Image Processing, Analysis, and Machine Vision

Milan Sonka
The University of Iowa, Iowa City

Vaclav Hlavac
Czech Technical University, Prague

Roger Boyle
University of Leeds, Leeds
# Contents

<table>
<thead>
<tr>
<th>List of algorithms</th>
<th>xiv</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>xvii</td>
</tr>
<tr>
<td>Possible course outlines</td>
<td>xxii</td>
</tr>
</tbody>
</table>

## 1 Introduction

1.1 Motivation  
1.2 Why is computer vision difficult?  
1.3 Image representation and image analysis tasks  
1.4 Summary  
1.5 References

## 2 The image, its representations and properties

2.1 Image representations, a few concepts  
2.2 Image digitization  
2.2.1 Sampling  
2.2.2 Quantization  
2.3 Digital image properties  
2.3.1 Metric and topological properties of digital images  
2.3.2 Histograms  
2.3.3 Entropy  
2.3.4 Visual perception of the image  
2.3.5 Image quality  
2.3.6 Noise in images  
2.4 Color images  
2.4.1 Physics of color  
2.4.2 Color perceived by humans  
2.4.3 Color spaces  
2.4.4 Palette images  
2.4.5 Color constancy  
2.5 Cameras: an overview  
2.5.1 Photosensitive sensors  
2.5.2 A monochromatic camera  
2.5.3 A color camera  
2.6 Summary  
2.7 References
3 The image, its mathematical and physical background

3.1 Overview
3.1.1 Linearity
3.1.2 The Dirac distribution and convolution

3.2 Linear integral transforms
3.2.1 Images as linear systems
3.2.2 Introduction to linear integral transforms
3.2.3 1D Fourier transform
3.2.4 2D Fourier transform
3.2.5 Sampling and the Shannon constraint
3.2.6 Discrete cosine transform
3.2.7 Wavelet transform
3.2.8 Eigen-analysis
3.2.9 Singular value decomposition
3.2.10 Principal component analysis
3.2.11 Other orthogonal image transforms

3.3 Images as stochastic processes

3.4 Image formation physics
3.4.1 Images as radiometric measurements
3.4.2 Image capture and geometric optics
3.4.3 Lens aberrations and radial distortion
3.4.4 Image capture from a radiometric point of view
3.4.5 Surface reflectance

3.5 Summary
3.6 References

4 Data structures for image analysis

4.1 Levels of image data representation
4.2 Traditional image data structures
4.2.1 Matrices
4.2.2 Chains
4.2.3 Topological data structures
4.2.4 Relational structures

4.3 Hierarchical data structures
4.3.1 Pyramids
4.3.2 Quadtrees
4.3.3 Other pyramidal structures

4.4 Summary
4.5 References

5 Image pre-processing

5.1 Pixel brightness transformations
5.1.1 Position-dependent brightness correction
5.1.2 Gray-scale transformation

5.2 Geometric transformations
5.2.1 Pixel co-ordinate transformations
5.2.2 Brightness interpolation
5.3 Local pre-processing
- Image smoothing 124
- Edge detectors 132
- Zero-crossings of the second derivative 138
- Scale in image processing 142
- Canny edge detection 144
- Parametric edge models 147
- Edges in multi-spectral images 147
- Local pre-processing in the frequency domain 148
- Line detection by local pre-processing operators 154
- Detection of corners (interest points) 156
- Detection of maximally stable extremal regions 160

5.4 Image restoration
- Degradations that are easy to restore 164
- Inverse filtration 165
- Wiener filtration 165

5.5 Summary 167
5.6 References 169

5.3.1 Image smoothing 123
5.3.2 Edge detectors 132
5.3.3 Zero-crossings of the second derivative 138
5.3.4 Scale in image processing 142
5.3.5 Canny edge detection 144
5.3.6 Parametric edge models 147
5.3.7 Edges in multi-spectral images 147
5.3.8 Local pre-processing in the frequency domain 148
5.3.9 Line detection by local pre-processing operators 154
5.3.10 Detection of corners (interest points) 156
5.3.11 Detection of maximally stable extremal regions 160

5.4 Image restoration
- Degradations that are easy to restore 164
- Inverse filtration 165
- Wiener filtration 165

6 Segmentation I

6.1 Thresholding
- Threshold detection methods 179
- Optimal thresholding 180
- Multi-spectral thresholding 183

6.2 Edge-based segmentation
- Edge image thresholding 185
- Edge relaxation 188
- Border tracing 191
- Border detection as graph searching 197
- Border detection as dynamic programming 207
- Hough transforms 212
- Border detection using border location information 221
- Region construction from borders 222

6.3 Region-based segmentation
- Region merging 225
- Region splitting 227
- Splitting and merging 229
- Watershed segmentation 233
- Region growing post-processing 235

6.4 Matching
- Matching criteria 238
- Control strategies of matching 240

6.5 Evaluation issues in segmentation
- Supervised evaluation 241
- Unsupervised evaluation 245

6.6 Summary 246
6.7 References 249
### 7 Segmentation II

7.1 Mean Shift Segmentation 257  
7.2 Active contour models—snakes 265  
7.2.1 Traditional snakes and balloons 265  
7.2.2 Extensions 269  
7.2.3 Gradient vector flow snakes 270  
7.3 Geometric deformable models—level sets and geodesic active contours 275  
7.4 Fuzzy Connectivity 283  
7.5 Towards 3D graph-based image segmentation 291  
7.5.1 Simultaneous detection of border pairs 292  
7.5.2 Sub-optimal surface detection 297  
7.6 Graph cut segmentation 298  
7.7 Optimal single and multiple surface segmentation 306  
7.8 Summary 318  
7.9 References 320

### 8 Shape representation and description

8.1 Region identification 332  
8.2 Contour-based shape representation and description 335  
8.2.1 Chain codes 335  
8.2.2 Simple geometric border representation 336  
8.2.3 Fourier transforms of boundaries 339  
8.2.4 Boundary description using segment sequences 341  
8.2.5 B-spline representation 344  
8.2.6 Other contour-based shape description approaches 347  
8.2.7 Shape invariants 347  
8.3 Region-based shape representation and description 351  
8.3.1 Simple scalar region descriptors 352  
8.3.2 Moments 357  
8.3.3 Convex hull 360  
8.3.4 Graph representation based on region skeleton 365  
8.3.5 Region decomposition 368  
8.3.6 Region neighborhood graphs 369  
8.4 Shape classes 370  
8.5 Summary 371  
8.6 References 373

### 9 Object recognition

9.1 Knowledge representation 381  
9.2 Statistical pattern recognition 386  
9.2.1 Classification principles 387  
9.2.2 Classifier setting 390  
9.2.3 Classifier learning 393  
9.2.4 Support Vector Machines 396  
9.2.5 Cluster analysis 402  
9.3 Neural nets 404  
9.3.1 Feed-forward networks 405
10 Image understanding

10.1 Image understanding control strategies
  10.1.1 Parallel and serial processing control
  10.1.2 Hierarchical control
  10.1.3 Bottom-up control
  10.1.4 Model-based control
  10.1.5 Combined control
  10.1.6 Non-hierarchical control

10.2 RANSAC: Fitting via random sample consensus

10.3 Point distribution models

10.4 Active Appearance Models

10.5 Pattern recognition methods in image understanding
  10.5.1 Classification-based segmentation
  10.5.2 Contextual image classification

10.6 Boosted cascade of classifiers for rapid object detection

10.7 Scene labeling and constraint propagation
  10.7.1 Discrete relaxation
  10.7.2 Probabilistic relaxation
  10.7.3 Searching interpretation trees

10.8 Semantic image segmentation and understanding
  10.8.1 Semantic region growing
  10.8.2 Genetic image interpretation

10.9 Hidden Markov models
  10.9.1 Applications
  10.9.2 Coupled HMMs
  10.9.3 Bayesian belief networks

10.10 Gaussian mixture models and expectation-maximization
# Contents

10.11 Summary 534
10.12 References 537

11 3D vision, geometry 546

11.1 3D vision tasks 547
    11.1.1 Marr’s theory 549
    11.1.2 Other vision paradigms: Active and purposive vision 551
11.2 Basics of projective geometry 553
    11.2.1 Points and hyperplanes in projective space 553
    11.2.2 Homography 555
    11.2.3 Estimating homography from point correspondences 558
11.3 A single perspective camera 561
    11.3.1 Camera model 561
    11.3.2 Projection and back-projection in homogeneous coordinates 565
    11.3.3 Camera calibration from a known scene 565
11.4 Scene reconstruction from multiple views 566
    11.4.1 Triangulation 566
    11.4.2 Projective reconstruction 568
    11.4.3 Matching Constraints 569
    11.4.4 Bundle adjustment 571
    11.4.5 Upgrading the projective reconstruction, self-calibration 571
11.5 Two cameras, stereopsis 573
    11.5.1 Epipolar geometry; fundamental matrix 573
    11.5.2 Relative motion of the camera; essential matrix 575
    11.5.3 Decomposing the fundamental matrix to camera matrices 577
    11.5.4 Estimating the fundamental matrix from point correspondences 578
    11.5.5 Rectified configuration of two cameras 579
    11.5.6 Computing rectification 581
11.6 Three cameras and trifocal tensor 583
    11.6.1 Stereo correspondence algorithms 584
    11.6.2 Active acquisition of range images 591
11.7 3D information from radiometric measurements 594
    11.7.1 Shape from shading 595
    11.7.2 Photometric stereo 598
11.8 Summary 600
11.9 References 601

12 Use of 3D vision 606

12.1 Shape from X 606
    12.1.1 Shape from motion 606
    12.1.2 Shape from texture 613
    12.1.3 Other shape from X techniques 614
12.2 Full 3D objects 617
    12.2.1 3D objects, models, and related issues 617
    12.2.2 Line labeling 618
    12.2.3 Volumetric representation, direct measurements 620
    12.2.4 Volumetric modeling strategies 622
    12.2.5 Surface modeling strategies 624
12.2.6 Registering surface patches and their fusion to get a full 3D model 626
12.3 3D model-based vision 632
12.3.1 General considerations 632
12.3.2 Goad's algorithm 633
12.3.3 Model-based recognition of curved objects from intensity images 637
12.3.4 Model-based recognition based on range images 639
12.4 2D view-based representations of a 3D scene 639
12.4.1 Viewing space 639
12.4.2 Multi-view representations and aspect graphs 640
12.4.3 Geons as a 2D view-based structural representation 641
12.4.4 Visualizing 3D real-world scenes using stored collections of 2D views 642
12.5 3D reconstruction from an unorganized set of 2D views—a case study 646
12.6 Summary 650
12.7 References 651

13 Mathematical morphology 657
13.1 Basic morphological concepts 657
13.2 Four morphological principles 659
13.3 Binary dilation and erosion 661
13.3.1 Dilation 661
13.3.2 Erosion 662
13.3.3 Hit-or-miss transformation 665
13.3.4 Opening and closing 665
13.4 Gray-scale dilation and erosion 667
13.4.1 Top surface, umbra, and gray-scale dilation and erosion 667
13.4.2 Umbra homeomorphism theorem, properties of erosion and dilation, opening and closing 670
13.4.3 Top hat transformation 671
13.5 Skeletons and object marking 672
13.5.1 Homotopic transformations 672
13.5.2 Skeleton, maximal ball 673
13.5.3 Thinning, thickening, and homotopic skeleton 675
13.5.4 Quench function, ultimate erosion 677
13.5.5 Ultimate erosion and distance functions 680
13.5.6 Geodesic transformations 681
13.5.7 Morphological reconstruction 682
13.6 Granulometry 684
13.7 Morphological segmentation and watersheds 687
13.7.1 Particles segmentation, marking, and watersheds 687
13.7.2 Binary morphological segmentation 687
13.7.3 Gray-scale segmentation, watersheds 689
13.8 Summary 691
13.9 References 692

14 Image data compression 694
14.1 Image data properties 696
14.2 Discrete image transforms in image data compression 696
### Contents

14.3 Predictive compression methods
14.4 Vector quantization
14.5 Hierarchical and progressive compression methods
14.6 Comparison of compression methods
14.7 Other techniques
14.8 Coding
14.9 JPEG and MPEG image compression
  14.9.1 JPEG—still image compression
  14.9.2 JPEG-2000 compression
  14.9.3 MPEG—full-motion video compression
14.10 Summary
14.11 References

15 Texture

  15.1 Statistical texture description
    15.1.1 Methods based on spatial frequencies
    15.1.2 Co-occurrence matrices
    15.1.3 Edge frequency
    15.1.4 Primitive length (run length)
    15.1.5 Laws’ texture energy measures
    15.1.6 Fractal texture description
    15.1.7 Multiscale texture description—wavelet domain approaches
    15.1.8 Other statistical methods of texture description
  15.2 Syntactic texture description methods
    15.2.1 Shape chain grammars
    15.2.2 Graph grammars
    15.2.3 Primitive grouping in hierarchical textures
  15.3 Hybrid texture description methods
  15.4 Texture recognition method applications
  15.5 Summary
  15.6 References

16 Motion analysis

  16.1 Differential motion analysis methods
  16.2 Optical flow
    16.2.1 Optical flow computation
    16.2.2 Global and local optical flow estimation
    16.2.3 Combined local–global optical flow estimation
    16.2.4 Optical flow in motion analysis
  16.3 Analysis based on correspondence of interest points
    16.3.1 Detection of interest points
    16.3.2 Correspondence of interest points
  16.4 Detection of specific motion patterns
  16.5 Video tracking
    16.5.1 Background modeling
    16.5.2 Kernel-based tracking
    16.5.3 Object path analysis
16.6 Motion models to aid tracking
   16.6.1 Kalman filters
   16.6.2 Particle filters
16.7 Summary
16.8 References

Acknowledgments

Index