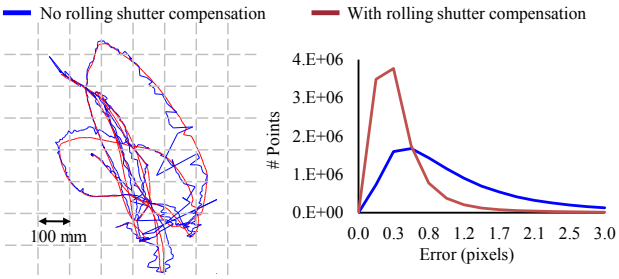
Spatiotemporal Bundle Adjustment for Dynamic 3D Reconstruction*
(Supplementary Material)

Minh Vo Srinivasa G. Narasimhan Yaser Sheikh
Carnegie Mellon University
{*mpvo*, *srinivas*, *yaser*}@cs.cmu.edu



(a) Ground projection of the camera trajectory (b) Reprojection error

Figure 1: Analysis of the spatial camera calibration for the Checkerboard sequence with different camera models.

1. Data Preprocessing

3D Corpus and Initial Camera Pose Estimation: To reduce drift in camera pose estimation, we use the key frames sampled from each video to build the 3D corpus of the scene using SfM algorithm [2]. We register other frames to this corpus using the r6P algorithm and refine their parameters using the Cayley transform model [1]. Fig. 1 shows the importance of modeling the rolling shutter effect for smartphone and action cameras.

Tracking: We detect and match SIFT features for all images at multiple time instances. We discard matches with low gradient score and track the remaining points both forward and backward in time using affine template matching. The backward-forward consistency check is used to discard erroneous optical flow during tracking. Finally, we check for the appearance consistency between patches of the first and the last frame using Normalized Cross Correlation and remove the entire trajectory if the score is below 0.8.

Trajectory classification: We exploit the fact that triangulation based methods work for static points but produce large errors for dynamic points in order to identify 2D trajectories of dynamic points. This is done using these two heuristics: (1) the reprojection error of a static point should be small regardless of which camera frame it is triangulated from. We randomly sample frames along the 2D trajectory to triangulate and consider the 2D trajectory as belonging to a static point if the reprojection threshold is smaller than 3 pixels for more than 80% of the time. (2) the reprojection error of a dynamic point forms a steep valley as the time offset passes by its true value. We reject any set of trajectories as belonging to a dynamic point if the minimum of the cost valley is not smaller than 80% of the average cost.

*http://www.cs.cmu.edu/ ILIM/projects/IM/STBA/

References

[1] C. Albl, Z. Kukelova, and T. Pajdla. R6p-rolling shutter absolute pose problem. In *CVPR*, 2015.

[2] C. Wu. Visualsfm: A visual structure from motion system. http://http://ccwu.me/vsfm, 2011.