Majority Based Ranking Approach in Web Image Retrieval

Gunhan Park, Yunju Baek, and Heung-Kyu Lee

Division of Computer Science
Department of Electrical Engineering & Computer Science
Korea Advanced Institute of Science and Technology
373-1 Kusung-Dong Yusong-Gu Taejon, 305-701, Republic of Korea
{gunhan,yunju,hklee}@rtlab.kaist.ac.kr

Abstract. In this paper, we address a ranking problem in web image retrieval. Due to the growing availability of web images, comprehensive retrieval of web images has been expected. Conventional systems for web image retrieval are based on keyword- based retrieval. However, we often find undesirable retrieval results from the keyword based web image retrieval system since the system uses the limited and inaccurate text information of web images ; a typical system uses text information such as surrounding texts and/or image filenames, etc. To alleviate this situation, we propose a new ranking approach which is the integration of results of text and image content via analyzing the retrieved results. We define four ranking methods based on the image contents analysis of the retrieved images; (1) majority-first method, (2) centroid-of-all method, (3) centroid-of-top K method, and (4) centroid-of-largest-cluster method. We evaluate the retrieval performance of our methods and conventional one using precision and recall graphs. The experimental results show that the proposed methods are more effective than conventional keyword-based retrieval methods.

1 Introduction

The rapid growing of web environment, and advances of technology have led us to access and manage huge images easily in various areas. The comprehensive retrieval of the image collections on the web become the important research and industrial issue.

The web image retrieval has different characteristics from typical content-based image retrieval(CBIR) systems. In general, web images have the related text annotations which could be obtained from the web pages where images are contained. So conventional web image retrieval systems utilize the text information of the images, and work as text(keyword) retrieval systems. Some systems use the texts and simple image information(e.g. image size, image format, graph/non-graph, etc.), and other systems provide the user input interface for relevance feedback.

Existing web image search systems allow users to search for images via keywords interface and/or via query by image example. Generally, the system

E. M. Bakker et al. (Eds.): CIVR 2003, LNCS 2728, pp. 111–120, 2003.

presents pages of representative thumbnail images to the user. The user then marks one or more images as relevant to the query. The visual image features for these images are then used in defining a visual query. However, it is often observed that there are many wrong results in high rank from the keyword-based image retrieval. Moreover, it is difficult to guarantee that there will be even one expected image shown in the initial page. Sclaroff called this the page zero problem[8].

To alleviate such a problem, we propose a new ranking approach that provides the better retrieval performance using image contents of retrieved results. Our approach is basically based on a integration of results of text and image contents via analyzing the retrieved results. The proposed approach determines the candidates using keyword first, and then automatically re-ranks images using visual features of retrieved results. We define four ranking methods based on the cluster analysis and majority of retrieved images. In experiments, we show that the proposed ranking approach improves retrieval performance of web image retrieval as compared to conventional one.

The paper is organized as follows. In Section 2, we briefly summarize the related work on web image retrieval. Our approach is described in Section 3. In Section 4, we present experimental results and discussions. Conclusions will be given in the last section.

2 Related Work

In recent years, there has been a number of research about CBIR systems. Most of the research has concentrated on feature extraction of an image, e.g., QBIC[2], VisualSeek[3], SIMPLicity[5], Netra[4], and Blobworld[6]. None of these systems provides a web search method; these systems are not based on textual cues. However, several systems have been developed for web image retrieval. These web image retrieval methods utilize different attributes; textual cues. PictoSeek[9] indexes images collected from the web. First, the system uses pure visual information, then it uses text information to classify the images. A similar system, Webseek[3] performs user helped classification. The system makes categories, and searches images within category, and provides category browsing and a search by example. Webseer[10] retrieves images using keywords, and additional image information that express the size of image, format of image, and simple classification information(e.g., graph, portrait, computer generated graphic, close-up, number of faces etc.). ImageRover[8] system allows the user to specify the image via keywords, an example image and relevance feedback. The ImageRover approach is most similar in spirit to that of WebSeer; however, ImageRover differs in that it allows searches of Web images based directly on image contents. ImageRover also proposed a method combining textual and visual cues using LSI(latent Semantic Indexing).

Generally, we can summarize the mechanism of conventional web image retrieval as follows : 1) the system retrieves the images using keywords or simple information about an image(not image contents such as color, texture, and

shape). 2) the system provides the interface to select relevant images from the first retrieved results(relevance feedback mechanism). 3) the system retrieves the images using selected images or/and keywords. 4) the system refines the results in repeating step 2 and step 3.

As shown in the previous systems keywords may help guide the search, and also become the important evidence in web image retrieval. Unfortunately, keywords may not accurately or completely describe image contents. Information about image contents directly from the image must also be added to retrieval processing. So in this paper, we propose a new approach that improve the retrieval performance using image contents analysis. Our approach will be described in detail in next section.

3 Majority-Based Ranking Approach

In this section, we describe a new ranking approach using image contents analysis. We also define four ranking methods based on the cluster analysis and majority of retrieved images. The difference from previous scheme with a relevance feedback mechanism is that we re-rank the results without assistance from the user(i.e. our approach is automatic). We will explain image features and clustering methods at first, and then we will explain the our approach using these features and the clustering methods.

3.1 Image Features and Clustering Methods

Various image features such as color, shape and texture have been developed in the literature. In a typical image retrieval model, image features are represented as a vector in a n-dimensional vector space. Color is an important attribute for describing the contents of image. Color histogram, that represents the proportion of specific colors in images, has been widely used among color representation methods. It has been known that color histogram in the CBIR provides reasonable retrieval performance when we use the HSV(Hue, Saturation, Value) color space, 128 quantization level, and the histogram intersection as a similarity function. The HSV color model is most frequently used for CBIR because it presents human perception well. The histogram intersection is calculated as follows; $H(I, I') = \sum_{i=1}^{n} min(f_i, f_i')/(\sum_{i=1}^{n} f_i')$. If the size of an image is same, histogram intersection is equivalent to the use of the sum of absolute differences or city-block metric[1]. In this paper, we use the city- block metric for similarity computation. City-block distance is defined as follows :

$$D_{city-block}(I, I') = \sum_{i=1}^{n} |f_i - f_i'| \tag{1}$$

Many clustering techniques for improving retrieval effectiveness have been proposed in the information retrieval literature, and also proposed in CBIR[11][12]. We use clustering methods to group the images and to select the representative image features for our approach.

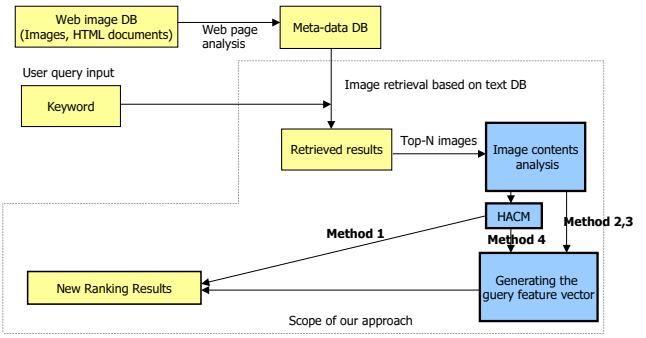


Fig. 1. Overview of the proposed approach

The hierarchical agglomerative clustering methods(HACM) are known to be effective in information retrieval applications. There are several methods to make a tree structure (it is usually called as a dendrogram[7]) for the HACM. We use group average link, and Ward's method among the clustering methods. The advantages of each method are well compared in the literature[13][14]. There are three approaches to the implementation of the general HACM. Among them, we use the stored matrix approach. We first calculate a N by N matrix containing all pairwise dissimilarity values using an association of measure function, and the Lance-Williams update formula makes it possible to re-calculate the dissimilarity between cluster centers using only the stored values. Eq. (2) shows the update formula, and Table 1 shows its parameters[7].

$$d_{C_{i,j}C_k} = \alpha_i d_{C_iC_k} + \alpha_j d_{C_jC_k} + \beta d_{C_iC_j} + \gamma |d_{C_iC_k} - d_{C_jC_k}| \tag{2}$$

3.2 Proposed Ranking Methods

In our web image retrieval, we utilize the retrieved results from keyword-based retrieval which is commonly used as web image retrieval systems; then the results are re-ranked with the proposed ranking approach. Fig 1 shows the architecture and the scope of our approach.

The brief explanation of proposed ranking approach is as follows; we analyze the top-N retrieval results, and re-ranks images according to the majority-based

Table 1. Lance-Williams parameters

HACM	α	β	γ
Group average link	$m_i/(m_i+m_j)$	0	0
Ward's method	$(m_i+m_k)/(m_i+m_j+m_k)$	$-m_k/(m_i+m_j+m_k)$	0

algorithms that we propose. Our basic hypothesis is that the more popular images have the higher probability to be desirable images. Based on this hypothesis, we propose the four methods that represent image contents for ranking as follows.

- **Method 1 (majority-first method)** : This method is using the majority property of retrieved images. For this method, we partition the retrieved images using HACM, and then we order the clusters according to the size of clusters. In other words, the largest cluster ranks first, and the sequence of clusters is determined as decreasing order of the size of cluster. After determining the order of clusters, we ranks the images within a cluster by distance to a centroid of the cluster.
- **Method 2 (centroid-of-all method)** : This method uses the centroid of the whole images of the retrieved results. Thus the centroid is represented as the average of retrieved images. Using this centroid as a query vector(a feature vector of a query image), the system is turned into conventional CBIR; the system ranks the images using a similarity function to this feature vector.
- **Method 3 (centroid-of-top-K method)** : This method uses the centroid of the K top-ranked images. Since there are many undesirable images in retrieved results, we only select some of top ranked images. We use 20 for K in the experiments. Like method 2, the centroid is used as a query vector for the CBIR system to re-rank the results.
- **Method 4 (centroid-of-largest-cluster method)** : In the fourth method, we use the centroid of the largest cluster as a query vector for image searching. In this method, we use the effect of the clustering to select a query vector. We assume the original rank is not important in this method as different from method 2 and 3.

In our methods, we define the centroid $C(A_I)$ of image set A_I as follows.

$$C(A_I) = \frac{\sum_{v \in A_I} \boldsymbol{v}}{|A_I|} \tag{3}$$

where $|A_I|$ is a size of A_I and $\boldsymbol{v}$ is a feature vector of an image.

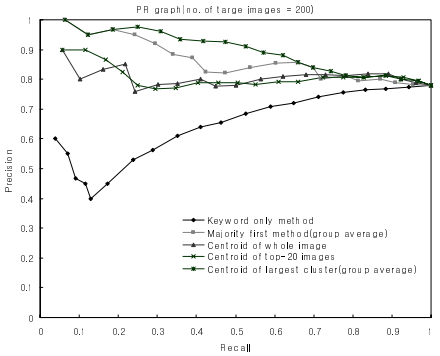
Using the four ranking methods and the two clustering methods, we evaluate the proposed approach in the next section.

4 Experimental Results

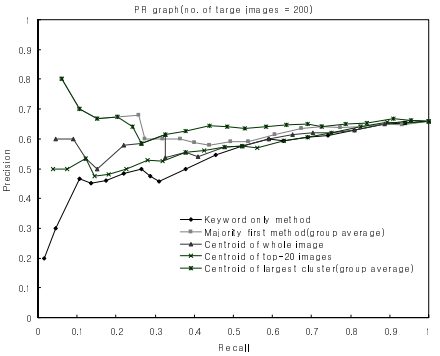
4.1 Experimental Environments: Test Collections

We conducted experiments using retrieved images from Naver[1], and Google[2] for some keywords : tiger, car, sea, etc. We gathered the top-200 images from results

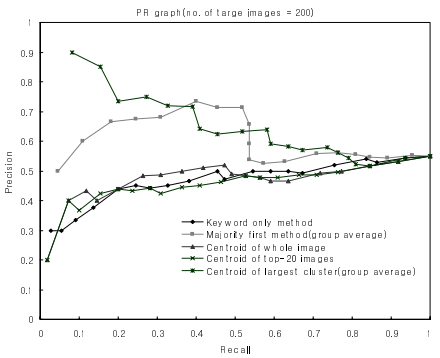
[1] http://www.naver.com is one of the most popular search engines in Korea. This search engine retrieve the relevant image among over 10 million ones on the web.
[2] see http://www.google.com



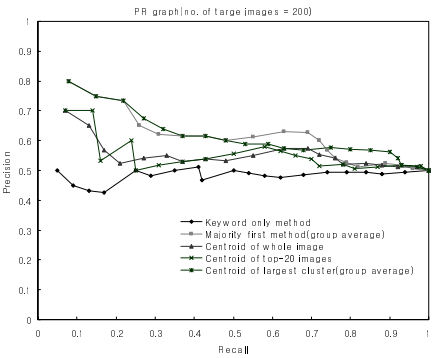
(a) Query "tiger" (data from Naver)



(b) Query "car" (data from Naver)



(c) Query "sea" (data from Naver)



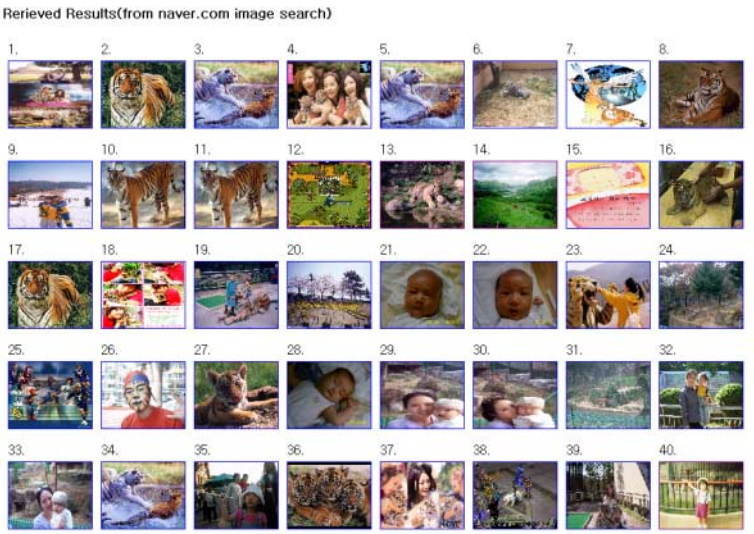
(d) Query "tiger" (data from Google)

Fig. 2. Comparison of retrieval effectiveness: Precision/Recall graphs

of a search engine for experiments, and evaluate the effectiveness in comparing the precision and recall value. Precision and recall are calculated as follows : precision = number of relevant retrieved images/total number of retrieved images, recall = number of relevant retrieved images/number of all relevant images. For the evaluation, we marked relevant images and irrelevant images manually about top-200 images. Naver basically use text annotations for image retrieval in web image album service, and Google use image filenames and frequency of user selection. Acquired images for our experiments are subject to change, but tendency of results is similar to our experimental data.

4.2 Results

The goal of the experiment is to evaluate the retrieval effectiveness of the proposed methods. The results of experiments are shown in Fig 2. The results show precision/recall graphs about initial method(keyword only retrieval) and our four



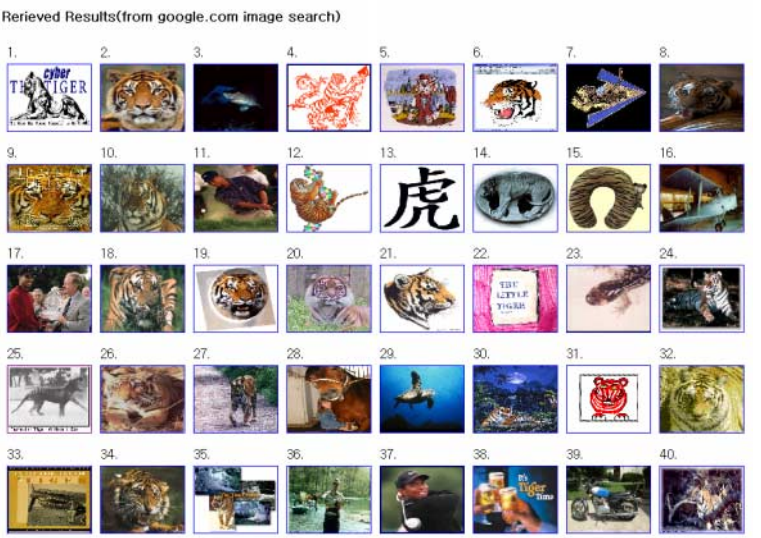
(a) the retrieved results from Naver

(b) the retrieved results using the proposed method(centroid-of-largest-cluster)

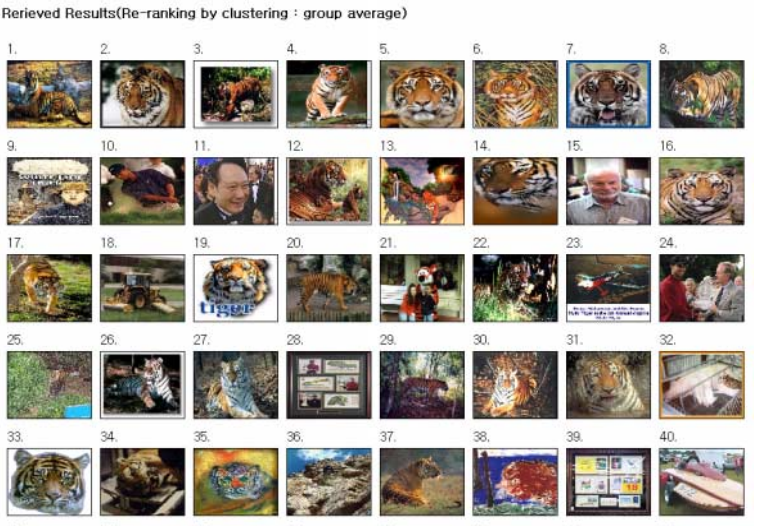
Fig. 3. The retrieval example using "tiger" (data from Naver)

methods : majority-first method, centroid-of-largest-cluster method, centroid-of-all method, and centroid-of-top-20 method. In the case of two methods based on HACM, the results are reported only for the method(group average link) with the better effectiveness.

The results in Fig. 2 show that the centroid-of-largest-cluster and majority-first methods have better effectiveness among four proposed methods. We think



(a) the retrieved results from Google



(b) the retrieved results using the proposed method(centroid-of-largest-cluster)

Fig. 4. The retrieval example using "tiger" (data from Google)

that the reason is because other methods contain the more wrong images in calculating the centroid, and the clustering methods reflect characteristics of images well.

The retrieval examples of the initial method and our methods are shown in Fig. 3, and Fig. 4. We used a centroid-of-largest-cluster method and group average link clustering for these examples. It is clear that relevant images to the

keyword query are ranked higher in the proposed methods than in the initial method.

It should be noticed that the effectiveness of the proposed method has improved significantly compared to the initial method. As shown in the results of experiments, in which case many relevant images are contained in the top-200 results the effectiveness of retrieval has improved significantly(Fig. 2(a)), while in the case of a few relevant images in the results the effectiveness has improved a little(Fig. 2(d)) relatively.

The overhead of this algorithm is that it has additional computation time for constructing the clusters. However, the algorithm has little added computational time(0.08 second for clustering, 0.02 second for ordering) since it performs on small number(200) of the images. We performed the experiments using Red Hat Linux 7.2 and a Pentium III 800 MHz system.

5 Conclusions

In this paper, we have proposed a new ranking approach which is the integration of results of text and image contents : majority based ranking approach. It has an advantage that it can use the contents of image in determining the rank of web images. We compared a keyword-based retrieval method and four proposed methods in our experiments. Experimental results show that the majority based approach, especially with the centroid-of-largest-cluster method, has better effectiveness than the initial method using only text evidence. Since our approach can use additional information of retrieved images, we believe that the majority based ranking approach will be a good effectiveness enhancement method compared to a general keyword-based retrieval method. In future work, we plan to apply other methods using various image features to our approach.

References

[1] M. J. Swain, and D. H. Ballad, "Color Indexing," International Journal of Computer Vision, Vol. 7, No. 1, pp. 11-32, 1991. 113
[2] C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W Niblack, D. Petkovic, and W. Equitz, "Efficient and Effective Querying by Image Content," Journal of Intelligent Information Systems, Vol. 3, No. 3/4, pp. 231-262, 1994. 112
[3] J. R. Smith, "Integrated Spatial and Feature Image Systems : Retrieval, Analysis and Compression," Doctoral Dissertations, Columbia University, 1997. 112
[4] W. Y. Ma and B. S. Manjunath, "NeTra: a toolbox for navigating large image databases", Multimedia Systems, Vol. 7, No. 3, pp. 184-198, 1999. 112
[5] J. Z. Wang, J. Li, and G. Wiederhold, "SIMPLIcity: Semantics-sensitive Integrated Matching for Picture Libraries,'" IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 9, pp. 947-963, 2001. 112
[6] C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Blobworld: Image Segmentation Using Expectation-Maximization and Its Application to Image Querying," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 8, pp. 1026-1038, 2002. 112

[7] W. B. Frakes, and R. Baeza-Yates, "Information Retrieval : Data Structures & Algorithms," Prentice Hall, 1992. 114
[8] S. sclaroff, M. la Cascia, S. Sethi, and L. Taycher, "Unifying Textual and Visual Cues for Content-Based Image Retrieval on the World Wide Web," Computer Vision and Image Understanding, Vol. 75, No.1/2, pp. 86-98, 1999. 112
[9] T. Gevers and A. W. M. Smeulders, "Pictoseek: A content-based image search engine for the WWW," Proceedings of International Conf. On Visual Information Systems, pp. 93-100, 1997. 112
[10] C. Frankel, M. J. Swain, and V. Athitsos, "WebSeer : An Image Search Engine for the World Wide Web," University of Chicago Technical Report TR-96-14, 1996. 112
[11] A. Tombros, R. Villa, R., and C. J. Van Rijsbergen, "The effectiveness of query-specific hierarchic clustering in information retrieval," Information Processing & Management, Vol. 38, No. 4, pp. 559-582, 2002. 113
[12] G. Park, Y Baek, and H. K. Lee, "A Ranking Algorithm Using Dynamic Clustering for Content-Based Image Retrieval,", the Challenge of Image and Video Retrieval(CIVR2002): International Conference on Image and Video Retrieval, pp. 316-324, 2002. 113
[13] E. M. Voorhees, "The cluster hypothesis revisited," Proceedings of 8th ACM SIGIR International Conference on Research and Development in Information Retrieval, pp. 188-196, 1985. 114
[14] P. Willett, "Recent trends in hierarchic document clustering : A critical review," Information Processing & Management, Vol. 24, No. 5, pp. 577-587, 1988. 114