MACHINE VISION
McGRAW-HILL SERIES IN COMPUTER SCIENCE

SENIOR CONSULTING EDITOR
C. L. Liu, University of Illinois at Urbana-Champaign

CONSULTING EDITOR
Allen B. Tucker, Bowdoin College

Fundamentals of Computing and Programming
Computer Organization and Architecture
Computers in Society/Ethics
Systems and Languages
Theoretical Foundations
Software Engineering and Database
Artificial Intelligence
Networks, Parallel and Distributed Computing
Graphics and Visualization
The MIT Electrical Engineering and Computer Science Series

ARTIFICIAL INTELLIGENCE

Bowen: Prolog and Expert Systems Programming
Fu: Neural Networks in Computer Intelligence
*Horn: Robot Vision
Jain, Kasturi, and Schunck: Machine Vision
Levine: Vision in Man and Machine
Rich and Knight: Artificial Intelligence

*Co-published by the MIT Press and McGraw-Hill, Inc.
MACHINE VISION

Ramesh Jain  
*University of California, San Diego*

Rangachar Kasturi  
*Pennsylvania State University*

Brian G. Schunck  
*Adept Technology*
About the Authors

Ramesh Jain is currently a Professor of Electrical and Computer Engineering, and Computer Science and Engineering at the University of California at San Diego. Before joining UCSD, he was a Professor of Electrical Engineering and Computer Science and the founding Director of the Artificial Intelligence Laboratory at the University of Michigan, Ann Arbor. He was also the founder and chairman of Imageware Inc. His current research interests are in multimedia information systems, image databases, machine vision, and intelligent systems.

Ramesh is a Fellow of IEEE, AAAI, and Society of Photo-Optical Instrumentation Engineers, and member of ACM, Pattern Recognition Society, Cognitive Science Society, Optical Society of America, and Society of Manufacturing Engineers. He is currently the Editor-in-Chief of IEEE Multimedia, and is on the editorial boards of Machine Vision and Applications, Pattern Recognition, and Image and Vision Computing. He received his Ph.D. from IIT, Kharagpur, in 1975 and his B.E. from Nagpur University in 1969.

Rangachar Kasturi joined Penn State University in 1982 after completing his graduate studies at Texas Tech University (Ph.D. 1982 and M.S.E.E. 1980). He received a B.E. (Electrical) degree from Bangalore University in 1968. His primary research focus in recent years has been in the area of Document Image Analysis (DIA). His group’s main contribution has been the
design of efficient algorithms to generate intelligent interpretations of engineering drawings and maps to facilitate automatic conversion from paper medium to computer databases. He is the Editor-in-Chief of IEEE Transactions on Pattern Analysis and Machine Intelligence. He was the managing editor of Machine Vision and Applications during 1993–94. He is a coauthor of the tutorial texts, Computer Vision: Principles and Applications and Document Image Analysis, both published by IEEE CS Press, and a coeditor of the book Image Analysis Applications (Marcel Dekker, 1990). During 1987–90 he delivered lectures at many chapters of the IEEE Computer Society through its Distinguished Visitor Program. He has served the International Association for Pattern Recognition in various capacities.

Brian G. Schunck has worked for several years on the development of systems for machine vision and image processing. He was educated in computer science at the University of California, Irvine, where he received the B.S. magna cum laude in 1976. He studied electrical engineering, systems theory, and artificial intelligence at M.I.T., where he received the Master’s and E.E. degrees in 1979 for work on control algorithms for robotic manipulators and the doctorate in 1983 for research on image flow. He was an assistant professor in the Department of Electrical Engineering and Computer Science and a member of the Artificial Intelligence Laboratory at the University of Michigan, Ann Arbor. Currently he is the Director of Vision Software at Adept Technology.

Brian’s current interests include statistical methods for machine vision and industrial inspection; contour, surface, and volume models for computer vision and medical image processing; structure and motion estimation for mobile robots; reverse engineering part models from range data; computer graphics; user interfaces; and marine navigation.

Brian Schunck is a member of the IEEE, ACM, the Society for Industrial and Applied Mathematics, the American Statistical Association, the American Society for Photogrammetry and Remote Sensing, the Society for Manufacturing Engineers, and the Society for Automotive Engineers.
To
Sudha
—Ramesh Jain
My Grandmother
—Rangachar Kasturi
Gizmo
—Brian G. Schunck
Contents

Preface xvii
Acknowledgments xix

1 Introduction 1
  1.1 Machine Vision 1
  1.2 Relationships to Other Fields 4
  1.3 Role of Knowledge 5
  1.4 Image Geometry
    1.4.1 Perspective Projection 8
    1.4.2 Coordinate Systems 9
  1.5 Sampling and Quantization 10
  1.6 Image Definitions 12
  1.7 Levels of Computation
    1.7.1 Point Level 14
    1.7.2 Local Level 15
    1.7.3 Global Level 16
    1.7.4 Object Level 17
  1.8 Road Map 18

2 Binary Image Processing 25
  2.1 Thresholding 28
  2.2 Geometric Properties 31
    2.2.1 Size 31
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2.2</td>
<td>Position</td>
<td>32</td>
</tr>
<tr>
<td>2.2.3</td>
<td>Orientation</td>
<td>33</td>
</tr>
<tr>
<td>2.3</td>
<td>Projections</td>
<td>35</td>
</tr>
<tr>
<td>2.4</td>
<td>Run-Length Encoding</td>
<td>38</td>
</tr>
<tr>
<td>2.5</td>
<td>Binary Algorithms</td>
<td>39</td>
</tr>
<tr>
<td>2.5.1</td>
<td>Definitions</td>
<td>40</td>
</tr>
<tr>
<td>2.5.2</td>
<td>Component Labeling</td>
<td>44</td>
</tr>
<tr>
<td>2.5.3</td>
<td>Size Filter</td>
<td>47</td>
</tr>
<tr>
<td>2.5.4</td>
<td>Euler Number</td>
<td>48</td>
</tr>
<tr>
<td>2.5.5</td>
<td>Region Boundary</td>
<td>50</td>
</tr>
<tr>
<td>2.5.6</td>
<td>Area and Perimeter</td>
<td>50</td>
</tr>
<tr>
<td>2.5.7</td>
<td>Compactness</td>
<td>51</td>
</tr>
<tr>
<td>2.5.8</td>
<td>Distance Measures</td>
<td>52</td>
</tr>
<tr>
<td>2.5.9</td>
<td>Distance Transforms</td>
<td>53</td>
</tr>
<tr>
<td>2.5.10</td>
<td>Medial Axis</td>
<td>55</td>
</tr>
<tr>
<td>2.5.11</td>
<td>Thinning</td>
<td>57</td>
</tr>
<tr>
<td>2.5.12</td>
<td>Expanding and Shrinking</td>
<td>60</td>
</tr>
<tr>
<td>2.6</td>
<td>Morphological Operators</td>
<td>61</td>
</tr>
<tr>
<td>2.7</td>
<td>Optical Character Recognition</td>
<td>70</td>
</tr>
<tr>
<td>3</td>
<td>Regions</td>
<td>73</td>
</tr>
<tr>
<td>3.1</td>
<td>Regions and Edges</td>
<td>73</td>
</tr>
<tr>
<td>3.2</td>
<td>Region Segmentation</td>
<td>76</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Automatic Thresholding</td>
<td>76</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Limitations of Histogram Methods</td>
<td>86</td>
</tr>
<tr>
<td>3.3</td>
<td>Region Representation</td>
<td>86</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Array Representation</td>
<td>88</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Hierarchical Representations</td>
<td>88</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Symbolic Representations</td>
<td>90</td>
</tr>
<tr>
<td>3.3.4</td>
<td>Data Structures for Segmentation</td>
<td>92</td>
</tr>
<tr>
<td>3.4</td>
<td>Split and Merge</td>
<td>96</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Region Merging</td>
<td>97</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Removing Weak Edges</td>
<td>100</td>
</tr>
<tr>
<td>3.4.3</td>
<td>Region Splitting</td>
<td>103</td>
</tr>
<tr>
<td>3.4.4</td>
<td>Split and Merge</td>
<td>104</td>
</tr>
<tr>
<td>3.5</td>
<td>Region Growing</td>
<td>105</td>
</tr>
</tbody>
</table>
4 Image Filtering
4.1 Histogram Modification 112
4.2 Linear Systems 115
4.3 Linear Filters 118
4.4 Median Filter 122
4.5 Gaussian Smoothing 123
  4.5.1 Rotational Symmetry 127
  4.5.2 Fourier Transform Property 128
  4.5.3 Gaussian Separability 129
  4.5.4 Cascading Gaussians 129
  4.5.5 Designing Gaussian Filters 132
  4.5.6 Discrete Gaussian Filters 134

5 Edge Detection 140
5.1 Gradient 143
5.2 Steps in Edge Detection 145
  5.2.1 Roberts Operator 146
  5.2.2 Sobel Operator 147
  5.2.3 Prewitt Operator 148
  5.2.4 Comparison 148
5.3 Second Derivative Operators 149
  5.3.1 Laplacian Operator 149
  5.3.2 Second Directional Derivative 156
5.4 Laplacian of Gaussian 157
5.5 Image Approximation 162
5.6 Gaussian Edge Detection 168
  5.6.1 Canny Edge Detector 169
5.7 Subpixel Location Estimation 173
5.8 Edge Detector Performance 176
  5.8.1 Methods for Evaluating Performance 177
  5.8.2 Figure of Merit 178
5.9 Sequential Methods 179
5.10 Line Detection 180

6 Contours 186
6.1 Geometry of Curves 188
6.2 Digital Curves 188
  6.2.1 Chain Codes 189
6.2.2 Slope Representation 191
6.2.3 Slope Density Function 191
6.3 Curve Fitting 192
6.4 Polyline Representation 194
  6.4.1 Polyline Splitting 196
  6.4.2 Segment Merging 196
  6.4.3 Split and Merge 198
  6.4.4 Hop-Along Algorithm 199
6.5 Circular Arcs 200
6.6 Conic Sections 203
6.7 Spline Curves 207
6.8 Curve Approximation 210
  6.8.1 Total Regression 212
  6.8.2 Estimating Corners 214
  6.8.3 Robust Regression 214
  6.8.4 Hough Transform 218
6.9 Fourier Descriptors 223

7 Texture 234
  7.1 Introduction 234
  7.2 Statistical Methods of Texture Analysis 236
  7.3 Structural Analysis of Ordered Texture 239
  7.4 Model-Based Methods for Texture Analysis 240
  7.5 Shape from Texture 241

8 Optics 249
  8.1 Lens Equation 250
  8.2 Image Resolution 250
  8.3 Depth of Field 251
  8.4 View Volume 253
  8.5 Exposure 254

9 Shading 257
  9.1 Image Irradiance 257
    9.1.1 Illumination 259
    9.1.2 Reflectance 261
  9.2 Surface Orientation 264
## CONTENTS

9.3 The Reflectance Map                                   267
  9.3.1 Diffuse Reflectance                               267
  9.3.2 Scanning Electron Microscopy                     268
9.4 Shape from Shading                                   269
9.5 Photometric Stereo                                   271

### 10 Color

10.1 Color Physics                                       276
10.2 Color Terminology                                   277
10.3 Color Perception                                    278
10.4 Color Processing                                    280
10.5 Color Constancy                                     284
10.6 Discussion                                          286

### 11 Depth

11.1 Stereo Imaging                                      289
  11.1.1 Cameras in Arbitrary Position and Orientation    291
11.2 Stereo Matching                                     293
  11.2.1 Edge Matching                                    294
  11.2.2 Region Correlation                              295
11.3 Shape from X                                        298
11.4 Range Imaging                                       300
  11.4.1 Structured Lighting                              301
  11.4.2 Imaging Radar                                   305
11.5 Active Vision                                       305

### 12 Calibration

12.1 Coordinate Systems                                  311
12.2 Rigid Body Transformations                           313
  12.2.1 Rotation Matrices                               316
  12.2.2 Axis of Rotation                                318
  12.2.3 Unit Quaternions                                318
12.3 Absolute Orientation                                320
12.4 Relative Orientation                                325
12.5 Rectification                                       331
12.6 Depth from Binocular Stereo                         332
12.7 Absolute Orientation with Scale                     334
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.8</td>
<td>Exterior Orientation</td>
<td>336</td>
</tr>
<tr>
<td>12.8.1</td>
<td>Calibration Example</td>
<td>340</td>
</tr>
<tr>
<td>12.9</td>
<td>Interior Orientation</td>
<td>341</td>
</tr>
<tr>
<td>12.10</td>
<td>Camera Calibration</td>
<td>346</td>
</tr>
<tr>
<td>12.10.1</td>
<td>Simple Method for Camera Calibration</td>
<td>347</td>
</tr>
<tr>
<td>12.10.2</td>
<td>Affine Method for Camera Calibration</td>
<td>352</td>
</tr>
<tr>
<td>12.10.3</td>
<td>Nonlinear Method for Camera Calibration</td>
<td>355</td>
</tr>
<tr>
<td>12.11</td>
<td>Binocular Stereo Calibration</td>
<td>357</td>
</tr>
<tr>
<td>12.12</td>
<td>Active Triangulation</td>
<td>359</td>
</tr>
<tr>
<td>12.13</td>
<td>Robust Methods</td>
<td>361</td>
</tr>
<tr>
<td>12.14</td>
<td>Conclusions</td>
<td>361</td>
</tr>
</tbody>
</table>

13 Curves and Surfaces  
13.1 Fields                                            366  
13.2 Geometry of Curves                                  367  
13.3 Geometry of Surfaces                                 369  
13.3.1 Planes                                            369  
13.3.2 Differential Geometry                             370  
13.4 Curve Representations                                373  
13.4.1 Cubic Spline Curves                                373  
13.5 Surface Representations                              374  
13.5.1 Polygonal Meshes                                   374  
13.5.2 Surface Patches                                    378  
13.5.3 Tensor-Product Surfaces                             380  
13.6 Surface Interpolation                                381  
13.6.1 Triangular Mesh Interpolation                       381  
13.6.2 Bilinear Interpolation                             382  
13.6.3 Robust Interpolation                               384  
13.7 Surface Approximation                                 385  
13.7.1 Regression Splines                                 387  
13.7.2 Variational Methods                                395  
13.7.3 Weighted Spline Approximation                       395  
13.8 Surface Segmentation                                 397  
13.8.1 Initial Segmentation                               398  
13.8.2 Extending Surface Patches                          399  
13.9 Surface Registration                                400  

14 Dynamic Vision

14.1 Change Detection  
14.1.1 Difference Pictures  
14.1.2 Static Segmentation and Matching  
14.2 Segmentation Using Motion  
14.2.1 Time-Varying Edge Detection  
14.2.2 Stationary Camera  
14.3 Motion Correspondence  
14.4 Image Flow  
14.4.1 Computing Image Flow  
14.4.2 Feature-Based Methods  
14.4.3 Gradient-Based Methods  
14.4.4 Variational Methods for Image Flow  
14.4.5 Robust Computation of Image Flow  
14.4.6 Information in Image Flow  
14.5 Segmentation Using a Moving Camera  
14.5.1 Ego-Motion Complex Log Mapping  
14.5.2 Depth Determination  
14.6 Tracking  
14.6.1 Deviation Function for Path Coherence  
14.6.2 Path Coherence Function  
14.6.3 Path Coherence in the Presence of Occlusion  
14.6.4 Modified Greedy Exchange Algorithm  
14.7 Shape from Motion  

15 Object Recognition

15.1 System Components  
15.2 Complexity of Object Recognition  
15.3 Object Representation  
15.3.1 Observer-Centered Representations  
15.3.2 Object-Centered Representations  
15.4 Feature Detection  
15.5 Recognition Strategies  
15.5.1 Classification  
15.5.2 Matching  
15.5.3 Feature Indexing
15.6 Verification
  15.6.1 Template Matching 482
  15.6.2 Morphological Approach 483
  15.6.3 Symbolic  483
  15.6.4 Analogical Methods 486

A Mathematical Concepts 492
  A.1 Analytic Geometry 492
  A.2 Linear Algebra 494
  A.3 Variational Calculus 498
  A.4 Numerical Methods 500

B Statistical Methods 502
  B.1 Measurement Errors 502
  B.2 Error Distributions 504
  B.3 Linear Regression 506
  B.4 Nonlinear Regression 510

C Programming Techniques 511
  C.1 Image Descriptors 511
  C.2 Mapping Operators 516
  C.3 Image File Formats 517

Bibliography 519

Index 542
Preface

This book grew out of our efforts to provide a balanced coverage of essential elements of machine vision systems to students in our undergraduate and early graduate classes. The field of machine vision, or computer vision, has been growing at a fast pace. The growth in this field, unlike most established fields, has been both in breadth and depth of concepts and techniques. To make the situation more confusing, the number of new applications has also been growing. Machine vision techniques are being applied in areas ranging from medical imaging to remote sensing, industrial inspection to document processing, and nanotechnology to multimedia databases.

As in most developing fields, not all aspects of machine vision that are of interest to active researchers are useful to the designers of a vision system for a specific application. A designer needs to know basic concepts and techniques to be successful in designing or evaluating a vision system for a particular application. It may not be necessary to know the latest, often controversial, results from leading research centers. On the other hand, the techniques learned by a designer should not be ephemeral.

This text is intended to provide a practical introduction to machine vision. We made efforts to provide all of the details to allow vision algorithms to be used in practical applications. Intentionally omitted are theories of machine vision that do not appear to have sufficient practical applications at this time. We want this to be a useful introduction to machine vision rather than a state-of-the-art collection of research on machine vision.
The text is intended to be used in an introductory course in machine vision at the undergraduate or early graduate level and should be suitable for students with no prior knowledge of computer graphics or signal processing. Students should have a working knowledge of mathematics through calculus of two variables, including matrices and linear spaces, and familiarity with basic probability theory, computer programming, and elementary data structures. Numerical and statistical methods and advanced algorithms are described as needed as well as material on geometry in two and three dimensions. For some sections in the book, more mathematical background is needed. Such sections can be omitted by readers not interested in the rigorous formalization. We have made efforts to provide intuitive concepts, even for mathematical sections, that will help a reader understand the basic elements without the details.

An introductory text is based on material from several sources. This book also contains material from research papers, books, and other places. We have made no attempt to exhaustively list all original sources. We do provide some pointers to readers who are interested in exploring topics more deeply in each chapter. The references at the end of the book provide a list of sources that were directly used in the preparation of the book.

We strongly encourage readers to send any comments and corrections by mail to one of the authors or electronically to jain@ece.ucsd.edu.

Ramesh Jain
Rangachar Kasturi
Brian G. Schunck
Acknowledgments

This book has benefited from the comments and suggestions of many people. Professors Kim Boyer, Glenn Healey, William Thompson, and Mohan Trivedi provided extensive comments and suggestions on the book. We would also like to thank Kevin Bowyer, Rattikorn Hewett, Ting-Chuen Pong, and Mubarak Shah for their reviews. We are also thankful to Anil Jain, Lawrence O’Gorman, Dorothea Blostein, Jason Daida, Robert Bolles, and John Barron for their help in various ways. Eric Munson played a key role in this book by maintaining its schedule; his persistent but very kind and gentle reminders were most important in our finishing this project in a reasonable time. Thanks also to the rest of the staff at McGraw-Hill, New York, and at Publication Services, Inc., of Champaign, Illinois.

Numerous students directly contributed to the preparation of this book. The most important help and contributions came from David Kosiba. David helped us in all aspects of the book, including many of the illustrations and the concept for the cover art. This book has been influenced significantly by his energy, interest in machine vision, and perseverance. Sue Lott and Dino Terzides also helped with the illustrations. Other students who went out of their way to provide help included Sandy Bartlett, James Han, Patrick Kelly, Dan Sebald, Nilesh Patel, Francis Quek, Todd Elvins, Arun Katkere, Saied Moezzi, and Jennifer Schlenzig. We sincerely appreciate the secretarial support provided by Suzie Mostoller at Penn State, Kathy Dewitt at
University of Michigan, and Edna Nerona at UC, San Diego. Brian Schunck would also like to thank his colleagues at Perceptron, Charles Wu at the Ford Scientific Research Laboratories, and Chuck Meyer at the University of Michigan for supporting his research and providing facilities while this book was being written. On a more personal side, the authors are very thankful to their families for their patience and understanding.