# Local Gradient Hexa Pattern: A Descriptor for Face Recognition and Retrieval

Soumendu Chakraborty, *Graduate Student Member, IEEE*, Satish Kumar Singh, *Senior Member, IEEE*, and Pavan Chakraborty, *Member, IEEE*

***Abstract*— Local descriptors used in face recognition are robust in a sense that these descriptors perform well in varying pose, illumination, and lighting conditions. The accuracy of these descriptors depends on the precision of mapping the relationship that exists in the local neighborhood of a facial image into microstructures. In this paper, a local gradient hexa pattern is proposed that identifies the relationship among the reference pixel and its neighboring pixels at different distances across different derivative directions. Discriminative information exists in the local neighborhood as well as in different derivative directions. The proposed descriptor effectively transforms these relationships into binary micropatterns discriminating inter-class facial images with optimal precision. The recognition and retrieval performance of the proposed descriptor has been compared with state-of-the-art descriptors, namely, local derivative pattern, local tetra pattern, multiblock local binary pattern, and local vector pattern over the most challenging and benchmark facial image databases, i.e., Cropped Extended Yale B, CMU-PIE, color-FERET, LFW, and Ghallager database. The proposed descriptor has better recognition as well as retrieval rates compared with state-of-the-art descriptors.**

***Index Terms*— Face recognition, local derivative pattern (LDP), local gradient hexa pattern (LGHP), local pattern descriptors, local tetra pattern (LTrP), local vector pattern (LVP), multiblock local binary pattern (MLBP).**

## I. Introduction

### A. Motivation

LOCAL descriptors have been gaining more and more recognition in recent years as these descriptors are becoming capable enough to identify the unique features, which suitably and uniquely describe any image for recognition and retrieval. The original problem was that of the facial region identification and extraction of facial features (eye, nose, and mouth) [1]–[2]. The identification problem has evolved as a classification problem in which an image is classified as a face or a nonface. Some of the mathematical models used for classification are based on singular value decomposition [3], neural networks [4], principle component analysis (PCA) [5], [6], variations of PCA [7]–[10], and linear discriminant analysis [5], [6], [11]–[13]. In recent literature, deep learning-based image descriptors have been proposed in which a convolutional neural network is trained to classify the facial images [32], [33]. Rotation invariant feature descriptor, local binary pattern (LBP), and its variants have been proposed for texture classification [14]–[16]. LBP was later extended to identify the spatial relationships among the pixels in the local neighborhood of the facial images [17]. Color feature-based descriptor proposed in [18] extracts $M$ color component features to describe a facial image for recognition. Color texture feature extraction proposed in [19] applies Gabor wavelet and LBP on different color channels to extract the facial features. Multiblock LBP (MLBP) computes average values of intensities using $2 \times 3$ mask, over which LBP is computed [31]. Multiblock LBP achieves 8% improvement over the original LBP feature [31]. Local directional number pattern [20] computes the edges in an image using compass mass and selects most positive and negative directions as the features. There exists some descriptors, namely, local derivative pattern (LDP) [21], local tetra pattern (LTrP) [22], and local directional gradient pattern (LDGP) [34], which explore higher order derivative space in the local neighborhood of the reference pixel. Most of the existing local descriptors tend to confine the local neighborhood to the eight pixels nearest to the reference pixel (current pixel under consideration). These descriptors tend to ignore the discriminating information that exists at different radial widths as well as across higher order derivative spaces.

### B. Related Work

More discriminative features that exist in higher order derivative directions were captured by the LDP [21], an improvement over LBP. LDP computes the derivative of a facial image in four directions and encodes the relationship of the reference pixel and its eight neighbors in different derivative directions. LTrP [22] divides the image in two derivatives along 0° and 90° and encodes the directions of the reference pixel and its eight neighbors computed from these derivatives. LTrP is an extension of LDP and shows some improvement over LDP with respect to retrieval rate. Local vector pattern (LVP) [27] computes vectors in four unique directions, which is similar to the derivatives computed as in the case of LDP. Pair-wise vectors are encoded using comparative space transform [27] to generate the micropatterns. LVP is an effective extension over LTrP. Most of these descriptors discard the information that exists at higher radial distances, which adversely affect the retrieval and recognition accuracies. The proposed local gradient hexa pattern (LGHP) encodes the

Manuscript received May 21, 2016; accepted August 24, 2016. Date of publication August 26, 2016; date of current version January 5, 2018. This paper was recommended by Associate Editor A. Loui.

The authors are with Indian Institute of Information Technology at Allahabad, Allahabad 211012, India (e-mail: soum.uit@gmail.com; sk.singh@iiita.ac.in; pavan@iiita.ac.in).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCSVT.2016.2603535

1051-8215 © 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

discriminative relationship among the reference pixel and its eight neighbors at different distances in four distinct derivative directions using a unique encoding scheme and achieves better recognition and retrieval rate compared with LDP, LTrP, MLBP, and LVP.

### C. Major Contribution

Descriptors that explore higher order derivative space tend to achieve better results under pose, expression, light, and illumination variations. LDP is a local descriptor that explores the higher order derivative space, but tends to discard the radial extension of the local neighborhood. LDP also discards the relationships that exist between the higher order derivative spaces (i.e., inter-derivative space relationships). LVP is another descriptor that ignores the relationships among pixels at greater radial widths. However, it captures the discriminating information between derivative spaces, i.e., between (0°, 45°), (45°, 90°), (90°, 135°), and (135°, 180°).

The proposed descriptor captures discriminating information across different derivative spaces, namely, $G^1_{0°,R}$, $G^1_{45°,R}$, $G^1_{90°,R}$, and $G^1_{135°,R}$ at different angular widths 0°, 45°, 90°, and 135°, respectively, at different radial widths $R$. It also captures discriminating information between derivative spaces, i.e., between (0°, 45°), (0°, 90°), (0°, 135°), (45°, 90°), (45°, 135°), and (90°, 135°). Unlike LVP and LDP, the proposed descriptor encodes most of the relevant information that exists in the local neighborhood of the reference pixel across different angular widths, radial widths, and higher order derivative spaces including inter-derivative space relationships. The result analysis objectively proves that the proposed descriptor encodes additional relevant information, which is ignored by the state-of-the-art descriptors, i.e., LDP, LTrP, MLBP, and LVP.

The organization of the rest of this paper is as follows. Section II elaborates the proposed descriptor. Performance measures are defined in Section III. The performance of the proposed descriptor has been analyzed and compared with the LVP, LTrP, MLBP and LDP in Section IV. We conclude this paper in Section V.

## II. Local Gradient Hexa Pattern

LGHP identifies the variations in the original pixels of a facial image in the higher order derivative space. The major differences between LDP, LTrP, MLBP, LVP, and LGHP are the encoding function and the structure of the micropattern.

Four first-order directional derivatives (gradients) $G^1_{\alpha,D}$ of the image $I(P)$ of size $M \times N$ in $\alpha = 0°, 45°, 90°$, and 135° directions at distance $D$ for an arbitrary reference point $P_0$ are computed as

$$G^1_{\alpha,R}(P_0) = I(P_0) - I(P_{(\frac{\alpha}{45}+1),D}). \tag{1}$$

Fig. 1 shows the reference pixel and eight neighbors at different distances. $P_0$ is the reference pixel, and pixel in 0° direction at a distance $D = 1$ is represented as $P_{((0/45)+1),1}$ that is $P_{1,1}$. Fig. 2 shows the flow diagram of the LGHP computation process.

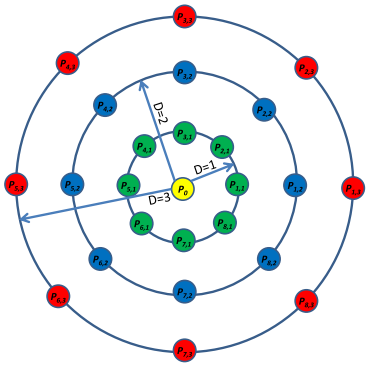


Fig. 1. Template of the local subregion of an image.

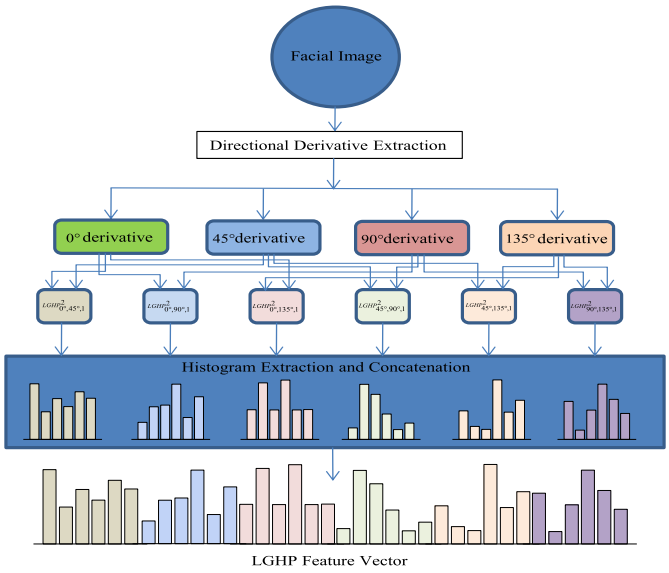


Fig. 2. Flow diagram of the LGHP computation.

Second-order $\text{LGHP}^2_{\alpha,\beta,D}(.)$ at $D$ distance is computed by encoding the pair-wise derivatives using encoding function $C(.,.)$ and concatenating these encoded patterns. $\text{LGHP}^2_{\alpha,\beta,D}(.)$ is defined as

$$\begin{aligned}&\text{LGHP}^2_{\alpha,\beta,D}(P_0)\\ &= \{C(G^1_{\alpha,D}(P_0), G^1_{\beta,D}(P_0)), C(G^1_{\alpha,D}(P_{1,D}), G^1_{\beta,D}(P_{1,D})),\\ &\quad \ldots C(G^1_{\alpha,D}(P_{8,D}), G^1_{\beta,D}(P_{8,D}))\}\end{aligned} \tag{2}$$

where $\alpha, \beta = 0°, 45°, 90°,$ and 135°.

The encoding function $C(.,.)$ for any point $P$ in the derivative space is defined as

$$C(G^1_{\alpha,D}(P), G^1_{\beta,D}(P)) = \begin{cases} 1, & \text{if } G^1_{\alpha,D}(P) > G^1_{\beta,D}(P) \\ 0, & \text{else.} \end{cases} \tag{3}$$

Second-order LGHP at a distance $D$ is calculated by concatenating $\text{LGHP}^2_{\alpha,\beta,D}(P_0)$ for different $(\alpha, \beta)$ pairs as

$$\begin{aligned}&\text{LGHP}^2_D(P_0)\\ &= \{\text{LGHP}^2_{\alpha,\beta,D}(P_0) | (\alpha,\beta) = (0°, 45°), (0°, 90°), (0°, 135°),\\ &\qquad (45°, 90°), (45°, 135°), (90°, 135°)\}.\end{aligned} \tag{4}$$

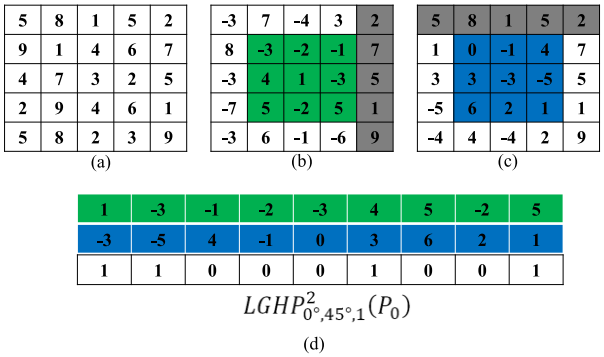


Fig. 3. (a) Original sample image. (b) $G^1_{0°,1}(P_0)$ of the original sample. (c) $G^1_{45°,1}(P_0)$ of the original sample. (d) Illustration of computation of $\text{LGHP}^2_{\alpha,\beta,D}(P_0)$ for $\alpha = 0°$ and $\beta = 45°$.

The proposed LGHP depends on another parameter $R \in [1, N]$ that specifies the upper limit of $D$ over which hexa patterns are computed.

Finally, second-order LGHP is computed as

$$\text{LGHP}^2_R(P_0) = \left\{\text{LGHP}^2_D(P_0) | D = 1, 2, 3 \ldots R\right\}. \quad (5)$$

LGHP encodes six binary patterns of 9 b each at a particular distance for different $(\alpha, \beta)$ pairs. These six patterns are converted to equivalent decimal values to generate six LGHP matrices. Spatial histograms of these six matrices are computed as

$$\begin{aligned}&\text{HLGHP}(\alpha, \beta, D)\\ &= \{H_{\text{LGHP}_{\alpha,\beta,D}} | (\alpha, \beta) = (0°, 45°), (0°, 90°), (0°, 135°),\\ &\qquad (45°, 90°), (45°, 135°), (90°, 135°)\} \quad (6)\end{aligned}$$

where $H_{\text{LGHP}_{\alpha,\beta,D}}$ is the histogram extracted from LGHP matrices corresponding to different pairs of $\alpha$ and $\beta$.

$L1$ is used to measure the similarity between two histograms as it performs better than other measures on the data sets used in the experiments. Similarity measure $S_{L1}(.,.)$ is defined as

$$S_{L1}(X, Y) = \sum_{i=0}^{q} |x_i - y_i| \quad (7)$$

where $S_{L1}(X, Y)$ is the $L1$ distance computed on two vectors $X = (x_1, \ldots, x_q)$ and $Y = (y_1, \ldots, y_q)$. Nearest one neighbor (1NN) classifier is used as used in [21] to compute the minimum $L1$ distance between the probe image and the gallery images, as similar regions of the probe and gallery images are effectively identified by 1NN classifier with optimal computational cost [21].

Fig. 3(a) shows a sample image, and its derivatives in 0° and 45° directions are shown in Fig. 3(b) and (c), respectively. Reference pixel and its eight neighbors in 0° and 45° derivative directions are shown in green and blue rows, respectively. Fig. 3(d) shows the computation of the 9-b binary pattern shown in the white row from 0° derivative shown in the green row and 45° derivative shown in the blue row. Similarly, the remaining five 9-b patterns for direction pairs $(0°, 90°)$, $(0°, 135°)$, $(45°, 90°)$, $(45°, 135°)$, and $(90°, 135°)$ are computed to generate the hexa pattern at a distance $D$.

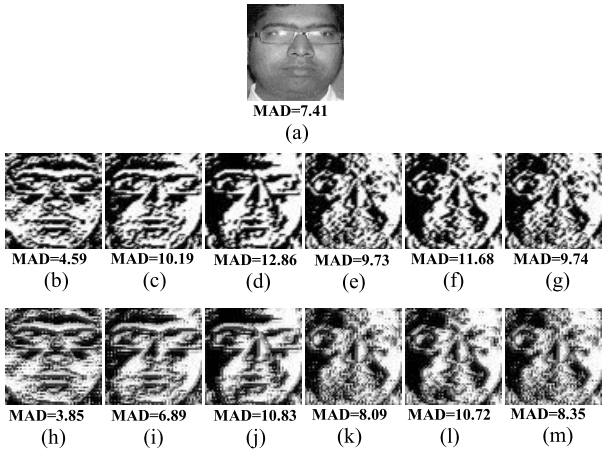


Fig. 4. FIs and MADs. (a) Original. (b) FI at $D = 1$ and direction pair $(0°, 45°)$ having MAD = 4.59. (c) FI at $D = 1$ and direction pair $(0°, 90°)$ having MAD = 10.19. (d) FI at $D = 1$ and direction pair $(0°, 135°)$ having MAD = 12.86. (e) FI at $D = 1$ and direction pair $(45°, 90°)$ having MAD = 9.73. (f) FI at $D = 1$ and direction pair $(45°, 135°)$ having MAD = 11.68. (g) FI at $D = 1$ and direction pair $(90°, 135°)$ having MAD = 9.74. (h) FI at $D = 2$ and direction pair $(0°, 45°)$ having MAD = 3.85. (i) FI at $D = 2$ and direction pair $(0°, 90°)$ having MAD = 6.89. (j) FI at $D = 2$ and direction pair $(0°, 135°)$ having MAD = 10.83. (k) FI at $D = 2$ and direction pair $(45°, 90°)$ having MAD = 8.09. (l) FI at $D = 2$ and direction pair $(45°, 135°)$ having MAD = 10.72. (m) FI at $D = 2$ and direction pair $(90°, 135°)$ having MAD = 8.35.

Feature images (FIs) computed using LGHP shown in Fig. 4 visually depict the discriminating information captured by the descriptor. As shown in Fig. 4, FIs computed at different distances encode useful relationships among neighborhood pixels, which help in discriminating inter-class facial images. FIs shown in Fig. 4(b)–(g) and (h)–(m) visually demonstrate that while considering the facial features at greater distances, the facial features at smaller distance should not be discarded. To statistically distinguish each FI mean absolute deviation (MAD) has been computed. Distinct values of MAD for individual FIs show that each image contributes uniquely to the descriptor design.

The novelty of the proposed descriptor over LVP is in a sense that the proposed descriptor identifies discriminating relationship between the local neighborhoods of the higher order derivatives of the original image with two distance parameters $D$ and $R$ and six different pairs of derivative directions. The novelty in the concept and advantages of the proposed LGHP are summarized as follows.

1) LDP identifies the discriminating relationship within the directional derivative space at a radius $R = 1$. LVP extends LDP and explores relationships existing in the inter-directional derivative space using directional derivative pairs $(0°, 45°)$, $(45°, 90°)$, $(90°, 135°)$, and $(135°, 180°)$. Even though LVP is computed at different radii, descriptor ignores neighboring pixels at smaller radii while computing it for larger radii (e.g., the descriptor ignores neighboring pixels at $R = 1$ and $R = 2$ while computing LVP for $R = 3$. The proposed LGHP overcomes the drawbacks of LDP and LVP by computing relationships existing in directional derivative pairs $(0°, 45°)$, $(0°, 90°)$, $(0°, 135°)$, $(45°, 90°)$, $(45°, 135°)$, and $(90°, 135°)$. LGHP is computed at

$R = 3$ without ignoring the fact that neighboring pixels at $R = 1$ and $R = 2$ also contribute to the discriminating information captured using the encoding function.

2) LGHP is able to extract the features at varying locality of the facial image. As the size of the locality ($R$) is increased, the discriminating power of the proposed descriptor is also increased. This shows that LDP and LVP ignore important relationships while computing the micropatterns.
3) Unlike LGHP, LTrP computes derivatives in only two directions ($0°, 90°$). Each pixel is assigned a decimal value between one and four based on the directions computed from the derivatives. LTrP computes a pattern of $13 \times 8$ b, whereas LGHP ($R = 1$) computes a pattern of $6 \times 9$ b.
4) Encoding functions of LDP, LVP, LTrP, and LGHP are vastly different from each other. LVP ignores the center pixel in the directional derivative space while computing the micropatterns, which is included in the micropatterns computed using LGHP.
5) There is a significant conceptual difference between proposed LGHP and LDGP. As LDGP [34] takes only the reference pixels of different derivatives into account to generate a 6-b pattern and ignores the local neighborhood, the micropattern computed by LDGP cannot be extended across radial widths. Hence, the length of LDGP is 6 b. In LGHP, a local window of radius $R$ has been used to compare the reference pixel and neighborhood pixels across different derivatives. A 9-b pattern is extracted by LGHP for each of the six pairs of different derivatives. Therefore, LGHP ($R = 1$) is $6 \times 9$-b long.
6) LDGP is a block-based method where the FI (decimal equivalents of 6-b patterns) is divided into $Z \times Z$ sub-blocks to compute the histogram. Therefore, the length of the histogram computed by LDGP is dependent on image size $M \times N$ (length of histogram $= (M/Z) \times (N/Z) \times$ number of bins). However, LGHP computes the histogram on the entire FI. Hence, the length of the histogram computed using LGHP is independent of size of the image (length of histogram = number of bins).
7) The proposed descriptor works even better in unconstrained environment as shown in the performance analysis.

## III. Performance Measures

The performance of the proposed LGHP has been analyzed with respect to retrieval and recognition accuracies. Performance measures used to evaluate retrieval accuracy are average precision of retrieval (APR) and average retrieval rate (ARR). Precision is computed as

$$P_r(I_q, n) = \frac{1}{n} \sum_{i=1}^{|\mathrm{DS}|} \Delta(\omega(I_q), \omega(I_i), \tau(I_q, I_i), n) | I_i \neq I_q \tag{8}$$

where $n$ is the number of images retrieved, $I_q$ is the query image, |DS| is the size of the data set, $\omega(.)$ returns the class of an image, and $\tau(I_q, I_i)$ is the rank of the $i$th image with respect to the query image $I_q$. Image rank is computed using similarity measure $S_{L1}(.,.)$ between the $i$th image in the data set and the query image. $\Delta(.)$ is a binary function defined as

$$\Delta(\omega(I_q), \omega(I_i), \tau(I_q, I_i), n) = \begin{cases} 1, & \omega(I_q) = \omega(I_i) \text{ and } \tau(I_q, I_i) \leq n \\ 0, & \text{else.} \end{cases} \tag{9}$$

Average precision per class is computed as

$$\mathrm{AP}(C_i, n) = \frac{1}{|C_i|} \sum_{q=1}^{|C_i|} P_r(I_q, n) \tag{10}$$

where $C_i$ denotes the $i$th class in the data set and $|C_i|$ denotes the number of images in the $i$th class. APR is calculated over the entire data set of $N_c$ distinct classes as

$$\mathrm{APR}(N_c) = \frac{1}{N_c} \sum_{i=1}^{N_c} AP(C_i, n). \tag{11}$$

Recall is defined as

$$R_e(I_q, C_i, n) = \frac{1}{|C_i|} \sum_{i=1}^{|\mathrm{DS}|} \Delta(\omega(I_q), \omega(I_i), \tau(I_q, I_i), n) | I_i \neq I_q. \tag{12}$$

Average recall per class and ARR over the entire data set are calculated as

$$\mathrm{AR}_e(C_i) = \frac{1}{|C_i|} \sum_{q=1}^{|C_i|} R_e(I_q, C_i) \tag{13}$$

$$\mathrm{ARR}(N_c) = \frac{1}{N_c} \sum_{i=1}^{N_c} \mathrm{AR}_e(C_i). \tag{14}$$

Recognition rate $\gamma$ is used to measure the recognition accuracy of the proposed LGHP. Recognition rate is defined as

$$\gamma = \left( \frac{1}{|\mathrm{DS}|} \sum_{\mathrm{probe}=1}^{|\mathrm{DS}|} \sum_{i=1}^{|\mathrm{DS}|} \hat{\Delta}(\omega(I_{\mathrm{probe}}), \omega(I_i), \tau(I_{\mathrm{probe}}, I_i)) \right) \times 100. \tag{15}$$

$\hat{\Delta}(.)$ computes the second best match for the probe image $I_{\mathrm{probe}}$ as

$$\hat{\Delta}(\omega(I_{\mathrm{probe}}), \omega(I_i), \tau(I_{\mathrm{probe}}, I_i)) = \begin{cases} 1, & \omega(I_{\mathrm{probe}}) = \omega(I_i) \text{ and } \tau(I_{\mathrm{probe}}, I_i) = 2 \\ 0, & \text{else.} \end{cases} \tag{16}$$

Recognition rate is computed over different disjoint sets of probe and gallery.

## IV. Performance Analysis

LGHP has been computed and analyzed on different publically available standard data sets, namely, Cropped Extended Yale B [23], [24], CMU-PIE [25], color-FERET [28], [29], LFW [30], and Gallagher database [35]. These databases have been used to test the robustness of the descriptor under severe illumination, pose, expression, and lighting conditions.

The performance of LGHP has been compared with state-of-the-art feature descriptors LDP, LTrP, MLBP, and LVP with respect to APR, ARR, and recognition rate. $L1$-norm distance measure is used as shown in (7) in all the experiments conducted on most challenging facial image databases. Precision recall space is used to show and compare the area under graph (AUG), as it presents a more detailed and clear picture of the performance of the descriptors for highly skewed data sets [26]. To compare the performance of the proposed method, different environmental parameters are considered, such as pose, illumination, different lighting conditions, and expressions. All the facial images have been resized to $64 \times 64$. Histograms of 256 bins are concatenated to generate the descriptors.

The experimental results are computed on a system with Intel Core i7 3.33-GHz CPU, 12-GB DDR3 RAM, and 64-b Windows 7 Professional operating system. MATLAB 7.11.0.584 is used to extract and match the features of different data sets.

### *A. Performance Analysis on Extended Yale B Database*

Extended Yale B is used in the experiments as it contains images with 64 different illumination variations of 38 subjects. It is a benchmark data set used to test the robustness of a descriptor against illumination variations. Cropped version of the Extended Yale B data set is used in the experiments.

APR and ARR of different descriptors for a maximum of eight retrieved images are shown in Fig. 5(a) and (b). LGHP (with $R = 3$ and 2) shows significant improvement with respect to ARR and APR over LDP, LTrP, MLBP, and LVP. Receiver Operating Characteristic (ROC) of the proposed descriptor for different $R$ in precision recall space has been shown in Fig. 5(c). AUG of the proposed LGHP is small compared with LDP and LVP.

Recognition rates of LGHP, LDP, and LVP are computed using (15) and (16) and has been shown in Fig. 5(d). Recognition rate of LGHP with $R = 3$ is 84.71%, whereas the recognition rates of LVP ($R = 3$), LDP, LTrP, and MLBP are 80.96%, 54.88%, 36.84%, and 26.72%, respectively. Average recognition rates shown in Fig. 5(e) are computed by randomly dividing the data set into disjoint probe and gallery sets of different sizes. Probe sets are prepared by randomly selecting 20%, 30%, 40%, 50%, and 60% images from the data sets, and the remaining images are used as corresponding gallery. Tenfold cross-validation is used to calculate the average recognition rates for each probe and gallery pair of the data sets. Average recognition rate is the average of recognition rates obtained in ten iterations of the experiment for a particular probe and gallery pair. Recognition rates of the descriptors are computed by modifying (15) and (16) as

$$\gamma = \left( \frac{1}{|\mathrm{PS}|} \sum_{\mathrm{probe}=1}^{|\mathrm{PS}|} \sum_{i=1}^{|\mathrm{GS}|} \hat{\Delta}(\omega(I_{\mathrm{probe}}), \omega(I_i), \tau(I_{\mathrm{probe}}, I_i)) \right) \times 100 \tag{17}$$

where |PS| and |GS| are the sizes of probe and gallery sets.

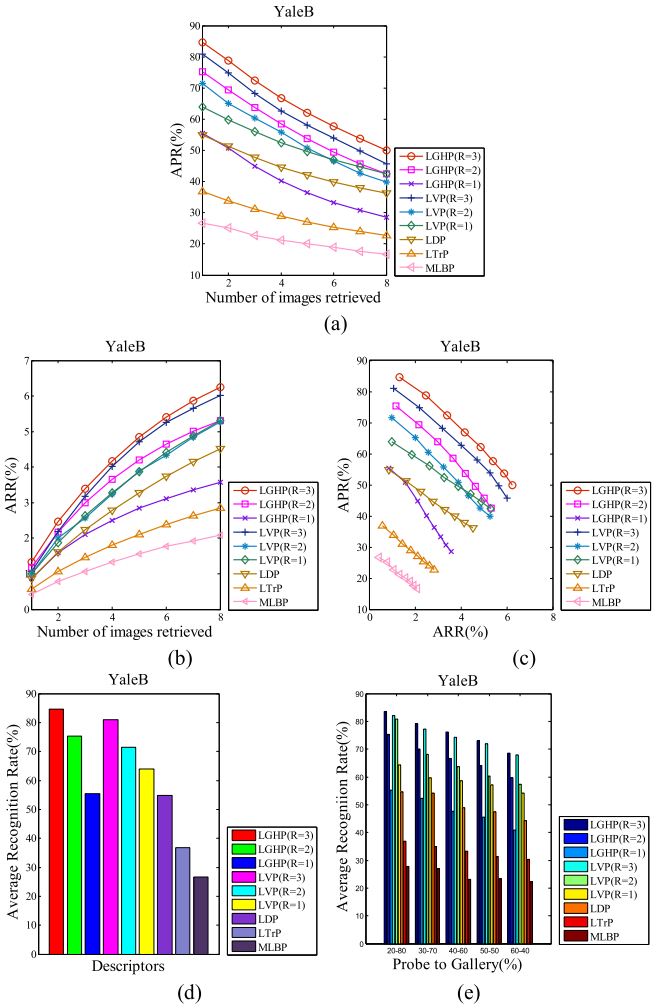


Fig. 5. Comparative results of LGHP, LVP, LTrP, MLBP, and LDP on Extended Yale B Face data sets. (a) APR. (b) ARR. (c) APR versus ARR. (d) Recognition rates of LDP, LTrP, MLBP, LVP, and LGHP with different values of $R$. (e) Comparative average recognition rates of LGHP, LVP, LTrP, MLBP, and LDP of tenfold cross-validation with different-sized probe and gallery.

$\hat{\Delta}(.)$ computes the first best match for the probe image $I_{\mathrm{probe}}$ as

$$\hat{\Delta}(\omega(I_{\mathrm{probe}}), \omega(I_i), \tau(I_{\mathrm{probe}}, I_i)) = \begin{cases} 1, & \omega(I_{\mathrm{probe}}) = \omega(I_i) \text{ and } \tau(I_{\mathrm{probe}}, I_i) = 1 \\ 0, & \text{else.} \end{cases} \tag{18}$$

LGHP shows significant improvement even for a very small sized gallery set. Average recognition rates of LGHP ($R = 3$), LVP ($R = 3$), LDP, LTrP, and MLBP for 60%–40% ratio of the probe and gallery set are 68.47%, 67.85%, 44.27%, 30.36%, and 22.27%, respectively.

### *B. Performance Analysis on CMU-PIE Database*

The database contains 38 556 images of 68 subjects taken under varying pose, illuminations, expression, and lighting. Expression data set contains 39 images with different expressions and poses for each subject. Illumination set of the database contains images with 24 different illuminations on 13 different poses of each individual; 72 images in 24 different

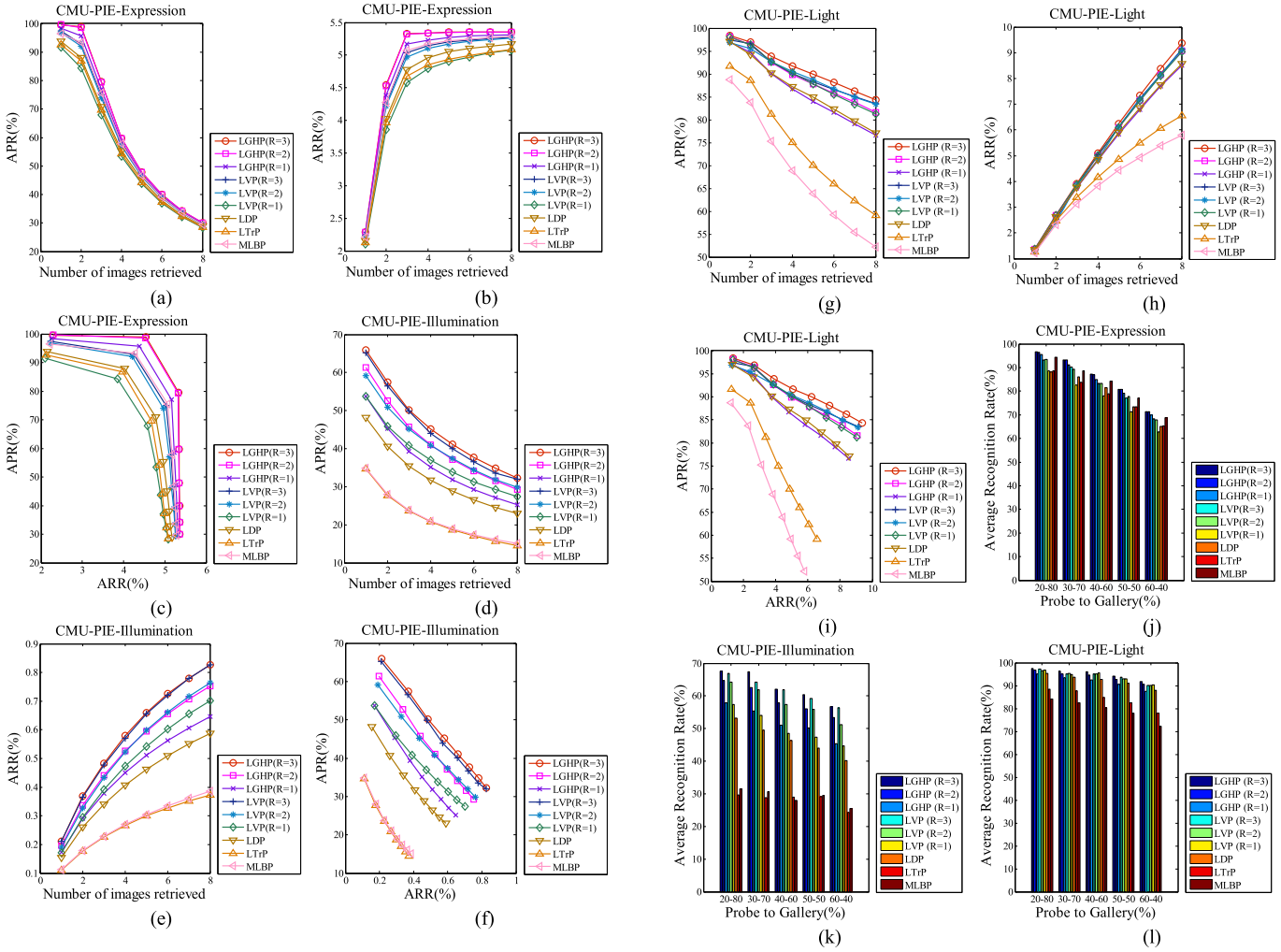


Fig. 6. Comparative results of LGHP, LVP, LTrP, MLBP, and LDP on CMU-PIE. Results on Expression data set (a) APR, (b) ARR, and (c) APR versus ARR. Results on Illumination data set (d) APR, (e) ARR, and (f) APR versus ARR. Results on Lights data set (g) APR, (h) ARR, and (i) APR versus ARR. Comparative average recognition rates of LGHP, LVP, LTrP, MLBP, and LDP of tenfold cross-validation with different-sized probe and gallery (j) Expression data set, (k) Illumination data set, and (l) Lights data set.

lighting conditions with three different pose are in the Lights data set of the database.

APR, ARR, and the ROC for different descriptors on Expression data set have been shown in Fig. 6(a)–(c). The results on Expression data set show that LGHP completely outperforms LVP, LTrP, MLBP, and LDP. LGHP at radius $R = 3$ performs better than LVP at radius $R = 3$, LDP, LTrP, and MLBP on illumination and Lights data sets as shown in Fig. 6(d)–(f) and (g)–(i). Maximum average precision rates of LGHP at $R = 3$, LVP at $R = 3$, LDP, LTrP, and MLBP on Expression data set are 99.59%, 97.39%, 93.83%, 92.69%, and 96.38%, respectively. Maximum average precision rates of LGHP at $R = 3$, LVP at $R = 3$, LDP, LTrP, and MLBP on Illumination data set are 65.97%, 65.22%, 48.18%, 34.63%, and 34.82%, respectively. Similarly, maximum average precision rates of LGHP at $R = 3$, LVP at $R = 3$, LDP, LTrP, and MLBP on Lights data set are 98.42%, 97.40%, 97.03%, 91.66%, and 88.76%, respectively.

Average recognition rates of descriptors on different data sets shown in Fig. 6(j)–(l) are computed on the same experimental setup as explained in Section IV-A. Probe and gallery sets are prepared as elaborated in Section IV-A. Average recognition rates of LGHP at $R = 3$, LVP at $R = 3$, LDP, LTrP, and MLBP on Expression data set for 20% probe size are 96.68%, 93.20%, 88.22%, 88.55%, and 94.36%, respectively. Average recognition rates of LGHP at $R = 3$, LVP at $R = 3$, LDP, LTrP, and MLBP on Illumination data set for 20% probe size are 68.14%, 67.72%, 53.25%, 28.86%, and 31.17%, respectively. Similarly, average recognition rates of LGHP at $R = 3$, LVP at $R = 3$, LDP, LTrP, and MLBP on Lights data set for 20% probe size are 97.44%, 97.28%, 95.40%, 88.35%, and 84.06%, respectively. It can be concluded from the above results that the proposed descriptor performs better than the state-of-the-art descriptors under varying expression, illumination, and lighting conditions.

### C. Performance Analysis on Color-FERET Database

"Portions of the research in this paper use the FERET database of facial images collected under the FERET program, sponsored by the DOD Counterdrug Technology Development Program Office." Color-FERET database is one of the most challenging facial image databases with severe variations in

TABLE I
DESCRIPTION OF DIFFERENT POSES IN IMAGES OF COLOR-FERET DATABASE

| Pose Name | Description |
|---|---|
| fa | regular frontal image |
| fb | alternative frontal image, taken shortly after the corresponding fa image |
| pl | profile left |
| hl | half left - head turned about 67.5 degrees left |
| ql | quarter left - head turned about 22.5 degrees left |
| pr | profile right |
| hr | half right - head turned about 67.5 degrees right |
| qr | quarter right - head turned about 22.5 degrees right |
| ra | random image - head turned about 45 degree left |
| rb | random image - head turned about 15 degree left |
| rc | random image - head turned about 15 degree right |
| rd | random image - head turned about 45 degree right |
| re | random image - head turned about 75 degree right |

pose and expression. The color-FERET database contains 11 338 facial images of 994 individuals at different angles. There are 13 different poses used in the images of the database [28], [29]. A brief description of pose has been given in Table I. The gallery used in the experiments contains images of those individuals having 20 or more images in the database. APR and ARR are shown for different descriptors along with their minimized version (computed using uniform two pattern minimization). Minimized versions of LGHP at $R = 3$, LVP at $R = 3$, LDP, LTrP, and MLBP are named LGHPU2, LVPU2, LDPU2, LTrPU2, and MLBPU2, respectively. Uniform 2 (U2) patterns are those binary patterns that have at most two transitions from 0 to 1 or vice versa [16]. APR and ARR shown in Fig. 7(a) and (b) illustrate that LGHP at different radii completely outperforms LVP, LDP, LTrP, and MLBP. Even the minimized version of LGHP performs better than LVP at $R = 3$, LVPU2 at $R = 3$, LDP, LTrP, MLBP, LDPU2, LTrPU2, and MLBPU2. Maximum APR of LGHP is almost 90% at $R = 3$ and $R = 2$. The APR of MLBP, which is nearest to the maximum APR of LGHP, is 80%. Maximum ARR of LGHP is 16.77%, which is more than the 16.55% ARR of its closest counterpart.

AUG for LGHP at $R = 3$ and $R = 2$ shown in Fig. 7(c) is less than the AUG of all other descriptors and their variations. Recognition rates shown in Fig. 7(d) are computed using experimental settings as elaborated in Section IV-A. Maximum recognition rate of LGHP is more than 80%, which is at least 7% more than the recognition rate of its nearest counterpart. Hence, the proposed descriptor shows significant improvement in recognition and retrieval rates compared with LDP and LVP and their other variations.

### *D. Performance Analysis on LFW Database*

There are 13 233 color facial images of 5749 individuals; 1680 individuals have two or more images, and the rest

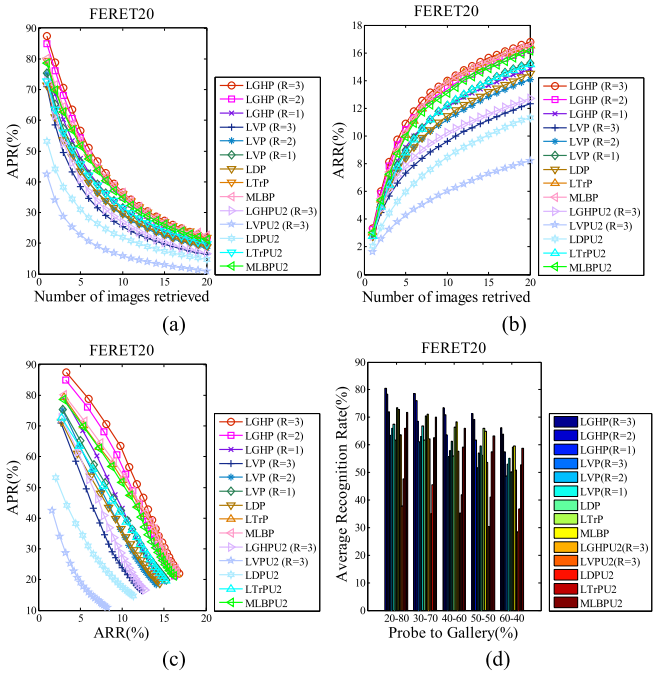


Fig. 7. Comparative results of LGHP, LVP, LTrP, MLBP, and LDP on FERET Face data sets along with U2 minimization. (a) APR. (b) ARR. (c) APR versus ARR. (d) Comparative average recognition rates of LGHP, LVP, LTrP, MLBP, and LDP of tenfold cross-validation with different-sized probe and gallery.

of the individuals have only one image. As the problem of unconstrained face recognition is one of the most general and fundamental face recognition problems, we test the proposed descriptors on the most challenging facial image database LFW. Experimental analysis is done on a subset of LFW database. In the experiments, images of those individuals having at least 20 images are taken. Similar settings are used for experiments as used in Section IV-A. APR, ARR, and recognition rates for LGHP, LVP, LDP, LTrP, MLBP, and their U2 versions have been shown in Fig. 8. Maximum APR of LGHP is approximately 2% more than the maximum APR of its immediate counterpart, whereas maximum ARR is approximately 0.5% more than the maximum ARR of its immediate counterpart.

AUG shown in Fig. 8(c) demonstrates that the proposed descriptor maintains its superiority over other related descriptors with increasing retrieval (recall) rates. LGHP performs better than most of the other descriptors and their variations even when U2 minimization has been used. Maximum recognition rate of LGHP shown in Fig. 8(d) is 3% more than the maximum recognition rate achieved by other descriptors.

A 2D Gabor filter [22] shown in (19) has been used to compute the Gabor response with two scales and two directions $(0°, 90°)$. A Gabor filter is a Gaussian response over sinusoid with frequency $f$ and standard deviations $\sigma_s$ and $\sigma_t$ [22]

$$\emptyset(s,t) = \frac{1}{2\pi\,\sigma_s\sigma_t} e^{\left[-\left(1/2\right)\left(s^2/\sigma_s^2 + t^2/\sigma_t^2\right) + 2\pi \mathrm{i} f s\right]}. \tag{19}$$

Gabor features are computed using different descriptors over Gabor response obtained by convolving Gabor filter with the original image.

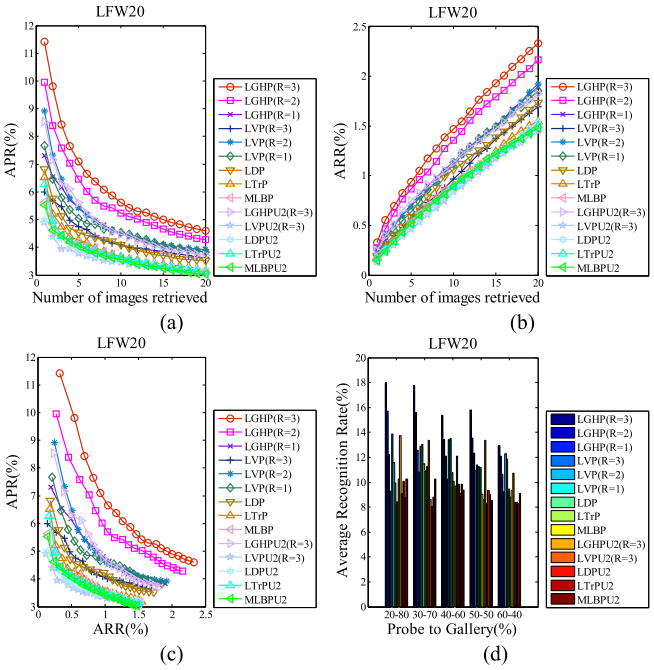


Fig. 8. Comparative results of LGHP, LVP, LTrP, MLBP, and LDP on LFW Face data set along with U2 minimization. (a) APR. (b) ARR. (c) APR versus ARR. (d) Comparative average recognition rates of LGHP, LVP, LTrP, MLBP, and LDP of tenfold cross-validation with different-sized probe and gallery.

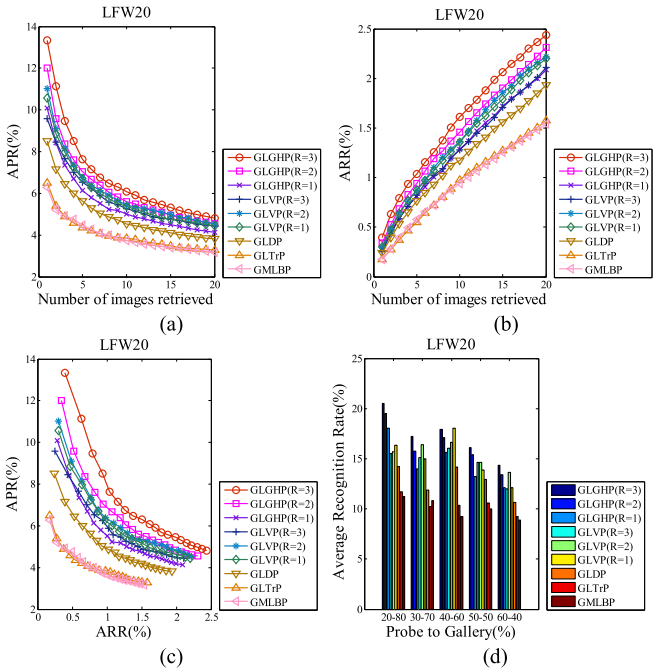


Fig. 9. Comparative results of Gabor features GLGHP, GLVP, GLTrP, GMLBP, and GLDP corresponding to LGHP, LVP, LTrP, MLBP, and LDP on LFW Face data sets. (a) APR. (b) ARR. (c) APR versus ARR. (d) Comparative average recognition rates of GLGHP, GLVP, GLTrP, GMLBP, and GLDP of tenfold cross-validation with different-sized probe and gallery.

APR, ARR, and recognition rates of Gabor features corresponding to different descriptors, namely, GLGHP (corresponding to LGHP), GLVP (corresponding to LGVP), and GLDP (corresponding to LGVP) are shown in Fig. 9. Gabor features improve the performance of all the descriptors even more so the proposed descriptor LGHP. Maximum APR of GLGHP shown in Fig. 9(a) is approximately 14%, which is 3% more than the APR of the immediate best counterpart. The maximum recognition rate of GLGHP shown in Fig. 9(d) is approximately 20%, which is 4% more than the maximum recognition rate of GLVP.

### *E. Performance Analysis Under Noise on LFW and FERET Databases*

The performance of the proposed descriptor has been analyzed and compared with LVP, LDP, LTrP, and MLBP under additive white Gaussian noise (AWGN) with zero mean and 0.05 variance. Features of the facial images in LFW and FERET databases are computed after adding the AWGN. Feature of each noisy image is taken as the query. APR and ARR values are computed for each queried feature. Computed APR and ARR values are shown in Fig. 10. Fig. 10(a) and (b) shows the results computed on LFW. The proposed descriptor shows approximately 3.5% and 1.5% improvement in APR and ARR, respectively, over its nearest counterpart MLBP under noise. The results computed over FERET database under noise are shown in Fig. 10(c) and (d). LGHP shows significant improvement of about 20% in APR and 3% in ARR as shown in Fig. 10(c) and (d). As the results show, the proposed descriptor is robust enough to achieve better retrieval accuracy under noisy conditions.

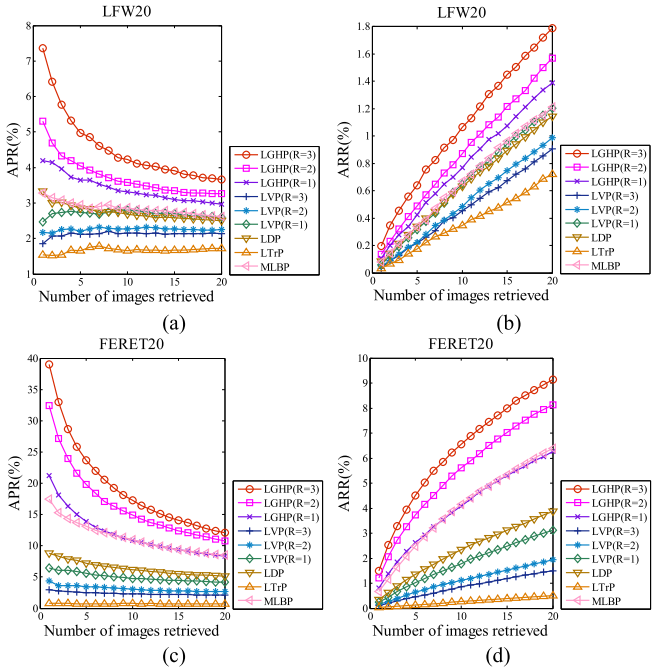


Fig. 10. Comparative results of the descriptors under AWGN with zero mean and 0.05 variance. (a) APR on LFW. (b) ARR on LFW. (c) APR on FERET. (d) ARR on FERET.

### *F. Performance Analysis on Gallagher Database*

Gallagher database [35] contains facial images with real facial expressions taken in the real-world conditions (uncontrolled light, illumination, and expression variations). The database is a very challenging database containing 931 facial

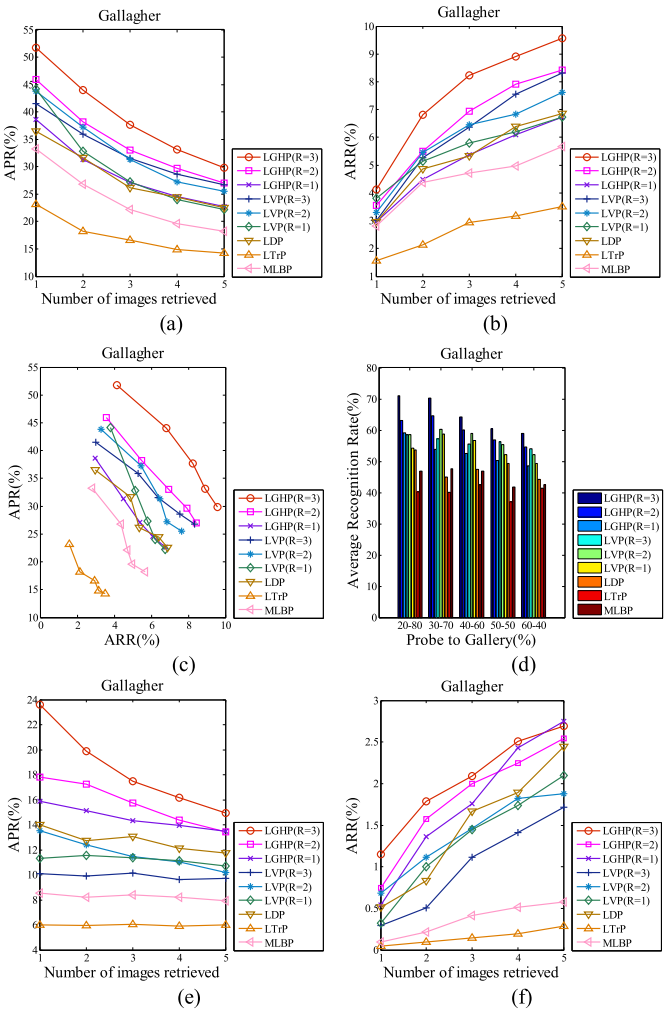


Fig. 11. Comparative results of LGHP, LVP, LTrP, MLBP, and LDP on Gallagher Face data set. (a) APR. (b) ARR. (c) APR versus ARR. (d) Comparative average recognition rates of LGHP, LVP, LTrP, MLBP, and LDP of tenfold cross-validation with different-sized probe and gallery. Comparative results of the descriptors under AWGN with zero mean and 0.05 variance (e) APR and (f) ARR.

images of 31 individuals. In the experiments, 896 images of those individuals having five or more images have been used.

Average precision and recall computed on the database for different descriptors are shown in Fig. 11(a) and (b). Average precision rates of LGHP ($R = 3$), LVP ($R = 3$), LDP, LTrP, and MLBP are 51.75%, 41.50%, 36.51%, 23.16%, and 30.20%, respectively. Maximum average recall rates of LGHP ($R = 3$), LVP ($R = 3$), LDP, LTrP, and MLBP are 4.13%, 3.01%, 2.96%, 1.56%, and 2.80%, respectively. Average recognition rates of LGHP ($R = 3$), LVP ($R = 3$), LDP, LTrP, and MLBP for 20%–80% probe to gallery ratio are 70.94%, 58.65%, 53.63%, 40.22%, and 46.92%, respectively.

Average precision and recall rates are computed over noisy images using AWGN with zero mean and 0.05 variance. Fig. 11(e) and (f) shows the retrieval accuracy (average precision and recall) under AWGN. Precision and recall rates of LGHP at $R = 1 - 3$ show significant improvement over LVP, LDP, LTrP, and MLBP.

Maximum precision rates of LGHP ($R = 3$), LVP ($R = 3$), LDP, LTrP, and MLBP are 23.61%, 10.11%, 13.97%, 5.97%, and 8.54%, respectively. Similarly, maximum recall rates of LGHP ($R = 3$), LVP ($R = 3$), LDP, LTrP, and MLBP are 2.68%, 1.71%, 2.44%, 0.28%, and 0.57%. It is evident from the precision and recall rates that the proposed descriptor works even better under noisy conditions.

## V. Conclusion

In this paper, a local descriptor has been proposed for face recognition and facial image retrieval. The performance of the proposed descriptor has been tested on benchmark databases and compared with the state-of-the-art descriptors to show the robustness of the proposed descriptor against pose, expression, illumination, and lighting variations. The experimental results demonstrate the superior performance over the existing state-of-the-art descriptors under huge variations in pose, expression, lighting, and illumination conditions. The proposed descriptor has the potential to be used in the real-world unconstrained facial recognition and retrieval application.

## References

[1] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," *Proc. IEEE*, vol. 83, no. 5, pp. 705–741, May 1995.

[2] A. Samal and P. A. Iyengar, "Automatic recognition and analysis of human faces and facial expressions: A survey," *Pattern Recognit.*, vol. 25, no. 1, pp. 65–77, 1992.

[3] Y. Kim and J. Aggarwal, "Determining object motion in a sequence of stereo images," *IEEE J. Robot. Autom.*, vol. RA-3, no. 6, pp. 599–614, Dec. 1987.

[4] H. A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 1, pp. 23–38, Jan. 1998.

[5] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.

[6] A. M. Martínez and A. C. Kak, "PCA versus LDA," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 2, pp. 228–233, Feb. 2001.

[7] X. Xie and K.-M. Lam, "Gabor-based kernel PCA with doubly nonlinear mapping for face recognition with a single face image," *IEEE Trans. Image Process.*, vol. 15, no. 9, pp. 2481–2492, Sep. 2006.

[8] D. Zhang and Z.-H. Zhou, "Two-directional two-dimensional PCA for efficient face representation and recognition," *Neurocomputing*, vol. 69, nos. 1–3, pp. 224–239, 2005.

[9] H. Kong, L. Wang, E. K. Teoh, X. Li, J.-G. Wang, and R. Venkateswarlu, "Generalized 2D principal component analysis for face image representation and recognition," *Neural Netw.*, vol. 18, nos. 5–6, pp. 585–594, 2005.

[10] D. Zhang, Z.-H. Zhou, and S. Chen, "Diagonal principal component analysis for face recognition," *Pattern Recognit.*, vol. 39, no. 1, pp. 140–142, 2006.

[11] K. Etemad and R. Chellappa, "Discriminant analysis for recognition of human face images," *J. Opt. Soc. Amer. A*, vol. 14, no. 8, pp. 1724–1733, 1997.

[12] S. Noushath, G. H. Kumar, and P. Shivakumara, "$(2D)^2$LDA: An efficient approach for face recognition," *Pattern Recognit.*, vol. 39, no. 7, pp. 1396–1400, 2006.

[13] S. Noushath, G. H. Kumar, and P. Shivakumara, "Diagonal Fisher linear discriminant analysis for efficient face recognition," *Neurocomputing*, vol. 69, nos. 13–15, pp. 1711–1716, 2006.

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

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

[16] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.

[17] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, Dec. 2006.
[18] J. Y. Choi, Y. M. Ro, and K. N. Plataniotis, "Boosting color feature selection for color face recognition," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1425–1434, May 2011.
[19] J. Y. Choi, Y. M. Ro, and K. N. Plataniotis, "Color local texture features for color face recognition," *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 1366–1380, Mar. 2012.
[20] A. R. Rivera, J. R. Castillo, and O. K. Chae, "Local directional number pattern for face analysis: Face and expression recognition," *IEEE Trans. Image Process.*, vol. 22, no. 5, pp. 1740–1752, May 2013.
[21] B. Zhang, Y. Gao, S. Zhao, and J. Liu, "Local derivative pattern versus local binary pattern: Face recognition with high-order local pattern descriptor," *IEEE Trans. Image Process.*, vol. 19, no. 2, pp. 533–544, Feb. 2010.
[22] S. Murala, R. P. Maheshwari, and R. Balasubramanian, "Local tetra patterns: A new feature descriptor for content-based image retrieval," *IEEE Trans. Image Process.*, vol. 21, no. 5, pp. 2874–2886, May 2012.
[23] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 643–660, Jun. 2001.
[24] K.-C. Lee, J. Ho, and D. J. Kriegman, "Acquiring linear subspaces for face recognition under variable lighting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 5, pp. 684–698, May 2005.
[25] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression database," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 12, pp. 1615–1618, Dec. 2003.
[26] J. Davis and M. Goadrich, "The relationship between precision-recall and ROC curves," in *Proc. 23rd Int. Conf. Mach. Learn.*, 2006, pp. 233–240, doi: 10.1145/1143844.1143874.
[27] K.-C. Fan and T.-Y. Hung, "A novel local pattern descriptor—Local vector pattern in high-order derivative space for face recognition," *IEEE Trans. Image Process.*, vol. 23, no. 7, pp. 2877–2891, May 2014.
[28] P. J. Phillips, H. Wechsler, J. Huang, and P. J. Rauss, "The FERET database and evaluation procedure for face-recognition algorithms," *Image Vis. Comput.*, vol. 16, no. 5, pp. 295–306, 1998.
[29] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 10, pp. 1090–1104, Oct. 2000.
[30] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," Dept. Comput. Sci., Univ. Massachusetts, Amherst, MA, USA, Tech. Rep. 07-49, Oct. 2007.
[31] L. Zhang, R. Chu, S. Xiang, S. Liao, and S. Z. Li, "Face detection based on multi-block LBP representation," in *Advances in Biometrics* (Lecture Notes in Computer Science), vol. 4642. Berlin, Germany: Springer, 2007, pp. 11–18.
[32] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: Closing the gap to human-level performance in face verification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1701–1708.
[33] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in *Proc. Brit. Mach. Vis.*, 2015, vol. 1. no. 3, p. 6.
[34] S. Chakraborty, S. K. Singh, and P. Chakraborty, "Local directional gradient pattern: A local descriptor for face recognition," *Multimedia Tools Appl.*, pp. 1–16, Dec. 2015.
[35] A. C. Gallagher and T. Chen, "Clothing cosegmentation for recognizing people," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.

**Soumendu Chakraborty** (S'15) received the B.Eng. degree in information technology from University Institute of Technology, University of Burdwan, Bardhaman, India, and the M.Tech. degree in computer science and engineering from GLA University, Mathura, India. He is currently working toward the Ph.D. degree from Indian Institute of Information Technology at Allahabad, Allahabad, India.

He has over eleven years of experience in teaching and research. His research interests include image processing, biometric systems, pattern recognition, and image steganography.

**Satish Kumar Singh** (M'11–SM'14) received the B.Tech., M.Tech., and Ph.D. degrees in 2003, 2005, and 2010, respectively.

He is currently an Assistant Professor with Indian Institute of Information Technology at Allahabad, Allahabad, India. He has over 11 years of experience in academic and research institutions. He has authored over 25 publications in reputed international journals and conference proceedings. His research interests include digital image processing, pattern recognition, multimedia data indexing and retrieval, and watermarking and biometrics.

Dr. Singh is a member of various professional societies, such as the Institution of Electronics and Telecommunication Engineers. He is also an Editorial Board Member and a Reviewer for various international journals.

**Pavan Chakraborty** (M'12) received the M.Sc. degree from IIT Kanpur, Kanpur, India, in 1993 and the Ph.D. degree from Indian Institute of Astrophysics, Bangalore, India, in 2001.

He is as an Associate Professor with the Department of Information Technology, Indian Institute of Information Technology at Allahabad, Allahabad, India. His research interests include artificial intelligence, human gait, computer vision, robotics, and instruments.