

Local Tri-directional median differential excitation co-occurrence pattern (LTriDMDECoP): A new feature descriptor for content based image retrieval

¹G V Satya Kumar, ²P G Krishna Mohan

¹Department of Electronics and Communication Engineering, JNT University, Hyderabad, India

Email: gvsatyakumar@gmail.com

² Department of Electronics and Communication Engineering, IARE, JNT University, Hyderabad, India

Abstract:

A novel feature descriptor called local tri directional median differential excitation co-occurrence pattern (LTriDMDECoP) for content based image retrieval is proposed in this paper. The LTriDMDECoP exploits the relationship between the focused or center pixel with its neighboring pixels using differential excitation instead of merely taking advantage of the gray level intensity difference which is sensitive to noise. Further, LTriDMDECoP considers a unique sampling strategy for computation of differential excitation. The proposed method considers median intensity of pixels in three directions to establish the relation between focused pixels with its neighbours using differential excitation. Further, co-occurrences of differential excitation values in local pattern map have been observed in different directions to accomplish fortified feature extraction. The performance of proposed feature descriptor has been tested for image retrieval on Corel-1000 and Corel-5000 bench mark databases. The experimental results demonstrate that LTriDMDECoP outperforms the well known LTCOP and CSLBCoP methods in terms of average precision and recall.

Keywords: Co-occurrence statistics, differential excitation, image feature extraction, image retrieval, local binary pattern, pattern recognition.

1.Introduction

Recent developments in image acquisition technology has resulted in accumulation of large volume of digital images both online and offline which in turn has led to gigantic image databases. Handling of these hefty image databases are extremely annoying rather impractical task. Many different types of retrieval methods have been proposed over the past decade for retrieving information from these massive databases based on a query image. Some of the earlier methods used text based image retrieval, which suffers from some serious drawbacks such as inappropriate metadata, lack of adequate information necessary for describing the image. This raises the need for new techniques which are able to describe the image based on the content of the image. This led to the emergence of content based image retrieval.

1.1 Motivation

The advancements in multimedia technology led to the accumulation of large image repositories requiring huge digital storage space. Instead of storing the gigantic image database, storing their extracted features in a feature database is more convenient and saves lot of memory. The features extracted from images are used to represent and index the database. Thus, the features extracted from images based on the content of an image play an important role in content based image retrieval (CBIR) system. Features can be of different types based on their visual content such as color, texture, shape and domain specific features such as finger print and human faces. Therefore, the feature extraction method or feature descriptor should be designed in such a way that it is able to retrieve information from images taken under different illumination and intensity conditions and able to capture the local microstructures present in an image. Apart from that, it is able to represent the local structural differences.

Many methods have been developed for efficient image retrieval in the recent past. However, most of these patterns have focused on the sign of gray level intensity difference between the center pixel and its neighboring pixels over a 3×3 patch. Most of the existing methods used only one

neighboring pixel at a time and ignored the original intensity stimulus in pattern map computation. In order to overcome some drawbacks of above mentioned methods, later works focused on the need of having separate sign and magnitude patterns to improve the discriminative power. Most of the local patterns did not consider the mutual relationship among the neighboring pixels for a given center pixel or focused pixel which resulted in comparatively weak performance.

This paper proposes a novel feature descriptor for image retrieval task. The striking advantage of the proposed method over other methods is that the proposed method computes the mutual relation between the focused or center pixel with its neighbours using differential excitation with unique sampling strategy. The proposed method also computes the co-occurrences of pattern mapped values in different directions for better feature extraction.

The remainder of the paper is systemized in the following manner. Section-2 discuss about the related work on content based image retrieval, main contributions of the paper and some popular local patterns concepts incorporated in the proposed method. The proposed feature extraction and similarity measurement is presented in section-3. Section- 4 presents the experimental results and discussion. Section-5 concludes the paper.

2. Related Works

A number of local patterns have been developed for content based image retrieval in the past decade. The popular local binary pattern (LBP) texture descriptor was introduced by Ojala et al [1]. The LBP finds application in various domains such as texture classification [2], face recognition [3], object tracking [4] etc. The authentic LBP method generates an eight bit binary string by comparing the center pixel with its neighboring pixels lying on the circumference of a circle of radius r and using a single threshold. The weakness of LBP to represent anisotropic structures due to circular sampling prompted the researchers to work extending the operator called as dominant local binary pattern (DLBP)[5], center symmetric local binary pattern (CSLBP)[6], Block based local binary pattern (BLK_LBP)[7], and completed local binary pattern (CLBP)[8]. Tan and Triggs [9] proposed a three value encoding strategy called local ternary pattern (LTP) to overcome the noise sensitivity of LBP.

Zhang et al [10] proposed a new descriptor based on the first order derivative of LBP called local derivative pattern (LDP) for face recognition application. Later, Murala et al [11] devised local tetra pattern (LTrPs) to represent spatial structure of image textures in different directions and also incorporated magnitude pattern into feature extraction. Xie et al [12] fused local patterns of Gabor magnitude and phase for the purpose of face recognition. Chen et al [13] devised robust local binary patterns (RLBP) considering both sign and magnitude information to reduce noise of LBP feature. Later, several feature descriptors using LBP concept are devised by various researchers in the recent past. Among these, some important implementations are Local mesh pattern (LMP)[14], Ellipse topology[15], local bit plane decoded pattern (LBDP)[16], average local binary pattern (ALBP)[17] etc. are worth mentioning.

A feature descriptor considering the intensity variations based on oriented neighbourhood called geometric local textural pattern (GLTP) presented in Orjuela Vargas et al[18]. Manisha et al [6] proposed a feature descriptor exploiting the co-occurrence of binary pattern values called center symmetric local binary co-occurrence pattern (CSLBCoP). Murala et al[19] proposed a method based on ternary edge co-occurrences called local ternary co-occurrence pattern (LTCOP) that combines the effectiveness of LBP and LDP.

2.1 Main contributions

The prime contributions of proposed work may be stated in brief as:

- First, the proposed work extracts the relationship between the center or focused pixel with its neighbor pixels using a unique sampling strategy on a 3×3 window.

- The relationship between focused pixel and its neighbors according to new sampling strategy is exploited using differential excitation instead of gray level intensity difference (like in LBP) as it is more robust to noise.
- Second, the differential excitation values are encoded to binary values using local binary pattern. Later, the co-occurrences of pattern mapped values are computed in different directions for enriched feature extraction to uplift the discriminative power of the descriptor.

2.2 Local binary patterns

The local binary pattern (LBP) was originally introduced by Ojala et al [1] for texture classification due to ease of computation and less complexity with high discriminative power. Later, LBP use was extended to different application domains like object tracking [4], facial expression recognition [3], and medical imaging. The LBP descriptor considers a small window of an image and computes the intensity difference between the center pixel and its P neighbours lying on the circumference of a circle of radius R. Later, the gray level intensity differences are encoded into binary form using a predefined threshold. These binary bits are summed up to produce a decimal value using specific weights. For a given image the local binary map of the image is generated by replacing each focused or center pixel with its binary pattern value. The feature vector is formed by computing the histogram of the local binary pattern values. The formulation of LBP for a given center or focused pixel x_c with P neighbour pixels at a radius R is as follows

$$LBP_{P,R} = \sum_{m=1}^P 2^{(m-1)} \times S(x_m - x_c) \quad (1)$$

$$S(x) = \begin{cases} 1, & x \geq 0 \\ 0, & \text{else} \end{cases} \quad (2)$$

x_m is the gray level intensity of its neighbours. Later, several feature descriptors based on LBP are presented in [1].

2.3 Local ternary pattern (LTP)

A three valued encoding scheme called local ternary pattern (LTP) is introduced by Tan and Triggs [9] as an extension to LBP overcoming the deficiencies associated with it. LTP encodes the gray level difference values into three zones. The zone of width $\pm T$ around the center pixel x_c to zero, those above $(x_c + T)$ are encoded to +1, and those below $(x_c - T)$ are encoded to -1. The mathematical formulation of LTP is as follows

$$LTP_{P,R} = \sum_{m=1}^P 2^{(m-1)} \times S_1(x_m - x_c) \quad (3)$$

$$S_1(x) = \begin{cases} 1, & x \geq +T \\ 0, & |x| < T \\ -1, & x \leq -T \end{cases} \quad \text{where } x = x_m - x_c \quad (4)$$

2.4 Differential excitation (η)

The differential excitation (η) is defined as “a function of ratio of gray level difference of focused pixel with its neighboring pixel to the focused pixel” by Chen et al [20]. Mathematically, η has been defined for a given focused pixel x_c with P neighbourhood pixels as

$$\eta(x) = \tau \left[\sum_{m=1}^{p-1} \frac{(x_m - x_c)}{x_c} \right] \quad (5)$$

Where τ is transformation function, x_m is neighbour pixel. The ratio term in the equation (5) $\frac{(x_m - x_c)}{x_c}$ is called Weber ratio. In the proposed case the arctan is used as transformation function τ . The arctan function restricts the Weber ratio values to vary from -1.5 to +1.5. Further, if x_c is zero, the Weber ratio will be undefined. To overcome this situation, x_c is assigned a small non zero positive value. More particulars about differential excitation and Weber ratio can be found in [20].

3. Proposed work

3.1 Local tri directional median differential excitation co- occurrence pattern (LTriDMDECoP)

The proposed feature extraction is inspired by the fact that the human perception of pattern depends not only on the change of gray scale intensity values but also on the original stimulus intensity. Most of the LBP variants obtain the local information structure based on gray level intensity difference and ignores the original stimulus intensity. The proposed LTriDMDECoP feature descriptor extracts the mutual relationship between the focused pixels with their neighbours using differential excitation which accounts for original stimulus intensity in an innovative way. The differential excitation (η) values are more robust to noise compared to gray level difference.

The LTriDMDECoP take advantage of all the neighboring pixels in the radius1 (8 neighbours) to accomplish the relation between the focused pixel and its neighbours on a 3×3 window unlike CSLBCoP which uses only the center symmetric pixels as shown in Figure1. Apart from that LTriDMDECoP engages the affiliation of focused pixel with its neighbourhood pixels using a novel sampling strategy as explained in Figure2.

The LTriDMDECoP computation is accomplished in two steps. First, the differential excitation values of focused pixels are obtained according to the predefined sampling strategy at radius1. This resulted in four differential excitation values for a given 3×3 window. Later, the four differential excitation values thus obtained are binary encoded using pre defined threshold and they are summed up to decimal values using specific weights. The obtained pattern mapped values are ranged from 0 to 15. In second step, the co-occurrences of pattern mapped values are obtained in different directions using gray level co-occurrence matrix (GLCM) which resulted in a co-occurrence matrix of size 16×16 . For feature extraction, each pixel in the image is considered as focused pixel and the co- occurrence of pattern mapped pairs are obtained in four directions $0^\circ, 45^\circ, 90^\circ, \text{ and } 135^\circ$ to obtain enhanced information about the image. The obtained histograms of the co-occurrence matrices in four directions are concatenated to form the feature vector. Hence $4 \times 16 \times 16 = 1024$ is the feature vector length.

The Figure 2a shows the sub-image for computation of pattern with Z as focused or center pixel with its eight neighbors $a, b, c, \dots, h$ at radius 1. The sampling strategy for the computation of differential excitation matrix is as shown in Figure 2b and 2c. Three neighbour pixels in three different directions are considered at a time for computation of four differential excitation values as shown in Figure 2c

The computation of LTriDMDECoP is explained as follows

- First, obtain the eight neighbourhood pixel candidates for a given center or focused pixel using a 3×3 window

- Now, compute the median intensity values by considering three neighbours in three directions at time as shown in Figure 2c indicated as m_1, m_2, m_3 , and m_4

$$m_1 = \text{median}(a, b, c) \quad (6)$$

Similarly, $m_2 = \text{median}(c, d, e)$, $m_3 = \text{median}(e, f, g)$, and $m_4 = \text{median}(a, h, g)$

- Therefore, the differential excitation values are given by

$$\eta_i = \arctan\left[\frac{m_i}{Z}\right], \quad \forall i = 1, 2, 3, 4 \quad (7)$$

- Now, differential excitation matrix $\eta = [\eta_1, \eta_2, \eta_3, \eta_4]$ is binary encoded and multiplied with specific weights to get the decimal equivalent pattern values. The pattern values are ranged from 0 to 15, total 16 different intensity values.
- Later, each pixel in the given image is treated as focused pixel and pattern mapped image is obtained
- For improved feature extraction, the co-occurrence of pattern mapped pixel values are computed in four directions $0^\circ, 45^\circ, 90^\circ$, and 135° using GLCM. Each direction results in a co-occurrence matrix having 16×16 size.
- Now, compute histograms of co-occurrence matrices in each direction and concatenate them to form the feature vector. Therefore, $4 \times 16 \times 16 = 1024$ is the length of feature vector. The proposed feature descriptor flow chart is as shown in Figure3.

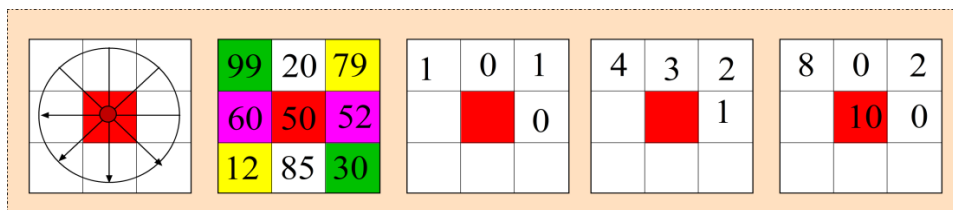


Figure1: Example computation of Center symmetric local binary pattern (CSLBP)

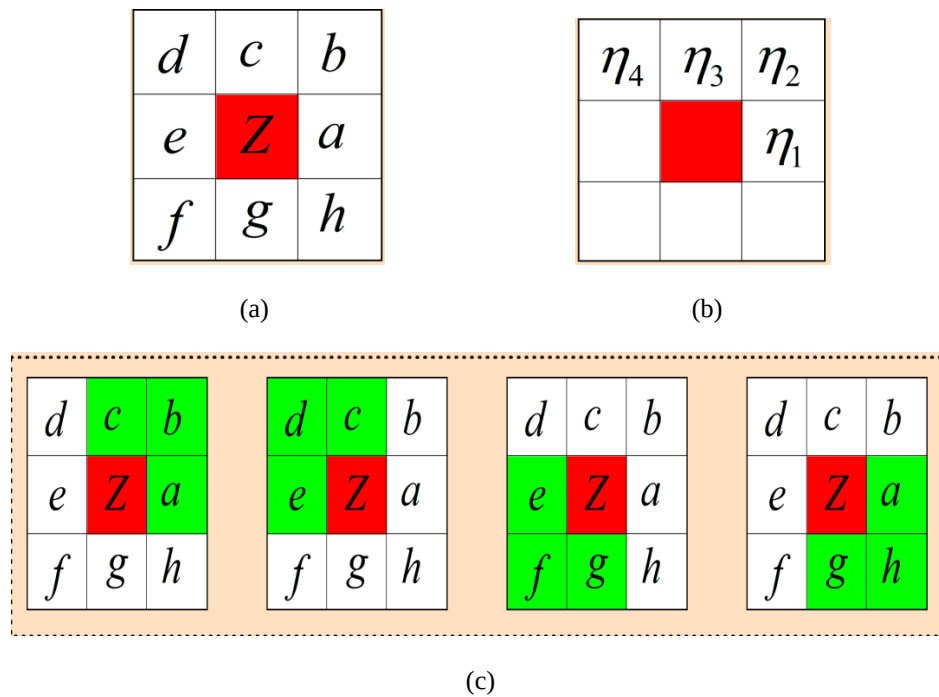


Figure2: a) Sub image b) Differential excitation matrix

c) Consideration of pixel candidates for computation of differential excitation

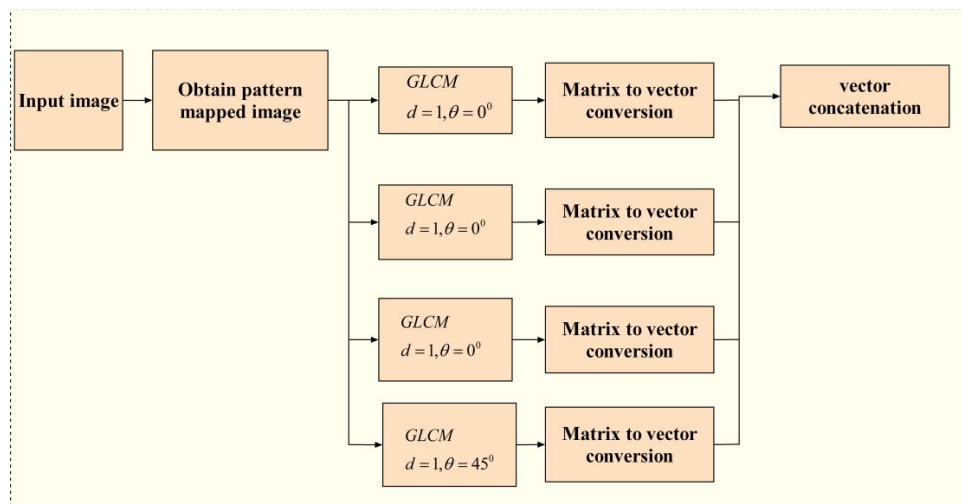


Figure3: The proposed feature descriptor flow chart

3.2 Similarity measurement

The feature vector of the query image is represented by $f_Q = (f_1, f_2, f_3, \dots, f_L)$ where L is feature vector length, 1×512 Vector. The number of images in the database is given by N and the

database feature vectors are represented as $f_{db} = (f_{db1}, f_{db2}, \dots, f_{dbN})$. The distance between the query image features to each image features in the database is measured and top n matches for the given query image are retrieved. For the similarity measure, d_1 distance metric is used as given below

$$d(Q, db) = \sum_{i=1}^L \left| \frac{f_{dbi} - f_Q}{1 + f_{dbi} + f_Q} \right| \quad (8)$$

Where f_Q the feature vector of query image is, f_{dbi} is the feature vector of i^{th} image in the database. The distance function is given by $d(Q, db)$

4. Experimental results

The performance of proposed method is evaluated on two benchmark databases followed by brief description about experimental setup. The Corel-1000[21] and Corel-5000[21] database comprises images of various contents ranging from human, animals and outdoor sports to natural images. The database is pre classified into different categories of size 100 by the domain professionals. Corel-1000 and Corel-5000 databases consist of 10 and 50 different categories respectively with each category having 100 images. The heterogeneous content of Corel-1000 and Corel-5000 make it ideal for evaluating the image retrieval systems.

The performance of the LTriDMDECoP is examined on two benchmark databases in terms of average precision rate (APR) and average recall rate (ARR). If N is the number of images present in the database and n is the number of top matches considered then

$$\text{Precision} - P_i(n) = \frac{\text{Number of relevant images retrieved}(R_v)}{\text{Total number of images retrieved}(R_r)} \quad (9)$$

$$\text{Recall} - R_i(n) = \frac{\text{Number of relevant images retrieved}(R_v)}{\text{Total number of images present in the database}(N)} \quad (10)$$

$$\text{Average Precision-} \quad AP_n(j) = \frac{1}{N_1} \sum_{i=1}^{N_1} P_i(n) \quad (11)$$

Where $AP_n(j)$ is the average precision of j^{th} category, N_1 - No. of images present in the corresponding category.

Average recall $AR_n(j)$ of the j^{th} category with N_1 number of images is given by

$$AR_n(j) = \frac{1}{N_1} \sum_{i=1}^{N_1} R_i(n) \quad (12)$$

The average precision rate (APR) and average recall rate (ARR) for a database with N_2 categories for n top matches is given below

$$APR = \frac{1}{N_2} \sum_{i=1}^{N_2} AP_n(i) \quad (13)$$

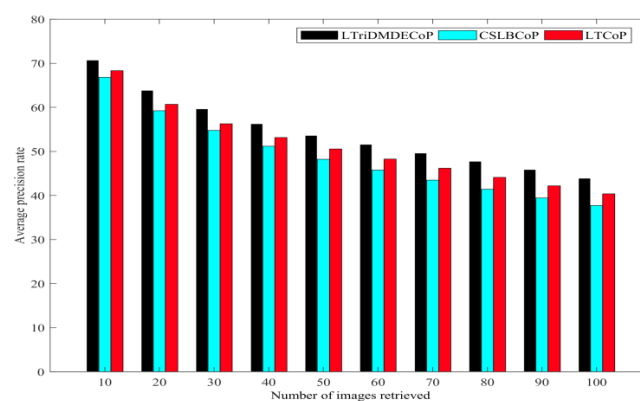
$$ARR = \frac{1}{N_2} \sum_{i=1}^{N_2} AR_n(i) \quad (14)$$

4.1 Experiment#1-Corel-1000 image database

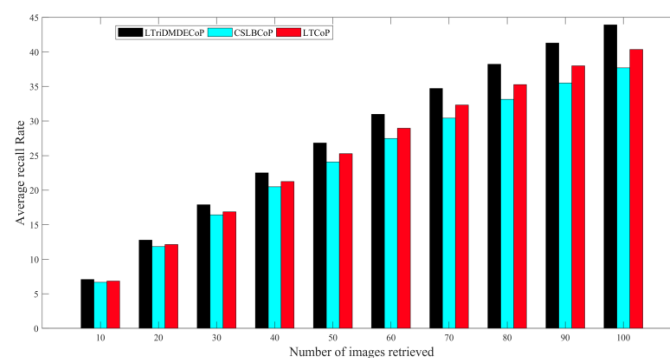
In this experiment, the proposed feature descriptor performance is evaluated on Corel-1000 database consisting of thousand images with 10 categories and each category consists of 100 images. For plotting the results, every image in the database is used as query image and for each query, the retrieved images are grouped into 10, 20, 30... 100. APR and ARR has been shown with number of images retrieved in Fig.4a and 4b. Table1 summarizes the precision and recall results of the proposed LTriDMDECoP and other existing methods.

The performance improvement of LTriDMDECoP in comparison with CSLBCoP and LTCoP on gray scale Corel-1000 image database are detailed below

- The average precision rate (APR) of LTriDMDECoP has significantly improved from 66.83% and 68.32% to 70.71% as compared with CSLBCoP and LTCoP respectively on Corel-1000 (n=10) as shown in Figure 4a and table1
- The average recall rate (ARR) has greatly improved from 37.7% and 40.36% to 43.9% (n=100) as compared to CSLBCoP and LTCoP.



(a)



(b)

Figure 4: Performance results on Corel-1000 image database

Table1 Performance comparison of proposed and other methods on Corel-1000 image database

Method	<i>APR(n=10)</i>	<i>ARR(n=100)</i>
LTriDMDECoP	70.71	43.9
CSLBCoP[6]	66.83	37.7
LTCoP[19]	68.32	40.36

4.2Experiment#2-Corel-5000 image database

The Corel-5000 image database consists of 5000 images with 50 categories and each category is having 100 images. The performance of proposed method with other existing methods is tabulated in Table2. The following conclusions are made from the experimental results

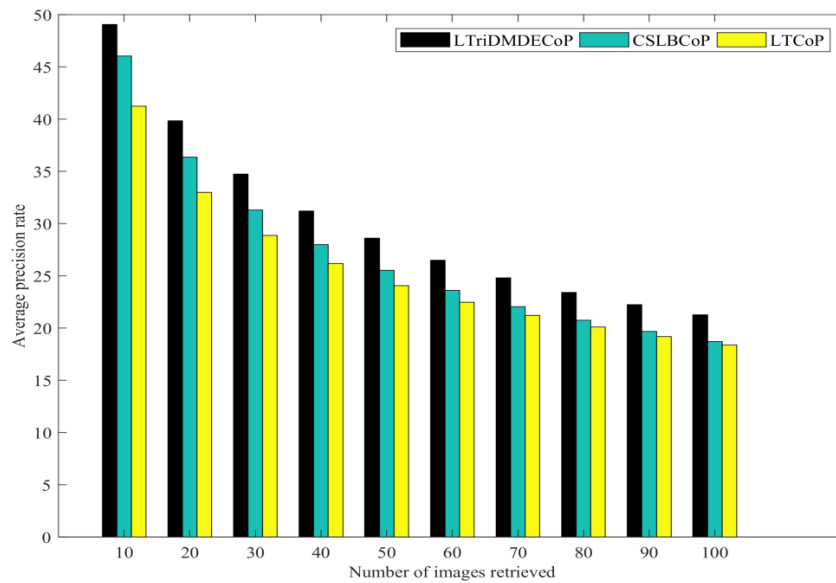
- The APR of LTriDMDECoP has been improved from 41.22% and 46.04% to 49.1% compared LTCoP and CSLBCoP as shown in Tabel2.
- From Table2 it is also clear that the proposed method outperformed the other two methods in terms of ARR. The performance comparison of proposed method with other existing methods in terms of APR and ARR Vs No. of images retrieved is as shown in Fig5.

Table 2: Performance comparison of proposed and other methods on Corel-5000 image database

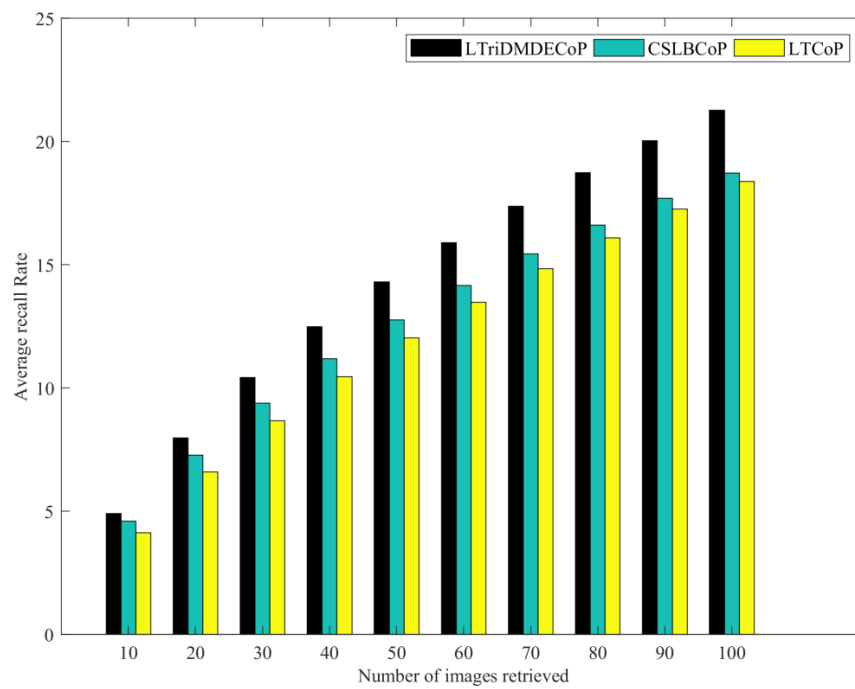
Method	<i>APR(n=10)</i>	<i>ARR(n=100)</i>
LTriDMDECoP	49.1	21.26
CSLBCoP[6]	46.04	18.71
LTCoP[19]	41.22	18.38

5. Conclusion

A novel feature extraction method called local tri directional median differential excitation co-occurrence pattern (LTriDMDECoP) is proposed in this work. The LTriDMDECoP feature exploited the relation between the center pixels with its neighbourhood pixel candidates at radius1 in an improved way. The human perception of pattern is also taken into account by considering the differential excitation values. Further, LTriDMDECoP encoded the co-occurrence of similar pattern pairs for feature vector formation. The experimental results indicated that the proposed LTriDMDECoP surpasses the existing LTCoP method in terms of APR and ARR on the two benchmark datasets and also shown significant improvement in terms of APR and ARR as compared to CSLBCoP on Corel-1000 and Corel-5000 image datasets. In conclusion the LTriDMDECoP extracts enriched features from the given image and shows enhanced discriminative power compared to the other two existing methods.



(a)



(b)

Figure5: Performance results on Corel-5000 image database

References

- [1] Ojala,T., Pietikäinen, Harwood, D., 1996. A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition*. 29(1), pp.51-59.
- [2] Cohen, F.S., Fan,Z., Attali, S.,1991. Automated inspection of textile fabrics using textural models. *IEEE Trans.Patrn.Alys.and Mach. Intell.* 13(8), pp.803-808.
- [3]Ahonen,T., Hadid,A., Pietikäinen,M.,2004. Face recognition with local binary patterns. *Computer Vision-ECCV-2004*, pp. 469-481.
- [4] Ning,J., Zhang,L., Zhang,D., 2009. Robust object tracking using joint color texture histogram. *Int. J. Pattern Recognit. Artif Intell*, 23 (7), pp. 1245-1263.
- [5]Liao,S., Law, M.W.K., Chung, A.C.S., 2009.Dominant local binary patterns for texture classification. *IEEE Trans. Image Processing*, 18(5), pp. 1107–1118.
- [6] Verma,M., Raman, B.,2015.Center symmetric local binary co-occurrence pattern for texture, face and bio-medical image retrieval. *J. Vis. Commun. Image Representation*, 32, pp. 224-236.
- [7]Takala,V.,Ahonen,T.,Pietikainen,M., Block-based methods for image retrieval using Local Binary Patterns. In *Lecture Notes in Computer Science*, 2005,(vol. 3540, pp. 882-891)
- [8] Zhao,Y., Jia,W., Hu,X.R.,Min,H.,2013.Completed robust local binary pattern for texture classification. *Journal of neuro computing*, 106, pp. 68-76.
- [9]Tan, X., Triggs, B., 2010.Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Trans. Image Process.*,19(6), pp. 1635-1650.
- [10] Zhang,B.,Gao,Y.,Zhao,S.,Liu,J.,2010.Local derivative pattern versus local binary pattern: Face recognition with high-order local pattern descriptor. *IEEE Trans. Image Processing*, 19(2), pp. 533-544.
- [11]Murala,S.,Maheshwari,R.P.,Balasubramanian,R.,2012.Local tetra patterns: A new feature descriptor for content-based image retrieval. *IEEE Trans. Image Processing*,21(5), pp. 2874-2886.
- [12] Xie, S., Shan,S., Chen,X., Chen,J.,2010.Fusing local patterns of gabor magnitude and phase for face recognition. *IEEE Trans. Image Processing*, 19(5), pp. 1349-1361.
- [13] Chen,J.,Kellokumpu,V.,RLBP: Robust local binary pattern. In *Proceedings of British machine vision conference*, 2016,(pp. 1-10).
- [14] Murala, S., Wu, Q.M.J., 2014.Local mesh patterns versus local binary patterns: Biomedical image indexing and retrieval. *IEEE J. Biomed. Heal. Informatics*, 18(3), pp. 929-938.
- [15] Liao, S., Chung, A., Face Recognition by Using Elongated Local Binary Patterns with Average Maximum Distance Gradient Magnitude. 2007,In *ACCV*, (vol.48, pp. 672-679).
- [16] Dubey,S.R.,Singh,S.K.,Singh,R.K.,2016.Local Bit-Plane Decoded Pattern: A Novel Feature Descriptor for Biomedical Image Retrieval. *IEEE J. Biomed. Heal. Informatics*, 20(4), pp. 1139-1147.
- [17] Hamouchene,I.,Aouat,S.,A New Texture Analysis Approach for Iris Recognition.In *AASRI Proceedings*, 2014,(vol. 9, pp. 2–7) .
- [18] Orjuela Vargas,S.A., Puentes,P.Y.J., Philips,W.,The Geometric local textural patterns (GLTP). *Local Binary Patterns: New Variants and Applications*, 2014,(Springer, pp.85-112).
- [19] Subrahmanyam, M., Wu, Q.M.J.,2013.Local ternary co-occurrence patterns: a new feature descriptor for MRI and CT image retrieval. *Journal of Neuro computing*,119 (7), pp.399–412.
- [20] Chen, J., Shinguang, Shan., Zhao, G., 2009.A robust descriptor based on Weber’s Law. *IEEE Trans. Pattern ana.Mach.Intell.*,32(9),pp 1705-1720.
- [21] Corel-10000 image database. [Online]. Available:<http://www.ci.gxnu.edu.cn/cbir/Dataset.aspx>.

Authors Biography

Mr. G V SATYA KUMAR obtained his B.Tech degree from JNT University, Hyderabad in year 2002 and M.Tech degree from ANU, Guntur in the year 2008. Presently he is pursuing Ph. D in image processing under the guidance of Dr P.G. Krishna Mohan. His areas of interests are image retrieval, object tracking.



Dr. P. G. Krishna Mohan presently working as Professor in Institute of Aeronautical Engineering, Hyderabad. He Worked as Head of ECE Dept. , Member of BOS for ECE faculty at University Level, Chairman of BOS of EIE group at University level, Chairman of BOS of ECE faculty for JNTUCEH, Member of selection committees for Kakitiya, Nagarjuna University, DRDL and convener for Universite a Hidian committees. He has more than 43 papers in various International and National Journals and Conferences. His areas of specialization are Signal Processing, Signal Estimation, Probability Random Variables and Communications.

