

Index

A

Absolute category rating (ACR), 225
Acronyms, list of, xi
Adaptive codebook, 93
Affricative, 13
Algebraic coding, 101
Algorithmic delay, 2, 78
Analysis by synthesis, 37, 39, 131, 212
 generalised, 205
Arithmetic coding, 145
Audio coding, 2

B

Bandwidth extension, 152, 158
Bark scale, 153
Bitrate, 2

C

Chebyshev transform, 64
Closed-loop approach, 94
Code-excited linear prediction (CELP)
 algebraic (ACELP), 101
 relaxed (RCELP), 205
Coding
 fixed or constant bitrate (CBR), 36
 lossless, 144
 lossy, 36
 variable bitrate (VBR), 36, 214
Coding gain, 32
Comfort noise, 172
Comfort noise generator (CNG), 215
Comparison category rating (CCR), 227
Computational complexity, 2
Concealment, *see* packet loss and concealment

Consonant, 18

Constant annoyance model, 59
Continuous quality scale (CQS), 230
Critical sampling, 32, 78, 138

D

Degradation category rating (DCR), 226
Discontinuous transmission (DTX), 215
Discrete cosine transform (DCT), 33
Discrete Fourier transform (DFT), 33, 140

E

Eigenvalue decomposition, 33
Entropy
 bitrate, 31
 coding, 31, 132, 144
 differential, 31
 random variable, 31
Equivalent rectangular bandwidth (ERB), 153
Evaluation model, 37
Expectation maximisation (EM), 38

F

Formants, 16
Fractional delay, 93
Frame, temporal processing-unit, 87
Frequency-masking model, 59, 144
Fricative, 13
Fundamental frequency, F_0 , 4, 91

G

Gains, 117

Glottal formant, 92
Glottis, 91

H

Half-sine window, 135
Harmonics to noise ratio (HNR), 119
Huffman coding, 145

I

Impedance spectrum frequencies (ISF), 61
Independent and identically distributed (IID), 47
Intelligibility, 2
Intonation, 4, 19

K

Kaiser-Bessel-derived (KBD) window, 136
Karhunen–Loève transform, 33

L

Lag-windowing of autocorrelation, 57
Lagrange multiplier, 52
Lapped transform, 143
Levinson-Durbin recursion, 52, 61, 72
Line spectral frequencies (LSF), 61
Linear prediction, 207
 anti-causal, 23
 linear predictive coding (LPC), 49
 normal equations, 51, 52, 70, 72
 prediction gain, 53
 reflection coefficient, 61, 73
 residual, 51
 stability, 52, 93
 Yule-Walker equations, 52, 71, 72
Linguistic content, 4
Linguistics, 11
Lip radiation, 25
Listening effort, 2
Log spectral distortion (LSD), 64
Long term prediction (LTP), 92, 207
Loudness, 4

M

Manner of articulation, 12
Mean opinion score (MOS), 225
Method for objective measurements of perceived audio quality (PEAQ), 234
Minimum mean square error (MMSE), 6

Modified discrete cosine transform (MDCT), 78, 142
Multicast protocol, 181
Multiple stimuli with hidden reference and anchor (MUSHRA), 228

N

Noise filling, 129
 in packet loss and concealment, 170

O

Obstruent, 24
Obstruent phoneme, 12
Open-loop approach, 94

P

Packet-loss, 2
Packet loss and concealment
 bit-errors, 162
 concealment, 163
 delayed packet, 162
 error correction, 163
 error resilience, 163
 error-robust transmission, 163
 forward error correction, 165
 interactive techniques, 163
 interleaving, 164
 interpolation, 172
 jitter buffer, 178
 late packet processing, 178
 lost packet, 162
 media-independent, 165
 media-specific, 165
 muting, 170
 noise substitution, 170
 packet management, 178
 receiver-based, 163
 recovery, 163
 repetition, 172
 resynchronization, 176
 retransmission, 161, 181
 sender-based, 163
 silence substitution, 170
 time scaling, 178
 waveform substitution, 172
Paralinguistic content, 4
Perceptual
 domain, 37
 model, 37, 132
 in packet loss and concealment, 168
 quality, 2

- transparency, 3, 64
- weighting function, 48
- windowing, 79
- Perceptual evaluation of speech quality (PESQ), 232
- Perceptual objective listening quality assessment (POLQA), 233
- Perfect reconstruction, 32, 78, 134, 138
- Performance evaluation model, 47
- Phonation, 4, 11
- Phone, 4
- Phoneme, 4, 11, 16
- Phonetic alphabet
 - international phonetic alphabet (IPA), 16
 - speech assessment methods phonetic alphabet (SAMPA), 16
- Pitch, 4, 19
 - estimation, 208
- Pleasantness, 2
- Pre-emphasis, 56
- Pre-/postfiltering
 - bass postfiltering, 126
 - excitation shaping, 124
 - formant enhancement, 124
 - phase scrambling, 128
 - pitch sharpening, 127
 - pulse dispersion, 128
- Princen-Bradley condition, 135, 138
- Principal component analysis, 33
- Prosody, 4

Q

- Quantisation
 - direct (uniform), 41
 - uniform, 34

R

- Randomised quantisation, 129, 168
- Rate-loop, 36, 147
- Residual codebook
 - algebraic, 101
 - complexity, 98
 - flexibility, 98
 - representation, 98
 - storage, 98
 - track structure, 103

S

- Scale factor, 60
- Service quality, 220
- Short-time Fourier transform (STFT), 140

- Signal model, 47
- Sonorant, 12
- Source coding, 31
- Source model, 31
 - fundamental frequency, 91
 - in packet loss and concealment, 168
 - spectral envelope model, 47
 - tube model of the vocal tract, 49
 - vocal tract, 20
- Spectral envelope model, 47
- Speech articulation, 11
- Speech content, 4
- Speech production organs
 - glottis, 12
 - larynx, 12
 - nasal cavities, 12
 - palate, 12
 - pharynx, 12
 - teeth, 12
 - tongue, 12
 - vocal folds, 12
 - vocal tract, 12
- Stop, 13, 18
- Subframe, temporal processing-unit, 88
- Subjective test methodology for evaluating speech communication systems that include noise suppression algorithm, 228

T

- Time domain aliasing cancellation (TDAC), 139
- Time-frequency transform, 132
- Transform coded excitation (TCX), 132, 147
- Trill, 19

U

- Unicast protocol, 181
- Unvoiced phoneme, 13
- User experience, 220

V

- Voicing, 12
- Vowel, 16

W

- White noise correction, 57
- Windowing
 - autocorrelation estimation, 53
 - overlapping, 133
 - signal and perceptual model, 77

Y

Yule-Walker equations, [52](#), [71](#)

Z

Zero input response (ZIR), [77](#), [84](#)