Local-to-Global Point Cloud Registration using a Dictionary of Viewpoint Descriptors

David Avidar, David Malah, and Meir Barzohar

Andrew and Erna Viterbi Faculty of Electrical Engineering, Technion – Israel Institute of Technology

Goal
- Registration (alignment) between a global, large-scale point cloud, and a single local scan.

Registration algorithm – Overview
- Main concepts:
  - Convert global cloud to a dictionary of viewpoint descriptors.
  - Find registration via dictionary search.

Viewpoint dictionary creation
- Detect ground (e.g., by region growing).
- Define a viewpoint grid over the global cloud (e.g., grid distance = 3m).
- Compute a panoramic range-image for each viewpoint.

Verification and registration refinement
- Select \( N_{\text{final}} \) (e.g., 3) coarse registrations with lowest RMSE (Root-Mean-Square Error between the local and global clouds).
- Refine selected registrations using ICP (Iterative Closest Point).
- Select final registration with lowest RMSE.

Panoramic range-images as viewpoint descriptors
- Use the view of the global scene from each viewpoint as a descriptor and match with local range-image.

Candidate viewpoint selection via phase-correlation
- Select a subset of grid viewpoints (e.g., 10) with largest phase-correlation peaks.
- Peak location \( \rightarrow \) coarse orientation estimate.

Motivation
- Combining airborne and terrestrial LiDAR scans for large-scale urban modeling.
- Localization in an urban environment even without GPS.

Challenges
- Airborne vs. terrestrial data characteristics:
  - Very different point density distributions (airborne \( \rightarrow \) more points on horizontal surfaces, terrestrial \( \rightarrow \) more points on vertical surfaces).
  - Missing data, different types of occlusion.
- Urban environments contain many flat surfaces and repetitive features.
- We found existing 3D feature-based methods (e.g., Spin-Images, Fast Point Feature Histograms - FPFH) to be unreliable under these conditions.

Results
- The registration algorithm was tested on a large-scale and challenging dataset:
  - 108 terrestrial LiDAR scans (Z+F Imager 5010).
  - Airborne LiDAR scan (global cloud) area: \(-1km^2\) (Leica ALS80).

<table>
<thead>
<tr>
<th>Localization error [m]</th>
<th>Relative Rotation Error [deg]</th>
<th>Registration runtime per cloud [sec]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>STD</td>
<td>Mean</td>
</tr>
<tr>
<td>Using GPS (3km search radius)</td>
<td>0.43</td>
<td>0.27</td>
</tr>
<tr>
<td>No GPS (search entire grid)</td>
<td>0.44</td>
<td>0.27</td>
</tr>
</tbody>
</table>

*Run on PC (27GHz CPU, 512GB RAM, MATLAB implementation)

- Comparison to FPFH - tested on a subset of 24 local clouds:
  - Using FPFH, only 6 out of 24 clouds had localization error lower than 3m (the other 18 registrations failed).
  - Using the proposed method, the maximal localization error was 0.79m (the max. error over all 108 clouds was 1.85m).

Conclusion
- 3D feature-based methods were found to be unreliable for local-to-global terrestrial-to-airborne registration.
- Registration performance of the proposed viewpoint-dictionary-based method is significantly better.