Web Image Learning for Searching Semantic Concepts in Image Databases

Chu-Hong Hoi
Computer Science and Engineering Department
The Chinese University of Hong Kong
Shatin, Hong Kong SAR
chhoi@cse.cuhk.edu.hk

Michael R. Lyu
Computer Science and Engineering Department
The Chinese University of Hong Kong
Shatin, Hong Kong SAR
lyu@cse.cuhk.edu.hk

ABSTRACT

Without textual descriptions or label information of images, searching semantic concepts in image databases is still a very challenging task. While automatic annotation techniques are yet a long way off, we can seek other alternative techniques to solve this difficult issue. In this paper, we propose to learn Web images for searching the semantic concepts in large image databases. To formulate effective algorithms, we suggest to engage the support vector machines for attacking the problem. We evaluate our algorithm in a large image database and demonstrate the preliminary yet promising results.

Categories and Subject Descriptors

H.3.3 [**Information Systems**]: Information Search and Retrieval—*Clustering, Query Reformation, Relevance Feedback*

General Terms

Design, Algorithms, Experimentation

Keywords

Web Image Learning, Semantic Searching, Image Retrieval, Relevance Feedback, Support Vector Machine

1. INTRODUCTION

Along with the rapid development of multimedia devices and the internet, the amount of images have been dramatically increased in the past decade. Although Content-Based Image Retrieval (CBIR) has been studied for many years [3], searching semantic concepts in image databases is still a formidable task. Earlier approaches for CBIR are usually based on the Query-By-Example (QBE) strategy [3]. These approaches are inflexible since users may have difficulties in describing the query concepts. In general, searching by keywords is more easier to describe the query concepts than the QBE strategy. Recent research work begins to study the annotation techniques for attaching the textual labels to images [1]. However, fully automatic annotation techniques are yet a long way off.

In order to search the semantic concepts in images databases, we propose a scheme to engage Web images for learning the semantic concepts searching since the Web images associated with textual descriptions can serve as an important knowledge base. Our strategy is to search the semantic concepts by words from the Web and learn the returned Web images associated with the words. The Web images after filtering out the noisy images serve as the training set for learning in the image databases. The idea of our scheme is similar to a previous study which also proposed to engage the Web images for image recognitions [5]. With a different purpose, our interest is to investigate the issue for searching semantic concepts in image databases. Moreover, we propose to employ the Support Vector Machine (SVM) techniques [2] for attacking the learning tasks.

The rest of the paper is organized as follows. Section 2 presents our proposed learning scheme and the associated techniques. Section 3 presents the experimental results and discussions. Section 4 concludes our work.

WWW2004, May 17–22, 2004, New York, New York, USA.
ACM 1-58113-912-8/04/0005.

2. A LEARNING SCHEME FOR SEARCHING SEMANTIC CONCEPTS

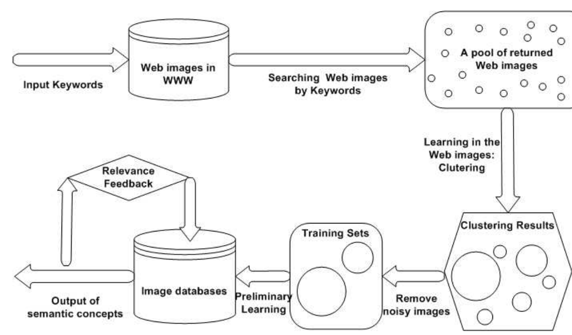

Figure 1: Overall Architecture of Our Proposed Scheme

Fig. 1 presents the overview of our proposed scheme for learning Web images to search the semantic concepts in image databases. We illustrate each step of our proposed system as follows.

2.1 Searching and clustering Web images

In our proposed system, a user first keys in words to represent their desired semantic concepts. Then, the proposed system searches the images on the Web which are associated with the related words. In our approach, we solicit Web image search engines to do this job. From the Web, we collect a pool of images which have textual descriptions related the semantic concepts. However, the image pool may contain many noisy images. Thus, we employ clustering techniques to remove the noises. Our strategy is to cluster the images into k clusters. Then, the top p clusters with the most images will be selected, and other clusters will be regarded as noises. The engaged clustering technique is based on the k-means algorithm.

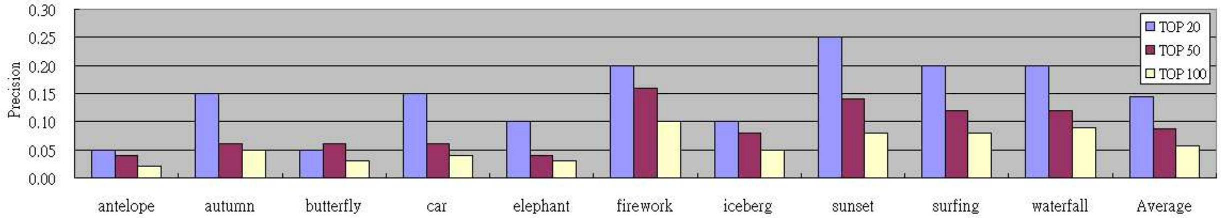

Figure 2: Experimental Results on 10 **Testing Semantic Concepts**

2.2 Learning semantic concepts by SVMs

After removing the noisy images, we can obtain a set of training images which roughly represent the semantic concepts. Then, we employ the SVM techniques to learn the semantic concepts in the image databases since SVMs provide good generalization performance and can achieve excellent results on pattern classifications problems [2].

In the preliminary searching round, we employ the One-class SVMs (1-SVM) to learn the training set of images in the database. 1-SVM is derived from classical SVMs for solving density estimation problems. After learning by 1-SVMs, we can obtain the preliminary searching results. Then, we employ the relevance feedback with two-class SVMs to improve the retrieval performance. Details for relevance feedback by SVMs can be found in [4].

3. EXPERIMENTAL RESULTS

In our image database, we collect $20,000$ images from the COREL image CDs which include 200 semantic categories, such as *antelope*, *car* and *sunset*, etc.

The image representation is an important step toward semantic learning in CBIR. We extract three features to represent the images: color, shape and texture. The color feature engaged is color moment, since it is closer to human natural perception. A 9-dimensional color moment is employed [3]. For the shape feature, edge direction histogram (EDH) is selected [3]. Canny edge detector is applied to obtain the edge images. The computed EDH from the edge image is quantized into 18 bins of 20 degrees each, hence an 18-dimensional EDH is used. For the texture feature, the wavelet-based texture is engaged. Discrete Wavelet Transformation (DWT) on the gray images employing a Daubechies-4 wavelet filter. We perform 3-level DWT decomposition and obtain ten subimages, in which nine subimages are selected to compute the entropies. Thus, a 9-dimensional wavelet texture feature is obtained.

To evaluate the performance of the proposed scheme in a large image database, we choose 10 semantic concepts, including *antelope*, *autumn*, *butterfly*, *car*, *elephant*, *firework*, *iceberg*, *sunset*, *surfing* and *waterfall*. To search Web images, we choose the Google Image Search Engine [1]. For each query semantic concept, top 40 returned imaged from Google were collected. For the clustering algorithm in our proposed scheme, we choose the parameters k=12 and p=4 in the k-mean algorithm. The kernel function used in SVMs is based on the Radial Basis Function [2]. Fig. 2 shows the experimental results. We observe that the average retrieval precision on TOP 20, TOP 50, and TOP 100 results is over 14%, 8%, and 5%, respectively.

The preliminary searching results are further improved by relevance feedbacks using SVMs. In each feedback round, 50 images are presented to users for judging their relevance. Table 1 shows the retrieval performance improved by 3-round relevance feedbacks. We can see that the average precision in TOP 20, TOP 50 and TOP 100 after 3-round feedbacks can achieve 58%, 32% and 18%, respectively.

Table 1: Average Retrieval Precision by Relevance Feedbacks

Feedback Round	TOP 20	TOP 50	TOP 100
No Feedback	14.5%	8.8%	5.7%
1 Feedback	29.0%	15.2%	15.4%
2 Feedback	47.0%	26.4%	16.1%
3 Feedback	58.5%	32.2%	18.3%

Although we have demonstrated promising results from the above experiments, we also notice limitation of our scheme. One disadvantage is that the preliminary retrieval performance will be sensitive to the quality of the collected images from the web, which is one of our future research tasks for investigation.

4. CONCLUSIONS

In this paper, we propose a scheme to learn Web images for searching semantic concepts in image databases. We suggest to employ the SVMs techniques to attack the learning tasks. We demonstrate the promising results and address the limitation of our scheme.

5. ACKNOWLEDGMENTS

The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CUHK4182/03E).

6. REFERENCES

[1] K. Barnard and D. Forsyth. Learning the semantics of words and pictures. In *Proc. International Conference on Computer Vision*, volume 2, pages 408–415, 2001.

[2] C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2):121–167, 1998.

[3] A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(12):1349–1380, 2000.

[4] S. Tong and E. Chang. Support vector machine active learning for image retrieval. In *Proc. ACM Multimedia*, pages 107–118, 2001.

[5] K. Yanai. Web image mining toward generic image recognition. In *Poster Proc. 12th International World Wide Web Conference*, 2003.

[1] http://images.google.com/