Face Recognition System Using Local Features Fusion for Multi-Masks

Mustafa Zuhaer Nayef Al-Dabagh\textsuperscript{1,2}* and Muhammad Imran Ahmad\textsuperscript{1}
\textsuperscript{1}Faculty of Electronic Engineering Technology, Universiti Malaysia Perlis, Perlis, Malaysia.
\textsuperscript{2}Department of Computer Science, College of Science, Knowledge University, Kirkuk Road, 44001 Erbil, Iraq.

*mustafa.zuhaer@knu.edu.iq

Abstract. Face recognition is a relatively novel research field, and its application is closely related to numerous other areas. Moreover, it is emerging as a critical research theme due to its broad range of applications. Thus, many face recognition methods use a variety of feature extraction approaches. Nonetheless, the issue continues to be challenging, particularly identifying non-biological entities. This paper proposes an extended descriptor for local features of an effectual facial recognition system using a local directional pattern operator. This technique combines the Frei-Chen and Robinson masks' strengths by fusion of the directional features of LDP for these two masks; this elicits a robust feature extraction method for distinguishing faces. Experimental results using the Yale database show that the extended descriptor considerably improved recognition rate and better performance than traditional local feature descriptors.

1. Introduction
As a typical application of image analysis and interpretation, face recognition has drawn considerable focus in various security applications [1] and research domains [2]. However, even though several commercial systems have surfaced, it is still a dynamic and stimulating theme, as diverse aspects like deviations could quickly impact the facial images in lighting conditions, aging, pose, alignment, and occlusion [3].

Face recognition consists of several significant steps, such as feature extraction, classification, and decision module. Most of the proposed face recognition system does not employ the feature selection algorithm; therefore, the classifier's features may include noise and redundant information. Different features can describe various characteristics in the image, and recognition of which features can be more helpful is challenging; therefore, necessitating many methods to overcome this challenge. Since the information about the feature representation influences the classification rate, the feature set must be distinguished because it is the most critical aspect for an effective facial recognition engine [4]. Even the finest classifier will not deliver a good performance given varying or insufficient features. Therefore, devising a robust feature extraction technique that could perform steadily in a fluctuating environment is still challenging.

The two methodologies for extracting facial features of images use geometric facial features and appearance-based features techniques. The calculation of the location and shapes of diverse facial components, such as mouth, nose, and eye, uses the geometric feature-based feature. Those facial elements generate the feature vector representing the face's geometry [5]. The redundancy and dimensionality of the facial features have a direct impact on face recognition precision. Of the various methods launched so far, a general geometric approach is the facial action coding system (FACS) [6],
which identifies facial expression with the aid of a suite of action units (AU). Every action unit matches up to the physical facet of a specific facial muscle. In addition, many researchers also examined fiducial point-based representations [7–8]. Nonetheless, the efficacy of geometric approaches is severely reliant on facial constituents' precise detection, which is challenging in varying environments. Therefore, geometric feature-based approaches are not easy to deploy in several real-world situations [9].

In appearance-based features, the input comprises the parameters specific to the extracted shape image [10]. A face is expressed as numerous raw intensity pictures of which one picture has a multidimensional vector form using appearance-based techniques. Image distribution is used to compute a feature space by employing statistical methods. However, not all the features in the feature vector space are beneficial. A few of the extensively utilized appearance-based approaches are principal component analysis (PCA) [11], independent component analysis (ICA) [12, 13], and Gabor wavelets [14, 15]. Even though ICA and PCA feature descriptors can efficiently encapsulate the training images' changeability, their performances wane in altering the environment [10, 16]. In contrast, identifying Gabor characteristics using several Gabor filters to convolution face pictures using different orientations and proportions is computationally intensive.

Recently, local binary patterns (LBP) [17] based on local appearance labels and their different forms [18] have been of great interest because of their performance effectiveness in uncontrolled settings. The LBP operator coded the local image texture using quantization of the grey levels corresponding to a neighbor in the local neighborhood relative to the central value; consequently, the binary pattern forms a model for micro-scale data like spots and corners edges. Nevertheless, in the case of random noise or significant illumination changes, the LBP method has poor performance [4] because a minor change in grey levels can lead to a high degree of change to the LBP code. The local ternary pattern (LTP) [16] was proposed later to enhance the flexibility of LBP specific to the near-uniform and uniform areas by augmenting with addition-al intensity discrimination and transforming the binary LBP code to a ternary form. Lately, the Sobel-LBP [19] technique has been suggested to enhance LBP performance by using the Sobel operator to augment edge-specific data before starting the LBP feature extraction process. Nevertheless, if near-uniform and uniform areas are processed through the Sobel operator, the patterns have discrepancies since the discrimination state is binary, like LBP. Local directional pattern (LDP) [20, 15] used a different technique for encoding textures; rather than using grey levels, the directional edge response data corresponding to a position is chosen. Despite this approach's higher recognition effectiveness, LDP typically outputs patterns with discrepancies in the uniform or near-uniform facial areas; furthermore, it depends extensively on choosing the prominent edge direction factor [21]. Weber's law-based local multiple patterns (LMP) [22] is demonstrated to produce robust and highly distinctive visual features.

The main contributions of this work can be summed up as follows:

- First, an extended LDP operator is proposed for face recognition: The proposed operator combines LDP's extraction feature of Robinson and Frei-Chen to get more discriminative features.
- This paper offers a broad comparison between the recommended descriptor and several state-of-the-art approaches.
- Yale database was selected for conducting comprehensive experiments. As a result, the proposed method has been given good results when matched against the state-of-the-art assessed approaches.

The rest of the paper is structured as follows: section 2 outlines a review of LBP and LDP methods. Then, the proposed system presents in section 3. The experimental outcomes and comparisons are stated in section 4. Finally, section 5 presents the conclusion.

2. A Review of LBP and LDP

Local binary pattern (LBP) is a modest yet effective local texture description method. LBP was presented initially by Ojala et al. [23] for grayscale and rotation-invariant texture analysis. Several researchers have since effectively espoused LBP in diverse face-related issues like face recognition [24] and facial expression analysis [17]. The basic LBP technique selects the local neighborhood
corresponding to all image pixels and uses the central reference to create thresholds for the neighboring grey levels. Subsequently, there is a binomial concatenation of the results; the resulting magnitude is assigned to the central pixel. The LBP operator can be signified as:

\[
LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(i_p - i_c)2^p
\]

(1)

\[
s(x) = \begin{cases} 
1, & x \geq 0 \\
0, & x < 0 
\end{cases}
\]

(2)

Here, \(i_c\) is the grey value of the center pixel (\(x_c, y_c\)), \(i_p\) is the grey value of the surrounding neighbors, \(P\) is the total number of neighbors, and \(R\) is the neighborhood’s radius. The grey level corresponding to a neighboring pixel is determined using bilinear interpolation in case the pixel position is not perfectly aligned. Subsequently, the image block or the LBP-encoded histogram of the image may be used for feature description. One restriction with the LBP encoding technique is the susceptibility of LBP codes to noise because a minute difference in the intensity value of the neighbors may totally change the corresponding binary code. Figure (1) describes the fundamental LBP encoding technique.

![Figure 1. Fundamental LBP encoding technique [25].](image)

Using a Local Directional Pattern (LDP) suggests a solution to this coding problem. The local directional pattern comprises the encoding of direction-specific data about facial textures, leading to a relatively compact discriminative code compared to the presently used techniques. First, compass masks are employed to ascertain directional knowledge and calculate the micro-pattern structure. Subsequently, this information is encoded with the use of conspicuous direction values. This technique facilitates the separation of homogenous patterns having distinct intensity transitions [26], [27].

### 3. Proposed System

In this work, the face recognition system is presented based on feature fusion using different compass masks. The system consists of four stages: preprocessing, feature extraction, feature fusion, and classification. Preprocessing aims to remove the noise from the input image. In the feature extraction stage, LDP features represent different and complementary characteristics in the image. In order to get a more robust feature, concatenative feature fusion is used in feature fusion. Finally, identifying the input face image is identified in the classification step based on feature vectors resulting from the previous stage. For doing the classification, K-nearest neighbor (KNN) is used. The design of the proposed system is displayed in the figure below:
3.1. Preprocessing
Before feature identification necessitates preprocessing, this step comprises several tasks formulated to handle local shadows, changes in illumination, and highlights while ensuring that the elements’ visual appearance remains primarily similar. In detail, the steps are as follows:

3.1.1. Gamma Correction
It corresponds to a nonlinear grey-level conversion that changes grey-level $I$ with $I^\gamma (\gamma \geq 0)$ or $\log(I)$ ($\gamma = 0$), where $\gamma \in [1,0]$ denotes a user-specified parameter. Thus, it improves the image's dynamic range from a local standpoint is less illuminated or shadowed areas while using compression corresponding to the brighter areas and highlights [19].

3.1.2. A Difference of Gaussian (DoG) Filtering
The effects of general intensity gradients such as shading are not eliminated using gamma correction. Shading provided by surface effects is likely to be a helpful visual indication; however, considering that it typically has low spatial frequency data, it is challenging to isolate it created by illumination gradients. High-pass filters allow the elimination of incidental and valuable information; consequently, noise and aliasing effects are reduced without distortion of a significant portion of the underlying image. The Difference of Gaussian (DOG) technique can enhance visibility around the edges and additional detail contained in a digital image. High-frequency comprise noise, and they can be moved using the Difference of Gaussian (DOG) method; the technique is appropriate for converting images that have substantial noise. The impulse response for DOG is specified as:

$$D(x, y) = \frac{1}{2\pi\sigma_1^2} e^{-\frac{x^2 + y^2}{2\sigma_1^2}} - \frac{1}{2\pi\sigma_2^2} e^{-\frac{x^2 + y^2}{2\sigma_2^2}}$$

(3)

Where the typical allocations of $\sigma_1$ and $\sigma_2$ are selected as 0.3 and 2.0, respectively. Considering that the effect causes the overall contrast to be reduced, it is essential to enhance subsequent steps [19].

3.1.3. Contrast Equalization
The objective is to change the image intensity scales to provide standardization to the intensity or contrast changes. It is crucial to employ a flexible estimator since the input may have outlier values

---

Figure 2. The Design of Proposed System.
created by minute dark areas or highlights created by aspects like nostrils or garbage near image contours, among other aspects. It is feasible to employ the absolute value of the signal using the median; however, in this case, a straightforward and rapid approximation is considered.

\[
I(x, y) = \frac{I(x, y)}{\left(\text{mean}(|I(x', y')|)^\alpha\right)}
\]

Here, \(\alpha\) denotes a decisive compressive factor that leads to the lesser influence of the relatively large values, \(\tau\) denotes a boundary value required to trim large values once the first normalization step is completed and the unmasked areas of the image comprise the mean value [19].

### 3.2. Feature Extraction

This step aims to determine a set of representative features from the images in feature vectors; to ascertain the face image's significant features, the features of the LDP operator depending on Frei-Chen and Robinson masks were combined in a single feature vector to get more discriminated and detailed information. The LDP corresponding to a pixel comprises an eight-bit binary sequence calculated by contrasting the edge response output for different vectors in the local 3 x 3 size neighborhood; considering a central pixel of the image, the edge response in the eight-directional form \(|m|\) \((i=0, 1\ldots 7)\). The LDP is calculated using Kirsch masks, Frei-Chen, or Robinson; the pixel position is considered central, and the eight orientations are centered on that position. Kirsch, Frei-Chen, or Robinson compass masks are used for finding edges in all eight directions of a compass. Subsequently, the first \(k\) responses \(|m|\) \((i=0, 1\ldots 7)\) are chosen, and the associated directional bits are provided with the value 1 while the other \((8-k)\) bits are provided with the value zero. Hence, the LDP code is derived by:

\[
LDP_K = \sum_{i=0}^{7} b_i(m_i - m_k)2^i
\]

\[
b_i(a) = \begin{cases} 
1, & a \geq 0, \\
0, & a < 0.
\end{cases}
\]

Here, the \(k^{th}\) most significant directional output is denoted by \(m_k [28]\).

### 3.3. Feature Fusion

LDP's feature vectors based on difference masks are fused into a single vector using concatenative feature fusion to obtain more discriminative features.

\[
V_{\text{compact}} = \begin{bmatrix} V_f^T \\ V_s^T \end{bmatrix}^T \in \mathbb{R}^D
\]

Where \(V_{\text{compact}}\) denotes the fused feature vector, \(V_f\) denotes the first feature vector, and \(V_s\) denotes the second feature vector, and \(D\) is the dimensionality of the fused feature vector. Thus, combining two feature vectors into a single one possesses higher discriminative power and improves results over methods that use either LDP based on Kirsch, Frei-Chen, or Robinson or alone [29].

### 3.4. Classification

The system uses the K-NN technique for feature classification based on k nearest neighbors’ corresponding to the feature space’s vectors; to comprehensively explain this technique, \(z\) is the chosen classified point. It is determined by computing the k neighbors that have the least distances concerning \(z\) and the other data elements, as depicted below:

\[
\{d_1(z, z_{c_1 i_1}), \ldots, d_k(z, z_{c_k i_k})\} = \min_{1 \leq c \leq 1, i \leq n_c} d(z, z_{ci})
\]

Considering the above expression, \(n_c\) represents the samples in every class. The calculation of distance is done using the following expression:
\[ \sum_{c=1}^{C} n_c \]  

(10)

Numerous parameters can be employed in the K-NN method, such as k that denotes the nearest neighbors’ number and other parameters pertaining to the distance model. In this system, the cosine distance is used in the experiments to get the most similar k number of images to the input image and predict the input ratings based on the information of the chosen neighbors [2].

4. Experiments and Results Analysis

A benchmark image database, namely the Yale database, was used to determine the formulated feature descriptor's effectiveness. First, images were selected to test the effectiveness of the proposed framework. Then, Yale University created the Yale face database [30], consisting of data specific to fifteen individuals for 165 GIF format greyscale images. Every subject has 11 representative images; there is one image for every facial expression or setting such as w/no glasses, w/glasses, center-light, normal, sad, sleepy, left-light, right-light, happy, wink, and surprised. The data is split into the training, testing, and evaluating sets. The database is randomly divided into training and test sets. The first five images from the first twelve classes are used for training (60 images), while the remaining six images from the first twelve classes, as well as all the images from classes (13, 14, 15), are applied for testing (105 images). In this way, 64% of the database images were used for training, and 36% were used for testing. The proposed technique and some commonly-used techniques demonstrate the salient features of this technique. Figure (3) shows a comparison of different methods using the Yale database.

![Figure 3. Comparison of Recognition rate (%) on Yale database.](image)

This proposal was implemented on a PC featuring a Core i7-9750H 2.6 GHz CPU and 16GB RAM and Windows 10 OS. using Matlab 2019b. All experiments comprised the use of the KNN classifier. These experiments show that the proposed system achieved the best accuracy, 88.57%. Table 1 lists the difference in recognition rates for the images in the Yale database. From the result, the feature fusion can increase separation power in the feature space. This is because one method can carry unique information; thus, combining two methods will produce better discrimination at the early stage.
Table 1. Recognition accuracy (%) for the Yale database

| Seq. | Method                          | Accuracy (%) |
|------|---------------------------------|--------------|
| 1    | LBP [23]                        | 81.9048      |
| 2    | LDP based (Kirsch)              | 82.8571      |
| 3    | LDP based (Frei-Chen)           | 87.6190      |
| 4    | LDP based (Robinson)            | 65.7143      |
| 5    | LDP based (Kirsch+Frei-Chen)    | 83.8095      |
| 6    | LDP based (Kirsch+Robinson)     | 81.9048      |
| 7    | LDP based (Frei-Chen+Robinson)  | 88.5714      |
| 8    | Monogenic Binary Coding[31]     | 79.0476      |
| 9    | Median Binary Pattern [32]      | 80.00        |
| 10   | Local Monotonic Pattern [33]    | 80.9524      |
| 11   | Local Frequency Descriptor[34]  | 79.0476      |
| 12   | Local Arc Pattern [35]          | 77.1429      |
| 13   | Improved Weber Binary Coding [36]| 80.00        |
| 14   | Local Gradient Pattern [37]     | 75.2381      |

5. Conclusion

This paper provides an extended local texture pattern for Facial Recognition. First, distinctive facial aspects are obtained using all the input images by employing the LDP-based masks having different types (Kirsch, Frei-Chen, and Robinson); these are subsequently integrated to yield a single vector. Finally, KNN processing is performed for additional classification. The Yale face database evaluation indicates that the proposed technique provides an acceptable recognition rate better than other techniques. Furthermore, the results prove that feature integration using proposed techniques leads to higher algorithmic effectiveness and enhanced recognition performance.

References

[1] Arafah M, Achmad A, Indrabayu and Areni I S 2020 Face Identification System Using Convolutional Neural Network for Low Resolution Image 2020 IEEE International Conference on Communication, Networks and Satellite, Compassat 2020 - Proceedings (Institute of Electrical and Electronics Engineers Inc.) pp 55–60
[2] Yee S Y, Rassem T H, Mohammed M F and Awang S 2020 Face recognition using Laplacian completed local ternary pattern (LapCLTP) Advances in Electronics Engineering (Springer) pp 315–27
[3] Nayef Al-Dabagh M Z, Imran Ahmad M, Md Isa M N and Amirul Anwar S 2020 Face Recognition System Based on Fusion Features of Local Methods Using CCA 2020 8th International Electrical Engineering Congress (iEECON) pp 1–4
[4] Arachchilage S P K W and Izquierdo E 2020 Deep-learned faces: a survey EURASIP J. Image Video Process. 2020 1–33
[5] M. M and A. M 2021 Facial geometric feature extraction based emotional expression classification using machine learning algorithms PLoS One 16 1–20
[6] Mo F, Zhang Z, Chen T, Zhao K and Fu X 2021 MFED: A Database for Masked Facial Expression IEEE Access 9 96279–87
[7] Malathi D, Mathangopi A and Rajinigirinath D 2019 Fiducial Point Location Algorithm for Automatic Facial Expression Recognition International Journal of Trend in Scientific Research and Development pp 290-293
[8] Salem E, Hassaballah M, Mahmoud M M and Ali A-M M 2021 Facial Features Detection: A Comparative Study The International Conference on Artificial Intelligence and Computer Vision (Springer) pp 402–12
[9] Shi Y, Zhang Y and Harik R 2020 Manufacturing feature recognition with a 2D convolutional neural network CIRP J. Manuf. Sci. Technol. 30 36–57
[10] Anwarul S and Dahiya S 2020 A comprehensive review on face recognition methods and factors affecting facial recognition accuracy Proc. ICRIC 2019 495–514
[11] Maheswari V, Sari C A and Rachmawanto E H 2020 Study Analysis of Human Face Recognition using Principal Component Analysis 2020 International Seminar on Application for Technology of Information and Communication (iSemantic) (IEEE) pp 55–60
[12] Tharwat A 2021 Independent component analysis: An introduction Appl. Comput. Informatics 17 222–49
[13] Kortli Y, Jridi M, Merzougui M, Alasiry A and Atri M 2020 Comparative Study of Face Recognition Approaches 2020 4th International Conference on Advanced Systems and Emergent Technologies (IC_ASET) (IEEE) pp 300–5
[14] Kortli Y, Jridi M, Al Falou A and Atri M 2020 Face recognition systems: A survey Sensors 20 342
[15] Ahmed S, Frikha M, Hussein T D H and Rahebi J 2021 Optimum Feature Selection with Particle Swarm Optimization to Face Recognition System Using Gabor Wavelet Transform and Deep Learning Biomed Res. Int. 2021
[16] Zhao S, Gao Y and Zhang B 2008 Sobel-LBP Proceedings - International Conference on Image Processing, ICIP
[17] Shan C, Gong S and McOwan P W 2009 Facial expression recognition based on local binary patterns: A comprehensive study Image Vis. Comput. 27 803–16
[18] Zhao G and Pietikäinen M 2009 Boosted multi-resolution spatiotemporal descriptors for facial expression recognition Pattern Recognit. Lett. 30 1117–27
[19] [Tan X and Triggs B 2010 Enhanced local texture feature sets for face recognition under difficult lighting conditions IEEE Trans. image Process. 19 1635–50
[20] Jabid T, Kabir M H and Chae O 2010 Robust facial expression recognition based on local directional pattern ETRI J. 32 784–94
[21] Ahmed F and Kabir M H 2012 Directional ternary pattern (DTP) for facial expression recognition 2012 IEEE International Conference on Consumer Electronics (ICCE) (IEEE) pp 265–6
[22] Yang W, Zhang X and Li J 2020 A local multiple patterns feature descriptor for face recognition Neurocomputing 373 109–22
[23] Ojala T, Pietikäinen M and Maenpaa T 2002 Multiresolution gray-scale and rotation invariant texture classification with local binary patterns IEEE Trans. Pattern Anal. Mach. Intell. 24 971–87.
[24] Ahonen T, Hadid A and Pietikäinen M 2006 Face description with local binary patterns: Application to face recognition IEEE Trans. Pattern Anal. Mach. Intell. 28 2037–41
[25] Lindahl T 2007 Study of Local Binary Patterns Sci. Technol
[26] Longmore C A, Liu C H and Young A W 2015 The importance of internal facial features in learning new faces J. Exp. Psychol. 68 249–60
[27] Rivera A R, Castillo J R and Chae O O 2012 Local directional number pattern for face analysis: Face and expression recognition IEEE Trans. image Process. 22 1740–52
[28] Luo Y-T, Zhao L-Y, Zhang B, Jia W, Xue F, Lu J-T, Zhu Y-H and Xu B-Q 2016 Local line directional pattern for palmprint recognition Pattern Recognit. 50 26–44
[29] Heracleous P, Even J, Ishi C T, Miyashita T and Hagita N 2012 Fusion of standard and alternative acoustic sensors for robust automatic speech recognition 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (IEEE) pp 4837–40
[30] Belhumeur P N, Hespanha J P and Kriegman D J 1997 Eigenfaces vs. fisherfaces: Recognition using class specific linear projection IEEE Trans. Pattern Anal. Mach. Intell. 19 711–20
[31] Xia X, Ying Z L and Chu W J 2014 Facial Expression Recognition Based on Monogenic Binary Coding Applied Mechanics and Materials vol 511 (Trans Tech Publ) pp 437–40
[32] Bashar F, Khan A, Ahmed F and Kabir M H 2014 Robust facial expression recognition based on median ternary pattern (MTP) 2013 International Conference on Electrical Information and Communication Technology (EICT) (IEEE) pp 1–5
[33] Mohammad T and Ali M L 2011 Robust facial expression recognition based on Local Monotonic Pattern (LMP) 14th International Conference on Computer and Information Technology (ICCIT 2011) (IEEE) pp 572–6
[34] Lei Z, Ahonen T, Pietikäinen M and Li S Z 2011 Local frequency descriptor for low-resolution face recognition Face and Gesture 2011 (IEEE) pp 161–6
[35] Islam M S and Auwatanamo S 2014 Facial expression recognition using local arc pattern *Trends Appl. Sci. Res* **9** 113–20

[36] Yang B, Zhang T, Gu C, Wu K and Guan X-P 2015 A novel face recognition method based on IWLD and IWBC *Multimed. Tools Appl.* **75** 6979–7002

[37] Islam M S 2014 Local gradient pattern-A novel feature representation for facial expression recognition *J. AI Data Min.* **2** 33–8