

Contents

1	Introduction	1
1.1	Motivation and Scope	1
1.2	Contribution	2
1.3	Outline	4
2	State of the Art in Object Recognition and Pose Estimation	7
2.1	Appearance-Based Methods	7
2.1.1	Global Approaches	8
2.1.2	Local Approaches	16
2.2	Model-Based Methods	27
2.2.1	Edge-Based Object Tracking	27
2.2.2	Edge-Based Object Recognition	35
2.2.3	Pose Estimation Based on Matched Feature Points	41
2.2.4	Object Recognition Based on 3D Point Clouds	43
2.2.5	Hybrid Approaches	45
2.3	Comparison	45
3	State of the Art in Human Motion Capture	49
3.1	VICON	50
3.2	Systems Using a Search Method	51
3.3	Systems Using a Minimization Method	52
3.3.1	Minimization Method for Articulated Objects	52
3.3.2	Systems Using a 3D-3D Minimization Method	55
3.3.3	Systems Using a 2D-3D Minimization Method	58
3.4	Systems Based on Particle Filtering	59
3.5	Pose Estimation Based on Silhouettes	62
3.6	Comparison	64

4	Fundamentals of Image Processing	67
4.1	Camera Model	67
4.1.1	Coordinate Systems	68
4.1.2	Intrinsic Camera Parameters of the Linear Mapping	69
4.1.3	Extrinsic Camera Parameters	70
4.1.4	Distortion Parameters	70
4.1.5	Overview	73
4.2	Segmentation	74
4.2.1	Thresholding	74
4.2.2	Background Subtraction	75
4.2.3	Color Segmentation	76
4.3	Correlation Methods	77
4.3.1	General Definition	78
4.3.2	Non-normalized Correlation Functions	79
4.3.3	Normalized Correlation Functions	79
4.4	Homography	81
4.4.1	General Definition	82
4.4.2	Least Squares Computation of Homography Parameters	82
4.5	Principal Component Analysis	84
4.5.1	Mathematical Definition	84
4.5.2	Eigenspace	85
4.5.3	Application	85
4.6	Particle Filtering	86
4.7	RANSAC	89
5	Guiding Principles	91
5.1	Ways of Using Calibrated Stereo Camera Systems	91
5.2	Eye Movements	92
5.3	Rectification of Stereo Image Pairs	93
5.4	Undistortion of Images	94
5.5	Grasp Execution	94
5.6	Imitation Learning	96
6	Stereo-Based Object Recognition and Pose Estimation System	99
6.1	Recognition and Pose Estimation Based on the Shape	99
6.1.1	Problem Definition	101
6.1.2	Basic Approach	102
6.1.3	Orientation Correction	107
6.1.4	Position Correction	108
6.1.5	Increasing Robustness	112
6.1.6	Summary of the Algorithm	115
6.2	Recognition and Pose Estimation Based on Texture	116
6.2.1	Accuracy Considerations	117

6.2.2	Feature Calculation	125
6.2.3	Recognition and 2D Localization	129
6.2.4	6-DoF Pose Estimation	136
6.2.5	Occlusions	141
6.2.6	Increasing Robustness	142
6.2.7	Runtime Considerations	144
6.2.8	Summary of the Algorithm	145
7	Stereo-Based Markerless Human Motion Capture System	147
7.1	Problem Definition	148
7.2	Human Upper Body Model	148
7.2.1	Kinematic Model	148
7.2.2	Geometric Model	148
7.3	General Particle Filtering Framework for Human Motion Capture	150
7.3.1	Edge Cue	151
7.3.2	Region Cue	153
7.3.3	Fusion of Multiple Cues	153
7.4	Cues in the Proposed System	154
7.4.1	Edge Cue	154
7.4.2	Distance Cue	154
7.4.3	Cue Comparison	155
7.4.4	Using a Calibrated Stereo System	161
7.5	Image Processing Pipeline	162
7.6	Hand/Head Tracking	164
7.7	Hierarchical Search	169
7.8	Fusing the Edge Cue and the Distance Cue	170
7.9	Adaptive Noise	173
7.10	Adaptive Shoulder Position	175
7.11	Incorporating Inverse Kinematics	178
7.12	Summary of the Algorithm	182
8	Software and Interfaces	185
8.1	Integrating Vision Toolkit	185
8.1.1	Implementation	185
8.1.2	The Class CByteImage	186
8.1.3	Implementation of Graphical User Interfaces	186
8.1.4	Connection of Image Sources	187
8.1.5	Integration of OpenCV	188
8.1.6	Integration of OpenGL	189
8.1.7	Camera Calibration and Camera Model	189
8.2	Master Motor Map	190
8.2.1	Specification	191
8.2.2	Framework and Converter Modules	193
8.2.3	Conversion to the Master Motor Map	195

8.3	Interfaces	196
8.3.1	Object Recognition and Pose Estimation	196
8.3.2	Human Motion Capture	197
9	Evaluation	199
9.1	Recognition and Pose Estimation System Based on the Shape	200
9.1.1	Accuracy	200
9.1.2	Real-World Experiments	207
9.1.3	Runtime	209
9.2	Recognition and Pose Estimation System Based on Texture	211
9.2.1	Accuracy	211
9.2.2	Real-World Experiments	221
9.2.3	Runtime	221
9.3	Markerless Human Motion Capture System	224
9.3.1	Automatic Initialization	224
9.3.2	Real-World Experiments and Accuracy	225
9.3.3	Runtime	228
10	Conclusion	231
10.1	Contribution	231
10.2	Example Applications	233
10.3	Discussion and Outlook	234
A	Mathematics	237
A.1	Singular Value Decomposition	237
A.2	Pseudoinverse	237
A.2.1	Using the Regular Inverse	237
A.2.2	Using the Singular Value Decomposition	238
A.3	Linear Least Squares	238
A.3.1	Using the Normal Equation	238
A.3.2	Using the QR Decomposition	239
A.3.3	Using the Singular Value Decomposition	239
A.3.4	Homogeneous Systems	239
A.4	Functions for Rotations	239
B	File Formats	243
B.1	Camera Parameters	243
	List of Figures	245
	List of Tables	251
	List of Algorithm	253
	References	255
	Index	267