International Journal of Control Theory and Applications

ISSN : 0974-5572

Volume 10 • Number 31 • 2017

Content Based Medical Image Retrieval using Scale Invariant Feature Transform and Speed-Up Robust Feature

S. Govindaraju[a] and G.P. Ramesh Kumar[b]

[a]*Asst. Professor, Dept. of Computer Science, S.N.R. Sons College, Coimbatore, India*
E-mail: kavigovindaraju@gmail.com
[b]*Asst. Professor, Department of Computer Science, Government Arts College, Kulithalai, India.*
E-mail: gpr.snr@gmail.com

Abstract: Content-based image recovery comprises of recovering the most outwardly comparative images to a given question image from a database of images. Therapeutic Image has been progressively connected in clinical determination and treatment. It is critical to make utilization of extensive quantities of images in therapeutic image administration framework keeping in mind the end goal to help clinician to break down and analyze. The customary data recovery methods are not fit for recovering huge scale therapeutic image databases. CBIR for restorative image databases does not mean to supplant the doctor by anticipating the infection of a specific case yet to help him/her in conclusion. This paper presents Content based Medical Image Retrieval (CBMIR) method into ordering therapeutic image databases. This paper proposes CBMIR utilizing neighborhood highlight descriptor and significance criticism. The neighborhood include descriptor SIFT, SURF are consolidated to concentrate highlights from the restorative images. Moreover, the significance input of client is taken to enhance execution of the recovery. The execution of the proposed strategy has been assessed through test comes about. The outcomes are promising, affirming that the created technique has better recovery precision over the contrasted with different strategies.

Keywords: *Medical Image Retrieval , SIFT, SURF, CBMIR.*

1. INTRODUCTION

The reason for nearby invariant components is to give a portrayal that permits to productively coordinating neighborhood structures between images. That is, we need to get a scanty arrangement of nearby estimations that catch the pith of the basic information images and that encode their fascinating structure. To meet this objective, the component extractors must satisfy two critical criteria. Feature descriptors are a vital part of numerous PC vision calculations. In expansive scale coordinating and substantial scale image recovery, the discriminative force of descriptors and their heartiness to image mutilations are a key consider the execution, we remove and standardize the district content and process a neighborhood descriptor for every locale. Highlight coordinating is them performed by looking at the neighborhood descriptors utilizing a reasonable closeness measure. • The element extraction process ought to be repeatable and exact, so that similar components are separated on two images demonstrating a similar question. • In the meantime, the elements ought to be particular, so that diverse image structures can be distinguished from each other. What's more, we commonly require an adequate number of highlight locales to cover the objective protest, with the goal that it can in any case be perceived under fractional impediment. Finding an arrangement of unmistakable key focuses. 2. Characterize an area around each key point in a scale-or relative invariant way. 3. Separate and standardize the locale content. 4. Register a descriptor from the standardized district. 5. Coordinate the neighborhood descriptors.

2. PREVIOUS WORKS

The proposed descriptor learning framework consists of two independent algorithms, one for learning descriptor pooling regions, and the other for discriminative dimensionality reduction. Most conventional feature descriptors are hand-crafted and use a fixed configurations of pooling regions, *e.g*. SIFT and its derivatives, use rectangular regions organized in a grid. In the Powell optimisation technique is employed to find the parameters of a DAISY like descriptor. The corresponding objective is not convex, making the optimisation prone to local extreme[1]. Recently, pooling region selection using boosting was proposed in . Since the optimisation is greedy, there is no guarantee to reach the global optimum. Discriminative dimensionality reduction can also be related to metric learning, on which a vast literature exists. While our ranking constraints are similar to those of, the authors themselves do not consider simultaneous dimensionality reduction.

A similar formulation with application to learning descriptors for image retrieval was used in dimensionality reduction is performed using the projections corresponding to the largest eigen values of the learnt Mahalanob is matrix. This method is ad hoc as the dimensionality reduction is not taken into account in the learning objective. In our case, we enforce a low rank of the Mahalanob is matrix by penalising its nuclear norm, which is a convex surrogate of the matrix rank[2]. We tackle the optimisation problem in a principled way and perform large-scale optimisation of the non-smooth objective using the recently developed Regularized Dual Averaging (RDA) method which we employ for both L -regularized learning of pooling regions and nuclear norm regularised learning of discriminative dimensionality reduction. BRIEF and BRISK descriptors are computed by comparing intensity values at patch locations, which are either randomly selected or hand-crafted[3]. A different approach was used in LDA Hash, where the binary descriptor is computed by thresholding the SIFT descriptor projected onto a subspace using a learnt projection matrix. Instead of SIFT, used the vectorised image patch.

3. LOCAL FEATURE DESCRIPTOR

Neighborhood Descriptors Once an arrangement of intrigue areas has been removed from a image, their substance should be encoded in a descriptor that is appropriate for discriminative coordinating. Lowe proposes the accompanying methodology for the introduction standardization step. For each identified intrigue locale, the district's scale is utilized to choose the nearest level of the Gaussian pyramid, with the goal that every single after calculation are performed in a scale invariant way[4]. We then develop an inclination introduction histogram with 36 receptacles covering the 360° scope of introductions. For every pixel in the district, the comparing inclination introduction is gone into the histogram, weighted by the pixel's angle extent and by a Gaussian window fixated on the key point with a size of 1.5σ[5,6,7]. The most noteworthy crest in the introduction histogram is taken as the overwhelming introduction, and a parabola is fitted to the 3 neighboring histogram qualities to interject the pinnacle position for better precision. By and by, it might happen that different similarly solid introductions are found for a solitary intrigue locale. In such cases, choosing just a single of them would imperil the acknowledgment methodology, since little changes in the image flag could bring about one of alternate introductions to be picked rather, which could prompt to fizzled matches[8,9]. Consequently, Lowe proposes to make a different intrigue area for every introduction crest that scopes no less than 80% of the predominant pinnacle's esteem. This procedure fundamentally enhances the district identifier's repeatability at a generally little extra cost by Low04b , just around 15% of the focuses are doled out various introductions)[10,11,12]. Other Feature descriptors - old and new are LBP, LTP and variations, HAAR; - PCA-SIFT, VLAD, MOSIFT, - profound elements, CNN, Fisher vector, - SV-DSIFT, BF-DSIFT, LL-MO1SIFT, 1SIFT, VM1SIFT, VLADSIFT, - DECAF, Fisher vector pyramid, IFV - Dirichlet Histogram - Simplex based STV (3-D), MSDR[13,14,15]. The stages of the proposed work.

1. Hybrid Local feature descriptors (SIFT, SURF)
2. Multiple Similarity measurement for features
3. Relevance feedback
4. Neighborhood image list
5. Top-*k* image retrieval

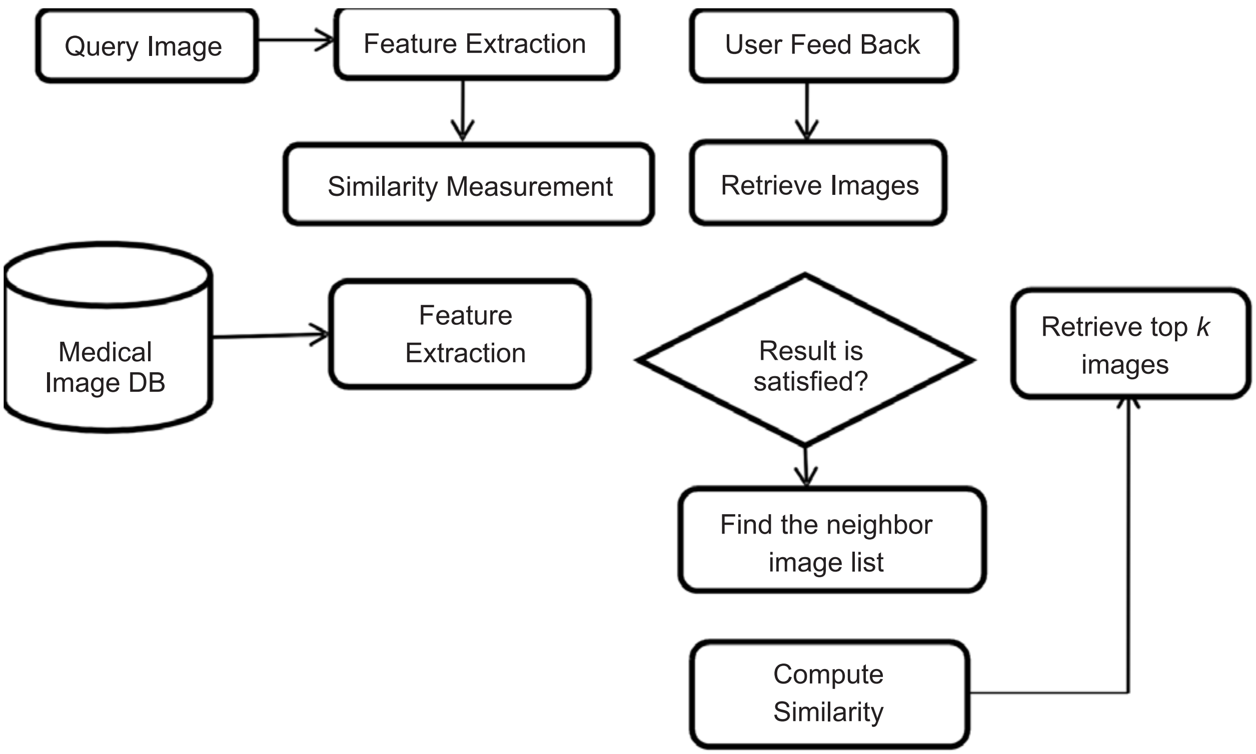

Figure 1: Architecture of proposed methodology

4. SCALE INVARIANT FEATURE TRANSFORM

Step 1: Scale-space extreme Detection - Detect interesting points (invariant to scale and orientation) using DOG.

Step 2: Key point Localization – Determine location and scale at each candidate location, and select them based on stability.

Step 3: Orientation Estimation – Use local image gradients to assigned orientation to each localized key point. Preserve theta, scale and location for each feature.

Step 4: Key point Descriptor - Extract local image gradients at selected scale around key point and form a representation invariant to local shape distortion and illumination them.

Accurate key point localization is to reject the low contrast points and the points that lie on the edge. Low contrast points elimination: Fit key point at to nearby data using quadratic approximation.

$$D(\underline{x}) = D + \frac{\partial D^T}{\partial \underline{x}}\underline{x} + \frac{1}{2}x^T\frac{\partial^2 D^T}{\partial \underline{x}^2}\underline{x}$$

$$D(x, \sigma) = G(x, k\sigma) - G(x,\sigma))*I(x)$$

Where Calculate the local maxima of the fitted function

$$D(\underline{x}) = D + \frac{\partial D^T}{\partial \underline{x}}\underline{x} + \frac{1}{2}x^T \frac{\partial^2 D^T}{\partial \underline{x}^2}\underline{x}$$

$$\frac{\partial D}{\partial \underline{x}} = \frac{\partial\left[D + \frac{\partial D^T}{\partial \underline{x}}\underline{x} + \frac{1}{2}x^T \frac{\partial^2 D^T}{\partial \underline{x}^2}\underline{x}\right]}{\partial \underline{x}}$$

$$\boxed{} = 0$$

$$\Rightarrow \qquad \hat{\underline{x}} = -\frac{\partial^2 D^{-1}}{\partial x^2}\frac{\partial D}{\partial \underline{x}}$$

Eliminating edge response: DOG gives strong response along edges – Eliminate those responses Solution for this is to check “cornerness” of each key point.

1. On the edge one of principle curvatures is much bigger than another.
2. High cornerness No dominant principle curvature component.
3. Consider the concept of Hessian and Harris corner

Hessian Matrix

$$H = \begin{bmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{bmatrix}$$

Harris corner criterion Matrix

$$\frac{Tr(H)^2}{Det(H)} < \frac{(r+1)^2}{r}$$

Step 3: Orientation Assignment

Assign constant orientation to each key point based on local image property to obtain rotational invari. The magnitude and orientation of gradient of an image patch $I(x, y)$ at a particular scale is:

$$m(x, y) = \sqrt{(I(x+1, y) - I(x-1, y))^2 + (I(x, y+1) - I(x, y-1))^2}$$

$$\theta(x, y) = \tan^{-1}\frac{I(x, y+1) - I(x, y-1)}{I(x+1, y) - I(x-1, y)}$$

Step 3: Orientation Assignment

1. Create weighted (magnitude +Gaussian) histogram of local gradient directions computed at selected scale.
2. Assign dominant orientation of the region as that of the peak of smoothed histogram
3. For multiple peaks create multiple key points

Step 4: Local image descriptor

Aim : Obtain local descriptor that is highly distinctive yet invariant to variation like illumination and affine change[16,17,18,19].

1. Consider a rectangular grid 16*16 in the direction of the dominant orientation of the region.
2. Divide the region into 4*4 sub-regions.
3. Consider a Gaussian filter above the region which gives higher weights to pixel closer to the center of the descriptor.
4. Create 8 bin gradient histograms for each sub-region Weighted by magnitude and Gaussian window (s is half the window size).

5. SPEED-UP ROBUST FEATURE

The SURF Detector/Descriptor As neighborhood highlight indicators and descriptors have turned out to be more boundless, productive usage are turning out to be increasingly vital. A few methodologies have therefore been proposed keeping in mind the end goal to accelerate the intrigue area extraction or potentially depiction stages [NC(08), BTV(06), BETV(08), RD(08)][20]. Among those, we need to select the SURF ("Speeded-Up Robust Features") approach, which has been outlined as an effective other option to SIFT [BTV(06), BETV(08)]. SURF consolidates a Hessian-Laplace area identifier with a possess slope introduction based element descriptor[21,22,23]. Rather than depending on Gaussian subsidiaries for its inward calculations, it is however in view of straightforward 2D box channels ("Haar wavelets"), as appeared in Figure[3,9]. Those container channels rough the impacts of the subsidiary channel pieces, yet can be effectively assessed utilizing basic images [VJ04][24,25]. Specifically, this assessment requires a similar steady number of queries paying little respect to the image scale, in this manner expelling the requirement for a Gaussian pyramid. In spite of this improvement, SURF has been appeared to accomplish practically identical repeatability as finders in view of standard Gaussian subordinates, while yielding speedups of more than an element of five contrasted with standard DoG[26]. The SURF descriptor is likewise roused by SIFT and seeks after a comparative spatial binning technique, isolating the component locale into a 4 × 4 network[27,28,29]. Be that as it may, rather than working up an inclination introduction histogram for each container, SURF just figures an arrangement of synopsis insights, bringing about a 64-dimensional descriptor, or a somewhat augmented set bringing about a 128-dimensional descriptor rendition. Spurred by the achievement of SURF, a further improved variant has been proposed in [NC(08)] that exploits the computational power accessible in current CUDA empowered illustrations cards[30,31]. This GPUSURF execution has been accounted for to perform highlight extraction for a 640 × 480 image at casing rates up to 200 Hz, along these lines making highlight extraction a really reasonable preparing step.

6. RESULTS AND DISCUSSIONS

Table 1
Mean Average Precision

	tf-idf	*tf-idf-sp*
SIFT-Baseline	0.712	0.69
SIFT -Proj	0.7245	0.71
Proposed Hybrid	0.812	0.79

Table 2
Improved Mean Average Precision

	tf-idf	*tf-idf-sp*
SIFT-Baseline	0.5	0.69
SIFT -Proj	0.56	0.71
Proposed Hybrid	0.61	0.8

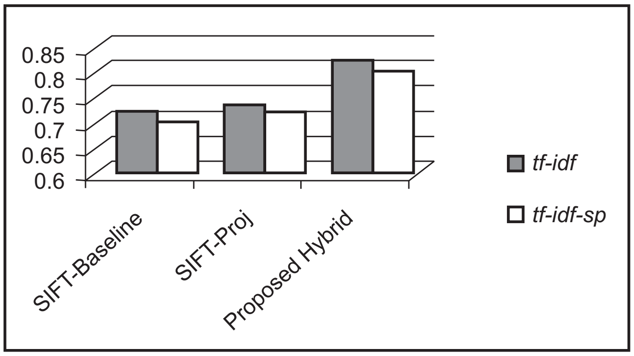

Figure 2: Mean Average Precision

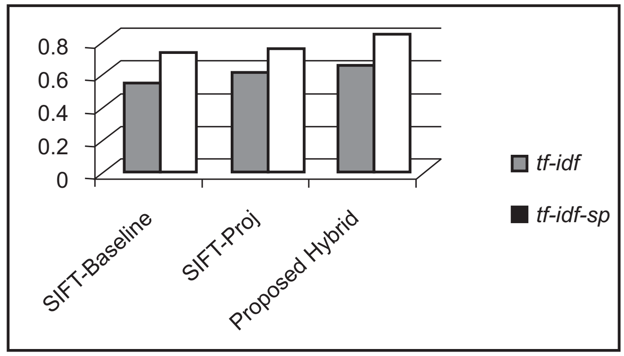

Figure 3: Improved Mean Average Precision

Figure 4: Matching points of the image 1

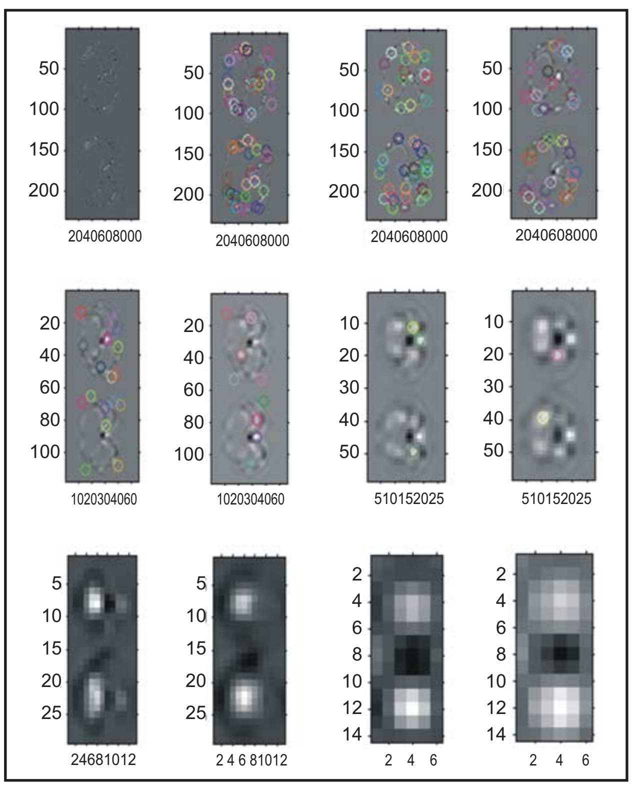

Figure 4: (*a*) Input Query image 1

Figure 5: Matching points of the image 2

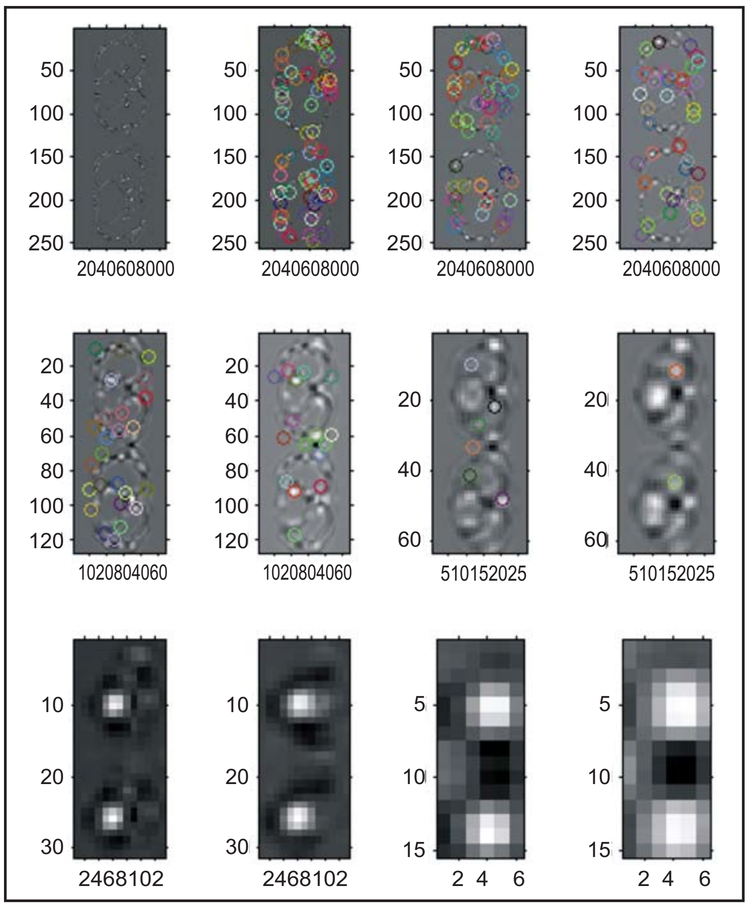

Figure 5: (*a*) Input Query image 1

The experiments with two different feature detection methods, presented in this section, indicate that the proposed learning framework brings consistent improvement irrespective of the underlying feature detector. The MAP improvements brought by the learnt descriptors are consistent for both datasets and retrieval engines, which indicates that our learnt models generalize well. The PR-proj descriptors evaluated above are inherently real-valued. To obtain a compact and fast-to-match representation, the descriptors can be compressed using either binarisation or product quantization. We call the resulting descriptors PR-projbin and PR-proj-pq respectively, and compare them with the state-of-the-art binary descriptors. The binary descriptor of is low-dimensional , while proposes a more accurate, but significant longer.

7. CONCLUSIONS

The improvement of neighborhood invariant elements has had a gigantic effect in numerous territories of PC vision, including wide-gauge stereo coordinating, image recovery, protest acknowledgment, and order. They have given the premise to many best in class calculations and have prompted to various new improvements. The learning definition to the instance of frail supervision and exhibited that the learnt descriptors are agreeable to binarisation. Thorough assessment demonstrated that the proposed calculation beats cutting edge genuine esteemed and parallel descriptors on testing datasets. This was accomplished by means of the utilization of raised learning details combined with expansive scale regularized advancement methods.

REFERENCES

[1] Kristen Grauman and Bastian Leibe, Visual Object Recognition, Synthesis Lectures on Artificial Intelligence and Machine Learning, April 2011, Vol. 5, No. 2, Pages 1- 181.

[2] Beaudet, "Rotationally invariant image operators", in International Joint Conference on Pattern Recognition, pp. 579-583., 1978.

[3] Förstner, W. and Gülch, E., "A fast operator for detection and precise location of distinct points, corners and centers of circular features", in ISPRS Inter commission Workshop', pp. 281-305, 1987.

[4] Harris, C. and Stephens, M., "A combined corner and edge detector", in 'Alvey Vision Conference', pp. 147–151, 1988.

[5] Lindeberg, T., 'Scale-space theory: A basic tool for analyzing structures at different scales', Journal of Applied Statistics 21(2), pp. 224–270, 1994.

[6] Lowe, D., 'Distinctive image features from scale-invariant keypoints', International Journal of Computer Vision 60(2), pp. 91–110, 2004.

[7] Mikolajczyk, K. and Schmid, C., 'A performance evaluation of local descriptors', IEEE Transactions on Pattern Analysis & Machine Intelligence 27(10), 31–37, 2005.

[8] R. Arandjelovi c and A. Zisserman. Three things everyone should know to improve object retrieval. In Proc. CVPR, 2012.

[9] X. Boix, M. Gygli, G. Roig, and L. Van Gool. Sparse quantization for patch description. In Proc. CVPR, 2013.

[10] M. Brown, G. Hua, and S. Winder. Discriminative learning of local image descriptors. IEEE PAMI, 33(1):43–57, 2011.

[11] M. Fazel, H. Hindi, and S. P. Boyd. A rank minimization heuristic with application to minimum order system approximation. In Proc. ACC, pages 4734–4739, 2001.

[12] Y. Gong and S. Lazebnik. Iterative quantization: A procrustean approach to learning binary codes. In Proc. CVPR, 2011.

[13] M. Guillaumin, J. Verbeek, and C. Schmid. Multiple instance metric learning from automatically labeled bags of faces. In Proc. ECCV, pages 634–647, 2010.

[14] Jelena Kovacevic and Amina Chebira. An introduction to frames. Foundations and Trends in Signal Processing, 2(1):1–94, 2008.

[15] M. P. Kumar, B. Packer, and D. Koller. Self-paced learning for latent variable models. In NIPS, pages 1189–1197, 2010.

[16] S. Leutenegger, M. Chli, and R. Siegwart. BRISK: Binary robust invariant scalable keypoints. In Proc. ICCV, pages 2548–2555, 2011.

[17] D. Lowe. Distinctive image features from scale-invariant keypoints. IJCV, 60(2):91–110, 2004.

[18] 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.

[19] K. Mikolajczyk and C. Schmid. An affine invariant interest point detector. In Proc. ECCV. Springer-Verlag, 2002.

[20] K. Mikolajczyk and C. Schmid. Scale & affine invariant interest point detectors. IJCV, 1(60):63–86, 2004.

[21] Y. Nesterov. Primal-dual subgradient methods for convex problems. J. Math. Prog., 120(1):221–259, 2009.

[22] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman. Object retrieval with large vocabularies and fast spatial matching. In Proc. CVPR, 2007.

[23] J. Philbin, M. Isard, J. Sivic, and A. Zisserman. Descriptor learning for efficient retrieval. In Proc. ECCV, 2010.

[24] J. D. M. Rennie and N. Srebro. Fast maximum margin matrix factorization for collaborative prediction. In Proc. ICML, pages 713–719, 2005.

[25] F. Schaffalitzky and A. Zisserman. Multi-view matching for unordered image sets, or "how do i organize my holiday snaps?. In Proc. ECCV, volume 1, pages 414–431. Springer-Verlag, 2002.

[26] K. Simonyan, A. Vedaldi, and A. Zisserman. Descriptor learning using convex optimisation. In Proc. ECCV, 2012.

[27] N. Snavely, S. Seitz, and R. Szeliski. Photo tourism: exploring photo collections in 3D. In Proc. ACM SIGGRAPH, volume 25, pages 835–846, 2006.

[28] C. Strecha, Bronstein A. M., M. M. Bronstein, and P. Fua. LDAHash: Improved matching with smaller descriptors. IEEE PAMI, 34(1), 2012.

[29] T. Trzcinski, M. Christoudias, P. Fua, and V. Lepetit. Boosting bnary key point descriptors. In Proc. CVPR, 2013.

[30] T. Trzcinski, M. Christoudias, V. Lepetit and P. Fua. Learning image descriptors with the boosting- trick. In NIPS, pages 278– 286, 2012.

[31] T. Trzcinski and V. Lepetit. Efficient discriminative projections for compact binary descriptors. In Proc. ECCV, 2012.