7-1-2007

Association-based image retrieval

Arun D. Kulkarni
University of Texas at Tyler, akulkarni@uttyler.edu

H. Gunturu

S. Dalta

Association-Based Image Retrieval

Arun Kulkarni

Computer Science Department
The University of Texas at Tyler, Tyler, TX 75799

Abstract
With advances in the computer technology and the World Wide Web there has been an explosion in the amount and complexity of multimedia data that are generated, stored, transmitted, analyzed, and accessed. In order to extract useful information from this huge amount of data, many content-based image retrieval (CBIR) systems have been developed in the last decade. A typical CBIR system captures image features that represent image properties such as color, texture, or shape of objects in the query image and try to retrieve images from the database with similar features. Recent advances in CBIR systems include relevance feedback based interactive systems. The main advantage of CBIR systems with relevance feedback is that these systems take into account the gap between the high-level concepts and low-level features and subjectivity of human perception of visual content. CBIR systems with relevance feedback are more efficient than conventional CBIR systems; however, these systems depend on human interaction. In this paper, we propose a new approach for image storage and retrieval called association-based image retrieval (ABIR). We try to mimic human memory. The human brain stores and retrieves images by association. We use a generalized bi-directional associative memory (GBAM) to store associations between feature vectors that represent images stored in the database. We propose three topologies of GBAM for the proposed ABIR system. As an illustration, we have considered a database with three sets of images. The results of our simulation are presented in the paper.

***Keywords**:* Association-based Image Retrieval, Feature vector, Generalized bi-directional Associative Memories, Multi Media Databases.

1. Introduction

The rapid growth in the number of large-scale image repositories in many domains such as medical image management, multimedia libraries, document archives, art collection, geographical information the systems, law enforcement management, environmental monitoring, biometrics, and journalism has brought the need for efficient CBIR mechanisms. Many ideas from fields such as computer vision, database, image processing, and information retrieval are used in CBIR. Effective retrieval of image data is an important building block for general multimedia information management. For an image to be searchable, it has to be indexed by its content. Earlier CBIR systems used keyword annotation for indexing images in the database. The keyword annotation method involves a large amount of manual effort. Furthermore, the keyword annotation depends upon human interpretation of image content, and it may not be consistent. To overcome these difficulties automatically extracted low-level features such as color, texture, and shape features are used for image indexing. Many CBIR systems with low-level features have been developed [2, 12, 15, 16]. Swain and Ballard [14] used color histograms for indexing images. Huang et al [3] used color correlograms instead of color histograms. Chang and Kuo [1] proposed the use of texture features via the wavelet transform. Shape features based on Fourier descriptors, moment invariants also have been used in conjunction with color, and texture features in CBIR systems [5, 13]. Mehrotra and Gary [10] have suggested a shape management system for retrieving similar shapes. Krishnapuram et al. [8] have proposed a fuzzy approach for CBIR. It may be noted that CBIR does not rely on describing the content of the image in its entirety. It may be sufficient that a retrieval system presents similar images in some user-defined sense. Smelders et al. [12] consider the description of image content in two steps. The first step deals with image processing operations that transpose the image data array into another spatial data array that can include methods over local color, local texture, or local geometry. The second step is to extract invariant features. The aim of invariant description is to identify objects no matter how and where they are observed. The strategy for earlier CBIR systems is to find the "best" representation for visual features and based on the selected features to find similar images to the user's query image. While this approach establishes the basis for CBIR, the performance is not satisfactory due to two reasons a) the gap between the high-level concepts and b) the subjectivity of human perception. In order to overcome these drawbacks, Rui et al. [11] proposed a relevance feedback based interactive retrieval approach. In their approach during the retrieval process, the user's high-level query and subjectivity are captured by dynamically updated weights based on the user's feedback. In this paper, we propose a new approach for image retrieval called association-based image retrieval (ABIR). We try to mimic the human brain. Association is one of the fundamental characteristics of the human brain. The human memory operates in an associative manner; that is, a portion of recollection can produce an associated stream of data from the memory. The human memory can retrieve a full image from a partial or noisy version of the image as the query image. Furthermore, given a query image as the input, the human brain can recall associated images that have been stored in the past. The human memory can respond to abstract queries. The main disadvantage of the current CBIR systems is their inability to respond to abstract queries. Abstract queries are based on a notion of similarity, the concept that is difficult to capture in a mathematical model. For example, if we see an image of a person, we can recall images of his house, spouse, and car. The associative storage and retrieval mechanism is not explored in the present CBIR systems. We propose architecture for the ABIR that is based on a GBAM. The rest of the paper is organized as follows: In Section 2, we provide overview of current CBIR systems. In Section 3, we describe the architecture of the proposed ABIR system, and introduce generalized bi-directional associative memory models. In Section 4, we present results of our simulation, and Section 5 deals with conclusions.

2. Background

Stages in a typical CBIR system include annotation, preprocessing, and feature extraction. In order to store an image it is first annotated. In the feature extraction stage, features based on attributes such as the color, texture, or shapes are extracted. Most CBIR systems use color histograms to compare the query image with images in the database. Often color histograms alone are not sufficient to retrieve desired images from the database, because a single color histogram may represent multiple images in the database. Many CBIR systems use texture or shape features in conjunction with color features. In order to retrieve similar images from the database the feature vector obtained from the query image is compared with feature vectors of images in the database. Smeulders, et al.[12] describe similarity of two feature vectors by Equation (1)

$$S\left(\mathbf{f}_q,\mathbf{f}_d\right) = g \circ d\left(\mathbf{f}_q,\mathbf{f}_d\right) \qquad (1)$$

where $\mathbf{f}_q$ represents a feature vector obtained from the query image and $\mathbf{f}_d$ represents a feature vector for an image in the database, g is a positive, monotonically non-increasing function, and d is a distance function. This assumption is consistent with a class of psychological models of human perception and requires that feature space be metric. The most commonly used similarity function is the Minkowski distance that is given by Equation (2).

$$d\left(\mathbf{f}_q,\mathbf{f}_d\right) = \left[\sum_{j=1}^{n} \left| f_{qj} - f_{dj} \right|^2\right]^{\frac{1}{2}} \qquad (2)$$

where n is the number of features. The query image is compared with images in the database and the images in the database are ranked based on the similarity measure. Images that are similar to the query image are retrieved and displayed. While it is feasible to retrieve a desired image from a small collection by exhaustive search, techniques that are more effective are needed with a large database. The well-known indexing technique is used for efficient retrieval. Other features of a typical CBIR system include defining query feature space and displaying query results. The first component of the query space is the selection of a subset of images from the large image archive. In order to reduce the query space images in the database can be clustered into a small number of categories. The next section describes the architecture for the proposed ABIR system.

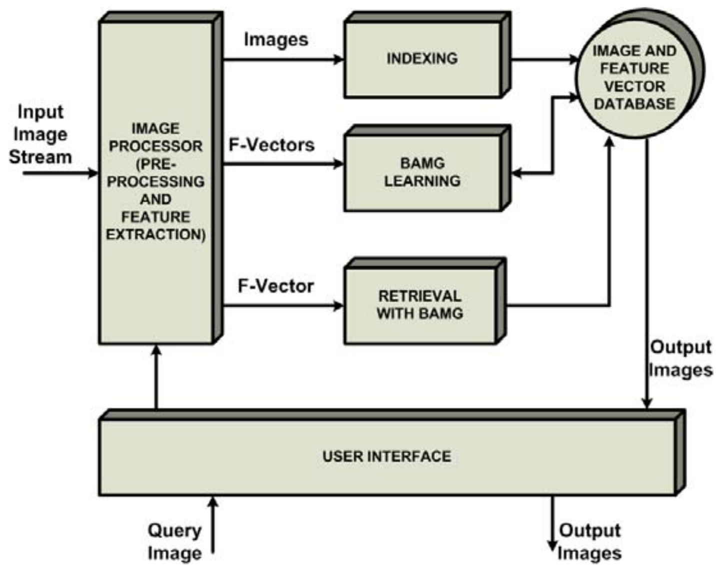

Figure 1. ABIR system architecture

3. Methodology

Architecture for the proposed ABIR system is shown in Figure 1. There are two data paths in the system. One deals with storing a set of images in the database, while other path deals with a query image. All images stored in the database are passed through the image-processing block that deals with preprocessing and feature extraction. The preprocessing stage deals with geometric and radiometric corrections, mapping the image from red, blue, green (RBG) color space to hue, saturation, and intensity (HIS) color space. The preprocessing stage may also include edge detection and transforms such as the Fourier transform. Extracted features represent color, texture, and shape characteristics of object in the image. The indexing stage deals with creating index table that contains feature vectors and pointers to the corresponding image locations. For faster retrieval, images in the database can be grouped in predefined categories. The user can select a query image via the user interface. The query image undergoes the same processing as the stored images. The feature vector obtained from the query image is used as the input to the GBAM, which produces the feature vectors of the associated images as the output. The corresponding images are then retrieved from the databases and presented to the user via the user interface. The most important feature of the proposed ABIR system is that we use associative memory for storage and retrieval of images. The GBAM is used to store association between feature vectors of associated images. The basic functions of the GBAM are to store associative pairs through a self-organizing process and to produce an appropriate response pattern on receipt of the associated stimulus input pattern. The stimulus and response patterns represent feature vectors that correspond to the query image and output images. Many associative memory models have been proposed to simulate human memory. However, the potential of these models has not been explored for CBIR. Linear associative memories have been studied extensively by Kohonen [6]. Bidirectional associative memories (BAMs) have studied by Kosko [7]. A two-layer network with feedback that simulates a BAM is shown Figure 2. The network is designed to map stimulus vectors

$\{\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_n\}$ to response vectors $\{\mathbf{y}_1, \mathbf{y}_2, \ldots \mathbf{y}_n\}$. In an auto-associative network, the response vector $\mathbf{y}_i$ and the corresponding identical. Associative memories are often able to produce correct response patterns even though stimulus patterns are distorted or incomplete. Conventional BAMs are used to store are retrieve pairs of stimulus and response patterns. However, if the number of associated inputs and/or outputs is more than two; that is, instead of pairs of vectors, if we want to store triplets or quadruplets of vectors, then we need to use generalized BAMs. Pairs of vectors $(\mathbf{x}_i, \mathbf{y}_i)$ can be stored with the BAM by summing bipolar correlation

If the input vectors are orthonormal then the recall is perfect.

$$\mathbf{x}_i \mathbf{x}_j = \begin{cases} 1 \text{ for } i = j \\ 0 \text{ for } i \neq j \end{cases} \tag{3}$$

If the input vectors are not orthonormal, then the output vector may contain cross talk. In a dual BAM, feedback is achieved with $\mathbf{W}^{\mathbf{T}}$ and is given by

$$\mathbf{W}^{\mathbf{T}} = \sum_{i=1}^{N} \left(\mathbf{y}_{\mathbf{i}} \mathbf{x}_{\mathbf{i}}^{\mathbf{T}}\right)^{\mathbf{T}} = \sum_{i=1}^{N} \mathbf{x}_{\mathbf{i}} \mathbf{y}_{\mathbf{i}}^{\mathbf{T}} \tag{4}$$

If we assume a nonlinear transfer function for neurons in the BAM, then the recalled output is a nonlinear function of a transformed input vector and is given by

$$\mathbf{y}_i = F(\mathbf{W}\mathbf{x}_i) \tag{5}$$

With the feedback the input vector $\mathbf{x}_i$ can be estimated as

$$\mathbf{x}_i = F\left(\mathbf{W}^T \mathbf{y}_i\right) \tag{6}$$

The simplest transfer function for the BAM is a step function. The stable reverberation corresponds to the system energy local minimum.

stimulus vector $\mathbf{x}_i$ are the same. In a hetero-associative memory, the stimulus and the response vector are not

matrices. In a $m \times n$ BAM, n neurons in layer L_1 represent field $\mathbf{F}_x = \{\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_n\}$, and m neurons in layer L_2 by field $\mathbf{F}_y = \{\mathbf{y}_1, \mathbf{y}_2, ..\mathbf{y}_n\}$. The two fields are interconnected by a $m \times n$ synoptic weight matrix **W**. The neuron states in field $\mathbf{F}_x$ and field $\mathbf{F}_y$ are the units of the short-term memory (STM). The connection matrix **W** is the unit of the long-term memory (LTM).

When the BAM neurons are activated, the network quickly evolves to a stable state of two-pattern reverberation or a
non-adaptive resonance. In order to improve recall accuracy, the output vector $\mathbf{y}_i$ can be synchronously fed back. The back-and-forth flow of distributed information quickly resonates on a fixed data pair. Humpert [4] has suggested generalization of the BAM that can store associated multiple input/output patterns. We suggest three models of the GBAM with topologies such as the bus, ring, and tree. The GBAM with tree topology is shown in Figure 3, and GBAMs with ring and bus topologies are shown in Figures 4 and 5, respectively. The bus and tree topologies are suitable for retrieving temporal images or image sequences. The generalization of the BAM to several vector fields raises the questions regarding the updating process. In a BAM, all units in a field are synchronously updated. By contrast, the sequence of updating weights in a generalized BAM is not obvious. The generalization of a BAM to several fields also raises question of interconnections. The generalized BAMs are extremely useful in content-based image storage and retrieval for multimedia applications. In addition, one needs to consider the storage capacity of the GBAM.

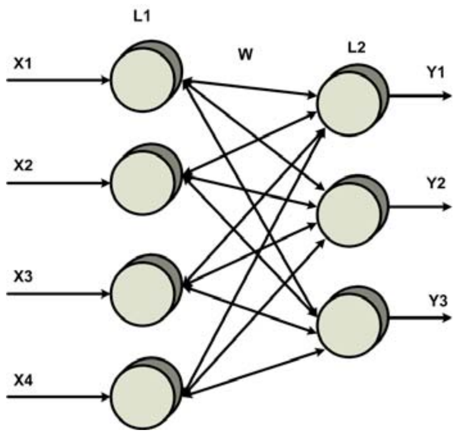

Figure 2. Bi-directional associative memory

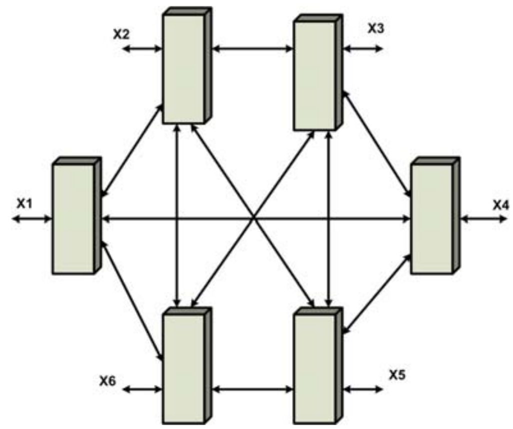

Figure 3. GBAM- Ring structure

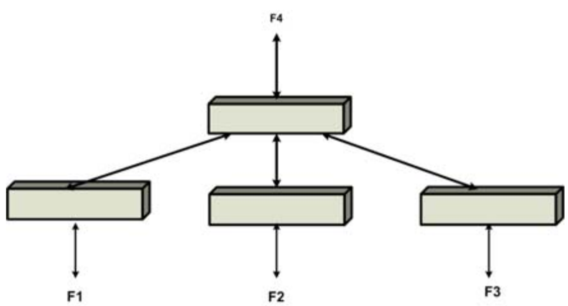

Figure 4. GBAM -Tree structure

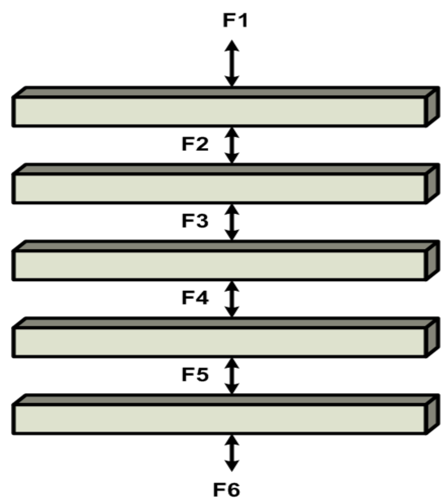

Figure 5. GBAM -Bus structure

4. Results and Discussions

We have developed software to simulate the GBAM. In the first example we have consider three sets of images [9]. The first set contained images of numbers {1, 2, 3}, the second set contained images of characters {A, B, C}, and the third set contained images of Greek characters {α, β, γ}. Each character was represented by a 12x12 matrix. Each image was represented by a feature vector of 144 elements. These feature vectors are stored in the GBAM. During retrieval, any one image (partial or noisy) from any set was used as the query image, and the corresponding images from the other sets were retrieved. The retrieved images are shown in Figures 6 and 7. The first row shows the query images and subsequent rows show corresponding retrieved images. The GBAM stores and recalls feature vectors. We can use feature vectors that may represent color, texture and/or shape of objects in the image. In the present simulation, we used a tree topology for storing three sets of images.

In the second example, the database contains three sets of images. The first set contains images flags of various countries, the second set contains images of monuments , and the third set contains images of outdoor scenes. We used the GBAM to retrieve images from the database. We assigned simulated feature vectors to index images in the database. The retried images are shown in Figures 8 and 9.

5. Conclusion

We have proposed a new image storage and retrieval system called ABIR. We try to mimic human memory. The human brain stores and retrieves images by association. We used a generalized bi-directional associative memory (GBAM) to retrieve images from a database. The method has been used successfully to retrieve associated images from the database. Our approach provides a new direction for CBIR.

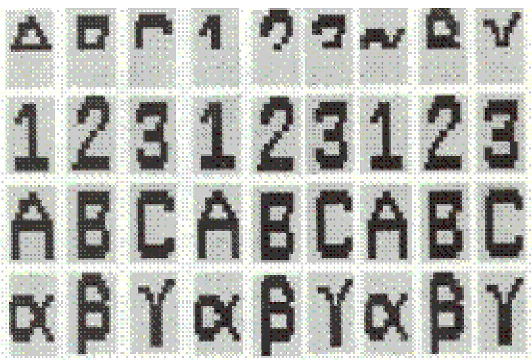

Figure 6. Query and retrieved images

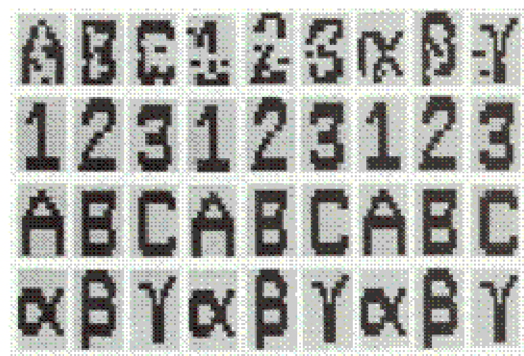

Figure 7. Query and retrieved images

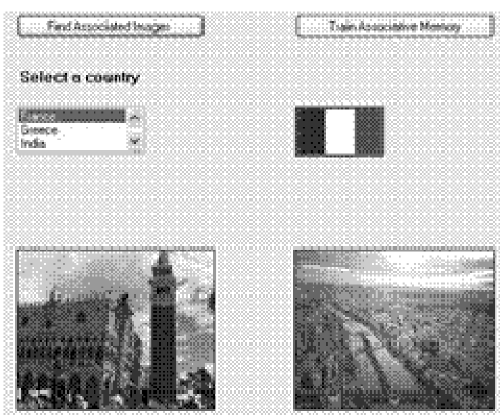

Figure 8. Query and retrieved images

Figure 9. Query and retrieved images

References

[1] T. Chang and C. C. J. Kuo, Texture analysis and classification with tree-structured wavelet transform, *IEEE Transactions on Image Processing*, vol. 2, no. 4, pp 429-441.

[2] M. Flickner, et al., Query by image and video content: The QBIC system, *IEEE Computer*, vol 9, pp 23-32.

[3] J. Haung, R. Kumar, M. Mitra, W. J. Zhu, and R. Zahib, Image indexing using color correlograms, *Proceedings of Conference on Computer Vision and Pattern Recognition*, San Juan and Puerto Rico, June 1997, pp 762-768.

[4] B. Humpert, Bi-directional associative memory with several patterns, *Proceedings of*

International Joint Conference on Neural Networks, vol. 1, San Diego, 1990, pp. 741-750.

[5] A. K. Jain and A. Vailaya, A Shape-based retrieval: A case study with trademark image databases, *Pattern Recognition*, vol. 31, no. 9, 1998, pp 1369-1390.

[6] T. Kohonen, *Self-Organization and Associative Memory*. Springer-Verlag, Berlin, 1988.

[7] B. Kosko, Bi-directional associative memories, *IEEE Transactions on Systems, Man and Cybernetics* vol. 18, 1988, pp 49-60.

[8] R. Krishnapuram, S. Medasani, S. Jung, Y. Choi, and R. Balasubramanian, Content-based image retrieval based on a fuzzy approach, *IEEE Transactions on Knowledge Engineering*, vol. 16, no. 10, 2004, pp 1185-1199.

[9] A. D. Kulkarni, *Computer Vision and Fuzzy-Neural Systems*. Prentice Hall, Upper Saddle River, NJ, 2001.

[10] R. Mehrotra and J. E. Gary, Similar-shape retrieval in shape data management. *IEEE Computer*, vol 9, pp 57-62.

[11] Y. Rui, T. S. Huang, M. Ortega, S. Mehrotra, Relevance feedback: A power tool for interactive content-based image retrieval. *IEEE Transactions on Circuits and Systems for Video Technology*, vol 8, no. 3, 1998, pp 644-655.

[12] A. W. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Recognition and Machine Learning*, vol. 22, no. 12, 2000, pp 1349-1380.

[13] R. Shekhar and S. Chaudhuri, Content-based image retrieval in presence of foreground disturbances, *Proceedings of EE, Vision Signal Processing*, vol. 153, no. 5, 2006, pp 625-638.

[14] M. J. Swain and D. H. Ballard, Color indexing, *International Journal of Computer Vision*, vol. 7, no. 1, 1991, pp 11-32.

[15] X. Wan and C. C. J. Kuo, A new approach to image retrieval with hierarchical color clustering. *IEEE Transactions on Circuits and Systems for Video Technology*, vol 8, no. 3, 1998, pp 628-643.

[16] T. Wang, Y. Rui, S. Hu, J. Sun, Adaptive three similarity learning for image retrieval, *Multimedia Systems*, vol 9, 2003, pp 131-143.