

Efficient image retrieval for 3D structures

Relja Arandjelović
relja@robots.ox.ac.uk

Andrew Zisserman
az@robots.ox.ac.uk

Department of Engineering Science,
University of Oxford
Parks Road
Oxford, OX1 3PJ, UK

Abstract

Large scale image retrieval systems for specific objects generally employ visual words together with a ranking based on a geometric relation between the query and target images. Previous work has used planar homographies for this geometric relation. Here we replace the planar transformation by epipolar geometry in order to improve the retrieval performance for 3D structures.

To this end, we introduce a new minimal solution for computing the affine fundamental matrix. The solution requires only two corresponding elliptical regions. Unlike previous approaches it does not require the rotation of the image patches, and ensures that the necessary epipolar tangency constraints are satisfied.

The solution is well suited for real time reranking in large scale image retrieval, since (i) elliptical correspondences are readily available from the affine region detections, and (ii) the use of only two region correspondences is very efficient in a RANSAC framework where the number of samples required grows exponentially with sample size. We demonstrate a gain in computational efficiency (over other methods of solution) without a loss in quality of the estimated epipolar geometry.

We present a quantitative performance evaluation on the Oxford and Paris image retrieval benchmarks, and demonstrate that retrieval of 3D structures is indeed improved.

1 Introduction

Systems for retrieving specific objects in large scale image datasets have seen tremendous progress over the last five years [9, 16, 19]. It is now possible to retrieve objects from datasets of millions of images [10, 18] and performance on standard retrieval benchmarks has improved significantly [5, 11, 18, 20]. Given a query image of the object, retrieval usually proceeds in two steps: an initial ranking/short list based on visual words [16, 21]; followed by a reranking based on a geometric relation. In order for the retrieval to be immediate both steps must be extremely efficient at run time. The visual words enable text retrieval like scalability and speed, but the reranking is potentially more costly as it involves computing a geometric relation between the query image and each target image in the short list. The geometric relation of choice has been a planar homography [9, 19].

In this work we show that epipolar geometry can be used for the ranking relation. Why is this important? As is well known [8], a planar homography is the exact relation between two images if (i) the scene is planar, or (ii) the scene is three dimensional but the camera centres of the two images coincide. In fact both these cases often apply for the type of flickr images in retrieval datasets since (i) often buildings are used for queries, and their facades

are near-planar, and (ii) tourists often take photos of famous landmarks from much the same place. However, where these cases don't apply, e.g. if the situation is a 3D structure acquired from different viewpoints, potential image matches may be lost. Epipolar geometry is the exact relation for this situation.

In this paper we develop a new *minimal* solution for the affine fundamental matrix (which represents the affine approximation of epipolar geometry) based on the correspondence of a pair of ellipses. This solution is tailored to the architecture of large scale retrieval systems since: first, ellipses are already produced by the affine covariant region detectors (e.g. affine-Harris [15]) used by such systems; and, second, as only two correspondences are required, computation of the relation (using RANSAC) is very efficient. We show that indeed there are matches missed by the homography relation which are correctly retrieved using epipolar geometry, and that the system performance in terms of standard benchmarks equals or exceeds a baseline using homographies. To our knowledge this is the first time that the ellipse pair solution has been given, and that epipolar geometry has been used to rerank a real time large scale retrieval system.

The original method for computing the affine fundamental matrix used four point correspondences [8]. With the development of affine covariant detectors, others have improved on this method by reducing the number of feature correspondences to two, but have not given *minimal* solutions: Perd'och, *et al.* [17] use two corresponding local affine frames (LAF) computed from matched maximally stable regions (MSER) [13]. Each LAF can be thought of as an elliptical region with a rotation associated with it. However, their method is not minimal, yielding up to 15 solutions [22] for each pair of LAF correspondences, and also there may be several LAFs associated with each MSER increasing further the number of possible solutions. Since evaluating hypotheses is the bottleneck of any RANSAC based system, having a smaller number of hypotheses (up to four in our case) is very beneficial. Furthermore, we do not require the rotation, which can be unstable. Goshen and Shimshoni [7] use two corresponding SIFT detections, but again their method is not minimal and requires rotation.

We present the new solution in section 2 and then evaluate it within a RANSAC estimation framework in section 3, where results are given on a number of standard image pairs. Section 4 then describes how the new relation is incorporated into the reranking step of a large scale retrieval system. Since epipolar geometry is a weaker geometric constraint than the homographies used previously, additional filtering steps are required to suppress false positives. We show that performance is improved on standard baselines without compromising on retrieval speed.

2 Computing the affine fundamental matrix from two ellipses

The affine fundamental matrix F_A represents the epipolar geometry in the case that the two camera projections are affine [8]. The problem we address is estimating F_A given a pair of corresponding ellipses. This is a minimal configuration for the problem, in a similar manner to the 4 point case. We will start with a counting argument and then derive the solution.

F_A has four degrees of freedom (dof) corresponding to the ratio of the five non-zero entries (see the form of (1)). In the case of points, each 3D point has 3 dof and each correspondence provides 4 measurements (from the position of the point in each image). Thus for n points there are $4n$ constraints, and $3n + 4$ unknowns, and for $n = 4$ the number of constraints equals the number of unknowns. In this minimal case there is a unique solution.

In the case of conics, each 3D conic has 8 dof (3 specifying the plane of the conic and

5 for the conic on that plane) and each correspondence provides 10 measurements (5 from each image conic). Thus for n conics there are $10n$ constraints and $8n + 4$ unknowns, and $n = 2$ is minimal. As will be seen, in this case there are 4 possible solutions.

As this is a minimal solution all constraints are exactly satisfied. Consequently the following necessary conditions must be satisfied: the centres of corresponding ellipses satisfy the epipolar constraint, and for each epipolar line tangent to an ellipse in one image there is a corresponding epipolar line tangent to the corresponding ellipse in the other image.

We will sketch the derivation of the solution here; full details are given in the report available at [2]. There are 4 steps: (i) the images are first transformed by an affine transformation to simplify the computation; (ii) given the image measurements in the transformed frame, the necessary conditions are used to obtain a quartic in a scalar variable; (iii) each solution of the quartic specifies a possible F_A which satisfies all the necessary conditions (there may be 0, 2 or 4 real solutions); finally, (iv) the F_A are transformed back to the original images.

Parametrisation of F_A . The affine fundamental matrix can be parametrised as [14]:

$$F_A = \begin{pmatrix} 0 & 0 & -b \sin \beta \\ 0 & 0 & b \cos \beta \\ -a \sin \alpha & a \cos \alpha & \sqrt{1 - a^2 - b^2} \end{pmatrix} \quad (1)$$

where α and β are the orientations of the epipolar lines in images 1 and 2 respectively. This can be easily seen by the fact that the right null vector (the epipole of image 1) is $(\cos \alpha, \sin \alpha, 0)^T$ and the left null vector (the epipole of image 2) is $(\cos \beta, \sin \beta, 0)^T$. Given a pair of corresponding epipolar lines separated by d and d' respectively in the two images, then [14]

$$-\frac{a}{b} = \frac{d'}{d} \quad (2)$$

i.e. the ratio of a to b specifies the ratio of epipolar line spacings between the images.

Choosing a canonical frame. We have the freedom to affine transform the images in order to simplify the calculation. We will drop the A subscript on F_A from here on for clarity. Given the two applied affine homographies H_A and H'_A for images 1 and 2 respectively, such that $\mathbf{x}_l = H_A \mathbf{x}$ and $\mathbf{x}'_l = H'_A \mathbf{x}'$, then the original F and transformed F_l are related as:

$$\mathbf{x}'_l{}^T F_l \mathbf{x}_l = 0, \quad \mathbf{x}'^T H'^T_A F_l H_A \mathbf{x} = 0, \quad \mathbf{x}^T F \mathbf{x} = 0 \quad \Rightarrow \quad F = H'^T_A F_l H_A \quad (3)$$

H_A and H'_A can be chosen such that one ellipse (call it ellipse number one) is transformed into a unit circle centred on the origin, and the x coordinate of the centre of the second ellipse is zero. An example illustrating the transformations is shown in Figure 1.

In the canonical frame the measurements (free parameters) are the four remaining parameters of the second ellipse (its centre is at $x = x' = 0$).

Necessary conditions in the canonical frame. By construction of H_A and H'_A the values of a_l and b_l are fully determined. Using the fact that ellipse one in images 1 and 2 are each centred at $(0, 0)^T$ and that their centres must obey the epipolar constraint:

$$(0 \quad 0 \quad 1) F_l \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0 \quad \Rightarrow \quad \sqrt{1 - a_l^2 - b_l^2} = 0 \quad (4)$$

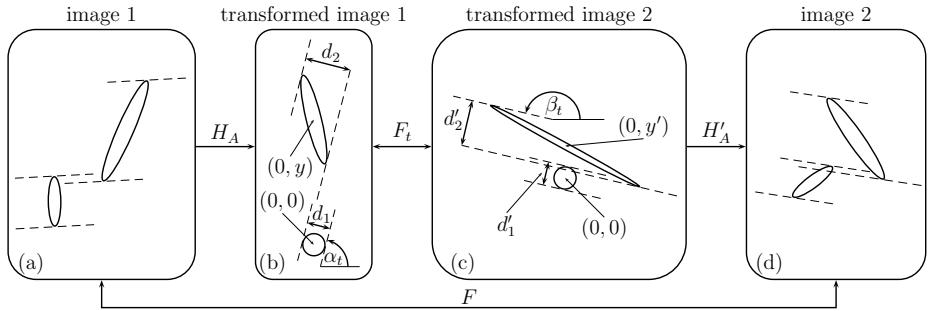


Figure 1: (a)(d) Two images with two ellipses each. (b)(c) the two transformed images as described in the text, and the relationships between the images. Epipolar lines are dashed.

From (2) and the fact that by construction $d'_1 = d_1$ (see figure 1) it is clear that $a_t = -b_t$. Combining this result with (4) gives $a_t = -b_t = \frac{1}{\sqrt{2}}$ and F_t can be written as:

$$F_t = \begin{pmatrix} 0 & 0 & \sin \beta_t \\ 0 & 0 & -\cos \beta_t \\ -\sin \alpha_t & \cos \alpha_t & 0 \end{pmatrix} \quad (5)$$

There are now only two remaining unknowns, α_t and β_t , and equations are obtained for these from the epipolar constraint for the centre of the second ellipse (the correspondence $(0, y)$ to $(0, y')$), and the distance between the epipolar lines tangent to this ellipse. The former gives $y \cos \alpha_t = y' \cos \beta_t$, while the latter and (2) give the remaining equation for the two unknowns α_t and β_t . This results in a quartic in a scalar variable t where $\cos \alpha_t = 1/\sqrt{1+t^2}$ and $\cos \beta_t = y/(y'\sqrt{1+t^2})$.

Finally, each pair of solutions for α_t and β_t can be substituted back into (5) and using (3) one can find F . An example with four solutions is given in Figure 2.

3 Estimation of affine epipolar geometry

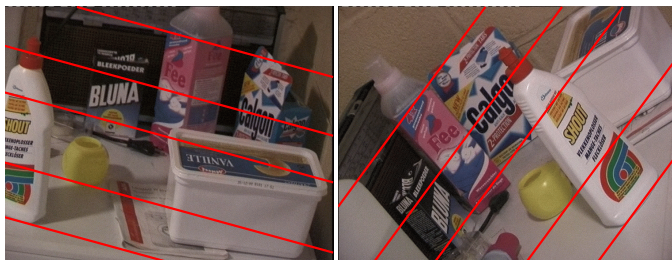
We describe here the estimation of epipolar geometry using the two ellipse solution. We use a standard framework of: computing affine covariant elliptical regions in each image; representing these by a SIFT descriptor for matching; and then estimating the affine fundamental matrix using the standard RANSAC algorithm [6]. Our interest here is in evaluating the performance of the two ellipse solution – for example are the detected ellipses stable enough to generate useful hypotheses? – and in comparing the number of hypotheses that RANSAC explores for the minimal two-ellipse solution compared to the minimal four-point solution. Note, more sophisticated RANSAC-like approaches [3, 4, 7, 23] could be used here, but that is not the main focus of this paper.

In this section we match on the raw SIFT descriptors. In section 4 we explore the case of visual word matching where the SIFT descriptors are vector quantized, and putative matches are simply formed from the descriptors quantized to the same visual word.

Implementation details. We use affine-Hessian interest point detectors [15] and describe the resulting elliptical region by a 128-dimensional SIFT descriptor [12]. Putative correspondences between the features of the two images are determined by Lowe’s second nearest neighbour test [12]: two nearest neighbour features are deemed to be putative matches only if their SIFT distance d_1 is less than kd_2 , where d_2 is the second nearest neighbour and $k \in (0, 1]$ (values for k are given below).



(a) Sol. #1: 7 inliers (b) Sol. #2: 79 inliers (c) Sol. #3: 18 inliers (d) Sol. #4: 9 inliers



(e) Final solution obtained using all the inliers of solution #2: 122 inliers

Figure 2: Estimating affine epipolar geometry from two ellipses for the Wash scene [25]. (a)-(d): the four solutions consistent with the two corresponding ellipses (shown in yellow), (b) is the correct one. Example epipolar lines are shown in red; note that they are tangent to the ellipses. (e): the final estimate. Note that there are significant perspective effects, but the method is capable of estimating a good affine approximation to the epipolar geometry.

RANSAC estimation proceeds by iteratively randomly sampling two ellipse correspondences from the list of putative matches, and computing all the hypotheses (affine fundamental matrices) compatible with the sample as described in Section 2. For each of the hypotheses (up to 4 different solutions) the number of putative matches which are consistent with it (inliers) are found; and the hypothesis with the largest support is kept. Iterations are terminated once the estimated probability of failing to find an all-inlier sample [8] falls below 0.01 or, for practical reasons, the number of iterations reaches 100000. Finally, once the sampling stage has finished, all the inliers supporting the best hypothesis are used to estimate a more accurate affine fundamental matrix using the Gold Standard algorithm described in [8]. The re-estimation is performed until the number of inliers converges.

3.1 Evaluation

Qualitative results are shown in Figures 2 and 3 and demonstrate the correctness of the proposed approach. For quantitative results we compare the two-ellipse minimal solution to a (reference method) four-point minimal solution [8] within the RANSAC estimation algorithm. The two approaches only differ in the sampling stage (since the required sample size is two or four).

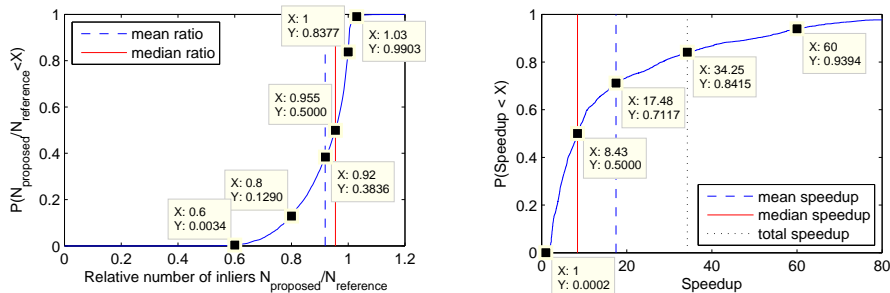
A set of 16 scenes is used for this performance evaluation. These are taken from standard sets Valbonne, Raglan, Keble, House and Wadham which are publicly available online[1], as well as the Wash[25] and China Wall[17] image pairs. The set was intentionally chosen to



(a) Image pair from the Paris dataset [20] (notredame 000447 and 000603)

(b) China Wall scene [17]

Figure 3: Examples of estimated affine epipolar geometry using the two-ellipse solution.



(a) Estimated cumulative density function of the inlier ratio between the two-ellipse and the reference method. For example, the point at $(0.6, 0.0034)$ shows that in 0.34% of the tests the support size found by the two-ellipse method is less than 60% of support size found by the reference method.

(b) Estimated cumulative density function of the obtained speedup, defined as the number of RANSAC samples required by the reference relative to the two-ellipse method. For example, the point at $(34.25, 0.8415)$ shows that in 84.15% of the tests the reference method's sample count was more than 34.25 times the two-ellipse method's.

Figure 4: Performance of the two-ellipse compared to the reference four-point method.

contain examples with significant perspective effects in order to demonstrate the fact that the affine epipolar assumption is not very limiting. The two approaches are tested on each of the scenes with a few appropriate values of k (0.9 and 0.95) and the threshold on the reprojection error [8] (2, 5, 10, 20, 40 pixels) used to define inliers. Each test is repeated 50 times with different random seeds used for the RANSAC sampling stage (i.e. the complete RANSAC estimation is repeated).

From figure 4(a) it is clear that the two-ellipse method manages to find a hypothesis with large support most of the time: on average it finds 92% of the support size found by the reference method while the median is 95.5%. Only 12.9% of the tests found less than 80% of the reference support size, while a negligible 0.3% of the times the two-ellipse method fails to find more than 60% of the reference support size. The two-ellipse method actually finds more inliers than the reference four-point method 16.2% of the time. Figure 4(b) shows the achieved speedup, defined as the number of samples required to estimate the epipolar geometry using the reference method divided by the corresponding number of samples for our method. The vast computational saving is apparent, with the median speedup of 8.43 and the mean of 17.48. 6.1% of the time our method is more than 60 times faster, while it is slower a negligible 0.02% of the time. Note also that even larger speedup would be achieved if we didn't limit the number of iterations to 100000 as 7.4% of the runs of the reference method reach this limit, while our method never comes even remotely close to it (the maximum number of iterations recorded in our tests was 15132 while only 7.5% reach 1000). Figure 4(b) also shows a "total speedup" of 34.25 which refers to the total number

of iterations of the reference method divided by the total number of iterations of our method across all the tests. This hints that the two-ellipse method provides a vast speedup in a large scale image retrieval set-up where many attempts at matching two images are made.

4 Large scale image retrieval

In a large scale image retrieval scenario one is interested in performing visual search in real time. We closely follow the approach described in [19] where image features are quantized into 1M visual words using approximate K-means clustering, and an inverted index is used to quickly find images in the database which contain the visual words of the query image. The retrieved images are ranked using the standard tf-idf weighting scheme. To ensure correct matches are ranked highly the top 200 tf-idf results are reranked using a more computationally demanding spatial verification.

As discussed in the introduction, spatial verification methods using an affine transformation perform very well when the geometric assumption applies (i.e. yields very few false positives), for example for building facades. For this reason we choose the following ordering for geometric verification: first an affine transformation is estimated, and verified images ranked by their number of inliers (as in [19]); then, for each of the remaining unverified image (in the top 200 results obtained from the tf-idf scheme), affine epipolar geometry is estimated, and the verified matches are listed next ordered by their number of inliers. Finally all unverified results are listed in the order of decreasing tf-idf scores.

Verification based on epipolar geometry has a tendency to yield a large number of false positives because the epipolar relation is quite loose as it provides only a point to line mapping, unlike the affine transformation which provides a point to point map. To filter out false positives while trying to keep as many true positives as possible a simple but effective heuristic is used: an extremely loose affine homography (reprojection errors of up to 200 pixels are tolerated) is fitted by RANSAC to the inliers of the fundamental matrix, and if less than 60% of these inliers are consistent with the homography then the image match is rejected. In the case of a true positive match most inliers are consistent. An example of a false positive filtered out by this heuristic is shown in figure 6.

4.1 Evaluation

The method is evaluated on the standard and publicly available Oxford [19] and Paris image datasets [20]. The performance is measured as the mean average precision (mAP) over a number of standard queries. For each dataset there are 11 landmarks and 5 queries for each of these. As a baseline we compare to the method of [19] which uses an affine transformation to verify and rerank matches.

The comparison of mAP scores is shown in Table 1. For Oxford we use the publicly available quantization into 1M words, this means that the baseline reproduces the results of [19]. For Paris (where there is not a publicly available vocabulary) we generate the 1M visual words using the approximate K-means method of [19]. In this case we run the entire process 5 times (i.e. generate the clustering 5 times using different seeds for the approximate K-means, and compute the mAP retrieval scores for each of these), and report also the standard deviation over the runs.

Our method outperforms the baseline; the margin looks small at a first glance, but when compared to the gain that the baseline method achieves over BoW (i.e. tf-idf with no spatial verification), it is clear that the improvement is significant. Figure 5 illustrates the type of images/objects that are missed by the affine homography but correctly retrieved by epipolar geometry. The table also gives the mean retrieval time required to perform each query in

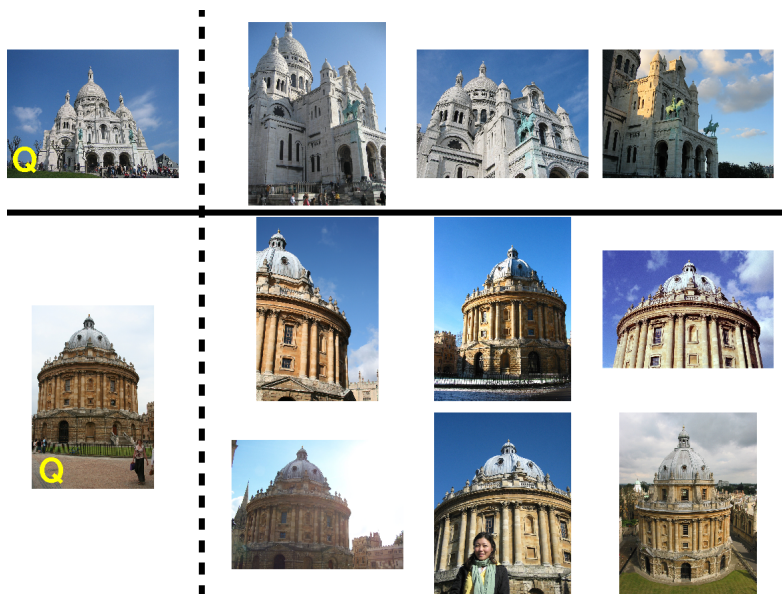


Figure 5: **Two queries from the Oxford and Paris datasets.** The results (shown right) are correctly verified with affine epipolar geometry, but are *not* verified using an affine transformation.

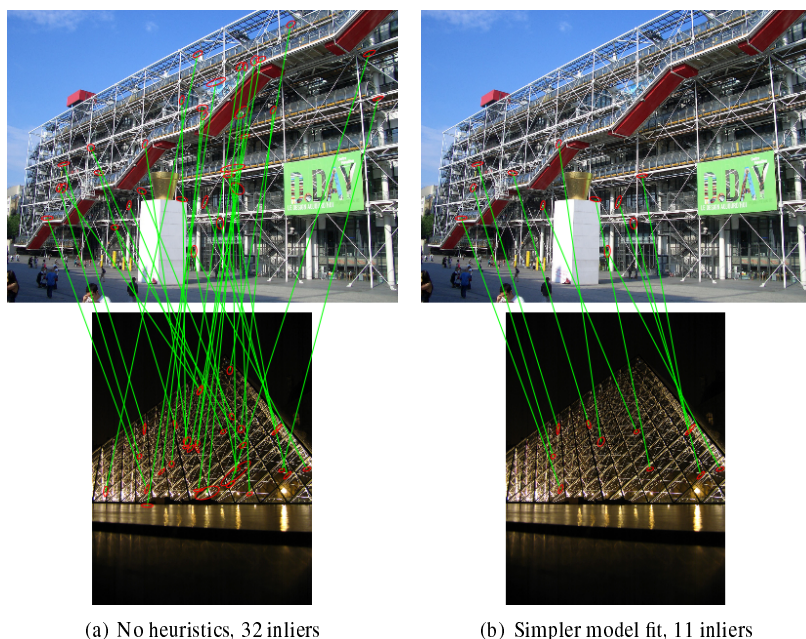


Figure 6: **A false positive object match correctly rejected by the loose homography heuristic.** (a) the inliers to a fundamental matrix fit. (b) the inliers to a homography fit to the inliers of (a). Because so few remain the image match is rejected.

Oxford 5k				Paris			
Query	BoW	H	F2	Query	BoW	H	F2
All Souls	0.535	0.576	0.578	Defense	0.414	0.426	0.433
Ashmolean	0.598	0.624	0.630	Eiffel	0.415	0.424	0.426
Balliol	0.600	0.609	0.601	Invalides	0.635	0.641	0.646
Bodleian	0.563	0.685	0.690	Louvre	0.444	0.451	0.459
Christ Church	0.624	0.650	0.658	Moulin Rouge	0.646	0.655	0.656
Cornmarket	0.568	0.597	0.625	Musee d'Orsay	0.533	0.607	0.610
Hertford	0.763	0.785	0.786	Notre Dame	0.845	0.851	0.852
Keble	0.706	0.807	0.872	Pantheon	0.905	0.909	0.910
Magdalen	0.157	0.174	0.177	Pompidou	0.892	0.888	0.886
Pitt Rivers	0.995	1.000	1.000	Sacre Coeur	0.853	0.859	0.861
Radcliffe Camera	0.675	0.685	0.687	Triomphe	0.520	0.523	0.523
mAP	0.617	0.654	0.664	mAP (mean)	0.646	0.658	0.660
				mAP (std)	0.001	0.001	0.001
Avg. query time	0.13s	0.50s	0.86s	Avg. query time	0.15s	1.03s	1.33s

Table 1: **Retrieval performance on the Oxford and Paris datasets.** BoW is the standard tf-idf weighting scheme, H is the verification using affine transformation, and F2 the new ranking on epipolar geometry based on the two-ellipse algorithm. The baseline H results on the Oxford set reproduce the performance of [19].

the Oxford and Paris datasets. Timings are on a 3GHz machine using a single core. The additional cost of performing the epipolar geometry reranking is minimal, showing that the method is suitable for real time retrieval. Finally, the approach was evaluated on the Oxford 100k dataset [19] where our method achieves a slight improvement over the baseline (mAP of 0.5663 compared to 0.5656).

5 Conclusions

We have shown that epipolar geometry can be efficiently integrated into the standard architecture of a large scale retrieval system, and that performance is improved – recovering matches on 3D structure missed by other geometric relations. Methods which have been developed to improve the performance over [19] such as Hamming embedding and similarity transformation filtering [9], run time query expansion [5], soft assignment [20], data set query expansion [24], and different feature detectors [18], could be incorporated here with similar benefits expected.

Acknowledgements. We are grateful for financial support from the Royal Academy of Engineering, Microsoft, and ERC grant VisRec no. 228180. We would also like to thank Michal Perd'och for providing his images.

References

- [1] <http://www.robots.ox.ac.uk/~vgg/data/data-mview.html>.
- [2] <http://www.robots.ox.ac.uk/~vgg/research/3D.html>.

- [3] O. Chum and J. Matas. Matching with PROSAC – Progressive sample consensus. In *Proc. CVPR*, 2005.
- [4] O. Chum, J. Matas, and J. Kittler. Locally optimized RANSAC. In G. Goos, J. Hartmanis, and J. van Leeuwen, editors, *DAGM 2003: Proceedings of the 25th DAGM Symposium*, number 2781 in LNCS, pages 236–243. Springer-Verlag, Sep 2003.
- [5] 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 *Proc. ICCV*, 2007.
- [6] M. A. Fischler and R. C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm. ACM*, 24(6):381–395, 1981.
- [7] L. Goshen and I. Shimshoni. Balanced exploration and exploitation model search for efficient epipolar geometry estimation. *IEEE PAMI*, 30(7):1230–1242, Jul 2008.
- [8] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.
- [9] H. Jégou, M. Douze, and C. Schmid. Hamming embedding and weak geometric consistency for large scale image search. In *Proc. ECCV*, 2008.
- [10] H. Jégou, M. Douze, and C. Schmid. Packing bag-of-features. In *Proc. ICCV*, Sep 2009.
- [11] H. Jégou, M. Douze, and C. Schmid. On the burstiness of visual elements. In *Proc. CVPR*, Jun 2009.
- [12] D. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2):91–110, 2004.
- [13] J. Matas, O. Chum, M. Urban, and T. Pajdla. Robust wide baseline stereo from maximally stable extremal regions. In *Proc. BMVC.*, pages 384–393, 2002.
- [14] P. R. S. Mendonça and R. Cipolla. Estimation of epipolar geometry from apparent contours: Affine and circular motion cases. In *Proc. CVPR*, 1999.
- [15] K. Mikolajczyk and C. Schmid. Scale & affine invariant interest point detectors. *IJCV*, 1(60):63–86, 2004.
- [16] D. Nister and H. Stewenius. Scalable recognition with a vocabulary tree. In *Proc. CVPR*, 2006.
- [17] M. Perd’och, J. Matas, and O. Chum. Epipolar geometry from two correspondences. In *Proc. ICPR*, 2006.
- [18] M. Perd’och, O. Chum, and J. Matas. Efficient representation of local geometry for large scale object retrieval. In *Proc. CVPR*, 2009.
- [19] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman. Object retrieval with large vocabularies and fast spatial matching. In *Proc. CVPR*, 2007.

- [20] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman. Lost in quantization: Improving particular object retrieval in large scale image databases. In *Proc. CVPR*, 2008.
- [21] J. Sivic and A. Zisserman. Video Google: A text retrieval approach to object matching in videos. In *Proc. ICCV*, volume 2, pages 1470–1477, Oct 2003.
- [22] H. Stewénius, D. Nistér, F. Kahl, and F. Schaffalitzky. A minimal solution for relative pose with unknown focal length. In *Proc. CVPR*, 2005.
- [23] P. H. S. Torr and A. Zisserman. MLESAC: A new robust estimator with application to estimating image geometry. *CVIU*, 78:138–156, 2000.
- [24] T. Turcot and D. G. Lowe. Better matching with fewer features: The selection of useful features in large database recognition problems. In *ICCV Workshop on Emergent Issues in Large Amounts of Visual Data (WS-LAVD)*, 2009.
- [25] T. Tuytelaars and L. Van Gool. Wide baseline stereo matching based on local, affinely invariant regions. In *Proc. BMVC.*, pages 412–425, 2000.