Index

μ-Law PCM  32–3
3.7 kb/s ACELP coder  319
4 kb/s ACELP coder  332
6 kb/s ACELP coder  332–3
8 kb/s ACELP coder  333–4
4.15 kb/s IMBE coder  270–1
2.4 kb/s MELP coder  113
4 kb/s MELP/CELP coder  283, 285
1.2 kb/s SB-LPC coder  128
4 kb/s SB-LPC coder  128, 271–5, 319
4.8 kb/s STC coder  268–70
1-tap pitch filter  81–2
3-tap pitch filter  82

ACELP mode error  348, 350
ACELP transition quantization  332–4
acoustic noise, performance of hybrid coder  336, 337–45
acoustic noise, robustness of SWPM 342
active noise suppression  2
adaptive codebook  52, 53
adaptive differential pulse code modulation (ADPCM)  1, 2, 8, 12, 13
adaptive multi-rate (AMR) coder  13, 14
adaptive multi-rate (ETSI) coding standard  360, 362–3, 364–8, 374–5
adaptive normalized least mean squared (ANLMS) algorithm  416–18
adaptive post-filtering  257–8
adaptive quantizer  33–6, 36–9
adaptive sequential LMS method  100–1
adaptive transversal filter in echo cancellation  413–14
ADPCM, see adaptive differential pulse code modulation
A-Law PCM  32–3
algebraic codebook excitation  247–51
algorithm
adaptive normalized least mean squared  416–18
Burg  74
Durbin  69, 72
frequency-domain pitch determination  155–8, 158–66, 177

A
AaS, see analysis-and-synthesis
Ab-LPC coder
  block diagram  201
  CELP  212, 213, 214, 219–60
  excitation signal  200, 206–8, 208–12
  generally  199–200
  MPLPC  215–17, 218
  perceptually-based error
  minimization procedure  200, 203–6
  procedure  201–2
  RPELPC  217–19
  SELP  212–15
  time-varying filter  200, 202–3
  weighting filter  204–5
AC, see autocorrelation
ACELP mode  298

Digital Speech. A. Kondoz
algorithm (continued)
K-means 43–4
normalized least mean squared 416
pitch measurement 81
time-domain pitch determination
151–5, 158–66, 177
voice activity detector 11, 281, 311,
341, 357–75

aliasing distortion 25
all-pole digital filter 90
all-pole modelling 269
all-pole synthesis filter 205
Al-Naimi 131, 416
AMBE speech coding standard 15, 16
AMDF, see average magnitude
difference function
AMR, see adaptive multi-rate
analogue speech signal bandwidth 25
analogue telephony system 1
analysis-and-synthesis (AaS) coder 199
analysis-by-synthesis coder, see
AbS-LPC coder
ANLMS, see adaptive normalized least
mean squared
anti-alias filter 130–46
APC 78
Application Specific Integrated Circuit
(ASIC) chip 11
Atal 78, 200, 204, 206
Atal, Singhal and 235
Atkinson 299
autocorrelation PDA 152–5
autocorrelation method 68–70
autocorrelation pitch measurement
algorithm 81
average magnitude difference function
(AMDF) PDA 81, 151–2
average spectral distortion 107

binary voicing, see hard decision voicing
bit rate conflict with speech quality 9
bit rate reduction xiii
bit rate, VAD and 359
Blackman window function 59, 60, 61
block quantization, see vector
quantization
Burg’s algorithm 74
burst error 336

C
capacity increase, VAD and 358, 359
cascaded codebook 48–9, 53
Cattermole 32
CCITT regulatory body 9
CCSR, see Centre for Communication
Systems Research
CDMA, see code division multiple access
CELP 8, 12
CELP coder
codebook excitation 240–54
excitation behaviour 212, 213, 214
excitation signal 208–12
generally 206, 207, 279
LPC prediction 221–2
multi-pulse excitation 230–2, 233
operation 219–21
pitch prediction 222–8
post-filtering 257–60
SNR 228–30
see also 4 kb/s MELP/CELP
Centre for Communication Systems
Research (CCSR), University of
Surrey xiii
centre-clipped codebook 241–2, 243
centre-clipping PDA 169–72
Cepstrum pitch measurement algorithm
81
channel dependent mode decision 9
channel error 10, 54, 130, 336, 345–50
Chebyshev Series method 100
Chen 258
Cho 371
Cholesky decomposition 71, 81, 234
closed-loop mode selection of voicing
311–12, 315–18
closed-loop optimization  199, 200
closed-loop prediction  235–6
CNG, see comfort noise generator
CNI, see comfort noise insertion
co-channel interference, VAD and  357
code division multiple access (CDMA)  14, 358
code-excited linear prediction, see CELP
codebook
  adaptive  52, 53
  binary search  46–8
  cascaded  48–9, 53
  centre-clipped  241–2, 243
  comparison of codebooks  117–21
  design  40, 43–4
  design of simultaneous joint  116
  full search  44–6
  gain-shape  50–2
  generally  40–2
  multi-stage vector quantization  110, 113–16
  optimization  43–6
  overlapping  241, 242, 243
  robustness  53–4
  split vector quantization  49–50, 110, 111–12
  testing  53
  training  52, 116, 246
  training database size  117
see also vector quantization

codebook excitation
  algebraic  247–51
  CELP coder  240–54
  Gaussian  241–3
  generally  206–7, 222
  LTP and  255–7
  pitch adaptive mixed  251–4
  vector sum  243–7
codebook vector  40
coder
  AbS-LPC  199–200, 200–2
  analysis-and-synthesis  199–200
  CELP  206, 207, 279
  combined low bit-rate  282–3
  enhanced variable rate  286
  harmonic  277–9
hybrid  8–9, 280–5
hybrid encoder  298–311
improved multi-band excitation  268, 270–1
4.15 kb/s INMARSAT-M  270–1
low bit-rate  2
MBE  189
MELP  189
4 kb/s MELP/CELP  283, 285
MPLPC  230–7
multi-band excitation  264–5, 277
parametric  6, 7–8
prototype waveform interpolation  282
4.8 kb/s SB-LPC  128, 271–5
selection of  15–16
sinusoidal transform  261–75, 268–70
speech-specific  2
split-band LPC  261, 268, 271–5
4.8 kb/s STC  268–70
variable rate  9
waveform approximating  6
see also CELP coder, hybrid coder
coding harmonic speech  261–75
coding delay  10
comb filter  156
combined low bit-rate coder  282–3
comfort noise generator (CNG)  357
comfort noise insertion (CNI)  357
companded quantizer  32–3
complex root method  95
compression
  µ-Law  32
  A-Law  32–3
compression of signal  1–2
covariance method  70–1
Cox  90

D
DFT, see discrete Fourier transform
DFT–LSF method  99
differential quantizer (DQ)  36–9, 122–4
differential vector quantization  52
digital signal processor (DSP) chip  xiii, 11
digital speech interpolation (DSI) 11

digitally-encoded speech, advantages of 1

direct expansion method 101–2

direct similarity measure 153

discrete cosine transform (DCT) 380

discrete Fourier transform (DFT) 380

distortion measure 42–3, 106

DoD speech coding standard 14–15

DQ, see differential quantizer

DSI, see digital speech interpolation

DSP chip, see Digital Signal Processor chip

Durbin’s algorithm 69, 72

E

echo cancellation

adaptive transversal filter 413–14
digital echo canceller 411–13
duplex connection 411

G.165 (ITU-T) 413, 417–18
generally 406–11

near-end speech detection 413, 415

performance 415–23

residual error suppression 413, 415

transversal filter 412

echo canceller with noise suppressor 418–9

EFR, see ETSI GSM enhanced full rate

EFR coder 251

EFR weighting method 109, 110, 120, 121

encoder, hybrid 298–311

enhanced variable rate coder (EVRC) 286

enhancing speech 379–426

environmental variability in signal 53–4

Ephraim and Malah 369, 381–387

European Telecommunications Standards Institute (ETSI)

GSM enhanced full rate (EFR) speech coding standard 13, 14, 360, 361–2, 364–8

GSM full rate (FR) speech coding standard 13, 14, 360, 361–2, 364–8

GSM half rate (HR) speech coding standard 13, 14, 360, 361–2, 364–8

speech coding standard 10, 13–14

UMTS speech coding standard 360

excitation of white noise 309–11

excitation signal, determination of optimum 208–12

excitation signal in AbS-LPC coder 200, 206–8

F

FEC, see forward error correction

filter, finite impulse response 283

filter, finite length 226–8

filter memory 203

finite impulse response (FIR) filter 283

finite length filter 226–8

forward adaptation 33, 34, 37

forward error correction (FEC) 10

fractional-delay LTP 225–6

frame energy of speech 185–6

frame-to-frame interpolation 90

frequency-domain analysis 57–8

frequency-domain pitch determination 155–8, 158–66, 177–8

frequency-domain voicing 263

frequency response of LPC filter 77

frequency, sampling 25

FS-1015 speech coding standard 14, 15

FS-1016 speech coding standard 14, 15

full search codebook 44–6

G

G.165 (ITU-T) speech coding standard 413, 417–18

G.711 (ITU-T) speech coding standard 12, 13

G.721 (ITU-T) speech coding standard 12, 13

G.722 (ITU-T) speech coding standard 13

G.722.1 (ITU-T) speech coding standard 13
Index

G.723.1 (ITU-T) speech coding standard 12, 13
G.723.1 Annex A (ITU-T) speech coding standard 360, 361, 364–8
G.723.1 coder 251
G.726 (ITU-T) speech coding standard 12, 13
G.728 (ITU-T) speech coding standard 12, 13
G.729 (ITU-T) speech coding standard 12, 13
G.729 coder 251
gain in SWPM 323
gain-shape codebook 50–2
Gaussian codebook excitation 241–3
generalized cubic phase interpolation 297–8
Gibson 380
Gram–Schmidt orthogonalization process 247
Griffin 299
group delay weighting method 109–10, 121
GSM, see ETSI speech coding standard
GSM enhanced full rate (EFR) ETSI speech coding standard 13, 14, 360, 361–2, 364–8
GSM full rate (FR) ETSI speech coding standard 13, 14, 360, 361–2, 364–8
GSM half rate (HR) ETSI speech coding standard 13, 14, 360, 361–2, 364–8
GSS, see spectral subtraction, generalized

H
Hamming window 59, 60–5, 165, 190, 262
Hard-decision noise adaptation 402–3
hard-decision voicing 150, 178–89
harmonic amplitude estimation 266–8
generally 272, 299–301
quantization in AbS 327–9
in SWPM 323
harmonic band 194
harmonic coder 8
see also coder
harmonic excitation mode 298
harmonic excitation quantization amplitude 327–9
gain 329–30
onset parameter 330–1
pitch 324, 325, 326
pitch pulse location 325–7
pitch pulse shape 327
transition detection 324–5
harmonic excitation, synchronized 299–301
harmonic gain quantization in AbS 329–30
harmonic memory initialization at onset 308–9
harmonic mode error 346–7, 348
harmonic peak detection PDA 156
harmonic speech coding 261–75
harmonic voicing 264–6
hierarchical clustering, see binary search
Hilbert envelope 286
human speech emulation 7
hybrid coder
ACELP transition excitation quantization 331
burst error 336
decoder 319–20
design 280–1
encoder 298–311
generally 6–7, 8–9
harmonic excitation quantization 323–31
limitations 284–5
LPC vocoder 281
performance 320–2
performance under acoustic noise 336, 337–45
performance under channel errors 336, 345–50
quantization issues 322–31
random channel error 336
speech classification 311–19
speech quality 334–5
transition detection 315–18
hybrid coder (continued)
unvoiced excitation quantization
voice 323, 324
voicing 281
hybrid mode selection of voicing
311–12

I
improved multi-band excitation (IMBE) coder 268, 270–1
improved multi-band excitation (IMBE) speech coding standard 15, 16
impulse train generator 65
INMARSAT speech coding standard 15, 16
INMARSAT-M coder 270–1
inter-frame correlation 121–30
interpolation
generalized cubic phase 297–8
generally 281
linear 262
filter 227
technique overlap and add 262
inverse filtering (LPC) 76, 77
inverse sine function 88, 90, 92
IS-54 (TIA/EIA) speech coding standard 14, 15
IS-96 (TIA/EIA) speech coding standard 360, 363–4, 364–8
IS-127 (TIA/EIA) speech coding standard 360, 363–4, 364–8
IS-641-A (TIA/EIA) speech coding standard 14, 15
IS-733 (TIA/EIA) speech coding standard 360, 363–4, 364–8
Itakura 90
Itakura–Saito distortion 389
iterative sequential optimization 79–80, 116
ITU regulatory body
G.165 speech coding standard 413, 417–18
G.711 speech coding standard 12, 13
G.721 speech coding standard 12, 13
G.722 speech coding standard 13
G.722.1 speech coding standard 13
G.723.1 Annex A speech coding standard 360, 361, 364–8
G.723.1 coder 251
G.723.1 speech coding standard 12, 13
G.726 speech coding standard 12, 13
G.728 speech coding standard 12, 13
G.729 speech coding standard 12, 13
G.729 Annex B speech coding standard 360, 361, 364–8, 374–5
G.729 coder 251
generally 9, 12–13
P.862 measure of quality 18

J
Jayant (one word memory) quantizer 34–6
JQ-MSVQ quantizer 128–31

K
Kaiser window function 59, 60, 61
Kalman filter 380
Karhunen–Loève transform (KLT) 380
Katugampala 285
Kleijn 282
K-means algorithm 43–4, 45–6
Kroon 227

L
lag, pitch 78, 81, 83, 175–7
LAR function, see log-area ratio function
lattice method 72–4
least mean square 67
likelihood ratio 368–75
line spectral frequency (LSF)
derivation 94–5
distribution plot 97
generally 87, 90
properties 103–5
line spectral pair (LSP) 87, 90
linear interpolation 262
linear prediction filter coefficient 7
linear prediction vocoder 7
linear predictive analysis, see LPC analysis
Index

log spectral distortion, see spectral distortion
log-area ratio (LAR) function 88, 90, 91
logarithmic scalar quantizer 32–3
Log-PCM system 1, 2
long term prediction 202, 203
long term predictor in CELP 222–8
low bit-rate coder 2
low-band to full-band energy of speech 184
low-pass filtering 134–46
LP filter coefficient 7
LPC analysis
  autocorrelation method 68–70
  covariance method 70–1
  generally 65–77, 90
  lattice method 72–4
  least mean square approach 67
  maximum likelihood approach 67
  performance 74–7
LPC difference equation 66
LPC filter 87–90
LPC frequency response 77
LPC inverse filtering 76, 77
LPC mode 298
LPC predictor 202–3, 221–2
LPC quantization process 87, 90, 94–5, 97
LPC residual 272
LPC spectral envelope 77
LPC synthesis 93–4, 102–3, 104
LPC transformation to LSF
  adaptive sequential LMS method 100–1
  Chebyshev Series method 100
  complex root method 95
  generally 90–101
  ratio filter method 98–100
  real root method 95–6
LSF estimation 130–46
LSF inverse distance weighting method 109, 110, 121
LSF prediction 122–4
LSF quantization process 105–10, 121–30, 128–30
LSF quantizer 107, 110–16
LSF transformation to LPC 101–2, 102–3, 104
LSF, see line spectral frequency
LSP, see line spectral pair
LTP analysis 228–30
LTP and codebook excitation 255–7
M
MA-MSVQ quantizer 129–31
Markov 380
Max quantizer 30–2
maximally smooth criterion 297
maximum likelihood 67
maximum likelihood pitch
  measurement algorithm 81
maximum likelihood STSA estimation 380, 384, 389–92, 398
MBE mixed voicing 190–3
MBE-based coder 189
M-best tree search 115–16
McAulay 261, 262, 265, 297, 299, 384
mean opinion score (MOS) scale 17
mean square error distortion measure 42
mean square error measurement 106, 107–10
MELP coder 189
  see also 4 kb/s MELP/CELP coder
memory, filter 203
meta-frame 128
minimum mean square error STSA
  (MMSE- STSA) estimation 380, 386–7, 389–92, 400, 401
MIRS, see modified intermediate response system
mixed decision noise adaptation 403–4
mode decision 9
mode error 345–50
modified intermediate response system (MIRS) 164
moving average (MA) predictor
  generally 123–4
  joint quantization and 129–31
  low-pass filtering and 135–46
  optimal order 124–5
  performance 126–8
moving average (MA) predictor
(continued)
prediction factor 125–6
training 125–6
MPLPC coder
amplitude, optimum excitation
232–5
excitation behaviour 215–17, 218
generally 230–2, 233
pitch prediction 235–7
pulse location 216, 217
MSVQ, see multi-stage vector quantization
MSVQ quantizer 125–6
multi-band excitation (MBE) coder
149, 264–5, 277
multi-band voicing 264
Multimedia Communications Research Group of CCSR xiii
multimode coder, see hybrid coder
multi-pulse excitation 207–8, 230–2, 233
multipulse LPC (MPLPC) 200, 207
multi-stage vector quantization
codebook training 116
comparison with SVQ 117–21
generally 111
M-best tree search 115–16
2.4 kb/s MELP coder 113
performance 117–21
search strategy 114–16

N
narrowband speech coding standard 12–13
NATO speech coding standard 15
near-end speech detection in echo cancellation 413, 415
network dependent mode decision 9
neural network 283
Nguyen 153
noise adaptation
hard decision 402–3
mixed decision 403–4
performance comparison of methods 404–6, 407, 408, 409, 410
soft decision 402, 403
voice activity detector 402
noise suppression 2
noise suppression rules 369
noise suppressor with echo canceller 418–9
noisy speech 189
non-uniform scalar quantizer 29–30
normalized least mean squared (NLMS) algorithm 416
Nyquist criterion xiii, 24, 25, 131

O
offset target modification in SWPM 304–8
one word memory adaptation 34–6
one-shot optimization 79, 80–1
onset harmonic memory initialization 308–9
onset harmonic parameter quantization in AbS 330–1
open-loop mode selection of voicing 311–14
open-loop search scheme 51
optimization closed-loop 199, 200
optimization of codebook 43–6, 116
optimization of pitch predictor 79–81
optimum scalar quantizer 29–32
ordering of LSF parameters 100, 103–5
outlier 107
see also performance
overlapping codebook 241, 242, 243

P
P.862 measure of quality 18
packet loss, VAD and 359
Paliwal–Atal weighting method 108, 110, 121
PAME, see pitch adaptive mixed excitation
pan-European digital mobile radio system, see GSM
Panter and Dite 30
parametric coder 6, 7–8
partial correlation (PARCOR) coefficient 73, 87–90, 93–4
pattern-matching quantization, see vector quantization
PCM, see pulse code modulation
PDA, see pitch determination algorithm
peak detector 177
peakiness of speech 179–80
peak-picking of the magnitude spectrum 266–7
perceptual evaluation of speech quality 18
perceptually-based error minimization procedure in AbS-LPC coder 200, 203–6
perceptually-determined distortion measure 42–3
performance of LTP analysis methods 228–30
performance of echo canceller 415–23
performance of hybrid coder 320–2
performance of JQ-MSVQ quantizer 129
performance of low-pass filtering 142–6
performance of LPC analysis 74–7
performance of MA-MSVQ quantizer 129, 130
performance of moving average predictor 126–8
performance of multi-stage vector quantization 117–21, 125–6
performance of noise adaptation methods 404–6, 407, 408, 409, 410
performance of pitch determination algorithms 164–6, 167, 168
performance of pitch tracking process 175
performance of speech coding standards 15–18
performance of speech enhancement methods 389–402
performance of voice activity detector (VAD) 364–8
periodicity in speech signal 77–8, 178–9
phase synchronization 281
pitch adaptive mixed excitation (PAME) 251–4
pitch determination 149, 150–78, 263
pitch determination algorithm (PDA) autocorrelation 152–5
average magnitude difference (AMDF) 151–2
centre-clipping 169–72
generally 149
harmonic peak detection 156
peak detector 177
performance comparison 164–6, 167, 168
spectral autocorrelation 158–62, 163–4
spectral synthesis 163–4
spectrum similarity 156–8
pitch determination preprocessing 166–77
pitch error 165, 177–8
pitch estimation, see pitch determination
pitch filter 81–2
pitch gain, optimum 153, 154
pitch lag 78, 81, 83, 175–7
pitch measurement algorithm 81
pitch period
generally 149, 150–1, 165
LP filter coefficient 7
SWPM 323
pitch prediction 235–7
see also long term prediction
pitch predictor 77–83
see also long term predictor
pitch pulse location in AbS 325–7
pitch pulse location in SWPM 286–91, 302–4, 323
pitch pulse shape in AbS 327
pitch pulse shape in SWPM 292–7, 302–4, 323
pitch quantization in AbS 324, 325, 326
pitch smoothing 172–7
pitch tracking 172–7
pitch–LPC formulation model 79
plosive detection 318–19
polyphase structure 227
post-filtering in a CELP coder 257–60
Index

power spectrum 99
power-saving, VAD and 358
PPL, see pitch pulse location
PPS, see pitch pulse shape
prediction gain 78, 80, 124–5, 140–1
prediction of LPC parameters in CELP 221–2
prediction of pitch in CELP 222–8
predictive vector quantization 52
predictor codebook 52
pre-emphasis of the signal 75
pre-emphasized energy of speech 183
probability density 33
prototype waveform interpolation (PWI) coder 282
public switched telephone network (PSTN) 5, 9, 10
pulse amplitude coding 237–8
pulse amplitude quantization, joint 238–40
pulse code modulation (PCM) xiii, 5, 32–3
pulse excitation 202
pulse location 211–12, 216, 217, 248
pulse position coding 237

Q

quality measurements 16
quantization 23, 238–40
see also types of quantization:
differential vector, LSF, multi-stage vector, predictive vector, scalar, split vector, vector
quantization error 27–8, 29, 106
quantization issues of hybrid coder 322–31
quantization noise 7
quantization process
scalar 26–39, 106
vector 39–50, 106
see also LPC quantization process, LSF quantization process
quantizer
adaptive 33–9
companded 32–3
differential 36–9
Jayant 34–6
JQ-MSVQ 128–31
logarithmic scalar 32–3
LPC 87, 90, 94–5, 97
LSF 107, 110–16
Max 30–2
see also scalar quantizer
quantizer input/output 31
quantizer step size 26

R

Rabiner 169, 172
random channel error 336
random noise excitation 202
random noise generator 65
ratio filter method 95–6, 98–100
real-time coder 108
real-time system 74
rectangular window function 58–9, 60–5, 165
Reeves 5
reference template 40
regular pulse excitation 207–8
regulatory body 9
residual error suppression in echo cancellation 413, 415
rms energy 317, 318, 333
robustness 10
RPELPC coder 217–19

S

sampling 23–5
satellite telephony 15
SB-LPC, see split-band LPC
scalar quantization process 26–39, 106
scalar quantizer 54
scalar quantizer, logarithmic 32–3
scalar quantizer, non-uniform 29–30
scalar quantizer, optimum 29–32
scalar quantizer, uniform 26–9
secure communication 14–15
segmental SNR 389
self-excitation 207
SELP 207
SELP coder 212–15
sequential optimization 116
Index

Shlomot 282
short term predictor, see LPC analysis
short-time spectral analysis 57–65
signal compression 1–2
signal power LP filter coefficient 7
signal processing 1–2
signal reconstruction 5
signal to noise ratio (SNR)
  CELP coder 228–30
generally 7
  RPELPC coder 218–19
segmental 389
signal variability 53–4
simultaneous joint codebook design 116
Singhal and Atal 235
sinusoidal analysis 262–3
sinusoidal coder 8, 261–75
sinusoidal model voicing 265–6
sinusoidal speech coder 149, 150, 156
sinusoidal speech-model matching 299
sinusoidal transform coder (STC) 149, 261–75
smoothed likelihood ratio (SLR) 371–2, 373, 374–5
SNR, see signal to noise ratio
soft-decision noise adaptation 402, 403
soft-decision voicing 150, 189–96
Sohn 368
Sondhi 169, 172
source dependent mode decision 9
source-filter model 65–7
speaker variability in signal 53
spectral analysis, short-time 57–65
spectral autocorrelation PDA 158–62, 163–4
spectral correlation 267–8
spectral distortion 106–7, 131–4
spectral envelope 77, 149, 202
spectral subtraction 380, 382–4, 389–92, 396, 397
spectral synthesis method 150
spectral synthesis PDA 163–4
spectral tilt of speech 182, 266
spectrum flattening 166–72
spectrum similarity PDA 156–8

speech characteristic
  frame energy 185–6
  low-band to full-band energy 184
  peakiness 179–80
  periodic similarity 178–9
  pre-emphasized energy 183
  spectrum tilt 182
  weighting 186–7
  zero crossing rate 180–1
speech classification in a hybrid coder 311–19
speech coder, see coder
speech coding standard
  DoD 14–15
  ETSI 13–14
  INMARSAT 15, 16
  ITU-T 12–13
  NATO 15
  performance 15–18
  TIA/EIA 14, 15
speech enhancement
  adaptive filtering 380
discrete cosine transform (DCT) 380
discrete Fourier transform (DFT) 380
echo cancellation 406–23, 424–6
generally 379–80
guidelines 402
Kalman filter 380
Karhunen–Loève transform (KLT) 380
maximum likelihood STSA
  estimation 380, 389–92, 398
minimum mean square error STSA
  estimation 380, 389–92, 400, 401
model-based 380
noise adaptation 402–6
performance comparison of methods 389–402
short-time spectral amplitude 381–402
spectral subtraction 389–92, 396, 397
transform domain 380
uncertainty of speech presence 387–9, 401
speech enhancement (continued)
  wavelet transform 380
  Wiener filtering 389–92, 399
speech presence, uncertainty of 387–9, 401
speech quality xiii, 9, 10, 334 – 5
speech signal
  average spectral distortion 107, 121
  LPC analysis 65–77
  outlier 107, 121
  periodicity 77–8
  requirements for good quality 107
  spectral analysis 57–65
  transition region 57
  unvoiced 57, 58, 65
  voiced 57, 58, 65
speech stationarity assumption 131
split-band LPC (SB-LPC) coder 128, 261, 268, 271–5, 342
split vector codebook 49–50
split vector quantization 111–12, 117–21
split-band mixed voicing 193–6
Stachurski 283
STANAG speech coding standard 15
statistical multiplexing, VAD and 359
STC, see sinusoidal transform coder
STP, see short term predictor
STSA, see short-time spectral analysis
STSA estimation 380, 389–92, 398, 400, 401
Sundberg 174
SWPM, see synchronized
  waveform-matched phase model
synchronized harmonic excitation 299–301
synchronized waveform-matched phase model (SWPM)
  advantages 301–4
  generally 285–98
  offset target modification 304–8
  robustness to acoustic noise 342

T
  tandem connection 11
  telephony system, analogue 1
threshold function, voicing 191–2, 196
TIA regulatory body
  enhanced variable rate coder (EVRC) 286
  IS-54 speech coding standard 14, 15
  IS-96 speech coding standard 360, 363–4, 364–8
  IS-127 speech coding standard 360, 363–4, 364–8
  IS-641-A speech coding standard 14, 15
  IS-733 speech coding standard 360, 363–4, 364–8
time division multiple access (TDMA) 14
time-domain pitch determination 151–5, 158–66, 177
time-varying codebook 52
time-varying filter in AbS-LPC coder 200, 202–3
Toeplitz matrix 69, 249
training a codebook 52, 116, 246
training a moving average predictor 125–6
Trancoso 282
transcoding 11
transition 298
transition detection 315–18
transition quantization in ACELP 331, 332–4
transition region 150
transmission channel errors 54
tree search, M-best 115–16
tree search codebook, see binary search codebook

U
UMLTS (ETSI) speech coding standard 360
uncertainty of speech presence 387–9, 401
uniform scalar quantizer 26–9, 33
unvoiced excitation 202
unvoiced speech signal 57, 58, 65, 150, 298
up-sampling 226–8
Index

V

VAD, see voice activity detector
variable bit-rate coding 331–5
variable rate coder 9

vector quantization
harmonic amplitude 272
multi-stage 111, 113–21
split 111–12
see also codebook

vector quantization process 39–50, 106

vector quantizer 54

vector sum codebook excitation 243–7

Villette 299

vocoder 6–7, 149, 150, 202

voice activity detector (VAD)
benefits 357–9
ETSI speech coding standards 360, 361–2, 362–3, 364–8, 374–5
hard decision noise adaptation 402
ITU-T speech coding standards 360, 361, 364–8, 374–5
likelihood ratio 368–75
performance 364–8
TIA/EIA speech coding standards 360, 363–4, 364–8
voicing decision 359

voice activity detector (VAD) algorithm
11, 281, 311, 341, 357

voiceband data handling 11–12

voiced excitation 202

voiced speech signal 57, 58, 65, 150, 298

voiced–unvoiced classification 149

voicing
frequency-domain 263
generally 281
harmonic 264–6
multi-band excitation 264

sinusoidal model 265–6
threshold function 191–2, 196

W

waveform coder 6–7, 8
waveform, equation for sampled 23
wavelet transform 380
weighted mean square error
measurement 106
weighted mean square error distortion
measure 42
weighting filter of AbS-LPC coder 204–5

weighting method
EFR 109, 110, 120, 121
group delay 109–10, 121
LSF inverse distance 109–10, 121
Paliwal–Atal 108, 110, 121
performance 119–21

white noise excitation 309–11
white noise excitation mode error 346, 347

wideband speech coding standard 13

wide-sense stationary assumption 133

Wiener filtering 380, 385–6, 389–92, 399

window length 81

window function
Bartlett 58–9, 60, 61
Blackman 59, 60, 61
generally 58–65, 75
Hamming 59, 60–5, 165, 190, 262
Kaiser 59, 60, 61
rectangular 58–9, 60–5, 165

window position test 132

Z

zero-crossing rate of speech 180–1, 313