Contents

1 Introduction 1
- 1.1 Different Types of Vision Tasks and Systems 1
- 1.2 Why Perception and Action? 3
- 1.3 Why Perception and Not Just Vision? 4
- 1.4 What Are Appropriate Interpretation Spaces? 5
  - 1.4.1 Differential Models for Perception 'Here and Now' 8
  - 1.4.2 Local Integrals as Central Elements for Perception 9
  - 1.4.3 Global Integrals for Situation Assessment 11
- 1.5 What Type of Vision System Is Most Adequate? 11
- 1.6 Influence of the Material Substrate on System Design: Technical vs. Biological Systems 14
- 1.7 What Is Intelligence? A Practical (Ecological) Definition 15
- 1.8 Structuring of Material Covered 18

2 Basic Relations: Image Sequences – "the World" 21
- 2.1 Three-dimensional (3-D) Space and Time 23
  - 2.1.1 Homogeneous Coordinate Transformations in 3-D Space 25
  - 2.1.2 Jacobian Matrices for Concatenations of HCMs 35
  - 2.1.3 Time Representation 39
  - 2.1.4 Multiple Scales 41
- 2.2 Objects 43
  - 2.2.1 Generic 4-D Object Classes 44
  - 2.2.2 Stationary Objects, Buildings 44
  - 2.2.3 Mobile Objects in General 44
  - 2.2.4 Shape and Feature Description 45
  - 2.2.5 Representation of Motion 49
- 2.3 Points of Discontinuity in Time 53
  - 2.3.1 Smooth Evolution of a Trajectory 53
  - 2.3.2 Sudden Changes and Discontinuities 54
- 2.4 Spatiotemporal Embedding and First-order Approximations 54
  - 2.4.1 Gain by Multiple Images in Space and/or Time for Model Fitting 56
  - 2.4.2 Role of Jacobian Matrix in the 4-D Approach to Vision 57

3 Subjects and Subject Classes 59
- 3.1 General Introduction: Perception – Action Cycles 60
- 3.2 A Framework for Capabilities 60
- 3.3 Perceptual Capabilities 63
  - 3.3.1 Sensors for Ground Vehicle Guidance 64
  - 3.3.2 Vision for Ground Vehicles 65
  - 3.3.3 Knowledge Base for Perception Including Vision 72
- 3.4 Behavioral Capabilities for Locomotion 72
  - 3.4.1 The General Model: Control Degrees of Freedom 73
  - 3.4.2 Control Variables for Ground Vehicles 75
  - 3.4.3 Basic Modes of Control Defining Skills 84
  - 3.4.4 Dual Representation Scheme 88
  - 3.4.5 Dynamic Effects in Road Vehicle Guidance 90
  - 3.4.6 Phases of Smooth Evolution and Sudden Changes 104
- 3.5 Situation Assessment and Decision-Making 107
- 3.6 Growth Potential of the Concept, Outlook 107
  - 3.6.1 Simple Model of Human Body as a Traffic Participant 108
  - 3.6.2 Ground Animals and Birds 110

4 Application Domains, Missions, and Situations 111
- 4.1 Structuring of Application Domains 111

4.2 Goals and Their Relations to Capabilities 117
4.3 Situations as Precise Decision Scenarios 118
4.3.1 Environmental Background 118
4.3.2 Objects/Subjects of Relevance 119
4.3.3 Rule Systems for Decision-Making 120
4.4 List of Mission Elements 121

5 Extraction of Visual Features 123
5.1 Visual Features 125
5.1.1 Introduction to Feature Extraction 126
5.1.2 Fields of View, Multifocal Vision, and Scales 128
5.2 Efficient Extraction of Oriented Edge Features 131
5.2.1 Generic Types of Edge Extraction Templates 132
5.2.2 Search Paths and Subpixel Accuracy 137
5.2.3 Edge Candidate Selection 140
5.2.4 Template Scaling as a Function of the Overall Gestalt 141
5.3 The Unified Blob-edge-corner Method (UBM) 144
5.3.1 Segmentation of Stripes Through Corners, Edges, and Blobs 144
5.3.2 Fitting an Intensity Plane in a Mask Region 151
5.3.3 The Corner Detection Algorithm 167
5.3.4 Examples of Road Scenes 171
5.4 Statistics of Photometric Properties of Images 174
5.4.1 Intensity Corrections for Image Pairs 176
5.4.2 Finding Corresponding Features 177
5.4.3 Grouping of Edge Features to Extended Edges 178
5.5 Visual Features Characteristic of General Outdoor Situations 181

6 Recursive State Estimation 183
6.1 Introduction to the 4-D Approach for Spatiotemporal Perception 184
6.2 Basic Assumptions Underlying the 4-D Approach 187
6.3 Structural Survey of the 4-D Approach 190
6.4 Recursive Estimation Techniques for Dynamic Vision 191
6.4.1 Introduction to Recursive Estimation 191
6.4.2 General Procedure 192
6.4.3 The Stabilized Kalman Filter 196
6.4.4 Remarks on Kalman Filtering 196
6.4.5 Kalman Filter with Sequential Innovation 198
6.4.6 Square Root Filters 199
6.4.7 Conclusion of Recursive Estimation for Dynamic Vision 202

7 Beginnings of Spatiotemporal Road and Ego-state Recognition 205
7.1 Road Model 206
7.2 Simple Lateral Motion Model for Road Vehicles 208
7.3 Mapping of Planar Road Boundary into an Image 209
7.3.1 Simple Beginnings in the Early 1980s 209
7.3.2 Overall Early Model for Spatiotemporal Road Perception 213
7.3.3 Some Experimental Results 214
7.3.4 A Look at Vertical Mapping Conditions 217
7.4 Multiple Edge Measurements for Road Recognition 218
7.4.1 Spreading the Discontinuity of the Clothoid Model 219
7.4.2 Window Placing and Edge Mapping 222
7.4.3 Resulting Measurement Model 224
7.4.4 Experimental Results 225

8 Initialization in Dynamic Scene Understanding 227
8.1 Introduction to Visual Integration for Road Recognition 227
8.2 Road Recognition and Hypothesis Generation 228

8.2.1 Starting from Zero Curvature for Near Range 229
8.2.2 Road Curvature from Look-ahead Regions Further Away 230
8.2.3 Simple Numerical Example of Initialization 231
8.3 Selection of Tuning Parameters for Recursive Estimation 233
8.3.1 Elements of the Measurement Covariance Matrix R 234
8.3.2 Elements of the System State Covariance Matrix Q 234
8.3.3 Initial Values of the Error Covariance Matrix P_0 235
8.4 First Recursive Trials and Monitoring of Convergence 236
8.4.1 Jacobian Elements and Hypothesis Checking 237
8.4.2 Monitoring Residues 241
8.5 Road Elements To Be Initialized 241
8.6 Exploiting the Idea of Gestalt 243
8.6.1 The Extended Gestalt Idea for Dynamic Machine Vision 245
8.6.2 Traffic Circle as an Example of Gestalt Perception 251
8.7 Default Procedure for Objects of Unknown Classes 251

9 Recursive Estimation of Road Parameters and Ego State While Cruising 253
9.1 Planar Roads with Minor Perturbations in Pitch 255
9.1.1 Discrete Models 255
9.1.2 Elements of the Jacobian Matrix 256
9.1.3 Data Fusion by Recursive Estimation 257
9.1.4 Experimental Results 258
9.2 Hilly Terrain, 3-D Road Recognition 259
9.2.1 Superposition of Differential Geometry Models 260
9.2.2 Vertical Mapping Geometry 261
9.2.3 The Overall 3-D Perception Model for Roads 262
9.2.4 Experimental Results 263
9.3 Perturbations in Pitch and Changing Lane Widths 268
9.3.1 Mapping of Lane Width and Pitch Angle 268
9.3.2 Ambiguity of Road Width in 3-D Interpretation 270
9.3.3 Dynamics of Pitch Movements: Damped Oscillations 271
9.3.4 Dynamic Model for Changes in Lane Width 273
9.3.5 Measurement Model Including Pitch Angle, Width Changes 275
9.4 Experimental Results 275
9.4.1 Simulations with Ground Truth Available 276
9.4.2 Evaluation of Video Scenes 278
9.5 High-precision Visual Perception 290
9.5.1 Edge Feature Extraction to Subpixel Accuracy for Tracking 290
9.5.2 Handling the Aperture Problem in Edge Perception 292

10 Perception of Crossroads 297
10.1 General Introduction 297
10.1.1 Geometry of Crossings and Types of Vision Systems Required 298
10.1.2 Phases of Crossroad Perception and Turnoff 299
10.1.3 Hardware Bases and Real-world Effects 301
10.2 Theoretical Background 304
10.2.1 Motion Control and Trajectories 304
10.2.2 Gaze Control for Efficient Perception 310
10.2.3 Models for Recursive Estimation 313
10.3 System Integration and Realization 323
10.3.1 System Structure 324
10.3.2 Modes of Operation 325
10.4 Experimental Results 325
10.4.1 Turnoff to the Right 326
10.4.2 Turnoff to the Left 328
10.5 Outlook 329

11 Perception of Obstacles and Other Vehicles 331
11.1 Introduction to Detecting and Tracking Obstacles 331
11.1.1 What Kinds of Objects Are Obstacles for Road Vehicles? 332
11.1.2 At Which Range Do Obstacles Have To Be Detected? 333
11.1.3 How Can Obstacles Be Detected? 334
11.2 Detecting and Tracking Stationary Obstacles 336
11.2.1 Odometry as an Essential Component of Dynamic Vision 336
11.2.2 Attention Focusing on Sets of Features 337
11.2.3 Monocular Range Estimation (Motion Stereo) 338
11.2.4 Experimental Results 342
11.3 Detecting and Tracking Moving Obstacles on Roads 343
11.3.1 Feature Sets for Visual Vehicle Detection 345
11.3.2 Hypothesis Generation and Initialization 352
11.3.3 Recursive Estimation of Open Parameters and Relative State 361
11.3.4 Experimental Results 366
11.3.5 Outlook on Object Recognition 375

12 Sensor Requirements for Road Scenes 377
12.1 Structural Decomposition of the Vision Task 378
12.1.1 Hardware Base 378
12.1.2 Functional Structure 379
12.2 Vision under Conditions of Perturbation 380
12.2.1 Delay Time and High-frequency Perturbation 380
12.2.2 Visual Complexity and the Idea of Gestalt 382
12.3 Visual Range and Resolution Required for Road Traffic Applications 383
12.3.1 Large Simultaneous Field of View 384
12.3.2 Multifocal Design 384
12.3.3 View Fixation 385
12.3.4 Saccadic Control 386
12.3.5 Stereovision 387
12.3.6 Total Range of Fields of View 388
12.3.7 High Dynamic Performance 390
12.4 *MarVEye* as One of Many Possible Solutions 391
12.5 Experimental Result in Saccadic Sign Recognition 392

13 Integrated Knowledge Representations for Dynamic Vision 395
13.1 Generic Object/Subject Classes 399
13.2 The Scene Tree 401
13.3 Total Network of Behavioral Capabilities 403
13.4 Task To Be Performed, Mission Decomposition 405
13.5 Situations and Adequate Behavior Decision 407
13.6 Performance Criteria and Monitoring Actual Behavior 409
13.7 Visualization of Hardware/Software Integration 411

14 Mission Performance, Experimental Results 413
14.1 Situational Aspects for Subtasks 414
14.1.1 Initialization 414
14.1.2 Classes of Capabilities 416
14.2 Applying Decision Rules Based on Behavioral Capabilities 420
14.3 Decision Levels and Competencies, Coordination Challenges 421
14.4 Control Flow in Object-oriented Programming 422
14.5 Hardware Realization of Third-generation EMS vision 426
14.6 Experimental Results of Mission Performance 427
14.6.1 Observing a Maneuver of Another Car 427
14.6.2 Mode Transitions Including Harsh Braking 429
14.6.3 Multisensor Adaptive Cruise Control 431

14.6.4 Lane Changes with Preceding Checks 432
14.6.5 Turning Off on Network of Minor Unsealed Roads 434
14.6.6 On- and Off-road Demonstration with Complex Mission Elements 437

15 Conclusions and Outlook 439

Appendix A
Contributions to Ontology for Ground Vehicles 443
A.1 General environmental conditions 443
A.2 Roadways 443
A.3 Vehicles 444
A.4 Form, Appearance, and Function of Vehicles 444
A.5 Form, Appearance, and Function of Humans 446
A.6 Form, Appearance, and Likely Behavior of Animals 446
A.7 General Terms for Acting "Subjects" in Traffic 446

Appendix B
Lateral dynamics 449
B.1 Transition Matrix for Fourth-order Lateral Dynamics 449
B.2 Transfer Functions and Time Responses to an Idealized Doublet in Fifth-order Lateral Dynamics 450

Appendix C
Recursive Least–squares Line Fit 453
C.1 Basic Approach 453
C.2 Extension of Segment by One Data Point 456
C.3 Stripe Segmentation with Linear Homogeneity Model 457
C.4 Dropping Initial Data Point 458

References 461

Index 473