



МИР

цифровой обработки

В.П. Дворкович,
А.В. Дворкович

Цифровые
видеоинформационные
системы
(теория и практика)

ТЕХНОСФЕРА
Москва
2012

УДК 004.421, 004.932, 519.722, 621.317.799, 621.397

ББК 32.81

Д24

Рецензенты: д.т.н., профессор Калмыков В.В.;

д.т.н., профессор Сагдуллаев Ю.С.

Д24 Дворкович В.П., Дворкович А.В.

Цифровые видеoinформационные системы (теория и практика)

Москва: Техносфера, 2012. – 1008с. ISBN 978-5-94836-336-3

Книга посвящена теоретическим основам и практической реализации цифровых видеoinформационных систем. Материал содержит девять частей, состоящих из 23 глав.

В Части I изложены проблемы цифрового преобразования статических и динамических изображений, оцениваются объемы информации при преобразовании изображений в цифровую форму. Часть II посвящена исследованиям проблем статистической и визуальной избыточности изображений, излагаются методы эффективного энтропийного кодирования, оцениваются возможности использования визуальной избыточности статических и динамических изображений. В Части III рассматриваются непосредственное кодирование, кодирование с предсказанием и групповое кодирование изображений, методы блочного и фрактального кодирования, кратномасштабная обработка изображений. В Части IV излагаются методы повышения качества внутрикадрового кодирования статических изображений, принципы анализа движения в динамических изображениях, идеология обработки изображений при кратномасштабном анализе. Часть V посвящена анализу методов и стандартов кодирования статических и динамических изображений, рассмотрены некоторые проблемы реализации кодирующих устройств и визуализации их параметров. В Части VI анализируются основные характеристики и цифровое представление речевой и звуковой информации, приводится описание основных стандартов кодирования речи и звука. Часть VII посвящена анализу основных методов канального кодирования и модуляции цифровой информации. В Части VIII приводится описание основных стандартов цифрового ТВ вещания. Часть IX посвящена изложению стандартов мультимедийного телерадиовещания, а также проблемам реализации систем видеоконференцсвязи на базе использования рекомендаций ITU-T.

Книга адресована студентам и специалистам телекоммуникационной отрасли, работающим в области обработки и передачи видео- и звуковой информации.

УДК 004.421, 004.932, 519.722, 621.317.799, 621.397

ББК 32.81

© 2012, Дворкович В.П., Дворкович А.В.

© 2012, ЗАО «РИЦ «Техносфера»,
оригинал-макет, оформление.

ISBN 978-5-94836-336-3

Содержание

Вступительное слово академика РАН А.С. Бугаева	13
Предисловие.....	16
Часть I. Видеоинформационные приложения и объем цифровой информации	
Введение	30
Глава 1	
Форматы изображений. Статические и динамические эталонные изображения	33
1.1. Аналоговые и цифровые растровые изображения.....	33
1.2. Цветовые системы	39
1.3. Статические и динамические эталонные изображения	47
Глава 2	
Системы формирования и передачи видеоинформации.....	52
2.1. Монохромное телевидение	52
2.2. Цветное телевидение.....	55
2.2.1. Система NTSC.....	57
2.2.2. Система PAL	59
2.2.3. Система SECAM	62
2.2.4. Цифровые компонентные телевизионные сигналы.....	67
2.3. Телевидение повышенной, высокой и сверхвысокой четкости, «цифровое кино»	70
2.4. Видеотелефония и видеоконференцсвязь, домашнее видео	71
2.5. Графические изображения, факсимильные системы и интерактивные системы передачи изображений.....	72
Глава 3	
Цифровое представление видеоинформации.....	75
3.1. Представление изображения конечным объемом данных.....	75
3.2. Фурье-преобразование, теорема о дискретизации	79
3.3. Изображение, изменяющееся во времени	82
3.4. Квантование. Импульсно-кодовая модуляция	84
Литература к части I.....	89

Часть II. Статистическая и визуальная избыточность изображений

Введение 93

Глава 4

Статистическая избыточность

дискретизированных данных 95

4.1. Информационная теорема 95

4.2. Виды статистического кодирования 96

4.2.1. Методы представления целых чисел 97

4.2.2. Алгоритм Шеннона–Фано 100

4.2.3. Алгоритм Хаффмана 101

4.2.4. Блочное и условное кодирование 105

4.2.5. Арифметическое кодирование 107

4.2.6. Словарные методы кодирования дискретной информации 109

4.2.7. Статистические методы моделирования дискретной информации 115

4.2.8. Контекстные методы энтропийного кодирования 127

4.3. Кодирование с преобразованием 136

4.3.1. Кодирование с предсказанием 136

4.3.2. Кодирование с дискретным преобразованием 138

4.4. Статистика монохромных и цветных изображений 142

4.4.1. Статистика монохромных изображений без преобразования 142

4.4.2. Статистика монохромных изображений с преобразованием 145

4.4.3. Статистика цветных изображений 146

4.5. Кодирование с сокращением объема передаваемой информации 147

Глава 5

Визуальная избыточность изображений 150

5.1. Заметность пространственных изменений яркости изображений 150

5.2. Заметность временных изменений яркости изображений 157

5.3. Влияние помех 158

5.4. Специфические особенности восприятия цветов 159

5.5. Методы субъективной оценки качества изображений 162

5.6. Компьютерные методы хранения изображений 166

Литература к части II 178

Часть III. Основные методы сжатия изображений

Введение 185

Глава 6

Непосредственное кодирование изображений и кодирование с предсказанием 188

6.1. Импульсно-кодовая модуляция 188

6.2. Кодирование изображений с предсказанием, дифференциальная импульсно-кодовая модуляция 191

6.3.	Основные виды предсказания.....	197
6.3.1.	Внутрикадровое предсказание.....	197
6.3.2.	Межкадровое предсказание.....	199
6.4.	Квантование сигналов, ошибки предсказания.....	201
6.5.	Помехоустойчивость кодирования с предсказанием.....	206
Глава 7		
Групповое кодирование изображений.....		208
7.1.	Дискретные линейные ортогональные преобразования.....	208
7.1.1.	Дискретное преобразование Фурье.....	210
7.1.2.	Преобразование Хаара.....	213
7.1.3.	Преобразование Уолша–Адамара.....	214
7.1.4.	Дискретное синусное преобразование.....	217
7.1.5.	Дискретные косинусные преобразования.....	217
7.1.6.	Преобразование Кархунена–Лоэва.....	221
7.1.7.	Другие виды преобразований.....	226
7.2.	Квантование коэффициентов преобразования.....	228
7.3.	Кодирование коэффициентов преобразования.....	231
7.4.	Межкадровое или трехмерное кодирование.....	233
Глава 8		
Другие методы кодирования изображений.....		235
8.1.	Кодирование с использованием гибридных преобразований.....	235
8.2.	Другие методы блочного и интерполяционного кодирования.....	236
8.3.	Фрактальные методы кодирования изображений.....	244
Глава 9		
Вейвлеты и кратномасштабная обработка изображений.....		252
9.1.	Непрерывное вейвлет-преобразование.....	253
9.2.	Дискретное вейвлет-преобразование.....	256
9.3.	Кратномасштабный вейвлет-анализ.....	258
9.4.	Основные сведения о субполосном кодировании.....	262
9.4.1.	КИХ-фильтры с нечетным числом отсчетов цифровой решетки и линейной (нулевой) ФЧХ.....	265
9.4.2.	КИХ-фильтры с четным числом отсчетов цифровой решетки....	269
9.4.3.	Квадратурно-зеркальные КИХ-фильтры.....	273
9.4.4.	Кратномасштабное дискретное вейвлет-преобразование изображений.....	274
9.4.5.	Использование банков трехполосных КИХ-фильтров при обработке изображений.....	275
Литература к части III.....		281

Часть IV. Внутрикадровая и межкадровая обработка изображений

Введение	287
Глава 10	
Внутрикадровое кодирование изображений	290
10.1. Изменения коэффициентов квантования блоков в JPEG-подобных алгоритмах видеокompрессии	290
10.2. Направленное пространственное внутрикадровое предсказание	295
10.2.1. Режим предсказания пикселей яркости блоков 4×4	297
10.2.2. Режим предсказания пикселей яркости блоков 8×8	301
10.2.3. Режим предсказания пикселей яркости блоков 16×16	304
10.2.4. Режим предсказания пикселей блоков цветности	305
10.3. Уменьшение влияния блокинг-эффекта	308
Глава 11	
Обработка изображений при кратномасштабном анализе	314
11.1. Оценка энергетических характеристик изображения в зависимости от вида вейвлет-базиса разложения	315
11.2. Выбор маски квантования высокочастотных компонент вейвлет-преобразования изображений	317
11.3. Обработка тестовых изображений двухполосными вейвлет-фильтрами	322
11.4. Обработка тестовых изображений трехполосными вейвлет-фильтрами	323
Глава 12	
Методы анализа и компенсации движения в динамических изображениях	329
12.1. Основные методы анализа движения	331
12.1.1. Метод полного перебора	332
12.1.2. Логарифмический, комбинированный по двум направлениям, трехшаговый, иерархический методы	332
12.1.3. Методы, основанные на оптическом уравнении	335
12.1.4. Использование временной и пространственной корреляции для поиска векторов движения	336
12.1.5. Использование временной и пространственной корреляции векторов с оптимизацией возмущения битового потока	337
12.1.6. Метод полного перебора с преобразованием Фурье	338
12.1.7. Методы фазовой корреляции	339
12.1.8. Оптимизация поиска векторов движения по стандарту MPEG-4 (метод быстрого поиска MVFAST)	340
12.2. Повышение эффективности анализа движения по опорным точкам ..	344
12.3. Дополнительные возможности компенсации движения деталей в динамических изображениях	355
12.4. Анализ смены сюжета в динамических изображениях	361
Литература к части IV	366

Часть V. Основные стандарты цифрового кодирования видеоинформации

Введение	369
Глава 13	
Стандарты кодирования статических изображений	371
13.1. Стандарт JPEG и его модификация	372
13.2. Стандарт JPEG-2000.....	377
Глава 14	
Стандарты кодирования динамических изображений	387
14.1. H.261 — первый стандарт видеокодирования МСЭ	390
14.2. H.263. Видеокодирование для систем связи с низкой пропускной способностью	395
14.3. MPEG-1 — кодирование видео- и аудиоинформации для цифровых систем с объемом данных приблизительно до 1,5 Мбитов/с	402
14.4. MPEG-2. Кодирование динамических изображений и связанной с ними звуковой информации	406
14.4.1. Общие положения	407
14.4.2. Совместимость стандартов MPEG-1 и MPEG-2	410
14.4.3. Масштабируемость в MPEG-2	411
14.4.4. Профили и уровни	414
14.4.5. Транспортный и программный потоки MPEG-2.....	415
14.4.6. Видеопоток стандарта MPEG-2	421
14.4.7. Структура битового потока видеопоследовательности стандарта MPEG-2	429
14.5. MPEG-4. Обобщенное кодирование аудиовизуальных объектов	454
14.5.1. Кодированное представление медийных объектов	455
14.5.2. Аудиообъекты	457
14.5.3. Видеообъекты	458
14.6. H.264/MPEG-4 часть 10 AVC — улучшенное видеокодирование	467
14.6.1. Профили и уровни стандарта H.264/AVC	468
14.6.2. Эффективность сжатия в стандарте H.264/AVC	471
14.6.3. Уровень сетевой абстракции	476
14.6.4. Уровень видеокодирования	478
14.6.5. Преобразование, масштабирование и квантование	490
14.6.6. Деблокирующая фильтрация	497
14.6.7. Энтропийное кодирование	498
14.7. Некоторые проблемы реализации видеокодирующих систем	503
14.7.1. Методика построения промежуточных кадров видеопоследовательности	503
14.7.2. Методика регулировки потока при сжатии динамических изображений	506
14.7.3. Анализ смены сюжета	516
14.7.4. Оценка быстродействия кодирования изображений	519

14.7.5. Возможности кодирования стереоскопических изображений в рамках совместимости со стандартами MPEG	521
14.8. Визуализация параметров видеокодирования	528
Литература к части V	537

Часть VI. Основные методы кодирования речевой и звуковой информации

Введение	543
-----------------------	-----

Глава 15

Основные характеристики и цифровое представление звуковой информации	547
---------------------------------------------------------------------------------------	-----

15.1. Статистика мгновенных значений и уровней звуковых сигналов.....	548
15.2. Спектральные характеристики звуковых сигналов	553
15.3. Огибающая и мгновенная частота звуковых сигналов.....	558
15.4. Пространственное восприятие звуковых сигналов	561
15.5. Основные принципы цифровых преобразований звуковых сигналов	564
15.6. Аналого-цифровое и цифро-аналоговое преобразование звуковых сигналов.....	568
15.6.1. Импульсно-кодовая модуляция (ИКМ)	568
15.6.2. Дифференциальная импульсно-кодовая модуляция (ДИКМ).....	569
15.6.3. Сигма-дельта-модуляция (СДМ).....	571
15.7. Принципы кодирования речевой и звуковой информации.....	574
15.7.1. Принципы кодирования речевой информации.....	574
15.7.2. Принципы кодирования звуковой информации	577
15.8. Предыскажения звуковых сигналов при цифровом кодировании	591

Глава 16

Основные стандарты кодирования речевой информации	592
----------------------------------------------------------	-----

16.1. G.711 — стандарт кодирования речи с применением ИКМ.....	593
16.2. G.722 — кодирование речи со спектром до 7 кГц в цифровой поток 64 кбитов/с	593
16.3. G.723.1 — речевой кодек с двумя скоростями потока — 5,3 и 6,3 кбитов/с	596
16.4. G.726 — речевой кодек АДИКМ со скоростями 40, 32, 24, 16 кбитов/с.....	602
16.5. G.727 — речевой кодек с алгоритмом вложенной АДИКМ.....	603
16.6. G.728 — речевой кодек с выходным потоком 16 кбитов/с и малой задержкой.....	605
16.7. G.729 — речевой кодек со скоростью 8 кбитов/с	607
16.8. GSM — речевой кодек для систем сотовой связи	611
16.9. GIPS iLBC — кодек, предназначенный для передачи речи по IP-сетям.....	616
16.10. Speex — свободный кодек, предназначенный для передачи речи по IP-сетям.....	621

Глава 17**Стандарты кодирования звуковой информации** 624

- 17.1. MPEG-1 Audio — кодирование аудиоинформации для цифровых систем с потоком до 1,5 Мбит/с..... 625
 - 17.1.1. MPEG-1 Audio Layer-1 — профессиональная система кодирования аудиоинформации с высоким качеством..... 626
 - 17.1.2. MPEG-1 Audio Layer-2 — система кодирования аудиоинформации для высококачественного телерадиовещания 632
 - 17.1.3. MPEG-1 Audio Layer-3 — система кодирования аудиоинформации для Интернет-вещания..... 634
- 17.2. MPEG-2 Audio — стандарт кодирования аудиоинформации для многоканального стереофонического вещания..... 640
- 17.3. MPEG-2 AAC — стандарт усовершенствованного аудиокодирования. 643
- 17.4. MPEG-4 Audio — стандарт аудиокодера для мультимедийных приложений..... 645
- 17.5. Dolby AC3 — стандарт аудиокодера для телевидения и других применений 647
- 17.6. Совершенствование алгоритмов сжатия звука на основе стандарта AAC 651
 - 17.6.1. Общие алгоритмы кодека AAC..... 654
 - 17.6.2. Метод спектрального дублирования SBR..... 658
 - 17.6.3. Параметрическое стереокодирование PS 663
 - 17.6.4. Конфигурации режимов работы кодеков AAC и HE-AAC 664

Глава 18**Подавление акустического эха в системах****телекоммуникационной связи**..... 667

- 18.1. Основные положения 667
- 18.2. Алгоритмы подавления акустического эха 670
- 18.3. Принципы линейной адаптивной фильтрации..... 674
- 18.4. Детектор двойного разговора 676
- 18.5. Подавление акустического эха на базе фильтра Калмана 679
- 18.6. Особые случаи 681
- 18.7. Метрологические требования к подавлению акустического эха..... 682

Литература к части VI..... 684**Часть VII. Основные методы и системы канального кодирования и модуляции цифровой информации****Введение** 688**Глава 19****Помехоустойчивое кодирование в системах передачи цифровой информации**..... 691

- 19.1. Принципы кодирования источника сообщений..... 691
 - 19.1.1. Теоретические предпосылки помехоустойчивого кодирования 692

19.1.2.	Классификация помехоустойчивых кодов	696
19.2.	Линейные блочные коды	698
19.2.1.	Математическое описание линейных блочных кодов	698
19.2.2.	Свойства линейных блочных кодов	701
19.2.3.	Практические примеры блочных кодов	704
19.3.	Циклические коды	708
19.3.1.	Понятие циклического кода	708
19.3.2.	Порождающая и проверочная матрицы циклического кода	710
19.3.3.	Систематическое кодирование циклическими кодами	712
19.3.4.	Обнаружение ошибок	714
19.3.5.	Исправление однократных ошибок	714
19.3.6.	Общий принцип декодирования циклических кодов	715
19.3.7.	Арифметика полей Галуа	716
19.3.8.	Циклические коды Хэмминга	722
19.3.9.	CRC-коды	723
19.3.10.	Двоичные BCH-коды	724
19.3.11.	Коды Рида–Соломона	726
19.4.	Сверточные коды	728
19.4.1.	Систематические сверточные коды	730
19.4.2.	Управление скоростью кодирования	732
19.4.3.	Представление сверточных кодеров в виде цифровых фильтров	733
19.4.4.	Диаграммы состояний и кодовые решетки	734
19.4.5.	Декодирование сверточных кодов	736
19.5.	Турбокодирование	741
19.5.1.	Основные положения	741
19.5.2.	Обобщенная схема турбокодера ТСС с параллельным каскадированием	742
19.5.3.	Сверточные турбокоды	742
19.5.4.	Декодирование турбокодов	744
19.5.5.	Характеристики помехоустойчивости сверточных турбокодов ТСС	749
19.5.6.	Блочное турбокодирование	751
19.6.	Низкоплотностные коды	753
19.6.1.	Основные положения	753
19.6.2.	Классификация LDPC-кодов	754
19.6.3.	Методы построения проверочных матриц	755
19.6.4.	Алгоритмы декодирования низкоплотностных кодов	761
19.6.5.	Оценка сложности алгоритмов декодирования	764

Глава 20

	Системы модуляции и сигнального кодирования цифровой информации	767
20.1.	Созвездия дискретной модуляции	768
20.1.1.	Одномерные и двумерные созвездия	768
20.1.2.	Сигнальные созвездия квадратурно-амплитудной модуляции	769
20.1.3.	Фильтрация сигналов в передающих и приемных устройствах	776



20.2. Система многочастотной модуляции	778
20.2.1. Структурные схемы OFDM-модуляторов и демодуляторов.....	779
20.2.2. Защитный интервал, устранение интерференции между несущими частотам.....	782
20.2.3. Иерархические режимы OFDM-модуляции	784
20.3. Алгоритмы синхронизации OFDM-сигналов в приемных устройствах	786
20.3.1. Методы демодуляции OFDM-сигнала.....	786
20.3.2. Оценка частотного смещения при приеме OFDM-сигнала	789
Литература к части VII	791

Часть VIII. Основные стандарты систем цифрового телевизионного вещания

Введение	794
-----------------------	------------

Глава 21

Стандарты цифрового телевизионного вещания	797
---------------------------------------------------------	------------

21.1. Стандарт цифрового телевидения ATSC.....	799
21.2. Система цифрового наземного телевизионного вещания DVB-T.....	808
21.3. Усовершенствованная система цифрового наземного ТВ-вещания DVB-T2	826
21.4. Система цифрового спутникового телевизионного вещания DVB-S...	867
21.5. Система высокоскоростного цифрового спутникового ТВ-вещания DVB-S2.....	870
21.6. Система цифрового кабельного телевизионного вещания DVB-C.....	881
21.7. Система высокоскоростного цифрового кабельного ТВ-вещания DVB-C2	883
21.8. Система наземного цифрового телевизионного вещания ISDB-T	893
21.9. Система цифрового спутникового многопрограммного ТВ-вещания ISDB-S.....	897
21.10. Система цифрового кабельного многопрограммного ТВ-вещания ISDB-C	902
21.11. Система цифрового наземного телевизионного вещания DTMB	905
21.12. Эффективность использования систем цифрового ТВ-вещания в России.....	909

Литература к части VIII.....	918
-------------------------------------	------------

Часть IX. Стандарты систем мобильного радиовещания и видеоконференцсвязи

Введение	921
-----------------------	------------

Глава 22

Стандарты цифрового мобильного радиовещания.....	930
---------------------------------------------------------	------------

22.1. Система мобильного телевизионного вещания DVB-H	930
22.2. Система мобильного телевизионного вещания MediaFLO	937
22.3. Система мобильного мультимедийного вещания T-DMB	940

22.4. Аудиовизуальная информационная система реального времени РАВИС ...	943
22.5. Цифровое радиовещание DRM+ в ОВЧ-диапазоне частот.....	963
Глава 23	
Системы видеоконференцсвязи	969
23.1. Основные стандарты систем видеоконференцсвязи.....	969
23.2. Возможности технологии видеоконференцсвязи.....	985
Литература к части IX.....	999

Вступительное слово академика РАН А.С. Бугаева

Внедрение информационных технологий в начале XXI века в основном определяется научными и практическими результатами развития техники цифровой обработки сигналов, обеспечивающей решение задач обработки, сокращения избыточности сигнала и передачи и приема информации в реальном масштабе времени.

Целью данной книги является подробное изложение современных достижений теории и практики видеоинформационных систем, реализующих цифровое кодирование и передачу динамических изображений, речи, звука, иных данных по каналам с различной пропускной способностью (видеотелефон, стационарная и мобильная видеоконференцсвязь, многофункциональные интерактивные телевизионные системы, радиолокационные системы и пр.).

Авторы книги — доктор технических наук, профессор Дворкович Виктор Павлович и его сын, доктор технических наук, профессор Дворкович Александр Викторович — хорошо известны специалистам большим вкладом в разработку систем и средств цифровой и компьютерной обработки изображений, а также метрологического обеспечения аналогового и цифрового телевизионного вещания.

С 2000 г. деятельность авторов связана с созданием принципиально новой системы мультимедийного вещания, получившей название РАВИС (RAVIS — Realtime AudioVisual Information System). Результаты разработки системы РАВИС отражены во многих научных статьях, в 18 вкладах России в Международные организации ИТУ и СЕРТ. В ходе разработок высокоэффективных систем кодирования изображений было сделано более 10 изобретений, защищенных патентами. Наряду с научной работой Дворковичи ведут активную научно-преподавательскую деятельность (Виктор Павлович — профессор МГТУ им. Н.Э. Баумана, Александр Викторович — профессор МЭИ), они авторы и соавторы ряда учебных пособий.

Авторы книги определяют в качестве важнейшей проблемы создания в России современных эффективных систем цифровой обработки, кодирования, передачи, приема изображений различного разрешения, звука и других видов информации полное овладение теорией и математическими методами обработки этих видов информации при использовании имеющихся отечественных и зарубежных материалов в качестве данных, задающих общее направление работ.

Внедрение видеоинформационных систем и их практическое использование послужит началом реализации планов обучения высококлассных специалистов данного профиля. В области организации науки и подготовки кадров важнейшим является использование научно-исследовательских центров по видеоинформационным проблемам, обладающих необходимым научным и техническим потенциалом для решения как фундаментальных проблем обработки сигналов, так и широкого круга прикладных задач.

Предлагаемый авторами материал книги дает возможность понять, с какими проблемами сталкиваются специалисты — разработчики цифровых систем обработки и передачи аудиовизуальной информации.

Следует отметить главное, что подтверждает актуальность издания этого материала, — важность мировой практики перехода к цифровым технологиям

в телекоммуникационных средствах передачи сигналов по каналам связи, наземного и спутникового вещания.

Книга содержит девять разделов, изложенных в 23 главах.

В первых двух разделах последовательно и подробно рассмотрены вопросы цифрового преобразования статических и динамических изображений, оцениваются объемы информации при преобразовании в цифровую форму изображений различного разрешения, приводятся результаты исследований статистической и визуальной избыточности изображений, излагаются методы и алгоритмы эффективного энтропийного кодирования различных видов информации, оцениваются возможности использования визуальной избыточности изображений для сокращения объемов передаваемой информации.

Следующие четыре раздела посвящены изложению методов кодирования изображений, включая их кратномасштабную обработку. Особое внимание уделяется оптимизации методов поиска векторов движения, дополнительным возможностям компенсации движения деталей в динамических изображениях, подробно изложены методы кодирования статических изображений JPEG и JPEG-2000, систем кодирования динамических изображений MPEG-1, MPEG-2, MPEG-4, а также наиболее эффективного стандарта кодирования — H.264/AVC, анализируются основные характеристики и основные стандарты цифрового кодирования речи, алгоритмы кодирования высококачественного звука, приведенные в стандартах MPEG-1, MPEG-2, MPEG-4, Dolby AC-3. Рассмотрены также проблемы подавления акустического эха в системах телеконференцсвязи.

Следующие три раздела посвящены анализу основных методов модуляции и помехоустойчивого кодирования цифровой информации, приводится достаточно подробное описание систем многочастотной OFDM-модуляции, основных современных стандартов вещания видео- и аудиоинформации. Основное внимание уделено изложению европейских стандартов цифрового телевидения DVB-T/DVB-T2, DVB-S/DVB-S2, DVB-C/DVB-C2, а также отечественной мультимедийной системе РАВИС и проблемам реализации систем видеоконференцсвязи.

Следует подчеркнуть, что отличительной особенностью данной книги является то, что она написана не сторонними исследователями приведенного объемного материала, а непосредственными разработчиками математических методов и аппаратных реализаций соответствующих систем.

Естественно, что участие в решении этих сложных задач и накопленный при этом опыт не могли не вызвать интереса к их анализу и систематизации, поскольку очевидно большое практическое значение таких работ. Это особенно важно в настоящее время — время неотвратимости слияния телекоммуникационных и компьютерных технологий.

Стремление авторов помочь читателям разобраться в весьма сложных проблемах цифровой обработки информации, значимости и эффективности стандартов кодирования и передачи аудиовизуальной информации является особенностью данной книги.

В этом фундаментальном труде авторов удивительно сочетается множество факторов, придающих ему исключительную значимость для дальнейших исследований и практического использования.

Книга, безусловно, будет способствовать активному внедрению цифровых телекоммуникационных систем и, с учетом широкого использования при этом ком-

пьютерных технологий, заинтересует не только специалистов в областях кодирования и передачи аудиовизуальной информации, но и в других областях знаний.

Поскольку в книге весьма полно отражены проблемы, обсуждению которых она посвящена, анализируется обширная литература по данным вопросам, предлагаются оригинальные решения, способствующие развитию науки и достижению эффективных результатов развития новых перспективных технологий, она может быть отнесена к категории монографий.



*Председатель экспертного совета
по электронике, измерительной технике,
радиотехнике и связи, академик РАН
А.С. Бугаев*

Предисловие

Технической основой создаваемой в России информационной среды становятся современные видеоинформационные системы. Они обеспечивают цифровую передачу динамических изображений, речи, звука, иных данных по каналам с различной пропускной способностью (видеотелефон, стационарная и мобильная видеоконференцсвязь, интерактивные телевизионные системы и пр.).

Состояние и перспективы развития информационных технологий в начале XXI века характеризуются становлением и широким практическим использованием техники цифровой обработки сигналов — одной из самых динамичных и быстро развивающихся технологий в мире телекоммуникаций и информатизации общества. Цифровая обработка сигналов (ЦОС) — это информатика реального времени, призванная решать задачи приема, обработки, сокращения избыточности сигнала и передачи информации в реальном времени.

Внедрение видеоинформационных систем различного назначения решает проблемы:

- создания высококачественных систем интерактивного цифрового телевизионного вещания при удовлетворении постоянно возрастающих запросов на частотные присвоения систем связи без пересмотра частотных планов;
- разработки и внедрения принципиально новых систем мобильного телерадиовещания;
- создания принципиально новых интерактивных систем опроса общественного мнения;
- обеспечения деятельности органов государственной власти;
- создания мобильной видеоконференцсвязи между центральными учреждениями с удаленными центрами и районами, а также удаленных районов между собой;
- обеспечения сбора и распространения информации различного экономического и политического характера, распространяемой органами власти среди населения, популяризации проводимых властями программ, акций, мероприятий;
- реализации оперативного контроля объектов и дистанционного управления по устранению аварий и чрезвычайных ситуаций;
- оптимизации лечебной и профилактической деятельности, создания систем мобильной телемедицины;
- создания систем дистанционного обучения на базе ведущих вузов, расширения системы подготовки абитуриентов и пр.;
- поддержки малого бизнеса, проведения рекламных мероприятий и др.

Проводимые во всех технически развитых странах разработки алгоритмов и аппаратуры сокращения объема, рационального пакетирования и передачи по каналам связи с различной пропускной способностью видео-, аудио- и сопутствующей информации являются основой эффективного использования телекоммуникационных систем и радиочастотного спектра, сохранения действующих частотных планов, высвобождения значительной части частотного пространства для передачи потребителям дополнительных видов услуг — видеотелефонии,

мобильной и стационарной видеоконференцсвязи, многопрограммного интерактивного телевидения, телевидения повышенной и высокой четкости, технологий трехмерного ТВ (3D-TV), телевидения ультравысокой четкости (ТУВЧ), многопрограммного звукового вещания, а также систем ТВ со многими (в будущем — с произвольным числом) точками наблюдения.

Внедрение видеоинформационных систем в России связано с решением ряда проблем:

- эффективного использования мирового опыта по созданию и применению систем цифровой обработки и передачи информации и разработке соответствующих стандартов России;
- разработки и внедрения собственных мультимедийных систем и соответствующей аппаратуры;
- подготовки специалистов для обеспечения разработок, производства и эффективной эксплуатации таких систем;
- создания средств метрологического обеспечения цифровых видеоинформационных систем;
- интегрирования в общемировую систему телекоммуникаций с учетом тенденций перехода к системам многопрограммного телевидения и телевидения высокой четкости в глобальном масштабе, передачи информации по наземным, спутниковым, кабельным и Интернет-сетям.

Размер кадров изображений, их частота и формат существенно изменяются в зависимости от вида приложений. Величина компрессии цифрового потока составляет от сотен (для систем с одной точкой наблюдения) до десятков тысяч раз (для многомерных систем).

В настоящее время наряду с улучшением и созданием новых систем сжатия традиционных видеосигналов, снимаемых с одной точки наблюдения, бурно развивается как стереоскопическое телевидение (бинокулярное), так и автостереоскопическое (или многокурсное, с несколькими точками наблюдения — до 9–16 видов). Также создано и развивается такое направление, как телевизионные системы с произвольной точкой наблюдения (ТСПН или Free Viewpoint TV — FTV).

На рис. П.1 поясняются основные тенденции развития современных видеоинформационных систем. Наряду с ростом числа пикселей в системах ТВ с одной точкой наблюдения, выражающемся в переходе от систем ТВ к ТВЧ и затем к ТУВЧ, наблюдается рост числа точек наблюдения — два вида для стереоскопических систем, до 16 видов для систем со многими видами и практически бесконечное число видов для систем ТСПН (FTV).

Кодирование видеоинформации можно охарактеризовать тремя факторами:

- оптимизированные методы сжатия, обладающие высокой эффективностью, но связанные с появлением артефактов на определенных типах изображений; при этом тратятся существенные усилия на улучшение кодирования с целью уменьшения их визуальной зависимости;
- вычислительная сложность кодирования зачастую становится важной проблемой, в том числе и из-за роста мобильных маломощных приложений, в то время как в прошлом это считалось не столь существенным фактором и приводило лишь к необходимости учета ограничений на сложность алгоритма обработки;

- активно рассматриваются новые типы контента (трехмерное телевидение, мультимедийное видео и т. д.), при этом технологии сжатия развиваются так, что особое внимание уделяется общей эффективности кодирования, но зачастую оказывается неясным, как пользователи смогут получить доступ к этим наборам данных и насколько хорошо сжатые форматы подходят к типичным сценариям их использования.

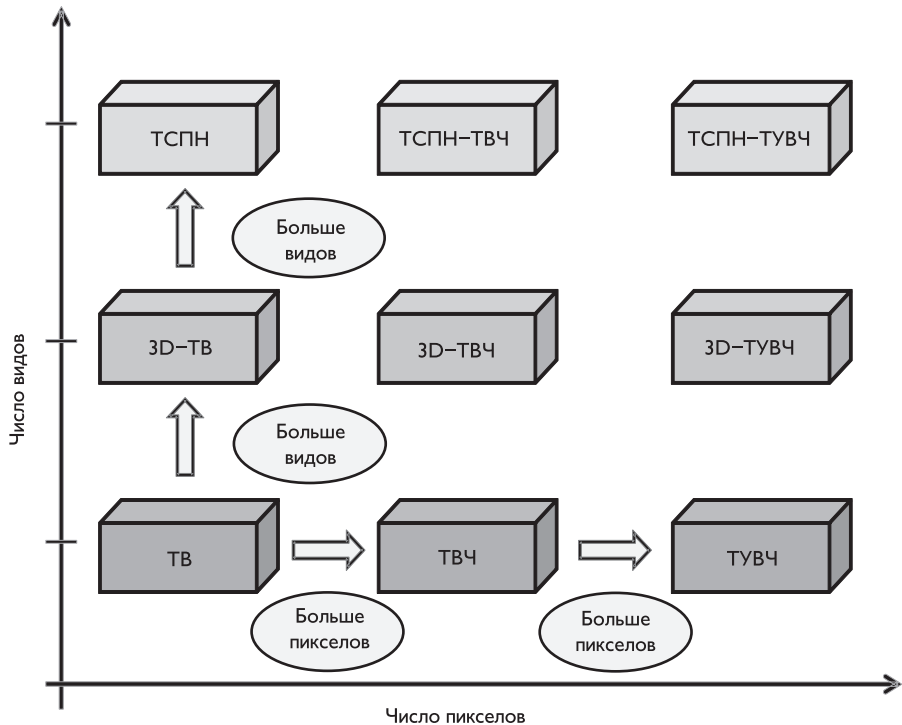


Рис. П.1. Тенденции развития современных видеоинформационных систем

Помимо устранения пространственной избыточности, эффективное временное предсказание движения деталей изображения и его компенсация всегда было ключевым фактором, определяющим общий коэффициент сжатия при кодировании видеоинформации.

Практически реализованные системы кодирования используют блочную оценку движения и его компенсацию, сопровождаемую блочным преобразованием изображения, квантованием и энтропийным кодированием. Даже после замены дискретного косинусного преобразования (ДКП) на его целочисленный вариант блочные технологии кодирования видеоинформации остаются подобными первоначально предложенным. Хотя очевидно, что «истинное» движение в кадре не является ни поступательным, ни блочно-постоянным, блочная природа этих алгоритмов весьма полезна в вычислительном отношении. Последние разработ-

ки существенно не изменили способов оценки движения. Расширение коснулось изменения размеров блока, повышения пиксельной точности и новых режимов предсказания при внутрикадровом кодировании.

Значительные успехи были достигнуты в создании метрик качества, которые лучше оценивают субъективно воспринимаемое (перцепционное) качество. Однако в большинстве случаев эти новые метрики были разработаны при кодировании статических изображений. Сопоставимые результаты для динамических изображений не так развиты. Развитие перцепционно ориентированных инструментов, которые включали бы временные качественные критерии, является ключевой задачей по улучшению рабочих характеристик систем кодирования видео. Заметим, что эти виды инструментов достаточно хорошо развиты для кодирования аудио.

Теоретически кодер будет более эффективным при применении соответствующих методов оптимизации, реализация которых приводит к использованию множества способов выбора режимов его работы в зависимости от структуры кодируемых динамических изображений.

В связи с резким увеличением числа возможных комбинаций режимов и его влиянием на вычислительную сложность таких кодеков анализируются различные возможности построения пространства решений, исследуются алгоритмы и синтаксис кодирования, которые понизят общее количество допустимых режимов (мод), ограничат выбор режимов блоков предсказания и компенсации движения деталей в нескольких кадрах динамических изображений.

В системах с произвольной точкой наблюдения пользователи практически всегда имеют неполный доступ к набору данных, что приводит к созданию кодеров, позволяющих декодерам обладать некоторой гибкостью. Предположим, сигналы от многих камер совместно кодировались, используя инструменты, развитые в контексте многоакурсного кодирования (MVC — Multiview Coding). Тогда пользователи могут изменять угол наблюдения экрана, выбирая только один из видов в потоке битов. Такие наборы данных должны позволять более комплексный доступ к видеoinформации, разрешая пользователям выбор точки наблюдения и управления произвольным образом. При этом декодеру нужно позволить использовать различные пути декодирования, каждый с различным порядком просмотра данных, таким образом соответствуя различным определениям того, что является будущим, а что — прошлым с точки зрения цели декодирования.

Основными организациями по стандартизации кодирования видео являются:

- ITU-T — Группа экспертов кодирования видео (Video Coding Experts Group — VCEG) в Международном союзе электросвязи — Телекоммуникационный сектор стандартизации (International Telecommunications Union — Telecommunications Standardization Sector — ITU-T, организация ООН, ранее МККТТ — ССИТТ), Исследовательская группа 16, Вопрос 6 (Study Group 16, Question 6);
- ISO/IEC — Группа экспертов подвижных изображений (Moving Picture Experts Group — MPEG) в Международной организации по стандартизации и Международной электротехнической комиссии, Объединенный технический комитет 1, Подкомиссия 29, Рабочая группа 11 (International Standardization Organization and International Electrotechnical Commission, Joint Technical Committee Number 1, Subcommittee 29, Working Group 11).

Движущей силой создания стандартов кодирования видеoinформации являются различные приложения и развитие их аппаратного обеспечения. Блочное гибридное кодирование является, по сути, ядром всех стандартов сжатия видео. Среди прошлых стандартов можно перечислить H.261, MPEG-1 часть 2, MPEG-2 часть 2/H.262, H.263 и MPEG-4 часть 2 (см. рис. П.2). Наиболее эффективным является стандарт MPEG 4 часть 10 AVC/H.264. К этому стандарту были приняты поправки по кодированию масштабируемого видео (SVC — Scalable Video Coding), кодированию многокадрового видео (MVC — Multiview Video Coding) и реконфигурируемого кодирования видео.

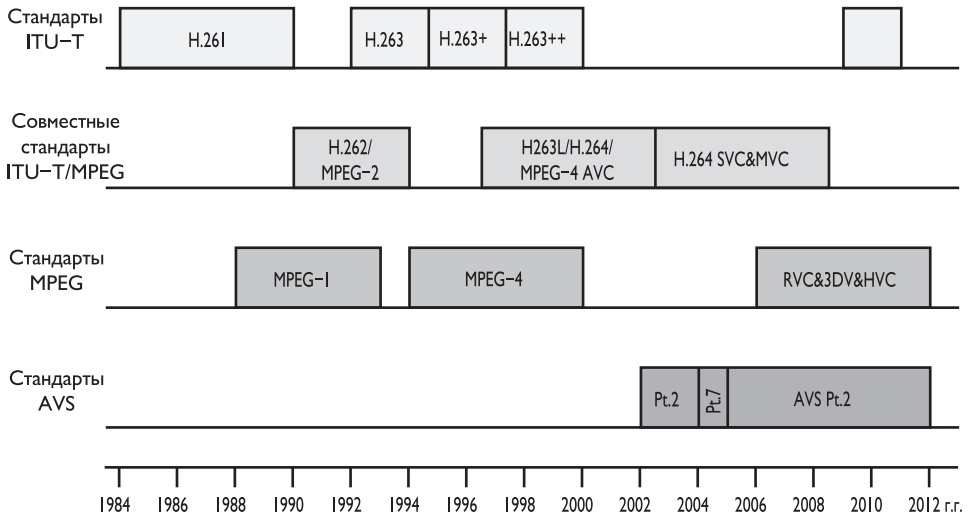


Рис. П.2. История развития методов кодирования видео

Гибридное кодирование комбинирует два метода: движение от кадра к кадру оценивается и компенсируется с помощью предсказания, основанного на ранее закодированных кадрах; остаточная разность после предсказания кодируется, данные декоррелируются в пространственной области посредством преобразования в 2-D-частотную (или вейвлет) область. Преобразованные данные квантуются, после чего данные энтропийно кодируются с помощью метода Хаффмана, арифметического кодера или иного метода.

Стандартный ряд гибридного кодирования H.26x ($x = 1 \dots 4$) рекомендован группой VCEG (Video Coding Experts Group) ITU-T. MPEG-x ($x = 1, 2, 4$) разрабатывается группой MPEG (Moving Picture Experts Group) и стандартизируется ISO/IEC.

Группа VCEG ответственна также за работу ITU-T над стандартами кодирования неподвижных изображений, включая JPEG (ITU-T T.80, T.81, T.83, T.84, и T.86), JBIG-1 (ITU-T T.80, T.82 и T.85), JBIG-2 (ITU-T T.88 и T.89), LS JPEG (ITU-T T.87 и T.870), JPEG 2000 (ITU-T T.800 — T.812), подобный JPEG ITU-T T.851, и JPEG XR (ITU-T T.832, T.834 и T.835). VCEG работает над большинством этих стандартов совместно с ISO/IEC JTC 1/SC 29/WG 1 (Совместная экспертная группа по фотографии/Объединенная группа экспертов по бинарным изображениям — Joint Photographic Experts Group/Joint Bi-level Image experts

Group). Исследовательская группа ответственна за исследования, касающиеся возможностей мультимедийной службы и возможностей приложений (включая поддержку сетей будущего поколения — NGN). Это охватывает мультимедийные терминалы, системы (сетевое оборудование обработки сигналов, многоточечные модули конференций, маршрутизаторы, гейткиперы, модемы, факсы), протоколы и обработку сигналов (кодирование источника).

Несмотря на широкое внедрение цифровых видеоинформационных систем, процесс разработки более эффективных алгоритмов сжатия визуальной и звуковой информации и их передачи по каналам с различной пропускной способностью продолжается. Поскольку имеющиеся стандарты определяют лишь цифровой поток кодируемых данных и не регламентируют принципы его формирования, имеется возможность более эффективной обработки сигналов в рамках этих стандартов.

В 2004 г. организация ITU-T VCEG начала изучение технологических решений, которые могли бы позволить создать новый стандарт сжатия видео, существенно превосходящий H.264/AVC, который получил условное название H.265.

В 2007 г. ISO/IEC (MPEG) запустила проект с экспериментальным названием HVC.

К июлю 2009 г. результаты экспериментов показали среднее понижение потока битов приблизительно на 20% по сравнению с AVC (Высокий профиль); эти результаты побудили MPEG инициировать сотрудничество с VCEG. Заявка на создание технологии сжатия видео была подготовлена VCEG и MPEG в январе 2010 г. Представленные предложения были рассмотрены на первой встрече Объединенной команды по кодированию видео (Joint Collaborative Team on Video Coding — JCT-VC) MPEG и VCEG, которая прошла в апреле 2010 г. JCT-VC была создана как группа экспертов по видеокодированию от Исследовательской группы ITU-T 16 (VCEG) и ISO/IEC JTC 1/SC 29/WG 11 (MPEG) с целью разработки стандарта видеокодирования нового поколения. Оценки показали, что некоторые из представленных 27 предложений могли обеспечить такое же визуальное качество, как у AVC, при уменьшении в два раза потока битов на многих тестовых последовательностях за счет увеличения вычислительной сложности в 2–10 раз. Некоторые предложения обеспечивали хорошее субъективное качество и значения потока битов с более низкой вычислительной сложностью, чем у AVC. Для совместного проекта было принято название HEVC (High Efficiency Video Coding — высокоэффективное видеокодирование). В настоящее время JCT-VC работает над объединением основных технологий лучших предложений в одно программное обеспечение. График работ нацелен на получение проекта конечного стандарта для HEVC приблизительно к январю 2013 г.

На рис. П.3 приведены зависимости отношения сигнал/шум MPEG-кодеров от реализуемых ими цифровых потоков. Из этих зависимостей следует, что стандарт H.264/AVC по эффективности существенно превосходит ранее разработанные кодеры стандартов MPEG-1, MPEG-2 и MPEG-4.

Мультивидовое (или многоакурсное) видео (MVC — Multiview Video Coding) обеспечивает обширную информацию о сцене и расширяет зрительские впечатления по сравнению с традиционным видео. Благодаря усовершенствованиям технологий захвата и отображения мультивидовое видео станет востребованным в потребительской сфере, включая 3DTV и телевидение с произвольной точкой

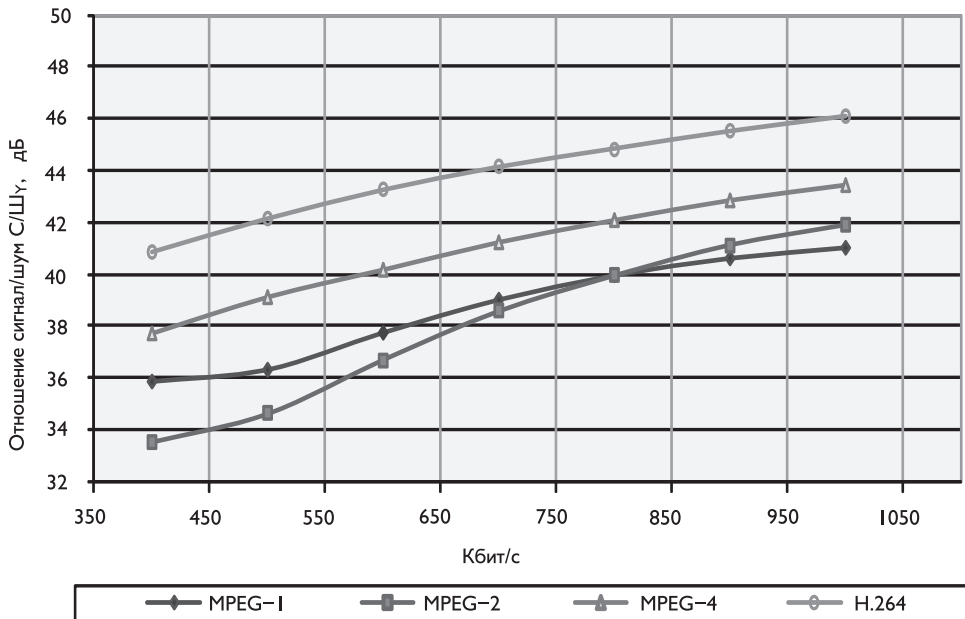


Рис. П.3. Сравнение характеристик MPEG-кодеров

наблюдения (ТСПН — FTV). Среди реализаций 3D-приложений MVC является одной из наиболее многообещающих технологий, поскольку огромный объем данных, пропорциональный числу камер, должен быть сжат до такой степени, чтобы его можно было передавать в пределах возможностей сетей связи.

Поправки MVC к структуре кодера H.264/AVC допускают широкий диапазон совместно применяемых структур временного и межвидового предсказания, что делает возможным достижение компромисса между эффективностью кодирования и сложностью декодирования, включая управление задержкой и буфером декодированного изображения. Все возможные структуры совместно применяемого временного и межвидового предсказания в основном относятся к особым случаям, определяемым поправкой масштабируемого кодирования видео (SVC).

Тенденции развития алгоритмов кодирования и полученные степени сжатия в их историческом развитии иллюстрируются рис. П.4, из которого следует, что стандарт H.264/AVC при высоком качестве кодирования обеспечивает формирование информационного потока при величине 0,17 бита на пиксел изображения.

Это означает, что при кодировании стандартного ТВ-сигнала (8×3 бит/пиксел, 720×576 пикселов/кадр, 25 кадров/с) удастся сформировать потоковое отображение менее 2 Мбит/с, а при кодировании сигнала ТВЧ — порядка 5 Мбит/с.

Предполагается, что новый стандарт H.265 обеспечит до 20% улучшения эффективных критериев и, кроме того, дополнительное уменьшение потока будет связано с изменением субъективной оценки, т. е. в стандарте H.265 будут допускаться большие потери, чем в H.264, но эти потери будут не так очевидны для восприятия.

После успешной разработки множества стандартов кодирования видео группа MPEG приступила к деятельности по стандартизации реконфигурируемого

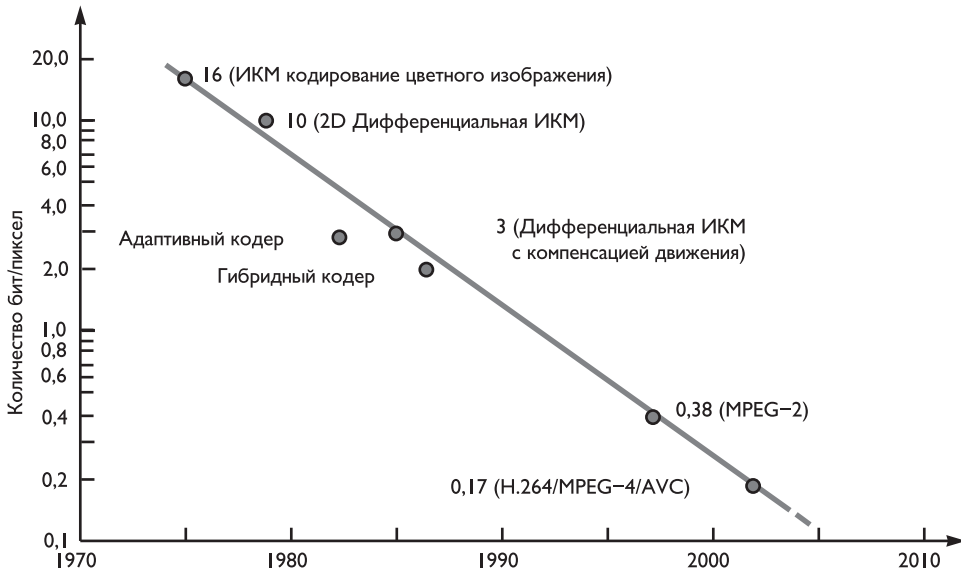


Рис. П.4. Достижения в кодировании видеосигналов

кодирования видео (RVC — Reconfigurable Video Coding). Стандарт RVC предлагает базовую структуру, подходящую для построения видеокодека с помощью конфигурации средств кодирования, таким образом способствуя динамической разработке, реализации и принятию решений по стандартизованному кодированию видео.

Совершенствование и развитие современных мультимедийных систем связано не только с реализацией эффективных методов сжатия различных видов информации, но и ее передачи в реальных сетях связи и вещания, предназначенных для доставки контента телерадиовещания, видеоконференцсвязи и ряда дополнительных услуг передачи данных. Такие сети, как правило, предназначены для достижения заданных качества обслуживания и режимов передачи/приема (например, с использованием наземных, спутниковых и кабельных сетей, Интернета).

В частности, цифровое телевизионное вещание в Европе основано на очень успешной серии стандартов DVB. Частотные распределения согласованы в рамках международного союза электросвязи ИТУ, в то время как сами сети вещания реализуются в соответствии с национальными спецификациями.

Как показано на рис. П.5, внедрение цифрового формата осуществляется в трех основных направлениях — спутникового, наземного и кабельного вещания. Первыми, последовательно внедряемыми, были спутниковые стандарты DVB-S, а затем DVB-S2. В дополнение к этим стандартам следует добавить стандарт DVB-RCS/RCS2 (RCS — Return Channel via Satellite), обеспечивающий реализацию обратного канала на основе режима многочастотного доступа с разделением по времени (MF-TDMA) для интерактивных спутниковых услуг, а также стандарт DVB-DSNG, используемый для сбора новостей через спутники.

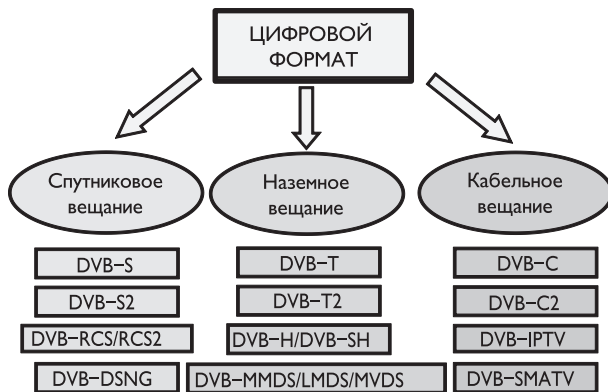


Рис. П.5. Научно-технический прогресс в сфере вещательных технологий

Стандарты DVB-T, а затем DVB-T2 определяют принципы трансляции эфирного цифрового телевизионного вещания. Стандарты DVB-H/DVB-SH используют эфирные и спутниковые каналы для передачи узкополосного телевизионного вещания на малогабаритные приемники с малыми экранами. Традиционно MMDS (Microwave Multipoint Distribution Service) — системы беспроводного многоканального телевидения ориентированы на микроволновые частоты в диапазоне 2,1 ГГц и от 2,5 ГГц до 2,7 ГГц. В последнее время все большее внимание уделяется системам беспроводной передачи на частотах выше 20 ГГц. В этой области стандартно используются диапазоны 25–32 ГГц и 40,5–42,5 ГГц. Нижний диапазон был первоначально освоен в США и Канаде, где системы высокочастотной передачи использовались вместо витой пары при построении телекоммуникационных сетей. В соответствии с назначением системы получили название LMDS (Local Multipoint Distribution Service). Сейчас технология LMDS при полосе каждого канала от 950 до 2150 МГц начала использоваться и в Европе, для построения телекоммуникационных, в том числе телевизионных сетей. Верхний диапазон первоначально был выделен в Европе для аналогового ТВ-вещания. В связи с этим системы, работающие на частотах 40,5–42,5 ГГц, получили название MVDS (Multipoint Video Distribution Systems). Наиболее привлекательным качеством систем MVDS является колоссальная ширина предоставляемого диапазона — 2 ГГц. Это в два с лишним раза превышает диапазон наземного вещания и в 10 раз — частотную полосу систем MMDS.

Системы DVB-C и DVB-C2 регулируют цифровые стандарты кабельного телевизионного вещания, использующие VHF и UHF-диапазоны частот на вторичной основе. К кабельным системам примыкают также стандарт телевизионного вещания по IP-сетям DVB-IPTV и стандарт DVB-SMATV (SMATV — Satellite Master Antenna Television), подразумевающий непосредственное распределение спутниковых сигналов по сетям кабельного телевидения. Действие европейского стандарта DVB-SMATV распространяется на кабельные сети до 3000 МГц, т. е. полностью — на весь диапазон промежуточных частот 950–2400 МГц.

С 2010 года внедряются стандарты цифрового вещания второго поколения для систем спутникового, эфирного и кабельного вещания — DVB-S2/DVB-T2/DVB-C2.

Эти стандарты применяются в службах вещания, распределения контента, электронного сбора новостей и подачи сигнала в студию, интерактивных службах. По сравнению с первым поколением стандартов DVB-S/DVB-T/DVB-C были изменены параметры обработки сигналов: использовано новое поколение прямой коррекции ошибок, что позволило вплотную приблизиться к пределу Шеннона; применяются более высокие созвездия, что привело к существенному повышению эффективности; увеличено возможное количество несущих OFDM; вводятся новые защитные интервалы; осуществлена минимизация количества рассеянных пилотов в зависимости от защитного интервала, в результате чего сокращены накладные расходы; расширена полоса пропускания; повышена эффективность обработки информации за счет расширенного временного и частотного перемежения битов и ячеек.

В табл. П.1 приведено сравнение достигаемых скоростей цифрового потока различных систем цифрового вещания первого и второго поколений для стандартных сетей связи.

Таблица П.1. Скорости цифрового потока, Мбит/с

Полоса канала	26 МГц		54 МГц		8 МГц			
	DVB-S	DVB-S2	DVB-S	DVB-S2	DVB-T	DVB-T2	DVB-C	DVB-C2
Минимальная	18,7	12,9	38,9	26,2	4,9	7,4	6,4	25,0
Максимальная	32,8	116,7	68,8	242,4	31,6	50,5	38,1	78,6

Из сравнения возможных максимальных скоростей цифрового потока, передаваемого по каналам цифрового телевизионного вещания систем первого и второго поколений стандартов, вытекает существенное увеличение эффективности использования стандартных каналов связи.

Так, например, если при использовании стандарта эфирного вещания DVB-T через канал с полосой пропускания 8 МГц возможно максимально передать до 14 программ стандартного телевидения или не более 6 программ телевидения высокой четкости, то при переходе на стандарт DVB-T2 эти данные могут быть увеличены примерно в 1,5 раза — до 25 и 10 соответственно.

Следует упомянуть несколько форматов мультимедийного вещания, ориентированных на прием с помощью мобильных терминалов. К ним относятся упомянутый при рассмотрении эфирных систем формат DVB-H, формат системы MediaFLO и группу форматов, базирующихся на системе радиовещания DAB (T-DAB, DAB+, T-DMB), японский стандарт эфирного вещания ISDB-T и китайский формат DTMB. Для сжатия информации стандартно в настоящее время используется технология MPEG-4 (H.264/AVC и HE-AAC), цифровые аудио- и видеосигналы мультиплексируются в один мультимедийный высокоскоростной транспортный поток.

Разработанная в России аудиовизуальная информационная система реального времени РАВИС (RAVIS) обеспечивает решение проблемы цифровизации сети мобильного радиовещания взамен аналоговых систем ЧМ-вещания (с полярной модуляцией и с пилот-тоном) в выделенных для этих целей в России диапазонах частот (66–74 МГц и 87,5–108 МГц соответственно) и внедрение собственных мультимедийных систем и соответствующей аппаратуры.

Система РАВИС позволяет в стандартном канале (200–250 кГц) передавать либо много программ стереозвука, включая поддержку объемного звукового ве-

щения, либо качественное цифровое телевизионное изображение с разрешением 352×288 пикселей на кадр и соответствующий звуковой стереосигнал.

Полностью отечественная разработка РАВИС имеет ряд ключевых преимуществ:

- возможность в стандартном канале ЧМ-вещания передачи нескольких стереофонических и многоканальных звуковых программ;
- трансляцию видеопрограмм для мобильного потребителя;
- реализацию телевизионного вещания в малонаселенных пунктах при удешевлении системы кодирования и трансляции;
- обеспечение устойчивого мобильного приема сигнала в условиях многолучевого распространения в среде с городской застройкой, в горной и лесистой местности, в водных акваториях;
- радиус покрытия в десятки раз больше, а затраты на создание инфраструктуры мобильного телевидения в десятки раз ниже (к примеру, покрытие мобильным телевидением всей территории Москвы может быть обеспечено одним либо максимум тремя передатчиками, работающими в одночастотном режиме);
- возможность создания одночастотных сетей мобильного вещания вдоль железнодорожных и автомобильных трасс;
- реализацию эфирных систем оповещения населения и организаций в условиях ЧС;
- организацию систем для использования различными спецслужбами и др.

В конце XX века одной из бурно развивающихся технологий в области телекоммуникаций стала видеоконференцсвязь (ВКС). Появление видеоконференцсвязи явилось технологическим развитием цифровой телефонии и видеотелефонии. Новые возможности, предоставляемые ВКС, привели к бурному росту количества пользователей этой технологии. Естественно, что развитие ВКС не могло идти без использования предшествующих стандартов ITU и создания новых рекомендаций, позволяющих унифицировать подходы к передаче мультимедийного контента и внедрения дополнительных сервисов.

Серьезной проблемой при внедрении видеоинформационных систем в России является подготовка специалистов данного профиля. Отечественная литература, посвященная изложению современных способов цифровой обработки мультимедийной информации, практически отсутствует, эта важнейшая область не отражена в ныне существующих программах обучения специалистов. Функцией обучения специалистов теории и практике мультимедийных систем, помимо создания представления о весьма сложных преобразованиях сигналов в процессе сокращения объема информации, должно стать привлечение молодых специалистов России к исследованиям еще не решенных проблем в этой перспективной области.

Хотелось бы надеяться, что начавшееся широкое внедрение в практику видеоинформационных систем послужит началом реализации планов обучения специалистов данного профиля. В области организации науки и подготовки кадров важнейшим является формирование сети научно-исследовательских центров по видеоинформационным проблемам, обладающих необходимым научным и тех-

ническим потенциалом для решения как фундаментальных проблем обработки сигналов, так и широкого круга прикладных задач.

Хотя проблемы метрологического обеспечения разработки и внедрения видеоинформационных систем в данной работе почти не затрагиваются, следует указать, что в России сложилась крайне тяжелая ситуация, вызванная повсеместным недостатком, а часто и отсутствием измерительной аппаратуры, необходимой для разработки, настройки и поддержания технических характеристик систем передачи информации в состоянии, обеспечивающем их качественное функционирование.

Особые проблемы возникают в связи с началом широкого внедрения систем и аппаратуры цифровой обработки и передачи компрессированных сигналов по каналам связи. Разрабатываемые системы требуют создания соответствующих контрольных и измерительных средств. С внедрением видеоинформационных систем потребность в использовании измерительной аппаратуры будет возрастать, поскольку требуется создание принципиально новых способов и средств контроля и измерений. При этом необходимо, чтобы эти средства были совместимы и с традиционными аналоговыми системами.

Метрологическая безопасность России — залог создания высококачественной видеоинформационной аппаратуры и ее эффективного использования.

На современном этапе развития техники выполнение требований метрологии связано с созданием виртуальных измерительных систем на базе использования персональных компьютеров в качестве устройств анализа и организации структуры систем формирования и обработки измерительной информации.

Для оценки соответствия качества передаваемой видеоинформации требованиям, предъявляемым к ней в различных системах и условиях передачи, а также для сравнения эффективности работы различных устройств видеокодирования и передачи видеоинформации, требуются методики и средства субъективной и объективной оценки качества видео, наблюдаемого потребителем. Эти методики должны учитывать характерные особенности и искажения как при аналоговой, так и при цифровой и смешанной системах обработки и передачи видеоинформации.

В настоящее время широко дебатировются вопросы, связанные с проблемами реализации и использования различных цифровых абонентских устройств. Самым рациональным решением проблемы приема является применение компьютерных систем, содержащих соответствующие тюнеры. Эффективность использования компьютеров связана также с возможностью постепенного наращивания мультимедийных функций путем замены плат компьютеров на более совершенные и дополнения их новыми платами и программами. При этом весьма привлекательным является использование программных декодеров, что обеспечивает возможность совершенствования стандартов кодирования, не изменяющих аппаратные средства абонентов.

Главным преимуществом применения компьютерной технологии является использование различных дисплеев: от настольных до настенных вариантов, от стандартного разрешения до разрешения высокой четкости и др. При этом изменение типа дисплея практически не требует замены аппаратной части.

В связи с этим важнейшей проблемой проведения фундаментальных работ в России по данной тематике является задача полного овладения теорией и ма-

тематическими методами обработки информации, используя имеющиеся зарубежные материалы в качестве информации, задающей общее направление работ. А далее с учетом полученного опыта по освоению стандартных алгоритмов следует разработка более эффективных систем.

Материал, изложенный в данной книге, состоит из девяти частей, изложенных в 22 главах.

В части I (главы 1–3) «Видеоинформационные приложения и объем цифровой информации» сжато изложены проблемы цифрового преобразования статических и динамических изображений, оцениваются объемы информации при преобразовании в цифровую форму изображений различного разрешения — от видеотелефонии до «цифрового кино», изложены принципы эффективного квантования различных видов информации.

Часть II (главы 4, 5) «Статистическая и визуальная избыточность изображений» посвящена исследованиям проблем статистической и визуальной избыточности монохромных и цветных изображений, излагаются методы и алгоритмы эффективного энтропийного кодирования различных видов информации об изображениях, оцениваются возможности использования визуальной избыточности статических и динамических изображений для сокращения объемов передаваемой информации.

В части III (главы 6–9) «Основные методы сжатия изображений» рассматриваются методы непосредственного кодирования и кодирования с предсказанием, групповое кодирование изображений, методы блочного и фрактального кодирования, кратномасштабная обработка изображений. Основное внимание уделено анализу дискретных линейных ортогональных преобразований и, в частности, широко используемым дискретному преобразованию Фурье, косинусному преобразованию, преобразованию Кархунена–Лоэва и др. Подробно анализируются методы обработки изображений с применением вейвлет-преобразований и реализации на их основе кратномасштабных преобразований изображений.

В части IV (главы 10–12) «Внутрикадровая и межкадровая обработка изображений» достаточно подробно излагаются методы повышения качества внутрикадрового кодирования статических изображений, принципы анализа движения деталей и его компенсации в динамических изображениях, идеология обработки изображений при кратномасштабном анализе. Особое внимание уделяется оптимизации методов поиска векторов движения, дополнительным возможностям компенсации движения деталей в динамических изображениях.

Часть V (главы 13, 14) «Основные стандарты цифрового кодирования видеоинформации» посвящена анализу методов кодирования статических изображений JPEG и JPEG-2000, системам кодирования динамических изображений MPEG-1, MPEG-2, обобщенному кодированию аудиовизуальных объектов MPEG-4, а также наиболее эффективному стандарту кодирования таких изображений H.264/AVC. Кроме того, в этой части рассмотрены некоторые проблемы реализации кодирующих устройств и визуализации их параметров.

В части VI (главы 15–18) «Основные методы кодирования речевой и звуковой информации» анализируются основные характеристики и цифровое представление речевой и звуковой информации, приводится достаточно подробное описание основных стандартов цифрового кодирования речи: G.711, G.722, G.723.1, G.728, G.729 и др., алгоритмы кодирования высококачественного звука, приведенные

в стандартах MPEG-1, MPEG-2, MPEG-4, Dolby AC-3. Рассмотрены также проблемы подавления акустического эха в системах телеконференцсвязи. Глава 18, посвященная подавлению акустического эха, написана совместно с Сараной Д.В.

Часть VII (главы 19, 20) «Основные методы и системы канального кодирования и модуляции цифровой информации» посвящена анализу основных методов модуляции и помехоустойчивого кодирования цифровой информации, приводится достаточно подробное описание систем многочастотной OFDM-модуляции, алгоритмы синхронизации OFDM-сигналов в приемных устройствах. Глава 19, посвященная канальному кодированию информации, составлена совместно с к. т. н. Бакке А.В.

В части VIII (глава 21) «Основные стандарты систем цифрового телевизионного вещания» приводится достаточно подробное описание основных современных стандартов передачи видео- и аудиоинформации: ATSC, DVB-S/S2, DVB-T/T2, DVB-C/C2, ISDB-S, ISDB-T, DTMB (DMB-T/H). Проводится сравнение европейских стандартов первого и второго поколений. Показано, сколь важен быстрый переход в России на системы вещания второго поколения DVB-S2, DVB-T2, DVB-C2.

Последняя часть IX (главы 22, 23) «Стандарты систем мобильного радиовещания и видеоконференцсвязи» посвящена изложению стандартов телерадиовещания DVB-H, MediaFLO, T-DMB/T-DAB, РАВИС и DRM+, а также проблемам реализации систем видеоконференцсвязи на базе использования трех разделов рекомендаций ITU-T:

- аудиовизуальные и мультимедийные системы (серия Н);
- передающие системы и средства, цифровые системы и цепи (серия G);
- терминалы для телематических сервисов (серия Т).

Основное внимание уделено изложению принципов построения российской аудиовизуальной системы реального времени — РАВИС, предназначенной для создания цифровых систем взамен аналоговых сетей УВЧ/ФМ радиовещания, а также реализации разработки российских программных решений видеоконференцсвязи VPhone и «ДАВ телекон». Глава 23, в которой рассмотрены проблемы построения систем видеоконференцсвязи, написана совместно с к. м. н. Федоровым В.Ф.

ЧАСТЬ I

ВИДЕОИНФОРМАЦИОННЫЕ ПРИЛОЖЕНИЯ И ОБЪЕМ ЦИФРОВОЙ ИНФОРМАЦИИ

Введение

Процесс цифровой обработки визуальной информации охватил различные виды видеоинформационных приложений и существенно расширил их спектр: кроме широко распространенного традиционного телевидения и различных экзотических разновидностей ТВ специального назначения появились и повсеместно внедряются самые различные видеоприложения — от видеотелефонии до телевидения высокой четкости (ТВЧ) и «цифрового кино» [1.1, 1.2].

Цифровые системы кодирования изображений, их передачи и приема позволяют достичь невиданного доселе уровня качества и предоставляют пользователю массу новых возможностей и новых видов услуг [1.1–1.4].

Неудивительно в связи с этим то, что цифровая обработка телевизионных и компьютерных изображений в виду ее особой важности выделилась в самостоятельную область техники, в которую входят [1.5–1.7]:

- коррекция изображений, их «препарирование», т. е. сознательное разделение на части цифровыми средствами, видоизменение этих частей и их обратная «сборка»;
- оценка параметров изображений с целью контроля качества их передачи и приема;
- преобразование и кодирование изображений для хранения или передачи по каналам связи;

- компьютерная графика, а также визуализация информации, т. е. представление массивов данных в виде различных изображений, что очень эффективно, так как облегчает решение многих задач, сложных именно своей абстрактностью.

Эта область включает также моделирование систем обработки, хранения и передачи визуальной информации по каналам связи, т. е. набор компьютерно-математических задач, необходимых для разработки новой цифровой телевизионной техники.

Разумеется, все перечисленные задачи связаны между собой — и по методам решения, и по используемым для этого техническим средствам.

Цифровая обработка изображений является, по сути, основной базой для создания нового поколения видеоинформационной техники. Без нее немыслима реализация систем видеотелефонии, видеоконференцсвязи, цифрового интерактивного телевизионного вещания, цифрового кино. Работы по созданию таких систем уже полным ходом ведутся сегодня в технически развитых странах, и привлеченные финансовые, технические и интеллектуальные ресурсы таковы, что становится совершенно ясно: переход к цифровым видеоинформационным системам в общемировом масштабе неизбежен и является делом близкого будущего.

Именно по всем этим причинам исследования методов цифровой обработки видеоинформации должны стать одной из важнейших частей курса обучения специалистов широкого профиля, и в частности, телевидения.

Однако и сама цифровая обработка изображений в связи с особенностями проблем, стоящих на пути создания видеоинформационных систем будущего, имеет свое ядро, т. е. свою главную задачу, без решения которой немыслимо решение всех остальных.

Это — задача так называемого сжатия изображений, т. е. уменьшения объема информации, необходимого для передачи информации по каналам связи путем сокращения содержащейся в составе изображения информационной избыточности. Таким образом, ключевыми вопросами, решение которых определяет эффективность развивающихся и вновь создаваемых систем, являются задачи компрессии видеоинформации и оценки ее качества.

Решение именно этой задачи имеет, в частности, важное значение для внедрения телевидения высокой четкости. Упомянутая «высокая четкость» в этих системах достигается увеличением как количества строк в кадре, так и количества элементов в каждой строке, а значит, резко возрастает и соответствующий каждому изображению (кадру) объем информации.

Между тем передавать эти изображения предстоит с использованием существующих каналов связи, пропускная способность которых, естественно, остается прежней. Становится очевидным, что как только будет создан эффективный и достаточно просто реализуемый алгоритм сжатия телевизионных изображений (а такие работы ведутся давно, и варианты решения этой задачи практически уже найдены), внедрение телевидения высокой четкости станет реальностью.

Однако не только для телевидения высокой четкости нужно сжатие изображений. Максимально эффективное использование существующих каналов связи является сегодня для телевидения проблемой номер один.

Частотное пространство телевидения во многих случаях ограничено, а, кроме того, крайне привлекательным представляется внедрять новые системы телеви-

дения, не меняя при этом существующую линейную каналообразующую аппаратуру. Поэтому решение такой задачи, как обеспечение возможности передачи хотя бы 4–6 программ телевидения в стандартном радиоканале, является сегодня исключительно важным, прежде всего с точки зрения экономической.

Таким образом, разработка эффективного способа сжатия изображений позволит решить сразу две задачи, равные по своей важности.

Главной функцией данного материала является создание ясного представления о сложных цифровых преобразованиях сигналов в процессе сокращения объема передаваемой визуальной информации.

Прежде чем перейти к изложению методов цифровой обработки видеoinформации, необходимо оценить ее объем для различных систем формирования и передачи изображений. Этому и посвящен первый раздел настоящего материала. В нем кратко изложены сведения об особенностях существующих стандартов телевидения, а также систем кодирования цвета, а кроме того, рассмотрены некоторые вопросы, связанные с представлением в цифровом виде сигналов различных систем передачи изображений. Следует отметить, что везде в этом разделе, где будет говориться о способах уменьшения объема информации при цифровой передаче изображений, имеются в виду широко известные тривиальные методы предварительного снижения объема цифрового потока, которые сами по себе не в состоянии обеспечить решения изложенных выше задач и не имеют ничего общего с собственно алгоритмами сжатия изображений, содержание которых изложено в последующих разделах.

ГЛАВА I

ФОРМАТЫ ИЗОБРАЖЕНИЙ. СТАТИЧЕСКИЕ И ДИНАМИЧЕСКИЕ ЭТАЛОННЫЕ ИЗОБРАЖЕНИЯ

1.1. Аналоговые и цифровые растровые изображения

Телевизионные и компьютерные изображения обычно воспроизводятся в виде растра, представимого непрерывной или дискретной функцией вдоль строк по горизонтали и конечным числом строк по вертикали. В аналоговом телевидении каждый кадр динамического изображения формируется путем чересстрочной развертки в виде двух последовательно передаваемых полей [1.8, 1.9].

На рис. 1.1 приведено схематичное представление чересстрочной развертки изображения. Сплошные черные линии характеризуют движение луча электронной трубки в одном поле, а пунктирные — в другом поле кадра. Тонкие серые штриховые линии характеризуют обратный ход луча по строкам и между полями.

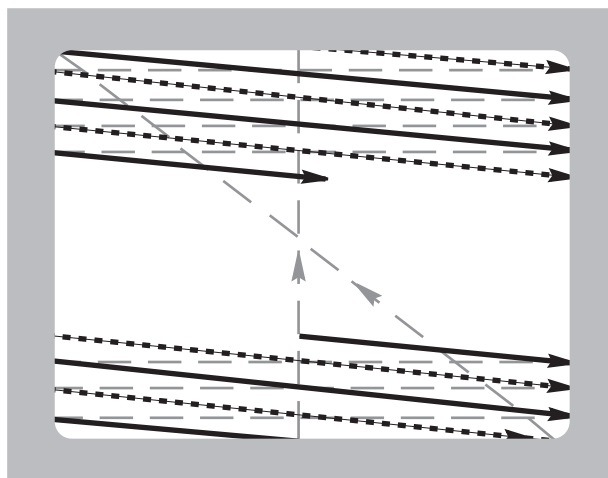


Рис. 1.1. Представление растра аналогового телевидения

Цифровые изображения разбиваются также и вдоль строк на дискретные отсчеты, называемые **пикселями** (от сокращения английских слов — **PI**Cture's **E**lement, элемент изображения). Количество пикселей, на которые разбито изображение, определяет его детальность. Если изображение представимо в виде M

строк, каждая из которых содержит N дискретных отсчетов, то массив кадра состоит из $N \times M$ пикселей.

В различных приложениях используются разные форматы кадров. Так, эталонные компьютерные изображения часто имеют равное количество строк и пикселей в каждой строке, т. е. NM . В стандартном телевидении используется формат кадра с соотношением $N : M = 4 : 3$, а в телевидении высокой четкости и в цифровом кино — с соотношением $N : M = 16 : 9$ [1.2]. Однако номенклатура используемых форматов кадра значительно более обширна, часть достаточно широко используемых форматов приведена в табл. 1.1 [1.10–1.18].

Таблица 1.1. Форматы кадра изображения

Наименование формата	Разрешение, $N \times M$	Соотношение сторон кадра	Применение
1	2	3	4
Общий формат обмена CIF и его производные			
QQCIF	88×72	11 : 9	Мобильные телефоны
SQCIF (Sub-QCIF)	128×96	4 : 3	- «-
QCIF	176×144	11 : 9	Видеоконференция (PAL/SECAM)
CIF	352×288	- «-	- «-
iCIF	352×576	11:18	Цифровое видео (PAL/SECAM)
2CIF	704×288	22:9	- «-
4CIF	704×576	11:9	ТВ PAL/SECAM
16CIF	1408×1152	- «-	Цифровое видео
Стандартный формат изображения SIF и его производные			
SQSIF	80×60	4:3	Мобильные телефоны
QSIF	160×120	- «-	- «-
SIF	320×240	- «-	Видеоконференция (NTSC)
iSIF	352×480	11:15	Цифровое видео(NTSC)
2SIF	704×240	44:15	-
4SIF	704×480	22:15	ТВ NTSC
Форматы графических подсистем			
QQVGA	80×60	4:3	Мобильные телефоны
QVGA	160×120	- «-	- «-
QVGA	320×240	- «-	Видеоконференция
VGA	640×480	- «-	ТВ NTSC
WVGA	800×480	5:3	Компьютерные мониторы
SVGA	800×600	4:3	- «-
XGA	1024×768	- «-	- «-
WXGA	1280×768	5:3	- «-
QuadVGA	1280×960	8:3	- «-
SXGA	1280×1024	5:4	- «-
SXGA+	1400×1050	4:3	- «-

Таблица 1.1 (окончание)

1	2	3	4
WSXGA	1440×900	8:5	-«-
UGA, UVGA	1600×1200	4:3	-«-
UXGA	1600×1280	5:4	-«-
WSXGA+	1680×1050	8:5	-«-
WUXGA	1920×1200	-«-	-«-
QXGA	2048×1536	4:3	-«-
WQXGA	2560×1600	8:5	-«-
QSXGA	2560×2048	5:4	-«-
WQSXGA	3200×2048	25:16	-«-
QUXGA	3200×2400	4:3	-«-
WQUXGA	3840×2400	8:5	-«-
HXGA	4096×3072	4:3	-«-
WHXGA	5120×3200	8:5	-«-
HSXGA	5120×4096	5:4	-«-
WHSXGA	6400×4096	25:16	-«-
HUXGA	6400×4800	4:3	-«-
WHUXGA	7680×4800	8:5	-«-
Видеоформаты и телевизионные форматы			
w288p	512×288	16:9	Широкоформатное видео
400p	528×400	4:3	Видео стандартной четкости
448p	576×448	-«-	-«-
w384p	672×384	16:9	Широкоформатное видео
ТВ стандарт (D1)	720×480	3:2	DVD-Video NTSC
ТВ стандарт (D1)	720×576	5:4	DVD-Video PAL/SECAM
w448p	768×448	16:9	Широкоформатное видео
w480p, ED	854×480	-«-	LCD/PDP ТВ-дисплеи
w576p	1024×576	-«-	Широкоформатное видео
w720p, HD	1280×720	-«-	ТВ повышенной четкости
w768p	1366×768	-«-	LCD/PDP ТВ дисплеи
w1080p, UHD	1920×1080	-«-	ТВЧ ATSC
w1152p, UHD	2048×1152	-«-	ТВЧ DVB-T
w4320p, UHDV	7680×4320	-«-	ТВ сверхвысокой четкости

Форматы кадров в этой таблице разбиты на 4 раздела.

Первый из них характеризует **Общий Формат Обмена CIF (Common Intermediate Format)**, обычно используемый в телеконференциях, и производ-

ные от него форматы. Второй раздел определяет параметры стандартного формата изображения **SIF (Standard Image/Interchange Format)** и его производных. Третий, наибольший раздел характеризует форматы графических подсистем, используемые в основном в компьютерных дисплеях. И, наконец, в четвертом разделе приведен перечень основных форматов, используемых в телевизионных системах стандартной и высокой четкости.

В зависимости от параметров каждого пиксела изображения можно разделить на бинарные, полутоновые, полноцветные и палитровые [1.6, 1.19].

Пикселы **бинарных изображений** имеют лишь два значения, характеризующие черные или белые точки. Такая информация используется для передачи простых изображений, например при передаче факсов.

Полутоновые изображения характеризуются значительным количеством уровней на каждый пиксел — его величина определяет яркость воспроизводимого элемента. Существуют различные форматы полутоновых изображений, определяющие динамический диапазон воспроизводимых яркостей элементов. Чаще всего используется формат элементов, который кодируется 8 битами на пиксел, т. е. может принимать значения от 0 до $2^8 - 1 = 255$. В ряде случаев, например в медицинских технологиях, используют полутоновые изображения, элементы которых кодируются до 16 битов на пиксел, т. е. число возможных градаций их яркости может составлять $2^{16} = 65\,536$.

Каждый пиксел **полноцветного изображения** представим в виде тройки — красной (**Red**), зеленой (**Green**) и синей (**Blue**) составляющих, каждая из которых обычно задается 8 битами — значениями от 0 до 255. Следовательно, цвет каждого пиксела определяется 24 битами, и изображение теоретически может содержать $(2^8)^3 = 16\,777\,216$ различных цветов.

Величины красного, зеленого и синего цветов являются *координатами цвета* и обозначаются соответственно X , Y и Z . Часто описание цвета выражают *координатами цветности* — относительными величинами, которые задаются соотношениями:

$$x = \frac{X}{X + Y + Z}; \quad y = \frac{Y}{X + Y + Z}; \quad z = \frac{Z}{X + Y + Z}. \quad (1.1)$$

Из (1.1) следует, что $x + y + z = 1$.

На рис. 1.2 приведено схематическое изображение цветового куба хуз. Точки, соответствующие красному, зеленому и синему цветам, расположены на трех вершинах куба с координатами $(1, 0, 0)$, $(0, 1, 0)$ и $(0, 0, 1)$. Точки на главной диагонали представляют собой оттенки серого цвета: от черного в начале координат $(0, 0, 0)$ до белого в точке $(1, 1, 1)$.

На рис. 1.3 представлены изображения цветового куба, состоящего из $(2^8)^3$ различных цветов. Иногда 24-битовое цветное изображение называют изображением в естественных цветах, поскольку при этом возможно воспроизведение практически всех оттенков реально существующих предметов.

Палитровые изображения используются в целом ряде приложений, в которых имеет смысл использовать ограниченное число цветов. При этом используемое количество воспроизводимых цветов ограничено величиной 256 или меньшим числом. Применительно к Интернет-приложениям это подмножество цветов называется палитрой фиксированных Web цветов или набором цветов, одинаково

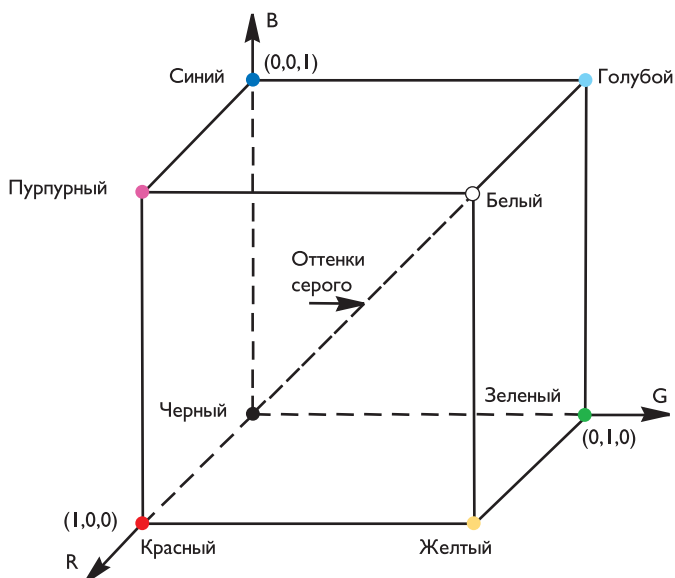


Рис. 1.2. Схематическое изображение цветового куба

воспроизводимых всеми программами просмотра Интернет-сайтов. Поскольку из 256 цветов только 216 воспроизводятся одинаково большинством операционных систем, эти $216 = 6^3$ цветов стали *de facto* стандартом фиксированных цветов (см. рис. 1.4).

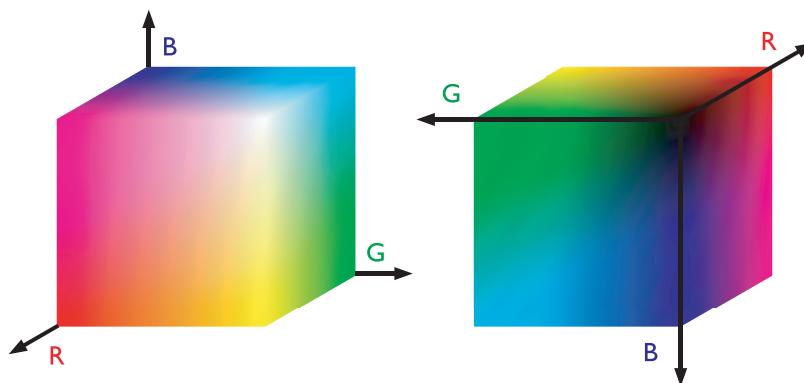


Рис. 1.3. 24-битовый полноцветный куб RGB: вид с двух сторон

В данном варианте каждая из трех RGB компонент может принимать лишь 6 возможных значений: 0, 51, 102, 153, 204 и 255. Эти значения обычно выражают в шестнадцатеричной системе счисления, как приведено в табл. 1.2.

Поскольку для формирования цвета требуется три числа, каждый цвет из палитры задается тремя двузначными числами (в шестнадцатеричной системе

счисления). Так, шесть воспроизводимых оттенков серого цвета, входящих в па-

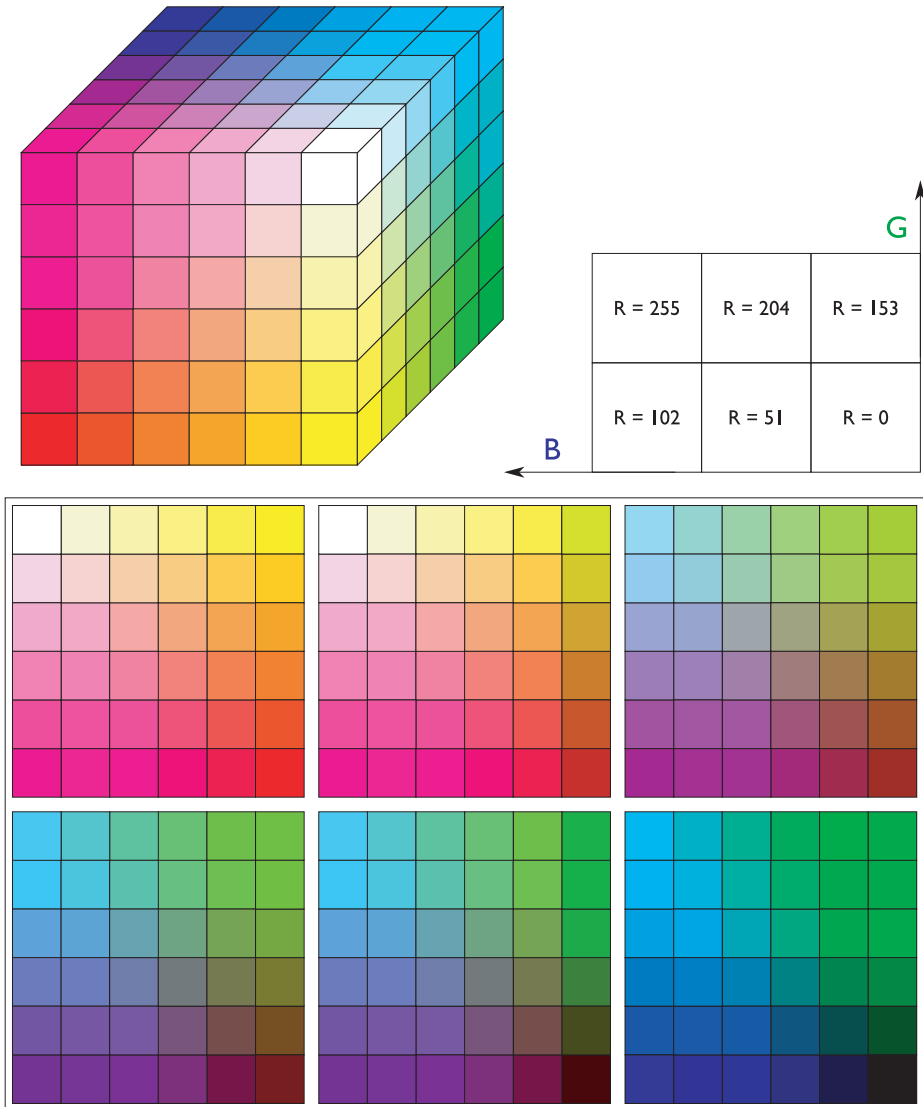


Рис. 1.4. Куб (а) и палитра (б) фиксированных RGB цветов

Таблица 1.2. Значения RGB компонент в палитре фиксированных частот

Система счисления	Значения RGB компонент					
Шестнадцатеричная	0	33	66	99	CC	FF
Десятичная	0	51	102	153	204	255

литру, определяются числами (00 00 00), (33 33 33), (66 66 66), (99 99 99), (CC CC CC),

(FF FF FF). Наиболее яркие красный, зеленый и синий цвета задаются, соответственно, значениями (FF 00 00), (00 FF 00) и (00 00 FF).

Основной закон смешения цветов устанавливает трехмерность цвета и возможность выражения практически всех цветов через смеси только трех линейно независимых цветов. При этом яркость смеси цветов равна сумме яркостей составляющих смеси [1.21–1.23].

Трехмерное цветовое пространство является удобной и наглядной формой геометрического представления множества цветов и действий над ними. В цветовом пространстве можно ввести сколько угодно координатных систем. Началом координат во всех этих системах обычно устанавливается точка нулевого цвета — черного.

1.2. Цветовые системы

Цветовая система полностью определяется заданием трех координатных осей, исходящих из начала координат, и соответствующим их масштабированием. В общем случае устанавливаемые по осям масштабы могут быть различными. Основные цвета, положенные в основу цветовой координатной системы, используются для выражения остальных цветов в данной координатной системе. Важным условием выбора основных цветов является их линейная независимость, т. е. ни один из трех основных цветов не должен получаться путем смешения двух других цветов.

В системе координат с заданными основными цветами X, Y, Z любой цвет F определяется соотношением

$$F = \alpha_x X + \alpha_y Y + \alpha_z Z. \quad (1.2)$$

Яркость F_B также является одной из возможных цветовых координат:

$$F_B = \alpha_{Bx} X + \alpha_{By} Y + \alpha_{Bz} Z. \quad (1.3)$$

Коэффициенты $\alpha_{Bx}, \alpha_{By}, \alpha_{Bz}$ характеризуют вклады каждой из цветовых координат в яркость воспроизводимого цвета и являются яркостными коэффициентами в данной цветовой системе координат. Термин «яркость» обозначает количественную меру интенсивности ахроматического (неокрашенного) света, которая принимает значения в диапазоне от черного до белого с промежуточными серыми оттенками.

Хроматический (окрашенный) свет характеризуется следующими параметрами [1.21, 1.22]:

- электромагнитным спектром, охватывающим диапазон длин волн примерно от 380 до 780 нм (рис. 1.5);
- потоком лучистой энергии, излучаемой в единицу времени и измеряемой в ваттах;
- световым потоком, оцениваемым по зрительному ощущению и измеряемым в люменах;
- светлотой, цветовым тоном и насыщенностью, которые отражают уровень производимого световым потоком зрительного ощущения и определяют цветное восприятие.

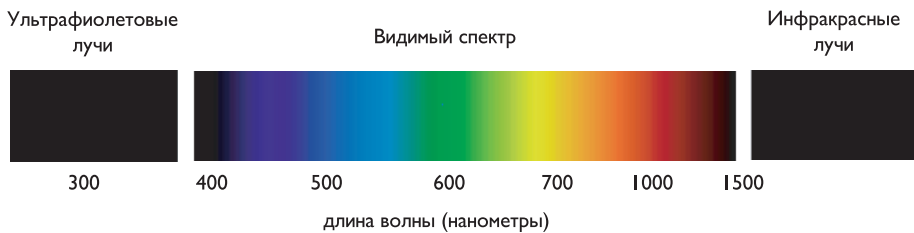


Рис. 1.5. Длины волн видимой части электромагнитного спектра

Цветовая система RGB

Субъективные измерения цветовых координат смешения трех чистых спектральных цветов (при одинаковой мощности монохроматических излучений с равной энергией) привели к разработке принятого Международной комиссией по освещению (МКО) стандартного набора монохроматических первичных основных цветов: красного (R) с длиной волны 700 нм, зеленого (G) — 546,1 нм и синего (B) — 435,8 нм.

В соответствии с теорией цветового зрения предполагается наличие в сетчатке глаза светочувствительных приемников (колбочек) трех видов: 65% из них воспринимают красный свет, 33% — зеленый свет и 2% — синий свет.

Совокупность реакций каждого из трех приемников на падающее излучение определяет восприятие цвета. Если обозначить спектральные чувствительности трех приемников $\bar{R}(\lambda)$, $\bar{G}(\lambda)$ и $\bar{B}(\lambda)$, то три составляющие светового потока R , G и B определяются соотношениями:

$$R = \int_{\lambda_{\min}}^{\lambda_{\max}} \varphi(\lambda) \bar{R}(\lambda) d\lambda, \quad G = \int_{\lambda_{\min}}^{\lambda_{\max}} \varphi(\lambda) \bar{G}(\lambda) d\lambda, \quad B = \int_{\lambda_{\min}}^{\lambda_{\max}} \varphi(\lambda) \bar{B}(\lambda) d\lambda, \quad (1.4)$$

где $\varphi(\lambda)$ — спектральная интенсивность лучистого потока на длине волны λ .

На рис. 1.6 представлены относительные кривые спектральной чувствительности трех светочувствительных приемников глаза [1.20]. Эти зависимости восприятия цвета показывают, что никакая монохроматическая составляющая не может характеризовать красный, зеленый или синий цвет. Следовательно, определенный набор монохроматических первичных составляющих не обеспечивает возможность получения всех цветов спектра. Разным цветовым ощущениям в одинаковых условиях наблюдения соответствуют разные спектральные распределения излучения. Однако обратное утверждение неверно, т. е. ощущение одного и того же цвета может возникать при различных спектральных распределениях излучения.

Трехцветное выражение цвета опирается на установленный опытным путем факт, что для большинства встречающихся в природе цветов можно выполнить зрительное уравнивание со смесями взятых в различных пропорциях только трех **основных** цветов. Эта связь взаимно однозначна независимо от спектрального распределения излучения данного цвета [1.23–1.29].

В основе RGB-модели лежит Декартова система координат, цветовое пространство в соответствии с соотношениями (1.1) и рис. 1.3 представляет собой

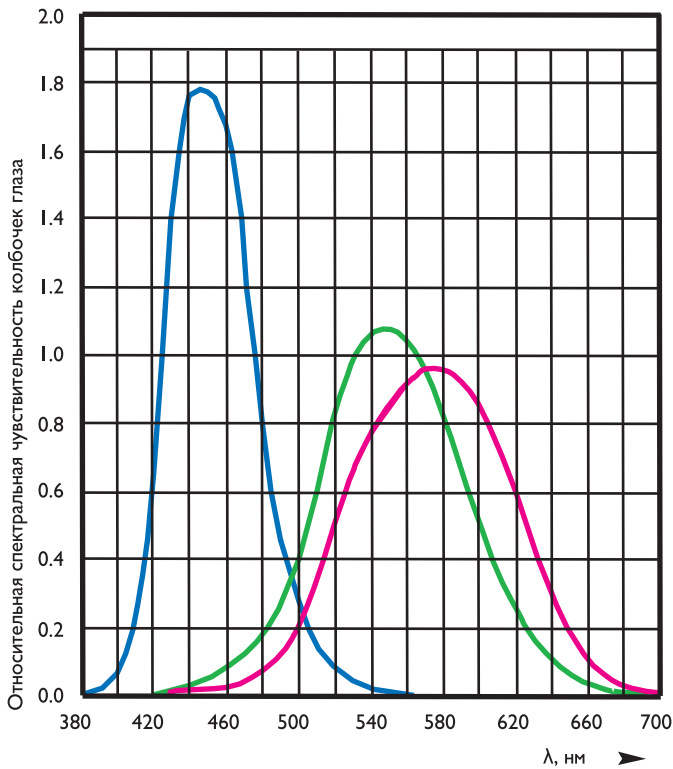


Рис. 1.6. Функции относительной спектральной чувствительности трех видов светочувствительных приемников глаза

куб. Другой способ задания цвета основан на применении двумерной диаграммы цветностей МКО с x (красной) и y (зеленой) координатами (рис. 1.7). Третья z (синяя) координата вычисляется однозначно: $z = 1 - x - y$.

На диаграмме цветностей МКО показано положение цветностей чистых спектральных цветов, образующих подковообразную кривую чистых спектральных цветностей. Линия, соединяющая точки чистых спектральных цветностей с длинами волн 380 и 780 нм, называется линией **пурпурных цветов**. На этой же диаграмме приведены точки стандартизованных величин источников света, которые воспроизводят условия освещения: А — искусственного — электрическими лампами накаливания; В — прямого солнечного света; С — рассеянного дневного; D_{65} — усредненного дневного. Точка Е — точка равной энергии при $x = y = z = 1/3$, она соответствует **цветности естественного дневного цвета**.

Источники света характеризуют также цветовой температурой, соответствующей абсолютной температуре идеального излучателя (абсолютно черного тела), цвет которого имеет ту же цветность, что и данное излучение (табл. 1.3) [1.29].

Выбранные на диаграмме МКО точки основных цветов определяют цветовую модель RGB. При этом сочетания двух из основных цветов: зеленый + синий, синий + красный, красный + зеленый — воспроизводятся в виде вторичных цветов — голубого, пурпурного и желтого цветов, как показано на рис. 1.8.

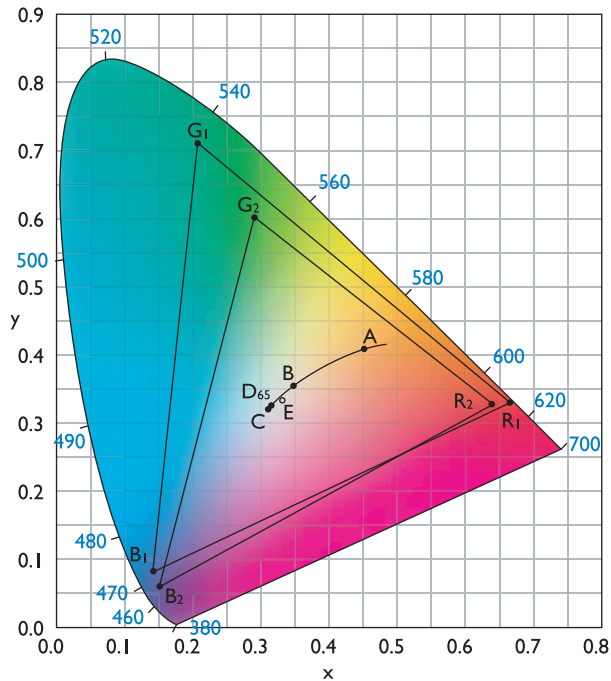


Рис. 1.7. Диаграмма цветностей МКО

Таблица 1.3. Координаты цветности и цветовая температура стандартных источников света

Источник излучения	Координаты цветности		Цветовая температура, К
	x	Y	
A	0,4512	0,4059	2856
B	0,3498	0,3527	4874
C	0,3104	0,3191	6774
D_{65}	0,3138	0,3310	6504
E	1/3	1/3	5400

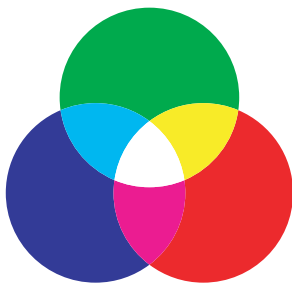


Рис. 1.8. Первичные и вторичные основные цвета световых источников

Таким образом, использование трех основных цветов световых источников — RGB, в которых каждый элемент (пиксел) имеет глубину представления 24 бита (по 8 битов на красную, зеленую и синюю составляющие), позволяет создавать на экране цветного монитора полноцветное (True Color) изображение, как показано на рис. 1.9.

Цветовые системы CMY и CMYK

Система CMY используется в цветных принтерах и копировальных устройствах, в которых цветные красители наносятся на бумагу.



Рис. 1.9. Формирование полноцветного изображения

Если в системе RGB голубой (**С**уан), пурпурный (**М**агента) и желтый (**Y**ellow) являются вторичными цветами, то в системе CMY они образуют первичные цвета красителей, реализуя субтрактивный цветовой синтез.

Очевидно, если поверхность, покрытая голубой краской, освещается белым светом, то красный цвет от нее не отражается. От поверхностей, покрытых пурпурной или желтой краской, не отражаются, соответственно, зеленый и синий цвета.

Преобразование данных из RGB в CMY осуществляется простой операцией:

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \quad (1.5)$$

где значения цветов нормированы и находятся в диапазоне $[0, 1]$.

В связи с тем, что смешивание в равном количестве голубого, пурпурного и желтого цветов приводит при печати к появлению не черного, а грязно-серого цвета, система CMY расширяется до системы CMYK, содержащей дополнительно ключевой (**К**ey) — черный цвет (рис. 1.10).

Цветовая система HSI

Цветовые системы RGB и CMYK не приспособлены для словесного описания цветов. Цветной объект обычно описывается такими понятиями, как цветовой тон, его насыщенность и светлота, соответствующая понятию интенсивности (полутоновой яркости).

В цветовой системе HSI (**H**ue — цветовой тон, **S**aturation — насыщенность, **I**ntensity — интенсивность) используется естественное описание цвета. Специфи-

кой системы HSI является тот факт, что яркостная информация (интенсивность) отделена от цветовой информации, характеризуемой цветовым тоном и насыщенностью.

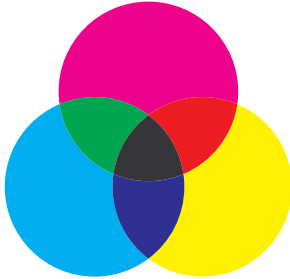


Рис. 1.10. Первичные и вторичные основные цвета красителей

Существуют различные формы геометрической интерпретации системы HSI. Одна из них представляет все цветовое пространство в виде фигуры из двух одинаковых и соединенных своими основаниями конусов, как показано на рис. 1.11.

Расстояние между вершинами этих конусов нормировано и равно 1, так что высота каждого конуса равна 0,5. Вершина нижнего конуса соответствует уровню черного (0, 0), а верхнего — уровню белого (1, 0). Таким образом, интенсивность (яркость) объекта определяется расстоянием от точки уровня черного до горизонтального сечения, характеризующего цветовую информацию. На рис. 1.11 приведено два сечения, соответствующие $I = 0,5$ (в соединении оснований двух конусов) и $I = 0,75$.

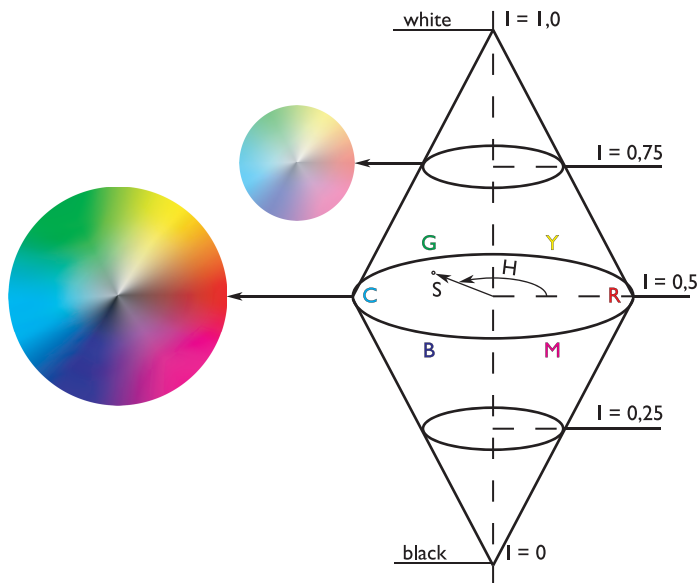


Рис. 1.11. Геометрическая интерпретация цветовой системы HSI

Цветовые параметры объекта характеризуются параметрами вектора, соединяющего точку с центром круга, образованного соединением оснований двух конусов, т. е. сечением фигуры на уровне интенсивности $I = 0,5$. Радиус этого круга равен 1. При этом цветовой тон определяется углом H° наклона вектора относительно ординаты, проведенной в направлении красного цвета. В представленной на рис. 1.11 плоскости сечения фигуры первичные основные (красный, зеленый и синий) цвета расположены относительно друг друга под углом 120° .

Вторичные (голубой, пурпурный и желтый) цвета смещены относительно друг друга также на углы 120° , а относительно первичных цветов — на углы 60° . Насыщенность цвета S имеет значения в диапазоне $[0, 1]$ и определяется длиной вектора, соединяющего центр круга с точкой, определяющей цветовые параметры воспроизводимой детали.

Преобразование цветов из системы RGB в систему HSI

Если координаты RGB нормированы таким образом, что их значения ограничены диапазоном $[0, 1]$, и угол H° отсчитывается от красной оси пространства HSI, то интенсивность (яркость), насыщенность и цветовой тон определяются соотношениями:

$$I = \frac{1}{3}(R + G + B), \quad (1.6)$$

$$S = \frac{\min(R, G, B)}{I}, \quad (1.7)$$

$$H^\circ = \begin{cases} \theta^\circ, & \text{при } B \leq G, \\ 360^\circ - \theta^\circ, & \text{при } B > G, \end{cases} \quad (1.8)$$

где $\theta^\circ = \left(\frac{180}{\pi}\right) \cdot \arccos \left\{ \frac{R - \frac{1}{2}(G+B)}{[(R-G)^2 + (R-B)(G-B)]^{1/2}} \right\}$.

Преобразование цветов из системы HSI в систему RGB

Формулы преобразования различны для трех секторов параметра H° .

RG сектор ($0^\circ \leq H^\circ < 120^\circ$):

$$\begin{cases} R = I \left[1 + \frac{S \cos H^\circ}{\cos(60^\circ - H^\circ)} \right], \\ B = I(1 - S), \\ G = 3I - (R + B). \end{cases} \quad (1.9)$$

GB сектор ($120^\circ \leq H^\circ < 240^\circ$):

$$\begin{cases} G = I \left[1 + \frac{S \cos H^\circ}{\cos(180^\circ - H^\circ)} \right], \\ R = I(1 - S), \\ B = 3I - (R + G). \end{cases} \quad (1.10)$$

BR сектор ($240^\circ \leq H^\circ < 360^\circ$):

$$\begin{cases} B = I \left[1 + \frac{S \cos H^\circ}{\cos(300^\circ - H^\circ)} \right], \\ G = I(1 - S), \\ R = 3I - (B + G). \end{cases} \quad (1.11)$$

Цветовые системы YUV, YIQ, YD_RD_B, YC_RC_B

Эти цветовые системы чаще всего используются в аналоговом и цифровом телевидении. Составляющая Y характеризует яркость элементов телевизионного изображения, ее величина определяется линейной комбинацией основных цветов:

$$Y = \alpha_R R + \alpha_G G + \alpha_B B, \quad (1.12)$$

где коэффициенты $\alpha_R, \alpha_G, \alpha_B$ определяют параметры исходного белого цвета, и при изменяющихся основных RGB цветах в диапазоне $[0, 1]$ $\alpha_R + \alpha_G + \alpha_B = 1$.

Остальные параметры этих систем $(U, V), (I, Q), (D_R, D_B), (C_R, C_B)$ — линейно зависят от двух так называемых цветоразностных составляющих R-Y и B-Y или являются их линейной комбинацией. Более подробное описание этих систем приводится в главе 2.

Цветовая система CIELAB

Унифицированное пространство восприятия цвета CIELAB является стандартом, принятым МКО в 1976 году [1.30]. Евклидово расстояние между двумя любыми одинаково визуально различимыми цветовыми точками этого пространства имеет практически равную величину.

В цветовом пространстве CIELAB (точнее, CIE 1976 L^*a^*b) значение **светлоты** отделено от значения хроматической составляющей цвета (**тон, насыщенность**). Светлота задана координатой L (изменяется от 0 до 100, т. е. от самого темного до самого светлого), хроматическая составляющая — двумя полярными координатами a и b. Первая обозначает положение цвета в диапазоне от зеленого до пурпурного, вторая — от синего до желтого. Это позволяет эффективно применять данную систему при цветовом анализе.

На рис. 1.12 приведены изображения системы L^*a^*b для двух значений светлоты L.

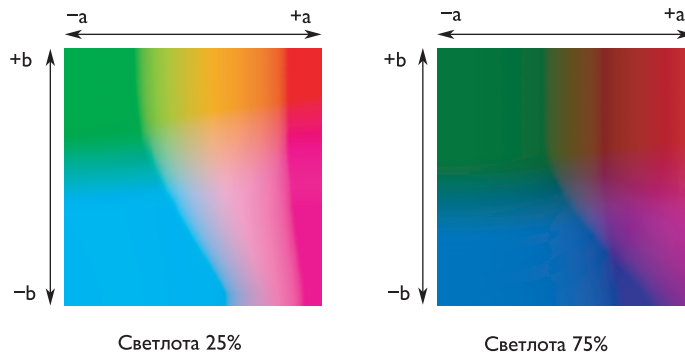


Рис. 1.12. Изображения изменений полярных координат a и b системы CIELAB при двух значениях светлоты L

В отличие от цветовых пространств RGB или CMYK, которые являются, по сути, набором аппаратных данных для воспроизведения цвета на экране монитора или на бумаге (цвет может зависеть от типа печатной машины, марки красок, влажности воздуха в цехе или производителя монитора и его настроек), CIELAB однозначно определяет цвет. Поэтому CIELAB нашел широкое применение в программном обеспечении для обработки изображений в качестве промежуточного цветового пространства, через которое происходит конвертирование данных между другими цветовыми пространствами. При этом особые свойства L^*a^*b сделали редактирование в этом пространстве мощным инструментом **цветокоррекции**.

Благодаря характеру определения цвета в $L^*a^*b^*$ появляется возможность отдельно воздействовать на яркость, контраст изображения и на его цвет. $L^*a^*b^*$ предоставляет возможность избирательного воздействия на отдельные цвета в изображении, усиления цветового контраста, незаменимыми являются и возможности, которые это цветовое пространство предоставляет для борьбы с шумом на изображении.

Учитывая соотношения

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0,412453 & 0,357580 & 0,180423 \\ 0,212671 & 0,715150 & 0,072169 \\ 0,019334 & 0,119193 & 0,950227 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \quad (1.13)$$

перевод RGB данных в компоненты $L^*a^*b^*$ осуществляется следующим образом:

$$\begin{cases} L^* = 116 \cdot f\left(\frac{Y}{Y_n}\right) - 16, \\ a^* = 500 \cdot \left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right) \right], \\ b^* = 200 \cdot \left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right) \right], \end{cases} \quad (1.14)$$

где

$$f(q) = \begin{cases} q^{1/3}, & \text{при } q < 0,008856; \\ 7,787 \cdot q + 16/116, & \text{иначе;} \end{cases} \quad q \in \left\{ \frac{X}{X_n}, \frac{Y}{Y_n}, \frac{Z}{Z_n} \right\}.$$

Величины X_n, Y_n, Z_n определяют источник белого цвета D_{65} стандарта МКО и могут быть получены при $R = G = B = 100$.

1.3. Статические и динамические эталонные изображения

Цифровая обработка изображений применяется для решения различных задач:

- пространственного и частотного улучшения изображений;
- восстановления изображений, пораженных различными видами помех;
- сокращения избыточности и эффективного кодирования изображений для хранения и передачи по каналам связи;
- распознавания и обнаружения объектов в изображениях и др.

При создании систем цифровой обработки изображений используются различные статические и динамические эталонные изображения, обладающие определенными свойствами [1.31–1.35].

Статические эталонные изображения применяются, в частности, для анализа эффективности методов сжатия изображений, а также для анализа внутрикадрового кодирования динамических изображений. Динамические эталонные изображения используются, например, для анализа эффективности поиска векторов и компенсации движения деталей в изображениях, и др.

Статические эталонные изображения

На рис. 1.13а приведено широко используемое классическое цветное изображение «Лена» («Lenna»), содержащее множество плавных цветовых переходов и наклонных границ, а на рис. 1.13б — черно-белое изображение «Элина» («Elaine»).



Рис. 1.13. Изображения «Лена» (а) и «Элина» (б)



Рис. 1.14. Цветные изображения «Девушка» (а) и «Перцы» (б)

Аналогичными свойствами обладают эталонные изображения «Девушка» («Girl») и «Перцы» («Peppers»), которые даны на рис. 1.14.

На рис. 1.15 представлены три черно-белых изображения. Первое из них — «Барбара» («Barbara») — содержит множество различно ориентированных полосок на скатерти, платке и брюках; при неэффективной обработке этого изображения часто возникает муар и другие артефакты. На втором изображении — «Золотой Холм» («Goldhill») — на переднем и заднем планах содержится множество мелких деталей, третье изображение — «Кинооператор» («Camerman») — содержит ряд протяженных разнонаправленных деталей.

Рис. 1.16 иллюстрирует два классических тестовых цветных изображения, содержащие вертикально и горизонтально ориентированные элементы, — «Гавань» («Harbour») и «Корабль» («Boat»).

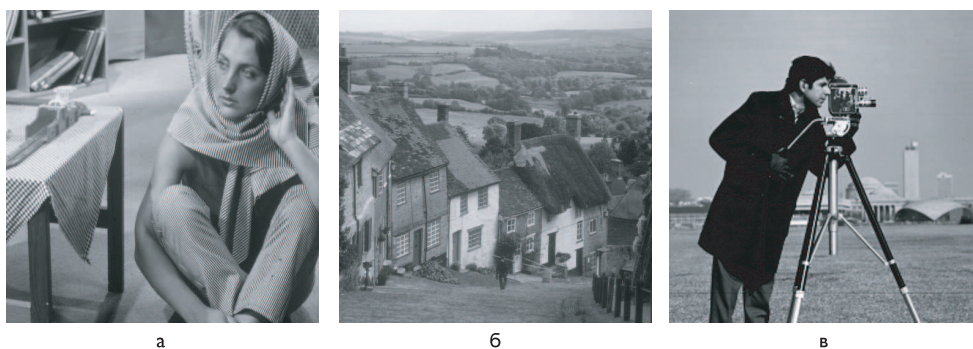


Рис. 1.15. Черно-белые эталонные изображения «Барбара» (а), «Золотой Холм» (б), «Кинооператор» (в)



Рис. 1.16. Цветные изображения «Гавань» (а), «Корабль» (б)

Динамические эталонные изображения

При анализе качества преобразований динамических изображений используются различные эталоны в зависимости от вида приложений.

Так, при разработке систем видеотелефонии и видеонаблюдения применяют малоподвижные изображения.

Такие последовательности можно увидеть при воспроизведении AVI файлов, содержащих цветные динамические изображения «Клер» («Claire», 335 кадров), черно-белые изображения «Сьюзи» («Susie», 149 кадров), цветные изображения «Монитор холла» («Hall_monitor», 599 кадров).

При создании систем видеоконференцсвязи используют изображения с более подвижными деталями. В качестве примеров приведены AVI файлы трех последовательностей изображений: «Теннис» («Tennis», 223 кадра), «Прораб» («Foreman», 599 кадров) и «В автомобиле» («Carphone», 191 кадр).

Для оценки качества кодирования телевизионных изображений чаще всего используются динамические изображения «Сад цветов» («FlowerGarden», 229 кадров), «Береговая охрана» («Coastguard», 300 кадров) и «Футбол» («Football», 249 кадров).



Рис. 1.17. Кадры из видеопоследовательности «Клер»



Рис. 1.18. Кадры из видеопоследовательности «Сьюзи»



Рис. 1.19. Кадры из видеопоследовательности «Монитор холла»



Рис. 1.20. Кадры из видеопоследовательности «Теннис»



Рис. 1.21. Кадры из видеопоследовательности «Прораб»

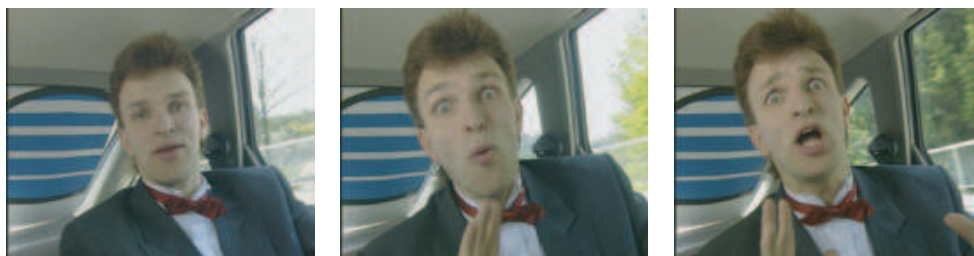


Рис. 1.22. Кадры из видеопоследовательности «В автомобиле»



Рис. 1.23. Кадры из видеопоследовательности «Сад цветов»



Рис. 1.24. Кадры из видеопоследовательности «Береговая охрана»



Рис. 1.25. Кадры из видеопоследовательности «Футбол»