Computer Processing of Pictorial Information Project 3

Huai-Jen Liang and Wei-An Lin
Department of Electrical and Computer Engineering
University of Maryland
College Park, Maryland 20740
Email: r81810g.umd@gmail.com, walin@terpmail.umd.edu

I. Fundamental matrix, Essential Matrix, and Camera Pose Estimation

We apply the procedure stated in the project description to estimate the fundamental matrix F and essential matrix E. When estimating the fundamental matrix, feature points are first processed to have zero mean and root-mean-square distance equal to $\sqrt{2}$. The computed F is then re-mapped to the original scale. This approach is known to be robust to noise. RANSAC is applied for each pair of images to choose appropriate feature matches. For the first two scenes, we extract four possible configurations of camera poses, and choose the one that satisfies the cheirality condition.

II. Triangulation

Given match points for the first two scenes, we first perform linear triangulation to estimate the corresponding 3D points in the world coordinate. The estimated 3D locations are shown in Fig. 1. It is clear that linear triangulation provides good initialization. This is because normalization is first performed in our implementation. Fig. 2 shows the reprojected results of linear and nonlinear triangulation for the second scene. Using the result from linear triangulation as initial estimation, we can minimize the geometric error by using Levenberg-Marquardt solver[1]. The result of nonlinear triangulation is shown in Fig. 2b. One can observe that the reprojected features in Fig. 2b have reduced error compared to Fig. 2a.

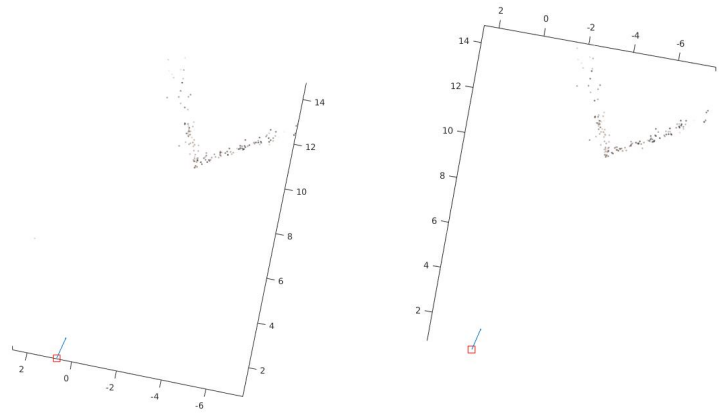
III. Register New Image

In order to choose a new image to register in our 3D-scene, we utilize greedy search to find the best image. The best image is defined as having the most features that are:

1) Visible in both this image and any registered images
2) Have been reconstructed in 3D scene

IV. Perspective-n-Points

Given reconstructed 3D points, we can estimate camera pose from $3D \leftrightarrow 2D$ correspondences by solving a linear least square optimization. Using the result from linearPnP as an initial estimation, we can minimize the geometric error by using Levenberg-Marquardt solver [1]. The result of linearPnP and nonlinearPnP for different scenes are shown in Fig. 3, Fig. 4, Fig. 5, and Fig. 6.



(a) Linear Triangulation (b) Nonlinear Triangulation

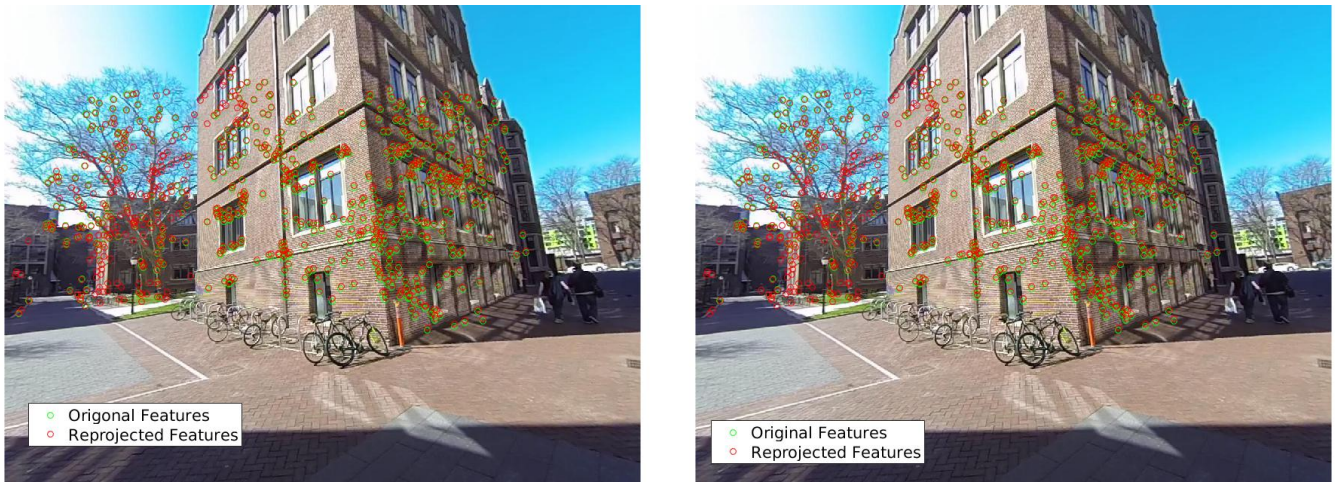
Fig. 1: Linear and nonlinear triangulation. The reprojection results are shown for scene 2.

V. Reconstruct New 3D Points

Once we have the camera pose, we can follow Sec. II by matching with new images and reconstruct additional 3D points in the scene.

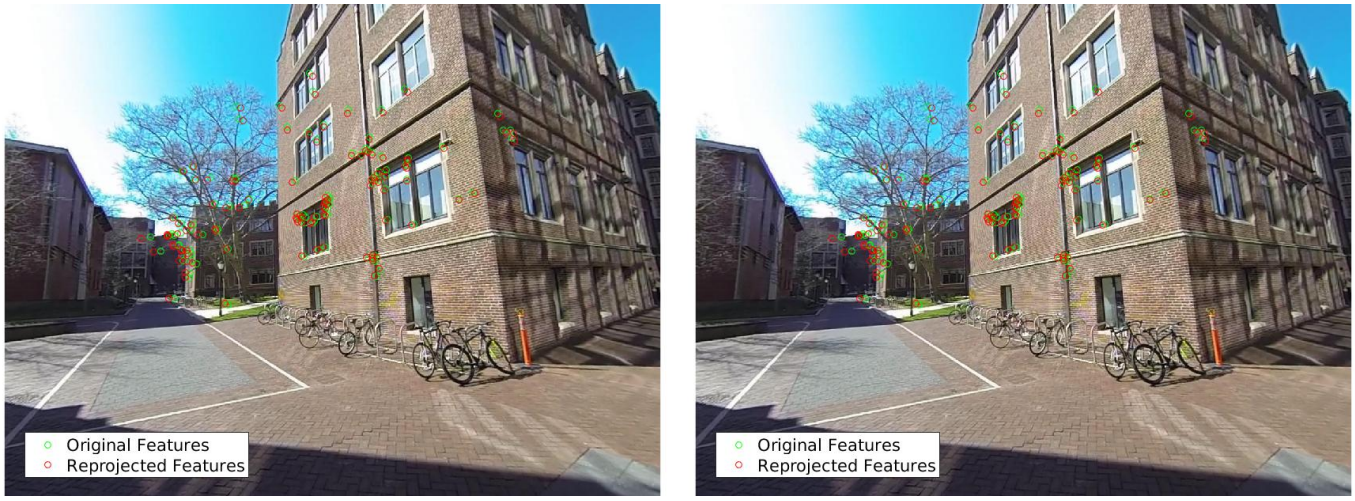
VI. Bundle Adjustment

Using camera poses and 3D points from previous sections as initial estimation, we can run a final bundle adjustment. We refine every poses and 3D points simultaneously by minimizing projection error. This will be extremely slow if we directly solve nonlinear optimization. However, since not every 3D points are visible to every image, we introduce Sparse Bundle Adjustment[2], which exploiting sparsity to efficiently solve the nonlinear optimization. The 3D reconstruction with and without bundle adjustment are shown in Fig. 7 and Fig. 8 respectively. It can be observed that the right side of Fig. 7 is wide. After applying bundle adjustment, the top view of Fig. 8 shows that the right wall has improved quality. The left side of the figure contains more points than Fig. 7. Note that we display all the 3D points after bundle adjustment. We do not remove the points with low reconstruction errors.



(a) Linear Triangulation (b) Nonlinear Triangulation

Fig. 2: Linear and nonlinear triangulation. The reprojection results are shown for scene 2.



(a) Linear Triangulation (b) Nonlinear Triangulation

Fig. 3: Linear and nonlinear PnP for scene 3.

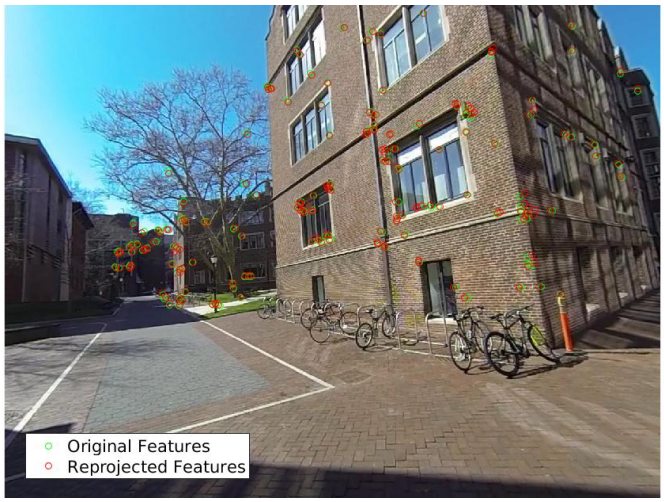
REFERENCES

[1] Donald W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Mathematics*, 1963.

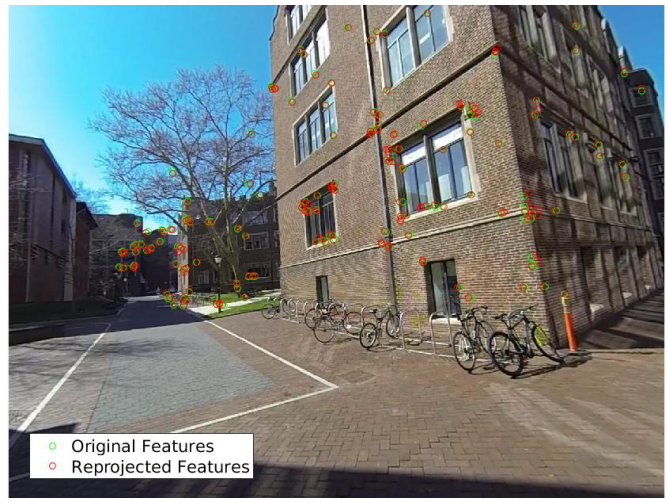
[2] M.I. A. Lourakis and A.A. Argyros. SBA: A Software Package for Generic Sparse Bundle Adjustment. *ACM Trans. Math. Software*, 2009.

[3] C. Wu. Towards linear-time incremental structure from motion. In *2013 International Conference on 3D Vision - 3DV 2013*, June 2013.

[4] C. Wu, S. Agarwal, B. Curless, and S. M. Seitz. Multicore bundle adjustment. In *CVPR 2011*, June 2011.

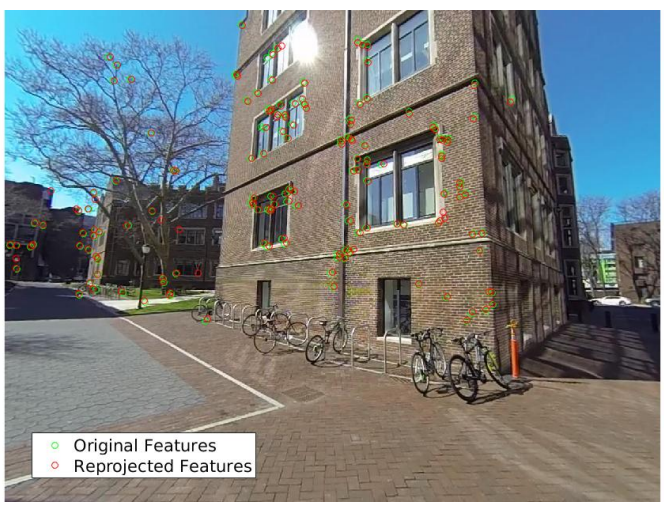


(a) Linear Triangulation

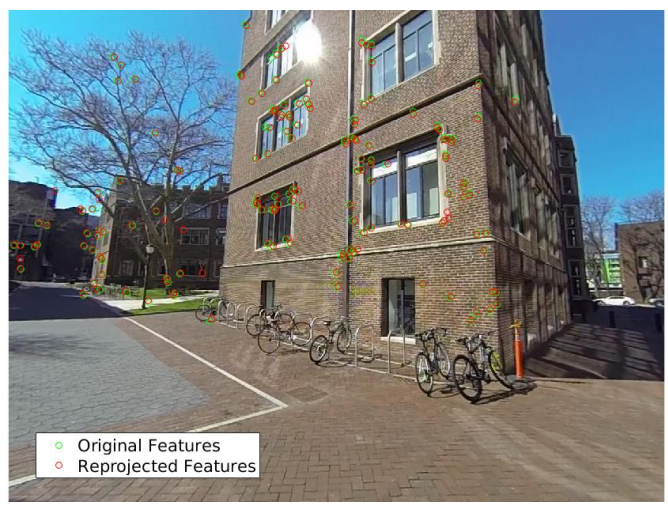


(b) Nonlinear Triangulation

Fig. 4: Linear and nonlinear PnP for scene 4.



(a) Linear Triangulation

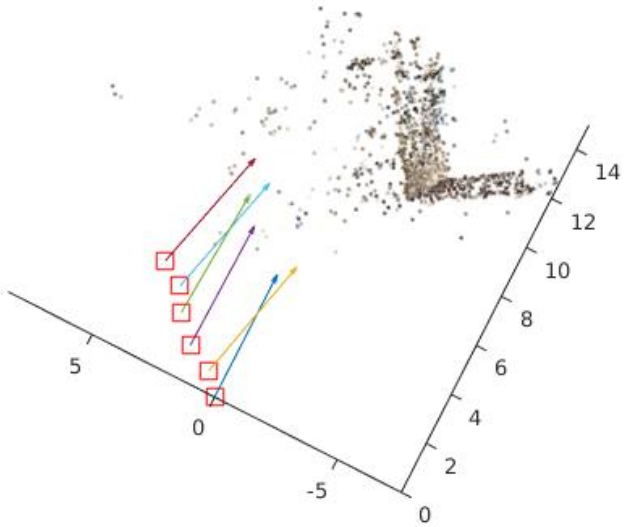


(b) Nonlinear Triangulation

Fig. 5: Linear and nonlinear PnP for scene 5.

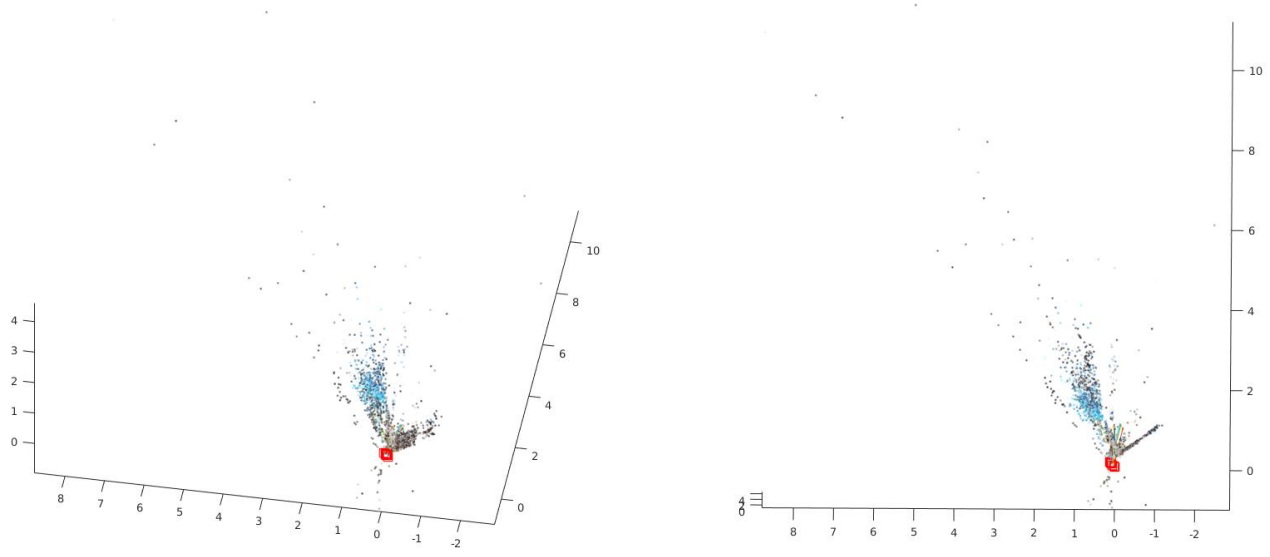
(a) Linear Triangulation (b) Nonlinear Triangulation

Fig. 6: Linear and nonlinear PnP for scene 6.



(a) Top view

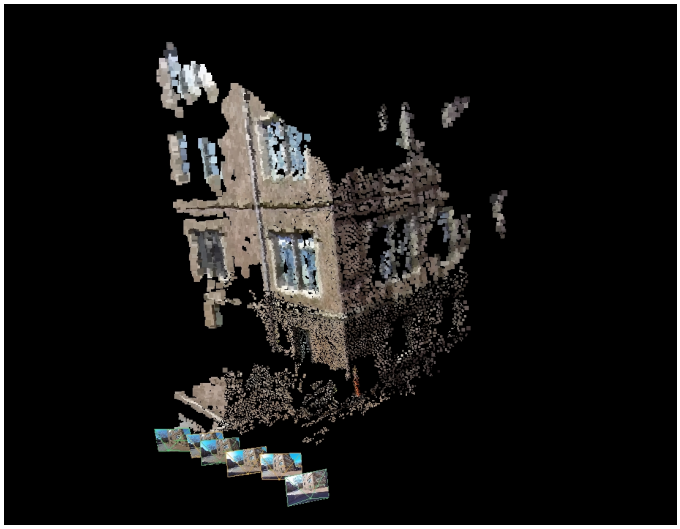
Fig. 7: 3D reconstruction without bundle adjustment. Red boxes represent camera locations. Orientations for each camera are represented using arrows.



(a) Profile view

(b) Top view

Fig. 8: 3D reconstruction with bundle adjustment. Red boxes represent camera locations. Orientations for each camera are represented using arrows.



(a) Dense Reconstruction

Fig. 9: 3D reconstruction using VisualSFM[3], [4]