Vision Aided 3D Laser Scanner Based Registration

ECMR, 2007, Freiburg, Germany

Henrik Andreasson, Achim Lilienthal

AASS Learning Systems Lab, Örebro University, Sweden
1. Motivation
2. Camera / Laser Range Finder (LRF) calibration
3. Registration method
4. Experimental results
5. Conclusions
1 Motivation

- Fundamental problem in Mapping / Localization
- Visual appearance based approach
- No initial pose estimate (loop closing)
- No tracking of visual features positions
- Suitable for *Graph-based* SLAM methods
2 Camera / LRF Calibration

- Robot Tjorven

Internal calibration (not covered here)

External calibration
- Rotation R
- Translation t
2 Camera / LRF Calibration

Reflective tape
2 Camera / LRF Calibration
2 Camera / LRF Calibration
2 Camera / LRF Calibration

- 3D points from LRF
- 3D points from camera

Determine R, t by optimizing the SSD
3 Registration Methods

2 Iterative Closest Point (ICP) based version

Plain ICP

Visual feature pair

\[ J(\mathbb{R}, \mathbf{t}) = \sum_{i=1}^{N} ||y_i - \mathbb{R}x_i - \mathbf{t}||^2 \]

Generalized Total Least Square ICP (GTLS-ICP)

\[ J(\mathbb{R}, \mathbf{t}) = \sum_{i=1}^{N} (q_i - y_i)^T C_{q_i}^{-1} (q_i - y_i) + \]

\[ q_i = \mathbb{R}x_i + \mathbf{t} \]

\[ \sum_{i=1}^{N} (y_i - q_i)^T C_{y_i}^{-1} (y_i - q_i), \]
3 Registration Methods

- Applying Covariance Estimate in the Registration
3 Registration Methods

- Determine the covariance

Laser readings

2D laser reading

Visual feature

Closest projected laser reading = Depth

Wrist movement
3 Registration Methods
3 Registration Methods

Wrong correspondences

\[ J(\mathbb{R}, t) = \sum_{i=1}^{\times} \| y_i - R x_i - t \|^2 \]
Use a ‘trimmed’ version: use 70% of the closest pairs (=iterative method)

\[ J(\mathbb{R}, t) = \sum_{i=1}^{0.7N} ||y_i - \mathbb{R}x_i - t||^2 \]
3 Registration Methods

- **ICP**

\[
J(\mathbf{R}, \mathbf{t}) = \sum_{i=1}^{N} ||y_i - \mathbf{R}x_i - \mathbf{t}||^2
\]

- **Generalized Total Least Square ICP (GTLS-ICP)**

\[
J(\mathbf{R}, \mathbf{t}) = \sum_{i=1}^{N} (q_i - y_i)^T C_{q_i}^{-1} (q_i - y_i) + \sum_{i=1}^{N} (y_i - q_i)^T C_{y_i}^{-1} (y_i - q_i),
\]

\(q_i = \mathbf{R}x_i + \mathbf{t}\)
3 Registration Methods

- **ICP**

\[ J(\mathbf{R}, \mathbf{t}) = \sum_{i=1}^{N} ||y_i - R x_i - t||^2 \]

Closed form solution – FAST
(obtain \( \mathbf{R} \) and \( \mathbf{t} \) in one step)
SVD approach by Arun et. al

- **Generalized Total Least Square ICP (GTLS-ICP)**

\[ J(\mathbf{R}, \mathbf{t}) = \sum_{i=1}^{N} (q_i - y_i)^T C^{-1}(q_i - y_i) + \]
\[ \sum_{i=1}^{N} (y_i - q_i)^T C^{-1}_{y_i}(y_i - q_i), \]

\( q_i = \mathbf{R} x_i + \mathbf{t} \)

No closed form solution – SLOW
(use regular optimization)
4 Experimental Results

3 Laser Scanner Sweeps
4 Experimental Results

3 Laser Scanner Sweeps & 7 Images per Robot Pose
4 Experimental Results

- 22 Robot Poses (3 scans + 7 images each)
- Incremental Registration
- Evaluation

Ground Truth
## 4 Experimental Results

### Evaluation

<table>
<thead>
<tr>
<th></th>
<th>$Tr. \text{ ICP}$</th>
<th>$Tr. \text{ GTLS – ICP}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>10 15 20</td>
<td>10 15 20</td>
</tr>
<tr>
<td>$d$</td>
<td>1.14 0.76 0.30</td>
<td>0.84 0.70 0.24</td>
</tr>
<tr>
<td>$\sigma_d$</td>
<td>0.54 0.83 0.11</td>
<td>0.33 0.85 0.14</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.30 0.17 0.05</td>
<td>0.25 0.18 0.06</td>
</tr>
<tr>
<td>$\sigma_\alpha$</td>
<td>0.18 0.24 0.02</td>
<td>0.11 0.22 0.04</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$Tr. \text{ ICP}$</th>
<th>$Tr. \text{ GTLS – ICP}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>30 40 60</td>
<td>30 40 60</td>
</tr>
<tr>
<td>$d$</td>
<td>0.09 0.14 0.13</td>
<td>0.11 0.19 0.15</td>
</tr>
<tr>
<td>$\sigma_d$</td>
<td>0.05 0.07 0.03</td>
<td>0.08 0.10 0.06</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.04 0.03 0.03</td>
<td>0.04 0.04 0.03</td>
</tr>
<tr>
<td>$\sigma_\alpha$</td>
<td>0.02 0.01 0.01</td>
<td>0.01 0.02 0.02</td>
</tr>
</tbody>
</table>

Number of features used (randomly selected)
## 4 Experimental Results

- Number of iterations for convergence using Tr-ICP (keeping 70% of the closest pairs)

<table>
<thead>
<tr>
<th>$N$</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>#iter</td>
<td>4.46</td>
<td>4.95</td>
<td>5.20</td>
<td>5.00</td>
<td>6.00</td>
<td>5.97</td>
</tr>
<tr>
<td>$\sigma_{iter}$</td>
<td>0.68</td>
<td>0.86</td>
<td>0.93</td>
<td>1.87</td>
<td>1.07</td>
<td>1.07</td>
</tr>
</tbody>
</table>
4 Additional Experimental Results

Outdoors
4 Additional Experimental Results

Outdoors
4 Additional Experimental Results
4 Additional Experimental Results

- Global optimization test:
4 Additional Experimental Results
4 Additional Experimental Results

Links / constraints - determine by the ‘similarity measure’ using visual features

Optimize all poses at once (21 * 6 params…)
4 Additional Experimental Results

No optimization

Optimization
5 Conclusions

- Similarity based approach
- Vision to handle correspondences
- No feature tracking
- Works well (indoors)

- More evaluation (outdoors)
- Comparison to other methods
  - (better results than 3D-NDT in the indoor data set providing initial pose estimates)