

Contents

Preface	x
I Introduction and Multimedia Data Representations	1
1 Introduction to Multimedia	2
1.1 What is Multimedia?	2
1.1.1 Components of Multimedia	2
1.2 Multimedia: Past and Present	3
1.2.1 Early History of Multimedia	3
1.2.2 Hypermedia, WWW, and Internet	6
1.2.3 Multimedia in the New Millennium	9
1.3 Multimedia Software Tools: A Quick Scan	11
1.3.1 Music Sequencing and Notation	11
1.3.2 Digital Audio	11
1.3.3 Graphics and Image Editing	12
1.3.4 Video Editing	12
1.3.5 Animation	13
1.3.6 Multimedia Authoring	13
1.4 Multimedia in the Future	14
1.5 Exercises	15
2 A Taste of Multimedia	18
2.1 Multimedia Tasks and Concerns	18
2.2 Multimedia Presentation	18
2.3 Data Compression	24
2.4 Multimedia Production	26
2.5 Multimedia Sharing and Distribution	26
2.6 Some Useful Editing and Authoring Tools	29
2.6.1 Adobe Premiere	29
2.6.2 Adobe Director	31
2.6.3 Adobe Flash	35
2.7 Exercises	38
3 Graphics and Image Data Representations	42
3.1 Graphics/Image Data Types	42
3.1.1 1-Bit Images	42
3.1.2 8-Bit Gray-Level Images	43
3.1.3 Image Data Types	45
3.1.4 24-Bit Color Images	45
3.1.5 Higher Bit-depth Images	46
3.1.6 8-Bit Color Images	47
3.1.7 Color Lookup Tables (LUTs)	48
3.2 Popular File Formats	51

3.2.1	GIF	51
3.2.2	JPEG	54
3.2.3	PNG	55
3.2.4	TIFF	55
3.2.5	Windows BMP	56
3.2.6	Windows WMF	56
3.2.7	Netpbm Format	56
3.2.8	EXIF	56
3.2.9	PS and PDF	56
3.2.10	PTM	57
3.3	Exercises	57
4	Color in Image and Video	60
4.1	Color Science	60
4.1.1	Light and Spectra	60
4.1.2	Human Vision	61
4.1.3	Spectral Sensitivity of the Eye	62
4.1.4	Image Formation	62
4.1.5	Camera Systems	64
4.1.6	Gamma Correction	64
4.1.7	Color-Matching Functions	66
4.1.8	CIE Chromaticity Diagram	67
4.1.9	Color Monitor Specifications	70
4.1.10	Out-of-Gamut Colors	70
4.1.11	White Point Correction	71
4.1.12	XYZ to RGB Transform	72
4.1.13	Transform with Gamma Correction	72
4.1.14	L*a*b* (CIELAB) Color Model	73
4.1.15	More Color Coordinate Schemes	74
4.1.16	Munsell Color Naming System	74
4.2	Color Models in Images	75
4.2.1	RGB Color Model for Displays	75
4.2.2	Multi-sensor Cameras	75
4.2.3	Camera-Dependent Color	75
4.2.4	Subtractive Color: CMY Color Model	76
4.2.5	Transformation from RGB to CMY	77
4.2.6	Undercolor Removal: CMYK System	77
4.2.7	Printer Gamuts	77
4.2.8	Multi-Ink Printers	78
4.3	Color Models in Video	78
4.3.1	Video Color Transforms	78
4.3.2	YUV Color Model	80
4.3.3	YIQ Color Model	81
4.3.4	YCbCr Color Model	82
4.4	Exercises	83
5	Fundamental Concepts in Video	87
5.1	Analog Video	87
5.1.1	NTSC Video	89
5.1.2	PAL Video	91
5.1.3	SECAM Video	92
5.2	Digital Video	92
5.2.1	Chroma Subsampling	92
5.2.2	CCIR and ITU-R Standards for Digital Video	93
5.2.3	High Definition TV (HDTV)	94
5.2.4	Ultra High Definition TV (UHDTV)	95

5.3	Video Display Interfaces	95
5.3.1	Analog Display Interfaces	95
5.3.2	Digital Display Interfaces	97
5.4	3D Video and TV	98
5.4.1	Cues for 3D Percept	98
5.4.2	3D Camera Models	99
5.4.3	3D Movie and TV Based on Stereo Vision	99
5.4.4	The Vergence-Accommodation Conflict	100
5.4.5	Autostereoscopic (Glasses-Free) Display Devices	101
5.4.6	Disparity Manipulation in 3D Content Creation	102
5.5	Exercises	103
6	Basics of Digital Audio	105
6.1	Digitization of Sound	105
6.1.1	What Is Sound?	105
6.1.2	Digitization	106
6.1.3	Nyquist Theorem	107
6.1.4	Signal-to-Noise Ratio (SNR)	109
6.1.5	Signal-to-Quantization-Noise Ratio (SQNR)	109
6.1.6	Linear and Nonlinear Quantization	111
6.1.7	Audio Filtering	113
6.1.8	Audio Quality versus Data Rate	114
6.1.9	Synthetic Sounds	115
6.2	MIDI: Musical Instrument Digital Interface	116
6.2.1	MIDI Overview	116
6.2.2	Hardware Aspects of MIDI	119
6.2.3	Structure of MIDI Messages	121
6.2.4	General MIDI	123
6.2.5	MIDI-to-WAV Conversion	123
6.3	Quantization and Transmission of Audio	124
6.3.1	Coding of Audio	124
6.3.2	Pulse Code Modulation	124
6.3.3	Differential Coding of Audio	126
6.3.4	Lossless Predictive Coding	126
6.3.5	DPCM	129
6.3.6	DM	131
6.3.7	ADPCM	132
6.4	Exercises	133
II	Multimedia Data Compression	137
7	Lossless Compression Algorithms	138
7.1	Introduction	138
7.2	Basics of Information Theory	139
7.3	Run-Length Coding	140
7.4	Variable-Length Coding (VLC)	141
7.4.1	Shannon–Fano Algorithm	141
7.4.2	Huffman Coding	143
7.4.3	Adaptive Huffman Coding	146
7.5	Dictionary-Based Coding	149
7.6	Arithmetic Coding	153
7.6.1	Basic Arithmetic Coding Algorithm	154
7.6.2	Scaling and Incremental Coding	157
7.6.3	Integer Implementation	160
7.6.4	Binary Arithmetic Coding	160

7.6.5	Adaptive Arithmetic Coding	161
7.7	Lossless Image Compression	163
7.7.1	Differential Coding of Images	163
7.7.2	Lossless JPEG	164
7.8	Exercises	165
8	Lossy Compression Algorithms	169
8.1	Introduction	169
8.2	Distortion Measures	169
8.3	The Rate-Distortion Theory	170
8.4	Quantization	170
8.4.1	Uniform Scalar Quantization	171
8.4.2	Nonuniform Scalar Quantization	173
8.4.3	Vector Quantization	174
8.5	Transform Coding	175
8.5.1	Discrete Cosine Transform (DCT)	175
8.5.2	Karhunen–Loève Transform*	186
8.6	Wavelet-Based Coding	189
8.6.1	Introduction	189
8.6.2	Continuous Wavelet Transform*	192
8.6.3	Discrete Wavelet Transform*	194
8.7	Wavelet Packets	201
8.8	Embedded Zerotree of Wavelet Coefficients	202
8.8.1	The Zerotree Data Structure	203
8.8.2	Successive Approximation Quantization	204
8.8.3	EZW Example	205
8.9	Set Partitioning in Hierarchical Trees (SPIHT)	207
8.10	Exercises	207
9	Image Compression Standards	211
9.1	The JPEG Standard	211
9.1.1	Main Steps in JPEG Image Compression	211
9.1.2	JPEG Modes	218
9.1.3	A Glance at the JPEG Bitstream	220
9.2	The JPEG2000 Standard	221
9.2.1	Main Steps of JPEG2000 Image Compression*	222
9.2.2	Adapting EBCOT to JPEG2000	228
9.2.3	Region-of-Interest Coding	228
9.2.4	Comparison of JPEG and JPEG2000 Performance	228
9.3	The JPEG-LS Standard	229
9.3.1	Prediction	232
9.3.2	Context Determination	232
9.3.3	Residual Coding	233
9.3.4	Near-Lossless Mode	233
9.4	Bilevel Image Compression Standards	233
9.4.1	The JBIG Standard	233
9.4.2	The JBIG2 Standard	234
9.5	Exercises	235
10	Basic Video Compression Techniques	238
10.1	Introduction to Video Compression	238
10.2	Video Compression Based on Motion Compensation	238
10.3	Search for Motion Vectors	239
10.3.1	Sequential Search	240
10.3.2	2D Logarithmic Search	241
10.3.3	Hierarchical Search	242

10.4	H.261	244
10.4.1	Intra-Frame (I-Frame) Coding	245
10.4.2	Inter-Frame (P-Frame) Predictive Coding	245
10.4.3	Quantization in H.261	245
10.4.4	H.261 Encoder and Decoder	246
10.4.5	A Glance at the H.261 Video Bitstream Syntax	248
10.5	H.263	249
10.5.1	Motion Compensation in H.263	250
10.5.2	Optional H.263 Coding Modes	251
10.5.3	H.263+ and H.263++	253
10.6	Exercises	253
11	MPEG Video Coding: MPEG-1, 2, 4, and 7	256
11.1	Overview	256
11.2	MPEG-1	256
11.2.1	Motion Compensation in MPEG-1	257
11.2.2	Other Major Differences from H.261	257
11.2.3	MPEG-1 Video Bitstream	260
11.3	MPEG-2	261
11.3.1	Supporting Interlaced Video	262
11.3.2	MPEG-2 Scalabilities	265
11.3.3	Other Major Differences from MPEG-1	268
11.4	MPEG-4	270
11.4.1	Overview of MPEG-4	270
11.4.2	Video Object-Based Coding in MPEG-4	271
11.4.3	Synthetic Object Coding in MPEG-4	282
11.4.4	MPEG-4 Parts, Profiles and Levels	288
11.5	MPEG-7	288
11.5.1	Descriptor (D)	289
11.5.2	Description Scheme (DS)	291
11.5.3	Description Definition Language (DDL)	293
11.6	Exercises	294
12	New Video Coding Standards: H.264 and H.265	298
12.1	H.264	298
12.1.1	Motion Compensation	299
12.1.2	Integer Transform	302
12.1.3	Quantization and Scaling	303
12.1.4	Examples of H.264 Integer Transform and Quantization	305
12.1.5	Intra Coding	305
12.1.6	In-loop Deblocking Filtering	308
12.1.7	Entropy Coding	308
12.1.8	Context-Adaptive Variable Length Coding (CAVLC)	310
12.1.9	Context-Adaptive Binary Arithmetic Coding (CABAC)	311
12.1.10	H.264 Profiles	313
12.1.11	H.264 Scalable Video Coding (SVC)	315
12.1.12	H.264 Multiview Video Coding (MVC)	315
12.2	H.265	315
12.2.1	Motion Compensation	316
12.2.2	Integer Transform	320
12.2.3	Quantization and Scaling	321
12.2.4	Intra Coding	321
12.2.5	Discrete Sine Transform (DST)	321
12.2.6	In-loop Filtering	322
12.2.7	Entropy Coding	323
12.2.8	Special Coding Modes	324

12.2.9	H.265 Profiles	324
12.3	Comparisons of Video Coding Efficiency	324
12.3.1	Objective Assessment	325
12.3.2	Subjective Assessment	325
12.4	Exercises	325
13	Basic Audio Compression Techniques	329
13.1	ADPCM in Speech Coding	329
13.1.1	ADPCM	329
13.2	G.726 ADPCM, G.727-9	329
13.3	Vocoders	332
13.3.1	Phase Insensitivity	332
13.3.2	Channel Vocoder	333
13.3.3	Formant Vocoder	333
13.3.4	Linear Predictive Coding	333
13.3.5	CELP	336
13.3.6	Hybrid Excitation Vocoders*	340
13.4	Exercises	342
14	MPEG Audio Compression	345
14.1	Psychoacoustics	345
14.1.1	Equal-Loudness Relations	346
14.1.2	Frequency Masking	346
14.1.3	Temporal Masking	351
14.2	MPEG Audio	352
14.2.1	MPEG Layers	352
14.2.2	MPEG Audio Strategy	353
14.2.3	MPEG Audio Compression Algorithm	354
14.2.4	MPEG-2 AAC (Advanced Audio Coding)	359
14.2.5	MPEG-4 Audio	359
14.3	Other Audio Codecs	360
14.3.1	Ogg Vorbis	360
14.4	MPEG-7 Audio and Beyond	361
14.5	Further Exploration	362
14.6	Exercises	362
III	Multimedia Communications and Networking	365
15	Network Services and Protocols for Multimedia Communications	366
15.1	Protocol Layers of Computer Communication Networks	366
15.2	Local Area Network (LAN) and Access Networks	367
15.2.1	LAN Standards	367
15.2.2	Ethernet Technology	368
15.2.3	Access Network Technologies	369
15.3	Internet Technologies and Protocols	373
15.3.1	Network Layer: IP	373
15.3.2	Transport Layer: TCP and UDP	374
15.3.3	Network Address Translation (NAT) and Firewall	378
15.4	Multicast Extension	379
15.4.1	Router-Based Architectures: IP Multicast	379
15.4.2	Non Router-Based Multicast Architectures	381
15.5	Quality-of-Service for Multimedia Communications	381
15.5.1	Quality of Service (QoS)	382
15.5.2	Internet QoS	384
15.5.3	Rate Control and Buffer Management	387

15.6	Protocols for Multimedia Transmission and Interaction	389
15.6.1	HyperText Transfer Protocol (HTTP)	389
15.6.2	Real-Time Transport Protocol (RTP)	390
15.6.3	RTP Control Protocol (RTCP)	391
15.6.4	Real-Time Streaming Protocol (RTSP)	392
15.7	Case Study: Internet Telephony	393
15.7.1	Signaling Protocols: H.323 and Session Initiation Protocol (SIP)	394
15.8	Further Exploration	396
15.9	Exercises	396
16	Internet Multimedia Content Distribution	401
16.1	Proxy Caching	402
16.1.1	Sliding-Interval Caching	403
16.1.2	Prefix Caching and Segment Caching	404
16.1.3	Rate-Split Caching and Work-ahead Smoothing	405
16.1.4	Summary and Comparison	407
16.2	Content Distribution Networks (CDNs)	407
16.2.1	Representative: Akamai Streaming CDN	409
16.3	Broadcast/Multicast Video-on-Demand	411
16.3.1	Smart TV and Set-Top Box (STB)	411
16.3.2	Scalable Multicast/Broadcast VoD	412
16.4	Broadcast/Multicast for Heterogeneous Users	416
16.4.1	Stream Replication	416
16.4.2	Layered Multicast	416
16.5	Application-Layer Multicast	418
16.5.1	Representative: End-System Multicast (ESM)	419
16.5.2	Multi-tree Structure	420
16.6	Peer-to-Peer Video Streaming with Mesh Overlays	421
16.6.1	Representative: CoolStreaming	421
16.6.2	Hybrid Tree and Mesh Overlay	424
16.7	HTTP-based Media Streaming	425
16.7.1	HTTP for Streaming	425
16.7.2	Dynamic Adaptive Streaming over HTTP (DASH)	426
16.8	Exercises	428
17	Multimedia over Wireless and Mobile Networks	433
17.1	Characteristics of Wireless Channels	433
17.1.1	Path Loss	433
17.1.2	Multipath Fading	434
17.2	Wireless Networking Technologies	435
17.2.1	1G Cellular Analog Wireless Networks	436
17.2.2	2G Cellular Networks: GSM and Narrow-Band CDMA	437
17.2.3	3G Cellular Networks: Wideband CDMA	439
17.2.4	4G Cellular Networks and Beyond	440
17.2.5	Wireless Local Area Networks (WLAN)	442
17.2.6	Bluetooth and Short Range Technologies	444
17.3	Multimedia over Wireless Channels	444
17.3.1	Error Detection	445
17.3.2	Error Correction	447
17.3.3	Error Resilient Coding	450
17.3.4	Error Concealment	454
17.4	Mobility Management	455
17.4.1	Network Layer Mobile IP	456
17.4.2	Link-Layer Handoff Management	458
17.5	Further Exploration	459
17.6	Exercises	459

18 Social Media Sharing	464
18.1 Representative Social Media Services	464
18.1.1 User-Generated Content (UGC) Sharing	464
18.1.2 Online Social Networking	465
18.2 User-Generated Media Content Sharing	465
18.2.1 YouTube Video Format and Meta-data	465
18.2.2 Characteristics of YouTube Video	466
18.2.3 Small-World in YouTube Videos	468
18.2.4 YouTube from a Partner's View	470
18.2.5 Enhancing UGC Video Sharing	472
18.3 Media Propagation in Online Social Networks	475
18.3.1 Sharing Patterns of Individual Users	476
18.3.2 Video Propagation Structure and Model	477
18.3.3 Video Watching and Sharing Behaviors	479
18.3.4 Coordinating Live Streaming and Online Storage	479
18.4 Further Exploration	481
18.5 Exercises	481
19 Cloud Computing for Multimedia Services	485
19.1 Cloud Computing Overview	485
19.1.1 Representative Storage Service: Amazon S3	488
19.1.2 Representative Computation Service: Amazon EC2	489
19.2 Multimedia Cloud Computing	491
19.3 Cloud-Assisted Media Sharing	492
19.3.1 Impact of Globalization	494
19.3.2 Case Study: Netflix	495
19.4 Computation Offloading for Multimedia Services	496
19.4.1 Requirements for Computation Offloading	497
19.4.2 Service Partitioning for Video Coding	497
19.4.3 Case Study: Cloud-Assisted Motion Estimation	498
19.5 Interactive Cloud Gaming	499
19.5.1 Issues and Challenges of Cloud Gaming	499
19.5.2 Real World Implementation	501
19.6 Further Exploration	504
19.7 Exercises	504
20 Content-Based Retrieval in Digital Libraries	507
20.1 How Should We Retrieve Images?	507
20.2 Synopsis of Early CBIR Systems	509
20.3 C-BIRD — A Case Study	510
20.3.1 Color Histogram	510
20.3.2 Color Density and Color Layout	512
20.3.3 Texture Layout	513
20.3.4 Search by Illumination Invariance	514
20.3.5 Search by Object Model	515
20.4 Quantifying Search Results	516
20.5 Key Technologies in Current CBIR Systems	519
20.5.1 Robust Image Features and Their Representation	519
20.5.2 Relevance Feedback	521
20.5.3 Other Post-processing Techniques	521
20.5.4 Visual Concept Search	522
20.5.5 The Role of Users in Interactive CBIR Systems	522
20.6 Querying on Videos	523
20.7 Querying on Videos Based on Human Activity	525

20.7.1	Modeling Human Activity Structures	525
20.7.2	Experimental Results	527
20.8	Quality-Aware Mobile Visual Search	528
20.8.1	Related Work	529
20.8.2	Quality-aware Method	530
20.8.3	Experimental Results	530
20.9	Exercises	532