LMZMPM: Local Modified Zernike Moment Per-unit Mass for Robust Human Face Recognition

Arindam Kar[1], Sourav Pramanik[2], *Student Member, IEEE*, Arghya Chakraborty[1], Debotosh Bhattacharjee,[2] *Senior Member, IEEE*, Edmond S. L. Ho[3], and Hubert P. H. Shum[3,*], *Senior Member, IEEE*

***Abstract*—In this work, we proposed a novel method, called Local Modified Zernike Moment per unit Mass (LMZMPM), for face recognition, which is invariant to illumination, scaling, noise, in-plane rotation, and translation, along with other orthogonal and inherent properties of the Zernike Moments (ZMs). The proposed LMZMPM is computed for each pixel in a neighborhood of size 3×3, and then considers the complex tuple that contains both the phase and magnitude coefficients of LMZMPM as the extracted features. As it contains both the phase and the magnitude components of the complex feature, it has more information about the image and thus preserves both the edge and structural information. We also propose a hybrid similarity measure, combining the Jaccard Similarity with the L1 distance, and applied to the extracted feature set for classification. The feasibility of the proposed LMZMPM technique on varying illumination has been evaluated on the CMU-PIE and the extended Yale B databases with an average Rank-1 Recognition (R1R) accuracy of 99.8% and 98.66% respectively. To assess the reliability of the method with variations in noise, rotation, scaling, and translation, we evaluate it on the AR database and obtain an average R1R higher than that of recent state-of-the-art methods. The proposed method shows a very high recognition rate on Heterogeneous Face Recognition as well, with 100% on CUFS, and 98.80% on CASIA-HFB.**

***Index Terms*—LMZMPM, Zernike Moments, face recognition, heterogeneous face recognition,similarity measure.**

I. Introduction

BESIDES being proficient in identifying faces of individuals, human beings are also capable of relating such faces with the right individual. Inherently, a human being can remember hundreds or even thousands of faces and can recognize a face with the state of different facial features and facial contours in different perspective variations, illuminations, ages, etc. As a consequence, any approach for automatic recognition of faces could be designed wherein the geometrical positions of the different facial features are measured first, and then the details of each feature are used for further matching [1], [2]. However, the problem becomes very challenging when the other barriers, like noise, disturbance, scaling, occlusion, rotational changes, come together. A serious interest in the domain of face recognition is, therefore, to develop an automatic face recognition system that is robust to variance in the pose, illumination, expression, and occlusion. Moreover, the system has to take into consideration the computational cost for real-time applications.

The growth of face recognition techniques over the past three decades has been substantial, as benchmarked by the National Institute of Standards and Technology (NIST) [3]. Earlier approaches proposed by the researchers are mainly application-oriented. Some of the popular approaches use the Eigenfaces [4], Fisherfaces [5], Laplacian faces [6], convolutional neural networks [7], [8] and others [9], [10] for recognition purpose. These algorithms work fine under a controlled environment, but most of them fail to achieve their reputations when the situation involves variation in illumination, pose, facial expression, aging, partial occlusion, etc. Moreover, they highly under-perform when there is a combination of such variations [11]. Such an observation has steered the researcher to work upon and develop more classy models that are tolerant to these variations.

To cope with these types of real-world problems, in this work, we propose a local moment-based approach for face recognition, which is invariant to changes in illumination, rotation, translation, and scale. Also, it is robust to noise and is computationally efficient. Moments provide useful series expansions for the representation of object shapes, while shape plays a significant role in face recognition, shape recognition, remote sensing, and classification. Inspired from the work of local characteristic features in G-Face [12] and the superior output results of Zernike moments(ZMs)-based methods [13], [14], and Legendre moments (LM) [15] in face and texture representation, we have proposed here a method called Local Modified Zernike Moments Per unit Mass (LMZMPM) for face recognition under uncontrolled environment that exploits the illumination invariant property of the resultant moment together with the invariant properties of orthogonal radial polynomials in a 3×3 neighborhood, which acts as a local characteristic and shape descriptor. This descriptor reveals the surface characteristics together with the edge (shape) of that neighborhood. The proposed LMZMPM method is not only illumination-invariant but is also rotation, shift, scale, noise and blur invariant. The key features of the proposed LMZMPM method are summarized in the following:

- The proposed LMZMPM improves the limitations of ZMs by making it invariant to scaling and illumination and at the same time, retaining the properties of rotation,

*Corresponding Author: H. P. H. Shum.

This work was supported in part by the Royal Society (Ref: IES\R2\181024 and IES\R1\191147).

A. Kar and A. Chakraborty are with Indian Statistical Institute, 203 B. T. Road, Kolkata-700108, India. (Emails: kgparindamkar@gmail.com and chakraborty.arghya9@gmail.com)

S. Pramanik and D. Bhattacharjee are with the Department of Computer Science and Engineering, Jadavpur University, Kolkata-700032, India. (E-mails: srv.pramanik03327@gmail.com, and debotosh@ieee.org.)

E. S. L. Ho and H. P. H. Shum are with the Department of Computer and Information Sciences, Northumbria University, Newcastle upon Tyne, UK. (E-mail: e.ho@northumbria.ac.uk and hubert.shum@northumbria.ac.uk.)

translation and blur invariant as well as robustness to noise.
- The inclusion of the angle between the center of mass of a neighborhood with respect to the center pixel, along with the orthogonality property of the ZMs not only makes the feature representations efficient but also keeps the relative positions of key facial features intact.
- It preserves maximum information of the image with the use of both magnitude and phase of the complex moment of LMZMPM, making it suitable for recognition problems.
- Our approach is based on a modified orthogonal moment using Zernike polynomials, which decreases intra-class variations and increases inter-class variations, resulting in a unique method favorable for face recognition.

II. LITERATURE REVIEW

The concept of Invariant Moment came from the concept of Algebraic Invariants. However, the use of moment invariants in applications of pattern recognition and image analysis was first introduced by Hu [16]. He has derived seven-moment invariants by employing the general theory of algebraic invariants under the affine transformation of the image plane. However, the approach was computationally expensive. Again, in the extreme case of rotationally symmetric objects, each of Hu's seven-moment invariants are zero. Also, the recovery of an image from the geometric moments is strongly ill-posed. To overcome the shortcomings of the geometric moments, Teague [17] was the first to introduce orthogonal moments in image analysis based on the Legendre, Gegenbauer [18], [19] and Zernike polynomials [20], [21]. The magnitude of the orthogonal moments, i.e., Zernike moments (ZMs), is invariant to rotation, scaling, and translation, but its phase changes with rotation. So in most of the works on invariant moments, the magnitudes of the orthogonal moments are considered neglecting their phase, which results in loss of information. Over the past years, various ZM-based methods have been developed in the field of face recognition, which are discussed as follows.

The authors in [22] proposed an illumination invariant method for texture classification based on invariant moments. Li et. al. [23] proposed shape-based image retrieval using the phase coefficient ZMs. Authors in [24] have proposed a feature descriptor by combining wavelet transforms with ZMs for face recognition. An adaptively weighted patch pseudo ZM-based method is proposed in [25] for face recognition. A combination of Gabor filters-based texture features and ZMs is used in [26] for face recognition. Fathi et. al. [27] introduced a multi-scale and rotation invariant global-Gabor-Zernike feature descriptor by applying the ZM to the Gabor filters outputs for face recognition. Recently, Dhekane et. al. [15] proposed an Illumination and Expression Invariant Face Recognition approach, which is mainly a fusion of Legendre moments (LM) and Uniform Local Binary Patterns (uLBP). Besides, Sariyanidi et. al. [13] have proposed local ZM (LZM) for the extraction of local features from face images and have shown that LZM features are more robust than the global ZMs. Authors in [28], have used Gauss scale space and LZM [13] for recognition of faces in low-resolution images. Further, a variation of LZM, called local Zernike XOR patterns (LZXP), is proposed in [29] for robust face recognition, where LZXPs are derived by encoding the LZM images using the XOR operator. Kahraman et al. [30] have used reduced LZM for face pair matching problem. Recently, Barasan et al. [14] proposed an efficient multi-scale scheme using local Zernike Moments (MSLZM) for Face Recognition. Also, Kaur et al. [31] proposed a method for Iris Recognition using ZMs and Polar harmonic transforms. In [32], a logarithmic total variation (LTV) model is proposed for the recognition of faces under varying illumination. In [33], the authors suggested another method based on Weber's law, called "Weber-face (W-face)" for illumination-invariant face representation. Rivera et al. [34] proposed a local directional number patterns (LDN) descriptor for the extraction of illumination-invariant features from face images. Recently, Kar et. al. [35] proposed a local Centre of Mass (LCMF) approach for Face recognition under illumination variation. The authors in [36] combined fractal analysis and a logarithmic function to extract logarithmic fractal analysis (LFA) features from face images for illumination-invariant face recognition. Lai et al. [37] proposed a multi-scale logarithm difference edge maps (MSLDE), to obtain illumination-invariant features. Fan et. al. [38] proposed another method for face recognition under varying illumination conditions that integrates modified homomorphic filtering (HF) and Histogram equalization (HQ) methods.

Existing methods for heterogeneous face recognition can be broadly categorized into three classes [39], [40]: common subspace projection-based techniques, synthesis-based techniques, and modality invariant feature-based techniques. In the first category, the heterogeneity of face images of different modalities is minimized by projecting them into a latent subspace [41], [42]. Besides, in synthesis-based techniques, face images of different modalities are transformed into the same modality. For example, Wang et. al. [43] proposed a sketch-photo synthesis technique by using a local patch-based multiscale Markov random field model. In [44], authors have used sparse neighbour selection (SNS) and sparse-representation-based enhancement (SRE) techniques for automatic sketch-photo synthesis. Further, Wang et al. [45] proposed a transductive face sketch-photo (TFSP) synthesis technique. In modality invariant feature-based techniques, face images of different modalities are encoded with local handcrafted features, which are then used for recognition. For example, Klare et al. [46] employed the scale-invariant feature transform (SIFT) and multi-scale local binary pattern (MLBP) features for forensic sketch recognition. In [47], authors have proposed a non-linear kernel prototype random subspace (KP-RS) learning for heterogeneous face recognition. Zhu et al. [48] proposed a transductive heterogeneous face matching (THFM) method for NIR-VIS face images. Hussain et. al. [49] proposed a novel image descriptor, named local quantized pattern for intensity images (I-LQP), for the representation of face images. Recently, an attractive method is proposed by Peng et. al. [40] for HFR based on sparse graphical representation-based discriminant analysis (SGR-DA). Roy et.al. have proposed

some interesting methods for HFR: MLGFP [50], zigzag pattern of local extremum logarithm difference (ZZPLELD) [51], quaternary pattern of local maximum quotient (QPLMQ) [39], local wavelet energy mesh pattern (LWEMeP) [39] and Local-Gravity-Face (LG-face) [11]. In [39], [40], the authors have presented some useful reviews of HFR. Hence, interested readers can refer these papers for further reading. Though the performances of these algorithms were tested on face images collected under well-controlled situations, they fail to handle the major challenges of face recognition problems. So new methods are needed to handle situations where there is not only variation in modality but also in pose, illumination, expression, and the other obstacles.

Over the past few years, several deep learning-based methods for both the standard and heterogeneous face recognition with good performances were developed and reported in the literature [8], [52]–[55]. However, the main drawback of such methods is the reliance on the training data and the computational cost needed for learning the network. Typically, it requires a large number of training samples and high training time. Also, due to the insufficient theoretical guidance, the problem of designing any deep learning model might be compounded.

The paper is organized as follows. In Section 3, the proposed LMZMPM method has been described in detail. The approach followed in Similarity Measurement is presented in Section 4. Experimental Set-up and the results on various databases are presented in Section 5, and finally, the paper is concluded in Section 6.

III. THE PROPOSED METHOD

In this section, we start with the basics of ZMs and orthogonal polynomials and then derived the proposed Local Modified Zernike Moment per unit mass (LMZMPM) method. Then, we move ahead with the theoretical proofs of the invariant properties of the proposed LMZMPM method. At the same time, we discuss how the feature vector is created from the complex moment of LMZMPM by keeping intact both the magnitude and phase coefficient of the complex feature. To facilitate the understanding of the reader, all the frequently used symbols are listed and described in Table-I.

A. The Basics of Zernike Moments

The Zernike polynomials [18] are a set of orthogonal polynomials defined over the unit disk $D = \{(x, y), x^2 + y^2 = 1\}$ as:

$$V_{pq}(\rho, \theta) = R_{pq}(\rho)e^{iq\theta} \tag{1}$$

where $p \geq 0$ and $0 \leq q \leq p$ signify the order and repetition of Zernike Moments (ZMs), respectively, with $p - |q|$ is even; $\rho = \sqrt{x^2 + y^2}$; $\theta = \tan^{-1}(\frac{y}{x})$; and $R_{pq}(\rho)$ is the radial polynomial and can be defined as:

$$R_{pq}(\rho) = \begin{cases} \sum_{l=0}^{(\frac{p-q}{2})} \frac{(-1)^l (p-q)! \rho^{p-2l}}{l![\frac{(p+q)}{2}-l]![\frac{(p-q)}{2}-l]!}, & for\ p-q\ even \\ 0, & for\ p-q\ odd \end{cases} \tag{2}$$

Now, the conjugate of $V_{pq}(\rho, \theta)$ can be defined as

$$[V_{pq}(\rho, \theta)]^* = R_{pq}(\rho)e^{-iq\theta} \tag{3}$$

Table I
FREQUENTLY USED SYMBOLS IN SECTION-III.

symbols	Description
V_{pq}	Zernike polynomial; p and q signify the order and repetition of Zernike Moments (ZMs)
R_{pq}	Radial polynomial
Z_{pq}	Zernike Moment
I	Face image
w	Neighborhood of size $k \times k$ defined in the image I
θ	Angle between pixels in w
V'_{pq}	Modified radial orthogonal polynomial
θ_{CoM}	Angle between the center pixel p_c and the Center-of-Mass (CoM) in w
Z'_{pq}	Local modified Zernike moment
C_{pq}	Local Modified Zernike Moment Per-unit Mass (LMZMPM)
R	Reflectance component
L	Luminance component
F	Complex LMZMPM feature vector
F_M, F_θ	Complex magnitude and phase feature vectors
F'_M, F'_θ	Standardized complex magnitude and phase feature vector

where $*$ denotes the conjugate. However, these polynomials must satisfy

$$\int\int [V_{pq}(\rho, \theta)][V_{p'q'}(\rho, \theta)]^* d\rho d\theta = \begin{cases} \frac{\pi}{p+1}, & for\ (p, q) = (p', q') \\ 0, & otherwise \end{cases} \tag{4}$$

Now, as for the image I, the ZM of order p with q repetition can be written as the projection of the image I onto the corresponding Zernike polynomial basis as:

$$Z_{pq} = \frac{(p+1)}{\pi} \int\int I(x, y)[V_{pq}(\rho, \theta)]^* dxdy; x^2 + y^2 \leq 1 \tag{5}$$

where $I(x, y)$ denotes the intensity of the image I at location (x, y). Now, for practical usage and easy computational work the integral in Eq. (5) is replaced by the summation and rewritten the Eq.(5) as:

$$Z_{pq} = \frac{(p+1)}{\pi} \sum_x \sum_y I(x, y)R_{pq}(\rho)e^{-iq\theta}; x^2 + y^2 \leq 1 \tag{6}$$

B. The LMZMPM Method

From the definition of ZM, it is seen that the ZMs are computed in the unit circle $D = \{(x, y), x^2 + y^2 = 1\}$. Hence, in this work, we have also defined our proposed Local Modified Zernike Moment Per-unit Mass (LMZMPM) in the neighborhood w of size $k \times k$ (here, $k = 3$) centered at pixel p_c. So, without loss of generality, it can be assumed that all the eight neighboring pixels $p_i (i = 1, 2, \ldots .8)$ of p_c in w lies on the circle and each of the consecutive neighboring pixels is sustaining the same $\theta = \frac{\pi}{4}$ angle, which is shown in Fig 1.

Now, at the very first step of the proposed LMZMPM method, we have defined a set of modified radial orthogonal polynomials as:

$$V'_{pq} = \frac{V_{pq}}{e^{iq\theta_{CoM}}} \tag{7}$$

where the term V_{pq} denotes the standard radial Zernike orthogonal polynomials [20] defined over a neighborhood of size 3×3, and θ_{CoM} is the angle between the center pixel p_c and the center-of-mass (CoM) in the neighborhood w.

From the theory of center-of-mass (CoM) [35], in a system of particles, the CoM is the point where all the mass of the

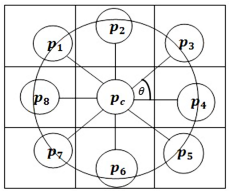

Figure 1. 3×3 neighborhood of pixels sustaining the same angle $\theta = \frac{\pi}{4}$.

system is concentrated. Let us assume that we have a system of n particles with masses $m_1, m_2, \ldots, m_n$ and position vectors $v_1, v_2 \ldots, v_n$. Now, more formally, the CoM of a specific mass distribution is a position vector $\overrightarrow{V}$ in the space, where the masses are uniformly spread around it. The position vector $\overrightarrow{V}$ is defined as:

$$\overrightarrow{V} = \frac{m_1 \overrightarrow{v}_1 + m_2 \overrightarrow{v}_2 + \cdots + m_n \overrightarrow{v}_n}{m_1 + m_2 + \cdots + m_n} \tag{8}$$

In the context of our problem, the pixel intensities in the neighborhood w can be considered as the mass distribution within w and the position vector $\overrightarrow{V}$ is computed as:

$$\overrightarrow{V} = \frac{p_1 \overrightarrow{v}_1 + p_2 \overrightarrow{v}_2 + \cdots + p_n \overrightarrow{v}_9}{p_1 + p_2 + \cdots + p_9} \tag{9}$$

Theorem 1. V'_{pq} is also orthogonal

Proof. Let us consider the eq. (4)

$$\int \int [V_{pq}(\rho, \theta)][V_{p'q'}(\rho, \theta)]^* d\rho d\theta \tag{10}$$

Now, putting eq. (7) in eq. (10), we have $\int_x \int_y [V'_{pq}(x, y)][V'_{p'q'}(x, y)]^* dxdy$

$$= \int_x \int_y \frac{[V'_{pq}(x, y)][V'_{p'q'}(x, y)]^* dxdy}{e^{iq\theta_{CoM}} e^{-iq'\theta_{CoM}}} \tag{11}$$

where (x, y) signifies the location of the pixels in w.

Since $\int \int [V_{pq}(\rho, \theta)][V_{p'q'}]^* d\rho d\theta = 0 \ for (p, q) \neq (p', q')$ (see eq. (4)), $\int_x \int_y [V'_{pq}(\rho, \theta)][V'_{p'q'}]^* dxdy = 0 \ for \ (p, q) \neq (p', q')$. This is because when $(p, q) \neq (p', q')$, the whole integral, i.e.,$\int_x \int_y \frac{[V'_{pq}(\rho,\theta)][V'_{p'q'}]^* dxdy}{e^{iq\theta_{CoM}} e^{-iq'\theta_{CoM}}}$ becomes 0.

Likewise, $\int_x \int_y [V'_{pq}(\rho, \theta)][V'_{p'q'}]^* dxdy = \frac{\pi}{p+1} \ for \ (p, q) = (p', q')$ (see eq. (4)).

Hence, V'_{pq} is a radial orthogonal polynomial and all the properties of orthogonal moments are inherited.

Now, the Local Modified Zernike Moment (LMZM) Z'_{pq} is defined as the projection of the neighborhood w of the image $I(x, y)$ onto the set of orthogonal polynomial basis as:

$$Z'_{pq} = \frac{\frac{(p+1)}{\pi} \int \int_{x,y \in w} I(x, y)[V_{pq}(x, y)]^* dxdy}{e^{-iq\theta_{CoM}}} = \frac{Z_{pq}}{e^{-iq\theta_{CoM}}} \tag{12}$$

Finally, the proposed LMZMPM,C_{pq} is defined in the specified neighborhood w as:

$$C_{pq} = \frac{Z'_{pq}}{\int \int_{x,y \in w} I(x, y)} = \frac{Z_{pq}}{e^{-iq\theta_{CoM}} \sum \sum_{x,y \in w} I(x, y)} \tag{13}$$

$$= \frac{\frac{(p+1)}{\pi} \sum \sum_{x,y \in w} I(x, y) R(\rho) e^{-iq\theta}}{e^{-iq\theta_{CoM}} \sum \sum_{x,y \in w} I(x, y)} \tag{14}$$

It is noted that the proposed LMZMPM, C_{pq}, is invariant to the geometric transformation such as rotation, translation, and scaling of the neighborhood w. Also, it is invariant to the illumination changes.

Theorem 2. For a neighborhood w of size 3×3, C_{pq} is invariant to the illumination changes in the image I for any p and q.

Proof. Taking into account the Illumination Reflectance Model (IRM) [56], any given image I can be expressed as $I(x, y) = \mathcal{R}(x, y) \times L(x, y)$, where the term $\mathcal{R}(x, y)$ refers to the reflectance component and the term $L(x, y)$ refers to the luminance component. Usually, the luminance component L represents the slow spatial variation without any discontinuity and the reflectance component $\mathcal{R}$ comprises data about the image's characteristics feature. So, it is quite meaningful to consider that $L(x, y)$, $L(x + \Delta x, y)$, $L(x, y + \Delta y)$ are almost the same, when Δx and Δy are small. As a result, if the L component is eliminated, the key facial attributes can still be presented with the help of the $\mathcal{R}$ component alone, which are considered to be the most crucial data for a face recognition strategy. Removal of the L component will result in the $\mathcal{R}$ component being illumination-invariant, which causes the image I to be illumination-invariant, and enhances the sharpness of image I. However, parting the $\mathcal{R}(x, y)$ and $L(x, y)$ components is an ill-posed issue [11]. The aim is to prove that in the neighborhood w of size 3×3 C_{pq} is free from L.

Now, the ZM of order p with q repetition of an image I in the neighborhood w can be written as:

$$Z_{pq} = [\frac{(p+1)}{\pi} \sum \sum_{x,y \in w} I(x, y)\{V_{pq}(x, y)\}^*] \tag{15}$$

$$= [\frac{(p+1)}{\pi} \sum \sum_{x,y \in w} \mathcal{R}(x, y) L(x, y)\{V_{pq}(x, y)\}^*] \tag{16}$$

$$\because I(x, y) = \mathcal{R}(x, y) \times L(x, y)$$

Since $L(x, y)$ is approximately constant over a small neighborhood [11], i.e., $L(x, y) = l$, the above equation can be modified as:

$$Z_{pq} = l[\frac{(p+1)}{\pi} \sum \sum_{x,y \in w} \mathcal{R}(x, y)\{V_{pq}(x, y)\}^*] \tag{17}$$

In a neighborhood of size 3×3, the value of $\rho = 1$. Hence, the above equation, i.e., eq. (17), can be rewritten as:

$$Z_{pq} = l[\frac{p+1}{\pi} \sum \sum_{x,y \in w} \mathcal{R}(x, y) R_{pq}(1) e^{-iq\theta}] \tag{18}$$

$$= l[\frac{(p+1)}{\pi} R_{pq}(1) \sum \sum_{x,y \in w} \mathcal{R}(x, y) e^{-iq\theta}] \tag{19}$$

Now, putting eq. (19) into eq. (13), we have

$$C_{pq} = \frac{l[\frac{(p+1)}{\pi} R_{pq}(1) \sum \sum_{x,y \in w} \mathcal{R}(x, y) e^{-iq\theta}]}{l[e^{-iq\theta_{CoM}} \sum \sum_{x,y \in w} \mathcal{R}(x, y)]} \tag{20}$$

$$\because I(x,y) = \mathcal{R}(x,y) \times L(x,y); L(x, y) = l$$

$$= \frac{[\frac{(p+1)}{\pi} R_{pq}(1) \sum \sum_{x,y \in w} \mathcal{R}(x, y) e^{-iq\theta}]}{[e^{-iq\theta_{CoM}} \sum \sum_{x,y \in w} \mathcal{R}(x, y)]} \tag{21}$$

From eq. (23), it can be observed that it is free of luminance component L. Hence, in the proposed LMZMPM, C_{pq} is invariant to the illumination change for any p and q.

Theorem 3. For a neighborhood w of size 3×3, C_{pq} is invariant to the uniform scaling for any p and q.

Proof. Suppose the image be scaled uniformly k times of its original size. Then, the coordinates (x, y) of a pixel in the neighborhood w will be (kx, ky). Now,

$$\theta_{scaled} = tan^{-1}(\frac{ky}{kx}) = tan^{-1}(\frac{y}{x}) = \theta_{initial} \tag{22}$$

Also, let the new coordinate of the CoM is (kx_{CoM}, ky_{CoM}). Therefore,

$$\theta_{CoM,scaled} = tan^{-1}(\frac{ky_{CoM}}{kx_{CoM}}) = tan^{-1}(\frac{y_{CoM}}{x_{CoM}}) = \theta_{CoM,initial} \tag{23}$$

Thus, C_{pq} remains unchanged after uniform scaling and hence it is proved that C_{pq} is invariant to the uniform scaling of the neighborhood w.

Theorem 4. For a neighborhood w of size 3×3, C_{pq} is invariant under plane rotation for any p and q.

Proof. Let us assume that θ_{CoM} be the angle between CoM and the center pixel p_c in w. Now, let w be rotated by an angle α and hence, the angle between the CoM and the p_c of w is now $\theta_{CoM} + \alpha$.

Since, w is assumed to be circular, when the rotation takes place, the angle between the CoM and p_c is also rotated by the same angle α. Now, let us assume that $C_{pq}^{(R)}$ be the rotated LMZMPM of the corresponding LMZMPM C_{pq}. Likewise, let the rotated Z_{pq} is denoted as $Z_{pq}^{(R)}$, and is defined as:

$$Z_{pq}^{(R)} = Z_{pq} e^{-iq\alpha} \tag{24}$$

Now, from eq. (15), we can write

$$C_{pq}^{(R)} = \frac{Z_{pq}^{(R)}}{e^{-iq(\theta_{CoM}+\alpha)} \sum\sum_{x,y\in w} I(x,y)} \tag{25}$$

$$= \frac{Z_{pq} e^{-iq\alpha}}{e^{-iq(\theta_{CoM}+\alpha)} \sum\sum_{x,y\in w} I(x,y)} \tag{26}$$

$$= \frac{Z_{pq} e^{-iq\alpha}}{e^{-iq\theta_{CoM}} e^{-iq\alpha} \sum\sum_{x,y\in w} I(x,y)} \tag{27}$$

$$= \frac{Z_{pq}}{e^{-iq\theta_{CoM}} \sum\sum_{x,y\in w} I(x,y)} = C_{pq} \tag{28}$$

Hence, from eq. (30), we can conclude that the proposed C_{pq} is invariant to the rotation of the neighborhood w.

Theorem 5. For a neighborhood w of size 3×3, C_{pq} is invariant to the translation for any p and q.

Proof. Suppose an image $I(x, y)$ be uniformly shifted by h units horizontally and v units vertically. Now, let us consider that the initial coordinates of the center pixel p_c and its any neighboring pixel p_i in w be (x_0, y_0) and (x_i, y_i) respectively. Therefore, initially

$$\theta_{initial} = tan^{-1}(\frac{y_i - y_0}{x_i - x_0}) \tag{29}$$

$$\rho_{initial} = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} \tag{30}$$

Now, after translation, the coordinates of the center pixel p_c and its any neighboring pixel in w are $(x_0 + h, y_0 + v)$ and $(x_i + h, y_i + v)$ respectively. Thus,

$$\theta_{trans} = tan^{-1}(\frac{([y_i + v] - [y_0 + v])}{([x_i + h] - [x_0 + h])}) = tan^{-1}(\frac{y_i - y_0}{x_i - x_0}) \tag{31}$$

$$\rho_{trans} = \sqrt{([x_i + h] - [x_0 + h])^2 + ([y_i + v] - [y_0 + v])^2} \tag{32}$$

$$= \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} \tag{33}$$

Thus, from eqs. (31) and (33), $\theta_{trans} = \theta_{initial}$, and from eqs. (32) and (35), $\rho_{trans} = \rho_{initial}$.

Also, let the coordinate of the CoM in the neighborhood is (x_{CoM}, y_{CoM}). In our previous work [35], we have already proved that the angle of the CoM remains unchanged concerning the center pixel p_c in w for translation, i.e., $\theta_{CoM} = tan^{-1}(\frac{y_{CoM}}{x_{CoM}})$ is also unchanged. At the same time, $\sum\sum_{x,y\in w} I(x, y)$ is also constant in the neighborhood [11]. Therefore, the proposed C_{pq} is invariant to the translation of the image I. Likewise, ρ, θ, θ_{CoM} are remain unchanged for uniform translation of the neighborhood w. Thus, C_{pq} is invariant to the translation of the neighborhood w.

Theorem 6. C_{pq} is invariant to blur.

Proof. Our proposed LMZMPM C_{pq} is, to some extent, invariant to blurring (convolution), which can be easily proved from work in Ref. [57]. The work in [57], is based on the conventional orthogonal Zernike Moments (ZMs). Besides, our proposed LMZMPM C_{pq} is the modification of it. Hence, from the theory of blur-invariant in [57], we can prove that LMZMPM C_{pq} is invariant to the blurring effect.

From Theorem-2 in [57], we have the blur-invariant as:

$$I(q + 2l, q)^{(f)} = Z_{q+2l,q}^{(f)} - \frac{1}{Z_{0,0}^{(f)}.\pi} \sum_{i=0}^{l-1} I(q + 2i, q)^{(f)} \sum_{j=0}^{l-i} Z_{2j,0}^{(f)}.A(q, l, i, j) \tag{34}$$

which satisfy the equation:

$$I(q + 2l, q)^{(g)} - I(q + 2l, q)^{(f)} = 0 \tag{35}$$

where, $g(x, y) = (f * h)(x, y)$ is the blurred image, $f(x, y)$ is the original image and $h(x, y)$ is the circularly symmetric point spread function (PSF) as in [57]. For the proof of above , refer to Appendix-A in [57].

Now, from (34), it can be deduced that

$$I(2l, 0)^{(f)} = (-1)^l (2l + 1) Z_{0,0}^{(f)} \tag{36}$$

and $I(0, 0)^{(f)} = Z_{0,0}^{(f)}$ will be used for the case $q = 0$ as an invariant (see Remark-2 in [57]).

Note that,

$$Z_{q+2l,q}^{(f)} = C_{q+2l,q}^{(f)} (\sum_x \sum_y f(x, y)) e^{-iq\theta_{CoM}^{(f)}} \tag{37}$$

where $\theta_{CoM}^{(f)}$ is the angle between the center-of-mass (CoM) and the center pixel p_c in the neighborhood w in f.

Hence,

$$Z_{2j,0}^{(f)} = C_{2j,0}^{(f)} (\sum_x \sum_y f(x, y)) \tag{38}$$

Algorithm 1: Computation of LMZMPM Face

Input: Face Image I.

Output: LMZMPM face matrix $F = [C_{pq}]$

1. Extract non-overlapping patches w of size (3×3) from image I.
2. For each patch, extract its Centre of Mass (CoM).
3. Compute a set of modified radial orthogonal polynomials V'_{pq} by
$$V'_{pq} = \frac{V_{pq}}{e^{iq\theta_{CoM}}}$$
4. Compute the local modified Zernike moments (LMZM) Z'_{pq} by
$$Z'_{pq} = \frac{Z_{pq}}{e^{-iq\theta_{CoM}}}$$
5. Compute the LMZMPM (C_{pq}) face: $C_{pq} = \frac{z'_{pq}}{\iint_{x,y \in w} I(x,y)}$

And thus,

$$C_{0,0}^{(f)} = \frac{Z_{0,0}^{(f)}}{\sum_x \sum_y f(x,y)} = \frac{1}{\pi} \tag{39}$$

Substituting the above in (34), we get:

$$\begin{aligned} I(q+2l,q)^{(f)} = C_{q+2l,q}^{(f)}.(\sum_x \sum_y f(x,y)).e^{-iq\theta_{CoM}^{(f)}} \\ - \sum_{i=0}^{l-1} I(q+2i,q)^{(f)}. \sum_{j=0}^{l-i} C_{2j,0}^{(f)}.A(q,l,i,j) \end{aligned} \tag{40}$$

such that eq. (35) holds. This is our required blur invariant in terms of lower order invariants and LMZMPM's.

It is shown that the whole complex component, i.e., both the magnitude and phase coefficient, of LMZMPM is invariant to illumination change, rotation, translation and uniform scaling in w. Typically, the ZMs [20] are not invariant to the illumination change, which is an important aspect of face recognition. Thus, we have modified it in our proposed LMZMPM method and make it invariant to the illumination change. Furthermore, in conventional ZMs [20], the magnitude remains unchanged while the phase would change with rotation. However, in our approach, we have efficiently tackled this problem with the inclusion of θ_{CoM}. Also, it carries the standard properties of the ZMs. Here, the relationship between a given image pixel and its local neighboring pixels is measured in the form of magnitude and phase coefficients of LMZMPM. For each pixel of the input image, we have computed LMZMPM in a local neighborhood of size 3×3, and then considered the complex tuple that contains both the magnitude and phase coefficients of LMZMPM as the extracted features. Typically, the magnitude of the conventional ZMs is invariant to rotation, but its phase changes with rotation. Due to this, in many applications [13], [14], [28]–[30] the phase component has been neglected. Thus, any ZM-based extracted features without the phase information could have relatively weak descriptive efficiency. Besides, our extracted feature element is quite informative, which is a complex moment containing both phase and magnitude of the LMZMPM denoted by $F = [C_{pq}]$. For an image I of size $M \times N$, where M and N are the height and width of the image respectively, the size of the feature vector will be $[M' \times N']$, $M' = M - 2$ and $N' = N - 2$. This is because the extreme rows and columns are not taken into account on the ground that they do not have a complete 3×3 neighborhood around them. The implementation of the LMZMPM face is summarized in Algorithm-1.

IV. Similarity Measure

For an image I, we obtain a complex LMZMPM feature vector F of size $M' \times N'$, which is of the form:

$$F = \begin{pmatrix} C_{11} & C_{12} & \dots & C_{1M'} \\ C_{21} & C_{22} & \dots & C_{2M'} \\ \dots & \dots & \dots & \dots \\ C_{N'1} & C_{N'2} & \dots & C_{N'M'} \end{pmatrix} \tag{41}$$

where C_{ij} is the complex moments of the LMZMPM calculated with eq. (14), in which i in $[1, N']$ and j in $[1, M']$. Since F is a complex feature, we further break it into two vectors: i) magnitude (F_M) and ii) phase (F_θ).

$$F_M = \begin{pmatrix} ||C_{11}|| & ||C_{12}|| & \dots & ||C_{1M'}|| \\ ||C_{21}|| & ||C_{22}|| & \dots & ||C_{2M'}|| \\ \dots & \dots & \dots & \dots \\ ||C_{N'1}|| & ||C_{N'2}|| & \dots & ||C_{N'M'}|| \end{pmatrix} \tag{42}$$

$$F_\theta = \begin{pmatrix} arg(C_{11}) & arg(C_{12}) & \dots & arg(C_{1M'}) \\ arg(C_{21}) & arg(C_{22}) & \dots & arg(C_{2M'}) \\ \dots & \dots & \dots & \dots \\ arg(C_{N'1}) & arg(C_{N'2}) & \dots & arg(C_{N'M'}) \end{pmatrix} \tag{43}$$

After that, we standardize these two vectors, i.e., F_M and F_θ, so that their range is comparable and the difference in this range of values does not affect the final feature vector.

The standardized magnitude value is given by:

$$F'_M = \frac{F_M - Min(F_M)}{Max(F_M) - Min(F_M)} \tag{44}$$

and the standardized phase value is given by:

$$F'_\theta = \frac{F_\theta - Min(F_\theta)}{Max(F_\theta) - Min(F_\theta)} \tag{45}$$

After computation of F'_M and F'_θ, the face images are classified using the proposed hybrid similarity that combines L_1 distance (scaled between 0 and 1) and the Jaccard distance (J) [58] of (F'_M) and (F'_θ), which is discussed in the following.

Suppose, for two face images, I_1 and I_2, the obtained magnitude and phase vectors are (F'_{M1}, F'_{M2}) and $(F'_{\theta 1}, F'_{\theta 2})$ respectively. Here, L_1 distance is essentially the sum of the standardized minimum magnitude difference and the standardized minimum phase difference between them. Here, lesser the distance between the two images, I_1 and I_2, more is the similarity between them, revised from the idea of Zhang et al. [59], shown in equation (46) and the Jaccard distance is shown in equation (47) is given by:

$$L_1 = \sum_{i=1}^{M' \times N'} Min(|F'_{M1i} - F'_{M2i}|, |F'_{\theta 1i} - F'_{\theta 2i}|) \tag{46}$$

$$J = 1 - [\sum_{i=1}^{M' \times N'} \frac{Min(F'_{M1i}, F'_{M2i})}{Max(F'_{M1i}, F'_{M2i})} + \sum_{i=1}^{M' \times N'} \frac{Min(F'_{\theta 1i}, F'_{\theta 2i})}{Max(F'_{\theta 1i}, F'_{\theta 2i})}] \tag{47}$$

The face classification is based on the minimum value of the similarity score S between I_1, I_2, reckoned according to (48):

$$S = \frac{1}{2}(\frac{L_1 - Min(L_1)}{Max(L_1) - Min(L_1)} + \frac{J - Min(J)}{Max(J) - Min(J)}) \tag{48}$$

V. EXPERIMENTAL RESULTS

The efficiency of the proposed LMZMPM method has been evaluated on various standard and challenging databases: CMU-PIE [60], Extended Yale B [61], [62], AR database [63], LFW [64], CUFS [43], and CASIA-HFB [65]. For the challenges of variation in illumination, LMZMPM has been tested on the CMU-PIE and the Extended Yale B databases. The AR database is considered for analyzing the robustness of LMZMPM method against noise, translation, scaling and in-plane rotation. To test the proficiency of the LMZMPM method under complicated variations of illumination, expression, pose, occlusion, and blurring, the "Labeled Faces in the Wild" (LFW) database has been considered. The CUHK Face Sketch (CUFS) and CASIA-HFB databases are used to test the efficacy of the proposed LMZMPM method in the Heterogeneous Face Recognition (HFR) scenario. The Rank-1 Recognition (R1R) accuracy of the proposed LMZMPM method, using the proposed hybrid similarity measure, was computed for each of these databases.

The performance of the proposed LMZMPM method on the datasets stated above is compared against several existing related state-of-the-art methods. In this regard, it is noted that we have provided six baselines performance on each scenario in order to make the experimental section more consistent and provide a comparison benchmark for the future. The six state-of-the-art methods used as baselines are: two state-of-the-art Zernike moments based methods, i.e., Multi-scale Local Zernike Moment (MSLZM) [14] and Local Zernike Moments (LZM) [13], Multi-scale Local Gradient fuzzy Patterns (MLGFP) [50], Multi-scale Logarithm Difference Edgemaps (MSLDE) [37], Local Center of Mass Face (LCMF) [35], and the convolutional neural network (CNN) [8]. For the sake of impartial comparison, all of the stated baselines (except CNN) are implemented in accordance with the description defined in the respective published papers, and then they were fine-tuned to obtain about the same performance as the published papers. Besides, for CNN, we have adopted here a transfer learning-based model using AlexNet [8]. For transfer learning the model, we have randomly chosen 50% of the images from each of the databases and the remaining images are used to test the model. Firstly, we have used raw face images for transfer learning the model and recorded the testing accuracy, which is named here as "Raw Fimage + CNN (RF-CNN)". Then, LMZMPM face images are used for transfer learning the model and the testing accuracy is recorded, which is represented as "LMZMPM Fimage + CNN (LMZMPM-CNN)".

In addition, for different face matching scenarios, the proposed LMZMPM method is compared with other existing state-of-the-art methods on the respective fields, which discussed in the following subsection.

A. Results on Illumination-variations Databases

In this subsection, we evaluate the robustness of the proposed LMZMPM method in recognizing faces under varying illumination conditions. For this purpose, we have conducted experiments on two standard face databases of illumination-variation, i.e., CMU-PIE [60] and Extended Yale B [61], [62], and compared the proposed method with above-mentioned baselines and other related state-of-the-art methods: HF+HQ [38], LTV [32], LFA [36], LDN [34], W-Face [33], ZZPLELD [51] and uLBP+LM [15].

Figure 2. Rows (a), (c) and (e) present face images of a subject of CMU-PIE database under varying illumination conditions; and rows (b), (d) and (f) show the corresponding LMZMPM faces.

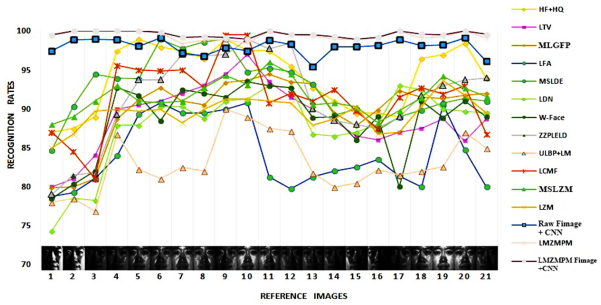

Figure 3. R1R accuracy using 68 different subjects, from the CMU-PIE database, considering each image (one at a time) as a reference image.

1) Results on CMU-PIE Face Database: As we are interested in measuring the illumination invariant property, we have considered the subset C27 of the CMU-PIE database in our experiment that consists of 41368 images of 68 different subjects (21 images per subject). Here, all the images have a frontal pose with a neutral expression, but they are all captured under varying illumination conditions. We have cropped and resized each of these images to obtain a 128×128 resolution. Fig. 2 presents an example of 21 faces of a subject from the subset C27 of the CMU-PIE database, along with their corresponding LMZMPM faces. Only one image per subject is randomly chosen as a reference image and the remaining images of that class are used as testing images. The average R1R accuracy results of the proposed LMZMPM method, six baselines, and other related state-of-the-art methods mentioned earlier are presented in Fig. 3. As can be seen from the experimental results that the proposed LMZMPM method is superior to that of the comparative methods. Also, it can be seen that the "LMZMPM Fimage + CNN" outperforms "Raw Fimage + CNN" and LMZMPM method with a hybrid similarity measure.

Figure 4. (a) Sample face images of a subject from Extended Yale B under varying illumination conditions, (b) The corresponding LMZMPM faces.

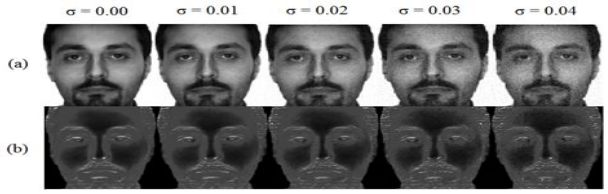

Figure 5. (a) Face image of an individual from the AR-database and its corresponding images with noise variations, (b) Corresponding LMZMPM faces.

Table II
RESULTS ON EACH OF THE FIVE EXTENDED YALE B DATABASE SUBSET.

Methods	Set I	Set II	Set III	Set IV	Set V	Average
HF+HQ	90.70	88.60	40.00	45.50	42.40	61.44
LTV	88.20	90.30	67.80	58.00	43.50	69.56
LFA	92.70	92.10	89.60	86.80	86.70	89.58
LDN	86.40	87.40	84.00	79.00	75.80	82.52
W-Face	88.47	86.89	83.57	83.11	81.72	84.75
ZZPLELD	99.00	100.00	96.70	92.50	94.30	96.50
uLBP+LM	98.50	98.30	97.20	96.50	91.80	96.46
MSLDE	97.20	98.00	93.40	92.40	88.90	93.98
MLGFP	99.80	98.30	97.00	91.00	90.00	95.22
LCMF	99.90	100.00	97.00	93.20	94.00	96.82
MSLZM	89.6	90.68	94.35	92.54	93.67	92.17
LZM	87.25	89.2	91.35	90.89	91.52	90.04
RF-CNN	100.00	99.83	98.70	97.67	96.57	98.54
LMZMPM	100.00	100.00	98.50	98.00	96.80	98.66
LMZMPM-CNN	100.00	100.00	99.50	98.75	98.24	99.29

2) Results on Extended Yale B Face Database: The Extended Yale B Face Database comprises 38 subjects with nine different poses and each pose is subjected to 64 different illumination angles. In our experiment, a total of $(64 \times 38 = 2432)$ images with frontal poses were used. All the images are cropped to extract only the facial portion and resized to 150×150 resolution. This database is divided into five different subsets based on the illumination angle using standard protocols [66]. Set I consists of 266 images (7 images per subject) with illumination angle of $0°$ to $12°$, Set II consists of 456 images each (12 images per subject) with illumination angle of $13°$ to $25°$, Set III also consists of 456 images each (12 images per subject) with illumination angle of $26°$ to $50°$, Set IV consists of 532 images (14 images per subject) with illumination angle of $51°$ to $77°$ and Set V consists of 722 images (19 images per subject) with illumination angle above $78°$. A few sample face images and their resultant LMZMPM faces from the Extended Yale B database are shown in Fig. 4. In the experiment, only one neutral image per subject was considered as the reference image, while the rest of the images were included in the test set for each class. The R1R accuracy of the proposed LMZMPM approach on the Extended Yale B database, along with that of other state-of-art techniques, is presented in Table-II, from where it can be observed that the proposed method (i.e., LMZMPM) is more remarkable than the other methods in all the individual subsets, as well as when the average over all the 5 subsets is taken. Again, "LMZMPM Fimage + CNN" gives a comparatively better result than that of the LMZMPM with hybrid similarity measure.

B. Results on varying noise, translation, scale and in-plane rotation on AR Face Database

The AR database, comprises more than $4,000$ color images of 126 subjects, of size 768×576 pixels. This database consists of images with not only varying lighting conditions but also images comprising expression changes. Two occlusion images of each subject (one with sunglasses and another with a scarf) are also a part of it. For our experiment, each image of this database is cropped with a resolution of 128×128 and then, converted to the grayscale counterparts. The reference image for a class was taken as the one having a neutral expression with equal lighting conditions. The rest of the images were taken as the test images for that class. In each of the cases, the Fisher Criteria (FC) [61] and Standard Deviation (SD) are evaluated to show the discriminative power and robustness of the proposed method compared to other state-of-the-art techniques.

1) Performance Against Additive Noise: The test images have been added with Gaussian noise with Standard Deviations (σ) as $\sigma = 0.01, \sigma = 0.02, \sigma = 0.03$ and $\sigma = 0.04$. Fig. 5(a) shows the original image of an individual from the AR database and its corresponding images with noise variations and Fig. 5(b) represents the equivalent LMZMPM images of each face presented in Fig. 5(a). Despite high noise variations of up to $\sigma = 0.04$ our proposed method is more invulnerable, when compared to the other methods like noise-resistant local binary pattern (NRLBP) [66] (which is a version of noise robust LBP), six baselines, and other related state-of-the-art methods such as HF+HQ [38], uLBP+LM [15], LFA [36] and ZZPLELD [51], the results of which can be shown in Table-III. Also, the proposed method has the highest R1R accuracy, maximum FC and minimum SD, which shows the discriminative power and vigor of LMZMPM over the other approaches. Further, it can be seen that "LMZMPM Fimage + CNN" gives comparatively better result than the LMZMPM method (i.e., LMZMPM + hybrid similarity measure).

2) Performance against rotational changes and scale changes: For rotational changes, the test images have been subjected to rotations, with angle of rotation as $10°$, $20°$, $30°$ and $40°$. Fig. 6(a) shows the original image of an individual and its corresponding images with different angle of rotation and Fig. 6(b) represents the equivalent LMZMPM images. For scaling changes the test images are scaled, with scaling factors taken as $S = 0.25$, $S = 1.5$, $S = 1.75$ and $S = 2$. Fig. 7(a) shows the original image of an individual and its corresponding images with different values of the scaling factor. Fig. 7(b) shows the equivalent LMZMPM images.

Table III
RESULTS ON THE AR DATABASE WITH NOISE VARIATIONS.

Methods	$\sigma = 0.00$			$\sigma = 0.01$			$\sigma = 0.02$			$\sigma = 0.03$			$\sigma = 0.04$		
	R1R	SD	FC	R1R	SD.	FC	R1R	SD	FC	R1R	SD	FC	R1R	SD	FC
HF+HQ	78.75	5.31	1.08	78.12	5.46	1.11	74.34	3.45	1.01	68.53	3.76	0.99	64.50	4.08	0.85
uLBP+LM	87.35	2.33	1.39	86.97	2.83	1.44	86.52	4.47	1.35	85.18	4.82	1.21	83.60	5.01	1.14
LFA	87.50	3.30	1.33	84.22	4.45	1.16	78.89	5.27	1.00	71.56	5.39	1.00	68.56	5.68	0.89
ZZPLELD	84.40	2.30	0.99	80.12	2.75	0.98	76.51	5.75	0.88	75.11	5.97	0.85	73.75	6.33	0.81
NRLBP	94.50	2.06	1.92	91.78	2.05	1.80	89.45	2.75	1.7	88.10	2.95	1.65	83.40	3.01	1.48
MSLDE	86.75	2.17	1.54	85.20	3.15	1.47	81.33	6.15	1.27	77.67	6.78	1.00	68.70	6.89	0.92
MLGFP	91.67	2.19	1.48	91.51	2.56	1.52	89.16	3.01	1.10	87.56	3.56	1.01	82.24	4.21	0.91
LCMF	91.60	2.26	1.70	90.83	2.25	1.65	88.56	3.00	1.55	86.33	3.10	1.44	85.00	3.75	1.35
MSLZM	92.65	4.27	1.09	89.65	3.75	1.00	88.44	3.67	1.17	88.24	3.67	1.25	87.67	3.48	1.57
LZM	90.57	4.58	0.98	88.45	3.89	0.87	88.21	4.87	1.05	87.25	3.25	1.16	85.67	3.82	1.14
RF-CNN	97.95	1.8	2.03	96.11	2.02	1.98	95.06	2.52	1.81	92.65	2.77	2.05	91.83	3.00	1.61
LMZMPM	98.74	1.71	2.47	96.72	1.76	2.21	95.17	1.8	2.19	94.51	2.25	2.10	92.10	2.25	2.01
LMZMPM-CNN	99.25	1.71	2.47	97.67	1.76	2.21	96.77	1.8	2.19	96.14	2.25	2.10	94.25	2.25	2.01

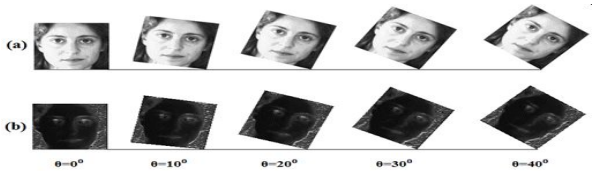

Figure 6. (a) Face image of an individual from the AR database and its corresponding images with different angle of rotation, (b) The corresponding LMZMPM faces.

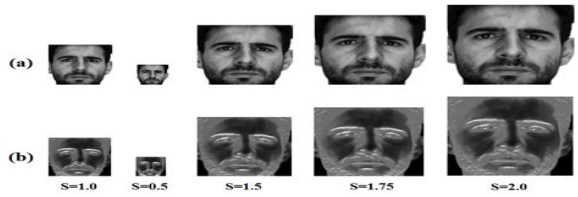

Figure 7. (a) Original image of an individual from the AR-database and its corresponding images with different values of the scaling factor, (b) corresponding LMZMPM faces.

Tables-IV and V shows the R1R accuracy, FC and SD values of LMZMPM along with other state-of-art techniques stated in Sec. V.C.1 for different values of the angle and scaling factor. From these tables, it is clear that LMZMPM method performs better than other approaches with the highest R1R accuracy, having maximum FC and minimum SD. Here, we can also seen that "LMZMPM Fimage + CNN" outperforms the LMZMPM method with a hybrid similarity measure.

3) Performance Against Translation: For the test images, the translations considered were $(h = 10, v = 10), (h = 10, v = 0), (h = 10, v = -10)$ and $(h = 20, v = -20)$. Fig. 8(a) shows the original image of an individual and its corresponding images with different translations. Fig. 8(b) represents the equivalent LMZMPM images. Here, also, only one image per subject without any translation is considered as the reference image. Table-VI shows the R1R accuracy, FC and SD values of LMZMPM along with other state-of-art methods for different translations, from where it is clear that the proposed LMZMPM method performs better than other approaches with the highest R1R accuracy, maximum FC and minimum SD. In this scenario, "LMZMPM Fimage + CNN" also gives comparatively better results than the LMZMPM method.

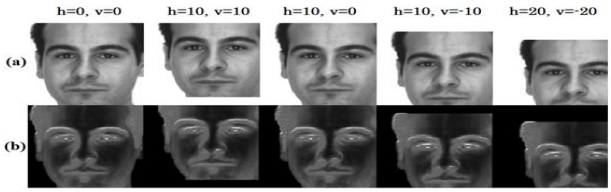

Figure 8. (a) Original image of an individual from the AR database and its corresponding images with translational variations, (b) corresponding LMZMPM faces.

C. Results on LFW Face Database

Presently LFW database is the most challenging database to work with complicated variations in illumination, expression, pose, occlusion, and blurring. The database contains 13,233 images of 5749 different famous persons having two Views, View1 is for fine-tuning and training and View2 is for testing. Usually, three different standards are used to evaluate this database. We followed the standard where neither parameter tuning nor training is required. From this database, face images of 30 individuals with 10 images per subject were randomly chosen to test the effectiveness of the proposed LMZMPM method. For each of the 30 individual subsets, the experiment has been conducted separately using a leave-one-out strategy. Fig. 9(a) shows some sample of face images from the LFW database, which are cropped to take only the facial portion. Fig. 9(b) shows their corresponding LMZMPM faces. A comparison of the R1R accuracy, together with FC and SD of the proposed method, six baselines, and other related state-of-the-art techniques such as uLBP+LM [15], NRLBP [66], ZZPLELD [51], PLS [41], and I-LQP[49] is presented in Table-VII, from where it is evident that LMZMPM method outperforms the other comparative methods (except "Raw Fimage + CNN"). Though the method of CNN shows a slight better performance than the proposed approach, i.e., LMZMPM, it has its own set of disadvantages- high computational complexity, time consuming and requires training. However, "LMZMPM Fimage + CNN" performs significantly better than "Raw Fimage + CNN".

D. Heterogeneous Face Recognition

In this subsection, the efficacy of the proposed LMZMPM method is evaluated in the Heterogeneous Face Recognition (HFR) scenario. For this purpose, we have conducted

Table IV
RESULTS ON THE AR DATABASE WITH ROTATIONAL CHANGES.

Methods	$\theta = 0^\circ$			$\theta = 10^\circ$			$\theta = 20^\circ$			$\theta = 30^\circ$			$\theta = 40^\circ$		
	R1R	SD	FC	R1R	SD.	FC	R1R	SD	FC	R1R	SD	FC	R1R	SD	FC
HF+HQ	78.75	5.31	1.08	70.40	3.42	0.97	65.50	3.45	0.96	63.50	3.99	0.89	59.30	4.74	0.82
uLBP+LM	87.35	2.33	1.39	84.50	4.43	1.33	82.00	4.45	1.16	81.60	4.66	1.03	71.65	4.65	0.97
LFA	87.50	3.30	1.33	73.25	5.19	0.79	68.67	5.25	0.68	63.56	5.56	0.62	55.50	5.56	0.59
ZZPLELD	84.40	2.30	0.99	79.50	5.40	0.74	77.00	5.87	0.70	71.74	6.30	0.68	71.44	6.91	0.64
NRLBP	93.50	1.96	1.92	85.82	2.74	1.74	84.00	2.91	1.63	80.40	3.00	1.43	76.66	3.24	1.40
MSLDE	86.75	2.17	1.54	76.33	3.97	1.25	72.60	4.17	0.98	67.65	4.46	0.91	66.25	4.76	0.88
MLGFP	91.67	2.19	1.48	86.00	2.98	1.05	84.55	3.54	1.00	77.24	4.17	0.92	75.45	4.19	0.89
LCMF	91.60	2.26	1.70	87.40	3.06	1.54	82.33	3.15	1.42	81.40	3.74	1.33	80.60	3.92	1.31
MSLZM	92.65	4.27	1.09	90.65	3.45	1.19	88.34	4.21	1.17	88.24	3.94	1.25	87.67	3.37	1.45
LZM	90.57	4.58	0.98	89.45	3.89	0.97	87.16	4.87	1.05	86.94	3.56	1.11	84.95	3.15	1.09
RF-CNN	97.75	1.90	2.22	95.34	1.90	2.03	94.00	2.46	1.88	92.57	2.47	1.80	91.82	2.53	1.78
LMZMPM	98.74	1.71	2.47	97.30	1.72	2.25	94.75	2.23	2.09	93.50	2.24	2.00	92.75	2.29	1.98
LMZMPM-CNN	99.25	1.71	2.47	97.53	1.72	2.25	94.75	2.23	2.09	95.42	2.24	2.00	93.67	2.29	1.98

Table V
RESULTS ON THE AR DATABASE WITH SCALE CHANGES.

Methods	$S = 1$			$S = 0.25$			$S = 1.5$			$S = 1.75$			$S = 2$		
	R1R	SD	FC	R1R	SD.	FC	R1R	SD	FC	R1R	SD	FC	R1R	SD	FC
HF+HQ	78.75	5.31	1.08	76.86	5.42	1.07	77.15	5.39	1.06	75.94	5.53	1.05	75.32	5.61	1.04
uLBP+LM	87.35	2.33	1.39	85.25	2.38	1.37	85.58	2.37	1.36	84.23	2.43	1.35	83.54	2.46	1.33
LFA	87.50	3.30	1.33	85.40	3.37	1.31	85.72	3.35	1.30	84.38	3.44	1.29	83.69	3.49	1.27
ZZPLELD	84.40	2.30	0.99	82.37	2.35	0.98	82.69	2.34	0.97	81.39	2.40	0.96	80.72	2.43	0.95
NRLBP	93.50	1.96	1.92	91.26	2.00	1.90	91.60	1.99	1.88	90.16	2.04	1.86	89.43	2.07	1.84
MSLDE	86.75	2.17	1.54	84.67	2.21	1.52	84.99	2.20	1.51	83.65	2.26	1.49	82.97	2.29	1.48
MLGFP	91.67	2.19	1.48	89.47	2.23	1.46	89.81	2.22	1.45	88.40	2.28	1.43	87.68	2.31	1.42
LCMF	91.60	2.26	1.70	89.40	2.30	1.68	89.74	2.29	1.66	88.33	2.36	1.65	87.61	2.39	1.63
MSLZM	92.65	4.27	1.09	91.461	3.45	1.19	88.34	4.21	1.17	88.24	3.94	1.25	87.67	3.37	1.45
LZM	90.57	4.58	0.98	89.45	3.89	0.97	87.16	4.87	1.05	86.94	3.56	1.11	84.95	3.15	1.09
RF-CNN	98.64	1.89	2.23	96.82	1.92	2.18	96.10	2.10	2.40	94.79	1.77	2.33	94.52	1.86	2.40
LMZMPM	98.74	1.71	2.47	96.92	1.75	2.44	95.78	1.75	2.42	94.79	1.79	2.39	94.53	1.81	2.37
LMZMPM-CNN	99.13	1.71	2.47	97.32	1.75	2.44	95.78	1.75	2.42	95.00	1.79	2.39	95.03	1.81	2.37

Table VI
RESULTS ON THE AR DATABASE WITH TRANSLATIONAL VARIATIONS.

Methods	$h = 0, v = 0$			$h = 10, v = 10$			$h = 10, v = 0$			$h = 10, v = -10$			$h = 20, v = -20$		
	R1R	SD	FC	R1R	SD.	FC	R1R	SD	FC	R1R	SD	FC	R1R	SD	FC
HF+HQ	78.75	5.31	1.08	76.39	5.47	1.06	77.33	5.45	1.07	75.60	5.51	1.04	75.05	5.61	1.03
uLBP+LM	87.35	2.33	1.39	84.73	2.43	1.36	85.78	2.39	1.37	83.86	2.42	1.34	83.24	2.46	1.33
LFA	87.50	3.30	1.33	84.88	3.40	1.32	85.93	3.39	1.32	84.00	3.42	1.28	83.39	3.49	1.27
ZZPLELD	84.40	2.30	0.99	82.63	2.33	0.97	82.88	2.36	0.98	81.02	2.39	0.95	80.43	2.43	0.95
NRLBP	93.50	1.96	1.92	91.54	2.01	1.88	91.82	2.01	1.90	89.76	2.03	1.85	89.11	2.07	1.84
MSLDE	86.75	2.17	1.54	84.15	2.24	1.51	85.19	2.23	1.52	83.28	2.25	1.48	82.67	2.29	1.47
MLGFP	91.67	2.19	1.48	88.92	2.26	1.45	90.02	2.25	1.46	88.00	2.27	1.43	87.36	2.31	1.42
LCMF	91.60	2.26	1.70	89.68	2.33	1.68	89.95	2.32	1.68	87.94	2.34	1.64	87.29	2.39	1.63
MSLZM	92.65	4.27	1.09	90.65	3.15	1.29	88.34	4.33	1.17	89.42	3.98	1.25	87.67	3.72	1.65
LZM	90.57	4.58	0.98	89.14	3.92	1.06	87.16	4.87	1.05	87.49	4.16	1.11	84.95	4.15	1.17
RF-CNN	98.73	1.87	2.07	96.71	1.96	2.19	96.56	1.67	2.17	96.00	1.97	2.11	95.88	1.88	2.28
LMZMPM	98.74	1.72	2.47	96.81	1.75	2.43	96.85	1.76	2.44	95.7	1.78	2.38	95.19	1.82	2.36
LMZMPM-CNN	99.19	1.72	2.47	97.21	1.75	2.44	97.25	1.76	2.44	98.12	1.78	2.38	98	1.82	2.36

Figure 9. (a) Sample face images from the LFW database, (b) the corresponding LMZMPM faces.

Table VII
PERFORMANCE COMPARISON ON THE LFW DATABASE.

Methods	R1R	Avg. SD	Avg. FC
uLBP+LM	77.1	2.344	1.362
NRLBP	82.67	2.125	2.421
ZZPLELD	76.4	2.056	1.912
PLS	85.47	2.34	2.23
I-LQP	76.42	3.016	2.278
MSLDE	73.6	3.342	1.725
MLGFP	85.47	3.125	1.716
LCMF	87.99	2.269	2.592
MSLZM	84.44	2.94	1.98
LZM	91.67	3.12	1.76
RF-CNN	92.72	2.126	2.785
LMZMPM	92.65	2.001	2.676
LMZMPM-CNN	94.75	2.00	2.72

experiments on two standard heterogeneous face databases such as CUHK Face Sketch Database (CUFS) [43] and the CASIA-HFB [65], and compared the proposed method with six baselines and other related state-of-the-art methods: SNS-SRE [44], TFSP [45], SIFT+MLBP [46], PLS [41], THFM [48], KP-RS [47], SGR-DA [40], I-LQP [49] and ZZPLELD [51].

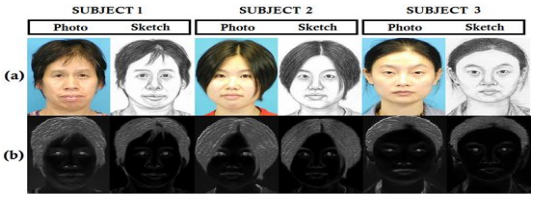

Figure 10. (a) Sample face-sketch pairs from the CUFS database, (b) the corresponding LMZMPM faces.

Table VIII
PERFORMANCE COMPARISON ON THE CUFS DATABASE.

Methods	R1R	Avg. SD	Avg. FC
SNS-SRE	96.5	2.945	4.564
TFSP	97.1	2.414	4.173
SIFT+MLBP	98.74	2.125	4.948
PLS	96.6	2.342	4.52
THFM	94.34	2.753	4.235
KP-RS	96.4	2.056	5.419
SGR-DA	98.12	2.75	4.67
I-LQP	84.46	2.856	5.137
ZZPLELD	93.72	3.192	3.453
MSLDE	90.1	2.629	3.32
MLGFP	98.7	2.345	4.882
LCMF	99.89	2.001	5.594
MSLZM	95.28	2.74	1.87
LZM	94.16	2.98	1.66
RF-CNN	100	2.826	2.785
LMZMPM	100	1.98	4.47
LMZMPM-CNN	100	1.97	4.72

1) Results on the CUHK Face Sketch (CUFS) Database: The CUFS database is a combination of three sketch-photo pair databases: The CUHK Student Face database [67], AR Sketch database [63] and the XM2VTS database [68]. In total, the CUFS database consists of 606 sketch and photo pairs that comprise 188 pairs from the CUHK Student Database, 295 pairs from the XM2VTS database and 123 pairs from AR Database. Each pair consists of one photo of a subject and its respective sketch image. In our analysis, the photograph has been considered as the reference picture, while its corresponding sketch is taken as the test image. Here, each image is cropped of size 100×100 that includes only the facial portion. Fig. 10(a) shows some samples of face and sketch images from the CUFS database. Fig. 10(b) shows their corresponding LMZMPM faces. A comparison of the R1R accuracy, FC and SD of the proposed method and other comparative methods stated earlier is presented in Table-VIII, from where it is evident that the LMZMPM method leaves behind the other recent techniques and tops the bench with a 100% accuracy, least SD and highest FC score compared to other approaches.

2) Results on the CASIA-HFB NIR-VIS Database: CASIA-HFB database is a database that has 200 subjects with gallery images captured in the visible light and test images captured in the near-infrared. Every single subject has 4 VIS pictures and 4 NIR pictures with posture and articulation variations. In each of these photos, the facial portion is cropped out to fit a 120×120 resolution in such a way that they are approximately at the same eye levels. The sample image NIR-VIS and its corresponding LMZMPM image, is shown in Fig. 11. The R1R accuracy, SD and FC values of the proposed LMZMPM is compared with other prevalent methods stated earlier, which is shown in Table-IX. As shown, the proposed method outperforms all the comparative techniques by achieving an average R1R accuracy of 98.80% with a maximum FC on this challenging database. Though "Raw Fimage + CNN" achieves a R1R accuracy of 98.89%, which is slightly better than the proposed LMZMPM method, it shows a higher SD and a lower FC score compared to that of the proposed LMZMPM method. In this scenario, "LMZMPM Fimage + CNN" also gives better accuracy than the LMZMPM method, but the FC is slightly higher than the LMZMPM method.

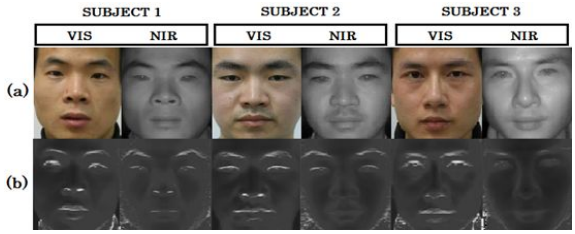

Figure 11. (a) Sample NIR-VIS face images from CASIA-HFB database, (b) corresponding LMZMPM faces .

Table IX
PERFORMANCE COMPARISON ON THE CASIA-HFB DATABASE.

Methods	R1R	Avg. SD	Avg. FC
SNS-SRE	91.54	3.02	3.87
TFSP	93.12	3.07	4.26
SIFT+MLBP	96.48	3.21	4.19
PLS	94.56	3.00	4.40
THFM	82.46	3.15	2.50
KP-RS	97.64	1.98	5.42
SGR-DA	97.1	2.93	4.78
I-LQP	89.46	2.33	5.23
ZZPLELD	89.90	2.65	3.59
MSLDE	75.6	2.96	3.32
MLGFP	98.96	2.69	4.88
LCMF	94.35	3.65	2.89
MSLZM	90.1.6	2.64	2.21
LZM	88.45	3.25	1.81
RF-CNN	98.89	2.11	5.78
LMZMPM	98.85	1.73	5.92
LMZMPM-CNN	99.57	1.52	6.14

E. Performance Comparison of Different Similarity Measures

In this sub-section, the effectiveness of our proposed hybrid similarity measure in recognizing faces is compared with three state-of-the-art similarity measures including structural similarity index (SSIM) [69], k-nearest neighbors (k-nn) [70], and L_1-norm distance measure [12]. We have conducted experiments on face images of AR-database. Table-X illustrates the performance of different similarity measures stated earlier in recognizing faces in terms of R1R accuracy. It can be observed that the proposed hybrid similarity measure performs significantly better than other comparative methods. Typically, our hybrid similarity measure considers common properties of two objects to measure similarity between them. In effect, it lowers the rate of error in recognition and thereby, improves the classification accuracy. It can also be observed that the R1R accuracy of L_1-norm distance measure is considerably lower than the proposed hybrid similarity measure. However, when we integrate the Jaccard distance with the L_1-distance

Table X
PERFORMANCE COMPARISON OF DIFFERENT SIMILARITY MEASURES ON THE AR-DATABASE.

Similarity Measures	R1R
SSIM	84.75
k-nn (k=3)	94.02
L_1-distance	96.68
Proposed hybrid similarity measure	98.74

(i.e., our hybrid similarity measure), the performance improves notably.

F. Unbiased Face Recognition- Accuracy, Specificity and Sensitivity Measure

Due to the biasedness of the R1R rate, it is not always a truthful measure for performance evaluation of an approach. Therefore, for unbiased performance assessment of the LMZMPM method, we have computed sensitivity (SN) and specificity (SP) [71] along with accuracy (AC) measure. These are computed as $AC = \frac{(T_P+T_N)}{(T_P+T_N+F_P+F_N)}$, $SN = \frac{T_P}{(T_P+F_N)}$, and $SP = \frac{T_P}{(T_P+F_P)}$ respectively, where (T_P), (F_P), (T_N) and (F_N) signify true positive , false positive , true negative , and false negative, respectively.

To indicate better performance, besides having good AC, both SN and SP should be high. In this experiment, datasets are created for each database, consisting of a number of classes. Each of the classes has a certain number of images in it, out of which 80% of the images are of a particular individual and the remaining 20% images are of other individuals selected randomly from the database, leaving that specific subject. From the 80% images of the particular individual, only one image of that individual is considered as the reference image, and the remaining images of the particular individual are used for positive testing. The negative testing is done using the rest 20% images of other individuals. The measure of AC, SN and SP of the proposed LMZMPM method, along with that of other state-of-art techniques are presented in Table-XI,for CMU-PIE and Extended YALE B databases; Table-XII for AR database with noise, rotation, scaling and translation; Table-XIII for LFW; and Table-XIV for CUFS and CASIA-HFB databases, from where it can be observed that besides having a high accuracy value, the proposed approach also shows a high SN and SP values on each of the databases and outperforms the other state-of-art techniques. Though the method of "Raw Fimage + CNN" (Table-XII) shows a better result than the proposed LMZMPM method in some cases, it is computationally very complex, time-consuming and needs training for performing efficiently. Again, "LMZMPM Fimage + CNN" outperforms the LMZMPM method with hybrid similarity measure.

Table XI
RESULTS ON CMU-PIE AND EXTENDED YALE B DATABASES.

Methods	CMU-PIE			Extended Yale B		
	SN	SP	AC	SN	SP	AC
HF+HQ	90.2	92.2	91.5	94.3	89.8	90.5
LTV	92.6	94.0	93.8	95.0	92.2	91.7
LFA	93.0	93.3	92.5	96.0	90.6	91.4
LDN	94.3	96.5	95.8	96.3	92.7	93.6
W-Face	90.0	94.3	92.7	89.9	86.3	87.0
ZZPLELD	95.3	96.9	98.7	97.6	92.2	94.0
uLBP+LM	93.5	95.3	94.3	97.1	92.0	93.2
MSLDE	91.0	92.6	94.4	93.2	90.1	91.6
MLGFP	95.4	95.1	94.3	96.9	92.5	94.4
LCMF	97.1	99.6	98.4	98.0	98.6	99.1
MSLZM	96.8	96.6	96.7	97.4	97.6	97.5
LZM	93.7	95.7	94.7	95.8	96.2	96.0
RF-CNN	99.5	99.8	99.8	99.2	99.4	99.5
LMZMPM	99.7	99.8	99.9	99.6	99.5	99.6
LMZMPM-CNN	100	100	100	100	99.75	99.75

VI. CONCLUSION AND FUTURE WORK

The paper puts forward a local orthogonal moment based modified approach for face recognition, called LMZMPM, which not only preserves the intrinsic properties of orthogonal moments, like translation and scaling invariant features, but is also invariant to illumination and rotation. Further, most of the work done on Zernike moment was done only taking into consideration its magnitude and neglecting the phase, as the phase is not rotation invariant, which results in information loss of the facial image. We not only modified it to include the phase with the magnitude, but also tuned it in such a way that besides magnitude, the phase becomes invariant to illumination, rotation, scale and translation changes. Hence the whole complex moment becomes illumination, rotation, scaling and translation invariant. Thus our LMZMPM feature preserves the maximum information of images in the form of magnitude and phase of the neighborhood. Besides being robust to noise, it also works well for heterogeneous face recognition. Due to the flexibility of LMZMPM, it can be beneficial for a range of problems as it is able to hold the basic structural information of faces, which serves as a high discriminative feature (as demonstrated by its Fisher score); and also it can be effectively used as a pre-processing algorithm. It can also be used as local invariant characteristic conserving algorithms in many biometric applications and also for hybrid approaches to meet the upcoming challenges of Face Recognition like plastic surgery, extreme rotation, disguise etc. Further, the method can be combined with other dimension reduction algorithms to compress the number of features even though retaining maximum information, with a lower order complexity.

REFERENCES

[1] S. Pankanti, R. M. Bolle, and A. Jain, "Biometrics: The future of identification [guest eeditors' introduction]," *Computer*, vol. 33, no. 2, pp. 46–49, 2000.
[2] A. Pentland and T. Choudhury, "Face recognition for smart environments," *Computer*, vol. 33, no. 2, pp. 50–55, 2000.
[3] P. Grother, G. Quinn, and P. Phillips, "Mbe 2010: Report on the evaluation of 2d still-image face recognition algorithms," *National Institute of Standards and Technology, NISTIR*, vol. 7709, p. 1, 2010.
[4] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of cognitive neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.
[5] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 7, pp. 711–720, 1997.
[6] X. He, S. Yan, Y. Hu, P. Niyogi, and H.-J. Zhang, "Face recognition using laplacianfaces," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 3, pp. 328–340, 2005.

Table XII
RESULTS ON THE AR DATABASE WITH VARIATIONS IN NOISE, ROTATION, SCALING AND TRANSLATIONS.

Methods	Noise			Rotation			Scaling			Translation		
	SN	SP	AC	SN	SP	AC	SN	SP	AC	SN	SP	AC
HF+HQ	87.8	89.5	91.5	88.7	88.9	90.3	87.8	89.5	90.3	87.8	89.5	91.5
uLBP+LM	92.3	94.3	94.8	90.7	90.3	93.7	91.9	96.4	94.3	92.3	94.3	94.8
LFA	92.5	91.5	92.5	89.4	89.2	92.5	90.2	91.5	90.0	92.5	91.5	92.5
ZZPLELD	95.7	96.5	96.8	91.0	91.4	95.7	92.5	97.7	95.7	95.7	96.5	96.8
NRLBP	98.9	96.7	97.5	90.7	93.8	93.3	91.5	96.4	95.8	95.9	98.7	97.5
MSLDE	90.5	90.0	93.4	87.5	89.3	92.4	91.0	91.8	91.6	90.5	90.0	93.4
MLGFP	97.2	94.6	94.3	89.6	89.6	94.5	92.9	93.0	93.5	97.2	94.6	94.3
LCMF	98.4	98.6	98.4	91.7	94.7	94.9	94.9	97.6	98.1	98.4	98.6	98.4
MSLZM	94.2	94.0	93.8	95.0	92.2	91.9	90.1	91.7	91.2	91.2	91.7	91.4
LZM	92.7	92.6	94.4	93.2	90.1	89.1	89.4	89.1	89.8	90.2	90.2	92.5
RF-CNN	100	99.4	100	98.1	100	98.1	96.7	98.9	100	99.3	98.6	98.9
LMZMPM	99.5	98.9	99.9	97.6	99.5	98.1	96.7	98.4	99.9	98.8	99.1	98.7
LMZMPM-CNN	100	100	100	99.1	100	99.6	98.4	99.7	100	100	100	100

Table XIII
RESULTS ON THE LFW DATABASE.

Methods	SN	SP	AC
uLBP+LM	90.6	91.7	89.2
NRLBP	90.6	91.0	89.9
ZZPLELD	89.9	91.7	91.0
PLS	89.9	90.1	89.5
I-LQP	88.2	91.7	90.5
MSLDE	85.5	88.3	87.6
MLGFP	88.4	92.0	89.4
LCMF	91.0	91.5	90.9
MSLZM	87.97	89.31	88.67
LZM	86.21	87.52	86.90
RF-CNN	96.5	96.2	96.1
LMZMPM	96.9	96.5	96.9
LMZMPM-CNN	97.12	97.24	97.44

Table XIV
RESULTS ON CUFS AND CASIA-HFB DATABASES.

Methods	CUFS			CASIA-HFB		
	SN	SP	AC	SN	SP	AC
TFSP	95.3	95.3	96.0	92.4	92.4	93.1
MLGFP	95.1	93.0	94.3	92.2	90.2	91.5
uLBP+LM	92.1	92.4	94.8	90.3	93.6	92.0
SNS-SRE	95.2	95.0	96.8	92.3	92.2	93.9
MSLDE	90.7	91.8	93.4	88.0	91.4	90.6
ZZPLELD	94.1	92.7	96.8	91.3	89.9	90.9
THFM	93.4	94.4	95.8	90.6	91.6	92.9
KP-RS	96.3	97.5	97.4	93.4	94.6	94.5
SGR-DA	95.7	97.1	97.2	92.8	94.2	94.3
SIFT+MLBP	97.5	97.9	98.1	94.6	95.0	95.2
PLS	98.2	96.8	98.4	95.3	93.9	95.4
I-LQP	97.3	97.1	97.5	94.4	94.2	94.6
LCMF	98.2	97.6	98.4	95.3	94.7	95.4
NRLBP	95.0	98.4	97.5	92.2	93.5	93.6
MSLZM	94.34	94.55	95.63	91.58	92.03	92.49
LZM	92.45	92.65	93.72	89.75	90.19	90.64
RF-CNN	99.8	99.9	100	98.1	98.9	98.5
LMZMPM	99.8	99.7	100	97.9	98.7	99.3
LMZMPM-CNN	100	99.7	100	98.51	99.44	99.56

[7] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: A convolutional neural-network approach," *IEEE transactions on neural networks*, vol. 8, no. 1, pp. 98–113, 1997.

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, pp. 1097–1105, 2012.

[9] B. Moghaddam, T. Jebara, and A. Pentland, "Bayesian face recognition," *Pattern recognition*, vol. 33, no. 11, pp. 1771–1782, 2000.

[10] H. Roy and D. Bhattacharjee, "Heterogeneous face matching using geometric edge-texture feature (getf) and multiple fuzzy-classifier system," *Applied Soft Computing*, vol. 46, pp. 967–979, 2016.

[11] H. Roy and D. Bhattacharjee, "Local-gravity-face (lg-face) for illumination-invariant and heterogeneous face recognition," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 7, pp. 1412–1424, 2016.

[12] T. Zhang, Y. Y. Tang, B. Fang, Z. Shang, and X. Liu, "Face recognition under varying illumination using gradientfaces," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2599–2606, 2009.

[13] E. Sarıyanidi, V. Dağlı, S. C. Tek, B. Tunc, and M. Gökmen, "Local zernike moments: A new representation for face recognition," in *2012 19th IEEE International Conference on Image Processing*, pp. 585–588, IEEE, 2012.

[14] E. Basaran, M. Gökmen, and M. Kamasak, "An efficient multiscale scheme using local zernike moments for face recognition," *Applied Sciences*, vol. 8, no. 5, p. 827, 2018.

[15] M. Dhekane, A. Seal, and P. Khanna, "Illumination and expression invariant face recognition," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 31, no. 12, p. 1756018, 2017.

[16] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE transactions on information theory*, vol. 8, no. 2, pp. 179–187, 1962.

[17] M. R. Teague, "Image analysis via the general theory of moments," *JOSA*, vol. 70, no. 8, pp. 920–930, 1980.

[18] G. Sansone, *Orthogonal functions*, vol. 9. Courier Corporation, 1959.

[19] G. Szeg, *Orthogonal polynomials*, vol. 23. American Mathematical Soc., 1939.

[20] A. Khotanzad and Y. H. Hong, "Invariant image recognition by zernike moments," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 12, no. 5, pp. 489–497, 1990.

[21] A. Bhatia and E. Wolf, "On the circle polynomials of zernike and related orthogonal sets," in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 50, pp. 40–48, Cambridge University Press, 1954.

[22] L. Wang and G. Healey, "Using zernike moments for the illumination and geometry invariant classification of multispectral texture," *IEEE Transactions on Image Processing*, vol. 7, no. 2, pp. 196–203, 1998.

[23] S. Li, M.-C. Lee, and C.-M. Pun, "Complex zernike moments features for shape-based image retrieval," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 39, no. 1, pp. 227–237, 2008.

[24] N. H. Foon, Y.-H. Pang, A. T. B. Jin, and D. N. C. Ling, "An efficient method for human face recognition using wavelet transform and zernike moments," in *Proceedings. International Conference on Computer Graphics, Imaging and Visualization, 2004. CGIV 2004.*, pp. 65–69, IEEE, 2004.

[25] H. R. Kanan, K. Faez, and Y. Gao, "Face recognition using adaptively weighted patch pzm array from a single exemplar image per person," *Pattern Recognition*, vol. 41, no. 12, pp. 3799–3812, 2008.

[26] H. Ouanan, M. Ouanan, and B. Aksasse, "Gabor-zernike features based face recognition scheme," *Int. J. Imaging Robot*, vol. 16, no. 2, pp. 118–131, 2015.

[27] A. Fathi, P. Alirezazadeh, and F. Abdali-Mohammadi, "A new global-gabor-zernike feature descriptor and its application to face recognition," *Journal of Visual Communication and Image Representation*, vol. 38, pp. 65–72, 2016.

[28] T. Alasag and M. Gokmen, "Face recognition in low resolution images by using local zernike moments," in *Proceedings of the International Conference on Machine Vision and Machine Learning, Beijing, China*, pp. 21–26, 2014.

[29] E. Basaran and M. Gokmen, "An efficient face recognition scheme using local zernike moments (lzm) patterns," in *Asian Conference on Computer Vision*, pp. 710–724, Springer, 2014.

[30] S. E. Kahraman and M. Gökmen, "Face pair matching with local zernike moments and l2-norm metric learning," in *2014 22nd Signal Processing and Communications Applications Conference (SIU)*, pp. 1524–1527, IEEE, 2014.
[31] B. Kaur, S. Singh, and J. Kumar, "Iris recognition using zernike moments and polar harmonic transforms," *Arabian Journal for Science and Engineering*, vol. 43, no. 12, pp. 7209–7218, 2018.
[32] Y. Zhang, X. Lu, and J. Li, "Single-sample face recognition under varying lighting conditions based on logarithmic total variation," *Signal, Image and Video Processing*, vol. 13, no. 4, pp. 657–665, 2019.
[33] B. Wang, W. Li, W. Yang, and Q. Liao, "Illumination normalization based on weber's law with application to face recognition," *IEEE Signal Processing Letters*, vol. 18, no. 8, pp. 462–465, 2011.
[34] A. R. Rivera, J. R. Castillo, and O. O. Chae, "Local directional number pattern for face analysis: Face and expression recognition," *IEEE transactions on image processing*, vol. 22, no. 5, pp. 1740–1752, 2012.
[35] A. Kar, S. Sarkar, and D. Bhattacharjee, "Local centre of mass face for face recognition under varying illumination," *Multimedia Tools and Applications*, vol. 76, no. 18, pp. 19211–19240, 2017.
[36] M. R. Faraji and X. Qi, "Face recognition under varying illumination with logarithmic fractal analysis," *IEEE Signal Processing Letters*, vol. 21, no. 12, pp. 1457–1461, 2014.
[37] Z.-R. Lai, D.-Q. Dai, C.-X. Ren, and K.-K. Huang, "Multiscale logarithm difference edgemaps for face recognition against varying lighting conditions," *IEEE transactions on image processing*, vol. 24, no. 6, pp. 1735–1747, 2015.
[38] C.-N. Fan and F.-Y. Zhang, "Homomorphic filtering based illumination normalization method for face recognition," *Pattern Recognition Letters*, vol. 32, no. 10, pp. 1468–1479, 2011.
[39] H. Roy and D. Bhattacharjee, "A novel quaternary pattern of local maximum quotient for heterogeneous face recognition," *Pattern Recognition Letters*, vol. 113, pp. 19–28, 2018.
[40] C. Peng, X. Gao, N. Wang, and J. Li, "Sparse graphical representation based discriminant analysis for heterogeneous face recognition," *Signal Processing*, vol. 156, pp. 46–61, 2019.
[41] A. Sharma and D. W. Jacobs, "Bypassing synthesis: Pls for face recognition with pose, low-resolution and sketch," in *CVPR 2011*, pp. 593–600, IEEE, 2011.
[42] Z. Lei, S. Liao, A. K. Jain, and S. Z. Li, "Coupled discriminant analysis for heterogeneous face recognition," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 6, pp. 1707–1716, 2012.
[43] X. Wang and X. Tang, "Face photo-sketch synthesis and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 11, pp. 1955–1967, 2008.
[44] X. Gao, N. Wang, D. Tao, and X. Li, "Face sketch–photo synthesis and retrieval using sparse representation," *IEEE Transactions on circuits and systems for video technology*, vol. 22, no. 8, pp. 1213–1226, 2012.
[45] N. Wang, D. Tao, X. Gao, X. Li, and J. Li, "Transductive face sketch-photo synthesis," *IEEE transactions on neural networks and learning systems*, vol. 24, no. 9, pp. 1364–1376, 2013.
[46] B. Klare, Z. Li, and A. K. Jain, "Matching forensic sketches to mug shot photos," *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 3, pp. 639–646, 2010.
[47] B. F. Klare and A. K. Jain, "Heterogeneous face recognition using kernel prototype similarities," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 6, pp. 1410–1422, 2012.
[48] J.-Y. Zhu, W.-S. Zheng, J.-H. Lai, and S. Z. Li, "Matching nir face to vis face using transduction," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 3, pp. 501–514, 2014.
[49] S. U. Hussain, T. Napoléon, and F. Jurie, "Face recognition using local quantized patterns," 2012.
[50] H. Roy and D. Bhattacharjee, "Face sketch-photo matching using the local gradient fuzzy pattern," *IEEE Intelligent Systems*, vol. 31, no. 3, pp. 30–39, 2016.
[51] H. Roy and D. Bhattacharjee, "Heterogeneous face matching using zigzag pattern of local extremum logarithm difference: Zzpleld," in *Advanced Computing and Systems for Security*, pp. 121–135, Springer, 2018.
[52] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1701–1708, 2014.
[53] O. M. Parkhi, A. Vedaldi, A. Zisserman, *et al.*, "Deep face recognition.," in *bmvc*, vol. 1, p. 6, 2015.
[54] X. Wu, R. He, Z. Sun, and T. Tan, "A light cnn for deep face representation with noisy labels," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 11, pp. 2884–2896, 2018.
[55] J. Lu, J. Hu, and Y.-P. Tan, "Discriminative deep metric learning for face and kinship verification," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4269–4282, 2017.
[56] B. K. P. Horn, "Robot vision," 2011.
[57] B. Chen, H. Shu, H. Zhang, G. Coatrieux, L. Luo, and J. L. Coatrieux, "Combined invariants to similarity transformation and to blur using orthogonal zernike moments," *IEEE Transactions on Image Processing*, vol. 20, no. 2, pp. 345–360, 2010.
[58] S. Niwattanakul, J. Singthongchai, E. Naenudorn, and S. Wanapu, "Using of jaccard coefficient for keywords similarity," in *Proceedings of the international multiconference of engineers and computer scientists*, vol. 1, pp. 380–384, 2013.
[59] W. Zhang, X. Wang, and X. Tang, "Coupled information-theoretic encoding for face photo-sketch recognition," in *CVPR 2011*, pp. 513–520, IEEE, 2011.
[60] T. Sim, S. Baker, and M. Bsat, "The cmu pose illumination and expression database teee trans. pattern analysis and machine intelligence," 2003.
[61] 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 Transactions on Pattern Analysis & Machine Intelligence*, no. 6, pp. 643–660, 2001.
[62] K.-C. Lee, J. Ho, and D. J. Kriegman, "Acquiring linear subspaces for face recognition under variable lighting," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 5, pp. 684–698, 2005.
[63] A. M. Martinez, "The ar face database," *CVC Technical Report24*, 1998.
[64] G. B. Huang, M. Mattar, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database forstudying face recognition in unconstrained environments," 2008.
[65] S. Z. Li, Z. Lei, and M. Ao, "The hfb face database for heterogeneous face biometrics research," in *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1–8, IEEE, 2009.
[66] J. Ren, X. Jiang, and J. Yuan, "Noise-resistant local binary pattern with an embedded error-correction mechanism," *IEEE Transactions on Image Processing*, vol. 22, no. 10, pp. 4049–4060, 2013.
[67] "Cuhk face sketch database (cufs)." http://mmlab.ie.cuhk.edu.hk/facesketch.html. Accessed: 2019-09-20.
[68] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre, "Xm2vtsdb: The extended m2vts database," in *Second international conference on audio and video-based biometric person authentication*, vol. 964, pp. 965–966, 1999.
[69] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.
[70] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *The American Statistician*, vol. 46, no. 3, pp. 175–185, 1992.
[71] R. Trevethan, "Sensitivity, specificity, and predictive values: foundations, pliabilities, and pitfalls in research and practice," *Frontiers in public health*, vol. 5, p. 307, 2017.

Arindam Kar is an Associate Scientist in Indian Statistical Institute, Kolkata, India. He received his MSc and MTech degrees from Indian Institute of Technology (Kharagpur), India in 1998 and 2000 respectively. He received his Ph.D. from Jadavpur University, India, in 2015. His research interests include computer vision, image processing, genetic algorithms, face recognition, medical imaging and watermarking techniques.

Sourav Pramanik (S'17) is a Visvesvaraya PhD-fellow in the Department of Computer Science and Engineering, Jadavpur University, Kolkata, India. He received his ME degree in Computer Science and Engineering from Jadavpur University, India in 2011. His current research interests are thermal breast image processing and Machine Learning.

Arghya Chakraborty is currently pursuing his B. Stat. degree in Statistics from the Indian Statistical Institute, Kolkata. His research interests include mathematical probability, supervised and unsupervised machine learning, recognition and classification.

Debotosh Bhattacharjee (SM'11) is working as a full professor in the Department of Computer Science and Engineering, Jadavpur University with fourteen years of post-PhD experience. His research interests pertain to the applications of machine learning techniques for Face Recognition, Gait Analysis, Hand Geometry Recognition, and Diagnostic Image Analysis. He has authored or coauthored more than 250 journals, conference publications, including several book chapters in the areas of Biometrics and Medical Image Processing. Two US patents have been granted on his works. Prof. Bhattacharjee has been granted sponsored projects by the Govt. of India with a total amount of around INR 2 Crore.

Edmond S. L. Ho is a Senior Lecturer with the Department of Computer and Information Sciences, Northumbria University, UK. He received the BSc degree in Computer Science from the Hong Kong Baptist University, in 2003, the MPhil degree in Computer Science from the University of Edinburgh, Scotland, in 2006 and the PhD degree in Informatics from the University of Edinburgh, Scotland, in 2011. He was a Research Assistant Professor with Hong Kong Baptist University in 2011-2016. His research interests include computer graphics and animation, computer vision, machine learning, and robotics.

Hubert P. H. Shum is an Associate Professor (Reader) and the Director of Research and Innovation of the Computer and Information Sciences Department at Northumbria University, UK. Before this, he worked as a Senior Lecturer at Northumbria University, UK, a Lecturer in the University of Worcester, UK, a post-doctoral researcher in RIKEN, Japan, as well as a research assistant in the City University of Hong Kong. He received his PhD degree from the School of Informatics in the University of Edinburgh, U.K. His research interests include computer graphics, computer vision, motion analysis and machine learning.