이미지 브라우징 처리를 위한 전형적인 의미 주석 결합 방법

주철화°, 왕령, 이양구, 류근호
충북대학교 데이터베이스/바이오인포매틱스 연구실
{thzhou, smile2867, leeyangkoo, khryu}@dblab.chungbuk.ac.kr

Clustering Representative Annotations for Image Browsing

Tie Hua Zhou°, Ling Wang, Yang Koo Lee, Keun Ho Ryu
Database/Bioinformatics Laboratory, Chungbuk National University, Korea

Abstract

Image annotations allow users to access a large image database with textual queries. But since the surrounding text of Web images is generally noisy, an efficient image annotation and retrieval system is highly desired, which requires effective image search techniques. Data mining techniques can be adopted to de-noise and figure out salient terms or phrases from the search results. Clustering algorithms make it possible to represent visual features of images with finite symbols. Annotation-based image search engines can obtains thousands of images for a given query; but their results also consist of visually noise. In this paper, we present a new algorithm Double-Circles that allows a user to remove noise results and characterize more precise representative annotations. We demonstrate our approach on images collected from Flickr image search. Experiments conducted on real Web images show the effectiveness and efficiency of the proposed model.

1. Introduction

Given huge amounts of image data, efficient and effective approaches for support image browsing and retrieval have become increasingly important. There are two general approaches for image retrieval: Text-based approach and Content-based approach. Many users prefer using keywords to conduct searches. Image annotation is an important problem that capabilities retrieving images from large image collections based on the semantic concepts or keyword annotations of images.

For Web images, there are rich textual features, such as name, alt text, Uniform Resource Locator (URL) and the surrounding text. In addition to the improved retrieval accuracy, another benefit for the multi-modal approaches is the added querying modalities. Web image search shares some common features with Web page search. For most of the queries, tens of thousands of images will be returned. For example as shown in Figure 1, few images retrieved by Flickr image search for the query "samsung", images of samsung are intermixed with images of samsung camera, samsung telephone, samsung TV, samsung computer, samsung car, samsung mp3 player, and so on. Annotation-based image search engines can obtains thousands of images for a given query; but their results also consist of visually noise.

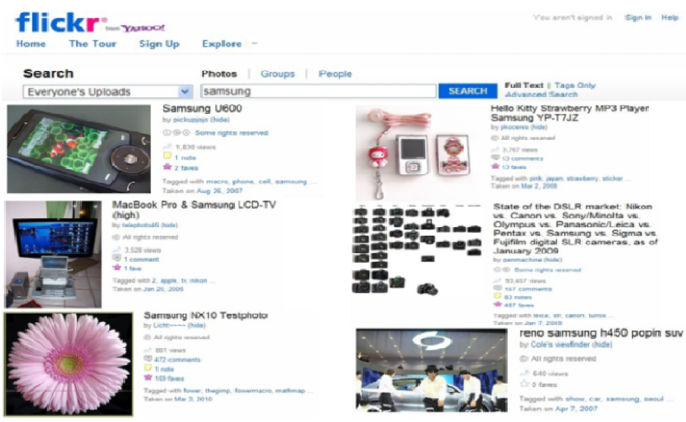


Figure 1: Web search by keyword "samsung"

These results show some common symptoms for web image retrieval: the image annotations are very noisy and the relevance of each result is highly variable. Since 2006, motivated by Web search technologies in many commercial systems, several search-based image annotation methods [1-3] have been developed, using Web-scale image database and unlimited vocabulary. However, most of work did not make a possible to represent visual features of images with finite symbols. Therefore, in order to solve the problems, we are focusing on second problem in this paper and propose the Double-Circles algorithm to find the most representative keywords as the annotations to image, and remove redundant keywords.

This paper is organized as follows: Section 2 discusses related work; Section 3 discusses the image retrieval scenario and shows proposed Double-Circles algorithm Section 4 shows the experimental analysis and results of this research. And finally concludes our work and future work.

2. Related Work

Several image search result clustering (SRC) algorithms have recently been proposed in the academic arena. A reinforcement clustering algorithm and a bipartite graph co-partitioning algorithm are proposed to integrate visual and textual features in [4, 5] respectively. Jing et al. [6] described IGroup that an image search engine based on image search results clustering, which only cluster the top few images using visual or textual features. The proposed algorithm first identifies several semantic clusters related to the query, based on a search result clustering algorithm for general Web search. It then assigns all the resulting images to the corresponding clusters. Wang et al. [7] proposed a novel annotation refinement algorithm to try to resolve the issues in Jin et al. [8]. In [7], by formulating the annotation refinement process as a Markov process and defining the candidate annotations as the states of a Markov chain, a content-based image annotation refinement (CIAR) algorithm is proposed to re-rank the candidate annotations.

Recently, Nasierding et al. [9] presents a novel multi-label classification framework that comprises an initial clustering phase that breaks the original training set into several disjoint clusters of data. It then trains a multi-label classifier from the data of each cluster. Given a new test instance, the framework first finds the nearest cluster and then applies the corresponding model. Morsillo et al. [10] presented a method for mining visual concepts from the web using minimal supervision, which is built upon a novel probabilistic graphical model which combines image features and text features from associated html documents. It introduced hybrid expectation maximization / expected gradient procedure for the model and showed that this semi-supervised approach gives better performance than a number of baseline tests. Schroff et al. [11] is another recent work which successfully retrieves clean image sets using image content alone.

3. Proposed Image Retrieval Algorithm

3.1 Image Retrieval scenario

Generally, the user will provide an example that represents their query concept, and the system will present some initial examples for the user to choose, the workflow as shown in Figure 2. This interaction with the users can also further refine the existing annotations. A major challenge faced how to mine annotations from the semantically and visually similar images, which is closely related to the efficiency of users' browsing.

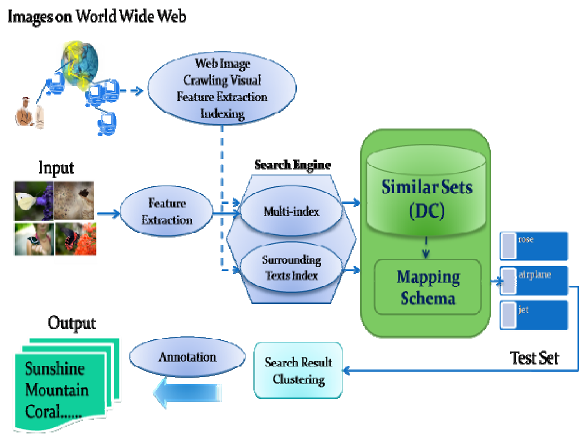


Figure 2: Proposed image retrieval scenario architecture

The main problems for image search consist of two components: the first one is not clear how much annotation is sufficient for a specific database, another one what the best subset of objects to annotate is. So in this paper, we are focusing on finding the most representative keywords as the annotations to the image, and remove redundant keywords.

3.2 Double-Circles Algorithm Analysis

We assumed there several subsets $\Omega = \{s_j \subseteq \Omega,\ j \in [1,u]\}$ partition by image category from image database Ω, so one keywords may be described different image in different subset. Under the Bayesian framework, the $P(w_i|s_j)$ (as shown in (1)) that w_i belongs to subset s is used as the discriminate function, and computed based on the class conditional density $f(s_j|w_i)$ and probability $P(w_i)$:

$$P(w_i|s_j) = \frac{f(s_j|w_i)p(w_i)}{f(s_j)} \qquad (1)$$

$$P(s_j|w_i) = \frac{f(w_i|s_j)p(s_j)}{f(w_i)} \qquad (2)$$

where we denote $p(s_j)$ as the probability density of sunset s_j and $f(s_j|w_i)$ as density of s_j conditional upon the assignment of annotation w_i. For any image feature X($P(w_i|s_j), P(s_j|w_i)$) as shown in (2), we consider $P(w_i|s_j)$ as the x-coordinate and $P(s_j|w_i)$ as the y-coordinate in the coordinate system and the annotated neighboring set as N_i (shown in Figure 3).

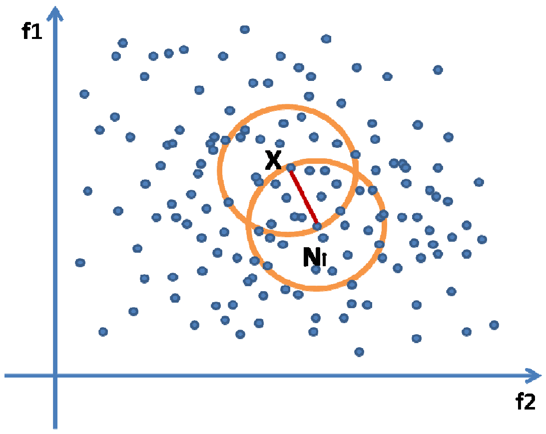


Figure 3: Annotated neighboring set of a test image X

Given an image set Ωs, we cluster the annotations of Ωs (i.e., title, color, etc.) to find the most annotations to I_{query}. The process of Double-Circles shown as follows:

Stage 1: give a certain threshold value to parameter k, and then base on this value to limit object size of each cluster.

Stage 2: chose random two points and make their distance as a certain radius to make two circles which the centre of each circle are two points mentioned before.

Stage 3: in each circle, if the size of cluster k' smaller than k, chose the most far point as a new centre tuple to make two new circles, otherwise restart stage 2.

Stage 4: make some new clusters by attribute relationships.

The Double-Circles-tree structure is shown in Figure 4.

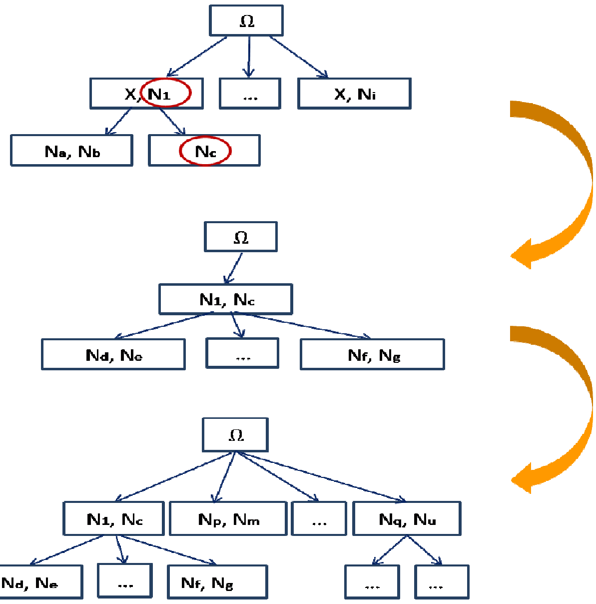


Figure 4: Double-Circles tree structure

4. Experiment and Evaluation

4.1 dataset acquisition

To evaluate the proposed algorithm, the dataset is a basic comparative dataset for recent research work in image annotation, which used in all experiments include images with annotations from online photo forums. There are 2,500 images from 25 Flickr CDs in this dataset, and divided into two parts: a training set of 2000 images and a test set of 500 images, and each CD includes 100 images on the same topic. An evaluation measure is proposed:

$$E = 2[(R_n + T_n)/C_n]$$

E value was used as the performance measure. C_n is defined as the number of correctly annotated images, R_n is defined as the number of retrieved images, and T_n is defined as the number of truly related images in test set.

4.2 Experimental Results

We vary the number of cluster DC equals to 2, 3, 5 and 7 respectively, series 1 presents the ratio that extracted classes account of all the original data sets; series 2 shows the ratio of most keywords set in the original datasets. As shown in Figure 5, the extracted classes decreased for the ratio of all the original datasets are became lower by DC increasing. But it's disadvantaged for semantic expression when the extracted classes more and effective relation keywords less. DC=7 have the other problem of the series 2. It contains so many keywords in one class, and these keywords maybe not only belong to one image type, it maybe has a lot of noisy data from other image types which are related to this image. So, in this case, when DC=3 and 5 are better results because of a smoothly changing.

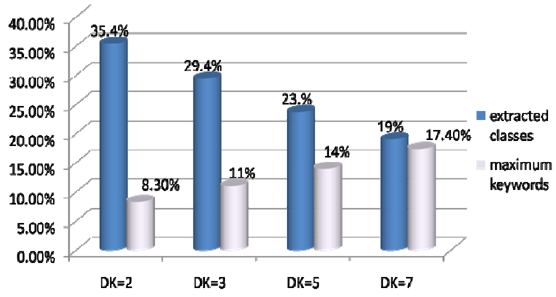


Figure 5: Extracted classes and maximum keywords

Figure 6 clearly shows the number of annotated words from one image by the Double-Circles algorithm. Obviously, our approach cluster correct annotations, and remove redundant keywords.

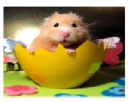
Image				
Original annotators	baby swim, iceland, baby, blue, water, swim, pool, diving, swimming	cute overload, cute hamster, hamsters, chmurka, syrian, animals, pets, rodents, chomik, hammie, pyza	orange,flowers, bathroom, wall, contrast, dyniss, orange flowers, orange bathroom	amanda,baby, girl, fouryears, flowerhat, brightpink, 50mm, child, girly, Babycappelli, kid,
Double Circles annotations	baby, swim, swimming, blue, water, diving	pets, animals, cute, hamster, hamsters	orange, flowers, wall	baby, girl, kid, child

Figure 6: A few examples of the annotation results yielded by Double-Circles algorithm

5. Conclusion

Image annotation is an important role in image retrieval system. Compared to the previous annotation approach, we are focusing on finding the most representative keywords as the annotations to the uncaptioned image, and remove redundant keywords. The experiments are conducted on Flickr Web image dataset, the experiments show Double-Circles algorithm is potentially applicable for the Web image annotations.

In our future work, we will try more Web search methods for image to refine the proposed approach and make the large-scale image annotation more possible and effective.

Acknowledgments. This research was supported by a grant (07KLSGC02) from Cutting-edge Urban Development – Korean Land Spatialization Research Project funded by Ministry of Land, Transport and Maritime Affairs, and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF No. 2010-0001732).

References

[1] C. Wang, F. Jing, L. Zhang, and H. J. Zhang. "ScalableSearch-Based Image Annotation of Personal Images. In Proc. ACM Inter. Workshop on Multimedia information Retrieval, pp. 26 – 27, 2006.

[2] X. J. Wang, L. Zhang, F. Jing, and W. Y. Ma. "AnnoSearch: Image Auto-Annotation by Search". In Proc. CVPR, pp. 17-22, 2006.

[3] C. H. Wang, F. Jing, et al. Image Annotation Refinement using Random Walk with Restarts. In proc. ACM MM, 2006.

[4] X. J. Wang, W.Y. Ma, L. Zhang and X. Li "Iteratively clustering Web images based on link and attribute reinforcements", Proc. of ACM Multimedia, pp. 122-131, 2005.

[5] B. Gao et al "Web image clustering by consistent utilization of visual features and surrounding texts", Proc. of ACM Multimedia, pp. 112-121, 2005.

[6] F. Jing, C.H. Wang, Y.H. Yao, K.F. Deng, L. Zhang, and W.Y. Ma "IGroup: Web Image Search Results Clustering" Proceedings of the 14th annual ACM international conference on Multimedia, pp: 377-384, 2006.

[7] C.Wang, F.Jing, L.Zhang, and H.J.Zhang "Image Annotation Refinement using Random Walk with Restarts", In Proceedings of the 14th Annual ACM international Conference on Multimedia, pp.647-650, 2006.

[8] Y.Jin, L.Khan, L.Wang and M.Awad "Image annotations by combining multiple evidence & wordNet", In Proceedings of the 13th Annual ACM international Conference on Multimedia, pp. 706-715, 2005.

[9] G. Nasierding, G. Tsoumakas and A. Z. Kouzani "Clustering Based Multi-Label Classification for Image Annotation and Retrieval", IEEE, 2009.

[10] N. Morsillo, C. Pal and R. Nelson "Mining the Web for Visual Concepts", Proceedings of the 9th International Workshop on Multimedia Data Mining, pp. 18-25, 2008.

[11] F. Schroff, A. Criminisi, and A. Zisserman, "Harvesting image databases from the web". In Proceedings of the 11th International Conference on Computer Vision, 2007.