An Efficient 3D Scene Modeling Algorithm Based on a Pan-Tilt-Driven RGB-D Camera

Chi-Yi Tsai, Chih-Hung Huang, and Shang-Jhih Jhang
Department of Electrical and Computer Engineering, Tamkang University
Outline

- Introduction
- Problem statement
- Algorithm
- Experimental result
- Conclusions
Introduction

3D registration

- Three-dimensional (3D) scene reconstruction is an important issue for several applications of robotic vision such as map construction, environment recognition, augmented reality, and Simultaneous Localization and Mapping (SLAM).
Introduction

- **3D registration method**
  - Local registration
    - Local registration methods have proven to be able to generate an accurate reconstruction model if multiple point clouds are close enough to each other.
  - Global registration
    - Global registration can be initiated with two point clouds in any status, without requiring them to be close enough to each other, and is able to obtain a result similar to local registration.
  - Local descriptors registration
    - Local descriptors registration requires a feature descriptors matching process to search for point correspondences in two point clouds, with the data of the two point clouds not required to be close to each other initially.
Problem statement

- Some of the modern point cloud alignment algorithms require a lot of computing power.
- The greater the amount of valid data in the point cloud, the amount of computing will become non-linear growth.
Proposed method

Device

Point data

Decision action

Offline calibration

Color and point data

Decision action

Online operation

Multi-view transformation matrix

Finish 3D sense model
Proposed method

- Offline calibration (camera calibration)

Calibrate color data

Multi-view extrinsic parameters
Proposed method

- Offline calibration (local registration)

- Multi-view extrinsic parameters
- Calibrate point data
- Iterative closest point
- Multi-view transformation matrix
Proposed method

- Online operation

Multi-view transformation matrix

runtime point data → Points coordinate transformation

Fine registration (Optional) ex. ICP

Finial 3D Sense model
Experimental results

Point cloud data 1

- Proposed method
- Super4PCS
- Fast global registration
- Proposed method + ICP
- Super4PCS + ICP
- Fast global registration + ICP
Experimental results

Room 1 Point cloud data

- Proposed method
- Super4PCS
- Fast global registration

- Proposed method + ICP
- Super4PCS + ICP
- Fast global registration + ICP
Experimental results

- **Average RMS results**

<table>
<thead>
<tr>
<th>Test dataset</th>
<th>Proposed method</th>
<th>Super 4PCS</th>
<th>FGR</th>
<th>Proposed method + ICP</th>
<th>Super4P CS+ ICP</th>
<th>FGR + ICP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1m dataset</td>
<td>19.6716</td>
<td>22.1680</td>
<td>19.7152</td>
<td>19.5366</td>
<td>20.3303</td>
<td>19.2841</td>
</tr>
<tr>
<td>3m dataset</td>
<td>19.9090</td>
<td>30.2470</td>
<td>20.1706</td>
<td>19.6401</td>
<td>29.2774</td>
<td>19.0025</td>
</tr>
<tr>
<td>5m dataset</td>
<td>32.0353</td>
<td>43.1703</td>
<td>33.9440</td>
<td>31.5003</td>
<td>40.6470</td>
<td>31.2734</td>
</tr>
<tr>
<td>Room 1</td>
<td>18.4981</td>
<td>27.0479</td>
<td>18.8040</td>
<td>18.3856</td>
<td>24.4929</td>
<td>18.4195</td>
</tr>
<tr>
<td>Room 2</td>
<td>26.3199</td>
<td>29.0571</td>
<td>29.5579</td>
<td>25.7527</td>
<td>25.6207</td>
<td>25.7157</td>
</tr>
</tbody>
</table>
## Experimental results

- **Average processing time (in milliseconds)**

<table>
<thead>
<tr>
<th>Test method</th>
<th>1m dataset</th>
<th>3m dataset</th>
<th>5m dataset</th>
<th>Room 1</th>
<th>Room 2</th>
<th>Living room</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed method</td>
<td>4.11</td>
<td>3.14</td>
<td>4.18</td>
<td>3.77</td>
<td>4.24</td>
<td>3.90</td>
</tr>
<tr>
<td>Proposed method + ICP</td>
<td>27378.52</td>
<td>17385.10</td>
<td>32572.52</td>
<td>40000.84</td>
<td>52331.24</td>
<td>46932.69</td>
</tr>
</tbody>
</table>
Conclusions

- One advantage of the proposed method is that it converges quickly and only needs to perform point cloud transformations without any iterative process.

- The proposed method can process in real time. The average processing time of the proposed method is only about 4ms to align two point clouds, which is nearly 138,000 times to the well-known Super4PCS approach and nearly 3,000 times to the recently published FGR approach.
Thanks for your attention