An enhanced Local tri-directional intensity pattern for face recognition

Praveena Kasu
Department of Computer Science and Engineering
Gokaraju Rangaraju Institute of Engineering and Technology
Hyderabad, India
kpraveenamary@gmail.com

Vijayalata Y
Department of Computer Science and Engineering
Gokaraju Rangaraju Institute of Engineering and Technology
Hyderabad, India
vijaya@griet.ac.in

Ashlin Deepa
Department of Computer Science and Engineering
Gokaraju Rangaraju Institute of Engineering and Technology
Hyderabad, India
deepa.ashlin@gmail.com

***Abstract*—With an increasing need for face, and emotion recognition, and potential biomedical, industrial and defense applications image retrieval techniques are being explored like never. Several new techniques and models have seen growth in recent years. However, very few techniques explore both direction and intensity information. In this paper, we present a local direction-based pattern for content-based image retrieval (CBIR). The proposed local feature descriptor captures potential and distinct features based on three directions and the relative magnitude difference of a pixel with its neighbors which makes the pattern robust to illumination changes. Experiments are conducted on four benchmark datasets- MIT Vistex database, Brodatz database, AT&T face database and IRIS thermal dataset. Analysis of results shows that the proposed method improves the retrieval rate.**

Keywords*— *Local binary pattern, feature extraction, texture descriptor, face recognition

I. INTRODUCTION

The evolution of diverse image acquisition techniques has resulted in rapid growth of digital image libraries and manual annotation has become tedious and expensive. This has heightened the need for a proficient automated technique for image retrieval from large databases. Content based image retrieval employs color, texture, shape, and other content-based information as opposed to metadata associated with an image. The potency of a system is determined by the type of feature extraction technique employed. Recently, a considerable literature has grown up around the theme of local feature descriptors, however, only a handful of them use both intensity and directional information of a pixel in its neighborhood.

Feature extraction in simple terms is used to reduce the dimensionality of an image. Texture represents the first level of spatial properties that can be extracted from an image. The current technique is a texture feature descriptor which retrieves potential sign and intensity information from three directions. The results presented demonstrate its effectiveness over previous pattern descriptors. The rest of the work is arranged as follows- Section 1 contains the related work, Section 2 describes two pattern descriptors, Section 3 has the proposed method, outcomes are presented in Section 4 and Section 5 concludes the work. are provided.

A. Related work

Feature extraction based on texture was extensively used in the past. As a result, an appreciable amount of literature has evolved which will be explored in this section. In [1], Gray-level co-occurrence matrix based on pixel pairs has been presented. In transformational domain methods like Discrete Wavelet Transform (DWT) [2], Gabor filters [3], rotated complex wavelet filter [4] have been proposed.

A powerful technique called Local Binary Pattern (LBP) [5] laid the foundation for several new patterns in the past two decades. Numerous methods were proposed with modifications on the traditional LBP such as [6, 7, 8, 9, 10]. In LBP only the first order local derivatives were considered, therefore it was extended to Local Derivative Pattern (LDP) [11] where nth order derivatives are included. For real-time applications which involve images from different environments, Local Ternary Pattern (LTP) [12] was introduced with an interval to compare the center and neighboring pixels. Local tetra pattern (LTrP) [13] incorporated the derivatives and results obtained are substantially better than LBP and LDP. Weber local pattern introduced in (WLBP) [14] combines differential excitation pattern with Local binary pattern. Local mesh patterns (LMeP) [15] were proposed which were employed in medical image retrieval. A non-linear fusion of LBP and LDP methods produced a rotation invariant Local Optimal Oriented Pattern (LOOP) [16].

Although many local patterns were proposed, the affiliation among local pixels was not considered until the introduction of Local Tri-directional Pattern (LTriDP) [17]. Later, local neighborhood intensity pattern (LNIP) [18] was proposed on similar lines.

Our present work, Local Tri-directional Intensity Pattern (LTriDIP) focuses on improving the performance of the magnitude component in the traditional $LTriDP_m$. As LTriDIP considers the relative magnitude variation in a neighborhood, it is unaffected by gray level changes. Experiments are conducted on four databases- Brodatz Texture database [19], MIT VisTex database [20], AT&T Face database [21], and IRIS Thermal face database [22] to observe the robustness of the proposed method.

II. LOCAL PATTERN DESCRIPTORS

Before LBP [10] describes the affiliation among pixels. Due to its simplicity, researchers employed LBP to solve problems like face recognition, fingerprint recognition, palm print recognition, classification, and object tracking. LBP can be explained mathematically by using eqn. (1), where *p* stands for neighbor pixels in a circle of radius r

$$LBPp,r = \sum_{l=0}^{p-1} 2l \times S1(Ii - Ic)$$

$$S_1(x) = \begin{cases} 1 & x \geq 0 \\ 0 & else \end{cases} \quad (1)$$

Here, center and neighborhood pixel intensities are represented by I_c and I_i respectively. The plot of the LBP map is generated using eqn. (2).

$$His(H)|_P = \sum_{a=1}^{u} \sum_{b=1}^{v} S_2(Pattern(a,b), H)$$

$$H \in [0, (2^p - 1)]$$

$$S_2(i,j) = \begin{cases} 1 & i = j \\ 0 & else \end{cases} \quad (2)$$

where *u* and *v* are dimensions of the image, respectively, Pattern here stands for local binary pattern, and *H* is the feature vector length.

LBP and several other descriptors capture only sign variation among the pixels but overlook the mutual affiliation among neighbor pixels. LTriDP [17] examined the mutual relationship between pixels in three directions along with their magnitude difference. Both direction and magnitude pattern histograms were combined to produce the feature vector.

A comparison of LTriDP and local tri-directional pattern with magnitude component ($LTriDP_m$) proved the excellency of LTriDP for image retrieval task [17]. Although tri-directional pattern itself is a powerful discriminator and more important than magnitude, [13] proves the significance of counting the magnitude pattern to create a superior feature vector. Also neglecting the magnitude component poses the danger of encoding different structural patterns into the same code. Hence, a magnitude pattern for better feature description is proposed in the present paper.

III. PROPOSED METHOD

After Local Tri-Directional Intensity Pattern (LTriDIP) is a consolidation of two components- the three directional pixel affiliation and the intensity variation among pixels. As both these constituents obtain different information, their concatenation produces a better feature descriptor. The proposed magnitude pattern is based on the concept of statistical dispersion. It calculates the average absolute deviations and therefore incorporates structural differences among the pixels which ensures resilience to light variations and curtails any noise in the image.

Pixels in proximity share higher correlation than pixels farther away. Consider a 3×3 patch with 9 pixels I_1, $I_2, \dots, I_8$ and I_c where I_c is at the center as in Fig. 2 (a). The pattern formation can be explained mathematically as follows:

$$D_1 = I_i - I_{i-1}$$

$$D_2 = I_i - I_{i+1}$$

$$D_3 = I_i - I_c \qquad \forall \quad i = 2,3 \dots, 7 \quad (3)$$

$$D_1 = I_i - I_8 \quad D_2 = I_i - I_{i+1} \quad D_3 = I_i - I_c$$

$$\text{for } i = 1 \quad (4)$$

$$D_1 = I_i - I_{i-1} \quad D_2 = I_i - I_1 \quad D_3 = I_i - I_c$$

$$\text{for } i = 8 \quad (5)$$

Here, D_1, D_2 and D_3 represent the three directional differences for each neighboring pixel. Based on the above differences a pattern number is assigned.

$$f(D_1, D_2, D_3) = \{\#(D_k < 0)\} \bmod 3 \quad \forall\, k = 1,2,3 \quad (6)$$

$\#(D_k < 0)$ is the number of D_k which are negative, for all $k = 1,2,3$ and it gives values ranging from 0 to 3. Later, mod of $\#(D_k < 0)$ is taken with 3 which results in 0, 1, or 2. For each neighbor pixel pattern value $f(D_1, D_2, D_3)$ is calculated using eqn. (6).

$$LTriDIP(I_c) = \{f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8\} \quad (7)$$

Eqn. (8) describes the generation of 8-bit ternary number from which two binary patterns are obtained.

$$LTriDIP_1(I_c) = \{S_3(f_1), S_3(f_2), \dots, S_3(f_8)\}$$

$$S_3(x) = \begin{cases} 1 & x = 1 \\ 0 & else \end{cases} \quad (8)$$

$$LTriDIP_2(I_c) = \{S_4(f_1), S_4(f_2), \dots, S_4(f_8)\}$$

$$S_4(x) = \begin{cases} 1 & x = 2 \\ 0 & else \end{cases} \quad (9)$$

$$LTriDIP(I_c)|_{i=1,2} = \sum_{l=0}^{7} 2^l LTriDIP_i(I_c)(l+1) \quad (10)$$

The histogram for binary patterns obtained in eqn. (8) and eqn. (9) are computed using eqn. (2).

When normal average of differences is used for calculating the magnitude pattern, the algebraic signs negate. Therefore, the magnitude pattern considered here uses the average absolute difference of two adjacent pixels from their common neighbor ($I_i\ \forall\ i = 1, 2, \dots 8$) as shown in Fig. 2. with respect to threshold is obtained by the average absolute difference of all 8 neighbor pixels (I_i) about the center pixel (I_c). The magnitude pattern can be explained mathematically as follows:

$$M_i = \frac{1}{2} \sum_{k=1}^{2} |S_i(k) - I_i| \quad (11)$$

$$T_c = \frac{1}{8} \sum_{i=1}^{8} |I_i - I_c| \quad (12)$$

Here, M_i stands for absolute mean deviation of i^{th} neighbor of the center from the two adjacent neighbors S_i (i = 1, 2, 3, …8), eqn. (11). This is calculated for all the 8 neighbors of I_c in a 3 x 3 patch. T_c is the average absolute difference of I_i (i=1,2,3, …8) from I_c, eqn. (12). Finally, the modified magnitude pattern is given by eqn. (14).

$$M(I_i, T_c) = S_5(M_i, T_c)$$

$$S_5(x,y) = \begin{cases} 1 & x \geq y \\ 0 & else \end{cases} \quad (13)$$

$$LTriDIP_3(I_c) = \sum_{i=1}^{8} 2^{i-1} \times M(I_i, T_c) \quad (14)$$

The histogram is generated using eqn. (2). The final combined histogram is as follows

$$Hist = [His|_{LTriDIP_1}, His|_{LTriDIP_2}, His|_{LTriDIP_3}] \quad (15)$$

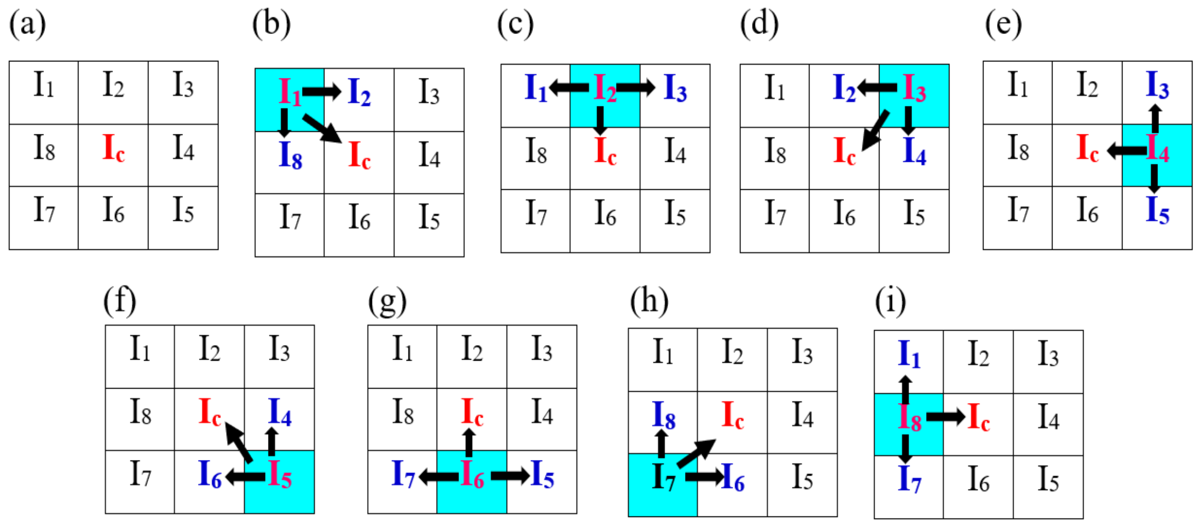


Fig. 1. Example illustration for LTriDIP. (a)- I_c and its neighbor pixels. (b)-(i): Tri-directional affiliation of pixels for each of the 8 neighborhood pixels (I_i, i=1,2,3,4,5,6,7,8) of center pixel I_c.

The local tri-directional pixel relation is illustrated in Fig. 2 through windows (a)-(i). Sample pattern map for LTriDIP is in Fig. 1. Sample pattern map for LTriDIP is shown in Fig. 2.

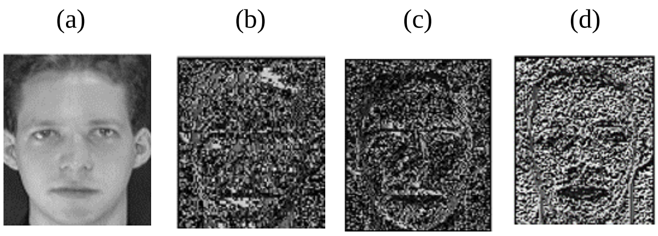


Fig. 2. (a): Example image of database C, (b): LTriDIP map generated using eqn. (8), (c): LTriDIP map generated using eqn. (9), (d): LTriDIP intensity pattern map generated using eqn. (14).

Implementation for the present method is same as in [16].

A. Query Matching

For query matching, d_1 similarity distance metric is used:

$$Dis(Q, DB^n) = \sum_{s=1}^{H} \left| \frac{F_{db}^n(s) - F_q(s)}{1 + F_{db}^n(s) + F_q(s)} \right| \quad (16)$$

Proper choice of distance metric empowers retrieval rate of a texture descriptor. The proposed method in combination with Euclidean distance measure or cosine distance measure shows a boost in the performance. Euclidean distance metric:

$$Dis(Q, DB^n) = \left(\sum_{s=1}^{H} \left|(F_{db}^n(s) - F_q(s))^2\right|\right)^{1/2} \quad (17)$$

Cosine similarity metric:

$$Dis(Q, DB^n) = 1 - \frac{dot\left(F_{db}^n(s), F_q(s)\right)}{norm\left(F_q(s)\right) norm\left(F_{db}^n(s)\right)} \quad (18)$$

here, H stands for feature vector length, F_{db}^n *and* F_q stand for the database nth image feature vector and the input image feature vector.

Further, although LTriDIP produces smaller feature vector than methods like PVEP, LMEBP and DLEP, its vector length is high. Therefore, Principal Component Analysis (PCA) is employed to describe the data without redundancies.

IV. RESULTS AND DISCUSSION

A. Performance metric

Precision and recall are used as retrieval performance evaluation metrics. They are calculated as described in [16].

B. Database A

MIT VisTex database [20] has 40 gray-scale texture images. Each image is divided into 16 images to obtain 40 image categories with 16 images each.

It is obvious from the results in Fig. 3. that the proposed LTriDIP outperforms LTriDP average precision (APP) and recall percentages (ARR) by 6.37% and 5.71% and $\mathrm{LTriDP_m}$ APP and ARP by 1.91% and 3.24% when cosine distance is used. Fig. 4 shows the effectiveness of cosine distance measure over Euclidean distance by 2.05% (average precision) and 2.06% (average recall).from the MS Word Standard toolbar and then select the correct number of columns from the selection palette.

C. Database B

Brodatz database [19] has 112 texture images. Each image is divided into 25 non-overlapping images. Therefore, we obtain 112 image categories with 25 images each.

Results from Fig. 5. outperforms LtriDP APP and ARP by 7.17% and 5.81%, and $\mathrm{LTriDP_m}$ APP and ARP by 2.71% and 2.61% when cosine distance is used as similarity measure. Fig. 6. shows cosine distance measure outperforms euclidean distance measure APP and ARP by 2.52% and 4.56%.database [19] has 112 texture images.

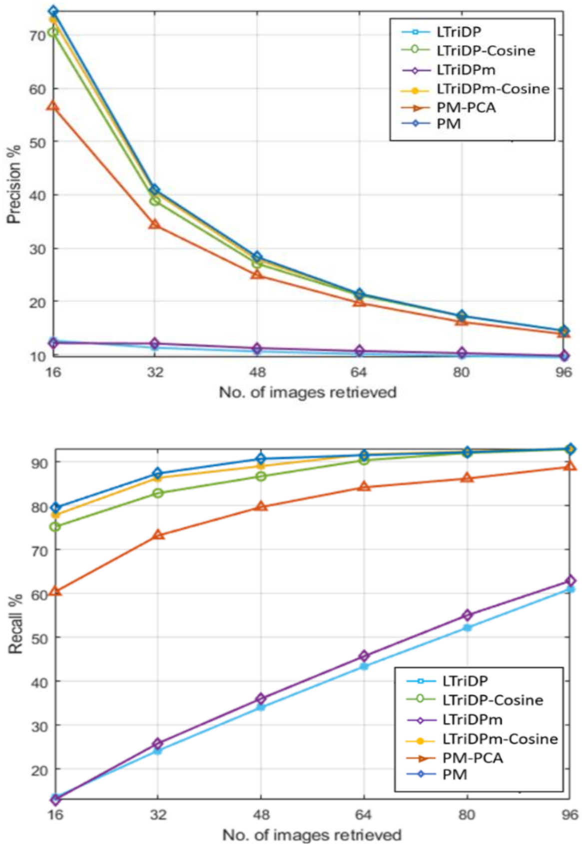


Fig. 3. Retrieval performance of LTriDIP on database A.

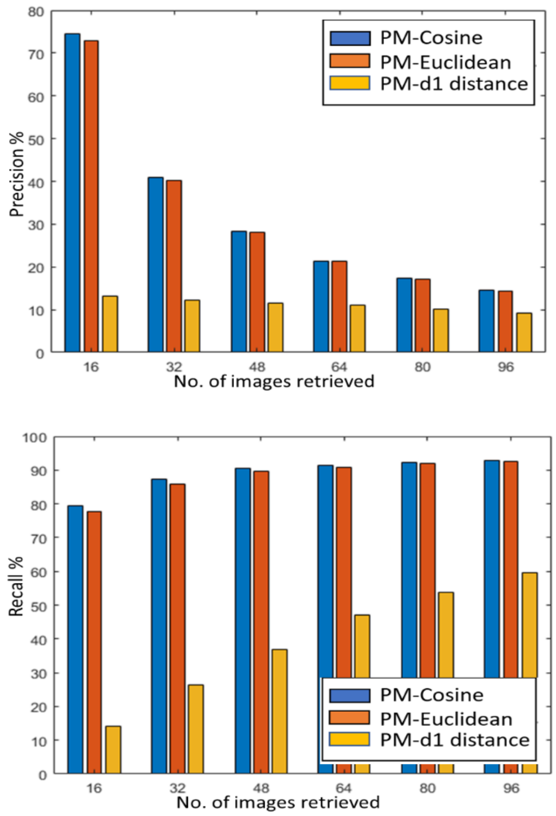


Fig. 4. Retrieval performance of different similarity measures on database A.

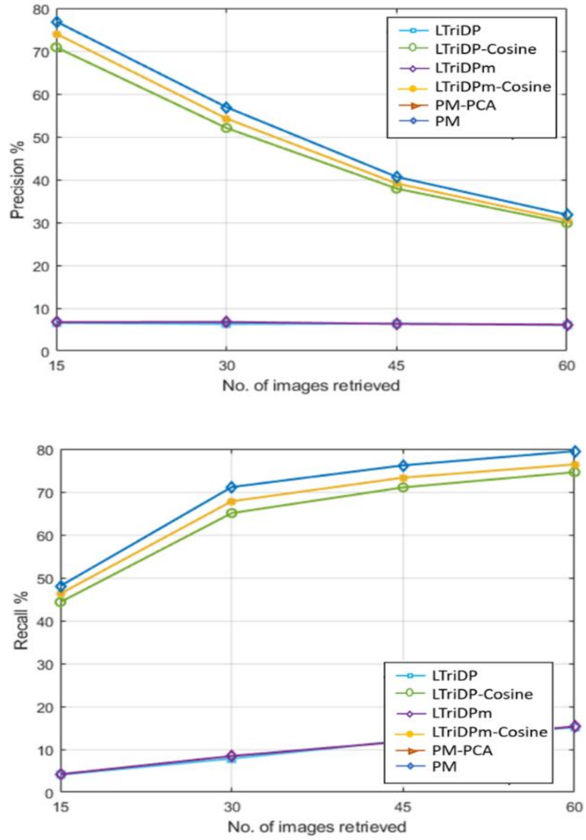


Fig. 5. Retrieval performance of LTriDIP on database B.

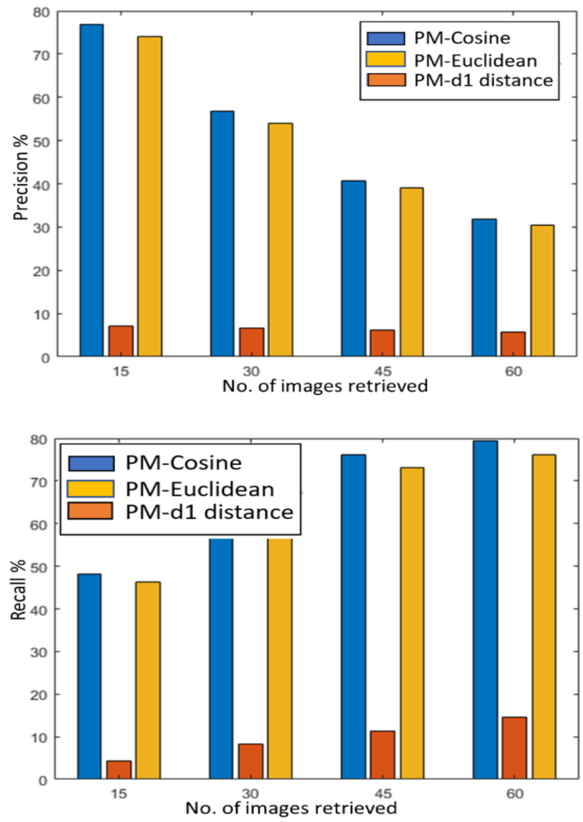


Fig. 6. Retrieval performance of different similarity measures on database B.

D. Database C

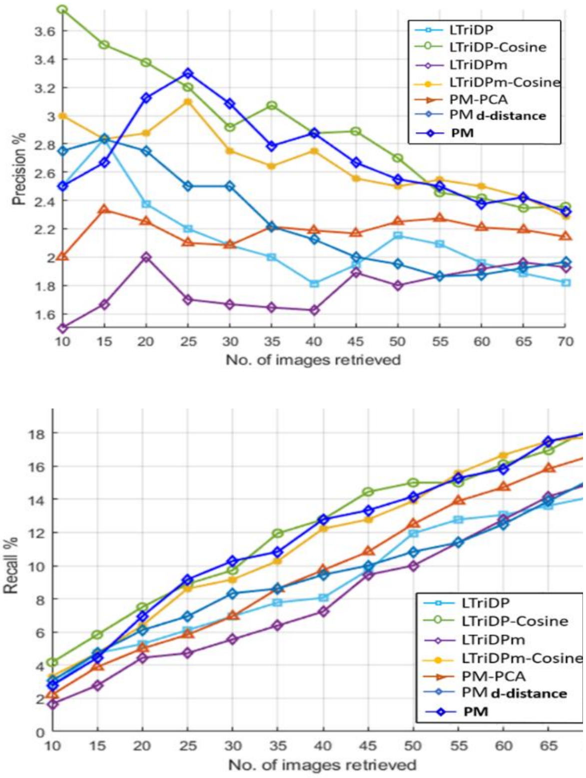


Fig. 7. Retrieval performance of LTriDIP on database C.

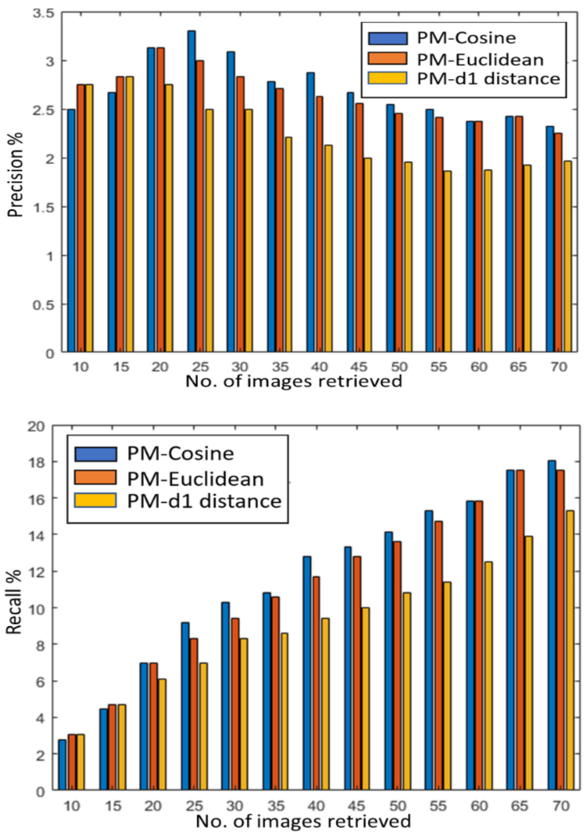


Fig. 8. Retrieval performance of different similarity measures on database C.

AT&T face database [21] is used in the third experiment. It has 400 images each of size 92 × 112.

The proposed method yields more satisfactory results than LTriDP$_m$, however LTriDP in combination with cosine distance measure proves to be more efficient. Cosine similarity measure outperforms Euclidean similarity measure APP and ARP by 8.85% and 7.15%

E. Database D

IRIS Thermal/Visible face database [22] has images with varying illumination, positions, and facial expressions.

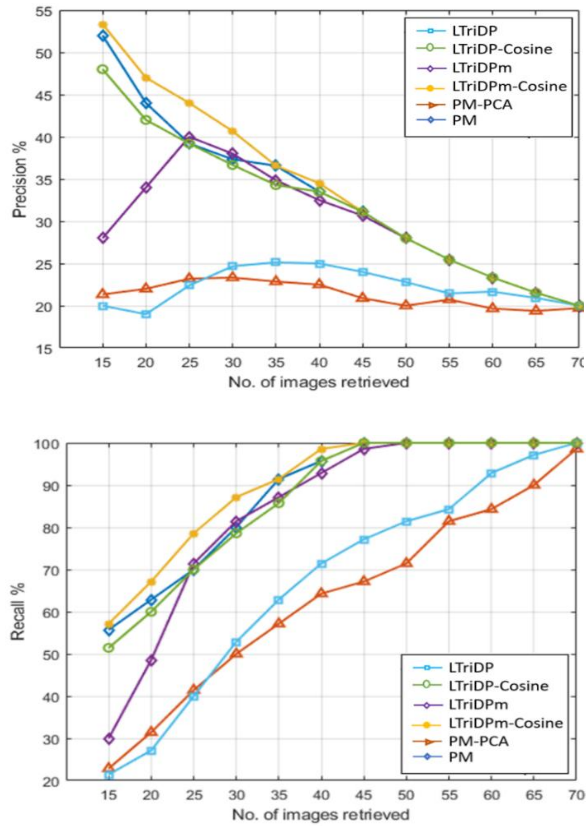
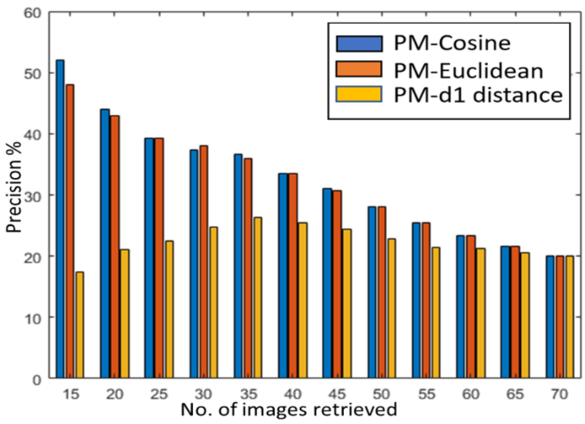


Fig. 9. Retrieval performance of LTriDIP on database D.



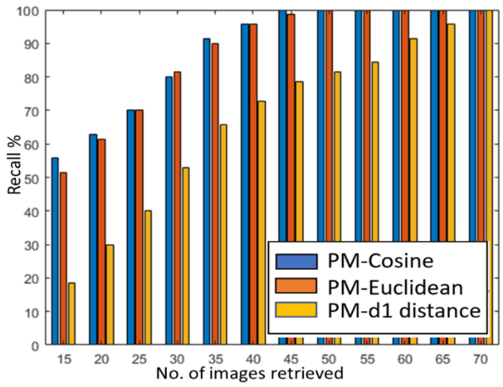


Fig. 10. Retrieval performance of different similarity measures on database .

LTriDIP shows higher retrieval potential than LTriDP when cosine distance is used as similarity measure and $\mathrm{LTriDP_m}$ when d1-distance is used as similarity measure. It shows an improvement in average precision and recall values by 12.99% and 8.21% for LTriDP. However, $\mathrm{LTriDP_m}$ with cosine distance as similarity measure shows slightly higher retrieval rate. Cosine distance as similarity measure improves the APP and ARP by 9.89% and 9.96% compared to Euclidean distance measure.

V. CONCLUSIONS AND FUTURE WORK

Here, a local descriptor entitled LTriDIP is proposed. This descriptor is powerful enough to capture potential features from noisy data and shows enriched discriminative power compared to several other local pattern descriptors. The relative magnitude variation among the pixels considered in this paper renders the model robust.

Although this method achieved good accuracy, due to the concatenation of three histograms the feature vector length is 768 which is undesirable. Hence, an improvement in the feature vector length will make the system more efficient. To achieve better precision and recall rates filters can be applied to the present technique. However, the present method is potential enough to be employed in various image retrieval applications.

REFERENCES

[1] R. M. Haralick, K. Shanmugam, I. Dinstein, SMC **3**, 6 (1973)

[2] A. Ahmadian, A. Mostafa, EMBC **1**, 930-933 (2003)

[3] B. S. Manjunath, W. Y. Ma, IEEE Trans. Pattern Anal. Mach. Intell. **18**, 8 (1996)

[4] M. Kokare, P. K. Biswas, B. N. Chatterji, IEEE Trans. Cybern. **35**, 6 (2005)

[5] T. Ojala, M. Pietikäinen, D. Harwood, Pattern Recognit. **29**, 1 (1996)

[6] T. Ojala, M. Pietikainen, T. Maenpaa, IEEE Trans. Pattern Anal. Mach. Intell. **24**, 7 (2002)

[7] S. Liao, M. W. K. Law, A. C. S. Chung, IEEE Trans. Image Process. **18**, 5 (2009)

[8] Z. Guo, L. Zhang, D. Zhang, IEEE Trans. Image Process. **19**, 6 (2010)

[9] Z. Guo, L. Zhang, D. Zhang, Pattern Recognit. **43**, 3 (2010)

[10] X. Qian, X. Hua, P. Chen, L. Ke, Pattern Recognit. **44**, 10-11 (2011)

[11] B. Zhang, Y. Gao, S. Zhao, J. Liu, IEEE Trans. Image Process. **19**, 2 (2010)

[12] X. Tan, B. Triggs, IEEE Trans. Image Process. **19**, 6 (2010)

[13] S. Murala, R. P. Maheshwari, R. Balasubramanian, IEEE Trans. Image Process. **21**, 5 (2012)

[14] S. Murala and Q. M. J. Wu, IEEE J. Biomed. Health Inform. **18**, 3 (2014)

[15] F. Liu, Z. Tang, J. Tang, Neurocomputing **120**, 325-335 (2013)

[16] T. Chakraborti, B. McCane, S. Mills, U. Pal, IEEE Signal Process. Lett. **25**, 5 (2018)

[17] M. Verma, B. Raman, Digit. Signal Process. **51**, 62-72 (2016)

[18] P. Banerjee, A. K. Bhunia, A. Bhattacharyya, P. P. Roy, S. Murala, Expert Syst. Appl. **113**, 100-115 (2018)

[19] A. Safia, D. He, ISRN Machine Vision, http://multibandtexture.recherche.usherbrooke.ca/original_brodatz.html (2013)

[20] MIT Vision and Modeling Group, Cambridge, Vision texture, http://vismod.media.mit.edu/pub/ (2010)

[21] AT&T Laboratories Cambridge, The AT&T database of faces, http://www.uk.research.att.com/facedatabase.html (2002)

[22] IEEE OTCBVS WS Series Bench; DOE University Research Program in Robotics under grant DOE-DE-FG02-86NE37968; DOD/TACOM/NAC/ARC Program under grant R01-1344-18; FAA/NSSA grant R01-1344-48/49; Office of Naval Research under grant #N000143010022.