Tracking and Mapping in Project Tango

Dr. Jürgen Sturm
Motivation

Many applications require a solid understanding of space and motion

- Personal assistants (smartphones, wearables)
- Autonomous robots, cars, drones
Relevant Information

- Where am I?
- How do I move?
- How does the world around me look like?
- How can I get from A to B?
- ...

Navigation

AR/VR games

Autonomous cars

Autonomous drones

Source: newsingeneral.com
Source: Magic Leap
Source: Google Self-Driving Car
Source: Amazon Prime Air
Motivating Question

How to solve this?

- Sensors
- Algorithms
- Compute

Source: Amazon Prime Air
Source: Magic Leap
Source: Google Self-Driving Car
Source: newsingeneral.com
My Research Profile

- Visual Navigation for Mobile Robots
  - RoboCup
  - Kinematic Learning
  - Articulated Objects
  - Quadcopters
  - MOOC Teaching

- Camera tracking, 3D reconstruction, Augmented Reality
  - RGB-D SLAM
  - Direct Methods
  - Large Scale Reconstruction
  - 3D Printing
  - Augmented Reality
Visual Navigation with a Quadcopter

[Engel, Sturm, Cremers; IROS ‘12]
Topic of this Talk

- Sparse maps are not suitable for autonomous navigation
- How can we generate a dense 3D model of the environment?
- Representation, estimation, refinement
Signed Distance Function (SDF)
[Curless and Levoy; SIGGRAPH 1996] [Newcombe et al; ISMAR 2011]
Signed Distance Function (SDF)

[Curless and Levoy; SIGGRAPH 1996] [Newcombe et al; ISMAR 2011]

\[ D(x) > 0 \]
\[ D(x) = 0 \]
\[ D(x) < 0 \]
Signed Distance Function (SDF)

[Curless and Levoy; SIGGRAPH 1996] [Newcombe et al; ISMAR 2011]

\[ D(x) > 0 \]
\[ D(x) = 0 \]
\[ D(x) < 0 \]
Signed Distance Function (SDF)
[Curless and Levoy; SIGGRAPH 1996] [Newcombe et al; ISMAR 2011]

- Compute SDF from a depth image
- Measure distance of each voxel to the observed surface
Signed Distance Function (SDF)
[Curless and Levoy; SIGGRAPH 1996] [Newcombe et al; ISMAR 2011]

- Assume known camera poses (for now)
- Calculate weighted average over all measurements

\[
D \leftarrow \frac{WD + wd}{W + w}
\]

\[
C \leftarrow \frac{WC + wc}{W + w}
\]

\[
W \leftarrow W + w
\]
Mesh Extraction with Marching Cubes

Find zero-crossings in the signed distance function by interpolation
3D Room Scanning with a Quadcopter

[Bylow, Sturm, Kahl, Cremers; RSS ’13] [Sturm, Bylow, Kahl, Cremers; UAV-g ’13]

- AscTec Pelican quadrocopter
- Real-time 3D reconstruction, position tracking and control (off-board, needs GPU)
More Examples
[Bylow, Sturm, Kahl, Cremers; RSS ’13] [Sturm, Bylow, Kahl, Cremers; UAV-g ’13]

Nice 3D models, but resolution could be higher.

**Problem:** Memory consumption grows cubically!

- $512^3$ voxels, 24 byte per voxel → 3.2 GB
- $1024^3$ voxels, 24 byte per voxel → 24 GB
- …
High-Resolution 3D Reconstruction
[Steinbrücker, Kerl, Sturm, Cremers; ICCV 2013] [Steinbrücker, Sturm, Cremers; ICRA 2014]

- Save data in multi-level oct-tree data structure
- Bricks ($8^3$ volumes) are only allocated when needed
- Store geometry at multiple resolutions
- Tree grows dynamically (no fixed size/origin)
High-Resolution 3D Reconstruction
[Steinbrücker, Kerl, Sturm, Cremers; ICCV 2013] [Steinbrücker, Sturm, Cremers; ICRA 2014]

- Multi-level oct-tree implementation, open-source!
- Runs at 45fps (640x480, 0.5cm grid resolution, 3GB)
Project Tango

Liked what you saw so far?

What if there was a device that can do (most of this) out-of-the-box?

Would this save you or your PhD students time?

What applications could you build on top of it?
Project Tango

Integrated hard- and software stack that provides:

- Low latency 6-DOF pose
- Map optimization and bundle adjustment
- Global re-localization
- Meshing
- C++/Java/Unity API
Use Cases

Emerging: Virtual Reality Robotics

Professional: Building-scale 3D scanning & measurement

Prosumer: Virtual Showroom Shopping Remodeling

Consumer: Centimeter-scale Indoor Navigation (no GPS/Wifi/Bluetooth)

Entertainment: Motion Gaming Geo-Social Activities
Tango: Visual-Inertial Odometry (VIO)

[Hesch, Kottas, Bowman, Roumeliotis; T-RO ’09] [Nerurkar, Wu, Roumeliotis; RSS ’13]
[Lynen, Sattler, Bosse, Hesch, Pollefeys, Siegwart; RSS ’15]

Sensor Data

IMU information
- accelerometer
- gyroscope

Visual information
- Frame-to-frame feature tracks
- Feature descriptors

VIO Frontend
Odometry

VIO On-device Backend
Online Mapping

VIO Cloud Backend
Bundle Adjustment

Visual Loop Closure

2D-3D feature matches

2D-2D feature matches

Initial state estimates

3D structure

2D-2D feature matches
Tango: Visual-Inertial Odometry (VIO)

[Hesch, Kottas, Bowman, Roumeliotis; T-RO '09] [Nerurkar, Wu, Roumeliotis; RSS ’13]
[Lynen, Sattler, Bosse, Hesch, Pollefeys, Siegwart; RSS ’15]

● Frontend: Sliding-Window Filter
  ○ Real-time w/ low computational complexity (10 Hz vision + 100 Hz IMU)
  ○ Accumulates drift over large distances (< 1% of distance travelled)

● On-device Backend: Online Bundle Adjustment with vision-based loop closure
  ○ Flexible raw + marginalized VI constraints
  ○ Real-time w/ < 5 cm RMSE on VICON

● Cloud Backend
  ○ MAP estimates with high accuracy (< 2 cm RMSE on VICON)
  ○ Offline w/ high computational complexity
Tango: 3D Reconstruction
[Klingensmith, Dryanovski, Srinivisa, Xiao; RSS ’15]

- Camera poses from online VIO or offline BA
- Fixed-resolution chunks of 16x16x16 voxels
- Spatial hashing
Tango: 3D Reconstruction

[Klingensmith, Dryanovski, Srinivisa, Xiao; RSS ’15]

- Timings on device:
  - Voxel updates: 128ms
  - Meshing: 102ms

- Default settings:
  - 3cm resolution
  - Raycasting
  - Carving
Tracking In Difficult Environments

- Time: 362.62
- FPS: 31.06
- Distance to origin:
- Position: 0.00, 0.12, 0.01
- Number of loop closures: 0
- Loop closure: 0.0 (0.7% of path)
- Open loop: 0.7 (102.7% of path)
- Number of keyframes: 160

Top-down view graph plot with grid size 1 m.
Conclusion

- Tango provides 6-DOF tracking and 3D reconstruction out-of-the-box
- Lightweight, small form factor, highly integrated
- Tango API for C++/Java/Unity
- Would this be useful for you or your PhD students in any of your projects?
- Free devices available, contact me after the talk
- Next steps: Improve 3D reconstruction, object detection & scene understanding
We are looking for academic partners and new colleagues!

Interns, Post-Docs, Visiting Faculty are very welcome.

Opportunities for funded research that help advance the platform.

Project Tango is located in Google offices in Mountain View, Zurich and Munich.

Contact us at: jsturm@google.com