Localization and Mapping of 3D interior space using Depth Camera

CS676: Project presentation

By
Anurag Sai
Sudhanshu Tamhankar

Problem Statement

- User holding and moving a standard Kinect camera should be able to rapidly create detailed 3D reconstructions of an indoor scene. [2]
- Two sub problems:
  - Track the trajectory of the camera
  - Reconstruct the surface
- Tracking is accomplished by just using the Depth images to find the necessary transformation between the camera and the global coordinate system.

Past Work and its drawbacks

- Previously before the use of Depth Cameras, localization in Visual SLAM was done using regular RGB cameras, in real time, but the mapping was very sparse and time consuming
- Many traditional techniques involve offline reconstruction
- In techniques like SfM, sparse scene feature detection is used for tracking.
- Already existing algorithm RGB-D SLAM[1]

Present day scenario

- Over the last decade, range images have grown in popularity and found increasing applications in fields including medical imaging (PET), object modelling, and robotics.
- In mobile robotics, the availability of range sensors capable of quickly capturing an entire 3Dscene has drastically improved the state of the art.
- Even in the challenging problems like autonomous driving, fast-scanning laser range sensors are used instead of cameras to perform obstacle avoidance, motion planning and mapping.

Our Approach

a) Acquiring Depth Map

b) Camera Tracking with ICP
(Testing Variants of ICP)[3]

c) Volumetric Integration[4]
(If time permits)

Dataset

- Depth images of a Kinect sensor[1]
- The ground-truth trajectory of the sensor.
- Data recorded at full frame rate (30 Hz) and sensor resolution (640×480).[1]
- Ground-truth trajectory obtained from a high-accuracy motion-capture system with eight high-speed tracking cameras (100 Hz).[1]

Current Status

Tracking the camera frame trajectory as the camera moves with respect to the ground frame

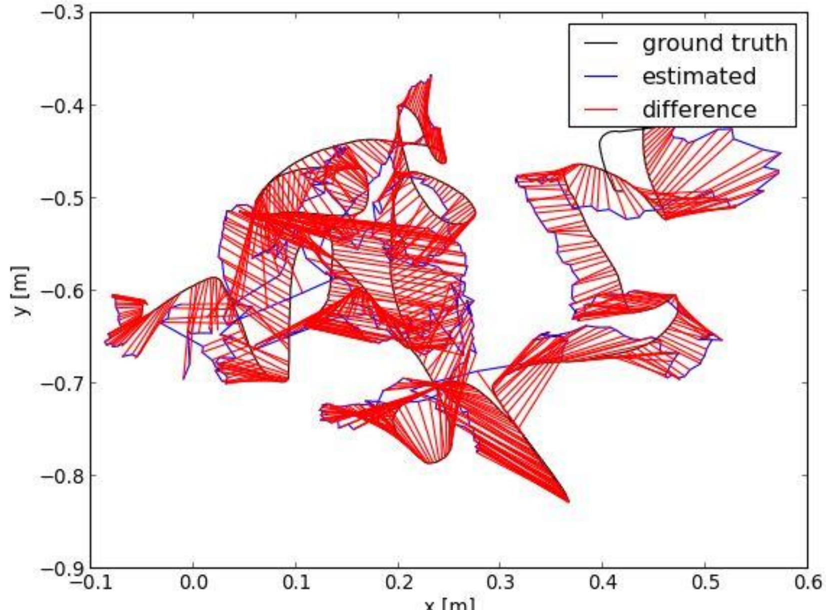
- http://www.youtube.com/watch?v=3hqBgHk6GtE&feature=youtu.be
  How the camera frame (green) and the camera frame predicted by RGBD-SLAM(green) move wrt the ground frame (blue)
- http://www.youtube.com/watch?v=c77Zt7-TZys&feature=youtu.be
  The actual RGB video for which the prediction has been done by RGBD-SLAM

Current Status

- Results for RGB-D SLAM

Output:

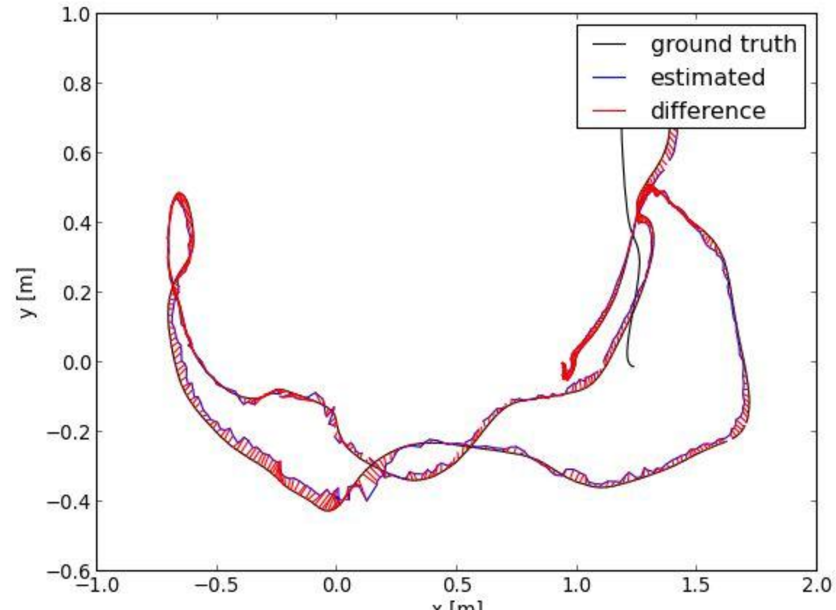
```
compared_pose_pairs 745 pairs
absolute_translational_error.rmse 0.070789 m
absolute_translational_error.mean 0.064573 m
absolute_translational_error.median 0.059535 m
absolute_translational_error.std 0.029006 m
absolute_translational_error.min 0.012246 m
absolute_translational_error.max 0.160127 m
```



firebug1_360

Output:

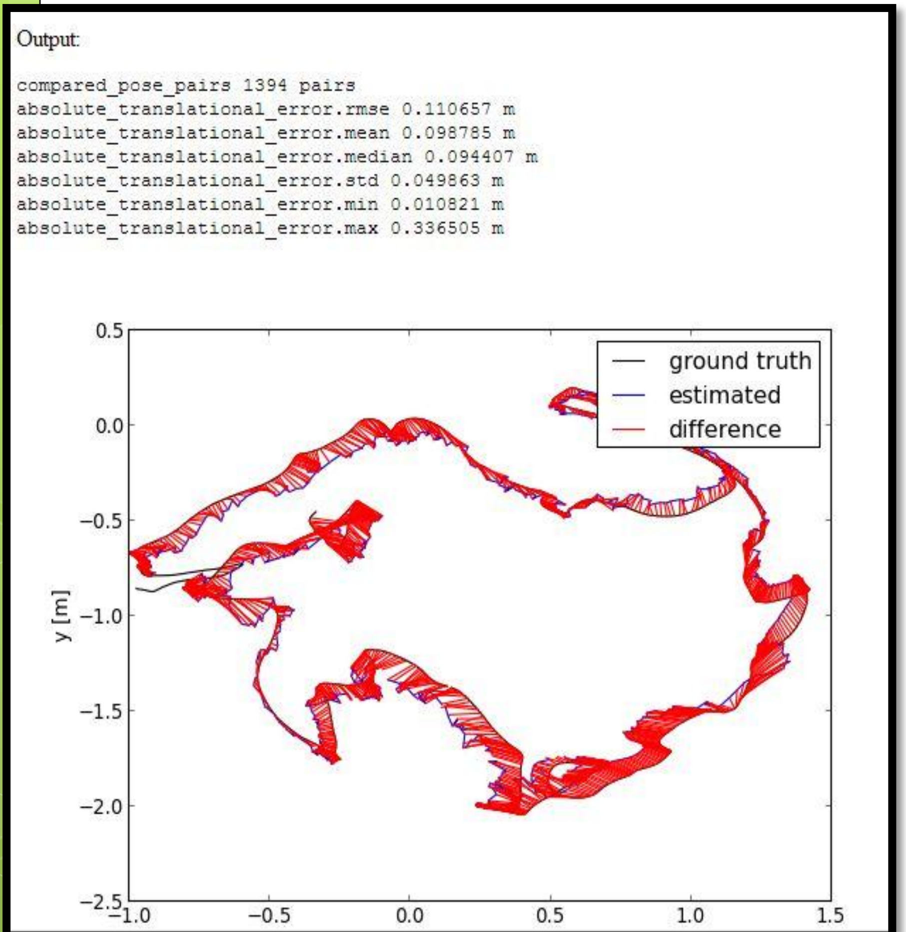
```
compared_pose_pairs 574 pairs
absolute_translational_error.rmse 0.025831 m
absolute_translational_error.mean 0.023132 m
absolute_translational_error.median 0.021388 m
absolute_translational_error.std 0.011497 m
absolute_translational_error.min 0.004203 m
absolute_translational_error.max 0.079256 m
```



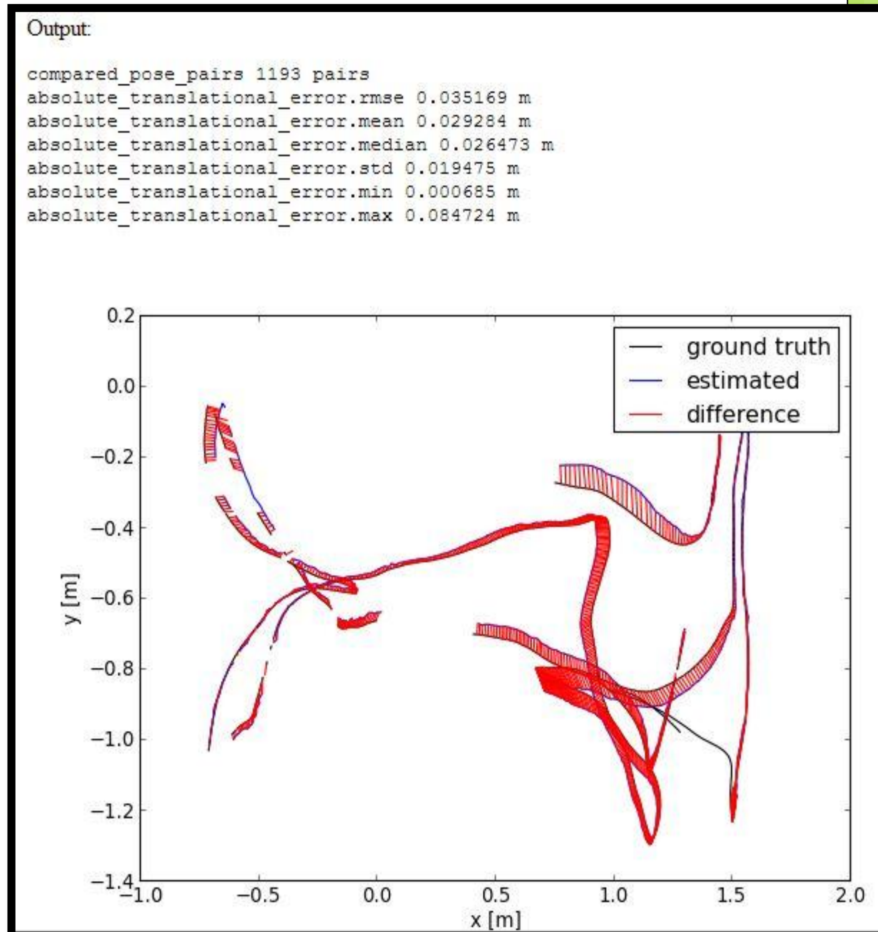
firebug1_desk

Current Status

- Results for RGB-D SLAM



firebug1_teddy



firebug1_floor

Over the next few weeks…

- Improving the performance of ICP
- Surface Representation using MeshLab or some other 3D representation software
- If time permits, Volumetric Integration to cleanup the surface produced.

References

- [1] http://cvpr.in.tum.de/data/datasets/rgbd-dataset
- **[2] "KinectFusion: Realtime 3D Reconstruction and Interaction Using a Moving Depth Camera".2011**
  -Shahram Izadi et al
- **[3] "Efficient Variants of the ICP Algorithm".2001**
  Szymon Rusinkiewicz and Marc Levoy.
- **[4] "A Volumetric Method for Building Complex Models from Range Images".1996**
  Brian Curless and Marc Levoy.