ABSTRACT
A two-level probabilistic content based image retrieval scheme for colored images is proposed. In the first level, the shape features computed through Color Angular Radial Transform (CART) are exploited for identifying the relevant images through the use of a Bayesian Framework. The features used in the second level are derived from Completed Robust Local Binary Pattern (CR-LBP) features. CR-LBP features are modified to capture texture information from colored images and are quantized for the computation of four-directional co-occurrence matrices to define a new color based texture descriptor. These new features (referred to as CR-LBP-Co) outperform their initial definition (implemented on colored images) in terms of retrieval accuracy. Further, they are computed only for a small set of images that have been identified to be closest to the query in the first level. The second level feature set comparison is made through a suitable similarity measure. When used in this probabilistic framework, the proposed CR-LBP-Co together with CART features achieve 51.96% retrieval accuracy in terms of Average Retrieval Rate (ARR) for Wang’s database.

KEY WORDS
Co-occurrence Matrix; Completed Robust LBP; Bayesian Sets; Angular Radial Transform; Average Retrieval Rate.

1. INTRODUCTION
Content Based Image Retrieval (CBIR) has gained a lot of popularity in the recent years. The number of digital images and videos has grown tremendously on the internet and are the impetus to strive for efficient retrieval results. Normally, the objective of CBIR system is the retrieval of relevant images from an image database whenever a query image is submitted to it. But, there are few interesting works [1], which describe how the user’s knowledge about the kind of images he/she wants to retrieve from the database can be utilized for effective image retrieval. This set of images facilitates training of the system which is then tested on a different test data. The work of Heller and Ghahramani [1] motivates us to use the Bayesian framework for two-level image retrieval. Another aspect of this paper is to use a set of features which describe the shape, texture and color information of images in the best possible manner. For this, we use the existing color ART (CART) features for encoding the shape information of colored images. We however propose to modify the Completed Robust Linear Binary Pattern (CR-LBP) features for colored images and use their four directional co-occurrence matrices as color based texture descriptor.

In the following sections, we describe our proposed two-level scheme in detail. Section 2 describes the existing and the proposed features. Section 3 explains the flow of the proposed scheme. In Section 4, we discuss the experimental results obtained from our system. Section 5 presents performance comparison of our CBIR system with the recent related works. Section 6 concludes the paper and states the future scope of this work.

2. EXISTING AND PROPOSED FEATURES
Angular Radial Transform (ART) is known to be MPEG-7 shape based descriptor. It is known to have an edge over other orthogonal moment based methods such as Zernike, pseudo-Zernike moments etc., because of its fast speed of computation and comparatively small feature size. Gray level co-occurrence matrix (GLCM) [2] is a well-known texture analysis method. The image properties are related to second-order statistics. The values in GLCM correspond to the number of occurrences of the gray level pairs which are at fixed distance from each other in the image. Haralick et. Al [2] proposed 14 statistical features to estimate the similarity between gray level co-occurrence matrices. A recent survey [3] on the usefulness of LBP features in recognition applications documents a number of LBP variants. Each of these variants has a property which can be utilized to suit the requirement of a specific application at hand. In one of the recent works by Zhao et. al[4], an improved and robust version of LBP i.e. CR-LBP has been proposed.

2.1 Color ART
Angular Radial Transform (ART) is a complex orthogonal unitary transform defined on a unit disk based on complex orthogonal sinusoidal basis functions in polar co-ordinates. The ART coefficients, $F_{nm}$ of order $n$ and $m$, are defined by:
\[ F_{nm} = \int_{0}^{2\pi} \int_{0}^{1} V_{n,m}^*(r,\theta) f(r,\theta) rdrd\theta \]  

where \( f(r,\theta) \) is an image intensity function in polar coordinates and \( V_{n,m}^*(r,\theta) \) is the ART basis function, which is complex conjugate of \( V_{n,m}(r,\theta) \) that is separable along the angular and radial directions as stated below

\[ V_{n,m}(r,\theta) = R_n(r)A_m(\theta) \]

with

\[ A_m(\theta) = \frac{1}{2\pi} e^{im\theta} \]

and

\[ R_n(r) = \begin{cases} 1 & (n = 0) \\ 2\cos(\pi nr) & (n > 0) \end{cases} \]

where \( n \) and \( m \) represent the order and repetition of ART, respectively. It has been observed in [5] that ART with outer circular disk implementation gives better result than inner circular disk. Therefore, in this paper, we use the outer circular disk.

Further as our images are colored, we compute Color ART features by converting an image from RGB to Lab color model. Luminance ('L') and chrominance ('a' & 'b') components are separated and 36 moments \( (n < 3, m < 12) \) for luminance component and \( 12 (n < 3, m < 4) \) each for chrominance components are computed [6]. Henceforth, we call these features as Color ART (CART) features.

### 2.2 Proposed Texture Descriptor

We propose a new texture descriptor that is based on completed robust local binary pattern (CR-LBP) [4]. This new descriptor is called completed robust local binary pattern co-occurrence (CR-LBP-Co) descriptor. It has been proved to possess higher discriminative power as compared to original LBP and CR-LBP. It is computed as follows:

1. First, the magnitude of RGB information of an image is computed as follows[7]:
   \[ \text{col} = \sqrt{R^2 + G^2 + B^2} \]  

2. We use the color information obtained in (5) to compute the CR-LBP features described in [4]. For this, we therefore utilize the color order relation described in [7].

3. We compute the pattern, magnitude and local central information of CR-LBP features using the color information computed in (5) as follows.

   3.1 The pattern is computed using (6)
   \[ \text{CR-LBP} = \sum_{p=0}^{P-1} s(\text{col}_p - \frac{G_c}{8 + \alpha})2^p \]
   where
   
   \[ G_c = \sum_{i=1}^{8} \text{col}_i + \alpha \text{col}_c, \quad s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \]

   Here, \( \text{col}_c \) represents the color value of the center pixel and \( \text{col}_i(i=1,\ldots,8) \) denotes the color value of the neighbor pixel of \( \text{col}_c \). \( P \) is the total number of neighbors. \( \text{col}_i \) is the color value of any neighbor of the center pixel on a circle of radius 1. \( \alpha \) is the parameter set to 8.

3.2 We further compute \( m_p \) as follows
   \[ m_p = \frac{1}{8+\alpha} |G_p - G_c| \]  

   where \( G_p = \sum_{i=1}^{8} \text{col}_{pi} + \alpha \text{col}_p \)

   Here, \( \text{col}_{pi}(i=1,\ldots,8) \) denotes the color value of the neighbor pixel of \( \text{col}_p \). \( G_p \) is computed as per (7).

3.3 The magnitude of these features is computed as
   \[ \text{CR-LBP}_{\text{mag}} = \sum_{p=0}^{P-1} s(m_p - c)2^p \]

   where \( c \) is the threshold set to the mean of \( m_p \) (computed in (8)) of the whole image. It measures the local variance of local color information.

3.4 The local central information is extracted through the following equation:
   \[ \text{CR-LBP}_{\text{cen}} = s(WG_c - c_i) \]

   Here, the threshold \( c_i \) is set as the average local color value of the whole image. \( G_c \) is computed as per (7).

4. These features are three dimensional and therefore are 3x256 in size. This size is not suitable for comparison purposes. Therefore, we perform uniform quantization of these features i.e. pattern, magnitude and central information features are quantized into 16 bins.

5. In the next step, we compute co-occurrence matrix of each of these quantized features in horizontal, vertical, diagonal and anti-diagonal directions. Thus, we generate \( 4 \times 3 = 12 \) co-occurrence matrices (i.e. \( \text{CR-LBP-C(i)}, \text{CR-LBP}_{\text{mag}}\text{-C(i)}, \text{CR-LBP}_{\text{cen}}\text{-C(i)}, \) for \( i=1,2,3,4 \)) out of these features. The dimension of each of these matrices is 16x16.

6. We then add these co-occurrence matrices (direction wise i.e. for \( i=1,2,3,4 \)) to finally obtain a set of four co-occurrence matrices (CR-LBP-Co). Since these matrices are symmetric about the diagonal, we safely fetch 136 non-redundant features from each of them.
and they constitute CR-LBP-Co descriptor of size 136x4. Thus, CR-LBP-Co can be represented as:

\[
\text{CR-LBP-Co}(i) = \text{CR-LBP-C}(i) + \text{CR-LBP}_{\text{mag}}(i) + \text{CR-LBP}_{\text{cen}}(i),
\]

where \( i = 1, 2, 3, 4 \)

(12)

Fig.1 depicts an example computation of CR-LBP features. After the computation of these features for whole image, their four directional co-occurrence matrices are computed to obtain CR-LBP-Co features using (12).

3. PROPOSED FRAMEWORK

In this section, we explain the steps involved in our Content Based Image Retrieval System. We describe it in two parts. First, we explain the probabilistic framework exploited in the first-level retrieval. This level utilizes CART features for extraction of related images from the test set. In the next level, the newly proposed CR-LBP-Co features are used. Fig.2 depicts the flow of the proposed framework. As can be observed from the figure, the main motivation of designing this framework is to reduce the search space for comparison of dense descriptor i.e. CR-LBP-Co.

3.1 Level-I Retrieval

Motivated by the results from Heller and Zoubin [1], we employ their Bayesian framework at the first level to identify a subset of relevant images from the entire search database. The probability distribution over features is computed for the queried image class using the training data set.

Then the algorithm uses a Bayesian criterion based on marginal likelihoods, to rank each image in the test space in the order of their relevance to given query. This is done as per the following expression:

\[
\text{score}(x^*) = \frac{p(x^*, D_q)}{p(x^*)p(D_q)},
\]

where \( D_q = \{x_1, x_2, \ldots, x_N\} \) are the training images corresponding to the query and \( x^* \) is the unlabelled image (belonging to test dataset) being scored. When the feature set is binary, computing the Bayesian score for all images in the test set gets reduced to a matrix-vector computation [1].

CART features are computed for all images in the database. These features are binarized following the preprocessing steps described in [1].

This turns the entire image dataset in to a sparse matrix, which is non-zero only for the features that distinguish the image from other images in the database. Bayesian score is computed for all the images using the above stated formula and only the top-k images (where \( k \) is half the size of test set) are used for retrieval in the second level. It is pertinent to mention here that the system requires the user to specify the type of images (using text query) he/she wants to retrieve from the database.

3.2 Level-II Retrieval

In the next level, we improve the retrieval accuracy by utilizing the proposed CR-LBP-Co features. These are now computed only for the database images retrieved in Level-I and the set of query images. Similarity measures like Euclidean distance and modified-Chi-Square distance are employed to find the top 10 matches. The modified Chi-square distance is computed as follows:

\[
\chi^2(x_i, x^*) = \frac{1}{n_f} \sum_{j=1}^{n_f} \left( \frac{x_i^j - \left( \frac{x_i^j + x^*_j}{2} \right)}{x_i^j + \left( \frac{x_i^j + x^*_j}{2} \right)} \right)^2
\]

(14)

Here, \( n_f \) is the number of features to be compared between query image and test image. The comparison is made direction wise for CR-LBP-Co features i.e. modified-Chi-square distance is computed between \( CR-LBP-Co(i), (i=1, 2, 3, 4) \) corresponding to each
direction $i$ and then added together to give the final distance between two images.

The proposed CBIR outperforms the individual performance of the existing and proposed descriptors through its effective retrieval mechanism wherein the first level retrieval is based on global shape features such as color ART features. The retrieval is then fine-tuned using the proposed CR-LBP-Co features in the second level of retrieval.

$$ARR = \frac{1}{N_q} \sum_{q=1}^{N_q} RR(q)$$

where $N_q$ represents the number of queries that are used for the purpose of verifying the descriptor in a dataset. $RR(q)$ represents the retrieval rate of a single query and is computed as per (16).

$$RR(q) = \frac{n_k}{n_q} \leq 1$$

where $n_k$ is the number of correct retrievals and $n_q$ is the total number of images (in the database) relevant to the query. Average retrieval precision (ARP) can also be computed by replacing the $n_q$ in (16) with number of retrieved images.

4.1 Retrieval Performance on Wang’s Color Dataset

Wang’s Color Dataset is a collection of 10 classes like Buildings, African people, Horses, Elephants, Food etc. Each class has 100 images. We first divide this dataset into labelled and unlabelled images. The labelled images serve as query whereas the unlabelled serve as test set. The user gives a textual query stating his/her images of interest e.g. a class name such as African people is specified for the desired retrieval. In this case, 90 images belonging to African people class act as query set and the remaining 10 from each class (i.e. total 100) comprise the test set.

Table 1 shows the experimental results of color LBP, color CR-LBP, CART and the proposed CR-LBP-Co descriptor. It is worth noting that the proposed descriptor outperforms the other features including CART in terms of ARR. These results have been obtained through the comparison of a query image’s features with all others in the database using modified Chi-square distance measure. CART features have been compared through Euclidean distance.

### Table 1. Classwise ARR of Proposed Descriptor on Wang’s Database

<table>
<thead>
<tr>
<th>Class Name</th>
<th>Existing Descriptor</th>
<th>Proposed Descriptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>African</td>
<td>33.38</td>
<td>41.00</td>
</tr>
<tr>
<td>Sea</td>
<td>28.00</td>
<td>34.63</td>
</tr>
<tr>
<td>Building</td>
<td>27.35</td>
<td>26.42</td>
</tr>
<tr>
<td>Bus</td>
<td>59.75</td>
<td>60.82</td>
</tr>
<tr>
<td>Dinosaur</td>
<td>81.42</td>
<td>91.41</td>
</tr>
<tr>
<td>Elephant</td>
<td>33.67</td>
<td>67.47</td>
</tr>
<tr>
<td>Flower</td>
<td>45.30</td>
<td>38.07</td>
</tr>
<tr>
<td>Horse</td>
<td>23.03</td>
<td>34.53</td>
</tr>
<tr>
<td>Mountain</td>
<td>23.58</td>
<td>26.29</td>
</tr>
<tr>
<td>Food</td>
<td>32.32</td>
<td>36.60</td>
</tr>
</tbody>
</table>

ARR | 38.78 | 44.91 | 32.29 | 45.72
The retrieval results in Table 2 justify the use of modified Chi-square similarity measure over the Euclidean distance with the proposed descriptor. It can be observed that approximately 10% accuracy improves with the choice of this similarity measure over Euclidean distance.

The retrieval results of the proposed CBIR system are better than the individual performance of CART and CR-LBP-Co features wherein Euclidean distance is employed for similarity comparison of CART features and modified Chi-square distance is used for comparison of CR-LBP-Co features. This is depicted through the bar chart in Fig.3. It is important to note that on an average (computed on 10 classes), the proposed scheme achieves 50.30% retrieval rate as compared to 32.29% and 45.72% retrieval rate achieved through CART and CR-LBP-Co features alone respectively. Note that in this case, last 90 images comprise training set and the remaining 10 are in the test set.

The retrieval results for a query stated as “Dinosaur” by the user are shown in Fig. 4. After first level retrieval through Bayesian framework, we obtain a set of 50 unlabelled images (based on their shape). In the second level, CR-LBP-Co features retrieve 10 closest results from these 50 images based on their color-texture information.

In order to further strengthen our claim, we change the 90 and 10 images of the training and test set respectively for 10 times. This always presents the system with new labelled dataset and test set corresponding to a user generated query. We then obtain an average precision for category searches for different recall values. Fig. 5 shows a plot of average retrieval precision for varying recall. For recall of 100%, we obtain an ARR of 51.96%.
5. COMPARISON WITH RECENT WORKS


We further give class-wise comparison of the performance of the proposed method with the results of Subrahmanyam et. al[9]. These results computed on Wang’s database are shown in Table 3. It can also be observed that the variance of our method’s retrieval accuracy is less than that of [9]. Further, Youssef [13] presents a CBIR system based on curvelet transform and dominant color extraction. This work reports ARR at 80% recall as 51% whereas 55.05% is achieved by our proposed method on Wang’s database at the same recall.

We compare our results with that of Liu and Yang[14] who propose the color difference histogram descriptor for capturing texture information from colored images. On the Corel 10k dataset, our proposed method achieves 57.80% average precision whereas the same is 45.24% with their method on the same database.

6. CONCLUSION AND FUTURE SCOPE

This paper proposes a new two-level probabilistic content based image retrieval system. It also proposes a new color based texture descriptor called CR-LBP-Co (CR-LBP co-occurrence features). These features prove to be better than the color counterpart of original CR-LBP by 0.81% in terms of ARR on Wang’s database. When used together with CART features in the Bayesian frame work, they appreciably improve the retrieval accuracy of the CBIR system. The efficacy of the proposed system has been demonstrated through a variety of experiments conducted on different datasets. Currently, a large part of the database has to be set aside for training of the system. Future work may target at finding an appropriate unsupervised method to cluster images based on similarity.
REFERENCES


