Journal of Southeast University (English Edition) Vol. 21, No. 3, pp. 287 – 292 Sept. 2005 ISSN 1003—7985

Auto-expanded multi query examples technology in content-based image retrieval

Wang Xiaoling[1,2] Xie Kanglin[1]

([1]Department of Computer Science and Engineering, Shanghai Jiaotong University, Shanghai 200030, China)

([2]Institute of Information Management, Shanxi Financial University, Taiyuan 030016, China)

Abstract: In order to narrow the semantic gap existing in content-based image retrieval (CBIR), a novel retrieval technology called auto-extended multi query examples (AMQE) is proposed. It expands the single one query image used in traditional image retrieval into multi query examples so as to include more image features related with semantics. Retrieving images for each of the multi query examples and integrating the retrieval results, more relevant images can be obtained. The property of the recall-precision curve of a general retrieval algorithm and the K-means clustering method are used to realize the expansion according to the distance of image features of the initially retrieved images. The experimental results demonstrate that the AMQE technology can greatly improve the recall and precision of the original algorithms.

Key words: content-based image retrieval; semantic; multi query examples; K-means clustering

Due to the rapid growth of the large capacity image databases, image retrieval has become an important issue involved in computer vision, image databases and knowledge mining in recent years. Image retrieval aims to retrieve similar or relevant images to the query images by means of the image features or the keywords related to the query images [1]. In the past, various approaches to image retrieval were proposed, most of which were content-based image retrieval (CBIR) that derived the image features such as color, texture and shape or any combination of them.

However, the semantic gap that exists between the high level semantic and the low level features of an image has degraded the efficiency of the CBIR technology[1]. Actually, images with the same semantics may have very different appearances (low level features). For example, "dog" images may include various breeds of dogs of different color, height and shape. Yoon and Jayant[2] also declared this situation with a "bear" images example existed in the image retrieval clearly. The traditional retrieval method hopes to retrieve all of the "dog" images based on one single query "dog" images, which can only retrieve those very strongly "dog" images similar to the query image and inevitably lose other breeds of "dog" images. Evidently, if more images of different "dog" can be used as the query images in the retrieval, more relevant "dog" images will definitely be retrieved. Field-oriented image retrieval methods use a so-called feature template to represent the essential feature of an object[3,4]. However, one problem with the feature template is that it lacks robustness to the feature variance, since one template cannot cover all the necessary content features. The multi query example retrieval technique is an effective method to solve this problem. Natsev and Smith[5] described the multi example retrieval technique in detail and studied methods for active selection of query examples and query features. Nastar et al. [6] introduced a system named surfimage which used a query-by-example approach for retrieving images. The key issues of the multi query example method include the selection of query examples and the fusion of the retrieval results. In this paper, we propose an auto-extended multi query examples approach to narrow the semantic gap. We make use of the property of the recall-precision curve of a general algorithm and the K-means clustering method to realize the automatic selection of query examples according to the distance of image features. Furthermore, we also consider the fusion of the retrieval results. We iterate the traditional retrieval to each of the multi query examples, which we call feedback retrieval. The experiments demonstrate the feasibility and efficiency of our proposed algorithm.

1 Principle of Auto-Extended Multi Examples Retrieval Technology

Most recall-precision curves of the current retrieval approaches have the following property: a low-

Received 2004-11-09.

Foundation item: The National High Technology Research and Development Program of China (863 Program) (No. 2002AA413420).

Biographies: Wang Xiaoling (1975—), female, graduate; Xie Kanglin (corresponding author), male, professor, xie-kl@mail.cs.sjtu.edu.cn.

er recall rate, for example between 0.1 and 0.3, often relates with a quite higher precision rate, for example about 0.9. It means that at the beginning of the retrieval result, most of the images are real relevant images to the query image. They share similar image features to the query image. At the same time, there exist some differences among these features. That is the key point for our proposed algorithm: the same parts of these features guarantee that they are relevant images to the query image while the different parts of the image features provide us the opportunity to retrieve other possible relevant images. We just select proper images from the beginning of the retrieval result to extend the query image since they are probably the real relevant images. Then we iterate the traditional retrieval based on the multi query examples so as to retrieve more relevant images.

The principle of our idea is illustrated in Fig. 1. The initial query image (represented by a circle in Fig. 1) has features A and B whose proportions are shown in the circle. The retrieved relevant images in the initial retrieval list have features C, D and E etc. apart from features A and B. We only need to select proper images from the initial retrieval list as the multi query examples and then conduct the retrieval again, which we call the "first time feedback retrieval". If necessary, we can retrieve images for those retrieved from the first time feedback retrieval, which we call the "second time feedback retrieval". Our proposed auto-extended multi query examples retrieval approach consists of the above two feedback retrievals. Through the feedback retrieval, we will acquire new relevant images such as the images denoted by the bold circles in Fig. 1, which have different proportions of contents features (we call them feature shift) from the images in the initial retrieval. The shift of the image features provides us the opportunity to find other relevant images. Of course, such shift must be confined within some restriction so as not to be too "far" away from the query image. How to select proper images from

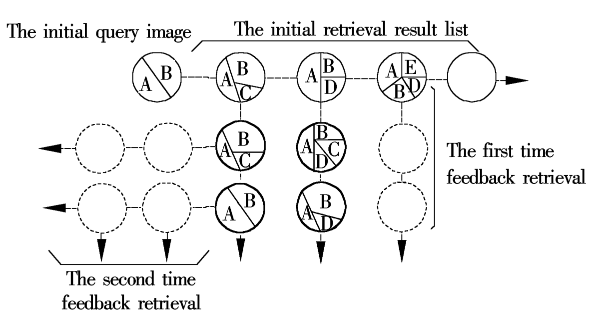

Fig. 1 Principle of the auto-extended multi query examples

the retrieval result and establish the final retrieval result will be discussed in the following sections.

2 Algorithm Descriptions

The auto-extended multi query examples approach deals with the following two problems.

2.1 How to select images to build the multi query images

In content-based image retrieval, the similarity is determined by the distance D between the query image Q and image I from the image database acoording to the feature $\boldsymbol{F}$.

$$D(I, Q) = d(\boldsymbol{F}_I, \boldsymbol{F}_Q) \tag{1}$$

The Euclidean distance is one of the most frequently used distance methods. The most similar image to the query image Q is the nearest one to it. In the experiments, we find that the distance D maintains a certain value especially for a group of strongly similar images. As an example, we adopt the color feature represented by the HSV color histogram to retrieve images.

In Fig. 2, the retrieved images numbered from 2 to 7 are strongly similar and their distances to the query image 5.jpg maintain about 0.7. The distances of the images numbered from 8 to 15, which also have similar appearances to each other, maintain about 0.6. Their distances to the query image 5.jpg are shown in Fig. 3. Here, we adopt the histogram intersection matching method. So the more the distance approaches to 1, the more similar the image to the query image.

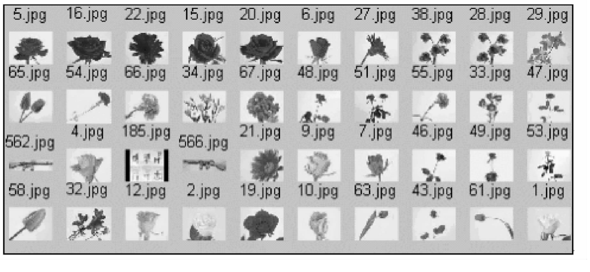

Fig. 2 Retrieval result of the image named 5.jpg using the HSV histogram

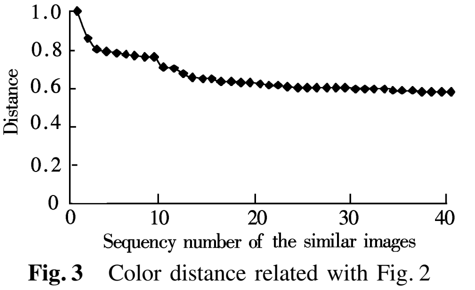

Fig. 3 Color distance related with Fig. 2

Such a property of the recall-precision curve of a general algorithm can be exploited to generate the multiquery examples. The followings are the details.

Step 1 Adopt the K-means clustering method to extend the query image. Suppose there are M images in the image database altogether and the retrieval result list L consists of N images determined by the feature distance D as in Eq. (3):

$$L = \{I_i \mid i = 1, 2, \ldots, N; N \leqslant M\} \tag{2}$$

$$D(I, Q) = \{d_j \mid j = 1, 2, \ldots, N; N < M\} \tag{3}$$

We adopt the K-means clustering method to select those images with a close distance to the query image Q. The K-means clustering method is one of the most frequently used clustering methods[7]. The K-means clustering method classifies a vector $\boldsymbol{X}_j (j = 1, 2, \ldots, n)$ into K classes $C_i (i = 1, 2, \ldots, k)$ by minimizing an objective function. Here, we use the square-error objective function as in Eq. (4):

$$J = \sum_{i=1}^{k} \sum_{x \in C_I} \|x - m_i\|^2 \tag{4}$$

where m_i is the mean of the class C_i. We select m images from the initial retrieval list L as the multi query examples Q_F:

$$Q_F = \{I_f \mid f = m, m \leqslant N\} \tag{5}$$

where m is determined by a two-clustering operation on D as in Eq. (6)

$$D(I, Q) = \{d_m, d_n \mid m = 2, 3, \ldots, k; n = k+1, k+2, \ldots, N\} \tag{6}$$

The objective function is

$$J = \sum_{i=1}^{2} \sum_{d \in C_i} \|d - m_i\|^2 \tag{7}$$

where m_i is the mean of the class $C_i (i = 1, 2)$. Based on the above analysis, the K-means clustering method needs to cluster N images into two classes: the most possibly relevant images and the not too relevant images to the query image according to the distance D. The most possibly relevant images will be used as the multi query examples. Since the retrieval result list has been sorted, the sequence number of the images in both clusters is continual.

Step 2 Next, we will retrieve M images to each image I_K numbered $K (K \leqslant N)$ in Q_F, which we call the first time feedback retrieval. Then we select s images from the retrieved images to form the retrieval list $L_F(I_K)$ related with image I_K.

$$L_F\left(\underset{I_k \in Q_F}{I_k}\right) = \{I_j \mid j = s, s \leqslant N, K \leqslant N\} \tag{8}$$

where s is determined by Eq. (9). The K-means clustering method classifies the distance D of the retrieved N images to image I_K into two classes. The reason why we choose the two clusters operations is the same as step 1.

$$D(I, I_K) = \{(d_s, d_t) \mid s = 1, 2, \ldots, l; t = l+1, l+2, \ldots, N\} \tag{9}$$

Since we hope to acquire "new" images in the feedback retrieval, if after the first time feedback retrieval, no new images are acquired and $k \leqslant 2$, then for the images numbered within [3, N], apply the K-means clustering method again as step 1. The meaning of the "new" image is explained in section 2.2.

Step 3 If $k \geqslant \eta$, we still have not gained any new images, so then we stop the feedback retrieval. This is because the images too far away form the query image in L are possible irrelevant ones as illustrated in Fig. 3. By the experiment, we find that $\eta = [N/2]$ is proper.

Step 4 From the above analysis, we know that those images close to the query image are possibly relevant ones. Therefore, we select those retrieved images numbered no more than δ in the first time feedback retrieval result to conduct a second time feedback retrieval. By the experiment, we found that $\delta = [N/4]$ is proper.

2.2 How to integrate the "new" images into the initial retrieval list

After we retrieve images to each of the the multi query examples, we need to integrate them properly. First, we compare them with the original retrieval result to determine the new relevant images and then insert them into the initial retrieval result.

1) Select the new images I_{new} by comparing the images numbered from 1 to t in the initial retrieval list L with the images in the feedbck retrieval list L_F.

$$I_{\text{new}}(I_K) = \{L \cap L_F\} = \{I_i \cap I_j \mid i = 1, 2, \ldots, t; j = s; t \leqslant N; s \leqslant N; K \leqslant N\} \tag{10}$$

where

$$t = \begin{cases} 10r\left(\dfrac{k+j}{10}\right) & k+j > 10 \\ 10 & k+j \leqslant 10 \end{cases} \tag{11}$$

r is a round off function, s is determined by Eq. (9), t is proportional to the summation of K and j, which guarantees that the comparison is conducted nearby the neighborhood of image I_K, see Eq. (10). New images may be different from the first M images in L.

2) Through feedback retrieval, we may get some new images that are different from the images in the initial retrieval list L. How can they be arranged into the final retrieval result? As we know, in essence, the content-based image retrieval is still dertermined by the distance of the content features. At the same time, the multi query examples used in the feedback retriev-

al come from the initial retrieval list L. Therefore, the new images I_{new} should be inserted into the initial retrieval list L following the k-th image to form the retrieval list L' after the first time feedback retrieval.

3) Since the new image $I_{new}(I_K)$ is obtained by comparing L with $L_F(I_K)$ individually, the new images related with different query images I_K may be duplicate images. So we must filter them in L' and keep the images with a smaller sequence number.

For the new images generated from the second time feedback retrieval, the integration is the same as the first time feedback retrieval.

We summarize the key steps of our proposed auto-extended multi query examples approach as follows:

① Use the K-means clustering method to select proper images automatically from the initial retrieval result as multi query examples;

② Retrieve images for each one of the multi query examples;

③ Re-use the K-means clustering method to select new images from the retrieved images in ②;

④ Integrate new images generated in ③ into the final retrieval result.

3 Experiments

In this paper, 1 500 images including 18 categories of images (rock, river, leaf, flower, cloud, car, mobile phone, desk, gun, face, cat, dog, plane, shoes, scenery, hand, lamp and teapot) downloaded from the web site http://photo.ayinfo.ha.cn/ are used in the experiment.

We choose the HSV color histogram retrieval method and the direction chain code retrieval method[8] in our experiment. The details of the color space conversion from RGB to HSV can be found in Ref. [9]. Our object is to verify the effectiveness of the auto-extended multi query examples approach for the two retrieval methods. One of the retrieval details is as follows:

1) The initial query image is named 5.jpg and its initial retrieval result list with 40 images ($N=40$) is shown in Fig. 2. According to the K-means clustering method, the images numbered from 2 to 7 ($s=2, 3, \ldots, 7$) are selected as the multi query examples to be used in the feedback retrieval, that is $Q_F = \{16.jpg, 22.jpg, 15.jpg, 20.jpg, 6.jpg, 27.jpg\}$.

2) Retrieve images for each image in the Q_F. For instance, when the query image is $I_k = 16.jpg$ ($K = 2$), we apply the K-means clustering method to cluster its retrieved images and obtain $j = s = 1, 2, \ldots, 8$. Then the images numbered from 1 to 8 in $L_F(I_2)$ are selected to compare with the images numbered from 1 to 10 (for $k+j \leqslant 10$, $t=10$) in the initial retrieval list L to generate new images. The retrieved new image is 67.jpg and it should be inserted into the initial retrieval list L following the second image for $K=2$. After the same processing as the other seven images, their respective new images are as follows: $I_{new}(I_2) = \{67.jpg\}$, $I_{new}(I_3) = \{17.jpg\}$, $I_{new}(I_4) = \{67.jpg, 19.jpg, 21.jpg, 31.jpg\}$, $I_{new}(I_5) = I_{new}(I_6) = \{58.jpg, 63.jpg\}$, $I_{new}(I_7) = \varnothing$, $I_{new}(I_8) = \{58.jpg\}$. Among them, 17.jpg and 31.jpg are completely new images while 67.jpg, 19.jpg and 21.jpg have appeared in the initial retrieval result L and they will be moved ahead in L. Update L by inserting all the new images into the corresponding position in L.

3) Filter the duplicated images in L' and keep the one with the smallest sequence number.

4) There are five new images with sequence number no more than 10 ($\delta = [N/4] = 10$) in L'. They are 67.jpg, 17.jpg, 19.jpg, 21.jpg and 31.jpg. Then in the second time feedback retrieval, we need to retrieve similar images to these images and obtain the following new images: $I_{new}(67.jpg) = \{21.jpg, 66.jpg, 19.jpg, 31.jpg, 9.jpg, 65.jpg\}$, $I_{new}(17.jpg) = \{8.jpg, 147.jpg, 22.jpg, 68.jpg, 134.jpg, 121.jpg, 7.jpg\}$, $I_{new}(19.jpg) = \{21.jpg, 67.jpg, 18.jpg, 10.jpg, 25.jpg\}$, $I_{new}(21.jpg) = \{67.jpg, 19.jpg, 1.jpg, 9.jpg\}$, $I_{new}(31.jpg) = \{21.jpg, 19.jpg, 18.jpg, 147.jpg, 559.jpg, 69.jpg, 56.jpg\}$.

5) Insert new images obtained in step 4) into the initial retrieval list L and filter them. The concrete processing is the same as the first time feedback retrieval.

Finally, the ultimate retrieval result after applying the auto-extended multi query examples approach is illustrated in Fig. 4. Evidently, the retrieval result includes more relevant images to 5.jpg than that in the initial list shown in Fig. 2. The relevant images include 19 images (5.jpg, 16.jpg, 67.jpg, 21.jpg, 66.jpg, 19.jpg, 17.jpg, 8.jpg, 141.jpg, 22.jpg, 134.jpg, 121.jpg, 7.jpg, 15.jpg, 18.jpg, 25.jpg, 20.jpg, 6.jpg, 27.jpg) while in Fig. 2, the number is 9 (5.jpg, 16.jpg, 22.jpg, 15.jpg, 6.jpg, 27.jpg, 67.jpg, 21.jpg, 19.jpg). Compared with the initially retrieved flower images, our newly retrieved images vary in the color and size of petals, for example some scarlet flowers with larger size in Fig. 4.

There are two principles to evaluate a retrieval

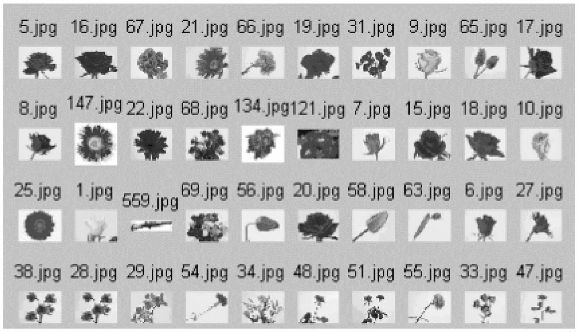

Fig. 4 Retrieval list after the second time feedback retrieval

approach: the precision and the recall measures. Suppose $R(q)$ is the set of images relevant to the query q and $A(q)$ is the set of retrieved images. The precision of the result is the fraction of retrieved images that is truly relevant to the query q,

$$P = \frac{|A(q)| \cap |R(q)|}{|A(q)|} \tag{12}$$

while the recall is the fraction of relevant images that are actually retrieved,

$$R = \frac{|A(q)| \cap |R(q)|}{|R(q)|} \tag{13}$$

Fig. 5 illustrates the performance improvement of the HSV histogram method and direction chain code method using our auto-extended multi query examples approach. The average retrieval performance has been improved by above 10% for the HSV histogram method. Although for direction chain code method the precision is degraded when recall is less than 0. 3, the average retrieval performance has also been improved by 15%. Their performance improvement is evident, which verifies the efficiency and feasibility of the auto-extended multi query examples approach. In the experiments, we have the following observations:

1) When the initially retrieved images are very strongly similar, it is difficult to find other new images by the auto-extended multi query examples approach for these images located very closely in the feature space.

2) When the initially retrieved images are weakly similar, by the auto-extended multi query examples approach, the performance can be improved greatly. The different parts of their features lead us to find more new relevant images.

3) When there are very few relevant images in the initial retrieval result, by the auto-extended multi query examples approach, we have the opportunity to retrieve more relevant images to improve the performance greatly. On the other hand, we may acquire similar performance as the original approach if we cannot find any new relevant images.

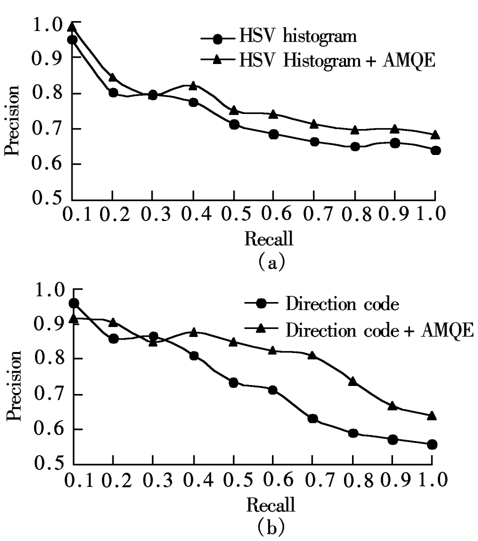

Fig. 5 Performance improvement using the auto-extended multi query examples approach. (a) HSV histogram; (b) Direction code

4 Conclusion and Future Work

The core of our proposed auto-extended multi query examples approach is that it makes use of the feature relations that exist among the relevant images in CBIR. The experiment demonstrates that with this approach, the retrieval ability of a simple feature such as the color or shape can be improved and the semantic gap can be narrowed to some extent.

Images with complicated semantics may have very different appearance. If these images share no mutual features completely, our proposed approach does not work. So the semantics suitable for our approach cannot be too complicated. Yoon and Jayant[2] give an example of two breeds of bear images, one of which is a white bear and the other a black bear. If retrieved by color, it is difficult with our proposed method to find them since they are located far away in the color feature space, especially when the database is large. In future, we consider importing relevance feedback technology[10] to point out the relevant images, which breaks the limit of the feature distance.

References

[1] Smeulders A W M, Worring M, Santini S, et al. Content-based image retrieval at the end of the early years [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, **22**(12): 1349 – 1380.

[2] Yoon Janghyun, Jayant Nikil. Relevance feedback for semantics based image retrieval [A]. In: *Proc of the IEEE International Conference on Image Processing* [C]. Thes-

saloniki, Greece, 2001. 42 - 45.

[3] Smith John R, Li Chung-Sheng. Image classification and querying using composite region templates [J]. *Computer Vision and Image Understanding*, 1999, **75**(1, 2): 165 - 174.

[4] Zhuang Y, Liu X, Pan Y. Apply semantic template to support content-based image retrieval [A]. In: *Proc of the SPIE Storage and Retrieval for Multimedia Database* [C]. San Jose, CA, USA, 2000. 442 - 449.

[5] Natsev Apostol, Smith John R. Active selection for multi-example querying by content [A]. In: *Proc of the IEEE International Conference on Multimedia and Expo* [C]. Baltimore, MD, USA, 2003. 445 - 448.

[6] Nastar C, Mitschke M, Meilhac C, et al. Surfimage: a flexible content-based image retrieval system [A]. In: *Proc of the ACM Multimedia* [C]. Bristol, England, 1998. 339 - 344.

[7] Han Jiawei, Kamber Micheline. *Data mining concepts and techniques* [M]. Translated by Fan Ming, Meng Xiaofeng. Beijing: China Machine Press, 2001. 232 - 235. (in Chinese)

[8] Shim Jae-Chang, Dorai Chitra. A generalized region labeling algorithm for image coding, restoration, and segmentation [A]. In: *Proc of the IEEE International Conference on Image Processing* [C]. Los Alamitos, CA, USA, 1999. 46 - 50.

[9] Hampapur Arun, Bolle Ruud M. Feature based indexing for media tracking [A]. In: *Proc of the IEEE International Conference on Multimedia and Expo* (Ⅲ) [C]. New York, NY, USA, 2000. 1709 - 1712.

[10] Zhou X S, Huang T S. Relevance feedback for image retrieval: a comprehensive review [J]. *Multimedia System*, 2003, **8**(6): 536 - 544.

基于内容自动扩展的多示例查询图像检索技术

王小玲[1,2]　　谢康林[1]

([1] 上海交通大学计算机科学与工程系,上海 200030)

([2] 山西财经大学信息管理学院,太原 030016)

摘要:为了缩短基于内容图像检索存在的“语义鸿沟”,提出了一种自动扩展的多示例查询技术. 该技术将传统检索使用的单一查询图像自动扩展为多个查询示例,从而包含了更多的与语义相关的图像特征. 对这些查询示例进行检索,并融合检索结果,可以获得更多相关图像. 扩展主要利用了一般检索算法的查准率-查全率曲线特点,对原始查询结果的图像特征距离应用 K-均值聚类算法,确定多个查询示例图像. 实验结果表明该方法可以显著提高原有检索算法的查全率和查准率.

关键词:基于内容的图像检索; 语义; 多示例查询; K 均值聚类

中图分类号:TN911. 73