A Novel Image Retrieval based on Representative Colors

Jianfeng Ren[1] , Yuli Shen[2] and lei Guo[1]

1. Department of Automatic Control, North Western Polytechnic University, P.R. CHINA

rjjff@163.com lguo@nwpu.edu.cn

2.Department of Electrical Engineering, Zhan Jiang Ocean University, P.R.CHINA

Abstract

In this paper, we present results of a project that seeks to transform the low level features to high- level meaning. Firstly we extract the low level features called as the representative colors from the images, and then we present a new approach called WordNet to establish the link from the low level feature vectors to the semantics. In order to improve the retrieval efficiency, the relevance feedback is also applied into our system. Experimental results show that our method is promising.

Keywords: representative colors, Wordnet, feedback

1 Introduction

Indexing diverse collections of multimedia data remains a challenging problem. Most current approaches to image retrieval mainly focused on two aspects: one is the visual features [1,2,3,4] such as the color, texture and shape, the other is the distance metrics [5,6,7]

Even though significant progress has been made toward developing effective content-based descriptors, such as the standard descriptors by MEPG-7 [8], there is a difficulty to narrow the gap between the low-level features in image analysis and image understanding at the semantic level. Because people analyze, understand and classify the image content according to its semantic features.

Due to the importance of semantics, some approaches have been provided to bridge the gap between the low-level features and semantic level features. Shi-F Chang [9] proposed the novel idea of Semantic Visual Templates (SVT) to narrow this gap. Aekasandra Mojsiovic and Bernice Rogowitz [10] propose a method for semantic categorization and retrieval of photographic images based on low-level image descriptors. The main drawback of this method is that we need to do a lot of psychophysical experiments. Wiiam I.Grosky and Rong zhao [11] presented the techniques, latent semantics indexing (SCI), to negotiate the gap. However, these methods seem to be inefficient due to the size of image collections.

In order to overcome the shortfalls mentioned above, in this paper, we propose a new approach based on the representative colors to narrow the gap between the low-level and semantic level, considering the intuitive characteristics of color. Firstly, the low-level features based on the color are presented. Then we use the WordNet in order to narrow the gap between the low-level feature and semantic level features through the training sample images and human interactions. In order to more efficiently retrieve images, relevance feedback has been adopted in our strategy. Experimental results demonstrate that this method could correctly retrieve images not only through the specified images but also keywords in a given domain.

This paper is organized as followed: we introduce the overview of our method in section2. In section 3, the low-level features based on the representative colors

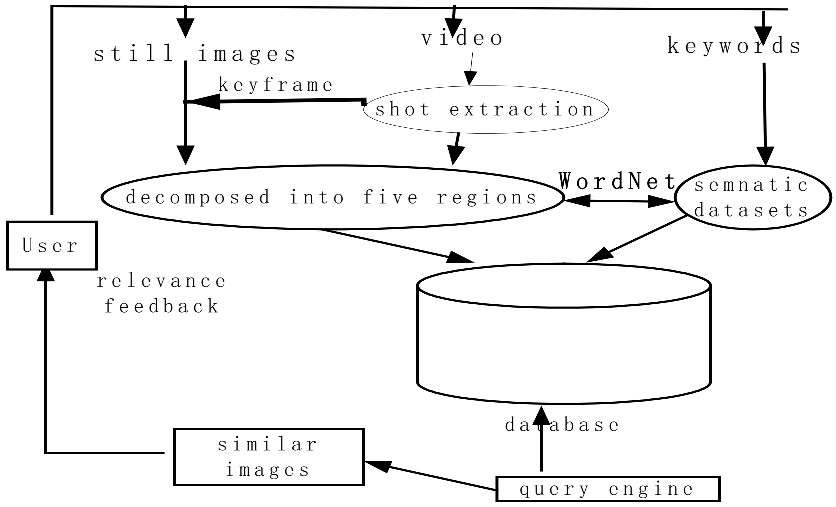

Figure 1 overview of our method

are developed. In section 4, we will present the WordNet to bridge the gap between low level features and high-level features. In section 5, the relevance feedback is applied into our strategy. In section 6, a detailed analysis based on the experiment results will be presented. Finally, Conclusions and remarks are given in section 7

2. Overview of our method

Our method mainly includes the following steps:
Step1: image decomposed into five regions
Step**2**: computing the frequency of eight domain color components in each region
Step3: establishing the WordNet from the color feature to semantics through the training samples and human interactions
Step4: re-define the query feature from the similar images using the relevance feedback technique until the user satisfies the results.
Figure 1 shows an overview of our provided method.

3 Color feature extraction

Color is perhaps the most expressive of all the visual features and has been extensively studied in image retrieval during the last decade. Some low-level features are developed by using the color histograms[2,12,13].

In this paper, we provide representative colors in a region of interest. The reason that we employ this method is as following:

- It is provided mainly based on the observation that a small number of colors are usually sufficient to characterize the color information in image region.
- Representative color approach can overcome the drawback of the traditional histogram. The number of bins in a typical color histogram range from few tens to a few hundreds. The high dimensionality of the feature vectors result in high computational cost in distance calculation for similarity retrieval, and inefficiency in indexing and search.
- It also considers the spatial information and relationships between the different regions in the images.

The process of our method is as fol1owing: Firstly the image can be decomposed into five fixed regions 14], which is shown in figure 2, then representative color histogram is computed in each region.

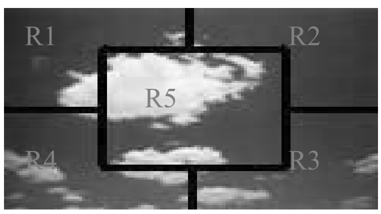

Figure 2 image decomposition

3.1 representative colors

According to human vision perception, a small number of image colors are sufficient to express the color information [15]. Therefore, in our method, we regard eight colors as the representative colors. Eight colors are illustrated as following: {red, green, blue, yellow, magenta, cyan, black and white}, which are mapped into corresponding points in the RGB space: {(255,0,0);(0,255,0);(0,0,255);(255,255,0);(255,0,255);(0,255,255);(0,0,0);(255,255,255)}.

3.2 color clustering

In order to compute the percentage of representative colors in the region, the colors in the image region should be clustered based on the nearest neighbor algorithm.

Now we assume eight representative colors $\{C_1,C_2,C_3,C_4,C_5,C_6,C_7,C_8\}$ as the clustering center, and **#C_i** denotes the total number of pixels which are grouped into the clustering center C_i, and $\# C$ denotes the total number of pixels in the image, so $\frac{\# C_i}{\# C}$ can be regarded as the percentage that representative color account for in the image. For example, in this paper, $\frac{\# C_1}{\# C}$ denotes that the percentage of the red component in the image.

Because the image are decomposed into five fixed regions such as the left-up, right-up, left-down, right-down and the center, we should express the spatial relationships between the different regions in the feature vectors. Here, let $S = [s1, s2, s3, s4, s5]$ expresses the corresponding region in the image. For example, in a given region $R5$, after color clustering, feature vector $[\frac{\#C_1}{\#C},\frac{\#C_2}{\#C},\frac{\#C_3}{\#C},\frac{\#C_4}{\#C},\frac{\#C_5}{\#C},\frac{\#C_6}{\#C},\frac{\#C_7}{\#C},\frac{\#C_8}{\#C},s5]$ is obtained.

4. Establishing the links

Now we have obtained the representative color features in the corresponding regions. However, these feature vectors could not be directly recognized by humans, just because people understand image at the semantic level. So a new approach called WordNet .is provided to solve this problem

In our paper, all the images are from the scene images, which have obvious color characters. Through the human interaction and manipulation, we can easily find the link between the representative color and the scenes. For example, blue can express many semantic words such as the blue sky and sea water.

A detailed link between the color and the corresponding keywords is shown in figure 3

Through the above links, we can easily narrow the gaps between the low-level and semantic level features. And more, we also see that the intermediate features play an essential role. Such links are stored in the database as the look-up table. In addition, the user can add the keywords in the look-up table mentioned above.

For a give region R5 located in the center of the image, the specified process from the low-level to semantic level is expressed as following:
If #C1 in the feature vector is maximum
Then the object can be {rose, sun….}
If #C3 in the feature vector is maximum
Then the object can be {see water, blue sky….}

●

Through the above algorithm, each region can be indexed one representative color, with which has a set of keywords associated. Accordingly, an image that is decomposed into five fixed regions will have not more than five sets of associated keywords, because the different regions may have the same representative color.

The weight W_{ij} associated on each link of the keywords with the image represents the degree of relevance in which the keyword describes the linked image's semantic content.

From above discussed, we see that each image object is associated with three different level features. respectively at the low-level feature, intermediate

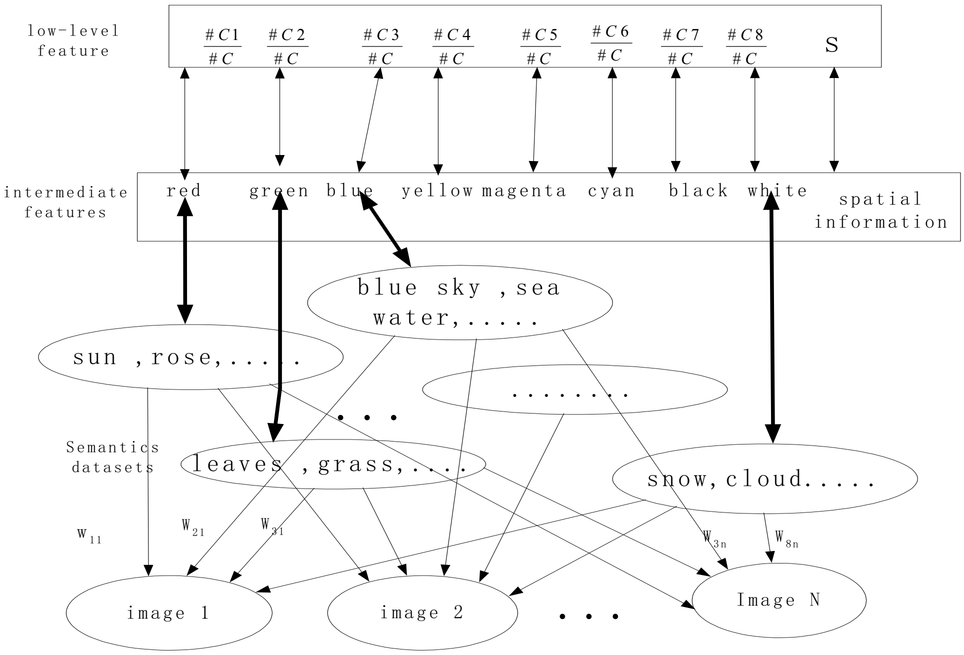

Figure 3 the link between representative colors and semantic keywords

level feature and semantic level feature

5. Relevance feedback

The approaches [16,17] performed relevance feedback at the low-level feature vector level, but failed to take account into the actual semantics for the image databases. The inherent problem with these approaches is that the low-level features are often not as powerful in representing complete semantic content of images as keywords in representing text documents.

In our paper, we provide relevance feedback not only at the low-level feature level but also at the semantic level.

We apply the relevance feedback strategy at the low-level features based on the formula given in [18].

However, semantic based relevance feedback can be performed relatively easily compared to low-level feature counterpart. The basic idea behind it is a simple scheme to update the weights W_{ij} associated with each link shown in Figure3. And the weight updating process is described as below:

1. Initialize all weights W_{ij} to 1. That is, every keyword has the same importance.
2. Collect the user query and the positive and negative feedback examples
3. For each keyword in the input query, check to see if any of them is not the keyword database. If so, add them into the database without creating any links
4. For each positive example, check to see if query keyword is not linked to it. If so, create a link with weight 1 from each missing keyword to this image. For all other keywords that already to this image, increment the weight by 1.

5. For each negative example, check to see if any query keyword is linked with it. If so, set the new weight $W'_{ij} = W_{ij}/2$. If the weight W_{ij} on any link is less than 1, delete that link.

6. Experimental result

We download about 10000 scene images from the website. Every representative color may have more than 1000 scene images. According to its semantic content of the image and the corresponding representative colors, each representative color can be associated with a few sets of keywords. Through the algorithms provided in section 3 and section 4, each image and associated keywords are put into the database.

The experimental research is concerned primarily with the retrieval time and the accuracy of the retrieved images.

The first experiment is conducted to evaluate the retrieval time. Table 1 shows the retrieval time at three different levels.

Table 1 the retrieval time

Query level	Retrieval time (s)
Query by Example image	132.13
Query by intermediate color	0.300
Query by keywords	0.270

From the above table, we can see that query by keywords can improve the retrieval speed greatly compared with the query by example image.

The second experiment is implemented in order to evaluate the retrieval accuracy based on the relevance feedback. A retrieved image is considered as a relevant one if it has the similar representative color to the query image. The retrieval accuracy is defined as

$$R = \frac{\text{relevant images retrieved in top N returns}}{\text{N}} \qquad (2)$$

In our experiment, N is usually set to 20. Thus, the retrieva1 accuracy is denoted by $R\ 20$

We performed four random queries on our system at the semantic level and at the low level feature respectively. Figure 4 shows the accuracy based on the relevance feedback at the semantic level, and figure 5 shows the accuracy based on the relevance feedback at the low level feature.

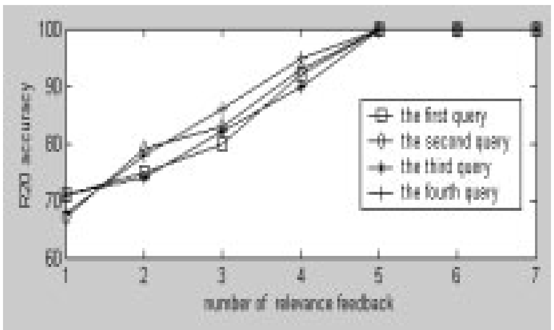

Figure4 Accuracy based on the relevance feedback at the semantic level

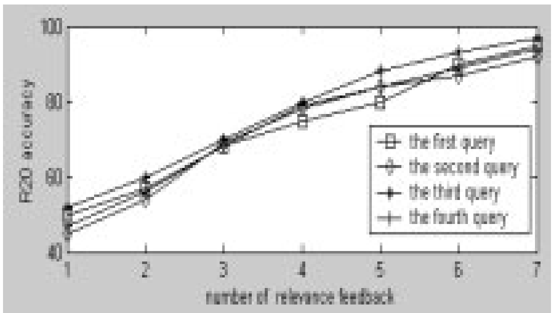

Figure5 Accuracy based on relevance feedback at the low level

As we can see from the results, our system achieves the high accuracy (>80%) after a few of relevance feedback for any given query. Unlike other methods where more user relevance feedback [17,18] may lead to lower the retrieval accuracy, experimental results demonstrate our method can be stable.

In addition to verifying the effectiveness of our proposed method, we also compared our method against the other techniques. We have chosen to compare our method with the retrieval technique used in[19]. The comparison is made through four random queries based on relevance feedback. Figure 6 shows the comparisons with the method in[19].

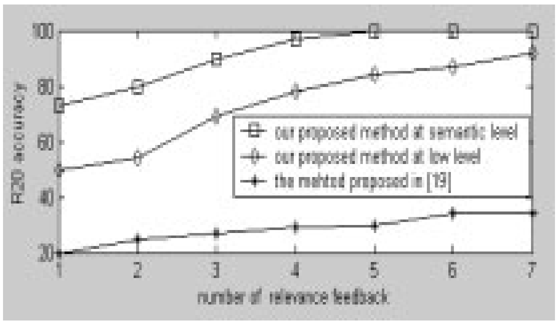

Figure6. Performance comparison

It is easily seen from results, our methods based on relevance feedback will improve the retrieval accuracy substantially.

7. Conclusions

In this paper, we provide the representative colors for indexing the images. Our method not only captures the color content of the images but also characterizes the spatial information of color in the image. In addition, we present a new approach called the WordNet to narrow the gap between the low-level features from the representative colors and semantic content of the images. Finally, in order to improve the retrieval performance, we applied the relevance feedback to retrieve images both at the semantic level and at low-level feature. Experimental results show this method is stable. In the future work, we will further narrow the gap from the shape descriptor to semantics after we have obtained the representative colors.

Reference

[1] M. Flicker, H, Sawhney, and W. Niblack, “query bay image and video: The QBIC system,” IEEE Computer, 28, 1995.

[2] M, Swain and D. Ballard, “ Color indexing”Vision,7,pp11-32,1991.

[3] C.C. Gotlieb and H.E. Kreyszig, “Texture descriptors based on co-occurrence martrices,” Computer Vision, Graphics, and Image Processing, 51,1990.

[4] J.Z. Wang et al, “Content-based image indexing and searching using Daubechies’s wavelets,” Int J. Digital libraries, 1, pp 311-328,1998.

[5] S. Santini and R. Jain, “Similarity measures,” IEEE Trans. On PAMI, 9,1999.

[6] A, Natses, R. Rasto, “Walrus: a similarity retrieval algorithm for image database,” SIGMOD Record, 28, 1999

[7] D, Bimbo and P, Pala, “Visual image retrieval by elastic matching of user sketches,” IEEE Trans on PAMI, 19, 1997.

[8] Text of ISO/IEC 15938-3 Multimedia Content Description Interface-Part 3: Visual, Final Committee Draft, ISO/IEC/JTC1/SC29/WG11, Doc, N4062, Mar, 2001.

[9] S.-F. Chang, W. Chen, and H. Sundaram, “Semantic Visual Templates: linking visual features to semantics Semantic Visual Templates: Linking Visual Features to Semantics,” Proceedings 1998 International Conference on Image Processing (ICIP'98), Chicago, Illinois, October 4-7, 1998.

[10] Mojsilovic, and B. Rogowitz, "Capturing image semantics with low-level descriptors", Proc. Int. Conference on Image Processing, ICIP 2001, Thessaloniki, Greece, September 2001.

[11] Rong Zhao and William I.Grosky, “Narrowing the Semantic Gap---Improved Text-Based Web Document Retrieval Using Visual Features,” IEEE Transaction on Multimedia, Vol,4, No.2, pp189-200, 2002

[12] W. Niblack, R.Barber, W.Equitz, M.Flickner, E. Glasman, D. Petkovic, P. Yanker, C.Faloutsos, and G.Taubin. “The QBIC project: Querying images by content, using color, texture, and shape.”. In SPIE Conference on Storage and Retrieval for Image and Video Databases. Volume 1908, pp 173-187,1993.

[13] T, Kato et al. “A sketch retrieval method for full color image database-query by visual example.” In proceedings, 11th IAMPR International Conference on Pattern Recognition, pp530-533, 1992.

[14] Markus Stricker, Alexander Dimai,” Color Indexing with Weak Spatial Constraints” SPIE Conference, Feb, 96,San jose.

[15] G.wyszecki and W. S. Stiles,Color Science: Concepts and Methods, John Wiley & Sons,1982.

[16] Y. Rui, T. S. Huang, and S. Mehrotra, “Content-based image retrieval with relevance feedback in MARS,” in Proc. IEEE Int. Conf. On Image Proc., 1997.

[17] Y. Ishikawa, R. Subramanya, and C.Faloutsos, “Mindread: Query Databases through multiple examples” in Proc. Of the 24th VlDB Conference,(New York),1998.

[18] RUI,Y, huang ,T.S, and Mehrotra, S ”content-based image retrieval with relevance feedback in MARS,” in proc. IEEE Int Conference. On image proceeding, 1997.

[19] Rui,Y., Huang, T.S. “A Novel Relevance Feedback Techniques in Image Retrieval”, ACM Multimedia,1999.