

Index

A

- Absolute Category Rating (ACR), 327
- Active Format Description (AFD), 110
- Additive Gaussian white noise (AGWN), 337
- Advanced Simple Profile (of MPEG-4), 14, 425
- Algorithm
 - diamond search, 283
 - HEXBS complexity, 284
 - NSS search, 281
 - TDL search, 280
 - video decoder
 - inter-mode decoder operation, 296
 - intra-mode decoder operation, 296
 - video encoding loop
 - inter-mode encoder operation, 295
 - intra-mode encoder operation, 295
- Arithmetic coding, 241
 - advantages, 246
 - binary arithmetic, 247
 - context-adaptive, 250
 - decoding algorithm, 244
 - encoding algorithm, 241
 - tag generation with scaling, 249
 - vs. Huffman coding, 251
- Aspect ratio
 - angles of retina, 106
 - field of view ratio, 109
 - of film and TV, 106–107
 - 16:9 format, 108
 - larger and closer screens, 108
 - QCIF and UHDTV formats, 106–107
 - widescreen format, 107–108
- Automatic Repeat reQuest (ARQ), 362, 375
 - advantage, 376
 - delay constrained retransmission, 376
 - problems, 376

B

- Backward prediction, 260, 292
- Bidirectional prediction, 261, 292
- Bi-predicted (B) pictures, 104, 292
- Binarization, 251
- Bit error rate (BER), 365
- Bit rates
 - and compression ratio, 124
 - management of, 351
 - requirement, 3

- Block-based hybrid video codec
 - deblocking filter operations
 - algorithm, 314
 - effect of, 314–315
 - one-dimensional edge sample, 314
 - performance of, 313
 - DFD signal, properties of, 293
 - intra-frame prediction
 - for large blocks, 300
 - for small luminance blocks, 298
 - motion estimation (*see* Motion estimation (ME))
 - operation of
 - video decoder, 295–296
 - video encoding loop, 294–295
 - picture types, 292
 - prediction modes, 292
 - tutorial problem, 491
 - variable sized transforms
 - DC coefficient transforms, 313
 - integer transforms, 310
- Block distortion measure (BDM), 270
- Block matching motion estimation (BMME), 264
 - full search algorithm, 272
 - matching criteria
 - block distortion measure, 270
 - SAD vs. SSD, 271
 - motion failure, 276–277
 - properties of
 - block motion field, 272
 - effect of block size, 272
 - motion residual error surface, 274
 - motion vector probabilities, 275
 - search grid size effect, 274
 - restricted and unrestricted vectors, 277
 - translational
 - motion vector orientation, 269–270
 - search window size, 269
- Boundary Matching Error (BME), 401

C

- Color processing, 56
 - CIE 1931 chromaticity chart, 39
 - opponent theories, 37–38
- Color spaces
 - for analog TV, 119
 - CIE definition, 117
 - for digital formats, 119–120
 - sub-sampled

- chroma sub-sampling, 121
 - common formats, 121–122
 - and compression, 123
 - luma processing, 121
 - transformations, 117, 122
 - Communicating pictures
 - compression, challenges for, 462
 - context-based video compression, 466
 - emerging formats, 452
 - extended video parameter space
 - creative continuum, 461–462
 - dynamic range, 459
 - generally, 451–452
 - influences, 453
 - parameter interactions, 461
 - spatial resolution, 453
 - temporal resolution, 455
 - immersive experiences, 452
 - overview, 467
 - parametric video compression, 464
 - rate-quality optimization, 451
 - synthesis-based coding, 465
 - texture synthesis, 464
 - visual experiences, 452
 - Complex cells, 30
 - Complexity *vs.* cost and performance, 10
 - Compression system
 - definition, 2
 - geometric interpretation of, 2–3
 - requirement of
 - bandwidth available, 4
 - picture formats and bit rate, 3
 - trade-off, 10
 - video (*see* Image and video compression)
 - Context Adaptive Binary Arithmetic Coding (CABAC), 241
 - binarization, 251
 - binary arithmetic coding, 251
 - context modeling, 251
 - H.264/AVC and HEVC standards, 250–251
 - Context adaptive variable length coding (CAVLC)
 - in H.264/AVC, 236
 - video coding layer, 433
 - Context-based video compression, 466
 - Context modeling, 251
 - Continuous wavelet transform (CWT), 175
 - Contrast sensitivity function (CSF) model, 40
 - chart, 41
 - definition, 41
 - for luminance and chrominance, 42
 - against spatial frequency, 43
 - uses of, 43
 - Creative continuum, 461–462
 - Critical flicker frequency (CFF), 51
 - Current frames, 260
 - Cyclic Redundancy Check (CRC), 375
- ## D
- Deblocking filter
 - algorithm, 314
 - effect of, 314–315
 - one-dimensional edge sample, 314
 - performance of, 313
 - Decoded picture buffer (DPB), 438
 - Degradation Category Rating method (DCR), 327
 - Delay *vs.* performance, 10
 - Diamond search (DS), 262, 282–283
 - Difference Mean Opinion Scores (DMOS),
 - 328, 335
 - vs.* PSNR predicted, 335–336
 - Difference of Gaussian functions (DoG) model, 27
 - Digital cellular network, 363
 - Digital picture formats
 - coding tree units, 104
 - color descriptions and HVS
 - Bayer filter mosaic, 123
 - chromaticity diagrams, 118
 - color spaces (*see* Color spaces)
 - trichromacy theory, 116
 - formats and aspect ratios (*see* Video format)
 - frame types
 - bi-predicted (B) pictures, 104
 - intra-coded (I) pictures, 104
 - predicted (P) pictures, 104
 - gamma correction
 - between camera capture and displays
 - content, 114
 - examples of, 115
 - gamma characteristic of device, 114–115
 - Rec.709 format, 116
 - groups of pictures (GOPs)
 - intra- *vs.* inter-frame coding, 105–106
 - typical structure of, 105
 - macroblocks
 - definition, 103
 - typical structure of, 103
 - measuring and comparing picture quality (*see* Picture quality)
 - moving pictures, 102
 - picture scanning (*see* Video scanning)
 - pixels/samples/pels
 - color image, 101
 - defined, 100
 - monochrome image, 100
 - at varying resolutions, 102

- rate distortions
 - characteristics of, 129–130
 - optimization, 131
- Dirac, 447
- Discrete cosine transform (DCT), 13, 82, 148
 - basis functions, 151, 154
 - coefficients, quantization of, 155, 159
 - complexity, 164
 - reduction, 164
 - derivation of, 149–151
 - extension to 2-D, 153–154
 - implementation
 - demonstration, 167–168
 - integer transforms, 167
 - interlaced sequences, field vs. frame encoding for, 166
 - McGovern algorithm, 164, 166
 - transform block size, 163
 - RDF of, 161–163
 - transforms comparison, 160
 - vs. DFT, 160
- Discrete Fourier transform (DFT), 82
 - and compression, 84
 - linear transforms and, 82–83
 - 2-D, 84
- Discrete-time analysis
 - filtering and transforms (*see* Image filtering and transformations)
 - images statistics
 - correlation in natural images, 71
 - histograms and distributions, 70
 - mean values, 70
 - information theory, 93
 - entropy, 94
 - self-information, 93
 - symbols and statistics, 96
 - linear predication
 - feedback-based, 90
 - feedforward analysis filter, 87
 - Wiener filter, 92
 - Wiener Hopf equations, 92–93
 - picture sampling
 - issues in, 64
 - multidimensional sampling lattices, 69
 - theorem, 65
 - quantization
 - basic theory, 84
 - HVS weighting, 87
 - non-uniform quantizers, 86
 - vector quantization, 87
 - tutorial problems, 480
- Discrete Time Fourier Series (DTFS), 83, 144
- Discrete Walsh-Hadamard transform (DWHT), 141

- basis functions, 142
 - 1-D, basis function for, 142
 - 2-D
 - basis function for, 143
 - compression, 144
- Discrete wavelet transform (DWT), 176
- Displaced frame difference (DFD) signal, 257, 260, 295
 - properties of, 293
- Double Stimulus Continuous Quality Scale (DSCQS), 129, 326–327, 330
- Double Stimulus Impairment Scale (DSIS), 327
- Dynamic range
 - generally, 459
 - high dynamic range (HDR), 459
 - human visual system (HVS), 459
 - perceptual compression methods and performance, 460

E

- Elecard Streameye analyzer, 310
- Embedded Block Coding with Optimal Truncation (EBCOT), 207
- Embedded zerotree coding (EZW), 207
 - structure for, 206
- Empiricists, 19
- End of Block (EOB) symbol, 230
- Erasure errors, 365
- Error resilience, 178
- Error resilient entropy coding (EREC)
 - operation principle, 389–390
 - decoding process, 390–391
 - encoding, 391
 - performance
 - benefits of, 393
 - decoder, 392
 - reconstructed images after, 392–393
 - vs. conventional sequential encoding, 392–393
- European Broadcasting Union (EBU), 111
- External Boundary Matching Error (EBME), 401
- Extended video parameter space
 - creative continuum, 461–462
 - dynamic range
 - generally, 459
 - high dynamic range (HDR), 459
 - human visual system (HVS), 459
 - perceptual compression methods and performance, 460
 - generally, 451–452
 - influences, 453
 - parameter interactions, 461

- spatial resolution
 - compression performance, 454
 - effective spatial resolution, 453
 - generally, 453
 - HEVC performance, 455
 - resolution vs. quality, 454
 - Ultra High Definition Television (UHDTV)
 - standard, 453–454
- Eye movements, 53–54

F

- Fast Fourier Transform (FFT), 164
- FEC channel coding, 376
 - cross packet FEC, 377–378
 - erasure codes, 377
 - rateless codes, 378
 - unequal error protection and data partitioning, 378
 - use of, 376
- Filter bank and wavelet compression
 - coefficient quantization and bit allocation
 - hierarchical coding, 207
 - zonal coding and, 204
 - multirate filtering, 182
 - downsampling, 183, 185, 187–188
 - perfect reconstruction, 189
 - system transfer function, 186
 - two-channel decomposition, spectral effects
 - of, 190–191
 - upsampling, 183–185
 - multiscale process
 - STFT and Gabor transform, 173
 - subband filtering and wavelet nomenclature, 173
 - wavelet, 174
 - perfect reconstruction
 - 1-D filter bank structure, 178
 - filter and decomposition requirement, 178
 - two-channel filter bank, development
 - of, 179
 - tutorial problem, 484
 - useful filters and
 - boundary extension, 201–202
 - characteristics of, 191
 - edge artifacts, 202
 - finite length sequences, 202
 - multistage decomposition, 197–199
 - performance, 203–204
 - quadrature mirror filters, 192
 - separability and extension to 2-D, 200
 - wavelet filters, 194

- Filter-rectify-filter (FRF) process, 44
- Finite impulse response (FIR), 77
- Fixed-length codewords (FLC), 368
- Flexible macroblock ordering (FMO), 403
- Forward Error Correction (FEC) coding, 375
- Forward prediction, 260, 292
- Fovea, 25
- Frame difference (FD), 260
 - vs. motion compensated prediction, 260
- Frame types
 - B-frames, 380
 - I-frames, 380
 - intra-refresh, 380
 - periodic reference frames, 381–382
 - P-frames, 380
 - reference picture selection, 381
- Frontal eye field (FEF), 31
- Fusiform face area (FFA), 31

G

- Gamma correction
 - between camera capture and displays content, 114
 - examples of, 115
 - gamma characteristic of device, 114–115
 - Rec.709 format, 116
- GIF (Graphics Interchange Format), 216
- Golomb coding
 - advantage, 239
 - exponential, 240
 - Golomb-Rice codes, 239
 - unary codes, 239
- Goodput, 4
- Grassman's Laws, 116
- Groups of blocks (GOBs), 415, 419
- Groups of pictures (GOPs), 292–293, 380
 - digital picture
 - Intra- vs. inter-frame coding, 105–106
 - typical structure of, 105

H

- H.120
 - conditional replenishment strategy, 415
 - history, 414
 - primary features of, 415
- H.261
 - frame format, 415
 - groups of blocks (GOBs), 415
 - history, 415
 - macroblock, 415
 - primary features of, 416

- H.263
 - advanced prediction mode, 421
 - extensions, 422
 - groups of blocks (GOBs), 419
 - history, 419
 - independent segment decoding mode, 422
 - macroblock, 419
 - motion vector coding, 421
 - overlapped motion compensation (OMC), 421
 - PB-frames mode, 422
 - picture format, 420
 - primary features of, 419
 - reference picture selection mode, 422
 - slice structured mode, 422
 - supplemental enhancement information mode, 423
 - syntax-based arithmetic coding, 421
 - temporal, SNR, and spatial scalability mode, 422
 - unrestricted motion vectors, 421
- Haar transform, 140
- H.264/AVC, 379
 - arithmetic coding, 250
 - bitstream syntax, 427
 - blocks, 429
 - block sizes supported by, 310
 - CABAC in, 250–251
 - in CAVLC, 236
 - Exp-Golomb coding, 240
 - Foreman* sequence coded using, 311
 - history, 426
 - improved coding efficiency, 426
 - instantaneous decoder refresh (IDR), 428
 - intra-prediction, 430
 - layer structure, 427
 - macroblock, 429
 - multiview extensions, 435
 - network abstraction Layer (NAL), 427–428
 - performance, 435
 - picture types, 429
 - primary features of, 426–427
 - profiles and levels, 434
 - scalable extensions, 435
 - simple syntax specification, 426
 - slices and slice groups, 429
 - syntax, 428
 - video coding layer (VCL), 428, 430
 - CABAC, 433
 - CAVLC, 433
 - coder control, 433
 - deblocking operations, 433
 - Exp-Golomb coding, 433
 - inter-coding, 431
 - intra-coding, 430
 - variable length coding, 433
- H.265/HEVC, 241, 310
 - arithmetic coding, 250
 - CABAC in, 250–251
 - clean random access (CRA) pictures, 439
 - coding tree blocks (CTBs), 439
 - coding tree units (CTUs), 439
 - decoded picture buffer (DPB), 438
 - Exp-Golomb coding, 240
 - extensions, 446
 - high level syntax, 437
 - intra-prediction, 298
 - levels, 446
 - main 10 profile, 446
 - main profile, 446
 - main still picture profile, 446
 - NAL units, 437
 - network abstraction, 437
 - overview, 436
 - performance gains for, 446
 - pictures and partitions, 438
 - prediction units (PUs), 439
 - primary features of, 436
 - quadtree CTU structure, 439
 - random access points, 439
 - reference picture set (RPS), 438
 - slice structures, 437
 - transform units (TUs), 439
 - video coding layer (VCL)
 - advanced motion vector prediction mode (AMVP), 441
 - CABAC, variable length coding with, 444
 - coefficient scanning, 443
 - contexts and significance, 444
 - deblocking filter (DBF), 445
 - generally, 440
 - in-loop filters, 445
 - inter-coding, 441
 - intra-coding, 440
 - quantization, 442
 - sample adaptive offset (SAO), 445
 - subpixel motion compensation, 441
 - transforms, 441
 - video encoder architecture, 437
 - video parameter set (VPS), 438
- Hexagon-based Search Method (HEXBS), 262, 284
- Hotelling or Eigenvector transform. *see* Karhunen-Loeve transform (KLT)
- Huffman coding
 - adaptive methods
 - elements, 226

- entropy coding, 226
 - limitations of, 225
 - tree evolution, as example, 226–228
 - basic algorithm, 219
 - decoding process, 223
 - minimum variance, 221
 - modified, 224
 - properties
 - Kraft inequality, 224
 - length of, 225
 - optimality, 225
 - tree-structured approach, 219
 - vs. arithmetic coding, 251
 - Human eye
 - cross-section of, 21
 - optical processing
 - cornea, 23
 - iris, 23
 - lens, 23
 - pupil, 23
 - retinal photoreceptors and distribution
 - cone cells, 24
 - fovea, 25
 - macula, 25
 - optic disk and nerve, 26
 - rod cells, 24
 - retinal tissue layers, 21
 - choroid, 22
 - ciliary body, 22
 - retina, 22
 - sclera, 21
 - visible spectrum, 20
 - visual processing in retina, 26
 - Human visual system (HVS), 18
 - color processing
 - CIE 1931 chromaticity chart, 39
 - opponent theories of color, 37–38
 - depth cues
 - and 3-D entertainment, 48
 - examples, 48
 - hollow mask illusion, 48–49
 - list of, 47
 - eye (*see* Human eye)
 - eye movements, 53–54
 - flicker and peripheral vision, 50
 - principles and theories, 19
 - saliency and attention, 52
 - size and scale, 46
 - spatial processing, 39–40
 - CSF model (*see* Contrast sensitivity function (CSF) model)
 - edge localization model, 44
 - JND/contrast/Weber's law, 40–41
 - phase information, 45–46
 - shape and object recognition, 44
 - textures perception, 44–45
 - spatio-temporal CSF, 50–51
 - temporal CSF, 50
 - tutorial problems, 479
 - visual cortex
 - adaptation processes, 29
 - biased competition, 29
 - dorsal and ventral streams, 30
 - extrastriate areas, 31
 - opponent coding, 29
 - prestriate cortex (V2), 30
 - primary visual cortex (V1), 30
 - structure, 28
 - visual fields and acuity
 - brightness, 34
 - field of view, 31
 - light level adaptation, 35
 - luma, 34
 - luminance, 32–34
 - radiant intensity and radiance, 33
 - visual masking (*see* Visual masking)
 - Hybrid ARQ (HARQ), 379
 - Hypercomplex cells, 30
 - Hypothetical Reference Decoder (HRD), 352
 - buffering and, 352–353
- I**
- Image and video compression
 - application drivers and markets
 - business, manufacturing and automation, 8
 - consumer video, 8
 - healthcare, 9
 - security and surveillance, 8
 - bandwidths availability, 4
 - basics of
 - encoding video, 12
 - measuring visual quality, 12
 - still image encoding system, 11
 - bit rate requirements, 3
 - digital solution benefits, 9
 - generic drivers, 6
 - human visual system (*see* Human Visual System (HVS))
 - influential factors, 56
 - picture formats (*see also* Video format), 3
 - requirements of, 10
 - standardization of, 12
 - basic facts, 12
 - chronology of, 13–14
 - transforms for (*see* Transform)

Image filtering and transformation
 convolution, 76
 discrete Fourier transform (DFT), 82–83
 and compression, 84
 2-D, 84
 discrete time linear systems
 linearity, 76
 shift invariance, 76
 frequency response, 79
 linear filters
 separability, 78
 2-D digital filters, 78
 non-linear filters, 81
 morphological filters, 82
 rank-order and median filters, 81
 practical filters examples
 H.264/AVC half-pixel interpolation filter, 80
 Le Gal wavelet analysis filters, 80
 Image statistics
 autocorrelation/autocovariance, 71–72
 spatial, 73
 for Stampe_SV4 image, 73
 temporal, 75
 histograms and distributions
 image example, 70
 spatial and subband distributions, 70
 mean values, 70
 Inferior temporal cortex (IT), 31
 Infinite impulse response (IIR), 77
 Information theory, 93
 entropy
 defined, 94
 and first order entropy, 95
 self-information, 93
 independent events, 94
 symbols and statistics, 96
 Instantaneous decoder refresh (IDR), 428
 Integer transforms, 310
 architecture for, 313
 reduced complexity, 313
 International Standards Organization (ISO), 13, 207
 International Telecommunications Union
 (ITU-T), 13
 Internet, 362
 Internet Protocol (IP), 362
 Interpolation methods
 interpolation-free methods, 306
 H.264/AVC
 half-pixel estimation, 303
 quarter-pixel estimation, 303–304
 Inter-prediction with subsequences, 381
 Intra-coded (I) pictures, 104, 292
 Intra-frame prediction

 for large blocks, 300
 for small luminance blocks, 298
 4×4 , 299–300
 in H.264/AVC, 299

J

JPEG, 11, 167
 historical bases, 230
 lossless mode
 defined, 235
 prediction modes, 238
 matrix encoding, 233
 performance of, 235–237
 QM Coder used in, 250
 run/size value distribution, 231
 symbol encoding in
 AC coefficient, 230, 232–233
 baseline image encoder
 architecture, 230
 coefficient size categories, 232, 234
 DC coefficient, 230, 232
 run/size values, 231
 JPEG2000
 aim of, 207
 benefits and status, 210
 bit planes and scalable coding
 architecture, 207
 coding performance, 208–209
 description, 207
 EBCOT (Embedded Block Coding with
 Optimized Truncation), 250
 ISO, 207
 region of interest coding, 208–209
 technical overview of, 208
 Just noticeable differences (JND), 40

K

Karhunen-Loeve transform (KLT), 147–148

L

Laboratory for Image and Video Engineering
 (LIVE) database, 331
 Lateral Geniculate Nucleus (LGN), 28
 Lempel-Ziv-Welch (LZW), 216
 Linear Correlation Coefficient (LCC), 332
 Linear predication
 feedback-based, 90
 feedforward analysis filter, 87
 predictor dynamic range, 88
 with quantization, 89

- Wiener filter, 92
- Wiener Hopf equations, 92
- Linear time-invariant (LTI) system, 76
- Lossless compression methods
 - arithmetic coding, 241
 - advantages, 246
 - binary arithmetic, 247
 - context-adaptive, 250
 - decoding algorithm, 244
 - encoding algorithm, 241
 - tag generation with scaling, 249
 - vs. Huffman coding, 251
 - Golomb coding
 - advantage, 239
 - exponential, 240
 - and Golomb-Rice codes, 239
 - unary codes, 239
 - Huffman coding
 - adaptive methods, 225–228
 - basic algorithm, 219
 - decoding process, 223
 - minimum variance, 221
 - modified, 224
 - properties of, 224
 - tree-structured approach, 219
 - vs. arithmetic coding, 251
- motivation for
 - applications, 215
 - approaches, 216
 - dictionary-based methods, 216
 - uses of, 214
- symbol encoding
 - entropy/efficiency/redundancy, 217
 - generic model for, 217
 - prefix codes and unique decodability, 218
- symbol formation and encoding
 - CAVLC in H.264/AVC, 236
 - in JPEG (*see* JPEG)
 - with sparse matrices, 229
- tutorial problem, 487
- Lossy compression
 - quantization, 155
 - vs. lossless compression, 10

M

- Mach band effect, 32–33
- Matlab® software, 167
- McGovern algorithm, 164
- Mean absolute difference (MAD), 128

- Mean Opinion Score (MOS), 128, 328
- Mean square error (MSE), 92, 125
- Medial superior temporal (MST), 31
- Medium Access Control (MAC), 363
- Modulation and coding schemes (MCS), 386–388
- Monochrome image, 100–102
- Morse code, 215
 - for English alphabet, structure of, 215
 - for MORSE CODE, 215
- Most Apparent Distortion (MAD) model, 318, 340
- Motion compensated prediction (MCP), 260
 - frame difference, 260
- Motion compensated temporal replacement (TEC_MCTR), 400–402, 404
- Motion compensation (MC), 260
- Motion estimation (ME), 260
 - approaches to, 262
 - block matching (*see* Block matching motion estimation (BMME))
 - motion models and
 - affine, bilinear and perspective model, 263
 - node-based warping, 264–265
 - pixel-recursive methods, 266
 - problem formulation, 263
 - translation model, 265
 - using phase correlation (PhC), 267
- multiple reference frame
 - performance and complexity, 308–309
 - properties, 307
 - reduced complexity, 308
 - set up of, 307
 - standardized in H.263, 307
- reduced complexity
 - vs. block matching performance, 286
 - diamond search, 282
 - full search complexity, 278
 - hexagonal search, 284–285
 - initialization criteria, 285
 - N-step search, 281
 - pixel grids and search grids, 277–278
 - reducing search complexity, 279
 - termination criteria, 285
 - 2-D logarithmic search (TDL), 280
- sub-pixel
 - interpolation-free methods, 306
 - interpolation methods, 303
 - matching, 301
 - performance, 305
- temporal correlation
 - apparent motion, 258

- prediction, formation of, 260
 - projected motion, 258
 - between two adjacent video frames, 259
 - tutorial problem, 489
 - variable block size
 - influence of, 309
 - supported by H.264/AVC, 310
- Motion tuned spatio-temporal quality assessment (MOVIE), 318, 339
- Motion vector
 - coding
 - entropy of, 287
 - prediction, 287
 - orientation, 269–270
 - predicted for initialization, 285
 - probabilities, 275–276
- Moving Pictures Experts Group (MPEG), 13, 417
- MPEG-2/DVB
 - encoder performance, 417
 - GOP structure, 418
 - history, 417
 - picture types, 417
 - primary features of, 417
 - profiles and levels, 418
- MPEG-4
 - advanced simple profile (ASP), 425
 - audiovisual scene, 424
 - coding framework, 425
 - history, 423
 - picture types, 424
 - primary features of, 424
 - video object plane (VOP) coding, 425
- Multiple description coding (MDC), 407
- Multiple input-multiple-output (MIMO) wireless technology, 408
- Multiple reference frame motion estimation (MRF-ME), 307
 - performance and complexity, 308–309
 - properties, 307
 - reduced complexity, 308
 - set up of, 307
 - standardized in H.263, 307
- Multiscale process
 - STFT and Gabor transform, 173
 - wavelet
 - comparison of time and frequency domain, 174
 - continuous wavelet transform, 175
 - discrete wavelet transform, 176
 - and filter bank properties, 177
 - theory, development of, 172–173
 - translation and dilation, 175

N

- Near-threshold strategy, 337–338
- Netflix, 113
- Network abstraction layer (NAL), 379, 428
- Network delivery
 - application layer solution
 - error tracking, 385
 - frame type, influence of (*see* Frame type)
 - network abstraction, 379
 - reversible VLC, 382
 - slice structuring, 384
 - synchronization codewords, 382
- bitstream error effects
 - error resilient solutions, 374–375
 - subjective appearance of loss, 373
- challenges and solution framework, 367
- cross layer optimization
 - definition, 386
 - link adaptation, 386
- effect of loss
 - header loss, 370
 - spatial error propagation, 370
 - synchronization failure, 368
 - temporal error propagation, 372
- error characteristics
 - encoding, types of, 366
 - error types, 365
 - test data, 365
- error concealment
 - benefits, 397
 - hybrid methods with mode selection, 403–404
 - missing information detection, 397
 - post-processing operation, 397
 - SEC, 398–399
 - temporal error concealment (*see* Temporal error concealment)
- error resilient entropy coding (*see* Error resilient entropy coding (EREC))
 - performance, 392
- error resilient, generic architecture of, 367
- modern network characteristics
 - fixed IP network, 362
 - wireless network, 363
- multiple description coding, 407
- operating constraints, 365
- pyramid vector quantization (*see* Pyramid vector quantization (PVQ))
- scalable video coding, 405–406
- transmission types
 - downloads and streaming, 364
 - interactive communication, 364

- multicast transmission, 364
- unicast transmission, 364
- transport layer solutions
 - ARQ technique (*see* Automatic Repeat reQuest (ARQ))
 - FEC channel coding (*see* FEC channel coding)
 - hybrid ARQ (HARQ), 379
 - packetization strategies, 379
- tutorial problem, 494
- Normalized cross correlation (NCCF), 262, 271
- N-Step Search (NSS), 262, 281
 - algorithm, 281
 - patterns, 281
 - three step search, example, 282

O

- Objective quality metrics, 319
 - classification
 - full-reference (FR) methods, 334
 - no-reference (NR) methods, 334
 - reduced-reference (RR) methods, 334
 - future of, 341
 - perception-based
 - MAD and STMAD, 340
 - MOVIE, 339
 - PVM, 341
 - reduced complexity metrics, 340
 - results comparison, 341–342
 - SSIM, 338
 - VDP and VDP-2, 340
 - VQM, 338
 - VSNR, 340
 - perceptual basis development, 337
 - PSNR characterization, 335
 - uses of
 - algorithm development and benchmarking, 333
 - rate-quality optimization, 333
 - streaming control, 334
- Overlapped motion compensation (OMC), 421

P

- Packet error rate (PER), 373
- Parametric video compression, 464
- Peak signal to noise ratio (PSNR), 126, 318
 - calculation of, 127
 - for color images and video, 128
 - quality comparisons for, 127
 - same PSNR value, comparisons of, 127
 - uses of, 126
- Perception-based Video Metric (PVM), 318, 341

- Perfect reconstruction
 - 1-D filter bank structure, 178
 - two-channel filter bank, development of, 179
 - filter and decomposition requirement, 178
 - multirate filtering, 189
- P-frames, 257
- Phase correlation (PhC), ME based
 - advantages of, 267
 - applications, 268
 - example of, 269
 - performance, 268
 - principles, 267
- Physical Layer (PHY), 363
- Picture grammars, 20
- Picture Parameter Set (PPS), 379
- Picture quality
 - compression ratio and bit rate, 124
 - influences on
 - artifact type, 320
 - content type, 320
 - evaluation of, 319
 - human visual perception, 319
 - viewing environment, 320
 - objective distortion and quality metrics
 - mean absolute difference, 128
 - mean square error, 125
 - peak signal to noise ratio (*see* Peak signal to noise ratio (PSNR))
 - objective quality metrics, 319
 - classification, 334
 - future of, 341
 - perception-based, 338
 - perceptual basis development basis, 337
 - PSNR characterization, 335
 - uses of, 333
- rate control
 - bit rate management, 350–351
 - buffering and HRD, 352
 - encoder and decoder, buffering at, 351
 - playout schedule, 351–352
 - in practice, 354
 - and regions of interest, 357
- rate-distortion optimization
 - coding modes and parameters, 348
 - practical, 345
 - to rate-quality optimization, 350
 - theory, 343
- subjective assessment, 128–129
- subjective test database
 - databases, 330
 - metrics evaluation using public/private, 332
 - MOS vs. objective metric, 331

- subjective testing, 319
 - consideration, 321
 - evaluation process, 329
 - methodology and recording of results, 324–325
 - statistical analysis and significance testing, 328
 - subject selection, 323
 - testing environment, 324
 - test sequences and conditions, 321
- tutorial problem, 494
- Picture sampling
 - issues in, 64
 - multidimensional sampling lattices, 69
 - theorem
 - extension to 2-D, 67
 - extension to 3-D, 67
 - one dimension, 65
- PNG (Portable Network Graphics), 216
- Predicted (P) pictures, 104, 292
- Prediction
 - intra-frame (*see* Intra-frame prediction)
 - motion (*see* Motion prediction)
- Principal Component Analysis (PCA), 147
- Principle of orthogonality, 92
- Progressive downloads, 364
- Pyramid vector quantization (PVQ)
 - operation principle
 - bit errors, resilient to, 394
 - encoding algorithm, 394
 - error propagation, prevention of, 394
 - intra-mode, 394
 - vector quantization, 394
 - performance, 394–396

Q

- Quadrature mirror filter (QMF), 172, 179, 192
 - aliasing elimination, 192
 - amplitude distortion, 192
 - practical, 193
- Quality of service (QoS), 362
- Quantization
 - of DCT coefficients, 155
 - HVS weighting, 87
 - non-uniform
 - benefits, 86
 - deadzone quantizer, 86
 - Lloyd Max quantizer, 87
 - theory of, 84
 - uniform quantization, 85
 - vector quantization, 87
- Quantizer parameter (QP), 351

R

- Random bit errors, 365
- Rate control
 - bit rate management, 350–351
 - buffering and HRD, 352
 - encoder and decoder, buffering at, 351
 - playout schedule, 351–352
 - in practice
 - buffer model, 354
 - coding unit bit allocation, 356
 - complexity estimation, 354
 - generic architecture, 354–355
 - GOP bit allocation, 356
 - QP initialization, 355
 - QP values, 354
 - rate-quantization model, 354
 - and regions of interest, 357
- Rate-distortion (RD) algorithm, 394
- Rate-distortion optimization (RDO), 130–131, 163
 - coding modes and parameters, 348
 - Lagrangian multipliers approach, 348
 - practical, 345
 - Lagrangian methods, 346–347
 - RDO complexity, 346
 - statistics to parameterizable code, 346
 - to rate-quality optimization, 350
 - rate vs. distortion, 343
 - theory, 343–344
 - distortion measures, 344
 - memoryless Gaussian source, 345
- Rate-distortion performance (RDF), 276, 367, 374
 - DCT comparisons, 161–163
- Rate vs. quality, 10
- Rationalists, 19
- Real-Time Control Protocol (RTCP), 362
- Real-Time Streaming Protocol (RTSP), 362
- Real-time Transport Protocol (RTP), 362
- RealVideo, 447
- Redundant slices (RS), 385
- Reed Solomon Erasure (RSE), 377
- Reference frames, 260
- Reference picture selection (RPS), 381
- Reference picture set (RPS), 438
- Residual bit error rate (RBER), 396
- Residual signal strength indication (RSSI), 386
- Retina, 22
 - healthy, image of, 22
 - photoreceptors distribution
 - cone cells, 24
 - fovea, 25
 - macula, 25
 - optic disk and nerve, 26
 - rod cells, 24

visual acuity of, 32
 visual processing in, 26
 Reversible VLC (RVLC), 382
 Run-length coding (RLC), 229

S

Scalability, 177
 Scalable video coding (SVC), 405–406
 Scene-grammar, 20
 Selective recovery technique, 385
 Sequence Parameter Set (SPS), 379
 Set partitioning in hierarchical trees (SPIHT)
 approach, 207
 Short-time Fourier transform (STFT), 173
 Simple cells, 30
 Single bit error
 on DCT/Huffman encoded image, effect of, 370
 loss of VLC synchronization, 368
 Single stimulus methods, 327
 Singular value decomposition (SVD), 408
 Slice structure, 384
 FMO approach, 384
 redundant slices, 385
 Soft Combining, 379
 Spatial activity (SA), 404
 Spatial error concealment (SEC), 398–399
 Spatial error propagation
 due to VLC errors, 371
 single bit error, 370
 Spatially multiplexed system, 408
 Spatial processing, 39–40
 CSF model (*see* Contrast sensitivity function (CSF) model)
 edge localization model, 44
 JND/contrast/Weber's law, 40–41
 phase information, 45–46
 shape and object recognition, 44
 textures perception, 44–45
 Spatial resolution
 compression performance, 454
 effective spatial resolution, 453
 generally, 453
 HEVC performance, 455
 resolution vs. quality, 454
 Ultra High Definition Television (UHDTV)
 standard, 453–454
 Spatio-temporal CSF, 50–51
 Spatio-temporal MAD (ST-MAD), 340
 Spearman Rank-Order Correlation Coefficient, 332
 Structural Similarity (SSIM), 338
 Subband filterbanks, 177
 Filterbanks and wavelet compression, 171
 Subjective Assessment Methodology for Video
 Quality (SAMVIQ), 324, 327
 Subjective test database
 databases, 330
 IRCCyN/IVC, 331
 LIVE, 331
 VQEG HDTV, 331
 VQEG multimedia Phase I, 331
 VQEQ FRTV, 330
 metrics evaluation using public/private
 linear correlation, 332
 outlier ratio, 333
 rank-order correlation, 332
 MOS vs. objective metric, 331
 uses of, 330
 Subjective testing, 319
 consideration, 321
 evaluation process, 329
 methodology and recording of results
 algorithm, 330
 double stimulus methods, 325
 general principles of, 324, 326
 pair comparison methods, 328
 single stimulus methods, 327
 triple stimulus methods, 328
 statistical analysis and significance
 testing, 328
 confidence interval, 328
 mean score calculation, 328
 screening of observer, 329
 subject selection, 323
 testing environment, 324
 test sequences and conditions
 activity/information levels, 322–323
 test conditions, 322
 test material, 321
 Sub-pixel motion estimation
 approaches, 302
 benefits of, 301
 concept of, 301
 half-pixel refinement, search grid with, 302
 interpolation-free methods, 306
 interpolation methods
 example, 304
 half-pixel and quarter-pixel in H.264, 303
 SAD values, 304
 with local refinement, 302
 performance
 full, half and quarter-pixel, comparison of,
 305–306
 on residual error, 306
 Sum of absolute differences (SAD), 262

- vs. sum of squared differences, 271
- Sum squared differences (SSD), 262
 - vs. sum of absolute difference, 271
- Suprathreshold strategy, 337–338
- Synchronization
 - codewords, 382
 - advantages, 382
 - disadvantages, 382
 - failure, 368
 - single bit error, 368
- T**
- Temporal activity (TA), 404
- Temporal correlation
 - apparent motion, 258
 - prediction, formation of, 260
 - projected motion, 258
 - between two adjacent video frames, 259
- Temporal CSF, 50
- Temporal Differential Pulse Code Modulation (DPCM), 13
- Temporal error concealment, 399
 - TEC_MCTR, 400–402, 404
 - temporal copying (TEC_TC), 400–401
- Temporal error propagation
 - due to prediction, 372
 - example of, 372–373
 - from single corrupted block, 372
- Temporal resolution
 - compression methods and performance, 457
 - frame rates, 455–457
 - generally, 455
 - shutter angles, 455–457
 - static vs. dynamic resolutions, 457–458
- Texture synthesis, 464
- Theories of vision, 19
- 3rd Generation Partnership Project (3GPP), 379
- 3-D imaging modalities, 9
- Transforms
 - basic
 - discrete Fourier transform, 144–145
 - Haar transform, 140
 - Walsh-Hadamard, 141
 - discrete cosine transform (DCT), 148
 - basis functions, 151
 - coefficient quantization of, 155, 159
 - comparison of transforms, 160
 - derivation of, 149
 - implementation, 163
 - rate-distortion performance of, 161–163
 - 2-D basis functions, 153–155
 - JPEG, 167

- optimum
 - discarding coefficients, 146
 - Karhunen-Loeve transform, 147–148
- performance comparisons
 - DCT vs. DFT revisited, 160
 - KLT vs. DCT, 161–162
 - on natural images, 160–161
- principles of decorrelating
 - image compression architecture, 135
 - principle components and axis rotation, 135
- tutorial problem, 482
- unitary
 - basis functions and linear combinations, 137
 - orthogonality and normalization, 138
 - two dimensional (2-D) signal, 139
- Translational block matching
 - motion vector orientation, 269–270
 - search window size, 269
- Transmission Control Protocol (TCP), 362
- Triple Stimulus Continuous Evaluation Scale
 - methodology (TSCES), 328
- Triple stimulus methods, 328
- Twitter, 112
- 2-D logarithmic search (TDL), 280

U

- Ultra High Definition Television (UHDTV)
 - standard, 4
 - generally, 453
 - compression ratio for, 6
 - parameters, 453–454
- Unitary transforms
 - basis functions and linear combinations, 137
 - orthogonality and normalization, 138
 - two dimensional (2-D) signal, 139
- UNIX compress, 216
- User Datagram Protocol (UDP), 362

V

- Variable length symbol encoding (VLC), 11
- VC1, 447
- Vector quantization (VQ), 87
- Video coding layer (VCL)
 - H.264/AVC, 428, 430
 - CABAC, 433
 - CAVLC, 433
 - coder control, 433
 - deblocking operations, 433
 - Exp-Golomb coding, 433
 - inter-coding, 431

- intra-coding, 430
- variable length coding, 433
- H.265/HEVC
 - advanced motion vector prediction mode (AMVP), 441
 - CABAC, variable length coding with, 444
 - coefficient scanning, 443
 - contexts and significance, 444
 - deblocking filter (DBF), 445
 - generally, 440
 - in-loop filters, 445
 - inter-coding, 441
 - intra-coding, 440
 - quantization, 442
 - sample adaptive offset (SAO), 445
 - subpixel motion compensation, 441
 - transforms, 441
- Video coding standards
 - chronology of, 412–413
 - dirac, 447
- H.120
 - conditional replenishment strategy, 415
 - history, 414
 - primary features of, 415
- H.261
 - frame format, 415
 - groups of blocks (GOBs), 415
 - history, 415
 - macroblock, 415
 - primary features of, 416
- H.263
 - advanced prediction mode, 421
 - extensions, 422
 - groups of blocks (GOBs), 419
 - history, 419
 - independent segment decoding mode, 422
 - macroblock, 419
 - motion vector coding, 421
 - overlapped motion compensation (OMC), 421
 - PB-frames mode, 422
 - picture format, 420
 - primary features of, 419
 - reference picture selection mode, 422
 - slice structured mode, 422
 - supplemental enhancement information mode, 423
 - syntax-based arithmetic coding, 421
 - temporal, SNR, and spatial scalability mode, 422
 - unrestricted motion vectors, 421
- H.264/AVC
 - bitstream syntax, 427
 - blocks, 429
 - history, 426
 - improved coding efficiency, 426
 - instantaneous decoder refresh (IDR), 428
 - layer structure, 427
 - macroblock, 429
 - multiview extensions, 435
 - network abstraction, 427
 - network abstraction Layer (NAL), 428
 - performance, 435
 - picture types, 429
 - primary features of, 426–427
 - profiles and levels, 434
 - scalable extensions, 435
 - simple syntax specification, 426
 - slices and slice groups, 429
 - syntax, 428
 - video coding layer (VCL), 428, 430
- H.265/HEVC
 - clean random access (CRA) pictures, 439
 - coding tree blocks (CTBs), 439
 - coding tree units (CTUs), 439
 - decoded picture buffer (DPB), 438
 - extensions, 446
 - high level syntax, 437
 - levels, 446
 - main 10 profile, 446
 - main profile, 446
 - main still picture profile, 446
 - NAL units, 437
 - network abstraction, 437
 - overview, 436
 - performance gains for, 446
 - pictures and partitions, 438
 - prediction units (PUs), 439
 - primary features of, 436
 - quadtree CTU structure, 439
 - random access points, 439
 - reference picture set (RPS), 438
 - slice structures, 437
 - transform units (TUs), 439
 - video coding layer (VCL), 440
 - video encoder architecture, 437
 - video parameter set (VPS), 438
- intellectual property and licensing, 414
- MPEG-2/DVB
 - encoder performance, 417
 - GOP structure, 418
 - history, 417
 - picture types, 417
 - primary features of, 417
 - profiles and levels, 418

- MPEG-4
 - advanced simple profile (ASP), 425
 - audiovisual scene, 424
 - coding framework, 425
 - history, 423
 - picture types, 424
 - primary features of, 424
 - video object plane (VOP) coding, 425
 - overview, 412, 447
 - RealVideo, 447
 - standard-compliant decoder, 413
 - standard-compliant encoder, 413
 - standardization
 - convergence, 414
 - divergence, 414
 - focus of, 412
 - process, 413
 - requirements definition, 414
 - verification, 414
 - VC1, 447
 - VP9, 447
 - Video encoder architecture, 437
 - Video parameter set (VPS), 438
 - VP9, 447
 - Video compression
 - algorithms, 362
 - block-based hybrid model (*see* Block-based hybrid model)
 - prediction, classes of, 260
 - Video format
 - aspect ratios
 - angles of retina, 106
 - field of view ratio, 109
 - of film and TV, 106–107
 - 16:9 format, 108
 - larger and closer screens, 108
 - QCIF and UHDTV formats, 106–107
 - widescreen format, 107–108
 - displaying different formats
 - 2.39:1 format, 110
 - 4:3 format, 110
 - pan and scan and AFD, 110
 - picture/frame types
 - B pictures, 104
 - I pictures, 104
 - P pictures, 104
 - Video Quality Experts Group (VQEG) FRTV
 - Phase 1 programme, 330
 - Video Quality Metric (VQM), 338
 - Video scanning
 - interlaced scanning
 - effects of, 112
 - vs. progressive scanning, 111
 - standards conversion
 - by motion compensation
 - method, 114
 - 3:2 pull-down, 112–113
 - Visual acuity, 32
 - Visual cortex
 - adaptation processes, 29
 - biased competition, 29
 - dorsal and ventral streams, 30
 - extrastriate areas, 31
 - opponent coding, 29
 - prestriate cortex (V2), 30
 - primary visual cortex (V1), 30
 - structure, 28
 - Visual masking, 53
 - edge masking, 55
 - temporal masking, 55–56
 - texture masking, 53–54
 - Visual short term memory (VSTM), 31
 - Visual Signal-to-Noise Ratio (VSNR), 318, 340
- ## W
- Walsh-Hadamard transform, 141
 - Wavelet filters, 194
 - Daubechies 9/7 filters, 196
 - Filterbanks and wavelet compression, 171
 - Le Gall 5/3 filters, 195–196
 - Weber's Law, 40
 - Wireless local area networks (WLANs), 363, 386
 - Wireless network, 363
- ## Z
- Zig-zag scanning, 229
 - ZIP compression, 216