

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Contributions	2
1.3	Structure of the Thesis	3
2	State of the Art	7
2.1	Image-based Tracking	7
2.2	Tracking-by-Detection	10
2.3	Stereo-based Tracking	14
3	Preliminaries	17
3.1	Object Detection	17
3.2	Stereo Estimation	20
3.3	Visual Odometry	23
3.4	Multi-Hypothesis Tracking	26
3.5	Real-Time Tracking-by-Detection	31
3.6	Camera Setup	36
3.7	Discussion	37
I	Hybrid High-Level/Low-Level Tracking	39
4	Hybrid High-Level/Low-Level Tracking	41
4.1	Motivation	41
4.2	Related Work	44
4.3	Integrated Tracking Frameworks	46
4.4	Hybrid Tracking with Level Sets	47
4.5	Hybrid Tracking with ICP	61
4.6	Discussion	73

II ROI based Object Detection and Tracking	77
5 Robust ROI Extraction and Segmentation	81
5.1 Point Cloud Labeling	82
5.2 ROI Extraction	84
5.3 ROI Segmentation	85
5.4 Experimental Results	87
5.5 Discussion	90
6 Close-Range Human Detection and Tracking	91
6.1 Related Work	92
6.2 Approach	93
6.3 Experimental Evaluation	97
6.4 Extensions	101
6.5 Discussion	103
7 Tracking with Time-Constrained Detection	107
7.1 Related Work	108
7.2 System Overview	109
7.3 Poisson Process Attention Model	110
7.4 Detailed Implementation	111
7.5 Experimental Results	115
7.6 Discussion	117
III Tracking People and Their Objects	121
8 Tracking Known and Unknown Objects	125
8.1 Related Work	126
8.2 Overview	128
8.3 3D Object Representation	129
8.4 Stereo Depth-Based Tracking-Before-Detection	130
8.5 Carried Item Detection	132
8.6 Experimental Results	135
8.7 Discussion	138
9 Person-Person and Person-Object Interaction	141
9.1 Related Work	143
9.2 Modeling Person-Object Interactions	143
9.3 Learning	145
9.4 Inference and Prediction	146
9.5 Robust 3D Data Association and Tracking	148
9.6 Experimental Results	151

9.7 Discussion	156
10 Conclusion	161
10.1 Contributions	162
10.2 Perspectives	163
Bibliography	167