Real-Time 3D Visual SLAM with a Hand-Held RGB-D Camera

Nikolas Engelhard$^1$, Felix Endres$^1$, Jürgen Hess$^1$, Jürgen Sturm$^2$, Wolfram Burgard$^1$

$^1$University of Freiburg, Germany
$^2$Technical University Munich, Germany
Motivation

- Our goal:
  Learn 3D models of (indoor) scenes

- Applications:
  - Architecture
  - Gaming
  - Archaeology
Related Work

- **Visual SLAM**
  - [Konolige, IROS 09], [Klein, ISMAR 09]

- **Laser scanners**
  - [Wurm, ICRA 10]

- **RGB-D**
  - [Pitzer, ICRA 10], [Henry, ISER 10]
RGB-D Sensor

- Principle: structured light
  - IR projector + IR camera
  - RGB camera
- Dense depth images
- Full video frame rate
Schematic Overview

Input: stream of RGB-D images

feature extraction and matching (SURF)

pose estimation (RANSAC)

pose refinement (ICP)

pose graph optimization (HOGMAN)

Output: 3D model (colored point cloud)
Feature Extraction and Matching

- **SURF**: scale and rotation invariant descriptor

- **Runtime**
  - 600-800 features
  - OpenCV: 1 sec
  - SurfGPU: 40 ms

- **Matching**
  - FLANN
  - < 15 ms
Features in 3D

- Associate features with 3D points
- Problem: missing data (glass, occlusion)
Estimate Relative Pose

- RANSAC: good for estimating models in the presence of outliers

- **Algorithm:**
  - Find correspondences
  - Repeatedly sample three correspondences, estimate pose, count inliers (and optimize)
  - Return pose with most inliers

- **Runtime:** < 5 ms
Correspondences in 2D

Image at time $t$  Image at time $t+1$
Pose Refinement

- Iterative Closest Point
- Generalized ICP [Segal, 2009]
  - Plane-to-Plane metric
- Local optimization (prone to local maxima)
- Needs good initialization

- Runtime: ~500 ms on a sampled subset
Pose Refinement (2)

Without ICP

With ICP
Loop Closing

- Summation of incremental error → drift
- Detect loops and average errors
- Model as pose graph
- Pose graph optimization [Grisetti, 2010]
  - System is over-determined
  - Needs to find the best globally consistent alignment
Example Pose Graph

- Axes = estimated camera poses
- White edges = relative transformations
Video (1/3)

RGBD SLAM with ROS + Kinect

[http://www.youtube.com/watch?v=XejNctt2Fcs]
Free-Hand 3D Model Reconstruction with Kinect and RGBDSLAM

by

Autonomous Intelligent Systems Lab
University of Freiburg, Germany

[http://www.youtube.com/watch?v=5qrBEPfEPaY]
Video 3/3

Free-Hand 3D Model Reconstruction with Kinect and RGBDSLAM

by

Autonomous Intelligent Systems Lab
University of Freiburg, Germany

[http://www.youtube.com/watch?v=NR-ycTNcQu0]
Conclusions

- Visual SLAM system
  - SURF feature matching
  - RANSAC pose estimation
  - ICP pose refinement
  - Pose graph optimization

- Real-time

- Open-source (in ROS) + Tutorial available: http://www.ros.org/wiki/openni/Contests/ROS 3D/RGBD-6DSLAM

→ Live demo after the coffee break!
Future Work

- Ground truth evaluation using a Motion Capturing system (in cooperation with ETH Zurich)
- Improve speed of ICP
- Parallelization
Thank you!