Contents

Preface xiii

Acknowledgments xv

Mnemonics for Proceedings and Special Collections Cited in the References xix

1 COMPUTER VISION 1

1.1 Achieving Simple Vision Goals 1
1.2 High-Level and Low-Level Capabilities 2
1.3 A Range of Representations 6
1.4 The Role of Computers 9
1.5 Computer Vision Research and Applications 12

Part I
GENERALIZED IMAGES
13

2 IMAGE FORMATION 17

2.1 Images 17
2.2 Image Model 18
- 2.2.1 Image Functions, 18
- 2.2.2 Imaging Geometry, 19
- 2.2.3 Reflectance, 22
- 2.2.4 Spatial Properties, 24
- 2.2.5 Color, 31
- 2.2.6 Digital Images, 35

2.3 Imaging Devices for Computer Vision 42
- 2.3.1 Photographic Imaging, 44
- 2.3.2 Sensing Range, 52
- 2.3.3 Reconstruction Imaging, 56

3 EARLY PROCESSING 63

3.1 Recovering Intrinsic Structure 63
3.2 Filtering the Image 65
- 3.2.1 Template Matching, 65
- 3.2.2 Histogram Transformations, 70
- 3.2.3 Background Subtraction, 72
- 3.2.4 Filtering and Reflectance Models, 73

3.3 Finding Local Edges 75
- 3.3.1 Types of Edge Operators, 76
- 3.3.2 Edge Thresholding Strategies, 80
- 3.3.3 Three-Dimensional Edge Operators, 81
- 3.3.4 How Good Are Edge Operators? 83
- 3.3.5 Edge Relaxation, 85

3.4 Range Information from Geometry 88
- 3.4.1 Stereo Vision and Triangulation, 88
- 3.4.2 A Relaxation Algorithm for Stereo, 89

3.5 Surface Orientation from Reflectance Models 93
- 3.5.1 Reflectivity Functions, 93
- 3.5.2 Surface Gradient, 95
- 3.5.3 Photometric Stereo, 98
- 3.5.4 Shape from Shading by Relaxation, 99

3.6 Optical Flow 102
- 3.6.1 The Fundamental Flow Constraint, 102
- 3.6.2 Calculating Optical Flow by Relaxation, 103

3.7 Resolution Pyramids 106
- 3.7.1 Gray-Level Consolidation, 106
- 3.7.2 Pyramidal Structures in Correlation, 107
- 3.7.3 Pyramidal Structures in Edge Detection, 109

PART II
SEGMENTED IMAGES
115

4 BOUNDARY DETECTION 119

4.1 On Associating Edge Elements 119

4.2 Searching Near an Approximate Location 121

4.2.1 Adjusting A Priori Boundaries, 121

4.2.2 Non-Linear Correlation in Edge Space, 121

4.2.3 Divide-and-Conquer Boundary Detection, 122

4.3 The Hough Method for Curve Detection 123

4.3.1 Use of the Gradient, 124

4.3.2 Some Examples, 125

4.3.3 Trading Off Work in Parameter Space for Work in Image Space, 126

4.3.4 Generalizing the Hough Transform, 128

4.4 Edge Following as Graph Searching 131

4.4.1 Good Evaluation Functions, 133

4.4.2 Finding All the Boundaries, 133

4.4.3 Alternatives to the A Algorithm, 136

4.5 Edge Following as Dynamic Programming 137

4.5.1 Dynamic Programming, 137

4.5.2 Dynamic Programming for Images, 139

4.5.3 Lower Resolution Evaluation Functions, 141

4.5.4 Theoretical Questions about Dynamic Programming, 143

4.6 Contour Following 143

4.6.1 Extension to Gray-Level Images, 144

4.6.2 Generalization to Higher-Dimensional Image Data, 146

5 REGION GROWING 149

5.1 Regions 149

5.2 A Local Technique: Blob Coloring 151

5.3 Global Techniques: Region Growing via Thresholding 152

5.3.1 Thresholding in Multidimensional Space, 153

5.3.2 Hierarchical Refinement, 155

5.4 Splitting and Merging 155

5.4.1 State-Space Approach to Region Growing, 157

5.4.2 Low-Level Boundary Data Structures, 158

5.4.3 Graph-Oriented Region Structures, 159

5.5 Incorporation of Semantics 160

6 TEXTURE 166

6.1 What Is Texture? 166

6.2 Texture Primitives 169

6.3 Structural Models of Texel Placement 170
6.3.1 Grammatical Models, 172
6.3.2 Shape Grammars, 173
6.3.3 Tree Grammars, 175
6.3.4 Array Grammars, 178
6.4 Texture as a Pattern Recognition Problem 181
6.4.1 Texture Energy, 184
6.4.2 Spatial Gray-Level Dependence, 186
6.4.3 Region Texels, 188
6.5 The Texture Gradient 189

7 MOTION 195

7.1 Motion Understanding 195
7.1.1 Domain-Independent Understanding, 196
7.1.2 Domain-Dependent Understanding, 196
7.2 Understanding Optical Flow 199
7.2.1 Focus of Expansion, 199
7.2.2 Adjacency, Depth, and Collision, 201
7.2.3 Surface Orientation and Edge Detection, 202
7.2.4 Egomotion, 206
7.3 Understanding Image Sequences 207
7.3.1 Calculating Flow from Discrete Images, 207
7.3.2 Rigid Bodies from Motion, 210
7.3.3 Interpretation of Moving Light Displays—A Domain-Independent Approach, 214
7.3.4 Human Motion Understanding—A Model-Directed Approach, 217
7.3.5 Segmented Images, 220

Part III
GEOMETRICAL STRUCTURES
227

8 REPRESENTATION OF TWO-DIMENSIONAL GEOMETRIC STRUCTURES 231

8.1 Two-Dimensional Geometric Structures 231
8.2 Boundary Representations 232
8.2.1 Polylines, 232
8.2.2 Chain Codes, 235
8.2.3 The Ψ-s Curve, 237
8.2.4 Fourier Descriptors, 238
8.2.5 Conic Sections, 239
8.2.6 B-Splines, 239
8.2.7 Strip Trees, 244

8.3 Region Representations 247
8.3.1 Spatial Occupancy Array, 247
8.3.2 *y* Axis, 248
8.3.3 Quad Trees, 249
8.3.4 Medial Axis Transform, 252
8.3.5 Decomposing Complex Areas, 253
8.4 Simple Shape Properties 254
8.4.1 Area, 254
8.4.2 Eccentricity, 255
8.4.3 Euler Number, 255
8.4.4 Compactness, 256
8.4.5 Slope Density Function, 256
8.4.6 Signatures, 257
8.4.7 Concavity Tree, 258
8.4.8 Shape Numbers, 258

9 REPRESENTATION OF THREE-DIMENSIONAL STRUCTURES 264

9.1 Solids and Their Representation 264
9.2 Surface Representations 265
9.2.1 Surfaces with Faces, 265
9.2.2 Surfaces Based on Splines, 268
9.2.3 Surfaces That Are Functions on the Sphere, 270
9.3 Generalized Cylinder Representations 274
9.3.1 Generalized Cylinder Coordinate Systems and Properties, 275
9.3.2 Extracting Generalized Cylinders, 278
9.3.3 A Discrete Volumetric Version of the Skeleton, 279
9.4 Volumetric Representations 280
9.4.1 Spatial Occupancy, 280
9.4.2 Cell Decomposition, 281
9.4.3 Constructive Solid Geometry, 282
9.4.4 Algorithms for Solid Representations, 284
9.5 Understanding Line Drawings 291
9.5.1 Matching Line Drawings to Three-Dimensional Primitives, 293
9.5.2 Grouping Regions Into Bodies, 294
9.5.3 Labeling Lines, 296
9.5.4 Reasoning About Planes, 301

Part IV
RELATIONAL STRUCTURES
313

10 KNOWLEDGE REPRESENTATION AND USE 317

10.1 Representations 317
10.1.1 The Knowledge Base—Models and Processes, 318

10.1.2 Analogical and Propositional Representations, 319
10.1.3 Procedural Knowledge, 321
10.1.4 Computer Implementations, 322
10.2 Semantic Nets 323
10.2.1 Semantic Net Basics, 323
10.2.2 Semantic Nets for Inference, 327
10.3 Semantic Net Examples 334
10.3.1 Frame Implementations, 334
10.3.2 Location Networks, 335
10.4 Control Issues in Complex Vision Systems 340
10.4.1 Parallel and Serial Computation, 341
10.4.2 Hierarchical and Heterarchical Control, 341
10.4.3 Belief Maintenance and Goal Achievement, 346

11 MATCHING 352

11.1 Aspects of Matching 352
11.1.1 Interpretation: Construction, Matching, and Labeling 352
11.1.2 Matching Iconic, Geometric, and Relational Structures, 353
11.2 Graph-Theoretic Algorithms 355
11.2.1 The Algorithms, 357
11.2.2 Complexity, 359
11.3 Implementing Graph-Theoretic Algorithms 360
11.3.1 Matching Metrics, 360
11.3.2 Backtrack Search, 363
11.3.3 Association Graph Techniques, 365
11.4 Matching in Practice 369
11.4.1 Decision Trees, 370
11.4.2 Decision Tree and Subgraph Isomorphism, 375
11.4.3 Informal Feature Classification, 376
11.4.4 A Complex Matcher, 378

12 INFERENCE 383

12.1 First-Order Predicate Calculus 384
12.1.1 Clause-Form Syntax (Informal), 384
12.1.2 Nonclausal Syntax and Logic Semantics (Informal), 385
12.1.3 Converting Nonclausal Form to Clauses, 387
12.1.4 Theorem Proving, 388
12.1.5 Predicate Calculus and Semantic Networks, 390
12.1.6 Predicate Calculus and Knowledge Representation, 392
12.2 Computer Reasoning 395
12.3 Production Systems 396
12.3.1 Production System Details, 398
12.3.2 Pattern Matching, 399

12.3.3 An Example, 401
12.3.4 Production System Pros and Cons, 406
12.4 Scene Labeling and Constraint Relaxation 408
12.4.1 Consistent and Optimal Labelings, 408
12.4.2 Discrete Labeling Algorithms, 410
12.4.3 A Linear Relaxation Operator and a Line-Labeling Example, 415
12.4.4 A Nonlinear Operator, 419
12.4.5 Relaxation as Linear Programming, 420
12.5 Active Knowledge 430
12.5.1 Hypotheses, 431
12.5.2 HOW-TO and SO-WHAT Processes, 431
12.5.3 Control Primitives, 431
12.5.4 Aspects of Active Knowledge, 433

13 GOAL ACHIEVEMENT 438

13.1 Symbolic Planning 439
13.1.1 Representing the World, 439
13.1.2 Representing Actions, 441
13.1.3 Stacking Blocks, 442
13.1.4 The Frame Problem, 444
13.2 Planning with Costs 445
13.2.1 Planning, Scoring, and Their Interaction, 446
13.2.2 Scoring Simple Plans, 446
13.2.3 Scoring Enhanced Plans, 451
13.2.4 Practical Simplifications, 452
13.2.5 A Vision System Based on Planning, 453

APPENDICES 465

A1 SOME MATHEMATICAL TOOLS 465

A1.1 Coordinate Systems 465
A1.1.1 Cartesian, 465
A1.1.2 Polar and Polar Space, 465
A1.1.3 Spherical and Cylindrical, 466
A1.1.4 Homogeneous Coordinates, 467
A1.2 Trigonometry 468
A1.2.1 Plane Trigonometry, 468
A1.2.2 Spherical Trigonometry, 469
A1.3 Vectors 469
A1.4 Matrices 471
A1.5 Lines 474
A1.5.1 Two Points, 474
A1.5.2 Point and Direction, 474
A1.5.3 Slope and Intercept, 474

A1.5.4 Ratios, 474
A1.5.5 Normal and Distance from Origin (Line Equation), 475
A1.5.6 Parametric, 476
A1.6 Planes 476
A1.7 Geometric Transformations 477
A1.7.1 Rotation, 477
A1.7.2 Scaling, 478
A1.7.3 Skewing, 479
A1.7.4 Translation, 479
A1.7.5 Perspective, 479
A1.7.6 Transforming Lines and Planes, 480
A1.7.7 Summary, 480
A1.8 Camera Calibration and Inverse Perspective 481
A1.8.1 Camera Calibration, 482
A1.8.2 Inverse Perspective, 483
A1.9 Least-Squared-Error Fitting 484
A1.9.1 Pseudo-Inverse Method, 485
A1.9.2 Principal Axis Method, 486
A1.9.3 Fitting Curves by the Pseudo-Inverse Method, 487
A1.10 Conics 488
A1.11 Interpolation 489
A1.11.1 One-Dimensional, 489
A1.11.2 Two-Dimensional, 490
A1.12 The Fast Fourier Transform 490
A1.13 The Icosahedron 492
A1.14 Root Finding 493

A2 ADVANCED CONTROL MECHANISMS 497

A2.1 Standard Control Structures 497
A2.1.1 Recursion, 498
A2.1.2 Co-Routining, 498
A2.2 Inherently Sequential Mechanisms 499
A2.2.1 Automatic Backtracking, 499
A2.2.2 Context Switching, 500
A2.3 Sequential or Parallel Mechanisms 500
A2.3.1 Modules and Messages, 500
A2.3.2 Priority Job Queue, 502
A2.3.3 Pattern-Directed Invocation, 504
A2.3.4 Blackboard Systems, 505

AUTHOR INDEX 509

SUBJECT INDEX 513