The Future of Real-Time SLAM

ElasticFusion: real-time dense SLAM without a pose graph

Stefan Leutenegger (representing Tom Whelan)
18th December 2015 (ICCV Workshop)
State of the art in real-time dense SLAM

Plenty have limitations

- No (online) loop closure
- Only estimates trajectory
  - Raw point cloud back projections
  - Key frames
- Non-scalable
- Non-robust pose estimation

Some of these include;

- Henry et al., Endres et al., Meilland & Comport, Kerl et al., Keller et al., Chen et al, Nießner et al., Newcombe et al., Steinbruecker et al., Stueckler et al.
State of the art

Offline approaches

- Qian-Yi Zhou
  Amazing results with RGB-D, strictly offline (>1hr processing time)
State of the art

Online approaches

• Kintinuous arguably the state of the art
  » Online loop closure
  » Estimates full 3D surface and trajectory
  » Scalable (100’s of metres)
  » Full RGB and Depth pose estimation
• However, still not the “perfect” system
  » A number of limitations
State of the art

Kintinuous

- Surface aliasing
State of the art

Kintinuous
We want a SLAM system that...

Robustly estimates camera pose
- Geometry + photometry

Reliably estimates the surface
- Fused representation to remove noise

Scales well
- Room, house scale

Is completely closed loop (updating)
- Update revisited areas

Real-time
- Globally consistent map available at any point in time

Non-restrictive of motion
- Happy to deal with extremely loopy motion and many such loop closures
Introducing ElasticFusion

Robustly estimates camera pose
  • Full RGB and Depth frame-to-model tracking

Reliably estimates the surface
  • Point-based fusion is good quality and a nice representation

Scales well
  • Room scale seems doable, at a minimum

Is completely closed loop (updating)
  • No separation between front end and back end

Real-time
  • Strictly

Non-restrictive of motion
  • Since the front end and back end are one and the same, it is less restrictive given the full frame-to-model tracking
How it works

1. Reconstruct surfel-based map of environment
2. Split into active/inactive regions
3. Directly register multiple passes of the same surface together
4. Reflect this in the map with a non-rigid space deformation
5. Use fern encoded key frames for global loop closures
How it works: Tracking

Data terms

Residuals

\[ E_{rgb} = \sum_{u \in \Omega} \left( I(u, C_t^l) - I\left( \pi(K \exp(\xi)T p(u, D_t^l)), C_{t-1}^a \right) \right)^2 \]

\[ E_{icp} = \sum_k \left( \left( v^k - \exp(\xi)T v_t^k \right) \cdot n^k \right)^2 \]
How it works: Building a Deformation Graph

Mapping left to right

Mapping right to left

Deformation graph

Time scale
How it works: Time Stretched Visualisation
How it works: Loop Closure

\[ E_{rot} = \sum_{l} \left\| \mathcal{G}_R^l \mathcal{T} \mathcal{G}_R^l - \mathbf{I} \right\|_F^2 \]  As-rigid-as-possible

\[ E_{reg} = \sum_{l} \sum_{n \in \mathcal{N}(\mathcal{G}^l)} \left\| \mathcal{G}_R^l (\mathcal{G}_g^n - \mathcal{G}_g^l) + \mathcal{G}_g^l + \mathcal{G}_t^l - (\mathcal{G}_g^n + \mathcal{G}_t^n) \right\|_2^2 \]  Smoothness regulariser

\[ E_{con} = \sum_{p} \left\| \phi(Q_s^p) - Q_d^p \right\|_2^2 \]  New-to-old loop closure constraints

\[ E_{pin} = \sum_{p} \left\| \phi(Q_d^p) - Q_d^p \right\|_2^2 \]  Old-to-old anchoring constraints

\[ E_{rel} = \sum_{p} \left\| \phi(R_s^p) - \phi(R_d^p) \right\|_2^2 \]  Relative constraints (previous loop closures)
Overview (Real-time)

Revisited inactive areas are detected and trigger local dense surface loop closures

https://www.youtube.com/watch?v=XySrhzpODYs
ElasticFusion – Extras

MIT_76_417b dataset (Real-time)

https://www.youtube.com/watch?v=-dz_VauPjEU
## Quantitative Results

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DVO SLAM</td>
<td>0.021m</td>
<td>0.018m</td>
<td>0.035m</td>
<td>0.018m</td>
</tr>
<tr>
<td>RGB-D SLAM</td>
<td>0.023m</td>
<td><strong>0.008m</strong></td>
<td>0.032m</td>
<td>0.017m</td>
</tr>
<tr>
<td>MRSMap</td>
<td>0.043m</td>
<td>0.020m</td>
<td>0.042m</td>
<td>2.018m</td>
</tr>
<tr>
<td>Kintinuous</td>
<td>0.037m</td>
<td>0.029m</td>
<td>0.030m</td>
<td>0.031m</td>
</tr>
<tr>
<td>Frame-to-model</td>
<td>0.022m</td>
<td>0.014m</td>
<td>0.025m</td>
<td>0.027m</td>
</tr>
<tr>
<td>ElasticFusion</td>
<td><strong>0.020m</strong></td>
<td>0.011m</td>
<td><strong>0.017m</strong></td>
<td><strong>0.016m</strong></td>
</tr>
</tbody>
</table>

### TUM RGB-D Trajectory Error

<table>
<thead>
<tr>
<th>System</th>
<th>kt0</th>
<th>kt1</th>
<th>kt2</th>
<th>kt3</th>
</tr>
</thead>
<tbody>
<tr>
<td>DVO SLAM</td>
<td>0.032m</td>
<td>0.061m</td>
<td>0.119m</td>
<td>0.053m</td>
</tr>
<tr>
<td>RGB-D SLAM</td>
<td>0.044m</td>
<td>0.032m</td>
<td>0.031m</td>
<td>0.167m</td>
</tr>
<tr>
<td>MRSMap</td>
<td>0.061m</td>
<td>0.140m</td>
<td>0.098m</td>
<td>0.248m</td>
</tr>
<tr>
<td>Kintinuous</td>
<td>0.011m</td>
<td>0.008m</td>
<td>0.009m</td>
<td>0.150m</td>
</tr>
<tr>
<td>Frame-to-model</td>
<td>0.098m</td>
<td><strong>0.007m</strong></td>
<td>0.011m</td>
<td>0.107m</td>
</tr>
<tr>
<td>ElasticFusion</td>
<td><strong>0.007m</strong></td>
<td><strong>0.007m</strong></td>
<td><strong>0.008m</strong></td>
<td><strong>0.028m</strong></td>
</tr>
</tbody>
</table>

### ICL-NUIM Surface Error
Quantitative Results

DVO SLAM

RGB-D SLAM

MRSMap

Kintinuous

Frame-to-model

ElasticFusion
Main Advances

Real-time deformation
  • Great for overcoming the drift problem in a dense map

Fully closed loop
  • No frontend/backend division opens up many possibilities

Open source
  • https://github.com/mp3guy/ElasticFusion
Light Source Estimation

Reflectance-driven
  • Detect speculars

Why?
  • It’s cool
  • Convincing AR effects
  • Can be used to improve tracking
  • Aids in path planning (i.e. avoid bright areas)
Light Source Estimation

Reconstruct diffuse appearance

Bright raw observations are reflected rays
  • Vote in voxel space (hough-like scheme)

Intersections of high votes and geometry are potentially light sources
  • Geometry helps determine extent of light source, directionality and removes need for comprehensive reflected ray coverage