CSfM: Community-based Structure-from-Motion

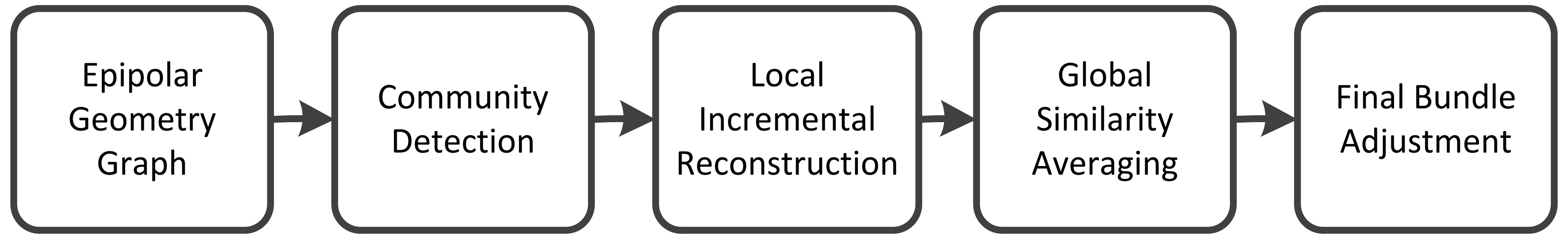
Hainan Cui, Shuhan Shen, Xiang Gao, Zhanyi Hu
NLPR, Institute of Automation, Chinese Academy of Sciences

IEEE International Conference on Image Processing
ICIP 2017

Motivation

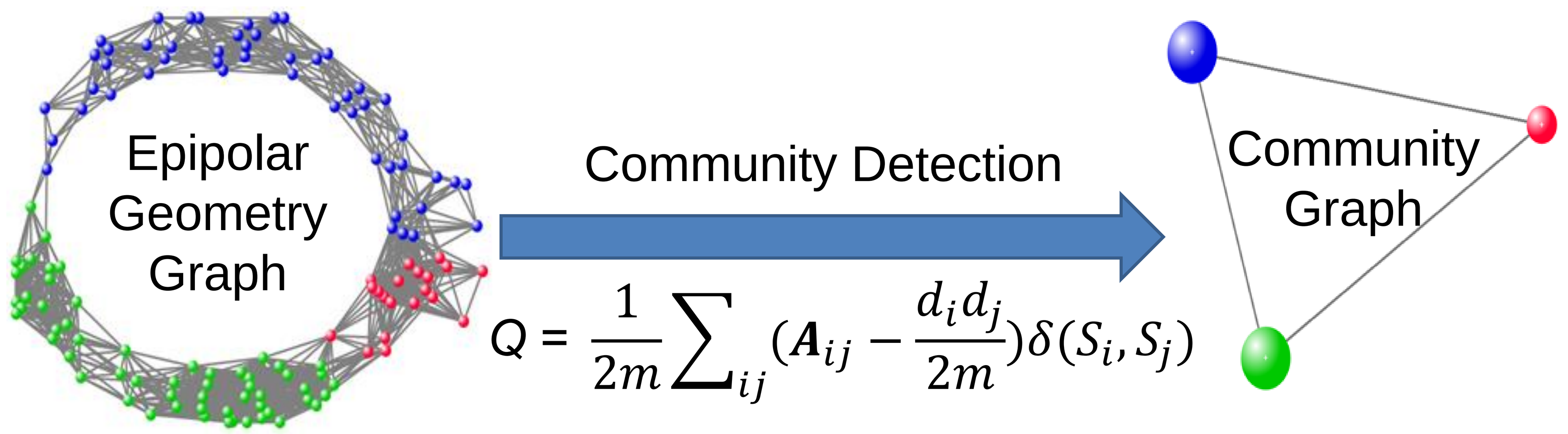
Incremental SfM has advanced in both accuracy and robustness, while it suffers from error accumulation and poor efficiency for large-scale reconstruction. Global SfM is efficient, but is sensitive to the image feature match outliers.

Proposed CSfM pipeline



CSfM inherits the advantages from both manners:
(1) Local incremental reconstruction is robust to feature match outliers;
(2) Global similarity averaging tackles the error accumulation problem.

Community Detection



$$Q = \frac{1}{2m}\sum_{ij}\left(A_{ij} - \frac{d_i d_j}{2m}\right)\delta(S_i, S_j)$$

Clustering the epipolar geometry graph into a community graph:
(1) Each node corresponds to a community;
(2) An edge links two communities when some epipolar edges exist between them.

Global Similarity Averaging

Given the scale factor s_i for each community and the relative scale s_{ij} between two connected communities:

$$s_{ij} = s_j \,/\, s_i \qquad \log(s_{ij}) = \log(s_j) - \log(s_i)$$

Given the rotation $\boldsymbol{R}_i$ for each community and relative rotations $\boldsymbol{R}_{ij}$ in the similarity transformation:

$$\boldsymbol{R}_{ij} = \boldsymbol{R}_j\boldsymbol{R}_i^T \qquad \boldsymbol{w}_{ij} = \boldsymbol{w}_j - \boldsymbol{w}_i$$

Given the translation $\boldsymbol{T}_i$ for each community and relative translations $\boldsymbol{T}_{ij}$ in the similarity transformation:

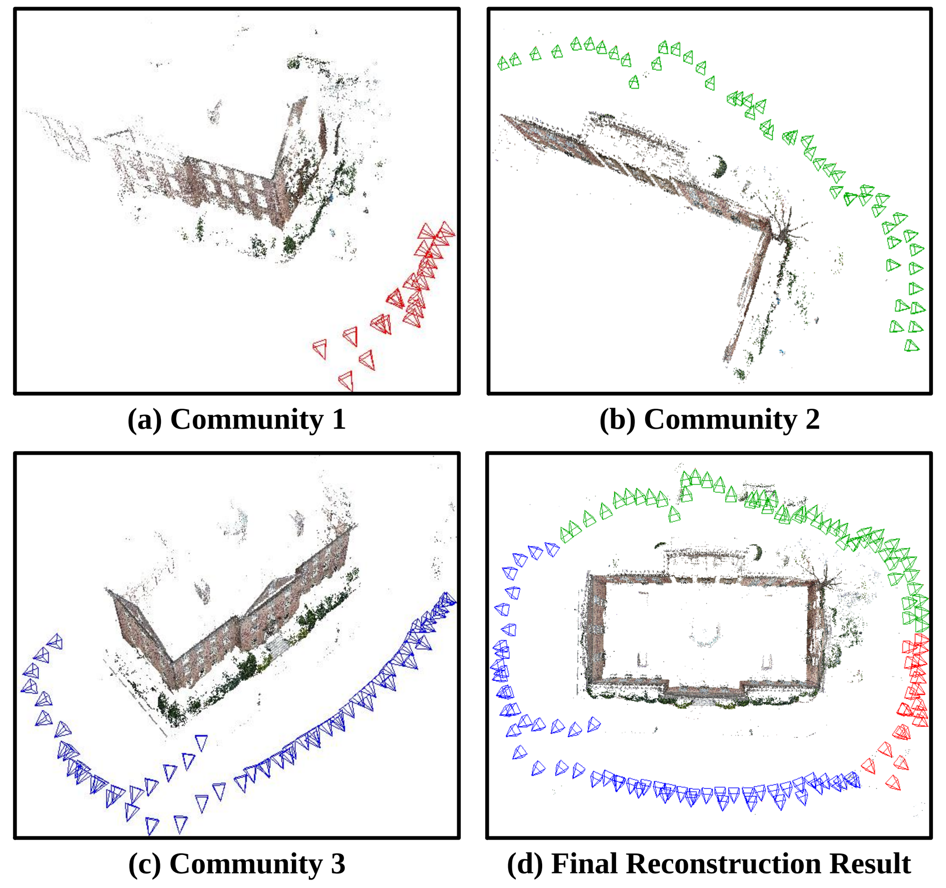
$$\boldsymbol{T}_{ij} = \boldsymbol{T}_j - \boldsymbol{T}_i$$

All these problems are convex L1 problems, and could be solved by [1].

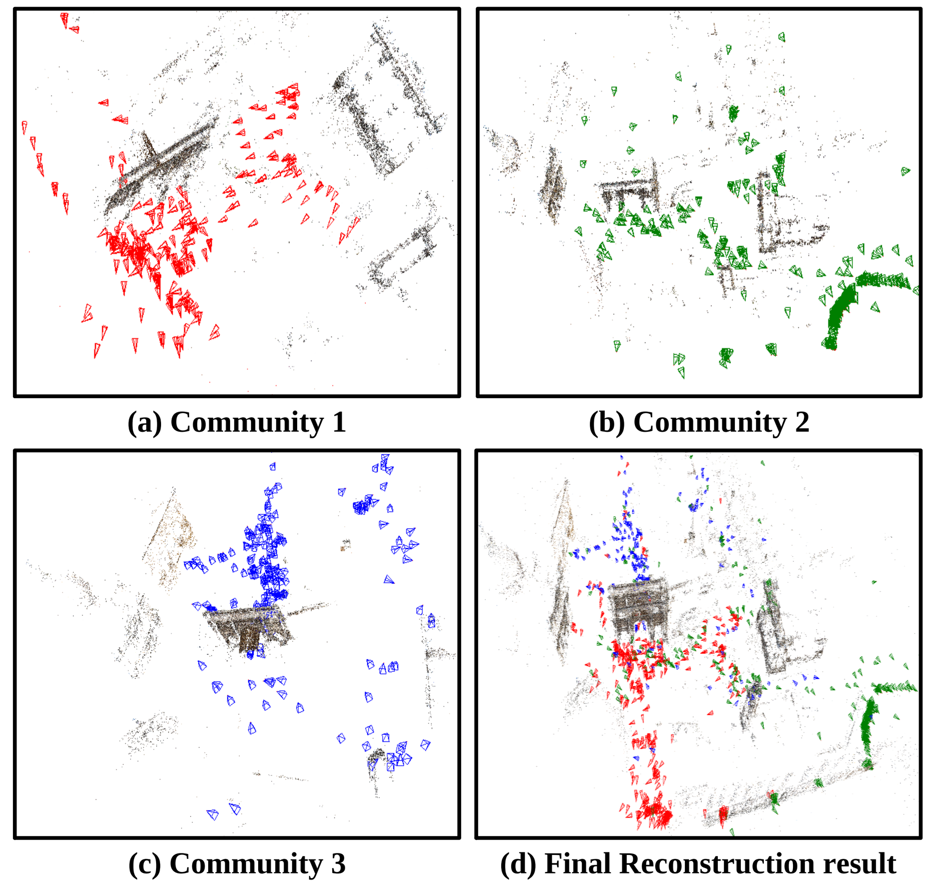
To mitigate the impact of feature match outliers, we perform the final bundle adjustment (BA) after global similarity averaging.

$$\min_{\boldsymbol{K}_i,\boldsymbol{R}_i,\boldsymbol{T}_i,\boldsymbol{X}_j} \sum_{i=1}^{N}\sum_{j=1}^{M} \delta_{ij}\left\|\boldsymbol{x}_{ij} - \gamma(\boldsymbol{K}_i,\boldsymbol{R}_i,\boldsymbol{T}_i,\boldsymbol{X}_j)\right\|_{huber}$$

Illustrative reconstruction process of CSfM system

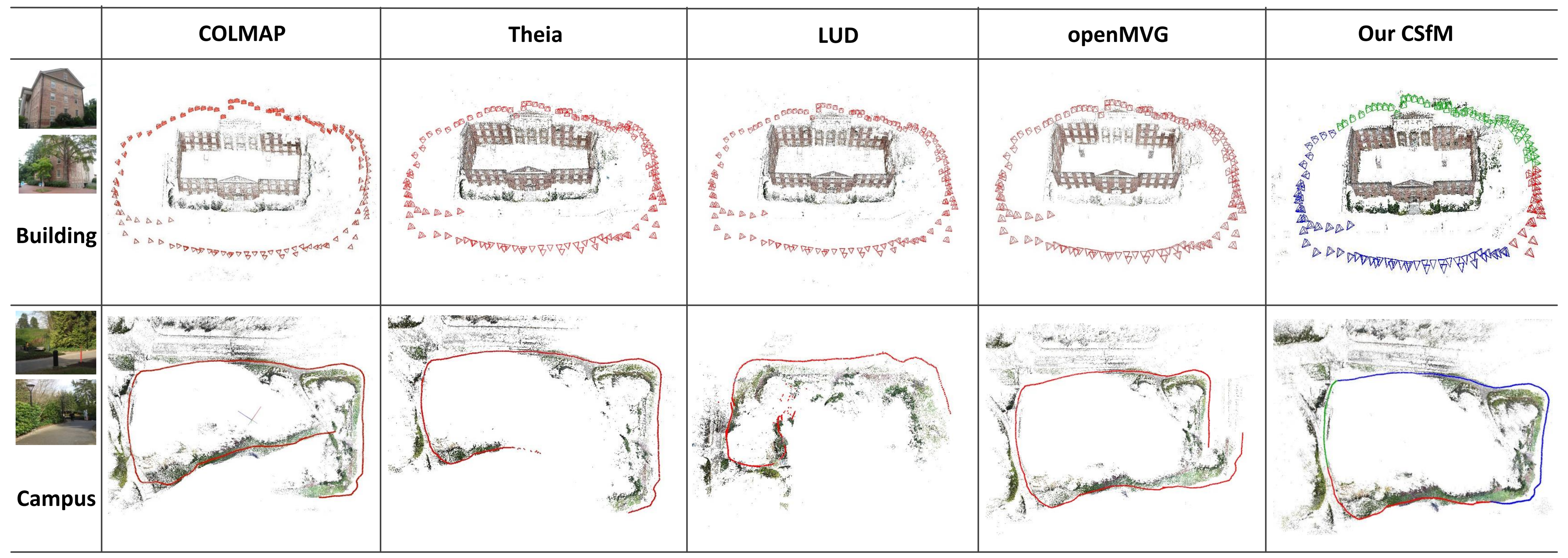


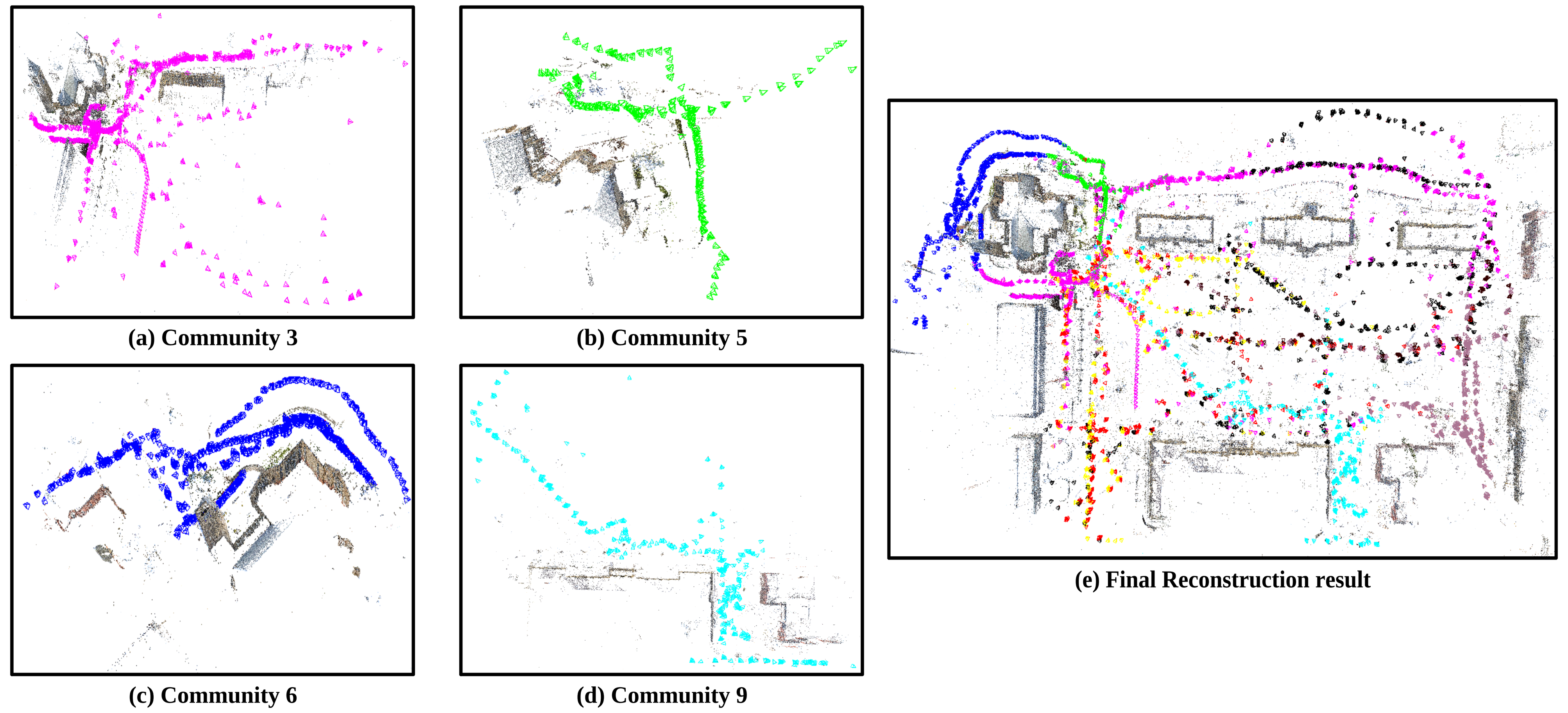
(a) Community 1 (b) Community 2 (c) Community 3 (d) Final Reconstruction Result



(a) Community 1 (b) Community 2 (c) Community 3 (d) Final Reconstruction result

Extensive Experiments





(a) Community 3 (b) Community 5 (c) Community 6 (d) Community 9 (e) Final Reconstruction result

For the dataset Quad [3], with a similar accuracy: DISCO [3] 1.16m, COLMAP [2] 0.85m, Ours 0.83m. The efficiency of our CSfM is 5 times faster than COLMAP [2], and 10 times faster than DISCO [3].

References

[1] Emmanuel Candes, et al. “l1-magic: Recovery of sparse signals via convex pro-gramming” www.acm.caltech.edu/l1magic/downloads/l1magic.pdf, vol. 4, pp. 46, 2005.
[2] Johannes Schonberger, et al. “Structure-from-motion revisited” CVPR2016
[3] D. J. Crandall, et al. “SfM with MRFs: Discrete-continuous optimization for large-scale structure from motion” IEEE TPAMI2013.