Image Processing, Analysis, and Machine Vision

Second Edition

Milan Sonka Vaclav Hlavac Roger Boyle

Contents

List of algorithms xiii

List of symbols and abbreviations xvii

Preface xix

Course contents xxiii

1 Introduction **1**
- 1.1 Summary 8
- 1.2 Exercises 8
- 1.3 References 9

2 The digitized image and its properties **10**
- 2.1 Basic concepts 10
  - 2.1.1 Image functions 10
  - 2.1.2 The Dirac distribution and convolution 13
  - 2.1.3 The Fourier transform 13
  - 2.1.4 Images as a stochastic process 15
  - 2.1.5 Images as linear systems 17
- 2.2 Image digitization 18
  - 2.2.1 Sampling 18
  - 2.2.2 Quantization 22
  - 2.2.3 Color images 23
- 2.3 Digital image properties 27
  - 2.3.1 Metric and topological properties of digital images 27
  - 2.3.2 Histograms 32
  - 2.3.3 Visual perception of the image 33
  - 2.3.4 Image quality 35
  - 2.3.5 Noise in images 35
- 2.4 Summary 37
- 2.5 Exercises 38
- 2.6 References 40

3 Data structures for image analysis 42
- 3.1 Levels of image data representation 42
- 3.2 Traditional image data structures 43
  - 3.2.1 Matrices 43
  - 3.2.2 Chains 45
  - 3.2.3 Topological data structures 47
  - 3.2.4 Relational structures 48
- 3.3 Hierarchical data structures 49
  - 3.3.1 Pyramids 49
  - 3.3.2 Quadtrees 51
  - 3.3.3 Other pyramidical structures 52
- 3.4 Summary 53
- 3.5 Exercises 54
- 3.6 References 55

4 Image pre-processing 57
- 4.1 Pixel brightness transformations 58
  - 4.1.1 Position-dependent brightness correction 58
  - 4.1.2 Gray-scale transformation 59
- 4.2 Geometric transformations 62
  - 4.2.1 Pixel co-ordinate transformations 63
  - 4.2.2 Brightness interpolation 65
- 4.3 Local pre-processing 68
  - 4.3.1 Image smoothing 69
  - 4.3.2 Edge detectors 77
  - 4.3.3 Zero-crossings of the second derivative 83
  - 4.3.4 Scale in image processing 88
  - 4.3.5 Canny edge detection 90
  - 4.3.6 Parametric edge models 93
  - 4.3.7 Edges in multi-spectral images 94
  - 4.3.8 Other local pre-processing operators 94
  - 4.3.9 Adaptive neighborhood pre-processing 98
- 4.4 Image restoration 102
  - 4.4.1 Degradations that are easy to restore 105
  - 4.4.2 Inverse filtration 106
  - 4.4.3 Wiener filtration 106
- 4.5 Summary 108
- 4.6 Exercises 111
- 4.7 References 118

5 Segmentation 123
- 5.1 Thresholding 124
  - 5.1.1 Threshold detection methods 127
  - 5.1.2 Optimal thresholding 128
  - 5.1.3 Multi-spectral thresholding 131
  - 5.1.4 Thresholding in hierarchical data structures 133

5.2 Edge-based segmentation 134
  5.2.1 Edge image thresholding 135
  5.2.2 Edge relaxation 137
  5.2.3 Border tracing 142
  5.2.4 Border detection as graph searching 148
  5.2.5 Border detection as dynamic programming 158
  5.2.6 Hough transforms 163
  5.2.7 Border detection using border location information 173
  5.2.8 Region construction from borders 174
5.3 Region-based segmentation 176
  5.3.1 Region merging 177
  5.3.2 Region splitting 181
  5.3.3 Splitting and merging 181
  5.3.4 Watershed segmentation 186
  5.3.5 Region growing post-processing 188
5.4 Matching 190
  5.4.1 Matching criteria 191
  5.4.2 Control strategies of matching 193
5.5 Advanced optimal border and surface detection approaches 194
  5.5.1 Simultaneous detection of border pairs 194
  5.5.2 Surface detection 199
5.6 Summary 205
5.7 Exercises 210
5.8 References 216

6 Shape representation and description 228
6.1 Region identification 232
6.2 Contour-based shape representation and description 235
  6.2.1 Chain codes 236
  6.2.2 Simple geometric border representation 237
  6.2.3 Fourier transforms of boundaries 240
  6.2.4 Boundary description using segment sequences 242
  6.2.5 B-spline representation 245
  6.2.6 Other contour-based shape description approaches 248
  6.2.7 Shape invariants 249
6.3 Region-based shape representation and description 254
  6.3.1 Simple scalar region descriptors 254
  6.3.2 Moments 259
  6.3.3 Convex hull 262
  6.3.4 Graph representation based on region skeleton 267
  6.3.5 Region decomposition 271
  6.3.6 Region neighborhood graphs 272
6.4 Shape classes 273
6.5 Summary 274
6.6 Exercises 276
6.7 References 279

7 Object recognition **290**
- 7.1 Knowledge representation 291
- 7.2 Statistical pattern recognition 297
  - 7.2.1 Classification principles 298
  - 7.2.2 Classifier setting 300
  - 7.2.3 Classifier learning 303
  - 7.2.4 Cluster analysis 307
- 7.3 Neural nets 308
  - 7.3.1 Feed-forward networks 310
  - 7.3.2 Unsupervised learning 312
  - 7.3.3 Hopfield neural nets 313
- 7.4 Syntactic pattern recognition 315
  - 7.4.1 Grammars and languages 317
  - 7.4.2 Syntactic analysis, syntactic classifier 319
  - 7.4.3 Syntactic classifier learning, grammar inference 321
- 7.5 Recognition as graph matching 323
  - 7.5.1 Isomorphism of graphs and sub-graphs 324
  - 7.5.2 Similarity of graphs 328
- 7.6 Optimization techniques in recognition 328
  - 7.6.1 Genetic algorithms 330
  - 7.6.2 Simulated annealing 333
- 7.7 Fuzzy systems 336
  - 7.7.1 Fuzzy sets and fuzzy membership functions 336
  - 7.7.2 Fuzzy set operators 338
  - 7.7.3 Fuzzy reasoning 339
  - 7.7.4 Fuzzy system design and training 343
- 7.8 Summary 344
- 7.9 Exercises 347
- 7.10 References 354

8 Image understanding **362**
- 8.1 Image understanding control strategies 364
  - 8.1.1 Parallel and serial processing control 364
  - 8.1.2 Hierarchical control 364
  - 8.1.3 Bottom-up control strategies 365
  - 8.1.4 Model-based control strategies 366
  - 8.1.5 Combined control strategies 367
  - 8.1.6 Non-hierarchical control 371
- 8.2 Active contour models—snakes 374
- 8.3 Point distribution models 380
- 8.4 Pattern recognition methods in image understanding 390
  - 8.4.1 Contextual image classification 392
- 8.5 Scene labeling and constraint propagation 397
  - 8.5.1 Discrete relaxation 398
  - 8.5.2 Probabilistic relaxation 400
  - 8.5.3 Searching interpretation trees 404

8.6 Semantic image segmentation and understanding 404
    8.6.1 Semantic region growing 406
    8.6.2 Genetic image interpretation 408
8.7 Hidden Markov models 417
8.8 Summary 423
8.9 Exercises 426
8.10 References 428

9 3D vision, geometry, and radiometry 441
9.1 3D vision tasks 442
    9.1.1 Marr's theory 444
    9.1.2 Other vision paradigms: Active and purposive vision 446
9.2 Geometry for 3D vision 448
    9.2.1 Basics of projective geometry 448
    9.2.2 The single perspective camera 449
    9.2.3 An overview of single camera calibration 453
    9.2.4 Calibration of one camera from a known scene 455
    9.2.5 Two cameras, stereopsis 457
    9.2.6 The geometry of two cameras; the fundamental matrix 460
    9.2.7 Relative motion of the camera; the essential matrix 462
    9.2.8 Fundamental matrix estimation from image point correspondences 464
    9.2.9 Applications of epipolar geometry in vision 466
    9.2.10 Three and more cameras 471
    9.2.11 Stereo correspondence algorithms 476
    9.2.12 Active acquisition of range images 483
9.3 Radiometry and 3D vision 486
    9.3.1 Radiometric considerations in determining gray-level 486
    9.3.2 Surface reflectance 490
    9.3.3 Shape from shading 494
    9.3.4 Photometric stereo 498
9.4 Summary 499
9.5 Exercises 501
9.6 References 502

10 Use of 3D vision 508
10.1 Shape from X 508
    10.1.1 Shape from motion 508
    10.1.2 Shape from texture 515
    10.1.3 Other shape from X techniques 517
10.2 Full 3D objects 519
    10.2.1 3D objects, models, and related issues 519
    10.2.2 Line labeling 521
    10.2.3 Volumetric representation, direct measurements 523
    10.2.4 Volumetric modeling strategies 525
    10.2.5 Surface modeling strategies 527

10.2.6 Registering surface patches and their fusion to get a full 3D model 529
10.3 3D model-based vision 535
10.3.1 General considerations 535
10.3.2 Goad's algorithm 537
10.3.3 Model-based recognition of curved objects from intensity images 541
10.3.4 Model-based recognition based on range images 543
10.4 2D view-based representations of a 3D scene 544
10.4.1 Viewing space 544
10.4.2 Multi-view representations and aspect graphs 544
10.4.3 Geons as a 2D view-based structural representation 545
10.4.4 Visualizing 3D real-world scenes using stored collections of 2D views 546
10.5 Summary 551
10.6 Exercises 552
10.7 References 553

11 Mathematical morphology 559
11.1 Basic morphological concepts 559
11.2 Four morphological principles 561
11.3 Binary dilation and erosion 563
11.3.1 Dilation 563
11.3.2 Erosion 565
11.3.3 Hit-or-miss transformation 568
11.3.4 Opening and closing 568
11.4 Gray-scale dilation and erosion 569
11.4.1 Top surface, umbra, and gray-scale dilation and erosion 570
11.4.2 Umbra homeomorphism theorem, properties of erosion and dilation, opening and closing 573
11.4.3 Top hat transformation 574
11.5 Skeletons and object marking 576
11.5.1 Homotopic transformations 576
11.5.2 Skeleton, maximal ball 576
11.5.3 Thinning, thickening, and homotopic skeleton 578
11.5.4 Quench function, ultimate erosion 581
11.5.5 Ultimate erosion and distance functions 584
11.5.6 Geodesic transformations 585
11.5.7 Morphological reconstruction 586
11.6 Granulometry 589
11.7 Morphological segmentation and watersheds 590
11.7.1 Particles segmentation, marking, and watersheds 590
11.7.2 Binary morphological segmentation 592
11.7.3 Gray-scale segmentation, watersheds 594
11.8 Summary 595
11.9 Exercises 597

11.10 References 598

12 Linear discrete image transforms 600
12.1 Basic theory 600
12.2 Fourier transform 602
12.3 Hadamard transform 604
12.4 Discrete cosine transform 605
12.5 Wavelets 606
12.6 Other orthogonal image transforms 608
12.7 Applications of discrete image transforms 609
12.8 Summary 613
12.9 Exercises 617
12.10 References 619

13 Image data compression 621
13.1 Image data properties 622
13.2 Discrete image transforms in image data compression 623
13.3 Predictive compression methods 624
13.4 Vector quantization 629
13.5 Hierarchical and progressive compression methods 630
13.6 Comparison of compression methods 631
13.7 Other techniques 632
13.8 Coding 633
13.9 JPEG and MPEG image compression 634
13.9.1 JPEG—still image compression 634
13.9.2 MPEG—full-motion video compression 636
13.10 Summary 637
13.11 Exercises 640
13.12 References 641

14 Texture 646
14.1 Statistical texture description 649
14.1.1 Methods based on spatial frequencies 649
14.1.2 Co-occurrence matrices 651
14.1.3 Edge frequency 653
14.1.4 Primitive length (run length) 655
14.1.5 Laws' texture energy measures 656
14.1.6 Fractal texture description 657
14.1.7 Other statistical methods of texture description 659
14.2 Syntactic texture description methods 660
14.2.1 Shape chain grammars 661
14.2.2 Graph grammars 663
14.2.3 Primitive grouping in hierarchical textures 664
14.3 Hybrid texture description methods 666
14.4 Texture recognition method applications 667
14.5 Summary 668
14.6 Exercises 670

14.7 References 672

15 Motion analysis 679
15.1 Differential motion analysis methods 682
15.2 Optical flow 685
  15.2.1 Optical flow computation 686
  15.2.2 Global and local optical flow estimation 689
  15.2.3 Optical flow computation approaches 690
  15.2.4 Optical flow in motion analysis 693
15.3 Analysis based on correspondence of interest points 696
  15.3.1 Detection of interest points 696
  15.3.2 Correspondence of interest points 697
  15.3.3 Object tracking 700
15.4 Kalman filters 708
  15.4.1 Example 709
15.5 Summary 710
15.6 Exercises 712
15.7 References 714

16 Case studies 722
16.1 An optical music recognition system 722
16.2 Automated image analysis in cardiology 727
  16.2.1 Robust analysis of coronary angiograms 730
  16.2.2 Knowledge-based analysis of intra-vascular ultrasound 733
16.3 Automated identification of airway trees 738
16.4 Passive surveillance 744
16.5 References 750

Index 755

List of algorithms

2.1 Chamfering 28
2.2 Computing the brightness histogram 32
2.3 Generate additive, zero mean Gaussian noise 36
3.1 Co-occurrence matrix $C_r(z, y)$ for the relation r 44
4.1 Histogram equalization 61
4.2 Smoothing using a rotating mask 74
4.3 Efficient median filtering 75
4.4 Canny edge detector 92
5.1 Basic thresholding 124
5.2 Iterative (optimal) threshold selection 129
5.3 Recursive multi-spectral thresholding 131
5.4 Hierarchical thresholding 134
5.5 Non-maximal suppression of directional edge data 135
5.6 Hysteresis to filter output of an edge detector 137
5.7 Edge relaxation 139
5.8 Inner boundary tracing 142
5.9 Outer boundary tracing 143
5.10 Extended boundary tracing 146
5.11 Border tracing in gray-level images 147
5.12 A-algorithm graph search 149
5.13 Heuristic search for image borders 157
5.14 Boundary tracing as dynamic programming 161
5.15 Curve detection using the Hough transform 167
5.16 Generalized Hough transform 172
5.17 Region forming from partial borders 175
5.18 Region merging (outline) 177
5.19 Region merging via boundary melting 179
5.20 Split and merge 182
5.21 Split and link to the segmentation tree 184
5.22 Single-pass split-and-merge 184
5.23 Removal of small image regions 190
5.24 Match-based segmentation 191
5.25 Three-dimensional graph searching 203
6.1 4-neighborhood and 8-neighborhood region identification 233
6.2 Region identification in run length encoded data 234
6.3 Quadtree region identification 235

6.4 Calculating area in quadtrees 255
6.5 Region area calculation from Freeman 4-connectivity chain code representation 255
6.6 Region convex hull construction 262
6.7 Simple polygon convex hull detection 265
6.8 Skeleton by thinning 267
6.9 Region graph construction from skeleton 270
7.1 Learning and classification based on estimates of probability densities assuming the normal distribution 306
7.2 Minimum distance classifier learning and classification 306
7.3 MacQueen k-means cluster analysis 308
7.4 Back-propagation learning 311
7.5 Unsupervised learning of the Kohonen feature map 313
7.6 Recognition using a Hopfield net 314
7.7 Syntactic recognition 316
7.8 Graph isomorphism 326
7.9 Maximal clique location 327
7.10 Genetic algorithm 332
7.11 Simulated annealing optimization 335
7.12 Fuzzy system design 343
8.1 Bottom-up control 365
8.2 Coronary border detection—a combined control strategy 371
8.3 Non-hierarchical control 372
8.4 Snake growing 378
8.5 Approximate alignment of similar training shapes 382
8.6 Fitting an ASM 387
8.7 Contextual image classification 395
8.8 Recursive contextual image classification 396
8.9 Discrete relaxation 399
8.10 Probabilistic relaxation 403
8.11 Updating a region adjacency graph and dual to merge two regions 405
8.12 Semantic region merging 406
8.13 Genetic image segmentation and interpretation 411
9.1 Ego-motion estimation 467
9.2 3D similarity reconstruction from two cameras 469
9.3 PMF stereo correspondence 483
9.4 Extracting shape from shading 497
10.1 Line labeling 522
10.2 Iterative closest reciprocal points 534
10.3 Goad's matching algorithm 539
14.1 Autocorrelation texture description 649
14.2 Co-occurrence method of texture description 653
14.3 Edge-frequency texture description 654
14.4 Primitive-length texture description 656
14.5 Shape chain grammar texture generation 661
14.6 Texture primitive grouping 665

15.1 Relaxation computation of optical flow from dynamic image pairs 687
15.2 Optical flow computation from an image sequence 688
15.3 Velocity field computation from two consecutive images 698

List of symbols and abbreviations

$\arg(x, y)$	angle (in radians) from x axis to the point (x, y)
$\mathrm{argmax}_i(\mathrm{expr}(i))$	the value of i that causes $\mathrm{expr}(i)$ to be maximal
$\mathrm{argmin}_i(\mathrm{expr}(i))$	the value of i that causes $\mathrm{expr}(i)$ to be minimal
div	integer division
mod	remainder after integer division
$\mathrm{round}(x)$	largest integer which is not bigger than $x + 0.5$
$\emptyset$	empty set
A^c	complement of the set A
$A \subset B$, $B \supset A$	set A is included in set B
$A \cap B$	intersection between sets A and B
$A \cup B$	union of sets A and B
$A \mid B$	difference between sets A and B
$\mathbf{A}$	(uppercase bold) matrices
$\mathbf{x}$	(lowercase bold) vectors
$\lvert \mathbf{x} \rvert$	magnitude (or modulus) of the vector $\mathbf{x}$
$\mathbf{x} \cdot \mathbf{y}$	scalar product between vectors $\mathbf{x}$ and $\mathbf{y}$
$\tilde{x}$	estimate of the value x
$\lvert x \rvert$	absolute value of a scalar
Δx	small finite interval of x, difference
$\partial f / \partial x$	partial derivative of the function f with respect to x
$\nabla \mathbf{f}$, $\mathrm{grad}\, \mathbf{f}$	gradient of $\mathbf{f}$
$\nabla^2 \mathbf{f}$	Laplace operator applied to $\mathbf{f}$
$f * g$	convolution between functions f and g
D_E	Euclidean distance (see Section 2.3.1)
D_4	city block distance (see Section 2.3.1)
D_8	chessboard distance (see Section 2.3.1)
F	complex conjugate of the complex function F
$\mathrm{rank}(A)$	Rank of a matrix A
T^*	transformation dual to transformation T
$\delta(x)$	Dirac function
$\mathcal{E}$	mean value operator
$\mathcal{L}$	linear operator

$\mathcal{O}$	origin of the coordinate system
$\#$	number of (e.g., pixels)
$\breve{B}$	point set symmetrical to point set B
$\oplus$	morphological dilation
$\ominus$	morphological erosion
$\circ$	morphological opening
$\bullet$	morphological closing
$\otimes$	morphological hit-or-miss transformation
$\oslash$	morphological thinning
$\odot$	morphological thickening

1D	one dimension(al)
2D	two dimension(al)
3D	three dimension(al)
AI	artificial intelligence
ASM	active shape model
B-rep	boundary representation
CAD	computer-aided design
CCD	charge-coupled device
CSG	constructive solid geometry
CT	computer tomography
dof	degrees of freedom
ECG	electro-cardiogram
EEG	electro-encephalogram
FFT	fast Fourier transform
FOE	focus of expansion
GA	genetic algorithm
HMM	hidden Markov model
IHS	intensity, hue, saturation
JPEG	Joint Photographic Experts Group
MR	magnetic resonance
MRI	magnetic resonance imaging
OCR	optical character recognition
OS	order statistics
PDM	point distribution model
PET	positron emission tomography
PMF	Pollard-Mayhew-Frisby (correspondence algorithm)
RGB	red, green, blue
SNR	signal-to-noise ratio
SVD	singular value decomposition
TV	television

Preface

Image processing, analysis and machine vision represent an exciting and dynamic part of cognitive and computer science. Following an explosion of interest during the 1970s, the 1980s and 1990s were characterized by the maturing of the field and the significant growth of active applications; remote sensing, technical diagnostics, autonomous vehicle guidance, medical imaging (2D and 3D) and automatic surveillance are the most rapidly developing areas. This progress can be seen in an increasing number of software and hardware products on the market, as well as in a number of digital image processing and machine vision courses offered at universities worldwide.

There are many texts available in the areas we cover—most (indeed, all of which we know) are referenced somewhere in this book. The subject suffers, however, from a shortage of texts which are 'complete' in the sense that they are accessible to the novice, of use to the educated, and up to date. Here we present the second edition of a text first published in 1993 in which we hope to include many of the very rapid developments that have taken and are still taking place, which quickly age some of the very good textbooks produced over the last two decades or so. The target audience is the undergraduate with negligible experience in the area through to the Master's and research student seeking an advanced springboard in a particular topic. Every section of this text has been updated since the first version (particularly with respect to references); additionally, wholly new sections are presented on: compression via JPEG and MPEG; fractals; fuzzy logic recognition; hidden Markov models; Kalman filters; point distribution models; three-dimensional vision; watershed segmentation; wavelets; and an entire chapter devoted to case studies. Additionally, each chapter now includes a concise Summary section. To help the reader to acquire practical understanding, newly added Exercise sections accompany each chapter; these are in the form of short-answer questions and problems of varying difficulty, frequently requiring practical usage of computer tools and/or development of application programs.

This book reflects the authors' experience in teaching one- and two-semester undergraduate and graduate courses in Digital Image Processing, Digital Image Analysis, Machine Vision, Pattern Recognition, and Intelligent Robotics at their respective institutions. We hope that this combined experience will give a thorough grounding to the beginner and provide material that is advanced enough to allow the more mature student to understand fully the relevant areas of the subject. We acknowledge that in a very short time the more active areas will have moved beyond this text.

This book could have been arranged in many ways. It begins with low-level processing and works its way up to higher levels of image interpretation; the authors have chosen this framework because they believe that image understanding originates from a common database of information. The book is formally divided into 16 chapters, beginning with low-level processing and working toward higher-level image representation, although this structure will be less apparent after Chapter 10, when we present transforms, compression, morphology, texture, and motion analysis which are very useful but often special-purpose approaches that may not always be included in the processing chain. The final chapter presents four live research projects which illustrate in practical use much of what has gone before.

Decimal section numbering is used, and equations and figures are numbered within each chapter. Each chapter is accompanied by an extensive list of references and exercises. A selection of algorithms is summarized formally in a manner that should aid implementation—not all the algorithms discussed are presented in this way (this might have doubled the length of the book); we have chosen what we regard as the key, or most useful or illustrative, examples for this treatment.

Each chapter presents material from an introductory level through to an overview of current work; as such, it is unlikely that the beginner will, at the first reading, expect to absorb all of a given topic. Often it has been necessary to make reference to material in later chapters and sections, but when this is done an understanding of material in hand will not depend on an understanding of that which comes later. It is expected that the more advanced student will use the book as a reference text and signpost to current activity in the field—we believe at the time of going to press that the reference list is full in its indication of current directions, but record here our apologies to any work we have overlooked. The serious reader will note that many references are very recent, and should be aware that before long more relevant work will have been published that is not listed here.

This is a long book and therefore contains material sufficient for much more than one course. Clearly, there are many ways of using it, but for guidance we suggest an ordering that would generate four distinct modules:

Digital Image Processing, an undergraduate course.

Digital Image Analysis, an undergraduate/graduate course, for which Digital Image Processing may be regarded as prerequisite.

Computer Vision I, an undergraduate/graduate course, for which Digital Image Processing may be regarded as prerequisite.

Computer Vision II, a graduate course, for which Computer Vision I may be regarded as prerequisite.

The important parts of a course, and necessary prerequisites, will naturally be specified locally; a suggestion for partitioning the contents follows this Preface.

Assignments should wherever possible make use of existing software; it is our experience that courses of this nature should not be seen as ‘programming courses’, but it is the case that the more direct practical experience the students have of the material discussed, the better is their understanding. Since the first edition was published, an explosion of World Wide Web-based material has been made available, permitting many of the exercises we present to be conducted without the necessity of implementing from scratch—we do not present explicit pointers to Web material, since they evolve so quickly; however, pointers to specific support materials for this book and others may be located via the publisher, *http://www.brookscole.com*.

The book has been prepared using the LaTeX text processing system. Its completion would have been impossible without extensive usage of the Internet computer network and electronic mail. We should like to acknowledge the University of Iowa, the Czech Technical University, and the School of Computer Studies at Leeds University for providing the environment in which this book was prepared.

Milan Sonka was a faculty member of the Department of Control Engineering, Faculty of Electrical Engineering, Czech Technical University, Prague, Czech Republic for ten years, and is now an Associate Professor at the Department of Electrical and Computer Engineering, the University of Iowa, Iowa City, Iowa, USA. His research interests include medical image analysis, knowledge-based image analysis, and machine vision. Vaclav Hlavac is an Associate Professor at the Department of Control Engineering, Czech Technical University, Prague. His research interests are knowledge-based image analysis and 3D model-based vision. Roger Boyle is a Senior Lecturer in Artificial Intelligence in the School of Computer Studies at the University of Leeds, England, where his research interests are in low-level vision and pattern recognition. The first two authors have worked together for some years, and have been co-operating with the third since 1991.

The authors have spent many hours in discussions with their teachers, colleagues, and students, from which many improvements to early drafts of this text resulted. Particular thanks are due to Tomáš Pajdla, Petr Kodl, Radim Šára at the Czech Technical University; Steve Collins at the University of Iowa; Jussi Parkkinen at the University of Lappeenranta; Guido Prause at the University of Bremen; David Hogg at the University of Leeds; and many others whose omission from this list does not diminish the value of their contribution. The continuous support and encouragement we received from our wives and families, while inexplicable, was essential to us throughout this project—once again, we promise that our next book will not be written outside standard office hours or during holidays (but this time we mean it).

All authors have contributed throughout—the ordering on the cover corresponds to the weight of individual contribution. Any errors of fact are the joint responsibility of all, while any errors of typography are the responsibility of Roger Boyle. Jointly, they will be glad to incorporate any corrections into future editions.

Milan Sonka (milan-sonka@uiowa.edu)
The University of Iowa, Iowa City, Iowa, USA

Vaclav Hlavac (hlavac@vision.felk.cvut.cz)
Czech Technical University, Prague, Czech Republic

Roger Boyle (roger@scs.leeds.ac.uk)
University of Leeds, Leeds, England

Course contents

In this section, one possible ordering of the material covered in the four courses proposed in the Preface is given. This coverage should not be considered the only possibility—on the contrary, the possibilities for organizing Image Processing and Analysis courses are practically endless. Therefore, what follows should only be regarded as suggestions, and instructors should tailor course content to fit the already acquired knowledge, abilities, and needs of the students enrolled.

Digital Image Processing. An undergraduate course.

1 Introduction

2 The digitized image and its properties

3 Data structures for image analysis

4 Image pre-processing (excluding 4.3.6–4.3.9, 4.4.3, limited coverage of 4.3.4, 4.3.5)

5 Segmentation
  - 5.1 Thresholding (excluding 5.1.3, 5.1.4)
  - 5.2 Edge-based segmentation (excluding 5.2.8, limited coverage of 5.2.4, 5.2.5)
  - 5.3 Region growing segmentation (excluding 5.3.4)
  - 5.4 Matching

12 Linear discrete image transforms

13 Image data compression

16 Case studies (selected topics)

Digital Image Analysis. An undergraduate/graduate course, for which Digital Image Processing may be regarded as prerequisite. Sections that were covered in the Digital Image Processing class and re-appear are intended to be discussed at more depth than is possible in the introductory course.

1 Introduction (brief review)

2 The digitized image and its properties (brief review)

5 Segmentation
  - 5.1.3 Multi-spectral thresholding
  - 5.1.4 Thresholding in hierarchical data structures

  - 5.2.4 Edge following as graph searching
  - 5.2.5 Edge following as dynamic programming
  - 5.3.4 Watershed segmentation
- **6** Shape representation and description (excluding 6.2.7, 6.3.4–6.3.6, 6.4)
- **7** Object recognition
  - 7.1 Knowledge representation
  - 7.2 Statistical pattern recognition
  - 7.3 Neural networks
  - 7.4 Syntactic pattern recognition
- **11** Mathematical morphology
- **14** Texture
- **16** Case studies (selected topics)

Computer Vision I. An undergraduate/graduate course, for which Digital Image Processing may be regarded as prerequisite.

- **1** Introduction (brief review)
- **2** The digitized image and its properties (brief review)
- **4** Image pre-processing
  - 4.3.3 Zero-crossings of the second derivative
  - 4.3.4 Scale in image processing
  - 4.3.5 Canny edge detection
  - 4.3.6 Parametric edge models
  - 4.3.7 Edges in multi-spectral images
  - 4.3.8 Other local pre-processing operators
  - 4.3.9 Adaptive neighborhood pre-processing
- **6** Shape representation and description
- **7** Object recognition
- **8** Image understanding
- **16** Case studies (selected topics)

Computer Vision II. A graduate course, for which Computer Vision I may be regarded as prerequisite.

- **5** Segmentation
  - 5.2.4 Edge following as graph searching
  - 5.2.5 Edge following as dynamic programming
  - 5.5 Advanced border and surface detection approaches
- **9** 3D Vision, geometry and radiometry
- **10** Use of 3D vision
- **15** Motion analysis
- Practical 3D vision projects

Index

Bold text refers to major or defining entries.

2.5D sketch, 445–446, 520
2D co-ordinate system, 510, 538
2D projection, 231, 232, 521, 522, 541, 544, 680
2D shape, 43, 228, 231
3D co-ordinate system, 537
3D information, 228, 509
3D interpretation, 444, 510, 522
3D model, 520
3D object, 519
3D representation, 11, 445, 446, 520, 521
3D shape, 43, 228, 232

A-algorithm, 148–156, 160, 161
ACRONYM, 373, 527
active perception, 447
active sensor, 484
active shape model, *see* ASM
active vision, 512
acuity, 34
adaptive neighborhood, 98–102
ADEOS, 621
albedo, 492
algorithm (in Marr's theory), 444
aliasing, 20
anti-aliasing, 22
anti-extensive transformation, 566
arc (of a graph), 47
area, 45, 51, 237, 254–256, 260, 292, 527, 560
area-based stereo, 480
ASM, 387–390, 722
aspect, 545
aspect graph, **545**
autocorrelation, 627

B-reps, **527**
back propagation, *see* neural nets
back-projection, 538, 539
back-tracking, 160, 161, 320, 321, 326, 327, 373, 404, 538, 539
background, 29
ball, 577
  geodesic, 585
  maximal, 577
  unit, 577
balloon, 378
baseline, 458
bay, 32
Bayes formula, 300, 301, 391, 394, 395
BDRF, 491
bin-picking, 481, 541
blackboard, 372, 373
blocks world, **521**
blur, 745
  Gaussian, 33, **84**, 90, 97
Boltzmann constant, 333
border, **30**, 34
  detection, **134**, 335
    optimal, **148–163**, 194
    simultaneous, 194, 731
  extended, 144–146
  inner, 30, 142
  inter-pixel, 144
  outer, 30, 142–145
border detection, 722
boundary, *see* border
  occlusion, 496
boundary representations, *see* B-reps
brightness, 2, 3, 5, **10**, 11, 12, 18, 22, 23, 27, 29, 32–34, 37, 42–44, 46, 47, 52, 57–59, 61, 63, 72–74, 77, 94, 98–102, 124, 126–128, 133, 134, 146, 147, 174, 176–178, 180, 181, 185, 259, 261, 394, 488, 495, 543, 621–

623, 630, 632, 649, 653, 655, 656, 659, 667, 682, 683
  correction, 58
  interpolation, 65–68
  transformation, 58–61
brightness interpolation
  bi-cubic, 67
  linear, 67
  nearest neighbor, 66

calculus, 27, 77, 329, 559
camera, 10, 36, 43, 58, 103, 105, 250, 253, 679, 680, 686
  extrinsic parameters, 452
  intrinsic parameters, 452
  self-calibration, 454
Canny edge detector, 80, 90–93
center of gravity, 260, 272
chain, **45**
  Markov, 659
chain code, 45, 146, **236–238**, 244, 255
chamfer matching, *see* matching, chamfer
chamfering, 27, 192, 193
characteristic strip, 495
characteristic view, **544**
chromaticity, 26
class, 297
  identifier, 298
classification
  contextual, **392–397**, 403
    recursive, 395
classifier, **297–306**, 319
  Φ, **300**, 302
  best approximation, 302
  learning, 302, **303–306**, 321
  linear, **299**
  maximum likelihood, 300
  minimum distance, 299, **306–308**, 724
  minimum error, 300–306
  non-linear, 300
  setting, 300–303
  syntactic, 319
clique, 327, 328
closing, 568–569
cluster analysis, 307, 308, 722
clustering, *see* cluster analysis
CMY, 25
co-lineation, 449
co-occurrence matrix, **44**
co-ordinate system
  2D, *see* 2D co-ordinate system
  3D, *see* 3D co-ordinate system
  object-based, 11, 446, 520
  polar, 235
  rectangular, 235
  tangential, 235, 237
COBIUS, 372
code
  chain, *see* chain code
  dictionary, 629, 633
  Freeman, 45
  leaf, 52
  run length, 46
coding
  Huffman, 633
  low and high frequencies, 632
  region border, 632
color, 34
  image, 23
  palette, 59, 633
  primary, 23
  secondary, 24
colorimetry, 12
compatibility function, **398**, 410
compression, 3, 5, 621–637
  application
    asymmetric, 629, 636
    symmetric, 629, 636
  dictionary-based, 633
  DPCM, **627**, 629
  fractal, 632
  hierarchical, 630–632
  hybrid, 621, 629, 630, **632**
  JPEG, 606, 612, 633–635
  Lempel-Ziv, 633
  MJPEG, 635, 637
  MPEG, 612, 634, 636–637
  predictive, 621, **624–629**, 632
  progressive, 630–631

  pyramid, 630
  ratio, **623**, 624, 627, 630
  region growing, 632
  smart, 630, 631
  transform, 621, **623**, 631, 632
  vector quantization, **629**, 632
  wavelet, 624
computed tomography, 738
computer graphics, 11, 43, 62, 245, 663
confidence, 395–413
conic, 251, 252
constraint
  epipolar, *see* epipolar constraint
  propagation, **397–401**, 405, 406, 521
constructive solid geometry, **525**
context, 123, 137–139, 364, 367, **392–397**, 405, 406
contour
  false, 23
  partitioning, 244
  shape from, *see* shape from contour
contrast, 33, 34, 59, 60, 97, 99, 130, 135, 189, 543, 652, 654, 683
  enhancement, 100–102
control strategy, 291, **363–373**, 399
  bottom-up, 157, 290, **365–366**, 367, 368, 371, 400
  combined, 367, 371
  hierarchical, **364**, 371
  hypothesize and verify, *see* hypothesize and verify
  model-based, 366–368
  non-hierarchical, 371, 372
  parallel, **364**, 366, 371
  serial, **364**
  top-down, 365–368, 371
convex hull, 31, 559
convolution, **13**, 192, 600, 602, 604, 608, 609
  mask, 69
  theorem, 14
core, 269
corner, 97
correlation, 16, **191–194**, 248, 621–624
correlation-based correspondence, 480
correspondence, 63, 97, 509, 510, 680, 681, 696, 697, 699, 704
  problem, 97, 482, 509, 510
  stereo, *see* stereo correspondence
correspondence problem, 476
cost function, 148–156, 722, 731, 735
  automated design, 163
cost transform, 154, 163
crack edge, 31, 138, 139, 141, 144, 148, 178, 179
criterion
  detection, 90
  localization, 90
  one response, 90
crossover, 331–333
CSG, *see* constructive solid geometry
CT imaging, 722
curvature, 237
  peak, 244
  primal sketch, 245
curvature primal sketch, 542
curve
  decomposition, 245
  detection, 167, 169, 171, 173
  granulometric, 589
cyclopean image, **481**
cyclopean separation, 482

daemon, 373, 399
data structure, **42**
  hierarchical, 49
  relational, 48
  traditional, 43
de-centering, 456
de-fuzzification, 339, 342
  composite maximum, 342
  composite moments, 342
decimation, 530
decision rule, 298, 300, 304
deficit of convexity, 32
deformable objects, 520
degradation, 105
  atmospheric turbulence, 105

  relative motion of the camera and object, 105
  wrong lens focus, 105
depth, 11, 515
depth map, 445, 484, 514
diagram
  Voronoi, 404
difference image, *see* image, difference
difference of Gaussians, 86
dilation, 563–565, 745
  conditional, 592
  geodesic, 586
  gray-scale, 569–574
Dirac distribution, 13
discrete topology, 30
discrimination function, **298–302**
disparity, **459**, 482
  gradient, **481**, 482
  gradient limit, 481
distance, **27**, 192
  chessboard (D_8), 27
  city block (D_4), 27
  Euclidean (D_E), 27
  geodesic, 585
  Levenshtein, 328
distance function, 584
DoG, *see* difference of Gaussians
duality (morphological), **561**
dynamic programming, 158–161, 730
  live lane, 163
  live wire, 162

Ebbinghaus illusion, 34
edge, 3, 4, **30**, 445
  chain, 134, 158
  crack, 31
  detector, 335
  direction, 77
  magnitude, 77
  relaxation, **137–142**, 156, 161
edge detection, 722
edge detector, 77–88, 445, 537, 730, 734
  Canny, 90–93, 655
  compass, 81
  facet model, 93
  in multi-spectral image, 94
  Kirsch, **83**
  Laplace, 78, **81**
  Marr-Hildreth, 83
  parametric, 80, 93
  Prewitt, **81**
  Roberts, **80**
  Robinson, **83**
  Sobel, **82**
  zero-crossing, 83
EGI, 543
ego-motion, 467
elastics, 367
entropy, 15, 621–623, 654
epipolar constraint, 459, **477**, 483
epipolar line, 458, 483
epipolar plane, 458
epipolar transfer, 476
epipole, 458
erosion, 565–567, 745
  geodesic, 586
  gray-scale, 569–574
  ultimate, 582
errors
  matching, 535
essential matrix, 462
Euler-Poincaré characteristic, 256, 560
evaluated graph, 47
event (aspect), 545
exemplar, **299**, 306–308, 313, 314
expansion
  isotropic, **563**
extended boundary, *see* border, extended
extended Gaussian image, 543
extrinsic parameters, 452

facet, 93
feature, 292
  discriminativity, 303
  informativity, 303
  space, 297
  vector, 292
feature synthesis, 91
feature-based correspondence, 481
feedback, 3, 135, 631
fill, **563**

filter, 57–107, 600–613
  Gable, 691
  Gabor, 691
  Gaussian, 445
  median, 74–76
filtering, 68
  band-pass, 609–611
  high-pass, 609, 611
  inverse, 106
  Kalman, 105
  low-pass, 609, 611
  Wiener, 106
fitness, 409
focal point, 449
focus
  shape from, *see* shape from focus
forward algorithm, 420
Fourier descriptor, *see* shape description, Fourier
Fourier transform, *see* transform, Fourier
fractal, 248, 661
  dimension, 237, 657
frame, 295, 296
free-form surface, 519
Freeman code, **45**
frequency
  spatial, 14
function, 10
  autocorrelation, 16
  autocovariance, 16
  cross correlation, 16
  cross covariance, 16
  Dirac, 13
  distance, 584
  distribution, 15
  point spread, 17
  quench (morphology), 581
fundamental matrix, 460
fuzzy
  complement, 339
  composition, 339
    min–max, 340
  correlation
    minimum, 340
    product, 341
  intersection, 339
  logic, 294, 336–344, 743–744
  membership function, 336
    maximum normal form, 337
    minimum normal form, 337
  reasoning
    monotonic, 340
  set, 336
    hedge, 337
  space, 336
  system, 336–344
    model, 339
  union, 339

ganglion cell, 88
Gaussian blur, *see* Gaussian, blur
Gaussian filter, 84, 86, 445
generalized cones, *see* generalized cylinders
generalized cylinders, **526**
genetic algorithm, 330–333, 344, 409–416
genus, 256
geodesic transformation, 585
Geographical Information Systems, 52
geometric signals, 483
geometric transformation, 2, 62–68, 722
geon, **545**
Gestaltist theory, 509
GIF, 633
GIS, 52
Golay alphabet, 579
gradient descent, 310
gradient operator, 68, 77–88
  approximated by differences, 79
  Kirsch, 83
  Laplace, 81
  Prewitt, 81
  Roberts, 80
  Robinson, 83
  Sobel, 82
gradient space, **490**
grammar, **292**, 316–322, 660–664
  context-free, 318
  context-sensitive, 318
  fuzzy, 319
  general, 318

  inference, 316, **321–323**
  non-deterministic, 318
  regular, 318
  stochastic, 319
granulometry (morphological), 589
graph, **47**, 144, 148, 158, 194, 254, 267, 292, 293, 316, 320, 323–328
  arc, 47
  assignment, 327, 328
  evaluated, 47, 144, 295, 323, 324
  isomorphism, 323–328
  neighborhood, 272
  node, 47
  region, 267, 270
  region adjacency, 47, 53, 124, 180, 182, 272, 401, 405, 406, 409–412, 742
  search, **148–156**, 161, 367, 368, 730, 734
    advanced approaches, 194
    heuristic, **151–157**, 161
    three-dimensional, 194, 731
  similarity, 323, 328
graph matching, 536
graph search, 722
gray-level, *see* brightness
gray-scale transformation, 59
grid, 22
  hexagonal, 22
  square, 22
group, 250, 252
  Lie, 250
  plane-projective, 250
grow, **563**

HEARSAY, 373
heuristic, **5**, 151, 152, 156, 161, 177–179, 190, 405, 406, 416
hidden Markov model, **417–423**
  Baum-Welch algorithm, 422, 423
  decoding, 418, 420–422
  evaluation, 418–420
  forward algorithm, 420
  Forward-Backward algorithm, 422
  learning, 418, 422
  Viterbi algorithm, 420, 423
histogram, **32**, 123, 127–129, 131, 178, 622, 623
  bi-modal, 127–129
  cumulative, 60
  equalization, 25, 60–61
  modification, 100
  multi-dimensional, 132
  multi-modal, 128
  smoothed, 131
  transformation, 128
hit-or-miss transformation, **568**
HMM, *see* hidden Markov model
hole, 29
homogeneity, 176, 177, 181–190
homogeneous co-ordinates, 448
homotopic substitute (of skeleton), 579
homotopic transformation, 576
Hopfield networks, *see* neural nets
horizon, 516
Hough transform, *see* transform, Hough
HSI, 25, 34
hue, 25
human visual system, 445
hypothesis, 362, 366, 409–416
hypothesize and verify, 244, 273, **366**, 367, 409, 536
hypothesize-and-verify, 538
hysteresis, 91, 92, 136

ICP algorithm, 533
IHS, *see* HSI
illumination, 488
image, **10**
  binary, 23, 44
  co-ordinates, 12
  color, 23
  compression, *see* compression, *see* compression
  cyclopean, **481**
  difference, 682–684, 745
  digitization, 18–26
  dynamic, 12
  enhancement, 57
  iconic, 42
  intensity, *see* intensity image

  interpretation, **363–417**, 722
  multi-spectral, 23
  pre-processing, 57–107
  quality, 35
  reconstruction, 621–624, 631
  restoration, *see* restoration
  scale-space, 89, 245
  segmented, 43
  sharpening, 79
  skew, 62, 64, 65, 722, 723
  smoothing, *see* smoothing
  static, 12
  transform, 600–613
  understanding, 5, **362–417**
image irradiance equation, 493
image plane, 449
image rectification, 466
image sharpening, 3
imaging
  ultrasound, 734
implementation (in Marr's theory), 444
impossible objects, 522
impulse
  Dirac, 13
  limited, 20
increasing transformation, 564, 566
inference, 363
intensity, *see* brightness, 25
intensity axis of symmetry, 269
intensity image, 11, 57, 83, 366, 496, 537, 542, 543, 696
interest point, 97
interpretation
  3D, *see* 3D interpretation
  genetic, 408
  tree, 404, 536
interval tree, 89, 245
intrinsic parameters, 452
invariants, 231, 249–252, 542
  scalar, 250
inverse filtering, 106
inverse transformation, 568
irradiance, 488
irradiance equation, 493
ISODATA, 308, 391
isotropic expansion, **563**

IYQ, *see* YIQ

Kalman filter, 105, **708–710**, 722, 747–749
Kalman gain matrix, **709**
knowledge, 3, 5, **291–296**, 330, 333
  a priori, **6**, 135, 148, 164, 173, 175, 182, 230, 368, 373, 391, 406, 409, 413, 735, 740
  base, 291, 293, 294
  procedural, 294
  representation, 291–296, 363
Kohonen feature maps, *see* neural nets
Kohonen networks, *see* neural nets

label, 232–235, 373, 391–404, 406, 407, 410–412, 740–742
  collision, 233–235
labeling, 232, 233, 255, 373, 391, 393, 395, 396, **397–417**, 722
  consistent, 397, 399
  discrete, **398**, 404
  probabilistic, 397, **400**
  semantic, 397
lacunarity, 657
Lagrange multipliers, 497, 687
lake, 32
Lambertian surface, 492
landmarks, 274
Landsat, 621
language, 316–322
Laplacian, 78, 445
Laplacian of Gaussian, 84
learning, 299, **303–307**, 311, 317, 322, 333
  from experience, 363
  unsupervised, 307
LIDAR, 484
light, 58, 88
  source, 11, 495
line
  detector, 537, 539
  finding, 94
  labeling, 400, 521–523, 536
  thinning, 96
linear system, 17
linguistic

  - variable, 294, 337, 340
- live lane, 163
- live wire, 162
- local pre-processing, 68–102
- local shading analysis, 497
- locus
  - visibility, 538–540
- logic
  - fuzzy
    - training, 743
- luminance, 25
- luminous efficacy, 487
- luminous flux, 487
- LZW, *see* Lempel-Ziv-Welch

- magnetic resonance, 738
- map
  - region, 47
- marker, 591
- Markov chain, 659
- Markov model, **417**, 423
- Marr (David), 5, 444
- Marr paradigm, *see* Marr's theory
- Marr's theory, 83, 366, 444–446, 520
- Marr-Hildreth edge detector, 80, 83, 90, 739
- matching, **190–194**, 328, 330, 363
  - chamfer, 27, 192
  - errors, 535
  - graphs, 191, 323
  - relational structures, 321
  - sub-graphs, 328
- mathematical morphology, 268
- matrix, 43
  - camera calibration, 451
  - co-occurrence, **44**
  - essential, 462
  - fundamental, 460
  - projective, 453
- maximal ball, 577
- MDL, 531
- medial axis, 271
- medial axis transform, 268
- median filter, *see* filter, median
- memory
  - long-term, 373
  - short-term, 373
- Mexican hat, 85
- Minkowski algebra, 563
- model, 123, 155, 156, 172, 176, 362, 363, **365–373**, 408
  - 3D, 520
  - active contour, 174, 367, **374–380**, 681
  - base, 536
  - deformable, 367, **374–380**
  - facet, 93, 95, 97
  - hidden Markov, *see* hidden Markov model
  - Markov, *see* Markov model
  - partial, 535
  - quadric surface, 529
  - surface, 520
  - volumetric, 520, **523**
- model-based vision, 535
- modes of variation, 383
- Moiré interferometry, 486
- moment
  - invariant, 260
    - affine, 261
- Moravec detector, 97
- morphological noise reduction, 722
- morphological transformation, 561
  - quantitative, **562**
- morphology, 559–595, 659
- motion, 508, 679–708
  - analysis, 679–708
    - correspondence of interest points, 680, 681, **696–704**
    - differential, 681–684
  - assumptions, 681, 700, 704
  - continuous, 512
  - correspondence of interest points, 705
  - cyclic, 705
  - description length, 702
  - events, 705
  - features, 705
  - field, 680
  - gesture interpretation, 705
  - lipreading, 705
  - object tracking, 700–708

  - path coherence, 700
    - deviation, 700
    - function, 700, 701
  - recognition, 705
  - relative, 705
  - rotational, 510, 513, 686
  - shape from, *see* shape from motion
  - trajectory
    - parametrization, 705
  - translational, 510, 513, 693, 694
  - verb recognition, 705
- multi-view representation, **544**
- mutation, 331–333

- Necker cube, 444
- neighbor, 28
- neural nets, 308–315, 344
  - adaptive resonance theory, 313
  - back-propagation, **310–311**
  - epoch, **311**
  - feed-forward nets, 310–311
  - gradient descent, 310
  - Hopfield, **313–315**
  - Kohonen networks, **312–313**
  - momentum, **311**
  - perceptron, 309
  - transfer function, 309
  - unsupervised learning, 312–313
- node (of a graph), 47
- noise, 3, 4, 35, 229, 236, 239, 255, 260, 262, 269, 392
  - additive, 36
  - Gaussian, 35, 708
  - impulsive, 37
  - multiplicative, 37
  - quantization, 37
  - salt-and-pepper, 37
  - suppression, 99
  - white, 35, 708
- non-maximal suppression, 91, 96, 135
- NURBS, 527

- object
  - coloring, 232
  - connected component labeling, 148, 232
  - description, 297
    - formal, 297
    - qualitative, 297, 315
    - quantitative, 297, 315
    - relational, 315
  - identification, 232–235
  - impossible, 522
  - labeling, 232, 233, 255
  - recognition, 290–335
  - reconstruction, 228, 229
- objective function, 402, 403, 407–410
- occlusion, 229, 232, 242, 244, 253, 273, 522, 681, 749, 750
- occupancy grid, 523
- OCR, 4, 62, 228, 418, 420, 423, 722, 724
- octrees, 52
- opening, 568–569
- operator
  - morphological, 562
  - Zadeh, 339
- optical axis, 449
- optical center, 449
- optical character recognition, *see* OCR
- optical flow, 497, 512, 513, 680, **685–696**, 699
  - computation, 680, 681
  - field, 512
  - global and local estimation, 689
- optimization, 158, 159, 193, 303, 306, 313, **328–335**, 386
- orthographic
  - projection, 11, 510, 512
  - view, 510

- palette, *see* look-up table, 24, 59, 633
- parallel implementation, 5, 44, 49, 141, 142, 146, 162, 173, 364, 397, 399, 403, 416, 417, 483, 564, 688, 690
- path, 29
  - simple, 146
- pattern, 297
  - space, 297
  - vector, 297
- pattern recognition, 290–323

  statistical, **292–308**, 315
  syntactic, 315–323
PDM, **380–390**, 722, 746
  alignment, 381
  covariance matrix, 383, 384
  eigen-decomposition, 383
  landmark, **381**, 382–385, 387–389
  modes of variation, **383**
  polar, 390
  polynomial regression, 390
perception, 33–35, 363, 660
  color, 34
  human, 22, 33, 100, 515
  visual, 33
perceptron, *see* neural nets
perimeter, 143, 178, 179, 232, 237, 560
perspective
  projection, 11, 512, 542
perspective projection, 448
photometric stereo, 498
photometry, 12, 487
picture element, **22**
pigment, 25
pixel, **22**
  adjacency, 28
pixel co-ordinate transformation, 63
planning, 363
plausibility (of a match), 539, 541
point
  representative, 560
  sampling, 22
  sets (morphological), **560**
point distribution model, *see* PDM
post-processing, 392, 393, 408, 416
power spectrum, 17
pre-processing, 57–107, 365, 393
  adaptive neighborhood, 98–102
  classification, 57
  edge detector, *see* edge detector
  local, 68–102
predicate logic, 293, 294
primal sketch, 445–446, 520
  curvature, 245
primitive
  texture, 515
  volumetric, 446
principal components analysis, 383
probability
  density, 304
    estimation, 304
production
  rules, 293
  system, 293, 294, 363, 373
projection, 256, 510, 512, 538, 560, 608
  2D, *see* 2D projection
  histogram, 256, 722, 724
  orthographic, 11, 510, 512
  parallel, 11
  perspective, 11, 512
projective matrix, 453
projective transformation, 449
PROLOG, 293
pseudo-color, 61
purposive vision, 448
pyramid, **49**, 133, 134, 182–184, 364, 630, 632
  equivalent window, 53
  irregular, 53
  Laplacian, 53
  M-pyramid, 49
  matrix, 49
  reduction factor, 52
  reduction window, 52
  regular, 52
  T-pyramid, 49
  tree, 49

quadric surface model, 529
quadrilinear constraint, 473
quadtree, **51**, 182, 189, 235, 237, 255, 272, 630
qualitative vision, 447
quantization, 22
quench function, 581

RADAR, 484
radial distortion, 456
radiance, 488
radiant flux, 487
radiometry, 486
random dot stereograms, 480
range image, 484, 529

receptive field, 88
reconstruction (morphological), 584, 586
rectification, 466
reduce, **565**
redundancy
  information, 621–624
reference view, 546
reflectance, 11, 12, 490
reflectance coefficient, 492
reflectance function, 492, 495
reflectance map, 492
region, **28**, 30
  concavity tree, 266
  decomposition, 254, **271–272**
  identification, 232–235
  skeleton, 248, 254, 267–270, 272
region adjacency graph, **47**, **53**
region growing, *see* segmentation, region growing
region map, 47
regional extreme, 582
registration, 529
relation
  neighborhood, 44
  spatial, 43
relational structure, 294, 315, 316, 321, 322, 363
relaxation, 128, 138, 139, 141, 189, **398–404**, 407, 417, 482, 497, 687, 689, 690
  discrete, 398–400
  probabilistic, 400, 403
reliability (of a match), 539
remote sensing, 62, 65, 131, 132, 393, 397, 406, 621, 623, 705
representation, 444
  3D, 446
  complete, 520
  geometric, 43
  iconic image, 42
  intermediate, 11, 42
  level of, 42
  multi-view, **544**
  relational model, 43
  segmented image, 43
  skeleton, 527
  unique, 520
reproduction, 331, 332
resolution
  radiometric, 12
  spatial, 5, 12
  spectral, 12
  time, 12
restoration, 102–107
  deterministic, 103
  geometric mean filtration, 107
  inverse filtering, 106
  power spectrum equalization, 107
  stochastic, 103
  Wiener filtering, 106
RGB, 23
rigidity, 510
rigidity constraint, 510
rim, 518
rotating mask, 73
rotational movement, 510, 513, 686
run length coding, **46**, 234, 272, 630

SAI, 543
sampling, 18–22
  interval, 19
  point, 22
saturation, 25
scale, 88–93, 229, 445, 646–648, 654, 660
scale-space, 88, 89, 229, 244, 245, 269, 274, 480
scene reconstruction, 471
script, 295
seed pixel, 98
segmentation, 4, 123–194, 365
  border detection, 148–156, 173, 727–737
    simultaneous, 731
  border tracing, 142–147, 161
    extended, 144–146
    inner, 142
    outer, 142–145
  classification-based, 390–393
  complete, 4, 123, 124, 174
  dynamic programming, 158–161
  edge thresholding, 135

  edge-based, 123, **134–175**, 176, 188
  global, 123, 124
  Hough transform, 162, **163–173**
    generalized, 165, 171, 172
  match-based, 190–194
  morphological, 590
  multi-thresholding, 128
  partial, 4, 123, 134, 135, 174
  region construction
    from borders, 174
    from partial borders, 175
  region growing, 123, 176–186, 188, 189, 404, 406, 408, 416, 722, 739
    color image, 176
    merging, 144, 177–181
    over-growing, 188
    semantic, 405, **406–408**
    split-and-merge, 181, 182
    splitting, 177, 178, 181
    under-growing, 188
  region-based, 123, 146, 176–186
  region-growing, 99
  semantic, 404
    region growing, 406–408
  semi-thresholding, 126
  superslice, 174
  texture, 335
  thresholding, **124–134**, 174, 175, 180, 181, 739
    hierarchical, 134
    minimum error, 129
    multi-spectral, 131
    p-tile, 127, 135
  tree, 182–184
  watersheds, 186, 408
self-calibrtion, 454
self-occlusion, 476
semantic net, 294–295, 363, 397
semantics, 291, 294, 400, 405, 408, 416, 722
sequential matching, 538
sequential thickening, *see* thickening, sequential
sequential thinning, *see* thinning, sequential
set difference, 560
shading, 494
  shape from, *see* shape from shading
Shannon sampling theorem, 20
shape, 228–273, 519
  2D, *see* 2D shape
  3D, *see* 3D shape
  class, 229, 273
  description
    area, 237, 241, **254–256**, 257, 258, 260, 271
    bending energy, **237**
    border length, **237–238**
    chord distribution, **239**
    compactness, 144, 254, **259**, 292, 412
    contour-based, 229, 232, **235–253**
    convex hull, 262, 266
    cross ratio, 250, 251
    curvature, 232, **237**, 242–248, 271
    direction, 254, **258**
    eccentricity, 256
    elongatedness, 229, 254, **257**, 258, 269
    Euler's number, **256**
    external, 229, 232
    Fourier descriptors, 238, **240–242**
    graph, 254, **267**
    height, 256
    internal, 229, 232
    invariants, 249–252
    moments, 241, 248, 258, **259–262**, 272
    moments, area-based, 262
    moments, contour-based, 261, 262
    perimeter, 143, 178, 179, 232, **237**
    polygonal, **242–244**, 245, 271
    projection-invariant, 231, 249
    projections, 256

        rectangularity, 254, **258**
        region concavity tree, 266
        region-based, 229, 232, 248, **254–273**
        run length, 234, 272
        segment sequence, 239, **242**
        signature, **239**
        sphericity, 258
        spline, **245–248**, 722, 745, 746
        statistical, 229, 254
        syntactic, 228, 229, 243, 272
        width, 256
shape from
    contour, 518
    de-focus, 517
    focus, 517
    motion, 508–515
    optical flow, 512
    shading, 494–497, 516
    stereo, 476–483
    texture, 515–517
    vergence, 517
    X, 445
shape from X, 508
shape primitive, 271
sharpening, 79
shrink, **565**
sieving analysis, 589
sifting, 13
signal-to-noise ratio, *see* SNR
silhouette, 518
simplex angle image, 543
simulated annealing, 333–335
singular point, 496
singular value decomposition (SVD), 456
skeleton, 248, 254, 267–270, 272, 527, 559, 576–579
    by influence zones, 585
    by maximal balls, 577
sketch
    2.5D, *see* 2.5D sketch
    primal, *see* primal sketch
skew, 62, 64, 65, 722, 723
skewing, *see* skew
SKIZ, skeleton by influence zones, 585
slant, **515**
smart snake, 387
smoothing, 69–77
    averaging, 69
    averaging according to inverse gradient, 72
    averaging with limited data validity, 71
    edge preserving, 69
    Gaussian, 244
    Gaussianr, 244
    median, *see* filter, median
    non-linear mean filter, 77
    order statistics, 76
    rank filtering, 76
    rotating mask, 73
snake, 174, 367, **374–380**, 681
    growing, 377
SNR, 36
spatial angle, 488
spectral density, 17
spectrum, 17
    band-limited, 20
    frequency, 589, 603, 611
    granulometric, 589
    phase, 603
    power, 603
spline, *see* shape, description, spline
state space search, 178
stereo
    shape from, *see* shape from stereo
stereo correspondence, 335, 509
stereo vision, 457, 458, 509
stereopsis, 457
stochastic process, 15–17
    ergodic, 17
    stationary, 16
    uncorrelated, 16
structure from motion theorem, **510**
structured light, 484
structuring element (morphological), **560**
super-quadrics, 525
supergrid, 144, 178, 179
surface
    detection, 194

  - free-form, 519
- surface features, 542
- surface models, *see* B-reps
- surface reflectance, 490
- surveillance, 744
- SVD, singular value decomposition, 456
- sweep representations, *see* generalized cylinders, 526
- symbol
  - non-terminal, 317
  - terminal, 317
- symmetric axis transform, 268
- syntactic analysis, **316–322**, 660, 662, 663, 722
- syntax, 291
- system approach, 88
- system theory, 443

- texel, 515, 646
- texton, 649, 654
- texture, 123, 127, 134, 177, 335, 390, 393, 397, 515, **646–668**, 705
  - Brodatz, 651
  - coarse, 646, 648–651, 655, 659
  - description
    - autocorrelation, 649, 660
    - autoregression model, 659
    - chain grammar, 661
    - co-occurrence, 651, 653, 660
    - discrete transform, 650
    - edge frequency, 653
    - fractal, 657, 661
    - grammar, 660
    - graph grammar, 663, 664, 666
    - hybrid, 649, 660, 666–667
    - Laws' energy measures, 656
    - morphology, 659
    - optical transform, 650
    - peak and valley, 659
    - primitive grouping, 664
    - primitive length, 655
    - run length, 655
    - shape rules, 661
    - statistical, 648–660, 664, 666
    - syntactic, 648, 660–666
    - texture properties, 654
    - texture transform, 659
  - element, 646
  - fine, 646, 648–651, 655, 659
  - generation, 661, 662
  - gradient, 515, 516
  - hierarchical, 664
  - primitive, 175, 515, **646**, 648, 649, 660, 661, 663, 665, 667
  - segmentation, 654, 659, 660, 664–666
  - shape from, *see* shape from texture
  - strong, 648, 666
  - structure, 646, 648
  - tone, 646, 648, 667
  - weak, 648, 666
- theory
  - computational, 444
  - Gestaltist, *see* Gestaltist theory
  - Marr's, *see* Marr's theory
- thickening, 559, 578–579
  - sequential, 579
- thinning, 254, 267, 268, 559, 578–579
  - sequential, 579
- three dimensions, *see* 3D
- threshold, 124
  - optimal, 129
  - selection, 124, 127, 129
    - mode method, 128
    - optimal, 128
- thresholding, 59
  - adaptive, 125
  - optimal, 128
  - with hysteresis, 91
- TIFF, 633
- tilt, **515**
- tolerance interval, 242
- top surface (morphological), 570
- top-down approach, 319, 321, 365, 366, 538
- topographic characterization, **543**
- topographic primal sketch, **543**
- topology, 31
- training set, **302–303**, 304, 306, 307, 310, 311, 322, 323, 380, 391, 395

- transform
  - binary, 624
  - cosine, **605**, 624, 632
  - distance, 269
  - Fourier, 12–20, 30, 68, 79, 105–107, 192, 229, 240, 241, 423, 600, **602**, 604, 606, 608–611, 613, 624, 650
    - inverse, 14
  - Gabor, 660
  - Haar, 608
  - Hadamard, **604–605**, 608, 624, 632
  - Hadamard-Haar, 608
  - hat, 85
  - Hough, 162, **163–173**, 248, 516, 608
  - image, 600–613
  - Karhunen-Loeve, 608, 623, 624
  - linear discrete, 600–613
  - orthogonal, 601, 605
  - Paley, 608
  - Radon, 608
  - recursive block coding, 624
  - sine, 608
  - Slant-Haar, 608
  - Walsh, 608, 624
  - Walsh-Hadamard, 605
  - wavelet, 660
  - wavelets, 606, 611
- transformation
  - affine, 64
  - anti-extensive, 566
  - bilinear, 64
  - brightness correction, 58
  - geodesic, 585
  - geometric, 62–68
    - brightness interpolation, 65
    - change of scale, 64
    - pixel co-ordinate, 63
    - rotation, 64
    - skewing, 65
  - gray-scale, 59–61
    - histogram equalization, 60–61
    - logarithmic, 61
    - pseudo-color, 61
    - thresholding, 59
  - hat, 67
  - homotopic, 576
  - increasing, 564, 566
  - inverse, 568
  - morphological, *see* morphological transformation, 561
  - pixel brightness, 58–61
  - projective, 449
  - top hat, 574
- translation (morphological), 561
- translational movement, 510, 513, 693, 694
- transmission
  - progressive, 630, 631
  - smart, 631, 632
- tree
  - interval, 245
- tree pruning, 154, 320
- trilinear constraint, 475
- two dimensions, *see* 2D
- two-dimensional, *see* 2D

- ultimate erosion, 582
- ultrasound, 734
  - intra-vascular, 734
- umbra (morphological), 570
- unit ball, 577
- unsharp masking, 79
- unsupervised learning, 307
- upper semi-continuinty, **563**

- vanishing points, 516
- velocity
  - field, 681
    - computation, 698
  - smoothness constraint, 686, 689, 690
  - vector, 680
- vergence, 517
  - shape from, *see* shape from vergence
- vertex, 521
- view
  - reference, 546
  - topographic (morphology), 570
  - virtual, 546
- viewing space, 544

viewing sphere, 544
viewpoint, 11
vignetting, 490
virtual view, 546
visibility, 540
visibility classes, 549
visibility locus, 538–540
vision
  active, 512
  model-based, 535
  stereo, 458, 509
  view-based, 544
VISIONS, 372, 374
visual
  potential, **545**
visual system
  human, 33, 445, 457, 510, 512, 631, 660
Viterbi algorithm, 420, 423
volume
  partial, 131
volumetric model, **523**
volumetric primitives, 446
voxel, 130, **523**

watersheds, 592
wavelength, 23
wavelets, 606–607, 611, 660
Wiener filtering, 106
WINSOM, 525

X
  shape from, *see* shape from X

YIQ, 25

Zadeh operators, 339
zero crossing, 445
zero-crossing, 80, **83**, 90, 244, 245, 445, 480, 605, 654, 690
Zuniga–Haralick operator, 97, 696