

Supplementary Material: Comparative Evaluation of Hand-Crafted and Learned Local Features

Johannes L. Schönberger¹ Hans Hardmeier¹ Torsten Sattler¹ Marc Pollefeys^{1,2}
¹ Department of Computer Science, ETH Zürich ² Microsoft Corp.
 {jsch,harhans,sattlert,pomarc}@inf.ethz.ch

1. Reconstruction Metrics

Figures 1, 2, 3, 4, 5, 6, 7, and 8 qualitatively present the metrics of the reconstruction results. We show the figures to allow for an easier comparison of the methods. We refer to Table 3 from the paper for the numerical values.

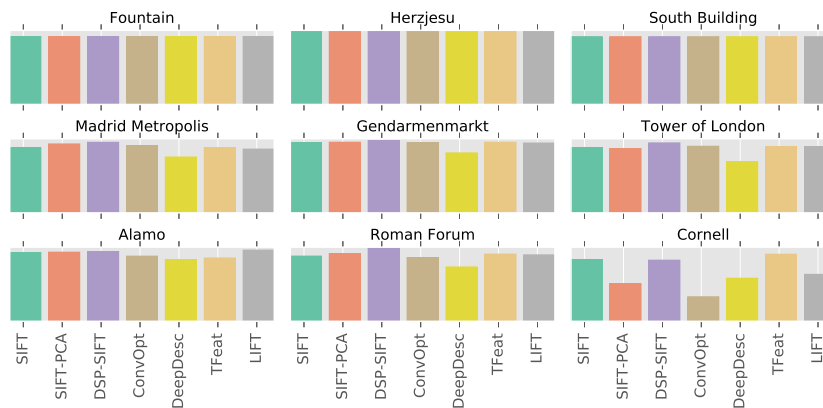


Figure 1. Number of registered images for the different methods.

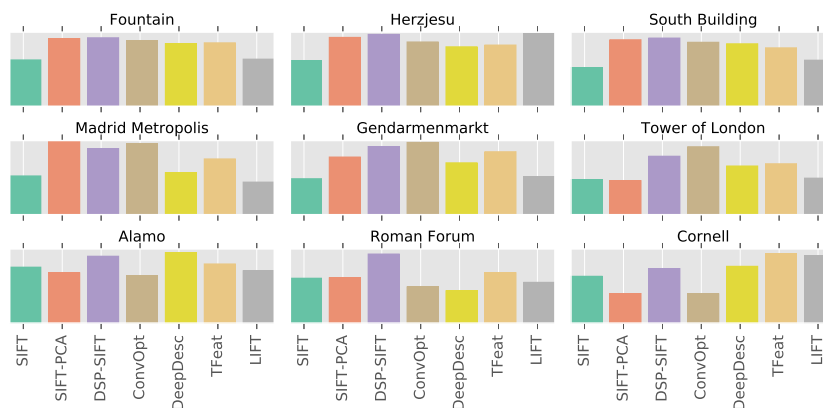


Figure 2. Number of sparse points for the different methods.

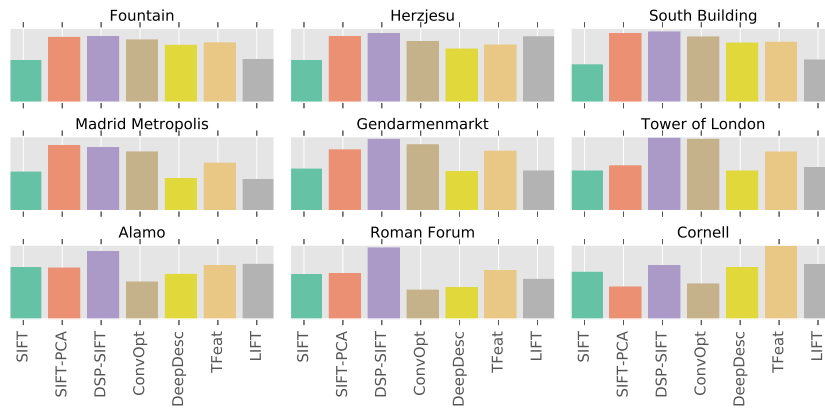


Figure 3. Number of observations for the different methods.

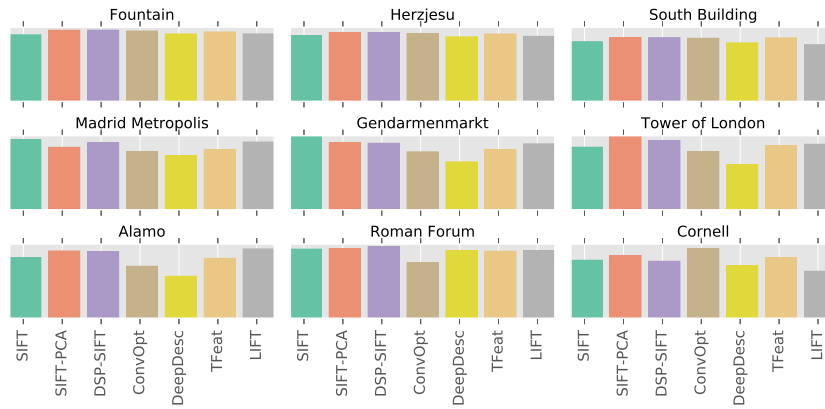


Figure 4. Mean track length for the different methods.

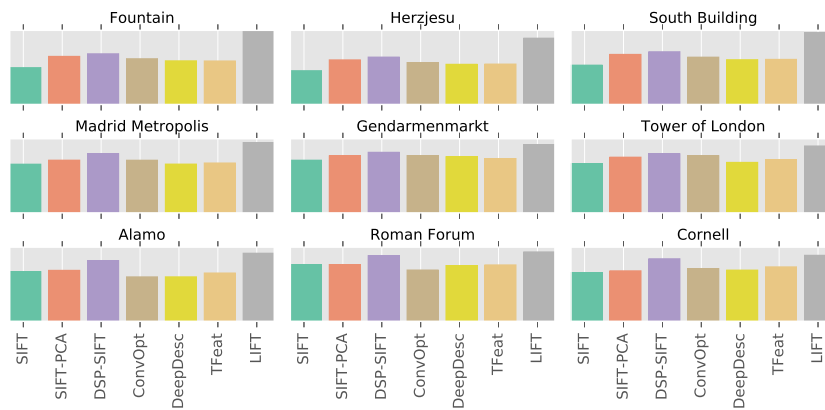


Figure 5. Mean reprojection error for the different methods.

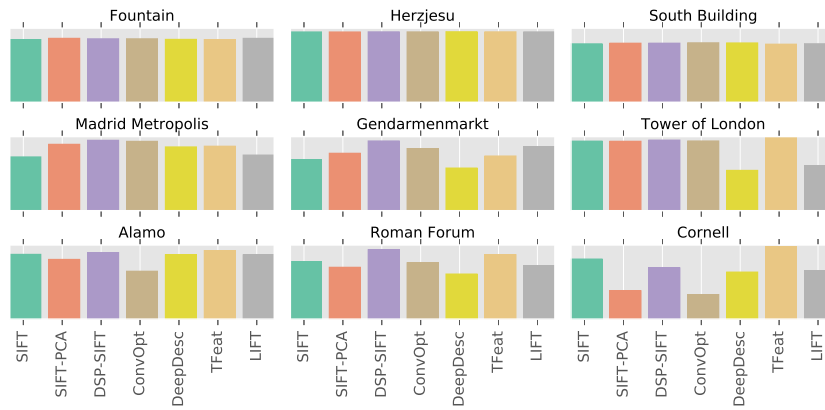


Figure 6. Number of dense points for the different methods.

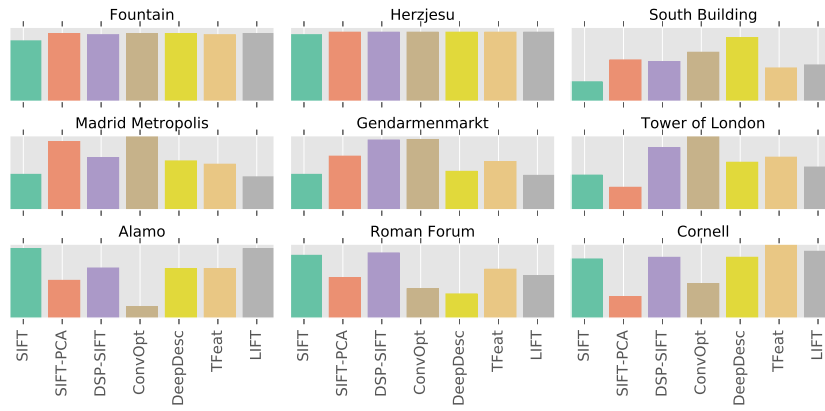


Figure 7. Number of inlier pairs for the different methods.

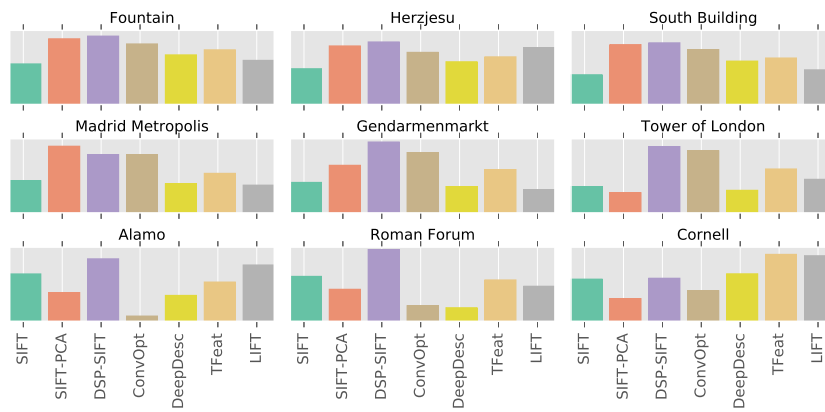


Figure 8. Number of inlier matches for the different methods.

2. Reconstruction Screenshots

Figures 9, 10, 11, 12, 13, 14, and 15 show the sparse and dense reconstructions for the Fountain dataset. We observe that all methods produce visually indistinguishable results, underlining the need for more challenging benchmark datasets for a meaningful comparison of local features.



Figure 9. Sparse and dense reconstruction of Fountain for SIFT.



Figure 10. Sparse and dense reconstruction of Fountain for SIFT-PCA.

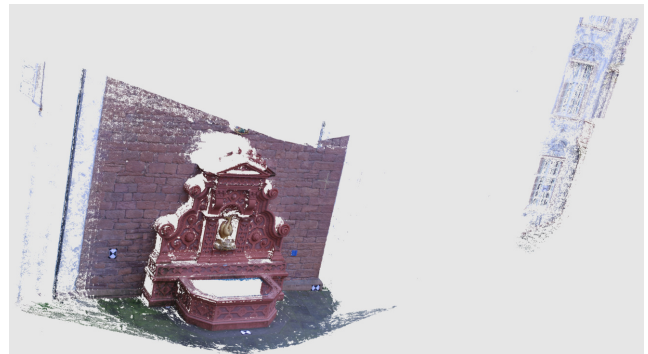


Figure 11. Sparse and dense reconstruction of Fountain for DSP-SIFT.

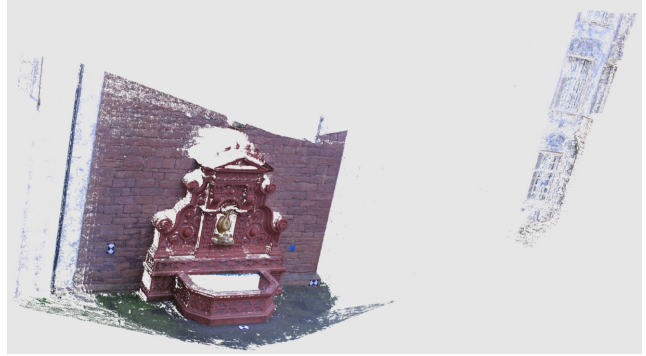
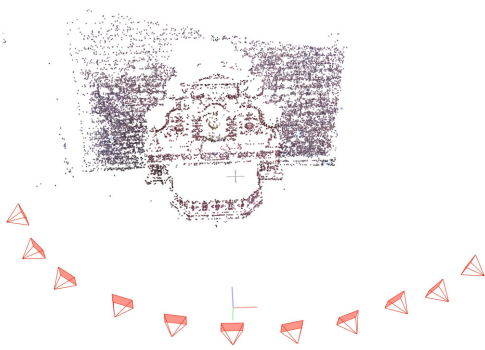


Figure 12. Sparse and dense reconstruction of Fountain for ConvOpt.



Figure 13. Sparse and dense reconstruction of Fountain for DeepDesc.



Figure 14. Sparse and dense reconstruction of Fountain for TFeat.



Figure 15. Sparse and dense reconstruction of Fountain for LIFT.

Figures 16, 17, 18, 19, 20, 21, and 22 show the sparse and dense reconstructions for the Gendarmenmarkt dataset. We observe that more challenging datasets result in significantly different reconstructions for the different local feature approaches. While the differences are not always prominent in the sparse reconstructions, the dense models often significantly differ. This evidences that dense reconstructions are a necessary indicator of the quality of a reconstruction. DSP-SIFT produces the most complete results. Notice the horizontal facades in the lower part of the screenshots, which are captured by comparatively few images. The reconstruction by TFeat contains some gross outlier structures due to incorrectly registered cameras. The results demonstrate that more challenging datasets are needed to meaningfully evaluate the performance of local features in practice.

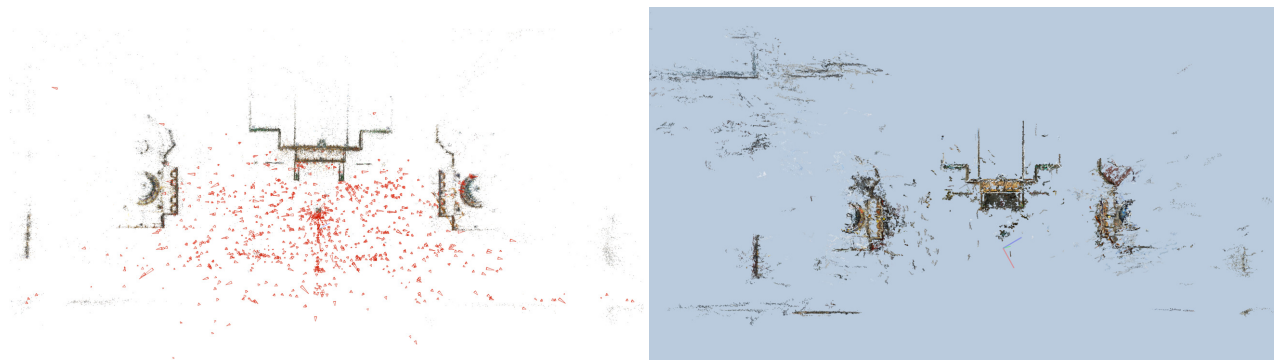


Figure 16. Sparse and dense reconstruction of Gendarmenmarkt for SIFT (950 registered images).

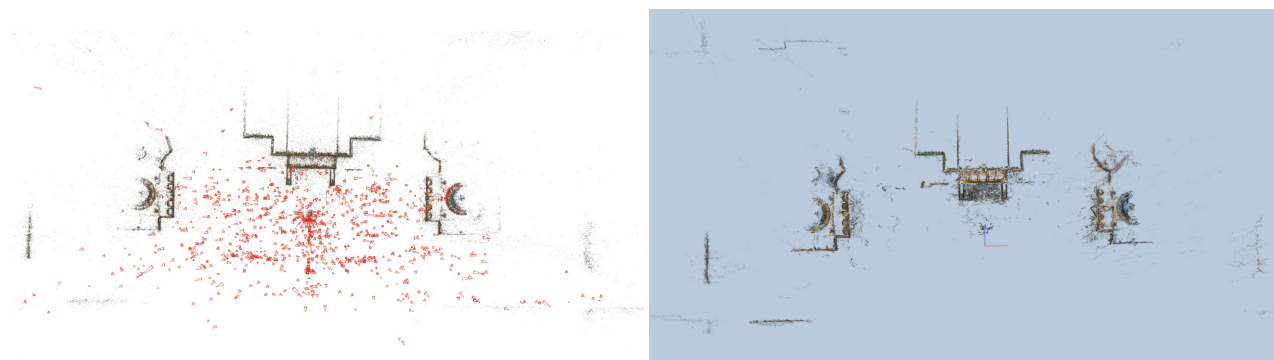


Figure 17. Sparse and dense reconstruction of Gendarmenmarkt for SIFT-PCA (953 registered images).

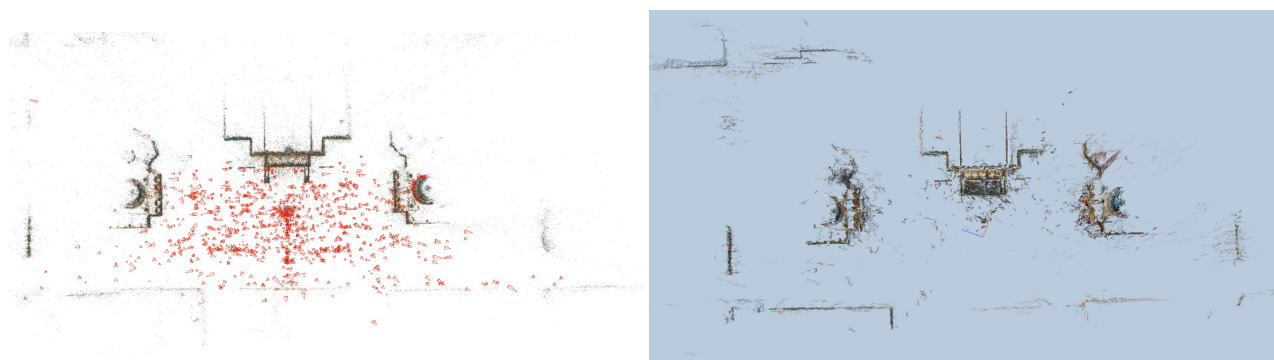


Figure 18. Sparse and dense reconstruction of Gendarmenmarkt for DSP-SIFT (975 registered images).

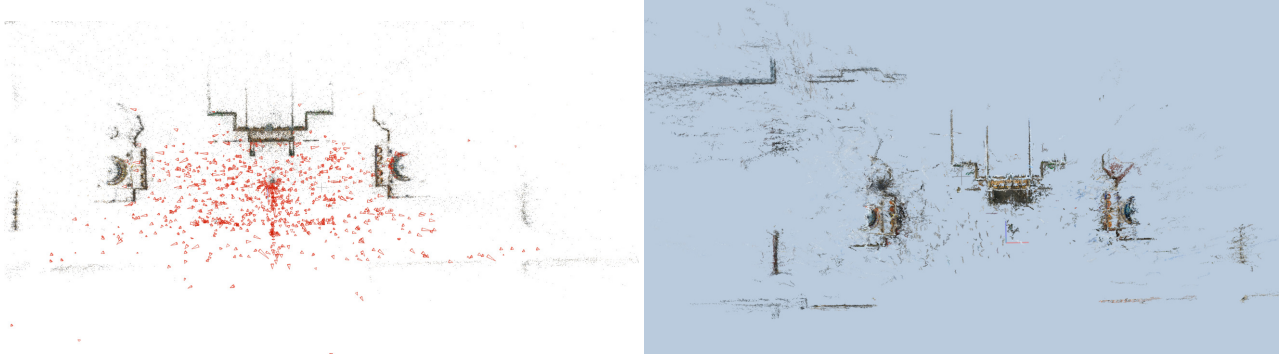


Figure 19. Sparse and dense reconstruction of Gendarmenmarkt for ConvOpt (945 registered images).

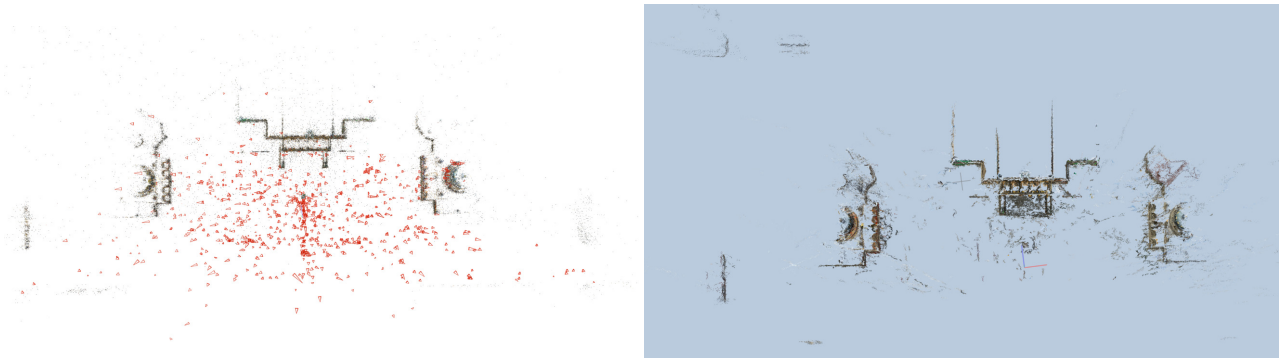


Figure 20. Sparse and dense reconstruction of Gendarmenmarkt for DeepDesc (809 registered images).

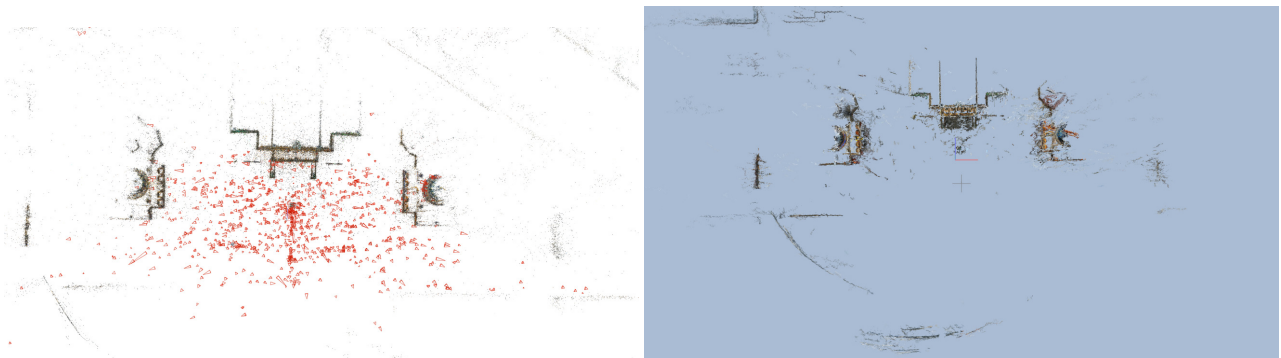


Figure 21. Sparse and dense reconstruction of Gendarmenmarkt for TFeat (953 registered images).

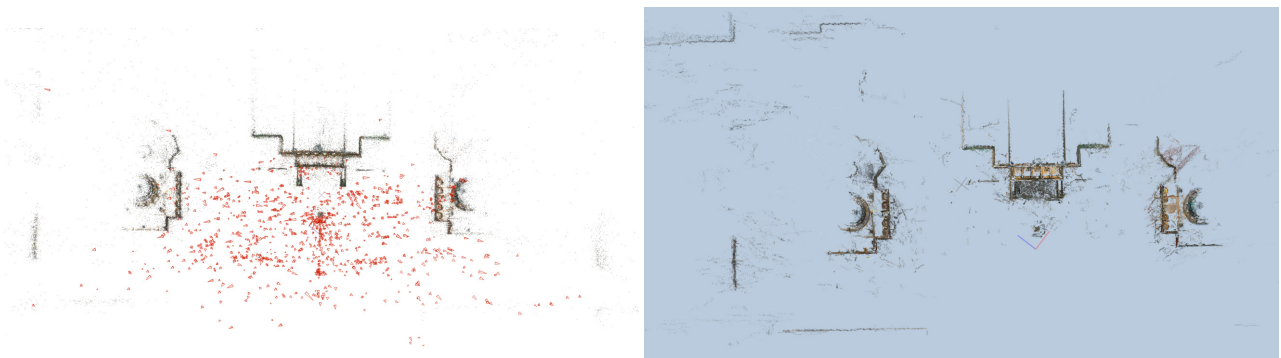


Figure 22. Sparse and dense reconstruction of Gendarmenmarkt for LIFT (942 registered images).