Contents

Chapter 1 Introduction 1

1.1 Introduction to Computer Vision 2

1.1.1 Visual essentials 2

1.1.2 The goal of computer vision 3

1.1.3 Related disciplines 4

1.2 Computer vision Theory and Framework 7

1.2.1 Visual computational theory 7

1.2.2 Framework issues and improvements 13

1.2.3 A discussion on Marr's reconstruction theory 15

1.2.4 Research on new theoretical framework 18

1.2.5 Discussion from the Perspective of Psychological Cognition 21

1.3 Introduction to Image Engineering 24

1.3.1 Three levels of technology in image engineering 25

1.3.2 Research and application of image engineering 27

1.4 Introduction to Deep Learning 28

1.4.1 Deep learning overview 28

1.4.2 Deep learning core technology 34

1.4.3 Deep learning in computer vision 36

1.5 Organizations and content of this book 38

References 40

Chapter 2 Camera Imaging and Calibration 41

2.1 Lightness Imaging Model 42

2.1.1 Photometric concepts 42

2.1.2 Basic Lightness Imaging Model 43

2.2 Space Imaging Model 44

2.2.1 Projection imaging geometry 44

2.2.2 Basic space imaging model 46

2.2.3 General space imaging model 50

2.2.4 Complete space imaging model 52

2.3 Camera Calibration Model 55

2.3.1 Linear camera model 55

2.3.2 Non-linear camera model 58

2.4 Camera Calibration Methods 62

2.4.1 Classification of calibration methods 62

2.4.2 Traditional calibration methods 64

2.4.3 Self-calibration methods 69

2.4.4 A calibration method for structured light active vision system 72

2.4.5 An on-line camera external parameter calibration method 79

References 82

Chapter 3 Depth Image Acquisition 85

3.1 Depth Image and Depth Imaging 86

3.1.1 Depth images 86

3.1.2 Depth imaging 88
3.2 Direct Depth Imaging 89
3.2.1 Introduction to laser scanning 89
3.2.2 Time of flight 94
3.2.3 LiDAR 97
3.2.4 Structured light 98
3.2.5 Moiré contour striping 101
3.3 Indirect Depth Imaging 105
3.3.1 Binocular horizontal mode 106
3.3.2 Binocular convergent horizontal mode 112
3.3.3 Binocular axial mode 115
3.4 Single-Pixel Depth Imaging 117
3.4.1 The principle of single-pixel imaging 117
3.4.2 Single-pixel camera 119
3.4.3 Single-pixel 3D imaging 121
3.5 Biocular Vision and Stereopsis Vision 123
3.5.1 Biocular vision and binocular vision 123
3.5.2 From monocular to binocular stereo 124
References 125

Chapter 4 3D Point Cloud Data Acquisition and Processing 127
4.1 Point Cloud Data Overview 128
4.1.1 Point cloud data acquisition modes 128
4.1.2 Point cloud data types 129
4.1.3 Point cloud data processing tasks 129
4.1.4 LiDAR test dataset 131
4.2 Point Cloud Data Preprocessing 132
4.2.1 Point cloud data trapping 132
4.2.2 Point cloud data denoising 133
4.2.3 Point cloud data ground area filtering 135
4.2.4 Point cloud data reduction/compression 135
4.2.5 Multi-platform point cloud data registration 138
4.2.6 Registration of Point cloud data and image data 142
4.3 Laser Point Cloud 3D Modeling 143
4.3.1 Delaunay triangulation 143
4.3.2 Patch fitting 144
4.4 Texture Mapping for 3D Models 148
4.4.1 Color texture mapping 149
4.4.2 Geometric texture mapping 150
4.4.3 Procedural texture mapping 151
4.5 Point Cloud Feature Description 152
4.5.1 Global and local feature descriptors 152
4.5.2 Signature of histogram of orientation 153
4.5.3 Rotation projection statistics 153
4.5.4 Tri-orthogonal local depth map 154
4.6 Point Cloud Understanding and Deep Learning 155

4.7 Bionic Optimized Registration of Point Clouds 156
4.7.1 Cuckoo search 157
4.7.2 Improved cuckoo search 158
4.7.3 Point cloud registration application 159
References 161

Chapter 5 Binocular Stereo Vision 165
5.1 Region-Based Binocular Stereo Matching 166
5.1.1 Template matching 166
5.1.2 Stereo matching 170
5.2 Feature-Based Binocular Stereo Matching 177
5.2.1 The basic steps 177
5.2.2 Scale invariant feature transformation 181
5.2.3 Speed up robustness features 184
5.2.4 Dynamic programming matching 191
5.3 Parallax Map Error Detection and Correction 192
5.3.1 Error detection 193
5.3.2 Error correction 194
5.4 Stereo Matching Based on Deep Learning 197
5.4.1 Stereo matching networks 197
5.4.2 Matching based on feature cascade CNN 199
References 200

Chapter 6 Multi-Ocular Stereo Vision 203
6.1 Horizontal Multi-Ocular Stereo Matching 204
6.1.1 Multi-ocular images and SSD 204
6.1.2 Inverse distance and SSSD 206
6.2 Orthogonal Trinocular Stereo Matching 209
6.2.1 Orthogonal trinocular 209
6.2.2 Orthogonal matching based on gradient classification 214
6.3 Multi-Ocular Stereo Matching 219
6.3.1 Matching of arbitrarily arranged trinocular stereo 220
6.3.2 Orthogonal multi-ocular stereo matching 225
6.4 Equal Baseline Multiple Camera Set 227
6.4.1 Image acquisition 227
6.4.2 Image merging methods 228
6.5 Single Camera Multi-Mirror Catadioptric System 231
6.5.1 Overall system structure 231
6.5.2 Imaging and calibration models 232
References 234

Chapter 7 Monocular Multi-Image Scene Restoration 237
7.1 Monocular Image Scene Restoration 238
7.2 Shape from Illumination 240
7.2.1 Scene brightness and image brightness 240
7.2.2 Surface reflection characteristics and brightness 244

7.2.3 Scene surface orientation 247
7.2.4 Reflectance map and image brightness constraint equation 249
7.2.5 Solving the image brightness constraint equation 253
7.3 Shape from Motion 256
7.3.1 Optical flow and motion fields 257
7.3.2 Optical flow field and optical flow equation 260
7.3.3 Solving the optical flow equation 262
7.3.4 Optical flow and surface orientation 269
7.3.5 Optical flow and relative depth 273
7.4 Shape from Segmented Contour 274
7.5 Photometric Stereo Review 276
7.5.1 Light source calibration 276
7.5.2 Non-Lambertian surface reflection models 277
7.5.3 Color photometric stereo 278
7.5.4 3D reconstruction methods 279
7.6 GAN-Based Photometric Stereo 279
7.6.1 Network structure 281
7.6.2 Loss function 282
References 283

Chapter 8 Monocular Single-Image Scene Restoration 287
8.1 Shape from Shading 288
8.1.1 Shades and shapes 288
8.1.2 Solving the image brightness constraint equation 292
8.2 Shape from Texture 300
8.2.1 Monocular imaging and texture distortion 300
8.2.2 Restore surface orientation from texture changes 303
8.2.3 Texture vanishing point detection 311
8.3 Shape from Focus 316
8.4 Estimating Pose from Perspective Three-Point 318
8.4.1 Perspective three-point problem 319
8.4.2 Iterative solution 320
8.5 Shape from Shading in Hybrid Surface Perspective Projection 321
8.5.1 Improved Ward reflection model 322
8.5.2 Image brightness constraint equation under perspective projection 323
8.5.3 Solving the image brightness constraint equation 325
8.5.4 Equations based on the Blinn-Phong reflection model 327
8.5.5 Solving the new image brightness constraint equation 328
References 330

Chapter 9 Generalized Matching 333
9.1 Introduction to Matching 334
9.1.1 Matching strategy 335
9.1.2 Matching classification 336
9.1.3 Matching evaluation 337
9.2 Object Matching 338

9.2.1 Matching metrics 338
9.2.2 Corresponding point matching 341
9.2.3 Inertia equivalent ellipse matching 343
9.3 Dynamic Pattern Matching 345
9.3.1 Matching process 345
9.3.2 Absolute pattern and relative pattern 346
9.4 Matching and Registration 349
9.4.1 Implementation of registration 349
9.4.2 Heterogeneous remote sensing image registration based on feature matching 351
9.4.3 Image matching based on spatial relationship reasoning 352
9.5 Relationship Matching 353
9.6 Graph Isomorphism Matching 358
9.6.1 Introduction to graph theory 358
9.6.2 Graph isomorphism and matching 361
9.7 Labelling and Matching of Line Drawing 364
9.7.1 Contour labelling 365
9.7.2 Structure reasoning 366
9.7.3 Backtracking labelling 367
9.8 Multimodal Image Matching 369
9.8.1 Region-based technology 370
9.8.2 Feature-based technology 372
References 374

Chapter 10 Simultaneous Location and Mapping 377
10.1 SLAM Overview 378
10.1.1 Laser SLAM 378
10.1.2 Visual SLAM 381
10.1.3 Comparison and combination 383
10.2 Laser SLAM Algorithms 385
10.2.1 Gmapping algorithm 385
10.2.2 Cartographer algorithm 388
10.2.3 LOAM algorithm 392
10.3 Visual SLAM Algorithms 394
10.3.1 ORB-SLAM algorithm series 395
10.3.2 LSD-SLAM algorithm 400
10.3.3 SVO algorithm 408
10.4 Swarm Robots and Swarm SLAM 410
10.4.1 Characteristics of swarm robots 411
10.4.2 Problems to be solved by swarm SLAM 412
10.5 Some New Trends of SLAM 413
10.5.1 Combination of SLAM and Deep Learning 413
10.5.2 Combination of SLAM and multi-agent 414
References 416

Chapter 11 Spatial-Temporal Behavior Understanding 419
11.1 Spatial-Temporal Technology 420

11.1.1 New research domains 420
11.1.2 Multiple levels 421
11.2 Action Classification and Recognition 422
11.2.1 Action classification 423
11.2.2 Action recognition 425
11.3 Actor and Action Joint Modeling 428
11.3.1 Single-label actor-action recognition 429
11.3.2 Multiple-label actor-action recognition 430
11.3.3 Actor-action semantic segmentation 432
11.4 Activity and Behavior Modeling 435
11.4.1 Action modeling 436
11.4.2 Activity modeling and recognition 442
11.4.3 Joint point-based behavior recognition 447
11.5 Abnormal Event Detection 450
11.5.1 Automatic activity analysis 450
11.5.2 Classification of abnormal event detection methods 455
11.5.3 Detection based on convolutional autoencoder block learning 457
11.5.4 Detection based on one-class neural network 459
References 460

Index 463