

Rapid Object Search Engine for Contextual Advertisement

Yuning Jiang Junsong Yuan* Jingjing Meng
School of EEE, Nanyang Technological University
50 Nanyang Avenue, Singapore 639798
{ynjiang, jsyuan, jingjing.meng}@ntu.edu.sg

ABSTRACT

Visual object search, with the goal to find and locate the target object in large image or video collections, is of great interest for many applications and hence has received intensive attentions in recent years. In this demo, we present a spatial context-aware large-scale visual object search system, which is robust to cluttered backgrounds and can well handle scale variations of the objects. Different from the traditional image retrieval systems only matching individual points or fixed-scale spatial contexts, the proposed system considers spatial contexts of varying sizes and shapes, in the form of *randomized spatial partition* (RSP), and hence provides more accurate search results. Moreover, compared to the computational expensive RANSAC algorithm used in the state-of-the-art retrieval systems, the RSP framework lends our system to easy parallelization and significant speedup for object localization. Consequently, our system works accurately and efficiently. In addition, an Android application has been developed for mobile tasks, by which the user can take a photo of the object he/she wants and then search the same products and their selling information.

Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: Retrieval models

Keywords

Visual Object Search, Spatial Context, Randomized Spatial Partition

1. INTRODUCTION

Visual object search, with the goal to find and locate the target object in large image or video collections, is of great interest for many applications, including product search and recommendation, contextual ads, landmark search, video surveillance, automatic annotation, etc. Also because of its extensive applications, visual object search has received intensive attentions in recent years [2, 5]. However, despite lots of previous work, visual object search remains a very challenging problem.

*This work is supported in part by the Nanyang Assistant Professorship (SUG M58040015) to Dr. Junsong Yuan.

Copyright is held by the author/owner(s).
MM'12, October 29–November 2, 2012, Nara, Japan.
ACM 978-1-4503-1089-5/12/10.

The main challenges come from the two aspects: firstly, different from the traditional whole-image retrieval system, the object search system needs to handle the target objects, e.g., logos, which usually occupy only a small portion of an image with cluttered background, and can differ significantly from the query in scale, viewpoint and color. These all lead to difficulties in matching, and matching individual points [5] or fixed-scale contexts [7] can not provide satisfactory results; secondly, object localization is too time consuming for a real-time system. Most state-of-the-art systems locate objects through a post-processing step such as RANSAC [6] or subimage search [3], which is usually computationally expensive and applied only to several top candidates. Besides, the iterative RANSAC or subimage search is not straightforward to be parallelized, and thereby limits its application to large-scale datasets.

Our visual object search system addresses these two challenges by *randomized spatial partition* (RSP). RSP is proposed in [1] and specifically designed for small object search. Compared with the state-of-the-arts, our RSP-based visual search system offers several advantages: on one hand, by matching spatial contexts of varying scales and shapes, our system is more robust to the cluttered backgrounds as well as the object variations, and thus can provide more accurate results; on the other hand, our system can locate the target objects by simple thresholding on the confidence map, which is more efficient than RANSAC or subimage search. Moreover, because of the independence of each round of partition, the RSP-based system can be further speeded up with parallel implementation, thus it is well suited for large-scale applications.

2. DEMO

2.1 Server Environment

We demonstrate our visual object search system on the BelgaLogo database [2]. An overview of the system is shown in Fig. 1. Belgalogo is a very challenging logo database containing 10,000 images covering various aspects of life and current affairs. And to further verify the scalability and effectiveness of our system, we build a 0.1 million image database by adding crawled Flickr images to the BelgaLogo database as distractors. From this database, more than 200 million SIFT features [4] are extracted, and then clustered into 1 million visual words using the Hierarchical K-Means (HKM) method [6]. The system on the server end is performed on an HP PC with 3.10 GHz Intel CPU and 8 GB of RAM, and we implement the parallelization in 4 threads on CPU.

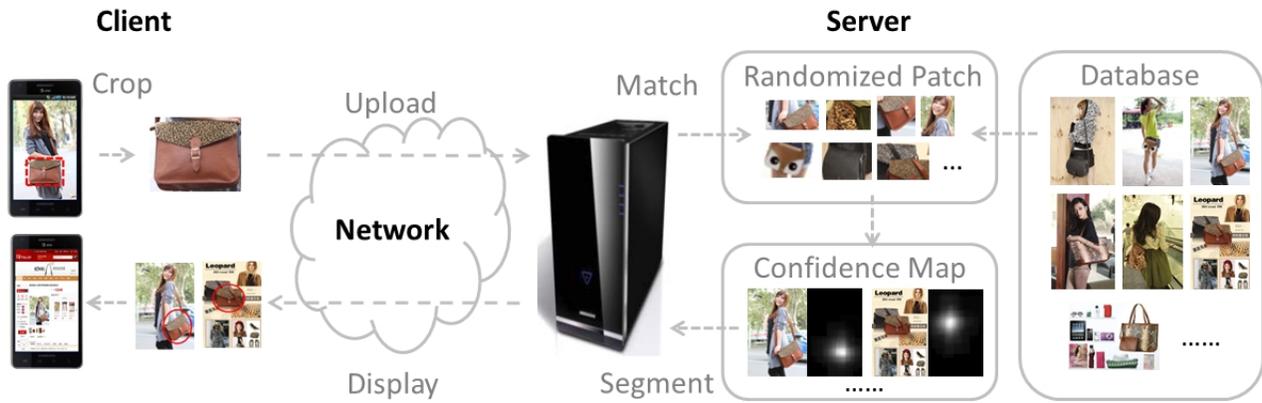


Figure 1: The framework of our RSP-based visual object search system.

2.2 Web-based Client

The interface of the web-based client is shown in Fig. 2. Given a query image as shown in Fig. 2(a), the user can crop any a region as the query object. After that, our system allows the user to make a trade-off between effectiveness and efficiency through adjusting the number of partition times on-line. In general, increasing the number of partition times leads to a more accurate search result but also more computational cost. A list of ranked images are returned as in Fig. 2(b). In each returned image, the detected object is highlighted in its original colors, while the backgrounds are denoted in gray. The confidence map of each returned image can be accessed via the *confidence map* link.



Figure 2: The interface of the web-based client. (a) is the query page while (b) is the results page.

2.3 Mobile-based Client

Since the mobile platforms have played a more and more important role in our daily life, we also develop a mobile-based client for our system, in the form of an Android application (see Fig. 3). With the mobile-based client application, the user can snap a picture containing the object he/she wants to search for, select the region of interest on the touch screen using simple gestures, and submit the query to our server to retrieve relevant images containing the same object. Coupled with our visual object search server, our mobile client can be the key enabler for various mobile applications such as product search and recommendation, education and tourism.

3. CONCLUSION

In this demo, we present a spatial context-aware visual

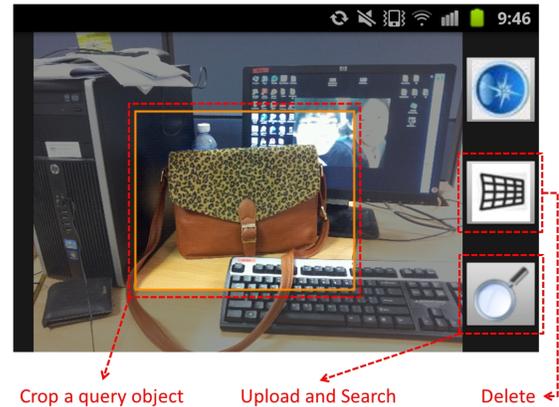


Figure 3: The interface of the mobile-based client.

object search system based on randomized spatial partition algorithm. Under the RSP framework, multi-scale spatial contexts are considered, which ensures the robustness of our system to cluttered backgrounds and object variations. Moreover, our system can efficiently locate the objects on the pixel-wise confidence map, and be further speeded up with parallel implementation. Since the proposed visual object search system has shown great performance in both accuracy and efficiency, we believe it has a good potential for other large-scale applications as well.

4. REFERENCES

- [1] Y. Jiang, J. Meng, and J. Yuan. Randomized visual phrases for object search. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2012.
- [2] A. Joly and O. Buisson. Logo retrieval with a contrario visual query expansion. In *Proc. ACM Multimedia*, 2009.
- [3] C. H. Lampert, M. B. Blaschko, and T. Hofmann. Beyond sliding windows: Object localization by efficient subwindow search. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2008.
- [4] D. Lowe. Distinctive image features from scale-invariant keypoints. *Intl. Journal of Computer Vision*, 2004.
- [5] J. Meng, J. Yuan, Y. Jiang, N. Narasimhan, V. Vasudevan, and Y. Wu. Interactive visual object search through mutual information maximization. In *Proc. ACM Multimedia*, 2010.
- [6] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman. Object retrieval with large vocabularies and fast spatial matching. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2007.
- [7] S. Zhang, Q. Tian, G. Hua, Q. Huang, and S. Li. Descriptive visual words and visual phrases for image applications. In *Proc. ACM Multimedia*, 2009.