A Novel Image Retrieval system Based on BP Neural Network

Jun-Hua Han[1], De-Shuang Huang[1], Tat-Ming Lok[2] and Michael R. Lyu[3]
Intelligent Computing Lab, Institute of Intelligent Machines,
Chinese Academy of Sciences, P.O.Box 1130, Hefei, Anhui, 230031, China[1]
Information Engineering Dept., The Chinese University of Hong Kong Shatin, Hong Kong[2]
Computer Science & Engineering Dept.,The Chinese University of Hong Kong Shatin, Hong Kong[3]
E-mails: jhhan@iim.ac.cn, dshuang@iim.ac.cn, tmlok@ie.cuhk.edu.hk, lyu@cse.cuhk.edu.hk

***Abstract*— This paper[1] presents a novel BP-based image retrieval (BPBIR) system , which is based on the observation that the images users need are often similar to a set of images with the same conception instead of one query image and the assumption that there is a nonlinear relationship between different features. If users aren't satisfied with the retrieved results, relevance feedback method is used to enhance the performance of the proposed system by changing the weights of the BP neural networks. In addition, we discuss some divisional methods to give rough information on the spatial color composition. Finally, we compare the performance of the proposed system with other systems. Experimental results show the efficacy of the proposed system.**

I. Introduction

With the rapid development of computing hardware, digital acquisition of information has become one popular method in recent years. Every day, G-bytes of images are generated by both military and civilian equipment. Consequently, how to make use of this huge amount of images effectively becomes a highly challenging problem [1], [2]. The traditional approach relies on image content manual annotation and Database Management System (DBMS) to accomplish the image retrieval through keywords. Although simple and straightforward, the traditional approach has two main limitations. First, the descriptive keywords of an image are inaccurate and incomplete. Second, manual annotation is time-consuming and subjective. Users with different backgrounds tend to describe the same object using different descriptive keywords resulting in the difficulties in image retrieval. To overcome the drawbacks of the traditional approach , content-based image retrieval (CBIR) was proposed to retrieve visual-similar images from an image database based on automatically-derived image features, which has been a very active research area[1], [2], [3].

There have been many projects performed to develop efficient systems for content-based image retrieval. The well-known CBIR system is probably IBM's QBIC [4]. Other notable systems include MIT's Photobook [5], Virage's VisualSEEK [6], SIMPLIcity[7], etc. However, in these systems, only one image is used to retrieve the images in image database. While the images that people need are often similar to a set of images with the same conception instead of one query image. In addition, most of the existing approaches assume a linear relationship between different features, and the usefulness of such systems was limited due to the difficulty in representing high-level concepts using low-level features.

In this paper, we present a novel BP neural network-based image retrieval system. Firstly, some typical images of each class should be selected to train the designed BP neural network that would be used to retrieve the images in an image database based on the difference of BP neural network's output between the query images and the images in the image database. If the user isn't satisfied with the retrieved results, relevance feedback method is used to enhance the performance of the proposed system by changing the weights of the BP neural networks. The proposed BPBIR system is based on the observation that the images that users need are often similar to a set of images with the same conception instead of one query image and the assumption that there is a nonlinear relationship between different features. In addition, we discuss some divisional methods so as to give rough information on the spatial color composition. Experimental results show that it is more effective and efficient to retrieve visual-similar images for a set of images with same conception can be retrieved.

The remainder of the paper is organized as follows: Section II describes the image content representation including color moments and wavelet decomposition energy and discusses some divisional methods in order to give rough information on spatial composition. Section III describes the proposed BP-based image retrieval system in detail. Section IV presents the experimental results and demonstrates the validity and feasibility of the proposed algorithm by some experiments. Finally, section V draws some conclusions.

II. Image Content Representation

In this section, we firstly describe some divisional methods in order to give rough information on the spatial color composition. Secondly, color moments and wavelet decomposition energy are introduced to represent the image content. It needs to be pointed out that feature selection is not restricted in this way for the proposed BPBIR system is a general framework for multi-purpose content-based image retrieval, and new features can be used in the proposed system.

[1]This work was supported by the National Science Foundation of China (Nos.60472111 and 60405002).

0-7803-9048-2/05/$20.00 ©2005 IEEE

Authorized licensed use limited to: Chinese University of Hong Kong. Downloaded on February 11,2021 at 11:42:14 UTC from IEEE Xplore. Restrictions apply.

A. Some Division Methods

In order to give rough information on the spatial color composition, we describe some divisional methods, which can be seen from Fig. 1. In (a), the image is divided into 9 parts with the same size of the image, while in (b), the image is divided into 9 different parts, and the size of center part is 9/25 of the area of the image. In (c), the image is divided into 5 parts including upper left part, upper right part, lower left part, lower right part, and center part. The size of each part is 1/4 of the area of the image. In (d), the image is also divided into 5 parts including upper half part, lower half part, left half part, right half part, and central part. In these five parts, the central part's size is 4/9 of the area of the image, the others are 1/2 of the area of the image.

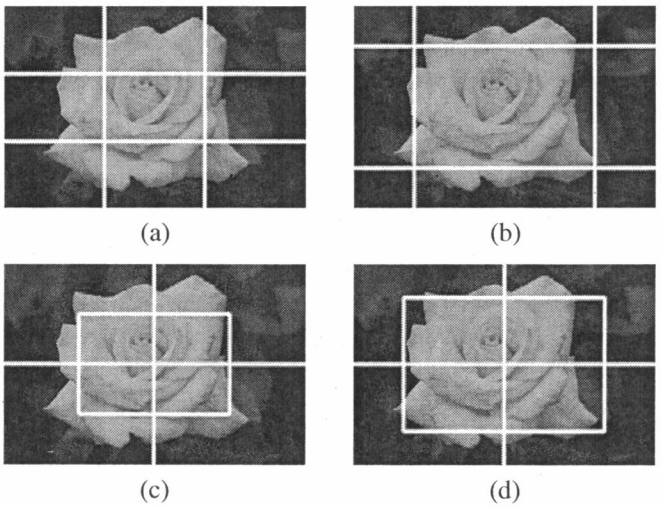

Fig. 1. The image divided into several parts with different methods

When computing the distance between the query images and the images in image database, we can set a larger weight of the central part for the interest object often lies in the center of the image.

B. Color Moments

Color moments have been successfully used in many color-based image retrieval systems[4], especially when the image contains just the object. The first order (mean), the second (variance) and the third order (skewness) color moments have been proved to be efficient and effective in representing color distributions of images[8]. Mathematically, the first three moments can be defined as:

$$\mu_k = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} p_{ij}^k \tag{1}$$

$$\sigma_k = \left(\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(p_{ij}-\mu_k)^2\right)^{1/2} \tag{2}$$

$$s_k = \left(\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(p_{ij}-\mu_k)^3\right)^{1/3} \tag{3}$$

where p_{ij}^k is the value of the k-th color component of the image ij-th pixel, and M is the height of the image, and N is the width of the image.

C. Wavelet Decomposition Energy

Wavelets can decompose a signal with a family of basis functions $\psi_{mn}(x)$ obtained through translation and dilation of a mother wavelet $\psi(x)$, i.e.,

$$\psi_{mn}(x) = 2^{-m/2}\psi(2^{-m}x - n) \tag{4}$$

where m and n are dilation and translation parameters. A signal $f(x)$ can be represented as:

$$f(x) = \sum_{m,n} c_{mn}\psi_{mn}(x) \tag{5}$$

The computation of the wavelet transforms of a 2D signal involves recursive filtering and sub-sampling. At each level, the signal is decomposed into four frequency sub-bands, LL, LH, HL, and HH, where L denotes the low frequency and H denotes the high frequency.

Since wavelets capture shape, texture, and location information of an image in a single unified framework, wavelet transform features are often used in content-based image retrieval systems. In this paper, we use the symlets that are nearly symmetrical wavelets proposed by Daubechies[9] as modifications to the db family. After the three-level wavelet decomposition, feature vectors can be constructed using Ea, which is the percentage of energy corresponding to the approximation, and vectors Eh, Ev, Ed, which contain the percentages of energy corresponding to the horizontal, vertical, and diagonal details, respectively.

III. BP-Based Image Retrieval System

In this section, we describe the structure of the proposed BP-based image retrieval system in detail. Fig. 2 shows the main components of the proposed BPBIR system and the control flows among them.

The proposed image retrieval system is based on Back-propagation neural network[10], [12]. There are three layers of BP neural network including input layer, hidden layer, and output layer. This is based on the observation that the images users need are often similar to a set of images with the same conception instead of one query image and the assumption that there is a nonlinear relationship between different features. The following is the proposed BP-based image retrieval algorithm.

Step 1: All the images in image database are divided into 5 and 9 parts using the divisional methods above. Color moments and wavelet decomposition energy of each part are computed as the features to represent image content. In addition, the feature vectors are normalized using the following equation:

$$f_i^{(j)} = \frac{f_i^{(j)} - min(f_i^{(j)})}{max(f_i^{(j)}) - min(f_i^{(j)})} \tag{6}$$

Authorized licensed use limited to: Chinese University of Hong Kong. Downloaded on February 11,2021 at 11:42:14 UTC from IEEE Xplore. Restrictions apply.

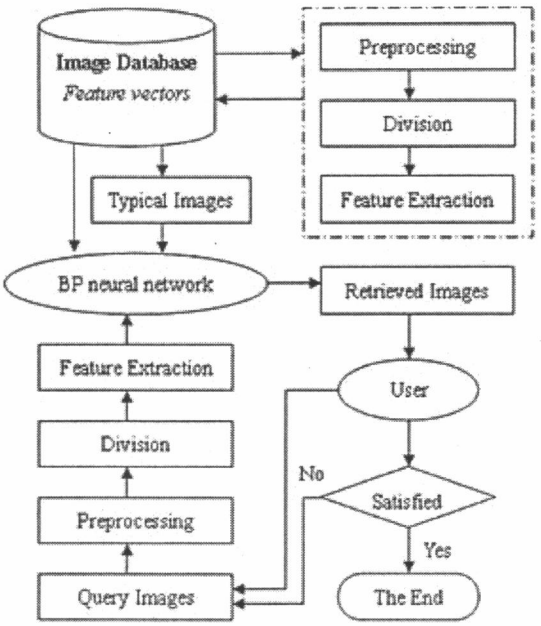

Fig. 2. The structure of the proposed BP-based image retrieval system

where $f_i^{(j)}$ is the j-th feature value of the i-th image in image database, $min(f_i^{(j)})$ is the minimum of the j-th feature value of all the images in image database, $max(f_i^{(j)})$ is the maximum of the j-th feature value of all the images in image database.

Step 2: Some typical images are selected as training sample to be used to train BP neural network using one step secant (OSS) method [11].

Step 3: The retrieved images are sorted based on the distance of BP neural network's output between the query image and the images in image database, which is defined as:

$$d_i = (\frac{1}{C}\sum_{j=1}^{C}(o_j^i - o_j^q)^2)^(1/2) \quad (7)$$

where o_j^i is the j-th output value of BP neural network with the i-th retrieval image, o_j^q is the j-th output value of BP neural network with the query image, and C is the class number of image database.

Step 4: If the user isn't satisfied with the retrieved results, some required images that are selected by users from the retrieved images can be used to train the BP neural network again with the same output of query image, which is different from other relevance feedback methods[15] for the user's feedback is not to change the weights of different feature but to change the weights of BP neural network.

IV. Experimental Results

In this section, we demonstrate some experimental results to show the performance of the proposed BPBIR system. These results are based on an implementation using a test image database of 1000 images including 10 classes, which can be download from the website: http://wang.ist.psu.edu/ jwang/test1.tar. Each class has 100 images.

TABLE I
Comparison of Average Precision Using Different Divisional Methods

Classes	1(%)	2(%)	3(%)	4(%)	5(%)
African people and villages	31.36	32.17	40.20	34.26	33.27
Beaches	30.40	25.69	28.42	29.33	30.71
Buildings	19.82	18.72	24.61	20.51	21.34
Buses	35.98	31.00	32.41	31.25	33.96
Dinosaurs	88.36	71.81	80.21	83.27	88.30
Elephants	33.94	38.97	34.99	37.15	35.28
Flowers	40.23	43.69	29.80	45.82	42.61
Horses	45.76	55.37	45.89	45.38	46.02
Mountains and glaciers	23.01	20.86	20.78	22.05	22.34
Food	26.22	27.58	35.19	30.52	29.67

A. Experiment 1

In this experiment, we compared some divisional methods using color moments feature. In Table I, column 1 is the average retrieval precision using no divisional method, and column 2, 3, 4, and 5 are the average retrieval precision using color moments with the different divisional methods in Fig. 1: (a), (b), (c), and (d), respectively. From Table I, we can see that the retrieval results using divisional methods would be better than no divisional method for the divisional methods could give rough information on the spatial color composition.

B. Experiment 2

In this experiment, we use the first 40 images of each class as typical images to train the designed BP neural network including 75 inputs, 20 hidden units, and 10 outputs that depend on the number of the classes in the image database. The feature vector has 75 dimension including 45-dimensional 5-block color moments and 30-demensional wavelet decomposition energy. Fig. 3 and Fig. 4 show the retrieved results of the proposed BPBIR system, in which the first image is the query image. In Fig. 3, the query image is a horse image, while in Fig. 4, the query is a flower image. We can see from Fig. 3 and Fig. 4 that the proposed system is very efficient for a set of image with same conception can be retrieved, which is more similar to the human visual system.

In order to verify the performance of the proposed system, we also compare the proposed BPBIR system with other image retrieval systems. Table II is the comparison of average precision with the SIMPLIcity system[7] and the RIBIR system[14], which shows that the proposed BPBIR system is more efficient to retrieve visual-similar images than the other two systems. The reason is that the trained BP neural network can memorize some prior information about each class.

Authorized licensed use limited to: Chinese University of Hong Kong. Downloaded on February 11,2021 at 11:42:14 UTC from IEEE Xplore. Restrictions apply.

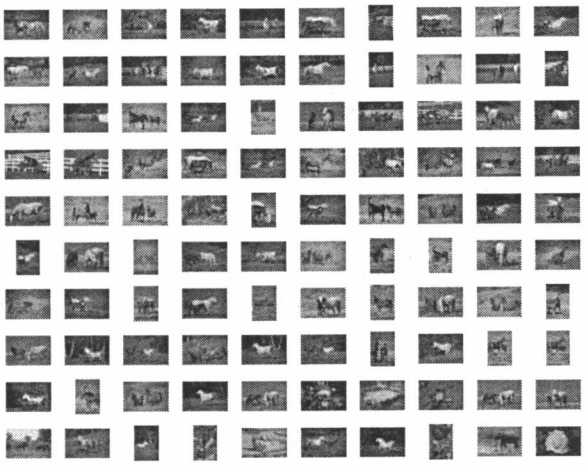

Fig. 3. The retrieved results of BP-based image retrieval system

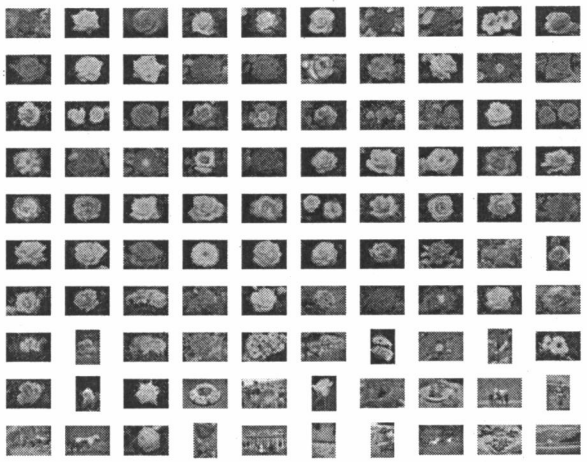

Fig. 4. The retrieved results of BP-based image retrieval system

TABLE II

COMPARISON OF AVERAGE PRECISION BETWEEN BPBIR SYSTEM AND OTHER SYSTEMS

Classes	RIBIRs'	SIMPLIcitys'	BPBIRs'
African people and villages	32%	48%	57.59%
Beaches	40%	32%	53.90%
Buildings	34%	33%	52.66%
Buses	43%	37%	81.23%
Dinosaurs	54%	98%	98.62%
Elephants	49%	40%	52.12%
Flowers	36%	40%	75.50%
Horses	54%	71%	80.14%
Mountains and glaciers	34%	32%	46.55%
Food	50%	31%	67.58%

V. CONCLUSIONS

In this paper, we presented a novel BP-based image retrieval system, which is based on the observation that the images users need are often similar to a set of images with the same conception instead of one query image and the assumption that there is a nonlinear relationship between different features. In addition, we discuss some division methods in order to give rough information on the spatial color composition. Finally, we compare the performance of the proposed system with other image retrieval system. Experimental results show that it is more effective and efficient to retrieve visual-similar images for a set of images with same conception can be retrieved.

ACKNOWLEDGMENT

The authors would like to thank Dr. Xiao-Ping Zhang at Ryerson University for help and some discussions about this paper.

REFERENCES

[1] Y. Rui and T. S. Huang, "Image Retrieval: Current Techniques, Promising Directions, and Open Issues", *Journal of Visual Communication and Image Representation*, Vol.10, pp.39-62, 1999

[2] R. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval*, Reading, MA: Addison-Wesley,June, 1999.

[3] J. P. Eakins, "Towards intelligent image retrieval", *Pattern Recognition* , Vol.35, No.1, pp.3-14, 2002.

[4] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by image and video content: The QBIC system." *IEEE Computer*, Vol.28, No.9, pp.23-32, Sept. 1995.

[5] A. Pentland, R.W. Picard, S. Sclaroff, "Photobook: content-based manipulation for image databases", *Int. J. Comput. Vision*, Vol.18, pp.233-254, 1996.

[6] J.R. Smith, and S.-F. Chang, "VisualSEEk: a fully automated content-based image query system", In:*Proceedings of the ACM Multimedia*, Vol.96, pp.87-98, 1996.

[7] J.Z. Wang, J.Li, G. Wiederhold, "SIMPLIcity: semantic-sensitive integrated matching for picture libraries, *IEEE Trans. Image Processing*, Vol.9, No.1, pp.947-963, 2000.

[8] M. Stricker, and M. Orengo, "Similarity of color images", *SPIE Storage and Retrieval for Image and Video Databases III*, vol. 2185, pp.381-392, Feb. 1995.

[9] Cohen, A.; I. Daubechies, B. Jawerth, P. Vial, "multiresolution analysis, wavelets and fast wavelet transform on an interval", CRAS Paris, Ser. A, t. 316, pp. 417-421, 1993.

[10] Rumelhart, D. E., G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors", *Nature*, Vol. 323, pp. 533-536, 1986.

[11] Battiti, R. "First and second order methods for learning: Between steepest descent and Newton's method", *Neural Computation*, vol. 4, no. 2, pp. 141-166, 1992.

[12] D. S. Huang, *Systematic Theory of Neural Networks for Pattern Recognition*, Beijing: Publishing House of Electronic Industry of China, 1996.

[13] Hagan, M. T., H. B. Demuth, and M. H. Beale, *Neural Network Design*, Boston, MA: PWS Publishing, 1996.

[14] B. Si, W. Gao, H. Lu, W. Zeng, L. D, "An Image Retrieval Method Based Regions of Interest", *High Technology Letters*, Vol.13, No.5, pp.13-18, 2003.

[15] H.J. Zhang, "Relevance Feedback in Content-Based Image Retrieval", in *Multimedia Information Retrieval and Management-Technological Fundamentals and Applications*, D. Feng, W.C. Siu, and H.J.Zhang. (ed.), Springer, 2002.

Authorized licensed use limited to: Chinese University of Hong Kong. Downloaded on February 11,2021 at 11:42:14 UTC from IEEE Xplore. Restrictions apply.