

Index

Note to index: An *f* following a page number denotes a figure on that page; an *n* following a page number denotes a note on that page; a *t* following a page number denotes a table on that page.

- absolute truth, 79n16
 - ACRONYM system, 4, 5 *f*, 14
 - Adaboost model, 254 *f*
 - adaptation, 149
 - adolescents
 - with autism, 358
 - category selectivity in, 120–121 *f*, 122 *f*
 - category selectivity in ventral stream, 349
 - face-related cortex development in, 350, 352–353, 354–355, 361–362
 - aerial images, 136
 - affordances, 138, 140–141
 - Agarwal, S., 203
 - Aggelopoulos, N. C., 266, 269
 - agnosia, 102
 - Ahuja, N., 9, 10 *f*
 - Aminoff, E., 393, 394
 - Amit, Y., 203
 - Amsterdam Library of Object Images, 23, 24
 - amygdala, 266, 268, 354–355, 359
 - analysis by synthesis, 74
 - AND/OR graph, 14, 242–244
 - example of, 243 *f*
 - inference algorithm, 250–252
 - recursive, 251 *f*
 - learning/estimation with, 245–247
 - maximum likelihood learning of Θ , 246
 - relation set, 246–247
 - learning/sampling experiments, 247–250
 - learning by random synthesis, 247–248 *f*
 - predicting missing parts, 248–249 *f*
 - small sample set generalization, 249–250 *f*
 - object recognition experiments, 252–254
 - occluded bicycle detection, 252, 253 *f*
 - ROC curves, 254 *f*
 - parse graph, 243, 244 *f*–245, 252
 - possible pairwise relationships, 244 *f*
 - relational constraints, 243
- annotation, textual, 167
 - appearance-based 3-D object recognition system, 7 *f*
 - appearance-based models, 241
 - shortcomings of, 431–432
 - arousal, 71n3, 460
 - articulated shape, 132–133, 142, 444
 - artificial intelligence (AI), 134–136
 - Aspect, A., 160
 - aspect graph, 6–7, 22
 - attractor network, 270
 - autism spectrum disorder (ASD), 356–357
 - auto-annotation, 118
 - auto-illustration, 118
 - autonomous driving, 57, 59
- bag-of-features model, 18, 369, 401–402
 - bag-of-words model, 404, 431
 - Bajcsy, R., 140
 - banana histogram, 468
 - Bar, M., 386–387, 392, 393, 394
 - barber pole illusion, 310
 - Barnard, K., 168, 169, 173–174, 175–176
 - Barrow, H. G., 503
 - bas-relief ambiguity, 495
 - Baumgartner, A., 136

- Bayes' circle, 152–153, 158
 Bayesian framework, 74
 Bayesian ideal observer theory, 150
 Bayes network, 74, 457
 Bayes' theorem, 151–152
 Belongie, S., 168
 benchmarks
 Amsterdam Library of Object Images, 23, 24
 Brown Dataset 1, 431 *f*
 Caltech-101 database, 23, 58, 60–61, 210, 254, 411, 430 *f*
 Caltech-256 database, 23, 58, 60–61
 categorization algorithms, 23
 Columbia COIL-100 database, 23, 24
 for core object recognition, 41–42, 50
 ETH database, 23, 342 *t*
 MIT “tiny” image database, 23
 need for new, 22–24
 PASCAL'06 dataset, 334–337, 335 *f*–336 *f*
 PASCAL Challenge, 61
 PASCAL Object Recognition Database Collection, 23, 58
 PASCAL Visible Object Classes Challenge, 411–412
 suite-0, 23
 suite-1, 23–24
 Berg, T. L., 172–173, 174–175
 Beymer, D., 507
 biased competition
 neurodynamical model, 275–277
 suppression model, 232
 bi-directional reflection function (BRDF), 495
 Biederman, I., 12, 58–59, 119, 137, 182, 188, 242, 503
 binding problem, 47
 Binford, T. O., 10–11, 503
 binocular disparity, 22n7
 biological vision
 computer vision parallels, 80–81
 just seeing vs. seeing as, 70–73, 80–81
 outer space example, 72
 pedestrian-avoidance example, 72
 mesencephalic vision, 80
 reconciling with computational vision, 192–193
 role of conceptual knowledge in, 78–79
 role of context in, 76–78
 seeing as
 overview, 73–74
 problems with, 74
 as structural, 184–189
 telencephalic vision, 80
 Vapnik-Chervonenkis dimension and, 71–72, 75
 visual task taxonomy, 74–76
 blasting caps, in X-ray images, 144
 Blei, D., 331, 338, 340
 blob graph, 12 *f*, 13
 blocks world scene, 4
 blood-oxygen-level dependent (BOLD)
 response
 confounds, 121 *f*
 contrast stimulus responses, 220–221, 222, 223, 224 *f*, 225
 frontal view responses, 116 *f*, 117
 and persons with autism, 359
 Blum, H., 12–13, 477
 Bogoni, L., 140
 bootstrapping
 cortical face processing, 355
 low-salience camouflaged targets, 418
 motion-based grouping, 316, 321
 region grouping, 308
 static grouping heuristics, 313–315, 314 *f*
 Bosch, A., 411
 Bowyer, K., 131, 137, 140
 brain-behavior correspondences, 353–354, 362
 brain imaging
 electroencephalography, 217
 electrophysiology, 103, 109, 110, 114, 220, 225–227
 magnetoencephalography, 358, 386–387. *See also* functional magnetic resonance imaging
 brain processing pathways, 39
 koniocellular (K) pathway, 385
 magnocellular (M) pathway, 385–386, 390–393, 395 *f*
 parvocellular (P) pathway, 385–386
 Brooks, R., 4, 14
 Brown Dataset 1, 431 *f*
 brown noise, 453
 Brunel, N., 277
 Bülthoff, H. H., 311, 490, 492, 493
 Burbeck, C. A., 480, 482, 484
 Burl, M. C., 326
 CAD models, 5, 144
 Caltech-101 database, 23, 58, 60–61, 210, 254, 411, 430 *f*
 Caltech-256 database, 23, 58, 60–61
 Cao, L., 328
 captions, 167

- categories
 - category sensitivity, 434
 - defining, 55–56
 - formation of, 64
 - labeling, 289
 - satisficing categories, 148–149. *See also*
 - multimodal categorization; object categorization; visual categorization; whole-image categorization
- categorization algorithms, 23
- category learning, 422–423
- causal relations. *See* nonaccidental relations
- causal representation, 139
- central and anterior inferotemporal cortex (CIT/AID), 188
- Chandrasekaran, B., 135
- Chang, S.-F., 412
- change blindness, 268
- Chapelle, O., 29, 176
- character recognition, in reading, 69
- Chelazzi, L., 275, 276
- Chen, H., 512
- children
 - with autism, 357–358
 - category selectivity in ventral stream, 120–123, 121 *f*, 122 *f*
 - category-specific topography of ventral visual pathway, 349–351, 352–356, 360–362
- Chorus of Prototypes, 75
- Chum, O., 341, 412
- Church-Turing Thesis, 71n2
- circular planar right generalized cylinders (circular PRGCs), 92, 95
 - generalized cylinder (GC), 92, 95
- circular PRGCs, 92, 95
- classification, in visual object recognition
 - benchmark datasets for, 58, 60–61
 - error increase with number of categories, 60, 61 *f*
 - of images in collection, 56–57, 59–60
 - scaling in, 60
- class-specific fragments, 432
- clock grammar, 15 *f*
- closed list, 251
- clutter, 46
 - clutter-tolerant object recognition, 47
 - detecting targets in cluttered scenes, 418–420, 419 *f*
 - invariant representations in cluttered backgrounds, 273–275
 - recognition of clutter, 228–229
 - reducing effects in machine vision, 231–232
 - robustness in hierarchical compositionality, 199–200
- coarse representation input, 384–385 *f*
- coarse-to-fine appearance
 - Marr's representation, 14n6
 - modeling with rewrite rules of grammar, 14
- coarse-to-fine model hierarchies, 9n4
- collinearity-based grouping, 315–316
- color/luminance similarity-based grouping, 313–315 *f*
- Columbia COIL-100 database, 23, 24
- common fate, 312
- competitive exclusion, 468
- competitive interaction, 275, 363, 456–457 *f*, 460 *f*, 461, 470
- composition. *See* hierarchical processing
- compositional boosting, 254
- compositionality
 - definition of, 198. *See also* hierarchical processing
- computational model, 73
 - biological vision parallels, 80–81
 - feature hierarchy model for object/face recognition, 260–262
 - and functionality, 136–138
 - inference of 3-D descriptions, 87–88
 - invariant global object motion in dorsal visual system, 279, 280 *f*
 - of invariant visual object/face recognition, 262–265
 - fifth-layer neurons, 281–282
 - response profiles for 2 fourth-layer neurons, 264 *f*, 281
 - VisNet network discrimination comparison, 264 *f*
- of object representation/attention in natural scenes, 269–273
- cortical architecture for
 - hierarchical/attention-based visual perception, 273, 274 *f*
 - inferior temporal cortex model, 270–272, 271 *f*
- probabilistic decision-making, 277–278
- 2 1/2-D sketch, 87
- computational vision, 189–192
 - hidden layer tuning in shape identification network, 189–191, 190 *f*
 - hidden unit structural coding score distribution, 191 *f*–192
 - reconciling with biological vision, 192–193
- conceptual labels, 78 *f*–79
- configural processing, 353, 358, 361–362
- conical computational strategy, 48

- consciousness, hard problem of, 76n12
- constellation model, 297, 431
- construction thesis, 153
- content-based image retrieval (CBIR), 14, 88
- context
- role in biological vision, 76–78. *See also*
 - contextual predictions, for object recognition
- context-based vision system. *See* knowledge-based vision system
- contextual predictions, for object recognition, 393–396
- episodic context frames, 394
 - high spatial frequency, 395, 396
 - low spatial frequency, 394–395 *f*, 396, 397
 - medical prefrontal cortex role, 395–396
 - parahippocampal cortex role, 393, 394, 395, 396
 - prototypical context frames, 394–395
 - retrosplenial complex role, 393, 394, 395, 396
- contour-based representations, 432
- contour feature recognition, simple, 371–377
- algorithm for, 372–374
 - ETHZ dataset example, 375–376 *f*
 - geometric perceptual cues, 379 *f*–380 *t*
 - grouping, 377–381
 - learning steps in, 374–375
 - pairwise geometric relationships, 371, 372 *f*, 373
 - pairwise grouping, 378–381
 - pairwise grouping constraints, 380 *f*–381
- conventional theory of perception
- Bayes' circle, 152–153, 158
 - Bayes' theorem, 151–152
 - construction thesis, 153
 - principle of faithful depiction, 149–150
 - principle of satisficing categories, 148–149, 154
 - reconstruction thesis, 153, 154. *See also*
 - interface theory of perception
- convexity prior, 494, 495 *f*
- Cooke, T., 497
- Cooper, D., 139
- cooperative interaction, 459–460 *f*
- core, 480
- inferior temporal (IT) cortex, 46–47
 - ventral visual stream, 44–46, 45 *f*
- Corel test set, 171 *f*–172, 176
- core object recognition, 41–42, 50
- benchmarks for, 41–42, 50
 - defining, 41
 - goal of brain's recognition machinery, 43–46
- correlations
- in natural scenes, 452–461
 - spatial, in neuronal connectivity, 457–461
- spike count, 460–461
- between words and pictures, 168–170
- correspondence strength, 369
- cortical magnification factor, 271–272, 273
- cortical visual system, 81
- Coyne, B., 168
- Criminisi, A., 505
- Cunningham, H., 174
- curve bundles, 432
- Cuturi, M., 404
- cylinders
- invariant properties of, 95
 - quasi-invariant properties of, 95. *See also*
 - generalized cylinder
- Dalal, N., 327, 330, 337, 341
- Dan, Y., 469
- Darrell, T., 404–405, 406, 412
- decision making
- categorical decisions and learning, 422–423
 - novel pattern categorization, 423, 424 *f*
 - prefrontal cortex role in, 422
 - temporal cortex role in, 422
 - detecting targets in cluttered scenes, 418–420, 419 *f*
 - enhanced neuronal sensitivity during
 - perceptual learning, 420
 - fMRI studies on visual analysis stages, 418
 - perceptual decisions and learning, 417–422
 - V1 learning effects, 417–418
- decision trees, 9n4
- Deco, G., 276
- Defense Advanced Research Projects Agency (DARPA), 18, 19
- deformable template, 98
- deformation paths, 435–437, 436 *f*, 440
- Demirci, M. F., 15–16
- depth cues, 22n7
- Desimone, R., 275, 276, 310
- detection
- clutter detection, 418–419
 - detection problem, 5
 - discriminative detection, 337–338
 - generative/discriminative training, 337
 - sliding-window approach to, 337–338
 - speeding up computation time for, 338
 - edge detection methods, 94
 - experiments, 338–342
 - on estimation/evaluation of approximate inference, 339–340 *f*
 - on ETH shape database, 342 *f*
 - on generative decomposition, 339 *f*
 - on PASCAL'06 VOC challenge, 340–342

- localization task and, 57
- occluded bicycle detection, 252, 253 *f*
- signal detection theory, 24
- Dickinson, S. J., 182–183
- Diehl, R. L., 150, 152, 161
- digital pictures. *See* words and pictures
- dimensionality expansion, 47–48
- Dimitrov, P., 482
- directed acyclic graph (DAG), 14–15
- Dirichlet process, 172, 173 *f*
- discriminant/generative training, 337
- disparity tuning, 455 *f*, 456, 457 *f*, 459–460 *f*
- distributed object-form topography, 119
- Doermann, D., 136
- domain-specific knowledge, 18–10
- dot product encoding, 258
- Duncan, J., 275, 310
- Duric, Z., 137
- Duygulu, P., 169
- Dyer, C., 21
- Dynamic Causal Modeling (DCM), 391–392 *f*
- dynamic inputs (Dylan framework), 311–321
 - appearance model/novel input matching, 320
 - bootstrapping static grouping heuristics, 313, 314 *f*
 - grouping regions into proto-objects, 312–316
 - collinearity-based grouping, 315–316
 - color/luminance similarity-based grouping, 313–315 *f*
 - keyframes, 319–320
 - orienting to interesting regions of visual field, 312
 - redundancy reduction, 317–318
 - temporally extended appearance model, 316–320, 317 *f*
 - tracking proto-objects, 316
- Earth Mover's Distance (EMD) algorithm, 16
- eccentricity bias, 110, 119
- ecological psychology, 141
- Edelman, S., 490
- edge detection methods, 94, 327
- Eikonal equation, 434
- Eklundh, J.-O., 21
- electroencephalography (EEG), 217
- electrophysiology, and monkeys, 103, 109, 110, 114, 220, 225–227
- Elliffe, M. C. M., 265
- embodied agent concept, 496
- epigenetic robotics, 499
- episodic context frames, 394
- Epshtein, B., 290
- ETH database, 23, 342 *t*
- evaluation benchmarks. *See* benchmarks
- exact object geometry, 88
- exemplar graph, 15
- experience-dependent learning, 123, 362–363, 416, 417–418, 421, 425, 468–469, 470
- experience-dependent plasticity, 425
- expert vision system. *See* knowledge-based vision system
- extrastriate body area (EBA), 105, 109, 118
- face detectors, 98
- face inversion effect (FIE), 220
- face-inversion task, 353–354
- face learning, 304–305
- face-processing network, 353–356
- face recognition
 - by age group, 350, 351 *f*, 352–353, 354–355, 361–362
 - benchmark test case, 219–221
 - bootstrapping, 355
 - computational model, 260–265
 - computational model for, 262–265
 - conclusions, 361–363
 - face inversion effect, 220
 - face-processing/face-related behavior, 353–356
 - fearful face recognition, 354–355
 - feature hierarchy model for, 260–262
 - fusiform face area, 105, 109, 110, 118–119, 120, 220, 221, 349
 - fusiform gyrus mechanisms, 351–353
 - invariant, 262–265, 264 *f*
 - Mikolajczyk's face detector, 174
 - occipital face area, 349, 351 *f*, 352, 354, 355, 356, 359, 360
 - by persons with autism, 356–361
 - specialized, 118
 - typical development of face-related cortex, 350–356
- face-related cortex, typical development of, 350–356
- factor graph, 292
- faithful depiction principle, 149–150
- fearful face recognition, 354–355
- feature abstraction. *See* image abstraction
- feature hierarchy model, 260–262, 297
- feature network hierarchy, 273–275
- feedback, 45 *f*, 94, 137, 201–202, 278, 391–392, 418, 421, 461
- feedforward, 188–189, 201–202, 203, 222n2, 227, 263
 - in animal/no-animal detection task, 230–232
 - bottom-up, 275
 - decision making and, 278
 - dynamic causal modeling and, 391–392

- feedforward (*cont.*)
 excitatory, 269
 hierarchical, 273, 279, 282, 354, 386
 and machine learning, 192
 in natural scenes, 217, 452, 461, 462
 in object recognition, 218 *f*–219
 in primates, 216, 217–218
 top-down, 281
 and tuning, 190 *f*, 469
- Fei-Fei, L., 297, 328, 407
- Fergus, R., 8 *f*, 327, 328, 333
- Ferrari, V., 342
- figure-ground segmentation, 107
- flattening, 48
- Fleuret, F., 203
- fMRI-adaption (fMRI-A), 106, 110
- fMRI-RA, 220, 223–225, 224 *f*
- fMRI voxels, 39
- Földiák, P., 261
- Forsyth, D., 169, 172–173
- fovea, 266, 268, 269, 271–272, 273
- FRAME model, 245, 246
- frames, 394
- François, A. R. J., 95, 136
- Freedman, D. J., 397
- Fritz, G., 140, 326, 340
- Fromovich, G., 138
- Frome, F. S., 479
- frontal cortex, 233, 423. *See also* orbitofrontal cortex; prefrontal cortex
- Fukumizu, K., 404
- Fukushima, K., 265
- full object interpretation, 289–295
 informative feature hierarchy
 advantages of hierarchy scheme, 292
 object/part recognition using, 291–292
 objects in motion, 293–294
 part variability/full interpretation, 292, 293 *f*
 semantic features, 293–295
 semantic hierarchy, 294–295
 informative feature hierarchy construction, 290–291
 configuration classification, 295
- functional brain development, 349–350, 351–353, 356, 362–363
- functional connectivity, 392 *f*, 458
- functional organization, in human ventral stream, 117–120
 distributed object-form topography, 119
 process maps, 118–119
 selectivity, 118
 topographic representation, 119
- functionality, recognition based on
 21/2-D, 138
 affordances, 138, 140–141
 functionality in related disciplines, 134–141
 artificial intelligence, 134–136
 cognitive psychology/human perception, 141
 computer vision, 136–138
 robotics, 138–141
- functional ontology for naive mechanics, 135
 Generic Recognition Using Form and Function, 129
 articulated motion of 3-D parts, 132–133
 development of, 131
 reasoning about static 3-D shapes, 131–132
 robotic interaction with objects, 133–134
 sensed range images, 133
- HIPES (history, intentional perspective, physical environment, and event sequences), 141
- motion-to-function mappings, 137
- near-misses, 135–136
- open problem areas, 141–144
 scalability and accuracy, 143
 scalability and efficiency, 142–143
 scalability and learning, 143–144
 successes and limitations to GRUFF, 142
- overview of, 129–131
- s-map, 138
- wayfinding, 141, 143
- functional magnetic resonance imaging (fMRI), 217
 on categorical decisions and learning, 422–423
 on clutter detection, 418–419
 to compare hemodynamic responses, 393
 on functional organization of human ventral stream
 HMAX, 18
 LO, 106, 108, 109, 111 *f*–112 *f*, 117
 on orbitofrontal cortex and top-down predictions, 387
 on visual analysis stages, 418
- functional object descriptions, 16
- functional ontology for naive mechanics (FONM), 135
- functional representation (FR), 135
- fusiform body area (FBA), 105, 109, 118
- fusiform face area (FFA), 105, 109, 110, 118–119, 120, 220, 221, 349
- future research, 123–124
- Gabor elements, 418, 419 *f*

- Gabor filters, 45 *f*, 204, 206, 209–210
 Gabor jets, 513
 Gabor patches (GPs), 480
 Garcia-Lazaro, J., 453
 Gaussian colored noise, 513
 Gaussian dispersion, 152
 Gaussian distribution, 116 *f*, 162, 184, 208, 263, 297, 458
 Gaussian filters, 279
 Gaussian scale space of aperture, 404
 Gaussian transformation, 218 *f*
 Gaussian tuning functions, 185 *f*, 186, 189, 190 *f*, 191
 Gauthier, I., 118–119
 GDL (factor graph) computation, 292
 Geisler, W. S., 150, 152, 161
 Geman, D., 203, 457–458, 461, 466, 471
 generalized cylinder (GC), 5 *f*, 10–11 *f*
 circular planar right generalized cylinders, 92, 95
 and generic object representation, 92
 planar right constant generalized cylinders, 92, 95
 restricted, 12
 generative decomposition, 330–332, 339 *f*
 generative model benefits, 325–326
 generative training, 337
 generic object recognition
 computation of part-based description, 92–93
 description component of, 89
 description issues, 90
 feature extraction, 93–94
 grouping, 94
 hierarchy for description process, 94–95
 perceptual edge detection, 94
 surface patches, 94–95
 learning component of, 89
 matching component of, 89
 object instance recognition *vs.*, 431
 organization of description process, 93 *f*
 part-based representations, 90–92
 principled approach to, 89–90
 limitations of, 98–99
 retrieval process, 94
 selection, 94
 tasks in, 87, 89 *f*, 93–94
 2-D appearance-based approaches, 88
 3-D model-based approaches, 88
 3-D shape-based approaches, 88
 using part-based descriptions, 95–98, 97 *f*
 volumetric part-based descriptions, 91 *f*–92, 95
 Generic Recognition Using Form and Function (GRUFF), 129
 articulated motion of 3-D parts, 132–133
 development of, 131
 reasoning about static 3-D shapes, 131–132
 robotic interaction with objects, 133–134
 sensed range images, 133
 success of, 142
 geodesic intensity histogram (GIH), 510
 geometric invariants, 90
 geometric matching, 370
 geons, 11 *f*, 12, 13, 108, 242
 Gestalt laws, 198, 309, 315, 316, 361, 451
 Gibbs sampling, 331–332
 Giblin, P. J., 435
 Gibson, J., 141
 Golarai, G., 353, 354
 Google re-ranking test, 333*t*–334 *f*
 grammar-based method, 14, 15 *f*
 grammars, 241
 graph
 And/OR (*see* AND/OR graph)
 aspect graph, 6–7, 22
 blob graph, 12 *f*, 13
 directed acyclic graph, 14–15
 exemplar graph, 15
 factor graph, 292
 for managing uncertainty, 74
 medial axis graph, 433–434 *f*
 parse graph, 243, 244 *f*–245, 252
 region adjacency graph, 15
 shock graph (*see* shock graph)
 graph grammars, 241
 grassfire flow approach, 477, 482–483
 Grauman, K., 404–405, 406, 412
 Griffiths, T. L., 331, 338, 339, 340
 Gurevich, N., 135–136
 Haar-feature-based Adaboost model, 254 *f*
 HabCam computer vision system, 80n18
 hallucinations, 470
 haptic modality, 497
 Harris-Affine feature detector, 430
 Harris-Laplace feature detector, 430
 hashing, 9n4
 Haxby, J. V., 119
 He, X., 17 *f*
 Hebbian-like mechanism, 461, 463
 Hebb rule, 261, 263–264 *f*
 herring gulls, 155
 hidden Markov model (HMM), 463, 465–466, 468, 469
 hidden state shape model (HSSM), 14

- Hierarchical Model, 18
- hierarchical processing, 196, 197, 198–201
 experimental results, 209–212
 expressive power, 200
 feedforward and feedback, 200–201
 indexing/matching scheme implementation, 199, 202–203
- learning compositional hierarchy of parts, 203–209
 category-specific higher layers, 209
 compositional library of parts, 204–205 *f*
 grouping labels by similarity/
 co-occurrence, 208–209
 indexibility constraint, 208
 indexing and matching scheme, 205–207;
 hierarchical recognition architecture, 206 *f*–207; processing with Layer 1, 205–207
 local inhibition, 207
 unsupervised learning of part compositions, 207–208
- processing complexity, 199
 receptive field, 199
 robustness to clutter, 199–200
 storage requirements, 199. *See also*
 statistical, bottom-up learning
- hierarchy
 description process, 94–95
 feature hierarchy model, 260–262, 297
 feature network hierarchy, 273–275
 informative feature hierarchy
 advantages of scheme, 292
 configuration classification, 295
 construction of, 290–291
 object/part recognition using, 291–292
 objects in motion, 293–294
 part variability/full interpretation, 292, 293 *f*
 semantic features, 293–295
 semantic hierarchy, 294–295
- IS-A hierarchy, 20
- segmentation hierarchy, 432. *See also*
 hierarchical processing
- high-level vision, 70
 high-order primitive shape abstraction, 16
 high spatial frequency (HSF), 395, 396
 Hillel, A. B., 326
 HIPES (history, intentional perspective, physical environment, and event sequences), 141
 histogram pyramid, 18
 Hodges, J., 135
 Hoffman, T., 150, 331
 Hofmann, K. P., 168
 HOG, 199, 254 *f*, 327, 330
 Holub, A. D., 326
- horizontal connections
 spatial relationships between parts, 466–469
 temporal association of views, 461–466
- human-machine interface, 64
- Hwang, V., 19 *f*
- iconographic senses, 176–177
- idealism, 159
- identity-preserving image variation, 49 *f*, 50
- image abstraction process, 1–3
- image collections, analyzing/indexing, 56–57, 59–60
- image gradient, 327, 343, 512–513
- image hypotheses, 470
- image-keyword search, 168
- inattentive blindness, 268
- incremental learning, 64, 202
- incremental model selection, 295–299
 closing performance gap using, 295–296
 cross-generalization and, 299
 family combination schemes, 297–298
 minimum description length (MDL) in, 298
 model family selection, 297
 structural risk minimization (SRM) in, 298
- indexing structures, 9–10, 9n4, 17
- indirect realism, 159
- ineffability, 78–79
- inference
 of 3-D descriptions, 87–88
 estimation/evaluation of approximate, 339–340 *f*
 inference algorithm, 250–252
 recursive inference algorithm, 251 *f*
- inferior temporal (IT) cortex model, 46–47, 270–272, 271 *f*
- inferior temporal (IT) neurons, 46–47, 49 *f*
- informative feature hierarchy
 advantages of scheme, 292
 configuration classification, 295
 construction of, 290–291
 object/part recognition using, 291–292
 objects in motion, 293–294
 part variability/full interpretation, 292, 293 *f*
 semantic features, 293–295
 semantic hierarchy, 294–295
- instantiated computational model, 50
- integrate-and-fire model, 276
- interaction signature, 137
- interactive specialization, 349, 352, 356
- interface theory of perception, 154–160
 future research issues, 160–161
 interface games, 161–163
 interface-world constraints, 158–160
 conventionalist objections to, 159–160

- user interfaces in nature, 155–158
 - conventionalist objections to, 156–158
 - herring gulls, 155, 156
 - jewel beetles, 151, 155, 156
 - jumping spiders, 156
 - intermediate-level representations, 432
 - interpretation tree, 369
 - invariance transforms, 470
 - invariant object recognition, 103, 108–109
 - invariant visual object and face recognition (VisNet), 262–265, 281–282
 - inverse optics, 73–74
 - IS-A hierarchy, 20

 - Jacobs, D. W., 508, 510
 - Jameson, K. A., 161
 - Janowicz, K., 143–144
 - jewel beetles, 151, 155, 156
 - Johnson, S., 157
 - Jones, M., 98
 - Josephson, J. R., 135
 - Joshi, S., 168, 444
 - Julesz, B., 480, 484
 - jumping spiders, 156
 - just seeing vs. seeing as, 70–73, 80–81
 - outer space example, 72
 - pedestrian-avoidance example, 72

 - Kalman filter, 316
 - Kanizsa, G., 308
 - Kanwisher, N., 118
 - kd-trees, 9n4
 - Kellman, P. J., 309
 - Kesselman, Y., 15
 - keyframes, 319–320, 492–494
 - keypoint-based representations, 432
 - k-fan model, 297, 370, 431
 - Khan, E. A., 494, 495
 - Kim, D., 138
 - Kimia, B. B., 435, 438
 - Kimmel, R., 508
 - Klatsky, R. L., 141
 - K*-means clustering, 468–469
 - knowledge-based vision system, 18–20
 - knowledge primitives, 131–132
 - Koenderink, J., 6, 404
 - Koffka, K., 308
 - Komarova, N. L., 161
 - koniocellular (K) pathway, 385
 - Kourtzi, Z., 418
 - Kovács, I., 480, 484
 - Kullback-Leibler divergence, 245

 - LabelMe, 79
 - Lafon, S., 468
 - Lambertian surfaces, 471, 512
 - Laplacian distributions, 458–459
 - Laptev, I., 327, 412
 - large-scale representations, 202
 - lateral connectivity, 465
 - lateral occipital complex (LO), 104, 106, 108, 109, 111*f*–112*f*, 117, 359
 - lateral occipital cortex (LOC), 104–105
 - cue-invariant responses, 105–108, 106*f*, 107*f*
 - distributed patterns, 110–112*f*, 111*f*
 - driven by global shape/surface information, 106–107
 - driven by shape/form, 106–107
 - subregion sensitivity to border ownership, 107
 - cue-invariant responses in, 105–108, 106*f*, 107*f*
 - distributed response patterns, 110–112*f*, 111*f*
 - driven by global shape vs. surface information, 106–107
 - driven by shape vs. form, 106–107
 - subregion sensitivity to border ownership, 107
- defining, 118
- object and position information in, 109–113
 - eccentricity effect, 110
 - insensitivity to, 110
 - polar angle effect, 110
 - reduced action, 110
 - size and position invariance, 109
- selectivity in, 223, 224*f*
- theoretical implications, 113
- viewpoint sensitivity across, 113–117
 - as function of rotation level, 114–115*f*
 - implications for object recognition theory, 117
 - simulated responses in voxel, 115–117, 116*f*
- Lazebnik, S., 402, 412
- learning, 328–329
 - bottom-up, 196, 197–198, 201–202
 - category, 422–423
 - experience-dependent, 123, 362–363, 416, 417–418, 421, 425, 468–469, 470
 - future research, 343
 - incremental, 64, 202
 - preconceptual, 417–418, 420
 - semi-supervised, 329
 - statistical, 196, 203
 - supervised, 328

- learning (*cont.*)
 unsupervised, 64, 197, 198, 201–202, 207–208, 328, 329
 weakly supervised, 328
- learning structured appearance model, 21 *f*
- Lebanon, G., 404
- Lee, A. B., 468
- Lee, M. W., 99, 482
- Lehar, S., 150
- Leibe, B., 327
- Levi, K., 327
- Leyton, M., 475–476
- lighting variations, in 2-D image/3-D object comparison, 511–513
- Ling, H., 508, 510
- linguistic knowledge, 20
- liquid-state machines, 82n20
- Liu, G., 412
- local features, 326–327
 abstraction mechanisms, 17–18
 bag-of-features model, 18, 369, 401–402
 challenges using, 318, 327
 grouping and abstracting, 13, 17 *f*, 197
 invariant, 403
 machine learning and, 492, 493
 mapping, 25
 neuronal coding of, 109
 pyramid matching and, 402, 406–407
 reliance on by child with autism, 357
 re-ranking images using, 333
 scope of, 8–9
 shaped-based, 1–3, 8
- local image statistics, 17–18
- locally orderless images, 404
- local realism, 160
- Loeff, N., 170
- Logothetis, N. K., 217
- long-lagged fMRI-adaption (fMRI-A), 114
- long-term memory, 80
- Long-term potentiation (LTP), 264
- Lotto, R. B., 150
- Lotus Hill Database, 254
- Lowe, D. G., 5–6 *f*, 202–203, 340
- lowest common abstraction, 15
- low-level vision, 70
- low spatial frequencies (LSF), 386, 388–390, 389 *f*, 394–395 *f*, 396, 397
- LSF filtering, 394–395
- machine learning
 prior work in, 303–307
 from semi-supervised models, 305–307
- machine vision
 extraction in, 230
- invariance and specificity in, 216–217
 reducing effects of clutter, 231–232
 weak activation in, 229. *See also*
 computational model; dynamic inputs
- magnetoencephalography (MEG), 358, 386–387
- magnocellular (M) pathway, 385–386, 390–393, 395 *f*
- Malach, R., 119
- Malik, J., 401
- Malladi, R., 508
- Maloof, M., 144
- many-to-many node correspondences, 15–16
- Markov random-field model, 458, 471
- Markov transition matrix, 466, 468–469
- Marr, D., 14n6, 40, 70, 73, 137, 182, 188, 455, 503
- Marszalek, M., 412
- material composition, 133
- mathematical learning theory, 297
- MAX, 218 *f*, 228, 282
- maximum likelihood learning of Θ , 246
- Mayer, H., 136
- medial axis transform (MAT), 12–13
- medial prefrontal cortex (MPFC), 393, 395–396
- medial representations of objects, 475–479
 Blum medial locus, 476
 external medial locus, 476, 477
 grassfire flow approach, 477, 482–483
 internal medial locus, 476, 477
 internal medial locus of 3-D object, 477–478 *f*
- local geometry of medial loci, 478–479 *f*
- maximal inscribed ball, 476
- medial loci in human vision, 479–482
- medial locus of 2-D objects, 476
- medial locus of 3-D objects, 476
- medial point of Ω , 476
- outward flux and object angle, 483–485 *f*
- Process Induced Symmetric Axis, 475–476
- shocks in, 477
- topological properties of medial loci, 477–478
 wiggle stimuli, 480–482, 481 *f*
- medial surface graph-based recognition system, 12–13
- Medioni, G., 95, 136
- MEG, 359, 388–389 *f*
- memory
 long-term, 80
 memory-intensive strategy, 48
 short-term, 261, 271, 278, 279, 282
 working memory, 276, 277. *See also*
 memory-based predictions

- memory-based predictions
 - advantages of top-down predictors, 396–398
 - coarse representation input, 384–385 *f*
 - context role in, 393–396
 - episodic context frames, 394
 - high spatial frequency, 395, 396
 - low spatial frequency, 394–395 *f*, 396, 397
 - medical prefrontal cortex role, 395–396
 - parahippocampal cortex role, 393, 394, 395, 396
 - prototypical context frames, 394–395
 - retrosplenial complex role, 393, 394, 395, 396
- Dynamic Causal Modeling, 391–392 *f*
- koniocellular (K) pathway, 385
- magnocellular (M) pathway, 385–386, 390–393, 395 *f*
- M and P condition comparison, 391 *f*
- orbitofrontal cortex
 - low spatial frequencies, 388–390, 389 *f*
 - role in top-down predictions, 386–387
 - top-down facilitation model, 387–388
- parvocellular (P) pathway, 385–386
- phase-locking analysis of MEG data, 388–389 *f*
- top-down facilitation model, 387–388
- memory-intensive strategy, 48
- Mercer kernels, 405
- mesencephalic vision, 80
- middle temporal (MT) visual area, 104, 106 *f*, 231, 274 *f*, 276, 279, 280 *f*, 310
- mid-level primitives, 17
- mid-level representations, 432
- Mikolajczyk, K., 174, 326, 327, 340, 510
- Mikolajczyk's face detector, 174
- Miller, E., 275, 397
- Milward, T., 264
- mimicry, 75
- minimax learning algorithm, 242
- minimum-length description (MDL) codes, 469–470
- Minsky, M., 129
- Mirmehdi, M., 136
- missing parts, predicting using learned models, 248–249 *f*
- MIT "tiny" image database, 23
- Miyashita, Y., 311
- model paradigm, 325–326
 - discriminative model benefits, 326
 - future research, 343
 - generative model benefits, 325–326
 - hybrid methods, 326
- Mohan, R., 327
- monkey
 - object representation neurophysiology, 257–260
 - structural encoding, 186, 188
- Moratz, R., 137–138, 141, 142–143
- motion
 - as depth cue, 22n7
 - as visual clue, 307–308
- motion-based grouping, 309–310
- motion-to-function mappings, 137
- MRP, 242. *See also* AND/OR graph
- Mukerjee, A., 135
- multidimensional scaling (MDS), 497
- multidimensional similarity space, 76n12
- multimodal categorization
 - computational issues, 499
 - embodied agent concept, 496
 - material properties perception, 494–496
 - bas-relief ambiguity, 495
 - bi-directional reflection function, 495
 - multisensory object processing, 496–499
 - perception-action cycle, 496 *f*
 - visuohaptic similarity ratings, 497–499
- object representations suitable for, 500
- perceptual research/computer graphics/computer vision fusion, 489 *f*
- spatiotemporal object representations, 490–494
 - exemplar-based keyframe framework, 492–494
 - exemplar-based *vs.* structural object representations, 490–491
 - Recognition-by-Components model, 490
 - temporal information for object recognition, 491–492
- multiscale blobs, 13
- multisensory object processing, 496–499
- multi-voxel pattern analysis, 110–111, 113
- Mumford, D., 15 *f*, 203, 469–470, 471
- Murase, H., 7 *f*
- Mutch, J., 202–203, 340
- naïve realism, 159
- Narens, L., 161
- natural noise, 453
- natural selection, 151
- natural vision. *See* biological vision
- Nayar, S. K., 7 *f*
- near-misses, 135–136
- neural coding
 - adaptive, 422. *See also* scene statistics, neural coding of
- neural connectivity, 361

- neural plasticity, 225, 233, 362, 416, 418–419, 421
- neurodynamical models for biased competition, 275–277
- in attention context, 276
 - biased competition hypothesis, 275
 - using integrate-and-fire model, 276
 - in working memory context, 276, 277
- neurophysiology, of object representation
- in complex natural scenes, 265–269
 - attention against blank background, 266, 267 *f*, 270–271
 - attention/selection, 265–266, 267 *f*
 - multiple objects in scene, 269, 270 *f*
 - representation/action interface, 266, 268
 - in inferior temporal visual cortex in macaque, 257–260
 - firing rates of neurons in IT, 257–258, 259 *f*
 - primate properties leading to IT, 257, 258 *f*
- Nevatia, R., 95, 138, 340
- Newell, F., 497
- Niebur, E., 276
- Nöe, A., 150
- noise
- brown noise, 453
 - in datasets, 167
 - Gaussian colored, 513
 - pink (natural) noise, 453
 - recognition within, 278, 384, 417, 420, 421
 - types of, 453
 - white noise, 48, 453
- nonaccidental relations, 1, 2
- non-action (*wu wei*), 72
- nonrigid object recognition, 88
- NP-complete subgraph isomorphism problem, 97–98
- null model of object representation, 43 *f*
- object angle, 478
- object categorization
- clutter recognition, 228–229
 - future research questions, 232–234
 - humans
 - benchmark test case of faces, 219–221
 - discrimination tasks, 223–224, 225 *f*
 - face inversion effect, 220
 - fMRI data, 222–225
 - selectivity, 223, 224 *f*
 - weak activation in, 229
 - machine vision
 - extraction in, 230
 - invariance and specificity in, 216–217
 - reducing effects of clutter, 231–232
 - weak activation in, 229
- primates
- feedforward processing, 217, 219
 - IT monkey electrophysiology experiments, 220
 - mediation of object recognition in, 217
 - monkey electrophysiology data, 225–227
 - Standard Model of object recognition in cortex, 217, 218 *f*, 219
- object category
- blasting caps, 144
 - chair, 14, 16, 58 *f*, 65 *f*, 129–130 *f*, 131–132, 135, 137–138, 139, 393, 409 *f*
 - cups, 131, 132, 432
 - defining, 241–242
 - documents, 136, 332, 404
 - scissors, 132–133
- object class discovery, 334–337, 335 *f*–336 *f*
- object class recognition. *See* generic object recognition
- object class representation, 229–230
- object discovery
- developmental perspective on, 301–303
 - human visual processing relevant to, 307–311
 - barber pole illusion, 310
 - grouping diverse object instances, 310–311
 - grouping region into larger assemblies, 308–310
 - motion-based grouping, 309–310
 - orienting to interesting location of visual field, 307–308
 - rod-and-frame task, 309 *f*
 - prior work in machine learning, 303–307
 - learning from normalized images, 304–305
 - learning from semi-supervised models, 305–307
 - problem of, 301
 - theoretical framework (Dylan), 311–321
 - bootstrapping static grouping heuristics, 313, 314 *f*
 - grouping regions into proto-objects, 312–316; collinearity-based grouping, 315–316; color/luminance similarity-based grouping, 313–315 *f*
 - keyframes, 319–320
 - matching appearance models against novel inputs, 320
 - orienting to interesting regions of visual field, 312
 - redundancy reduction, 317–318
 - temporally extended appearance model, 316–320, 317 *f*
 - tracking proto-objects, 316

- object recognition
 - appearance-based 3-D system, 7*f*
 - aspect graph, 6–7
 - Brooks' ACRONYM system, 4, 5*f*, 14
 - component elements, 55
 - evolution of, 3–10, 4*f*, 182–184
 - functional goal of, 41
 - Lowe's SCERPO system, 5–6*f*
 - overall scene, 55
 - region-based hierarchical object model, 9, 10*f*
 - scale-invariant parts-based object model, 8*f*
 - scale-invariant patch, 8–9
- object representation
 - computational model
 - and attention in natural scenes, 269–273
 - hierarchical/attention-based visual perception, 273, 274*f*
 - inferior temporal cortex model, 270–272, 271*f*
 - generic object representation, 92
 - multimodal categorization
 - exemplar-based keyframe framework, 492–494
 - exemplar-based vs. structural object representations, 490–491
 - object representations suitable for, 500
 - Recognition-by-Components model, 490
 - spatiotemporal object representations, 490–494
 - temporal information for object recognition, 491–492
 - neurophysiology in complex natural scenes, 265–269
 - attention against blank background, 266, 267*f*, 270–271
 - attention/selection, 265–266, 267*f*
 - relative positions of multiple objects in scene, 269, 270*f*
 - representation/action interface, 266, 268
 - neurophysiology in inferior temporal visual cortex, 257–260
 - firing rates of neurons, 257–258, 259*f*
 - null model of object representation, 43*f*
 - primate ventral visual system properties leading to IT, 257, 258*f*
 - surface/object inference, 461–469
 - object structure, 140, 183, 184, 188, 189, 192–193. *See also* hierarchical processing
 - occipital face area (OFA), 349, 351*f*, 352, 354, 355, 356, 359, 360
 - 1/*f* noise, 452*f*, 453
 - one-shot learning, 64
 - one-to-one correspondence assumption, 2–3, 6
 - OpenCV, 98
 - open-ended representational system, 82n20
 - open list, 251
 - open source named entity generator, 174
 - optic flow map, 312*f*
 - orbitofrontal cortex (OFC)
 - low spatial frequencies, 388–390, 389*f*
 - role in top-down predictions, 386–387
 - top-down facilitation model of object recognition, 387–388
 - orderless texture histogram, 401
 - pairwise geometric relationships, 371, 372*f*, 373
 - pairwise grouping, 378–381
 - constraints, 380*f*–381
 - pairwise relationships, 244*f*
 - Paletta, L., 140
 - Palmer, P. L., 150–151
 - “paperclip” neurons, 217, 228
 - paradigm integration, 328–329, 338–342
 - parahippocampal cortex (PHC), 393, 394, 395, 396
 - parahippocampal place area (PPA), 105, 359
 - parameterized geometric models, 135
 - parametric approach, 114
 - parametric models for object classification, 297
 - parse graph (pg), 243, 244*f*–245, 252
 - part-based descriptions, 95–98
 - hierarchical database organization, 96, 97*f*
 - object description retrieval, 96–98
 - retrieval process, 97–98
 - transaction costs, 96–97
 - symbolic, hierarchical descriptions, 96
 - volumetric, 91*f*–92, 95
 - part-based extraction, 88
 - segmentation into parts, 92
 - volumes vs. surfaces, 91–92
 - partial occlusion, 437*f*
 - part recognition, in industrial setting, 69
 - parts-based abstraction, 1
 - parts-based approach, 431
 - parvocellular (P) pathway, 385–386
 - PASCAL'06 dataset, 334–337, 335*f*–336*f*
 - PASCAL Challenge, 61, 338, 340*f*, 376, 412
 - PASCAL Object Recognition Database Collection, 23, 58
 - PASCAL Visible Object Classes Challenge, 411–412
 - Passarotti, A. M., 353
 - pedestrian-avoidance system, installed in car, 72

- perception, primary function of, 81–82
 perception-action loop, 64–65
 perceptual categorization theories, 221–222
 perceptual grouping, 437–441
 perceptual learning, 417–418, 420
 Perdeck, A. C., 155
 performance evaluation, 23, 133, 171–172, 332, 339–340, 343
 Perona, P., 407
 Peursum, P., 137
 p/FUS/OTS, 104
 phase-locking analysis of MEG data, 388–389 *f*
 philosophical investigations
 of Wittgenstein, 70, 73
 Of Poggio, 70, 203, 310, 455, 507
 photographs
 analyzing/indexing, 56–57. *See also* words and pictures
 pink noise, 453
 PIT tuning for object-centered position, 186, 188 *f*
 Pizer, S., 444, 482, 483
 planar right constant generalized cylinders (PRCGCs), 92, 95
 Plantinga, H., 21
 Platonism, 79n16
 pLSA model, 328
 Poggio, T., 70, 203, 310, 455, 507
 Poisson PDE, 444
 polar angle effect, 110
 polyhedral shapes, 5–6 *f*, 131, 132, 484
 Ponce, J., 95
 population pattern of activity, 44
 position-tolerant object recognition, 47
 possible pairwise relationships, 244 *f*
 posterior inferior temporal cortex (TEO), 260
 Potter, M. C., 386
 Pourtois, G., 355
 prefrontal cortex (PFC), 217
 medial prefrontal cortex, 393, 395–396
 role in decision making, 422
 PRGC, 95
 primates
 feedforward processing, 217, 219
 IT monkey electrophysiology experiments, 220
 mediation of object recognition in, 217
 monkey electrophysiology data, 225–227
 view-independent recognition, 262
 primitives
 high-order primitive shape abstraction, 16
 knowledge, 131–132
 mid-level, 17
 principle of satisficing categories, 148–149
 priors
 compositional binding, 202
 Dirichlet, 331, 332
 learning and encoding, 461, 469, 471–472, 494
 mid-level shape, 13
 verifying, 152–153, 169
 probabilistic topic model, 330–332
 probability theory, 158–159
 Probabilistic Boosting Tree, 254
 probabilistic models, 327
 Process Induced Symmetric Axis (PISA), 475–476
 process maps, 118–119
 prosopagnosia, 102
 proto-objects
 grouping regions into, 312–316
 tracking, 317
 prototype-tuned filters, 75
 prototypical context frames, 394–395
 Psotka, J., 479–480
 psychophysical methods, 24
 purposive vision, 74
 Purves, D., 150
 Puzicha, J., 168
 pyramid matching, 402, 404–405. *See also* spatial pyramid matching
 pyramid match kernel, 405
 quadratic assignment problem (QAP), 372, 373–374
 random forest classification, 411
 random synthesis, 247–248 *f*
 range images, 137–138
 rarefaction waves, 434
 ratio template, 304
 rat tracing, 152
 Raubal, M., 141, 142–144
 Ray, D., 411
 read-out neurons, 44
 read-out tools, 44
 Recognition-by-Components (RBC) model, 490
 Recognition by Components theory, 242
 recognition using simple contour features, 371–377
 algorithm for, 372–374
 ETHZ dataset example, 375–376 *f*
 geometric perceptual cues, 379 *f*–380 *t*
 grouping, 377–381
 learning steps in, 374–375
 pairwise geometric relationships for, 371, 372 *f*, 373
 pairwise grouping, 378–381

- pairwise grouping constraints, 380*f*–381
 PASCAL dataset example, 375*t*
 quadratic assignment problem, 372,
 373–374
 reconstruction thesis, 153, 154
 recurrent connections, 452, 458, 461
 recursive inference algorithm, 251*f*
 redundancy reduction, 317–318
 Regan, J. K., 150
 region adjacency graph, 15
 regional integration, 308*f*–309
 regional labels, 167
 region-based hierarchical object model, 9,
 10*f*
 region-based representations, 432
 region entropy, 173–174
 region grouping, 308
 into larger assembly, 308–310
 into proto-object, 312–316
 region-based hierarchical object model, 9,
 10*f*
 Reitman, W., 81–82
 relational constraints, 243
 Renninger, L., 401
 representational gap, 4, 8, 10
 representation paradigm, 326–328
 dense representation, 327–328
 future research, 343
 global representation, 326–327
 local representation, 326–327
 semilocal representation, 327
 sparse features, 327
 response
 attention-reduced, 232
 cue-invariant, 105–108, 106*f*, 107*f*
 simulated putative neural, 115–117, 116*f*
 restricted generalized cylinder (GC), 12
 retinotopic area, 104*f*, 112*f*, 186, 188, 271,
 386, 418, 462, 468
 retinotopic maps, 112
 retrosplenial complex (RSC), 393, 394, 395, 396
 rewrite rules of grammar, 14
 Reynolds, J. H., 276
 rhesus monkey visual system, 45*f*
 Riemannian metric, 435
 Riesenhuber, M., 203
 right lateral PFC (rLPFC), 223
 Rivlin, E., 139
 robotic interaction, 131, 133–134, 139, 141,
 142
 robotics, 73, 136–138
 ROC curves, 254*f*
 rod-and-frame task, 309*f*
 Rolls, E. T., 264, 266, 269, 270–271
 Rosenfeld, A., 139
 Russell, C., 328
 Sahin, E., 140
 saliency map, 232
 satisficing categories, 148–149, 154
 scale-invariant parts-based object model, 8*f*
 scale-invariant patch, 8–9
 scene interpretation, 20, 74*n*9, 458
 scene reconstruction
 and seeing as, 74
 scenes, 170
 scene segmentation, 79*f*
 scene statistics
 global, 411, 412–413
 scene statistics, neural coding of
 computational implications on object
 representation, 461–469
 constraints in neuronal interaction, 455–457
 correlations in natural scenes, 452–461
 spatial correlations in neuronal connectivity,
 457–461
 spatial relationships between parts, 466–469
 structures in tuning properties, 452–455
 temporal association of views, 461–466
 scene understanding, 64–66
 SCERPO system, 5–6*f*
 SCFG, 242, 245
 schemata, 394
 Scherf, K. S., 354
 Schmid, C., 26, 510
 scientific physicalism, 159
 scripts, 394
 search by content, 14
 second-order geometric relationships, 370
 seeing as
 just seeing vs., 70–73, 80–81
 overview, 73–74
 problems with, 74
 segmentation, figure-ground, 107
 segmentation hierarchy, 432
 segmentation model, 17*f*, 18, 183
 semantic model, 20
 sensitive period, for brain development,
 362–363
 Serre, T., 230
 shape abstraction, 10–14
 content-based image retrieval, 14
 functional object descriptions, 16
 generalized cylinder, 5*f*, 10–11*f*, 11
 geons, 11*f*, 12, 13
 high-order primitive, 16
 medial axis transform, 12–13
 shock graph, 12*f*, 13

- shape abstraction (*cont.*)
 superquadric ellipsoids, 11*f*, 13
 symmetry-based abstraction, 13
- shape concepts, 69–70
- shaped-based local features, 1–3, 8
- shaped-based models
 contour-based representations, 432
 keypoint-based representations, 432
 region-based representations, 432
- shape grammar, 15*f*, 16
- shape perception, 64, 163, 479–482
- sharpening, 223, 418, 456, 460
- SHGC, 95
- shock graph (SG), 12*f*, 13
- shock-graph (SG) representation, 432–433
 bottom-up recognition of shapes embedded in
 figures, 441–442
 category sensitivity, 434
 contact shocks, 434
 deformation paths, 435–437, 436*f*, 440
 dimensionality of model, 443
 dynamic vs. static representation, 435
 expressiveness of model, 443
 of figure-ground segregated shape, 433–437
 gap transform, 439
 loop transform, 439
 medial axis (MA) graph vs., 433–434*f*
 partial occlusion, 437*f*
 perceptual grouping, 437–441
 rarefaction waves, 434
 shock curve segments, 434
 shock graph loops, 440–441*f*
 shocks, 434
 spurious edges, 439
 top-down organization using, 443–445
 visual fragment, 438
- Shokoufandeh, A., 14–15
- short-lagged fMRI-adaption (fMRI-A), 114
- short-term memory, 261, 271, 278, 279, 282
- Siddiqi, K., 482, 483
- SIFT descriptor, 199, 254, 327, 407, 430
- SIGMA image understanding system, 19*f*
- Sigman, M., 418
- signal detection theory, 24
- Sinha, 310
- SLAM, 88
- sliding-window approach, 337–338
- s-map, 138
- Smart, J. J. C., 76n12
- Sobel operators, 315
- Sochen, N., 508
- sparse feature vector, 306
- sparseness constraint, 48
- spatial pyramid kernel, 411
- spatial pyramid matching, 403–404, 405–407,
 406*f*
 extensions/applications
 adaptive weighting scheme, 411
 combining multiple feature “channels,”
 411, 412
 extending to video, 411, 412
 feature and space covariant kernel, 412
 image subwindow precision, 411, 412
 random forest classification, 411
 temporally aligned pyramid matching, 412
 limitations of, 412–413
- spatial pyramid representation, 402*f*
- spatiotemporal object representations, 490–494
- spectral graph theory, 14
- Spelke, E. S., 309
- spike count correlation, 460–461
- spike-timing dependent plasticity learning rule,
 463
- spiking patterns, 39
- Sproat, R., 168
- Standard Model, of object recognition in cortex,
 217, 218*f*, 219
- Stark, L., 131, 133, 137, 140
- star model, 297, 370
- static grouping heuristics, 313, 314*f*
- statistical, bottom-up learning, 196, 197–198,
 201–202
- statistical correlations, 451, 452–461
- statistical priors, 494, 495*f*
- stereo, 106*f*, 107*f*, 108, 139, 143, 503,
 505–507*t*
- stereoscopy, 63–64
- Steyvers, M., 331, 338, 339, 340
- stochastic grammar, 15*f*, 74n9
- Stringer, S. M., 264, 270–271, 279
- structural abstraction, 14–16
 ACRONYM system, 14
 AND/OR graph, 14
 directed acyclic graph, 14–15
 exemplar graph, 15
 grammar-based method of, 14, 16
 hidden state shape models, 14
- structural representation of object shape in brain
 parts-based structural coding in brain,
 184–189
 PIT tuning for object-centered position,
 186, 188*f*
 shape reconstruction from V4 neural
 population response, 186, 187*f*
 V4 shape tuning, 184–186, 185*f*
- structure-based models, 241
- subdivide-and-disorder techniques, 403–404
- suite-0, 23

- suite-1, 23–24
 sum of square differences (SSDs), 510
 superior colliculus, 80
 superior temporal sulcus (STS), 349, 351*f*, 354, 355, 356, 359, 360
 supernormal stimulus, 155
 superpixels, 432
 superquadric ellipsoids, 11*f*, 13
 surface-based representation, 91
 surface smoothness constraint, 458
 suspicious coincidence, 468
 Sutton, M. A., 133
 symmetry-based abstraction, 13
 symmetry-curvature duality principle, 475–476
 synthesis, random, 247–248*f*
 synthesis methods, 74
 systems identification framework, 39

 Tamraker, A., 438
 Tanaka, K., 261–262
 target recognition problem, 5
 Tarr, M. J., 118–119
 task-independent image description, 65
 taxonomization, 64, 66–67
 telencephalic vision, 80
 template, object, 69
 temporal association, 310–311
 temporal association matrix, 466–468, 467*f*
 temporal image evolution, 48–49
 temporally aligned pyramid matching, 412
 temporally extended appearance model (TEAM), 316–320, 317*f*
 Tenbrink, T., 138, 143
 Tenenbaum, J. M., 503
 texture patterns, 105, 106
 3-D, 20–22
 appearance-based object recognition system, 7*f*
 articulated motion of 3-D parts, 132–133
 compactness constraint, 22
 computer vision inference of 3-D descriptions, 87–88
 inferring 3-D shape from 2-D appearance, 22*f*
 internal medial locus of 3-D object, 477–478*f*
 medial locus of 3-D objects, 476
 model-/shape-based approaches, 88
 static 3-D shapes, 131–132
 Tinbergen, N., 155
 Todorovic, S., 9
 top-down organization, 443–445
 Torralba, A., 326, 334
 trace-learning rule, 273, 463

 transfer function, 39
 transformation voting alignment, 369
 Trappenberg, 270–271
 Trappenberg, T. P., 270–271
 TRECVID dataset, 412
 Triggs, B., 203, 327, 330, 337, 341
 Trivers, R. L., 151
 Tsotsos, J. K., 196
 2-D image/3-D object comparison
 deformations and articulations, 508–511
 drawbacks to, 503–504
 early examples of, 503
 images with pose variation, 505–508
 epipolar constraint, 505
 occlusion, 505
 ordering constraint, 505
 lighting variations, 511–513
 using 2-D images to reconstruct 3-D properties, 503–504
 using prior experience and machine learning, 504
 21/2-D, 138

 UIUC single-scale database, 338
 Ullman, S., 290, 369, 394
 Umwelt (self-world), 156
 uniqueness constraint, 469
 unsupervised learning, 64, 197, 198, 201–202, 207–208, 328, 329
 Usher, M., 276
 utilitarian theory of perception, 159

 V1
 attention-based visual perception, 274*f*
 attention-reduced response, 232
 benchmarking and, 42, 43*f*
 binocular disparity/luminance correlation, 453–455*f*, 458
 correlation structures in natural scenes, 452*f*
 depth interference, 459*f*, 469
 encoding, 263, 470
 learning effects, 48, 417–419*f*, 424*f*
 motion-sensitive neurons, 279
 noise and, 453
 object recognition, 218*f*, 260, 272, 386
 receptive field size, 462
 task-dependent modulations, 231
 tuning, 183, 417
 V2
 attention-based visual perception, 274*f*
 attention-reduced response, 232
 learning effects, 48
 motion-sensitive neurons, 279
 object recognition, 218*f*, 260

- V2 (*cont.*)
 orientation tuning, 183
 task-dependent modulations, 231
- V4
 attentional modulations in, 232
 attention-based visual perception, 274 *f*
 object recognition, 218 *f*, 260
 receptive field size, 462
 shape reconstruction, 186, 187 *f*
 shape tuning, 184–186, 185 *f*
- van Doorn, A. J., 6, 404
- Vapnik-Chervonenkis (VC) dimension, 71–72, 73, 75
- variational inference method, 338
- Varma, M., 411
- VC dimension, 71–72, 73
- ventral visual pathway, 349–351, 352–356, 360–362
- ventral visual stream, 44–46, 45 *f*
 cononical computational strategy, 48
 differential development of, 120–122 *f*
 fMRI measurements, 120–122 *f*, 121 *f*
 implications, 122–123
 dimensionality expansion strategy, 47–48
 extrastriate body area, 105, 109, 118
 functional organization of human, 103–105
 fusiform body area, 105, 109, 118
 fusiform face area, 105, 109, 110, 118–119, 120, 220, 221, 349
 hierarchical organization across, 113
 inferior temporal (IT) cortex, 46–47
 lateral fusiform regions, 105
 lateral occipital complex, 104–105
 lateral occipital complex (*see* lateral occipital complex)
 LO, 106, 108, 109, 111 *f*–112 *f*, 117, 359
 medial fusiform regions, 105
 object-, face-, place-selective cortex, 104 *f*
 occipital face area, 349, 351 *f*, 352, 354, 355, 356, 359, 360
 parahippocampal place area, 105, 110, 114, 118, 119
 pFUS/OTS, 106, 108, 109
 superior temporal sulcus, 349, 351 *f*, 354, 355, 356, 359, 360
 temporal image evolution, 48–49
 transfer functions, 48–49
 visual word form area, 105, 110
- veridical perception, 150, 151, 162, 163
- vibrotactile frequency discrimination, 278
- video, analyzing/indexing, 56–57, 59
- view-independent recognition, 262
- viewpoint sensitivity, 115–117, 116 *f*
- viewpoint-tolerant descriptions, 91, 95
- view-tuned model unit (VTU), 218 *f*, 220, 228, 230
- Viola, P., 98
- VisNet, 262–265, 269, 275, 281–282
- visual building blocks, 201
- visual categorization
 detection experiments, 338–342
 on estimation/evaluation of approximate inference, 339–340 *f*
 on ETH shape database, 342 *f*
 on generative decomposition, 339 *f*
 on PASCAL'06 VOC challenge, 340–342
- discriminative detection of, 337–338
 generative/discriminative training, 337
 sliding-window approach to, 337–338
 speeding up computation time for, 338
- generative decomposition of, 330–332
 data representation, 330
 Gibbs sampling, 331–332
 Latent Dirichlet Allocation model of, 331–332 *f*
 probabilistic topic models of localized gradients, 330–332
- learning paradigm, 328–329
 future research, 343
 semi-supervised learning, 329
 supervised learning, 328
 unsupervised learning, 328
 weakly supervised learning, 328
- model paradigm, 325–326
 discriminative model benefits, 326
 future research, 343
 generative model benefits, 325–326
 hybrid methods, 326
- paradigm integration, 329
 combining generative and discriminative, 329
 combining local and global, 329
 combining unsupervised and supervised, 329
- paradigm integration experiments, 338–342
 Google re-ranking task, 333 *t*–334 *f*
 object class discovery, 335 *f*–336 *f*
 unsupervised discovery, 332–337
- representation paradigm, 326–328
 dense representation, 327–328
 future research, 343
 global representation, 326–327
 local representation, 326–327
 semilocal representation, 327
 sparse features, 327
- selectivity, 360–361
- unsupervised discovery, 332–337
 Google re-ranking task, 333 *t*–334 *f*

- object class discovery, 334–337, 335*f*–336*f*
- visual cortex
 - in macaque, 257–260. *See also* ventral visual stream
- visual development
 - atypical, of category-specific topography, 356–361
 - differential, from childhood to adulthood, 120–123
 - fMRI measurements, 120–122*f*, 121*f*
 - implications of, 122–123
 - typical, of face-related cortex, 350–356
 - category-specific topography by age group, 351*f*
 - mechanisms in fusiform gyrus, 351–353
- visual experience, 77, 81
 - in animals, 49, 192
 - and horizontal connections, 462–463, 466
 - and human development, 362
 - information organization in, 308
 - just seeing and, 73, 75
 - object discovery through, 303, 311, 320
- visual learning, 63–64, 328–329
 - future research, 343
 - semi-supervised learning, 329
 - supervised learning, 328
 - unsupervised learning, 328
 - weakly supervised learning, 328. *See also* visual object recognition
- visual object recognition
 - categorization, 55–56
 - classification task (*see* classification, in visual object recognition)
 - description task, 57–58
 - detection and localization task, 57
 - limitations to current approaches in field of, 63
 - naming task, 57, 58, 65
 - number of categories in, 58*f*–59
 - progress in field of, 61–63, 62*f*
 - reasons to study, 56–57
 - recognition task, 57–58
 - strategy for understanding, 50–51
 - verification task, 57
- visual task taxonomy, 74–76
- visual vocabulary paradigm, 406
- visual word form area (VWFA), 105, 110
- visuoceptual expertise, for face recognition, 355, 357–358, 361, 362
- volume-based representation, 91
- volumetric part-based descriptions, 91*f*–92, 95
- von Uexküll, J., 156
- voxels, 39, 222–223
 - multi-voxel pattern analysis, 110–111, 113
 - simulated putative neural responses, 115–117, 116*f*
- Vuilleumier, P., 355
- Wallis, G., 311, 492, 493
- Wallraven, C., 492
- Wang, G., 14, 140, 277
- wayfinding, 141
- Weber, M., 328
- weighted kernel framework, to combine multiple feature “channels,” 411, 412
- Weiss, Y., 327
- whitening, 513
- white noise, 48, 453
- whole-image categorization
 - appearance-based methods, 403
 - bags of features, 401–402
 - contextual factors, 403
 - experiments
 - classification results, 408*t*, 410*t*
 - example images, 408*f*
 - scene category recognition, 407–409
 - setup, 407
 - summary of, 410–411
 - using Caltech-101 database, 409–410*f*
 - local invariant features, 403
 - locally orderless images, 404
 - locally weighted bags of words, 404
 - orderless texture histogram model, 401
 - pyramid matching, 402, 404–405
 - related work on, 403–404
 - spatial pyramid matching (*see* spatial pyramid matching)
 - spatial pyramid representation, 402*f*
- wiggle stimuli, 480–482, 481*f*
- Wittgenstein, L., 70, 73
- Wixson, L. E., 139
- Worboys, M., 141
- words and pictures
 - annotations, 167
 - attributes, 175–176
 - captions, 167, 174–174
 - Corel test set, 171*f*–172, 176
 - correlations between words and pictures, 168–170
 - auto-annotation, 168
 - auto-illustration, 168
 - clustered image-keyword searches, 168
 - joint image-keyword searches, 168
 - layout, 168

Cambridge University Press

978-0-521-88738-0 - Object Categorization: Computer and Human Vision Perspectives

Edited by Sven J. Dickinson, Ales Leonardis, Bernt Schiele and Michael J. Tarr

Index

[More information](#)

536

INDEX

- words and pictures (*cont.*)
- region label prediction, 168–169
 - word annotation prediction methods, 169*t*–170
 - depiction, 174–175
 - iconography, 176–177
 - Latent Dirichlet Allocation to obtain images, 172, 173*f*
 - modifiers, 175
 - nearby text, 167
 - online knowledge sources to obtain images, 172–173*f*
 - regional labels, 167
 - scenes and annotations, 170–172
 - word and picture datasets, 167–168
 - working memory, 276, 277
 - Wu, B., 340
 - Wunstel, M., 137–138
 - Xu, D., 412
 - Yanai, K., 173–174, 175–176
 - Yuille, A., 443, 468
 - Yushkevich, P., 444
 - Zerroug, M., 95, 136
 - Zheng, F., 266
 - Zhu, X., 15*f*, 203, 329, 468–469, 471
 - Zihl, J., 276
 - Zisserman, A., 341, 412