Component-based face recognition using statistical pattern matching analysis

Sushil Kumar Paul¹ · Saida Bouakaz² · Chowdhury Mofizur Rahman³ · Mohammad Shorif Uddin⁴

Received: 5 July 2019 / Accepted: 9 June 2020 / Published online: 18 July 2020
© Springer-Verlag London Ltd., part of Springer Nature 2020

Abstract

The aim of this research is to develop a fusion concept to component-based face recognition algorithms for features analysis of binary facial components (BFCs), which are invariant to illumination, expression, pose variations and partial occlusion. To analyze the features, using statistical pattern matching concepts, which are the combination of Chi-square (CSQ), Hu moment invariants (HuMIs), absolute difference probability of white pixels (AbsDifPWPs) and geometric distance values (GDVs) have been proposed for face recognition. The individual grayscale face image is cropped by applying the Viola–Jones face detection algorithm from a face database having variations in illumination, appearance, pose and partial occlusion with complex backgrounds. Doing illumination correction through histogram linearization technique, the grayscale face components such as eyes, nose and mouth regions are extracted using the 2D geometric positions. The binary face image is created by applying cumulative probability distribution function with Otsu adaptive thresholding method and then extracted BFCs such as eyes, nose and mouth regions. Five statistical pattern matching tools such as the standard deviation of CSQ values with probability of white pixels (PWPs), standard deviation of HuMIs with Hu’s seven moment invariants, AbsDifPWPs and GDVs are developed for recognition purpose. GDVs are determined between two similar facial corner points (FCPs) and nine FCPs are extracted from binary whole face and BFCs. Pixel Intensity Values (PIVs) which are determined using L_2 norms from grayscale values of the whole face and grayscale values of the face components. Experiment is performed using BioID Face Database on the basis of these pattern matching tools and appropriate threshold values with logical and conditional operators and gives the best expected results from true positive rate perspective.

Keywords Binary facial component · Histogram linearization technique · Otsu thresholding · Probability of white pixels · Hu’s moment invariants · Facial corner points

✉ Sushil Kumar Paul
paulksushil@yahoo.com

Saida Bouakaz
saida.bouakaz@univ-lyon1.fr

Chowdhury Mofizur Rahman
cmruu@gmail.com

Mohammad Shorif Uddin
shorifuddin@gmail.com

¹ Controller of Examinations, Bangladesh Technical Education Board, Agargaon, Dhaka 1207, Bangladesh

² LIRIS Lab, Université Claude Bernard Lyon1, 69622 Villeurbanne Cedex, France

³ Department of Computer Science and Engineering, United International University, Badda, Dhaka 1212, Bangladesh

⁴ Department of Computer Science and Engineering, Jahangirnagar University, Dhaka 1342, Bangladesh

1 Introduction

Human face recognition has the most significant role for identifying a person in many real-world applications in computer vision like identification, authentication, security, surveillance system, human–computer interaction, anti-terrorism, psychology and so on. It is not an intrusive technique (i.e., not convey any health risks like corona virus), and it does not need to touch anything during the acquisition level. Moreover, face is a rich source of non-verbal information about human behavior [1]. Due to complex and multidimensional structure of face, it requires enormous process and computation. On the other hand, face perception is the most developed visual perceptual ability in human beings. Infants desire to look at faces and try to recognize faces from shortly after birth [2]. Most people

consume more time looking at faces than at any other type of objects. Humans have not the ability to recognize unlimited number of different faces, but computer can do it because computer has a bulk amount of memory and processing capabilities. Although automatic face recognition is not a new topic, the challenge of developing an appropriate face recognition algorithm remains unsolved. The main problems of face recognition are illuminations, different expressions, lighting variations, different poses, partial occlusion, complex and multidimensional structure of face and cluttered background.

Face recognition is the technique to identify one or more person in images. Algorithms for face recognition typically extract facial features such as facial corner points and facial component values and compare them to a database to find the best match. Facial features extraction task is the initial stage for face recognition in the field of computer vision. The most significant facial features are facial corner points, eyes, nostrils and mouth areas. Facial corner points are eyes corners, nostrils, nose tip and mouth corners. Eyes, nostrils and mouth areas are also the key feature regions for face recognition. Using all of these information with some appropriate pattern, matching tools are accomplished the recognition decision [3–8]. Face recognition algorithm complies with appropriate matching to identify a similar face and reject the dissimilar face images from the face database.

Face recognition techniques are mainly classified into two principal categories: appearance based and model based. In appearance-based techniques, facial feature extraction concept, are holistic (global), facial components information (local) and fusion (hybrid) approaches play the vital roles for the classification and recognition of faces. In global (holistic) approach, entire information is extracted with a single vector from the whole face image. In local feature-based methods, facial component features such as eyes, nose, mouth, etc. are extracted from face images. Eyes are robust to facial expressions and occlusions because of its constant inter-ocular distance among people and unaffected by mustache or beard [7, 9]. Nose indicates the head pose and its nostrils and nose tip are the symmetry points of both right and left side face regions and also robust to facial expressions. Mouth also is a key component of face recognition, while mouth length and lips convey facial expression information [7].

Feature-based approach can perform with diverse imaging conditions to achieve better performance. Fusion (hybrid) method is the combination of both global and local concepts and applied for achieving the desired outcome. So, holistic features may not be used if the exact and accurate facial component features extraction is possible [1].

Various holistic approaches used for face recognition, such as Eigenfaces [10], Fisherfaces [11], Independent Component Analysis (ICA) [12], Moment Invariants [13],

Discrete Cosine Transform (DCT) [14], etc. Moreover, component feature-based techniques are described by facial components with various ideas, for examples, components with support vector machine (SVM) [15], LDA [16], 3D models [17], etc.

Zhang et al. [18] proposed a holistic-based illumination invariant face recognition approach used only high-frequency components and some useful identity information are lost due to discarding of the low-frequency components. Wada et al. [19] proposed a holistic-based expression invariant face recognition system, which used a constrained optical flow algorithm. It needs high computational cost as well as the obtained results are not shown satisfactory.

As a model-based approach, Weyrauch et al. [17] constructed a component feature-based algorithm using 3D morphable models. It took fourteen component features but applied only nine components due to high computational complexity. Bonnen et al. [20] developed a component-based technique utilizing heterogeneous concepts, which is too much intricate. Hua [21] developed a pose invariant component-based face recognition technique, but it needs high computational cost and does not produce good performance. Turk and Pentland [10] used Eigenfaces, but it cannot incorporate additional training data into an existing PCA projection matrix and not invariant to change in shape, pose and expression [11, 22]. Rajiv Kapoor and Pallavi Mathur [23] developed three kinds of moment invariants including Hu moment invariants and achieved poor result. Harguess and Aggarwal [24] demonstrated average-half-face with multiple concepts used full frontal view face applied six types of face recognition algorithms and shown not so good results. Nabatchian et al. [13] applied nine moment invariants (MIs) but got poor results except Pseudo-Zernike Moment Invariants (PZMI). Brunelli and Poggio [25] applied geometrical feature and template-matching concepts in frontal view face with the same illumination condition and applied facial component features. Sushil et al. developed the component-based face recognition algorithms using statistical pattern matching concepts and does not show good result as the TPR perspective [3, 4].

The above-mentioned holistic and model-based techniques need high computational cost and are not suitable due to variations in illumination, expression, pose and partial occlusion. Therefore, it is very much indispensable to develop a less computation cost component-based face recognition algorithms that are able to include new feature, discard or modify existing feature to summarize similarity decision as a fusion concept if any one facial component feature of a test face is matched with corresponding facial component feature of reference database to establish an invariant properties of different illumination, expression, pose variation and partial occlusions for better performance.

The main objective of this research activity is that the different pattern matching algorithms are recognized the same images as well as different images of the same individual using entire face database. Therefore, combination of multiple algorithms will be provided the same percent of TPR values for each individual.

The purposes of this research work are: (1) to extract some crucial facial features such as eyes, nostrils and mouth regions as the key components and also to apply shape invariant concept of these components, (2) to develop an invariant technique of multiple imaging situations such as illumination, expression, pose variations and partial occlusion, (3) to analyze of facial features that which feature need to be included or discarded for achieving better performance as TPR perspective, (4) to need simple mathematical calculation for constructing desired pattern matching tools and (5) to use two or more pattern matching tools can be formed new or expected recognition algorithms. As a fusion-based similarity technique, the five statistical pattern matching tools such as standard deviation of CSQ values [3], standard deviation of Hu moment invariants (HuMIs) [4], Geometric Distance Values (GDVs) [5–8], absolute difference probability of white pixels (AbsDifPWPs) [3, 4] and Pixel Intensity Values (PIVs) [3, 4] are constructed for face recognition algorithms.

In this research activity the facial component feature extraction concept is based on BFCs that are mainly employed on the three independent methods such as probability of white pixels (PWPs) with CSQ formula [3], shape similarities with Hu moment invariants (HuMIs) [4] and geometric distance values (GDVs) of nine similar FCPs between test face and reference database images using each of individual facial component and whole face [5–8].

The PWPs depend on facial appearance and expression but independent on illumination variations. The PWPs are calculated from the number of white pixels divided by total number of pixels of each binary face component. To establish matching technique without complex processing, only normalization idea (sum of probability of white and black pixels of each binary facial component is equal to one) is applied to confirm more invariance of facial appearance and expression. Matching of each component is done with the help of Chi-square (CSQ) formula [3].

The shape of all images of each individual (person) is same, but the shape of different people is different. The shape similarity as a statistical pattern matching concepts have been established using Hu's seven nonlinear moment functions for recognition purpose due to its invariance characteristics. A BFC will be invariant if its moment value shows constant to alter in scale, rotation, translation or/and reflection in each image component to ensure the independent of appearance, expression, pose and lighting variations, and it is computationally efficient to get better

outcome because it does not require any prior information about the face model [4].

The geometric distance values (GDVs) is used for recognition activity, and it is calculated from the automatic extraction of nine similar FCPs, such as both eyes corner points, mouth corner points, nostrils and nose tip between test face and reference database images [5–8]. The benefit of FCPs is that it is completely independent of facial appearance, expression and illumination, pose variations and partial occlusion and requires the least computational cost [8].

The absolute difference of probability of white pixels (AbsDifPWPs) is calculated using binary test face and binary reference database images. The pixel intensity values (PIVs) between grayscale test face and reference database are determined using $\mathcal{L}_2$ norm is known as Euclidean distance formula [3, 4, 8].

In this paper, we have shown six types of recognition algorithms such as: (i) Chi-Square (CSQ) [3], (ii) Chi-Square with Facial Corner Points (CSQ + FCPs) [8], (iii) Hu Moment Invariants (HuMIs) [4], (iv) Hu Moment Invariants with Facial Corner Points (HuMIs + FCPs), (v) Chi-Square and Hu Moment Invariants (CSQ + HuMIs) and (vi) Chi-Square and Hu Moment Invariants with Facial Corner Points (CSQ + HuMIs + FCPs) methods. Five components such as whole face, both eyes, nose and mouth regions are used from both grayscale and binary images [3, 4]. Algorithms (i) to (iii) are existing methods and algorithms (iv) to (vi) are to be developed for features analysis.

The major contributions of this paper are: (1) Construction of an invariant technique of multiple imaging conditions and shape invariant concept using binary facial components, (2) Development of a new recognition algorithm by the combination of two or more pattern matching tools, (3) Making a decision framework for finding an appropriate pattern matching tool on the basis of TPR perspective and (4) Performance analysis to use of each pattern matching tool.

The remainder of this work is organized accommodating the following six sections. Sections 2 and 3 describe the preprocessing and processing works, respectively. Determination of five types of statistical pattern matching tools is described in Sect. 4. The similarity evaluation and recognition decisions are performed in Sect. 5. In Sect. 6, implementation and results are described and finally, conclusions and future work are mentioned in Sect. 7.

2 Preprocessing work

Detection of face image, normalization face size, eliminating forehead portion and conditionally, histogram linearization technique (HLT) are done in this section. Localizing and cropping the exact face portion

from the cluttered background is included in the preprocessing work. Since not all the detected faces are the same size, so it is essential to build all images into same size is known as normalization task for uniformity purpose. The Viola–Jones face detection algorithm is used for localizing and cropping the exact face area [26]. There is no information available on forehead portion in the binary face image, therefore, it is necessary to eliminate forehead portion for speed up processing activity [3–6, 25]. The complete proposed work is shown in Fig. 1.

2.1 Face detection and cropping

Face detection in an input face image is done by Paul Viola and Michael Jones called Viola–Jones face detector. It is based on four key concepts such as rectangular features like Haar features, an integral image for fast feature detection, classifier training and feature using the AdaBoost machine-learning algorithm and a cascade classifier to combine many features efficiently [26].

Rectangular feature (Haar-like features) image is represented by two-, three- and four-rectangular features. A two-rectangle feature value is computed by the difference between the sums of the pixels within two vertically or horizontally adjacent rectangular regions. A three-rectangle feature value is computed by the sum within two outside rectangles subtracted from the sum in a center rectangle and a four-rectangle feature value is computed by the difference between diagonal pairs of rectangles.

The integral image at pixel location $p(x, y)$ is the sum of all the pixels above it and to its left. It is calculated in one pass over the original image. Therefore, feature can be

determined rapidly because the value of each rectangle needs only four array references.

A machine-learning algorithm called AdaBoost is applied to construct a strong classifier through a weighted combination of weak classifiers. Combination of a series of AdaBoost classifier is called as a filter chain and each filter, which is a separate AdaBoost classifier, contains a fairly small number of weak classifier. If any one of these filters fails to pass an image region is immediately classify as “Not Face”. If image regions that pass through all filters in the chain are classify as “Face” and this filtering chain is known as cascade [26].

Finally, the detected face is cropped (face size = $W \times H$ pixels) from an input image. The detected and cropped face are shown in Fig. 3a, b, respectively.

2.2 Face size normalization

All the detected faces are not of the same size during face detection task, therefore, it is indispensable to convert all images into equal size (normalized face size = $W1 \times H1 = 128 \times 128$ pixels)) for uniformity purpose (see Fig. 3c).

2.3 Face forehead portion elimination

There is no information (i.e., no white pixels are present in the forehead portion) is available in the forehead region during binary image conversion task. So, it is necessary to discard forehead portion (face size without forehead region = $W2 \times H2 = 0.75W1 \times 0.60H1 = 96 \times 76$ pixels) for speed up processing activity (see Fig. 3d).

Fig. 1 Face Recognition Overview

2.4 Illumination uniformity activity using histogram linearization technique (HLT)

Face image having weak illumination can be reduced the recognition performance. Therefore, illumination adjustment activity is necessary in the preprocessing stage. HLT transforms the intensity values so that the histogram of the output image approximately appears the flat (uniform) histogram. We have applied HLT for illumination uniformity activity on a grayscale face image if its average pixel intensity value is less than 170 [27] (see Fig. 3e).

3 Processing work

Conversion of binary face image using cumulative probability distribution function (CPDF) with Otsu's optimal global thresholding value, and cropping of four BFCs (such as both eyes, nose and mouth components), four grayscale facial components (GFCs) (such as both eyes, nose and mouth components) and nine FCPs extraction are the principal tasks in this section. Details procedures are shown in Fig. 3. Cropping of facial components is explained in details in [5–7] (see Fig. 3f, h). The following mathematical concepts are applied to explain the Otsu's adaptive thresholding, binary image conversion, cropping of four BFCs (such as both eyes, nose and mouth components), four grayscale facial components (GFCs) (such as both eyes, nose and mouth components and nine FCPs extraction techniques [3–7, 28]. These are as follows.

3.1 Otsu's optimal global thresholding

Otsu's thresholding is a nonlinear, nonparametric and unsupervised statistical discriminant analysis image segmentation method known as clustering-based adapting thresholding that transforms a grayscale image into a binary image. Its fundamental concept is to split the image's pixels into two classes and confirms the optimal threshold to maximize between-class (inter-class) variance or minimize the weighted sum of within-class (intra-class) variances.

An image, $I(x, y)$, is a 2D grayscale intensity function, and contain $N = W \times H$ pixels with gray levels from 0 to $L - 1$. The number of pixels with gray level i is denoted n_i giving a probability of gray level i in an image of $\rho_i = \frac{n_i}{N}$. Where, $\sum_{i=0}^{L-1} \rho_i = 1$, $\rho_i \geq 0$, $0 \leq i \leq L - 1$ and $L = 256$.

In the case of bi-level thresholding of an image, $I(x, y)$, the pixels are divided into two classes, C_1 with gray levels $[0, 1, 2, \dots, t]$ and C_2 with gray levels

$[t + 1, t + 2, t + 3, \dots, L - 1]$ and suppose, a threshold value, $\tau_{\text{threshold}}(t) = t$, and $0 < t < L - 1$. Then, the probabilities, $\omega_1(t)$ and $\omega_2(t)$ that the pixels are assigned to the two classes C_1 and C_2 are given by the following cumulative sum [28]:

$$\omega_1(t) = \sum_{i=0}^t \rho_i \quad (1)$$

and

$$\omega_2(t) = \sum_{i=t+1}^{L-1} \rho_i = 1 - \omega_1(t) \quad (2)$$

The mean intensity values, $\mu_1(t)$ and $\mu_2(t)$, of the pixels assigned to the two classes C_1 and C_2 using Bayes' rule are:

$$\mu_1(t) = \sum_{i=0}^t iP(i/C_1) = \sum_{i=0}^t iP(C_1/i) \frac{P(i)}{P(C_1)} = \frac{\sum_{i=0}^t i\rho_i}{\omega_1(t)} \quad (3)$$

$$\begin{aligned} \text{And similarly for class } C_2, \mu_2(t) &= \sum_{i=t+1}^{L-1} iP(i/C_2) \\ &= \frac{\sum_{i=t+1}^{L-1} i\rho_i}{\omega_2(t)} \end{aligned} \quad (4)$$

The cumulative mean (average intensity), $\mu_C(t)$ up to level t and the global mean of the entire image, μ_G up to level $L - 1$ are expressed by the following equations:

$$\mu_C(t) = \sum_{i=0}^t i\rho_i \quad (5)$$

and

$$\mu_G = \sum_{i=0}^{L-1} i\rho_i \quad (6)$$

The global mean, μ_G , can be written as sum of weighed mean of the two classes C_1 and C_2 up to level $L - 1$ is:

$$\mu_G = \omega_1(t)\mu_1(t) + \omega_2(t)\mu_2(t) \quad (7)$$

and

$$\omega_1(t) + \omega_2(t) = 1 \quad (8)$$

Global variance (total variance i.e., the intensity variance of all pixels in the image), σ_G^2 , is the sum of within-class variance, σ_W^2 and between-class variance, σ_B^2 .

$$\sigma_G^2 = \sigma_W^2(t) + \sigma_B^2(t) \quad (9)$$

where,

$$\sigma_G^2 = \sum_{i=0}^{L-1} (i - \mu_G)^2 \rho_i \quad (10)$$

which is constant and independent of threshold value

Fig. 2 Applications of Otsu's optimal global thresholding method: **a** original image of grains of rice, **b** thresholded image of grains of rice, **c** Original image of grayscale human face and **d** thresholded binary image of human face [27]

Fig. 3 Procedures of proposed work: **a** input image, **b** detected and extracted the actual face region, **c** normalized face size ($W1 \times H1 = 128 \times 128$ pixels), **d** eliminated forehead region (size = $W2 \times H2 = 0.75W1 \times 0.60H1 = 96 \times 76$ pixels), **e** conditionally applied histogram linearization technique (HLT), **f** extracted grayscale both

eyes, nose and mouth regions, **g** converted binary image taking **(d)** or **(e)**, **(h)** extracted binary both eyes, nose and mouth regions, **(i)** to **(k)** constructed five statistical pattern matching tools such as σ_{CSQ} , σ_{HuMIs} , δ_{ABS} , d_{EDF} , and $G DVs$.

$$\sigma_W^2(t) = \sum_{i=0}^t (i - \mu_1(t))^2 \frac{\rho_i}{\omega_1(t)} + \sum_{i=t+1}^{L-1} (i - \mu_2(t))^2 \frac{\rho_i}{\omega_2(t)} \quad (11)$$

Indicate weighted within-class variance for classes C_1 and C_2

$$\begin{aligned}\sigma_B^2(t) &= \omega_1(t)(\mu_1(t) - \mu_G)^2 + \omega_2(t)(\mu_2(t) - \mu_G)^2 \\ &= \omega_1(t)\omega_2(t)(\mu_1(t) - \mu_2(t))^2 = \frac{(\mu_G\omega_1(t) - \mu_C(t))^2}{\omega_1(t)(1 - \omega_1(t))},\end{aligned}\quad (12)$$

indicate between-class variance for classes C_1 and C_2 .

Since, total variance (global variance) is constant and does not depend on threshold value, so that the algorithm must concentrate on minimizing $\sigma_W^2(t)$ or maximizing $\sigma_B^2(t)$. Calculating within-class variance for each of the two classes for each possible threshold needs a lot of computation, but it is easy way to maximize $\sigma_B^2(t)$ because $\sigma_B^2(t)$ depends only global mean, cumulative mean and probability of class C_1 which are easy to compute.

The optimal threshold value t^* , which is maximized σ_B^2 can be shown the relation:

$$\sigma_B^2(t^*) = \max_{0 \leq t \leq L-1} \sigma_B^2(t) \quad (13)$$

and finally, we get the OTSU threshold value,

$$\tau_{\text{Otsu}} = t^* \quad (14)$$

(see Fig. 2b, d).

3.2 Binary image conversion using OTSU thresholding

The cumulative probability density function (CPDF) idea is applied on a cropped grayscale face image with Otsu's thresholding to get the desire binary face image. Finally, the four ROIs (Region of Interests) such as binary both eyes, nose and mouth regions are extracted using the 2D geometrical positions (Cartesian coordinate system on xy-plane) concept [3–7] (see Fig. 3g, h).

If $\mathcal{I}_{\text{grayscale}}(\mathcal{X}, \mathcal{Y})$ is a cropped grayscale face image having size $\mathcal{W} \times \mathcal{H}$, the binary face image, $\mathcal{I}_{\text{binary}}(\mathcal{X}, \mathcal{Y})$ is performed by using the following mathematical equations.

$$\mathcal{P}_{\mathcal{Z}}^{\mathcal{I}_{\text{grayscale}}(\mathcal{X}, \mathcal{Y})=\mathcal{Z}} = \frac{n_{\mathcal{Z}}}{\mathcal{W} \times \mathcal{H}} \quad (15)$$

where $0 \leq \mathcal{Z} \leq (L-1)$, $L = 256$ and $0 \leq \mathcal{P}_{\mathcal{Z}} \leq 1.0$

$$cpdf_{\mathcal{Z}}^{\mathcal{I}_{\text{grayscale}}(\mathcal{X}, \mathcal{Y})} = \sum_{j=0}^{\mathcal{Z}} \mathcal{P}_j^{\mathcal{I}_{\text{grayscale}}(\mathcal{X}, \mathcal{Y})} \quad (16)$$

$$\mathcal{I}_{\text{binary}}(\mathcal{X}, \mathcal{Y}) = \begin{cases} 255, & \text{if } cpdf_{\mathcal{Z}}^{\mathcal{I}_{\text{grayscale}}(\mathcal{X}, \mathcal{Y})} \leq \tau_{\text{Otsu}} \\ 0 & \text{otherwise} \end{cases} \quad (17)$$

where $\mathcal{P}_{\mathcal{Z}}^{\mathcal{I}_{\text{grayscale}}(\mathcal{X}, \mathcal{Y})}$ = Probability Density Function of pixel value is $\mathcal{Z}$ ($0 \leq \mathcal{Z} \leq 255$), $\mathcal{W} \times \mathcal{H}$ = Image Size = Width $\times$ Height, $n_{\mathcal{Z}}$ = Total number of pixels having same pixel value $\mathcal{Z}$, $cpdf_{\mathcal{Z}}^{\mathcal{I}_{\text{grayscale}}(\mathcal{X}, \mathcal{Y})}$ = Cumulative probability density function up to the intensity value $\mathcal{Z}$ τ_{otsu} = Otsu's adaptive threshold value

3.3 Cropping grayscale and binary facial components

Viewing from the human frontal face structure concept for both grayscale and binary images, eyes, nose and mouth regions are situated in upper, middle and lower portions of the face image, respectively. Again, the upper portion is divided vertically into left and right segments for isolating right and left eyes, respectively. For cropping nose region, the center portion is extracted from the middle portion of the human face structure with discarding equal size of the leftmost and rightmost regions, and finally, the mouth portion is extracted from the lower portion of the human face structure with discarding equal size of the leftmost and rightmost regions (see Fig. 3f, h). Details are explained in [5–7]. The details of ROI (Region of Interest) size are mentioned below:

Normalized Face Image Size

$$= W1 \times H1 = 128 \times 128 \text{ pixels},$$

Face Image Size (without forehead area)

$$= W2 \times H2 \text{ pixels},$$

$$= (1.0 - 2 \times 0.125) W1 \times (1.0 - 0.25(\text{upper portion}) - 0.15(\text{lower portion}))$$

$$H1 = 0.75W1 \times 0.60H1 = 96 \times 76 \text{ pixels},$$

Upper Portion Image Size

$$= W2 \times 0.42H2 \text{ and } \mathbf{Left \text{ and } Right \text{ Eyes Size}}$$

$$= 0.50W2 \times 0.42H2,$$

Middle Portion Image Size

$$= W2 \times 0.31H2 \text{ and } \mathbf{Nose \text{ size}}$$

$$= (1.0 - 2 \times 0.17)W2 \times 0.31H2 = 0.66W2 \times 0.31H2,$$

Lower Portion Image Size

$$= W2 \times 0.27H2 \text{ and } \mathbf{Mouth \text{ size}}$$

$$= (1.0 - 2 \times 0.17)W2 \times 0.27H2 = 0.66W2 \times 0.27H2.$$

3.4 Human facial corner points (FCPs) extraction

Four binary ROI (Region of Interest) images such as **Right-Eye**, **Left-Eye**, **Nose** and **Mouth** regions are used to detect corner points automatically (see Fig. 3h). Simple linear search concept is applied on Right Eye, Left Eye and Mouth binary ROIs to detect the first white pixel locations and the contour algorithm is applied on binary Nose ROI to detect nostrils locations. Finally, a Nose Tip is calculated from the nostrils locations. Details are explained in [5–7]. The position of nine facial corner points is summarized and shown in Table 1.

Results of detected nine human facial corner points are shown in Fig. 4.

Table 1 Summary of facial corner points (FCPs) detection

Name of Feature Areas	Binary ROIs	Corner Points	Position of Facial Corner points
Right-Eye			1. <i>Right-Corner Point</i> _{Right-Eye} (x_{rcp-re}, y_{rcp-re}) 2. <i>Left-Corner Point</i> _{Right-Eye} (x_{lcp-re}, y_{lcp-re})
Left-Eye			3. <i>Right-Corner Point</i> _{Left-Eye} (x_{rcp-le}, y_{rcp-le}) 4. <i>Left-Corner Point</i> _{Left-Eye} (x_{lcp-le}, y_{lcp-le})
Nose			5. <i>Right-Nostril</i> _{Last-Contour} (x_{rn-lc}, y_{rn-lc}) 6. <i>Left-Nostril</i> _{previous Last-Contour} (x_{ln-plc}, y_{ln-plc}) 7. <i>Nose Tip</i> ($x_{Nose\ Tip}, y_{Nose\ Tip}$)
Mouth			8. <i>Right-Corner Point</i> _{Mouth} ($x_{rcp-mouth}, y_{rcp-mouth}$) 9. <i>Left-Corner Point</i> _{Mouth} ($x_{lcp-mouth}, y_{lcp-mouth}$)

4 Determination statistical pattern matching tools

The five statistical pattern matching tools such as standard deviation of five CSQ values, standard deviation of fifteen HuMIs values, five AbsDifPWPs values between binary test face and binary reference database images and five PIVs ($\mathcal{L}_2$ norm) between grayscale test face and grayscale reference database images and nine GDVs using automatic extracted nine FCPs between test face and reference database images are constructed mathematically to describe the six different types of face recognition algorithms [3–8].

4.1 Chi-SQUARE (CSQ) method

Chi-square (CSQ) statistic calculates the goodness-of-fit of the data to the model. That is, Chi-square is the sum of the squared difference between observed ($\mathbf{O}$) and the expected (ε) data (or the deviation, δ), divided by the sum of observed and expected data in all possible categories.

If the **observed** values in each of β bins are $\mathbf{O}_i$, and the **expected** values from the model are ε_i , then χ^2_{CSQ} statistic can be written by the following formula using Eqs. (19) and (20) [29, 30]:

$$\chi^2_{CSQ} = \sum_{i=1}^{\beta=2} \frac{(\mathbf{O}_i - \varepsilon_i)^2}{(\mathbf{O}_i + \varepsilon_i)} \quad (18)$$

$$\chi^2_{CSQ} = \frac{(P_{Ref} - P_{Test})^2}{(P_{Ref} + P_{Test})} + \frac{(Q_{Ref} - Q_{Test})^2}{(Q_{Ref} + Q_{Test})} \quad (19)$$

$$\delta_{CSQ} = \chi^2_{CSQ} = \frac{2(P_{Ref} - P_{Test})^2}{(P_{Ref} + P_{Test})(2 - P_{Ref} - P_{Test})} \quad (20)$$

where P = Probability of White Pixels (PWPs), $\mathbf{O} = P_{Ref}$, $\varepsilon = P_{Test}$, Q = Probability of Black Pixels (PBPs) = $1 - P$, P_{Ref} and P_{Test} are the PWPs of reference and test images, respectively.

A low value of $\delta = \chi^2_{CSQ}$ indicates a better match than a high score. An exact match is 0(zero) and a total mismatch is unbounded and it depends on the size of the bin.

4.2 Standard deviation of Chi-Square value

Standard deviation of five CSQ values for five BFCs between the test face and a reference database image is given in Eq. (21) [3].

$$\sigma_{CSQ} = CSQ_i(\mathcal{I}_{Test_i}(x, y), \mathcal{I}_{Ref_i}(x, y)) = \frac{\sqrt{\sum_{i=1}^5 (\delta_{CSQ_i} - \overline{\delta_{CSQ}})^2}}{5} \quad (21)$$

where, σ_{CSQ} is a standard deviation of CSQ value, δ_{CSQ_i} are CSQ values for five BFCs, $\overline{\delta_{CSQ}}$ is an mean CSQ value, and $i = 1, 2, \dots, 5$

Fig. 4 Geometric distance between the two similar corner points of test face image and reference

4.3 Moments and moment invariants

Moment concept is mainly used for shape descriptor of a probability distribution function and use to many real-world applications such as computer vision, image processing and pattern recognition areas for object matching, recognition, classification and identification purposes. Mathematically, moments are “projection” of a function onto a polynomial basis. The image shape feature performs a vital role in image classification, identification and recognition activities. So, effectively and efficiently extraction of shape features is the key element of the image representation and comparison. Moment invariants mean invariant to certain class of image degradations such as invariant to translation, rotation, scaling (TRS) or affine transform, etc. In this research work, we use moment invariant concept as shape descriptor on five binary facial components for face recognition [31].

4.4 Hu moment invariants (HuMIs)

Hu [32] first introduced two-dimensional geometric moment invariants concept to apply for shape recognition task. A set of seven nonlinear moment functions are derived from the second and third order moment, which are translation, scale and rotation invariants.

A digital image $f(a, \ell)$ having size $\mathcal{W} \times \mathcal{H}$, the 2D traditional geometric moments of order $(\alpha + \beta)$ are expressed by:

$$\mathcal{K}_{\alpha\beta} = \sum_{a=0}^{\mathcal{W}-1} \sum_{\ell=0}^{\mathcal{H}-1} f(a, \ell) a^{\alpha} \ell^{\beta} \quad (22)$$

where $\alpha, \beta = 0, 1, 2, \dots$ are integer values.

The double integrals are to be considered over the entire area of the image including its boundary.

In the image plane, the image centroids are used to define the central moments to normalize for translation. The central moment of order $(\alpha + \beta)$ is defined as:

$$\lambda_{\alpha\beta} = \sum_{a=0}^{\mathcal{W}-1} \sum_{\ell=0}^{\mathcal{H}-1} f(a, \ell) (a - \bar{a})^{\alpha} (\ell - \bar{\ell})^{\beta} \quad (23)$$

where $\bar{a} = \frac{\mathcal{K}_{10}}{\mathcal{K}_{00}}$, and $\bar{\ell} = \frac{\mathcal{K}_{01}}{\mathcal{K}_{00}}$

Since central moments are origin independent and therefore they are translation invariant. But these moments are not invariant to scale or rotation in their original form. When a scaling normalization is applied the central moments become as:

$$\gamma_{\alpha\beta} = \frac{\lambda_{\alpha\beta}}{\lambda_{00}^{\frac{(\alpha+\beta+2)}{2}}} \quad \text{for } \alpha + \beta = 2, 3, \dots \quad (24)$$

Hu’s [32], seven scalar values, determined by normalizing central moments through order three, that are invariant to object scale, position and orientation. In terms of the central moments, a set of the seven moments are derived by Eqs. (25–31):

$$\psi_1 = r_{20} + r_{02} \quad (25)$$

$$\psi_2 = (r_{20} - r_{02})^2 + 4r_{11}^2 \quad (26)$$

$$\psi_3 = (r_{30} - 3r_{12})^2 + (3r_{21} - r_{03})^2 \quad (27)$$

$$\psi_4 = (r_{30} + r_{12})^2 + (r_{21} + r_{03})^2 \quad (28)$$

$$\begin{aligned} \psi_5 = & (r_{30} - 3r_{12})(r_{30} + r_{12})((r_{30} + r_{12})^2 - 3(r_{21} + r_{03})^2) \\ & + (3r_{21} - r_{03})(r_{21} + r_{03})(3(r_{30} + r_{12})^2 - (r_{21} + r_{03})^2) \end{aligned} \quad (29)$$

$$\begin{aligned} \psi_6 = & (r_{20} - r_{02})((r_{30} + r_{12})^2 - (r_{21} + r_{03})^2) \\ & + 4r_{11}(r_{30} + r_{12})(r_{21} + r_{03}) \end{aligned} \quad (30)$$

$$\begin{aligned} \psi_7 = & (3r_{21} - r_{03})(r_{30} + r_{12})((r_{30} + r_{12})^2 - 3(r_{21} + r_{03})^2) \\ & + (3r_{12} - r_{30})(r_{21} + r_{03})(3(r_{30} + r_{12})^2 - (r_{21} + r_{03})^2) \end{aligned} \quad (31)$$

These seven invariant moments, $\psi_z, 1 \leq z \leq 7$ (Eqs. 25–31), are independent of scale, translation and rotation. We have applied binary five images (the binary four ROIs such as right eye, left eye, nostrils and mouth component and binary face image discarding forehead portion) of both for test and reference images by ψ_z as matching the shapes.

4.5 Standard deviation of fifteen Hu moment invariants (HuMIs)

The three types of Hu's invariant moment values such as ε_i , φ_i and ζ_i are computed using the following equations [4] (32–36):

$$\varepsilon_i(I_{\text{Ref}_i}(a, \ell), I_{\text{Test}_i}(a, \ell)) = \sum_{z=1 \dots 7} \left| \frac{1}{m_z^{\text{Ref}_i}} - \frac{1}{m_z^{\text{Test}_i}} \right| \quad (32)$$

$$\varphi_i(I_{\text{Ref}_i}(a, \ell), I_{\text{Test}_i}(a, \ell)) = \sum_{z=1 \dots 7} |m_z^{\text{Ref}_i} - m_z^{\text{Test}_i}| \quad (33)$$

$$\zeta_i(I_{\text{Ref}_i}(a, \ell), I_{\text{Test}_i}(a, \ell)) = \sum_{z=1 \dots 7} \frac{|m_z^{\text{Ref}_i} - m_z^{\text{Test}_i}|}{|m_z^{\text{Ref}_i}|} \quad (34)$$

where $i = 1, 2, 3, 4, 5$,

$$m_z^{\text{Ref}} = \text{sign}(\psi_z^{\text{Ref}}) \cdot \log(\psi_z^{\text{Ref}}) \quad (35)$$

$$m_z^{\text{Test}} = \text{sign}(\psi_z^{\text{Test}}) \cdot \log(\psi_z^{\text{Test}}) \quad (36)$$

ψ_z^{Ref} and ψ_z^{Test} are the HuMIs of binary form of test and reference images, respectively. It is invariant to scale, rotation and reflection [4].

Using three groups of Hu's invariant moments such as ε_i , φ_i and ζ_i , the Standard deviation of fifteen moment-s (HuMIs) is [4]:

$$\begin{aligned} \sigma_{\text{HuMIs}} = & \text{StdDev}_{\text{HuMIs}}(I_{\text{Binary-Ref}}(a, \ell), I_{\text{Binary-Test}}(a, \ell)) \\ = & \frac{\sqrt{\sum_{k=1}^{15} (d_k - \bar{d})^2}}{15} \end{aligned} \quad (37)$$

where $d_{k=1,2,\dots,15} = \{\varepsilon_{i=1,2,\dots,5}, \varphi_{i=1,2,\dots,5}, \zeta_{i=1,2,\dots,5}\}$;
 $\bar{d}$ is the mean value of d_k 's ($k = 1, 2, 3, \dots, 15$)

4.6 Absolute difference of probability of white pixels (AbsDifPWPs)

Absolute difference of probability of white pixels (AbsDifPWPs) between the binary form of test image and a reference database image is shown in Eq. (38) [3, 4].

$$\begin{aligned} \delta_{\text{ABS}_i} = & \text{AbsDifPWPs}_i(I_{\text{Test}_i}(x, y), J_{\text{Ref}_i}(x, y)) \\ = & |\mathcal{P}_{\text{Test}_i}(I_{\text{Test}_i}(x, y)) - \mathcal{P}_{\text{Ref}_i}(J_{\text{Ref}_i}(x, y))| \end{aligned} \quad (38)$$

$\delta_{\text{ABS}_i} \geq 0$ and $\delta_{\text{ABS}_i} = 0$ indicate perfect match, where $i = 1, 2, \dots, 5$.

4.7 Pixel intensity values (PIVs) using L_2 norms

Similarity (or closeness) is calculated by using a Euclidean distance formula (EDF). The formula for the square root of sum of the squares differences between the corresponding pixel values of the same size ($\mathcal{W} \times \mathcal{H}$) of two grayscale images $I_{\text{Test}_i}(x, y)$ and $J_{\text{Ref}_i}(x, y)$ is known as $\mathcal{L}_2$ norm. The equation of $\mathcal{L}_2$ (39) is [3, 4]:

$$\begin{aligned} d_{\mathcal{EDF}_i} = & \mathcal{EDF}_i(I_{\text{Test}_i}(x, y), J_{\text{Ref}_i}(x, y)) \\ = & \frac{\sqrt{\sum_{x,y} (I_{\text{Test}_i}(x, y) - J_{\text{Ref}_i}(x, y))^2}}{\mathcal{W} \times \mathcal{H}} \end{aligned} \quad (39)$$

where $i = 1, 2, \dots, 5$, $d_{\mathcal{EDF}_i} \geq 0.0$ and $d_{\mathcal{EDF}_i} = 0.0$ indicates perfect match.

4.8 Geometric distant values (GDVs) using nine facial corner points (FCPs) between two faces

The Geometric Distance Values (GDVs) between two Cartesian coordinate points on the xy-plane is defined by the geometric distance formula.

The distance between the two similar corner points of $Q_i^{\text{Test}}(x_k, y_\ell)$ and $\mathcal{R}_i^{\text{Ref}}(x_u, y_v)$ of the same size (size: 128×128) of two face images $\text{Img}_{\text{Test}_i}(x, y)$ and $\text{Img}_{\text{Ref}_i}(x, y)$, respectively is (see Fig. 4) [5–8]:

$$\begin{aligned} \text{GDVs}_i(Q_i^{\text{Test}}(x_k, y_\ell), \mathcal{R}_i^{\text{Ref}}(x_u, y_v)) \\ = & \sqrt{(x_{\text{Test}_k} - x_{\text{Ref}_u})^2 + (y_{\text{Test}_\ell} - y_{\text{Ref}_v})^2} \end{aligned} \quad (40)$$

where $i = k = \ell = u = v = 1, 2, 3, \dots, 9$ and $\text{GDVs}_i \geq 0.0$ and $\text{GDVs}_i = 0.0$ confirm perfect match

5 Similarity evaluation and recognition decisions

Recognition decision comes from the idea of matching in any one facial component value of a test face image is same as corresponding facial component value of reference

database images using any two or three statistical tools out of five statistical tools such as σ_{CSQ} , σ_{HuMIs} , δ_{ABS_i} , d_{EDF_i} , and any three points out of nine points of GDVs_i with the help of logical and conditional operators and appropriate threshold values such as T_{CSQ} , T_{HuMIs} , T_{ABS} , T_{PIVs} and T_{GDVs} , respectively. The six different types of recognition decisions are explained below:

CSQ (Method 1):

if $((\sigma_{\text{CSQ}} < T_{\text{CSQ}}) \& \& (\text{any one out of five values of } \delta_{\text{ABS}_i} \leq T_{\text{ABS}}) \& \& (\text{any one out of five values of } d_{\text{EDF}_i} < T_{\text{PIVs}}))$

{The test and reference images are the same face};

CSQ + FCPs (Method 2):

if $((\sigma_{\text{CSQ}} < T_{\text{CSQ}}) \& \& (\text{any one out of five values of } \delta_{\text{ABS}_i} \leq T_{\text{ABS}}) \& \& (\text{any one out of five values of } d_{\text{EDF}_i} < T_{\text{PIVs}}))$

{The test and reference images are the same face};

elseif $((\text{any three out of nine values of } \text{GDVs}_r < T_{\text{GDVs}}) \& \& (\text{any one out of five values of } d_{\text{EDF}_i} < T_{\text{PIVs}}))$

{The test and reference images are the same face};

HuMIs (Method 3):

if $((\sigma_{\text{HuMIs}} \leq T_{\text{HuMIs}}) \& \& (\text{any one out of five values of } \delta_{\text{ABS}_i} \leq T_{\text{ABS}}) \& \& (\text{any one out of five values of } d_{\text{EDF}_i} < T_{\text{PIVs}}))$

{The test and reference images are the same face};

HuMIs + FCPs (Method 4):

if $((\sigma_{\text{HuMIs}} \leq T_{\text{HuMIs}}) \& \& (\text{any one out of five values of } \delta_{\text{ABS}_i} \leq T_{\text{ABS}}) \& \& (\text{any one out of five values of } d_{\text{EDF}_i} < T_{\text{PIVs}}))$

{The test and reference images are the same face};

elseif $((\text{any three out of nine values of } \text{GDVs}_r < T_{\text{GDVs}}) \& \& (\text{any one out of five values of } d_{\text{EDF}_i} < T_{\text{PIVs}}))$

{The test and reference images are the same face};

CSQ + HuMIs (Method 5):

if $((\sigma_{\text{CSQ}} < T_{\text{HuMIs}}) \& \& (\text{any one out of five values of } \delta_{\text{ABS}_i} \leq T_{\text{ABS}}) \& \& (\text{any one out of five values of } d_{\text{EDF}_i} < T_{\text{PIVs}}))$

{The test and reference images are the same face};

elseif $((\sigma_{\text{HuMIs}} \leq T_{\text{CSQ}}) \& \& (\text{any one out of five values of } \delta_{\text{ABS}_i} \leq T_{\text{ABS}}) \& \& (\text{any one out of five values of } d_{\text{EDF}_i} < T_{\text{PIVs}}))$

{The test and reference images are the same face};

CSQ + HuMIs + FCPs (Method 6):

if $((\sigma_{\text{CSQ}} < T_{\text{CSQ}}) \& \& (\text{any one out of five values of } \delta_{\text{ABS}_i} \leq T_{\text{ABS}}) \& \& (\text{any one out of five values of } d_{\text{EDF}_i} < T_{\text{PIVs}}))$

{The test and reference images are the same face};

elseif $((\sigma_{\text{HuMIs}} \leq T_{\text{HuMIs}}) \& \& (\text{any one out of five values of } \delta_{\text{ABS}_i} \leq T_{\text{ABS}}) \& \& (\text{any one out of five values of } d_{\text{EDF}_i} < T_{\text{PIVs}}))$

{The test and reference images are the same face};

elseif $((\text{any three out of nine values of } \text{GDVs}_r < T_{\text{GDVs}}) \& \& (\text{any one out of five values of } d_{\text{EDF}_i} < T_{\text{PIVs}}))$

{The test and reference images are the same face};

6 Implementation and results

By applying the combination of five pattern matching tools to establish the six different types of recognition algorithms are described each of the individual(person) results on the basis of the best true positive rate(TPR) and average result in the tabular form, the entire results in graphical form and some true output face images(true alarms) using only method 6(CSQ + HuMIs + FCPs) are shown in pictorial form, and all methods are compared to the five basic performance measurement parameters such as the best True Positive Rate (TPR)/Recall/Sensitivity and corresponding to Recognition Rate (RR)/Accuracy, Precision, F-Score and False Positive Rate (FPR)/1-Specificity are derived from the confusion matrix [33]. Only result of false positive rate(FPR) is arranged in descending order and the rest of the four performance measurement parameters such as TPR, precision, F-Score and RR are arranged in ascending order (see Tables 3, 4 and 5 and Figs. 6, 7).

6.1 Performance measures from the confusion matrix

The confusion matrix has four outcomes: True Positives (T_p), True Negatives (T_n), False Positives (F_p) and False Negatives (F_n). The five basic performance measures from the confusion matrix are described in Table 2 [33].

Where, T_p = Correctly predicted an authorize person (True Alarms), T_n = Correctly predicted a non-authorize

person (True No-Alarms), F_p = Incorrectly accepted a non-authorize person (False Alarms), F_n = Incorrectly rejected an authorize person (Missed Alarms), N = Total Number of Images of all person in a database = 1306

6.2 Face database

In this paper, cropping only face area is done using a popular face detection algorithm known as Viola–Jones face detector [26]. BioID face database [34] is used for experimental purpose, which is consists of 1521 grayscale images of 25 persons having the properties of diverse illuminations, expression, pose variations and partial occlusions; face area is located in different positions and complex background with a resolution of 384×286 pixels. Only $N_{\text{Total}} = 1306$ faces are detected and rest of the images are discarded due to detect false region (not face) by the face detector. Each face image is considered as a test face and compare to all faces ($N_{\text{Total}} = 1306$) as a reference database during the assessment activities.

7 Results

The proposed task is implemented and tested in OpenCV platform using c/c++ and GNU GCC compiler with Code::Blocks. Face detection and localization, cropping accurate face region, face size normalization,

Table 2 The five basic performance measures from the confusion matrix

Sl. No	Measure	Formula	Descriptions	Remarks (For the best case)
1.	Recognition Rate(RR)/ Accuracy	$\frac{N - (F_p + F_n)}{N} \times 100\%$	It is the percentage of correct classification and RR is inversely related to F_p .	RR = 100% (Shown in Table 3, Table 5, Table 6 and Fig. 7)
2.	True Positive Rate (TPR)/ Recall/Sensitivity	$\frac{T_p}{(T_p + F_n)}$	Probability of, given positive example, outcome will be a positive test result. TPR is independent on F_p .	TPR = 1.0 (Shown in Table 3, Table 5, Table 6 and Fig. 6a)
3.	Precision	$\frac{T_p}{(T_p + F_p)}$	Probability that, given a positive test result, sample will be positive. Precision is inversely related to F_p .	Precision = 1.0 (Shown in Table 4, Table 5, Table 6 and Fig. 6b)
4.	F-Score	$\frac{2 * (\text{Recall} * \text{Precision})}{(\text{Recall} + \text{Precision})}$ $= \frac{2 * T_p}{(2 * T_p + F_p + F_n)}$	Combine both precision and recall into a single measures that conveys both properties. F-Score is inversely related to F_p .	F-Score = 1.0 (Shown in Table 4, Table 5, Table 6 and Fig. 6b)
5.	False Positive Rate (FPR)/(1-Specificity)	$\frac{F_p}{(F_p + T_n)}$	Probability of, given a negative example, outcome will be a negative test result. FPR is directly related to F_p .	FPR = 0.0 (Shown in Table 5, Table 6 & Fig. 6a)

Table 3 25 the best TPR(Recall) and corresponding recognition rate of 25 persons using BioID Face database

Persons (Individual)	No. of images of each persons (individual)	The best True Positive Rate (TPR)				Corresponding recognition rate (RR)			
		CSQ	CSQ + FCPs	HMI FCPs	HMI + FCPs	CSQ + HMI	CSQ + FCPs	HMI FCPs	CSQ + HMI + FCPs
Per1	25	0.6400	0.6000	0.7200	0.5600	0.5200	0.5200	0.5200	0.5200
Per2	27	0.8889	0.8889	0.9259	0.9259	0.9630	0.9630	0.9630	0.9630
Per3	67	0.5075	0.6269	0.5075	0.5672	0.5075	0.5075	0.5075	0.5075
Per4	74	0.3649	0.4595	0.3108	0.5000	0.3649	0.3649	0.5000	0.5000
Per5	98	0.5408	0.8571	0.5918	0.8673	0.8673	0.8673	0.8673	0.8673
Per6	46	0.5435	0.5435	0.5000	0.5000	0.6304	0.6304	0.6304	0.6304
Per7	84	0.7738	0.9048	0.5833	0.6071	0.7381	0.7381	0.7381	0.7381
Per8	59	0.5763	0.6102	0.5593	0.5932	0.5763	0.5763	0.6949	0.6949
Per9	46	0.6739	0.8043	0.8043	0.6087	0.8913	0.8913	0.9565	0.9565
Per10	36	0.6667	0.6944	0.6944	0.7222	0.7500	0.7500	0.7500	0.7500
Per11	102	0.5098	0.5588	0.6078	0.6373	0.6078	0.6078	0.6078	0.6078
Per12	89	0.8090	0.8539	0.5955	0.5730	0.6629	0.6629	0.8876	0.8876
Per13	130	0.5000	0.5000	0.5231	0.5231	0.6154	0.6154	0.5077	0.5077
Per14	30	0.5000	0.5667	0.7000	0.8333	0.9667	0.9667	0.9667	0.9667
Per15	49	0.9592	0.9796	0.5918	0.9592	0.9388	0.9388	0.9796	0.9796
Per16	71	0.5493	0.5070	0.6056	0.5070	0.8592	0.8592	0.5211	0.5211
Per17	48	0.5208	0.5417	0.5417	0.5000	0.5833	0.5833	0.5833	0.5833
Per18	34	0.7059	0.7059	0.6765	0.7647	0.9706	0.9706	0.9706	0.9706
Per19	36	0.6389	0.6944	0.6944	0.7222	0.7778	0.7778	0.7778	0.7778
Per20	39	0.7436	0.7179	0.5897	0.8718	0.8462	0.8462	0.8718	0.8718
Per21	46	0.8696	0.8478	0.5435	0.8913	0.9348	0.9348	0.9783	0.9783
Per22	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Per23	46	0.5652	0.8478	0.7391	0.8261	0.8261	0.8261	0.8478	0.8478
Per24	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Per25	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Average	–	0.6819	0.7324	0.6643	0.7224	0.7759	0.7759	0.7859	0.7859

Table 4 25 the best TPR(Recall) and corresponding Precision and F-Score of 25 persons using BioID Face database

Persons (individual)	No. of images of each persons (individual)	The best TPR(Recall) and corresponding Precision and F-Score											
		Precision						F-Score					
		CSQ	CSQ + FCPs	HMI	HMI + FCPs	CSQ + HMI	CSQ + HMI + FCPs	CSQ	CSQ + FCPs	HMI	HMI + FCPs	CSQ + HMI	CSQ + HMI + FCPs
Per1	25	0.6957	0.2885	0.2368	0.2059	0.2955	0.2889	0.6667	0.3896	0.3564	0.3011	0.3768	0.3714
Per2	27	0.7273	0.5714	0.3788	0.3425	0.3377	0.2921	0.8000	0.6957	0.5376	0.5000	0.5000	0.4483
Per3	67	0.2125	0.2333	0.1219	0.2533	0.1878	0.1744	0.2996	0.3401	0.1965	0.3502	0.2742	0.2595
Per4	74	0.2368	0.1726	0.1081	0.1298	0.1080	0.1128	0.2872	0.2509	0.1730	0.2061	0.1667	0.1841
Per5	98	0.5889	0.4828	0.6105	0.5000	0.3498	0.3283	0.5638	0.6176	0.6010	0.6343	0.4985	0.4793
Per6	46	0.7813	0.7576	0.8214	0.7931	0.5800	0.5686	0.6410	0.6329	0.6216	0.6133	0.6042	0.5979
Per7	84	0.6019	0.6496	0.3475	0.5543	0.6200	0.6200	0.6771	0.7562	0.4356	0.5795	0.6739	0.6739
Per8	59	0.2086	0.2057	0.1352	0.1417	0.1164	0.1323	0.3063	0.3077	0.2178	0.2288	0.1937	0.2222
Per9	46	0.5849	0.3814	0.2517	0.1931	0.2971	0.2366	0.6263	0.5175	0.3834	0.2932	0.4457	0.3793
Per10	36	0.6857	0.4237	0.5208	0.3611	0.3699	0.3699	0.6761	0.5263	0.5952	0.4815	0.4954	0.4954
Per11	102	0.4860	0.4488	0.4429	0.4333	0.3147	0.3039	0.4976	0.4978	0.5124	0.5159	0.4147	0.4052
Per12	89	0.4865	0.4810	0.3897	0.6071	0.3491	0.3435	0.6076	0.6154	0.4711	0.5896	0.4574	0.4953
Per13	130	0.4545	0.4545	0.4964	0.4964	0.5063	0.4962	0.4762	0.4762	0.5094	0.5094	0.5556	0.5019
Per14	30	0.1364	0.1298	0.1469	0.1563	0.1086	0.1074	0.2143	0.2112	0.2428	0.2632	0.1953	0.1933
Per15	49	0.2271	0.2330	0.1871	0.3507	0.1357	0.1404	0.3672	0.3765	0.2843	0.5137	0.2371	0.2455
Per16	71	0.5000	0.8000	0.3772	0.8780	0.3987	0.7551	0.5235	0.6207	0.4649	0.6429	0.5446	0.6167
Per17	48	0.3049	0.1955	0.1150	0.2330	0.2171	0.2171	0.3846	0.2873	0.1898	0.3179	0.3164	0.3164
Per18	34	0.9231	0.9231	0.3898	0.2955	0.4400	0.2946	0.8000	0.8000	0.4946	0.4262	0.6055	0.4521
Per19	36	0.3286	0.2577	0.2155	0.2047	0.1564	0.1556	0.4340	0.3759	0.3289	0.3190	0.2605	0.2593
Per20	39	0.4085	0.2947	0.2323	0.2378	0.2143	0.2179	0.5273	0.4179	0.3333	0.3736	0.3420	0.3487
Per21	46	0.3960	0.3223	0.1214	0.5694	0.2324	0.2381	0.5442	0.4671	0.1984	0.6949	0.3723	0.3830
Per22	18	0.1818	0.1343	0.1286	0.1098	0.0559	0.0517	0.3077	0.2368	0.2278	0.1978	0.1059	0.0984
Per23	46	0.2766	0.2120	0.2720	0.2209	0.1357	0.1279	0.3714	0.3391	0.3977	0.3486	0.2331	0.2222
Per24	4	0.0625	0.0303	0.0265	0.0173	0.0139	0.0115	0.1176	0.0588	0.0516	0.0340	0.0275	0.0228
Per25	2	0.0222	0.0179	0.0175	0.0140	0.0059	0.0058	0.0435	0.0351	0.0345	0.0276	0.0118	0.0116
Average	-	0.4207	0.3641	0.2837	0.3320	0.2619	0.2636	0.4704	0.4340	0.3544	0.3985	0.3564	0.3473

Table 5 Summary of six different types of Recognition Algorithms showing average performance evaluation parameters (using Tables 3 and 4)

Serial No.	Average						Threshold values (TVs)
	Methods	TPR(Recall)	RR%	FPR	Precision	F-Score	
1	CSQ [3]	0.6819	93.9387	0.0476	0.4207	0.4704	$\frac{\tau_{OTSU}}{6.00}$, $\mathcal{T}_{CSQ} = 0.008$
2	CSQ + FCPs [8]	0.7324	92.4288	0.0660	0.3641	0.4340	$\mathcal{T}_{HuMIs} = 0.08$
3	HuMIs [4]	0.6643	90.7106	0.0799	0.2837	0.3544	$\mathcal{T}_{ABS} = 0.005$
	ours						$\mathcal{T}_{GDVs} = 3.25$
4	HuMIs + FCPs	0.7224	91.2649	0.0767	0.3320	0.3985	$\mathcal{T}_{PIVs} = 1.0$
5	CSQ + HuMIs	0.7759	87.4395	0.1187	0.2619	0.3564	
6	CSQ + HuMIs + FCPs	0.7859	86.6156	0.1275	0.2636	0.3473	
		0.7859					

Table 6 Summary of six different types of Recognition Algorithms with number of the optimal TPR, Precision and FPR values (using the entire Database)

Serial No.	Methods		N_{TPR} (TPR = 100%)	$N_{Precision}$ (Number of Precision Values=1.00)	N_{FPR} (FPR = 0%)
1	CSQ [3]		23	24	24
2	CSQ + FCPs [8]		43	12	12
3	HuMIs [4]		25	14	14
4	HuMIs + FCPs	ours	34	8	8
5	CSQ + HuMIs	ours	73	7	7
6	CSQ + HuMIs	ours	95	7	7

determination of OTSU's dynamic thresholding and shape matching tasks are implemented by OpenCV library functions [3, 4].

For BioID face database, the five thresholds such as $\mathcal{T}_{CSQ}$, $\mathcal{T}_{HuMIs}$, $\mathcal{T}_{ABS}$, $\mathcal{T}_{PIVs}$ and $\mathcal{T}_{GDVs}$ for five pattern matching tools such as σ_{CSQ} , σ_{HuMIs} , δ_{ABS_i} , d_{EDF_i} and $GDVs_i$ are taken as 0.008, 0.08, 0.005, 1.0 and 3.25, respectively (see Table 5).

Tables 3 and 4 are shown the results of performance evaluation parameters such as TPR, RR, precision, F-Score on the basis of the best true positive rate (TPR) values using the 25 individuals (person) with the concept of the confusion matrix.

Using six different types of recognition algorithms: the true positive rate (TPR), recognition rate (RR) are shown in Table 3 and precision, F-Score are shown in Table 4 using the concept of the best TPR values, respectively.

Table 5 shows the summary of average performance evaluation parameters using Tables 3 and 4 and Table 6 shows the summary of optimal number of TPR, Precision and FPR values (TPR = precision = 1.0 and FPR = 0.0 are the optimum values) using the entire database (see Fig. 6a, b). The best average TPR and optimal number of TPR values for method 6 are 0.7859 and 95, respectively, as compare to the rest of the five methods which is satisfied our proposed work.

The different pattern matching algorithms recognized the same images as well as different images of the same individual (person) using entire face database. Therefore, combination of multiple algorithms has achieved the same percent of TPR values (optimal number of TPR = 95) for each individual which is fulfilled our research objective, but combination of multiple algorithms support the large numbers of false alarms (F_p).

The other performance evaluation parameters such as RR, Precision, F-Score and FPR are shown in little bit deteriorated due to support the false alarms (F_p). Because, these parameters (RR, Precision, F-Score and FPR) are related to false alarms (F_p) except True Positive Rate (TPR/Recall). TPR is not depend on F_p and $(T_p + F_n)$ is always shown constant for each individual (see the Table 2, second column of Tables 3 or 4 and Table 5). So, system is not effect the other performance evaluation parameters such as RR, Precision, F-Score and FPR to increase the T_p (true alarms). Since, individually, every algorithm is supported false alarms (F_p), therefore, fusion algorithms are also supported a large number of alarms (F_p) using the same threshold values setting as compare to individual pattern matching algorithm.

Figure 5 shows that algorithm (method 6) is invariants to illumination, expression, pose variation and partial occlusion. The left most column indicates test face as an

Fig. 5 Different Imaging Conditions: System supports different imaging conditions such as variations of illumination, expression, pose and partial occlusion. The left column indicates test face(input)

and the rest right five columns(used reference database) indicate recognized faces(output)

input image and the rest right five columns indicate recognized faces as an output images which is supported the invariants properties of multiple imaging conditions.

Figures 6 and 7 shows the performance and accuracy curves using six different types of recognition algorithms, respectively, and it is shown that the method 6 (CSQ + MuMIs + FCPs) has achieved the best performance of TPR(Recall), and the rest of the curves such as FPR, Precision, F-Score and RR are laid little bit below for other five methods. The sorted TPR(Recall), Precision, F-Score,

FPR and RR using six algorithms are shown below using Table 5, Figs. 6a, b and 7:

1. $TPR_{CSQ+HuMIs+FCPs} > TPR_{CSQ+HuMIs} > TPR_{CSQ+FCPs} > TPR_{HuMIs+FCPs} > TPR_{CSQ} > TPR_{HuMIs}$ (according to Table 5 and Fig. 6a),

2(a). $FPR_{CSQ} < FPR_{CSQ+FCPs} < FPR_{HuMIs+FCPs} < FPR_{HuMIs} < FPR_{CSQ+HuMIs} < FPR_{CSQ+HuMIs+FCPs}$ (according to Table 5),

Fig. 6 Performance curves of BioID face database: (a) TPR and FPR. **b** Performance Curves of BioID Face Database: (b) Precision and F-Score

- 2(b). $FPR_{CSQ} < FPR_{HuMIs} < FPR_{CSQ+FCPs} < FPR_{HuMIs+FCPs} < FPR_{CSQ+HuMIs} < FPR_{CSQ+HuMIs+FCPs}$ (according to Fig. 6a),
- 3(a). $Pre_{CSQ} > Pre_{CSQ+FCPs} > Pre_{HuMIs+FCPs} > Pre_{HuMIs} > Pre_{CSQ+HuMIs+FCPs} > Pre_{CSQ+HuMIs}$ (according to Table 5),
- 3(b). $Pre_{CSQ} > Pre_{CSQ+FCPs} > Pre_{HuMIs} > Pre_{HuMIs+FCPs} > Pre_{CSQ+HuMIs} > Pre_{CSQ+HuMIs+FCPs}$ (according to Fig. 6b),
- 4(a). $F - Sc_{CSQ} > F - Sc_{CSQ+FCPs} > F - Sc_{HuMIs+FCPs} > F - Sc_{CSQ+HuMIs} > F - Sc_{HuMIs} > F - Sc_{CSQ+HuMIs+FCPs}$ (according to Table 5),
- 4(b). $F - Sc_{CSQ} > F - Sc_{CSQ+FCPs} > F - Sc_{HuMIs+FCPs} > F - Sc_{HuMIs} > F - Sc_{CSQ+HuMIs} > F - Sc_{CSQ+HuMIs+FCPs}$ (according to Fig. 6b),
- 5(a). $RR_{CSQ} > RR_{CSQ+FCPs} > RR_{HuMIs+FCPs} > RR_{HuMIs} > RR_{CSQ+HuMIs} > RR_{CSQ+HuMIs+FCPs}$ (according to Table 5),
- 5(b). $RR_{CSQ} > RR_{CSQ+FCPs} > RR_{HuMIs} > RR_{HuMIs+FCPs} > RR_{CSQ+HuMIs} > RR_{CSQ+HuMIs+FCPs}$ (according to Fig. 7).

On the other hand, using simple nine corner points, (even though, Geometric Distant Values (GDVs) concept needs the least computational cost), GDVs are required only distance values, but it is reduced the system performance due to assist to support false alarms(F_p). But it supports the invariants properties of shape and different imaging conditions (see Fig. 5). So, it is not good enough to use simple nine corner points. But rest of the four pattern matching tools such as σ_{CSQ} , σ_{HuMIs} , δ_{ABS_i} , d_{EDF_i} and larger number of corner points are indispensable features for face recognition (see Tables 3, 4 and Figs. 6a, b and 7).

Fig. 8 Some true recognition results using method 6

Since the two or more pattern matching tools are to be constructed a recognition algorithm, it is easy to add or remove any tool to form a new algorithm for getting optimum performance.

Setting of threshold values of each individual algorithm and combination of multiple algorithms are the same. As a result, performance of TPR(Recall) is increased, but the other performance evaluation measurement parameters such as RR, Precision, F-Score and FPR are decreased due to support false alarms(F_p). Therefore, it is necessary to reset(decreased) the threshold values to control the false alarms(F_p) using fusion algorithms.

Some true recognition results as a pictorial form using method 6 are shown in Fig. 8.

8 Conclusion and future work

The propose of this research activity as a facial component feature extraction analysis is based on BFCs that are basically constructed from three independent concepts such as PWP with CSQ formula, shape similarities with HuMIs and GDVs with Cartesian coordinate system and the texture concept and employed five pattern matching tools on the basis on performance analysis with TPR perspective. Combination of any two or three tools to construct the six types of independent recognition algorithms. So, algorithms are supported to add, remove or modify any facial component feature to get the optimal outcome.

Fig. 7 Accuracy curves of BioID Face Database

Fig. 8 continued

The advantage of component-based face recognition algorithm is that it is invariant to shape and also different imaging conditions such as illumination, expression, pose and partial occlusion, and the simple mathematical calculation is also done to construct the desired pattern matching tool.

Combination of multiple algorithms supported a large number of false alarms, therefore, it is necessary to reset the threshold values to restrict the false alarms using fusion algorithms. Nine FCPs are not enough to get the better outcome. The rest of the four pattern matching tools such as σ_{CSQ} , σ_{HuMs} , δ_{ABS_i} and d_{EDF_i} and large number of FCPs are the crucial features for face recognition.

The main contributions of this work are achieved the same percent of TPR values using fusion algorithms and performance analysis of each crucial facial pattern matching tool on the basis of confusion matrix.

Excluding GDVs and including Local Binary Pattern or Edge Detection Concepts as the new features and changing

the threshold values to control the false alarm with multiple databases will be the focus on our future activity.

References

1. Zhao W, Chellappa R, Phillips PJ, Rosenfeld A (2003) Face recognition: a literature survey. *ACM Comput Surv* 35(4):399–458
2. Morton J, Johnson MH (1991) CONSPEC and CONLERN: a two-process theory of infant face recognition. *Psychol Rev* 98(2):164–181
3. Paul SK, Uddin MS, Bouakaz S (2014) Face recognition using eyes, nostrils and mouth features. In: The proceedings of the 16th international conference on computer and information technology, ICCIT2013, pp 117–120, 08–10 March 2014, Khulna, Bangladesh
4. Paul SK, Uddin MS, Bouakaz S (2016) Component based face recognition using feature matching through Hu moment invariants. *Int J Comput Sci Inf Secur (IJCSIS)* 14(6):589–604
5. Paul SK, Uddin MS, Bouakaz S (2012) Extraction of facial feature points using cumulative distribution function by varying single threshold group. In: Proceedings of the IEEE/OSA/IAPR

- international conference on informatics, electronics and vision, ICEV2012, 18–19 May, 2012, Dhaka, Bangladesh, pp 806–8011
6. Paul SK, Uddin MS, Bouakaz S (2012) Extraction of facial feature points using cumulative histogram. *IJCSI Int J Comput Sci Issues* 9(1):44–51
 7. Paul SK, Bouakaz S, Shorif Uddin M (2011) Automatic adaptive facial feature extraction using CDF analysis. In: Cherifi H, Zain JM, El-Qawasmeh E (eds) *Digital information and communication technology and its applications. DICTAP 2011. Communications in computer and information science*, vol 166. Springer, Berlin
 8. Paul SK, Uddin MS, Bouakaz S (2015) Face recognition using facial features. In: *SOP transaction on signal processing (STSP) Pub. Date (Web): 2015-03-23, ISSN (Print): 2377-0538 ISSN (Online): 2377-0546, Scientific Online Publishing Co., Ltd, USA.* <http://www.scipublish.com/journals/STSP/>
 9. Rao KS, Rajagopalan AN (2005) A probabilistic fusion methodology for face recognition. *EURASIP J Adv Signal Process* 2005:813102. <https://doi.org/10.1155/ASP.2005.2772>
 10. Turk Matthew, Pentland Alex (1991) Eigenfaces for recognition. *J Cogn Neurosci* 3(1):71–86
 11. Belhumeur PN, Hespanha JP, Kriegman DJ (1997) Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *IEEE Trans Pattern Anal Mach Intell* 19(7):711–720
 12. Comon P (1994) Independent component analysis; A new concept? *Signal Process* 36(3):287–314
 13. Nabatchian A, Abdel-Raheem E, Ahmadi M (2008) Human face recognition using different moment invariants: a comparative study, image and signal processing, 2008. In: *CISP '08. Congress on*, 27–30 May 2008, China
 14. Ekenel HK, Stiefelhagen R (2005) Local appearance based face recognition using discrete cosine transform. In: *Proceedings of 13th European signal processing conference (EUSIPCO 2005)*, Antalya, Turkey, September 2005
 15. Heisele B, Serre T, Poggio T (2007) A component-based framework for face detection and identification. *Int J Comput Vis* 74(2):167–181
 16. Zhang W, Shan S, Gao W, Chang Y, Cao B (2005) Component-based cascade Linear Discriminant Analysis for face recognition. In: *Advances in biometric person authentication lecture notes in computer science*, vol 3338, pp 19–79
 17. Weyrauch B, et al. (2004) Component-based face recognition with 3D Morphable Models. In: *2004 IEEE computer society conference on computer vision and pattern recognition workshop (CVPRW'04)*, June 27–July 2, 2004, Washington DC, USA
 18. Zhang J, Xie X (2012) A study on the effective approach to illumination-invariant face recognition based on a single image. In: Zheng WS, Sun Z, Wang Y, Chen X, Yuen PC, Lai J (eds) *Biometric recognition. CCBP 2012. Lecture notes in computer science*, vol 7701. Springer, Berlin
 19. Hsieh CK, Lai SH, Chen YC (2009) Integrated expression-invariant face recognition with constrained optical flow. In: Wada T, Huang F, Lin S (eds) *Advances in image and video technology. PSIVT 2009. Lecture notes in computer science*, vol 5414. Springer, Berlin
 20. Bonnen K, Klare B, Jain AK (2013) Component-based representation in automated face recognition. *IEEE Trans Inf Forensics Secur* 8(1):239–253
 21. Hua G (2015) Probabilistic elastic part model for real-world face recognition. In: Ji Q, Moeslund BT., Hua G, Nasrollahi K (eds) *Face and facial expression recognition from real world videos. FFER 2014. Lecture Notes in Computer Science*, vol 8912. Springer, Cham
 22. Zhao H, Yuen PC, Kwok JT (2006) A novel incremental principal component analysis and its application for face recognition. *IEEE Trans Syst Man Cybern Part B Cybern* 36(4):872–886
 23. Kapoor R, Mathur P (2013) Face recognition using moments and wavelets. *Int J Eng Res Appl IJERA* 3(4):82–95
 24. Harguess J, Aggarwal JK (2009) A case for the average-half-face in 2D and 3D for face recognition. In: *Proceedings of the IEEE computer society conference on computer vision and pattern recognition (CVPR'09)*, Miami, FL, USA, June 2009, pp 7–12
 25. Brunelli R, Poggio T (1993) Face recognition: features versus templates. *IEEE Trans Pattern Anal Mach Intell* 15(10):1042–1052
 26. Viola P, Jones MJ (2004) Robust real-time object detection. *Int J Comput Vision* 57(2):137–154
 27. Gonzalez RC, Woods RE (2002) *Digital image processing*, 2nd edn. Prentice Hall, Upper Saddle River
 28. Otsu N (1979) A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern* 9(1):62–66
 29. Cochran WG (1952) The χ^2 test of goodness of fit. *Ann Math Stat* 23:315
 30. An online Book of “Learning OpenCV” by Gary Bradski and Adrian Kaehler, Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472. Website: <http://it-ebooks.info/book/299/>
 31. Flusser J (2006) Moment invariants in image analysis. In: *Proceedings of world academy of science, engineering and technology volume 11, Issn 1307-6884*
 32. Hu M (1962) Visual pattern recognition by moment invariants. *IRE Trans Inf Theory* 8(2):179–187
 33. Fawcett Tom (2006) An introduction to ROC analysis. *Pattern Recogn Lett* 27(2006):861–874
 34. BioID Face Database. <http://www.bioid.com/downloads/facedb/index.php>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.