Overview

- Keyframe-based SLAM
- 3D rendering for Augmented Reality
- Problems with traditional keyframe-based SLAM
- Solution: Deferred Triangulation SLAM
KeyFrame-based SLAM

2D Tracking [Optical Flow]
Stereo Initialization [Triangulation]
3D Mapping [Bundle Adjustment]
3D Tracking [Pose Estimation]
Rendering [Overlaying (AR)]
[Scene Reconstruction]

Time →

Tracking
Optical flow
Stereo triangulation
Bundle Adjustment
3D Tracking and pose estimation
Incremental mapping and camera pose refinement
Adding Keyframes, data association, and recovery
Rendering objects with the camera poses and geometry (map)

Mapping

Rendering
SimpleClothing - 8 - Male / Trenchcoat 800 - CPU, localspace
FPS = 29
DTAM
We have done Kinectfusion-type of processing using SoftKinetic range scanners, the quality and framerate of the depth is better than on Tango.
How to deal with the rotation?

**Live Tracking and Mapping from Both General and Rotation-Only Camera Motion**
Steflen Gauglitz*, Chris Sweeney*, Jonathan Ventura*, Matthew Turk*, Tobias Höllerer*
Department of Computer Science, University of California, Santa Barbara
ISMAR 2012

**Handling Pure Camera Rotation in Keyframe-Based SLAM**
Christian Pirchheim, Dieter Schmalstieg, Gerhard Reitmayr*
Graz University of Technology
ISMAR 2013
DT-SLAM: Deferred Triangulation for Robust SLAM

Daniel Herrera C.†, Kihwan Kim‡, Juho Kannala†, Kari Pulli‡, and Janne Heikkilä†

†University of Oulu ‡NVIDIA Research

3DV 2014
How to deal with the rotation?

- Deferred triangulation

0.5x Speed for visualization
- Deferred 2D points
- Triangulated 3D points
How to deal with the rotation?

- Deferred triangulation
- Jointly (2D/3D) constrain a pose

0.5x Speed for visualization
- Deferred 2D points
- Triangulated 3D points
How to overcome the rotation?

- Deferred triangulation
- Jointly (2D/3D) constrain a pose
- Region merging
Pose estimation

\[
E_{3D} = \| \phi \left( [R_k \mid t_k] \tilde{x} \right) - m_k \|^2.
\]

\[
\arg\min_{R_k, t_k} \sum_i \rho \left( E_{3D, i} \right)
\]
Epipolar segment

Infinite

Epipole

0.5

4 2 1
Epipolar segment
Pose estimation

\[ E_{3D} = \| \phi ([R_k | t_k] \tilde{x}) - m_k \|^2. \]

\[ \arg\min_{R_k, t_k} \sum_i \rho(E_{3D,i}) + \sum_j \rho(E_{2D,j}) \]
Bundle Adjustment

$$\arg\min_{R, T, X} \sum_{k \rightarrow K} \left( \sum_{i \rightarrow M} \rho(E_{3D,k,i}) + \sum_{j \rightarrow N} \rho(E_{2D,k,j}) \right)$$
Translate the camera slowly sideways, and press spacebar again to perform stereo init.
Quantitative evaluation

DT-SLAM

PTAM
Klein et al.

Trajectory
DT-SLAM

Trajectory
PTAM
x2 speed
Comparison with Hybrid SLAM and PTAM

Results from Hybrid SLAM and PTAM taken directly from Pirchheim et al. The footage contains captions from the original video.

DT-SLAM (Ours)
Another scene reconstruction example

Hand-held camera scene in City of Sights dataset x2 speed
Keyframe-based SLAM is efficient
- and can run in real time on mobile devices

But it has problems
- A separate initialization phase is annoying
- Breaking with pure rotations is a critical failure

Both can be addressed by
- tracking first in 2D
- deferring triangulation until there is enough baseline between the keyframes

Bonus: we plan to open source the implementation