

Contents

Abbreviations and Symbols	xiii
List of Figures	xv
List of Tables	xxi
1. Introduction	1
1.1. Motivation	1
1.2. Problem description	2
1.3. Contributions	4
1.4. Requirements	5
1.5. Outline	8
2. State of the art	9
2.1. Data acquisition	9
2.1.1. Stereo vision	10
2.1.2. Time of flight	14
2.1.3. Low-cost structured light	15
2.1.4. Sensor fusion	16
2.2. Object modeling	18
2.2.1. Feature descriptors	19
2.2.2. Data association	22
2.2.3. Object modeling procedures	24
2.2.4. Simultaneous localization and mapping (SLAM)	25
2.3. Object recognition	26
2.3.1. 3D-3D pose estimation	27
2.3.2. Robust estimators	27
2.3.3. Object recognition frameworks	28
3. System design	30
3.1. Sensor fusion for data acquisition	30
3.2. Object modeling	31

3.3. Object recognition	33
3.4. Hardware setup	34
3.5. Software design	38
4. Sensor Fusion	40
4.1. Calibration	42
4.2. Pre-processing	44
4.3. Projection	45
4.4. Pixelwise matching costs	46
4.5. Cost aggregation	47
4.6. Post-processing	49
4.7. Evaluation	49
5. Object Modeling	55
5.1. Data acquisition	58
5.2. Modeling of textured objects	60
5.2.1. Local feature descriptor	60
5.2.2. Data association	63
5.2.3. Model-based object modeling	64
5.2.4. Post-processing	68
5.2.5. Evaluation	69
5.3. Modeling of texture-less objects	79
5.3.1. Local feature descriptor	81
5.3.2. Dynamic programming for rapid descriptor computation	85
5.3.3. Appearance-based object modeling	86
5.3.4. Evaluation	90
6. Object Recognition	91
6.1. Data acquisition	93
6.2. Recognition of textured objects	93
6.2.1. Pre-processing	93
6.2.2. Data association and clustering	94
6.2.3. Pose estimation	96
6.2.4. Evaluation	96
6.3. Recognition of texture-less objects	105
6.3.1. Adaptive sliding window for data association	106
6.3.2. Voting map for pose estimation	107
6.3.3. Evaluation	108

7. Conclusion and Outlook	117
A. Glossary	121
Bibliography	123