Weighted Multi Feature Based Image Retrieval with Orthogonal Polynomials Model and Genetic Algorithm

Krishnamoorthi Ramasamy and Sathiya Devi Shanmugam

Vision Lab, Department of Computer Science and Engineering, Anna University of Technology, Thiruchirappalli, Tamilnadu, India
rkrish26@hotmail.com, sathyadevi_s@yahoo.com

Abstract. This paper proposes a new image retrieval method with weighted multi feature in multi resolution enhanced orthogonal polynomials model and genetic algorithm. In the proposed method, initially the orthogonal polynomials model coefficients are computed and reordered into multiresolution subband like structure. Then the statistical and invariant texture and shape features such as mean, standard deviation and moments are directly extracted from the subband coefficients. The extracted texture and shape features are integrated into linear multi feature set and the significance of each feature in the multi feature set is determined by assigning appropriate weight. This paper also proposes a method to compute the optimized weight for each feature in the integrated linear multi feature set using genetic algorithm. Then the obtained optimized weight is multiplied with the corresponding features in the multi feature set and the weighted Manhattan distance metric is used for retrieving similar images. The efficiency of the proposed method is experimented on the standard subset of COREL database images and yields promising results.

Keywords: Orthogonal Polynomials, Shape, Texture, Genetic Algorithm.

1 Introduction

Advances in data storage and image acquisition technologies have allowed for creating large multimedia data sets. In order to deal with this huge volume of data, it is necessary to develop appropriate information systems or tools, which can support different services such as searching, retrieving, browsing and indexing. One such information system that provides the above-mentioned services is called Content Based Image Retrieval (CBIR) system. Basically, CBIR systems try to retrieve images similar to a user defined specification or pattern based on low level content properties such as shape, texture and color. Though some of the image retrieval applications are implemented based on single feature, but the single feature is found to be insufficient for natural, web based image retrieval applications as it affects the retrieval performance. Hence recently general-purpose CBIR systems concentrate on multiple features along with some domain specific features for improving the performance of image retrieval. The combination of structure, color, shape and texture called mutifeatures for image retrieval is presented in [1-2]. Wavelet transform based multi feature extraction for image retrieval is reported in [3-4]. Since the physical meaning, importance and value ranges of each feature are different in the

D.-S. Huang et al. (Eds.): ICIC 2011, LNCS 6838, pp. 370–376, 2011.

linear combination of multi feature set, the similarity score computation with single distance metric becomes a series problem and degrades the retrieval accuracy. This problem can be solved using two methods viz., (i) relevance feedback and (ii) appropriate weight generation. Relevance feedback is computationally high demanding and difficult to incorporate human into the loop. The latter method can be viewed as an optimization problem and a suitable optimization technique has to be incorporated to generate the weight in an adaptive manner for effective discrimination and retrieval with less computational cost. Hence this paper proposes a method for optimized weight generation using genetic algorithm for multi feature representation with multi resolution enhanced orthogonal polynomials model for efficient image retrieval. This paper is organized as follows. The multi feature extraction from orthogonal polynomials enhanced multiresolution subband is presented in section 2. The genetic algorithm based optimized weight generation process is presented in section 3. The performance metric and experiments and results are discussed in section 4. Conclusion is drawn in section 5.

2 Proposed Feature Vector Extraction

In this section, texture and shape feature extraction processes for the image under analysis with orthogonal polynomials model coefficients are presented. The orthogonal polynomials model based transformation and reordering of the model coefficients into multiresolution subband like structure is presented in [5].

2.1 Texture Feature Extraction

The texture feature extraction process from the orthogonal polynomials enhanced multiresolution reordered subbands is presented in this sub section. The statistical property of the texture is characterized by the well known second order measures. In the orthogonal polynomials model, it is observed that when the block of an image is rotated the coefficient's magnitude remain unaltered but their position and the sign vary. At the same time the absolute difference between the corresponding coefficients of the original and the rotated block in the zig - zag sequence is zero. Hence the feature extraction considers the absolute value of the subband coefficients for extracting rotation invariant texture feature. Since most of the signal energy is contained in the low frequency coefficients of the reordered logarithmically spaced subbands, the texture features are extracted from them. In the proposed work, the image is reordered into four multi resolution subbands and the statistical texture features viz mean (μ_k), standard deviation (σ_k) and energy (E_k) are extracted from them. The obtained feature vector F_s of dimension 12 (i.e 4 subbands (4*3 = 12)) as

$$F_s = (E_0, \mu_0, \sigma_0, E_1, \mu_1, \sigma_1, \ldots\ldots\ldots\ldots E_{3\log_2 N+1}, \mu_{3\log_2 N+1}, \sigma_{3\log_2 N+1}) \tag{1}$$

2.2 Shape Feature Extraction

The shape features are extracted based on salient points. The salient points are extracted from edges and are used for extracting important shape feature for rough shape modeling. The shape feature extraction consists of two stages. (i) Salient point extraction with gradient and (ii) Moment feature computation.

(i) Salient Point Extraction with Gradient: The salient points are extracted in two steps: (i) Non-maximum suppression and (ii) Adaptive threshold. Conventionally, the non maximum suppression is performed in high computational cost. In order to avoid this, this paper proposes a simple method for non maximum suppression performed either horizontal (x) or vertical (y) direction depending on the magnitude of the gradient points. The coefficients of S_2 and S_3 subband posses the gradient information in horizontal (x) and vertical (y) directions and are used in non maximum suppression process. If the gradient magnitude of S_2 subband is larger and not greater than its neighbors on both the sides of x direction, then the coefficients of S_2 and S_3 are suppressed. Similar comparisons and suppressions are performed in the y direction if the coefficients of S_3 has larger magnitude. After performing the non maximum suppression on the coefficients of S_2 and S_3 subbands, the gradient magnitude G_{ij} is computed and is defined as:

$$G_{ij} = \sum_{i=0}^{X} \sum_{j=0}^{Y} (| \beta_{ij}^{'S_2} | + | \beta_{ij}^{'S_3} |) \tag{2}$$

where X and Y are the size of the subbands and $\beta_{ij}^{'S_2}$, $\beta_{ij}^{'S_3}$ are the coefficients of S_2 and S_3 subbands respectively. The adaptive threshold T, which is a sum of the mean of pixel values and the gradient magnitude is used to extract the salient points and is defined as:

$$T = \frac{1}{MN} \left(\sum_{i=0}^{M} \sum_{j=0}^{N} \beta_{ij}^{'S_1} \right) + \frac{1}{XY} \left(\sum_{i=0}^{X} \sum_{j=0}^{Y} G_{ij} \right) \tag{3}$$

where M and N are the row and column values of the S_1 subband. In the proposed method, the salient points s_i are extracted from the gradient magnitude G_{ij} based on the adaptive threshold T. Thus the adaptive threshold is effectively applied to the gradient magnitude which removes the majority of non edge responses and the salient points (s_1, s_2,…s_n) are extracted. These salient points form the salient map for extracting the shape features.

(ii) Moment Feature Computation: The image shape features are directly extracted from the salient map to characterize the shape information with Hu moment [6] which is invariant to scaling, rotation and translation. From the salient points, the seven invariant Hu's moments are computed resulting in the feature vector FV of dimension seven. The extracted feature vector FV and texture feature vector F_s both requires normalization so as to avoid domination of features having high magnitude over the others. Hence this paper uses the min-max normalization and the new maximum and minimum value of the features, which are 1 and 0 respectively in this proposed work.

2.3 Proposed GA Based Weight Generation

In this section, GA based feature weight generation process using texture and shape is presented. In the proposed weight generation method, initially the database (DB) images are subjected to two portions: (i) Training and (ii) Testing. In the training portion, training pair (TP) is generated as: A *training Pair* (TP) is the pair which consists of query image I_Q and the user defined best matched image I_M and is denoted as:

$$TP = (I_Q, I_M) \tag{4}$$

where $I_Q, I_M \varepsilon DB$. During experimentation, m training pairs are considered and each of this training pair the appropriate weight is generated iteratively (generation) using integrated weighted dissimilarity function Dt, uniform order based crossover and order based mutation with genetic algorithm. The Dt is defined below: *The Integrated Weighted Dissimilarity Function* $Dt(I_1, I_2)$ between two images I_1 and I_2 is defined as

$$Dt(I_1, I_2) = \sum_{i=1}^{n} diff(FV_{I_1}, FV_{I_2}) \frac{w_i}{\sum_{i=1}^{n} w_i} \tag{5}$$

where $diff(FV_{I1}, FV_{I2})$ is the Manhattan distance between the features of the query image I_1 and the target image I_2 and w_i is the weight associated with the features and n is the number of features. For each iteration or generation, the uniform order based crossover, order based mutation and roulette wheel selection are used. At the end of the final iteration, the chromosome that gives optimal weight in which the dissimilarity value Dt is minimum, is considered as the optimal weight. By utilizing this weight the similarity is computed with the query image and the images in the testing database including the user defined best matched image in the training pair TP based on weighted Manhattan distance. Then the obtained distances are sorted in ascending order and the rank of the best user defined image is identified.

The ranking score is computed as per *score function S(r)*, which is defined as: The *ranking score function S(r)* for the given training pair $TP = (I_Q, I_M)$ is defined as:

$$S(r) = \begin{cases} (k + 1 - r)/k & \text{if } r \le k \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

where r is the ranking of the target and k is the lowest ranking that has a score. The same process is repeated for m training pair and total count $TC(w)$ is computed as defined below: *Total Count:* The total count $TC(w)$ is the number of correct hits given by an integrated weighted dissimilarity function Dt with the set of weights w for searching in DB against training pair and is given below.

$$TC(w) = \sum_{i=1}^{m} S(r_i) \tag{7}$$

where S is the ranking score function, m is the number of training pairs, r_i is the ranking for the target i using Dt and w is the set of weight in Dt. The optimal weight for each feature is obtained in such a way that, in which training pair, the set of weight will give the maximum $TC(W)$ value. In other words, the optimal weight w is obtained when $TC(W)$ is maximized and is defined as:

$$\text{argmax}_w TC(w) \tag{8}$$

4 Experiments and Results

The retrieval efficiency of the proposed method is experimented with a subset of popular image database COREL [7] and the experimental results are presented in this section. We have considered five classes of images such as Dinosaur, Elephant, Rose, Bus and Waterfall, each class containing 100 images. The images in the Corel database are of color images and hence are converted into gray scale. These images are of size (256 x 256) with the pixel values in the range 0 – 255. During experimentation, the image under analysis is divided into (2 x 2) non overlapping blocks and each block is subjected to the orthogonal polynomials transformation. The transformed coefficients

Table 1. Weight for all the features using GA

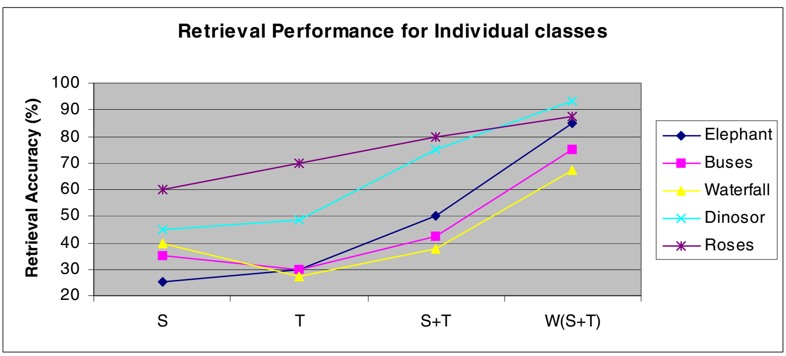
Texture + Shape Features	Optimized Weight value	Texture + Shape Features	Optimized Weight value	Texture + Shape Features	Optimized Weight value	Texture + Shape Features	Optimized Weight value	Texture + Shape Features	Optimized Weight value
μ_0	0.264	σ_1	0.623	E_2	0.491	Φ1	0.787	Φ6	0.220
σ_0	0.003	E_1	0.220	μ_3	0.264	Φ2	0.003	Φ7	0.230
E_0	0.623	μ_2	0.538	σ_3	0.268	Φ3	0.787		
μ_1	0.264	σ_2	0.003	E_3	0.220	Φ4	0.538		

are reordered into multi resolution subband like structure and the same process is repeated for all the blocks in the image. The image is then decomposed into one level and each subband is named as S_1, S_2, S_3 and S_4. The statistical texture features are extracted from these subbands resulting in the feature vector $F_{s.}$ The S_2 and S_3 subband coefficients posses the first order derivative information in horizontal and vertical directions and are used to extract edge information with non maximum suppression. After the non maximum suppression process, gradient magnitude G_{ij} is obtained using equation (2) and adaptive threshold is applied on G_{ij} to extract the salient map. The moment feature is extracted from the salient map so as to result a rough shape feature vector of dimension 7. Then the global multi feature vector of dimension 19 is obtained by combining texture and shape feature vector and is normalized. These features are stored in the feature database. The same process is repeated for all the images in the database and the corresponding multifeatures are stored in the feature database. Then from each class fifty percentage of images are taken for training to compute the optimized weight using GA. The weight assignment process with genetic algorithm is performed as described in section 3. During the weight generation process, the following parameters are used: (i) *PopSize* (*P*) = 50 (ii) *Training Pair* (*TP*) = 35 (iii) *Cross over Probability* = 0.7 (iv) *Mutation Probability* = 0.02 (v) *Maximum no. iterations* = 100 (vi) *Database size* (*D*) = 500 and (vii) k = 50. The genes in the population take the value between 0 and 1. (i.e., the weights are assigned between 0 and 1). The value of k in the ranking score function is set to 50, because the total number of images considered in this experiment is 500 and 10% of total number of images is 50. Therefore, if the target is ranked in the first position, total count $TC(w)$ is increased by 1. If the target is ranked in the second position, $TC(w)$ is increased by 49 / 50 and so on. The above mentioned process is repeated for the all training image pairs. The experiments are repeated 10 times with different random seeds. The optimized weights for all the 19 features are computed and are shown in the table 1. The weights are converged in the 135th iterations and the weights also exhibit human visual perception. All the training pairs are ranked between 3th and 16th places in the ranked retrieval except the waterfall pair.

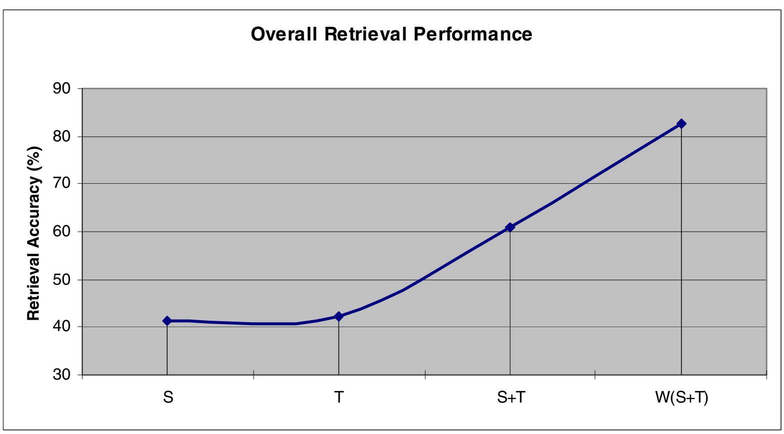
In order to strengthen the experimental result of the proposed weighted multi feature set, the experiment is conducted with four categories of features (i) multi feature with optimized weight ((W(S+T)), (ii) multi feature (S+T) (Combination of shape and texture features), (iii) Shape feature (S) and (iv) Texture feature (T). The performance of the proposed method is measured in terms Average Retrieval Ratio (ARR) which is defined as:

$$ARR = \frac{1}{N} \sum_{i=1}^{N} \frac{m_i}{N} \tag{9}$$

where N is the total number of similar images in one category and m_i is the number of retrieved relevant images. The retrieval accuracy of each class with the weighted multi feature set is obtained by considering each image in the database as query image and the above mentioned process is repeated to extract the shape and texture features. Then the obtained multifeature set is multiplied by optimized weight as shown in table 2. The similarity is computed with integrated weighted dissimilarity function as described in the equation (5) between the query image and the images in the database. Then the performance is evaluated using average precision. The obtained retrieval accuracy of all the five classes are plotted as a graph with the categories in X axis and the retrieval accuracy in Y axis and the same is presented in figure 1(a). Then the overall retrieval accuracy of all the five classes are computed. The obtained result is plotted as a graph and is shown in the figure 1(b). The overall accuracy for the



(a)



(b)

Fig. 1. (a) Retrieval Accuracy of individual classes with five categories. (b) Overall Retrieval Accuracy of individual classes with four categories.

proposed weighted multi feature set is 82.73%. In order to strengthen the experimental result of the proposed method, the experiment is conducted with multi feature set without assigning weight for the features. Similarly the retrieval accuracy of each class using the multi feature set is obtained and the results are incorporated in the same figure 1(a). The overall retrieval accuracy is presented in figure 1(b). From the result we inferred that, the retrieval performance increases sharply for the multi feature set with optimized weight category than multi feature set alone. The experiments are performed by considering the shape and texture features individually and the retrieval accuracy is measured. The obtained retrieval result for all the classes are incorporated in the same figure 1(a). From the figure, it is observed that texture feature outperforms compared to shape feature. The overall retrieval accuracy for these two categories are 41.36% and 42.28% respectively and the same are also plotted as a graph and is shown in figure 1(b). From the graph, it can be observed that the multi feature set with weight yields 82.73% of retrieval accuracy compared with other categories. Hence from the various categories of experimental result, it is inferred that, considering more than one feature with appropriate weight yield better retrieval result. The multifeature is the more discriminative power than single feature in general purpose image retrieval system.

5 Conclusion

In this paper a new image retrieval method with weighted multi feature in multi resolution enhanced orthogonal polynomials model and genetic algorithm is proposed. The transformed coefficients are reordered into multiresolution subband like structure and the statistical and invariant texture and shape features are directly extracted from them. The extracted features are integrated into linear multi feature set and the significance of each feature in the multi feature set is determined by assigning appropriate weight with genetic algorithm. The proposed method gives promising result.

References

1. Iqbal, Q., Aggarwal, J.K.: Combining Structure, Color and Texture for Image Retrieval: A Performance Evaluation. In: 16th International Conference on Pattern Recognition, vol. 2, pp. 438–443 (2002)
2. Howe, N.R., Huttenlocher, D.P.: Integrating Color, Texture, and Geometry for Image Retrieval. In: IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 239–246 (2000)
3. Yumin, T., Lixia, M.: Image Retrieval Based on Multiple Features Using Wavelet. In: Fifth International Conference on Computational Intelligence and Multimedia Applications, pp. 137–142 (2003)
4. Huang, R.B., Du, M.H.: Object Region Based Color Image Retrieval Integrating Multi features. In: Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition, pp. 1729–1734 (2007)
5. Krishnamoorthy, R., Sathiya Devi, S.: Design of Fusion Texture Feature with Orthogonal Polynomials Model and Co-occurrence Property for Content Based Image Retrieval. Lecture Notes in Engineering and Computer Science, vol. 2173(1), pp. 1172–1177 (2008)
6. Hu, M.K.: Visual Pattern Recognition by Moment Invariants. IRE Transactions on Information Theory 8(2), 179–187 (1962)
7. `http://www.corel.com`