

# Table of Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Non-Intrusive Acquisition of Human Action</b>	<b>7</b>
2.1	Hand Gesture Commands	7
2.1.1	Image Acquisition and Input Data	8
2.1.1.1	Vocabulary	8
2.1.1.2	Recording Conditions	9
2.1.1.3	Image Representation	10
2.1.1.4	Example	11
2.1.2	Feature Extraction	12
2.1.2.1	Hand Localization	14
2.1.2.2	Region Description	20
2.1.2.3	Geometric Features	22
2.1.2.4	Example	27
2.1.3	Feature Classification	29
2.1.3.1	Classification Concepts	29
2.1.3.2	Classification Algorithms	31
2.1.3.3	Feature Selection	32
2.1.3.4	Rule-based Classification	33
2.1.3.5	Maximum Likelihood Classification	34
2.1.3.6	Classification Using Hidden Markov Models	37
2.1.3.7	Example	47
2.1.4	Static Gesture Recognition Application	48
2.1.5	Dynamic Gesture Recognition Application	50
2.1.6	Troubleshooting	56
2.2	Facial Expression Commands	56
2.2.1	Image Acquisition	59
2.2.2	Image Preprocessing	60
2.2.2.1	Face Localization	61
2.2.2.2	Face Tracking	67
2.2.3	Feature Extraction with Active Appearance Models	68
2.2.3.1	Appearance Model	73
2.2.3.2	AAM Search	75
2.2.4	Feature Classification	78
2.2.4.1	Head Pose Estimation	78
2.2.4.2	Determination of Line of Sight	83
2.2.4.3	Circle Hough Transformation	83
2.2.4.4	Determination of Lip Outline	87
2.2.4.5	Lip modeling	88
2.2.5	Facial Feature Recognition – Eye Localization Application	90
2.2.6	Facial Feature Recognition – Mouth Localization Application	90

References .....	92
<b>3 Sign Language Recognition .....</b>	<b>95</b>
3.1 Recognition of Isolated Signs in Real-World Scenarios .....	99
3.1.1 Image Acquisition and Input Data .....	101
3.1.2 Image Preprocessing .....	102
3.1.2.1 Background Modeling .....	102
3.1.3 Feature Extraction .....	105
3.1.3.1 Overlap Resolution .....	106
3.1.3.2 Hand Tracking .....	107
3.1.4 Feature Normalization .....	109
3.1.5 Feature Classification .....	110
3.1.6 Test and Training Samples .....	110
3.2 Sign Recognition Using Nonmanual Features .....	111
3.2.1 Nonmanual Parameters .....	111
3.2.2 Extraction of Nonmanual Features .....	113
3.3 Recognition of continuous Sign Language using Subunits .....	116
3.3.1 Subunit Models for Signs .....	116
3.3.2 Transcription of Sign Language .....	118
3.3.2.1 Linguistics-orientated Transcription of Sign Language ..	118
3.3.2.2 Visually-orientated Transcription of Sign Language ...	121
3.3.3 Sequential and Parallel Breakdown of Signs .....	122
3.3.4 Modification of HMMs to Parallel Hidden Markov Models ....	122
3.3.4.1 Modeling Sign Language by means of PaHMMs .....	124
3.3.5 Classification .....	125
3.3.5.1 Classification of Single Signs by Means of Subunits and PaHMMs .....	125
3.3.5.2 Classification of Continuous Sign Language by Means of Subunits and PaHMMs .....	126
3.3.5.3 Stochastic Language Modeling .....	127
3.3.6 Training .....	128
3.3.6.1 Initial Transcription .....	129
3.3.6.2 Estimation of Model Parameters for Subunits .....	131
3.3.6.3 Classification of Single Signs .....	132
3.3.7 Concatenation of Subunit Models to Word Models for Signs. .	133
3.3.8 Enlargement of Vocabulary Size by New Signs .....	133
3.4 Performance Evaluation .....	134
3.4.1 Video-based Isolated Sign Recognition .....	135
3.4.2 Subunit Based Recognition of Signs and Continuous Sign Language .....	136
3.4.3 Discussion .....	137
References .....	137

<b>4</b>	<b>Speech Communication and Multimodal Interfaces</b>	141
4.1	Speech Recognition	141
4.1.1	Fundamentals of Hidden Markov Model-based Speech Recognition	142
4.1.2	Training of Speech Recognition Systems	144
4.1.3	Recognition Phase for HMM-based ASR Systems	145
4.1.4	Information Theory Interpretation of Automatic Speech Recognition	147
4.1.5	Summary of the Automatic Speech Recognition Procedure	149
4.1.6	Speech Recognition Technology	150
4.1.7	Applications of ASR Systems	151
4.2	Speech Dialogs	153
4.2.1	Introduction	153
4.2.2	Initiative Strategies	155
4.2.3	Models of Dialog	156
4.2.3.1	Finite State Model	156
4.2.3.2	Slot Filling	157
4.2.3.3	Stochastic Model	158
4.2.3.4	Goal Directed Processing	159
4.2.3.5	Rational Conversational Agents	160
4.2.4	Dialog Design	162
4.2.5	Scripting and Tagging	163
4.3	Multimodal Interaction	164
4.3.1	In- and Output Channels	165
4.3.2	Basics of Multimodal Interaction	166
4.3.2.1	Advantages	167
4.3.2.2	Taxonomy	167
4.3.3	Multimodal Fusion	168
4.3.3.1	Integration Methods	169
4.3.4	Errors in Multimodal Systems	172
4.3.4.1	Error Classification	172
4.3.4.2	User Specific Errors	172
4.3.4.3	System Specific Errors	173
4.3.4.4	Error Avoidance	173
4.3.4.5	Error Resolution	174
4.4	Emotions from Speech and Facial Expressions	175
4.4.1	Background	175
4.4.1.1	Application Scenarios	175
4.4.1.2	Modalities	176
4.4.1.3	Emotion Model	176
4.4.1.4	Emotional Databases	177
4.4.2	Acoustic Information	177
4.4.2.1	Feature Extraction	178
4.4.2.2	Feature Selection	179
4.4.2.3	Classification Methods	180

4.4.3	Linguistic Information	180
4.4.3.1	N-Grams	181
4.4.3.2	Bag-of-Words	182
4.4.3.3	Phrase Spotting	182
4.4.4	Visual Information	183
4.4.4.1	Prerequisites	184
4.4.4.2	Holistic Approaches	184
4.4.4.3	Analytic Approaches	186
4.4.5	Information Fusion	186
4.4.6	Discussion	187
	References	187
<b>5</b>	<b>Person Recognition and Tracking</b>	<b>191</b>
5.1	Face Recognition	193
5.1.1	Challenges in Automatic Face Recognition	193
5.1.2	Structure of Face Recognition Systems	196
5.1.3	Categorization of Face Recognition Algorithms	200
5.1.4	Global Face Recognition using Eigenfaces	201
5.1.5	Local Face Recognition based on Face Components or Template Matching	206
5.1.6	Face Databases for Development and Evaluation	212
5.1.7	Exercises	214
5.1.7.1	Provided Images and Image Sequences	214
5.1.7.2	Face Recognition Using the Eigenface Approach	215
5.1.7.3	Face Recognition Using Face Components	221
5.2	Full-body Person Recognition	224
5.2.1	State of the Art in Full-body Person Recognition	224
5.2.2	Color Features for Person Recognition	226
5.2.2.1	Color Histograms	226
5.2.2.2	Color Structure Descriptor	227
5.2.3	Texture Features for Person Recognition	228
5.2.3.1	Oriented Gaussian Derivatives	228
5.2.3.2	Homogeneous Texture Descriptor	229
5.2.4	Experimental Results	231
5.2.4.1	Experimental Setup	231
5.2.4.2	Feature Performance	231
5.3	Camera-based People Tracking	232
5.3.1	Segmentation	234
5.3.1.1	Foreground Segmentation by Background Subtraction	234
5.3.1.2	Morphological Operations	236
5.3.2	Tracking	237
5.3.2.1	Tracking Following Detection	238
5.3.2.2	Combined Tracking and Detection	243
5.3.3	Occlusion Handling	248
5.3.3.1	Occlusion Handling without Separation	248

5.3.3.2 Separation by Shape .....	248
5.3.3.3 Separation by Color .....	249
5.3.3.4 Separation by 3D Information .....	250
5.3.4 Localization and Tracking on the Ground Plane .....	250
5.3.5 Example .....	253
References .....	259
<b>6 Interacting in Virtual Reality .....</b>	<b>263</b>
6.1 Visual Interfaces .....	265
6.1.1 Spatial Vision .....	265
6.1.2 Immersive Visual Displays .....	266
6.1.3 Viewer Centered Projection .....	268
6.2 Acoustic Interfaces .....	272
6.2.1 Spatial Hearing .....	272
6.2.2 Auditory Displays .....	273
6.2.3 Wave Field Synthesis .....	274
6.2.4 Binaural Synthesis .....	275
6.2.5 Cross Talk Cancellation .....	276
6.3 Haptic Interfaces .....	279
6.3.1 Rendering a Wall .....	281
6.3.2 Rendering Solid Objects .....	283
6.4 Modeling the Behavior of Virtual Objects .....	286
6.4.1 Particle Dynamics .....	286
6.4.2 Rigid Body Dynamics .....	288
6.5 Interacting with Rigid Objects .....	292
6.5.1 Guiding Sleeves .....	294
6.5.2 Sensitive Polygons .....	296
6.5.3 Virtual Magnetism .....	296
6.5.4 Snap-In .....	297
6.6 Implementing Virtual Environments .....	298
6.6.1 Scene Graphs .....	298
6.6.2 Toolkits .....	301
6.6.3 Cluster Rendering in Virtual Environments .....	303
6.7 Examples .....	305
6.7.1 A Simple Scene Graph Example .....	307
6.7.2 More Complex Examples: SolarSystem and Robot .....	308
6.7.3 An Interactive Environment, Tetris3D .....	310
6.8 Summary .....	310
6.9 Acknowledgements .....	312
References .....	312

<b>7 Interactive and Cooperative Robot Assistants</b> .....	315
7.1 Mobile and Humanoid Robots .....	316
7.2 Interaction with Robot Assistants .....	321
7.2.1 Classification of Human-Robot Interaction Activities .....	323
7.2.2 Performative Actions .....	323
7.2.3 Commanding and Commenting Actions .....	325
7.3 Learning and Teaching Robot Assistants .....	325
7.3.1 Classification of Learning through Demonstration .....	326
7.3.2 Skill Learning .....	326
7.3.3 Task Learning .....	327
7.3.4 Interactive Learning .....	328
7.3.5 Programming Robots via Observation: An Overview .....	329
7.4 Interactive Task Learning from Human Demonstration .....	330
7.4.1 Classification of Task Learning .....	331
7.4.2 The PbD Process .....	333
7.4.3 System Structure .....	335
7.4.4 Sensors for Learning through Demonstration .....	336
7.5 Models and Representation of Manipulation Tasks .....	338
7.5.1 Definitions .....	338
7.5.2 Task Classes of Manipulations .....	338
7.5.3 Hierarchical Representation of Tasks .....	340
7.6 Subgoal Extraction from Human Demonstration .....	342
7.6.1 Signal Processing .....	343
7.6.2 Segmentation Step .....	343
7.6.3 Grasp Classification .....	344
7.6.3.1 Static Grasp Classification .....	344
7.6.3.2 Dynamic Grasp Classification .....	346
7.7 Task Mapping and Execution .....	348
7.7.1 Event-Driven Architecture .....	348
7.7.2 Mapping Tasks onto Robot Manipulation Systems .....	350
7.7.3 Human Comments and Advice .....	354
7.8 Examples of Interactive Programming .....	354
7.9 Telepresence and Telerobotics .....	356
7.9.1 The Concept of Telepresence and Telerobotic Systems .....	356
7.9.2 Components of a Telerobotic System .....	359
7.9.3 Telemanipulation for Robot Assistants in Human-Centered Environments .....	360
7.9.4 Exemplary Teleoperation Applications .....	361
7.9.4.1 Using Teleoperated Robots as Remote Sensor Systems	361
7.9.4.2 Controlling and Instructing Robot Assistants via Teleoperation .....	362
References .....	364

<b>8 Assisted Man-Machine Interaction</b> .....	369
8.1 The Concept of User Assistance .....	370
8.2 Assistance in Manual Control .....	372
8.2.1 Needs for Assistance in Manual Control .....	373
8.2.2 Context of Use in Manual Control .....	374
8.2.3 Support of Manual Control Tasks .....	379
8.3 Assistance in Man-Machine Dialogs .....	386
8.3.1 Needs for Assistance in Dialogs .....	386
8.3.2 Context of Dialogs .....	387
8.3.3 Support of Dialog Tasks .....	390
8.4 Summary .....	394
References .....	395
<b>A LTI-LIB — A C++ Open Source Computer Vision Library</b> .....	399
A.1 Installation and Requirements .....	400
A.2 Overview .....	402
A.2.1 Duplication and Replication .....	402
A.2.2 Serialization .....	403
A.2.3 Encapsulation .....	403
A.2.4 Examples .....	403
A.3 Data Structures .....	404
A.3.1 Points and Pixels .....	405
A.3.2 Vectors and Matrices .....	406
A.3.3 Image Types .....	408
A.4 Functional Objects .....	409
A.4.1 Functor .....	410
A.4.2 Classifier .....	411
A.5 Handling Images .....	412
A.5.1 Convolution .....	412
A.5.2 Image IO .....	413
A.5.3 Color Spaces .....	414
A.5.4 Drawing and Viewing .....	416
A.6 Where to Go Next .....	418
References .....	421
<b>B IMPRESARIO — A GUI for Rapid Prototyping of Image Processing Systems</b> .....	423
B.1 Requirements, Installation, and Deinstallation .....	424
B.2 Basic Components and Operations .....	424
B.2.1 Document and Processing Graph .....	426
B.2.2 Macros .....	426
B.2.3 Processing Mode .....	429
B.2.4 Viewer .....	430
B.2.5 Summary and further Reading .....	431
B.3 Adding Functionality to a Document .....	432

B.3.1 Using the Macro Browser . . . . .	432
B.3.2 Rearranging and Deleting Macros . . . . .	434
B.4 Defining the Data Flow with Links . . . . .	434
B.4.1 Creating a Link . . . . .	434
B.4.2 Removing a Link . . . . .	436
B.5 Configuring Macros . . . . .	436
B.5.1 <i>Image Sequence</i> . . . . .	437
B.5.2 <i>Video Capture Device</i> . . . . .	440
B.5.3 <i>Video Stream</i> . . . . .	442
B.5.4 Standard Macro Property Window . . . . .	444
B.6 Advanced Features and Internals of IMPRESARIO . . . . .	446
B.6.1 Customizing the Graphical User Interface . . . . .	446
B.6.2 Macros and Dynamic Link Libraries . . . . .	448
B.6.3 Settings Dialog and Directories . . . . .	449
B.7 IMPRESARIO Examples in this Book . . . . .	450
B.8 Extending IMPRESARIO with New Macros . . . . .	450
B.8.1 Requirements . . . . .	450
B.8.2 API Installation and further Reading . . . . .	451
B.8.3 IMPRESARIO Updates . . . . .	451
References . . . . .	451
<b>Index</b> . . . . .	<b>453</b>





<http://www.springer.com/978-3-540-30618-4>

Advanced Man-Machine Interaction  
Fundamentals and Implementation

Kraiss, K.-F. (Ed.)

2006, XIX, 461 p. With CD-ROM., Hardcover

ISBN: 978-3-540-30618-4