HSV Color Motif Co-Occurrence Matrix for Content based Image Retrieval

K. N. Prakash¹, K. Satya Prasad²
¹Research Scholar, ²Rector & Professor
Department of Electronics and Communication Engineering
Jawaharlal Nehru Technological University, Kakinada
Andhra Pradesh, India

ABSTRACT

In this paper, HSV based color motif co-occurrence matrix (HSV-Motif) is proposed for content based image retrieval (CBIR). The HSV-Motif is proposed in contrast to the RGB based color motif co-occurrence matrix (RGB-Motif). First the RGB (red, green, and blue) image is converted into HSV (hue, saturation, and value) image, then the H and S images are used for histogram calculation by quantizing into Q levels and the local region of V (value) image is represented by seven motif, which are evaluated by taking into consideration of local difference between the pixels. Motif extracts the information based on distribution of edges in an image. Two experiments have been carried out for proving the worth of our algorithm. It is further mentioned that the database considered for experiments are Corel 1000 database (DB1), and MIT VisTex database (DB2). The results after being investigated show a significant improvement in terms of their evaluation measures as compared to RGB-Motif.

Keywords
Color; Texture; Feature Extraction; Local Binary Patterns; Image Retrieval.

1. INTRODUCTION

Since the 1990s, content-based image retrieval (CBIR) has become an active and fast-advancing research area in image retrieval. In a typical CBIR, features related to visual content such as shape, color, and texture are first extracted from a query image, the similarity between the set of features of the query image and that of each target image in a DB is then computed, and target images are next retrieved which are most similar to the query image. Extraction of good features which compactly represent a query image is one of the important tasks in CBIR. Shape is a visual feature that describes the contours of objects in an image, which are usually extracted from segmenting the image into meaningful regions or objects. However, since it is difficult to achieve such image segmentation for natural images, the use of shape features in image retrieval has been limited to special applications where the extraction of object contours is readily available such as in trademark images. Comprehensive and extensive literature survey on CBIR is presented in [1]-[4].

Swain et al. proposed the concept of color histogram in 1991 and also introduced the histogram intersection distance metric to measure the distance between the histograms of images [5]. Stricker et al. used the first three central moments called mean, standard deviation and skewness of each color for image retrieval [6]. Pass et al. introduced color coherence vector (CCV) [7]. CCV partitions the each histogram bin into two types, i.e., coherent, if it belongs to a large uniformly colored region or incoherent, if it does not. Huang et al. used a new color feature called color correlogram [8] which characterizes not only the color distributions of pixels, but also spatial correlation of pair of colors. Lu et al. proposed color feature based on vector quantized (VQ) index histograms in the discrete cosine transform (DCT) domain. They computed 12 histograms, four for each color component from 12 DCT-VQ index sequences [9].

Texture is another salient and indispensable feature for CBIR. Smith et al. used the mean and variance of the wavelet coefficients as texture features for CBIR [10]. Moghaddam et al. proposed the Gabor wavelet correlogram (GWC) for CBIR [11, 12]. Ahmadian et al. used the wavelet transform for texture classification [13]. Moghaddam et al. introduced new algorithm called wavelet correlogram (WC) [14]. Saadatmand et al. [15, 16] improved the performance of WC algorithm by optimizing the quantization thresholds using genetic algorithm (GA). Birgale et al. [17] and Subrahmanym et al. [18] combined the color (color histogram) and texture (wavelet transform) features for CBIR. Subrahmanym et al. proposed correlogram algorithm for image retrieval using wavelets and rotated wavelets (WC+RWC) [19].

Ojala et al. proposed the local binary patterns (LBP) for texture description [20] and these LBPs are converted to rotational invariant for texture classification [21]. Pietikainen et al. proposed the rotational invariant texture classification using feature distributions [22]. Ahonen et al. [23] and Zhao et al. [24] used the LBP operator facial expression analysis and recognition. Heikkila et al. proposed the background modeling and detection by using LBP [25]. Huang et al. proposed the extended LBP for shape localization [26]. Heikkila et al. used the LBP for interest region description [27]. Li et al. used the combination of Gabor filter and LBP for texture segmentation [28]. Zhang et al. proposed the local derivative pattern for face recognition [29]. They have considered LBP as a nondirectional first order local pattern, which are the binary results of the first-order derivative in images. Abdullah et al. [30] proposed fixed partitioning and salient points schemes for dividing an image into patches, in combination with low-level MPEG-7 visual descriptors to represent the patches with particular patterns. A clustering technique is applied to construct a compact representation by grouping similar patterns into a cluster codebook. The codebook will then be used to encode the patterns into visual keywords. In order to obtain high-level information about the relational context of an image, a correlogram is constructed from the spatial relations between visual keyword indices in an image. For classifying images k-nearest neighbors (k-NN) and a support vector machine (SVM) algorithm are used and compared.

Jhanwar et al. [31] have proposed the motif co-occurrence matrix (MCM) for content based image retrieval. The MCM is derived from the motif transformed image which is calculated by dividing the whole image into non-overlapping 2x2 pixel patterns. They also proposed the color MCM which is...
calculated by applying MCM on individual red (R), green (G), and blue (B) color planes. C H Lin et al. [32] combined the color feature, k-mean color histogram (CHKM) and texture features, motif co-occurrence matrix (MCM) and difference between the pixels of a scan pattern (DBPSP). MCM is the conventional pattern co-occurrence matrix that calculates the probability of the occurrence of same motif between each motif and its adjacent ones in each motif transformed image, and this probability is considered as the attribute of the image. According to the sequence of motifs of scan patterns, DBPSP calculates the difference between pixels and converts it into the probability of occurrence on the entire image. Each pixel color in an image is then replaced by one color in the common color palette that is most similar to color so as to classify all pixels in image into k-cluster, called the CHKM feature.

To improve the retrieval performance in terms of retrieval accuracy, in this paper, we combined color (color histogram on H and S spaces) and texture (motif co-occurrence matrix on V space) feature on HSV image. Two experiments have been carried out on Corel database and MIT VisTex database for proving the worth of our algorithm. The results after investigation show a significant improvement in terms of their evaluation measures as compared to RGB-Motif spaces.

The organization of the paper as follows: In section 1, a brief review of image retrieval and related work is given. Section 2, presents a concise review of local binary patterns. Section 3, presents the proposed system framework and similarity measure. Experimental results and discussions are given in section 4. Based on above work conclusions are derived in section 5.

2. MOTIF CO-OCURRENCE MATRIX

In this paper, motif co-occurrence matrix (MCM) is used to represent the traversal of adjacent pixel color difference in an image. As each pixel corresponds to four adjacent pixel colors, each image can be presented by four images of motifs of scan pattern, which can be further constructed into four two-dimensional matrices of the image size. Based on these four matrices, the attribute of the image will then be computed with the motifs of scan pattern and a color motif co-occurrence matrix (CMCM) can be obtained, which is the feature used in this paper.

For each pixel G(x, y), a 3x3 convolution mask can be generated as shown in Fig. 1. This 3x3 convolution mask can be further divided into four blocks of 2x2 grids (pixels) with each including pixel G(x, y).

In general, there are 25 different scan patterns in a grid if the traversal goes from four angles. Here, we only consider the scan starting from the top left corner pixel p1 as shown in Fig. 1(a), because it represents a complete family of space filling curve, reducing the number of patterns to only 7 as shown in Fig. 2. Among them motif number 0 signifies the situation wherein a motif cannot be formed due to the equivalence.

These 2x2 grids are then replaced by motifs of scan pattern which would traverse the grid in an optimal sense. The optimality of the scan is related to the incremental difference in intensity along the scan line minimizing the variation in the intensities in a local neighborhood.

Let G(x, y): N×N → Z be the gray levels of an N×N image I for Z = {0, 1, ..., 255}. Pixel G(x, y) is divided into four blocks, each of which has 2x2 grids (pixels), to form four different motifs shown by number m of motif. These four motifs will be saved in four N×N two-dimensional motifs of scan pattern matrix P[x, y], in which i=1, 2, 3, 4, x=1, 2, ..., Nx, y=1, 2, ..., Ny, and P[i, x, y]: N×N → W denotes an N×N matrix for W= {0, 1, ..., 6}.

The motif co-occurrence matrix (CCM) calculates the distribution within the two-dimensional motifs of scan pattern matrix P[x, y]. That is, it takes into account the probability of the co-occurrence between the two motifs respectively corresponding to (x, y) and its adjacent (x + dx, y + dy). This probability is then the attribute of image color variation used in this paper. The coordinate that distances from (x, y) on the x axis in dx and on y axis in dy, then the total number of co-occurring motifs of scan pattern pairs (u, v) (where u=0, 1, ..., 6 and v=0, 1, ..., 6) is determined by

\[
M_i(u, v) = M_i(u, v)[dx, dy] = M_i(P[i, x, y], P[x + dx, y + dy])
\]

where \(P[i, x, y] = u, P[x + dx, y + dy] = v, 1 \leq u \leq 4, 1 \leq x \leq Nx, 1 \leq y \leq Ny, 1 \leq x + dx \leq Nx, \) and \(1 \leq y + dy \leq Ny)\.

3. PROPOSED SYSTEM FRAMEWORK

In this paper, we proposed the new technique by using color and texture features for image retrieval. First the RGB (red, green, and blue) image is converted into HSV (hue, saturation, and value) image, then the H and S images are used for histogram calculation by quantizing into Q levels and the local region of V (value) image is represented by motif patterns, which are evaluated by taking into consideration of local difference between the pixel. Further, co-occurrence matrix is constructed between the motifs. Finally, feature vector is constructed by concatenating the features collected on HSV spaces and then these are used for image retrieval.
The flowchart of the proposed system is shown in Fig. 3 and algorithm for the same is given below:

**Algorithm:**

**Input:** Image; **Output:** Retrieval results.

1. Load the input image.
2. Convert RGB image into HSV image.
3. Perform the quantization on H and S spaces.
4. Calculate the histograms.
5. Calculate the motif co-occurrence matrix on V space.
6. Form the feature vector by concatenating the HSV features.
7. Calculate the best matches using Eq. (2).
8. Retrieve the number of top matches.

### 3.1. Similarity Measurement

In the presented work four types of similarity distance metric are used as shown below:

#### 3.1.1. Manhattan or L1 or city-block Distance

This distance function is computationally less expensive than Euclidean distance because only the absolute differences in each feature are considered. This distance is sometimes called the city block distance or L1 distance and defined as

$$D(Q,T) = \sum_i |f_i(Q) - f_i(T)|$$  \hspace{1cm} (2)

#### 3.1.2. Euclidean or L2 Distance

For \(p=2\) in the equation (1.1) give the Euclidean distance and defined as:

$$D(Q,T) = \left( \sum_i [f_i(Q) - f_i(T)]^2 \right)^{1/2}$$  \hspace{1cm} (3)

The most expensive operation is the computation of square root.

#### 3.1.3. \(D_1\) Distance

$$D(Q,T) = \sum_{i=1}^{L_g} \frac{f_{T,i} - f_{Q,i}}{1 + f_{T,i} + f_{Q,i}}$$  \hspace{1cm} (4)

#### 3.1.4. Canberra Distance

$$D(Q,T) = \sum_{i=1}^{L_g} \frac{f_{T,i} - f_{Q,i}}{|f_{T,i} + f_{Q,i}|}$$  \hspace{1cm} (5)

where \(Q\) is query image, \(L_g\) is feature vector length, \(T\) is image in database; \(f_{T,i}\) is \(i^{th}\) feature of image \(T\) in the database, \(f_{Q,i}\) is \(i^{th}\) feature of query image \(Q\).
4. EXPERIMENTAL RESULTS AND DISCUSSIONS

For the work reported in this paper, retrieval tests are conducted on two different databases (Corel 1000, and MIT VisTex) and results are presented separately.

4.1. Database DB1

Corel database [33] contains large amount of images of various contents ranging from animals and outdoor sports to natural images. These images are pre-classified into different categories of size 100 by domain professionals. Some researchers think that Corel database meets all the requirements to evaluate an image retrieval system, because of its large size and heterogeneous content. In this paper, we collected the database DB1 contains 1000 images of 10 different categories (groups G). Ten categories are provided in the database namely Africans, beaches, buildings, buses, dinosaurs, elephants, flowers, horses, mountains and food. Each category has 100 images \( (G^N = 100) \) and these have either \( 256 \times 384 \) or \( 384 \times 256 \) sizes. Fig. 4 depicts the sample images of Corel 1000 image database (one image from each category).

The performance of the proposed method is measured in terms of average precision and average recall by Eq. (6) and (7) respectively.

\[
\text{Precision} = \frac{\text{No. of Relevant Images Retrieved}}{\text{Total No. of Images Retrieved}} \times 100 \quad (6)
\]

\[
\text{Recall} = \frac{\text{No. of Relevant Images Retrieved}}{\text{Total No. of Relevant Images in Database}} \times 100 \quad (7)
\]

Table I and II summarizes the retrieval results of the proposed method (HSV-Motif) and RGB-Motif in terms of average retrieval precision and recall respectively. From Table I and Table II, and Fig. 5 (a)&(b), it is clear that the proposed method showing better performance compared to RGB-Motif [31] in terms of average retrieval precision and recall. Table III and Fig. 5 (c) illustrate the performance of proposed method with different distance measures. From Table III and Fig. 5 (c), it is found that the City Block distance showing better performance as compared to other distance measures. Fig. 6 illustrates the query results of proposed method on Corel 1000 database.

<table>
<thead>
<tr>
<th>Category</th>
<th>Precision (n=10)</th>
<th>Recall (n=100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africans</td>
<td>60.00</td>
<td>36.19</td>
</tr>
<tr>
<td>Beaches</td>
<td>54.70</td>
<td>33.32</td>
</tr>
<tr>
<td>Buildings</td>
<td>49.10</td>
<td>22.55</td>
</tr>
<tr>
<td>Buses</td>
<td>89.20</td>
<td>50.24</td>
</tr>
<tr>
<td>Dinosaurs</td>
<td>93.70</td>
<td>81.39</td>
</tr>
<tr>
<td>Elephants</td>
<td>46.10</td>
<td>22.29</td>
</tr>
<tr>
<td>Flowers</td>
<td>86.50</td>
<td>50.50</td>
</tr>
<tr>
<td>Horses</td>
<td>71.60</td>
<td>39.15</td>
</tr>
<tr>
<td>Mountains</td>
<td>38.30</td>
<td>23.56</td>
</tr>
<tr>
<td>Food</td>
<td>58.30</td>
<td>29.72</td>
</tr>
<tr>
<td>Total</td>
<td>64.75</td>
<td>38.89</td>
</tr>
</tbody>
</table>

n—No. top matches considered

Table I Results Of All Techniques In Terms Of Precision (%) On DB1 Database

<table>
<thead>
<tr>
<th>Distance</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>City Block (L1)</td>
<td>77.29</td>
<td>71.13</td>
</tr>
<tr>
<td>Euclidian (L2)</td>
<td>50.023</td>
<td>45.259</td>
</tr>
<tr>
<td>Canberra</td>
<td>75.71</td>
<td>45.543</td>
</tr>
<tr>
<td>d1</td>
<td>75.82</td>
<td>45.682</td>
</tr>
</tbody>
</table>

Fig. 4: Sample images from Corel 1000 (one image per category)
Table IV: Results of proposed method using different distance measures on DB2 database

<table>
<thead>
<tr>
<th>Distance</th>
<th>ARR(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>City Block (L1)</td>
<td>85.0</td>
</tr>
<tr>
<td>Euclidian (L2)</td>
<td>78.37</td>
</tr>
<tr>
<td>Canberra</td>
<td>83.34</td>
</tr>
<tr>
<td>$d_1$</td>
<td>84.75</td>
</tr>
</tbody>
</table>

Fig. 5: Comparison of proposed method (HSV-Motif) with RGB-Motif in terms: (a) average retrieval precision, (b) average retrieval rate.

Fig. 6: Query Results of proposed method on Corel database

Fig. 7. Comparison of proposed method (CLBP) with LBP in terms of: (a) average retrieval precision, (b) average retrieval rate.

Fig. 8: Query Results of proposed method on MIT VisTex database
4.2. Database DB2
The database DB2 used in our experiment consists of 40 different textures [35]. The size of each texture is 512×512. Each, 512×512 image is divided into sixteen 128×128 non-overlapping sub-images, thus creating a database of 640 (40×16) images. The performance of the proposed method is measured in terms of average retrieval rate (ARR) is given by Eq. (8).

\[
ARR = \frac{\text{No. of Relevant Images Retrieved}}{\text{Total No. of Relevant Images in Database}} \times 100 \quad (8)
\]

The database DB2 is used to compare the performance of the proposed method (HSV-Motif) with RGB-Motif. Fig. 7 (a) illustrates the retrieval results of proposed method (HSV-Motif) with RGB-Motif in terms of average retrieval rate. From Fig. 7 (a), it is evident that the proposed method is outperforming the RGB-Motif.

The results of the proposed method are also compared with the different distance measures as shown in Table IV and Fig. 7 (b). From Table IV and Fig. 7 (b), it is found that the City Block distance is outperforming the other distances.

5. CONCLUSIONS AND FUTURE SCOPE
A new image indexing and retrieval algorithm is proposed in this paper by combining color histogram features on H and S spaces and motif co-occurrence matrix on V space. Two experiments have been carried out on Corel database and MIT VisTex database for proving the worth of our algorithm. The results after being investigated shows a significant improvement in terms of their evaluation measures as compared to RGB-Motif.

In this paper, we combine the color and texture features for image retrieval. Further, the retrieval results can be improved by cinderering shape along with color and texture.

6. ACKNOWLEDGMENTS


7. ABOUT AUTHORS

K. SATYA PRASAD received his Ph.D degree from IIT Madras, India. He is presently working as professor in the Department of Electronics and Communication Engineering, JNTU college of Engineering Kakinada and Rector of Jawaharlal Nehru Technological University, Kakinada, India. He has more than 31 years of teaching and 20 years of research experience. He published 30 research papers in international and 20 research papers in National journals. He guided 8 Ph.D theses and 20 Ph.D theses are under his guidance. He authored Electronic Devices and Circuits text book. He held different ositions in his carrer like Head of the Department, Vice Principal, Principal for JNTU Engg College. His area of interests includes Digital Signal and Image Processing, Communications, Ad-hoc networks etc.