Fast Object Retrieval Using Direct Spatial Matching -Additional Results

Zhiyuan Zhong and Jianke Zhu *Member, IEEE,*

Table I shows the comparison results on Oxford5K and Paris with various DoFs. It can be seen that DoF model affects little for both FSM and our proposed DSM scheme.

TABLE I
PERFORMANCE COMPARISON ON AFFINE TRANSFORMATION MODELS WITH VARIOUS DOFS.

DoF	3 DoF	4 DoF	5 DoF
DSM on Oxford5K	0.850	0.849	**0.852**
DSM on Paris	**0.814**	0.813	0.812
FSM [1] on Oxford5K	0.644	0.646	**0.648**

Table II shows the details of the performance on Oxford dataset. +AQE means performing AQE after DSM, +k-NN means performaning k-NN algorithm after DSM and AQE. There are some badly performing queries of the landmark Magdalen, Ashmolean, Balliol and Cornmarket where the mAP are below 0.8. However, after performing the k-NN algorithm, they all achieve over 0.8, which is quite hard as discussed in [2], [3]. We observe that for *cormarket_000105, oxford_001115, balliol_000187* the AP are $0.49, 0.08, 0.389$ using DSM, but after k-NN ranking the AP are $0.86, 0.68, 0.924$ respectively, while [2] report that all methods perform badly for these cases due to the significant view point changes. Under those extreme circumstances DSM can only guarantee the very top ranked results are correct. However, combining the effective AQE and k-NN ranking can significantly improve the performance as long as the top ranked object are correctly located. We believe the reason why our overall mAP on Oxford5k is higher than that of [2] is because we achieve quite a performance boost even for those badly case after k-NN ranking.

TABLE II
PERFORMANCE ON OXFORD DATASET IN DETAIL.

Ground Truth	Oxford5K			Oxford105K		
	DSM	+AQE	+k-NN	DSM	+AQE	+k-NN
All Souls	0.926	0.993	0.982	0.916	0.985	0.973
Ashmolean	0.789	0.887	0.917	0.760	0.764	0.865
Balliol	0.756	0.893	0.957	0.758	0.897	0.962
Bodleian	0.958	0.995	0.998	0.954	0.975	0.997
Christ Church	0.888	0.981	0.993	0.872	0.959	0.980
Cornmarket	0.765	0.882	0.874	0.748	0.754	0.822
Hertford	0.958	0.999	0.997	0.940	0.985	0.978
Keble	0.973	0.996	0.996	0.972	0.980	0.992
Magdalen	0.501	0.702	0.824	0.480	0.626	0.723
Pitt Rivers	1.000	1.000	1.000	1.000	1.000	1.000
Radcliffe Cam.	0.853	0.879	0.918	0.805	0.814	0.863
mAP	0.850	0.928	0.950	0.836	0.883	0.924

Fig. 1 shows visual comparison of various methods. For a specific Average Precision, the Y-axis indicates the number of queries achieve the AP or higher.

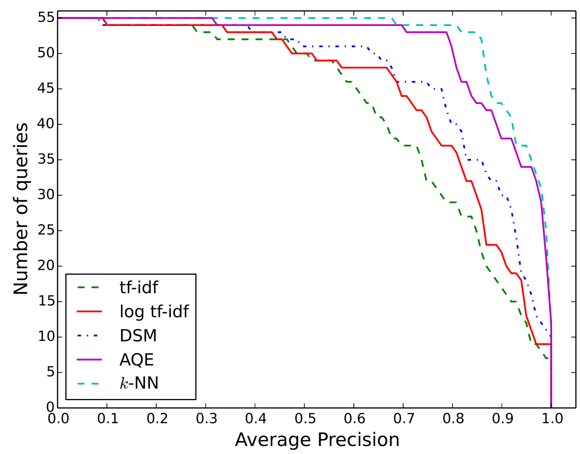

Fig. 1. Performance comparisons of various methods on Oxford5K.

Fig. 2 shows the three challenging queries from the landmark Magdalen, Ashmolean, Balliol and Cornmarket.

Fig. 3 shows visual examples for the cases, where 'AQE+k-NN' reranking greatly boost the retrieval performance. In each case, the five precision recall curves show the result of the five queries for the corresponding landmark.

Fig. **??** shows some object localization examples of DSM. The first 4 queries come from Paris6K dataset, and the last 4 queries come from Oxford5K dataset. We can see the DSM can handle scale variation, deformation and viewpoint change to a certain extent, which demonstrate the robustness and flexibility of the proposed direct spatial matching method. All of the queries are taken from Oxford5K and Paris dataset. Since our system is based on gravity vector assumption [1], [4], the rotation is simply ignored. As discussed in [4], gravity vector assumption holds surprisingly well, especially for the building images.

REFERENCES

[1] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman, "Object retrieval with large vocabularies and fast spatial matching," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2007.

[2] R. Arandjelović and A. Zisserman, "Three things everyone should know to improve object retrieval," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2012.

[3] O. Chum, J. Philbin, J. Sivic, M. Isard, and A. Zisserman, "Total recall: Automatic query expansion with a generative feature model for object retrieval." in *IEEE International Conference on Computer Vision*, 2007.

[4] M. Perdoch, O. Chum, and J. Matas, "Efficient representation of local geometry for large scale object retrieval," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2009.

Fig. 2. Three challenging queries from the landmark Magdalen, Ashmolean, Balliol and Cornmarket.

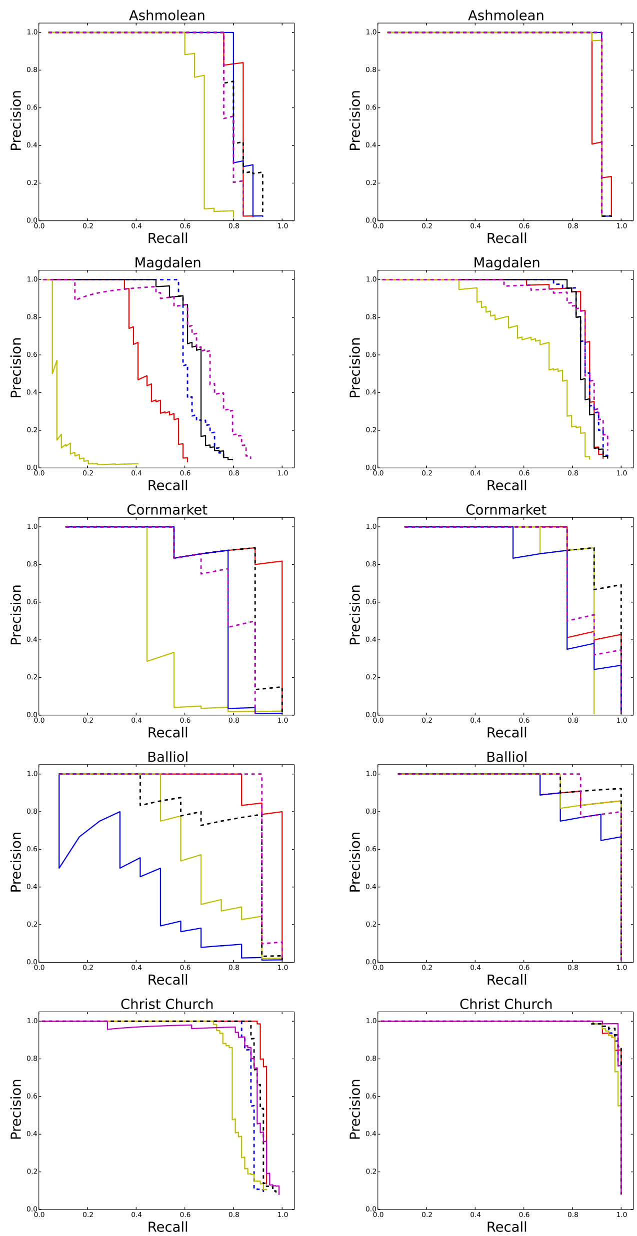

Fig. 3. Precision recall curves before (left) and after (right) k-NN reranking on Oxford5K experiment.

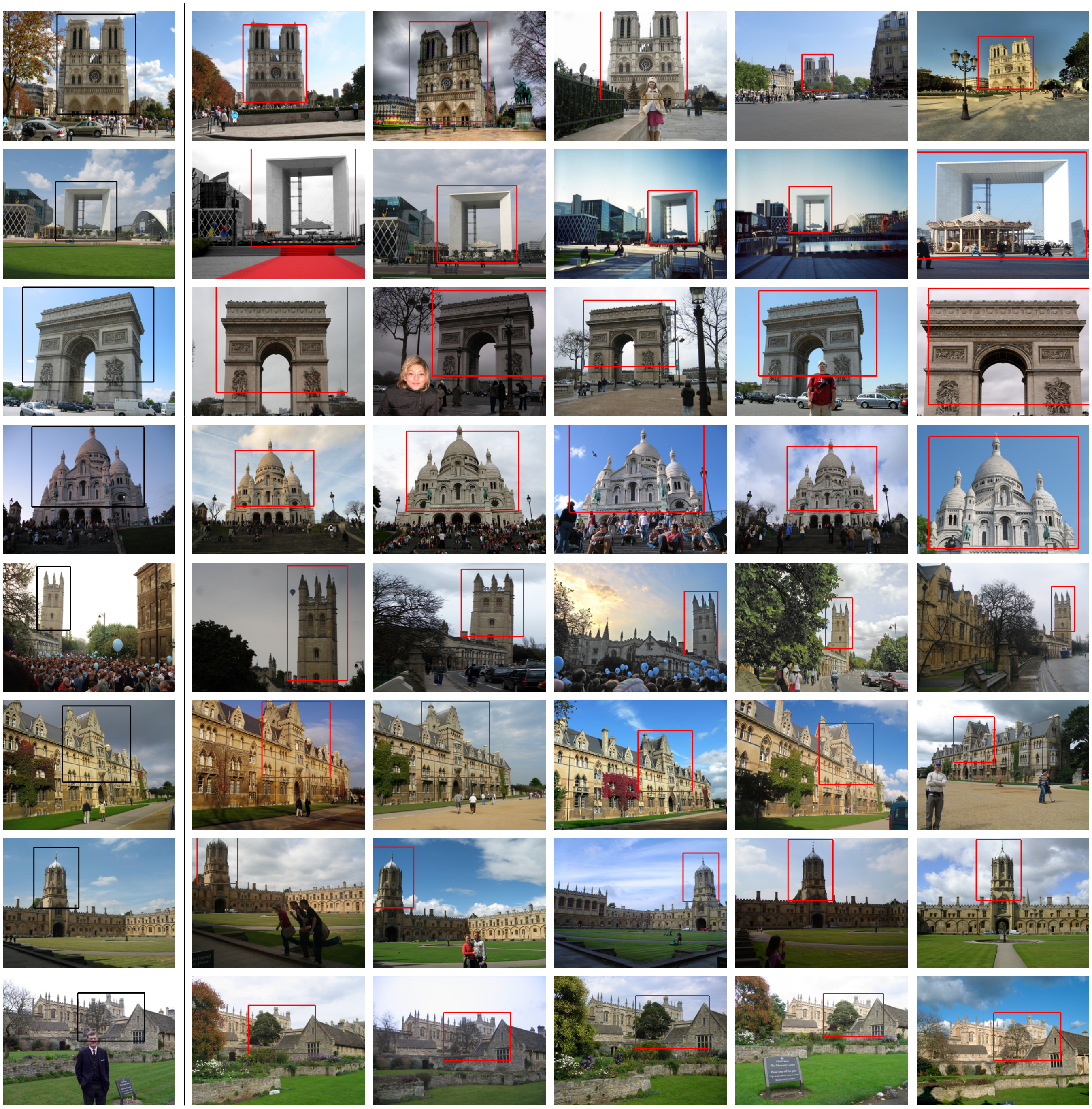

Fig. 4. Example of object localization. The black rectangles on the left are the queries. The images on the right are the top ranked images.

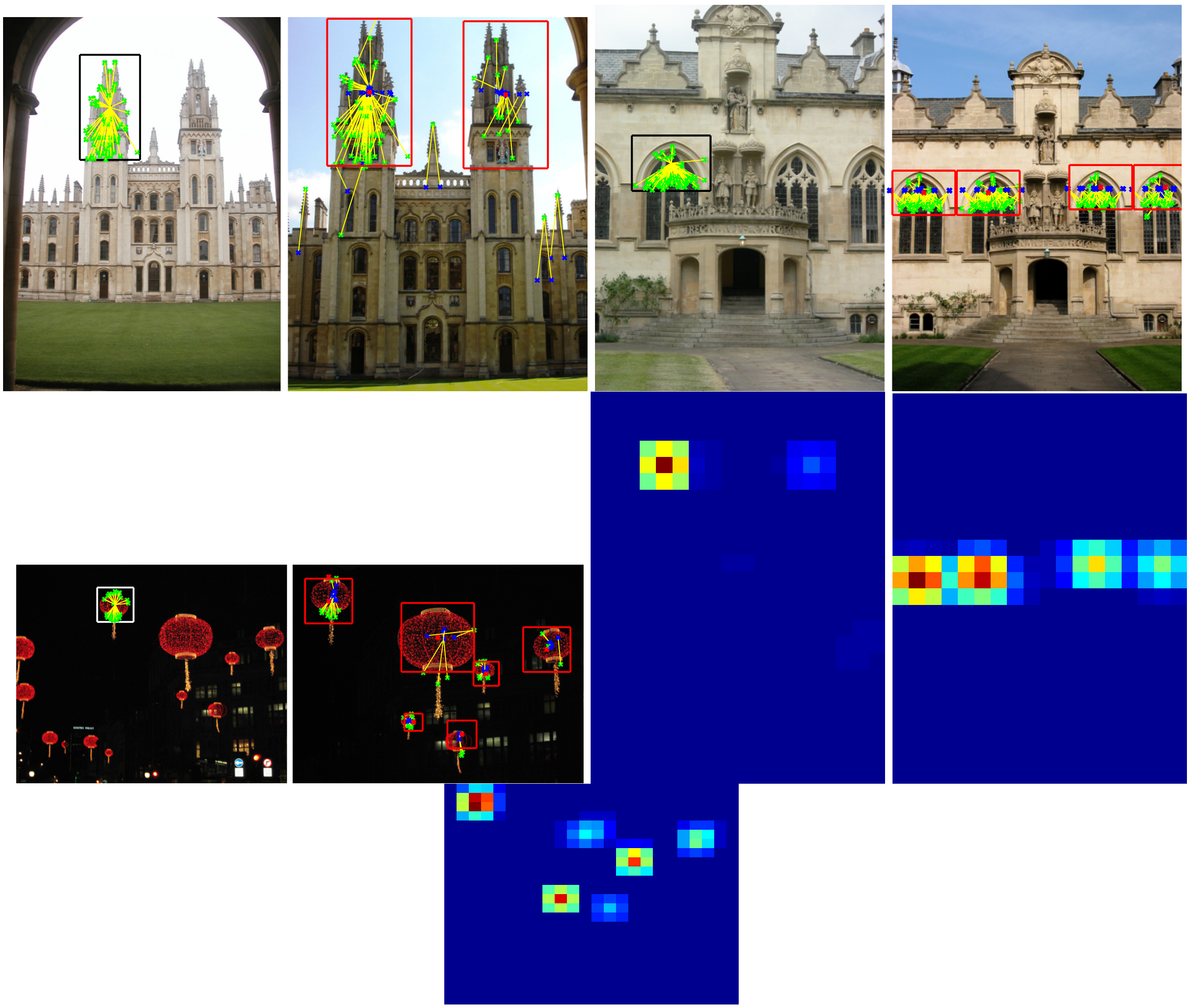

Fig. 5. Locate the near-duplicate structures using direct spatial matching by performing non-maximal suppression on voting map. These are multiple objects localization examples and their corresponding voting maps.