

Contents

PREFACE	xxi
GLOSSARY OF NOTATIONS	xxv
1 VIDEO FORMATION, PERCEPTION, AND REPRESENTATION	1
1.1 Color Perception and Specification 2	
1.1.1 <i>Light and Color</i> , 2	
1.1.2 <i>Human Perception of Color</i> , 3	
1.1.3 <i>The Trichromatic Theory of Color Mixture</i> , 4	
1.1.4 <i>Color Specification by Tristimulus Values</i> , 5	
1.1.5 <i>Color Specification by Luminance and Chrominance Attributes</i> , 6	
1.2 Video Capture and Display 7	
1.2.1 <i>Principles of Color Video Imaging</i> , 7	
1.2.2 <i>Video Cameras</i> , 8	
1.2.3 <i>Video Display</i> , 10	
1.2.4 <i>Composite versus Component Video</i> , 11	
1.2.5 <i>Gamma Correction</i> , 11	
1.3 Analog Video Raster 12	
1.3.1 <i>Progressive and Interlaced Scan</i> , 12	
1.3.2 <i>Characterization of a Video Raster</i> , 14	

1.4	Analog Color Television Systems	16
1.4.1	<i>Spatial and Temporal Resolution</i>	16
1.4.2	<i>Color Coordinate</i>	17
1.4.3	<i>Signal Bandwidth</i>	19
1.4.4	<i>Multiplexing of Luminance, Chrominance, and Audio</i>	19
1.4.5	<i>Analog Video Recording</i>	21
1.5	Digital Video	22
1.5.1	<i>Notation</i>	22
1.5.2	<i>ITU-R BT.601 Digital Video</i>	23
1.5.3	<i>Other Digital Video Formats and Applications</i>	26
1.5.4	<i>Digital Video Recording</i>	28
1.5.5	<i>Video Quality Measure</i>	28
1.6	Summary	30
1.7	Problems	31
1.8	Bibliography	32
2	FOURIER ANALYSIS OF VIDEO SIGNALS AND FREQUENCY RESPONSE OF THE HUMAN VISUAL SYSTEM	33
2.1	Multidimensional Continuous-Space Signals and Systems	33
2.2	Multidimensional Discrete-Space Signals and Systems	36
2.3	Frequency Domain Characterization of Video Signals	38
2.3.1	<i>Spatial and Temporal Frequencies</i>	38
2.3.2	<i>Temporal Frequencies Caused by Linear Motion</i>	40
2.4	Frequency Response of the Human Visual System	42
2.4.1	<i>Temporal Frequency Response and Flicker Perception</i>	43
2.4.2	<i>Spatial Frequency Response</i>	45
2.4.3	<i>Spatiotemporal Frequency Response</i>	46
2.4.4	<i>Smooth Pursuit Eye Movement</i>	48
2.5	Summary	50
2.6	Problems	51
2.7	Bibliography	52
3	VIDEO SAMPLING	53
3.1	Basics of the Lattice Theory	54
3.2	Sampling over Lattices	59
3.2.1	<i>Sampling Process and Sampled-Space Fourier Transform</i>	60
3.2.2	<i>The Generalized Nyquist Sampling Theorem</i>	61
3.2.3	<i>Sampling Efficiency</i>	63

Contents

xi

- 3.2.4 *Implementation of the Prefilter and Reconstruction Filter, 65*
- 3.2.5 *Relation between Fourier Transforms over Continuous, Discrete, and Sampled Spaces, 66*
- 3.3 Sampling of Video Signals 67
 - 3.3.1 *Required Sampling Rates, 67*
 - 3.3.2 *Sampling Video in Two Dimensions: Progressive versus Interlaced Scans, 69*
 - 3.3.3 *Sampling a Raster Scan: BT.601 Format Revisited, 71*
 - 3.3.4 *Sampling Video in Three Dimensions, 72*
 - 3.3.5 *Spatial and Temporal Aliasing, 73*
- 3.4 Filtering Operations in Cameras and Display Devices 76
 - 3.4.1 *Camera Apertures, 76*
 - 3.4.2 *Display Apertures, 79*
- 3.5 Summary 80
- 3.6 Problems 80
- 3.7 Bibliography 83

4 VIDEO SAMPLING RATE CONVERSION

84

- 4.1 Conversion of Signals Sampled on Different Lattices 84
 - 4.1.1 *Up-Conversion, 85*
 - 4.1.2 *Down-Conversion, 87*
 - 4.1.3 *Conversion between Arbitrary Lattices, 89*
 - 4.1.4 *Filter Implementation and Design, and Other Interpolation Approaches, 91*
- 4.2 Sampling Rate Conversion of Video Signals 92
 - 4.2.1 *Deinterlacing, 93*
 - 4.2.2 *Conversion between PAL and NTSC Signals, 98*
 - 4.2.3 *Motion-Adaptive Interpolation, 104*
- 4.3 Summary 105
- 4.4 Problems 106
- 4.5 Bibliography 109

5 VIDEO MODELING

111

- 5.1 Camera Model 112
 - 5.1.1 *Pinhole Model, 112*
 - 5.1.2 *CAHV Model, 114*
 - 5.1.3 *Camera Motions, 116*
- 5.2 Illumination Model 116
 - 5.2.1 *Diffuse and Specular Reflection, 116*

5.2.2	<i>Radiance Distribution under Differing Illumination and Reflection Conditions, 117</i>	
5.2.3	<i>Changes in the Image Function Due to Object Motion, 119</i>	
5.3	Object Model 120	
5.3.1	<i>Shape Model, 121</i>	
5.3.2	<i>Motion Model, 122</i>	
5.4	Scene Model 125	
5.5	Two-Dimensional Motion Models 128	
5.5.1	<i>Definition and Notation, 128</i>	
5.5.2	<i>Two-Dimensional Motion Models Corresponding to Typical Camera Motions, 130</i>	
5.5.3	<i>Two-Dimensional Motion Corresponding to Three-Dimensional Rigid Motion, 133</i>	
5.5.4	<i>Approximations of Projective Mapping, 136</i>	
5.6	Summary 137	
5.7	Problems 138	
5.8	Bibliography 139	
6	TWO-DIMENSIONAL MOTION ESTIMATION	141
6.1	Optical Flow 142	
6.1.1	<i>Two-Dimensional Motion versus Optical Flow, 142</i>	
6.1.2	<i>Optical Flow Equation and Ambiguity in Motion Estimation, 143</i>	
6.2	General Methodologies 145	
6.2.1	<i>Motion Representation, 146</i>	
6.2.2	<i>Motion Estimation Criteria, 147</i>	
6.2.3	<i>Optimization Methods, 151</i>	
6.3	Pixel-Based Motion Estimation 152	
6.3.1	<i>Regularization Using the Motion Smoothness Constraint, 153</i>	
6.3.2	<i>Using a Multipoint Neighborhood, 153</i>	
6.3.3	<i>Pel-Recursive Methods, 154</i>	
6.4	Block-Matching Algorithm 154	
6.4.1	<i>The Exhaustive Block-Matching Algorithm, 155</i>	
6.4.2	<i>Fractional Accuracy Search, 157</i>	
6.4.3	<i>Fast Algorithms, 159</i>	
6.4.4	<i>Imposing Motion Smoothness Constraints, 161</i>	
6.4.5	<i>Phase Correlation Method, 162</i>	
6.4.6	<i>Binary Feature Matching, 163</i>	
6.5	Deformable Block-Matching Algorithms 165	
6.5.1	<i>Node-Based Motion Representation, 166</i>	
6.5.2	<i>Motion Estimation Using the Node-Based Model, 167</i>	

Contents

xiii

6.6	Mesh-Based Motion Estimation	169
	6.6.1 <i>Mesh-Based Motion Representation</i> ,	171
	6.6.2 <i>Motion Estimation Using the Mesh-Based Model</i> ,	173
6.7	Global Motion Estimation	177
	6.7.1 <i>Robust Estimators</i> ,	177
	6.7.2 <i>Direct Estimation</i> ,	178
	6.7.3 <i>Indirect Estimation</i> ,	178
6.8	Region-Based Motion Estimation	179
	6.8.1 <i>Motion-Based Region Segmentation</i> ,	180
	6.8.2 <i>Joint Region Segmentation and Motion Estimation</i> ,	181
6.9	Multiresolution Motion Estimation	182
	6.9.1 <i>General Formulation</i> ,	182
	6.9.2 <i>Hierarchical Block Matching Algorithm</i> ,	184
6.10	Application of Motion Estimation in Video Coding	187
6.11	Summary	188
6.12	Problems	189
6.13	Bibliography	191
7	THREE-DIMENSIONAL MOTION ESTIMATION	194
7.1	Feature-Based Motion Estimation	195
	7.1.1 <i>Objects of Known Shape under Orthographic Projection</i> ,	195
	7.1.2 <i>Objects of Known Shape under Perspective Projection</i> ,	196
	7.1.3 <i>Planar Objects</i> ,	197
	7.1.4 <i>Objects of Unknown Shape Using the Epipolar Line</i> ,	198
7.2	Direct Motion Estimation	203
	7.2.1 <i>Image Signal Models and Motion</i> ,	204
	7.2.2 <i>Objects of Known Shape</i> ,	206
	7.2.3 <i>Planar Objects</i> ,	207
	7.2.4 <i>Robust Estimation</i> ,	209
7.3	Iterative Motion Estimation	212
7.4	Summary	213
7.5	Problems	214
7.6	Bibliography	215
8	FOUNDATIONS OF VIDEO CODING	217
8.1	Overview of Coding Systems	218
	8.1.1 <i>General Framework</i> ,	218
	8.1.2 <i>Categorization of Video Coding Schemes</i> ,	219

8.2	Basic Notions in Probability and Information Theory 221	
	8.2.1 <i>Characterization of Stationary Sources, 221</i>	
	8.2.2 <i>Entropy and Mutual Information for Discrete Sources, 222</i>	
	8.2.3 <i>Entropy and Mutual Information for Continuous Sources, 226</i>	
8.3	Information Theory for Source Coding 227	
	8.3.1 <i>Bound for Lossless Coding, 227</i>	
	8.3.2 <i>Bound for Lossy Coding, 229</i>	
	8.3.3 <i>Rate-Distortion Bounds for Gaussian Sources, 232</i>	
8.4	Binary Encoding 234	
	8.4.1 <i>Huffman Coding, 235</i>	
	8.4.2 <i>Arithmetic Coding, 238</i>	
8.5	Scalar Quantization 241	
	8.5.1 <i>Fundamentals, 241</i>	
	8.5.2 <i>Uniform Quantization, 243</i>	
	8.5.3 <i>Optimal Scalar Quantizer, 244</i>	
8.6	Vector Quantization 248	
	8.6.1 <i>Fundamentals, 248</i>	
	8.6.2 <i>Lattice Vector Quantizer, 251</i>	
	8.6.3 <i>Optimal Vector Quantizer, 253</i>	
	8.6.4 <i>Entropy-Constrained Optimal Quantizer Design, 255</i>	
8.7	Summary 257	
8.8	Problems 259	
8.9	Bibliography 261	
9	WAVEFORM-BASED VIDEO CODING	263
9.1	Block-Based Transform Coding 263	
	9.1.1 <i>Overview, 264</i>	
	9.1.2 <i>One-Dimensional Unitary Transform, 266</i>	
	9.1.3 <i>Two-Dimensional Unitary Transform, 269</i>	
	9.1.4 <i>The Discrete Cosine Transform, 271</i>	
	9.1.5 <i>Bit Allocation and Transform Coding Gain, 273</i>	
	9.1.6 <i>Optimal Transform Design and the KLT, 279</i>	
	9.1.7 <i>DCT-Based Image Coders and the JPEG Standard, 281</i>	
	9.1.8 <i>Vector Transform Coding, 284</i>	
9.2	Predictive Coding 285	
	9.2.1 <i>Overview, 285</i>	
	9.2.2 <i>Optimal Predictor Design and Predictive Coding Gain, 286</i>	
	9.2.3 <i>Spatial-Domain Linear Prediction, 290</i>	
	9.2.4 <i>Motion-Compensated Temporal Prediction, 291</i>	

Contents

xv

- 9.3 Video Coding Using Temporal Prediction and Transform Coding 293
 - 9.3.1 *Block-Based Hybrid Video Coding, 293*
 - 9.3.2 *Overlapped Block Motion Compensation, 296*
 - 9.3.3 *Coding Parameter Selection, 299*
 - 9.3.4 *Rate Control, 302*
 - 9.3.5 *Loop Filtering, 305*
- 9.4 Summary 308
- 9.5 Problems 309
- 9.6 Bibliography 311

10 CONTENT-DEPENDENT VIDEO CODING

314

- 10.1 Two-Dimensional Shape Coding 314
 - 10.1.1 *Bitmap Coding, 315*
 - 10.1.2 *Contour Coding, 318*
 - 10.1.3 *Evaluation Criteria for Shape Coding Efficiency, 323*
- 10.2 Texture Coding for Arbitrarily Shaped Regions 324
 - 10.2.1 *Texture Extrapolation, 324*
 - 10.2.2 *Direct Texture Coding, 325*
- 10.3 Joint Shape and Texture Coding 326
- 10.4 Region-Based Video Coding 327
- 10.5 Object-Based Video Coding 328
 - 10.5.1 *Source Model F2D, 330*
 - 10.5.2 *Source Models R3D and F3D, 332*
- 10.6 Knowledge-Based Video Coding 336
- 10.7 Semantic Video Coding 338
- 10.8 Layered Coding System 339
- 10.9 Summary 342
- 10.10 Problems 343
- 10.11 Bibliography 344

11 SCALABLE VIDEO CODING

349

- 11.1 Basic Modes of Scalability 350
 - 11.1.1 *Quality Scalability, 350*
 - 11.1.2 *Spatial Scalability, 353*
 - 11.1.3 *Temporal Scalability, 356*
 - 11.1.4 *Frequency Scalability, 356*

- 11.1.5 *Combination of Basic Schemes, 357*
- 11.1.6 *Fine-Granularity Scalability, 357*
- 11.2 Object-Based Scalability 359
- 11.3 Wavelet-Transform-Based Coding 361
 - 11.3.1 *Wavelet Coding of Still Images, 363*
 - 11.3.2 *Wavelet Coding of Video, 367*
- 11.4 Summary 370
- 11.5 Problems 370
- 11.6 Bibliography 371

12 STEREO AND MULTIVIEW SEQUENCE PROCESSING 374

- 12.1 Depth Perception 375
 - 12.1.1 *Binocular Cues—Stereopsis, 375*
 - 12.1.2 *Visual Sensitivity Thresholds for Depth Perception, 375*
- 12.2 Stereo Imaging Principle 377
 - 12.2.1 *Arbitrary Camera Configuration, 377*
 - 12.2.2 *Parallel Camera Configuration, 379*
 - 12.2.3 *Converging Camera Configuration, 381*
 - 12.2.4 *Epipolar Geometry, 383*
- 12.3 Disparity Estimation 385
 - 12.3.1 *Constraints on Disparity Distribution, 386*
 - 12.3.2 *Models for the Disparity Function, 387*
 - 12.3.3 *Block-Based Approach, 388*
 - 12.3.4 *Two-Dimensional Mesh-Based Approach, 388*
 - 12.3.5 *Intra-Line Edge Matching Using Dynamic Programming, 391*
 - 12.3.6 *Joint Structure and Motion Estimation, 392*
- 12.4 Intermediate View Synthesis 393
- 12.5 Stereo Sequence Coding 396
 - 12.5.1 *Block-Based Coding and MPEG-2 Multiview Profile, 396*
 - 12.5.2 *Incomplete Three-Dimensional Representation of Multiview Sequences, 398*
 - 12.5.3 *Mixed-Resolution Coding, 398*
 - 12.5.4 *Three-Dimensional Object-Based Coding, 399*
 - 12.5.5 *Three-Dimensional Model-Based Coding, 400*
- 12.6 Summary 400
- 12.7 Problems 402
- 12.8 Bibliography 403

13 VIDEO COMPRESSION STANDARDS 405

- 13.1 Standardization 406
 - 13.1.1 *Standards Organizations, 406*
 - 13.1.2 *Requirements for a Successful Standard, 409*
 - 13.1.3 *Standard Development Process, 411*
 - 13.1.4 *Applications for Modern Video Coding Standards, 412*
- 13.2 Video Telephony with H.261 and H.263 413
 - 13.2.1 *H.261 Overview, 413*
 - 13.2.2 *H.263 Highlights, 416*
 - 13.2.3 *Comparison, 420*
- 13.3 Standards for Visual Communication Systems 421
 - 13.3.1 *H.323 Multimedia Terminals, 421*
 - 13.3.2 *H.324 Multimedia Terminals, 422*
- 13.4 Consumer Video Communications with MPEG-1 423
 - 13.4.1 *Overview, 423*
 - 13.4.2 *MPEG-1 Video, 424*
- 13.5 Digital TV with MPEG-2 426
 - 13.5.1 *Systems, 426*
 - 13.5.2 *Audio, 426*
 - 13.5.3 *Video, 427*
 - 13.5.4 *Profiles, 435*
- 13.6 Coding of Audiovisual Objects with MPEG-4 437
 - 13.6.1 *Systems, 437*
 - 13.6.2 *Audio, 441*
 - 13.6.3 *Basic Video Coding, 442*
 - 13.6.4 *Object-Based Video Coding, 445*
 - 13.6.5 *Still Texture Coding, 447*
 - 13.6.6 *Mesh Animation, 447*
 - 13.6.7 *Face and Body Animation, 448*
 - 13.6.8 *Profiles, 451*
 - 13.6.9 *Evaluation of Subjective Video Quality, 454*
- 13.7 Video Bit Stream Syntax 454
- 13.8 Multimedia Content Description Using MPEG-7 458
 - 13.8.1 *Overview, 458*
 - 13.8.2 *Multimedia Description Schemes, 459*
 - 13.8.3 *Visual Descriptors and Description Schemes, 461*
- 13.9 Summary 465
- 13.10 Problems 466
- 13.11 Bibliography 467

14	ERROR CONTROL IN VIDEO COMMUNICATIONS	472
14.1	Motivation and Overview of Approaches	473
14.2	Typical Video Applications and Communication Networks	476
	14.2.1 <i>Categorization of Video Applications</i> ,	476
	14.2.2 <i>Communication Networks</i> ,	479
14.3	Transport-Level Error Control	485
	14.3.1 <i>Forward Error Correction</i> ,	485
	14.3.2 <i>Error-Resilient Packetization and Multiplexing</i> ,	486
	14.3.3 <i>Delay-Constrained Retransmission</i> ,	487
	14.3.4 <i>Unequal Error Protection</i> ,	488
14.4	Error-Resilient Encoding	489
	14.4.1 <i>Error Isolation</i> ,	489
	14.4.2 <i>Robust Binary Encoding</i> ,	490
	14.4.3 <i>Error-Resilient Prediction</i> ,	492
	14.4.4 <i>Layered Coding with Unequal Error Protection</i> ,	493
	14.4.5 <i>Multiple-Description Coding</i> ,	494
	14.4.6 <i>Joint Source and Channel Coding</i> ,	498
14.5	Decoder Error Concealment	498
	14.5.1 <i>Recovery of Texture Information</i> ,	500
	14.5.2 <i>Recovery of Coding Modes and Motion Vectors</i> ,	501
	14.5.3 <i>Syntax-Based Repair</i> ,	502
14.6	Encoder-Decoder Interactive Error Control	502
	14.6.1 <i>Coding-Parameter Adaptation Based on Channel Conditions</i> ,	503
	14.6.2 <i>Reference Picture Selection Based on Feedback Information</i> ,	503
	14.6.3 <i>Error Tracking Based on Feedback Information</i> ,	504
	14.6.4 <i>Retransmission without Waiting</i> ,	504
14.7	Error-Resilience Tools in H.263 and MPEG-4	505
	14.7.1 <i>Error-Resilience Tools in H.263</i> ,	505
	14.7.2 <i>Error-Resilience Tools in MPEG-4</i> ,	508
14.8	Summary	509
14.9	Problems	511
14.10	Bibliography	513
15	STREAMING VIDEO OVER THE INTERNET AND WIRELESS IP NETWORKS	519
15.1	Architecture for Video Streaming Systems	520
15.2	Video Compression	522

Contents

xix

15.3 Application-Layer QoS Control for Streaming Video 522
15.3.1 Congestion Control, 522
15.3.2 Error Control, 525

15.4 Continuous Media Distribution Services 529
15.4.1 Network Filtering, 529
15.4.2 Application-Level Multicast, 531
15.4.3 Content Replication, 532

15.5 Streaming Servers 533
15.5.1 Real-Time Operating System, 534
15.5.2 Storage System, 537

15.6 Media Synchronization 539

15.7 Protocols for Streaming Video 542
15.7.1 Transport Protocols, 543
15.7.2 Session Control Protocol: RTSP, 545

15.8 Streaming Video over Wireless IP Networks 546
15.8.1 Network-Aware Applications, 548
15.8.2 Adaptive Service, 549

15.9 Summary 554

15.10 Bibliography 555

APPENDIX A: DETERMINATION OF SPATIAL-TEMPORAL GRADIENTS 562

A.1 First- and Second-Order Gradient 562

A.2 Sobel Operator 563

A.3 Difference of Gaussian Filters 563

APPENDIX B: GRADIENT DESCENT METHODS 565

B.1 First-Order Gradient Descent Method 565

B.2 Steepest Descent Method 566

B.3 Newton's Method 566

B.4 Newton-Raphson Method 567

B.5 Bibliography 567

APPENDIX C: GLOSSARY OF ACRONYMS 568

APPENDIX D: ANSWERS TO SELECTED PROBLEMS 575

