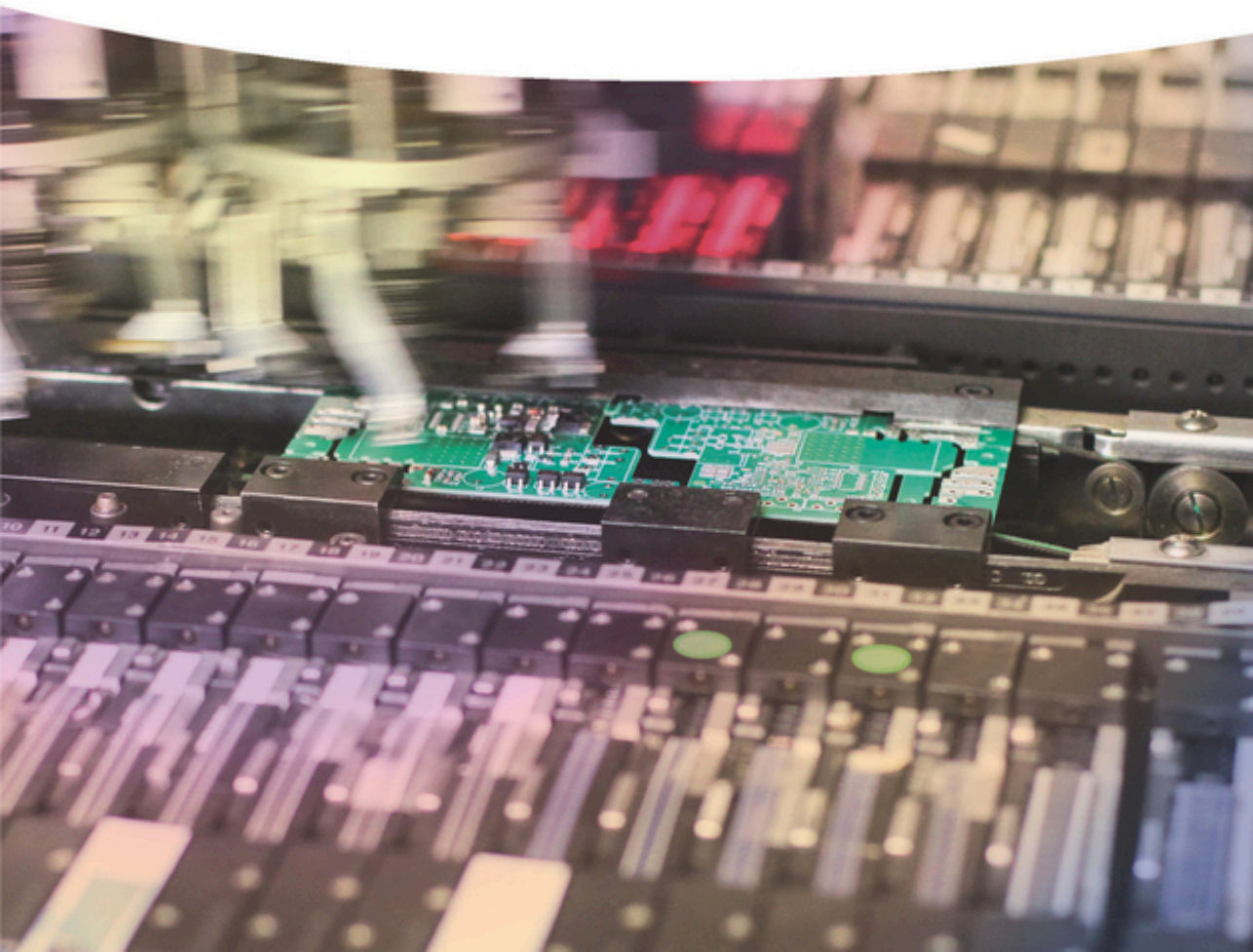
WILEY-VCH

Carsten Steger, Markus Ulrich,
and Christian Wiedemann

Machine Vision Algorithms and Applications

Second Completely Revised and Enlarged Edition



Machine Vision Algorithms and Applications

Carsten Steger, Markus Ulrich, and Christian Wiedemann

Machine Vision Algorithms and Applications

2nd, completely revised and enlarged Edition

WILEY-VCH

Authors

Dr. Carsten Steger
MVTec Software GmbH
Arnulfstraße 205
80634 München
Germany

Dr. Markus Ulrich
MVTec Software GmbH
Arnulfstraße 205
80634 München
Germany

Dr. Christian Wiedemann
MVTec SoftwareGmbH
Arnulfstraße 205
80634 München
Germany

All books published by **Wiley-VCH** are carefully produced. Nevertheless, authors, editors, and publisher do not warrant the information contained in these books, including this book, to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

Library of Congress Card No.: applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <http://dnb.d-nb.de>.

© 2018 Wiley-VCH Verlag GmbH & Co. KGaA, Boschstr. 12, 69469 Weinheim, Germany

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – by photoprinting, microfilm, or any other means – nor transmitted or translated into a machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

Print ISBN: 978-3-527-41365-2
ePDF ISBN: 978-3-527-81290-5
ePub ISBN: 978-3-527-81289-9
Mobi ISBN: 978-3-527-81291-2

Cover Design Adam Design, Weinheim, Germany
Typesetting SPi Global, Chennai, India
Printing and Binding

Printed on acid-free paper

10 9 8 7 6 5 4 3 2 1

Contents

List of Abbreviations xv

Preface to the Second Edition xix

Preface to the First Edition xxi

1 Introduction 1

2 Image Acquisition 5
2.1 Illumination 5
2.1.1 Electromagnetic Radiation 5
2.1.2 Types of Light Sources 7
2.1.3 Interaction of Light and Matter 8
2.1.4 Using the Spectral Composition of the Illumination 10
2.1.5 Using the Directional Properties of the Illumination 12
2.1.5.1 Diffuse Bright-Field Front Light Illumination . . . 13
2.1.5.2 Directed Bright-Field Front Light Illumination . . . 14
2.1.5.3 Directed Dark-Field Front Light Illumination . . . 15
2.1.5.4 Diffuse Bright-Field Back Light Illumination . . . 16
2.1.5.5 Telecentric Bright-Field Back Light Illumination . 17
2.2 Lenses 18
2.2.1 Pinhole Cameras 18
2.2.2 Gaussian Optics 19
2.2.2.1 Refraction 19
2.2.2.2 Thick Lens Model 20
2.2.2.3 Aperture Stop and Pupils 22
2.2.2.4 Relation of the Pinhole Model to Gaussian Optics . 24
2.2.3 Depth of Field 25
2.2.4 Telecentric Lenses 28
2.2.4.1 Object-Side Telecentric Lenses 28
2.2.4.2 Bilateral Telecentric Lenses 30
2.2.4.3 Image-Side Telecentric Lenses 31
2.2.4.4 Projection Characteristics of Lenses 32

2.2.5 Tilt Lenses and the Scheimpflug Principle 32
2.2.6 Lens Aberrations 36
2.2.6.1 Spherical Aberration 36
2.2.6.2 Coma 36
2.2.6.3 Astigmatism 37
2.2.6.4 Curvature of Field 37
2.2.6.5 Distortion 37
2.2.6.6 Chromatic Aberration 39
2.2.6.7 Edge-Spread Function 39
2.2.6.8 Vignetting 40
2.3 Cameras 41
2.3.1 CCD Sensors 41
2.3.1.1 Line Sensors 41
2.3.1.2 Full Frame Array Sensors 43
2.3.1.3 Frame Transfer Sensors 43
2.3.1.4 Interline Transfer Sensors 44
2.3.1.5 Readout Modes 46
2.3.2 CMOS Sensors 46
2.3.2.1 Sensor Architecture 46
2.3.2.2 Rolling and Global Shutters 47
2.3.3 Color Cameras 48
2.3.3.1 Spectral Response of Monochrome Cameras 48
2.3.3.2 Single-Chip Cameras 49
2.3.3.3 Three-Chip Cameras 50
2.3.3.4 Spectral Response of Color Cameras 50
2.3.4 Sensor Sizes 50
2.3.5 Camera Performance 52
2.3.5.1 Noise 52
2.3.5.2 Signal-to-Noise Ratio 53
2.3.5.3 Dynamic Range 54
2.3.5.4 Nonuniformities 54
2.4 Camera–Computer Interfaces 55
2.4.1 Analog Video Signals 56
2.4.1.1 Analog Video Standards 56
2.4.1.2 Analog Frame Grabbers 58
2.4.2 Digital Video Signals 60
2.4.2.1 Camera Link 61
2.4.2.2 Camera Link HS 62
2.4.2.3 CoaXPress 64
2.4.2.4 IEEE 1394 65
2.4.2.5 USB 2.0 67
2.4.2.6 USB3 Vision 68
2.4.2.7 GigE Vision 69
2.4.3 Generic Interfaces 72
2.4.3.1 GenICam 72

2.4.3.2 GenICam GenTL 77
2.4.4 Image Acquisition Modes 79
2.5 3D Image Acquisition Devices 82
2.5.1 Stereo Sensors 82
2.5.2 Sheet of Light Sensors 84
2.5.3 Structured Light Sensors 86
2.5.3.1 Pattern Projection 87
2.5.3.2 Gray Codes 88
2.5.3.3 Fringe Projection 90
2.5.3.4 Hybrid Systems 91
2.5.4 Time-of-Flight Cameras 91
2.5.4.1 Continuous-Wave-Modulated Time-of-Flight Cameras 92
2.5.4.2 Pulse-Modulated Time-of-Flight Cameras 93

3 Machine Vision Algorithms 97
3.1 Fundamental Data Structures 97
3.1.1 Images 97
3.1.2 Regions 98
3.1.3 Subpixel-Precise Contours 101
3.2 Image Enhancement 101
3.2.1 Gray Value Transformations 102
3.2.1.1 Contrast Enhancement 102
3.2.1.2 Contrast Normalization 102
3.2.1.3 Robust Contrast Normalization 103
3.2.2 Radiometric Calibration 105
3.2.2.1 Chart-Based Radiometric Calibration 105
3.2.2.2 Chartless Radiometric Calibration 106
3.2.3 Image Smoothing 110
3.2.3.1 Temporal Averaging 111
3.2.3.2 Mean Filter 112
3.2.3.3 Border Treatment of Filters 113
3.2.3.4 Runtime Complexity of Filters 114
3.2.3.5 Linear Filters 115
3.2.3.6 Frequency Response of the Mean Filter 116
3.2.3.7 Gaussian Filter 117
3.2.3.8 Noise Suppression by Linear Filters 118
3.2.3.9 Median and Rank Filters 119
3.2.4 Fourier Transform 120
3.2.4.1 Continuous Fourier Transform 120
3.2.4.2 Discrete Fourier Transform 123
3.3 Geometric Transformations 126
3.3.1 Affine Transformations 126
3.3.1.1 Projective Transformations 127

3.3.2 Image Transformations 128
3.3.2.1 Nearest-Neighbor Interpolation 128
3.3.2.2 Bilinear Interpolation 129
3.3.2.3 Bicubic Interpolation 130
3.3.2.4 Smoothing to Avoid Aliasing 132
3.3.3 Projective Image Transformations 133
3.3.4 Polar Transformations 133
3.4 Image Segmentation 135
3.4.1 Thresholding 135
3.4.1.1 Global Thresholding 135
3.4.1.2 Automatic Threshold Selection 136
3.4.1.3 Dynamic Thresholding 138
3.4.1.4 Variation Model 141
3.4.2 Extraction of Connected Components 144
3.4.3 Subpixel-Precise Thresholding 147
3.5 Feature Extraction 149
3.5.1 Region Features 149
3.5.1.1 Area 149
3.5.1.2 Moments 150
3.5.1.3 Ellipse Parameters 151
3.5.1.4 Enclosing Rectangles and Circles 153
3.5.1.5 Contour Length 154
3.5.2 Gray Value Features 154
3.5.2.1 Statistical Features 154
3.5.2.2 Moments 155
3.5.2.3 Ellipse Parameters 155
3.5.2.4 Comparison of Region and Gray Value Moments . 155
3.5.3 Contour Features 158
3.5.3.1 Contour Length, Enclosing Rectangles and Circles . 158
3.5.3.2 Moments 158
3.6 Morphology 159
3.6.1 Region Morphology 159
3.6.1.1 Set Operations 159
3.6.1.2 Minkowski Addition and Dilation 161
3.6.1.3 Minkowski Subtraction and Erosion 163
3.6.1.4 Region Boundaries 166
3.6.1.5 Hit-or-Miss Transform 167
3.6.1.6 Opening and Closing 168
3.6.1.7 Skeleton 172
3.6.1.8 Distance Transform 173
3.6.2 Gray Value Morphology 175
3.6.2.1 Minkowski Addition and Dilation 175
3.6.2.2 Minkowski Subtraction and Erosion 176
3.6.2.3 Opening and Closing 177
3.6.2.4 Morphological Gradient 179

3.7 Edge Extraction 180
3.7.1 Definition of Edges 180
3.7.1.1 Definition of Edges in 1D 180
3.7.1.2 Definition of Edges in 2D 181
3.7.2 1D Edge Extraction 183
3.7.2.1 Discrete Derivatives 184
3.7.2.2 Smoothing Perpendicular to a Profile 185
3.7.2.3 Optimal Edge Filters 186
3.7.2.4 Pixel-Accurate Edge Extraction 188
3.7.2.5 Subpixel-Accurate Edge Extraction 189
3.7.3 2D Edge Extraction 189
3.7.3.1 Discrete Derivatives 190
3.7.3.2 Optimal Edge Filters 191
3.7.3.3 Non-Maximum Suppression 192
3.7.3.4 Hysteresis Thresholding 194
3.7.3.5 Subpixel-Accurate Edge Extraction 194
3.7.4 Accuracy and Precision of Edges 196
3.7.4.1 Definition of Accuracy and Precision 197
3.7.4.2 Analytical Edge Accuracy and Precision 198
3.7.4.3 Edge Accuracy and Precision on Real Images . . . 199
3.8 Segmentation and Fitting of Geometric Primitives 203
3.8.1 Fitting Lines 204
3.8.1.1 Least-Squares Line Fitting 204
3.8.1.2 Robust Line Fitting 205
3.8.2 Fitting Circles 208
3.8.2.1 Least-Squares Circle Fitting 208
3.8.2.2 Robust Circle Fitting 209
3.8.3 Fitting Ellipses 210
3.8.3.1 Least-Squares Ellipse Fitting 210
3.8.3.2 Robust Ellipse Fitting 210
3.8.4 Segmentation of Contours 211
3.8.4.1 Segmentation of Contours into Lines 211
3.8.4.2 Segmentation of Contours into Lines, Circles, and Ellipses 213
3.9 Camera Calibration 215
3.9.1 Camera Models for Area Scan Cameras with Regular Lenses 216
3.9.1.1 Exterior Orientation 217
3.9.1.2 Projection From 3D to 2D 219
3.9.1.3 Lens Distortions 219
3.9.1.4 Image Coordinates 221
3.9.2 Camera Models for Area Scan Cameras with Tilt Lenses . . 222
3.9.2.1 Lens Distortions 222
3.9.2.2 Modeling the Pose of the Tilted Image Plane 222
3.9.2.3 Image-Space Telecentric Lenses 223
3.9.2.4 Image-Space Perspective Lenses 224

3.9.3 Camera Model for Line Scan Cameras 225
3.9.3.1 Camera Motion 225
3.9.3.2 Exterior Orientation 226
3.9.3.3 Interior Orientation 227
3.9.3.4 Nonlinearities of the Line Scan Camera Model . . . 229
3.9.4 Calibration Process 230
3.9.4.1 Calibration Target 230
3.9.4.2 Single-Image Camera Calibration 232
3.9.4.3 Degeneracies When Calibrating With a Single Image 233
3.9.4.4 Multi-Image Camera Calibration 234
3.9.4.5 Degeneracies Occurring With Tilt Lenses 234
3.9.4.6 Excluding Parameters From the Optimization . . . 235
3.9.5 World Coordinates from Single Images 235
3.9.5.1 Telecentric Cameras 236
3.9.5.2 Perspective Cameras 236
3.9.5.3 Line-Scan Cameras 238
3.9.5.4 Image Rectification 238
3.9.6 Accuracy of the Camera Parameters 238
3.9.6.1 Influence of the Number of Calibration Images on the Accuracy 239
3.9.6.2 Influence of the Focus Setting on the Camera Parameters 240
3.9.6.3 Influence of the Diaphragm Setting on the Camera Parameters 240
3.10 3D Reconstruction 241
3.10.1 Stereo Reconstruction 241
3.10.1.1 Stereo Geometry 242
3.10.1.2 Stereo Calibration 243
3.10.1.3 Epipolar Geometry 244
3.10.1.4 Image Rectification 247
3.10.1.5 Disparity 249
3.10.1.6 Stereo Matching 250
3.10.1.7 Effect of Window Size 252
3.10.1.8 Robust Stereo Matching 253
3.10.1.9 Spacetime Stereo Matching 254
3.10.2 Sheet of Light Reconstruction 254
3.10.2.1 Extraction of the Laser Line 255
3.10.2.2 Sensor Calibration and 3D Reconstruction 256
3.10.3 Structured Light Reconstruction 257
3.10.3.1 Decoding the Stripes 257
3.10.3.2 Sensor Calibration and 3D Reconstruction 258
3.11 Template Matching 262
3.11.1 Gray-Value-Based Template Matching 263
3.11.1.1 Similarity Measures Based on Gray Value Differences 263

3.11.1.2 Normalized Cross-Correlation 264
3.11.1.3 Efficient Evaluation of the Similarity Measures . . 266
3.11.2 Matching Using Image Pyramids 267
3.11.2.1 Image Pyramids 268
3.11.2.2 Hierarchical Search 270
3.11.3 Subpixel-Accurate Gray-Value-Based Matching 271
3.11.4 Template Matching with Rotations and Scalings 272
3.11.5 Robust Template Matching 273
3.11.5.1 Mean Squared Edge Distance 274
3.11.5.2 Hausdorff Distance 276
3.11.5.3 Generalized Hough Transform 277
3.11.5.4 Geometric Hashing 281
3.11.5.5 Matching Geometric Primitives 283
3.11.5.6 Shape-Based Matching 286
3.12 3D Object Recognition 292
3.12.1 Deformable Matching 293
3.12.1.1 Principle of Deformable Matching 293
3.12.1.2 Model Generation 295
3.12.1.3 Similarity Measure 295
3.12.1.4 Hierarchical Search 297
3.12.1.5 Least-Squares Pose Refinement 297
3.12.1.6 3D Pose Estimation 297
3.12.1.7 Recognition of Locally Deformed Objects 300
3.12.2 Shape-Based 3D Matching 302
3.12.2.1 View-Based Approach 303
3.12.2.2 Restricting the Pose Range 306
3.12.2.3 Hierarchical Model 307
3.12.2.4 2D Model Generation 308
3.12.2.5 Perspective Correction 310
3.12.2.6 Least-Squares Pose Refinement 311
3.12.2.7 Examples 312
3.12.3 Surface-Based 3D Matching 313
3.12.3.1 Global Model Description 314
3.12.3.2 Local Parameters 316
3.12.3.3 Voting 318
3.12.3.4 Least-Squares Pose Refinement 319
3.12.3.5 Extension for Recognizing Deformed Objects . . . 321
3.12.3.6 Extension for Multimodal Data 321
3.13 Hand–Eye Calibration 323
3.13.1 Introduction 323
3.13.2 Problem Definition 325
3.13.3 Dual Quaternions and Screw Theory 327
3.13.3.1 Quaternions 327
3.13.3.2 Screws 329
3.13.3.3 Dual Numbers 330

3.13.3.4 Dual Quaternions 330
3.13.4 Linear Hand–Eye Calibration 331
3.13.5 Nonlinear Hand–Eye Calibration 334
3.13.6 Hand–Eye Calibration of SCARA Robots 335
3.14 Optical Character Recognition 337
3.14.1 Character Segmentation 338
3.14.2 Feature Extraction 339
3.15 Classification 342
3.15.1 Decision Theory 343
3.15.1.1 Bayes Decision Rule 343
3.15.1.2 Classifier Types 345
3.15.1.3 Training, Test, and Validation Sets 345
3.15.1.4 Novelty Detection 345
3.15.2 Classifiers Based on Estimating Class Probabilities 346
3.15.2.1 *k* Nearest-Neighbor Classifiers 347
3.15.2.2 Gaussian Mixture Model Classifiers 347
3.15.3 Classifiers Based on Constructing Separating Hypersurfaces 350
3.15.3.1 Single-Layer Perceptrons 350
3.15.3.2 Multilayer Perceptrons 352
3.15.3.3 Support Vector Machines 358
3.15.3.4 Convolutional Neural Networks 365
3.15.4 Example of Using Classifiers for OCR 369

4 Machine Vision Applications 371
4.1 Wafer Dicing 371
4.1.1 Determining the Width and Height of the Dies 372
4.1.2 Determining the Position of the Dies 374
4.1.3 Exercises 376
4.2 Reading of Serial Numbers 377
4.2.1 Rectifying the Image Using a Polar Transformation 377
4.2.2 Segmenting the Characters 380
4.2.3 Reading the Characters 382
4.2.4 Exercises 382
4.3 Inspection of Saw Blades 383
4.3.1 Extracting the Saw Blade Contour 384
4.3.2 Extracting the Teeth of the Saw Blade 385
4.3.3 Measuring the Angles of the Teeth of the Saw Blade 386
4.3.4 Exercise 388
4.4 Print Inspection 388
4.4.1 Creating the Model of the Correct Print on the Relay 389
4.4.2 Creating the Model to Align the Relays 390
4.4.3 Performing the Print Inspection 391
4.4.4 Exercises 392
4.5 Inspection of Ball Grid Arrays 392
4.5.1 Finding Balls with Shape Defects 393

4.5.2 Constructing a Geometric Model of a Correct BGA 395
4.5.3 Finding Missing and Extraneous Balls 397
4.5.4 Finding Displaced Balls 398
4.5.5 Exercises 400
4.6 Surface Inspection 400
4.6.1 Segmenting the Doorknob 401
4.6.2 Finding the Surface to Inspect 402
4.6.3 Detecting Defects 405
4.6.4 Exercises 407
4.7 Measurement of Spark Plugs 408
4.7.1 Calibrating the Camera 409
4.7.2 Determining the Position of the Spark Plug 410
4.7.3 Performing the Measurement 411
4.7.4 Exercises 413
4.8 Molding Flash Detection 414
4.8.1 Molding Flash Detection Using Region Morphology 414
4.8.2 Molding Flash Detection with Subpixel-Precise Contours . . 418
4.8.3 Exercise 421
4.9 Inspection of Punched Sheets 421
4.9.1 Extracting the Boundaries of the Punched Sheets 422
4.9.2 Performing the Inspection 424
4.9.3 Exercises 425
4.10 3D Plane Reconstruction with Stereo 425
4.10.1 Calibrating the Stereo Setup 426
4.10.2 Performing the 3D Reconstruction and Inspection 428
4.10.3 Exercise 432
4.11 Pose Verification of Resistors 432
4.11.1 Creating Models of the Resistors 433
4.11.2 Verifying the Pose and Type of the Resistors 436
4.11.3 Exercises 438
4.12 Classification of Non-Woven Fabrics 438
4.12.1 Training the Classifier 438
4.12.2 Performing the Texture Classification 440
4.12.3 Exercise 443
4.13 Surface Comparison 443
4.13.1 Creating the Reference Model 443
4.13.2 Reconstructing and Aligning Objects 445
4.13.3 Comparing Objects and Classifying Errors 445
4.13.4 Exercise 450
4.14 3D Pick-and-Place 451
4.14.1 Performing the Hand–Eye Calibration 452
4.14.2 Defining the Grasping Point 455
4.14.3 Picking and Placing Objects 457
4.14.4 Exercises 458

References 461

Index 475

List of Abbreviations

ADC analog-to-digital converter. 42, 47, 52, 53, 58
AOI area of interest. 47, 74, 75
API application programming interface. 62, 72, 74, 77
APS active pixel sensor. 46

BCS base coordinate system. 324, 325, 336, 337
BGA ball grid array. 148, 157, 209, 269, 392, 393, 395–400
BRDF bidirectional reflectance distribution function. 9

CAD computer-aided design. 298, 303, 304, 308–310, 312, 314, 315, 322
CCD charge-coupled device. 11, 41–52, 54
CCIR Comité consultatif international pour la radio. 56, 57
CCS camera coordinate system. 217, 226, 228, 229, 236, 237, 242, 244, 247, 298, 305, 306, 312, 324, 325, 337, 412
CD compact disk. 134, 135, 377, 378, 380, 383
CFA color filter array. 49
CLProtocol Camera Link Protocol. 62, 73, 77
CMOS complementary metal-oxide semiconductor. 11, 41, 46–52, 54, 82
CNN convolutional neural network. 366–369
CPU central processing unit. 15, 58, 369
CWM continuous-wave-modulated. 91–95

DCS distributed control system. 2
DFT discrete Fourier transform. 123, 124
DHCP Dynamic Host Configuration Protocol. 70
DLP digital light processing. 87
DMA direct memory access. 58, 80
DMD digital micromirror device. 87–89, 258, 259
DN digital number. 42, 47, 53, 61
DPS digital pixel sensor. 47
DR dynamic range. 54
DSNU dark signal nonuniformity. 54, 55
DSP digital signal processor. 2

EIA Electronic Industries Alliance. 56, 57
EM expectation maximization. 348
EMVA European Machine Vision Association. 52, 55

FFT fast Fourier transform. 124
FPGA field-programmable gate array. 2, 61

GenApi Generic application programming interface for configuring cameras. 63, 65, 69, 73–75, 77, 79
GenCP Generic Control Protocol. 63, 69, 73, 76, 77
GenICam Generic Interface for Cameras. 62–65, 69, 71, 73–79
GenTL Generic Transport Layer. 65, 73, 77–79
GHT generalized Hough transform. 277, 279, 280, 286
GMM Gaussian mixture model. 348–350, 355, 357, 364
GPIO general-purpose input/output. 63, 64
GPU graphics processing unit. 369
GUI graphical user interface. 74, 75
GVCP GigE Vision Control Protocol. 71, 76, 77
GVSP GigE Vision Streaming Protocol. 71, 72

HTTP Hypertext Transfer Protocol. 70, 73
HVS human visual system. 48, 49, 82

I/O input/output. 2, 3, 63, 64, 71, 74
IC integrated circuit. 136, 137, 139–141, 245, 246, 252, 253, 291, 294, 369, 370
ICP iterative closest point. 319–321
ICS image coordinate system. 221, 227, 228, 236, 237, 242
IDE integrated development environment. 74
IEEE Institute of Electrical and Electronics Engineers. 2, 55, 65, 66, 70
IP Internet Protocol. 70, 72
IPCS image plane coordinate system. 219, 221, 227, 228, 232, 236, 237
IPv4 Internet Protocol, version 4. 70
IPv6 Internet Protocol, version 6. 70
IR infrared. 5, 6, 8, 11, 48, 50, 51, 92, 93, 98
IRLS iteratively reweighted least-squares. 207, 320
ISO International Organization for Standardization. 38, 39, 52

kNN k nearest-neighbor. 347, 449

LCD liquid-crystal display. 87
LCOS liquid crystal on silicon. 87
LED light-emitting diode. 8, 13–16, 86, 401
LLA Link-Local Address. 70
LUT lookup table. 28, 102, 107, 109, 275, 278, 279, 416
LVDS low-voltage differential signaling. 61

MCS model coordinate system. 298, 299, 305, 306, 312
MLP multilayer perceptron. 352, 354–358, 362, 364–370, 382, 438–440, 443

NCC normalized cross-correlation. 251–254, 265–267, 271, 272, 413, 438
NN nearest-neighbor. 347
NTSC National Television System Committee. 56

OCR optical character recognition. 104, 126, 127, 149, 150, 238, 302, 338, 340–343, 345, 348, 355, 365, 367, 369, 370, 377, 380, 382, 383, 438

PAL phase alternating line. 56
PC personal computer. 2, 280
PCB printed circuit board. 1, 10–12, 42, 102, 136, 137, 148, 172–175, 178, 245, 246, 264–266, 269, 270, 278, 392, 432, 434
PFNC pixel format naming convention. 64, 65, 69, 71, 73, 76, 77
PLC programmable logic controller. 2, 3
PLL phase-locked loop. 58, 59
PM pulse-modulated. 91–95
PRNU photoresponse nonuniformity. 54, 55
PTP Precision Time Protocol. 71, 72

RANSAC random sample consensus. 208, 209, 211
ReLU rectified linear unit. 366, 367
ROI region of interest. 99, 100, 102, 103, 126, 128, 136, 137, 139–141, 153, 163, 174–176, 195, 196, 263–265, 271, 272, 276, 288, 291, 292, 298, 338, 373, 375, 389, 395, 400, 401, 405, 408, 418, 422, 423, 435, 436

SAD sum of absolute gray value differences. 250–254, 263–267, 271, 275
SCARA Selective Compliant Arm for Robot Assembly. 335, 336
SED mean squared edge distance. 275, 276
SFNC standard features naming convention. 63, 65, 69, 71, 73–77, 79
SGD stochastic gradient descent. 368, 369
SLR single-lens reflex. 36
SNR signal-to-noise ratio. 52–54, 90–92, 94, 187, 188, 198, 199, 280
SSD sum of squared gray value differences. 250–254, 263, 264, 267, 271, 275
SVD singular value decomposition. 333
SVM support vector machine. 359, 361–364

TCP Transmission Control Protocol. 70
TCS tool coordinate system. 324, 325, 336
TOF time-of-flight. 82, 91–95, 322

U3VCP USB3 Vision Control Protocol. 69
U3VSP USB3 Vision Streaming Protocol. 69
UDP User Datagram Protocol. 70, 71

USB Universal Serial Bus. 2, 5, 55, 65, 67–70, 72
UV ultraviolet. 5, 6, 8, 48

WCS world coordinate system. 217, 226, 231, 232, 236–238, 256, 257, 298, 324, 398, 399, 410
WWW World Wide Web. 70

XML extensible markup language. 63, 65, 69, 71, 73, 74, 79

Preface to the Second Edition

It has been almost exactly ten years since the first edition of this book was published. Many things that we stated in the preface to the first edition of this book have remained constant. Increasing automation has continued to provide the machine vision industry with above-average growth rates. Computers have continued to become more powerful and have opened up new application areas.

On the other hand, many things have changed in the decade since the first edition was published. Efforts to standardize camera–computer interfaces have increased significantly, leading to several new and highly relevant standards. MVTec has participated in the development of many of these standards. Furthermore, sensors that acquire 3D data have become readily available in the machine vision industry. Consequently, 3D machine vision algorithms play an increasingly important role in machine vision applications, especially in the field of robotics. Machine learning (classification) is another technology that has become increasingly important.

The second edition of this book has been extended to reflect these changes. In Chapter 2, we have added a discussion of the latest camera–computer interface and image acquisition standards. Furthermore, we have included a discussion of 3D image acquisition devices. Since many of these sensors use Scheimpflug optics, we have also added a discussion of this important principle. In Chapter 3, we have extended the description of the algorithms that are used in 3D image acquisition devices to perform the 3D reconstruction. Furthermore, we describe camera models and calibration algorithms for cameras that use Scheimpflug optics. The growing importance of 3D processing is reflected by new sections on hand–eye calibration and 3D object recognition. Furthermore, the section on classification has been extended by algorithms that have become increasingly important (in particular, novelty detection and convolutional neural networks). In Chapter 4, we have added two new application examples that show how the 3D algorithms can be used to solve typical 3D applications. Overall, the book has grown by more than 35%.

The applications we present in this book are based on the machine vision software HALCON, developed by MVTec Software GmbH. To make it possible to also publish an electronic version of this book, we have changed the way by which HALCON licenses can be obtained. MVTec now provides the HALCON Student Edition for selected universities and academic research institutes. Please contact your lecturer or local distributor to find out whether you are entitled to par-

ticipate in this program. Note that the student version of HALCON 8.0 is no longer available. To download the applications discussed in Chapter 4, please visit www.machine-vision-book.com.

The first edition of this book has been used extensively in the lectures "Image understanding I: Machine vision algorithms" given by Carsten Steger at the Department of Informatics of the Technical University of Munich, "Industrial Photogrammetry" given by Markus Ulrich at the Department of Civil, Geo, and Environmental Engineering of the Technical University of Munich, and "Industrielle Bildverarbeitung und Machine Vision" given by Markus Ulrich at the Institute of Photogrammetry and Remote Sensing of the Karlsruhe Institute of Technology. We have integrated the feedback we have received from the students into this edition of the book. A substantial part of the new material is based on the lecture "Image understanding II: Robot vision" given by Carsten Steger since 2011 at the Department of Informatics of the Technical University of Munich.

We would like to express our gratitude to several of our colleagues who have helped us in the writing of the second edition of this book. Jean-Marc Nivet provided the images in Figures 3.129–3.131 and proof-read Sections 2.5 and 3.10. Julian Beitzel supported us by preparing the pick and place example described in Section 4.14. We are also grateful to the following colleagues for proof-reading various sections of this book: Thomas Hopfner (Section 2.4), Christoph Zierl (Section 2.4), Andreas Hofhauser (Section 3.12.1), Bertram Drost (Section 3.12.3), Tobias Böttger (Section 3.13), Patrick Follmann (Sections 3.13 and 3.15.3.4), and David Sattlegger (Section 3.15.3.4). Finally, we would like to thank Martin Preuß and Stefanie Volk of Wiley-VCH who were responsible for the production of this edition of the book.

We invite you to send us suggestions on how to improve this book. You can reach us at authors@machine-vision-book.com.

München, July 2017 Carsten Steger, Markus Ulrich, Christian Wiedemann

Preface to the First Edition

The machine vision industry has enjoyed a growth rate well above the industry average for many years. Machine vision systems currently form an integral part of many machines and production lines. Furthermore, machine vision systems are continuously deployed in new application fields, in part because computers get faster all the time and thus enable applications to be solved that were out of reach just a few years ago.

Despite its importance, there are few books that describe in sufficient detail the technology that is important for machine vision. While there are numerous books on image processing and computer vision, very few of them describe the hardware components that are used in machine vision systems to acquire images (illuminations, lenses, cameras, and camera–computer interfaces). Furthermore, these books often only describe the theory, but not its use in real-world applications. Machine vision books, on the other hand, often do not describe the relevant theory in sufficient detail. Therefore, we feel that a book that provides a thorough theoretical foundation of all the machine vision components and machine vision algorithms, and that gives nontrivial practical examples of how they can be used in real applications, is highly overdue.

The applications we present in this book are based on the machine vision software HALCON, developed by MVTec Software GmbH. To enable you to get a hands-on experience with the machine vision algorithms and applications that we discuss, this book contains a registration code that enables you to download, free of charge, a student version of HALCON as well as all the applications we discuss. For details, please visit `www.machine-vision-book.com`.

While the focus of this book is on machine vision applications, we would like to emphasize that the principles we will present can also be used in other application fields, e.g., photogrammetry or medical image processing.

We have tried to make this book accessible to students as well as practitioners (OEMs, system integrators, and end-users) of machine vision. The text requires only a small amount of mathematical background. We assume that the reader has a basic knowledge of linear algebra (in particular, linear transformations between vector spaces expressed in matrix algebra), calculus (in particular, sums and differentiation and integration of one- and two-dimensional functions), Boolean algebra, and set theory.

This book is based on a lecture and lab course entitled "Machine vision algorithms" that Carsten Steger has given annually since 2001 at the Department of Informatics of the Technical University of Munich. Parts of the material have also been used by Markus Ulrich in a lecture entitled "Close-range photogrammetry" given annually since 2005 at the Institute of Photogrammetry and Cartography of the Technical University of Munich. These lectures typically draw an audience from various disciplines, e.g., computer science, photogrammetry, mechanical engineering, mathematics, and physics, which serves to emphasize the interdisciplinary nature of machine vision.

We would like to express our gratitude to several of our colleagues who have helped us in the writing of this book. Wolfgang Eckstein, Juan Pablo de la Cruz Gutiérrez, and Jens Heyder designed or wrote several of the application examples in Chapter 4. Many thanks also go to Gerhard Blahusch, Alexa Zierl, and Christoph Zierl for proofreading the manuscript. Finally, we would like to express our gratitude to Andreas Thoß and Ulrike Werner of Wiley-VCH for having the confidence that we would be able to write this book during the time HALCON 8.0 was completed.

We invite you to send us suggestions on how to improve this book. You can reach us at `authors@machine-vision-book.com`.

München, May 2007 Carsten Steger, Markus Ulrich, Christian Wiedemann

1
Introduction

Machine vision is one of the key technologies in manufacturing because of increasing demands on the documentation of quality and the traceability of products. It is concerned with engineering systems, such as machines or production lines, that can perform quality inspections in order to remove defective products from production or that control machines in other ways, e.g., by guiding a robot during the assembly of a product.

Some of the common tasks that must be solved in machine vision systems are as follows (Fraunhofer Allianz Vision, 2003):

- Object identification is used to discern different kinds of objects, e.g., to control the flow of material or to decide which inspections to perform. This can be based on special identification symbols, e.g., character strings or bar codes, or on specific characteristics of the objects themselves, such as their shape.
- Position detection is used, for example, to control a robot that assembles a product by mounting the components of the product at the correct positions, such as in a pick-and-place machine that places electronic components onto a printed circuit board (PCB). Position detection can be performed in two or three dimensions, depending on the requirements of the application.
- Completeness checking is typically performed after a certain stage of the assembly of a product has been completed, e.g., after the components have been placed onto a PCB, to ensure that the product has been assembled correctly, i.e., that the right components are in the right place.
- Shape and dimensional inspection is used to check the geometric parameters of a product to ensure that they lie within the required tolerances. This can be used during the production process but also after a product has been in use for some time to ensure that the product still meets the requirements despite wear and tear.
- Surface inspection is used to check the surface of a finished product for imperfections such as scratches, indentations, protrusions, etc.

Figure 1.1 displays an example of a typical machine vision system. The object (1) is transported mechanically, e.g., on a conveyor belt. In machine vision applications, we would often like to image the object in a defined position. This requires mechanical handling of the object and often also a trigger that triggers the image acquisition, e.g., a photoelectric sensor (4). The object is illuminated by a suitably chosen or

Machine Vision Algorithms and Applications, Second Edition. Carsten Steger, Markus Ulrich, and Christian Wiedemann.
© 2018 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2018 by Wiley-VCH Verlag GmbH & Co. KGaA.

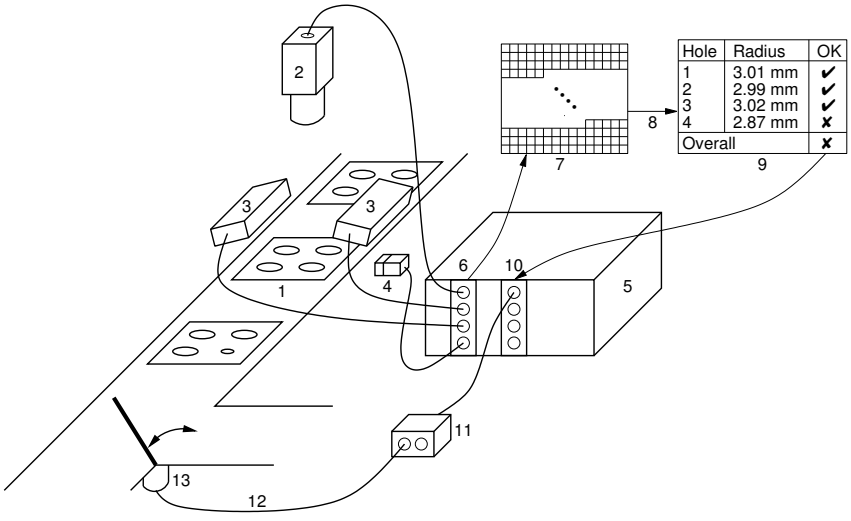


Figure 1.1 The components of a typical machine vision system. An image of the object to be inspected (1) is acquired by a camera (2). The object is illuminated by the illumination (3). A photoelectric sensor (4) triggers the image acquisition. A computer (5) acquires the image through a camera–computer interface (6), in this case a frame grabber. The photoelectric sensor is connected to the frame grabber. The frame grabber triggers the strobe illumination. A device driver assembles the image (7) in the memory of the computer. The machine vision software (8) inspects the objects and returns an evaluation of the objects (9). The result of the evaluation is communicated to a PLC (11) via a digital I/O interface (10). The PLC controls an actuator (13) through a fieldbus interface (12). The actuator, e.g., an electric motor, moves a diverter that is used to remove defective objects from the production line.

specially designed illumination (3). Often, screens (not shown) are used to prevent ambient light from falling onto the object and thereby lowering the image quality. The object is imaged with a camera (2) that uses a lens that has been suitably selected or specially designed for the application. The camera delivers the image to a computer (5) through a camera–computer interface (6), e.g., a frame grabber. The device driver of the camera–computer interface assembles the image (7) in the memory of the computer. If the image is acquired through a frame grabber, the illumination may be controlled by the frame grabber, e.g., through strobe signals. If the camera–computer interface is not a frame grabber but a standard interface, such as IEEE 1394, USB, or Ethernet, the trigger will typically be connected to the camera and illumination directly or through a programmable logic controller (PLC). The computer can be a standard industrial PC or a specially designed computer that is directly built into the camera. The latter configuration is often called a smart camera. The computer may use a standard processor, a digital signal processor (DSP), a field-programmable gate array (FPGA), or a combination of the above. The machine vision software (8) inspects the objects and returns an evaluation of the objects (9). The result of the evaluation is communicated to a controller (11), e.g., a PLC or a distributed control

system (DCS). Often, this communication is performed by digital input/output (I/O) interfaces (10). The PLC, in turn, typically controls an actuator (13) through a communication interface (12), e.g., a fieldbus or serial interface. The actuator, e.g., an electric motor, then moves a diverter that is used to remove defective objects from the production line.

As can be seen from the large number of components involved, machine vision is inherently multidisciplinary. A team that develops a machine vision system will require expertise in mechanical engineering, electrical engineering, optical engineering, and software engineering.

To maintain the focus of this book, we have made a conscious decision to focus on the aspects of a machine vision system that are pertinent to the system until the relevant information has been extracted from the image. Therefore, we will forgo a discussion of the communication components of a machine vision system that are used after the machine vision software has determined its evaluation. For more information on these aspects, please consult Caro (2003); Berge (2004); Mahalik (2003).

In this book, we will try to give you a solid background on everything that is required to extract the relevant information from images in a machine vision system. We include the information that we wish someone had taught us when we started working in the field. In particular, we mention several idiosyncrasies of the hardware components that are highly relevant in applications, which we had to learn the hard way.

The hardware components that are required to obtain high-quality images are described in Chapter 2: illumination, lenses, cameras, and camera–computer interfaces. We hope that, after reading this chapter, you will be able to make informed decisions about which components and setups to use in your application.

Chapter 3 discusses the most important algorithms that are commonly used in machine vision applications. It is our goal to provide you with a solid theoretical foundation that will help you in designing and developing a solution for your particular machine vision task.

To emphasize the engineering aspect of machine vision, Chapter 4 contains a wealth of examples and exercises that show how the machine vision algorithms discussed in Chapter 3 can be combined in non-trivial ways to solve typical machine vision applications.

2
Image Acquisition

In this chapter, we will take a look at the hardware components that are involved in obtaining an image of the scene we want to analyze with the algorithms presented in Chapter 3. Illumination makes the essential features of an object visible. Lenses produce a sharp image on the sensor. The sensor converts the image into a video signal. Finally, camera–computer interfaces (frame grabbers, bus systems like USB, or network interfaces like Ethernet) accept the video signal and convert it into an image in the computer's memory.

2.1
Illumination

The goal of illumination in machine vision is to make the important features of the object visible and to suppress undesired features of the object. To do so, we must consider how the light interacts with the object. One important aspect is the spectral composition of the light and the object. We can use, for example, monochromatic light on colored objects to enhance the contrast of the desired object features. Furthermore, the direction from which we illuminate the object can be used to enhance the visibility of features. We will examine these aspects in this section.

2.1.1
Electromagnetic Radiation

Light is electromagnetic radiation of a certain range of wavelengths, as shown in Table 2.1. The range of wavelengths visible for humans is 380–780 nm. Electromagnetic radiation with shorter wavelengths is called ultraviolet (UV) radiation. Electromagnetic radiation with even shorter wavelengths consists of X-rays and gamma rays. Electromagnetic radiation with longer wavelengths than the visible range is called infrared (IR) radiation. Electromagnetic radiation with even longer wavelengths consists of microwaves and radio waves.

Monochromatic light is characterized by its wavelength λ. If light is composed of a range of wavelengths, it is often compared to the spectrum of light emitted by a

Machine Vision Algorithms and Applications, Second Edition. Carsten Steger, Markus Ulrich, and Christian Wiedemann.
© 2018 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2018 by Wiley-VCH Verlag GmbH & Co. KGaA.

Table 2.1 The electromagnetic spectrum relevant for optics and photonics. The names of the ranges for IR and UV radiation correspond to ISO 20473:2007. The names of the colors for visible radiation (light) are due to Lee (2005).

Range	Name	Abbreviation	Wavelength λ
Ultraviolet	Extreme UV	–	1 nm–100 nm
	Vacuum UV	UV-C	100 nm–190 nm
	Deep UV		190 nm–280 nm
	Mid UV	UV-B	280 nm–315 nm
	Near UV	UV-A	315 nm–380 nm
Visible	Blue-purple		380 nm–430 nm
	Blue		430 nm–480 nm
	Green-blue		480 nm–490 nm
	Blue-green		490 nm–510 nm
	Green		510 nm–530 nm
	Yellow-green		530 nm–570 nm
	Yellow		570 nm–580 nm
	Orange		580 nm–600 nm
	Red		600 nm–720 nm
	Red-purple		720 nm–780 nm
Infrared	Near IR	IR-A	780 nm–1.4 µm
		IR-B	1.4 µm–3 µm
	Mid IR	IR-C	3 µm–50 µm
	Far IR		50 µm–1 mm

black body. A black body is an object that absorbs all electromagnetic radiation that falls onto it and thus serves as an ideal source of purely thermal radiation. Therefore, the light spectrum of a black body is directly related to its temperature. The spectral radiance of a black body is given by Planck's law (Planck, 1901; Wyszecki and Stiles, 1982):

$$I(\lambda, T) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{hc/(\lambda k T)} - 1} \tag{2.1}$$

Here, $c = 2.997\,924\,58 \times 10^8\,\mathrm{m\,s^{-1}}$ is the speed of light, $h = 6.626\,0693 \times 10^{-34}\,\mathrm{J\,s}$ is the Planck constant, and $k = 1.380\,6505 \times 10^{-23}\,\mathrm{J\,K^{-1}}$ is the Boltzmann constant. The spectral radiance is the energy radiated per unit wavelength by an infinitesimal patch of the black body into an infinitesimal solid angle of space. Hence, its unit is $\mathrm{W\,sr^{-1}\,m^{-2}\,nm^{-1}}$.

Figure 2.1 displays the spectral radiance for different temperatures T. It can be seen that black bodies at 300 K radiate primarily in the middle and far IR range. This is the radiation range that is perceived as heat. Therefore, this range of wavelengths is also called thermal IR. The radiation of an object at 1000 K just starts to enter the visible range. This is the red glow that can be seen first when objects are heated. For $T = 3000$ K, the spectrum is that of an incandescent lamp (see Section 2.1.2). Note that it has a strong red component. The spectrum for $T = 6500$ K is used to represent