Autonomous Navigation for Flying Robots

Lecture 8.3: Direct Methods for Visual SLAM
Jürgen Sturm
Technische Universität München
Feature-Based Visual SLAM

- Video feed from quadrocopter
Feature-Based Visual SLAM

- What PTAM actually sees
Direct Visual Odometry
[Steinbrücker, Sturm, Cremers; ICCV LDRMC 2011]

- How can we use more/all information from the image?
- Idea

- Photo-consistency constraint

\[ I_1(x) = I_2(\pi(g_\xi(z \cdot x))) \quad \text{holds for all pixels } x \]
How to deal with noise?

[Steinbrücker, Sturm, Cremers; ICCV LDRMC 2011]

- Photo-consistency constraint will not perfectly hold
  - Sensor noise
  - Pose error
  - Reflections, specular surfaces
  - Dynamic objects (e.g., walking people)
- Residuals will be non-zero
  \[ r = I_1(x) - I_2(\pi(g_{\xi}(z \cdot x))) \]
- Residual distribution \( p(r) \)
Residual Distribution

[Steinbrücker, Sturm, Cremers; ICCV LDRMC 2011]

- Zero-mean, peaked distribution
- Example: Correct camera pose
Residual Distribution
[Steinbrücker, Sturm, Cremers; ICCV LDRMC 2011]

- Zero-mean, peaked distribution
- Example: Correct camera pose
Residual Distribution

[Steinbrücker, Sturm, Cremers; ICCV LDRMC 2011]

- **Goal:** Find the camera pose that maximizes the observation likelihood
Motion Estimation
[Steinbrücker, Sturm, Cremers; ICCV LDRMC 2011]

- **Goal:** Find the camera pose that maximizes the observation likelihood

\[ \xi^* = \arg \max \prod_i p(r_i(\xi)) \]

- Assume pixel-wise residuals are conditionally independent
- How can we solve this optimization problem?
Approach

[Steinbrücker, Sturm, Cremers; ICCV LDRMC 2011]

- Take negative logarithm

\[ \xi^* = \arg \min_{\xi} \sum_i - \log p(r_i(\xi)) \]

- Set derivative to zero

\[ \sum_i \frac{\partial \log p(r_i(\xi))}{\partial \xi} = \sum_i \frac{\partial \log p(r_i)}{\partial r_i} \frac{\partial r_i(\xi)}{\partial \xi} \overset{!}{=} 0 \]
Robust Cost Functions

[Kerl, Sturm, Cremers; ICRA 2013]

- Quadratic cost term is not robust
- Rewrite as a weighted least squares problem

\[ \xi^* = \arg \min_{\xi} \sum_i w(r_i)(r_i(\xi))^2 \]

with weights
\[ w(r_i) = \frac{\partial \log p(r_i)}{\partial r_i} \frac{1}{r_i} \]

- \( r_i(\xi) \) is non-linear in \( \xi \)
- Need to linearize, solve, and iterate (Gauss-Newton method)
Weighted Error
[Kerl, Sturm, Cremers; ICRA 2013]

- normal distribution
- Tukey weights
- t-distribution

Weighted error $w(r)^2$ vs. residual $r$
Example
[Kerl, Sturm, Cremers; ICRA 2013]
Coarse-to-Fine
[Steinbrücker, Sturm, Cremers; ICCV LDRMC 2011]

- Linearization only holds for small motions
- Coarse-to-fine scheme
- Image pyramids
Direct Visual Odometry: Results

[Steinbrücker, Sturm, Cremers; ICCV LDRMC 2011]

Direct SLAM
[Kerl, Sturm, Cremers; IROS 2013]

- Dense Visual Odometry
  - Input: Two RGB-D frames
  - Output: Relative pose
  - Runs in real-time on single CPU core
- Use this in pose graph SLAM
  - Select keyframes
  - Detect loop-closures
  - Build and optimize pose graph (using g2o)
Direct SLAM
[Kerl, Sturm, Cremers; IROS 2013]

The blue cameras show the position of keyframes.

Large-Scale 3D Reconstruction
[Steinbrücker, Kerl, Sturm, Cremers; ICCV 2013]

- **We have:** Optimized pose graph
- **We want:** High-resolution 3D map

- **Problem:** High-resolution voxel grids consume much memory (grows cubically)
  - $512^3$ voxels, 24 byte per voxel $\rightarrow$ 3.2 GB
  - $1024^3$ voxels, 24 byte per voxel $\rightarrow$ 24 GB
  - …
Large-Scale 3D Reconstruction
[Steinbrücker, Kerl, Sturm, Cremers; ICCV 2013]

- Save data in oct-tree data structure
- Only allocate cells that are close to the surface
- Store geometry at multiple resolutions
- Tree can grow dynamically (no fixed size)
Large-Scale 3D Reconstruction
[Steinbrücker, Kerl, Sturm, Cremers; ICCV 2013]

Volumetric 3D Mapping in Real-Time on a CPU

Frank Steinbrücker, Jürgen Sturm, Daniel Cremers

ICRA 2014
Submission 636

Computer Vision and Pattern Recognition Group
Department of Computer Science
Technical University of Munich

http://youtu.be/7s9JoFSlm-M
Direct SLAM with a Monocular Camera
[Engel, Sturm, Cremers; ICCV 2013]

Semi-Dense Visual Odometry for a Monocular Camera
Jakob Engel, Jürgen Sturm, Daniel Cremers
International Conference on Computer Vision (ICCV)
December 2013, Sydney

Computer Vision Group
Department of Computer Science
Technical University of Munich

http://youtu.be/LZChzEcLNzl
Lessons Learned

- Direct visual odometry
- Photoconsistency constraint
- Loop closing
- Large-scale 3D reconstruction
- Software available as open-source
  https://github.com/tum-vision