

# List of figures

Fig. 1.1	A geometric interpretation of compression.	3
Fig. 1.2	The multimedia communications jigsaw puzzle.	7
Fig. 1.3	Simplified high level video compression architecture.	11
Fig. 1.4	The scope of standardization.	13
Fig. 1.5	A chronology of video coding standards from 1990 to the present date.	14
Fig. 2.1	The visible spectrum.	20
Fig. 2.2	Cross-section of the human eye. (Public domain: <a href="http://commons.wikimedia.org/wiki/File:Schematic_diagram_of_the_human_eye_en.svg">http://commons.wikimedia.org/wiki/File:Schematic_diagram_of_the_human_eye_en.svg</a> .)	21
Fig. 2.3	Fundus image of a healthy retina. (Public domain from: <a href="http://commons.wikimedia.org/wiki/File:Fundus_photograph_of_normal_right_eye.jpg">http://commons.wikimedia.org/wiki/File:Fundus_photograph_of_normal_right_eye.jpg</a> .)	22
Fig. 2.4	The focal length of the lens.	23
Fig. 2.5	Photoreceptor distribution in the retina. (Reproduced with permission from: Mustafia et al. [12].)	24
Fig. 2.6	Normalized rod and cone responses for the human visual system. (Reproduced with permission from: Bowmaker and Dartnall [33]. <a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1279132/">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1279132/</a> . Avail Wikimedia Commons.)	25
Fig. 2.7	Retinal cell architecture (Public domain image adapted from <a href="http://commons.wikimedia.org/wiki/File:Retina-diagram.svg">http://commons.wikimedia.org/wiki/File:Retina-diagram.svg</a> ).	27
Fig. 2.8	Spatial opponency, showing center-surround on cell and its firing pattern due to excitation.	28
Fig. 2.9	The visual cortex.	28
Fig. 2.10	Mach band effect.	33
Fig. 2.11	Adelson's grid. (Reproduced with permission from: <a href="http://web.mit.edu/persci/people/adelson/checkershadow_illusion.html">http://web.mit.edu/persci/people/adelson/checkershadow_illusion.html</a> .)	33
Fig. 2.12	CIE luminous efficiency curve. (Public domain image: <a href="http://en.wikipedia.org/wiki/File:CIE_1931_Luminosity.png">http://en.wikipedia.org/wiki/File:CIE_1931_Luminosity.png</a> .)	34
Fig. 2.13	Dark adaptation of rods and cones.	35
Fig. 2.14	Increased immersion from color images.	36
Fig. 2.15	Opponent processing of color.	38

Fig. 2.16	Color dependence on context. The bottom picture is just a squashed version of the top one yet the green stripes become blue.	38
Fig. 2.17	The CIE 1931 chromaticity chart. (Reproduced with permission from Ref. [21].)	39
Fig. 2.18	Just-noticeable differences at different contrast increments.	40
Fig. 2.19	JND curve for human vision.	41
Fig. 2.20	Contrast sensitivity chart.	41
Fig. 2.21	Luminance and chrominance CSF responses.	42
Fig. 2.22	Luminance contrast sensitivity function.	43
Fig. 2.23	Texture change blindness. (Images courtesy of Tom Troscianko.)	45
Fig. 2.24	The importance of phase information in visual perception. Left: original. Right: phase distorted version using the complex wavelet transform (Reproduced with permission from Vilankar et al. [14]).	46
Fig. 2.25	Perspective-based depth cues can be very compelling and misleading.	47
Fig. 2.26	Pits and bumps—deceptive depth from lighting.	48
Fig. 2.27	The hollow mask illusion.	49
Fig. 2.28	Spatio-temporal CSF. (Adapted from Kelly [18].)	51
Fig. 2.29	Variation of critical flicker frequency (Reproduced with permission from Tyler [23]).	52
Fig. 2.30	Eye movements in response to task (Public domain image from: <a href="http://commons.wikimedia.org/wiki/File:Yarbus_The_Visitor.jpg">http://commons.wikimedia.org/wiki/File:Yarbus_The_Visitor.jpg</a> ).	54
Fig. 2.31	Example of texture masking.	54
Fig. 2.32	Edge masking for high and low dynamic range content.	55
Fig. 2.33	Temporal masking effects for various edge step sizes. (Reproduced with permission from Girod [30].)	55
Fig. 3.1	Spectral characteristics of sampling and aliasing.	66
Fig. 3.2	Demonstration of aliasing for a 1-D signal. Top: sinusoid sampled below Nyquist frequency. Bottom: Fourier plot showing spectral aliasing.	67
Fig. 3.3	2-D spectral characteristics of sampling and aliasing. Left: Top—original signal spectrum; Bottom—sampled signal spectrum with no aliasing. Right: Top—original signal spectrum; Bottom—sampled signal spectrum with aliasing due to sub-Nyquist sampling.	68
Fig. 3.4	Hexagonal sampling lattice and its reciprocal as defined by equation (3.8).	69
Fig. 3.5	Example image histogram for $256 \times 256$ image Stampe_SV4.	70

Fig. 3.6	Autocorrelation plots for Acer image ( $512 \times 512$ ). Top left to bottom right: original image; autocorrelation function for row 100; autocorrelation function for whole image; 2-D autocorrelation surface.	72
Fig. 3.7	Autocorrelation plots for Stampe_SV4 image ( $512 \times 512$ ). Top left to bottom right: original image; autocorrelation function for row 100; autocorrelation function for whole image; 2-D autocorrelation surface.	73
Fig. 3.8	Autocorrelation plots for Stampe_SV4 image ( $256 \times 256$ ). Top left to bottom right: original image; autocorrelation function for row 100; autocorrelation function for whole image; 2-D autocorrelation surface.	74
Fig. 3.9	Temporal autocorrelation plots for Foreman (30 fps). Top to bottom right: sample frame showing selected $16 \times 16$ block used; temporal correlation for a single pixel; temporal correlation for a $16 \times 16$ block.	75
Fig. 3.10	Filterbank responses for the LeGall low-pass and high-pass analysis filters.	80
Fig. 3.11	Filter response for H.264 half-pixel interpolation filter.	81
Fig. 3.12	Common uniform quantizer characteristics.	85
Fig. 3.13	Common non-uniform quantizers. Left: center deadzone. Right: Lloyd Max quantizer.	86
Fig. 3.14	Feedforward linear prediction. Top: encoder. Bottom: decoder.	88
Fig. 3.15	Prediction signal dynamic range. Top: input signal. Bottom left: distribution of 1000 samples of input signal. Bottom right: distribution of 1000 samples of prediction residual.	89
Fig. 3.16	Feedback-based linear prediction.	90
Fig. 3.17	Feedback-based linear predictor with quantization noise modeling.	91
Fig. 3.18	Self-information and probability. Left: plot of self-information vs probability for a single event. Right: plot of the self-information of an event weighted by its probability.	95
Fig. 4.1	Image sample array.	100
Fig. 4.2	Image samples.	101
Fig. 4.3	Pixelation at varying resolutions. Top left to bottom right: $256 \times 256$ ; $64 \times 64$ ; $32 \times 32$ ; $16 \times 16$ .	102
Fig. 4.4	Typical macroblock structure.	103
Fig. 4.5	Typical group of pictures structure and prediction modes.	105
Fig. 4.6	Aspect ratios of common formats, normalized according to resolution.	107
Fig. 4.7	Widescreen formats.	108
Fig. 4.8	Variation of field of view with viewing distance (aspect ratio = 16:9 here).	109
Fig. 4.9	Interlaced vs progressive frame scanning.	111

Fig. 4.10	Example of effects of interlaced scanning with poor deinterlacing.	112
Fig. 4.11	3:2 pull-down example.	113
Fig. 4.12	Gamma curves for $\gamma = \{0.45, 1, 2.2\}$ .	114
Fig. 4.13	Examples of gamma correction.	115
Fig. 4.14	CIE 1931 chromaticity diagram. (Public domain image from: <a href="http://commons.wikimedia.org/wiki/File:CIEy1931_Rec_2020_and_Rec_709.svg">http://commons.wikimedia.org/wiki/File:CIEy1931_Rec_2020_and_Rec_709.svg</a> .)	118
Fig. 4.15	Color difference images using the $YC_bC_r$ decomposition. Top left: original. Right: luma channel. Bottom left: $C_b$ channel. Right: $C_r$ channel.	120
Fig. 4.16	Common chroma sub-sampling formats.	121
Fig. 4.17	Bayer filter mosaic.	124
Fig. 4.18	Quality comparisons for the same PSNR (16.5 dB). Top left to bottom right: original; AWGN (variance 0.24); grid lines; salt and pepper noise; spatial shift by 5 pixels vertically and horizontally.	127
Fig. 4.19	Rate-distortion plot for various coding parameter choices.	130
Fig. 4.20	Rate-distortion plot showing quantizer controlled operating points.	130
Fig. 4.21	Example rate-quality comparison curves for typical standard definition entertainment content.	131
Fig. 5.1	Typical transform-based image compression architecture.	135
Fig. 5.2	Plot of correlated adjacent pixel data (left) and decorrelation through rotation of principal axes (right).	136
Fig. 5.3	Relationship between original and transformed data.	136
Fig. 5.4	DWHT basis functions for $N = 4$ .	142
Fig. 5.5	Basis functions for the 2-D DWHT.	143
Fig. 5.6	Spectral leakage due to windowing and periodic extension with the DFT: the rectangular window function and its magnitude spectrum (top); the input signal showing periodic extension and its spectrum (bottom).	145
Fig. 5.7	Symmetrical signal extension for the DCT.	150
Fig. 5.8	DCT basis functions for $N = 8$ .	153
Fig. 5.9	2-D DCT basis functions for $N = 8$ .	154
Fig. 5.10	Separable computation of the forward DCT.	154
Fig. 5.11	Effects of coefficient quantization on various types of data block.	159
Fig. 5.12	Comparison of DFT and DCT for compression. Top: input signal $x[n]$ and the symmetrically extended sequence $x_1[n]$ . Middle: eight-point DFT $\{x[n]\}$ (magnitude only) and the eight-point DCT $\{x[n]\}$ . Bottom: IDFT $\{X(k)\}$ (magnitude only) and IDCT $\{C(k)\}$ .	160
Fig. 5.13	Comparison of KLT, DFT, DWHT, DST, and DCT on a natural image.	161
Fig. 5.14	Comparison of KLT and DCT performance.	162

Fig. 5.15	DCT performance. Left: Lena original $256 \times 256$ . Right: compressed at 0.3 bpp.	162
Fig. 5.16	DCT performance. Left: Lena $256 \times 256$ 0.3 bpp. Right: Lena $512 \times 512$ 0.3 bpp.	163
Fig. 5.17	DCT rate–distortion performance for $256 \times 256$ Lena image using Huffman encoding.	163
Fig. 5.18	Efficient DCT implementation: McGovern’s algorithm.	166
Fig. 5.19	DCT demo GUI.	168
Fig. 5.20	JPEG baseline system diagram.	168
Fig. 6.1	Equivalences between subband filtering and wavelet nomenclature.	173
Fig. 6.2	Comparison of time domain (top right), frequency domain (bottom left), and time–frequency analysis (bottom right) applied to the combination of a sinusoidal signal and an impulse function (top left). (Example adapted from Vetterli and Kovacevic [12].)	174
Fig. 6.3	Wavelet translation and dilation.	175
Fig. 6.4	Two-channel perfect reconstruction filter bank.	176
Fig. 6.5	Complementary half-band filters from equations (6.4) (solid) and (6.5) (dotted).	180
Fig. 6.6	Basic system diagram for the two-channel filter bank.	181
Fig. 6.7	The two-channel filter bank.	181
Fig. 6.8	Upsampling by a factor $L$ .	183
Fig. 6.9	Spectral effect of upsampling: (a) input; (b) output.	184
Fig. 6.10	Two-channel synthesis filter bank.	184
Fig. 6.11	Spectral effects of upsampling where $L = 2$ .	185
Fig. 6.12	Downsampling by a factor of $L$ .	185
Fig. 6.13	Spectral effects of downsampling: (a) input; (b) output.	187
Fig. 6.14	Two-channel analysis filter bank.	187
Fig. 6.15	Spectral effects of downsampling with $M = 2$ .	188
Fig. 6.16	Spectral relationships within the two-channel perfect reconstruction filter bank. The spectral characteristics at signal points $A$ – $H$ are shown.	190
Fig. 6.17	QMF filter pair frequency response.	194
Fig. 6.18	LeGall filter frequency responses.	196
Fig. 6.19	1-D three stage wavelet decomposition. Top: block diagram. Bottom: associated time–frequency tiling.	197
Fig. 6.20	Time–frequency tiling for progressive stages of decomposition in a 1-D wavelet filter bank: (a) input of eight samples; (b) after the first stage; (c) after the second stage; (d) after the third stage.	198
Fig. 6.21	1-D three-stage wavelet decomposition—modified downsampling structure.	199
Fig. 6.22	2-D two-channel, two-stage wavelet decomposition. Data transposition (not shown) is required between row and column processing.	200

Fig. 6.23	2-D wavelet—frequency plane tiling.	200
Fig. 6.24	Frame from Akiyo sequence (left) wavelet transformed (right) using a three-stage decomposition with LeGall 5/3 filters.	201
Fig. 6.25	Illustration of boundary extension.	201
Fig. 6.26	Compression performance, DCT vs wavelet, for the $256 \times 256$ luma Lena image.	204
Fig. 6.27	Comparisons of wavelet-coded images, $256 \times 256$ Lena. Left: LeGall 5/3 wavelets at 0.3 bpp. Right: LeGall 5/3 wavelets at 0.5 bpp.	204
Fig. 6.28	Subband weighting.	205
Fig. 6.29	Scanning of wavelet subbands.	205
Fig. 6.30	Tree structures for embedded wavelet coding.	206
Fig. 6.31	JPEG2000 encoding of <i>Puppet</i> test image ( $512 \times 512$ ). Top left to bottom right: JPEG at 64:1; JPEG2000 at 64:1; JPEG at 128:1; JPEG2000 at 128:1.	209
Fig. 6.32	JPEG2000 ROI coefficient coding using Maxshift.	209
Fig. 7.1	Symbol-code mappings for the 26 letters from the English alphabet in Morse code.	215
Fig. 7.2	Generic model of lossless coding.	217
Fig. 7.3	Adaptive entropy coding.	226
Fig. 7.4	Example of adaptive Huffman encoding tree evolution.	227
Fig. 7.5	Example of adaptive Huffman encoding showing node swapping (bottom) after weight updating symbol <i>a</i> by five counts.	228
Fig. 7.6	Zig-zag scanning prior to variable length coding.	229
Fig. 7.7	JPEG baseline image encoder architecture.	230
Fig. 7.8	Distribution of <i>run/size</i> values in a typical JPEG encoding.	231
Fig. 7.9	JPEG results for various compression ratios for $512 \times 512$ <i>maple</i> image. Top left: original. Top right: 0.16 bpp. Middle left: 0.22 bpp. Middle right: 0.5 bpp. Bottom left: 1 bpp. Bottom right: 2 bpp.	236
Fig. 7.10	JPEG results for various compression ratios for $512 \times 512$ <i>puppet</i> image. Top left: original. Top right: 0.16 bpp. Middle left: 0.22 bpp. Middle right: 0.5 bpp. Bottom left: 1 bpp. Bottom right: 2 bpp.	237
Fig. 7.11	Prediction samples for lossless JPEG.	238
Fig. 7.12	A graphical illustration of the arithmetic encoding and decoding processes. Top: probability intervals for three iterations. Bottom: codeword intervals. Shows $P(a) = 0.25$ and $P(b) = 0.75$ with example sequences <i>abb</i> and <i>bbb</i> . Adapted from Ref. [20].	247
Fig. 7.13	Context-Adaptive Binary Arithmetic Coding. Adapted from Ref. [17].	251
Fig. 7.14	Huffman vs arithmetic coding.	252
Fig. 8.1	Variation in the number of bits per frame and frame coding mode.	257

Fig. 8.2	Temporal correlation between two adjacent video frames.	259
Fig. 8.3	Frame differencing (left) vs motion-compensated prediction (right).	260
Fig. 8.4	Probability distributions for an original frame and FD and DFD frames from the <i>Football</i> sequence.	261
Fig. 8.5	Motion-based image deformations.	264
Fig. 8.6	Mesh-based warping.	265
Fig. 8.7	Phase correlation-based motion estimation.	267
Fig. 8.8	Example of phase correlation-based motion estimation. Top: input frames of a tarmac road surface taken with an approximately translating camera. Bottom: phase correlation surface.	269
Fig. 8.9	Translational block matching motion estimation (BMME).	270
Fig. 8.10	Effect of matching function: SAD vs SSD.	271
Fig. 8.11	Autocorrelation values for typical block motion fields ( <i>Foreman</i> at 25 fps). (Adapted from Ref. [5].)	272
Fig. 8.12	The effect of block size on motion estimation performance.	273
Fig. 8.13	The influence of search range on prediction accuracy and coding bit rate.	274
Fig. 8.14	Typical motion residual error surface for a $16 \times 16$ block from the <i>Foreman</i> sequence (CIF at 25 fps, $\pm 16 \times \pm 16$ search window). Top left: search window. Top right: current block.	275
Fig. 8.15	Motion vector pdfs for <i>Akiyo</i> , <i>Coastguard</i> and <i>Football</i> sequences.	276
Fig. 8.16	Example of motion model failure for <i>Football</i> sequence. Top left: current frame, right: reference frame. Bottom left: full-pixel estimation, right: half-pixel estimation.	277
Fig. 8.17	The pixel grid and the search grid.	278
Fig. 8.18	2-D logarithmic search patterns.	280
Fig. 8.19	<i>N</i> -Step Search patterns.	281
Fig. 8.20	Diamond search patterns.	283
Fig. 8.21	Hexagonal search patterns.	284
Fig. 8.22	Predicted motion vectors for initialization.	285
Fig. 8.23	Fast motion estimation performance comparison for the CIF Bream sequence at 25 fps: 4PMV+0 vs ADZS-ER vs FS.	286
Fig. 9.1	Typical group of pictures structure.	293
Fig. 9.2	Video encoder structure (intra-frame mode). Dotted lines depict control signals relating to quantizer step size, etc.	294
Fig. 9.3	Video encoder structure (inter-frame mode).	294
Fig. 9.4	Video decoder structure.	296
Fig. 9.5	Intra-prediction modes in H.264/AVC.	299
Fig. 9.6	$4 \times 4$ intra-prediction modes.	299
Fig. 9.7	Sub-pixel search with local refinement.	302
Fig. 9.8	Motion estimation search grid with half-pixel refinement.	302

Fig. 9.9	H.264 sub-pixel interpolation. (a) Half-pixel interpolation. (b) Quarter-pixel interpolation.	303
Fig. 9.10	Comparison of full, half-, and quarter-pixel motion estimation.	306
Fig. 9.11	Influence of sub-pixel motion estimation on residual error. Left: integer-pixel ME. Right: half-pixel ME.	306
Fig. 9.12	Multiple reference frame motion estimation.	308
Fig. 9.13	MRF-ME performance. Left: PSNR. Right: complexity.	309
Fig. 9.14	Variable block sizes supported by H.264/AVC. Top: $16 \times 16$ modes 1–4. Bottom: $8 \times 8$ modes 1–4.	310
Fig. 9.15	Block size distribution (Foreman sequence coded using H.264/AVC). Produced using Ref. [14].	311
Fig. 9.16	One-dimensional edge example.	314
Fig. 9.17	Illustration of the effect of a deblocking filter. Left: without deblocking. Right: with deblocking filter applied.	314
Fig. 10.1	Spatial and temporal information coverage for a range of test sequences.	323
Fig. 10.2	Generic arrangement for subjective video evaluation.	325
Fig. 10.3	A section of the DSCQS assessment form.	326
Fig. 10.4	Interface for SAMVIQ testing. Courtesy: ProVision Communications.	326
Fig. 10.5	Typical timing of a DSCQS subjective test.	327
Fig. 10.6	Reduced reference video quality assessment.	334
Fig. 10.7	Scatter plots of DMOS vs PSNR predicted DMOS for various classes of sequence from the VQEG database [42].	336
Fig. 10.8	Comparison of MOVIE, ST-MAD, and PVM on LIVE and VQEG FRTV database [42].	342
Fig. 10.9	Rate–distortion plot for various coding parameter choices. The associated convex hull provides the operational R-D characteristic for the given codec–source combination.	344
Fig. 10.10	Lagrangian optimization for a coding unit, $i$ .	347
Fig. 10.11	Managing bit rate in the context of changing source complexity.	351
Fig. 10.12	Transmission buffering at encoder and decoder.	351
Fig. 10.13	Streaming playout schedule on a picture by picture basis, corresponding to the points labeled in Figure 10.12.	352
Fig. 10.14	Example decoder buffer timing, indicating HRD parameters.	353
Fig. 10.15	Generic rate controller.	355
Fig. 11.1	Logged radio link parameter variations with channel conditions for a wireless video trial.	363
Fig. 11.2	Radiowave propagation modeling using ray tracing. (Courtesy of A. Nix.)	366
Fig. 11.3	Generic error-resilient video transmission architecture.	367
Fig. 11.4	The effect of a single bit error on a DCT/Huffman encoded Lena image.	370



Fig. 11.5	Impact of increased error rates on a DCT image codec. Left: 0.01% BER. Middle: 0.1% BER. Right: 1% BER.	371
Fig. 11.6	Temporal error propagation from a single corrupted block in frame $k - 1$ to multiple corrupted blocks in frame $k$ .	372
Fig. 11.7	Example of temporal propagation; Top: across four frames of the table tennis sequence. Bottom: across 190 frames of Foreman.	373
Fig. 11.8	Subjective appearance of loss. PVQ encoded Riverbed sequence at 30 Mb/s. Left: encoded at 30 Mb/s (intra). Right: reconstructed after 25% PER. Bottom: magnified artifact from box shown in top right image.	374
Fig. 11.9	Idealized comparison of performance.	375
Fig. 11.10	Cross packet FEC.	377
Fig. 11.11	Performance of cross packet FEC for a coding depth of 8 (left) and 32 (right) for various coding rates.	378
Fig. 11.12	Error resilience based on reference picture selection.	381
Fig. 11.13	Periodic reference frames.	382
Fig. 11.14	Example slice structures.	384
Fig. 11.15	Use of dispersed mode slice groups as a basis for error concealment.	385
Fig. 11.16	802.11g MCS switching characteristics based on throughput. (Reproduced with permission from Ref. [20]).	387
Fig. 11.17	802.11g MCS switching characteristics based on video distortion. (Reproduced with permission from Ref. [20]).	387
Fig. 11.18	MCS switching characteristic comparison: video quality based (left) vs throughput based (right). (Reproduced with permission from Ref. [20]).	388
Fig. 11.19	MCS link adaptation results. (Reproduced with permission from Ref. [20]).	388
Fig. 11.20	EREC operation.	390
Fig. 11.21	EREC performance examples of reconstructed DCT/Huffman encoded images after erroneous transmission. Left: 0.1% BER. Right: 1% BER. (Courtesy of D. Redmill.)	393
Fig. 11.22	Graph of EREC performance for increasing BER compared to conventional DCT/VLC encoding. (Courtesy of D. Redmill.)	393
Fig. 11.23	RD optimized PVQ performance for Tennis. Left: RD performance in a clean channel. Right: performance in a lossy 802.11n channel with correlated errors (Reproduced with permission from Ref. [26]).	395
Fig. 11.24	Frame shots for Tennis at 25 Mb/s and 25% BER. Top: H.264/AVC (HPS). Bottom: RD optimized PVQ (Reproduced with permission from Ref. [26]).	395
Fig. 11.25	Variability in frame distortions for PVQ and H.264/AVC (Reproduced with permission from Ref. [26]).	396

Fig. 11.26	Packet error rate vs residual bit error rate for an 802.11n channel (Reproduced with permission from Ref. [26]).	396
Fig. 11.27	Spatial error concealment using nearest neighboring pixels.	398
Fig. 11.28	Spatial error concealment applied to a lost row of macroblocks in the Foreman sequence. Left: original with error shown. Middle: result of spatial error concealment. Right: amplified difference signal. MSE = 46.6. (Courtesy of D. Agrafiotis.)	399
Fig. 11.29	Temporal error concealment based on temporal copying (top) and motion compensated prediction (bottom).	400
Fig. 11.30	Temporal error concealment results for Foreman based on temporal copying. MSE = 19.86. (Courtesy of D. Agrafiotis).	401
Fig. 11.31	Boundary Matching Error measures. Left: Boundary Matching Error (BME). Right: External Boundary Matching Error (EBME).	401
Fig. 11.32	Temporal error concealment results for <i>Foreman</i> based on motion compensated replacement. MSE = 14.61. (Courtesy of D. Agrafiotis.)	402
Fig. 11.33	Enhanced error concealment with mode selection [27].	404
Fig. 11.34	Scalable (spatial or temporal) video codec architecture.	406
Fig. 11.35	Frame types for scalable encoding.	406
Fig. 11.36	Multiple description codec architecture.	407
Fig. 12.1	A chronology of video coding standards from 1990 to the present date.	413
Fig. 12.2	The scope of standardization.	413
Fig. 12.3	H.261 macroblock, GOB, and CIF frame format.	416
Fig. 12.4	MPEG-2 encoder performance improvements between 1994 and 2004. (Adapted from an original presented by Tandberg.)	417
Fig. 12.5	MPEG-2 GOP structure.	418
Fig. 12.6	H.263 picture format (QCIF).	420
Fig. 12.7	Motion vector coding in H.263.	421
Fig. 12.8	An example of an MPEG-4 audiovisual scene.	424
Fig. 12.9	Generic representation of MPEG-4 Video Object Plane (VOP) coding.	425
Fig. 12.10	H.264/AVC layer structure.	427
Fig. 12.11	H.264 syntax.	428
Fig. 12.12	Example slice structures.	430
Fig. 12.13	Hadamard transformation of prediction residuals and coded block ordering.	431
Fig. 12.14	Variable block sizes supported by H.264/AVC. Top: $16 \times 16$ modes 1–4. Bottom: $8 \times 8$ modes 1–4.	432
Fig. 12.15	Example H.264/AVC macroblock partition for inter-coding.	432

Fig. 12.16	Hierarchical B-pictures in H.264.	433
Fig. 12.17	Selected H.264 profiles.	434
Fig. 12.18	Typical H.264/AVC (MP) performance relative to MPEG-2 for standard definition entertainment content.	435
Fig. 12.19	HEVC video encoder architecture.	437
Fig. 12.20	CTU partitioning in HEVC—an example.	439
Fig. 12.21	HEVC intra-prediction modes (mode 0 = planar and mode 1 = DC).	440
Fig. 12.22	CB partitions for inter-prediction PBs.	442
Fig. 12.23	Coefficient scanning in HEVC.	443
Fig. 12.24	Significance flag contexts for $4 \times 4$ and $8 \times 8$ HEVC transform blocks.	444
Fig. 13.1	Does increased resolution mean increased quality?	454
Fig. 13.2	Frame rate and shutter angle. The figure illustrates the effect of a ball thrown from left to right across a static scene, captured at various shutter angles.	456
Fig. 13.3	The influence of frame rate ( <i>Outdoor</i> sequence). Left: 25 fps. Right: 600 fps.	456
Fig. 13.4	Static vs dynamic resolution—the relationship between frame rate and spatial resolution (adapted from an original by Richard Salmon, BBC).	457
Fig. 13.5	Motion vector PDF comparisons for 25 fps and 600 fps.	458
Fig. 13.6	Motion vectors and residual content for high frame rate content ( <i>Mouse</i> sequence). Top: 400 fps. Bottom: 25 fps. Left: DFD signal. Right: spatial distribution of motion vectors.	458
Fig. 13.7	Intensity-dependent quantization for HDR extensions to HEVC.	460
Fig. 13.8	The creative continuum.	462
Fig. 13.9	Parametric video compression architecture (Reproduced with permission from Zhang and Bull [15]).	465
Fig. 13.10	Texture classification results for parametric video coding. Left: original frame. Right: classified regions for coding: Red—static, blue—dynamic, black—structural, or non-textured (Reproduced with permission from Ref. [15]). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this book.)	465
Fig. 13.11	Coding results from parametric video coding. Left: H.264 frame. Right: proposed method. Center: coding mode employed (from Ref. [15]).	466
Fig. 13.12	SURF-based feature matching for context-based coding of planar regions. Left: reference frame. Right: current frame. Circles represent RANSAC inliers used for planar motion modeling.	467
Fig. 13.13	Context-based video compression results. Left: H.264 coded. Right: context-based coding.	467

For colour versions of figures please refer to the electronic version or the website  
<http://booksite.elsevier.com/9780124059061/>