Max-SIFT: Flipping Invariant Descriptors for Web Logo Search

Lingxi Xie[1], Qi Tian[2] and Bo Zhang[1]

[1]Department of Computer Science and Technology, Tsinghua University, Beijing, China

[2]Department of Computer Science, University of Texas at San Antonio, Texas, USA

ABSTRACT

Logo search is widely required in many real-world applications. As a special case of near-duplicate images, logo pictures have some particular properties, for instance, suffering from flipping operations, e.g., geometry-inverted and brightness-inverted operations. Such operations completely change the spatial structure of local descriptors, such as SIFT, so that image search algorithms based on Bag-of-Visual-Words (BoVW) often fail to retrieve the flipped logos.

We propose a novel descriptor named Max-SIFT, which finds the maximal SIFT value sequence for detecting flipping operations. Compared with previous algorithms, our algorithm is extremely easy to implement yet very efficient to carry out. We evaluate the improved descriptor on a large-scale Web logo search dataset, and demonstrate that our method enjoys good performance and low computational costs.

NOVELTY

It is well known that SIFT is scale-and-rotation invariant but not flipping invariant. This might cause incorrect feature matches in some images, e.g., logos. In this paper, we propose a novel descriptor named Max-SIFT. Max-SIFT achieves flipping invariance by observing the impact on SIFT by geometry-inverted, brightness-inverted and geometry-and-brightness-inverted operations, and canceling out the flipping operation by performing a MAX operation on all the four candidates.

The advantages of the proposed descriptor are summarized as follows:

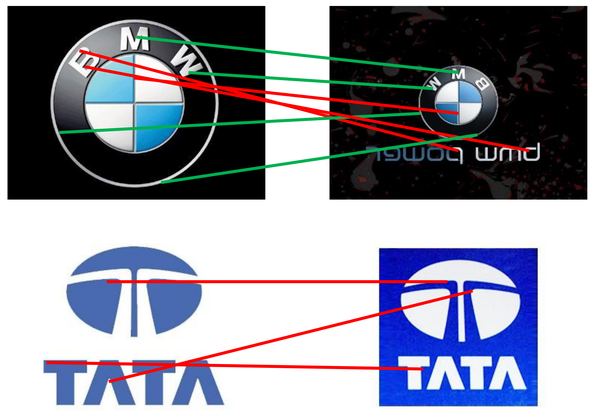
1. It is possible to find much more efficient matches using Max-SIFT, especially on the logo images which are often flipped.
2. Max-SIFT is fast to carry out, even when compared with previous candidates [3-5]. Only +1% extra time beyond SIFT is required.

Experimental results on a Web logo search dataset verify that the proposed descriptor is effective in finding true matches between original and flipped images, which is not able to be captured by original SIFT. Consequently, our algorithm produces consistent accuracy gain over a number of queries, evaluated by the mAP score.

The success of our algorithm suggests that flipping invariance is indeed important in image search tasks. It also implies the possibilities of applying flipping invariant descriptors onto other problems such as image classification.

THE PROPOSED ALGORITHM

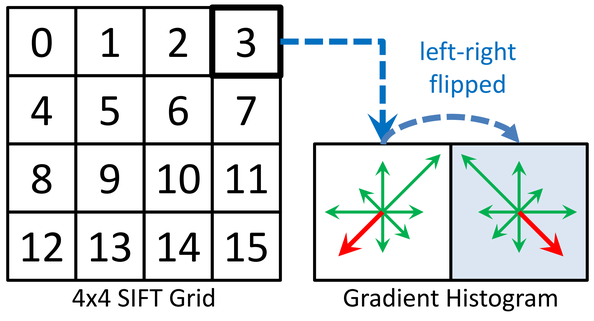
SIFT does not work!



The Max-SIFT Descriptor

- Compute original SIFT descriptors
- For each descriptor, computer its three (3) flipped copies by permutation on the 128 dimensions (see “Flipping a SIFT”)
- Obtaining the Max-SIFT descriptor by performing MAX operation on the four candidates (selecting the MAX one by **alphabetical order**)

Flipping a SIFT



0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Original SIFT Descriptor

3	2	1	0	7	6	5	4	11	10	9	8	15	14	13	12

Geometry-Inverted SIFT Descriptor

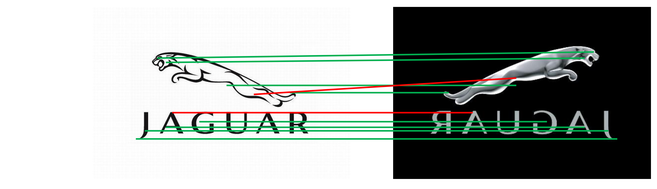
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0

Brightness-Inverted SIFT Descriptor

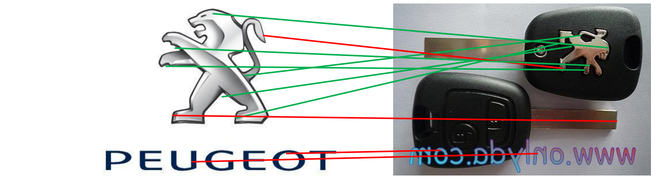
12	13	14	15	8	9	10	11	4	5	6	7	0	1	2	3

Geometry-and-Brightness-Inverted SIFT Descriptor

Representative Search Queries



114 (3) 33 (7) 17 (2) 317 (26) 183 (19)

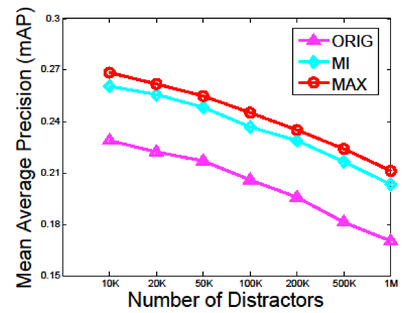


15 (0) 119 (6) 142 (12) 42 (7) 62 (27)

RESULTS

The CarLogo-51 Dataset

- A challenging dataset collected from the Web
- 51 car logos (brands), 11908 images
- Available at the first author’s homepage



Consistent accuracy gain over:
ORIG (original SIFT) and MI-SIFT [3].

REFERENCES

Key references are numbered as they appear in the paper.

[1] D.G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints,” International Journal on Computer Vision, 2004.

[3] R. Ma, J. Chen, and Z. Su, “MI-SIFT: Mirror and Inversion Invariant Generalization for SIFT Descriptor,” ACM International Conference on Image and Video Retrieval, 2010.

[4] X. Guo and X. Cao, “FIND: A Neat Flip Invariant Descriptor,” International Conference on Pattern Recognition, 2010.

[5] W.L. Zhao and C.W. Ngo, “Flip-Invariant SIFT for Copy and Object Detection,” IEEE Transactions on Image Processing, 2013.

[6] L. Xie, Q. Tian, W. Zhou, and B. Zhang, “Fast and Accurate Large-Scale Web Image Search via Graph Propagation and Search Process Tradeoff,” Computer Vision and Image Understanding, 2014.

[14] W. Zhou, Y. Lu, H. Li, and Q. Tian, “Scalar Quantization for Large Scale Image Search,” ACM Multimedia, 2012.

ACKNOWLEDGE.

This work was supported by the National Basic Research Program (973 Program) of China (Grant Nos. 2013CB329403 and 2012CB316301), the National Natural Science Foundation of China (Grant Nos. 61128007, 61332007, 61273023 and 91120011), the Beijing Natural Science Foundation (Grant No. 4132046), and the Tsinghua University Initiative Scientific Research Program (Grant No. 20121088071). This work was also supported in part to Dr. Qi Tian by ARO grant W911NF-12-1-0057, Faculty Research Awards by NEC Laboratories of America, and 2012 UTSA START-R Research Award, respectively.