Image Retrieval Based On Color and Full Texton Matrix Histogram (C&Ftmh) Features

G. BinduMadhavi, V. Vijaya Kumar, K. Sasidhar

Abstract One of the challenging tasks of image processing is image retrieval. Image retrieval is performed by matching the features of the image. In the literature many researchers derived local texture features and produced diverse feature descriptors for efficient image retrieval. Most of the local features are derived based on the relationship between centre pixels versus with boundary or neighboring pixels over a local neighborhood. In this paper we propose a new local descriptor based on full texton index pattern using HSV color space. The HSV color space is used to derive color, intensity and brightness features in the form of histograms. This paper transforms the V-plane image into full texton index (FT) image. A co-occurrence matrix is derived on FT, image and this results full texton matrix (FTM). This paper derives the co-occurrence structural features in the form of histogram and color histogram features are concatenated to derive feature vector. This feature vector is named as "color and full texton matrix histograms" (C&Ftmh). The proposed C&Ftmh framework is tested on the five popular databases and results are compared with state-of-art methods with color features.

Index Terms: HSV, texton, Histograms, structural, color features.

I. INTRODUCTION

The content based image retrieval (CBIR) requires the knowledge of mathematics, computer science and image processing. Image matching is the basic principle behind image retrieval. And image matching in turn requires matching of the extracted image features. The image features contains color information, distribution of textural information, structural, statistical information etc. The image features may be derived globally, locally or on a region based. The existence of huge number of images in online and offline make retrieval of images attractive, exciting and significant.

Out of the various features, the texture and color are the major and significant features of the image. Texture is a surface property and its feature basically depends upon the local intensity levels. The statistical, local neighborhood, motif and texton features represent the most significant features of texture. The distribution of intensities in different color channel is represented as a color histogram and it is one of the crucial factors in CBIR. The gray level co-occurrence matrix (GLCM) is one of the popular methods for extracting statistical features of the image texture [1]. GLCM is also used for retrieval of color images [2]. The GLCM features also used for the retrieval of rock images. The GLCM is broadly used for different applications in texture classification, analysis, face recognition etc. [3,4,5,6].

The neighborhood features, the region based features, the local features derived on a 3x3 grid or on a 2x2 grid with GLCM features have attained high results in various image processing applications [7-9]. The enhanced versions of motif co-occurrence matrix are proposed in the literature for CBIR [5, 10] and they have exhibited high retrieval rate. Recently a new variant to motif is proposed in the literature for image retrieval [11].

The local binary pattern (LBP) proposed by Ojala et al.[12] has become popular local descriptor. The LBP and its variants achieved more accurate results than its counterparts [13, 14]. The LBP derives only circular or isotropic information. Recently the isotropic and anisotropic information on a neighborhood is derived using circular and extend LBP and its variants [15]. The CELBP and its variants achieved high classification results [16]. The rest of the paper is organized as follows. The section 2 outlines proposed method. The section 3 deals with experimental setup and results and discussions. The section 4 presents conclusions.

II. PROPOSED METHOD

Images can be broadly divided into three types. The first one is black and white image. The black and white images only contain two intensity levels i.e. black and white. The second type of images is known as gray level images. The gray level image can have range of intensities from 0 to 255 under one band. The third types of images are known as color images. The color images will have multiple bands for each pixel and contain a range of intensity for each band. The popular and the most generally used color model is RGB. The RGB has three color bands called red, green and blue. And the RGB used additive color mixing. The other type of color model is HSV color space where H, S and V stand for Hue, Saturation and Value respectively. HSV is often used by artists because color, intensity and brightness can be extracted individually in HSV color space than in terms of additive or subtractive color components as in the case of RGB and other color models. This paper constructed histograms for all three components of HSV individually. The histogram features extracts the global information. The textons represents the simple patterns on a micro grid of size 2 x 2 and they have played a key role in many image processing applications. The complex patterns of an image surface can be easily represented by textons. The texton co-occurrence matrix (TCM) [17] and multi-texton histogram (MTH) [18] were derived for CBIR in the
A texton is formed whenever two or more pixels of the grid have exactly similar identical gray level values. The TCM derives only five texton types on a 2x2 micro grid as shown in Fig. 1.

Fig. 1: The five types of textons used in TCM.

Fig. 2: Frame work of TCM approach.

The TCM identifies the texton pattern in an overlapped manner on a 2x2 grid, thus the TCM scans the texture image five times and in each scan the TCM identifies one of the texton patterns. The TCM forms the final texton image by fusing the five texton type images. If a texton type is identified then TCM approach places a zero in the pixel location of the 2x2 grid, which is not part of the texton pattern, otherwise the grey levels are unchanged. The MTH approach derived texton patterns with two identical pixels on a 2x2 grid as shown in Fig. 3 (T6 to T9). The MTH overcomes the fusing operation of TCM approach by dividing the image into micro grids of size 2x2. The texton patterns are identified in each 2x2 grid. The MTH approach retains the 2x2 grid without changing the pixel gray level intensities whenever a texton pattern of MTH is identified. The entire 2x2 grid is assigned a value zero if no texton type of MTH is identified.

Fig. 3: The four types of textons used on MTH.

The frame work of MTH approach is shown in Fig. 4 on a sub image of size 6x6.

Fig. 4: The texton detection mechanism of MTH.
the following disadvantages in TCM and MTH: The TCM ignored completely the patterns with two identical pixels and it requires a complex fusing operation for constructing the final texton image; The MTH approach ignored the texton patterns with three identical pixels and only defined few textons with two identical pixels and this has lead lot of ambiguity in identification of textons as shown in (Fig.5); The existing texton methods derived a framework mainly to assign a zero value to the 2 x 2 grids which are not forming any textons or those pixels which are not part of textons; Most of the texton based approaches have not replaced the 2x2 texton micro grid with texton indexes. The reason for this, they have derived only few number or partial number of texton patterns. Further, the replacement with texton indexes is not possible in TCM due to fusing operation. This paper addressed these disadvantages by deriving full textons and replaces the 2 x 2 grid with full texton indexes.

![Sample image patch.](image1)

<table>
<thead>
<tr>
<th>22</th>
<th>22</th>
<th>215</th>
<th>215</th>
</tr>
</thead>
<tbody>
<tr>
<td>W X</td>
<td>F F</td>
<td>F F</td>
<td>F F</td>
</tr>
<tr>
<td>Z Y</td>
<td>F F</td>
<td>A F</td>
<td>F A</td>
</tr>
<tr>
<td>F F</td>
<td>F B</td>
<td>F A</td>
<td>A F</td>
</tr>
<tr>
<td>A B</td>
<td>A F</td>
<td>B F</td>
<td>A B</td>
</tr>
<tr>
<td>6 7</td>
<td>8 9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>F F</td>
<td>A A</td>
<td>A A</td>
<td>A A</td>
</tr>
</tbody>
</table>

![FTi Image of (a).](image2)

<table>
<thead>
<tr>
<th>9</th>
<th>1</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 3 13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 14 11</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

![FTi Image of (a).](image3)

| 22 65 |
| 215 25 |

(a) Texton type T7 and T8 MTH.
(b) Texton type T6 and T9 MTH.

![Ambiguity in identifying texton types in MTH.](image4)

Fig. 5: Ambiguity in identifying texton types in MTH.

This paper derives full texton indexes that include all the patterns with four, three and two identical pixels on a micro grid of size 2x2. The V-component holds the gray scale information; therefore value component is used for extraction of local features in this paper. The proposed method transforms the V-plane of the texture image into a full texton index (FTi) image in the following way. The proposed FTi derives 14 texton shapes on the micro grid of size 2x2 (Fig.6) and completely overcomes the problem of ambiguity.

![Proposed Full textons patterns.](image5)

Fig.6: Proposed Full textons patterns.

The FTi framework initially divides the image into sub grids of size 2x2 (Fig.7). The sub grids are replaced by the corresponding FTi (Fig.7).

This paper quantized the gray level intensities of V-plane by transforming, it into an FTi image. The proposed FTi framework reduces the original image of size NxM into N/2 x M/2 and replaces each micro grid with FTi ranging from 0 to 14.

![The derivation of FTi image.](image6)

Fig. 7: The derivation of FTi image.

The GLCM is computed on FTi image and this derives a Full texton matrix (FTM). These features represents the texton (structural) and statistical information of the image texture. The co-occurrence matrix (CM) measures the occurrence of pixel pairs, located at a particular distance and specific direction. The size of CM depends on the size of gray levels. The values in the GLCM ranges from 0 to m where m is the maximum number of times a specific pixel pair appeared with a distance d and the rotation angle $\alpha$. The
process of computation of co-occurrence pixel pair is shown in Fig.8 (a) and 8(b). The Fig.8 (a) displays the original image grey level matrix. The Fig.8 (b) displays the GLCM of Fig.8(a) with distance 1 and \( \alpha = 0^\circ \). In Fig. 8(b) i.e. in GLCM the topmost row and leftmost column which are displayed in red color, indicates the grey level values of the original matrix 8(a) (0 to 4). For each pair (0,0), (0,1), (0,2), (0,3), (0,4),..(1,0),..(1,4),..(4,4),..(4,4) the occurrence has been calculated. For the pair (3,3) and (4,4) with a distance 1 and with an angle of rotation 0o, the occurrence pairs in original matrix are shown in red and blue colors respectively and such occurrences are reported four and three times respectively. Therefore in GLCM for the pixel pair (3,3) and (4,4) the values 5 and 6 are placed.

The FTM derives co-occurrence pairs of full texton image indexes and this derives the complete structural texture information. Finally the feature vector “C&FTMH” is generated by concatenating the histograms of HSV color space with histogram of FTM.

![Fig.8](a): Original matrix.
![Fig.8](b): GLCM for Fig.8(a).

This paper has deliberated a new method called “C&FTMH” for CBIR based on histograms of color, co-occurrences of full textons. The main contributions in the current paper are as follows:

1. A new feature vector called full texton image index is derived to overcome the ambiguity and fusing problems of earlier texton based methods.
2. The proposed method replaces the micro grids with full texton index which is missing in earlier methods of textons.
3. The derivation of co-occurrence matrix on FTM extracts the co-occurrence frequencies of powerful local structural features in the form of full textons located at a distance d.
4. Color information in the form of histograms is derived on Hue (H), Saturation (S) and Value (V) components of HSV.
5. The integration of color histogram feature with FTM histogram derives the strong and discriminate information of global and local features of texture “C&FTMH”.

### III. RESULTS AND DISCUSSIONS

To test the efficacy of the proposed C&FTMH descriptor, this paper used five bench mark color databases images: Corel-1k [19] and Corel-10K databases [20], MIT-VisTex database [21], CMU-PIE database [22] and Holidays dataset [23], and the description about this database are given in table 1 and sample imagers are displayed in Fig.9 to Fig. 13 respectively.

A. Similarity measure and query matching

This paper derived feature vector of all database images and query image. The similarity measure between query image versus the database images are computed using: Euclidean and Manhattan distance measure.

**Manhattan distance measure:**

\[
D(d_{bi},q) = \sum_{d=1}^{D} \left| F_{dbi}(l) - F_q(l) \right|
\]

**Euclidean distance measure:**

\[
D(d_{bi},q) = \sum_{d=1}^{D} \left( F_{dbi}(l) - F_q(l) \right)^2
\]

Where the distance between database image i (dbi) and query image q is measured by D(dbi,q). V represents the total number of features in the features vector. Fdbi(l) and Fq(l) are the feature vector ‘l’ of ith database image and query image respectively. The best relevant image will have the shortest distance.

B. Evaluation measures

The precision and recall are used as evaluation measures. In each experiment each
image of the database is used as query image and image retrieval is performed and retrieval performance is evaluated. And average performance is noted down. For each retrieval 20 most similar images are retrieved. The precision measures the ratio between the number of relevant images retrieved versus the total number of images retrieved (n) (Eqn.3). The recall measures the ratio between number of relevant images retrieved versus the total number of relevant images in the database (Nic) i.e. number of images in each category ‘c’ of the database for the query image (Eqn.3).

\[
P(i, n) = \frac{\text{Number of relevant images retrieved}}{n} \quad (3)
\]

\[
R(i, n) = \frac{\text{Number of relevant images retrieved}}{N_{ic}} \quad (4)
\]

Average precision and recall are given in equation 5 and 6.

\[
P_{\text{avg}}(i, n) = \frac{1}{N_c} \sum_{i=1}^{N_c} P(i, n) \quad (5)
\]

\[
R_{\text{avg}}(j, n) = \frac{1}{N_c} \sum_{i=1}^{N_c} R(i, n) \quad (6)
\]

Where, J denotes the number of categories. The total precision and total recall for the entire database are calculated as.

Table 2: APR values using Euclidean and Manhattan distances when number of top matches are 10 for proposed C&FTMH method.

<table>
<thead>
<tr>
<th>Distances/ Database</th>
<th>Corel-1k</th>
<th>Corel-10k</th>
<th>MIT-VisTex</th>
<th>Holidays</th>
<th>CMU-PIE</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manhattan</td>
<td>89.32</td>
<td>45.68</td>
<td>86.15</td>
<td>85.64</td>
<td>90.21</td>
<td></td>
</tr>
<tr>
<td>Euclidean</td>
<td>85.36</td>
<td>41.20</td>
<td>83.23</td>
<td>83.25</td>
<td></td>
<td>88.32</td>
</tr>
</tbody>
</table>

Table 3: ARR values using Euclidean and Manhattan distances when number of top matches are 10 for proposed C&FTMH method.

<table>
<thead>
<tr>
<th>Distances/ Database</th>
<th>Corel-1k</th>
<th>Corel-10k</th>
<th>MIT-VisTex</th>
<th>Holidays</th>
<th>CMU-PIE</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manhattan</td>
<td>0.201</td>
<td>0.125</td>
<td>0.201</td>
<td>0.189</td>
<td>0.221</td>
<td></td>
</tr>
<tr>
<td>Euclidean</td>
<td>0.200</td>
<td>0.118</td>
<td>0.207</td>
<td>0.182</td>
<td>0.215</td>
<td></td>
</tr>
</tbody>
</table>

For validation the proposed C&FTMH method is compared with the existing methods by concatenating their features with color histograms. This work concatenated RGB color histogram of 18 bins for each band i.e. 24 bins with the feature vectors of LBP[12], CS-LBP [24], LDP[25], LTP[26], BLK-LBP[27], TCM [17] and MTH [18] separately. The results are plotted and explained in terms of APR and ARR on each affordable database.

This paper retrieved for each query image different number of images from the database and performance is computed using the performance measures APR and ARR. The Fig. 15, 16, 17, 18 and 19 shows the performance of the proposed C&FTMH method versus the existing methods using similarity measures APR and ARR on the above databases. The precision and recall graphs clearly indicated the high performance of the proposed C&FTMH method over the existing methods and the research concatenated the color histograms with the features of the existing methods to maintain the comparison levels. From these results it is noted that the presented frame work of this paper is more advantageous in terms of average precision, recall and accuracy than the existing methods.

C. Results and discussions

This paper measured the capabilities of the proposed C&FTMH method on an aforementioned five color database using the above performance measures and graphs are plotted. The present C&FTMH method derived features from color and full texton co-occurrence features of texture. This paper measured the APR and ARR on the derived feature vector using the Manhattan and Euclidean distance similarity measures on each database and listed out in Table 2 and Table 3. The Manhattan similarity measure, shown good results when compared to Euclidean distance measure. In rest of the paper Manhattan distance measure is used.
Fig. 9: Corel-1K database sample images.

Fig. 10: Corel-10K database sample images.

Fig. 11: The sample textures from Holidays database.
Fig. 12: The sample textures from MIT-VisTex texture.

Fig. 13: The sample facial images from CMU-PIE database.
The top 20 retrieved images for one of the query image from each database is displayed in Fig. 14.

(a) Corel 1k.
Fig. 14: Top 20 retrieved images of considered databases.
Fig. 15: Comparison of over Corel-1k database using (a) APR (b) ARR.
Fig.16: Comparison over Corel-10k database using (a) APR (b) ARR.

(a) APR.

(b) ARR.

Fig.16: Comparison over Corel-10k database using (a) APR (b) ARR.
Fig. 17: Comparison over Holidays database using (a) APR (b) ARR.

(A) APR.

(B) ARR.

Fig. 17: Comparison over Holidays database using (a) APR (b) ARR.
Image Retrieval Based On Color and Full Texton Matrix Histogram (C&FTMH) Features

Fig. 18: Comparison over MIT-VisTex database using (a) APR (b) ARR.

(A) APR.

(B) ARR.

Fig. 18: Comparison over MIT-VisTex database using (a) APR (b) ARR.
Fig. 19: Comparison over CMU-PIE database using (a) APR (b) ARR.

(A) APR.

(B) ARR.

Fig. 19: Comparison over CMU-PIE database using (a) APR (b) ARR.
The proposed C&FTMH frame work utilizes the global features in the form of individual H,S,V color histograms and local features using FTM. This paper replaces each 2x2 grid of the V-plane with the full texton index and derives a co-occurrence matrix. The HSV color space is used to extract color, brightness and value features respectively. The full texton derives all possible texton patterns on a 2x2 grid and it overcomes the fusing problems associated with TCM and ambiguity issues arises in MTH. The TCM and MTH approaches have not replaced the 2x2 grids with texton indexes due to the limited number of textons. Further, the derivation of GLCM on FT indexes derive the strong and discriminate information of global and local features of texture and this has made the proposed C&FTMH descriptor as superior than existing methods with color features.

REFERENCES

7. B. Vamsee Mohan, V. Vijaya Kumar, Face Recognition Using Robs Fuzzy Based Fundamental Texture Matrix (RFFTM), Jour of Adv Research in Dynamical & Control Systems, Vol. 10, 15-Special Iss 2018
8. B. Kishore , V. Vijaya Kumar, Local Texton Centre Symmetric Patte Matrix (LTCPM) On Wavelet Domain for Texture Classification International Journal of Innovative Technology and Explori Engineering (IJITEE), ISSN: 2278-3075, Volume-8 Issue:- December, 2018
18. Guang-Hai Liu a, Lei Zhang b, Ying-Kun Hou d, Zuo-Yong Li c, Jing-Yu Yang, Image retrieval based on multi-texton-histogram, Pattern Recognition, 43(2010),2380–2389
19. Corel Photo Collection Color Image Database, online available on http://wang.ist.psu.edu/docs/related.shtml
24. and weak geometric consistency for large scale image search.

AUTHORS PROFILE

Mrs. G. Bindu Madhavi is working as an Assistant Professor in Anurag Group of Institutions. She has 8 years of experience in teaching. She completed her M.Tech in Bioinformatics Engineering College, INTU. She taught Data Mining, Computer Organization, Digital Logic Design, Object Oriented Programming courses. She guided 15 projects at B.Tech and M. Tech level. She organized and participated in many workshops and Faculty Development Programs. She published 10 papers in various national and international journals. Her areas of interest are Image Processing, Networking and Data mining. She is member of ACM and IEEE.

Dr. V. Vijaya Kumar is working as Professor & Dean Department of Computer Science & Engineering and Information Technology in Anurag Group of Institutions (Autonomous) Hyderabad and Director-Center for Advanced Computational Research (CACR). He received integrated M.S.Engg, in CSE from USSR in 1989. He received his Ph.D. from Jawaharlal Nehru Technological University (JNTU), Hyderabad, India in 1998 in CSE and guided 33 research scholars for Ph.D. He acted as principle investigator for various R&D projects.
He has served JNT University for 13 years as Assistant Professor, Associate Professor and Professor. He has received the Distinguished Professor award from Computer Science of India (CSI), Mumbai, best researcher and best teacher award from JNT University, Kakinada, India, Leading Scientist of the WORLD - 2009 and Top 100 Scientists award in 2010 from International Biographical Centre, Cambridge, England. His research interests include Big data and image analytics, image retrieval, texture analysis, author attribution, digital water marking. At present he is also acting as BoS member for various universities and institutions. He is the life member of CSI, ISCA, ISTE, IE (I), IETE, ACCS, CRSI, IRS and REDCROSS. He published more than 120 research publications till now in various national, international journals and conferences. He has delivered key note addresses at various international conferences.

Dr. K.Sasidhar has 20.8 years of teaching & research experience in Computer Science Engineering courses. His research interests are in the areas of Biometrics, Data Mining and Big Data. He received best teacher award for thrice. Dr. Sasidhar published and presented 20 research papers in highly reputed journals, conferences at international and national level. He is also recognized as a co supervisor for research scholars. He guided more than 60 Projects at UG and PG level. He is a life member in ISTE and ASCI. He is also external BOS member for Electronics and Computer Engineering department. He delivered around 16 invited talks, guest lecturers for computer science students at various engineering colleges and AICTE sponsored FDP programs. He organized and participated in many workshops, conferences and Faculty Development Programs.