### Index

#### A
- Accumulators, 59, 62–66, 70
- Acoustic echo cancellation, 136–140
- Active noise cancellation, 129, 135–136
- Adaptive codebook, 122
- Adaptive filters, 131, 132, 135, 138, 139
- Adaptive noise cancellation, 130–133
- Algebraic codebook, 122
- Analog signal processing, 2
- Analog-to-digital converter (ADC), 75–82, 87
- Antialiasing filter, 13, 14, 28
- Audio codec, 82–85, 87
- Auto-correlation, 17–19

#### B
- Band pass filters, 25–28
- Band stop filters, 25–28
- Barrel Shifter, 70
- Bit-reversed addressing, 59, 67–70

#### C
- Central processing unit (CPU), 55–74
- Cepstrum, 152, 153
- Circular buffers, 59, 67–68
- Code excited linear prediction (CELP), 118–124
- Companding, 103, 104
- Concatenative speech synthesis, 157–159
- Consumer appliances, 5
- Controller area network (CAN), 89
- Convolution, 9, 15–19, 21, 22, 24
- Cross-correlation, 17

#### D
- Data normalization, 70
- Decimation, 22, 23, 35–36
- Decoding, 108
- Differential pulse code modulation (DPCM), 106
- Digital filters, 25, 28–30
- Digital signal controllers (DSC), 60–71
- Digital signal processing, 2, 4, 11, 15, 25
- Digital signal processor (DSP), 2, 4, 6, 59–73
- Digital systems, 1, 2, 4
- Digital-to-analog converter (DAC), 76, 80–82, 87
- Direct form, 33–35
- Direct memory access (DMA), 79, 81, 83, 90
- Direct quantization, 97
- Discrete Fourier transform (DFT), 20–22, 24
- Double talk detector (DTD), 139
- Dynamic time warping (DTW), 149

#### E
- Embedded speech processing, 167
- Embedded systems, 3–4
- Encoding, 106, 109
- Euclidean distance, 147–149
- Exception processing, 73–74
- Excitation, 41, 42, 47

#### F
- False acceptance rate, 154
- False rejection rate, 154
- Fast Fourier transform (FFT), 22–25, 30
- Feature vectors, 147–150, 152, 153
- Filtering, 25, 26, 28, 30
- Finite impulse response (FIR), 30–33
- Fixed point, 59–62
- Floating point, 59–62, 73
- Frequency domain, 21, 22, 24, 30
- Front-end analysis, 152–153
- Full-duplex configuration, 93–95
- Fundamental frequency, 164
Index

G
G.167, 137
G.168, 137
G.711, 104–108, 110
G.722, 108–110
G.723.1, 119, 122–124
G.726, 105–110
G.728, 119–122, 124
G.729, 119, 122–124
G.726A, 105–110

H
Half-duplex configuration, 93–95
Hidden Markov models (HMM), 150–151, 153–155
High pass filters, 25–26
Human auditory system, 49–51
Human speech production, 37–39, 43

I
Industrial control, 5
Infinite impulse response (IIR), 30–35
Input sampling rate, 96
Intercom systems, 128, 136, 138
Inter-integrated circuit (I2C), 87–90
Interpolation, 35–36

L
Lattice form, 44, 45
Least mean square (LMS), 131, 132, 135, 138–141
Linear prediction, 114, 118, 121
Linear predictive coding (LPC), 113–118, 120, 121, 124
Line echo cancellation, 136, 137, 140, 141
Linguistic analysis, 163
Logarithmic quantization, 103, 104
Low delay code excited linear prediction (LD-CELP), 120, 121
Low pass filters, 13, 25–28, 36

M
Magnitude difference function, 49
Mean opinion score (MOS), 53
Memory caches, 72–73
Microcontroller units (MCU), 4, 57–60, 63, 71–73
Microphone, 127–136, 138–141
Microprocessors, 3, 4
Million instructions per second (MIPS), 98, 99

N
Mobile hands-free kits, 128, 138
Modulo addressing, 67, 68
Morphological analysis, 162, 163
Multi-band re-synthesis, 161, 162
Multiply-accumulate (MAC), 59, 62–66

O
Noise cancellation, 127–141
Noise suppression, 127–130, 133–135, 137, 140–141
Nyquist–Shannon sampling theorem, 12, 13, 20, 36

P
Parallel processing, 166
Parametric quantization, 97
Peripherals, 75–90
Phonetic transcription, 163
Pipelining, 71–72
Pitch period, 47–49
Pitch-synchronous overlap-add (PSOLA), 161, 164
Program flow control, 66–67
Prosody generation, 163
Pulse code modulation (PCM), 103–106
Pulse width modulation (PWM), 81–82

Q
Quadrature mirror filters (QMF), 108
Quantization, 11–15, 31, 33

R
Reflection coefficients, 44–46
Relative spectral (RASTA), 153

S
Sampling, 10–16, 20, 35, 36
Sampling rate, 12–14, 20, 35, 36
Saturation, 59, 64
Scalar quantization, 101–102, 105
Serial peripheral interface (SPI), 87, 88
Short-time energy, 47, 48
Signal conditioning, 79–80
Signal-to-noise ratio (SNR), 52, 53
Simplex, 93, 95
Speaker, 128, 129, 135–139, 141
Speaker-dependent speech recognition, 145, 149, 152
Speaker identification, 146–149, 154
Speaker-independent speech recognition, 144, 145, 149, 152, 153, 155
Speaker normalization, 153
Speakerphones, 128, 136, 138
Speaker verification, 146, 147, 149, 154
Spectral subtraction, 130
Speech coders, 105–110
Speech compression, 93–99
Speech decoding, 96
Speech decompression, 96, 98
Speech encoding, 93, 97, 98
Speech playback, 80–82
Speech processing, 4–7, 37–53
Speech recognition, 143–155
Speech sampling, 75–76, 82
Speech synthesis, 157–164
Speex, 118–120, 122–124
Subbands, 108–110
Superscalar, 166
Syntactic analysis, 163

Text-to-speech (TTS) systems, 159, 160, 162–164
Time domain, 21, 22, 25, 30

U
Uniform quantization, 102–103
Universal asynchronous receiver/transmitter (UART), 85–88
Unvoiced sounds, 41, 42, 47

V
Vector codebook, 118, 121, 122
Vector quantization, 118–119, 124
Very long instruction word (VLIW), 166
Viterbi algorithm, 151–152
Vocal tract, 38–44
Vocoders, 113, 114, 116–120
Voice activity detection (VAD), 133, 134, 139
Voice coders, 97, 100, 113–124
Voiced sounds, 39–42

W
Waveform coders, 97, 100–110
Window functions, 24, 25, 31, 32

Z
Zero-crossing rate, 48
Zero-input zero-state filtering, 46