

# Index

- accuracy 41, 117–118, 120–121, 123, 137, 139, 272, 316
- acoustic echo
  - cancellation 2, 21–22, 52, 307–309, 311, 313–315, 320–321, 329, 331, 334, 336–338, 351
  - reduction 2–3, 5, 21–22, 52, 307–308, 313, 316, 320, 326–328, 334–335, 337–338, 341, 356
  - suppression 3, 5, 22, 52, 308, 311–313, 323, 325–327, 330, 333–338, 341, 353–354, 356
    - multichannel 327–328, 330, 333, 336, 338, 355
- acoustic echo cancellation
  - multichannel 329–330, 336, 338
  - decorrelation 329
  - reducing the degrees of freedom 331, 338
  - non-linear 307, 320–322, 329
- acoustic echo reduction
  - center clipper 334
  - feedback prevention 335
  - multichannel 327–329, 336, 338
  - shadow filters 334
- acoustic echo suppression 22, 52, 311–312, 323, 330, 338
- active noise cancellation 20
- ambient noise gain 194, 213
- array gain 193–195, 230, 256
- array processing
  - medicine 168
  - radars 165–166, 169
  - radio astronomy 166
  - audio applications 169
  - under water 166
- attenuation
  - sound 19–20, 63–65, 85–86, 93, 169, 220–221, 236
- audio
  - subjective quality 49, 52, 345
    - differential MOS 49–50, 345
    - mean opinion score 49, 52
    - measurement 49–50, 52
- audio processing
  - bandwidth limiting 45
  - cepstral analysis 43, 45
  - frequency domain 4, 7, 12, 32–34, 37, 40–43, 52, 96–97, 156, 159, 161, 233
  - lapped transformations 42
  - spectrogram 40–41, 156
- automatic speech recognition 1, 21, 51, 169
- A-weighting 37, 66, 119, 276
- bandwidth limiting
  - audio processing 45
- Bark scale 144
- beamformer
  - adaptive 198, 231–237, 247–250, 252, 260, 336–338, 356
  - beamwidth 193
  - delay-and-sum 182, 185–188, 191, 202, 207–208, 211, 213, 228–230, 250, 252, 280, 288
  - direct optimization 210–212, 225, 230, 255
  - Frost 198, 232–233, 250

- beamformer (*Continued*)
  - generalized 187, 200, 224, 233–236, 247–248, 253, 282, 288
  - generalized side-lobe canceller 233–236, 247
  - Griffiths-Jim 233
  - MPDR 203, 231
    - adaptive 231
  - MPDR 203, 231
    - adaptive 231
  - MVDR 198, 201–203, 207–208, 210–211, 225, 230–232, 238, 282
    - with instrumental noise
  - optimization 199, 207–208, 210–212, 216, 218, 224–226, 229–230, 253
  - pattern synthesis 203, 207, 209, 211
    - cosine function 203
    - Dolph-Chebyshev polynomial
    - robust generalized side-lobe canceller 235
    - time invariant 198, 200, 203, 208, 231, 233–236, 248, 253, 255, 260, 336
    - Widrow 232–233, 248
  - beamwidth 193
- beamforming 3, 174, 188–189, 198, 213, 234, 251, 266, 336, 354
- beamsteering 3, 189, 263, 281
- beamwidth 193
- B-weighting 66
- calibration
  - microphone array 165, 218–219, 221–224, 230–231, 333
- carbon microphone 69
- cardioid microphone 75, 90, 191, 254
- channel mismatch
  - microphone array 214, 224–226, 230
- C-message weighting 67, 121
- condenser microphone 69
- convergence time
  - echo cancellation 313–315, 320
- C-weighting 66–67, 121
- decision directed approach 128–129, 162, 242
- delay-and-sum beamformer 182, 185–188, 191, 202, 208, 211, 228–230, 252, 280, 288
- de-reverberation
  - using cepstrum 352
  - using speech signal properties 353, 356
  - via de-convolution 351
  - via suppression 353
  - with LP residual 352
- diffraction
  - sound 60–61, 63–65, 93, 170–171, 174, 267
- direction of arrival 61, 70, 74, 165, 189–190, 196–197, 218, 221–222, 241, 244, 264, 272–273, 275–276, 278
- directivity index
  - microphone 75, 77–78, 82–84, 93, 190–192, 194–196, 211, 213, 216–217, 246–247, 252–254, 342, 354
- discretization
  - errors 29
    - averaging 29–31, 291
    - finite duration 31, 37
    - non-ideal sampling function 29
- Dolph-Chebyshev polynomial
- Doppler effect
  - sound 65–66, 93
- double talk
  - detectors 316–317, 338
  - coherence algorithms 319
  - cross-correlation algorithms 317
  - Geigel algorithm 317
- D-weighting 67
- echo cancellation
  - affine projection algorithm 315, 338
  - fast affine projection algorithm 338
  - LMS 313–316, 321
  - NLMS 315–316, 320, 322
  - non-linear 307, 320–322, 329
    - in frequency domain 323
    - power filters 322
    - Volterra filters 321
  - RLS 315–316, 329
- echo reduction 2–3, 5, 21–22, 52, 307–308, 313, 326–328, 334–335, 337–338, 341, 355
- Echo Return Loss Enhancement 313
- echo suppression 22, 52, 311–313, 323, 326–327, 330, 338
- electret microphone 74, 83, 182
- electrodynamic microphone 69
- Ephraim and Malah
  - suppression rules 103–105, 114–115, 129, 150, 242
- ERLE 313, 315, 320–323, 326

- false negatives 117–118, 120, 316, 320  
 false positives 116–120, 316  
 far end 22, 309, 315–317, 320–321, 324–325,  
 327–329, 331–333, 335  
 frequency domain  
   audio processing 4, 7, 12, 32–33, 37, 40, 42,  
   49, 52, 96, 156, 160, 336  
   properties 4, 7, 11–13, 23, 33–35, 43, 52,  
   96, 151  
   conjugation 34  
   convolution 24, 34  
   correlation 34  
   derivatives 34  
   energy 11–13, 34, 52, 151  
   linearity 34  
   Parseval's theorem 34  
   Plansherel's theorem 34  
   scaling 34  
   time shift 34  
   Wiener-Khinchin's theorem 34  
 short-time transformation 36  
 weighting 36
- gain calibration  
   microphone array 219, 221  
 gain self-calibration  
   microphone array 219  
 generalized beamformer 253  
 G-weighting 67
- head related transfer function  
   (HRTF) 267
- human hearing  
   Bark scale 144  
   ear 19–20, 57, 142  
   loudness 62, 143  
   masking effects 144, 147–148, 162  
   organ 142  
 Huygens' principle 63, 65  
 hypercardioid microphone 342
- instantaneous direction of arrival (IDOA)  
   196, 241
- instrumental noise gain 201, 208, 213  
 intensity  
   sound 56–57, 61–64, 73–74, 90, 93, 267  
 interaural intensity difference 267  
 interaural time difference 267  
 ITU-R 468–4 67
- JMAP SAP  
   suppression rule 103
- longitudinal wave 55, 58, 92
- loudness  
   human hearing 62, 143
- manufacturing tolerances  
   microphone array 165, 189–190, 214,  
   216–219, 223–226, 229  
   robust design 225, 227, 230  
   microphones 70, 78–80, 82–83, 165, 189,  
   214–219, 223–224, 227, 260
- MAP SA  
   suppression rule 103
- Martin  
   suppression rule 105
- masking effects  
   human hearing 144, 147, 162
- maximum likelihood  
   suppression rule 100–102, 108–109,  
   114–115, 128–129, 136
- McAulay and Malpass  
   suppression rule 100–102, 124
- MEMS microphones 70
- microphone  
   carbon 69, 93  
   cardioid 75, 79, 82, 84, 90, 93, 171,  
   176, 191, 214, 216, 220, 251–254,  
   256  
   close talk 2, 85, 111, 250, 313, 341,  
   345, 347  
   condenser 69–70, 78, 93  
   directivity index 75, 77, 82–83, 93,  
   190–191, 194–195, 211–212, 216,  
   246–247, 252, 254, 342, 354  
   electret 70, 74, 78–79, 83, 93, 182, 214, 216  
   electrodynamical 69, 93  
   hypercardioid 82, 84, 93, 176, 254, 342  
   manufacturing tolerances 70, 78–79, 82,  
   165, 189, 214–218, 223–225,  
   227–230, 260
- MEMS 70  
   noise cancelling 141  
   omnidirectional 70, 72, 77, 82–83, 87, 90,  
   93, 172, 176, 183, 185, 193–195, 200,  
   204, 219–220, 224, 240, 251–252, 272,  
   276, 280, 342  
   piezoelectric 69, 93

- microphone (*Continued*)
- pressure 19, 55–56, 68–75, 78, 82–86, 92–93, 142–143, 214, 218, 221, 264, 342, 350
  - pressure gradient 70–73, 75, 82
  - proximity effect 84–85, 93
  - sensitivity 56, 73–74, 77–79, 87–88, 93, 142–143, 166, 185, 188–190, 214–217, 219
  - subcardioid 82, 84, 93, 176
  - supercardioid 82, 84, 93, 176
  - unidirectional 4, 74, 83, 171–173, 185, 200, 204, 211, 219, 244, 251–252, 254, 276
- microphone array
- adaptive beamformer 232, 236, 248, 250
  - ambient noise gain 194
  - array gain 193–195, 230
  - beamforming 3, 174, 188–189, 198, 251, 266, 336, 354
  - beamsteering 3, 189, 263
  - beamwidth 193
  - calibration 165–166, 218–219, 221–223, 230, 333
  - channel mismatch 224, 226, 230
  - circular 166, 172–173, 202, 219
  - delay-and-sum beamformer 182, 185–188, 191, 202, 208, 211, 229–230, 252, 280, 288
  - direction of arrival 165, 189–190, 196–197, 218, 221–222, 241, 244, 264, 272, 276, 278
  - far field model 177–178, 185, 190
  - Frost beamformer 232
  - gain calibration 219, 221
  - gain self-calibration 219
  - generalized beamformer 253
  - generalized side-lobe canceller 236, 247
  - instrumental noise gain 201, 208
  - linear microphone array 171–172, 180, 185, 219, 243–244, 278, 285, 288
  - near field model 176, 178
  - noise gain 194–195, 197, 201, 208–209, 211, 216–217, 224, 226, 229
  - non-correlated noise gain
  - planar 173–174, 197, 204, 221, 279
  - post-processor 190, 238–239, 241, 246, 270–271, 302
    - IDOA 241–242, 247
    - MMSE estimator 238
  - noise field coherence 240
  - power densities 238
  - spatial filter 246
- robust design 225, 227, 230
- direct optimization 225, 230
  - MVDR beamformer 225, 230
- robust generalized side-lobe canceller 236
- self-calibration 218–219, 221
- small microphone array 253–254, 264
- sound source localization 3, 165, 174, 189, 196–197, 244, 263–267, 270–272, 278, 280, 286–289, 297–299, 301
- spatial aliasing 178–180, 185
- spatial ambiguity 180
- spatial filtering 190, 197, 244, 251, 254
- specialized 170, 174, 218
- total noise gain 195, 197, 208, 216
- volumetric 173, 181, 197, 204
- Widrow beamformer 232, 248
- microphones
- measuring 68–69, 93, 175
  - models 4, 33, 55, 73, 78, 84, 93, 166, 178, 195, 220, 241–243, 255
- MINT algorithm 354
- MMSE
- suppression rule 103
- MMSE
- suppression rule 102–106, 111–112, 114–115, 131, 134, 136–138, 149, 162
- musical noise 99–101, 109, 115, 156, 159, 162, 249, 260
- near end 309, 311, 314–320, 324–328, 333, 335
- noise
- active cancellation 20
  - cancellation 1–2, 9, 19–22, 52, 139–141, 153–154, 156, 162, 307–309, 311, 313–315, 320, 329, 336
  - definition 4, 7, 9, 12, 43, 48, 66, 68, 97, 99, 190, 195
  - properties 4, 7, 9, 11–13, 16–18, 23, 33, 43, 52, 95–96, 151, 245, 338, 353, 356
  - spatial 11, 18, 52, 245
  - spectral 9–10, 12, 16, 18–19, 43–44, 52, 96–97, 151, 354, 356
  - statistical 7–8, 11, 13–14, 52, 95
  - temporal 11, 17, 52, 245
- noise cancellation 20–22, 139–141, 162

- noise gain
  - ambient 77, 194–197, 201, 208, 211–213, 225, 231
  - instrumental 201–202, 208, 211–213, 225, 229
  - non-correlated
  - total 77–78, 195, 197, 201, 208, 211–213, 216, 224
- noise model
  - decision directed approach 162
  - estimation 45, 130–131, 134, 138, 160, 162, 190, 237, 243–244, 323
- noise reduction
  - cancellation 20–21, 52, 139–140, 162
  - psychoacoustic based
  - suppression 20, 52, 85, 95–100, 139, 161
- noise suppression
  - a posteriori SNR 97, 99, 101–102, 104, 127–128, 136, 138
  - a priori SNR 97, 99, 102, 104, 127–128, 138–139, 162, 242
  - decision directed approach 128, 162, 242
  - musical noise 99–101, 115, 159, 162, 260
  - noise model 130–131, 136–137, 139, 159, 162, 189, 243
  - perceptually balanced 149
  - predictable components 150
  - speech modeling 157
  - suppression rule 20, 97, 99–105, 114–115, 128, 130–131, 136–138, 149, 160–162, 242
  - uncertain presence 101, 115, 136, 160, 162
- noise suppressor
  - architecture 5, 130
  - optimization 137–138, 161
- non-correlated noise gain
- normalized misalignment 313
- nullforming 189
- nullsteering 189
- O.41 weighting 77
- omnidirectional microphone 77, 82, 185, 193–195, 200, 342
- overlap-add process 37, 40, 43, 52, 310
- piezoelectric microphones 69
- post-processor
  - microphone array 190, 238, 241–242, 246–247, 270–271, 294, 297, 302, 304
  - IDOA 241–243, 247
  - MMSE estimator 238
  - noise field coherence
  - power densities 238
  - spatial filter 246
  - spatial filter in side-lobe canceller 247
  - spatial noise suppressor 243
- pressure gradient microphone 71, 73, 75, 82
- pressure level
  - sound 61–64, 66, 74–75, 78, 93, 103, 143–144, 342
- pressure microphone 70–71, 73, 75, 83, 86
- proximity effect
  - microphone 84–86, 93
- quantization 2, 23, 25–27, 52, 158, 195, 297
- receiver operating characteristics 117
- reflection
  - sound 21–22, 63–66, 93, 266–267, 269–270, 278–279, 293–294, 307–309, 341–342, 350
- residual energy
  - estimation 307, 313–314, 323
  - suppression 308, 312–313, 323, 325
- reverberation
  - absorption ratio 344, 350
  - critical distance 236, 342, 344–345, 354
  - direct path signal 182, 341
  - early reflections 348
  - human perception 349, 352
  - image method 345, 350, 356
  - speech recognition 2–4, 21, 45, 169, 251, 336–337, 347–348, 352
  - tail 22, 342, 345, 347, 350, 353
- reverberation effect 341
- reverberation time 344–345, 347–349, 356
- ROC 116–118, 120, 123, 272, 316, 320
- room impulse response 21, 266, 331–332, 341–342, 345, 347–348, 351, 354, 356
- sampling theorem 23, 25, 27–28, 31, 52, 181
- Schroeder formula 348
- self-calibration
  - microphone array 218–219, 221
- short term log-MMSE
  - suppression rule 103
- short term MMSE
  - suppression rule 102–104, 149

- signal
  - definition 4, 7, 9, 12, 43, 48, 68, 97, 99, 190
  - properties 4, 7, 9, 11–13, 16–18, 23, 33–35, 43, 51–52, 95–96, 151, 245, 338, 351–353, 356
  - spatial 11, 18, 52, 245
  - spectral 9–10, 12, 16, 18–19, 43–44, 51–52, 96–97, 151, 354, 356
  - statistical 7–8, 11, 13–14, 52, 95
  - temporal 11, 17, 52, 245
  - quantization 2, 23, 25–27, 52, 158
  - reconstruction 20, 27–28, 39–41, 43, 52, 352
- signal-to-noise ratio 32, 48, 52, 68, 74, 93, 95, 121, 182
- signal-to-reverberation ratio 341
- SNR
  - a posteriori 97, 99, 101–112, 115, 122–123, 125, 127–128, 130–131, 136, 138–139, 162, 242, 326
  - a priori 96–97, 99, 102, 104, 106–112, 115, 122–123, 125–131, 136, 138–139, 158, 162, 242–243, 326
  - estimation 96–100, 102, 105, 107, 112, 115, 125–131, 138, 162, 242–244, 326
- sound
  - attenuation 19, 63–64, 85, 93, 145, 168, 220, 236
  - diffraction 60, 63–64, 93, 170, 174, 267
  - Doppler effect 65, 93
  - Huygens' principle 63, 65
  - intensity 56, 61–63, 73, 90, 93, 267
  - pressure level 61–63, 66, 74, 78, 93, 103, 143–144, 342
  - measuring 61, 66–67, 93, 143
  - weighting curves 66
  - reflection 21, 63–65, 93, 266, 269, 278, 293, 307–308, 341, 350
  - speed of sound 58–60, 65, 71, 93, 176, 200, 225, 264, 272–273, 344, 351
    - in air 59, 61, 71
    - in liquids 58
    - in solids 58
- sound pressure level
  - weighting curves 66
    - A-weighting 66
    - B-weighting 66
    - C-message weighting 67
    - C-weighting 66
    - D-weighting 67
    - G-weighting 67
    - ITU-R 468–4 67
    - O.41 67
    - Z-weighting 67
- sound propagation 59, 62–63, 165, 167, 176, 181, 204, 219, 273
- sound source localization
  - least commitment 278
  - localization error 271
  - particle filtering 288, 304
  - post-processing 190, 263, 293, 302
    - clustering 294, 302
    - confidence levels 302
  - steered response power 280–281, 286–288, 290, 304
  - time delay estimation 272–273, 304
- sound source tracking
  - particle filtering
  - post-processing
    - confidence levels 271, 302
    - fuzzy c-means clustering 298
    - k-means clustering 297
    - simple clustering 294, 302
- spatial filter
  - in side-lobe canceller 247
  - microphone array 244, 246–247, 250, 337
- spectral subtraction
  - suppression rule 100–102, 114–115, 124–125, 131, 136, 330
- speech enhancement 1, 3–5, 21, 51–52, 114, 137–138, 141, 162, 337, 356
- speech recognition 1–4, 13, 21, 41, 45, 51, 95, 114, 137–138, 169, 241, 251, 327, 336–337, 347–348, 352
- speed of sound
  - in air 59, 61, 71
  - in liquids 58
  - in solids 58
- steered response power
  - beamsteering 281
  - comparison 285, 304
  - maximum likelihood 282, 285
  - MUSIC algorithm 282, 285
  - PHAT weighted 285
  - weighted 280–281, 285, 291
- subcardioid 82, 84, 93, 176
- supercardioid 80, 82, 84, 93, 176

- suppression rule
  - JMAP SAP 103
  - MAP SA 103
  - Martin 105
  - maximum likelihood 100–101, 108, 114, 128, 135
  - McAulay and Malpass 100–101, 124
  - MMSE SP 103
  - MMSE 102–105, 111, 113–114, 131, 134–136, 138, 149, 162
  - short term log-MMSE 103
  - short term MMSE 102–103, 149
  - spectral subtraction 100–101, 113–114, 124, 131, 135, 330
  - Wiener 98–99, 101, 103, 105, 108–110, 112–114, 128–129, 135–136, 150, 162, 242, 325
- suppression rules
  - comparison 101, 109–113, 131, 134
  - Ephraim and Malah 97, 102–104, 114, 125, 129, 150, 242
  - modifier 125, 136, 160
  - probabilistic 105, 109, 111, 114, 129, 134, 136
  - Wolfe and Godsill 103
- system distance 312
- time delay estimation
  - comparison 304
  - weighting 273
    - Eckhart 275
    - PHAT 274
    - Roth 274
    - SCOT 274
- total noise gain 195, 197, 208, 213, 216
- transverse waves 63
- true negatives 117, 120
- true positives 116–117, 120, 320
- unidirectional microphone 83, 211, 251, 254
- VAD
  - accuracy 117–118, 120–121, 123, 137, 316
  - binary 48, 115–117, 119–121, 220, 235, 255, 276, 278
  - false negatives 117–118, 120, 316, 320
  - false positives 116–120, 316
  - hangover scheme 122
  - likelihood ratio test 122
  - true negatives 117, 120
  - true positives 116–117, 120, 320
- voice activity detector 51, 115, 120–121, 131, 134, 153, 162, 220, 255, 270, 304, 315–316, 323, 337, 354
- wavelength of sound 57, 60, 179
- weighting curves
  - A-weighting 66
  - B-weighting 66
  - C-message weighting 67
  - C-weighting 66
  - D-weighting 67
  - G-weighting 67
  - ITU-R 468–4 67
  - O.41 67
  - Z-weighting 67
- Wiener
  - suppression rule 98–106, 108–112, 114–115, 128–130, 136–137, 162, 242–243, 325
- Wolfe and Godsill
  - suppression rules 103
- Z-weighting 67

