Large Scale Partially Duplicated Web Image Retrieval

Wengang Zhou[1], Yijuan Lu[2], Houqiang Li[1], Yibing Song[1], Qi Tian[3]

Dept. of EEIS, University of Science and Technology of China[1], Hefei, P.R. China
Dept. of Computer Science, Texas State University at San Marcos[2], Texas, TX 78666
Dept. of Computer Science, University of Texas at San Antonio[3], Texas, TX 78249

zhwg@mail.ustc.edu.cn[1], yl12@txstate.edu[2], lihq@ustc.edu.cn[1], ybsong@mail.ustc.edu.cn[1], qitian@cs.utsa.edu[3]

ABSTRACT
The state-of-the-art image retrieval approaches represent images with a high dimensional vector of visual words by quantizing local features, such as SIFT, in the descriptor space. The geometric clues among visual words in an image is usually ignored or exploited for full geometric verification, which is computationally expensive. In recent years, partially duplicated images are prevalent on the web. In this demo, we focus on partial-duplicated web image retrieval, and propose a retrieval system based on a novel scheme, spatial coding, to encode the spatial information among local features in an image. Our spatial coding is both efficient and effective to discover false matches of local features between images, and can greatly improve retrieval performance.

Categories and Subject Descriptors
I.2.10 [Vision and Scene Understanding]: VISION

General Terms
Algorithms, Experimentation, Verification.

Keywords
Image retrieval, partial-duplicate, large scale.

1. INTRODUCTION
Due to the convenience of assisted image editing tools, partially duplicated images are prevalent on the web. Partial-duplicate web images are usually obtained by editing the original 2D image with changes in color, scale, rotation, partial occlusion, *etc*. Partial duplicates exhibit different appearance but still share some duplicated patches [4]. There are many applications of such a system to detect such duplicates, for instance, finding out where an image is derived from and getting more information about it, tracking the appearance of an image online, detecting image copyright violation, discovering modified or edited versions of an image, and so on.

In this paper, we will demonstrate a prototype system *DupSearch* to find the partial-duplicate versions of a query image in a large web image database. The framework of our system is illustrated in Fig. 1. We adopt SIFT features [1] and use Bag-of-Visual-Words model to represent images [2, 3]. Candidate target images are found though looking up the index, which is an inverted file structure [2] as shown in Fig. 2. For each SIFT feature, we store its orientation value and spatial position in the image. Features from different images quantized to the same visual word are considered as a match pair across the images. To remove those false matches as shown in Fig. 3, we propose a spatial coding scheme to encode the spatial information among local features and a spatial verification strategy to efficiently check the spatial consistency. Consequently, false feature matching can be identified to be removed and image similarities will be more accurately defined.

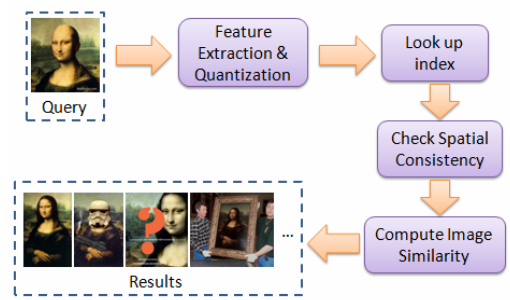


Figure 1. Framework of our DupSearch system.

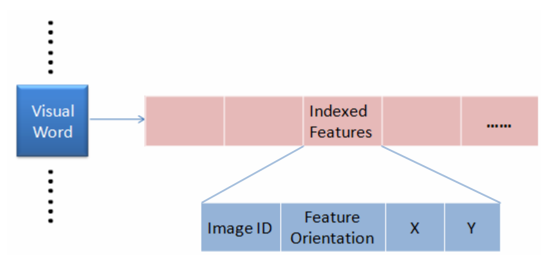


Figure 2. The inverted file structure.

2. DESCRIPTION OF DUPSEARCH
In our *DupSearch* system, after a query image is uploaded by a user, partially duplicated images will be efficiently identified and returned to the user. Fig. 3 illustrates the user interface for our duplicates retrieval system. The query image is shown on the left of the splitter bar, while the returned results are shown on the right, each with a ranking order shown bellow. The dataset size and the time cost for the corresponding query are shown in the status bar on the bottom. The demo is run on a laptop, with 400 thousand images indexed.

MM'10, October 25–29, 2010, Firenze, Italy.
ACM 978-1-60558-933-6/10/10.

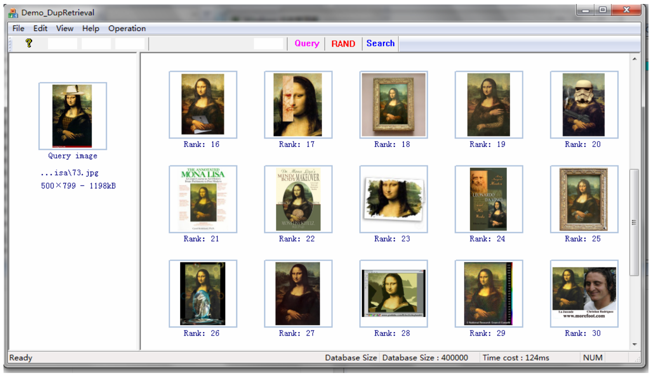
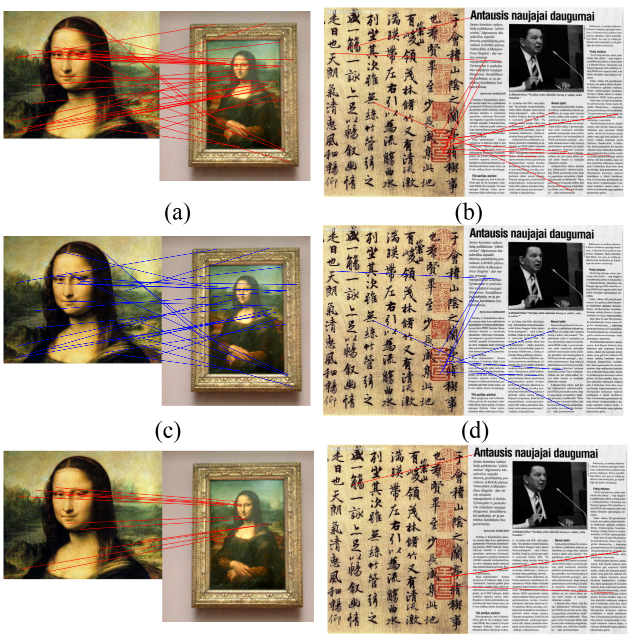


Figure 3. The user interface for partial-duplicates retrieval.

In our retrieval system, the similarity between query image and candidate target image is determined by the number of truly matched feature pairs. We remove those false feature matches by our spatial coding scheme.



(a) (b)

(c) (d)

(e) (f)

Figure 4. An illustration of spatial verification with spatial coding on a relevant pair (left column) and an irrelevant pair (right column). (a) (b)Initial matched feature pair after quantization; (c) (d) False matches detected by spatial verification; (e) (f) True matches after spatial verification.

Fig. 4 shows two instances of the spatial verification with spatial coding on a relevant image pair and an irrelevant image pair. Both image pairs have many matches of local features after quantization. For the left "Mona Lisa" instance, after spatial verification via spatial coding, those false matches are discovered and removed, while true matches are satisfactorily kept. For the right instance, although they are irrelevant in content, 12 matched feature pairs are still found after quantization. However, by doing spatial verification, most of the mismatching pairs are removed and only 3 pairs of matches are kept. Moreover, it can be observed that those 3 pairs of features do share high geometric similarity. In practice, by setting a threshold, images with irrelevant content can be easily filtered.

Fig. 5 shows the sample results from the 36th to the 50th with query "Star Bucks". Although the query image is polluted in the middle, target images with severe partial occlusion (such as the 39th and the 42th), scale changes. Editing due to rotation changes can also be identified with our approach.

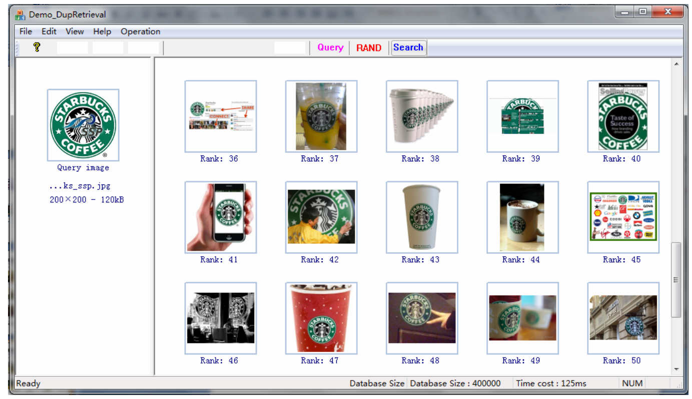


Figure 5. The sample search results with query "Star Bucks".

We perform the experiments on a server with 2.0G Hz CPU and 16G memory to index 1 million images that are most frequently clicked on a commercial image-search engine. Following the Tineye demo results (http://www.tineye.com/cool_searches), we collected and manually labeled 1100 partially duplicated web images of 23 groups from both Tineye [5] and Google Image search to build our ground truth dataset. For our approach the average query time cost is 0.49 second.

3. CONCLUSION

We have demonstrated *DupSearch*, a prototype system for large scale partially duplicated web image retrieval. Our DupSearch can overcome partial-occlusion and achieve invariance of scale rotation changes. Our future work will focus on better quantization strategy and distance metric learning, to further improve the search performance.

4. ACKNOWLEDGMENTS

This work was supported in part by NSFC under contract No. 60632040 and 60672161, Program for New Century Excellent Talents in University (NCET), Research Enhancement Program (REP) and start-up funding from the Texas State University.

5. REFERENCES

[1] D. Lowe. Distinctive image features from scale-invariant key points. *IJCV*, 60(2):91-110, 2004.

[2] J. Sivic and A. Zisserman. Video Google: A text retrieval approach to object matching in videos. In *Proc. ICCV*, 2003.

[3] D. Nister and H. Stewenius. Scalable recognition with a vocabulary tree. In Proc. CVPR, pages 2161-2168, 2006.

[4] Zhong Wu, Qifa Ke, Michael Isard, and Jian Sun. Bundling Features for Large Scale Partial-*Duplicate* Web Image Search. In *Proc. CVPR*, 2009.

[5] http://www. Tineye.com