Supplemental Material (2): Experiments on Real Images

Chi Xu, Tingrui Guo, Yuan Huang, Li Cheng, Senior Member, IEEE

Abstract

It is the supplemental material of paper "A Hough Voting based 2-Point RANSAC Solution to the Perspective-n-Point Problem". Detailed information of the real image experiment is provided in this document.

1 Overview

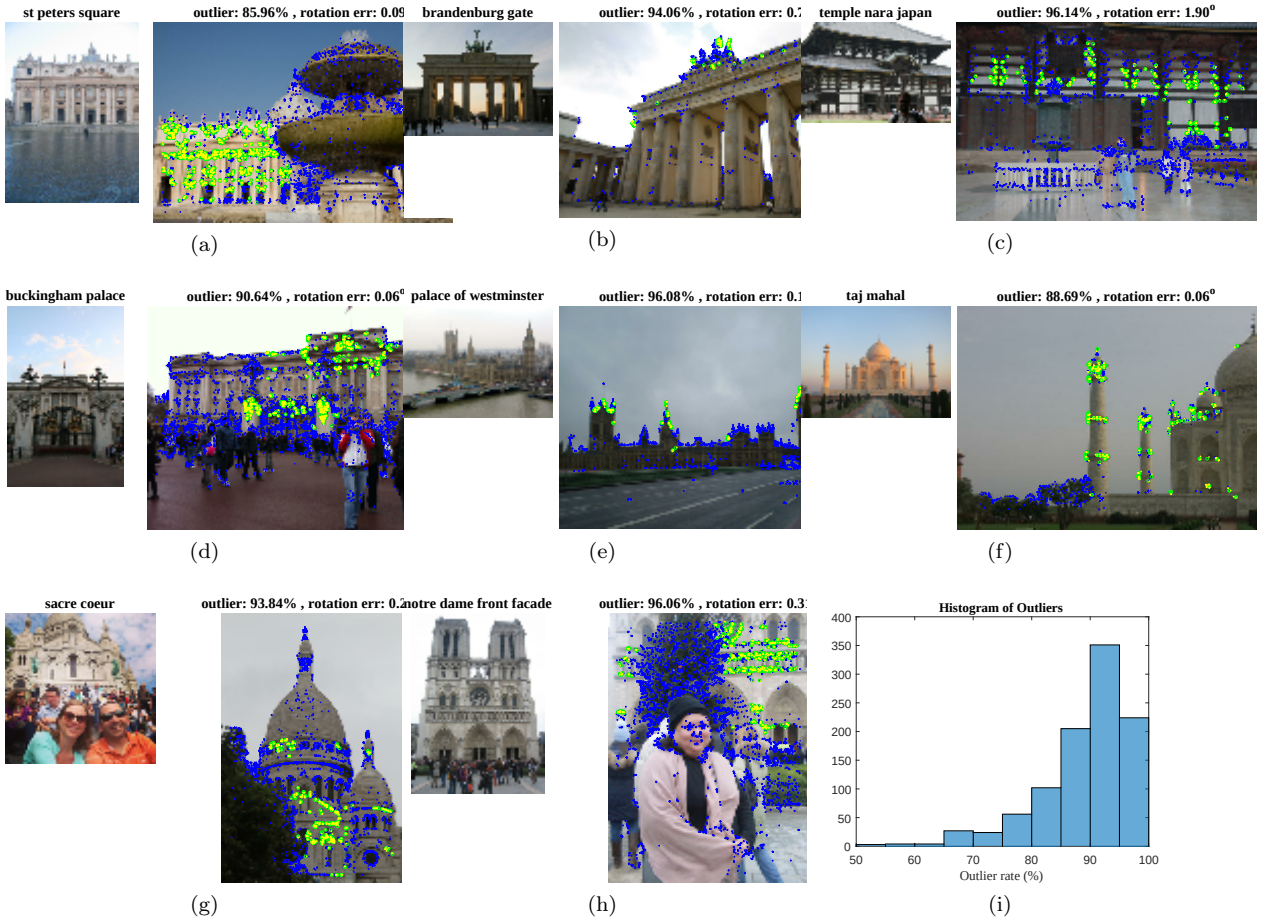


Figure 1: Exemplar pose estimation results of our solution, obtained from challenging real-life environment with high outlier rate. Panels (a) to (h) are examples taken from wide-baseline image matching dataset [1]. In each panel, on the left shows the reference image, and on the right is the target image. The feature points are extracted by ORB. Outlier rate and estimated rotation error of our solution are shown on top of each target image. Blue dots denote 2D points matched, green dots denote inliers detected, and yellow dots denote re-projection of the inliers using the estimated pose. In the experiment, we randomly pick pairs of reference and target images from [1], and the histogram of the outlier rate is shown in (i).

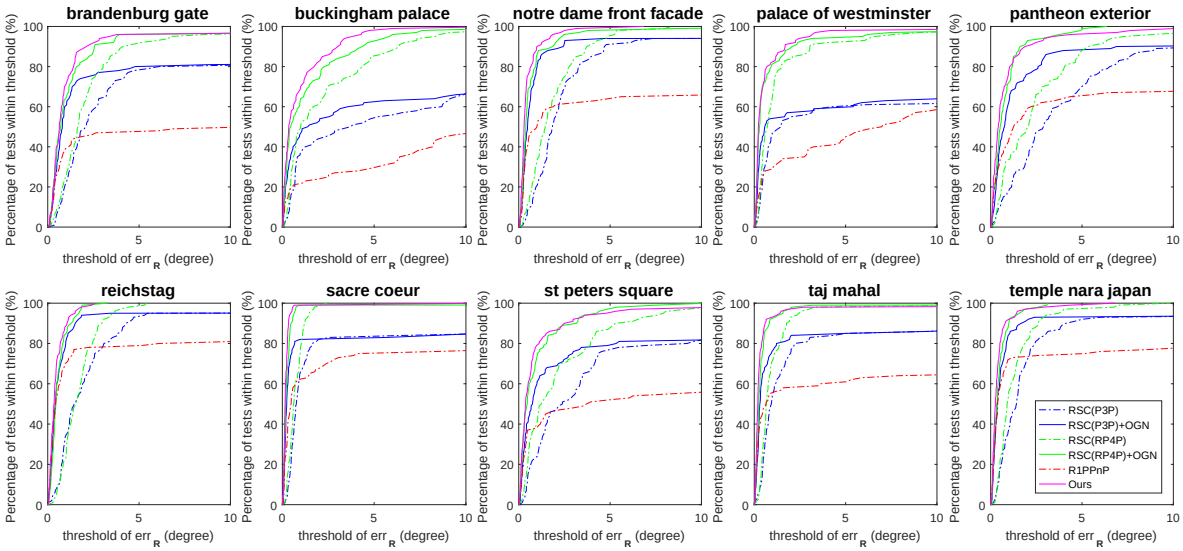


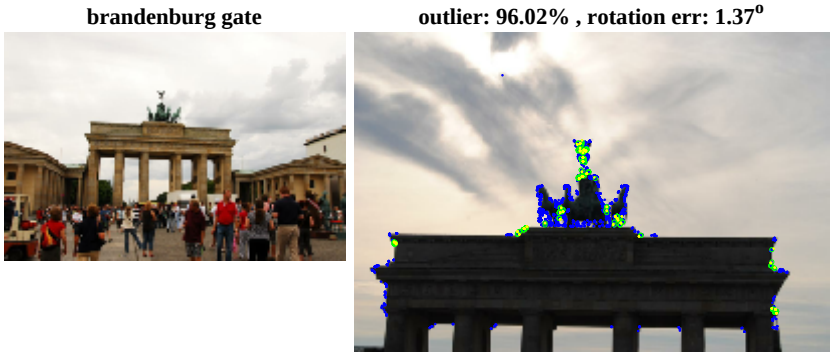
Figure 2: Results of each evaluated scene in real dataset [1].

The images are taken from wide-baseline image matching dataset [1]. Ten scenes are evaluated, which are "brandenburg gate", "buckingham palace", "notre dame front facade", "palace of westminster", "pantheon exterior", "reichstag", "sacre coeur", "st peters square", "taj mahal", and "temple nara japan". The visual examples are shown in Fig. 1. The experimental results of the compared methods on ten scenes are shown in and 2.

2 Visual Examples

In each figure, on the left shows the reference image, and on the right is the query image. The feature points are extracted using ORB. Outlier rate and estimated rotation error of our solution are shown on top of each target image. **Blue dots** denote 2D points matched, **green dots** denote inliers detected, and **yellow dots** denote re-projection of the inliers using the estimated pose. The exemplar estimation results of our solution are list as follows:

2.1 Brandenburg Gate



brandenburg gate

outlier: 79.04% , rotation err: 0.92°

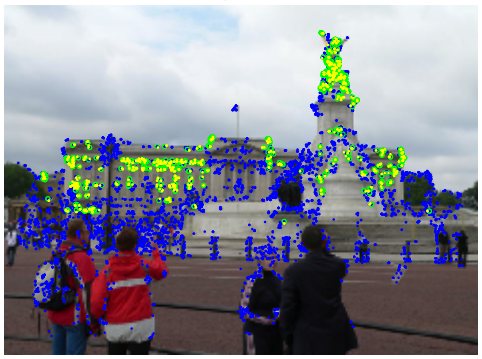
2.2 Buckingham Palace

buckingham palace

outlier: 94.48% , rotation err: 0.23°

buckingham palace

outlier: 90.41% , rotation err: 0.08°

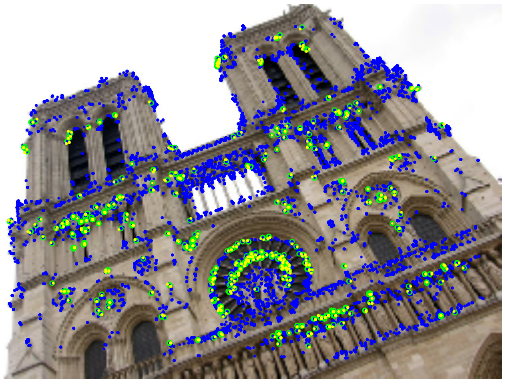


2.3 Notre Dame Front Facade

notre dame front facade

outlier: 89.24% , rotation err: 0.11°

notre dame front facade **outlier: 96.57% , rotation err: 0.13^o**

2.4 Palace of Westminster

palace of westminster **outlier: 95.79% , rotation err: 1.18^o**


palace of westminster **outlier: 96.50% , rotation err: 1.17^o**

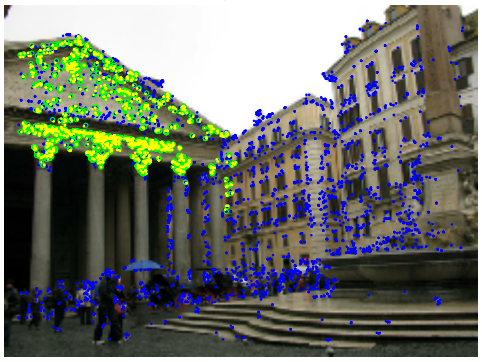

palace of westminster **outlier: 93.43% , rotation err: 0.23^o**


palace of westminster **outlier: 87.43% , rotation err: 0.86°**

2.5 Pantheon Exterior

pantheon exterior **outlier: 95.49% , rotation err: 0.12°**

pantheon exterior **outlier: 77.28% , rotation err: 0.06°**



2.6 Reichstag

reichstag **outlier: 95.92% , rotation err: 0.55°**

buckingham palace

outlier: 96.39% , rotation err: 6.03^{o}

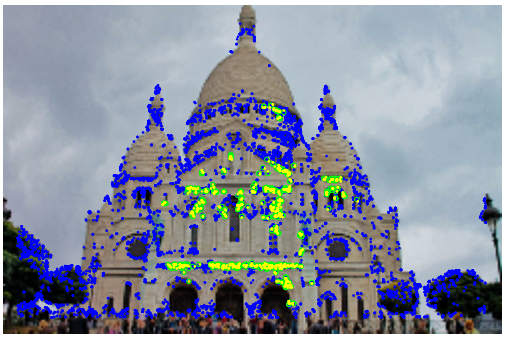
2.7 Sacre Coeur

sacre coeur

outlier: 94.73% , rotation err: 0.06^{o}

sacre coeur

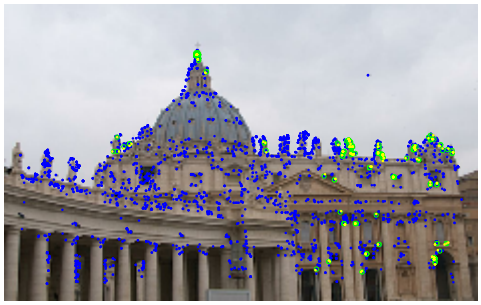
outlier: 93.25% , rotation err: 0.11^{o}



2.8 St Peters Square

st peters square

outlier: 95.96% , rotation err: 0.20^{o}



st peters square

outlier: 95.51% , rotation err: 0.32^{o}

2.9 Taj Mahal

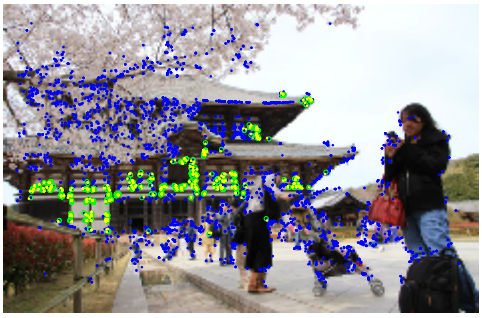
taj mahal

outlier: 95.06% , rotation err: 0.28^{o}

2.10 Temple Nara Japan

temple nara japan

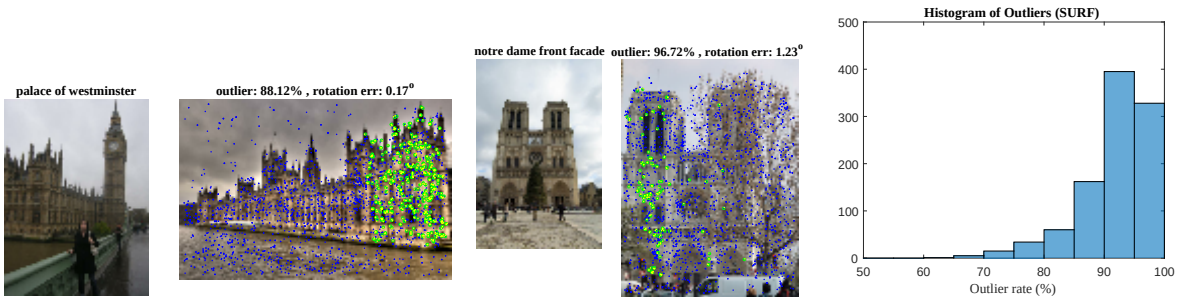
outlier: 92.34% , rotation err: 0.63^{o}



3 Failure Cases

4 SURF and SIFT features

The examples of our estimation results using SURF feature are provided in Fig. 3 (a)-(b), and the histogram of outlier rate using SURF feature is shown in Fig. 3 (c). The estimated rotation errors of all ten scenes are shown in Fig. **??**. Similarly, the corresponding results of using SIFT feature are provided in Fig. 5 and Fig. **??**.



(a) (b) (c)

Figure 3: Results of our solution using SURF feature.

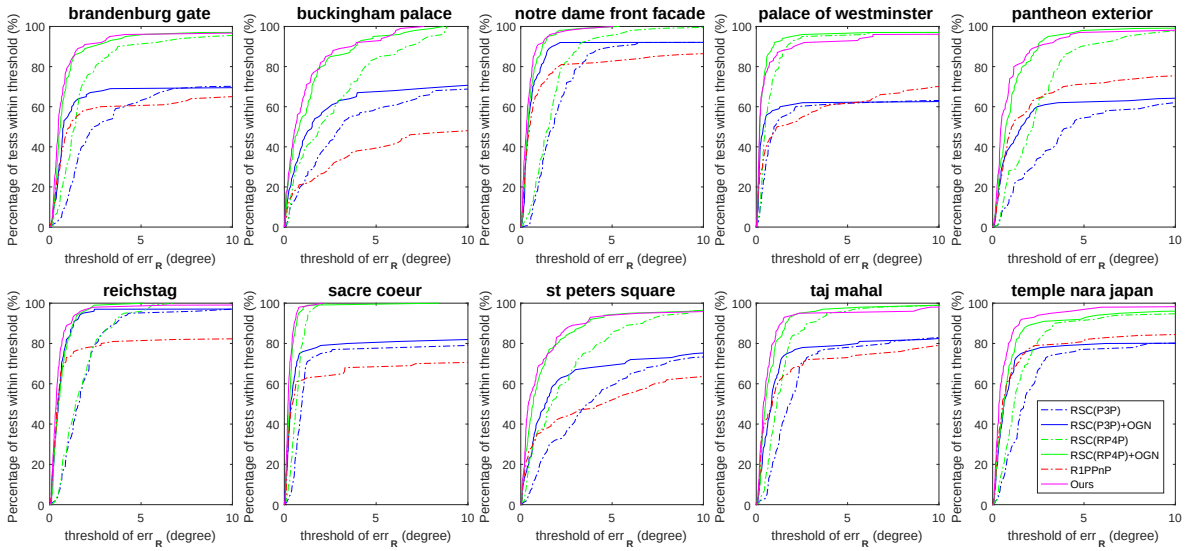
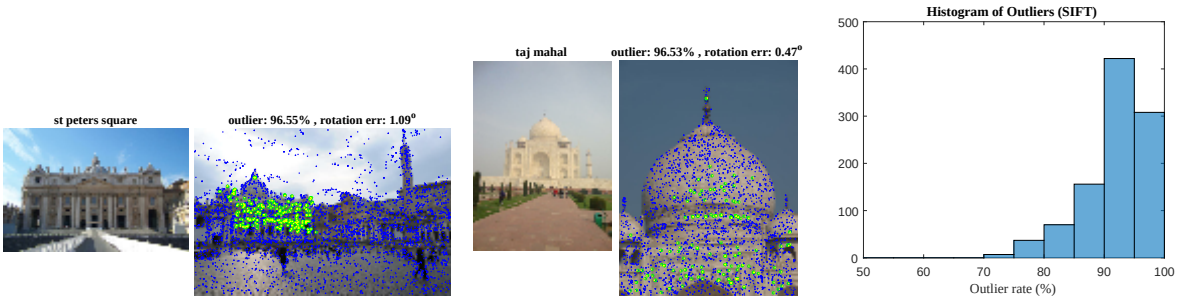


Figure 4: Results of each evaluated scene in real dataset [1]. SURF is used for feature extraction.



(a) (b) (c)

Figure 5: Results of our solution using SIFT feature.

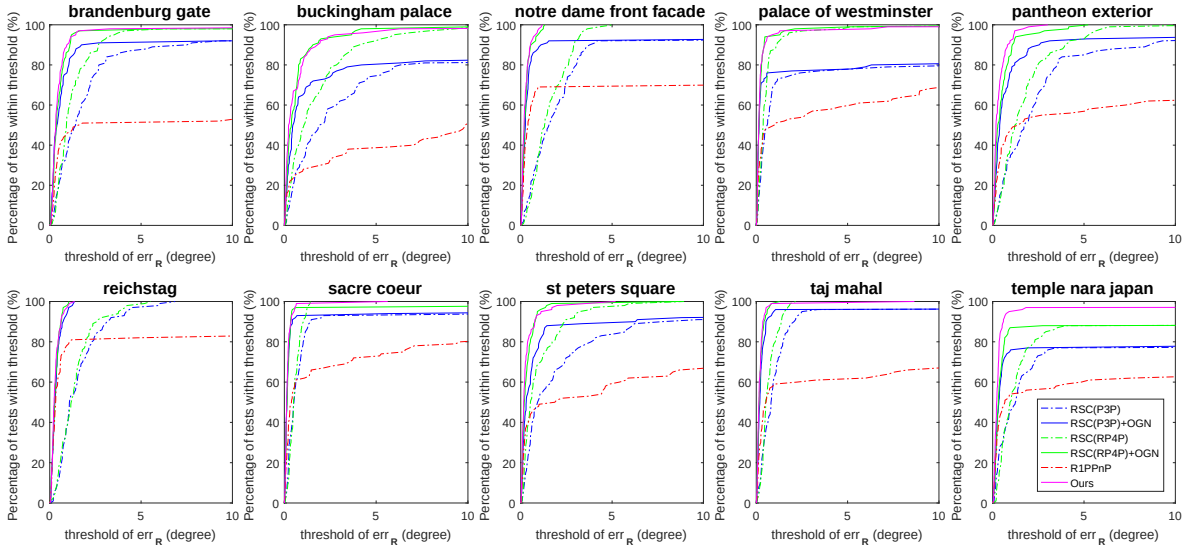


Figure 6: Results of each evaluated scene in real dataset [1]. SIFT is used for feature extraction.

References

[1] Y. Jin, D. Mishkin, A. Mishchuk, J. Matas, P. Fua, K. M. Yi, and E. Trulls, "Image matching across wide baselines: From paper to practice," *International Journal of Computer Vision*, vol. 129, no. 2, pp. 517–547, 2021.