
Contents

List of algorithms	xxi
Preface	xxv
Possible course outlines	xxxi
1 Introduction	1
1.1 Motivation	1
1.2 Why is computer vision difficult?	3
1.3 Image representation and image analysis tasks	5
1.4 Summary	9
1.5 Exercises	10
1.6 References	10
2 The image, its representations and properties	11
2.1 Image representations, a few concepts	11
2.2 Image digitization	14
2.2.1 Sampling	14
2.2.2 Quantization	15
2.3 Digital image properties	16
2.3.1 Metric and topological properties of digital images	16
2.3.2 Histograms	23
2.3.3 Entropy	24
2.3.4 Visual perception of the image	25
2.3.5 Image quality	27
2.3.6 Noise in images	28
2.4 Color images	30
2.4.1 Physics of color	30
2.4.2 Color perceived by humans	32
2.4.3 Color spaces	36
2.4.4 Palette images	38
2.4.5 Color constancy	39
2.5 Cameras: An overview	40
2.5.1 Photosensitive sensors	40
2.5.2 A monochromatic camera	42
2.5.3 A color camera	44
2.6 Summary	45

2.7	Exercises	46
2.8	References	48
3	The image, its mathematical and physical background	50
3.1	Overview	50
3.1.1	Linearity	50
3.1.2	The Dirac distribution and convolution	51
3.2	Linear integral transforms	52
3.2.1	Images as linear systems	53
3.2.2	Introduction to linear integral transforms	53
3.2.3	1D Fourier transform	54
3.2.4	2D Fourier transform	59
3.2.5	Sampling and the Shannon constraint	62
3.2.6	Discrete cosine transform	65
3.2.7	Wavelet transform	66
3.2.8	Eigen-analysis	72
3.2.9	Singular value decomposition	73
3.2.10	Principal component analysis	74
3.2.11	Radon transform	77
3.2.12	Other orthogonal image transforms	78
3.3	Images as stochastic processes	79
3.4	Image formation physics	82
3.4.1	Images as radiometric measurements	82
3.4.2	Image capture and geometric optics	83
3.4.3	Lens aberrations and radial distortion	86
3.4.4	Image capture from a radiometric point of view	89
3.4.5	Surface reflectance	92
3.5	Summary	95
3.6	Exercises	97
3.7	References	98
4	Data structures for image analysis	100
4.1	Levels of image data representation	100
4.2	Traditional image data structures	101
4.2.1	Matrices	101
4.2.2	Chains	104
4.2.3	Topological data structures	106
4.2.4	Relational structures	107
4.3	Hierarchical data structures	108
4.3.1	Pyramids	108
4.3.2	Quadrees	109
4.3.3	Other pyramidal structures	111
4.4	Summary	112
4.5	Exercises	113
4.6	References	115
5	Image pre-processing	116
5.1	Pixel brightness transformations	117

5.1.1	Position-dependent brightness correction	117
5.1.2	Gray-scale transformation	117
5.2	Geometric transformations	120
5.2.1	Pixel co-ordinate transformations	121
5.2.2	Brightness interpolation	123
5.3	Local pre-processing	125
5.3.1	Image smoothing	125
5.3.2	Edge detectors	133
5.3.3	Zero-crossings of the second derivative	139
5.3.4	Scale in image processing	143
5.3.5	Canny edge detection	144
5.3.6	Parametric edge models	147
5.3.7	Edges in multi-spectral images	148
5.3.8	Local pre-processing in the frequency domain	148
5.3.9	Line detection by local pre-processing operators	155
5.3.10	Detection of corners (interest points)	156
5.3.11	Detection of maximally stable extremal regions	160
5.4	Image restoration	162
5.4.1	Degradations that are easy to restore	163
5.4.2	Inverse filtering	163
5.4.3	Wiener filtering	164
5.5	Summary	165
5.6	Exercises	167
5.7	References	174
6	Segmentation I	178
6.1	Thresholding	179
6.1.1	Threshold detection methods	181
6.1.2	Optimal thresholding	183
6.1.3	Multi-spectral thresholding	186
6.2	Edge-based segmentation	187
6.2.1	Edge image thresholding	188
6.2.2	Edge relaxation	190
6.2.3	Border tracing	191
6.2.4	Border detection as graph searching	196
6.2.5	Border detection as dynamic programming	206
6.2.6	Hough transforms	210
6.2.7	Border detection using border location information	217
6.2.8	Region construction from borders	218
6.3	Region-based segmentation	220
6.3.1	Region merging	221
6.3.2	Region splitting	224
6.3.3	Splitting and merging	225
6.3.4	Watershed segmentation	229
6.3.5	Region growing post-processing	232
6.4	Matching	232
6.4.1	Template matching	233
6.4.2	Control strategies of templating	235

6.5	Evaluation issues in segmentation	236
6.5.1	Supervised evaluation	237
6.5.2	Unsupervised evaluation	240
6.6	Summary	241
6.7	Exercises	245
6.8	References	248
7	Segmentation II	255
7.1	Mean shift segmentation	255
7.2	Active contour models—snakes	263
7.2.1	Traditional snakes and balloons	264
7.2.2	Extensions	267
7.2.3	Gradient vector flow snakes	268
7.3	Geometric deformable models—level sets and geodesic active contours	273
7.4	Fuzzy connectivity	280
7.5	Towards 3D graph-based image segmentation	288
7.5.1	Simultaneous detection of border pairs	289
7.5.2	Suboptimal surface detection	293
7.6	Graph cut segmentation	295
7.7	Optimal single and multiple surface segmentation— LOGISMOS	303
7.8	Summary	317
7.9	Exercises	319
7.10	References	321
8	Shape representation and description	329
8.1	Region identification	333
8.2	Contour-based shape representation and description	335
8.2.1	Chain codes	336
8.2.2	Simple geometric border representation	337
8.2.3	Fourier transforms of boundaries	341
8.2.4	Boundary description using segment sequences	343
8.2.5	B-spline representation	346
8.2.6	Other contour-based shape description approaches	348
8.2.7	Shape invariants	349
8.3	Region-based shape representation and description	353
8.3.1	Simple scalar region descriptors	353
8.3.2	Moments	358
8.3.3	Convex hull	360
8.3.4	Graph representation based on region skeleton	365
8.3.5	Region decomposition	370
8.3.6	Region neighborhood graphs	372
8.4	Shape classes	373
8.5	Summary	373
8.6	Exercises	375
8.7	References	379
9	Object recognition	385
9.1	Knowledge representation	386

9.2	Statistical pattern recognition	390
9.2.1	Classification principles	392
9.2.2	Nearest neighbors	393
9.2.3	Classifier setting	395
9.2.4	Classifier learning	398
9.2.5	Support vector machines	400
9.2.6	Cluster analysis	406
9.3	Neural nets	407
9.3.1	Feed-forward networks	409
9.3.2	Unsupervised learning	411
9.3.3	Hopfield neural nets	412
9.4	Syntactic pattern recognition	413
9.4.1	Grammars and languages	415
9.4.2	Syntactic analysis, syntactic classifier	417
9.4.3	Syntactic classifier learning, grammar inference	420
9.5	Recognition as graph matching	421
9.5.1	Isomorphism of graphs and subgraphs	421
9.5.2	Similarity of graphs	425
9.6	Optimization techniques in recognition	426
9.6.1	Genetic algorithms	427
9.6.2	Simulated annealing	430
9.7	Fuzzy systems	432
9.7.1	Fuzzy sets and fuzzy membership functions	432
9.7.2	Fuzzy set operators	434
9.7.3	Fuzzy reasoning	435
9.7.4	Fuzzy system design and training	438
9.8	Boosting in pattern recognition	439
9.9	Random forests	442
9.9.1	Random forest training	444
9.9.2	Random forest decision making	446
9.9.3	Random forest extensions	448
9.10	Summary	448
9.11	Exercises	452
9.12	References	459
10	Image understanding	464
10.1	Image understanding control strategies	466
10.1.1	Parallel and serial processing control	466
10.1.2	Hierarchical control	466
10.1.3	Bottom-up control	467
10.1.4	Model-based control	468
10.1.5	Combined control	469
10.1.6	Non-hierarchical control	472
10.2	SIFT: Scale invariant feature transform	474
10.3	RANSAC: Fitting via random sample consensus	477
10.4	Point distribution models	481
10.5	Active appearance models	492
10.6	Pattern recognition methods in image understanding	503

10.6.1	Classification-based segmentation	503
10.6.2	Contextual image classification	505
10.6.3	Histograms of oriented gradients—HOG	509
10.7	Boosted cascades of classifiers	513
10.8	Image understanding using random forests	517
10.9	Scene labeling and constraint propagation	524
10.9.1	Discrete relaxation	525
10.9.2	Probabilistic relaxation	527
10.9.3	Searching interpretation trees	530
10.10	Semantic image segmentation and understanding	531
10.10.1	Semantic region growing	532
10.10.2	Genetic image interpretation	534
10.11	Hidden Markov models	543
10.11.1	Applications	548
10.11.2	Coupled HMMs	549
10.11.3	Bayesian belief networks	551
10.12	Markov random fields	553
10.12.1	Applications to images and vision	555
10.13	Gaussian mixture models and expectation–maximization	556
10.14	Summary	564
10.15	Exercises	568
10.16	References	572
11	3D geometry, correspondence, 3D from intensities	582
11.1	3D vision tasks	583
11.1.1	Marr’s theory	585
11.1.2	Other vision paradigms: Active and purposive vision	587
11.2	Basics of projective geometry	589
11.2.1	Points and hyperplanes in projective space	590
11.2.2	Homography	592
11.2.3	Estimating homography from point correspondences	594
11.3	A single perspective camera	598
11.3.1	Camera model	598
11.3.2	Projection and back-projection in homogeneous coordinates	601
11.3.3	Camera calibration from a known scene	602
11.4	Scene reconstruction from multiple views	602
11.4.1	Triangulation	603
11.4.2	Projective reconstruction	604
11.4.3	Matching constraints	605
11.4.4	Bundle adjustment	607
11.4.5	Upgrading the projective reconstruction, self-calibration	608
11.5	Two cameras, stereopsis	609
11.5.1	Epipolar geometry; fundamental matrix	610
11.5.2	Relative motion of the camera; essential matrix	612
11.5.3	Decomposing the fundamental matrix to camera matrices	613
11.5.4	Estimating the fundamental matrix from point correspondences	614
11.5.5	Rectified configuration of two cameras	615
11.5.6	Computing rectification	617

11.6	Three cameras and trifocal tensor	619
11.6.1	Stereo correspondence algorithms	621
11.6.2	Active acquisition of range images	627
11.7	3D information from radiometric measurements	630
11.7.1	Shape from shading	631
11.7.2	Photometric stereo	635
11.8	Summary	636
11.9	Exercises	637
11.10	References	639
12	Use of 3D vision	644
12.1	Shape from X	644
12.1.1	Shape from motion	644
12.1.2	Shape from texture	651
12.1.3	Other shape from X techniques	652
12.2	Full 3D objects	655
12.2.1	3D objects, models, and related issues	655
12.2.2	Line labeling	656
12.2.3	Volumetric representation, direct measurements	658
12.2.4	Volumetric modeling strategies	660
12.2.5	Surface modeling strategies	662
12.2.6	Registering surface patches and their fusion to get a full 3D model	663
12.3	2D view-based representations of a 3D scene	670
12.3.1	Viewing space	670
12.3.2	Multi-view representations and aspect graphs	670
12.4	3D reconstruction from an unorganized set of 2D views, and Structure from Motion	671
12.5	Reconstructing scene geometry	674
12.6	Summary	677
12.7	Exercises	678
12.8	References	680
13	Mathematical morphology	684
13.1	Basic morphological concepts	684
13.2	Four morphological principles	686
13.3	Binary dilation and erosion	687
13.3.1	Dilation	688
13.3.2	Erosion	689
13.3.3	Hit-or-miss transformation	692
13.3.4	Opening and closing	692
13.4	Gray-scale dilation and erosion	694
13.4.1	Top surface, umbra, and gray-scale dilation and erosion	694
13.4.2	Umbra homeomorphism theorem, properties of erosion and dilation, opening and closing	697
13.4.3	Top hat transformation	698
13.5	Skeletons and object marking	699
13.5.1	Homotopic transformations	699

13.5.2	Skeleton, medial axis, maximal ball	699
13.5.3	Thinning, thickening, and homotopic skeleton	701
13.5.4	Quench function, ultimate erosion	704
13.5.5	Ultimate erosion and distance functions	706
13.5.6	Geodesic transformations	707
13.5.7	Morphological reconstruction	709
13.6	Granulometry	711
13.7	Morphological segmentation and watersheds	713
13.7.1	Particle segmentation, marking, and watersheds	713
13.7.2	Binary morphological segmentation	714
13.7.3	Gray-scale segmentation, watersheds	716
13.8	Summary	717
13.9	Exercises	718
13.10	References	720
14	Image data compression	722
14.1	Image data properties	723
14.2	Discrete image transforms in image data compression	724
14.3	Predictive compression methods	727
14.4	Vector quantization	730
14.5	Hierarchical and progressive compression methods	730
14.6	Comparison of compression methods	732
14.7	Other techniques	733
14.8	Coding	733
14.9	JPEG and MPEG image compression	734
14.9.1	JPEG—still image compression	734
14.9.2	JPEG-2000 compression	736
14.9.3	MPEG—full-motion video compression	738
14.10	Summary	740
14.11	Exercises	742
14.12	References	744
15	Texture	747
15.1	Statistical texture description	750
15.1.1	Methods based on spatial frequencies	750
15.1.2	Co-occurrence matrices	752
15.1.3	Edge frequency	754
15.1.4	Primitive length (run length)	755
15.1.5	Laws' texture energy measures	757
15.1.6	Local binary patterns—LBPs	757
15.1.7	Fractal texture description	762
15.1.8	Multiscale texture description—wavelet domain approaches	764
15.1.9	Other statistical methods of texture description	768
15.2	Syntactic texture description methods	769
15.2.1	Shape chain grammars	770
15.2.2	Graph grammars	772
15.2.3	Primitive grouping in hierarchical textures	773

15.3	Hybrid texture description methods	775
15.4	Texture recognition method applications	776
15.5	Summary	777
15.6	Exercises	779
15.7	References	782
16	Motion analysis	787
16.1	Differential motion analysis methods	790
16.2	Optical flow	794
16.2.1	Optical flow computation	794
16.2.2	Global and local optical flow estimation	797
16.2.3	Combined local–global optical flow estimation	800
16.2.4	Optical flow in motion analysis	801
16.3	Analysis based on correspondence of interest points	804
16.3.1	Detection of interest points	805
16.3.2	Lucas–Kanade point tracking	805
16.3.3	Correspondence of interest points	807
16.4	Detection of specific motion patterns	810
16.5	Video tracking	814
16.5.1	Background modeling	815
16.5.2	Kernel-based tracking	820
16.5.3	Object path analysis	826
16.6	Motion models to aid tracking	831
16.6.1	Kalman filters	831
16.6.2	Particle filters	837
16.6.3	Semi-supervised tracking—TLD	840
16.7	Summary	843
16.8	Exercises	846
16.9	References	848
Index		853