Abstract - This paper proposes a simple color image retrieval method based on multi resolution enhanced orthogonal polynomials model. In the proposed method, a set of orthogonal polynomials has been chosen and the model coefficients are reordered into multiresolution subband like structure. The subband coefficients are quantized into three levels. Then the weighted autocorrelogram is computed from the quantized subbands in R, G and B color space. The obtained weighted autocorrelogram is termed as global color feature vector and are normalized with Z - score normalization. This feature vector is used for retrieving similar images with weighted Manhattan distance metric. The efficiency of the proposed method is experimented on a subset of standard COREL database and the results are compared with existing techniques. The proposed method yields significant retrieval results.

Keywords: Multiresolution, Orthogonal Polynomials, weighted autocorrelogram, color image retrieval.

1 Introduction

With the development of computer vision technology and the increased emphasis on multimedia applications, huge volume of digital images has been accumulated. Searching and retrieving of images from these collections containing user-specified characteristics have become an important and challenging task. The classical image searching and retrieving systems are text-based meaning that searches for images based on one or more keywords specified by the user. However, there are cases in which a query request cannot be easily described by keywords and manual indexing of keywords are also cumbersome. To solve this dilemma, the Content-Based Image Retrieval (CBIR) has been emerged for effectively searching and retrieving images based on a given query image. This approach retrieves images on the basis of visual content such as color, texture, shape or spatial relationship of objects within the image. Thus the visual content becomes the features and these features are stored in an image feature database for future use. When a query image is given, these features of the query image are extracted to match the features in the feature database by a pre-established algorithm, so that a group of similar images to the query image can be returned as the retrieval images. The features of the images in the CBIR systems include color [1 - 3], texture [4 - 5], shape [6 - 8] and the combination of these features [9 – 11]. In general, human eye is more sensitive to color than texture and shape and color alone acts as a more discriminative feature for some retrieval applications. Also the color feature is robust to background complication and is independent of image size and orientation. Thus, retrieval systems using color have been popular for a long time.

Early color based retrieval systems have used the global RGB histogram information such as the Local Color histogram [12], histogram difference approach [13], histogram intersection [1 – 2, 14] and the quadratic histogram comparison [15]. Though the color histogram based approaches are extremely easy to compute and insensitive to small changes in viewing positions and partial occlusion, they do not capture local spatial color information. Hence this approach is liable to false positives and is not robust to large appearance changes. Another approach called Color moment [16 – 17] has been used in retrieval systems for extracting colors, especially for retrieval of images only containing the objects of user’s interest. Stricker et al. [18] showed that characterizing one dimensional color distributions with the first three moments is more robust and efficient than working with color histograms. Several recent schemes viz, Color Coherence Vector [19], Color Correlogram [20] and Binary Color Set [21 – 22] incorporate spatial correlation of color regions as well as the global distribution of local spatial correlation of colors to improve upon the histogram method. Though these techniques perform better than traditional histograms, they require intensive computation.

Generally color spatial techniques are classified into three categories [23] for image retrieval: (i) partition based approach (ii) signature based approach and (iii) cluster based approach. T. Chua et al. [24] have established a signature based approach in which an image is represented by its major dominant colors. The dominant colors consist of the colors that have highest frequencies in the global color histogram of the image. In order to represent the spatial distribution of colors, the image is partitioned into (n x m) cells of equal size, where each cell is assigned an index k in the range [1, 2, ..., (n x m)]. A bit-string is assigned to each dominant color to
describe its spatial distribution. A bit \( k \) is set to 1 if the cell number \( k \) contains significant number of pixels of that color. The image representation consists of the set of all bit-strings (also called bit signatures) and this representation has interesting advantages when compared to traditional histogram based representation, but it is not invariant to rotation or scaling.

J. Seongtaek et al. [25] produced a better result than T. Chua et al. [24] approach by introducing two basic representations: (i) Dominant Color Composition (DCC) and (ii) Dominant Color Distribution (DCD). The DCC and DCD describe the dominant colors and the spatial distribution of the dominant colors of the image respectively. The DCD signature consists of two bit strings called DCD\(_a\) and DCD\(_b\), and these bit strings record the dominant colors that are found in the horizontal and vertical axis of the image. This approach can be used for the retrieval of rotated images but like T. Chua et al.'s [24] approach, it is not also suitable for retrieving scaled images. V. Chitkarraicu et al. [26] introduced a compact representation of colors found in the image and this method has considered both dominant and less dominant colors. The percentage of each color in the image is calculated and stored in bins that accommodate varying percentage of compositions. This representation is called Variable Bin Allocation (VBA) and the spatial distribution of color information is captured by the \((n \times m)\) cells that constitute the image partition. Each cell of the image is described by a bit-string, representing its compact signature. The similarity between two images are calculated as the cumulative similarities of their respective cells. L. Cinque et al. [21] have introduced partition based approach, in which the image is partitioned into \((n \times m)\) regions and represent it using a 2D coordinate system. The spatial distribution of each color is represented by mean and standard deviation of pixels having that color. This approach is mainly suitable for continuous and homogeneous regions and it is not invariant to rotation and scaling. M. S. Kankanhalli et al. [27] have developed a cluster based approach for color image retrieval. The dominant colors (called color clusters) are first extracted and their spatial distribution is then described by performing a connected component labeling. Each connected component is called spatial cluster. The Euclidean distance is used both for clustering and similarity measurement process. A. Abdesselam et al. [28] have introduced a cluster based approach that produces better result than M. S. Kankanhalli et al. approach [27]. In this, a predefined HSV color set is constructed instead of the RGB color set. Each pixel in the image is assigned to one color among the \(n\) predefined color clusters using clustering process. The image is partitioned into \((m \times m)\) sub areas to get the spatial distribution of the color. For each subarea, dominant clusters are obtained to form the so called Color Cluster Distribution (CCD) image that captures the spatial distribution of the colors. The image similarity is defined by the cumulative distance between all the corresponding sub-regions in each orientation. This approach is also capable of retrieving rotated images (main rotations only such as 90°, 180°, etc.) but cannot retrieve scaled images.

According to K. L. Tan et al. [23], signature based approach is generally better than partition and cluster based approaches in terms of retrieval effectiveness (which is measured by precision and recall) and efficiency (which is measured by retrieval time). And the image partitioning approach depends on pixel position and hence it is unlikely to tolerate large image appearance changes. The same problem occurs in the histogram refinement method, which depends on local properties to further refine color buckets in histograms. Recently, frequency domain methods, such as wavelet based methods, which provide better local spatial information in transform domain, have been used [29 – 30] for extracting color feature for effective image retrieval. Among different approaches, Color correlogram have been shown to be robust to large image appearance changes and to be more effective in image retrieval than histograms both in spatial and frequency domains. The color correlogram method, however, takes into account the local spatial correlation between colors as well as the global distribution of this spatial correlations. Since color correlogram has high dimension and requires more computation, a simplified and subset of correlogram version called autocorrelogram is often used. Hence in this paper a simple multiresolution enhanced orthogonal polynomials based weighted autocorrelogram feature is proposed for effective color image retrieval.

This paper is organized as follows: Section 2 describes the quantization and weighted autocorrelogram computation process based on the multiresolution enhanced orthogonal polynomials model. Similarity and performance measure is discussed in section 3. In section 4, the experimental results of the proposed method and comparison with other existing methods are presented. Finally conclusion is given in Section 5.

2 Proposed Color Feature Extraction

The orthogonal polynomials model and reordering of these coefficients into a multiresolution like structure are described in [31]. Based on the reordered coefficients, this section, describes a color feature extraction called Weighted Multiresolution Enhanced Orthogonal Polynomials Autocorrelogram (WMEOPA). The color feature extraction consists of two steps: (i) Quantization (ii) Weighted Multiresolution Enhanced Orthogonal Polynomials Autocorrelogram (WMEOPA) extraction. Space - frequency decomposition of the input signal is an important property of the wavelet transform and it encourages for applying the spatial (pixel) domain technique such as autocorrelogram in the multiresolution enhanced orthogonal polynomials coefficients in RGB color space. The subband \(S_0\) of the reordered multiresolution structure is a coarse version of the original image and is modeled to carry out the color frequency and energy distribution. The remaining subbands \(S_1, S_2, S_7,\)
2.2 Weighted Multiresolution Enhanced Orthogonal 
Q. 

where \( \mu \) is mapped from \( \Gamma \) of that window is termed as root pixel (\( r \)). The value for \( r \) is assigned in the range \{ 0, 1,2 \} as defined as follows:

\[
Q = \begin{cases} 
0 & r \leq \mu - t \ast \sigma \\
1 & \mu - t \ast \sigma < r < \mu + t \ast \sigma \\
2 & r \geq \mu + t \ast \sigma 
\end{cases} 
\]  

(1)

where \( r \) is the central pixel (root pixel), \( \mu, \sigma \) are the mean and standard deviation of the window and \( t \) is the constant. After this quantization procedure the value of \( r \) is mapped from \( r \) to \( Q \).

2.2 Weighted Multiresolution Enhanced Orthogonal Polynomials Autocorrelogram (WMEOPA) Computation

Having quantized the subband coefficients into three levels in the previous sub section, the procedure for computing the autocorrelogram features from these coefficients is presented in this sub section. Since color correlogram is an extension of correlogram, first color correlogram is defined. A color correlogram (henceforth correlogram) expresses how the spatial correlation of pair of colors changes with distance. Generally a correlogram for an image is a table indexed by color pairs, where the \( k \) th entry for the row specifies the probability of finding a pixel of color \( j \) at a distance \( k \) from a pixel of color \( i \) in the image and \( k = \{1, 2, 3,4 \} \). Let \( I \) be an image of size \( (N \times N) \) quantized into \( m \) colors \( C_1, ..., C_m \). For a pixel \( p = (x, y) \in I \), let \( I(p) \) denotes its color value and \( I(C_i) = \{ p \mid I(p) = C_i \} \). Hence the correlogram \( \Gamma_{c_i,c_j}^k \), for the quantized color pair \((C_i, C_j)\) and a pixel distance \( k \leq d \) can be defined as:

\[
\Gamma_{c_i,c_j}^k (I) = \Pr_{p_1 \in (C_i), p_2 \in (C_j)} [p_2 \in I(C_j) \mid |p_1 - p_2| = k] 
\]  

(2)

where \( C_i, C_j \in \{C_1, ..., C_m \} \), \( k = \{1, ..., d \} \) and \(|p_1 - p_2| \) is the distance between the pixels \( p_1 \) and \( p_2 \) in \( L_1 \) norm. Since the feature vector size of correlogram is \((O(m^2d))\), a simplified version of the feature called the color correlogram \( \alpha_{C_i}^k (I) \) is used. It captures the spatial correlation between identical colors and thus reduces the feature dimension into \((O(md))\), which is defined as:

\[
\alpha_{C_i}^k (I) = \Gamma_{C_i,C_j}^k (I) 
\]  

(3)

The autocorrelogram features are weighted with respect to the decomposition level from which it is derived. Hence the weight assignment to the autocorrelogram feature from each subband is defined as:

\[
w_i = \frac{1}{2^i} \quad i = 1,2, ..., l 
\]  

(4)

where \( l \) is the number of subband decomposition levels.

In the proposed work, the orthogonal polynomials model coefficients are reordered into three levels, resulting in 10 subbands. The subbands \( S_0, S_1, S_2 \) and \( S_3 \) are in 3rd level and \( S_4, S_5 \) and \( S_6 \) are in 2nd level whereas \( S_7, S_8 \) and \( S_9 \) subbands are in 3rd level. Since most of the color related information are contained in the \( S_0 \) subband and the WMEOPA features are extracted by considering the neighborhood of subband coefficients in all the directions. The WMEOPA feature of \( S_0 \) subband in \( R \) channel is defined with respect to equation (3) as:

\[
AC_{C_i}^k (S_0) = \frac{|\{ (x,y) | S_0(x,y) = C_i, S_0(x+k,y+z+k) = C_i \}|}{8k \times |\{ (x,y) | S_0(x,y) = C_i \}|} 
\]  

(5)

where \( C_i \) is the quantized color and \( k \) is the correlation distance. Similarly the vertical autocorrelogram is computed from the \( S_1, S_4 \) and \( S_7 \) subband coefficients as

\[
AC_{C_i}^k (S_1) = \frac{|\{ (x,y) | S_1(x,y) = C_i, S_1(x+k,y) = C_i \} \mid 2k \times |\{ (x,y) | S_1(x,y) = C_i \}|} 
\]  

(6)

The coefficients of \( S_2, S_3 \) and \( S_8 \) subband are used to compute the horizontal autocorrelogram as

\[
AC_{C_i}^k (S_2) = \frac{|\{ (x,y) | S_2(x,y) = C_i, S_2(x,y+k) = C_i \} \mid 2k \times |\{ (x,y) | S_2(x,y) = C_i \}|} 
\]  

(7)

The denominator in the equations (5), (6) and (7) denotes the total number of pixels at a distance \( k \) from any pixel of color \( C_i \) and the factor \( 8k \) and \( 2k \) are due to the properties of \( L_1 \) norm used to compute distance between pixels. Since the local correlations between colors are more significant than global correlations in an image, a small value of \( k \) is sufficient to capture the spatial correlation. Hence in the proposed method the value of \( k \) is considered as unity. Since there are three quantized levels the resultant feature vector of dimension 21 (\( 7 \times 3 = 21 \)) is obtained for \( R \) channel. The obtained features are multiplied with the weight according to the equation (4). The above mentioned process is repeated for the remaining \( G \) and \( B \) channels for extracting the color autocorrelogram feature. Thus the feature vector of \( R, G \) and \( B \) channels are
combined to form the global feature vector \( GFV \) of dimension 63 and is represented as:

\[
GFV = \{ Cor_1, Cor_2, Cor_3, ..., Cor_{63} \}
\] (8)

The feature vector values are normalized using Z-score normalization and is defined as:

\[
Cor' = \frac{Cor - mean_{GFV}}{stand_{dev}_{GFV}}
\] (9)

where \( Cor' \) is the normalized value, \( Cor \) is the original feature value, \( mean_{GFV} \), \( stand_{dev}_{GFV} \) are the mean and standard deviation of the feature vector elements. Thus the normalized global feature vector \( GFV_n \) is termed as:

\[
GFV_n = \{Cor'_1, Cor'_2, Cor'_3, ..., Cor'_{63} \}
\] (10)

3 similarity and performance measure

Having extracted the color autocorrelation in the previous section, this section describes the retrieval of relevant images from the database images against the query image using the similarity measure since it is a key component of the content based image retrieval system. In the proposed retrieval scheme, similarity between the query image and the images present in the database are calculated using the well known weighted Manhattan distance metric. Let \( CFV_n(Q) = \{Cor'^{Q}_1, Cor'^{Q}_2, Cor'^{Q}_3, ..., Cor'^{Q}_{63} \} \) and \( CFV_n(D) = \{Cor'^{D}_1, Cor'^{D}_2, Cor'^{D}_3, ..., Cor'^{D}_{63} \} \) be the normalized WMEOPA of a query image \( Q \) and database image \( D \) respectively. The distance between \( Q \) and \( D \) is measured using weighted Manhattan distance as:

\[
d(Q,D) = \frac{\sum_{i=1}^{n} | Cor'^{Q}_i - Cor'^{D}_i |}{\sum_{i=1}^{l} W_i}
\] (11)

where \( n \) and \( l \) are the length of the feature vector, number of decomposition levels respectively and \( W_i \) is weight defined as in equation (4).

The performance of the proposed method is measured in terms Average Retrieval Ratio (ARR) which is defined as:

\[
ARR = \frac{1}{N} \sum_{i=1}^{N} \frac{m_i}{N}
\] (12)

where \( N \) is the total number of similar images in one category and \( m_i \) is the number of retrieved relevant images. The performance is also measured with popular measures such as precision and recall rate. Recall rate is the ratio of number of relevant images retrieved and the total number of relevant images in the collection. The precision rate is the ratio of the number of relevant images retrieved and total number of images in the collection.

4 Experiments and Results

The retrieval efficiency of the proposed method is experimented with a subset of standard COREL [32] image database and the experimental results are presented in this section. The images in the Corel database are of color images. They are also resized into \((256 \times 256)\). In this experiment seven classes of images such as Dinosaurs, Elephant, Roses, Buses, Vintage Cars, Penguins and Pigeons are considered and each class contains 100 images. During experimentation, the R, G and B color pallets are extracted from the image under analysis. For R channel, the image is divided into \((8 \times 8)\) blocks and the orthogonal polynomials transformation is applied to each block. Then the transformed coefficients are reordered into three level multiresolution like structure resulting in ten subbands. The subbands are named as \(S_0, S_1, ..., S_9\). Then the subbands are quantized as described in subsection 2.1. For quantization, the proposed method considers the overlapping window of size \((3 \times 3)\) and the root values are mapped into 0, 1 or 2 according to the equation (1). Based the experiments, the value of \( t \) is considered as 0.3 in equation (1). The same process is repeated for all the subbands in the image. After quantization, the Weighted Multiresolution Enhanced Orthogonal Polynomials Autocorrelation (WMEOPA) of \( S_0 \) subband in all the directions, the horizontal autocorrelation using the subbands \( S_7, S_8 \) and \( S_9 \) and vertical autocorrelation using the subbands \( S_2, S_5 \) and \( S_8 \) are computed as described in sub section 3.2. The above mentioned process is repeated for G and B channels and the WMEOPA is computed. Thus the feature vector of R, G and B channels are combined to form the global feature vector \( GFV \) of dimension 63. The GFV is normalized using Z-Score normalization as described in equation (9) and the normalized global feature vector \( GFV_n \) is obtained. The same process is repeated for all the images in the database and the weighted autocorrelation features are stored in the feature database.

In order to measure the performance of the proposed weighted autocorrelation scheme, the sample query image rose is considered from Corel database and is shown in figure 1(a). The normalized global color feature vector \( CFV_n \) for this query image is computed similar to the process of database images as described above. Then the distances are computed with weighted Manhattan Distance between the query image and the images that are present in the database. The obtained distances are sorted in ascending order. The top 10 retrieval
results by the proposed method for this query image are shown in figure 1(b). The retrieval rate is 90% for this query image and from this figure it can be observed that the proposed method could retrieve the same class of images as query image and perceptually similar images.

Figure 1. (a) Query Image (b) Retrieval results obtained with the proposed weighted autocorrelogram corresponding to the query image.

Instead of considering one query image, we conduct experiments by taking each of the database images as query image and the performance of the proposed method is evaluated in terms of average precision and recall. We obtain the average recall and precision as 0.89 and 0.68 respectively with the proposed technique and these results are presented in table 1.

Table 1. Performance measure of the proposed and other existing methods

<table>
<thead>
<tr>
<th>Methods</th>
<th>Recall</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed Method (WMEOPA)</td>
<td>0.89</td>
<td>0.68</td>
</tr>
<tr>
<td>Local Color Histogram (LCH)</td>
<td>0.72</td>
<td>0.32</td>
</tr>
<tr>
<td>Scalable Color Descriptor (SCD)</td>
<td>0.73</td>
<td>0.34</td>
</tr>
<tr>
<td>Color Correlogram (CCQ)</td>
<td>0.68</td>
<td>0.31</td>
</tr>
<tr>
<td>Wavelet Correlogram (WC)</td>
<td>0.75</td>
<td>0.35</td>
</tr>
</tbody>
</table>

In order to evaluate the effectiveness and efficiency of the proposed weighted autocorrelogram feature based color image retrieval, performance comparisons are made with Local Color Histogram (LCH) [12], Scalable Color Descriptor (SCD) [33], Color Correlogram (CCQ) [20] and Wavelet Correlogram (WC) [34] methods in terms average precision and recall, feature extraction and retrieval time and feature vector dimension. The average recall and precision of the existing methods are measured using the above mentioned procedure and the results are incorporated in the same table 1. It is evident from the tabulated result that the proposed method is found to perform well when compared to the other four existing methods. The feature vector dimensions for the proposed and the existing methods are tabulated in table 2.

Table 2. Feature vector (FV) dimensions of proposed and other methods

<table>
<thead>
<tr>
<th>Methods</th>
<th>FV Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed Method (WMEOPA)</td>
<td>63</td>
</tr>
<tr>
<td>Local Color Histogram (LCH)</td>
<td>96</td>
</tr>
<tr>
<td>Scalable Color Descriptor (SCD)</td>
<td>64</td>
</tr>
<tr>
<td>Color Correlogram (CCQ)</td>
<td>96</td>
</tr>
<tr>
<td>Wavelet Correlogram (WC)</td>
<td>96</td>
</tr>
</tbody>
</table>

From the table it is observed that the proposed method has less dimension when compared to other existing methods.

Table 3. Performance of various methods color feature retrieval time (On Pentium IV machine)

<table>
<thead>
<tr>
<th>Methods</th>
<th>Feature Extraction Time (for query image)</th>
<th>Searching Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed Method (WMEOPA)</td>
<td>0.364</td>
<td>0.231</td>
</tr>
<tr>
<td>Local Color Histogram (LCH)</td>
<td>0.463</td>
<td>0.234</td>
</tr>
<tr>
<td>Scalable Color Descriptor (SCD)</td>
<td>0.729</td>
<td>0.359</td>
</tr>
<tr>
<td>Color Correlogram (CCQ)</td>
<td>3.416</td>
<td>0.320</td>
</tr>
<tr>
<td>Wavelet Correlogram (WC)</td>
<td>8.050</td>
<td>0.244</td>
</tr>
</tbody>
</table>

6. CONCLUSION

In this paper, a new multiresolution enhanced orthogonal polynomials based weighted autocorrelogram method for color image retrieval has been proposed. The transformed coefficients are reordered into multiresolution subband like structure and are quantized into three levels. The weighted autocorrelogram features in horizontal and vertical directions are extracted from the subbands in RGB color space and are normalized. The proposed scheme is experimented with a subset of COREL database and is compared with some existing methods. The proposed scheme gives better retrieval result.
REFERENCES:


32. www.corel.com
