A Novel Technique for Fast Content-Based Image Retrieval using Dual-Cross Patterns.

Bhavana Kakde, Manish Okade
Department of Electronics & Communication Engineering,
National Institute of Technology, Rourkela, 769008, India.
Email: bhavanakakde06@gmail.com, okadem@nitrkl.ac.in

Abstract– **This proposed work introduces a novel technique for Fast Content-Based Image Retrieval (CBIR) using Dual-Cross Patterns (DCP). DCP encodes second order information in the vertical, horizontal and diagonal direction, by performing the encoding of sample points in the local surrounding region of every center pixel in an image. The local binary pattern (LBP), an efficient visual texture descriptor, performs a comparison between the center pixel and its neighboring pixel. Local tetra patterns (LTrP) is a method which acquires more detailed information by using four possible directions of every center pixel in an image, and is calculated from first-order derivatives in horizontal and vertical directions. To analyze its effectiveness, the proposed technique, LBP, and LTrP are compared. Our extensive simulation on Corel database shows that the proposed technique has better time complexity compared to these methods.**

Index terms– CBIR, query image, local sampling, pattern recognition, dual-cross patterns(DCP), pattern encoding, local tetra patterns (LTrP).

I. Introduction

Content-Based Image Retrieval (CBIR) has a useful role in image retrieval framework. It is a commonly received solution for an efficient and effective method that can look up for the required image from the large database without human interaction. There occurs a need of CBIR because the development digital images, due to widespread capturing of images using web cameras, mobile phones enable with the camera, and digital cameras is rendering the management of image database tedious. It can also be used in other application such as web engines and social media which stores a large number of images and requires fast retrieval of the image selected by the user. The extraction of a feature in CBIR is a noticeable step whose viability is dependent upon the strategy used for feature extraction from given images. These features can be arranged in classes like as histogram, spatial layout, shape, texture, color, etc. The CBIR uses these features to retrieve and index the image database. The objective is to find a unique representation for all different variation because the user captures images in various conditions such as occlusion and varying illumination etc. An all-inclusive and considerable literature study on CBIR is given in [1]–[4].

The texture is well appropriated for the recognition of items such as marble, slabs, parquet, ceramic tiles, etc., and has drawn wide consideration from texture retrieval industries and recherches. Texture analysis has potential in extracting the outstanding features due to which it is being used in computer vision technology and pattern recognition application. The concept of wavelet correlogram is introduced in Moghaddam et al [5], [6]. Discrete wavelet transform (DWT) [7] have been used for texture classification. DWT generalized Gaussian density plus KullbackLeibler distance, an application of DWT, has given efficient results for retrieval of the texture image [8] and image segmentation [9]. But, DWT can extricate data from horizontal direction, diagonal direction, and vertical direction, in an image. A method using local tetra pattern (LTrP) retrieval algorithm is used for direction limitation. The local tetra patterns (LTrP) method is able to acquire more detailed information by encoding the relationship between the center pixel and its surrounding pixels, with four different values based on the directions that are estimated using the first-order derivatives in horizontal and vertical directions [10]. Various extensions of the LBP [11], completed LBPs [12], dominant LBPs [13], etc., also introduced for rotational invariant and multi resolution gray scale texture classification and other descriptors, LBPs [14] are used for CBIR [15], [16].

This paper shows a novel technique for fast CBIR using Dual-Cross Patterns (DCP) [17]. DCP is a novel descriptor, which captures useful information by encoding second order discriminative data in the vertical, horizontal and diagonal image components. Sampling strategy adapted by DCP extracts twice the pixel information as compared to LBP. The sampled pixels are grouped in accordance with maximum joint Shannon entropy, thereby keeping a reasonable feature vector size. A sub- DCP cross encoder (denoted herein as E-1 and E-2) having exactly the same memory cost and time complexity as LBP, gives a significantly better performance. The DCP is computationally efficient.

II. Review of Dual-Cross Patterns

Image filtering, local sampling, and pattern encoding are the three principle parts in the formation of a DCP image descriptor. The key of DCP is to carry out pattern encoding and local sampling in the most descriptive necessary direction, and this encrypts second-order statistical information.

The sampled points are encoded in two stages. First, encoding of textural details along each of the eight directions is done, which is followed by the combining of the patterns

978-1-5386-4273-3/18/$31.00 2018 IEEE

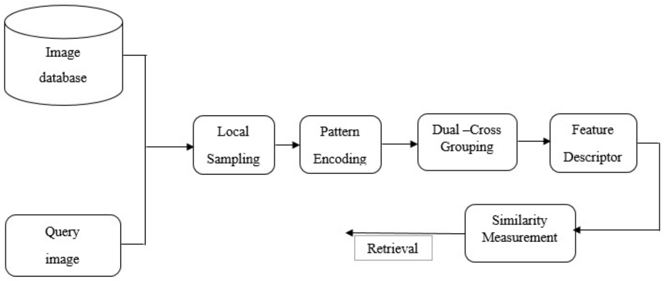


Fig. 1: Framework of proposed technique.

obtained to form the DCP codes. A grouping strategy is used in DCP to reduce the computational complexity but due to the information lost, grouping is done according to joint Shannon entropy to minimize it.

III. PROPOSED TECHNIQUE

Fig. 1 shows the framework of the technique presented in this paper and an algorithm is explained below.

Algorithm:

I/P: Query image; O/P: Retrieved images

1) Take an image, and change its color model to grayscale model.
2) Perform symmetrical sampling in eight directions in the local neighborhood for each pixel of the image.
3) In each sampling direction, to quantize the textural information pattern encoding is used.
4) Dual-Cross Grouping is performed to minimize information loss.
5) Two cross encoder generates the encoded images and derive the feature vector for each images.
6) Combine both the feature vector.
7) The images in given database are compared with the query image using (7).
8) Images are retrieved based on the best matches.

The detailed description of each stage is given below:

A. local sampling

Dual-cross pattern encrypts higher order useful information in the direction of image components; vertical, horizontal, and diagonal directions of local neighborhood. In Fig.2, a local sampling of the Dual cross pattern is shown. For each center pixel O ,symmetrically sample the points in eight directions (0, $\pi/4$, $\pi/2$, $3\pi/4$, π, $5\pi/4$, $3\pi/2$, and $7\pi/4$) in the local neighborhood. In each direction two pixels gets sampled and the final sampled pixels are represented as $\{A_0, B_0; A_1, B_1; ...; A_7, B_7\}$. As shown in Fig. 2, on the inner circle having radius R_{in} points $A_0, A_1, ..., A_7$ are uniformly distributed, while $B_0, B_1, ..., B_7$ points are evenly spaced on the outer circle having radius R_{ex}.

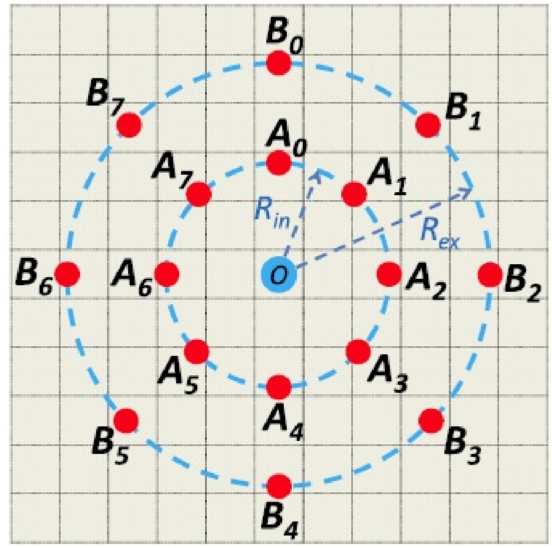


Fig. 2: Local sampling of DCP. Sixteen pixels are sampled around the local neighborhood central pixel O.

B. Pattern Encoding

Two steps are used to encode the sampled points. First, encoding of textural details along each of the eight directions is done, which is followed by the combining of the patterns obtained to form the DCP codes. A unique decimal number is assigned in each sampling direction to quantize the textural information:

$$Dcp_k = F(A_{A_k} - A_O) \times 2 + F(A_{B_k} - A_{A_k}), 0 \leq k \leq 7 \quad (1)$$

where

$$F(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (2)$$

and A_0, A_{A_k}, and A_{B_k} gray level value of center pixel O, and neighboring pixels A_k, and B_k, respectively. Accordingly, second-order statistics along each direction are encoded by the four patterns and each of the four patterns signifies one kind of textural information.

When all eight directions are considered simultaneously the total number of possible dual cross pattern code is $4^8 = 65{,}536$. This value is too high for real-time recognition application, so to reduce computational complexity, eight-direction are divided into two group, and each group called one encoder. By this manner, the accumulative number of local patterns is minimized to $4^4 \times 2 = 512$, which helps in decreasing the computational complexity. Although this technique results in data loss, robustness and compactness of the descriptor are promoted.

C. Dual-Cross Grouping

The total thirty-five combinations are produced on partitioning eight directions by grouping technique explained in the earlier subsection. Joint Shanon entropy criteria is used for optimal grouping of eight directions to preserve the necessary information required for image retrieval. With the above analysis, Dcp_k $(0 \leq k \leq 7)$ can take one of the four possible values: 0, 1, 2, and 3. The joint Shannon entropy for the set $\{Dcp_0, Dcp_1, Dcp_2, Dcp_3\}$ is represented as

$$H(Dcp_0, Dcp_1, Dcp_2, Dcp_3) = -\sum_{dcp_0} ... \sum_{dcp_3} P(dcp_0, ..., dcp_3) \log_2 P(dcp_0, ..., dcp_3) \quad (3)$$

where dcp_0, dcp_1, dcp_2, and dcp_3 are particular values of Dcp_0, Dcp_1, Dcp_2, and Dcp_3 respectively. $P(dcp_0, ..., dcp_3)$ is the probability of dcp_0, dcp_1, dcp_2, and dcp_3 values occurring simultaneously. For four variables, joint Shannon entropy is maximum when they have statistical independence.

In images when the pixels are more sparsely scattered they are more independent of each other.Hence, when the sample points are separated by a maximum distance; maximum joint Shannon entropy in each subgroup is obtained. As a result $\{Dcp_0, Dcp_2, Dcp_4, Dcp_6\}$ forms the first subset and $\{Dcp_1, Dcp_3, Dcp_5, Dcp_7\}$ forms the second subset.

D. DCP Feature Descriptor

These two sub Dcp encoders are called E-1 and E-2, respectively. For each pixel *O*, code generated by sub DCP encoder E-1:

$$E - 1 = \sum_{k=0}^{3} Dcp_{2k} \times 4^k, \quad (4)$$

and code generated by the second sub DCP encoder E-2:

$$E - 2 = \sum_{k=0}^{3} Dcp_{2k+1} \times 4^k, \quad (5)$$

By combining the codes generated by the two sub DCP encoder E-1 and E-2; final DCP descriptor for pixel *O* is given as:

$$DCP = \left\{ \sum_{k=0}^{3} Dcp_{2k} \times 4^k, \sum_{k=0}^{3} Dcp_{2k+1} \times 4^k \right\}. \quad (6)$$

Using the sub DCP encoders, two coded images are generated that are divided into nonoverlapping regions and in each individual region, histogram is computed. To generate the holistic representation all the histograms are combined. By using the similarity metrics such as query matching, a similarity can be measured using this holistic representation between a pair of images.

IV. Query Matching

F_q is the feature vector obtained from feature extraction for the given query image q, and it is represented as$F_q = \left(F_{q_1}, F_{q_2}, ..., F_{q_{l_g}}\right)$. Similarly for each image in the database, a feature vector using DCP is derived, and the overall feature vector set F_D of the database $|D|$ is represented as: $F_D = \left(F_{D_{i1}}, F_{D_{i2}}, ..., F_{D_{ilg}}\right)$, $i = 1, 2, ..., |D|$. The objective is to get the *m* top images from the database that have similar categories as that of query image. This is carried out by selecting *m* top-matched images by estimating similarity distance metric d_1 between images in Corel database $|D|$ with the query image using (7).

Fig. 3: Sample images from database DB.

$$d_1(q, D) = \sum_{j=1}^{l_g} \left| \frac{F_{D_{ij}} - F_{q_j}}{1 + F_{D_{ij}} + F_{q_j}} \right| \quad (7)$$

where $F_{D_{ij}}$ is the *j*th feature of the *i*th image in database $|D|$.

V. Experimental Results

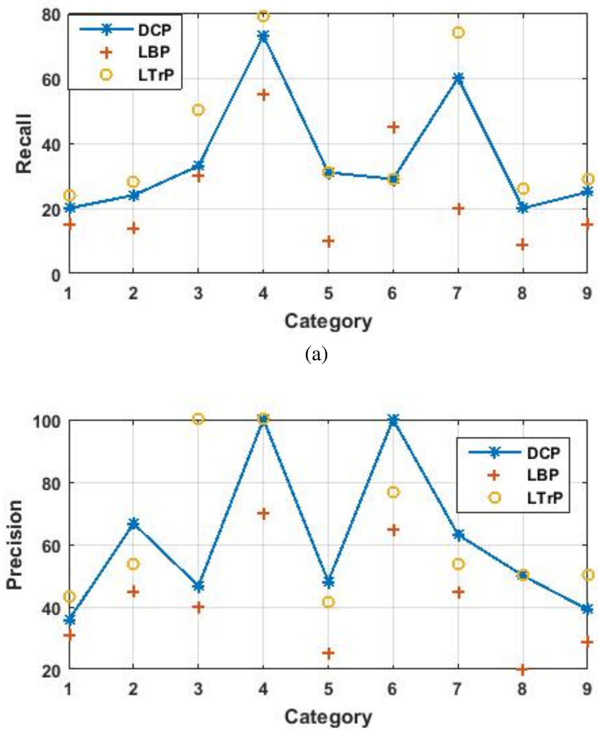
A. Experiment

Simulation is performed in Matlab 8.6 to evaluate the performance of retrieval framework using DCP, and Corel database is used for this purpose. It comprises of huge number of different images, contents varying from natural scenes to animals. The database is categorized into multiple categories by domain professionals, each category have 100 images. As it has a collection of diverse images and huge size, researchers have the view that it fulfills all the requirement to assess an image retrieval framework. This experiment used 900 images from Corel database D, and images are collected from nine distinctive domain, namely *Buildings, Elephants, Mountains, Flowers, Buses, Beaches, Food*, and *Horses*. N_g=100 number of images in each category, with a resolution of either 384×256 or 256×384. Fig.3 shows the samples images from different category of database D. In the simulation work each image is used as a query image from the database D. Framework should retrieve those m images $Y = (y_1, y_2, ..., y_m)$ from the database, which has small and sufficient matching distance from the query image, evaluated using (7). Framework retrieval accuracy depends upon the category of the query image and retrieval image y_i=1,2,...,*m* if the category is same then expected image is suitably recognized by it otherwise it has failed to do so.

The simulation performance of presented technique is evaluated in terms of average recall and average precision. The precision for query image Aq is given as follows:

$$Pr(A_q, m) = \frac{1}{m} \sum_{j=1}^{|D|} \left| \theta\left(\psi\left(A_j\right), \psi\left(A_q\right)\right) \right| Rank\left(A_j, A_q\right) \leq m \quad (8)$$

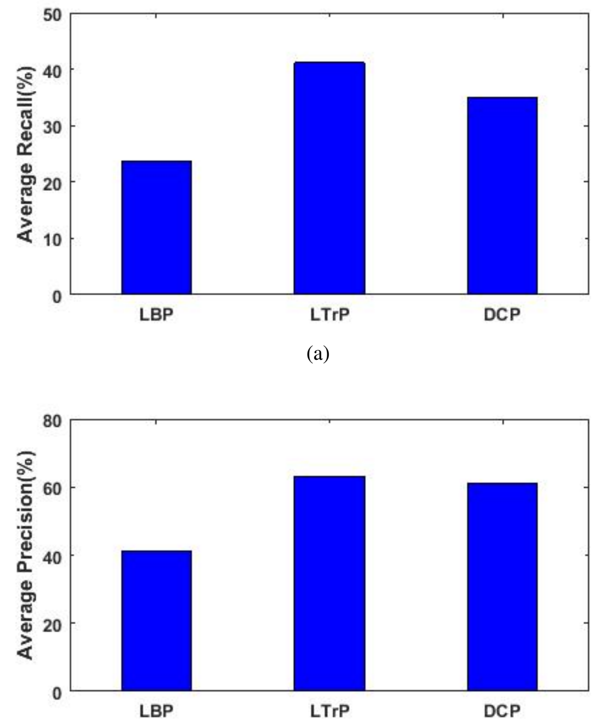
where $\psi(y)$ is the category of "y", "m" represents the number of retrieved images, $|D|$ is the total number of images in the



(a)

(b)

Fig. 4: Performance comparison of the DCP with LTrP, and LBP methods on database D, Recall vs. Category (a). Precision vs. Category (b).



(a)

(b)

Fig. 5: Performance comparison of the DCP with LTrP, and LBP techniques on database D. Average recall % (a). Average precision % (b).

database and *Rank*(A_i, A_q) returns the rank of query image A_i among all images of database $|D|$, and

$$\theta\left(\psi\left(A_j\right), \psi\left(A_q\right)\right) = \begin{cases} 1, & \phi(A_j) = \psi(A_q) \\ 0, & else. \end{cases} \tag{9}$$

Recall is given as

$$Re(A_q, m) = \frac{1}{N_g} \sum_{j=1}^{|D|} \left|\theta\left(\psi\left(A_j\right), \psi\left(A_q\right)\right)\right| Rank\left(A_j, A_q\right) \leq m \tag{10}$$

The average precision of the *i*th similar category of the reference image database is given by

$$Pr^i_{ave}(m) = \frac{1}{N_g} \sum_{j \in N_g} Pr(A_i, m) \tag{11}$$

In the similar manner, average recall can be defined.

B. Result

For Corel database, Fig. 4 shows category wise recall and precision performance for the DCP, LTrP, and LBP. Recall and precision values are the useful parameter for quantitative performance measurement of CBIR technique. Performance is supposed to be good if these parameters have high value. Fig. 4 depicts, DCP has better precision for five categories (2,4,5,6 and 7) out of nine, compared to LTrP and outperforms LBP in all categories in terms of both recall and precision.

Fig. 5 shows the performance of the DCP, LTrP, and LBP techniques which is estimated in terms of average recall percentage and average precision percentage on the Corel database. The performance of proposed technique gives better result compared to LBP technique and approximately similar to the LTrP.

TABLE I: Time Complexity

Method	Time for one image
LBP	377.94. sec
LTrP	1926.48 sec
DCP	34.75 sec

Table I shows time complexity of LBP, LTrP and DCP methods as we can see DCP required very less time compared to other two method i.e. DCP has good time complexity.

VI. CONCLUSION

The technique presented in this paper, referred to as DCP for Fast CBIR.The essence of dcp is to carry out pattern encoding and local sampling in the most descriptive direction within the image. DCP encrypts second order information using two cross encoders.

The significant improvement of the proposed technique is in time complexity as compared to LTrP and LBP methods is shown in Table 1. The average recall for DCP is **35%** and the average precision for DCP is **61.05%** which is similar to LTrP and an improvement upon LBP.

As proposed technique has encouraging results, it can be applied to other pattern classification applications like fingerprint recognition, bioinformatics, etc. In the future ,we will try to implement our proposed technique with combination of other descriptors.

REFERENCES

[1] Y. Rui, T. S. Huang, and S.-F. Chang, "Image retrieval: Current techniques, promising directions, and open issues," *Journal of visual communication and image representation*, vol. 10, no. 1, pp. 39–62, 1999.

[2] A. W. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 12, pp. 1349–1380, 2000.

[3] M. Kokare, B. Chatterji, and P. Biswas, "A survey on current content based image retrieval methods," *IETE Journal of Research*, vol. 48, no. 3-4, pp. 261–271, 2002.

[4] Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma, "A survey of content-based image retrieval with high-level semantics," *Pattern recognition*, vol. 40, no. 1, pp. 262–282, 2007.

[5] H. A. Moghaddam, T. T. Khajoie, and A. H. Rouhi, "A new algorithm for image indexing and retrieval using wavelet correlogram," in *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, vol. 3. IEEE, 2003, pp. III–497.

[6] H. A. Moghaddam and M. Saadatmand-Tarzjan, "Gabor wavelet correlogram algorithm for image indexing and retrieval," in *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, vol. 2. IEEE, 2006, pp. 925–928.

[7] A. Ahmadian and A. Mostafa, "An efficient texture classification algorithm using gabor wavelet," in *Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE*, vol. 1. IEEE, 2003, pp. 930–933.

[8] M. N. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized gaussian density and kullback-leibler distance," *IEEE transactions on image processing*, vol. 11, no. 2, pp. 146–158, 2002.

[9] A. Laine and J. Fan, "Texture classification by wavelet packet signatures," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 15, no. 11, pp. 1186–1191, 1993.

[10] S. Murala, R. Maheshwari, and R. Balasubramanian, "Local tetra patterns: a new feature descriptor for content-based image retrieval," *IEEE Transactions on Image Processing*, vol. 21, no. 5, pp. 2874–2886, 2012.

[11] Z. Guo, L. Zhang, and D. Zhang, "Rotation invariant texture classification using lbp variance (lbpv) with global matching," *Pattern recognition*, vol. 43, no. 3, pp. 706–719, 2010.

[12] ——, "A completed modeling of local binary pattern operator for texture classification," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1657–1663, 2010.

[13] S. Liao, M. W. Law, and A. C. Chung, "Dominant local binary patterns for texture classification," *IEEE transactions on image processing*, vol. 18, no. 5, pp. 1107–1118, 2009.

[14] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern recognition*, vol. 29, no. 1, pp. 51–59, 1996.

[15] M. Pietikäinen, T. Ojala, and Z. Xu, "Rotation-invariant texture classification using feature distributions," *Pattern Recognition*, vol. 33, no. 1, pp. 43–52, 2000.

[16] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

[17] C. Ding, J. Choi, D. Tao, and L. S. Davis, "Multi-directional multi-level dual-cross patterns for robust face recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 3, pp. 518–531, 2016.

978-1-5386-4273-3/18/$31.00 2018 IEEE