2018 December Reading Reports

3D Modeling on the Go: Interactive 3D Reconstruction of Large-Scale Scenes on Mobile Devices

This paper presented and evaluated a system for large-scale 3D reconstruction of outdoor scenes that runs at interactive frame rates on modern mobile devices. It used monochrome fisheye images. The authors showed the importance of filtering outliers in the depth maps and propose multiple filtering steps.

The trajectory is given and the depth estimation and uncertainty is calculated by a cost function (one image). Then the estimation is updated by the consecutive images, followed by several filters.

The filtered depth maps are integrated into a volume storing a TSDF using the **voxel hashing**.

More details are shown in [1] (plane-sweeping stereo) and [2] (voxel hashing).

[1] Häne, C., Heng, L., Lee, G. H., Sizov, A., & Pollefeys, M. (2014, December). Real-time direct dense matching on fisheye images using plane-sweeping stereo. In *3D Vision (3DV), 2014 2nd International Conference on* (Vol. 1, pp. 57-64). IEEE.

[2] Klingensmith, M., Dryanovski, I., Srinivasa, S., & Xiao, J. (2015, July). Chisel: Real Time Large Scale 3D Reconstruction Onboard a Mobile Device using Spatially Hashed Signed Distance Fields. In *Robotics: science and systems* (Vol. 4).

Real-Time Direct Dense Matching on Fisheye Images Using Plane-Sweeping Stereo

This paper introduced a method to use fisheye images for plane-sweeping stereo. It contains camera projection model (MEI or FOV), homography (with ZNCC matching scores), cost aggregation and depth extraction and sub-pixel interpolation.

It is slightly slow but have larger coverage.

No pretty sure about **sub-pixel interpolation** mentioned in the paper.

Real-Time Plane-Sweeping Stereo with Multiple Sweeping Directions

It presented a real-time multi-view stereo algorithm based on plane-sweeping which correctly handles slanted surfaces. Identifying sweeping directions (by uniform sampling 2D points), plane selection (range), incorporating plane priors (SfM), and selecting depths from multiple sweeps (containing the matching cost, surface normal smoothness, and depth smoothness, **three-label graph cut**).

ElasticFusion: Dense SLAM Without A Pose Graph

(i) use photometric and geometric **frame-to-model** predictive tracking in a fused surfel-based dense map;

(ii) perform dense **model-to-model** local surface loop closures with a **non-rigid space deformation** and

(iii) utilise a **predicted surface appearance-based place recognition method** to resolve global surface **loop closures** and hence capture globally consistent dense surfel-based maps ***without a pose graph***.

This paper uses RGB-D camera to build consistent map. It contains four parts, fused predicted tracking (geometric pose estimate + (added, full color) photometric pose estimation), deformation graph as [1] (added a pin item, time related), local loop closure and global loop closure [2] (fern encoded frame).

[1] R. W. Sumner, J. Schmid, and M. Pauly. Embedded deformation for shape manipulation. In Proceedings of SIGGRAPH, 2007.

[2] RealTime RGB-D Camera Relocalization via Randomized Ferns for Keyframe Encoding. IEEE Transactions on Visualization and Computer Graphics, 21(5):571–583, 2015.