

Contents

| | |
|---|-----------|
| 1. Introduction | 1 |
| 1.1. Key Contributions | 3 |
| 1.2. Publications | 4 |
| 1.3. Open-Source Software Releases | 8 |
| 1.4. Collaborations | 8 |
| | |
| I. RGB-D Representation and Registration Methods | 9 |
| | |
| 2. RGB-D Image Representation in Multi-Resolution Surfel Maps | 11 |
| 2.1. RGB-D Sensor Characteristics | 12 |
| 2.2. Multi-Resolution Surfel Maps | 15 |
| 2.2.1. Modeling Measurement Errors | 17 |
| 2.2.2. Shape-Texture Descriptor | 18 |
| 2.2.3. Efficient RGB-D Image Aggregation | 19 |
| 2.2.4. Handling of Image and Virtual Borders | 20 |
| 2.3. Experiments | 20 |
| 2.3.1. Single RGB-D Image Aggregation | 21 |
| 2.3.2. Multi-View Map Aggregation | 23 |
| 2.4. Related Work | 25 |
| 2.5. Summary | 26 |
| | |
| 3. Rigid Registration | 29 |
| 3.1. Background | 29 |
| 3.1.1. Non-Linear Function Optimization | 29 |
| 3.1.2. Non-Linear Least Squares Optimization | 31 |
| 3.2. Efficient Rigid Registration of Multi-Resolution Surfel Maps | 34 |
| 3.2.1. Multi-Resolution Surfel Association | 34 |
| 3.2.2. Pose Estimation | 36 |

| | |
|---|-----------|
| 3.3. Experiments | 41 |
| 3.3.1. Evaluation Measure | 42 |
| 3.3.2. Accuracy | 42 |
| 3.3.3. Robustness | 46 |
| 3.3.4. Run-Time | 47 |
| 3.4. Related Work | 50 |
| 3.5. Summary | 51 |
| 4. Rigid Multi-Body Registration | 53 |
| 4.1. Background | 54 |
| 4.1.1. Expectation-Maximization | 54 |
| 4.1.2. Probabilistic Graphical Models for Image Labeling Tasks | 56 |
| 4.2. Efficient Rigid Multi-Body Registration of RGB-D Images | 65 |
| 4.2.1. An Expectation-Maximization Framework for Dense 3D Motion Segmentation of Rigid Parts | 65 |
| 4.2.2. Image Labeling Posterior | 67 |
| 4.2.3. Efficient Approximate Solution of the Expectation-Maximization Formulation | 69 |
| 4.2.4. Model Complexity | 71 |
| 4.2.5. Sequential Segmentation | 72 |
| 4.2.6. Image Representation | 73 |
| 4.3. Experiments | 77 |
| 4.3.1. Evaluation Measures | 79 |
| 4.3.2. Run-Time | 79 |
| 4.3.3. Segmentation Accuracy | 80 |
| 4.3.4. Motion Estimate Accuracy | 81 |
| 4.4. Related Work | 82 |
| 4.5. Summary | 83 |
| 5. Deformable Registration | 85 |
| 5.1. Background: Coherent Point Drift | 85 |
| 5.1.1. Mixture Model for Observations | 86 |
| 5.1.2. Registration through Expectation-Maximization | 86 |
| 5.1.3. Regularized Deformation Field | 88 |
| 5.1.4. Regularized Maximization Step | 89 |
| 5.2. Efficient Coarse-To-Fine Deformable Registration of Multi-Resolution Surfel Maps | 91 |
| 5.2.1. Per-Resolution Initialization | 91 |
| 5.2.2. Resolution-Dependent Kernel with Compact Support | 92 |
| 5.2.3. Handling of Resolution-Borders | 93 |
| 5.2.4. Convergence Criteria | 96 |
| 5.2.5. Color and Contour Cues | 96 |

| | |
|--|-----|
| 5.3. Local Deformations | 96 |
| 5.3.1. Local Deformations from Model to Scene | 96 |
| 5.3.2. Local Deformations from Scene to Model | 98 |
| 5.4. Experiments | 99 |
| 5.4.1. Quantitative Evaluation | 99 |
| 5.4.2. Deformable Registration and Local Transformations | 101 |
| 5.5. Related Work | 102 |
| 5.6. Summary | 106 |

II. Scene and Object Perception 109

| | |
|---|------------|
| 6. Modeling and Tracking of Rigid Scenes and Objects 111 | 111 |
| 6.1. Background | 112 |
| 6.1.1. Simultaneous Localization and Mapping | 112 |
| 6.1.2. SLAM Graph Optimization as Sparse Non-Linear Least Squares | 113 |
| 6.1.3. Particle Filters | 114 |
| 6.2. Scene and Object Modeling with Multi-Resolution Surfel Maps | 115 |
| 6.2.1. Constraint Detection | 117 |
| 6.2.2. Key-View Pose Graph Optimization | 119 |
| 6.2.3. Obtaining Scene and Object Models from Key View Graphs | 119 |
| 6.3. Object Detection and Real-Time Tracking | 120 |
| 6.3.1. Detecting Objects and Estimating Pose with Multi-Resolution Surfel Maps | 120 |
| 6.3.2. Tracking through Registration | 124 |
| 6.3.3. Object Tracking with Particle Filters | 124 |
| 6.3.4. Joint Object Detection, Pose Estimation, and Tracking in a Particle Filter Framework | 131 |
| 6.4. Experiments | 132 |
| 6.4.1. Evaluation Measures | 133 |
| 6.4.2. SLAM in Indoor Scenes | 134 |
| 6.4.3. Learning 3D Object Models | 135 |
| 6.4.4. Object Detection and Pose-Estimation | 137 |
| 6.4.5. Object Tracking | 139 |
| 6.4.6. Joint Object Detection, Pose Estimation, and Tracking | 142 |
| 6.4.7. Public Demonstration | 144 |
| 6.5. Related Work | 146 |
| 6.5.1. SLAM with RGB-D Sensors | 146 |
| 6.5.2. Object Detection and 6-DoF Pose Estimation | 148 |
| 6.5.3. Object Tracking | 149 |
| 6.5.4. Joint Object Detection, Pose Estimation, and Tracking | 151 |
| 6.6. Summary | 151 |

| | |
|---|------------|
| 7. Non-Rigid Scene and Object Perception | 153 |
| 7.1. Discovery and Dense Modeling of Object Hierarchies in Dynamic Scenes | 154 |
| 7.1.1. Discovery of Objects and Relations in RGB-D Video | 155 |
| 7.1.2. Simultaneous Localization and Mapping of Singularized Objects | 161 |
| 7.1.3. Out-Of-Sequence Relations | 162 |
| 7.1.4. Dense Models of Singularized Objects | 163 |
| 7.2. Shape Matching for Object Manipulation Skill Transfer | 164 |
| 7.2.1. Grasp Transfer | 165 |
| 7.2.2. Motion Transfer | 165 |
| 7.3. Experiments | 167 |
| 7.3.1. Hierarchical Object Discovery and Dense Modelling | 167 |
| 7.3.2. Object Manipulation Skill Transfer | 181 |
| 7.4. Related Work | 182 |
| 7.4.1. Hierarchical Object Discovery and Dense Modelling | 182 |
| 7.4.2. Object Manipulation Skill Transfer | 183 |
| 7.5. Summary | 183 |
| 8. Semantic Object-Class Perception | 185 |
| 8.1. RGB-D Object-Class Segmentation with Random Decision Forests | 185 |
| 8.1.1. Structure of Random Decision Forests | 185 |
| 8.1.2. RGB-D Image Features | 186 |
| 8.1.3. Training Procedure | 188 |
| 8.2. Dense Real-Time Semantic Mapping of Object-Classes | 189 |
| 8.2.1. Probabilistic 3D Mapping of Object-Class Image Segmentations | 189 |
| 8.2.2. Integrated Real-Time Semantic Mapping | 191 |
| 8.3. Experiments | 191 |
| 8.3.1. NYU Depth v2 Dataset | 192 |
| 8.3.2. AIS Large Objects Dataset | 195 |
| 8.4. Related Work | 196 |
| 8.5. Summary | 198 |
| 9. Conclusions | 199 |
| Acronyms | 203 |
| Bibliography | 228 |