

# Index

- F*-measures, *see* image classification
- $F_1$ -measure, *see* image classification
- $F_\beta$ -measure, *see* image classification
- $P_b$ , *see* probability of boundary
- $\chi$ -squared kernel, *see* classifier
- 0-1 loss, 459, 511
- $2\frac{1}{2}$ -dimensional sketch, 210
- aberrations, 10–12
  - astigmatism, 11
  - chromatic aberration, 12
  - circle of least confusion, 11
  - coma, 11
  - distortion, 11
    - radial, 27
  - field curvature, 11
  - primary aberrations, 11
  - spherical aberration, 10
- absolute
  - conic, 249
  - dual quadric, 250
- absolute reconstruction, 611
- accumulator array, *see* fitting
- actions from the web, *see* datasets
- active markers, *see* people
- active sensors, 422
- active vision, 218, 221
- activity recognition, *see* people
- affine, 230
- affine geometry
  - affine shape, 232
  - affine transformation, 231
- affine motion model, *see* optical flow
- affine shape, 232
- affine structure from motion
  - ambiguity, 232
  - definition, 230
  - Euclidean upgrade, 238–240
    - multiple views, 239–240
  - from binocular correspondences, 233–236
    - from multiple images
      - factorization, 237–238
- affine transformation, 15
- affinity, *see* clustering
- affordances, *see* object recognition
- AIC, *see* model selection
- airlight, *see* color sources
- albedo, 34
  - low dynamic range, 90
  - reflectance, 76
  - spectral albedo, 76
  - spectral reflectance, 76
- algebraic surface, 454
- aliasing, *see* sampling
- alignment, 377
- all-vs-all, *see* classifier
- alpha expansion, 214, 680
- ambient illumination, 35
- ambiguity
  - affine, *see* affine structure from motion
  - Euclidean, 223, *see* structure from motion, 246
  - projective, *see* projective structure from motion
- aperture problem, *see* corner
- apparent curvature, 402, 406
- apparent motion, 315
- arc length, 400, 404
- area source, 36
- area sources
  - shadows, 36, 37
- aspect, 416, 531, 547
- aspect graph, 392, 416–417
- astigmatism, 11
- asymptotic
  - bitangent, 415
  - curve, 409–410
    - red and blue, 409
  - directions, 399, 403–405, 425
  - spherical image, 412
  - spherical map, 407, 410–411

- tangent, 409, 412, 414
- atlas, *see* medical imaging
- attached shadow, 403
- attributes, 550, *see* object recognition
- AUC, 466
- average pooling, *see* pooling
- average precision, *see* image classification
  
- background subtraction, 261–263, 314
- backward variable, *see* people
- barycentric coordinates, 379
- baseline, 199
- basis pursuit, 184
- Bayes classifier, *see* classifier
- Bayes information criterion, *see* model selection
- Bayes risk, *see* classifier
- beak-to-beak, 412
- beat frequency, 424
- bed-of-nails function, 125
- best bin first, *see* nearest neighbors
- between-class variance, 495
- between-class variation, *see* image classification
- Bhattacharyya coefficient, 336
- bias, 515
- binary terms, *see* tree-structured models
- binocular fusion, 197
- binormal
  - vector, 403
- birds, *see* image classification
- bitangent, 413
  - line, 414
    - asymptotic, 415
    - limiting, 415
  - plane, 415, 417
  - ray manifold, 407, 413–414
- black body, *see* color sources
  - color, physical terminology, 75
- blind spot, 14
- blurring, *see* smoothing
- BM3D, 183–184
- boosting, 475
- bootstrapping, *see* classifier
- bottles, *see* datasets
- brightness, 32, 44, *see* color spaces
- browsing, 643
- browsing images, 627
  - desiderata for systems, 643
- BSDS, *see* datasets
- Buffy, *see* datasets
  
- bundle adjustment, 245, *see* for mosaics
  
- Caltech 101, *see* image classification
- Caltech 256, *see* image classification
- Caltech-101, *see* datasets
- Caltech-256, *see* datasets
- camels, *see* datasets
- camera
  - obscure, 3
  - pinhole, *see* pinhole camera
- camera calibration
  - linear, 23–26
    - camera parameters, 25–26
    - degeneracies, 24–25
    - projection matrix, 23–24
  - photogrammetry, 27–29
  - weak, *see* weak calibration
- camera consistency, 367
- camera model
  - affine, 20–22
    - extrinsic parameters, 22
    - intrinsic parameters, 22
  - projection equation, 6, 7, 22
  - perspective, 14–20
    - extrinsic parameters, 18–19
    - intrinsic parameters, 16–18
    - projection equation, 18
  - weak perspective, 20–22
- canonical variates, *see* classifier
- capacity, *see* graphs
- CART, 450
- carved visual hull, 572–573
- cascade, *see* detection
- cast shadow, 403
- catadioptric optical systems, 8
- category, *see* object recognition
- center of curvature, 395
- central limit theorem, 430
- central perspective, *see* projection model →
  - pinhole perspective
- chaff, 349
- Chamfer matching, 388
- chromatic aberration, 12
- CIE, *see* color spaces
- CIE LAB, *see* color spaces
- CIE u'v' space, *see* color spaces
- CIE xy, *see* color spaces
- CIE xy color space, *see* color spaces
- CIE XYZ, *see* color spaces
- CIE XYZ color space, *see* color spaces
- circle of curvature, 395

- circle of least confusion, 11
- circuit, *see* graph
- class-confusion matrix, *see* classifier
- classifier, 457, *see also* image classification
  - Bayes classifier, 459, 460
  - Bayes risk, 459
  - boosting, 475
    - decision stump, 475
    - decision stump, training, 476
    - discrete adaboost, 477
    - fast face detection by, 524
    - real adaboost, 478
    - weak learner, 475
  - class confusion matrix, 465
  - class-confusion matrix, 464
  - correlated image keywords, 649, 650
  - cost of misclassification, 457
  - cross-validation, 464
  - decision boundary, 458
    - chosen to minimize total risk, 458, 459
  - definition, 457
  - error rate, 463
  - estimating performance, 459, 464
    - detection rate, 466
    - leave-one-out cross-validation, 464
    - receiver operating characteristic curve, 466
    - receiver operating curve, 466, 501
    - true positive rate, 466
  - examples
    - finding faces using a neural network, 522, 523
  - false negative, 457
  - false positive, 457
  - feature selection
    - by canonical variates, 494, 497–499
    - by principal component analysis or PCA, 493–495
    - problems with principal component analysis, 496
  - from class histograms
    - skin finding examples, 466, 500, 501
  - general methods for building
    - directly determining decision boundaries, 467
    - from explicit probability models, 467
  - kernel machine, 473
    - kernel, 474
  - kernel machines
    - $\chi$ -squared kernel, 488
    - histogram intersection kernel, 489
    - spatial pyramid kernel, 489
  - linking faces to names, 651, 652
  - logistic regression, 461
    - logistic loss, 461
  - loss, 457
  - Mahalanobis distance, 468
  - multiclass
    - all-vs-all, 479
    - one-vs-all, 479
  - naive Bayes, 469
  - nearest neighbor classifier
    - k-nearest neighbor classifier, 469
  - nearest neighbors, 469, 470
  - neural network
    - finding faces using, 522, 523
  - plug-in classifier
    - example, 467, 468
  - practical tips, 475
    - bootstrapping, 477
    - hard negative mining, 478
    - manipulating training data, 477
  - predicting keywords for images, 646–648, 654, 656
  - regularization, 460
    - choosing weight with cross-validation, 463
    - for numerical reasons, 463
    - possible norms, 463
    - to improve generalization, 463
  - risk function, 458
  - software, 480
    - LIBSVM, 480
    - multiple kernel learning, 480
    - PEGASOS, 480
    - SVMLight, 480
  - support vector machine, 526
    - for linearly separable datasets, 470–472
    - for non-separable data, 472
    - hinge loss, 472
  - total risk, 458
- closing point, 564
- closure, *see* segmentation
- clustering, 268
  - as a missing data problem, *see* missing data problems
  - complete-link clustering, 270
  - dendrogram, 270
  - graph theoretic, 277
    - agglomerative, 280, 281

- divisive, 281–283
  - normalized cuts, 284, 285
- group average clustering, 270
- grouping and agglomeration, 269
- image pixels using agglomerative or divisive methods, 270
- image pixels using K-means, 272, 273
- image pixels, as a missing data problem, 308
- normalized cut, 284
- partitioning and division, 269
- single-link clustering, 270
- using hierarchical k-means, 640
- using K-means, 172
- clutter, 377
- CMU Quality of Life dataset, *see* datasets
- CMY space, *see* color spaces
- coarse-to-fine matching, 137
- coding, *see* image classification
- collections of images, browsing
  - MDS for layout, 644, 645
- color bleeding, 59, 101
- color constancy, 95
  - finite-dimensional linear model, 96
    - recovering surface color by gamut mapping, 98
    - recovering surface color from average reflectance, 98
    - recovering surface color from gamut, 99
    - recovering surface color from specular reflections, 98
  - human color constancy, 95, 96
  - lightness computation, 44
- color matching functions, *see* color spaces
- color perception, 68
  - cone, 72, 73
  - Grassman's laws, 70
  - lightness
    - computing lightness, 44
  - photometry does not explain, 95, 96
  - primaries, 68
  - principle of univariance, 71
  - rod, 72
  - subtractive matching, 69
  - surface color, 96
  - surface color perception, 95
  - test light, 68
  - trichromacy, 69
- color sources
  - airlight, 73
  - black body, 75
    - color temperature, 75
  - daylight, 73
  - fluorescent light, 74
  - incandescent light, 74
  - mercury arc lamps, 75
  - Mie scattering, 74
  - Rayleigh scattering, 73
  - skylight, 73
  - sodium arc lamps, 75
- color spaces, 77
  - brightness, 83
  - by matching experiments, 78
  - CIE, 79
  - CIE LAB, 85
  - CIE u' v' space, 87
  - CIE u'v' space, 85
  - CIE xy color space, 80
  - CIE XYZ color space, 79, 81
    - CIE xy, 80, 82, 83
  - CMY space, 81
    - mixing rules, 81
    - use of four inks, 82
  - color differences, 86
  - color matching functions, 78
    - negative values, 78
    - obtaining weights from, 78
  - cyan, 81
  - HSV space, 84, 85
  - hue, 83
  - imaginary primaries, 78
  - just noticeable differences, 84
  - lightness, 83
  - Macadam ellipses, 84, 86
  - magenta, 81
  - opponent color space, 80
  - RGB color space, 80
  - RGB cube, 80
  - saturation, 83
  - subtractive matching, required, 78
  - uniform color space, 85, 87
  - uniform color spaces, 86
  - value, 83
  - yellow, 81
- color temperature, *see* color sources
- color, modeling image, 87
  - color constancy using model, 96
  - color depends on surface and on illuminant, 89
  - complete equation, 88
  - diffuse component, 89

- finding specularities using model, 92
  - finite-dimensional linear model, 96, 97
    - computing receptor responses, 98
    - recovering surface color by gamut mapping, 98
    - recovering surface color from average reflectance, 98
    - recovering surface color from gamut, 99
    - recovering surface color from specular reflections, 98
  - illuminant color, 88, 89
  - color, physical terminology
    - spectral energy density, 68
  - coma, 11
  - comb function, 125
  - combinatorial optimization, 663
  - common fate, *see* segmentation
  - common region, *see* segmentation
  - compact support, 116
  - complete data log-likelihood, *see* missing data problems
  - composition across the body, *see* people
  - compound lenses, 12
  - computational molecules, 429
  - computed tomography imaging, *see* medical imaging
  - concave point, 405
    - plane curve, 397
    - surface, 399
  - cone strip, 561, 567
  - cones, 13
  - conjugate directions, 403, 405
  - connected, *see* graph
  - connected components, *see* graph
  - connected graph, *see* graph
  - consecutive, *see* graph
  - content based image retrieval, *see* digital libraries
  - continuity, *see* segmentation
  - contour
    - image, *see* image contour
    - occluding, *see* occluding contour
  - convex point
    - planar curve, 397
    - surface, 399
  - convolution, 107
    - associative, 117
    - commutative, 117
    - continuous, 115, 117
      - 1D derivation, 115
      - 2D derivation, 117
    - impulse response, 117
    - point spread function, 117
    - properties, 117
  - discrete, 108
    - convention about sums, 108
    - effects of finite input datasets, 118
  - examples
    - finite differences, 110, 112, 141, 142, 144, 145
    - local average, 107
    - ringing, 108, 109
    - smoothing, *see* smoothing
    - weighted local average, *see* smoothing
  - gives response of shift invariant linear system
    - discrete 1D derivation, 113
    - discrete 2D derivation, 114
  - kernel, 107, 117
  - like a dot product, 131–133
  - notation, 114, 115
- convolution theorem, 120
  - convolutional nets, 216
  - coordinate system
    - local, *see* differential geometry
  - copyright, *see* near duplicate detection
  - corner
    - aperture problem, 148
    - covariant windows, 156
    - detection, 149
    - estimating scale and orientation, 151
    - estimating scale with Laplacian of Gaussian, 151, 153
    - finding pattern elements, 154, 155
    - Harris corner detector, 150
    - interest point, 148
    - self-similar, 149
  - correspondence problem, 197
  - cosine similarity, 633
  - cost-to-go function, *see* tree-structured models
  - covariance, 154
  - cross-validation, *see* model selection, *see* classifier
  - crowdsourcing, 515
  - crust algorithm, 454
  - curvature, 396, 402, 404
    - curve, 395, 397, 428
      - apparent, 402
      - sign, 397

- surface
  - Gaussian, 401, 404, 406
  - normal, 398, 401, 405
  - principal, 41, 399, 401
- cusp, 391, 394, 399, 402, 407–409, 411–413, 416
- crossing, 415, 416
- of Gauss, 408, 410, 411
- of the first kind, 394
- of the second kind, 394
- point, 391, 404
- cut, *see* graph
- cyan, *see* color spaces
- cyclopean retina, 205
- cylindrical panorama, 371
- data association, 349
- Data-driven texture representations, 164
- datasets
  - activity
    - actions from the web, 625
    - CMU Quality of Life dataset, 626
    - IXMAS dataset, 626
    - KTH action dataset, 625
    - movies and scripts datasets, 626
    - MSR action dataset, 625
    - Olympic sports dataset, 626
    - PMI dataset, 626
    - UCF activity datasets, 626
    - VIRAT dataset, 626
    - Weizmann activity, 625
  - BSDS, 287
  - crowdsourcing, 515
  - dataset bias, 515
  - face detection, 539
  - general object detection, 539
  - human interactions, 624
  - human signers, 625
  - image classification, 513
    - bottles, 514
    - Caltech-101, 513
    - Caltech-256, 513
    - camels, 514
    - flowers, 514
    - Graz-02, 514
    - Imagenet, 514
    - LabelMe, 513
    - Lotus Hill data, 514
    - materials, 515
    - Pascal challenge, 514
    - repositories, 515
    - scenes, 515
  - LHI, 287
  - near duplicate detection, 640
  - parsing
    - Buffy, 624
    - HumanEva dataset, 626
  - pedestrian detection, 538
  - people, 624
  - pose
    - Buffy, 624
  - segmentation, 287
- daylight, *see* color sources
- decision boundary, *see* classifier
- decision stump, *see* classifier
- decision tree, 448
- Decision trees, 448–450
- decoding, *see* object recognition
- definite patch neighbor, 577
- definitely visible patch, 576
- deformable objects
  - detection, 530, 533, 534
    - software, 535
- deformation, 531
- degree, *see* graph
- Delaunay triangulation, 454
- delta function, *see* convolution
- dendrogram, *see* clustering
- dense depth map, 47
- depth map, 47, 422
- depth of field, 9
- depth of focus, 9
- derivative of Gaussian filters, *see* gradient,
  - estimating
- derivatives
  - using finite differences, 430
- derivatives, estimating
  - differentiating and smoothing with one
    - convolution, 142
  - using finite differences, 110
    - noise, 112
    - smoothing, 141, 142, 144, 145
- detection
  - cascade, 524
  - deformable objects
    - model, 530, 533, 534
    - software, 535
  - evaluation
    - overlap test, 535
  - face detection, 520, 524
  - non-maximum suppression, 519
  - occluding boundaries, 527, 529

- evaluation, 530
  - method, 531
- pedestrian detection, 524–526, 528
  - evaluation, 527, 537
  - sliding windows, 520
  - state of the art, 536
- detection rate, *see* classifier
- dictionary, 183, 184
- dielectric surfaces, 90
- differential geometry, 391, 441
  - analytical, 424–426
  - descriptive, 393–401
  - plane curves, 393–397
    - concave point, 397
    - convex point, 397
    - curvature, 395, 396, 428
    - cuspid of the first kind, 394
    - cuspid of the second kind, 394
    - Gauss map, 396
    - Gaussian image, 396
    - inflection, 394, 396
    - local coordinate system, 393
    - normal line, 393
    - normal vector, 393
    - regular point, 394
    - singular point, 394
    - tangent line, 393
    - tangent vector, 393
  - space curves, 397
    - curvature, 397
    - Gauss map, 397
    - normal plane, 397
    - osculating plane, 397
    - parametric curve, 420
    - principal normal line, 397
    - tangent line, 397
  - surfaces, 397–401
    - asymptotic directions, 399
    - concave point, 399
    - conjugate directions, 403, 405
    - convex point, 399
    - differential of the Gauss map, 400, 425
    - elliptic point, 399, 405
    - first fundamental form, 425
    - Gauss map, 399
    - Gaussian curvature, 401, 404
    - Gaussian image, 399
    - hyperbolic point, 399
    - local coordinate system, 399
    - Monge patch, 425
    - normal curvature, 398
    - normal line, 398
    - normal section, 398
    - normal vector, 398
    - parabolic point, 399
    - parametric surface, 41
    - principal curvatures, 399
    - principal directions, 398, 403
    - second fundamental form, 400, 425
    - second-order model, 399
    - tangent plane, 398
- differential of the Gauss map, 400, 425
- diffraction, 8
- diffuse reflection, 33
- diffusion equation, 138
- digital libraries, 627
  - browsing, 627
    - desiderata for systems, 643
    - MDS for layout, 644, 645
  - correlated image keywords, 649, 650
  - evaluation, 629
    - news search example, 506
    - patent search example, 506
    - precision, 506
    - recall, 506
    - web filtering example, 506
  - example applications, 628
    - image browsing, 629
    - image search, 628
    - near duplicate detection, 628
    - trademark evaluation, 628
  - image search by keyword, 645
  - linking faces to names, 651, 652
  - predicting keywords for images, 646–648, 654, 656
  - user behavior, 630, 631
  - user needs, 629
- dilation, 499, 500
- diopter, 13
- directed graph, *see* graph
- discrete Adaboost, 475
- discrete cosine transform, 184
- discrete wavelet transform, 184
- disparity, 197
- distance minimization
  - geometric, 225
- distance transformation, 388
- distant point light source, 34
- distortion, 11
- distributional semantics, *see* information retrieval

- dual problem, 473
- dynamic programming, 210, *see* hidden Markov models
- dynamics, 327
- E-step, *see* missing data problems
- ecologically valid, 260
- edge
  - detection, 424
  - roof edges, 428–429
  - step edges, 427–428
- edge detection, 144
  - gradient based, 145
  - examples, 148
  - finding maxima of gradient magnitude, 145, 146
  - hysteresis, 146
  - nonmaximum suppression, 147
  - poor behavior at corners, 149
  - rectifying, *see* corner
  - roof edges, 161
  - step edges, 161
- edge points, 145
- edge-preserving smoothing, 138
- edges, 141, 145
- egg-shaped, *see* differential geometry → surfaces → elliptic
- egomotion, 315
- eigenvalue problem, 666
  - generalized, 666
- eight-point algorithm, *see* weak calibration
- elastic net, 675
- elliptic point, 399, 405
- EM, *see* missing data problems
- empirical cost function, 673
- entropy, 94, 387
- entropy coding, 586
- envelope, 410
- epipolar
  - geometry, 198–199
  - line, 198, 199, 202, 210
  - plane, 198
- epipolar constraint, 199
  - affine fundamental matrix, 233–235
  - parameterization, 235
  - essential matrix
    - characterization, 201
    - parameterization, 200
  - fundamental matrix, 201
  - characterization, 201
  - Longuet–Higgins relation
    - uncalibrated, 201
  - Longuet–Higgins relation
    - calibrated, 200
    - instantaneous, 218
- epipoles, 199
- erosion, 500, 501
- error rate, *see* classifier
- essential matrix, *see* epipolar constraint
- estimating scale with Laplacian of Gaussian
  - corner
    - Laplacian, 151
- Euclidean geometry
  - Euclidean shape, 223, 246
  - rotation matrix
    - exponential representation, 219
    - quaternions, 434
    - Rodrigues formula, 455
- Euclidean structure from motion
  - ambiguity, 224
  - definition, 222
  - from binocular correspondences, 224–228
- Euclidean upgrade
  - from affine structure, 238–240
  - multiple views, 239–240
  - from projective structure, 246–248
  - partially-calibrated cameras, 246–248
- Euler angles, 15
- exemplars, 609
- EXIF tags, 17
- exitance, 55
- expectation-maximization, *see* EM
- exterior orientation, 27
- extremal points, 564
- extrinsic parameters, 18
  - affine camera, 22
  - perspective camera, 18–19
- f-number, 9, 206
- false negative, *see* classifier
- false positive, *see* classifier
- false positive rate, 466
- familiar configuration, *see* segmentation, *see* segmentation
- feature tracking, 137
- feature vectors, 457
- field curvature, 11
- field of view
  - camera, 9
  - human eye, 13
- figure-ground, *see* segmentation



- filtering, *see* linear filtering
- finite difference, 110
- finite differences, 112, 142
  - choice of smoothing, 145
  - derivative of Gaussian filters, 142, 144, 145
  - differentiating and smoothing with one convolution, 142
  - smoothing, 141, 142, 144
- finite-dimensional linear model, *see* color, modeling image
- first fundamental form, 425
- first-order geometric optics, *see* paraxial  $\rightarrow$  geometric optics
- fitting, *see* segmentation, 290
  - curves, 297
    - parametric curves, distance from a point to, 299
  - Hough transform, 290
    - accumulator array, 292
    - for lines, 290–292
    - practical issues, 292
  - identifying outliers with EM, 312, 313
  - layered motion models with EM, 313, 318–320
  - least squares, 294, 295
    - outliers, 299
    - sensitivity to outliers, 299, 300
  - lines
    - by least squares, 294, 295
    - by total least squares, 294, 295
    - fitting multiple lines, 296
    - identifying outliers with EM, 312, 313
    - incremental line fitting, 296
    - k-means, 296
  - outliers, *see* robustness
  - planes, 295
  - robust, *see* robustness
  - tokens, 293
  - total least squares, 294, 295
  - using k-means, 296
- flecnodal
  - curve, 411, 412, 414, 417
  - point, 411, 412
- flow, *see* graph
- flow-based matching, 334
- flowers, *see* datasets
- FM beat, 424
- focal points, 9
- focus of expansion, *see* optical flow
- fold
  - of the Gauss map
    - plane curve, 396
    - surface, 400, 404
  - point, 391, 404
- footskate, *see* people
- for mosaics
  - bundle adjustment, 374
- forest, *see* graph
- forward selection, 673
- forward variable, *see* people
- Fourier transform, 118, 119
  - as change of basis, 118, 119
  - basis elements as sinusoids, 119
  - definition for 2D signal, 119
  - inverse, 120
  - is linear, 120
  - of a sampled signal, 126
  - pairs, 120
  - phase and magnitude, 120
    - magnitude spectrum of image uninformative, 121, 122
  - sampling, *see* sampling
- fovea, 13
- frame-bearing group, 369
- free-form surface, 441
- Frobenius norm, 247, 669
- frontier point, 561, 564
- fronto-parallel plane, 6
- fundamental matrix, *see* epipolar constraint
- gamut, 81
- gate, 350
- Gauss map
  - differential of the, 400, 425
  - fold, 396, 400, 404
  - plane curve, 396
  - space curve, 397
  - surface, 399
- Gauss sphere, 399, 439
- Gauss–Newton, *see* nonlinear least squares
- Gaussian curvature, 401, 404, 406, 425
- Gaussian image
  - curve, 396, 404
  - surface, 399, 401
- Gaussian kernel, 474
- Gaussian smoothing, 426
- generalized
  - eigenvalue problem, 666
  - eigenvector, 666
- generalized eigenvalue problem, 496
- generalized polyhedron, 561

- generalizing badly, 460
- generative model, 306
- generic
  - surface, 409
- geoconsistency, 572
- geometric consistency
  - in pedestrian detection, 528
- geometrical modes, 65
- geometry
  - differential, *see* differential geometry
- gestalt, *see* segmentation
- GIST features, *see* image classification
- global shading model, 52
  - color bleeding, 59
  - comparing black and white rooms, 56
  - governing equation, 56
  - interreflections, 35
  - reflexes, 59
  - smoothing effect of interreflection, 57
- gradient, estimating
  - differentiating and smoothing with one convolution, 142
  - using derivative of Gaussian filters, 142, 145
  - using finite differences, 110
    - noise, 112
    - smoothing, 141, 144
- graph, 277
  - agglomerative clustering using, 280, 281
  - circuit, 278
  - connected, 278
  - connected components, 279
  - connected graph, 278
  - consecutive, 278
  - cut, 279
  - degree, 277
  - directed graph, 278
  - divisive clustering using, 281–283
  - flow, 279
  - forest, 279
  - min-cut, 283
  - path, 278
  - self-loop, 278
  - spanning tree, 279
  - tree, 278
  - undirected graph, 278
  - weighted graph, 278
- graph cuts, *see* min-cut/max-flow problems and combinatorial optimization
- graphs
  - capacity, 279
- Grassman’s laws, *see* color perception
- Graz-02, *see* datasets
- grouping, *see* segmentation, *see* fitting
- gutterpoint, 408
- gzip, 586
- half-angle direction, 40
- hard negative mining, *see* classifier, 535
- hard thresholding, 183
- Harris corner detector, *see* corner
- height map, 47
- Hessian, 671
- hidden Markov models, *see also* matching
  - on relations, *see also* people, *see also* tracking, *see also* tree structured energy models
  - dynamic programming, 594
  - homogeneous Markov chain, 591
  - Markov chain, 591
  - state transition matrix, 591
  - stationary distribution, 592
  - Viterbi algorithm, 594
- hierarchical k-means, *see* clustering
- high dynamic range image, 39
- highlights, *see* specular
- hinge loss, *see* classifier
- histogram equalization, 521
- histogram intersection kernel, *see* classifier
- HOG feature, 155, 157, 159, 524–526
  - software, 160
- HOG features
  - difficulties with, 546
- homogeneous
  - projection matrices, 19
- homogeneous Markov chain, *see* hidden Markov models
- homography, *see* projective transformation, 372, 589
- homotopy continuation, 419
- horopter, 204
- Hough transform, *see* fitting
- HSV space, *see* color spaces
- hue, *see* color spaces
- human
  - eye, 12–14
    - blind spot, 14
    - cones, 13
    - fovea, 13
    - Helmoltz’s schematic eye, 13
    - macula lutea, 13
    - rods, 13

- stereopsis, 203–205, 217
      - cyclopean retina, 205
      - horopter, 204
      - monocular hyperacuity threshold, 204
      - random dot stereogram, 204
  - human parser, *see* people
  - HumanEva dataset, *see* datasets
  - hyperbolic point, 391, 399, 405
  - hypernyms, 511
  - hyponyms, 511
  - hysteresis, *see* edge detection
  - ICP algorithm, *see* iterated closest-point algorithm
  - illumination cone, 65
  - illusory contour, 260
  - illusory contours, *see* segmentation
  - image browsing, 629
  - image classification, *see also* classifier
    - as information retrieval, 639
    - between-class variation, 482
    - classifying images of single objects, 504
      - evaluation, 505
      - general points, 505
    - correlated image keywords, 649, 650
    - dataset bias, 515
    - datasets, 512, 513
      - birds, 515
      - bottles, 514
      - Caltech 101, 510
      - Caltech 256, 510
      - Caltech-101, 513
      - Caltech-256, 513
      - camels, 514
      - crowdsourcing, 515
      - flowers, 514
      - Graz-02, 514
      - Imagenet, 514
      - LabelMe, 513
      - Lotus Hill data, 514
      - materials, 515
      - Pascal challenge, 509, 514
      - repositories, 515
      - scenes, 515
    - evaluation
      - $F$ -measures, 506
      - $F_1$ -measure, 506
      - $F_\beta$ -measure, 506
      - average precision, 507
      - news search example, 506
      - patent search example, 506
      - precision, 506
      - recall, 506
      - web filtering example, 506
  - example applications
    - explicit images, 482, 498
    - material classification, 483, 502
    - scene classification, 484, 502
  - features
    - coding, 544
    - contour features, 546
    - general points, 484
    - geometric representations, 548
    - GIST features, 486
    - histogram equalization, 521
    - HOG feature, 524–526
    - pooling, 545
    - preclustering, 546
    - shading features, 547
    - spatial pyramid kernel, 489
    - visual words, 488, 639
  - linking faces to names, 651, 652
  - output
    - affordances, 549
    - attributes, 550–553
  - predicting keywords for images, 646–648, 654, 656
  - software, 512
    - color descriptor, 513
    - course software, 513
    - GIST, 513
    - link repository, 513
    - pyramid match, 513
    - VLFeat, 513
  - specialized problems, 511
  - state of art
    - number of classes, 509
    - performance on fixed classes, 508
  - within-class variation, 482
- image completion, 176
    - by matching, 180
    - by texture synthesis, 181
    - methods, 179
    - state of the art, 182
  - image contour, 391
    - convexities and concavities, 405
    - curvature, 404
    - cusplike, 391, 403, 404
    - inflection, 403, 404
    - Koenderink’s theorem, 404–407
    - T-junction, 391
  - image denoising, 182–186

- BM3D, 183
  - learned sparse coding, 184
  - non-local means, 183
  - results, 186
- image hole filling, 176
  - by matching, 180
  - by texture synthesis, 181
  - methods, 179
  - state of the art, 182
- image irradiance equation, 59
- image mosaic, 370
- image plane, 4
- image pyramid, 134, *see also* scale, 135
  - coarse scale, 135
  - Gaussian pyramid, 135
    - analysis, 135, 136
    - applications, 136, 137
- image rectification, 202–203, 205
- image search, *see* digital libraries
  - correlated image keywords, 649, 650
  - linking faces to names, 651, 652
  - predicting keywords for images, 646–648, 654, 656
  - using keywords, 645
- image stabilization, 330
- image-based modeling and rendering, 221, 559
  - PMVS, 573
  - visual hulls, 559
- Imagenet, *see* datasets
- impulse response, 114, *see* convolution
- incremental fitting, *see* fitting
- index of refraction, 9, 11
- inertia
  - axis of least, 667
  - second moments, 667
- inflection, 394, 396, 399, 408, 412, 413
- information retrieval, 632
  - distributional semantics, 634
  - latent semantic analysis, 634
  - latent semantic indexing, 634
  - pagerank, 638
    - for image layout, 642
  - query expansion, 639
  - ranking by importance, 638
  - stop words, 632
  - strategies applied to image classification, 639
  - tf-idf weighting, 633
  - word counts, 632
    - smoothing, 633
- integrability, 50
  - in lightness computation, 45
  - in photometric stereo, 50
- integral image, 522
- interactive segmentation, *see* segmentation
- interest point, *see* corner
- interior orientation, 27
- interior-point methods, 677
- interpretation tree, 438, 454
- interreflections, *see* global shading model
- intrinsic parameters, 17
  - affine camera, 22
  - perspective camera, 16–18
- intrinsic representations, 43
- invariant image, *see* shadow removal
- inverted index, 632
- isotropy, *see* texture
- iterated closest points, 369
- iterated closest-point algorithm, 434–436
- IXMAS dataset, *see* datasets
- Jacobian, 670
- joint angles, *see* people
- joint positions, *see* people
- k-d tree, *see* nearest neighbors, 637
  - software, 638
- k-d trees, 435
- k-means, *see* clustering
- k-nearest neighbor classifier, *see* classifier
- Kalman filter, 345
- Kalman filtering, *see* tracking
- kernel, *see* convolution, *see* classifier
  - classifier
- kernel profile, 274, 336
- kernel smoother, 336
- kernel smoothing, 273
- key frame, *see* shot boundary detection
- Kinect, 446–453
- Koenderink’s theorem, 404–407
- KTH action dataset, *see* datasets
- Ky-Fan norm, 649
- LabelMe, *see* datasets
- Lagrange multiplier, 673
- Lambert’s cosine law, 34
- Lambertian+specular model
  - image color model, 91, 92
- lambertian+specular model, 36
- Laplacian, *see* estimating scale with Laplacian of Gaussian

- Lasso, 184
- latent semantic analysis, *see* information retrieval
- latent semantic indexing, *see* information retrieval
- latent variable, 531
- layered motion, 318, *see also* fitting, *see also* optical flow
  - support maps, 319
- learned sparse coding, 184–185
- least squares, *see* fitting, 663
  - linear, 201
    - homogeneous, 436
    - non-homogeneous, 441
  - nonlinear, 202
- least-angle regression, 673
- leave-one-out cross-validation, 322, *see* classifier
- lenses
  - depth of field, 9
  - f number, 9
- level set, 436
- Levenberg–Marquardt, *see* nonlinear least squares
- LHI, *see* datasets 287
- LIBSVM, *see* classifier
- light field, 559, 584–586
- light slab, 585
- lightness, 44, *see* color spaces
- lightness computation, 44
  - algorithm, 43, 45
  - assumptions and model, 44
  - constant of integration, 45
- lightness constancy, 44
- limiting bitangent developable, 415
- line space, 291
- linear, *see* properties, 113
- linear filtering, *see* convolution, *see* linear systems, shift invariant
- linear least squares, 663–669
  - homogeneous, 665–669
    - eigenvalue problem, 666
    - generalized eigenvalue problem, 666
  - nonhomogeneous, 664–665
    - normal equations, 664
    - pseudoinverse, 665
- linear systems, shift invariant
  - convolution like a dot product, 131–133
  - filtering as output of linear system, 107
  - filters respond strongly to signals they look like, 131
  - impulse response, 117
  - point spread function, 117
  - properties, 112
    - scaling, 113
    - superposition, 113
  - response given by convolution, 115
    - 1D derivation, 115
    - 2D derivation, 117
    - discrete 1D derivation, 113
    - discrete 2D derivation, 114
- lines of curvature, 421
- lip, 412
- local shading model, 36
- Local texture representations, 164
- local visual events, 407, 412–413
- locality sensitive hashing, *see* nearest neighbors, 636
  - software, 638
- logistic loss, *see* classifier
- logistic regression, *see* classifier
- Longuet–Higgins relation, *see* epipolar constraint, 200
- loss, *see* classifier
- Lotus Hill data, *see* datasets
- luminaires, 33
- M-estimator, *see* robustness
- M-step, *see* missing data problems
- macula lutea, 13
- magenta, *see* color spaces
- magnetic resonance imaging, *see* medical imaging
- magnification, 6
- magnitude spectrum, *see* Fourier transform
- Mahalanobis distance, *see* classifier
- manifold, 414
- MAP inference, 213
- marching cubes, 437
- markerless motion capture, 589
- Markov chain, 339, *see* hidden Markov models
- Markov models, hidden, *see* hidden Markov models, *see also* matching on relations, *see also* people, *see also* tracking
- matching on relations
  - hidden Markov models, 590
    - backward variable, 598
    - dynamic programming, 594

- dynamic programming algorithm, 595
- dynamic programming figure, 596
- example of Markov chain, 592
- fitting a model with EM, 595, 597, 598
- forward variable, 598
- node value, 594
- trellis, 592, 593
- Viterbi algorithm, 594, 595
- pictorial structure models, 602, 604, 605, 608, 610
- tree-structured energy models, 600
- material properties, 164
- materials, *see* datasets
- matrix
  - nullspace, 24
  - range, 664
  - rank, 664
- matte, 266
- max pooling, *see* pooling
- MDL, *see* model selection
- MDS, *see* multidimensional scaling
- Mean average precision, 508
- mean shift
  - tracking with, 335
- medical imaging
  - applications of registration, 383, 387
  - atlas, 384
  - imaging techniques, 384
    - computed tomography imaging, 385
    - magnetic resonance imaging, 385
    - nuclear medical imaging, 385
    - ultra-sound imaging, 385
- metric shape, *see* Euclidean shape
- Meusnier's theorem, 401
- Mie scattering, *see* color sources
- min-cut/max-flow, 214
- min-cut/max-flow problems and combinatorial optimization, 675–682
- min-cuts, 663
- minimum description length, *see* model selection
- missing data problem, 307
- missing data problems, 307
  - EM algorithm, 310, 311
    - background subtraction example, 314
    - complete data log-likelihood, 309
    - E-step, 311
    - fitting HMM with, 595, 597, 598
    - incomplete data log-likelihood increases at each step, 311
    - M-step, 311
    - motion segmentation example, 313, 318–320
    - outlier example, 312, 313
    - practical difficulties, 312
    - soft weights, 311
  - image segmentation, 308
  - iterative reestimation strategy, 310
  - layered motion, 318
  - mixture model, 309
    - mixing weights, 309
    - mixture, 309
- mixing weights, *see* missing data problems
- mixture, *see* missing data problems
- mixture model, *see* missing data problems
- mobile robot
  - navigation, 197, 221
- model selection, 319
  - AIC, 321
  - Bayes information criterion, 321
  - Bayesian, 321
  - BIC, 321
  - cross-validation, 322
  - MDL, 321
  - minimum description length, 321
  - overfit, 321
  - selection bias, 320
  - test set, 320
  - training set, 320
- model-based vision
  - alignment, 377
  - application in medical imaging, 383, 384, 387
  - verification, 377
    - by edge proximity and orientation, 378
    - by edge proximity is unreliable, 378, 379
- Mondrian worlds, 96
- Monge patch, 47, 425
- motion capture, 451, *see* people
- motion field, 218
- motion graph, *see* people
- motion primitives, *see* people
- movies and scripts datasets, *see* datasets
- MSR action dataset, *see* datasets
- Muller-Lyer illusion, 256
- multidimensional scaling, 644
  - for image layout, 645
- multilocal visual events, 407, 414–416
- multiple kernel learning, 475, *see* classifier

- multiple-view stereo, *see* stereopsis → multiple views
- Munsell chips, 100
- mutual information, 387
- N-cut, *see* clustering
- naive Bayes, *see* classifier
- narrow-baseline stereopsis, 217, 573
- near duplicate detection, 628, 639
  - using hierarchical k-means, 640
  - copyright, 628
  - trademark, 628
  - using LSH, 640
  - using visual words, 639
- nearest neighbor classifier, *see* classifier
- nearest neighbors, 350
  - approximate algorithms, 635, 637
  - best bin first, 637
  - k-d tree, 636
  - locality sensitive hashing, 635
  - software, 638
  - correlated image keywords, 649, 650
  - linking faces to names, 651, 652
  - near duplicate detection, 640
  - predicting keywords for images, 647, 648
- neural net, 522
- Newton's method
  - convergence rate, 670
  - nonlinear equations, 670
  - nonlinear least squares, 670–671
- node value, *see* matching on relations
- noise
  - additive stationary Gaussian noise, 142, 143
  - choice of smoothing filter
    - effect of scale, 145
  - smoothing to improve finite difference estimates, 141, 142, 144
- non-local means, 183
- non-maximum suppression, *see* detection
- non-square pixels, 124
- nonlinear least squares, 669–672
  - Gauss–Newton, 671–672
  - convergence rate, 672
  - Levenberg–Marquardt, 672
  - Newton, *see* Newton's method
- nonmaximum suppression, *see* edge detection
- normal
  - curvature, 398
  - line, 393, 398
  - plane, 397
  - principal, 397
  - section, 398
  - vector, 393, 398
- normal equations, 665
- normal section, 399
- normalized correlation, 133, 206, 445, 576
- normalized cut, *see* clustering, 285
- normalized image plane, 16
- nuclear medical imaging, *see* medical imaging
- Nyquist's theorem, 126
- object model acquisition
  - from range images, 436–438
- object recognition
  - affordances, 549
  - aspect, 547
  - attributes, 550–553
  - categorization, 542–543
  - current strategies, 542
  - desirable properties, 540–541
  - from range images, 438–446
  - geometric representations, 548
  - part representations, 553
  - poselets, 553, 554
  - selection, 544
  - visual phrases, 554, 555
    - decoding, 555, 556
- observations, 327
- occluding boundaries
  - detecting, 527, 529
  - evaluation, 530
  - method, 531
- occluding contour, 141, 391
  - cuspid point, 391, 404
  - fold point, 391, 404
- Olympic sports dataset, *see* datasets
- one-vs-all, *see* classifier
- OpenCV, 30
- opening point, 564
- opponent color space, *see* color spaces
- optical axis, 8
- optical flow, 313, 589
  - focus of expansion, 315
  - layered motion, 318
  - parametric models of
    - affine motion model, 316
    - more general, 317
  - segmentation by, 316, 318–320

- yields time to contact, 315
- ordering constraint, 210
- orientation, 147
- orientations, 144, 147
  - affected by scale, 150
  - differ for different textures, 151
  - do not depend on intensity, 149
  - in HOG features, 159
  - in SIFT descriptors, 157
- orthogonal matching pursuit, 673
- osculating plane, 397
- outliers, *see* robustness
- outline, *see* image contour
- overcomplete dictionaries, 672
- overfit, *see* model selection
- overfitting, 460
- overlap test, *see* detection
- oversegmentation, 268
  
- Pagerank, 639
- pagerank, *see* information retrieval
  - for image layout, 642
- parabolic
  - curve, 407–409, 411, 417
  - point, 410, 412
- parabolic point, 391, 399, 404, 405
- paraboloid, 399
- parallelism, *see* segmentation
- parametric
  - curve, 420
  - surface, 41
- parametric models of
  - optical flow, 316
- parametric surface, 424
- part representations, *see* object recognition
- parts, 532, 551
- Pascal challenge, *see* image classification,
  - see* datasets
- passive markers, *see* people
- patch-based multi-view stereopsis, 573–584
- path, *see* graph
- pattern elements
  - describing neighborhoods, 155, 157
  - finding with corner detector, 154
  - finding with Laplacian of Gaussian, 155
  - shape context, 613
  - software, 160
  - yield covariant windows, 156
- PCA, *see* classifier
- PEGASOS, *see* classifier
- penumbra, 37
  
- people
  - 3D from 2D, 611, 612, 614
    - ambiguities, 613
    - snippets, 616
  - activity is compositional, 619
    - composition across the body, 620
    - motion primitives, 619
  - activity recognition, 617, 619
    - by characteristic poses, 618
    - by poselets, 618, 619
    - by spacetime features, 621–623
    - from compositional models, 621, 624, 625
  - datasets, *see* datasets
  - detecting, 525, 526, 528
    - evaluation, 527, 537
  - hidden Markov models, 590
    - backward variable, 598
    - dynamic programming, 594
    - dynamic programming algorithm, 595
    - dynamic programming figure, 596
    - example of Markov chain, 592
    - fitting a model with EM, 595, 597, 598
    - forward variable, 598
    - trellis, 592, 593
    - Viterbi algorithm, 594, 595
  - human parser, 602
    - pictorial structure models, 602, 604, 605
  - motion capture, 617
    - active markers, 617
    - computed edges, 620
    - footskate, 617
    - joint angles, 617
    - joint positions, 617
    - motion graph, 620
    - passive markers, 617
    - skeleton, 617
  - pictorial structure models, 608, 610
  - software, *see* software
  - tracking, 606
    - by appearance, 608, 610
    - by templates, 609
    - is hard, 606
    - tree-structured energy models, 600
- perceptual organization, *see* segmentation
- perspective
  - camera, *see* camera model → perspective
  - effects, 5



- projection matrix, *see* projection matrix  $\rightarrow$  pinhole perspective
- phase spectrum, *see* Fourier transform
- photoconsistency, 572
- photogrammetry, 27, 197
- Photometric stereo, 47
- photometric stereo
  - depth from normals, 49
  - formulation, 49
  - integrability, 45, 50
  - normal and albedo in one vector, 48
  - recovering albedo, 49
  - recovering normals, 49
- pictorial structure models, *see* people
- pinhole, 3
  - camera, 4–6
- pinhole perspective, 4
- planes, representing orientation of
  - slant, 188, 189
  - tilt, 188, 189
  - tilt direction, 188, 189
- plenoptic function, 584
- PMI dataset, *see* datasets
- PMVS, *see* patch-based multi-view stereopsis, 574
- point spread function, *see* convolution
- pooled texture representations, 164
- pooling, *see* image classification
  - average pooling, 545
  - max pooling, 545
- pose, 367
- pose consistency, 367
- poselet, 552
- poselets, *see* object recognition
  - for activity recognition, 618, 619
- posterior, 340
- potential patch neighbor, 577
- potentially visible patch, 576
- pragmatics, 657
- precision, *see* image classification
- preclustering, *see* image classification
- predictive density, 340
- primaries, *see* color perception
- primary aberrations, 11
- principal
  - curvatures, 399, 426
  - directions, 398, 403, 425
- principal component analysis, *see* classifier
- principal points, 10
- principle of univariance, *see* color perception
- prior, 339
- probabilistic data association, 350
- probability distributions
  - normal distribution
    - important in tracking linear dynamic models, 344
  - sampled representations, 351
- probability of boundary
  - $P_b$ , 528
- probability, formal models of
  - expectation
    - computed using sampled representations, 351, 352
  - integration by sampling
    - sampling distribution, 351
  - representation of probability distributions by sampled representations, 352
    - marginalizing a sampled representation, 353
    - prior to posterior by weights, 354, 355
- projection equation
  - affine, 22
    - orthographic, 7
    - weak-perspective, 6
  - pinhole perspective, 6
    - points, 18
- projection matrix
  - affine, 22
    - characterization, 22
    - weak-perspective, 22
  - pinhole perspective
    - characterization, 19–20
    - explicit parameterization, 19
    - general form, 18
- projection model
  - affine, 6–7
    - orthographic, 7
    - paraperspective, 21
    - weak-perspective, 6
  - pinhole perspective
    - planar, 4–6
    - weak perspective, 20–22
- projective, 230
- projective coordinate system, 242
- projective geometry
  - projective shape, 242
  - projective transformation, 241
- projective projection matrix, 241
- projective shape, 241

- projective structure from motion
  - ambiguity, 241
  - definition, 241
  - Euclidean upgrade, 246–248
    - partially-calibrated cameras, 246–248
  - from multiple images
    - bilinear method, 244–245
    - bundle adjustment, 245
    - factorization, 244
  - from the fundamental matrix, 242–243
- projective transformations, 15
- properties
  - linear systems, shift invariant
    - linear, 107
    - shift invariant, 107, 113
- proximity, *see* segmentation
- pseudoinverse, 665
- pulse time delay, 424
- pyramid kernel, *see* image classification
  
- QR decomposition, 665, 666
- quaternions, 433–434, 436, 440, 454
- query expansion, *see* information retrieval
- QuickTime VR, 588
  
- radial
  - curvature, 405
  - curve, 404
  - distortion, 27
- radiance
  - definition, 52
  - units, 53
- radiometric calibration, 38
- radiosity, 54
  - of a surface whose radiance is known, 54
  - definition and units, 54
- radius of curvature, 395
- random dot stereogram, 204
- random forest, 448
- Random forests, 448–450
- random forests, 446
- range finders, 422–424
  - acoustico-optical, 424
  - time of flight, 423
  - triangulation, 422
- range image, 422
- ranking, *see* Pagerank
- RANSAC, *see* robustness
- ratio, 232
- Rayleigh scattering, *see* color sources
  
- recall, *see* image classification
- receiver operating characteristic curve, *see* classifier
  - estimating performance
    - ROC, 466
- reciprocity, 38
- reflectance, *see* albedo
  - color, physical terminology, 76
- reflexes, 58
- region, 164
- region growing, 431
- regional properties, 58
- regions, 256
- registration
  - from planes, 439–441
  - from points, 434–436
- regular point, 394
- regularization, 213, *see* classifier
- regularization term, 672
- regularizer, 463
- relative reconstruction, 611
- render, 377
- repetition, 164
- repositories, *see* datasets
- rest positions, 619
- retargeting, 451
- Retinex, 63
- RGB color space, *see* color spaces
- RGB cube, *see* color spaces
- rigid transformation, 15
- rigid transformations and homogeneous coordinates, 14–16
- rim, *see* occluding contour
- ringing, *see* convolution
- risk, *see* classifier
- risk function, *see* classifier
- robustness, 299
  - identifying outliers with EM, 312, 313
  - M-estimator, 300, 303, 304
    - influence function, 301
  - M-estimators, 301
    - scale, 302
  - outliers
    - causes, 299
    - sensitivity of least squares to, 299, 300
- RANSAC, 302
  - how many points need to agree?, 305
  - how many tries?, 303
  - how near should it be?, 305
  - searching for good data, 302

- ROC, *see* receiver operating characteristic curve
- Rodrigues formula, 455
- rods, 13
- roof, 426
- roof edge, 161
- root, 532
- root coordinate system, 611
- rotation matrices, 15
- Rotoscoping, 266
- ruled surface, 411
- saddle-shaped, *see* differential geometry → surfaces → hyperbolic
- sample impoverishment, 357
- sampling, 121, 122
  - aliasing, 123, 125–130
  - formal model, 122, 123, 125
  - Fourier transform of sampled signal, 126
  - illustration, 124
  - non-square pixels, 124
  - Nyquist's theorem, 126
  - poorly causes loss of information, 123
- sampling distribution, *see* probability, formal models of
- saturation, *see* color spaces
- scale, 134, *see* smoothing
  - anisotropic diffusion or edge preserving smoothing, 138
  - applications, 136
  - coarse scale, 135
  - effects of choice of scale, 145
  - of an M-estimator, 302
- scale ambiguity, *see* ambiguity → Euclidean scale space, 426
- scaled orthography, *see* weak perspective
- scaling, *see* linear systems, shift invariant
- scan conversion, 443
- scene classification, *see* scenes
- scenes, 483, *see* datasets
  - scene classification, 484
- searching for images, *see* digital libraries
- secant, 393
- second fundamental form, 400, 425
- segmentation, 164, 255, *see also* clustering, *see also* fitting, 424
  - as a missing data problem, 308, *see* missing data problems
  - by clustering, general recipe, 268
  - example applications, 261
    - background subtraction, 261–263
    - shot boundary detection, 264
- gestalt, 257
- human, 256
  - closure, 258
  - common fate, 258
  - common region, 258
  - continuity, 258
  - examples, 258–261
  - factors that predispose to grouping, 257–261
  - familiar configuration, 258, 260
  - figure and ground, 256, 257
  - gestalt quality or gestaltqualität, 256, 257
  - illusory contours, 260
  - parallelism, 258
  - proximity, 257
  - similarity, 257
  - symmetry, 258
- in humans
  - examples, 255
- interactive segmentation, 261
- range data, 424–432
- trimaps, 281
- watershed, 271
- selection bias, *see also* model selection, 460
- self-calibration, 250
- self-loop, *see* graph
- self-similar, *see* corner
- self-similarities, 183
- semi-local surface representation, 441
- SFM, 221
- shading, 33
- shading primitives, 65
- shadow, 35
- shadow removal, 92
  - color temperature direction, 94
  - estimating color temperature direction, 94
  - examples, 95
  - general procedure, 93
  - invariant image, 94
- shadows
  - area sources, 36, 37
  - penumbra, 37
  - umbra, 36
- shape
  - affine, *see* affine geometry
  - Euclidean, *see* Euclidean geometry
  - projective, *see* projective geometry
- shape context, *see* pattern elements

- shape from shading, 59–61
- shape from texture
  - for curved surfaces, 190
  - repetition of elements yields lighting, 190
- shift invariant, *see* properties
- shift invariant linear system, *see* linear systems, shift invariant
- shot boundary detection, 261, 264
  - key frame, 264
  - shots, 264
- shots, *see* shot boundary detection
- shrinkage, 183
- SIFT descriptor, 155, 157–159
  - software, 160
- SIFT descriptors
  - difficulties with, 546
- silhouette, *see* image contour
- similarity, 223, *see* segmentation
- simplex method, 576
- simulated annealing, 683
- singular point, 394
- singular value decomposition, 237, 244, 666–669
- skeleton, 563, *see* people
- skinning, 451
- skylight, *see* color sources
- slack variables, 472
- slant, 188, 189
  - normal is ambiguous given, 189, 191
- smoothing, 108
  - as high pass filtering, 126, 128–130
  - Gaussian kernel, 109
    - discrete approximation, 110
  - Gaussian smoothing, 108, 109
    - avoids ringing, 108, 109
    - discrete kernel, 110
    - effects of scale, 110, 145
    - standard deviation, 109
    - suppresses independent stationary additive noise, 111
  - scale, 143
  - to reduce aliasing, 126, 128–130
  - weighted average, 107
  - word counts, *see* information retrieval
- Snell's law, 8
- snippets, *see* people
- soft thresholding, 183
- soft weights, *see* missing data problems
- software
  - active appearance models, 383
  - approximate nearest neighbors, 638
  - classifier
    - LIBSVM, 480
    - multiple kernel learning, 480
    - PEGASOS, 480
    - SVMLight, 480
  - deformable object detection, 535
  - deformable registration, 383
  - face detection, 539
  - FLANN, 638
  - general object detection, 539
  - homography estimation, 373
  - image classification
    - color descriptor, 513
    - course software, 513
    - GIST, 513
    - link repository, 513
    - pyramid match, 513
    - VLFeat, 513
  - image segmenters, 285
  - pattern elements, 160
    - color descriptors, 160
    - HOG feature, 160
    - PCA-SIFT, 160
    - toolbox, 160
    - VLFEAT, 160
  - people, 624
- sources
  - source colors, 88, 89
- space carving, 587
- spanning tree, *see* graph
- sparse coding, 663, 672
- sparse coding and dictionary learning, 672–675
  - dictionary learning, 673–675
  - sparse coding, 672–673
  - supervised dictionary learning, 675
- sparse model, 183
- spatial frequency
  - see* Fourier transform, 118
- spatial frequency components, 119
- spectral albedo, *see* albedo
  - color, physical terminology, 76
- spectral colors, 68
- spectral energy density, *see* color, physical terminology
- spectral locus, 83
- spectral reflectance, *see* albedo
  - color, physical terminology, 76
- specular
  - dielectric surfaces, 90

- metal surfaces, 90
- specularities, 90
- specularity
  - finding, 90, 91
- specular albedo, 34
- specular direction, 34
- specular reflection, 34
- specularities, *see* specular
- specularity, 34
- spherical
  - aberration, 10
- spherical panorama, 371
- spin
  - coordinates, 442
  - images, 441–446
  - map, 442
- SSD, *see* sum-of-squared differences
- ssd, *see* sum of squared differences
- standard deviation, *see* smoothing
- state, 327
- state transition matrix, *see* hidden Markov
  - models
- stationary distribution, *see* hidden Markov
  - models
- step, 426
- step edges, 161
- stereolithography, 438
- stereopsis
  - binocular fusion
    - combinatorial optimization, 211–214
    - dynamic programming, 210–211
    - global methods, 210–214
    - local methods, 205–210
    - multi-scale matching, 207–210
    - normalized correlation, 205–207
  - constraints
    - epipolar, 198
    - ordering, 210
  - disparity, 197, 202, 203
  - multiple views, 214–215
  - random dot stereogram, 204
  - reconstruction, 201–203
  - rectification, 202–203
  - robot navigation, 215–216
  - trinocular fusion, 214
- stop words, *see* information retrieval
- structured light, 422
- stuff, 658
- submodularity, 213, 663
- subtractive matching, *see* color perception
- sum of absolute difference, 207
- sum of squared differences, 177
- sum-of-squared differences, 330
  - SSD, 330
- summary matching, 334
- superpixels, 268
- superposition, *see* linear systems, shift in-variant
- superquadrics, 454
- support maps, *see* layered motion
- surface color, *see* color perception
- SVMLight, *see* classifier
- swallowtail, 412
- symmetric Gaussian kernel, *see* smoothing
- symmetry, *see* segmentation
- system, *see* linear systems, shift invariant
- system identification, 362
- T-junction, 391, 404, 407, 412–416
- tangent
  - crossing, 415
  - line, 393
  - plane, 398
  - vector, 393
- template matching
  - filters as templates, 131
- test error, 459
- test set, *see* model selection
- texton, 166
- texture
  - examples, 164, 165
  - isotropy, 188
  - local representations, 166–171
  - pooled representations, 171, 173–175
  - representing with filter outputs, 166
    - algorithm, 169
    - example, 169–171
    - published codes, 170
    - scheme, 167
    - typical filters, 168
  - representing with vector quantization,
    - 172
    - algorithm, 172
    - example, 174, 175
    - scheme, 173
  - scale, 164
  - shape from texture, 187
    - for planes, 187–189
  - synthesis, 176, 178
    - algorithm, 177
    - example, 178, 179
  - texton, 166

- texture mapping, 559, 569, 585
- texture synthesis
  - for image hole filling, 181
- tf-idf weighting, *see* information retrieval
- thick lenses, 10
  - principal points, 10
- thin lenses, 9
  - equation, 9
  - focal points, 9
- tilt, 188, 189
- tilt direction, 188
- topics, 634
- total least squares, *see* fitting
- total risk, *see* classifier
- tracking
  - applications
    - motion capture, 326
    - recognition, 326
    - surveillance, 326
    - targeting, 326
  - as inference, 339
  - definition, 326
  - hidden Markov models
    - backward variable, 598
    - dynamic programming, 594
    - dynamic programming algorithm, 595
    - dynamic programming figure, 596
    - example of Markov chain, 592
    - fitting a model with EM, 595, 597, 598
    - forward variable, 598
    - trellis, 592, 593
    - Viterbi algorithm, 594, 595
  - Kalman filters, 344
    - example of tracking a point on a line, 344, 345
    - forward-backward smoothing, 345, 347, 348
  - linear dynamic models
    - all conditional probabilities normal, 344
    - are tracked using a Kalman filter,  $qv$ , 344
    - constant acceleration, 342, 343
    - constant velocity, 341, 342
    - drift, 341
    - periodic motion, 343
  - main problems
    - correction, 340
    - data association, 340
    - prediction, 339
  - measurement
    - measurement matrix, 341
    - observability, 341
  - particle filtering, 350
    - practical issues, 360
    - sampled representations, 351–355
    - simplest, 355
    - simplest, algorithm, 356
    - simplest, correction step, 356
    - simplest, difficulties with, 357
    - simplest, prediction step, 355
    - working, 358–361
    - working, by resampling, 358
  - smoothing, 345, 347, 348
  - tracking by detection, 327
  - tracking by matching, 327
- tracking by detection, *see* tracking
- tracking by matching, *see* tracking
- trademark, *see* near duplicate detection
- trademark evaluation, 628
- Training error, 459
- training set, *see* model selection
- transformation groups
  - affine transformations, 231
  - projective transformations, 241
  - similarities, 223
- tree, *see* graph
- tree-structured models
  - binary terms, 600
  - cost-to-go function, 601
  - unary terms, 600
- trellis, *see* matching on relations
- trichromacy, *see* color perception
- trimaps, *see* segmentation
- trinoocular fusion, *see* stereopsis  $\rightarrow$  trinocular fusion
- triple point, 415, 416
- tritangent, 415
- true positive rate, *see* classifier
- twisted cubic, 25
- twisted curves, *see* differential geometry  $\rightarrow$  space curves
- UCF activity datasets, *see* datasets
- ultra-sound imaging, *see* medical imaging
- umbra, 36
- unary terms, *see* tree-structured models
- undirected graph, *see* graph
- undulation, 410
- unode, 416
- value, *see* color spaces

- Vector quantization, 172
- vector quantization, 586
- vergence, 204, 217
- viewing
  - cone, 391
  - cylinder, 391
- viewing sphere, 416
- viewpoint
  - general, 404
- vignetting, 12
- VIRAT dataset, *see* datasets
- virtual image, 4
- visual events, 392, 411
  - curves, 411
  - equations, 421
  - local, 412–413
    - beak-to-beak, 412
    - lip, 412
    - swallowtail, 412
  - multilocal, 414–416
    - cusp crossing, 415, 416
    - tangent crossing, 415
    - triple point, 415, 416
- visual hull, 417, 559–573
- visual phrases, *see* object recognition
- visual potential, *see* aspect graph
- visual words, 487, *see* image classification, 639
  - recovering suppressed detail, 640
- Viterbi algorithm, *see* hidden Markov models
  - see* hidden Markov models, 594
- voxel, 437
  
- watershed, *see* segmentation
- wavelet shrinkage, 183
- weak calibration, 224–226
  - eight-point algorithm
    - minimal, 225
    - normalized, 225
    - overconstrained, 225
  - nonlinear, 225
- weak learner, *see* classifier
- weak perspective, 7
  - projection matrix, *see* projection matrix  $\rightarrow$  affine  $\rightarrow$  weak-perspective
- weighted graph, *see* graph
- Weizmann activity, *see* datasets
- wide-baseline, 215
- wide-baseline stereopsis, 217, 574
- window, *see* chaff
  
- within-class variance, 495
- within-class variation, *see* image classification
- word counts, *see* information retrieval
  
- yellow, *see* color spaces
  
- zero-skew projection matrix, 19
- zippered polygonal mesh, 454

# List of Algorithms

2.1	Determining the Lightness of Image Patches. . . . .	45
2.2	Photometric Stereo. . . . .	50
4.1	Subsampling an Image by a Factor of Two. . . . .	129
4.2	Forming a Gaussian Pyramid. . . . .	136
5.1	Gradient-Based Edge Detection. . . . .	146
5.2	Obtaining Location, Radius and Orientation of Pattern Elements Using a Corner Detector. . . . .	154
5.3	Obtaining Location, Radius, and Orientation of Pattern Elements Using the Laplacian of Gaussian. . . . .	155
5.4	Computing a SIFT Descriptor in a Patch Using Location, Orientation and Scale. . . . .	158
5.5	Computing a Weighted $q$ Element Histogram for a SIFT Feature. . . . .	159
6.1	Local Texture Representation Using Filters. . . . .	172
6.2	Texture Representation Using Vector Quantization. . . . .	173
6.3	Clustering by K-Means. . . . .	176
6.4	Non-parametric Texture Synthesis. . . . .	177
7.1	The Marr–Poggio (1979) Multi-Scale Binocular Fusion Algorithm. . . . .	208
7.2	A Dynamic-Programming Algorithm for Establishing Stereo Correspondences Between Two Corresponding Scanlines. . . . .	212
8.1	The Longuet-Higgins Eight-Point Algorithm for Euclidean Structure and Motion from Two Views. . . . .	228
8.2	The Tomasi–Kanade Factorization Algorithm for Affine Shape from Motion. . . . .	238
9.1	Background Subtraction. . . . .	263
9.2	Shot Boundary Detection Using Interframe Differences. . . . .	264
9.3	Agglomerative Clustering or Clustering by Merging. . . . .	269
9.4	Divisive Clustering, or Clustering by Splitting. . . . .	269
9.5	Finding a Mode with Mean Shift. . . . .	275
9.6	Mean Shift Clustering. . . . .	276
9.7	Mean Shift Segmentation. . . . .	277
9.8	Agglomerative Clustering with Graphs. . . . .	280
10.1	Incremental Line Fitting. . . . .	296
10.2	K-means Line Fitting. . . . .	297
10.3	Using an M-Estimator to Fit a Least Squares Model. . . . .	302
10.4	RANSAC: Fitting Structures Using Random Sample Consensus. . . . .	305
11.1	Tracking by Detection. . . . .	330
11.2	Tracking with the Mean Shift Algorithm. . . . .	335
11.3	The Kalman Filter. . . . .	346
11.4	Forward-Backward Smoothing. . . . .	347
11.5	Obtaining a Sampled Representation of a Probability Distribution. . . . .	352
11.6	Computing an Expectation Using a Set of Samples. . . . .	352
11.7	Obtaining a Sampled Representation of a Posterior from a Prior. . . . .	355
11.8	A Practical Particle Filter Resamples the Posterior. . . . .	359
11.9	An Alternative Practical Particle Filter. . . . .	360
14.1	The Model-Based Edge-Detection Algorithm of Ponce and Brady (1987). . . . .	427
14.2	The Iterative Closest-Point Algorithm of Besl and McKay (1992). . . . .	435
14.3	The Plane-Matching Algorithm of Faugeras and Hebert (1986). . . . .	440



14.4	Pointwise Matching of Free-Form Surfaces Using Spin Images, after Johnson and Hebert (1998, 1999).	445
14.5	Training a Decision Tree.	449
14.6	Training a Random Forest.	450
15.1	The Bayes Classifier.	459
15.2	Cross-Validation	464
15.3	Multi-class Classification Assuming Class-Conditional Densities are Normal	468
15.4	$(k, l)$ Nearest Neighbor Classification	470
15.5	Training a Two-Class Decision Stump	476
15.6	Discrete Adaboost	477
15.7	Real Adaboost	478
16.1	Principal Components Analysis	495
16.2	Canonical Variates	499
16.3	Dilation	499
16.4	Erosion	501
17.1	Sliding Window Detection.	520
19.1	Visual Hull Construction.	564
19.2	A Curve-Tracing Algorithm.	565
19.3	The Strip Triangulation Algorithm.	567
19.4	The PMVS Algorithm.	575
19.5	The Feature-Matching Algorithm of PMVS.	579
19.6	The Patch-Expansion Algorithm of PMVS.	581
20.1	The Viterbi Algorithm	595
20.2	Fitting Hidden Markov Models with EM	595
20.3	Computing the Forward Variable for Fitting an HMM	597
20.4	Computing the Backward Variable for Fitting an HMM	597
20.5	Updating Parameters for Fitting an HMM	598
21.1	Nearest Neighbor Tagging.	647
21.2	Greedy Labeling Using Kernel Similarity Comparisons.	648
22.1	The Alpha Expansion Algorithm of Boykov <i>et al.</i> (2001).	681

