A ROBUST IMAGE RETRIEVAL SYSTEM USING WEIGHTED COMPOSITION OF FEATURES

Tohid Sedghi

Department of Electrical Engineering, Urmia Branch, Islamic Azad University, Urmia, Iran
Sedghi.tohid@gmail.com

ABSTRACT

In this paper a new and reliable method of improving retrieval performance, and which complement feature weighting is proposed. More over a framework for combination of the feature image which results in distinguishes content of in image is presented. Image retrieval based on content is one of the active research directions. In this study, the combination of Gabor filter features and tile method provide effective feature set for image retrieval. Experimental results exhibit that the proposed method yields higher retrieval accuracy than some conventional methods. The system has the potential of developing into a semantic based CBIR system by proper mathematical modeling of the texture features.

KEYWORDS: *Image Analysis, feature composition, feature extraction, Gabor filter.*

I. INTRODUCTION

Content-based image retrieval systems demonstrate excellent performance at computing low-level features from pixel representations but its output does not reflect the overall desire of the user. The systems perform poorly in extracting high-level features that include objects and their meanings, actions and feelings. This phenomenon, referred to as the semantic gap, has necessitated current research in CBIR systems towards retrieving images by the type of object or scene depicted. Analyzing and interpretation of image data in large and diverse image database, as in a CBIR system is obviously difficult because there is no prior information on the size or scale of individual structures within the images to be analyzed. The image processing and computer vision community had developed scale-space theory to deal with this problem. Scale-space theory incorporates multi-scale representation of a signal, which itself is an ordered set of derived signals intended to represent the original signal at different levels of scale. Scale-space theory is based on the fact that real-world objects exist as meaningful entities over certain ranges of scale, and human perceive them at coarse or fine scales depending on the scale of observation. According to the principle of scale-space theory, to be able to extract information from image data, a probe, sensor or operator is required to interact with the actual image structure. The information extracted is dependent on the relationship between size of the image structures and the size of the operators. In the computer age all domains of human life are in need of, and use of images for efficient services. An active research area that is in the heart of all the above domains is the extraction or retrieval of images from a database given a query (image retrieval system). An image retrieval system is a system for searching and retrieving images from a large database of digital images. The most common method of image retrieval utilizes some method of annotation such as keywords, or descriptions to the images so that retrieval can be performed over the labels. Unfortunately manual annotation is time-consuming and expensive. The answer to the previous difficulty is termed CBIR. CBIR describes the process of retrieving desired images from the image database on the basis of syntactical image features. Early research comprise of systems such as [1][2][3]. More research can be found in [4]. Most current CBIR techniques are geared towards retrieval by some aspect of image appearance, depending on the automatic extraction and comparison of image features judged most likely to convey that appearance. The features most often used include color, texture, shape, spatial information and multi-resolution pixel intensity transformations such as

wavelets or multi-scale Gaussian filtering [6]. In this paper, An robust, flexible and effective image retrieval system using weighted combination of image retrieval features, is proposed. The proposed system is comprehensive because it incorporates Gabor filters of different grid sizes and flexible because the feature weights can be adjusted to achieve retrieval refinement according to user's need and robust because the system's algorithm is applicable to retrieval in all kinds of image database. The rest of the paper is organized as follows. In Section 2, we explain the proposed system. In Section 3, we present experiments and discuss about the results, and finally Section 4 concludes the paper.

II. PROPOSED METHOD

A repeated pattern of pixels over the spatial domain is called Texture; the pattern can possibly be contaminated by noise. The repetition frequencies might result in textures that appear to be random and unstructured. Textures are the visual patterns in an image that have properties of homogeneity that do not result from the presence of only a single color or intensity. Today, the most commonly used methods for texture feature description are statistical and transform based methods. In the present work a transformed based method is used. The state-of-the-art in transformed based texture feature extraction uses Gabor wavelets. This is due to physiological research evidence that Gabor filters model the neurons in the visual cortex of the human visual system. Furthermore, Manjunath *et al.* in [5_MM96] showed that Gabor features performs better than using pyramid structured, tree-structured wavelet transform features and multi-resolution simultaneous autoregressive model. A total of twenty four wavelets were generated from the "mother" Gabor function using four scales of frequency and six orientations. Redundancy, which is the consequence of the nonorthogonality of Gabor wavelets, was addressed by choosing the parameters of the filter bank to be set of frequencies and orientations that cover the entire spatial frequency space so as to capture texture information as much as possible. The lower and upper frequencies of the filters were set at 0.04 octaves and 0.5 octaves respectively, the orientations were at intervals of 30 degrees, and the half-peak magnitudes of the filter responses in the frequency spectrum are constrained to touch each other. Each image $I(x, y)$ in the database is convolved with each wavelet in the filter bank according to the convolution equation $G_{mn}(x, y) = \int I(x - s, y - t)\overline{\phi}(s, t)$ where s and t are the dimensions of the filter and $\overline{\phi}(s, t)$ is the complex conjugate of the Gabor wavelet. Furthermore, $m \in \{1, 2, 3, 4\}$, $n \in \{1, 2, \ldots, 6\}$ correspond to the scales of frequency and orientations respectively. By assuming spatial homogeneity of texture regions the mean and the std. deviation of the magnitude of the transformed coefficients was computed according to:

$$\mu_{mn} = \int |G_{mn}(x, y)| dxdy \quad (1)$$

$$\sigma_{mn} = \sqrt{\iint (G_{mn}(x, y) - \mu_{mn})^2 dxdy} \quad (2)$$

Finally, the texture feature vector for each image is constructed using the computed values for the mean μ_{mn} and std. deviation σ_{mn} according to: $I_{fr}(x, y) = [\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \ldots, \mu_{35}\sigma_{35}]$.

Shape of an object is the characteristic surface configuration as represented by the outline or contour. Shape recognition is one of the modes through which human perception of the environment is executed. Shape is important in CBIR systems because it corresponds to region of interests in images. In CBIR system designed for specific domain such as trademarks and silhouettes of tools, shape segmentation can be automatic and effective.
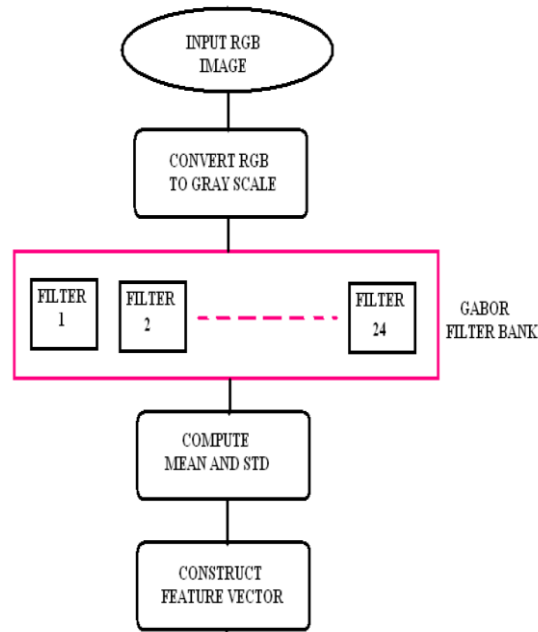
Fig. 1 Block diagram of texture feature extraction method

However this is not the case for a system having heterogeneous database. In this case shape segmentation may be difficult or sometimes impossible. In our proposal the shape features are extracted using local mean and std. deviation in a search 5×5 neighborhood employing the following formulas:

$$\mu_I(x, y) = \sum_{g=-2 \& h=-2}^{2} I(x-g, y-h) \times Z(5,5) \quad (3)$$

$$\sigma_I(x, y) = \sum_{g=-2 \& h=-2}^{2} I(x-g, y-h) \times K(5,5) \quad (4)$$

where Z and K are the impulse responses of the mean and std. deviation filters respectively. For that, each image in the database passes through a 5×5 grid size Gabor filter bank. Twenty four output images are then obtained. Afterwards, similarly to the pre-filtered image the local mean and std. deviation of each output of the filter bank is also calculated using a 5×5 neighborhood according to:

$$\mu_{pmn}(x, y) = \sum_{g=-2 \& h=-2}^{2} G_{pmn}(x-g, y-h) \times Z(5,5) \quad (5)$$

$$\sigma_{pmn}(x, y) = \sum_{g=-2 \& h=-2}^{2} G_{pmn}(x-g, y-h) \times K(5,5) \quad (6)$$

Thus for each pixel in the image there are twenty four reference pixels. Consider the pixel in the image and the twenty four corresponding pixels in the output of the filter bank, the distance between the image feature vector and any of the corresponding pixels feature vector is computed as:

$$D_{pmn}(x, y) = \left| \mu_{pmn} - \mu_I \right| + \left| \sigma_{pmn} - \sigma_I \right| \quad (7)$$

The computed distance is a measure of similarity of the texture between a pixel in the original image and the corresponding pixel at the output of the Gabor filter bank. For every pixel for each distance $D_{pmn}(x, y)$ with corresponding pixels in the output of the Gabor filter, the pixel whose texture is most likely similar to that of the database image is

$$G_{qmn}(x, y) \approx I(x, y), q = \arg \min_{i=1}^{24} \{ D_{imn}(x, y) \} \quad (8)$$

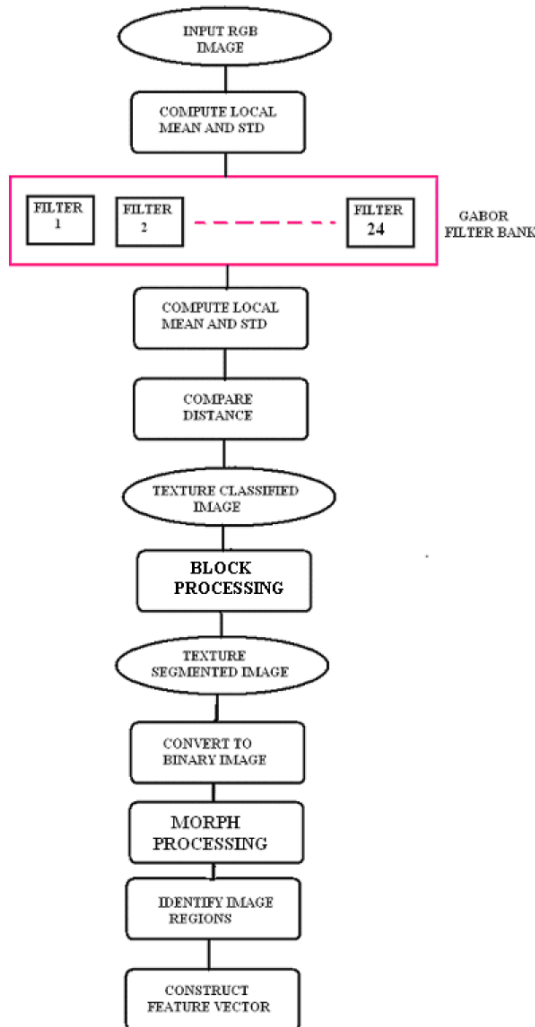These results in a filtered image called "texture classified image."

Fig. 2 Block diagram of shape feature extraction method

III. RESULT AND DISCUSSION

Have the above feature descriptors as the basis for CBIR we proceed to develop the system. The proposed system uses weighted combination of integrated Gabor texture features, shape features of texture regions and spatial information features of the texture regions. The similarity measure between the query image and each image in the database for each of texture, shape and spatial information features is carried out in the Euclidean space according to the equation

$$D_{ik} = \left| \sqrt{\{J_{ik}\}^2 - \{Q_i\}^2} \right|$$ (9)

D is the Euclidean distance
J is the database image feature vector
Q is the query image feature vector
K=1,2,..., P
P is the number of images in the database
I=1,2, 3
i=1 is the index for texture
i=2 is the index for shape
i=3 is the index for spatial information

The computed Euclidean distance between query image and database images for each of the feature vectors are normalized so it lies between 0 and 1. Normalization becomes necessary because the feature vectors have to be weighted in similarity distance calculation. Weighting of the feature vectors is necessary for a diverse database because since a particular feature cannot adequately describe an image a weighted combination will give an optimal description. In a diverse database the features suitable for retrieval of the images similar to a query image varies with the class of images with which the database is modeled or constituted. Weight assignment is the degree of relevance the similarity matching process assigns to a particular visual feature. For example if the query image is the silhouette of machine parts and the database is diverse but including images similar to the query image, it is more suitable to assign high weight to shape feature for retrieval as this is the most suitable feature for retrieval. It is useless assigning high weight to texture or spatial information except the database is narrow such that the other features particularly spatial information can be used for discrimination. This system is designed so that the user can retrieve images using flexible weight combinations. Flexible weight in a sense provides the much needed retrieval refinements and robust characteristics to the system. By flexibly combining the different weight features the retrieval process can be refined several times to satisfy the user's demand. For a query image Q, and k th number of images in the image database, if i numbers of visual feature vectors are considered for retrieval, different distances $d_{ik}(Q(x,y), I_k(x,y))$ will be obtained. The effective feature distance obtained from the weighted sum of each feature distance is given by

$$D_k = \sum_p w_p d_{ik}(Q(x,y), I_k(x,y))$$
(10)

$$\sum_p w_p = 1$$
(11)

Equation. 11 imply that the sum of the weights must be equal to unity. A weight assignment of 1 is the highest degree of relevance that can be assigned to a particular feature while assignment of zero is the highest degree of irrelevance. The retrieved images are the twelve most matched images whose feature distances are the first twelve in ascending order according to texture, shape. Spatial information and weighted features.
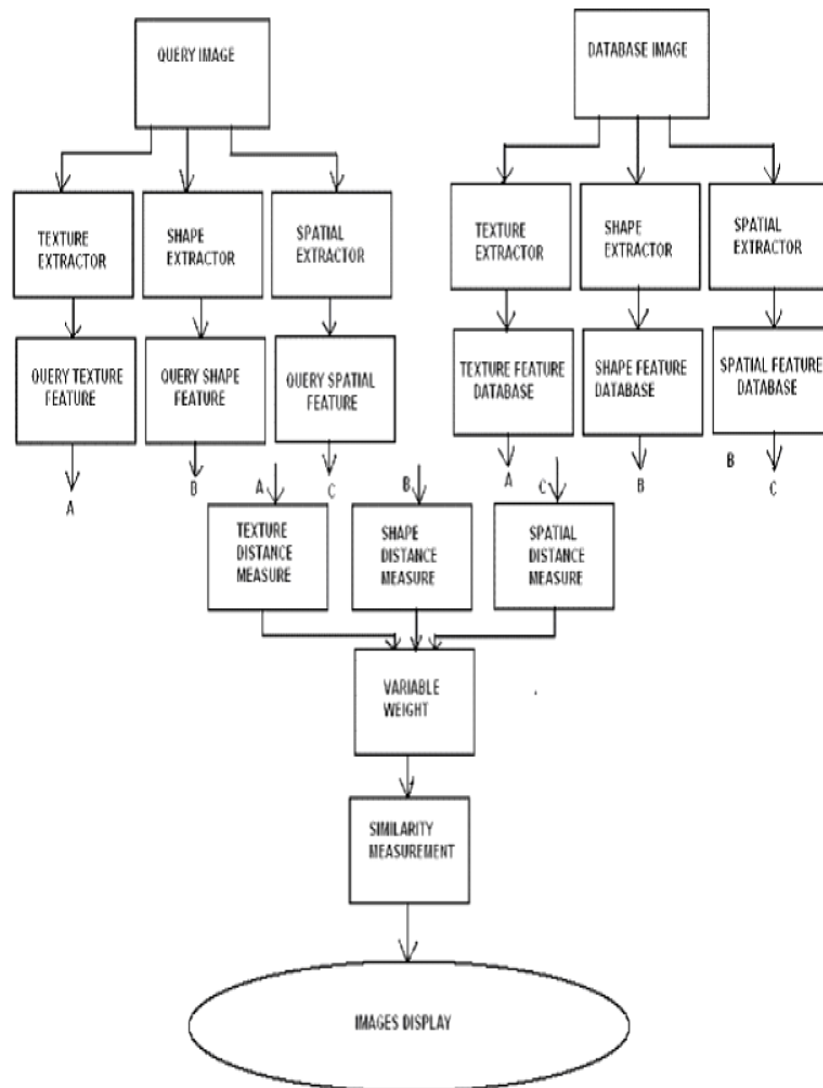
Fig. 3 Block diagram of proposed CBIR system

IV. EXPERIMENT RESULT

The proposed CBIR system was tested using the test database of Pennsylvania state university image database that consists of 800 images. The images are manually annotated in 110 categories. The performance of the proposed system is assessed using recall precision curves with the help of randomly selected query-images [6_Kor97]. *Recall* is defined as the fraction of relevant objects that are retrieved whereas *precision* is the fraction of retrieved objects that are relevant to the query. In the case under consideration the relevance or not of an image to a query-image is assessed using the annotation assigned to each individual image. Implementing a CBIR system is a painstaking process. The reason is that the Gabor filter dictionary adopted for the system design indicates the frequency of operation and the number of filters for optimal performance but it does not have a readymade answer for the filter grid size that gives optimal performance. On that, it is generally acceptable that larger Gabor grids are capable of capturing slowly varying levels than a lower grid size filter. Therefore, this aspect has been taken care by computing texture feature of the images in the database using Gabor filters of grid sizes 5×5, 15×15, 25×25, 35×35, 45×45 and 55×55. In assessing the performance of the proposed system a series of query-images are given to the system and then one precision curve is computed per query. Afterwards all the curves are averaged and this yields the so-called average recall precision curve. Figure 4 depicts the average curve for the CBIR system when only texture features were used. On the other hand Fig. 5depict the average curves when all three features are utilized. In the case of Fig. 5 the weights are 0.7, 0.15 and 0.15 respectively.
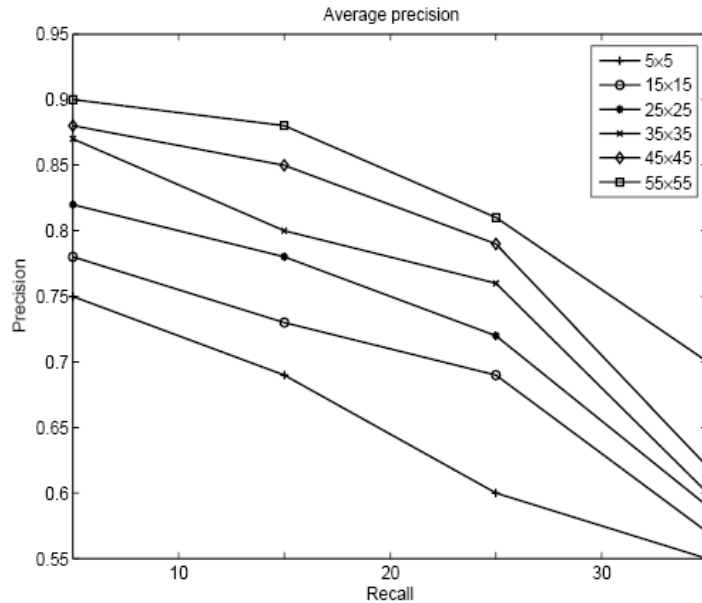
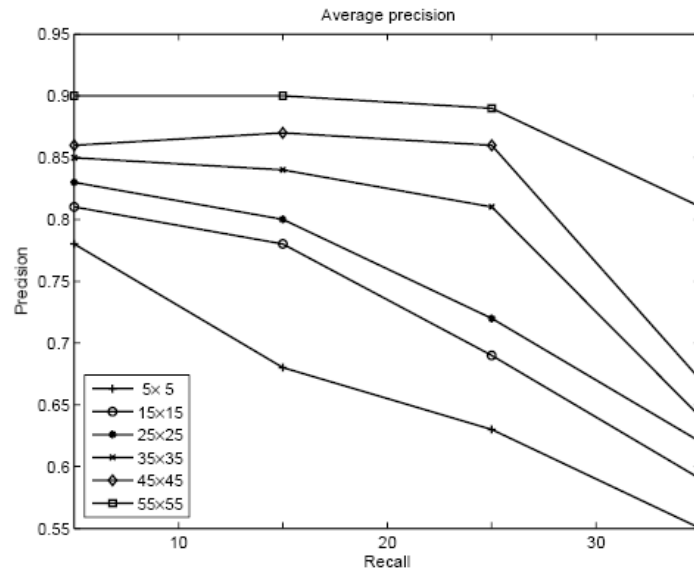**Figure 4:** The average precision of the CBIR system, using only texture.

Figure 5: The average precision when all three features are utilized with weights 0.7, 0.15 and 0.15.

V. CONCLUSION

Based on results obtained from this paper, we hereby state that the key to a breakthrough in current research in semantic image retrieval lies in the use of Gabor texture feature. Its benefits of Fourier as well as local analysis of images enable analysis of gradual changes of texture and texture variations which are essential properties of real-world scenes. The shape and spatial features are quite simple to derive and effective, and can be extracted in real time. The system is simple because of the ease with which the system can be operated and display results. The system is flexible because the feature weights can be adjusted to achieve retrieval refinement according to user's need. It is robust because the system's algorithm is applicable to retrieval in virtually all kinds of image database. In this paper a new and reliable method of improving retrieval performance, and which complement feature weighting is proposed. Since the system use Gabor filter for texture feature extraction, the proposal is weighting the features of the system as derived from various sizes of Gabor filter.

REFERENCES

[1] Sedghi T., "A Fast and Effective Model for cyclic Analysis and its application in classification" Arabian Journal for Science and Engineering Vol. 38 October 2012.

[2] T. Sedghi, "*Non Linear Transform for Retrieval System in Consideration of Feature Combination Technique*" International Journal of Natural and Engineering Sciences, Vol 6 No 3, 2012.

[3]T. Sedghi, "*Indexing of Shape Images based on Complementary Composited Features*" Australian Journal of Basic and Applied Sciences, Vol 5 No. 11, 2011.

[4] Sedghi T., Amirani C. M., Fakheri M.," *Robust and Effective Frame work for Image Retrieval Scheme using Shift Invariant Texture and Shape Features*" International Journal of Natural and Engineering Sciences 4 (1): 95-101, Vol. 3, Dec, 2010.

[5] Fakheri M., Sedghi T., Amirani M." *Region and Object Based Image Retrieval Technique Using Textural and Color Expectation Maximization Method*", International Journal of Natural Engineering science 5 (13), 19-25, Jan, 2011.

[6] B. Mashoufi, M. M. Fakheri, and T. Sedghi " *A Novel Approach for Fingerprinting Recognition*", International Review on Computers and Software, Vol. 5, No. 1, January 2010.(SCI Journal)

[6]T. Sedghi, M. Fakheri, H. Shirzad, J. Pourahmadazar, "*New Generation of Radar waveforms based Chaos Theory With Loss Detection to Targets*", International Review of   Electrical Engineering, Vol. 5, N, 4, pp 1817_1823, July, 2010.

AUTHORS

Tohid Sedghi was born in Urmia, Iran, in 1986. He received the B.S degree in Electronic engineering from IAU, Uurmia Branch, Urmia, Iran, in 2007. He received M.S degree in electrical engineering in University of Urmia, Iran in 2009. Now he is Phd student in IAU science research branch. His research interests include Localized, regional Content Based Image Retrieval and intelligent computing and Steg analysis. He has published more than twenty papers in hid referred international journal and more than 10 papers in national/international conferences.