия, не присутствующая в кодовой книге, а значит, не отображаемая ни одним из имеющихся кодовых слов. При этом в кодовую последовательность записывается код предыдущей, ранее уже известной серии, и осуществляется пополнение кодовой книги новой серией. Затем кодер начинает анализировать следующую последовательность символов, начиная с то-

го, на котором оборвалась предыдущая серия. Перед началом кодирования строится короткая тривиальная и заранее известная кодовая книга, а алгоритме пополнения дает возможность и кодеру и декодеру строить одинаковые кодовые книги и передавать по каналу лишь уже известные к настоящему моменту кодовые слова.

При запуске процесса кодирования строится начало кодовой книги или «словарь», содержащий лишь кодируемые символы источника. Для 8-битового монохромного изображения словарь имеет размеры в 256 слов и отображает значения яркостей 0, 1, 2, ..., 255. Кодер последовательно анализирует символы источника (т. е. значения пикселей), и при появлении отсутствующей в словаре серии, она помещается в определяемую алгоритмом (следующую свободную) позицию словаря. Если первые два пикселя изображения, например, были белыми (255-255), эта серия может быть приписана позиции 256, являющейся следующей свободной после зарезервированных для уровней яркостей позиций с 0 по 255. В следующий раз, когда встретится серия из двух белых пикселей, для их представления будет использовано кодовое слово 256, как адрес позиции, содержащей серию 255-255. В случае 9-битового словаря, содержащего 512 кодовых слов, исходные $8 + 8 = 16$ битов, требуемые для представления двух пикселей, будут заменены одним 9-битовым кодовым словом. Ясно, что допустимый размер словаря является важнейшим параметром. Если он слишком мал, то обнаружение совпадающих серий яркостей будет маловероятно; если слишком велик, то размер кодового слова будет ухудшать характеристики сжатия.

В качестве примера рассмотрим кодирование блока 8-битового изображения размером 4x4, матрица значений (отсчетов) которого приведена ниже:

39	39	126	126
39	39	126	126
39	39	126	126
39	39	126	126

Изначально подготавливается словарь на 512 кодовых слов (таблица 2.7). В начальный момент позиции с 256 по 511 еще не используются. При кодировании пиксели изображения обрабатываются слева направо и сверху вниз. Осуществляется присоединение каждого следующего значения яркости к имеющейся на данный момент серии, называемой «распознанная серия», которая приведена в позиции 1 таблицы 2.8.

Как можно увидеть, вначале эта переменная обнулена или пуста. Словарь просматривается на обнаружение совпадения с каждой очередной серией, и если таковая обнаруживается, что и отмечено в первой строке таблицы, то серия заменяется кодом (номером позиции) совпадающей и распознанной (т. е. имеющейся в словаре) серии, что отмечено в первой колонке второй строки. При этом еще не порождается никакого кода и не происходит обновления словаря. Если же совпадения серии и словаря не обнаруживается (что отмечено во второй строке таблицы), то номер позиции распознанной к настоящему моменту серии (39) подается на выход в качестве очередного кода: текущая

нераспознанная серия пополняет словарь, а состояние распознанной серии иницируется последним поступившим символом.

Таблица 2.7 – Исходное состояние таблицы кодирования

Позиция в словаре	Содержание (серия)
0	0
1	1
2	2
...	...
255	255
256	-
...	
511	-

Таблица 2.8 – Пример LZW-кодирования

Распознанная серия	Обрабатываемый пиксель	Выход кодера	Позиция слова в словаре	Входы словаря
	39			
39	39	39	256	39-39
39	126	39	257	39-126
126	126	12	258	126-126
126	39	12	259	126-39
39	39			
39-39	126	25	260	39-39-126
126	126			
126-126	39	25	261	126-126-39
39	39			
39-39	126			
39-39-	126	26	262	39-39-126-
126	39			
126-39	39	25	263	126-39-39
39	126			
39-126	126	25	264	39-126-126
126		12		

Последние две колонки таблицы описывают коды и серии яркостей, которые последовательно добавляются к словарю при кодировании всего изображения размерами 4x4 элемента. Добавляются девять дополнительных кодовых слов. По завершении кодирования словарь содержит 265 кодовых слов; при этом LZW-алгоритм успешно обнаружил несколько повторяющихся серий яркостей, что позволило ему сократить исходное 128-битовое изображение до 90-битового изображения (т. е. до 10 кодов из 9 битов). Кодовая последовательность на выходе образуется при чтении третьей колонки (Выход кодера) сверху вниз. Результирующий коэффициент сжатия равен 1,42:1.

Уникальным качеством только что продемонстрированного LZW-кодирования является то, что кодовая книга (словарь) создается в процессе кодирования данных. Примечательно, что LZW-декодер строит идентичный словарь восстановления, если он декодирует поток данных синхронно с кодером.

2.3 Кодирование с предсказанием

2.3.1 Классификация методов кодирования с предсказанием

Методы кодирования с предсказанием основаны на устранении статистической межэлементной избыточности – как пространственной, так и временной. Их общий принцип заключается в передаче не исходных значений сигнала (отсчетов либо спектральных коэффициентов), а разности между текущим и предсказанным значением. Кодер с предсказанием может использоваться для компрессии изображений как непосредственно, так и входить в состав комплексного гибридного кодека (см. подраздел 2.5).

Операция кодирования с предсказанием имеет массу разновидностей (рисунок 2.13) и заложена в большинство современных алгоритмов компрессии изображений.



Рисунок 2.13 – Классификация методов кодирования с предсказанием

Предсказание может выполняться как для каждого элемента изображения, так и для блоков. Поэлементное предсказание использовалось в первых методах сокращения скорости передачи данных цифровых ТВ сигналов [7] на

базе классической ДИКМ. В современных алгоритмах компрессии динамических изображений поэлементное предсказание выполняется для вычисления разности между отсчетами соседних кадров (временное предсказание). В иных случаях выполняется блочное предсказание, при котором обработке подвергается фрагмент изображения фиксированного либо переменного размера.

Среди используемых моделей предсказания наиболее простой является линейная, в которой прогнозирование осуществляется только на основе предыдущего элемента – соседнего по горизонтали (по строке), по вертикали либо по времени (имеющего одинаковые координаты в соседних кадрах). Большую точность обеспечивает линейное предсказание по нескольким предыдущим элементам, расположенным в одинаковом направлении с предсказываемым.

Более сложной в реализации, но и более эффективной с точки зрения повышения точности предсказания является двумерная пространственная модель, в которой в качестве прогноза используется комбинация элементов, расположенных по разные стороны от предсказываемого.

Предсказание может быть одно- и двунаправленным. В первом случае прогноз строится только на основе предыдущих значений, во втором – на основе предыдущих и последующих. Примером использования двунаправленного временного предсказания является вычисление разностных В-кадров в стандартах MPEG-1, MPEG-2 и MPEG-4 AVC.

Еще одной мерой повышения точности предсказания является компенсация движения. При ее использовании дополнительно оценивается и передается разность между позициями предсказанного и фактического блоков изображения – вектор движения.

2.3.2 Общие принципы кодирования с предсказанием

Упрощенная структурная схема системы кодирования с предсказанием показана на рисунке 2.14. В такой системе значение каждого m -разрядного элемента изображения $f_{\text{ПР}}(x, y)$ предсказывается на основании сведений о предшествующих по ходу развертки элементах $f(x-k, y)$, $f(x-k+1, y)$, ..., $f(x-2, y)$, $f(x-1, y)$. Количество K элементов, участвующих в процесс предсказания, влияет на точность работы и сложность реализации блока предсказания и на практике обычно не превышает 2...3, а в простейшем случае прогноз строится на основании одного предшествующего элемента $f(x-1, y)$.

При использовании двумерной (пространственной) модели предсказания в построении прогноза участвуют также и элементы, расположенные в предыдущих строках (рисунок 2.15). В общем виде модель предсказания может быть представлена в виде:

$$f_{\text{ПР}}(x, y) = \text{round} \left[\sum_{k=1}^K \sum_{l=1}^L \alpha_{k,l} f(x-k, y-l) \right], \quad (2.14)$$

где K и L – порядок предсказания (количество элементов, по которым ведется предсказание по осям x, y соответственно;

$\alpha_{k,l}$ – весовые коэффициенты отсчетов, использующихся для предсказания;

round – операция округления до ближайшего целого.

В случае одномерного предсказания функция $f_{\text{ПР}}(x, y)$ может быть упрощена до $f_{\text{ПР}}(t)$, если предсказываются элементы вдоль строки, либо до $f_{x,y}(nT_K)$, где x, y – координаты элемента, T_K – период кадра.

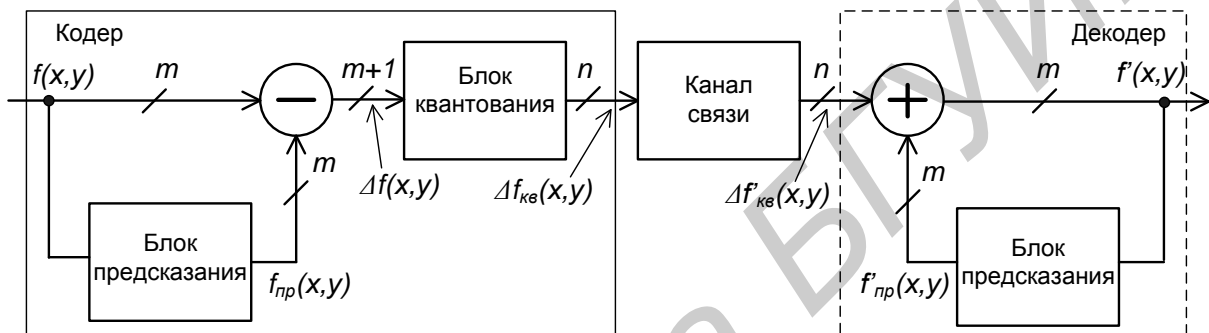


Рисунок 2.14 – Упрощенная структурная схема системы кодирования с предсказанием

Предсказанное значение $f_{\text{ПР}}(x, y)$ вычитается из фактического значения $f(x, y)$, затем разностный сигнал $\Delta f(x, y)$ подвергается повторному квантованию с уменьшением разрядности до n и передается по каналу связи. Коэффициент компрессии в данном случае равен отношению разрядностей входного и выходного сигналов m/n . На приемной стороне по квантованному разностному сигналу $\Delta f'_{\text{кв}}(x, y)$ восстанавливается исходный сигнал.

При построении предсказателя стремятся к минимизации СКО:

$$E\{\varepsilon^2(x, y)\} = E\{[f'(x, y) - f(x, y)]^2\}, \quad (2.15)$$

где $E\{*\}$ – математическое ожидание, при условии, что ошибка восстановления сигнала $f'(x, y)$ на выходе декодера пренебрежимо мала – т. е. выполняется соответствие:

$$f'(x, y) = \varepsilon(x, y) + f_{\text{ПР}}(x, y) \approx \varepsilon_{\text{КВ}}(x, y) + f_{\text{ПР}}(x, y) = f(x, y), \quad (2.16)$$

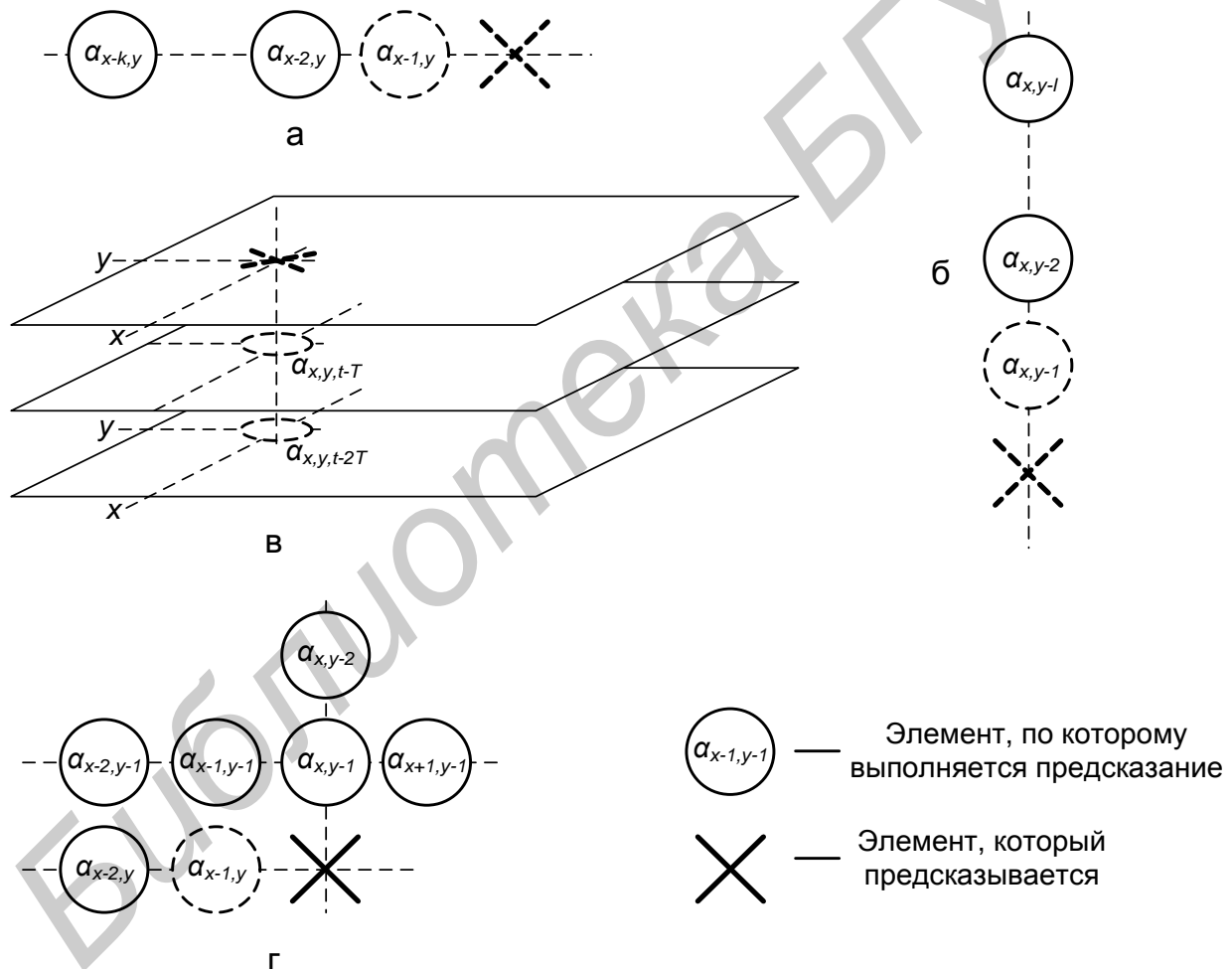
где $\varepsilon_{\text{КВ}}(x, y)$ – ошибка квантования, а значение $f_{\text{ПР}}(x, y)$ вычислено по (2.14).

Еще одним условием является ограничение суммы коэффициентов предсказания $\alpha_{k,l}$ значением 1: $\sum_{k=1}^K \sum_{l=1}^L \alpha_{k,l} = 1$.

Данное ограничение гарантирует, что значение сигнала на выходе предсказателя не превысит допустимый диапазон, а также снижает влияние помех, возникающих в канале связи, воздействие которых проявляется в виде тянущихся продолжений за пораженными элементами. Для сокращения времени восстановления при воздействии помех сумма коэффициентов предсказания должна быть меньше 1 [16].

При выполнении условия (2.16) проблема построения оптимального предсказателя сводится к задаче выбора коэффициентов предсказателя $\alpha_{k,l}$, минимизирующих выражение

$$E\{\varepsilon^2(x, y)\} = E\left\{ \left[f(x, y) - \sum_{k=1}^K \sum_{l=1}^L \alpha_{k,l} f(x-k, y-l) \right]^2 \right\}. \quad (2.17)$$



:

а – внутрикадровое одномерное предсказание по горизонтали (строке);
 б – внутрикадровое одномерное предсказание по вертикали; в – межкадровое предсказание; г – двумерное пространственное предсказание

Рисунок 2.15 – Варианты моделей предсказания

Для минимизации среднеквадратической ошибки предсказания следует вычислить частные производные функции $E\{\varepsilon^2(x, y)\}$ по каждой из переменных $\alpha_{k,l}$, приравнять каждую производную нулю, после чего решить систему из $K \cdot L + 1$ уравнений [46]. Несмотря на внешнюю простоту данного способа, практическая реализация его является чрезвычайно громоздкой и поэтому применяется крайне редко. В большинстве случаев оптимизация выполняется подбором коэффициентов.

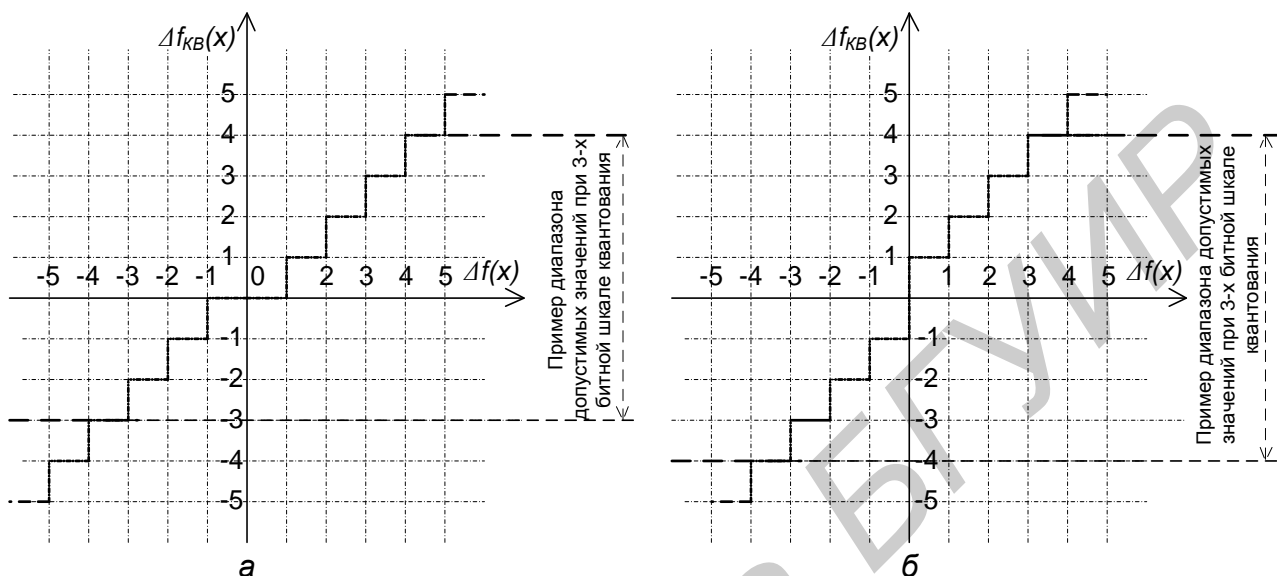
Еще одной не менее важной задачей является построение квантователя и, в частности, выбор шкалы квантования – соотношения между значениями выходного и входного сигналов. В задачах компрессии изображений на вход квантователя поступает уже квантованный сигнал, поэтому задача сводится к преобразованию кодовых комбинаций. Существует большое количество реализаций квантователя, мы же рассмотрим наиболее простой случай – линейный, при котором значения выходного сигнала равны значением входного до наступления порога ограничения. Разностный сигнал $\Delta f(x, y)$ на входе квантователя является двуполярным, что приводит к двум вариантам реализации шкалы квантования – с наличием нулевого значения (рисунок 2.16, а) и без него (рисунок 2.16, б)¹. В первом случае значения разностного сигнала, лежащие в пределах интервала $(-1, 1)$ принимаются равными 0. Для цифрового сигнала $\Delta f(x, y)$ это означает, что значению «0» на входе квантователя соответствует «0» на выходе, «1» – «1» и т. д. С учетом того, что количество уровней квантования кратно 2^m , при полном использовании шкала становится несимметричной – например, при 3-разрядном квантовании ($M = 2^3 = 8$ уровней) может быть закодировано 8 значений – либо $[-4; -3; -2; -1; 0; 1; 2; 3]$, либо $[-3; -2; -1; 0; 1; 2; 3; 4]$, что ведет к различной перегрузочной способности системы к отрицательным и положительным перепадам сигнала, особенно заметной при малых разрядностях, а при $m = 1$ (дельта-модуляция) такая шкала вовсе неприменима.

Во втором варианте шкалы значение «0» в выходном сигнале отсутствует, сигнал в интервале $(0; 1]$ преобразуется в «1» на выходе, в интервале $(-1; 0]$ – в «-1» и т. д., а при квантовании разностного сигнала, равного «0», возникает неопределенность выбора соответствующего выходного кода «-1» или «1». Практически неопределенность может решаться любым из двух вариантов, но одинаково в кодере и декодере.

С параметрами шкал квантования связаны специфические искажения сигнала – перегрузка по крутизне и гранулярный шум (рисунок 2.17). Первая объясняется ограничением шкалы (см. примеры на рисунках 2.16, а, б) и возникает на резких перепадах сигнала, второй обусловлен отсутствием нулевого значения выходного сигнала (рисунок 2.16, б) и проявляется на плавных переходах и участках сигнала с постоянной амплитудой. Применительно к

¹ Для упрощения анализа и для краткости на данном рисунке и некоторых последующих двумерный сигнал $f(x, y)$ заменен одномерным $f(x)$.

кодированию яркостного и цветоразностных сигналов перегрузка по крутизне вызывает размытие фронтов на резких перепадах яркости (либо резких изменениях цветности), что в результате приводит к снижению четкости изображений. Гранулярный шум бывает заметен на участках с плавно меняющейся либо постоянной яркостью и ухудшает субъективное восприятие изображения.



а – шкала квантования с нулевым значением;
б – шкала квантования без нулевого значения

Рисунок 2.16 – Варианты реализации шкал квантования



Рисунок 2.17 – Искажения сигнала в системах с предсказанием

Возможны два принципиально различных метода реализации систем с предсказанием, которые следует рассмотреть с точки зрения их применимости. Один из этих методов – прямое предсказание по разомкнутой схеме (feedforward), а другой – предсказание с обратной связью по замкнутой схеме (feedback).

Система с предсказанием по разомкнутой схеме соответствует приведенной на рисунке 2.14 и строит оценку в виде линейной комбинации уров-

ней предшествующих элементов, обращая за этими данными к запоминающему устройству. Основная трудность при построении систем с прямым предсказанием состоит в том, что декодирующий блок системы в отличие от кодирующего блока в этом случае не имеет доступа к зарегистрированной в запоминающем блоке совокупности фактических уровней предыдущих элементов; декодер может оперировать только восстановленными уровнями, значения которых в результате квантования определены с погрешностью. Таким образом, значения $f(x, y)$ и $f'(x, y)$ элементов, полученные в передатчике и приемнике, могут различаться между собой, что в конечном счете приведет систему к потере устойчивости – появлению ошибки восстановления, которая не может быть скомпенсирована.

Эта трудность может быть устранена предсказанием с обратной связью. Данная схема реализует принцип классической дифференциальной ИКМ поэтому далее для краткости будем называть ее ДИКМ-системой. В ней квантователь находится внутри петли обратной связи (рисунок 2.18), в результате в передатчике получается оценка в виде линейной комбинации прошлых ошибок предсказания. В результате передатчик и приемник будут давать одинаковые оценки, поскольку в этом случае работу предсказателей как передатчика, так и приемника будет определять одинаковый квантованный разностный сигнал. Контур обратной связи, построенный по такой схеме, может быть с одинаковым успехом спроектирован в расчете на достижение минимума как ошибки квантования, так и дисперсии разностного сигнала.

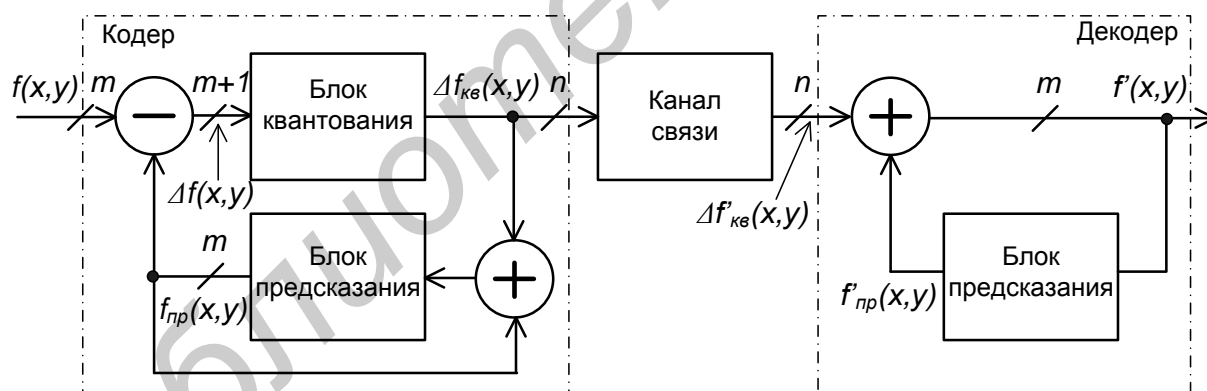


Рисунок 2.18 – Структурная схема системы кодирования с предсказанием с квантователем внутри петли обратной связи

Анализ работы схем с прямым предсказанием и с квантователем в петле обратной связи приводится на рисунках 2.19 – 2.21. В обеих схемах использована простейшая модель предсказания – по одному предыдущему элементу – и линейная шкала квантования. Результаты анализа показывают, что точное восстановление сигнала в любом случае возможно только при разрядности двоичного кодирования, превышающей разрядность входного сигнала минимум на 1 бит (рисунок 2.19). Следовательно, для компрессии без потерь

системы с предсказанием непосредственно (без использования дополнительных методов кодирования) не применимы.

При разрядностях квантования, равных или меньших разрядности входного сигнала, между результатами работы систем с квантователем вне и внутри контура обратной связи начинают проявляться серьезные различия. При восстановлении сигнала в первом случае появляются ошибки, которые не компенсируются – особенно это заметно при 3-х битном квантовании (рисунок 2.22, б). Частичная компенсация ошибок на рисунках 2.19, б и 2.20, а объясняется разным знаком ошибок при их одинаковом значении – данное совпадение случайно.

При расположении квантователя внутри петли обратной связи (классическая схема ДИКМ) ошибка восстановления постепенно компенсируется (при условии, что разностный сигнал $\Delta f(x)$ не выходит за пределы шкалы квантования. Если используется шкала квантования без нулевого значения (см. рисунок 2.16, б), на участках сигнала с медленно меняющейся либо постоянной амплитудой проявляется гранулярный шум (см. восстановленные сигналы $f'(x)$ на рисунке 2.21, б). Что касается ошибки восстановления, то она проявляется в любом случае, когда разностный $\Delta f(x)$ сигнал превышает динамический диапазон квантователя.

В системах с квантователем внутри петли обратной связи происходит компенсация ошибок квантования и, если разность между соседними элементами входного сигнала не превышает пределов шкалы квантования, происходит постепенное восстановление правильного сигнала на выходе декодера. На рисунке 2.22 представлены результаты непосредственного использования кодирования с предсказанием для сокращения избыточности реальных изображений.

Разрядность квантования разностного сигнала в первом случае выбрана 4 бита на отсчет, во втором – 3 бита на отсчет, что соответствует коэффициентам компрессии 2 и 2,67. Видно, что качество восстановленного изображения в обоих случаях достаточно низкое, а различия между исходным и восстановленным изображениями – существенны. Искажения в данном случае обусловлены исключительно перегрузкой по крутизне ДИКМ-кодера.

Анализ искажений удобнее проводить по испытательным изображениям с регулярной структурой, например, по изображению, представленному на рисунке 2.23. На этом же рисунке представлены разностные изображения, отображающие ошибку восстановления исходного изображения в зависимости от принятой модели предсказания – одномерной горизонтальной, одномерной вертикальной и двумерной. Заметно, что точность восстановления выше, а ошибка, соответственно, менее заметна, если структура изображения (в данном случае направление линий) совпадает с направлением предсказания. Также можно сделать вывод о невозможности построения модели предсказания, одинаково оптимальной для изображений с различной структурой.

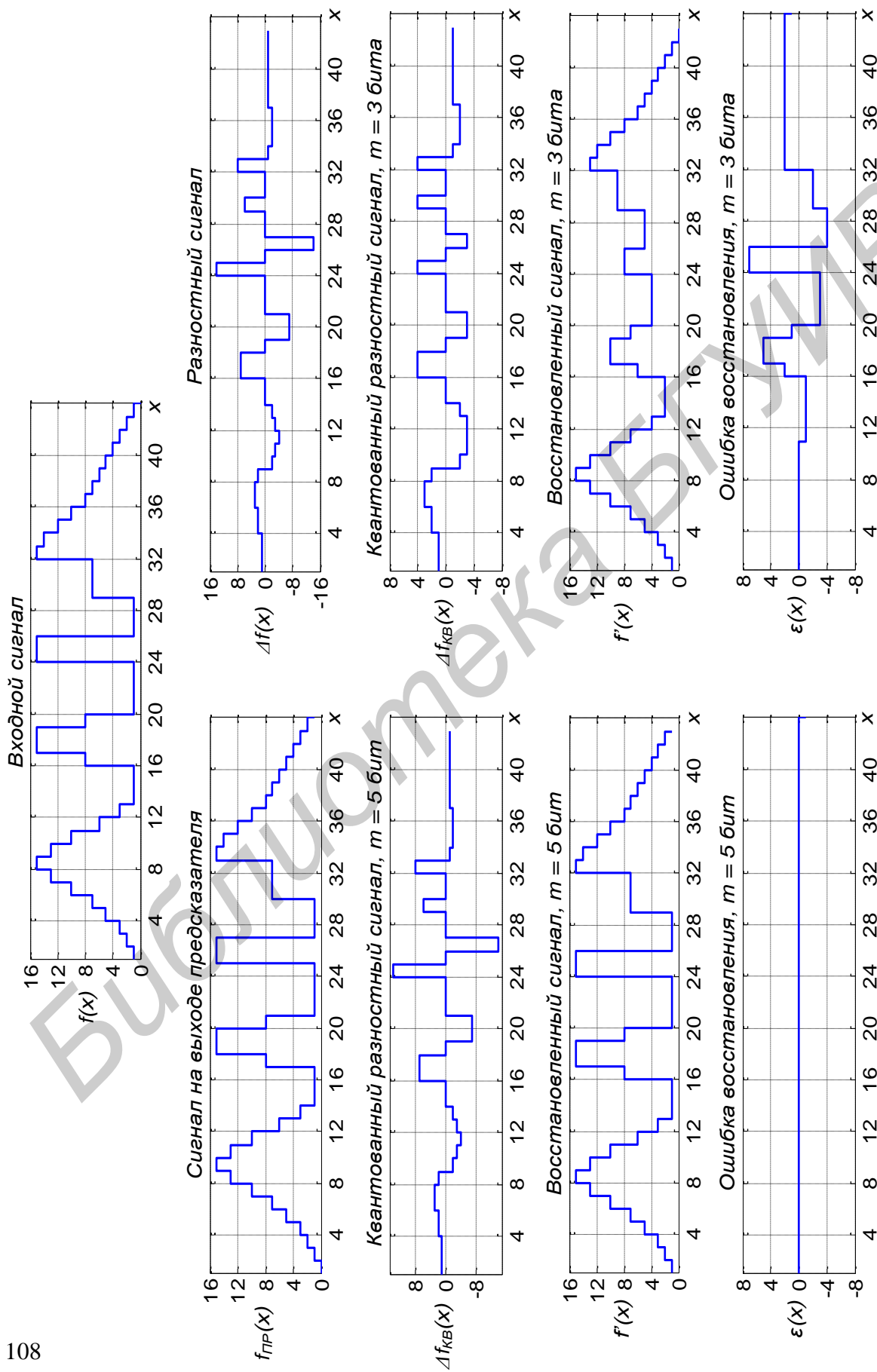
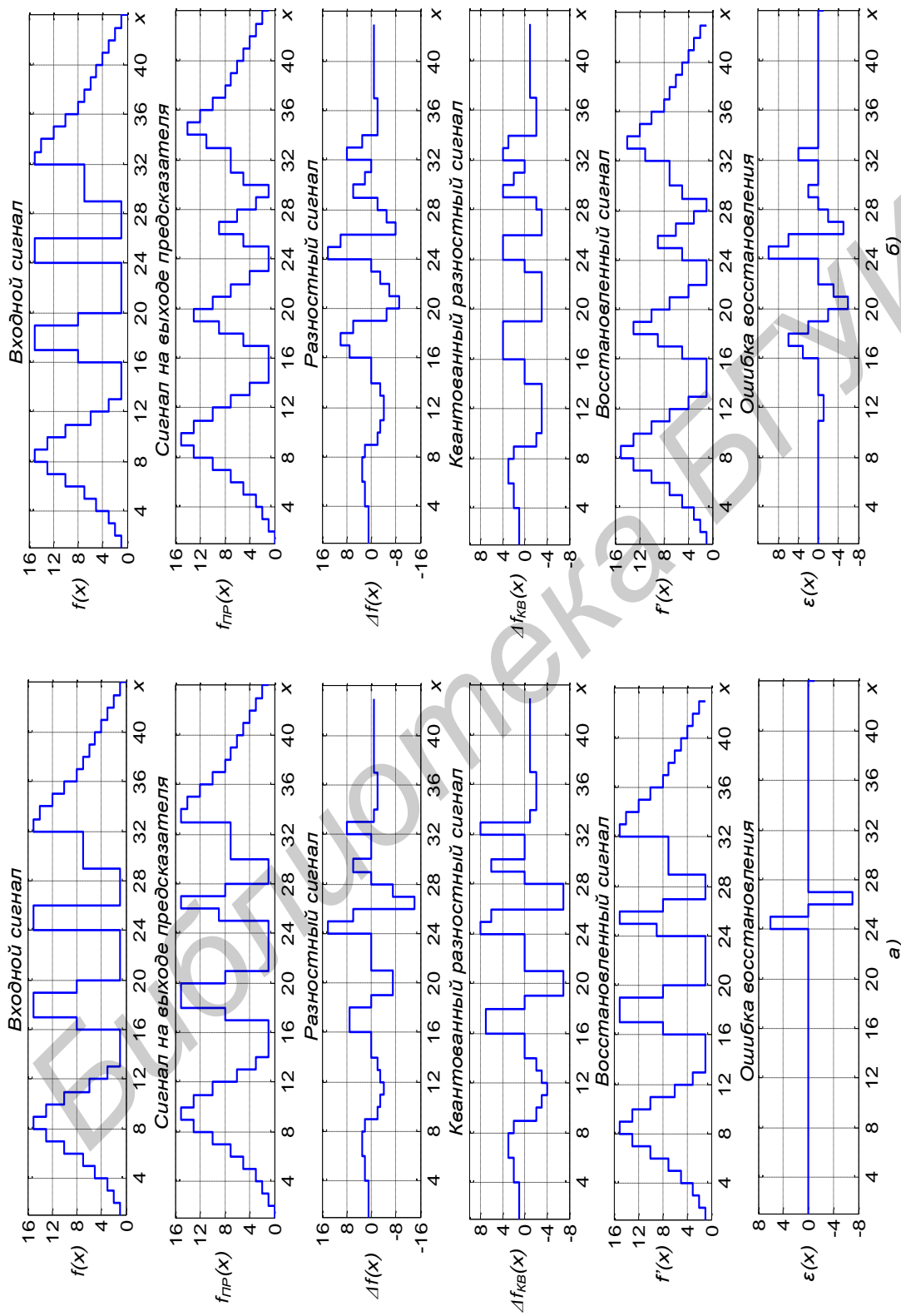
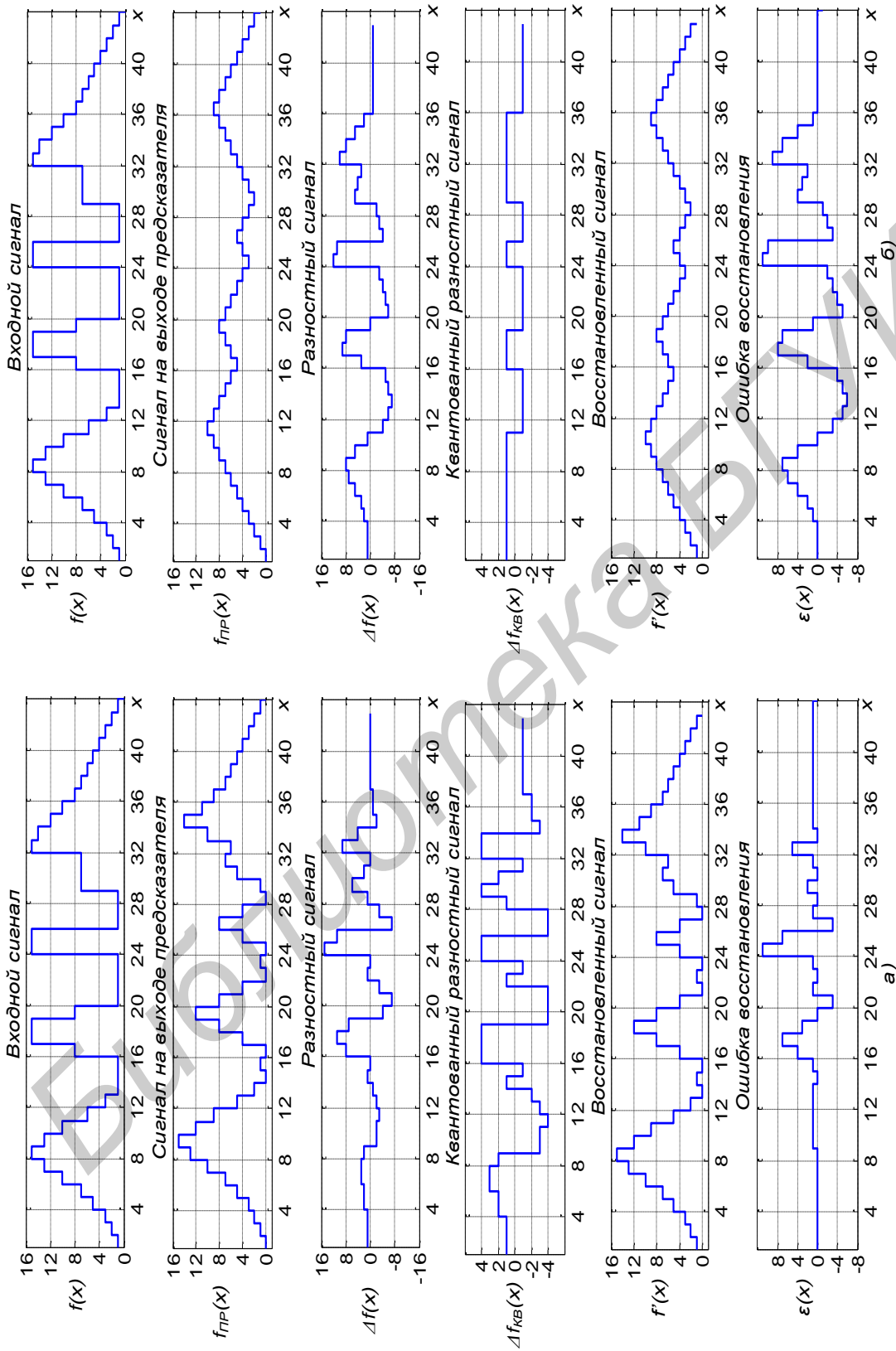


Рисунок 2.19 – Диаграммы работы системы кодирования с предсказанием с квантователем вне контура обратной связи



а – разрядность двоичного кодирования 4 бита/отсчет; б – разрядность двоичного кодирования 3 бита/отсчет

Рисунок 2.20 – Диаграммы работы системы кодирования с предсказанием с квантователем внутри контура обратной связи и шкалой квантования с нулевым значением



а – разрядность двоичного кодирования 3 бита/отсчет; б – разрядность двоичного кодирования 1 бит/отсчет

Рисунок 2.21 – Диаграммы работы системы кодирования с предсказанием с квантованием и внутри контура обратной связи и шкалой квантования без нулевого значения

Исходное изображение



Восстановленное изображение, $t = 4$ бита



Восстановленное изображение, $t = 3$ бита



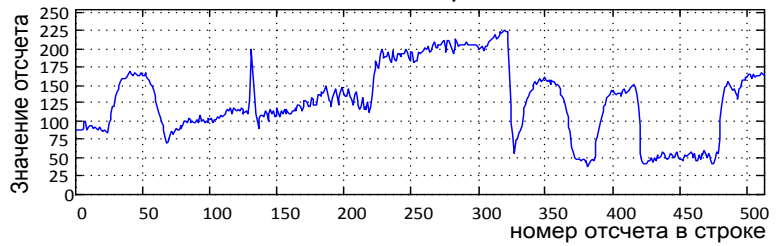
Ошибка восстановления, $t = 4$ бита



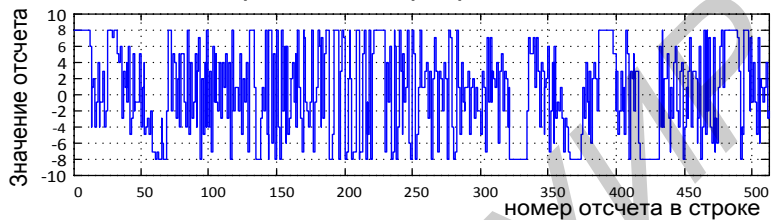
Ошибка восстановления, $t = 3$ бита



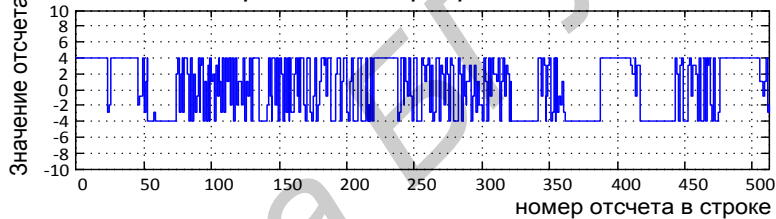
Значения отсчетов сигнала одной строки исходного изображения



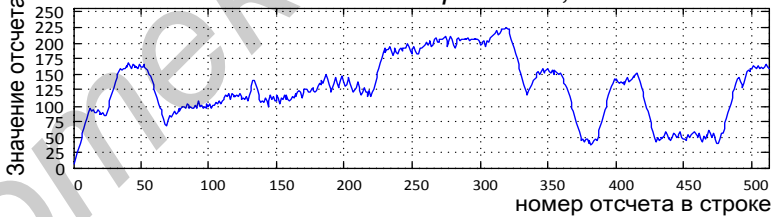
Сигнал, кодированный с разрядностью $t = 4$ бита



Сигнал, кодированный с разрядностью $t = 3$ бита



Значения отсчетов сигнала одной строки восстановленного изображения, $t = 4$ бита



Значения отсчетов сигнала одной строки восстановленного изображения, $t = 3$ бита

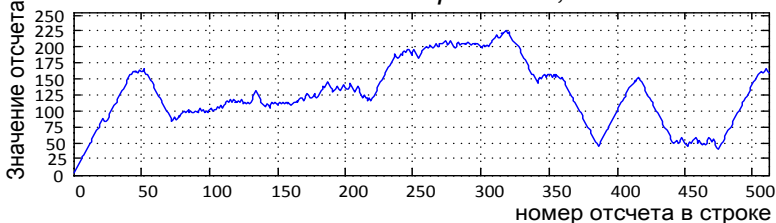


Рисунок 2.22 – Результаты использования кодирования с предсказанием

Таким образом, непосредственному использованию ДИКМ для внутрикадровой компрессии изображений присущи следующие недостатки:

- малая степень компрессии при удовлетворительном качестве изображений;
- невозможность компрессии без потерь.

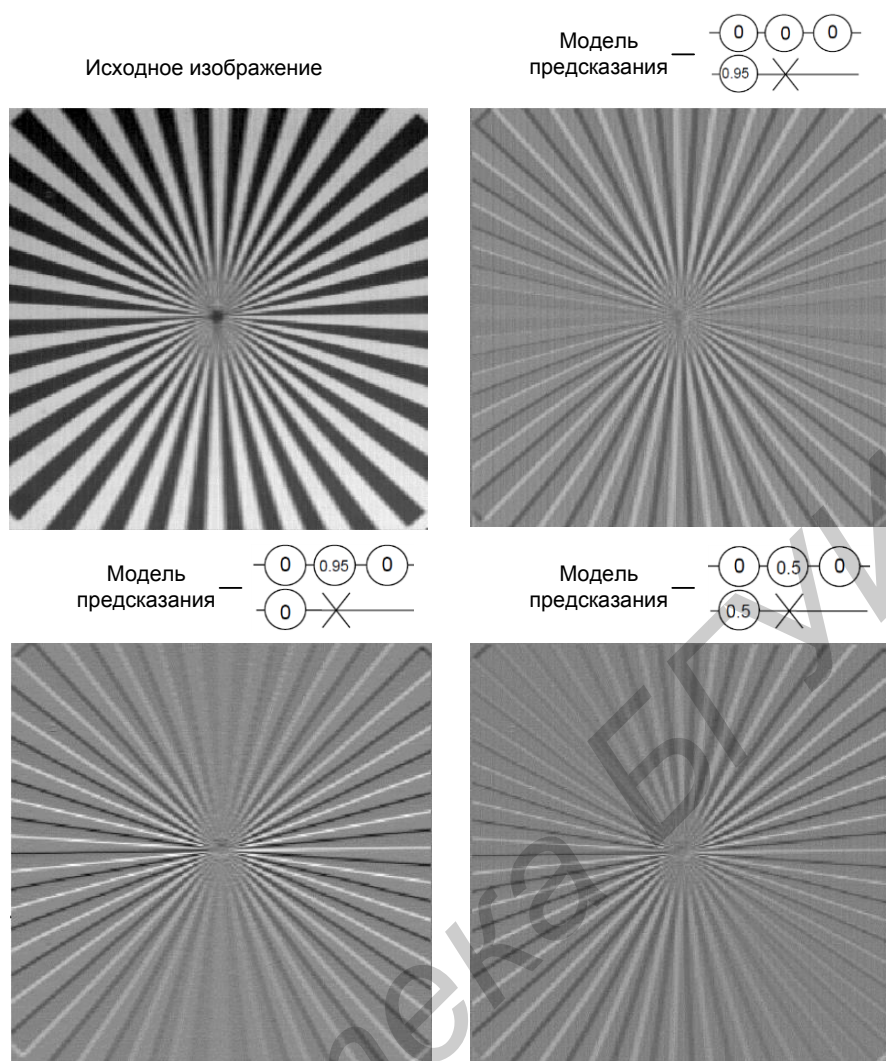


Рисунок 2.23 – Влияние модели предсказания на ошибку восстановления изображения

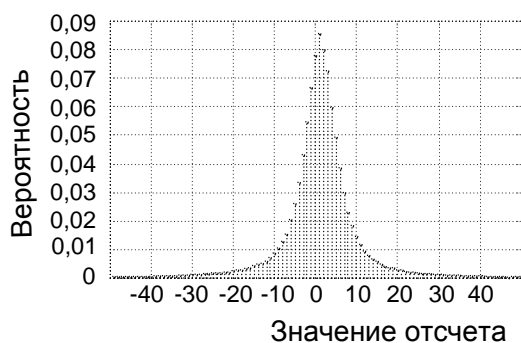
Для нахождения путей эффективного использования дифференциального кодирования при компрессии изображений проанализируем разностное изображение $\Delta f(x, y)$ на выходе вычитателя ДИКМ-кодера (см. рисунок 2.18) при отсутствии в схеме квантователя. В качестве тестового удобно использовать уже рассмотренные ранее изображения, например, приведенное на цветной вклейке (рисунок Б.2). Анализ разностного изображения (рисунок 2.24, а, б) показывает существенно сузившиеся гистограммы распределения вероятности его мгновенных значений (рисунок 2.24, в, г) как для горизонтальной модели предсказания (см. рисунок 2.15, а), так и для вертикальной (см. рисунок 2.15, б). Соответственно, снизилась и энтропия – до 4,82 при горизонтальной модели и до 4,66 – при вертикальной. Для сравнения, энтропия сигнала яркости исходного изображения равнялась 7,45 (см. рисунок Б.2).



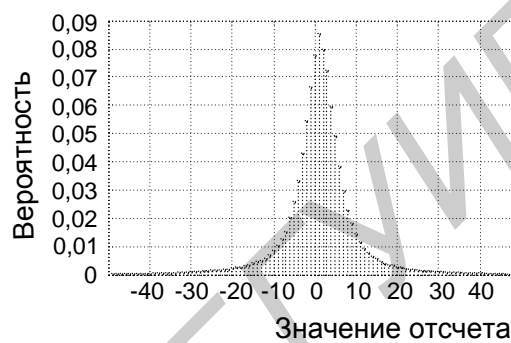
а



б



в



г

а, б – разностные изображения при использовании горизонтальной и вертикальной моделей предсказания соответственно; в, г – распределение вероятности значений отсчетов разностных изображений при использовании горизонтальной и вертикальной моделей предсказания соответственно

Рисунок 2.24 – Статистические характеристики разностного изображения на выходе вычитателя ДИКМ-кодера

Таким образом, напрашивается замена квантователя статистическим (энтропийным) кодером, которая позволит реализовать сжатие без потерь. В связи с тем, что в процессе статистического кодирования/декодирования потеря информации не происходит, кодер может быть установлен как внутри, так и вне петли обратной связи (рисунок 2.25). Однако если статистический кодер располагается внутри петли обратной связи, то в состав ДИКМ-кодера должен входить и статистический декодер, восстанавливающий отсчеты разностного изображения для блока предсказания. ДИКМ-декодер в обоих случаях остается одинаковым.

Рассматриваемая система компрессии позволяет обеспечить сжатие без потерь в 1,25...2 раза, однако необходимость передачи каждого отсчета изображения не позволяет даже при кодировании с потерями добиться эффективности сжатия, лучшей чем 1 бит/отсчет.

Преодолеть рубеж в 1 бит/отсчет возможно при объединении кодирования с предсказанием с кодированием с преобразованием. Для этого внутри петли ОС ДИКМ-кодека вводится блок кодирования с преобразованием (рисунок 2.26). Подробно принцип кодирования с преобразованием рассматривается в следующих подразделах, здесь же рассмотрим изменения в реализации блока предсказания.

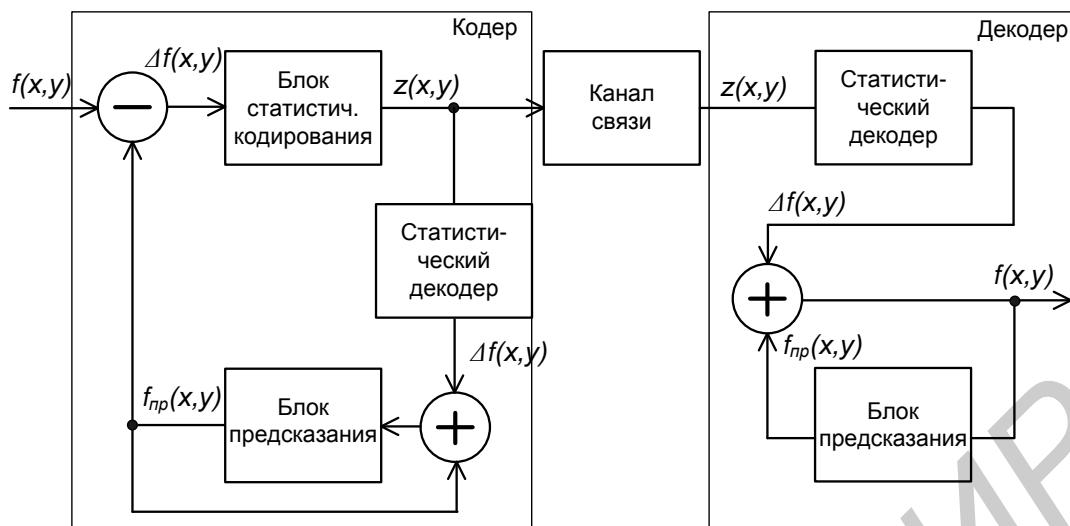
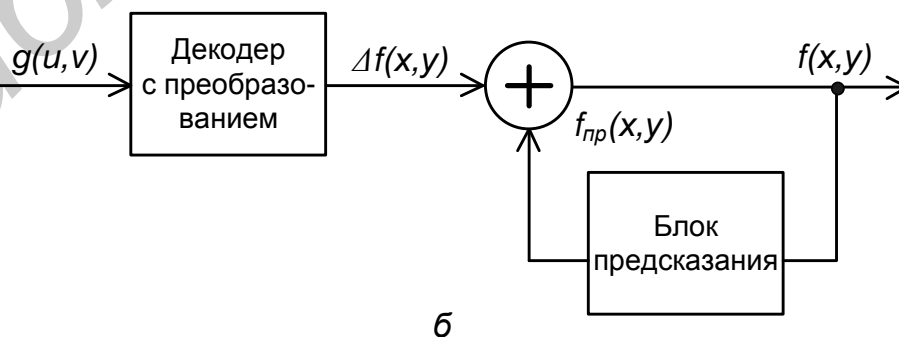
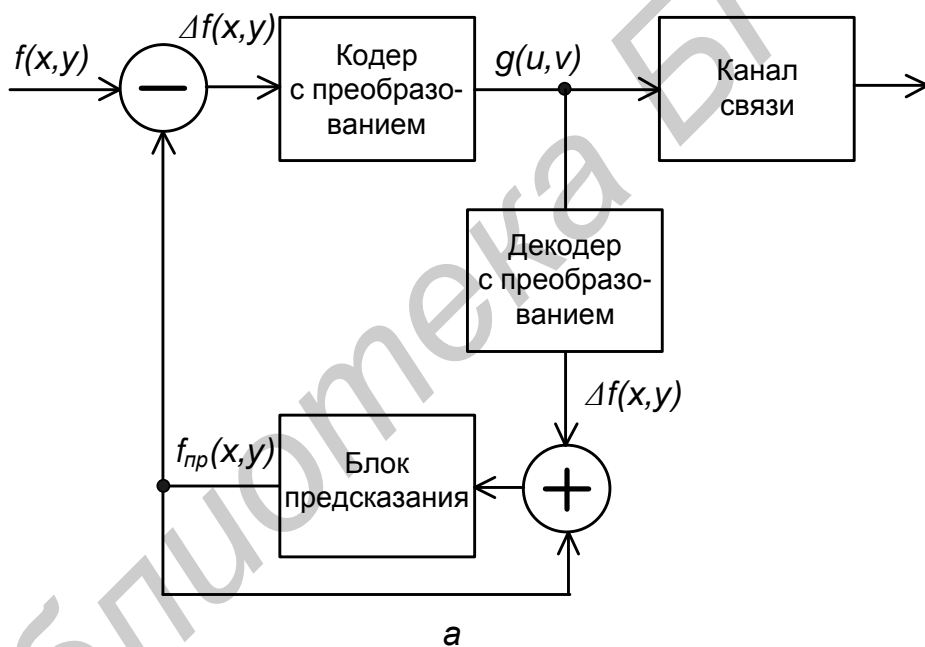


Рисунок 2.25 – Комбинированная схема компрессии изображений на базе дифференциального и статистического кодирования



а – кодер; б – декодер

Рисунок 2.26 – Общая схема объединения кодирования с предсказанием и кодирования с преобразованием

Проблемой блочного предсказания является гораздо меньшая корреляция между соседними блоками изображения по сравнению с корреляцией между соседними элементами, причем чем больше размер блока, тем она меньше, т. е. обратно пропорциональна размеру блока. Также существенно усложняется модель предсказания. В модели (2.15) для поэлементного предсказания предсказывается только один элемент, модель для блочного предсказания будет представлять систему из $N = X_{\text{бл}} \times Y_{\text{бл}}$ функций:

$$f_{\text{пр}}(x_{\text{бл}}, y_{\text{бл}}) = \begin{cases} \text{round} \left[\sum_{k=1}^K \sum_{l=1}^L \alpha_{k,l} f(x_1 - k, y_1 - l) \right] \\ \vdots \\ \text{round} \left[\sum_{k=1}^K \sum_{l=1}^L \alpha_{k,l} f(x_{X_{\text{бл}}} - k, y_{Y_{\text{бл}}} - l) \right] \end{cases}, \quad (2.18)$$

где K и L – порядок предсказания (количество элементов, по которым ведется предсказание по осям x , y соответственно;

$\alpha_{k,l}$ – весовые коэффициенты отсчетов, использующихся для предсказания;

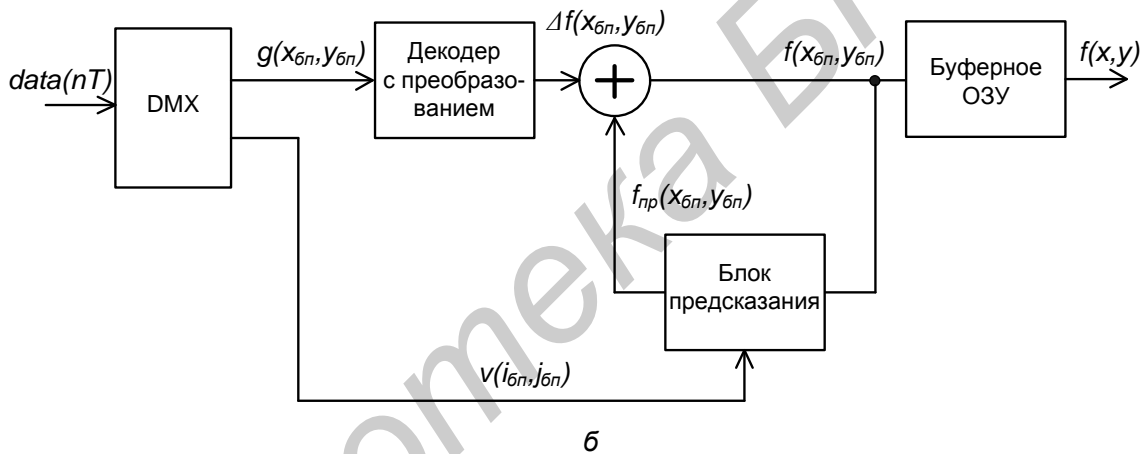
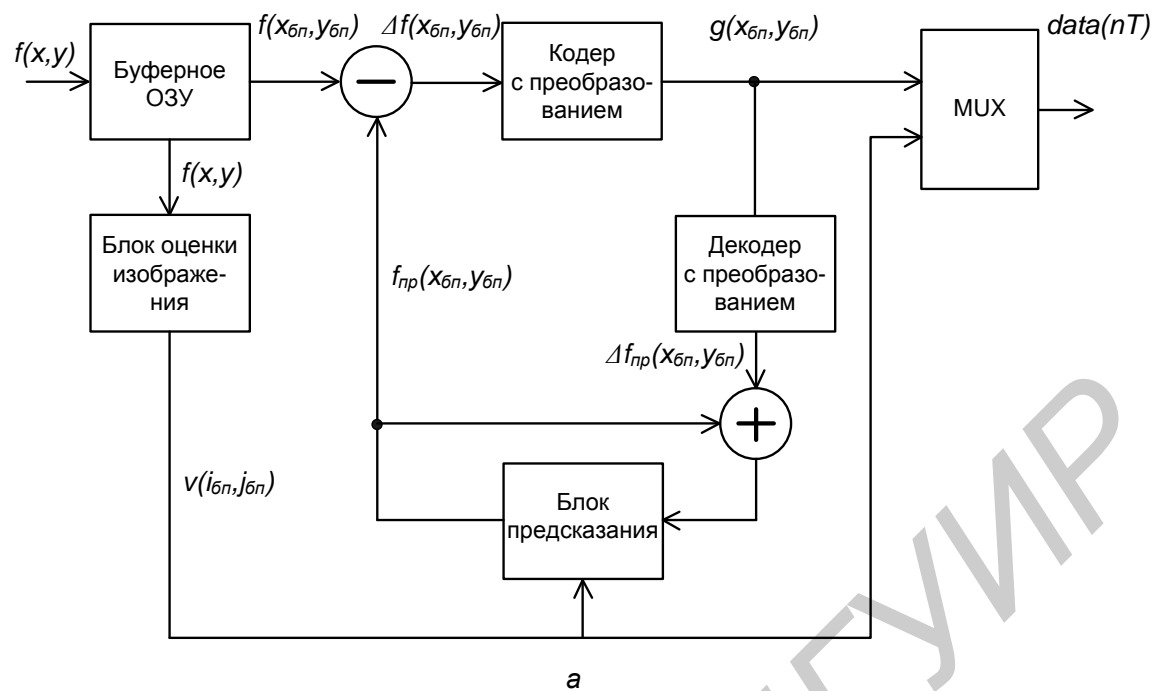
$X_{\text{бл}}, Y_{\text{бл}}$ – размеры предсказываемого блока по горизонтали и вертикали;

round – операция округления до ближайшего целого.

Соответственно, значительно уменьшается вероятность того, что модель предсказания будет одинаково оптимальной для всех блоков изображения. В результате приходится использовать несколько моделей предсказания, из которых для каждого предсказываемого блока выбирается та, которая дает наименьшую ошибку предсказания. Наиболее распространенным критерием является минимум СКО разностного блока, хотя могут использоваться и другие (см. подраздел 2.4).

Информация о выбранной модели предсказания должна быть передана декодеру. Для сокращения ее количества перечень используемых размеров блоков и моделей предсказания ограничивается, например в стандарте MPEG-4 AVC/H.264 [67] используются 3 размера блоков предсказания (4×4 , 8×8 и 16×16) и 9 моделей предсказания (4 для блоков размером 16×16). Количество доступных моделей зависит от положения предсказываемого блока на изображении.

Выбор модели предсказания производится во входящем в состав кодера блоке оценки изображения (рисунок 2.27). Алгоритмов выбора модели предсказания разработано большое количество, начиная от простых, с малым количеством шагов поиска, которые обеспечивают высокую скорость обработки, но имеют низкую вероятность выбора наиболее оптимальной модели, до сложных многошаговых, вероятность выбора наилучшей модели у которых максимальна, однако объем производимых вычислений велик, соответственно скорость обработки низкая.



а – кодер; б – декодер

Рисунок 2.27 – Гибридный кодек изображений с предсказанием и преобразованием

Выбор алгоритма поиска оптимальной модели предсказания, как правило, остается за разработчиком кодера. Декодер при этом обеспечивает однозначное восстановление блока в соответствии с полученной моделью предсказания.

Комбинирование кодирования с предсказанием и кодирования с преобразованием позволяет обеспечить эффективность сжатия на уровне 1 бит/отсчет при практически полной визуальной неотличимости восстановленного изображения от оригинала. При менее жестких требованиях к качеству можно достичь эффективности сжатия в 0,1 бит/отсчет и даже лучше.

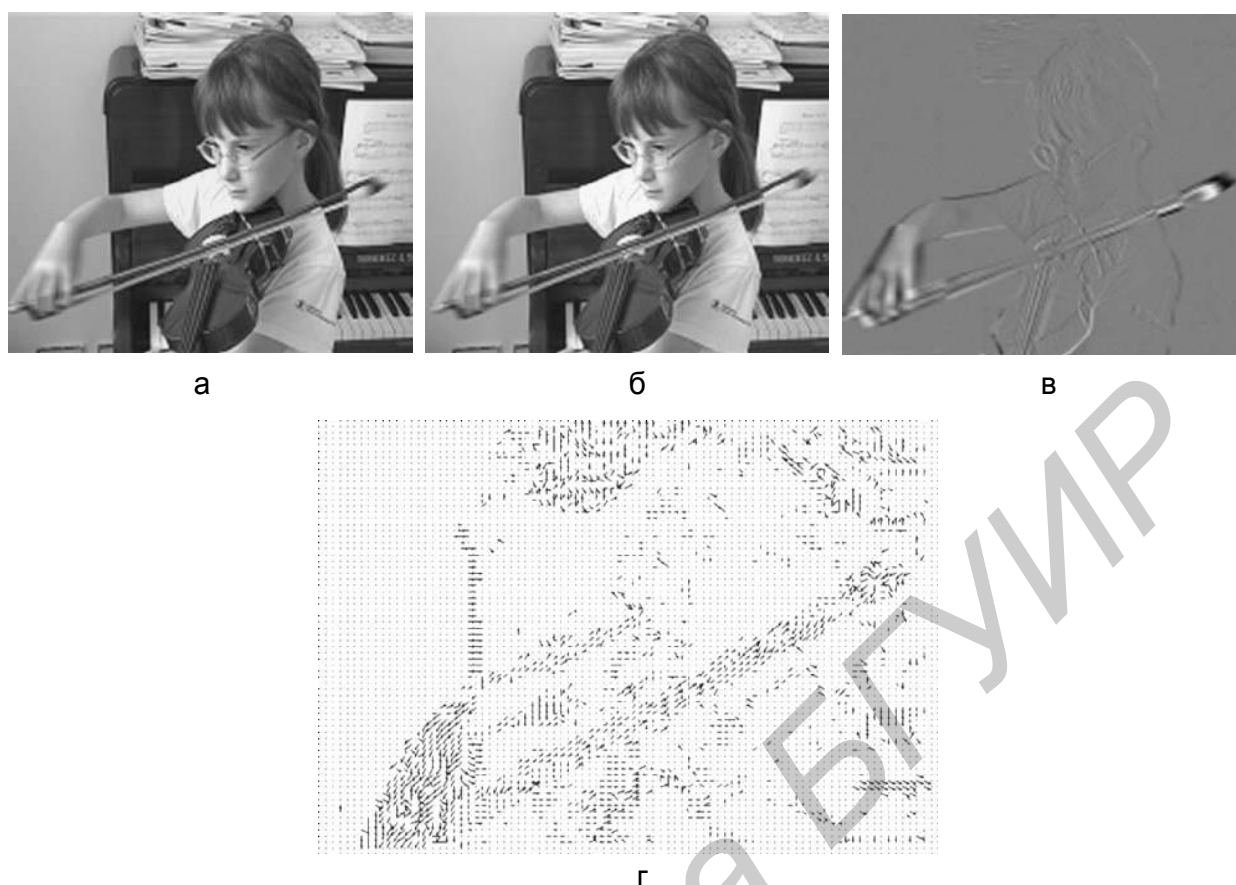
В то же время не меньшими значениями характеризуется и межкадровая (временная) избыточность, обусловленная корреляцией соседних кадров (см. рисунок 2.5). Далее рассмотрим принципы использования кодирования с

предсказанием для ее устранения. Для этого создается кадр-прогноз, который вычитается из текущего кадра. Результатом является остаточный кадр (разность кадров). При этом чем лучше прогноз, тем меньше энергии содержится в остаточном кадре. Остаточный кадр кодируется и посылается декодеру, который воспроизводит кадр-прогноз по некоторым уже построенным кадрам, складывает его с остаточным кадром и, таким образом, восстанавливает текущий кадр. Кадр-прогноз формируется по одному или нескольким прошлым или будущим кадрам («опорным кадрам»). Как правило, точность прогноза может быть улучшена компенсированием движения между ссылочными кадрами и текущим кадром.

2.3.3 Предсказание с компенсацией движения

Простейшие методы временного прогнозирования, рассмотренные ранее, основаны на использовании в качестве прогноза предыдущего кадра. Для примера, на рисунке 2.28 (см. также рисунок 2.5, в) показаны два последовательных кадра видеопоследовательности и остаточный кадр, полученный вычитанием прогноза (кадра $N-1$) из текущего (кадр N). Видно, что при простейшем прогнозе остаточный кадр содержит значительный объем информации (на что указывают яркие и темные области разностного кадра), которую необходимо будет передавать. Это обусловлено движением объектов в кадре, следовательно, лучший прогноз можно построить на основе компенсации движения между двумя этими кадрами.

Отличия между кадрами могут появиться из-за движения объектов – не меняющих свою форму, например автомобилей – и деформируемых, например рук или ног прохожих), перемещения видеокамеры (панорамная съемка, наклоны, наезды, изменения фокуса, повороты и т. п.), изменения перекрытия объектов (например, части сцены, закрытой движущимся предметом), изменения освещенности и в силу многих других причин. За исключением изменений перекрытия и освещенности объектов съемки, эти отличия соответствуют перемещению групп пикселей по кадру. Траектории перемещения каждого пикселя между последовательными кадрами образуют поле траекторий пикселей (рисунок 2.28, г) – оптический поток [47, 48]. По нему можно построить точный прогноз для большинства пикселей текущего кадра, перемещая каждый пиксель опорного кадра вдоль его вектора оптического потока. Однако точное вычисление оптического потока является довольно долгим, кроме того, придется пересылать все векторы оптического потока декодеру для декодирования и восстановления кадра-прогноза, в результате чего повышается объем пересылаемой информации, сводящий на нет преимущества разностного кадра с минимальной информацией.



а – предыдущий кадр; б – текущий; в – разностный; г – поле траекторий элементов изображения между кадрами

Рисунок 2.28 – Оценка разности между соседними кадрами

На практике широко используется метод компенсации движения, который компенсирует перемещение прямоугольных областей или «блоков» текущего кадра. При этом выполняется следующая процедура для каждого фиксированного блока, состоящего из $M \times N$ пикселей обрабатываемого кадра:

1 Поиск на опорном кадре (предыдущем или следующем, ранее закодированном и переданном декодеру) «подходящего» блока из $M \times N$ пикселей. Осуществляется путем сравнения фиксированного $M \times N$ блока с некоторыми или со всеми $M \times N$ блоками области поиска (обычно она представляет собой некий регион с центром в середине этого выбранного блока) и нахождение «самого подходящего» блока. Популярным критерием схожести блоков служит энергия остатка, получаемого вычитанием блока-кандидата из фиксированного $M \times N$ блока, т. е. выбирается блок-кандидат, минимизирующий энергию остатка. Этот процесс поиска подходящего блока называется *оценкой движения*.

2 Выбранный кандидат становится прогнозом текущего $M \times N$ блока и его вычитают из этого блока для получения остаточного $M \times N$ блока – *компенсации движения*.

3 Остаточный блок кодируется и вместе с координатами вектора сме-

щения текущего блока по отношению к позиции блока-кандидата (вектор движения) передается декодеру.

Декодер использует вектор движения для нахождения блока-прогноза, декодирует остаточный блок и складывает его с прогнозом для реконструкции версии исходного блока. Компенсация движения на основе блоков очень популярна по многим причинам. Этот метод достаточно прост и легко поддается программной реализации. Он хорошо вписывается в прямоугольный формат кадров, а также в процедуры преобразования изображений на основе блоков (например, в дискретное косинусное преобразование, см. далее). Компенсация движения определяет достаточно приемлемую временную модель для многих видеопоследовательностей.

Реализацию предсказания с компенсацией движения затрудняет ряд проблем. Например, реальные объекты редко имеют четкие прямоугольные границы. Объекты часто перемещаются на нецелое число пиксельных позиций между кадрами, и многие типы перемещений трудно скомпенсировать, используя методы на основе блоков (например, деформацию объектов, их повороты, изменение масштаба, т. е. приближение или удаление объектов от видеокамеры, сложные движения, такие как клубы дыма). Несмотря на это, компенсация движения на основе блоков является неотъемлемой частью кодирования с предсказанием во всех современных стандартах компрессии изображений.

Базисной единицей для прогноза с компенсацией движения во многих стандартах компрессии изображений, включая MPEG-2, MPEG-4 Visual и MPEG-4 AVC / H.264, является макроблок размером 16×16 элементов. Для формата цифрового представления 4:2:0 структура макроблока показана на рисунке 2.29. Он состоит из 256 отсчетов сигнала яркости, сгруппированных в 4 блока размером 8×8 и 64 отсчетов (блок 8×8) для каждого из цветоразностных сигналов, что дает в сумме 6 блоков размером 8×8 .

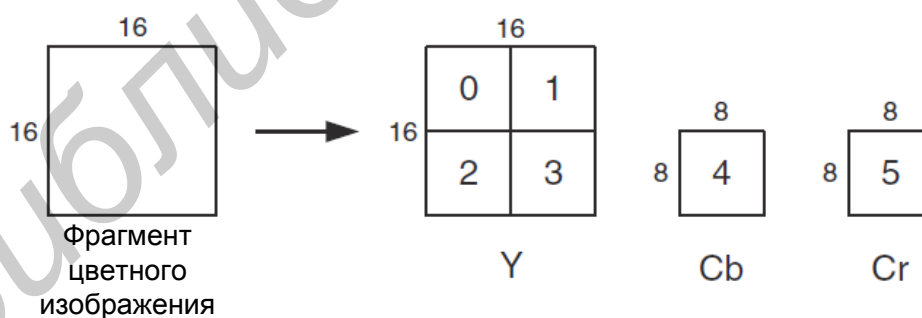


Рисунок 2.29 – Структура макроблока для формата 4:2:0

Оценка движения макроблока состоит из поиска области 16×16 элементов на опорном кадре, наиболее похожей на текущий макроблок. Опорным кадром является уже декодированный кадр, который хронологически может следовать раньше или позже текущего. Область поиска на опорном кадре имеет центр в середине положения текущего макроблока. Максималь-

ные значения смещения макроблока по горизонтали R_x и вертикали R_y определяют размер зоны поиска. Найденный блок 16 x 16 пикселей, имеющий наименьшее расхождение с текущим макроблоком, выбирается в качестве «подходящего» макроблока (рисунок 2.30).

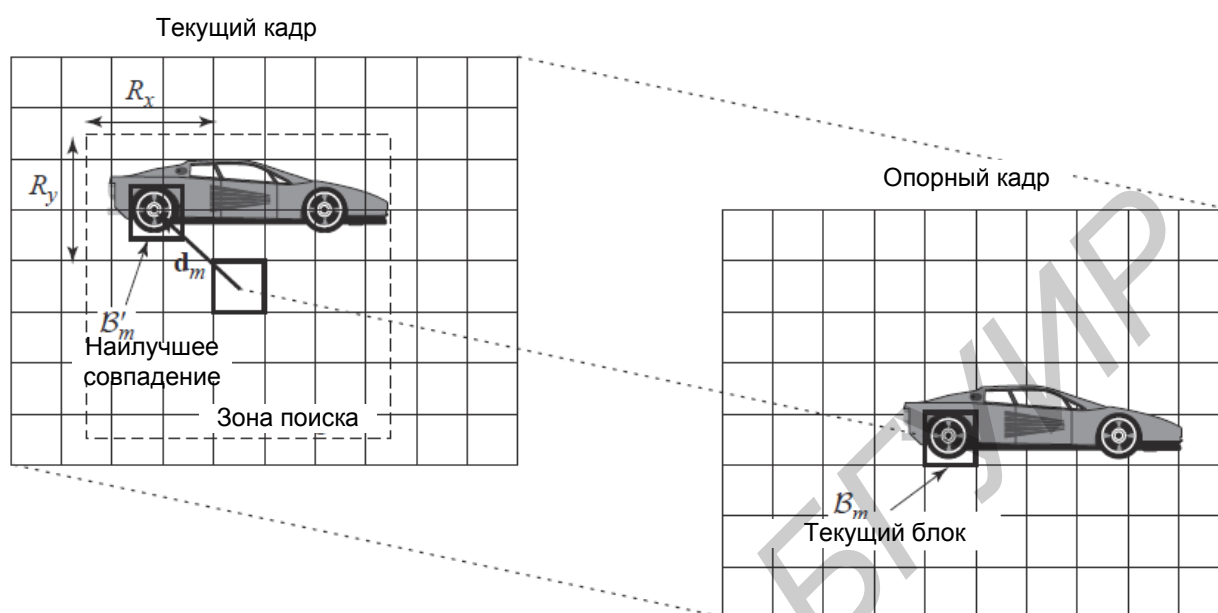


Рисунок 2.30 – Оценка движения

Компенсация движения. Выбранный «наилучший» макроблок на ссылочном кадре вычитается из текущего блока, и их разность, называемая остаточным макроблоком (яркости и цветности), кодируется и передается вместе с вектором движения (рисунок 2.31), определяющим позицию наилучшего макроблока по отношению к текущему.



а – прогнозируемый кадр; б – предыдущий (опорный) кадр с нанесенными векторами движения

Рисунок 2.31 – Формирование векторов движения

Существует много вариаций базового процесса оценки движения и компенсации движения. Опорным кадром может быть предыдущий кадр (в хронологическом порядке), будущий кадр или комбинация двух или более ранее декодированных кадров. Если в качестве опорного выбран будущий кадр, то необходимо его закодировать до текущего кадра (т. е. кадры должны быть закодированы не в хронологическом порядке). Там, где имеется существенное расхождение между опорным и текущим кадром (например при смене сцен), можно кодировать макроблок без компенсации движения.

Как уже отмечалось, движущиеся объекты редко следуют строго вдоль границ блоков пикселей 16×16 , поэтому более эффективным может быть использование переменных размеров блоков для оценки и компенсации движения. Объекты могут также перемещаться на нецелое число пикселей между кадрами. Поэтому лучший прогноз можно построить интерполяцией опорного кадра по позициям подпикселей перед поиском наилучшего положения прогноза.

Размер блока также влияет на точность прогноза. Очевидно, что чем меньше блок (в пределе – 1×1 элемент), тем лучше результат компенсации движения (рисунок 2.32). Однако уменьшение размера блока влечет за собой увеличение сложности вычислений (необходимо совершить больше операций поиска) и количества векторов движения, необходимых для декодирования.

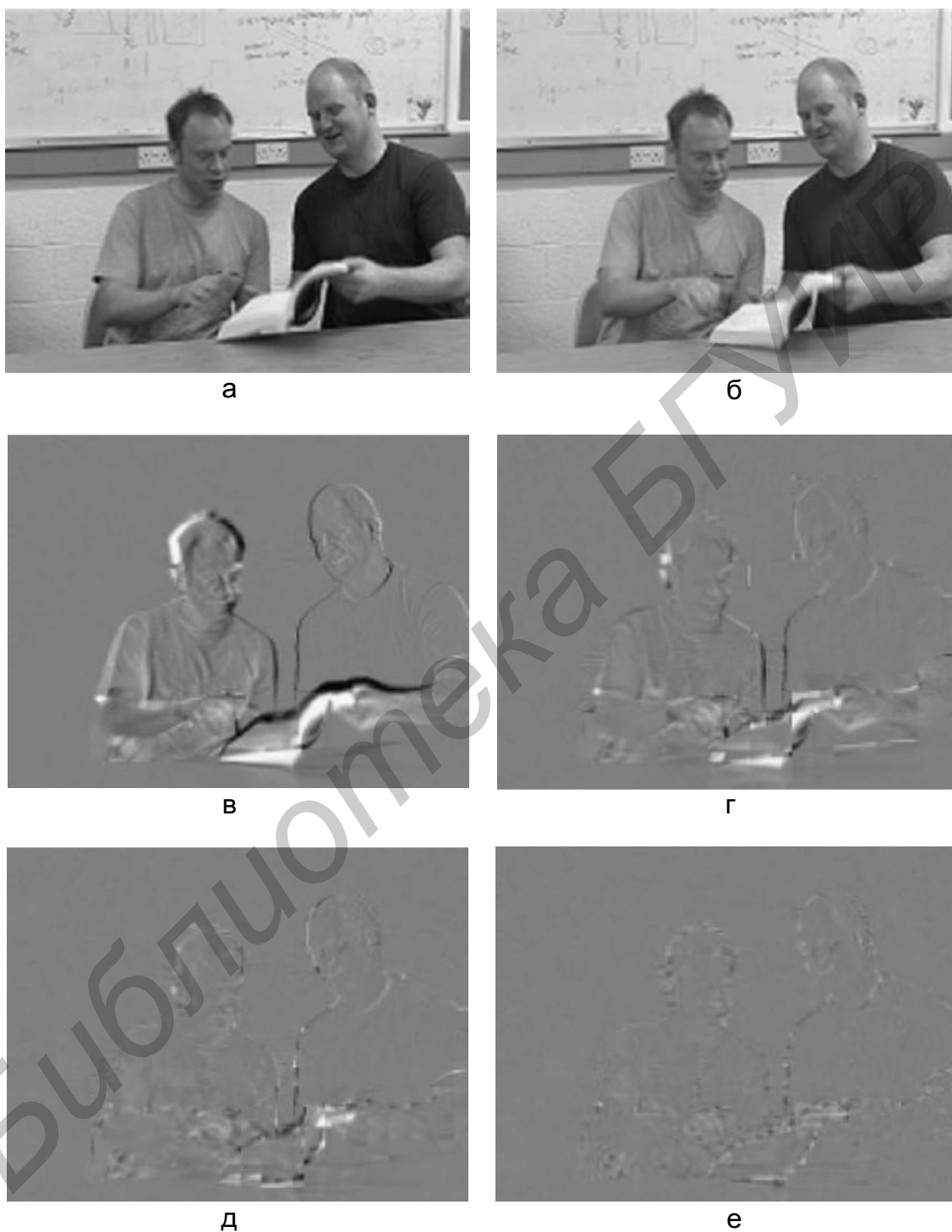
Передача каждого нового вектора движения требует дополнительных бит, поэтому увеличение числа векторов движения сведет на нет выигрыш от сокращения энергии остаточного кадра. Эффективный компромисс заключается в приспособлении размеров блоков к характеристикам изображений, т. е. можно выбирать большие размеры блоков для однородных областей кадра, а для областей с мелкими деталями и со сложным движением следует уменьшать размер блока (рисунок 2.33).

В заключение подраздела приведем общие выводы по использованию кодирования с предсказанием для компрессии изображений:

- кодирование с предсказанием, вне зависимости от его типа, не избавляет от необходимости передавать информацию о каждом элементе изображения, точнее, о его разности с соседним, и поэтому не позволяет достичь эффективности использования битового потока для передачи данных изображения, лучшей, чем 1 бит/отсчет (для монохроматических изображений). На практике удовлетворительное качество изображений, кодированных только на основе предсказания, сохраняется при разрядности кодированного потока данных, не меньшей 2...4 бит/элемент. Соответственно, непосредственное использование ДИКМ для передачи отсчетов изображения малоэффективно;

- кодирование с предсказанием уменьшает энтропию сигналов (изображений), поэтому каскадное его использование со статистическими методами или с кодированием с преобразованием позволяет достичь большей степени сжатия либо реализовать сжатие без потерь;

– компенсация движения способствует еще большему снижению количества информации в остаточных (разностных) изображениях, соответственно, ее использование повышает эффективность использования систем с предсказанием для компрессии динамических изображений.



а – опорный кадр; б – прогнозируемый кадр; в – разностный кадр без компенсации движения; г – разностный кадр с компенсацией движения и размером блока 16×16 ; д – разностный кадр с размером блока 8×8 ; е – разностный кадр с размером блока 4×4

Рисунок 2.32 – Влияние выбора размера блока на точность предсказания



Рисунок 2.33 – Пример разбиения изображения на блоки разных размеров

Таким образом, при использовании кодирования с предсказанием в составе компрессора изображений, результатом его являются:

- вычисленное разностное (иногда также называемое остаточным) изображение;
- векторы движения при межкадровом предсказании с компенсацией движения;
- режимы предсказания блоков или макроблоков – при внутрикадровом предсказании.

Критерием оценки эффективности кодирования с предсказанием в данном случае является минимизация суммарного объема передаваемых данных.

2.4 Кодирование с преобразованием

2.4.1 Общие принципы кодирования с преобразованием

Методы кодирования изображений с преобразованием основаны на том, что изображение, в общем случае представляющее собой матрицу отсчетов $f(x, y)$, путем соответствующих преобразований приводится к виду, наиболее удобному с точки зрения сокращения избыточной информации или снижения чувствительности к ошибкам кодирования и передачи, а после прохождения тракта передачи (консервации) путем обратных преобразований возвращается в исходную форму (рисунок 2.34).

В задачах компрессии изображений выполняется преобразование видеоинформации из пространственно-временной области в спектральную. Спектральная область при кодировании с преобразованием понимается в смысле обобщенных спектров. Преобразование ставит в соответствие сигналу его обобщенный спектр в выбранном базисе. Например, преобразование Фурье ставит в соответствие сигналу традиционный частотный спектр.

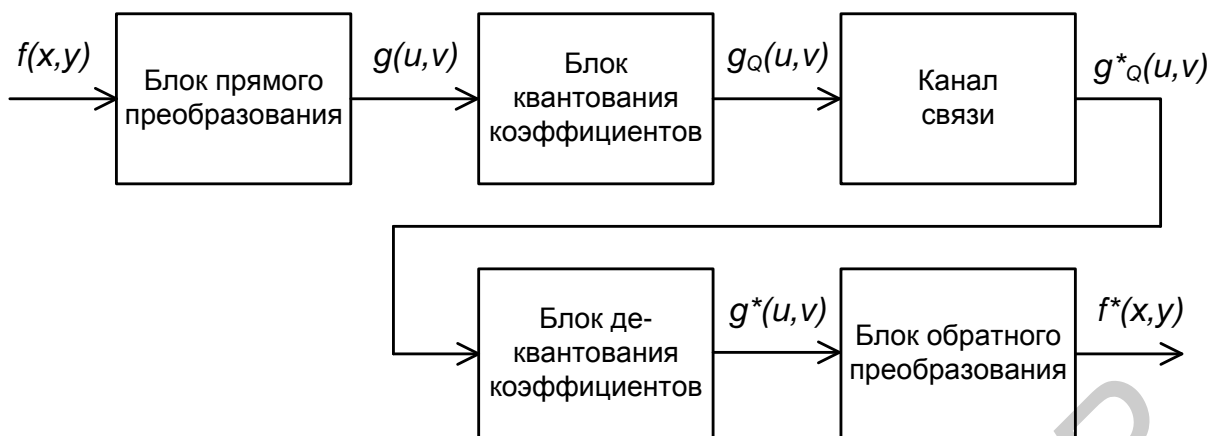


Рисунок 2.34 – Обобщенная структурная схема системы передачи изображений с преобразованием

Для эффективного кодирования необходимо, чтобы параметры преобразования согласовывались как с параметрами источника сигналов (изображением), так и с параметрами получателя (зрительной системой человека). Эффективность кодирования зависит от того, как в процессе преобразования происходит перегруппировка энергии составляющих сигнала. Необходимо стремиться к тому, чтобы в результате преобразования в спектральную область энергия составляющих была сосредоточена в минимальном количестве спектральных составляющих (трансформант). При линейных ортогональных преобразованиях указанное условие частично выполняется в случае, когда составляющие векторов матрицы преобразования соответствуют типичным временным характеристикам видеосигнала в горизонтальном и вертикальном направлении. Следует также отметить, что не существует преобразования, одинаково оптимального для всех типов изображений, поэтому для каждой задачи оно должно выбираться индивидуально, с учетом усредненных параметров обрабатываемых изображений.

Из различных линейных преобразований, позволяющих эффективно сокращать избыточность изображений, в современных алгоритмах компрессии, как уже отмечалось, наиболее часто используется дискретно-косинусное преобразование (ДКП). Достоинством его является хорошая согласованность с параметрами изображения, а также с характеристиками зрительного анализатора (пространственной разрешающей способностью). Последнее дает возможность проводить квантование коэффициентов (трансформант) ДКП с учетом их вклада в зрительное восприятие изображения – низкочастотные составляющие должны передаваться точно, высокочастотные, к которым чувствительность зрения минимальна – со сниженной точностью либо не передаваться вовсе, в зависимости от требований к качеству преобразуемого изображения. В прикладных задачах обработки изображений могут использоваться и другие преобразования – ДПФ, ПКЛ, ДПА и др.

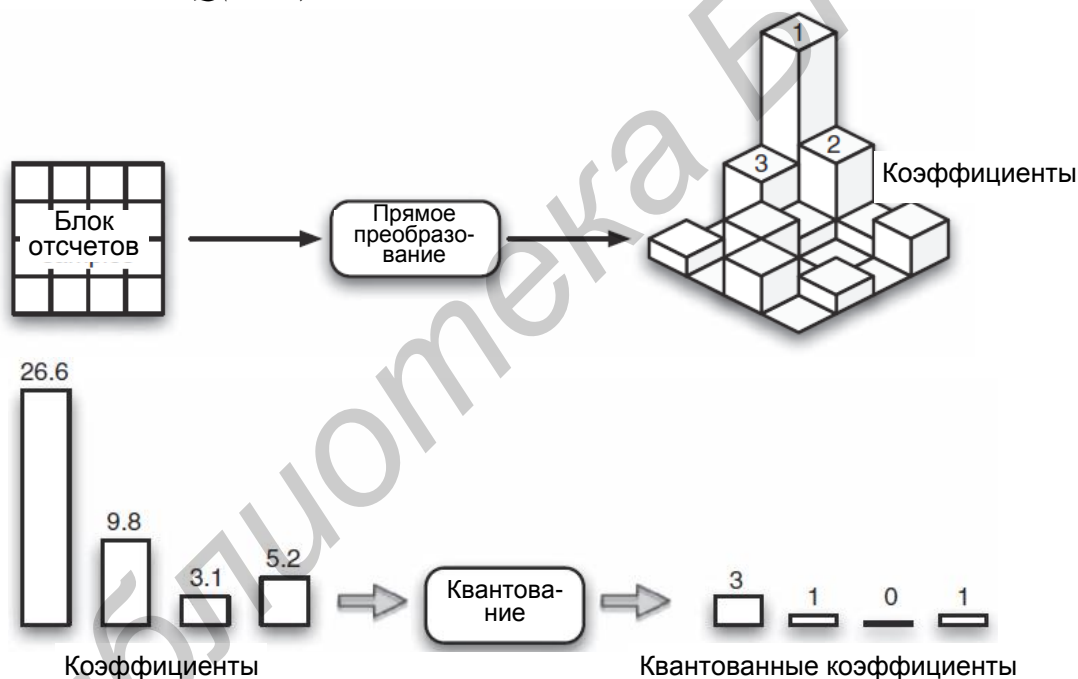
Преобразование выполняется не над всем изображением одновременно – первоочередной причиной тому является большой динамический диапа-

зон получающихся спектральных коэффициентов, а также вычислительная сложность. Поэтому обработка ведется поблочно (рисунок 2.35), причем каждый блок является квадратным и имеет достаточно малые размеры – в большинстве случаев 4×4 либо 8×8 элементов (отсчетов).

Следует отметить, что с математической точки зрения отдельно взятое ДКП (как и любое другое ортогональное преобразование) не дает сокращения избыточности! Сокращение достигается за счет последующего квантования (с уменьшением разрядности) передаваемых спектральных составляющих – трансформант – и уменьшения их количества за счет тех из них, которыми можно пренебречь ввиду малой их амплитуды и, соответственно, незначительности для зрительного восприятия.

Квантование осуществляется путем почленного деления полученной матрицы коэффициентов преобразования $g(u, v)$ на матрицу квантования $Q(u, v)$ с округлением до ближайшего целого:

$$g_Q(u, v) = \frac{g(u, v)}{Q(u, v)} \quad (2.19)$$



73	87	64	13	181.3	47.4	-65.8	4.0
40	63	23	2	5.7	29.6	16.4	-2.2
36	24	68	26	40.3	18.3	-28.8	-13.8
29	98	67	12	13.1	-20.2	13.8	33.9
Блок отсчетов 4×4				Коэффициенты преобразования			
18	5	-7	0	9	2	-3	0
1	3	2	0	0	1	1	0
4	2	-3	-1	2	1	-1	-1
1	-2	1	3	1	-1	1	2
Квантованные коэффициенты (шаг квантования = 10)				Квантованные коэффициенты (шаг квантования = 20)			

Рисунок 2.35 – Иллюстрация процесса кодирования с преобразованием

Результатом этого действия является матрица квантованных коэффициентов $g_Q(u, v)$, новый шаг квантования каждого из которых равен соответствующему значению из матрицы квантования $Q(u, v)$. Для упрощения понимания в примерах, приведенных на рисунках 2.35 и 2.36, использованы матрицы квантования с одинаковыми значениями $Q_1(u, v)$ и $Q_2(u, v)$, равными соответственно 10 и 20:

$$Q_1 = \begin{bmatrix} 10 & 10 & 10 & 10 \\ 10 & 10 & 10 & 10 \\ 10 & 10 & 10 & 10 \\ 10 & 10 & 10 & 10 \end{bmatrix}; \quad Q_2 = \begin{bmatrix} 20 & 20 & 20 & 20 \\ 20 & 20 & 20 & 20 \\ 20 & 20 & 20 & 20 \\ 20 & 20 & 20 & 20 \end{bmatrix}.$$

При этом следует помнить, что восстановленное изображение побитно уже не будет совпадать с исходным – см. таблицы на рисунках 2.33 и 2.34.

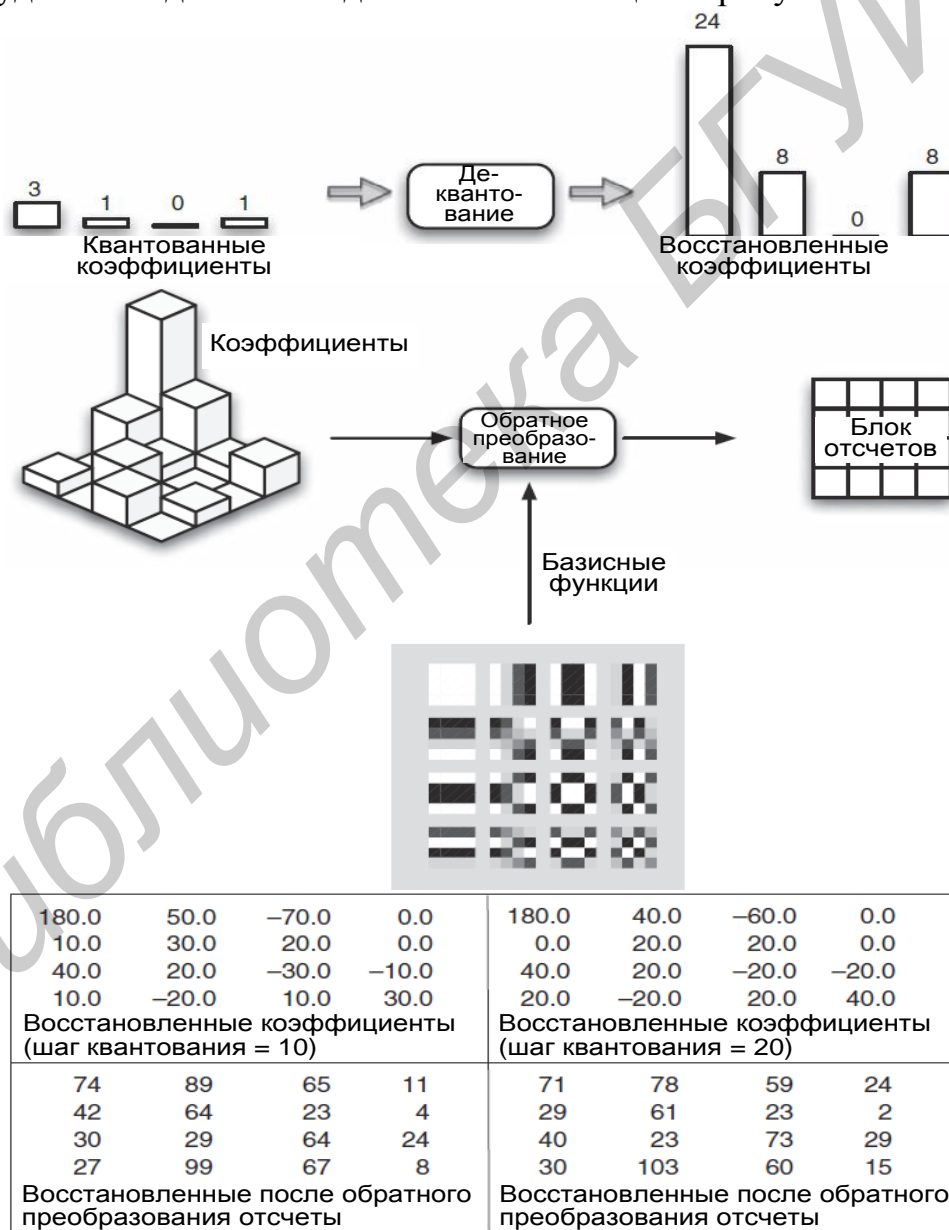


Рисунок 2.36 – Восстановление отсчетов изображения после кодирования с преобразованием

Полное совпадение возможно только в случае передачи всех трансформант и отсутствия ограничений на точность вычислений. В то же время субъективно (зрительно) восстановленное и исходное изображения могут быть практически неотличимы.

Квантование коэффициентов имеет чрезвычайно важное значение, ибо позволяет учесть особенности зрительного восприятия крупных и мелких деталей изображения и, соответственно, дает путь к сокращению его визуальной избыточности. Дело в том, что коэффициенты преобразования однозначно определяют удельный вес в блоке изображения составляющих с различными пространственными частотами. Коэффициент с номером $(0,0) - g(0,0)$ всегда определяет постоянную составляющую (либо среднюю яркость, если преобразованию подвергается блок отсчетов сигнала яркости либо сигналов основных цветов). Коэффициент $g(0,1)$ определяет первую гармонику горизонтальной пространственной частоты, $g(1,0)$ – первую гармонику вертикальной пространственной частоты и т. д.

Помня, что постоянная составляющая и гармоники с малыми номерами характеризуют медленные изменения яркости (либо цветности, если преобразуются отсчеты цветоразностных сигналов), к которым глаз более чувствителен, а гармоники с большими номерами несут информацию о резких перепадах, на которых глаз плохо различает градации яркости, можно сделать вывод, что преобразование блока отсчетов в блок коэффициентов позволяет обеспечить оптимальную точность представления низко- и высокочастотных составляющих сигнала. Таким образом, именно кодирование с преобразованием вносит основной вклад в сокращение визуальной избыточности изображений.

В качестве примера на рисунке 2.37 приведена матрица квантования яркостного сигнала, используемая в стандарте MPEG-2. Она обеспечивает квантование постоянной составляющей с шагом 8, первых гармоник – с шагом 16, для гармоник с большими номерами шаг постепенно повышается.

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

Рисунок 2.37 – Матрица квантования коэффициентов преобразования MPEG

Заключительным этапом кодирования является преобразование матрицы квантованных коэффициентов в вектор, осуществляемое путем зигзагообразного считывания. Последовательность считывания также имеет вполне определенное значение. Дело в том, что при квантовании достаточно большая часть коэффициентов с большими номерами становится равной нулю, и считывание должно быть организовано таким образом, чтобы с наибольшей вероятностью обеспечивалось бы получение максимально длинных серий нулей, которые хорошо поддаются уплотнению статистическими методами кодирования, например, кодированием длин серий. Поэтому процесс считывания всегда начинается с коэффициента с номером (0,0) и зигзагом продолжается в сторону увеличения номеров по горизонтали и вертикали (рисунок 2.38).

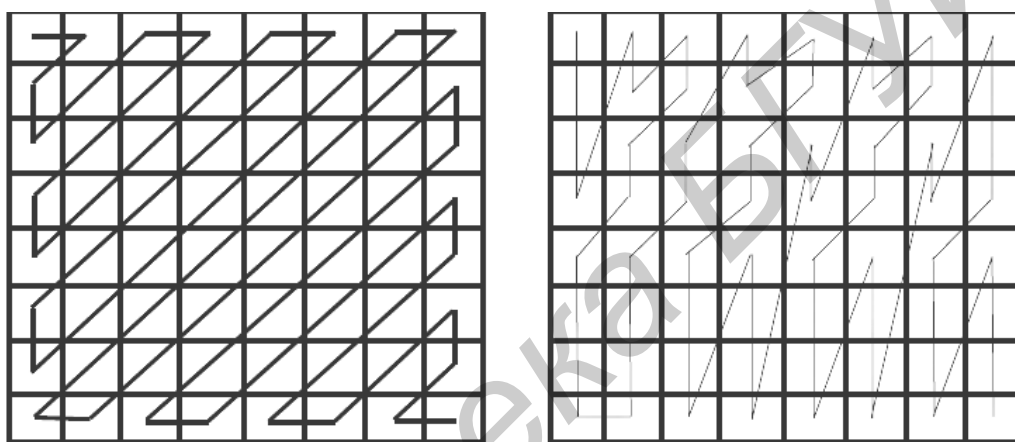


Рисунок 2.38 – Порядок считывания квантованных коэффициентов при построчной (слева) и чересстрочной (справа) развертке

2.4.2 Математические основы ортогональных преобразований

Дискретизация исходного изображения позволяет описать подвергающееся преобразованию изображение либо его фрагмент (блок) с помощью матрицы отсчетов сигнала. В дальнейшем будем рассматривать матрицу-строку $(1 \times X)$ или квадратную матрицу $(X \times Y, X=Y)$ отсчетов сигнала $f(x)$ или $f(x, y)$ соответственно.

Пусть φ_v – полная система одномерных дискретных функций

$$\varphi_v = [\varphi_v(0), \varphi_v(1), \dots, \varphi_v(X-1)], v \in \{0, \dots, X-1\}, \quad (2.20)$$

что можно представить в виде матрицы

$$[\varphi] = \begin{bmatrix} \varphi_0(0) & \varphi_0(1) & \dots & \varphi_0(N-1) \\ \varphi_1(0) & \varphi_1(1) & \dots & \varphi_1(N-1) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{N-1}(0) & \varphi_{N-1}(1) & \dots & \varphi_{N-1}(N-1) \end{bmatrix}. \quad (2.21)$$

Тогда для одномерного прямого преобразования отдельных сигнальных векторов выполняется соотношение

$$g(v) = \sum_{x=0}^{N-1} \varphi_v(x) f(x), \quad (2.22)$$

или в матричной форме записи

$$[g] = [\varphi][f], \quad (2.23)$$

где $[g]$, $[f]$ – матрицы-столбцы, соответствующие векторам f, g с N элементами; $[\varphi]$ – квадратная матрица ($N \times N$).

В результате решения получим N трансформант (спектральных коэффициентов), определяющих амплитуды набора ортогональных функций. Трансформант $g(v)$ всегда определяется соответствующей v -ой ортогональной функцией $\varphi_v(x)$ и величинами всех отсчетов сигнала в исходной временной области $f(x)$. С помощью трансформант преобразования исходный сигнал можно представить в виде

$$\begin{aligned} f(x) &= g(0)\varphi_0(x) + g(1)\varphi_1(x) + \dots + g(N-1)\varphi_{N-1}(x) = \\ &= \sum_{v=0}^{N-1} g(v)\varphi_v(x), \end{aligned} \quad (2.24)$$

что в матричной форме записывается как обратное преобразование

$$[f] = [\varphi]^T [g], \quad (2.25)$$

где $[\varphi]^T$ – матрица, транспонированная по отношению к $[\varphi]$.

Так как при одномерном обратном преобразовании в общем случае также выполняется соотношение

$$[f] = [\varphi]^{-1} [g] = [\varphi]^{-1} [\varphi][f], \quad (2.26)$$

при ортонормальном преобразовании обратная матрица определяется транспонированной матрицей

$$[\varphi]^{-1} = [\varphi]^T, \quad (2.27)$$

и одновременно выполняется условие

$$[\varphi]^{-1}[\varphi]^T = \mathbf{E}, \quad (2.28)$$

где \mathbf{E} – единичная матрица.

Если транспонированная матрица удовлетворяет, кроме того, условию симметрии по отношению к главной диагонали, то дополнительно выполняется условие

$$[\varphi]^{-1} = [\varphi]^T = [\varphi], \quad (2.29)$$

в соответствии с которым ядро преобразования одинаково для прямого и обратного преобразования.

Ортонормальному одномерному преобразованию присущи следующие свойства:

- сумма произведений элементов любой строки (столбца) на соответствующие элементы другой строки (столбца) транспонированной матрицы равна нулю;
- сумма квадратов элементов любой строки (столбца) соответствует энергии основного (инвертированного) сигнала и равна единице;
- обратная матрица преобразования получается в результате транспонирования исходной матрицы.

Аналогично и двумерное прямое преобразование будет определяться формулой

$$g(u, v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \varphi_{u,v}(x, y) f(x, y), \quad (2.30)$$

в результате вычисления которого получим N^2 трансформант (спектральных коэффициентов), определяющих амплитуды набора двумерных ортогональных функций. В общем случае трансформанта определяется (u, v) -й ортогональной функцией $\varphi_{u,v}(x, y)$ и величинами отсчетов $f(x, y)$ на плоскости. Исходная двумерная функция $f(x, y)$ с использованием трансформанта может быть представлена система двумерных ортогональных функций

$$f(x, y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} g(u, v) \varphi_{u,v}(x, y). \quad (2.31)$$

В соответствии с этой формулой осуществляется обратное преобразование.

Запись матрицы трансформанта и следовательно, формулы преобразования можно упростить, если двумерную систему ортогональных функций представить как произведение одномерных ортогональных функций, например вектора-столбца φ_u^T и вектора-строки φ_v :

$$\varphi_{u,v} = \varphi_u^T \cdot \varphi_v. \quad (2.32)$$

Формула одновременно отражает способ определения базисных изображений двумерного преобразования с разделяемым ядром. Для элементов матрицы трансформант $g(u, v)$ в матричной форме выполняется соотношение

$$g(u, v) = [\varphi_v][f][\varphi_u]^T,$$

или в общем случае для матрицы трансформант $g(u, v)$

$$[g] = [\varphi][f][\varphi]^T. \quad (2.33)$$

В соответствии с приведенным выражением двумерное прямое преобразование можно представить как две последовательно выполняемые одномерные операции. Первая операция заключается в умножении сигнальной матрицы $[f]$ слева на вектор-строку матрицы $[\varphi_v]$ и соответствует одномерному преобразованию столбцов матрицы $[f]$. Вторая операция заключается в последующем умножении справа результата первой операции на вектор-столбец матрицы $[\varphi_u]^T$, что соответствует одномерному преобразованию строк матрицы $[\varphi_v][f]$. Разделение ядра преобразования аналогичны образом упрощает и формулу обратного преобразования:

$$[f] = [\varphi]^T [f][\varphi]. \quad (2.34)$$

2.4.3 Дискретное преобразование Фурье

Дискретное преобразование Фурье (ДПФ) основано на разложении в ряд Фурье дискретного сигнала $f_D(t)$, представленного на интервале $T = N \cdot \Delta T$ последовательностью из N отсчетов $f(n\Delta T)$, $n \in \{0, 1, 2, \dots, N-1\}$, и ограниченного по частоте так, что в нем не содержатся гармоники частоты $\Omega_0 = 2\pi/T$ с номерами по модулю, превышающими $N/2$:

$$f_D(t) = \sum_{n=0}^{N-1} f(n\Delta T) \cdot \delta(t - n\Delta T) = \sum_{k=-N/2}^{N/2-1} c_k \exp(jk\Omega_0 t) \quad (2.35)$$

с соответствующими коэффициентами разложения

$$c_k = \frac{1}{T} \sum_{n=0}^{N-1} f(n\Delta T) \exp(-jk\Omega_0 n\Delta T) = \frac{1}{N\Delta T} \sum_{n=0}^{N-1} f(n\Delta T) \exp(-j \frac{2\pi kn}{N}). \quad (2.36)$$

В выражениях (2.35) и (2.36) реальный масштаб частоты и времени учитывается через множитель ΔT . При рассмотрении дискретных последовательностей обычно оперируют номерами отсчетов и спектральных гармоник без привязки к действительному масштабу времени и частоты, принимая частоту дискретизации равной 1 (т. е. нормируя по параметру $\Delta T = 1$).

На основе этого одномерное ДПФ определяется следующим образом. Если $\{ f(n) \}$ – последовательность $f(n)$ отсчетов сигнала (конечных действительных значений) для всех $n \in \{0, 1, 2, \dots, N-1\}$, то прямое ДПФ определяется формулой

$$c_k = \frac{1}{N} \sum_{n=0}^{N-1} f(n) W^{kn}, \quad (2.37)$$

где $W^{kn} = \exp\left(-j \frac{2\pi}{N} kn\right)$ – комплексное ядро преобразования, удовлетворяющее условию ортогональности

$$\sum_{n=0}^{N-1} W^{kn} \cdot W^{-ln} = \begin{cases} 0 & \text{при } k \neq l \\ N & \text{при } k = l \end{cases}. \quad (2.38)$$

Используя это условие и (2.22), запишем

$$\sum_{k=0}^{N-1} c_k W^{-kn} = \frac{1}{N} \sum_{k=0}^{N-1} [f(0) + f(1)W^k + \dots + f(n)W^{kn} + \dots + f(N-1)W^{k(N-1)}] W^{-kn},$$

откуда можно выразить $f(n)$ и определить обратное ДПФ

$$f(n) = \sum_{k=0}^{N-1} c_k W^{-kn}, \quad n \in \{0, 1, 2, \dots, N-1\}. \quad (2.39)$$

Следует отметить, что в размещении множителя $1/N$ в формулах (2.37) и (2.39) нет полного единства. В ряде литературных источников и математических пакетах компьютерных программ этот множитель внесен в выражение для обратного ДПФ.

Ядро преобразования W^{kn} является периодическим с периодом N , т. е. $W^{kn} = W^{k(n+N)} = W^{n(k+N)}$ для $k, n \in \{0, 1, 2, \dots, N-1\}$. Поэтому и последовательности $f(n)$, c_k также являются периодическими и выполняются соотношения $f(\pm n) = f(sN \pm n)$, $c_{\pm k} = c_{sN \pm k}$ для $s \in \{0, \pm 1, \pm 2, \dots\}$.

При использовании ДПФ для кодирования изображений с преобразованием важно распространение ДПФ на функции двух переменных. Для этого перейдем от одномерной временной последовательности $f(n)$ отсчетов сигнала $\{ f(n) \}$ к двумерному массиву $f(x, y)$ конечных значений (отсчетов) сигнала изображения, удовлетворяющих условию $x \in \{0, 1, 2, \dots, X-1\}$, $y \in \{0, 1, 2, \dots, Y-1\}$. В этом случае, по аналогии с (2.37), двумерное прямое ДПФ записывается формулой

$$g(u, v) = \frac{1}{X \cdot Y} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} f(x, y) W_1^{ux} W_2^{vy}, \quad (2.40)$$

где комплексное ядро преобразования $W(u, v, x, y)^{vx} = W_1^{vx} W_2^{vy}$,

$$W_1 = \exp\left(-j \frac{2\pi}{X}\right), u \in \{0, 1, 2, \dots, X-1\},$$

$$W_2 = \exp\left(-j \frac{2\pi}{Y}\right), v \in \{0, 1, 2, \dots, Y-1\},$$

разделимо и удовлетворяет условию ортогональности.

Аналогичным образом, если всю константу нормирования ввести в соотношение для прямого преобразования, то двумерное обратное ДПФ можно записать в виде

$$f(x, y) = \sum_{u=0}^{U-1} \sum_{v=0}^{V-1} g(u, v) W^{-u \cdot x} W^{-v \cdot y}, \quad (2.41)$$

для $x \in \{0, 1, 2, \dots, X-1\}$, $y \in \{0, 1, 2, \dots, Y-1\}$. Так как ядро двумерного ДПФ разделимо, то прямое преобразование можно записать в матричной форме

$$[g(u, v)] = [A] [f(x, y)] [A]_R, \quad (2.42)$$

где $[g(u, v)]$ – матрица спектральных коэффициентов (трансформант) ДПФ;

$[f(x, y)]$ – матрица значений (отсчетов) сигнала;

$[A]$ – квадратная матрица преобразования размером $(X \times X)$ для $X = 2^x$, $x = 1, 2, \dots$:

$$[A] = \frac{1}{\sqrt{X}} \begin{bmatrix} W_1^0 & W_1^0 & \dots & W_1^0 \\ W_1^0 & W_1^1 & \dots & W_1^{(X-1)} \\ W_1^0 & W_1^2 & \dots & W_1^{2(X-1)} \\ \vdots & \vdots & \dots & \vdots \\ W_1^0 & W_1^{X-1} & \dots & W_1^{(X-1)(X-1)} \end{bmatrix}, \quad (2.43)$$

$[A]_R$ – квадратная матрица преобразования размером $(Y \times Y)$ для $Y = 2^y$, $y = 1, 2, \dots$:

$$[A]_R = \frac{1}{\sqrt{Y}} \begin{bmatrix} W_2^0 & W_2^0 & \dots & W_2^0 \\ W_2^0 & W_2^1 & \dots & W_2^{(Y-1)} \\ W_2^0 & W_2^2 & \dots & W_2^{2(Y-1)} \\ \vdots & \vdots & \dots & \vdots \\ W_2^0 & W_2^{Y-1} & \dots & W_2^{(Y-1)(Y-1)} \end{bmatrix}, \quad (2.44)$$

Интерпретируя выражение (2.42), получаем спектральные коэффициенты (трансформанты) двумерного ДПФ как результат последовательного выполнения двух одномерных преобразований двумерного сигнала по строкам и столбцам. Первое одномерное преобразование (умножение на матрицу $[A]$ слева) соответствует ДПФ столбцов сигнальной матрицы $[f(x, y)]$. Второе одномерное преобразование (умножение матрицы спектральных коэффициентов $[A] [f(x, y)]$ справа на матрицу $[A]_R$) соответствует преобразованию по строкам, в процессе которого определяются спектральные коэффициенты $[g(u, v)]$ ДПФ.

Так как матрица ядра ДПФ симметрична по отношению к ее главной диагонали, матрицы $[A]$, $[A]_R$ являются унитарными и по определению

$$\begin{aligned} [A]^{-1} &= [A^*]^T = [A], \\ [A]_R^{-1} &= [A^*]_R^T = [A]_R, \end{aligned}$$

где $[A^*]$, $[A^*]_R$ комплексно сопряженные, а $[A]^T$, $[A]_R^T$ – транспонированные матрицы $[A]$, $[A]_R$. На основе этого обратное двумерное ДПФ в матричной форме записи имеет вид

$$[f(x, y)] = [A] [g(u, v)] [A]_R. \quad (2.45)$$

2.4.4 Дискретно-косинусное преобразование

Недостаток широко используемого в задачах цифровой обработки сигналов дискретного преобразования Фурье (ДПФ) с точки зрения кодирования с преобразованием заключается в том, что трансформанты, общее количество которых составляет $2N$, а период равен $T = N \Delta T$, носят комплексный характер. Учитывая состав трансформанта, целесообразнее осуществлять такое преобразование, в котором используется только реальная или мнимая часть ядра преобразования, т. е. только члены \cos или \sin . Однако в этом случае необходимо обеспечить полноту ортогональной системы преобразования.

Если перейти от общего ДПФ к более специфическому ДКП, основанному на базисе ортогональных косинусных составляющих, то выбором соответствующего интервала T необходимо добиться того, чтобы множество ортогональных функций образовало полную систему. В качестве ортогональных функций нельзя выбирать обычные косинусные функции $\cos(k\Omega_0 t)$, так как это связано с использованием функций $\sin(k\Omega_0 t)$. Удовлетворить условие ортогональности проще всего выбором ортогональных составляющих вида $\cos(k\frac{\pi}{T})$, что в более полной форме можно записать как

$$\Phi_k = \cos \frac{2n+1}{2N} k\pi, \quad k, n \in \{0, \dots, N-1\}. \quad (2.46)$$

Так как для ортогональной системы выполняется соотношение

$$\sum_{n=0}^{N-1} \cos \left(\frac{2n+1}{2N} k\pi \right) \cdot \sin \left(\frac{2n+1}{2N} l\pi \right) = \begin{cases} 0 & \text{для } k = l, \\ K_{k,l} & \text{для } k \neq l, \end{cases} \quad (2.47)$$

то ортогональная система в соответствии с выражением (2.46) образует полную систему и ее можно использовать в качестве основы ДКП.

Соответствующие постоянные нормирования для образования ортогональной системы находятся из условия ортогональности на интервале N :

$$\sum_{n=0}^{N-1} \cos\left(\frac{2n+1}{2N}k\pi\right) \cdot \cos\left(\frac{2n+1}{2N}l\pi\right) = \begin{cases} N & \text{для } k=l=0, \\ N/2 & \text{для } k=l \neq 0, \\ 0 & \text{для } k \neq l. \end{cases} \quad (2.48)$$

Если ввести константу нормирования в выражения как для прямого, так и для обратного преобразований, то при одномерной полной системе ортогональных функций ДКП получим

$$\varphi_k = \left[\frac{1}{\sqrt{N}}, \sqrt{\frac{2}{N}} \cos \frac{2n+1}{2N} k\pi \right], \quad \begin{cases} k \in \{1, 2, \dots, N-1\}, \\ n \in \{0, 1, \dots, N-1\}. \end{cases} \quad (2.49)$$

Собственно ядро преобразования при переходе от ДПФ к ДКП определяется как

$$\cos \frac{2n+1}{2N} k\pi = \operatorname{Re} \left\{ \exp \left[-j \frac{\pi}{2N} k \right] \cdot \exp \left[-j \frac{\pi}{2N} kn \right] \right\}. \quad (2.50)$$

С учетом этого, после преобразований получаем выражение для прямого преобразования:

$$C(0) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} f(n), \quad (2.51)$$

$$C(k) = \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} f(n) \cos \frac{2n+1}{2N} k\pi, \quad k \in \{1, 2, \dots, N-1\}.$$

Упрощенная стандартная формула обратного одномерного ДКП имеет вид

$$f(n) = \frac{1}{\sqrt{N}} C(0) + \sqrt{\frac{2}{N}} \sum_{k=1}^{N-1} C(k) \cos \frac{2n+1}{2N} k\pi. \quad (2.52)$$

Приведенные выше формулы для одномерного ДКП можно путем формального преобразования привести к замкнутому виду. Тогда прямое преобразование будет определяться выражением

$$C(k) = \frac{2}{N} c(k) \sum_{n=0}^{N-1} f(n) \cos \frac{2n+1}{2N} k\pi, \quad k \in \{0, 1, \dots, N-1\}, \quad (2.53)$$

а обратное преобразование –

$$f(n) = \sum_{k=0}^{N-1} c(k)C(k)\cos\frac{2n+1}{2N}k\pi, \quad n \in \{0, 1, \dots, N-1\}, \quad (2.54)$$

где

$$c(k) = \begin{cases} \frac{1}{\sqrt{2}} & \text{для } k = 0, \\ 1 & \text{для } k \in \{1, 2, \dots, N-1\}. \end{cases}$$

Матричная форма записи прямого ДКП имеет вид

$$[C(k)] = [\varphi(k, n)][f(n)], \quad (2.55)$$

где $[C(k)]$ – матрица-столбец спектральных коэффициентов ДКП размером $(1 \times N)$;

$[\varphi(k, n)]$ – матрица дискретного множества ортогональных функций ДКП размером $(N \times N)$ в соответствии с (10.19);

$[f(n)]$ – матрица-столбец дискретных значений (отсчетов) сигнала.

Соответствующее обратное одномерное преобразование в матричной форме записи представляется в виде

$$[f(n)] = [\varphi(k, n)]^{-1} [C(k)]. \quad (2.56)$$

А поскольку условие симметричности матрицы преобразования по отношению к ее главной диагонали не выполняется, то

$$[\varphi(k, n)]^{-1} = [\varphi(k, n)]^T, \quad (2.57)$$

и соответственно

$$[f(n)] = [\varphi(k, n)]^T [C(k)], \quad k, n \in \{0, 1, \dots, N-1\}. \quad (2.58)$$

Для преобразования двумерных фрагментов используется двумерное преобразование, ядро которого (без учета константы нормирования и с учетом замены одномерной последовательности отсчетов сигнала $f(n)$ двумерной матрицей отсчетов изображения $f(x, y)$) имеет вид

$$\widehat{\varphi}(u, x, v, y) = \cos\frac{2x+1}{2X}u\pi \cdot \cos\frac{2y+1}{2Y}v\pi, \quad \begin{cases} x, u \in \{0, 1, \dots, X-1\}, \\ y, v \in \{0, 1, \dots, Y-1\}. \end{cases} \quad (2.59)$$

Таким образом, прямое ДКП двумерного фрагмента форматом $(N \times M)$ запишется как

$$[g(u, v)] = [\varphi(u, x)][f(x, y)][\varphi(v, y)]^T, \quad (2.60)$$

где $[g(u, v)]$ – матрица спектральных коэффициентов ДКП размером $(N \times M)$;
 $[f(x, y)]$ – сигнальная матрица размером $(N \times M)$.
 $[\varphi(u, x)]$ – квадратная матрица функций ДКП размером $(N \times N)$:

$$\varphi(u, x) = \sqrt{\frac{2}{X}} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \cos \frac{2x+1}{2X} u\pi \end{bmatrix}, \quad \begin{cases} u \in \{1, 2, \dots, X-1\}, \\ x \in \{0, 1, \dots, X-1\}. \end{cases} \quad (2.61)$$

$[\varphi(v, y)]^T$ – квадратная матрица функций ДКП размером $(M \times M)$, транспонированная по отношению к матрице

$$\varphi(v, y) = \sqrt{\frac{2}{Y}} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \cos \frac{2y+1}{2Y} v\pi \end{bmatrix}, \quad \begin{cases} v \in \{1, 2, \dots, Y-1\}, \\ y \in \{0, 1, \dots, Y-1\}. \end{cases} \quad (2.62)$$

Соответствующая матричная форма записи двумерного обратного ДКП имеет вид

$$[f(x, y)] = [\varphi(u, x)]^T [g(u, v)][\varphi(v, y)]. \quad (2.63)$$

Если одномерное преобразование последовательности N отсчетов сигнала $f(n)$, $n \in \{0, 1, \dots, N-1\}$, во времени алгебраически интерпретировать как преобразование координат векторов в N -мерном метрическом пространстве, то исходя из основной формулы преобразования с ядром $[C_N]$

$$[c] = [C_N][f], \quad (2.64)$$

где $[c], [f]$ – матрицы-столбцы векторов c спектра и f исходного сигнала, векторам базиса пространства сигналов ставятся в соответствие векторы в спектральной области. В связи с ортогональностью ДКП также характеризуется тем, что векторы преобразования, за исключением первого, равны нулю. Координаты векторов преобразования представляют собой элементы соответствующей строки матрицы преобразования $[C_N]$.

Графическая иллюстрация координат векторов преобразования $C_j(i)$, $j, i \in \{0, 1, \dots, N-1\}$, для $N = 4$ приведена на рисунке 2.39.

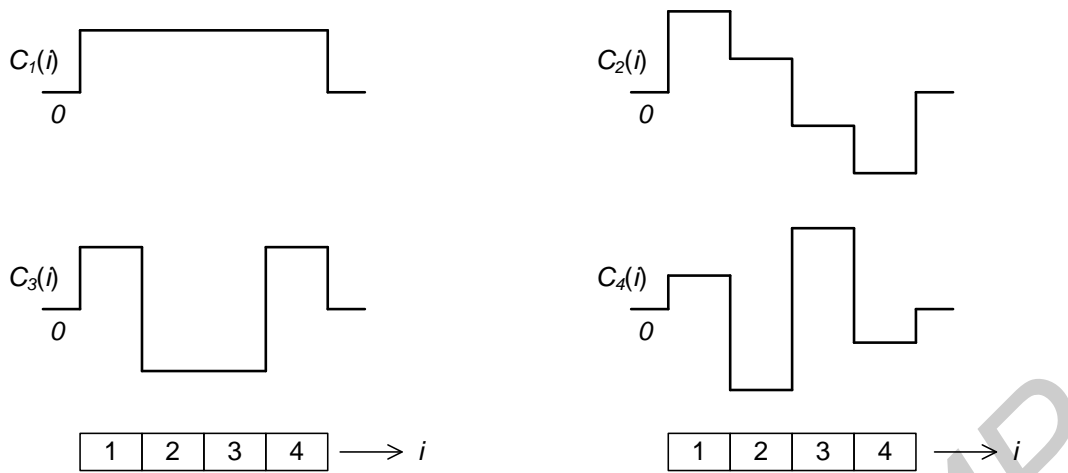


Рисунок 2.39 – Координаты векторов одномерного ДКП для $N = 4$

Если двумерное преобразование последовательности N^2 элементов изображения $f(x, y)$, $x, y \in \{0, 1, \dots, N-1\}$, алгебраически интерпретировать как преобразование координат векторов в N^2 -мерном пространстве, то в соответствии с обобщенной формулой

$$[c] = [{}^N C_{u,v}(x, y)][f], \quad (2.65)$$

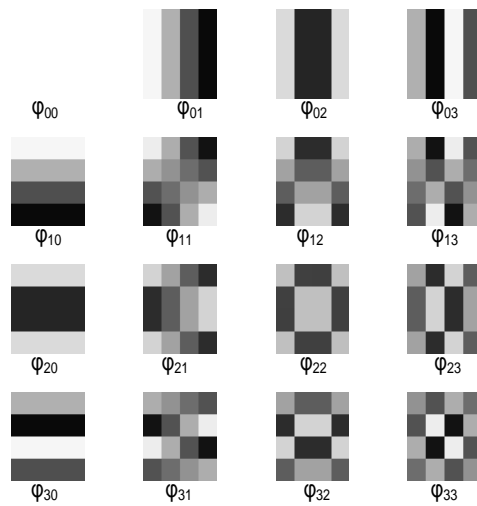
где $[c]$, $[f]$ – спектральная и сигнальная матрицы размером $(N \times N)$,

$[{}^N C_{u,v}(x, y)]$, $x, y, u, v \in \{0, 1, \dots, N-1\}$ – матрица преобразования размером $(N^2 \times N^2)$, отдельным векторам базиса пространства сигналов $f(x, y) = 1$, $x, y \in \{0, 1, \dots, N-1\}$, ставятся в соответствие векторы базиса в спектральной области.

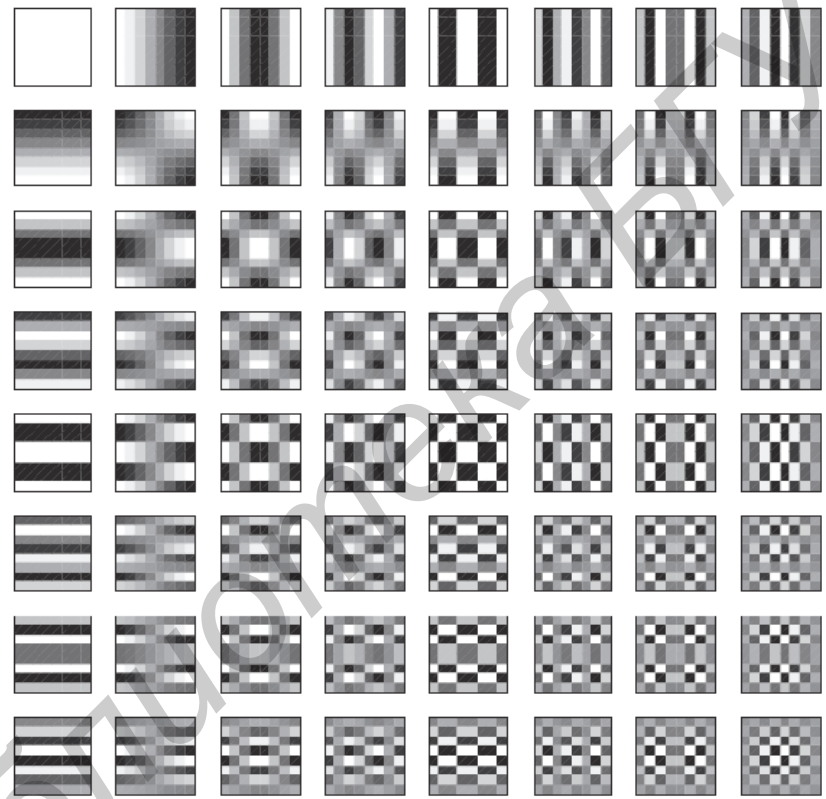
Координатам векторов преобразования, число которых составляет X^2 , в этом случае непосредственно соответствуют отдельные матрицы ${}^N C_{u,v}(x, y)$, принадлежащие множеству матриц преобразования $[{}^N C_{u,v}(x, y)]$ для $u, v \in \{0, 1, \dots, N-1\}$. Их графические представления зависят от типа ядра преобразования $[C_N]$, и для порядка преобразования $N = 4$ приведены на рисунке 2.41, а. Представленные двумерные функции являются так называемыми базисными изображениями ДКП и образованы двумерной дискретной системой ортогональных функций ДКП. Для разделяемого ядра преобразования они находятся как результат перемножения вектора-столбца \mathbf{c}_u^T и вектора-строки \mathbf{c}_v соответствующей матрицы преобразования

$$\mathbf{c}_{u,v} = \mathbf{c}_u^T \cdot \mathbf{c}_v. \quad (2.66)$$

Множество двумерных базисных изображений ДКП представляет собой множество двумерных функций, по которым раскладывается данный фрагмент изображения $f(x, y)$, показанных на рисунке 2.40, а (для размерности преобразования $N = 4$) и 2.40, б (для размерности преобразования $N = 8$).



а



б

а – размер преобразования 4×4 ; б – размер преобразования 8×8

Рисунок 2.40 – Графическое представление базисных изображений двумерного ДКП

2.4.5 Преобразование Уолша – Адамара

Наиболее известным и важным несинусоидальным ортогональным преобразованием является преобразование Уолша – Адамара (ПУА), одно из отличительных свойств которого состоит в том, что его расчет основан лишь

на операциях сложения (вычитания) и деления на 2^n , не требующих существенных вычислительных мощностей как при программной, так и при аппаратной реализации.

ПУА базируется на использовании квадратной матрицы Адамара, элементы которой могут иметь только значения $+1$ или -1 , а соответствующие строки и столбцы взаимно ортогональны. Строки матрицы можно рассматривать как дискретные функции двух переменных v (номер строки соответствует номеру функции) и u (номер элемента в строке соответствует дискретному моменту времени), определенных в целочисленных точках $0, 1, \dots, N-1$. Эти функции соответствуют дискретным функциям Уолша, упорядоченным по Адамару.

В общем случае матрица Адамара $[H_{2N}]$ порядка $2N$ образуется из матриц порядка N по закону

$$[H_{2N}] = \begin{bmatrix} H_N & H_N \\ H_N & -H_N \end{bmatrix}. \quad (2.67)$$

Матрица Адамара первого порядка $[H_1] = 1$, таким образом, для $[H_2]$ выполняется равенство

$$[H_2] = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

и т. д. Графическое представление матрицы Адамара размера 8×8 приведено на рисунке 2.41.

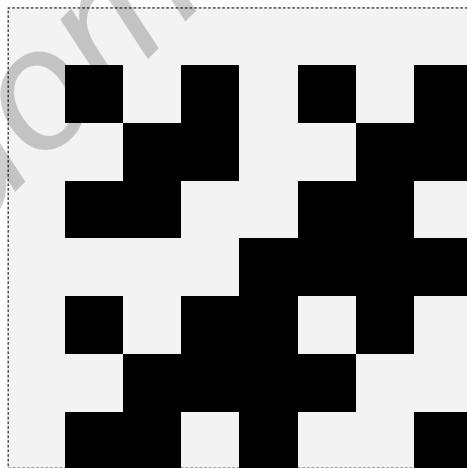


Рисунок 2.41 – Графическое представление матрицы Адамара размера 8×8

Прямое дискретное преобразование Уолша – Адамара исходного фрагмента изображения размером $(N \times N)$ элементов записывается в матричном виде в соответствии с формулой

$$[g(u, v)] = \frac{1}{N} [H(u, v)] [f(x, y)] [H(u, v)], \quad (2.68)$$

где $[g(u, v)]$ – матрица трансформанта ПУА;

u, v – обобщенные частоты в области трансформанты;

$[H(u, v)]$ – матрица Адамара порядка N ;

$[f(x, y)]$ – матрица отсчетов сигнала, соответствующего фрагменту изображения;

x, y – координаты, определяющие положение отсчета в плоскости изображения.

Выражение (2.68) представляет собой разложение исходного сигнала $f(x, y)$ по множеству прямоугольных функций Уолша и, соответственно, определяет их амплитуды. Применительно к преобразованию изображения осуществляется представление исходного изображения в системе базисных изображений преобразования (рисунок 2.42, а, б).

Обратное преобразование Уолша – Адамара определяется выражением

$$[f(x, y)] = \frac{1}{N} [H(u, v)] [g(u, v)] [H(u, v)]. \quad (2.69)$$

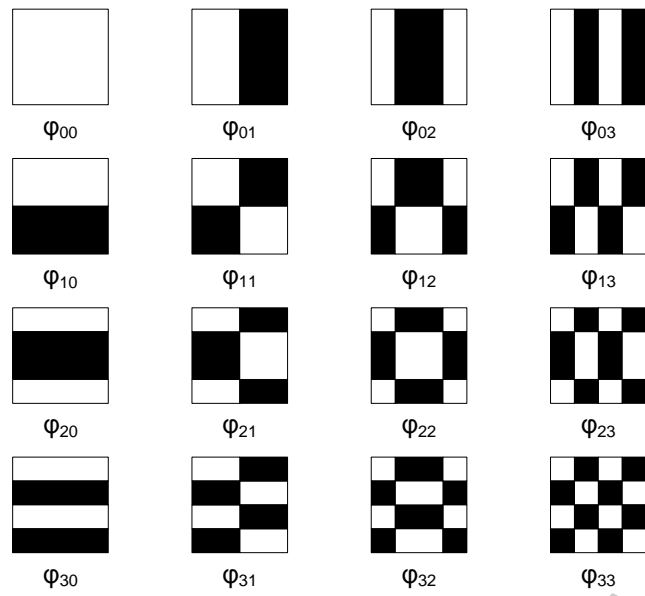
Обобщим основные свойства преобразования Уолша – Адамара с точки зрения цифрового кодирования ТВ-сигналов.

1 Трансформанта $F(0, 0)$ является мерой средней яркости исходного фрагмента. Это непосредственно следует из выражения (2.53), если принять, что $f(x, y)$ – константа.

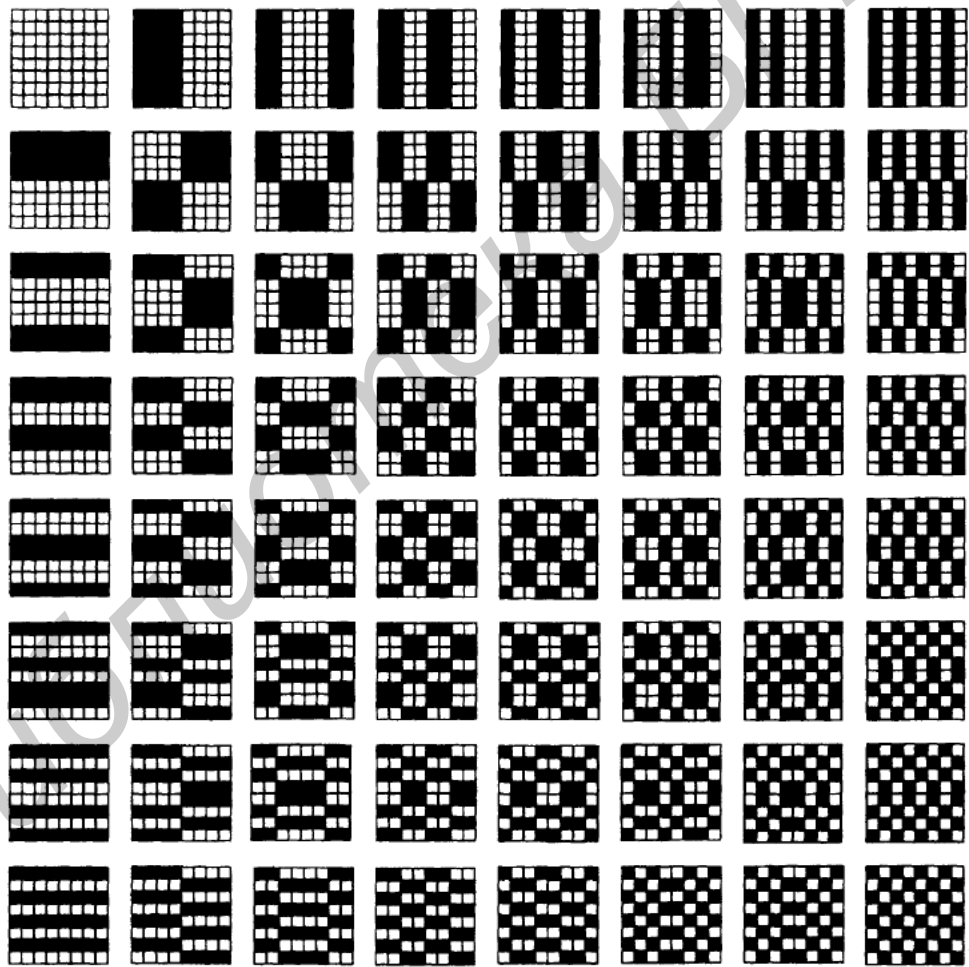
2 Если $f(x, y)$ – сигнал с определенной энтропией, то энтропия, соответствующая функции $g(u, v)$, будет такой же. Это позволяет осуществлять преобразование исходного сигнала без информационных потерь.

3 Выполняется равенство энергий исходного и преобразованного фрагментов (аналогично соотношению Парсеваля для преобразования Фурье). Поэтому, если трансформант имеют несколько большую амплитуду, то остальные должны характеризоваться обязательно малой амплитудой.

4 Динамический диапазон трансформанта значительно больше, чем динамический диапазон исходного сигнала, что повышает эффективность кодирования. Если исходный сигнал $0 \leq f(x, y) \leq 1$, то трансформанты лежат в диапазоне: $0 \leq F(0, 0) \leq N^2$ и $-0,5 N^2 \leq F(k, l) \leq 0,5 N^2$ для $k, l \geq 1$.



a



б

а – базисные изображения размером 4×4 ;
 б – базисные изображения размером 8×8

Рисунок 2.42 – Графическое представление базисных изображений двумерного преобразования Уолша – Адамара

2.4.6 Вейвлетное преобразование

Вейвлетное (wavelet) преобразование, широко используемое в настоящее время при сжатии изображений, основано на семействах фильтров с коэффициентами, которые эквивалентны дискретным вейвлетным функциям [16; 49]. Основная операция дискретного вейвлетного преобразования применяется к дискретному сигналу, состоящему из N последовательных отсчетов. Пара фильтров применяется к сигналу для его разделения на две составляющие: низкочастотную субполосу (L) и высокочастотную субполосу (H). Каждая субполоса прореживается через один элемент, и в каждой последовательности остается по $N/2$ отсчетов. При правильном выборе фильтров – их АЧХ должны быть взаимнообратными (иногда их называют зеркальными) такая операция является обратимой.

Этот подход можно обобщить для применения к двумерным сигналам, соответствующим дискретным изображениям (рисунок 2.43). Каждая строка двумерного изображения разделяется соответствующими горизонтальными фильтрами на низкочастотные и высокочастотные составляющие, которые прореживаются в 2 раза. Результатом являются два субизображения – L и H. После этого над ними повторяется вышеописанная операция, но с применением вертикальных фильтров, при этом получается 4 субизображения (LL, LH, HL и HH).

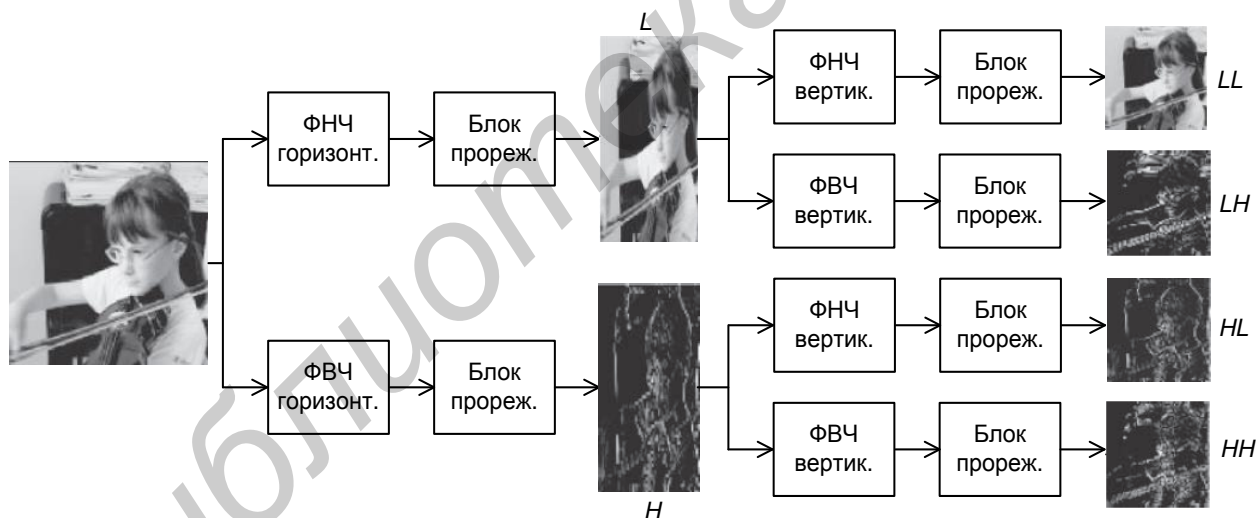


Рисунок 2.43 – Процесс двумерного вейвлетного разложения изображения

Четыре «субполосных» изображения для наглядности можно объединить как показано на рисунке 2.44. LL – исходное изображение, из которого удалены высокочастотные составляющие по горизонтали и вертикали и сделано прореживание с шагом 2 по каждому направлению – теперь оно несет информацию о крупных деталях основного изображения; HL – высокочастотная составляющая по горизонтали и оставшаяся низкочастотная составляющая по вертикали; LH – низкочастотная составляющая по горизонтали и высокочастотная составляющая по вертикали и, наконец, HH –

высокочастотные составляющие и по горизонтали, и по вертикали. Эти четыре субполосы содержат в себе всю информацию об исходном изображении, но разреженная структура LL, LH, HL и HH делает их удобными для сжатия, причем точность представления – и, соответственно, степень сжатия – для каждой субсоставляющей выбирается различной.

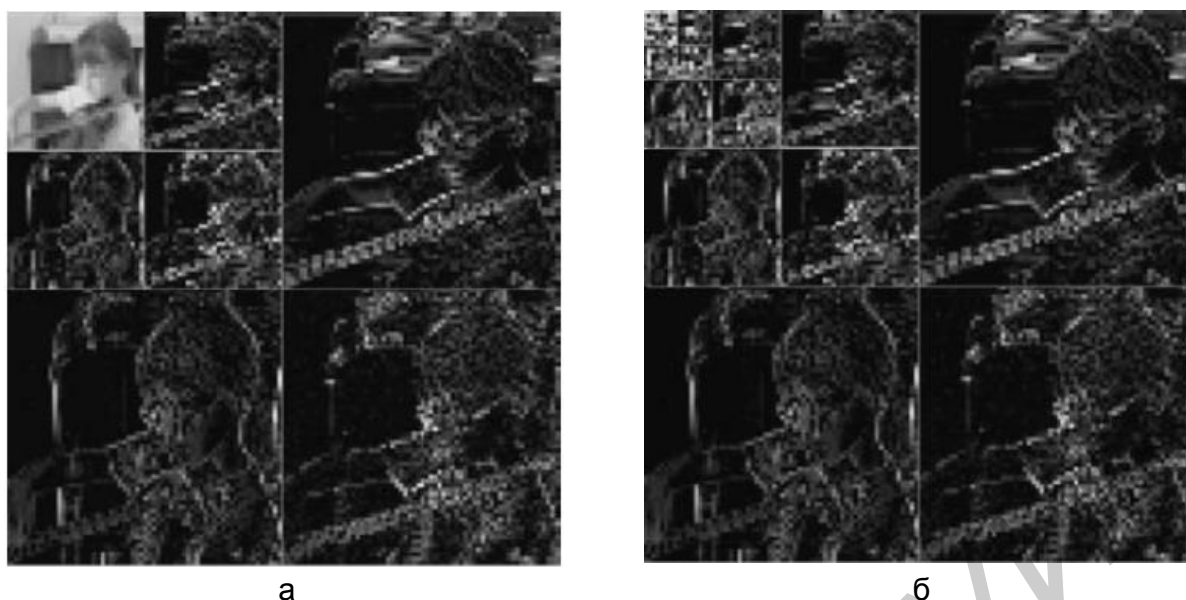


Рисунок 2.44 – Объединенные субизображения после первого шага вейвлетного разложения

При использовании в приложениях компрессии данных описанное выше двумерное вейвлетное разложение повторно применяется к части LL, из которой формируется четыре новых субполосных изображения. Эта процедура многократно повторяется применительно к низкочастотной составляющей двумерного сигнала (это всегда верхний левый квадрант). Результатом ее применения служит дерево субполосных изображений. На рисунке 2.45 приведено это дерево после второй (а) и пятой (б) стадий разложения.

Большинство отсчетов (коэффициентов) высокочастотных составляющих близки к нулю (почти черный цвет на рисунке), и эффект сжатия можно получить, удалив эти малозначащие коэффициенты либо применив к ним более грубое квантование.

В декодере для восстановления изображения выполняются операции, аналогичные показанным на рисунке 2.44, но в обратную сторону. Соответственно, операция прореживания заменяется на интерполяцию.



а – субизображения после второго шага; б – субизображения после пятого шага
 Рисунок 2.45 – Объединенные субизображения вейвлетного разложения

2.5 Гибридное кодирование

Произведенный в подразделах 2.2 – 2.4 анализ методов сокращения избыточности изображений показал, что наибольшая эффективность сжатия наблюдается при их совместном использовании. Кодеры изображений, реализующие комплекс перечисленных функций, называют *гибридными*.

Типовая структурная схема гибридного кодера приведена на рисунке 2.46, а. Представленный кодер реализует все три операции кодирования изображений – кодирование с предсказанием, кодирование с преобразованием и статистическое кодирование. Также в схему кодера входят блоки деквантования и обратного преобразования, необходимые для восстановления закодированного изображения, которое будет в дальнейшем использоваться для вычисления разностного изображения.

Работу кодера рассмотрим в двух режимах компрессии – внутрикадровой и межкадровой. Для упрощения предположим, что внутрикадровое предсказание не используется. Во внутрикадровом режиме, который используется при кодировании опорных кадров (I-кадры), а также для сжатия статических изображений, выполняется только кодирование с преобразованием и статистическое. Поскольку цепь предсказания в данном случае разомкнута, на блок прямого преобразования поступают блоки отсчетов изображения $f(x, y)$. Типичный размер преобразуемого блока – 4×4 либо 8×8 элементов. В стандартных алгоритмах компрессии (MPEG, JPEG) используется либо дискретно-косинусное, либо вейвлетное преобразования, теоретически может быть применено любое ортогональное преобразование.

В результате преобразования получаются k -разрядные коэффициенты $g(u, v)$, причем $k > t$, в противном случае преобразование будет необратимым за счет погрешности вычислений. Коэффициент $g(0, 0)$ определяет постоянную составляющую преобразуемого блока (либо среднюю яркость), остальные характеризуют степень изменений его отсчетов по горизонтали и вертикали.

В блоке квантования осуществляется изменение разрядности представления каждого из отсчетов (переквантование) в более грубую сторону, что приводит к сокращению количества бит данных для представления блока изображения. После переквантования матрица коэффициентов посредством зигзагообразного считывания (развертки) преобразуется в одномерную последовательность (вектор). Завершает процесс компрессии статистический кодер.

В межкадровом режиме цепь предсказания замыкается, в результате чего кодированию с преобразованием подвергается уже не исходное изображение $f(x, y)$, а обладающее существенно меньшей энтропией разностное $\Delta f(x, y)$. Оно получается путем вычитания из исходного изображения $f(x, y)$ реконструированного блоком предсказания изображения $f_{\text{пр}}(x, y)$. Исходными данными для реконструкции являются восстановленные блоками деквантования и обратного преобразования отсчеты одного либо нескольких предыдущих (в порядке поступления на кодер) кадров $f'(x, y)$. В зависимости от принятой модели предсказания в блоке памяти может храниться от одного до нескольких реконструированных кадров.

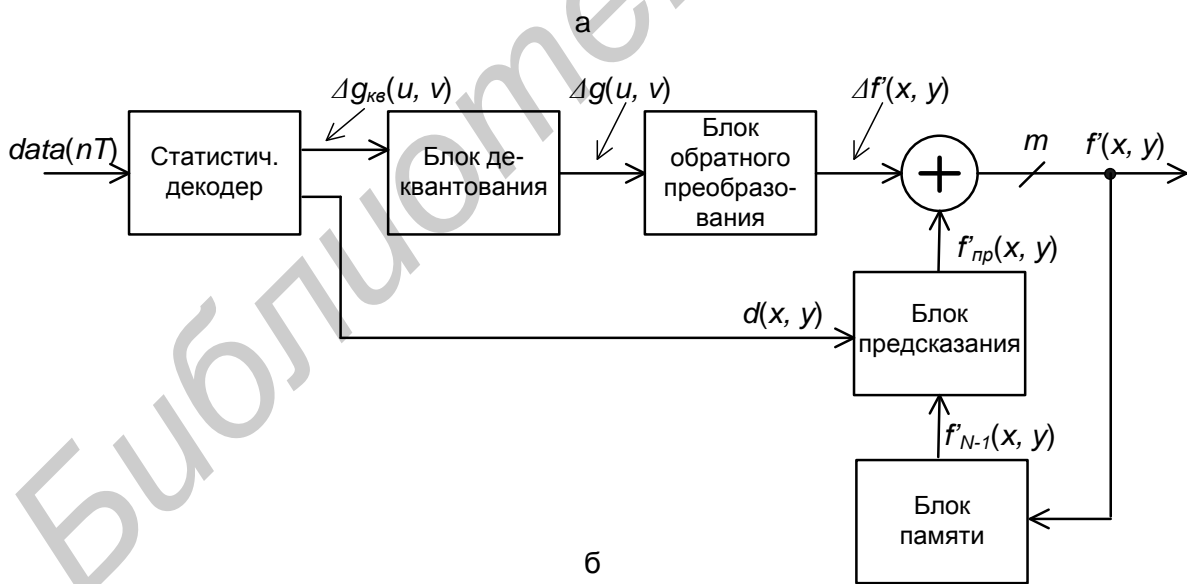
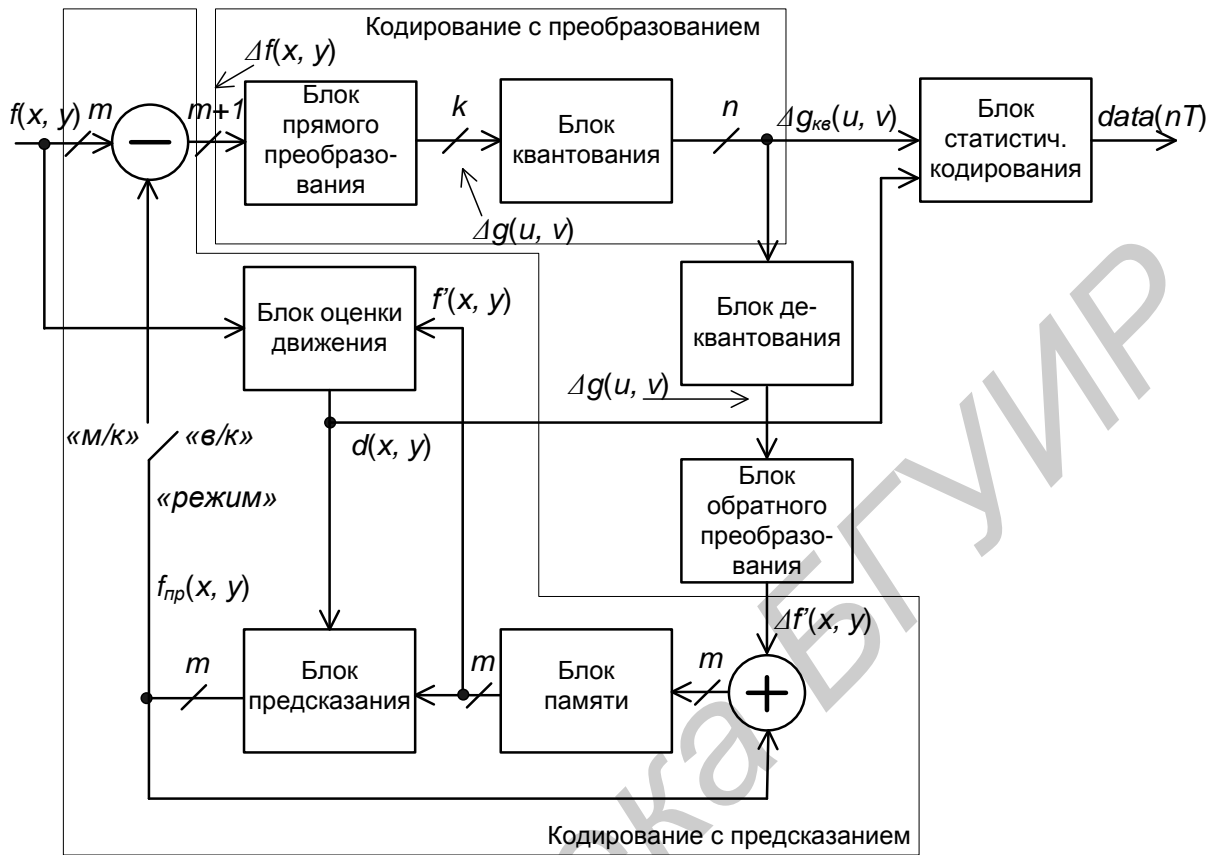
Путем сравнения предыдущего кадра с текущим в блоке оценки движения вычисляются вектора движения $d(x, y)$ для каждого макроблока (комбинации из нескольких блоков) изображения, которые передаются статистическому кодеру и далее декодеру приемной стороны, а также блоку предсказания для построения прогноза кадра.

Чем точнее оказывается прогноз, тем меньшей информационной емкостью обладает разностное изображение, соответственно, уменьшается количество бит данных, необходимых для передачи квантованных коэффициентов.

В декодере (см. рисунок 2.46, б) вначале восстанавливаются квантованные коэффициенты преобразования $\Delta g_{\text{кв}}(u, v)$ и вектора движения $d(x, y)$, после чего первые подаются на блок деквантования, вторые – на блок предсказания. Для правильного построения прогноза $f'_{\text{пр}}(x, y)$ предсказатель декодера должен работать идентично предсказателю кодера. Переданное изображение $f'(x, y)$ восстанавливается в сумматоре путем сложения спрогнозированного на базе предыдущих кадров изображения $f'_{\text{пр}}(x, y)$ с разностным изображением $\Delta f'(x, y)$, восстановленным на выходе блока обратного преобразования.

Описанная комбинация кодирования с преобразованием и статистического кодирования позволяет добиться степени сжатия порядка 3...5 раз при практически полном визуальном сходстве исходного и восстановленного

изображений. Дальнейшее повышение степени сжатия возможно за счет использования кодирования с предсказанием.



а – кодирование; б – декодирование

Рисунок 2.46 – Структурная схема типичного гибридного кодека изображений

Описанный алгоритм применяется в стандартах компрессии изображений MPEG-1 и MPEG-2 и позволяет при кодировании сигналов телевизион-

ного вещания стандартной четкости добиться сжатия потока данных в 10...15 раз при практически полной идентичности исходных и восстановленных изображений (студийное качество). Для целей вещания изображения компрессируются сильнее – в 20...80 раз, до достижения скорости потока данных на выходе кодера в 2...10 Мбит/с (в зависимости от характера изображений), субъективное качество при этом остается в пределах от отличного до очень хорошего.

Дальнейшее повышение эффективности кодирования может быть достигнуто за счет следующих факторов:

- повышения точности межкадрового предсказания;
- использования внутрикадрового предсказания;
- использования предсказания при передаче векторов движения, квантованных коэффициентов преобразования и информации о режимах прогнозирования;
- выполнения повторного ортогонального преобразования над сгруппированными определенным образом коэффициентами, полученными после первого преобразования.

Повышение точности межкадрового предсказания, как уже отмечалось, минимизирует количество информации в разностном изображении. Достичь его можно путем уменьшения размера блока или макроблока, для которого выполняется предсказание, а также усложнением модели предсказания – вычислению прогноза не по одному предыдущему кадру, а по нескольким, причем не только предыдущим, но и последующим (двунаправленное предсказание). Для реализации двунаправленного предсказания последовательность кадров изображения, подаваемых на декодер, должна быть изменена по отношению к исходной последовательности, для чего в состав кодера вводится входной буфер – блок памяти (рисунок 2.47), емкость которого определяется максимальной задержкой кадра изображения.

Оба рассмотренных метода усложняют схему кодера и декодера, но если недостатки усложненной модели предсказания на этом заканчиваются, то уменьшение размера предсказываемого блока увеличивает количество этих блоков, и, соответственно, количество передаваемых векторов движения. Данная проблема решается двумя путями – использованием блоков переменного размера, адаптивного к содержанию изображения, и передаче не абсолютных значений векторов движения, а разностей между ними.

Структура большинства реальных изображений неравномерна с точки зрения информационной составляющей – в них присутствуют как области постоянной либо слабо изменяющейся яркости либо цветности (фон, другие крупные детали), так и участки с резкими перепадами и мелкими деталями. В первом случае выгоднее обрабатывать блоки больших размеров и минимизировать количество передаваемых векторов движения, во втором, наоборот, предсказывать положения достаточно маленьких блоков, но с высокой точностью, тем самым уменьшая энтропию разностного изображения.

Если вернуться к анализу движения блоков изображения (см. подраздел 2.3.3), то можно отметить, что достаточно часто крупные фрагменты изображения смещаются в одном направлении и примерно на одинаковое расстояние (см. рисунок 2.32). Соответственно, и вектора движения $d(x_{MB}, y_{MB})$ (см. рисунок 2.46) будут в данном случае коррелированы (практически одинаковы), что позволяет и для них использовать кодирование с предсказанием – передавать разностные значения $\Delta d(x_{MB}, y_{MB})$, что потребует меньшего объема бит.

Следующим методом повышения эффективности компрессии, применимым, в отличие от предыдущего, как для динамических, так и для статических изображений, является использование внутрикадрового предсказания. Оно выполняется в пространственной области – текущие блоки изображения предсказываются на основе ранее закодированных соседних, расположенных, как правило, слева и сверху. В данном случае также возможно использование блоков переменного размера, адаптированного к содержанию изображения. Декодеру (рисунок 2.47) при этом должна быть передана дополнительная информация о режиме предсказания каждого блока $md(x_{MB}, y_{MB})$.

На блоки прямого преобразования и квантования при этом поступает разностное изображение $\Delta f(x, y)$, вычисленное путем вычитания из исходного изображения $f(x, y)$ прогноза $f_{\text{ПР}}(x, y)$. Отличие от межкадрового предсказания заключается в том, что для кадра с номером N прогноз строится по этому же кадру, а не по предыдущему, следующему либо их комбинации. Результатом внутрикадрового кодирования с предсказанием опять же является снижение количества информации в передаваемом разностном кадре.

При кодировании с преобразованием областей изображения с постоянной либо слабо меняющейся яркостью наблюдается заметная корреляция между значениями коэффициентов преобразования соседних либо близко расположенных блоков, что является еще одним резервом улучшения компрессии. Практическая реализация возможна двумя путями – использования кодирования с предсказанием применительно к коэффициентам преобразования, либо повторным ортогональным преобразованием сгруппированных коэффициентов из соседних блоков. Примером использования первого метода являются стандарты сжатия статических изображений JPEG/JPEG-2000, второй используется в MPEG-4 AVC/H.264 [67]. Повторному кодированию подвергаются коэффициенты $g(0, 0)$, несущие информацию о постоянной составляющей блоков отсчетов.

Использование перечисленных методов повышения эффективности компрессии дает дополнительный выигрыш в 1,5...2 раза. Например, при использовании кодирования по стандарту MPEG-4 AVC/H.264 можно добиться очень хорошего либо отличного качества восстановленных изображений при скорости потока данных 1...3 Мбит/с при передаче цифровых телевизионных сигналов стандартной четкости и 5...20 Мбит/с – для телевидения высокой четкости. При передаче изображений меньшего разрешения и частоты кадров возможно снижение скорости вплоть до 64 кбит/с, что дает возможность использования стандартного цифрового телефонного канала.

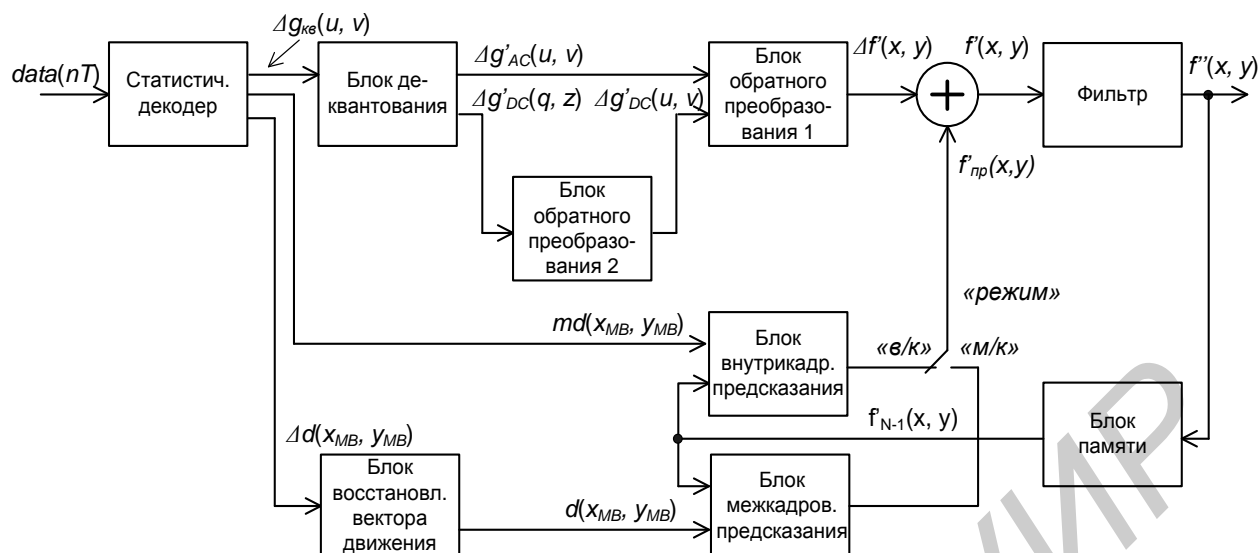


Рисунок 2.48 – Структурная схема гибридного декодера изображений

Новый стандарт компрессии изображений H.265 [69], принятый ITU-T в начале 2013 г. на основе стандарта HEVC [70, 71], также использует концепцию гибридного кодирования, однако допускает по сравнению с H.264 еще более гибкую сегментацию изображения на кодируемые блоки, причём как при кодировании с предсказанием, так и при кодировании с преобразованием. В результате эффективность компрессии повышается в сравнении с тем же стандартом H.264 на 25–50 %, в зависимости от используемого стандарта разложения изображений.

ЗАКЛЮЧЕНИЕ

В соответствии с общими принципами построения цифровых систем передачи информации они включают в себя три фундаментальных процесса: кодирование источника сообщений, кодирование канала (помехоустойчивое кодирование) и цифровую модуляцию.

В первой части учебно-методического пособия «Цифровое телевидение» рассмотрены вопросы, относящиеся к кодированию источника сообщений, т. е. сигналов телевизионных изображений. С общих позиций проанализированы и обоснованы операции цифрового представления сигналов телевизионного вещания стандартной и высокой чёткости, детально, с математическим обоснованием, рассмотрены основные алгоритмы сокращения избыточности, приведены рекомендации по их практическому применению.

Во второй части учебно-методического пособия будут освещены и другие проблемы цифрового телевизионного вещания.

ПРИЛОЖЕНИЕ А

(справочное)

Временные диаграммы и уровни сигналов

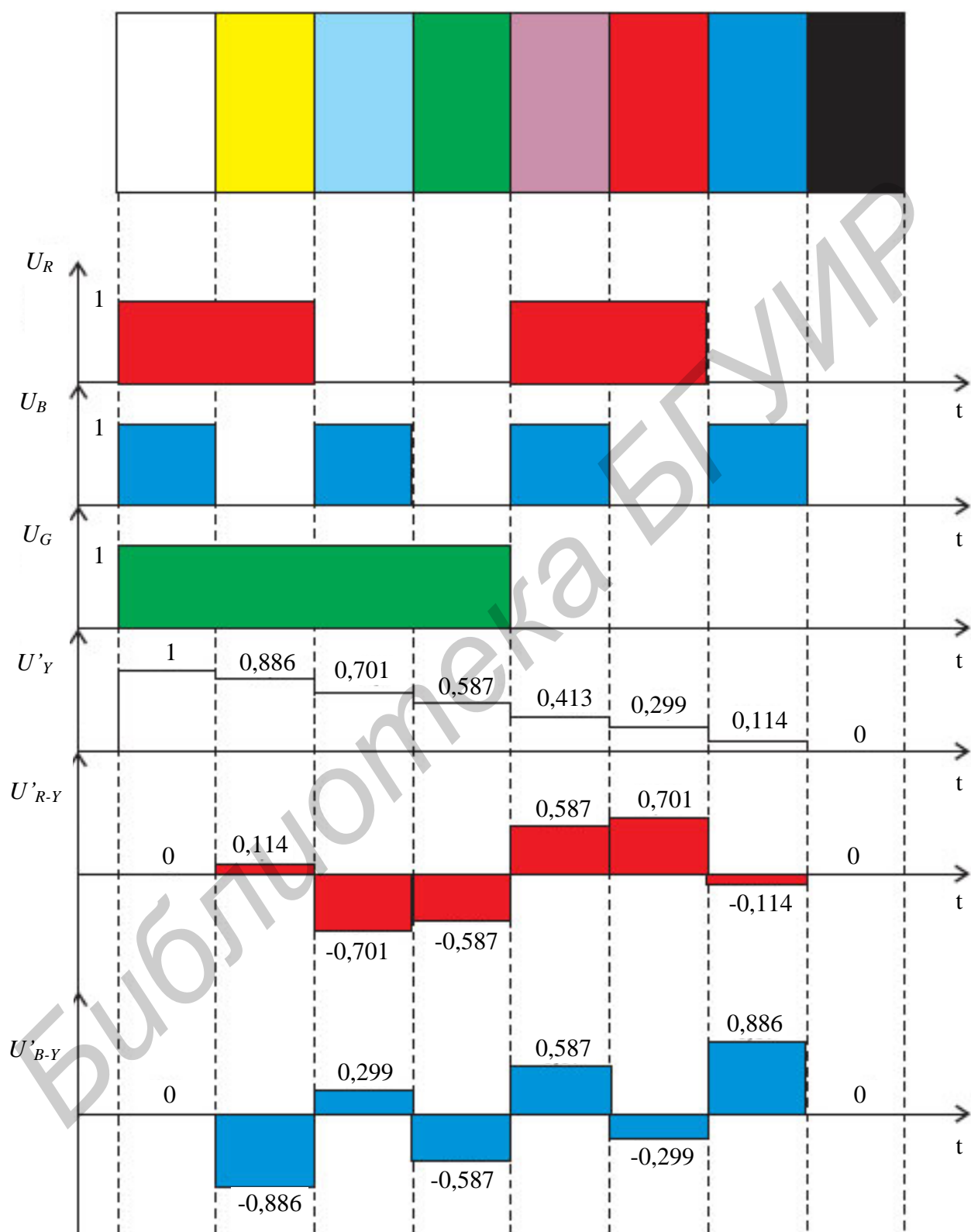
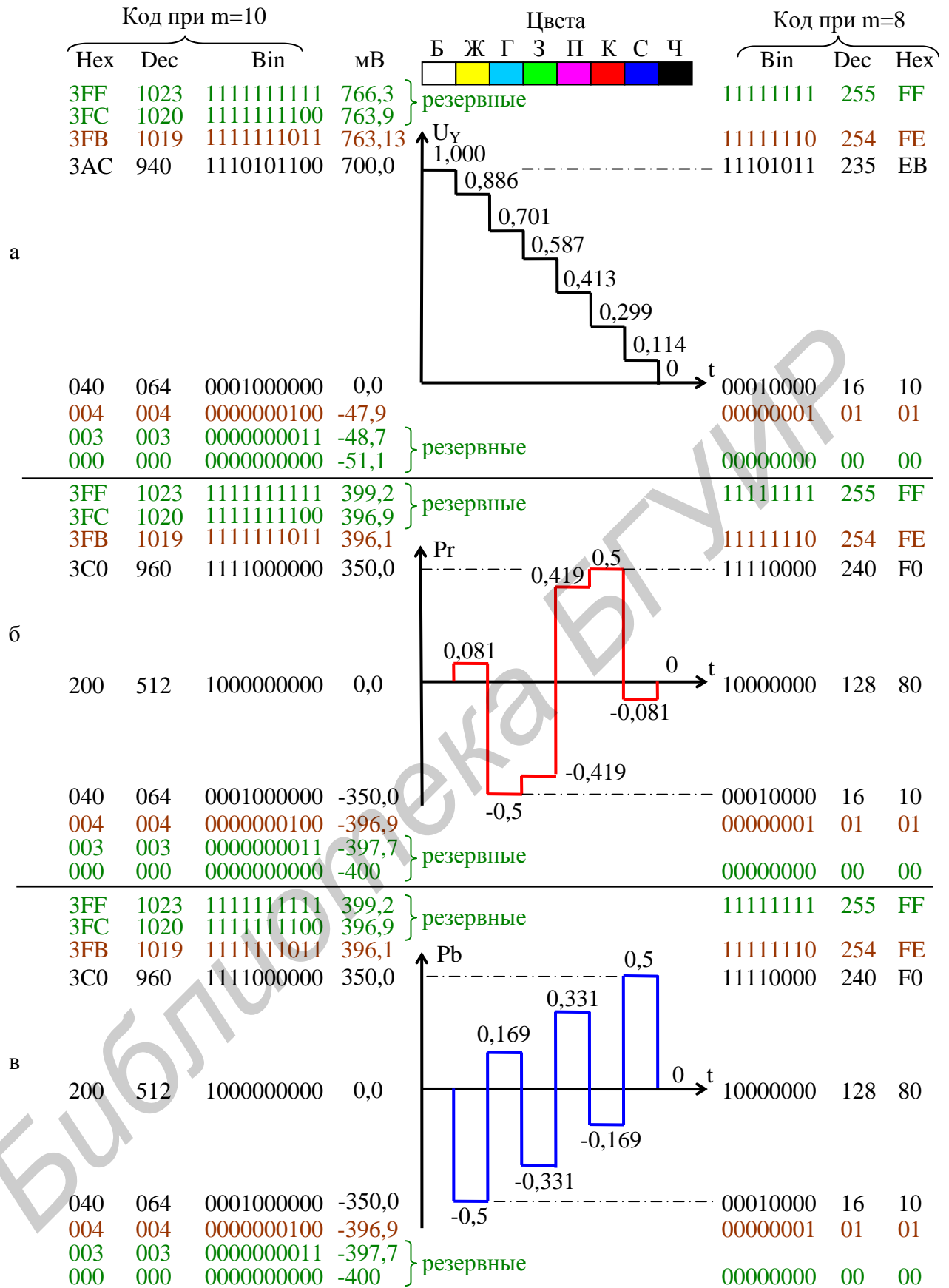


Рисунок А.1 – Временные диаграммы сигналов вертикальных цветных полос номенклатуры 100/0/100/0



а – сигнал яркости; б, в – нормированные ЦРС

Рисунок А.2 – Уровни сигналов при 8- и 10-разрядном кодировании и двоичном (Bin), десятичном (Dec), шестнадцатеричном (Hex) представлениях

ПРИЛОЖЕНИЕ Б

(справочное)

Примеры испытательного и реального изображений, для сигналов которых построены гистограммы распределения вероятности значений

и вычислена энтропия

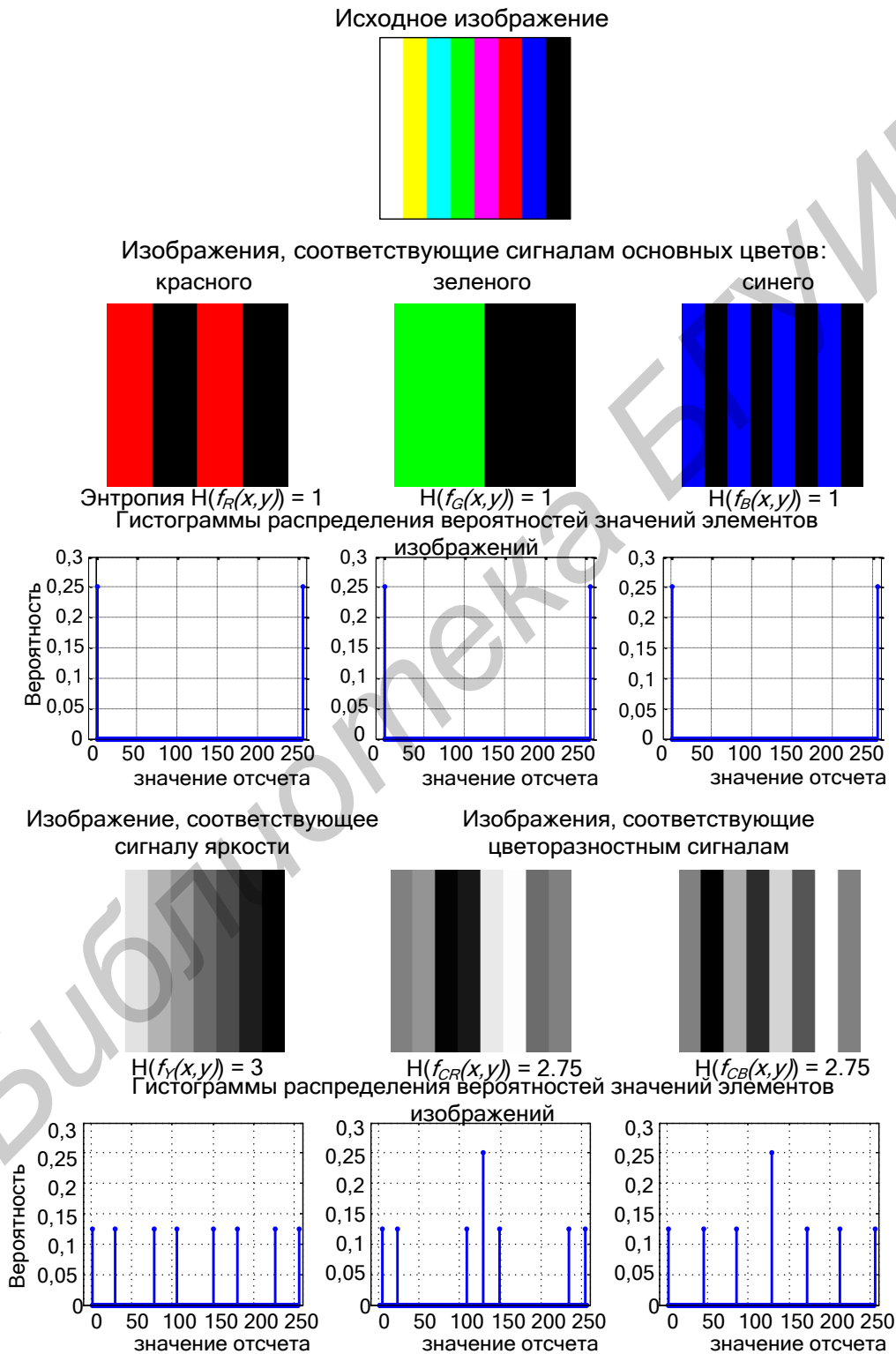


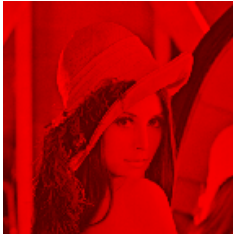
Рисунок Б.1 – Оценка кодовой избыточности испытательного изображения

Исходное изображение



Изображения, соответствующие сигналам основных цветов:

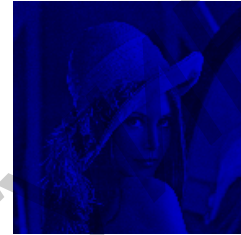
красного



зеленого



синего



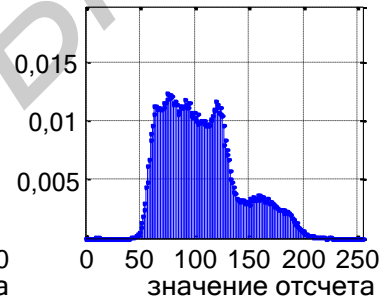
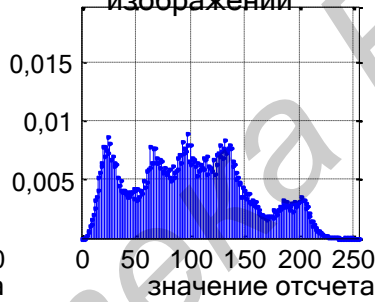
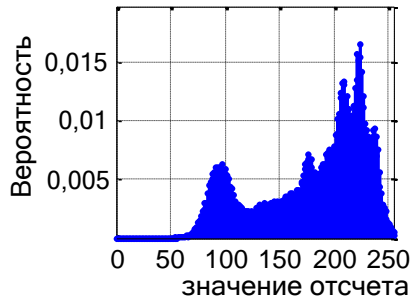
Энтропия $H(f_R(x,y)) = 7.25$

$H(f_G(x,y)) = 7.59$

$H(f_B(x,y)) = 6.97$

Гистограммы распределения вероятностей значений элементов

изображений



Изображение, соответствующее
сигналу яркости



$H(f_Y(x,y)) = 7.45$

Изображения, соответствующие
цветоразностным сигналам



$H(f_{CR}(x,y)) = 5.60$



$H(f_{CB}(x,y)) = 5.50$

Гистограммы распределения вероятностей значений элементов

изображений

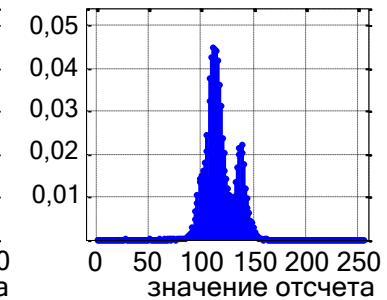
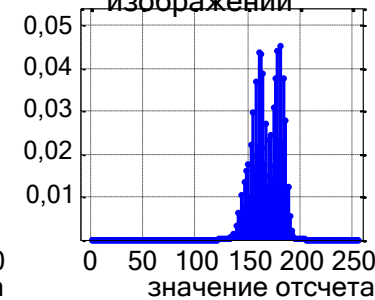
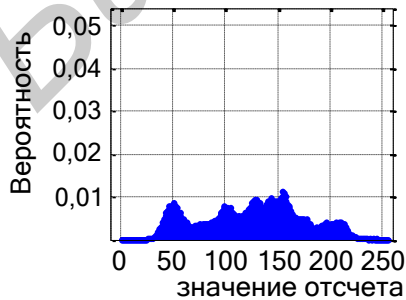


Рисунок Б.2 – Оценка кодовой избыточности реального изображения

ЛИТЕРАТУРА

- 1 Зубарев, Ю. Б. Цифровое телевизионное вещание: основы, методы, системы / Ю. Б. Зубарев, М. И. Кривошеев, И. Н. Красносельский. – М. : НИИР, 2001. – 568 с. ; Эко-Трендз, 2003. – 568 с.
- 2 Локшин, Б. А. Цифровое вещание – от студии к телезрителю / Б. А. Локшин. – М. : Компания Сайрус Системс, 2001. – 448 с.
- 3 Телевидение : учебник для вузов / В. Е. Джакония [и др.] ; под ред. В. Е. Джаконии. – 3-е изд. – М. : Радио и связь, 2004. – 616 с.
- 4 Быков, Р. Е. Основы телевидения и видеотехники : учебник для вузов / Р. Е. Быков. – М. : Горячая линия – Телеком, 2006. – 399 с.
- 5 Ткаченко, А. П. Цифровое представление сигналов изображения и звукового сопровождения : учеб. пособие по телевизионным дисциплинам / А. П. Ткаченко, П. А. Капура, А. Л. Хоминич. – Минск : БГУИР, 2003. – 56 с.
- 6 ГОСТ 7845-92. Система вещательного телевидения. Основные параметры. Методы измерений.
- 7 Птачек, М. Цифровое телевидение. Теория и техника : пер. с чешск. / М. Птачек. – М. : Радио и связь, 1990. – 528 с.
- 8 Ткаченко, А. П. Цветное телевидение / А. П. Ткаченко. – Минск : Беларусь, 1981. – 254 с.
- 9 Кириллов, В. И. Телевидение и передача изображений : учеб. пособие для вузов / В. И. Кириллов, А. П. Ткаченко. – Минск : Выш. шк. 1988. – 312 с.
- 10 Бытовая радиоэлектронная техника : энциклопед. справочник / А. П. Ткаченко [и др.] ; под общ. ред. А.П. Ткаченко. – Минск : БелЭн, 1995. – 832 с.
- 11 Ткаченко, А. П. Информатика. Информация. Скорость передачи информации / А. П. Ткаченко // Энциклопедия для школьников и студентов. В 12 т. Том 1. Информационное общество. XXI век / под общ. ред. В. И. Стражева. – Минск : Беларус. Энцыкл. імя П. Бровкі, 2009. – 528 с. (С. 307–308, 315–317, 374).
- 12 Ткаченко, А. П. Телевидение. Цветное телевидение. Цифровое телевидение / А.П. Ткаченко // Энциклопедия для школьников и студентов. В 12 т. Том 1. Информационное общество. XXI век / под общ. ред. В. И. Стражева. – Минск : Беларус. Энцыкл. імя П. Бровкі, 2009. – 528 с. (С.380–386).
- 13 Хоминич, А. Л. Пиксел // А. Л. Хоминич // Энциклопедия для школьников и студентов. В 12 т. Том 1. Информационное общество. XXI век / под общ. ред. В. И. Стражева. – Минск : Беларус. энцыкл. імя П. Броўкі, 2009. – 528 с. (С. 351–352).
- 14 Ткаченко, А. П. Повышение качества изображения и звукового сопровождения : учеб. пособие для студ. / А. П. Ткаченко, А. Л. Хоминич. В 2 ч. Ч. 1: Тракты промежуточной частоты изображения и звукового сопровождения. – Минск : БГУИР, 2001. – 55 с.; Ч. 2: Повышение качества изображения при формировании ТВ-сигнала. – 2004. – 70 с.

- 15 Килин, С. Я. Информация / С. Я. Килин, Д. Б. Хорошко // Энциклопедия для школьников и студентов. В 12 т. Том 2. Физика. Математика / под общ. ред. Н. А. Поклонского. – Минск : Беларус. Энцыкл. імя П. Броўкі, 2010. – 528 с. (С. 91–96)
- 16 Гонсалес, Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс; пер. с англ. – М. : Техносфера, 2006. – 1072 с.
- 17 Гонсалес, Р. Цифровая обработка изображений в среде MATLAB / Р. Гонсалес, Р. Вудс, С. Эддинс ; пер. с англ. – М.: Техносфера, 2006. – 616 с.
- 18 Яне, Б. Цифровая обработка изображений / Б. Яне ; пер. с англ. – М. : Техносфера, 2007. – 584 с.
- 19 Уразова, С. Телевидение ультравысокой четкости уже реальность / С. Уразова // 625. Науч.-техн. журнал. –2009. –№ 5. – С. 44–48.
- 20 Бителева, А. За пределами HDTV / А. Бителева // Телеспутник. – 2013. – №2. – С. 82–84.
- 21 Капуру, П. А. Телевидение. Телевидение и отображение информации : лаб. практикум / П. А. Капуру, А. Л. Хоминич. – Минск : БГУИР, 2010. – 43 с.
- 22 Стандарты цифрового телевизионного вещания // 625. Науч.-техн. журнал. – 2006. – №5. – С. 82–86.
- 23 Кондар, Р. Организация данных для инфраструктуры 3 Гбит/с / Ренди Конрад, П. Брискау // 625. Науч.-техн. журнал. – 2009. – №6. – С. 62–68.
- 24 Липкович, Э. Б. Проектирование и расчет систем цифрового спутникового вещания: метод. пособие / Э. Б. Липкович, Д. В. Кисель. – Минск : БГУИР, 2006. – 135 с.
- 25 Липкович, Э. Б. Наземные цифровые системы телевизионного вещания: метод. пособие / Э. Б. Липкович. – Минск : БГУИР, 2010. – 83 с.
- 26 Серов, А. DVB–T2 – цифровое телевидение второго поколения / А. Серов // 625. Науч.-техн. журнал. –2009. – №7. – С.70–73.
- 27 Ефимов, А. В. Оценка эффективности использования стандарта DVB–S2 для спутниковых систем связи / А. В. Белькович, В. П. Ефимов // Электросвязь. – 2008. – №10. – С. 65–68.
- 28 Robert, J. DVB–C2–стандарт передачи по сетям КТВ следующего поколения / J. Robert, C. Schaaf, L. Stadelmeëer // Телеспутник. – 2009. – №12.
- 29 Гриб, Д. В. Эффективность использования телевизионного канала связи / Д. В. Гриб, А. П. Ткаченко // Современные средства связи. Материалы XIV междун. науч.-техн. конф. (29 сент.–1 окт. 2009 г.). – Минск : ВГКС, 2009. – С. 59–60.
- 30 ITU-R Recommendation BT.601-6. Studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios. – ITU-R, 2007.
- 31 ITU-R Recommendation BT.1361. Worldwide unified colorimetry and related characteristics of future television and imaging systems. – ITU-R, 1998.
- 32 ITU-R Recommendation BT.656-5. Interface for digital component video signals in 525-line and 625-line television systems operating at the 4:2:2 level of Recommendation ITU-R BT.601. – ITU-R, 2007.

- 33 ITU-R Recommendation BT.500-11. Methodology for the subjective assessment of the quality of television pictures. – ITU-R, 2002.
- 34 ITU-T Recommendation P.910. Subjective video quality assessment methods for multimedia applications. – ITU-T, 2008.
- 35 ITU-T Recommendation J.247. Objective perceptual multimedia video quality measurement in the presence of a full reference. – ITU-T, 2008.
- 36 Watson, A. B. Digital video quality metric based on human vision / A. B. Watson, J. Hu and J. F. McGowan // Journal of Electronic imaging. – 2001. – Vol. 10. – №1. – P. 20–29.
- 37 Winkler, S. Digital video quality: vision models and metrics / Stefan Winkler. – John Wiley & Sons Ltd, 2005. – 175 p.
- 38 Shannon, C. E. A Mathematical Theory of Communication / C. E. Shannon // The Bell Sys. Tech. J. – 1948. – Vol. XXVII. – №3. – P. 379–423.
- 39 Huffman, D. A. A Method for the Construction of Minimum Redundancy Codes / D. A. Huffman // Proc. IRE. – 1952. – Vol. 40. – №9. – P. 1098–1101.
- 40 Golomb, S.W. Run-length encoding / S.W. Golomb // IEEE Trans, on Inf. Theory. – 1966. – IT-12. – P. 399–401.
- 41 ISO/IEC 14496-2. Information technology – Coding of audio-visual objects. Part 2: Visual.
- 42 Langdon G. C. Compression of Black-White Images with Arithmetic Coding / G. C. Langdon, J. J. Rissanen // IEEE Trans. Comm. – Vol. COM-29. – №6. – P. 858–867.
- 43 ITU-T Recommendation T.82 (JBIG), Information technology coded representation of picture and audio information progressive bi-level image compression.
- 44 Schwarz H. CABAC and slices / H. Schwarz, D. Marpe, T. Wiegand // JVT document JVT-D020, Klagenfurt, Austria, July 2002.
- 45 Marpe D. Context-Based Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standard / D. Marpe, H. Schwarz, T. Wiegand // IEEE Transactions on Circuits and Systems for Video Technology. – 2003. – Vol. 13. – P. 620–636.
- 46 Прэтт, У. Цифровая обработка изображений / У. Прэтт; пер. с англ. – М. : Мир, 1982. – Кн. 1. – 312 с.
- 47 Ричардсон, Я. Видеокодирование. H.264 и MPEG-4 – стандарты нового поколения / Я. Ричардсон; пер. с англ. – М. : Техносфера, 2005. – 368 с.
- 48 Richardson, I. The H.264 advanced video compression standard / Iain E. Richardson. – 2-nd ed., John Wiley & Sons, Ltd, 2010. – 316 p.
- 49 Mallat, S. A Wavelet Tour of Signal Processing / S. Mallat. – Academic Press, 1999. – 216 p.
- 50 Теория электрической связи : учебник для вузов / под ред. Д. Д. Кловского. – М. : Радио и связь, 1999. – 432 с.
- 51 Gauger, E. Actual standardization activities for coaxial cable networks / E. Gauger // Standardization – Buildings' Cable Infrastructure Evolution. – ANACOM Workshop, Carnexide – Portugal, 29 Oct. 2008.

52 Ткаченко, А. П. Тенденции развития систем цифрового телевизионного вещания / А. П. Ткаченко, М. И. Зорько, Д. А. Хатьков // Международная научно-техническая конференция, приуроченная к 50-летию МРТИ–БГУИР (Минск, 18–19 марта 2014 г.) : материалы конф. В 2 ч. Ч. 1. – Минск : БГУИР, 2014. – С. 261–263.

53 Кривошеев, М. И. О новой концепции развития ТВ-вещания / М. И. Кривошеев // Broadcasting. Телевидение и радиовещание. – 2011. – №2. – С.8–11; №3. – С.33–37.

54 Зубарев, Ю. Тенденции развития видеотехнологий и систем телевидения / Ю. Зубарев, Ю. Сагдуллаев // Broadcasting. Телевидение и радиовещание. – 2011. – №2. – С.22–25; №3. – С. 49–53.

55 ГОСТ Р 53533–2009. Цифровое телевидение высокой четкости. Основные параметры цифровых систем ТВЧ. Общие требования.

56 ГОСТ Р 53535–2009. Цифровое ТВЧ. Измерительные сигналы. Методы измерений. Общие требования.

57 ГОСТ Р 53534–2009. Цифровое телевидение высокой четкости. Аналоговое и цифровое представления сигналов. Цифровые интерфейсы. Технические требования.

58 ГОСТ Р 53536–2009. Цифровое телевидение повышенной четкости. Основные параметры цифровой системы с построчным разложением. Аналоговые и цифровые представления сигналов. Параллельный цифровой интерфейс.

59 ГОСТ Р 53540–2009. Цифровое телевидение. Широкоформатные цифровые системы. Основные параметры. Аналоговые и цифровые представления сигналов. Параллельный цифровой интерфейс.

60 ГОСТ Р 54309–2011. Аудиовизуальная информационная система реального времени (РАВИС). Процессы формирования кадровой структуры, канального кодирования и модуляции для системы цифрового наземного узкополосного радиовещания в ОВЧ-диапазоне. Технические условия.

61 Дворкович, В. Отвечает ли отечественная система РАВИС потребностям локального радиовещания / В. Дворкович, А. Дворкович // Broadcasting. Телевидение и радиовещание. – 2011. – №2. – С. 34–38; № 3. – С. 42–46.

62 Дворкович, А. В. / Способ трансляции информационного телевидения / В. В. Дворкович, В. П. Дворкович, Ю. Б. Зубарев, А. Ю. Соколов, Ю. А. Чернов // Патент РФ №2219676. – 08.11.2000.

63 Оборудование и технологии для IP-вещания: Обзор // 625. Науч.-техн. журнал. – 2011. – №8. – С.26–39.

64 Кривошеев, М. И. Прогресс в изучении стереоскопического и объемного телевизионного вещания в МСЭ-R и инициативы России / М. Кривошеев, В. Федунин // Broadcasting. Телевидение и радиовещание.

65 Уэлстид, С. Фракталы и вейвлеты для сжатия изображений в действии : учеб. пособие / С. Уэлстид. – М. : Издательство Триумф, 2003. – 320 с.

66 Калинин, А. Глобальный подход к будущему наземного ТВ вещания / А. Калинин // 625. Науч.-техн. журнал. – 2012. – №18. – С.42–47.

67 ITU-T Recommendation H.264. Advanced video coding for generic audio-visual services / Telecommunication Standardization Sector of ITU, 2010.

68 Стриго, С. Н. Обоснование новых координат цветности и правила формирования сигнала яркости / С. Н. Стриго, А. П. Ткаченко // Современные средства связи : материалы XVIII МНТК (15–16 октября 2013 г., Минск, Республика Беларусь). – Минск : ВГКС, 2013. – С. 61–63.

69 ITU-T Recommendation H.265. High Efficiency Video Coding / Telecommunication Standardization Sector of ITU, 04.2013.

70 Хоминич, А. Л. HEVC – новый стандарт компрессии изображений / А. Л. Хоминич // Современные средства связи : материалы XVIII МНТК (18–19 октября 2012 г., Минск, Республика Беларусь). – Минск : ВГКС, 2012. – С. 87–88.

71 Sullivan, G. J. Next-Generation High Efficiency Video Coding (HEVC) Standard. Presentation for HPA/ATSC. 15 February 2011. – 15 p.

Библиотека БГУИР

Учебное издание

Ткаченко Анатолий Пантелеевич
Хоминич Александр Леонидович

ЦИФРОВОЕ ТЕЛЕВИДЕНИЕ

В двух частях

Часть 1

КОДИРОВАНИЕ ИСТОЧНИКА СООБЩЕНИЙ В СИСТЕМАХ ЦИФРОВОГО ТЕЛЕВИЗИОННОГО ВЕЩАНИЯ

УЧЕБНО-МЕТОДИЧЕСКОЕ ПОСОБИЕ

Редактор *Г. С. Корбут, Е. С. Чайковская*
Корректор *А. В. Бас*

Компьютерная правка, оригинал-макет *Е. Г. Бабичева*

Подписано в печать 27.04.2015. Формат 60x84 1/16. Бумага офсетная. Гарнитура «Таймс».

Отпечатано на ризографе. Усл. печ. л. 9,65. Уч.-изд. л. 9,5. Тираж 200 экз. Заказ 279.

Издатель и полиграфическое исполнение: учреждение образования
«Белорусский государственный университет информатики и радиоэлектроники».

Свидетельство о государственной регистрации издателя, изготовителя,
распространителя печатных изданий №1/238 от 24.03.2014,

№2/113 от 07.04.2014, №3/615 от 07.04.2014.

ЛП №02330/264 от 14.04.2014.

220013, Минск, П. Бровки, 6