Simultaneous Localization and Mapping

- input: sequence of measurements
- output: trajectory and map
2D & 3D Mapping and Localization

Christoph Stiller
Andreas Geiger
Frank Moosmann

Institut für Mess- und Regelungstechnik

IROS San Francisco Sep. 2011
Trends in SLAM

- 2D sensors, flat world assumption
- sparse maps
- stop vehicle for scanning, off-line computation
- highly accurate laser sensors
- additional ego-motion sensors (odometer etc.)
- 3D world models
- dense maps
- on-the-fly scanning and real-time computation
- noise modelling, visual SLAM
- single sensor SLAM
Graph SLAM

\[ p(x_k, m|z_k^k, u_k^k) \quad \text{visual odometry (Online SLAM)} \]

\[ p(x_k^k, m|z_k^k, u_k^k) \quad \text{bundle adjustment (Full SLAM)} \]
Graph SLAM

\[ p(x_k, m | z^k) = \frac{p(z_k | x_k, m, z^{k-1}) p(x_k | m, z^{k-1})}{p(z_k | z^{k-1})} \]

\[ = \frac{p(z_k | x_k, m, z^{k-1}) \int p(x_{k-1}, m | z^{k-1}) p(x_k | x_{k-1}, m, z^{k-1}) dx_{k-1}}{p(z_k | z^{k-1})} \]
On-the-Fly Scanning

Continuous pose required
Estimate pose dynamics

Scan angle \( \phi \)

Scan #1

Scan #2

0°  90°  180°  270°

0°  90°  180°  270°  0°

Time [sec]
Features

- normal vectors
  - calculated as cross product from 4 neighbours
- flatness measure $f_i$
  - between 0 (non-flat) and 1 (flat)
  - probability if sampling theorem holds
Noise reduction

- shift points towards neighbouring planes
  - in "flat" areas only

\[
p'_i(a) = p_i + a \cdot n_i
\]

\[
a = \arg \min_a \sum_j w_{ij} (n_j^T (p'_i(a) - p_j))^2
\]

\[
w_{ij} = f_i f_j n_i^T n_j
\]

- online: by using existing, already shifted surfaces of map
- offline: by using originally scanned surfaces
Results

- evaluation of
  - map quality
  - localization precision
- new data set:
  http://www.mrt.kit.edu/z/publ/download/velodyneslam
Results: generated map
Simultaneous monoscopic and stereoscopic image sequence analysis:
+ Instantaneous object detection in near field
+ Range increasing with observation time

**stereo basis:**
- constant over time
- favorable orientation
- range uncertainty ~ range$^2$/resolution

**motion stereo basis:**
- unfavorable orientation
- increasing with time
3D SLAM Exploiting Multiple Visual Keys

- **disparity**
- **optical flow**
- (d)SLAM, ego-trajectory
- 3D static scene,
- 3D object trajectories
Correspondences

Feature matching:
- Detect interest points using non-maximum-suppression
- Match 4 images in a 'space-time' circle
- Use epipolar constraints for left ↔ right matching
- Accept if last feature coincides with first feature
Correspondences

Feature matching:
- Detect interest points using non-maximum-suppression
- Match 4 images in a 'space-time' circle
- Use epipolar constraints for left $\leftrightarrow$ right matching
- Accept if last feature coincides with first feature

Sep. 2011
Correspondences

Feature matching:
- Detect interest points using non-maximum-suppression
- Match 4 images in a 'space-time' circle
- Use epipolar constraints for left ↔ right matching
- Accept if last feature coincides with first feature

Sep. 2011
Correspondences

Feature matching:
- Detect interest points using non-maximum-suppression
- Match 4 images in a ‘space-time’ circle
- Use epipolar constraints for left $\leftrightarrow$ right matching
- Accept if last feature coincides with first feature
Correspondences

Feature matching:
- Detect interest points using non-maximum-suppression
- Match 4 images in a 'space-time' circle
- Use epipolar constraints for left $\leftrightarrow$ right matching
- Accept if last feature coincides with first feature
Fast feature matching:

- 1st: Match a sparse set of interest points within each class
- Build statistics over likely displacements within each bin
- Use this statistics for speeding up 2nd matching stage
- Rejection outliers (Delaunay triangulation)
Single CPU Implementation

CPU with SSE3 instruction set

Core 1 12 fps
Scene flow
Egomotion

Core 2 4 fps
Dense stereo
Fusion + 3d reprojection

Sep. 2011
Ego-Motion

\[
\min_{r, t} \sum_{i=1}^{N} \left\| \mathbf{x}_i^{(l)} - \pi^{(l)}(\mathbf{X}_i; r, t) \right\|^2 + \left\| \mathbf{x}_i^{(r)} - \pi^{(r)}(\mathbf{X}_i; r, t) \right\|^2
\]

- Minimize reprojection errors (Gauss-Newton + RANSAC)
- Kalman Filter (constant acceleration model)
Results

video
Conclusions

- 3d map generation & ego-trajectory estimation
  - lidar or video camera
  - large scale
  - real time on single CPU
  - large scale

- challenges
  - local bundle adjustment
  - object trajectories
  - scene understanding
  - reliability, plausibility

[Geiger, Ziegler, Stiller, IEEE Intelligent Vehicles Symposium 2011]
[Moosmann, Stiller, IEEE Intelligent Vehicles Symposium 2011]