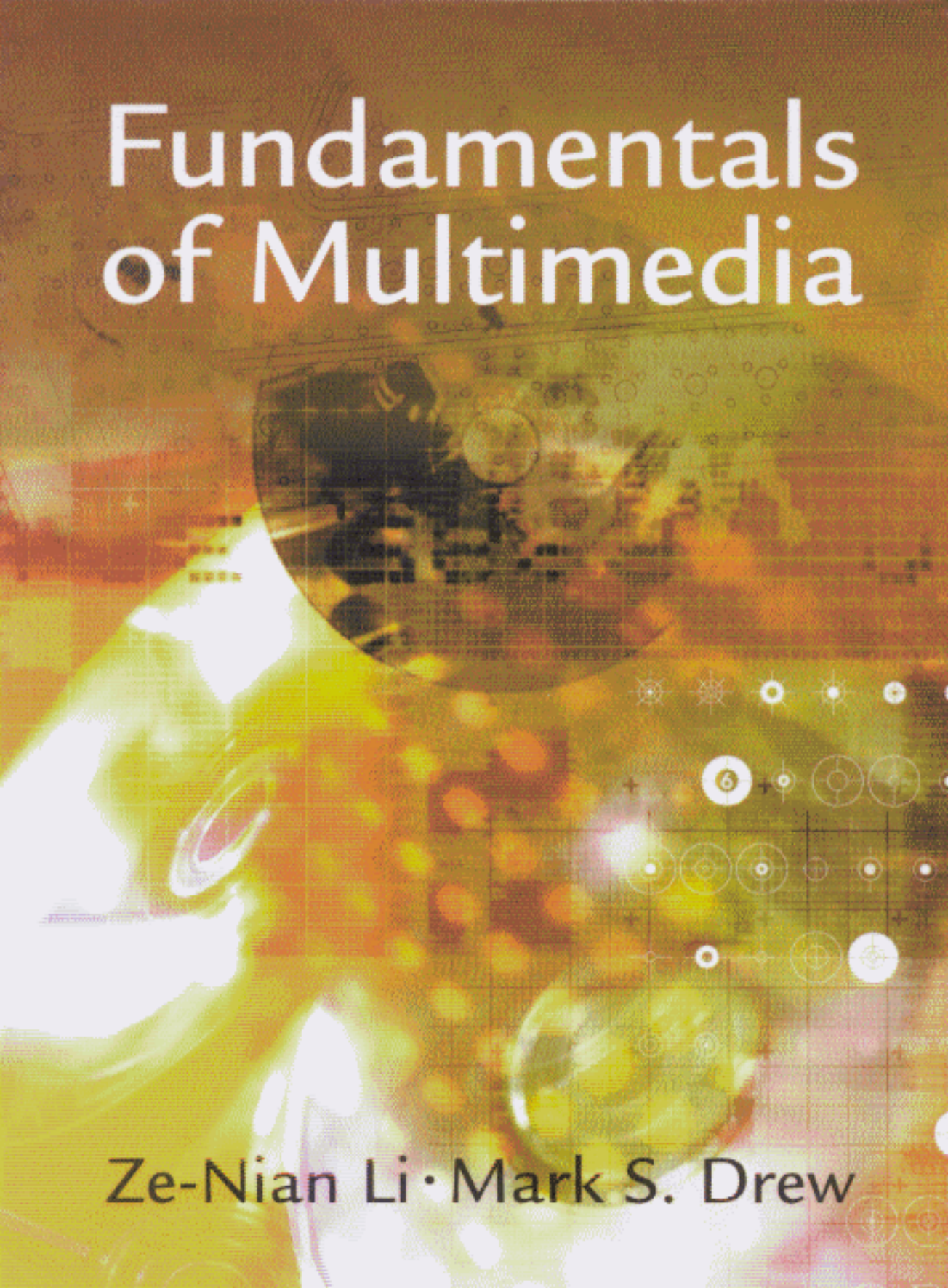


Fundamentals of Multimedia

The background is a complex digital composition. It features a grid pattern overlaid on a warm, golden-yellow color palette. A prominent circular inset in the upper center shows a globe with a grid of latitude and longitude lines. To the left, there's a bright, glowing, abstract shape that resembles a stylized letter 'L' or a similar character. The right side of the image is filled with various geometric patterns, including small circles, some with internal crosshairs, and larger circles with numbers inside. The overall aesthetic is futuristic and technical, representing multimedia and digital technology.

Ze-Nian Li • Mark S. Drew

Contents

Preface	xiv
I Multimedia Authoring and Data Representations	1
1 Introduction to Multimedia	3
1.1 What is Multimedia? 3	
1.1.1 Components of Multimedia 3	
1.1.2 Multimedia Research Topics and Projects 4	
1.2 Multimedia and Hypermedia 5	
1.2.1 History of Multimedia 5	
1.2.2 Hypermedia and Multimedia 7	
1.3 World Wide Web 8	
1.3.1 History of the WWW 8	
1.3.2 HyperText Transfer Protocol (HTTP) 9	
1.3.3 HyperText Markup Language (HTML) 10	
1.3.4 Extensible Markup Language (XML) 11	
1.3.5 Synchronized Multimedia Integration Language (SMIL) 12	
1.4 Overview of Multimedia Software Tools 14	
1.4.1 Music Sequencing and Notation 14	
1.4.2 Digital Audio 15	
1.4.3 Graphics and Image Editing 15	
1.4.4 Video Editing 15	
1.4.5 Animation 16	
1.4.6 Multimedia Authoring 17	
1.5 Further Exploration 17	
1.6 Exercises 18	
1.7 References 19	
2 Multimedia Authoring and Tools	20
2.1 Multimedia Authoring 20	
2.1.1 Multimedia Authoring Metaphors 21	
2.1.2 Multimedia Production 23	
2.1.3 Multimedia Presentation 25	
2.1.4 Automatic Authoring 33	
2.2 Some Useful Editing and Authoring Tools 37	
2.2.1 Adobe Premiere 37	
2.2.2 Macromedia Director 40	
2.2.3 Macromedia Flash 46	
2.2.4 Dreamweaver 51	
2.3 VRML 51	
2.3.1 Overview 51	
2.3.2 Animation and Interactions 54	
2.3.3 VRML Specifics 54	
2.4 Further Exploration 55	
2.5 Exercises 56	
2.6 References 59	

3 Graphics and Image Data Representations

60

- 3.1 Graphics/Image Data Types 60
 - 3.1.1 1-Bit Images 61
 - 3.1.2 8-Bit Gray-Level Images 61
 - 3.1.3 Image Data Types 64
 - 3.1.4 24-Bit Color Images 64
 - 3.1.5 8-Bit Color Images 65
 - 3.1.6 Color Lookup Tables (LUTs) 67
- 3.2 Popular File Formats 71
 - 3.2.1 GIF 71
 - 3.2.2 JPEG 75
 - 3.2.3 PNG 76
 - 3.2.4 TIFF 77
 - 3.2.5 EXIF 77
 - 3.2.6 Graphics Animation Files 77
 - 3.2.7 PS and PDF 78
 - 3.2.8 Windows WMF 78
 - 3.2.9 Windows BMP 78
 - 3.2.10 Macintosh PAINT and PICT 78
 - 3.2.11 X Windows PPM 79
- 3.3 Further Exploration 79
- 3.4 Exercises 79
- 3.5 References 81

4 Color in Image and Video

82

- 4.1 Color Science 82
 - 4.1.1 Light and Spectra 82
 - 4.1.2 Human Vision 84
 - 4.1.3 Spectral Sensitivity of the Eye 84
 - 4.1.4 Image Formation 85
 - 4.1.5 Camera Systems 86
 - 4.1.6 Gamma Correction 87
 - 4.1.7 Color-Matching Functions 89
 - 4.1.8 CIE Chromaticity Diagram 91
 - 4.1.9 Color Monitor Specifications 94
 - 4.1.10 Out-of-Gamut Colors 95
 - 4.1.11 White-Point Correction 96
 - 4.1.12 XYZ to RGB Transform 97
 - 4.1.13 Transform with Gamma Correction 97
 - 4.1.14 $L^*a^*b^*$ (CIELAB) Color Model 98
 - 4.1.15 More Color-Coordinate Schemes 100
 - 4.1.16 Munsell Color Naming System 100
- 4.2 Color Models in Images 100
 - 4.2.1 RGB Color Model for CRT Displays 100
 - 4.2.2 Subtractive Color: CMY Color Model 101
 - 4.2.3 Transformation from RGB to CMY 101
 - 4.2.4 Undercolor Removal: CMYK System 102
 - 4.2.5 Printer Gamuts 102
- 4.3 Color Models in Video 104
 - 4.3.1 Video Color Transforms 104
 - 4.3.2 YUV Color Model 104

4.3.3	YIQ Color Model	105	
4.3.4	YCbCr Color Model	107	
4.4	Further Exploration	107	
4.5	Exercises	108	
4.6	References	111	
5	Fundamental Concepts in Video		112
5.1	Types of Video Signals	112	
5.1.1	Component Video	112	
5.1.2	Composite Video	113	
5.1.3	S-Video	113	
5.2	Analog Video	113	
5.2.1	NTSC Video	116	
5.2.2	PAL Video	119	
5.2.3	SECAM Video	119	
5.3	Digital Video	119	
5.3.1	Chroma Subsampling	120	
5.3.2	CCIR Standards for Digital Video	120	
5.3.3	High Definition TV (HDTV)	122	
5.4	Further Exploration	124	
5.5	Exercises	124	
5.6	References	125	
6	Basics of Digital Audio		126
6.1	Digitization of Sound	126	
6.1.1	What Is Sound?	126	
6.1.2	Digitization	127	
6.1.3	Nyquist Theorem	128	
6.1.4	Signal-to-Noise Ratio (SNR)	131	
6.1.5	Signal-to-Quantization-Noise Ratio (SQNR)	131	
6.1.6	Linear and Nonlinear Quantization	133	
6.1.7	Audio Filtering	136	
6.1.8	Audio Quality versus Data Rate	136	
6.1.9	Synthetic Sounds	137	
6.2	MIDI: Musical Instrument Digital Interface	139	
6.2.1	MIDI Overview	139	
6.2.2	Hardware Aspects of MIDI	142	
6.2.3	Structure of MIDI Messages	143	
6.2.4	General MIDI	147	
6.2.5	MIDI-to-WAV Conversion	147	
6.3	Quantization and Transmission of Audio	147	
6.3.1	Coding of Audio	147	
6.3.2	Pulse Code Modulation	148	
6.3.3	Differential Coding of Audio	150	
6.3.4	Lossless Predictive Coding	151	
6.3.5	DPCM	154	
6.3.6	DM	157	
6.3.7	ADPCM	158	
6.4	Further Exploration	159	
6.5	Exercises	160	
6.6	References	163	

7 Lossless Compression Algorithms

167

- 7.1 Introduction 167
- 7.2 Basics of Information Theory 168
- 7.3 Run-Length Coding 171
- 7.4 Variable-Length Coding (VLC) 171
 - 7.4.1 Shannon–Fano Algorithm 171
 - 7.4.2 Huffman Coding 173
 - 7.4.3 Adaptive Huffman Coding 176
- 7.5 Dictionary-Based Coding 181
- 7.6 Arithmetic Coding 187
- 7.7 Lossless Image Compression 191
 - 7.7.1 Differential Coding of Images 191
 - 7.7.2 Lossless JPEG 193
- 7.8 Further Exploration 194
- 7.9 Exercises 195
- 7.10 References 197

8 Lossy Compression Algorithms

199

- 8.1 Introduction 199
- 8.2 Distortion Measures 199
- 8.3 The Rate-Distortion Theory 200
- 8.4 Quantization 200
 - 8.4.1 Uniform Scalar Quantization 201
 - 8.4.2 Nonuniform Scalar Quantization 204
 - 8.4.3 Vector Quantization* 206
- 8.5 Transform Coding 207
 - 8.5.1 Discrete Cosine Transform (DCT) 207
 - 8.5.2 Karhunen–Loève Transform* 220
- 8.6 Wavelet-Based Coding 222
 - 8.6.1 Introduction 222
 - 8.6.2 Continuous Wavelet Transform* 227
 - 8.6.3 Discrete Wavelet Transform* 230
- 8.7 Wavelet Packets 240
- 8.8 Embedded Zerotree of Wavelet Coefficients 241
 - 8.8.1 The Zerotree Data Structure 242
 - 8.8.2 Successive Approximation Quantization 244
 - 8.8.3 EZW Example 244
- 8.9 Set Partitioning in Hierarchical Trees (SPIHT) 247
- 8.10 Further Exploration 248
- 8.11 Exercises 249
- 8.12 References 252

9 Image Compression Standards

253

- 9.1 The JPEG Standard 253
 - 9.1.1 Main Steps in JPEG Image Compression 253
 - 9.1.2 JPEG Modes 262
 - 9.1.3 A Glance at the JPEG Bitstream 265
- 9.2 The JPEG2000 Standard 265
 - 9.2.1 Main Steps of JPEG2000 Image Compression* 267

- 9.2.2 Adapting EBCOT to JPEG2000 275
- 9.2.3 Region-of-Interest Coding 275
- 9.2.4 Comparison of JPEG and JPEG2000 Performance 277
- 9.3 The JPEG-LS Standard 277
 - 9.3.1 Prediction 280
 - 9.3.2 Context Determination 281
 - 9.3.3 Residual Coding 281
 - 9.3.4 Near-Lossless Mode 281
- 9.4 Bilevel Image Compression Standards 282
 - 9.4.1 The JBIG Standard 282
 - 9.4.2 The JBIG2 Standard 282
- 9.5 Further Exploration 284
- 9.6 Exercises 285
- 9.7 References 287

10 Basic Video Compression Techniques

288

- 10.1 Introduction to Video Compression 288
- 10.2 Video Compression Based on Motion Compensation 288
- 10.3 Search for Motion Vectors 290
 - 10.3.1 Sequential Search 290
 - 10.3.2 2D Logarithmic Search 291
 - 10.3.3 Hierarchical Search 293
- 10.4 H.261 295
 - 10.4.1 Intra-Frame (I-Frame) Coding 297
 - 10.4.2 Inter-Frame (P-Frame) Predictive Coding 297
 - 10.4.3 Quantization in H.261 297
 - 10.4.4 H.261 Encoder and Decoder 298
 - 10.4.5 A Glance at the H.261 Video Bitstream Syntax 301
- 10.5 H.263 303
 - 10.5.1 Motion Compensation in H.263 304
 - 10.5.2 Optional H.263 Coding Modes 305
 - 10.5.3 H.263+ and H.263++ 307
- 10.6 Further Exploration 308
- 10.7 Exercises 309
- 10.8 References 310

11 MPEG Video Coding I — MPEG-1 and 2

312

- 11.1 Overview 312
- 11.2 MPEG-1 312
 - 11.2.1 Motion Compensation in MPEG-1 313
 - 11.2.2 Other Major Differences from H.261 315
 - 11.2.3 MPEG-1 Video Bitstream 318
- 11.3 MPEG-2 319
 - 11.3.1 Supporting Interlaced Video 320
 - 11.3.2 MPEG-2 Scalabilities 323
 - 11.3.3 Other Major Differences from MPEG-1 329
- 11.4 Further Exploration 330
- 11.5 Exercises 330
- 11.6 References 331

12 MPEG Video Coding II — MPEG-4, 7, and Beyond

332

- 12.1 Overview of MPEG-4 332
- 12.2 Object-Based Visual Coding in MPEG-4 335
 - 12.2.1 VOP-Based Coding vs. Frame-Based Coding 335
 - 12.2.2 Motion Compensation 337
 - 12.2.3 Texture Coding 341
 - 12.2.4 Shape Coding 343
 - 12.2.5 Static Texture Coding 346
 - 12.2.6 Sprite Coding 347
 - 12.2.7 Global Motion Compensation 348
- 12.3 Synthetic Object Coding in MPEG-4 349
 - 12.3.1 2D Mesh Object Coding 349
 - 12.3.2 3D Model-based Coding 354
- 12.4 MPEG-4 Object types, Profiles and Levels 356
- 12.5 MPEG-4 Part10/H.264 357
 - 12.5.1 Core Features 358
 - 12.5.2 Baseline Profile Features 360
 - 12.5.3 Main Profile Features 360
 - 12.5.4 Extended Profile Features 361
- 12.6 MPEG-7 361
 - 12.6.1 Descriptor (D) 363
 - 12.6.2 Description Scheme (DS) 365
 - 12.6.3 Description Definition Language (DDL) 368
- 12.7 MPEG-21 369
- 12.8 Further Exploration 370
- 12.9 Exercises 370
- 12.10 References 371

13 Basic Audio Compression Techniques

374

- 13.1 ADPCM in Speech Coding 374
 - 13.1.1 ADPCM 374
- 13.2 G.726 ADPCM 376
- 13.3 Vocoders 378
 - 13.3.1 Phase Insensitivity 378
 - 13.3.2 Channel Vocoder 378
 - 13.3.3 Formant Vocoder 380
 - 13.3.4 Linear Predictive Coding 380
 - 13.3.5 CELP 383
 - 13.3.6 Hybrid Excitation Vocoders* 389
- 13.4 Further Exploration 392
- 13.5 Exercises 392
- 13.6 References 393

14 MPEG Audio Compression

395

- 14.1 Psychoacoustics 395
 - 14.1.1 Equal-Loudness Relations 396
 - 14.1.2 Frequency Masking 398
 - 14.1.3 Temporal Masking 403
- 14.2 MPEG Audio 405
 - 14.2.1 MPEG Layers 405

- 14.2.2 MPEG Audio Strategy 406
- 14.2.3 MPEG Audio Compression Algorithm 407
- 14.2.4 MPEG-2 AAC (Advanced Audio Coding) 412
- 14.2.5 MPEG-4 Audio 414
- 14.3 Other Commercial Audio Codecs 415
- 14.4 The Future: MPEG-7 and MPEG-21 415
- 14.5 Further Exploration 416
- 14.6 Exercises 416
- 14.7 References 417

III Multimedia Communication and Retrieval

419

15 Computer and Multimedia Networks

421

- 15.1 Basics of Computer and Multimedia Networks 421
 - 15.1.1 OSI Network Layers 421
 - 15.1.2 TCP/IP Protocols 422
- 15.2 Multiplexing Technologies 425
 - 15.2.1 Basics of Multiplexing 425
 - 15.2.2 Integrated Services Digital Network (ISDN) 427
 - 15.2.3 Synchronous Optical Network (SONET) 428
 - 15.2.4 Asymmetric Digital Subscriber Line (ADSL) 429
- 15.3 LAN and WAN 430
 - 15.3.1 Local Area Networks (LANs) 431
 - 15.3.2 Wide Area Networks (WANs) 434
 - 15.3.3 Asynchronous Transfer Mode (ATM) 435
 - 15.3.4 Gigabit and 10-Gigabit Ethernet 438
- 15.4 Access Networks 439
- 15.5 Common Peripheral Interfaces 441
- 15.6 Further Exploration 441
- 15.7 Exercises 442
- 15.8 References 442

16 Multimedia Network Communications and Applications

443

- 16.1 Quality of Multimedia Data Transmission 443
 - 16.1.1 Quality of Service (QoS) 443
 - 16.1.2 QoS for IP Protocols 446
 - 16.1.3 Prioritized Delivery 447
- 16.2 Multimedia over IP 447
 - 16.2.1 IP-Multicast 447
 - 16.2.2 RTP (Real-time Transport Protocol) 449
 - 16.2.3 Real Time Control Protocol (RTCP) 451
 - 16.2.4 Resource ReSerVation Protocol (RSVP) 451
 - 16.2.5 Real-Time Streaming Protocol (RTSP) 453
 - 16.2.6 Internet Telephony 455
- 16.3 Multimedia over ATM Networks 459
 - 16.3.1 Video Bitrates over ATM 459
 - 16.3.2 ATM Adaptation Layer (AAL) 460
 - 16.3.3 MPEG-2 Convergence to ATM 461
 - 16.3.4 Multicast over ATM 462

- 16.4 Transport of MPEG-4 462
 - 16.4.1 DMIF in MPEG-4 462
 - 16.4.2 MPEG-4 over IP 463
- 16.5 Media-on-Demand (MOD) 464
 - 16.5.1 Interactive TV (ITV) and Set-Top Box (STB) 464
 - 16.5.2 Broadcast Schemes for Video-on-Demand 465
 - 16.5.3 Buffer Management 472
- 16.6 Further Exploration 475
- 16.7 Exercises 476
- 16.8 References 477

17 Wireless Networks

479

- 17.1 Wireless Networks 479
 - 17.1.1 Analog Wireless Networks 480
 - 17.1.2 Digital Wireless Networks 481
 - 17.1.3 TDMA and GSM 481
 - 17.1.4 Spread Spectrum and CDMA 483
 - 17.1.5 Analysis of CDMA 486
 - 17.1.6 3G Digital Wireless Networks 488
 - 17.1.7 Wireless LAN (WLAN) 492
- 17.2 Radio Propagation Models 493
 - 17.2.1 Multipath Fading 494
 - 17.2.2 Path Loss 496
- 17.3 Multimedia over Wireless Networks 496
 - 17.3.1 Synchronization Loss 497
 - 17.3.2 Error Resilient Entropy Coding 499
 - 17.3.3 Error Concealment 501
 - 17.3.4 Forward Error Correction (FEC) 503
 - 17.3.5 Trends in Wireless Interactive Multimedia 506
- 17.4 Further Exploration 508
- 17.5 Exercises 508
- 17.6 References 510

18 Content-Based Retrieval in Digital Libraries

511

- 18.1 How Should We Retrieve Images? 511
- 18.2 C-BIRD — A Case Study 513
 - 18.2.1 C-BIRD GUI 514
 - 18.2.2 Color Histogram 514
 - 18.2.3 Color Density 516
 - 18.2.4 Color Layout 516
 - 18.2.5 Texture Layout 517
 - 18.2.6 Search by Illumination Invariance 519
 - 18.2.7 Search by Object Model 520
- 18.3 Synopsis of Current Image Search Systems 533
 - 18.3.1 QBIC 535
 - 18.3.2 UC Santa Barbara Search Engines 536
 - 18.3.3 Berkeley Digital Library Project 536
 - 18.3.4 Chabot 536
 - 18.3.5 Blobworld 537
 - 18.3.6 Columbia University Image Seekers 537

- 18.3.7 Informedia 537
- 18.3.8 MetaSEEk 537
- 18.3.9 Photobook and FourEyes 538
- 18.3.10 MARS 538
- 18.3.11 Virage 538
- 18.3.12 Viper 538
- 18.3.13 Visual RetrievalWare 538
- 18.4 Relevance Feedback 539
 - 18.4.1 MARS 539
 - 18.4.2 iFind 541
- 18.5 Quantifying Results 541
- 18.6 Querying on Videos 542
- 18.7 Querying on Other Formats 544
- 18.8 Outlook for Content-Based Retrieval 544
- 18.9 Further Exploration 545
- 18.10 Exercises 546
- 18.11 References 547