

Contents

Preface.....	xix
List of Figures	xxix
List of Tables.....	xli
List of Algorithms	xliii
Acknowledgments.....	xlvi

CHAPTER 1 Introduction	1
1.1 Communicating pictures: the need for compression	2
1.1.1 What is compression?	2
1.1.2 Why do we need compression?.....	3
1.2 Applications and drivers	6
1.2.1 Generic drivers	6
1.2.2 Application drivers and markets	8
1.3 Requirements and trade-offs in a compression system.....	9
1.3.1 The benefits of a digital solution.....	9
1.3.2 Requirements	10
1.3.3 Trade-offs.....	10
1.4 The basics of compression.....	11
1.4.1 Still image encoding	11
1.4.2 Encoding video	12
1.4.3 Measuring visual quality.....	12
1.5 The need for standards.....	12
1.5.1 Some basic facts about standards.....	12
1.5.2 A brief history of video encoding standards	13
1.6 Summary.....	14
References.....	15
 CHAPTER 2 The Human Visual System.....	 17
2.1 Principles and theories of human vision	19
2.1.1 Theories of vision	19
2.2 Acquisition: the human eye	20
2.2.1 Retinal tissue layers	21
2.2.2 Optical processing.....	23
2.2.3 Retinal photoreceptors and their distribution.....	24
2.2.4 Visual processing in the retina	26
2.3 The visual cortex.....	28
2.3.1 Opponent processes	29
2.3.2 Biased competition.....	29

2.3.3	Adaptation processes	29
2.3.4	V1—the primary visual cortex.....	30
2.3.5	V2—the prestriate cortex	30
2.3.6	Dorsal and ventral streams.....	30
2.3.7	Extrastriate areas.....	31
2.4	Visual fields and acuity	31
2.4.1	Field of view	31
2.4.2	Acuity.....	32
2.4.3	Light, luminance, and brightness	32
2.4.4	Light level adaptation.....	35
2.5	Color processing	36
2.5.1	Opponent theories of color.....	37
2.5.2	CIE 1931 chromaticity chart.....	39
2.6	Spatial processing	39
2.6.1	Just noticeable difference, contrast, and Weber's law.....	40
2.6.2	Frequency-dependent contrast sensitivity	40
2.6.3	Multiscale edges.....	44
2.6.4	Perception of textures	44
2.6.5	Shape and object recognition.....	44
2.6.6	The importance of phase information	45
2.7	Perception of scale and depth	46
2.7.1	Size or scale	46
2.7.2	Depth cues.....	47
2.7.3	Depth cues and 3-D entertainment.....	48
2.8	Temporal and spatio-temporal response	49
2.8.1	Temporal CSF	50
2.8.2	Spatio-temporal CSF.....	50
2.8.3	Flicker and peripheral vision	50
2.9	Attention and eye movements.....	52
2.9.1	Saliency and attention	52
2.9.2	Eye movements	53
2.10	Visual masking.....	53
2.10.1	Texture masking.....	53
2.10.2	Edge masking.....	55
2.10.3	Temporal masking.....	55
2.11	Summary: a perceptual basis for image and video compression	56
2.11.1	Influential factors	56
2.11.2	What have we learnt?.....	56
	References.....	59

CHAPTER 3	Discrete-Time Analysis for Images and Video	63
3.1	Signal and picture sampling.....	64
3.1.1	The sampling theorem.....	65
3.1.2	Multidimensional sampling lattices	69
3.2	Statistics of images	70
3.2.1	Histograms and distributions	70
3.2.2	Mean values	70
3.2.3	Correlation in natural images.....	71
3.3	Filtering and transforms.....	76
3.3.1	Discrete-time linear systems	76
3.3.2	Convolution.....	76
3.3.3	Linear filters	77
3.3.4	Filter frequency response	79
3.3.5	Examples of practical filters.....	80
3.3.6	Non-linear filters	81
3.3.7	Linear transforms and the DFT.....	82
3.4	Quantization.....	84
3.4.1	Basic theory of quantization	84
3.4.2	Adaptation to signal statistics	86
3.4.3	HVS weighting.....	87
3.4.4	Vector quantization	87
3.5	Linear prediction.....	87
3.5.1	Basic feedforward linear predictive coding	87
3.5.2	Linear prediction with the predictor in the feedback loop.....	90
3.5.3	Wiener Hopf equations and the Wiener filter.....	92
3.6	Information and entropy	93
3.6.1	Self-information.....	93
3.6.2	Entropy.....	94
3.6.3	Symbols and statistics	96
3.7	Summary.....	97
	References.....	97
CHAPTER 4	Digital Picture Formats and Representations	99
4.1	Pixels, blocks, and pictures.....	100
4.1.1	Pixels, samples, or pels	100
4.1.2	Moving pictures	102
4.1.3	Coding units and macroblocks.....	103
4.1.4	Picture types and groups of pictures	104

4.2	Formats and aspect ratios.....	106
4.2.1	Aspect ratios.....	106
4.2.2	Displaying different formats	110
4.3	Picture scanning.....	111
4.3.1	Interlaced vs progressive scanning	111
4.3.2	Standards conversion	112
4.4	Gamma correction	114
4.5	Color spaces and color transformations.....	116
4.5.1	Color descriptions and the HVS	116
4.5.2	Sub-sampled color spaces	121
4.5.3	Bayer filtering	123
4.6	Measuring and comparing picture quality	124
4.6.1	Compression ratio and bit rate	124
4.6.2	Objective distortion and quality metrics	125
4.6.3	Subjective assessment.....	128
4.7	Rates and distortions.....	129
4.7.1	Rate–distortion characteristics	129
4.7.2	Rate–distortion optimization.....	131
4.8	Summary	132
	References.....	132

CHAPTER 5 Transforms for Image and Video Coding 133

5.1	The principles of decorrelating transforms.....	134
5.1.1	The basic building blocks	134
5.1.2	Principal components and axis rotation	135
5.2	Unitary transforms	137
5.2.1	Basis functions and linear combinations.....	137
5.2.2	Orthogonality and normalization	138
5.2.3	Extension to 2-D	139
5.3	Basic transforms	140
5.3.1	The Haar transform	140
5.3.2	The Walsh–Hadamard transform	141
5.3.3	So why not use the discrete Fourier transform?.....	144
5.4	Optimum transforms.....	146
5.4.1	Discarding coefficients.....	146
5.4.2	The Karhunen–Loeve transform (KLT)	147
5.4.3	The KLT in practice	148
5.5	Discrete cosine transform (DCT).....	149
5.5.1	Derivation of the DCT	149
5.5.2	DCT basis functions.....	151
5.5.3	Extension to 2-D: separability	153

5.6	Quantization of DCT coefficients	155
5.7	Performance comparisons	160
5.7.1	DCT vs DFT revisited	160
5.7.2	Comparison of transforms	160
5.7.3	Rate-distortion performance of the DCT	161
5.8	DCT implementation	163
5.8.1	Choice of transform block size	163
5.8.2	DCT complexity reduction	164
5.8.3	Field vs frame encoding for interlaced sequences	166
5.8.4	Integer transforms	167
5.8.5	DCT demo	167
5.9	JPEG	167
5.10	Summary	169
	References	169

CHAPTER 6 Filter Banks and Wavelet Compression 171

6.1	Introduction to multiscale processing	172
6.1.1	The short-time Fourier transform and the Gabor transform	173
6.1.2	So what is a wavelet?	174
6.1.3	Wavelet and filter bank properties	177
6.2	Perfect reconstruction filter banks	178
6.2.1	Filter and decomposition requirements	178
6.2.2	The 1-D filter bank structure	178
6.3	Multirate filtering	182
6.3.1	Upsampling	183
6.3.2	Downsampling	183
6.3.3	System transfer function	186
6.3.4	Perfect reconstruction	189
6.3.5	Spectral effects of the two-channel decomposition	191
6.4	Useful filters and filter banks	191
6.4.1	Quadrature mirror filters	192
6.4.2	Wavelet filters	194
6.4.3	Multistage (multiscale) decompositions	197
6.4.4	Separability and extension to 2-D	200
6.4.5	Finite length sequences, edge artifacts, and boundary extension	202
6.4.6	Wavelet compression performance	203
6.5	Coefficient quantization and bit allocation	204

6.5.1	Bit allocation and zonal coding.....	204
6.5.2	Hierarchical coding.....	207
6.6	JPEG2000	207
6.6.1	Overview	207
6.6.2	Architecture—bit planes and scalable coding.....	207
6.6.3	Coding performance.....	208
6.6.4	Region of interest coding	208
6.6.5	Benefits and status.....	210
6.7	Summary	210
	References.....	210

CHAPTER 7 Lossless Compression Methods 213

7.1	Motivation for lossless image compression.....	214
7.1.1	Applications	215
7.1.2	Approaches	216
7.1.3	Dictionary methods	216
7.2	Symbol encoding	217
7.2.1	A generic model for lossless compression.....	217
7.2.2	Entropy, efficiency, and redundancy	217
7.2.3	Prefix codes and unique decodability	218
7.3	Huffman coding	219
7.3.1	The basic algorithm.....	219
7.3.2	Minimum variance Huffman coding.....	221
7.3.3	Huffman decoding.....	223
7.3.4	Modified Huffman coding.....	224
7.3.5	Properties of Huffman codes.....	224
7.3.6	Adaptive methods for Huffman encoding.....	225
7.4	Symbol formation and encoding.....	229
7.4.1	Dealing with sparse matrices	229
7.4.2	Symbol encoding in JPEG	230
7.4.3	JPEG performance examples	235
7.4.4	Lossless JPEG mode	235
7.4.5	Context-Adaptive Variable Length Coding (CAVLC) in H.264/AVC.....	236
7.5	Golomb coding	239
7.5.1	Unary codes	239
7.5.2	Golomb and Golomb–Rice codes	239
7.5.3	Exponential Golomb codes	240
7.6	Arithmetic coding	241
7.6.1	The basic arithmetic encoding algorithm.....	241

7.6.2	The basic arithmetic decoding algorithm.....	244
7.6.3	Advantages of arithmetic coding	246
7.6.4	Binary arithmetic coding.....	247
7.6.5	Tag generation with scaling	249
7.6.6	Context-adaptive arithmetic coding	250
7.7	Performance comparisons—Huffman vs arithmetic coding.....	251
7.8	Summary.....	252
	References.....	252

CHAPTER 8 Coding Moving Pictures: Motion Prediction..... 255

8.1	Temporal correlation and exploiting temporal redundancy	256
8.1.1	Why motion estimation?	256
8.1.2	Projected motion and apparent motion	258
8.1.3	Understanding temporal correlation	259
8.1.4	How to form the prediction	260
8.1.5	Approaches to motion estimation	262
8.2	Motion models and motion estimation	263
8.2.1	Problem formulation	263
8.2.2	Affine and high order models	263
8.2.3	Translation-only models	265
8.2.4	Pixel-recursive methods	266
8.2.5	Frequency domain motion estimation using phase correlation (PhC).....	267
8.3	Block matching motion estimation (BMME)	269
8.3.1	Translational block matching.....	269
8.3.2	Matching criteria.....	270
8.3.3	Full search algorithm	272
8.3.4	Properties of block motion fields and error surfaces	272
8.3.5	Motion failure	276
8.3.6	Restricted and unrestricted vectors	277
8.4	Reduced complexity motion estimation	277
8.4.1	Pixel grids and search grids	277
8.4.2	Complexity of full search.....	278
8.4.3	Reducing search complexity	279
8.4.4	2-D logarithmic search (TDL)	280
8.4.5	<i>N</i> -Step Search (NSS).....	281
8.4.6	Diamond search	282
8.4.7	Hexagonal search	284
8.4.8	Initialization and termination criteria.....	285

8.4.9	Reduced complexity block matching algorithm performance comparisons.....	286
8.5	Motion vector coding.....	287
8.5.1	Motion vector prediction.....	287
8.5.2	Entropy coding of motion vectors.....	287
8.6	Summary.....	288
	References.....	288
CHAPTER 9 The Block-Based Hybrid Video Codec		291
9.1	The block-based hybrid model for video compression.....	292
9.1.1	Picture types and prediction modes	292
9.1.2	Properties of the DFD signal.....	293
9.1.3	Operation of the video encoding loop.....	294
9.1.4	Operation of the video decoder.....	295
9.2	Intra-frame prediction.....	298
9.2.1	Intra-prediction for small luminance blocks	298
9.2.2	Intra-prediction for larger blocks	300
9.3	Sub-pixel motion estimation.....	301
9.3.1	Sub-pixel matching	301
9.3.2	Interpolation methods	303
9.3.3	Performance	305
9.3.4	Interpolation-free methods.....	306
9.4	Multiple reference frame motion estimation	307
9.4.1	Justification	307
9.4.2	Properties, complexity, and performance of MRF-ME	308
9.4.3	Reduced complexity MRF-ME.....	308
9.5	Variable block sizes for motion estimation.....	309
9.5.1	Influence of block size	309
9.5.2	Variable block sizes in practice.....	310
9.6	Variable sized transforms.....	310
9.6.1	Integer transforms	310
9.6.2	DC coefficient transforms	313
9.7	In-loop deblocking operations	313
9.8	Summary.....	315
	References.....	315
CHAPTER 10 Measuring and Managing Picture Quality.....		317
10.1	General considerations and influences.....	318
10.1.1	What do we want to assess?.....	319
10.1.2	Influences on perceived quality.....	319

10.2	Subjective testing	321
10.2.1	Justification	321
10.2.2	Test sequences and conditions	321
10.2.3	Choosing subjects	323
10.2.4	Testing environment	324
10.2.5	Testing methodology and recording of results	324
10.2.6	Statistical analysis and significance testing	328
10.2.7	The complete evaluation process	329
10.3	Test data sets and how to use them	330
10.3.1	Databases	330
10.3.2	The relationship between Mean Opinion Score and an objective metric	331
10.3.3	Evaluating metrics using public (or private) databases..	332
10.4	Objective quality metrics	333
10.4.1	Why do we need quality metrics?	333
10.4.2	A characterization of PSNR	335
10.4.3	A perceptual basis for metric development	337
10.4.4	Perception-based image and video quality metrics	338
10.4.5	The future of metrics	341
10.5	Rate–distortion optimization	343
10.5.1	Classical rate–distortion theory	343
10.5.2	Practical rate–distortion optimization	345
10.5.3	The influence of additional coding modes and parameters	348
10.5.4	From Rate–Distortion to Rate–Quality Optimization....	350
10.6	Rate control	350
10.6.1	Buffering and HRD	352
10.6.2	Rate control in practice	354
10.6.3	Regions of interest and rate control	357
10.7	Summary	357
	References	358

CHAPTER 11 Communicating Pictures: Delivery Across Networks 361

11.1	The operating environment	362
11.1.1	Characteristics of modern networks	362
11.1.2	Transmission types	364
11.1.3	Operating constraints	365
11.1.4	Error characteristics	365
11.1.5	The challenges and a solution framework	367
11.2	The effects of loss	368

11.2.1	Synchronization failure	368
11.2.2	Header loss	370
11.2.3	Spatial error propagation	370
11.2.4	Temporal error propagation	372
11.3	Mitigating the effect of bitstream errors	373
11.3.1	Video is not the same as data!	373
11.3.2	Error-resilient solutions	374
11.4	Transport layer solutions	375
11.4.1	Automatic Repeat request (ARQ)	375
11.4.2	FEC channel coding	376
11.4.3	Hybrid ARQ (HARQ)	379
11.4.4	Packetization strategies	379
11.5	Application layer solutions	379
11.5.1	Network abstraction	379
11.5.2	The influence of frame type	380
11.5.3	Synchronization codewords	382
11.5.4	Reversible VLC	382
11.5.5	Slice structuring	384
11.5.6	Error tracking	385
11.6	Cross layer solutions	386
11.6.1	Link adaptation	386
11.7	Inherently robust coding strategies	389
11.7.1	Error-Resilient Entropy Coding (EREC)	389
11.7.2	Pyramid Vector Quantization (PVQ)	394
11.8	Error concealment	397
11.8.1	Detecting missing information	397
11.8.2	Spatial error concealment (SEC)	398
11.8.3	Temporal error concealment	399
11.8.4	Hybrid methods with mode selection	403
11.9	Congestion management and scalable video coding	405
11.10	Summary	408
	References	408

CHAPTER 12 Video Coding Standards 411

12.1	The need for and role of standards	412
12.1.1	The focus of video standardization	412
12.1.2	The standardization process	413
12.1.3	Intellectual property and licensing	414
12.2	H.120	414
12.2.1	Brief history	414
12.2.2	Primary features	415

12.3	H.261	415
12.3.1	Brief history	415
12.3.2	Picture types and primary features.....	415
12.4	MPEG-2/DVB.....	417
12.4.1	Brief history	417
12.4.2	Picture types and primary features.....	417
12.4.3	MPEG-2 profiles and levels	418
12.5	H.263	419
12.5.1	Brief history	419
12.5.2	Picture types and primary features.....	419
12.6	MPEG-4.....	423
12.6.1	Brief history	423
12.6.2	Picture types and primary features.....	424
12.7	H.264/AVC	426
12.7.1	Brief history	426
12.7.2	Primary features	426
12.7.3	Network abstraction and bitstream syntax	427
12.7.4	Pictures and partitions.....	429
12.7.5	The video coding layer.....	430
12.7.6	Profiles and levels	434
12.7.7	Performance	435
12.7.8	Scalable extensions	435
12.7.9	Multiview extensions	435
12.8	H.265/HEVC	436
12.8.1	Brief background	436
12.8.2	Primary features	436
12.8.3	Network abstraction and high level syntax	437
12.8.4	Pictures and partitions.....	438
12.8.5	The video coding layer (VCL).....	440
12.8.6	Profiles and levels	446
12.8.7	Extensions	446
12.8.8	Performance gains for HEVC over recent standards	446
12.9	Other de-facto standards and proprietary codecs.....	447
12.9.1	VP9	447
12.9.2	VC-1	447
12.9.3	RealVideo.....	447
12.9.4	Dirac.....	447
12.10	Summary	447
	References.....	448

CHAPTER 13 Communicating Pictures—The Future	451
13.1 The motivation: more immersive experiences	452
13.2 Emerging formats and extended video parameter spaces	452
13.2.1 Influences	453
13.2.2 Spatial resolution	453
13.2.3 Temporal resolution	455
13.2.4 Dynamic range	459
13.2.5 Parameter interactions and the creative continuum	461
13.3 Challenges for compression	462
13.4 Parametric video compression	464
13.5 Context-based video compression	466
13.6 Summary	467
References	468
 Appendix A: Glossary of Terms	 471
 Appendix B: Tutorial Problems	 479
 INDEX	 499