Check for
updates

Fuzzy triangulation signature for detection of change in human emotion from face video image sequence

Md Nasir[1] · Paramartha Dutta[1] · Avishek Nandi[1]

Abstract

The present article proposes a geometry-based fuzzy relational technique for capturing gradual change in human emotion over time available from relevant face image sequences. As associated features, we make use of fuzzy membership arising out of five triangle signatures such as - (i) Fuzzy Isosceles Triangle Signature (FIS), (ii) Fuzzy Right Triangle Signature (FRS), (iii) Fuzzy Right Isosceles Triangle Signature (FIRS), (iv) Fuzzy Equilateral Triangle Signature (FES), and (v) Other Fuzzy Triangles Signature (OFS) to achieve the task of appropriate classification of facial transition from neutrality to one among the six expressions viz. anger (AN), disgust (DI), fear (FE), happiness (HA), sadness (SA) and surprise (SU). The effectiveness of the Multilayer Perceptron (MLP) classifier is tested and validated through 10 fold cross-validation method on three benchmark image sequence datasets namely Extended Cohn-Kanade (CK+), M&M Initiative (MMI), and Multimedia Understanding Group (MUG). Experimental outcomes are found to have achieved accuracy to the tune of 98.47%, 93.56%, and 99.25% on CK+, MMI, and MUG respectively vindicating the effectiveness by exhibiting the superiority of our proposed technique in comparison to other state-of-the-art methods in this regard.

✉ Md Nasir
  nasir.vb@gmail.com

  Paramartha Dutta
  paramartha.dutta@gmail.com

  Avishek Nandi
  avisheknandi10@gmail.com

[1] Visva-BharatiUniversity, Santiniketan, WestBengal, India

1 Introduction

Effectual person-robot interaction is a recent significant aspect in the affective computing field. In the interaction between people, verbal cues (such as spoken words, voice) and non-verbal cues (such as body gesture, facial expression) are used to describe their feeling [22]. The improvement of human-robot interaction can be ensured if a robot can interpret accurately the internal meaning of human emotion from facial expression. According to Pantic et al. [25], facial expression has a higher contribution than other cues to express the mental state of humans. There are several applications of facial expression recognition with high demand in the human healthcare system. The recognition system can help to identify the actual health condition of patients by examining the emotional behavior [30]. Here, the system takes into cognizance the patient's facial expression to interpret whether it belongs to positive expressions or negative expressions. If it displays any positive expressions such as happiness, gladness then patients can be considered to be in a healthy condition while the unhealthy condition is considered by observing any negative expressions such as sadness, anger. The reflection of human emotion in the face is undeniable. Even though the appropriate nature of association prevalent between the emotional states of mind with its reflection in the face is premature. The type of such association needs the application of cognitive science on one hand and multimedia framework on the other. So far, several works on the identification of emotion have been reported in other studies on the affective computing field. Among them, the work of Ekman and Friesen [11] established the idea for the detection of human emotion from facial muscle movements. In their literature study, it is reported that there is a correlation between the deformation of face components and expression. They introduced six different universal human emotions viz. anger, disgust, fear, happiness, sadness, and surprise by characterizing the similarity of expressions under different cultures, ages, and sex. Due to a large amount of variability in face appearance, lots of online applications for face modeling has been developed by the researchers in the recent era. Aging, rejuvenation, facial expressions are a few face appearances, considered in online system [8] to understand about social aspects of the face like biometric information of face appearance. Moreover, authors in [8] tried to predict special effects of facial appearance with facial expressions, aging, and face rejuvenation by their proposed web-based online system. In our proposed work, we are dealing with the issue of understanding the dynamic behavior of human emotion through the application of the proposed recognition system.

Several footsteps involving the traditional emotion recognition system have been utilized by the researchers for the study about the activities associated with emotion. These footsteps can be categorized into three major processes: facial components finding, exploration of feature, and identification of feature [34]. In the first step, the entire face region with specifying major face components: eyes, eyebrows, mouth are first detected from input face image (static images or video clips). In the step of feature exploration, appearance-based information, or geometry-based information is collected from face regions [17]. Appearance-based information includes texture, pixel intensity variation with respect to facial expression and these appearance elements are collected either from the whole area of the face or areas of major components [14]. It is described in [7] that somehow facial deformation is affected by the different face appearance elements such as changes in skin texture, changes in muscle points in terms of intensity values. Authors in [7], applied the facial rejuvenation process on facial images to describe facial distortion through the correction of the above mentioned automatic changes, observed on the face. In the case of geometry-based information, it includes shape information of major face regions rather than the entire face region [38].

Finally, explored features are exploited in the classification modules to identify the emotion separately.

In this article, we propose Fuzzy membership features such as fuzzy Isosceles triangle signature (FIS), fuzzy right triangle signature (FRS), fuzzy isosceles and right triangle signature (FIRS), fuzzy equilateral triangle signature (FES) and other fuzzy triangles signature (OFS) for better understanding about the facial transition. These features are individually fed into MLP recognizer to justify the task of differentiation of facial transitions separately. Landmarks on major face parts are found very useful to capture the geometric shape divergence on the face plane in emotional transition. Sequential face images are utilized by the appearance model to identify landmark points. We utilize only significant landmarks such as eight points on both eyes, six points on both eyebrows, three points on the nose, and six points on the mouth portion of the face. The fuzzy triangulation technique is applied to these landmarks to construct triangle shapes by considering each combination of three landmarks and we identify membership values of five different fuzzy triangles to obtain five different fuzzy triangle signatures discussed above. After utilizing these five signatures into MLP recognizer separately, five several recognition results are obtained. We exclusively report all results obtained through the implementation of our system on three benchmark face video datasets viz. CK+, MMI, and MUG and compare the results with others. It is observed that five proposed signatures perform separately better than others by showing higher recognition results on three datasets.

Motivation A face is an amalgamation of muscle articulations that are distorted from a neutral expression to any universal facial expressions because of the contraction of muscle on the face. During this muscle contraction, nonrigid movements are found around the eyes, eyebrows, lips, and nose on the face plane [33]. The identification of emotional class depends on the measurements obtained from feature points on those face components. In the literature study [4–6], authors used static images of highest intensity facial expression to identify the expressional class. It is difficult to estimate quantitative information about the motion of geometric shapes (eye, nose, lips, and eyebrows) while feature points move from frame to frame in a transition of emotion. To fix this problem, we introduce an automatic recognition system that uses a sequence of face frames denoting the transition of emotion from neutral to universal expression. Here we utilized the fuzzy triangulation technique to generate a geometric triangle shape derived from feature points. These shapes are used to understand the nonrigid motion of different face components for a dynamic expression by producing fuzzy membership signatures.

Contribution of our proposed work includes

I   Presenting a frame-based facial expression analysis to estimate flow control of behavioral changes in the transition of human emotion available in terms of video clips.
II Tracking landmark positions from face frame sequence with the application of AAM [35] to trace the frame-wise deviation of feature points.
III Introducing a fuzzy triangulation technique that is used to build fuzzy triangle shapes by taking a trio combination of landmark points to assess the fuzzy relationship among major portions of the face (eyes, eyebrows, nose, and lips).
IV Discussing the separately significant influence of five different fuzzy triangle membership signatures on recognition of dynamic changes in human emotion such as fuzzy Isosceles triangle signature (FIS), fuzzy right triangle signature (FRS), fuzzy isosceles and right triangle signature (FIRS), fuzzy equilateral triangle signature (FES) and other fuzzy triangles signature (OFS).

V   Briefly examining the system performance on three benchmark face sequence datasets: CK+ [19], MMI [36] and MUG [2] with the Multilayer Perceptron (MLP) classification task.

The paper structure of the remaining parts is as follows. Surveys on different previous works are narrated in Section 2. The proposed approach with the explanation of Landmark Points Detection, Fuzzy Triangle Based Geometric Feature Exploration, and Feature Learning by Multilayer Perceptron Network Module is represented in Section 3. Section 4 discusses experimental setup and different results with various benchmark image sequence datasets description. Section 5 reports comparative outcomes of our proposed approach with other works. Section 6 draws conclusion.

2 Literature survey

Several endeavors figure in literature distinguishing one facial expression from another. The task of prominent feature discovery has been mainly focused on the traditional expression identification systems. Feature invention procedure can be considered in two different ways: geometry-based invention and appearance-based invention. The most challenging task in the geometry-based invention is to locate the proper facial landmarks on the face image. More accuracy in finding the locations of landmarks ensures adequacy of feature mapping into the appropriate facial expression. In study [13], landmark identification is done with the combined application of the elastic bunch graph matching (EBGM) algorithm and Kanade–Lucas–Tomaci (KLT) tracker on face images. Authors in [13] first initialized landmarks by using the EBGM algorithm and after that, they tracked the location of landmarks points. Such effort of landmark identification generated distinguishable geometric features from selected points, lines, and triangles and they found different recognition results for different representations. Active shape model (ASM) [10] is applied to find landmarks points using a matching algorithm that works with point distribution information. ASM model constructs a statistical model for deformable shape objects to extract landmarks from face regions. But it is difficult to identify stable landmarks in the sequential frames due to the movement of the position. Authors in [1] presented a framework to generate time-varying features from sequential face frames by using ASM with the help of Lucas-Kanade (LK) optical flow application. In the literature study [9], authors have observed that AAM is found as a more useful landmark tracker to obtain principle landmark locations and these landmark locations are used for producing prominent geometric feature to recognize facial expression properly.

In [41], authors used only two face frames from each sequence to create discriminative features, one is a normal face frame and another is an emotional face frame with maximum expression. Here authors also focused on important facial points to generate differential geometric features by considering differences between distances of facial points in the normal frame and maximum expressional frame. Authors in [27] presented a novel framework to build an efficient recognition system for distinguishing various facial expressions. They also applied geometric features computed by using different statistical measurement techniques within their system and this illustrates the ability to capture the information about the deformation of facial features with the help of hidden Markov models (HMM) recognizer. Authors in [32] compared the recognition capability of combined geometric features with the individual features. They tested their system on individual features that include landmark information and their relative distances as well as on a combination of these two individual features. The combined feature is found more reliable than individual features to

differentiate different facial expressions with the help of an ensemble neural network recognizer. On the other side, a statistical measurement of pixel intensities is used for computing appearance information from the face image. Authors in [24] first established the local binary pattern (LBP) operator as an effective texture information retriever from an image. LBP operator encodes each pixel of an image by taking 3×3 neighborhood pixels of a particular center pixel. Here the binary number is assigned on each pixel after thresholding the template of size 3×3 and a histogram is generated from such binary labeled context. Such histograms help to understand the local pattern distribution of edges, flat areas, and spots, etc. Authors in [31], have found that only a basic LBP operator falls short to describe image features. They used an extended version of LBP with the flexibility in usage of different number of neighborhoods to acquire foremost features. Apart from LBP feature, histogram orientation gradient [12], local Gabor binary pattern [43] are used as appearance features for the detection of facial expression.

In [15], authors tried to extract hybrid features from important facial patches having a major role in changes of facial expressions. They used both geometric shape features as well as appearance features. They also reported the novelty of their work by reducing the computational cost for the extraction of discriminative features. Apart from the task of feature invention, there is another important step of classification learning which is mandatory for the emotion recognition system to discriminate extracted features into different expressional groups. Several classifiers such as support vector machine (SVM) [16], hidden Markov models (HMM) [40], K-Nearest Neighbor (KNN) [21], Multi-Layer Perceptron Neural Network (MLPNN) [26] are employed into the recognition system to estimate recognition results. The choice of an effective classification module plays a crucial role to ensure the robustness of the system because different classifiers may compute different recognition accuracies based on their capabilities.

3 Proposed approach

It is a very challenging task to capture the shape deformation of the face due to vagueness in the border of geometric shape [37]. Our proposed approach tries to deal with this problem by defining the triangle shapes with the fuzzy relationship prevalent among them. The emotional transition recognition process reported in this article is segmented into three important sub-processes namely identification of landmark points, computation of various fuzzy membership signatures corresponding to various triangles under consideration, and classification of the appropriate nature of the transition. Figure 1 displays the workflow of our system.

3.1 Landmark points detection

Uniform allocation of geometric positions into the face frame plays important role in the exploration of prominent geometric features. In our proposed system, a well-known landmark allocation algorithm namely active appearance model (AAM) [35] is put into use to track the reckonable coordinate points on the face plane. AAM model can be stated as a union of two different statistical models which unify them into a single model by collecting both geometric shape information and colors/intensity-based information from deformable face object. The model starts working with some training image samples which are annotated with initial landmark points. Here Procrustes analysis process is utilized to allocate initial landmarks on those training samples and each annotated training sample is denoted
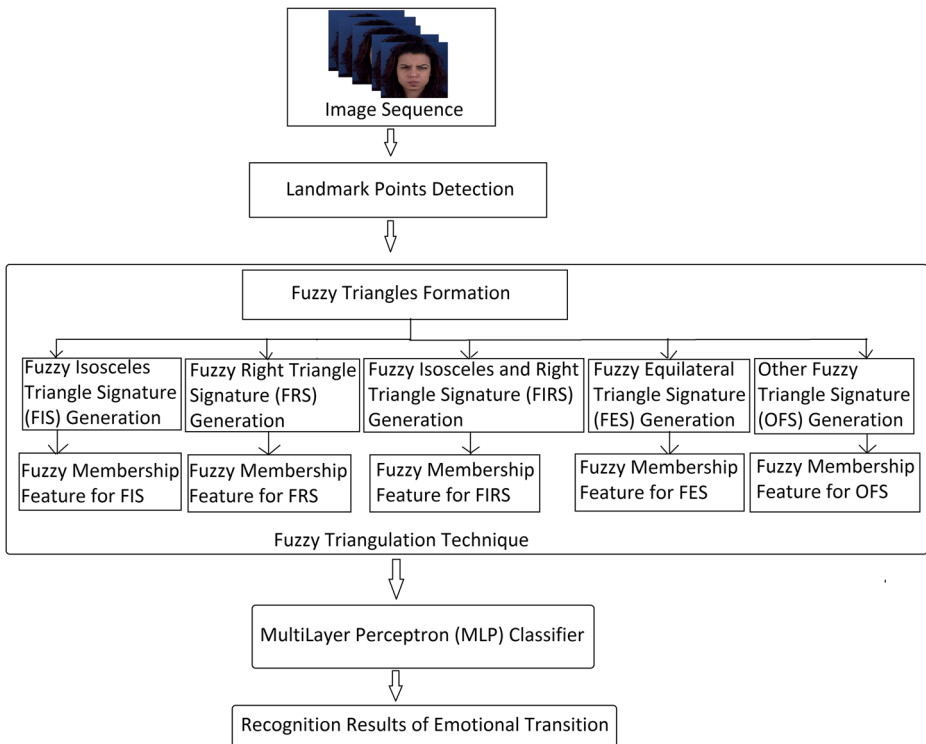
Fig. 1 Work flow of our proposed system

by the vector s containing landmarks as an element thereby building a shape model. This model is repeatedly trained to get better matching of initial landmark points with the mean shape $\bar{s}$. Such a contribution of the work is made from the application of an active shape model (ASM) [10]. On the other hand, the process of Eigen-analysis is executed to build a texture model that describes the local pattern of patches available in the shape region. After doing normalization of such texture information, it is stored into the vector g. At last, the correlation between the shape model and texture model is computed through the learning process of the parameters in the (1) and (2) to build an appearance model.

$$s = \bar{s} + Q_s C \tag{1}$$

$$g = \bar{g} + Q_g C \tag{2}$$

Where $\bar{s}$ and $\bar{g}$ are mean shape and mean texture respectively, c is control parameter. Q_s and Q_g are shape variation and texture variation respectively. After executing the application of such an appearance model on consecutive face frames we obtain a total of 68 various geometric coordinates describing the entire face region for each individual frame. Among them, only 23 coordinates are found very informative and as the expression changes, those coordinates are dislocated accordingly [4]. All informative points are taken from major face regions viz. eyes (8 points), eyebrows (6 points), nose (3 points), and lips (6 points). Figure 2 illustrates informative landmarks detection from a single face frame.
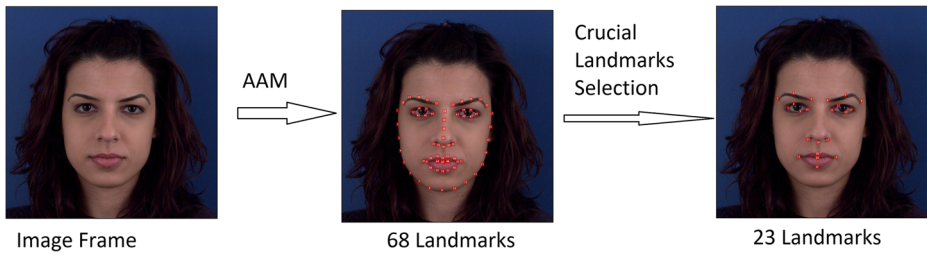
Fig. 2 Informative landmarks detection from single face frame

3.2 Fuzzy triangle based geometric feature exploration

Here, the fuzzy triangulation technique is introduced to capture the changing behavior of human emotion. This technique starts taking the combined information of every three landmark points from the set of 23 landmarks one by one and generates $^{23}C_3 = 1771$ many triangle shapes for each frame in the sequence. The dependency of displacement among landmarks is very high when facial expression evolves with the time span. The characteristics of such dependency are measured in the form of various components of the triangle. Let, $(a_l, b_l, c_l, A_l, B_l, C_l)^m_{i,j,k}$ represent three sides and three angles of triangle formed by i, j, k the three vertex points chosen as landmark in l^{th} frame in m^{th} sequence. First three components are calculated by using Euclidean distance metric given in (3), (4), (5) and last three components are computed by using following (6), (7) and (8).

$$a_l = \sqrt{(y_k - y_j)^2 + (x_k - x_j)^2} \tag{3}$$

$$b_l = \sqrt{(y_k - y_i)^2 + (x_k - x_i)^2} \tag{4}$$

$$c_l = \sqrt{(y_j - y_i)^2 + (x_j - x_i)^2} \tag{5}$$

$$A_l = \cos^{-1}(\frac{b_l^2 + c_l^2 - a_l^2}{2 \times b_l \times c_l}) \tag{6}$$

$$B_l = \cos^{-1}(\frac{a_l^2 + c_l^2 - b_l^2}{2 \times a_l \times c_l}) \tag{7}$$

$$C_l = \cos^{-1}(\frac{a_l^2 + b_l^2 - c_l^2}{2 \times a_l \times b_l}) \tag{8}$$

Formation of triangle and computation of angles are demonestrated in Fig. 3.

Now, we consider the universe of discourse U by taking three angles A_l, B_l, C_l as member and it is formulated as per (9).

$$U = \{(A_l, B_l, C_l) | A_l \geq B_l \geq C_l \geq 0; A_l + B_l + C_l = 180^o\} \tag{9}$$

Next, we fuzzyfy these angle components into different fuzzy triangle families such as fuzzy isosceles triangle (I), fuzzy right triangle (R), fuzzy isosceles and right triangle (IR), fuzzy equilateral triangle (E) and other fuzzy triangle (T) by using following fuzzy membership rules shown in (10), (11), (12), (13) and (14) respectively.

$$\{\mu_I(A_l, B_l, C_l)\}^m_{i,j,k} = 1 - \frac{1}{60} \min\{(A_l - B_l, B_l - C_l)\}^m_{i,j,k} \tag{10}$$
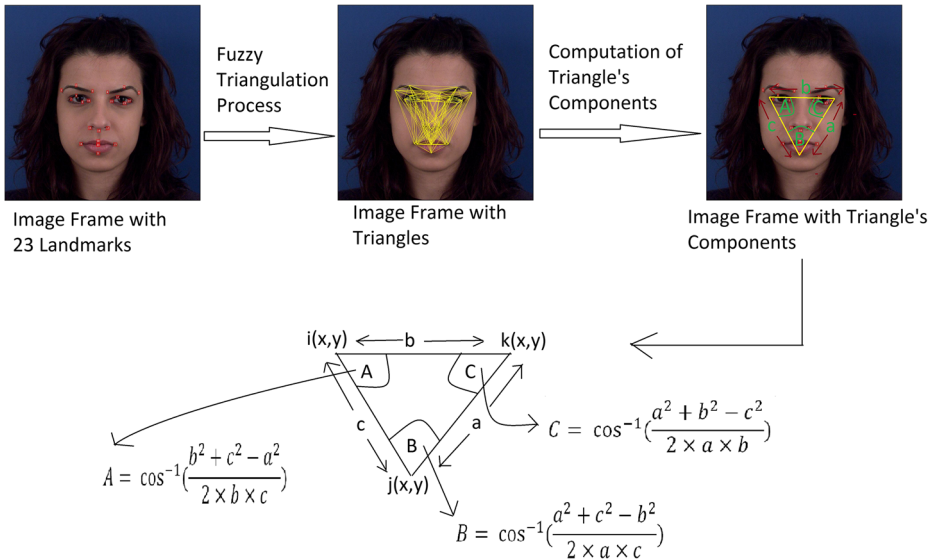
Fig. 3 Pictorial representation for computation of angle components from single triangle of a perticular frame

$$\{\mu_R(A_l, B_l, C_l)\}_{i,j,k}^m = 1 - \frac{1}{90} \mid A_l - 90 \mid_{i,j,k}^m \tag{11}$$

$$\{\mu_{IR}(A_l, B_l, C_l)\}_{i,j,k}^m = \min\{\mu_I(A_l, B_l, C_l), \mu_R(A_l, B_l, C_l)\}_{i,j,k}^m \tag{12}$$

$$\{\mu_E(A_l, B_l, C_l)\}_{i,j,k}^m = 1 - \frac{1}{180}(A_l - C_l)_{i,j,k}^m \tag{13}$$

$$\{\mu_T(A_l, B_l, C_l)\}_{i,j,k}^m = 1 - \max\{\mu_I(A_l, B_l, C_l), \mu_R(A_l, B_l, C_l), \mu_E(A_l, B_l, C_l)\}_{i,j,k}^m \tag{14}$$

Finally, five different fuzzy triangle membership signatures are formalized to signify a perticular face video sequence in various ways. These are defined by (15), (16), (17), (18) and (19) respectively.

Face sequence representation by fuzzy isosceles triangle signature

$$(FIS)_{i,j,k}^m = [\{\mu_I(A_0, B_0, C_0)\}_{i,j,k}^m, \{\mu_I(A_1, B_1, C_1)\}_{i,j,k}^m,$$
$$............, \{\mu_I(A_n, B_n, C_n)\}_{i,j,k}^m] \tag{15}$$

Face sequence representation by fuzzy right triangle signature

$$(FRS)_{i,j,k}^m = [\{\mu_R(A_0, B_0, C_0)\}_{i,j,k}^m, \{\mu_R(A_1, B_1, C_1)\}_{i,j,k}^m,$$
$$............, \{\mu_R(A_n, B_n, C_n)\}_{i,j,k}^m] \tag{16}$$

Face sequence representation by fuzzy isosceles and right triangle signature

$$(FIRS)_{i,j,k}^m = [\{\mu_{IR}(A_0, B_0, C_0)\}_{i,j,k}^m, \{\mu_{IR}(A_1, B_1, C_1)\}_{i,j,k}^m,$$
$$............, \{\mu_{IR}(A_n, B_n, C_n)\}_{i,j,k}^m] \tag{17}$$

Face sequence representation by fuzzy equilateral triangle signature

$$(FES)^m_{i,j,k} = [\{\mu_E(A_0, B_0, C_0)\}^m_{i,j,k}, \{\mu_E(A_1, B_1, C_1)\}^m_{i,j,k},$$
$$............, \{\mu_E(A_n, B_n, C_n)\}^m_{i,j,k}] \tag{18}$$

Face sequence representation by Other Fuzzy Triangles Signature

$$(OFS)^m_{i,j,k} = [\{\mu_T(A_0, B_0, C_0)\}^m_{i,j,k}, \{\mu_T(A_1, B_1, C_1)\}^m_{i,j,k},$$
$$............, \{\mu_T(A_n, B_n, C_n)\}^m_{i,j,k}] \tag{19}$$

Sequences have n frames. $n = 10$ is considered in present context. Thus $1771 \times 10 = 17710$ is feature vector size. The computation of five fuzzy triangle memberships is pictorially explained in Fig. 4.

We used the combination formula in (20) that calculates the number of combinations of triangles is used in our system to get prominent features. While using the twenty-three number of landmark points instead of the sixty-eight number we computed feature components only for the number of triangles that are found prominent. It actually reduces the number of computation through discarding all other triangles which are not prominent. For a better understanding of computational overhead, we described Case1 and Case2 in (21) and (22) respectively and compared them.

$$T = C(l, p)$$
$$= \frac{l!}{(p!(l - p)!)} \tag{20}$$

Number of triangles in single frame is T. l is landmark number available in single frame. p landmarks are used to make a single triangle.

Frame Index = l
Triangle Index = i, j, k
Sequence Index = m

$\{\mu_I(A_l, B_l, C_l)\}^m_{i,j,k}$

Fuzzy Isosceles Triangle Membership

$\{\mu_R(A_l, B_l, C_l)\}^m_{i,j,k}$

Fuzzy Right Triangle Membership

$\{\mu_{IR}(A_l, B_l, C_l)\}^m_{i,j,k}$

Fuzzy Isosceles and Right Triangle Membership

$\{\mu_E(A_l, B_l, C_l)\}^m_{i,j,k}$

Fuzzy Equilateral Triangle Membership

$\{\mu_T(A_l, B_l, C_l)\}^m_{i,j,k}$
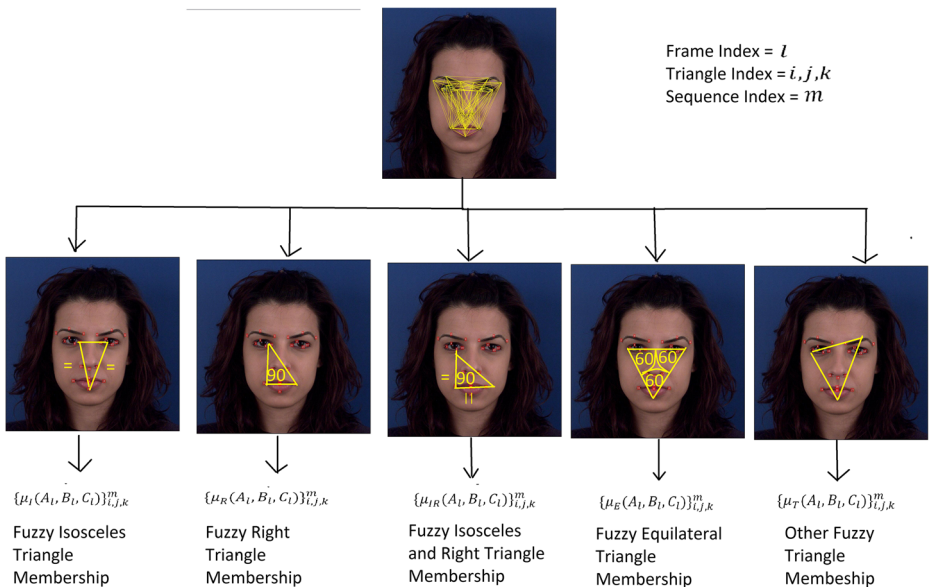
Other Fuzzy Triangle Membeship

Fig. 4 Five different fuzzy triangle membership values are computed from a single frame

Case1:
$l = 68$, $p = 3$, then

$$
\begin{aligned}
T_{case1} &= C(68, 3) \\
&= \frac{68!}{(3!(68-3)!)} \\
&= 50116
\end{aligned}
\tag{21}
$$

Case2:
$l = 23$, $p = 3$, then

$$
\begin{aligned}
T_{case2} &= C(23, 3) \\
&= \frac{23!}{(3!(23-3)!)} \\
&= 1771
\end{aligned}
\tag{22}
$$

From (21) and (22), it is found that computation complexity of $T_{case2} < T_{case1}$

3.3 Feature learning by multilayer perceptron network module

Our extracted geometric features are finally used in the proposed Multi-Layer Perceptron (MLP) classifier [3] to be recognized into different emotional classes with the representation of their transitional behaviors. Here, MLP uses such network architecture which is modeled by considering a set of nodes coming from three different layers (input layer, hidden layer, and output layer) and connections among them. Such nodes of the input layer take features as input followed by processing with connection waits and bias value so as to pass the input signal to the next layer called the hidden layer. Similarly, a signal from the hidden layer is transmitted to the output layer to estimate the classification outcomes. Each node is associated with activation function $\tan h$ to forward the signals. At this point, the important task is to check and control the error rate between the estimated outcome and the target outcome. Thus the learning rule for the network training is required for adjusting the error rate. Here, the backpropagation learning rule is applied to modify the weight value of the network connections. The learning rule uses the scaled conjugate gradient algorithm to reduce the error rate by setting up the network parameter by calibrating the learning rate at a suitable level. A lower learning rate makes a greater accuracy in the classification. The network stops learning when it finds a minimal error rate. The error is calculated by using (23).

$$
\epsilon = \frac{1}{2} \sum_k (\tau_k - \gamma_k)^2
\tag{23}
$$

Here, τ_k is the target outcome, and γ_k is the estimated outcome at k^{th} output node. The (24) shows how Partial differentiation is utilized for this error function to control the weight values and make the error minimization.

$$
\delta \omega_{jk} = -\alpha \frac{\partial \epsilon}{\partial \omega_{jk}}
\tag{24}
$$

Here, α is learning rate, $\delta \omega_{jk}$ is connection weight between nodes j and k. Our proposed network consists of 17710 many input nodes (as the number is equal to our feature vector size), 10 hidden nodes (at this number our system is found to yield good results) and 6 output nodes (as we considered six basic emotions). Feature learning process is described in Algorithm 1.

Algorithm 1 Aalgorithm for feature learning by using MLP.

Require: Face video sequence with labeled expression images, S
Ensure: Emotional transition detection
 1: **for** each sequence $s \in S$ **do**
 2:     **for** each frame $f \in s$ **do**
 3:         Execute AAM landmark tracker on frame f and generate set of informative landmarks, $L = [(x_1, y_1), (x_2, y_2), ..., (x_23, y_23)]$
 4:         Apply fuzzy triangulation technique on the set L to get set of triangles, $T = [(x_i, y_i), (x_j, y_j), (x_k, y_k)]$ where, $i, j, k \in L$
 5:         **for** each triangle $t_{i,j,k} \in T$ **do**
 6:             Calculate triangle components $(a_f, b_f, c_f, A_f, B_f, C_f)^s_{i,j,k}$, where $(a_f, b_f, c_f) \rightarrow$ three side lengths and $(A_f, B_f, C_f) \rightarrow$ three angles of triangle $t_{i,j,k}$
 7:             Consider Universe of discourse $U = \{(A_f, B_f, C_f)|A_f \geq B_f \geq C_f \geq 0; A_f + B_f + C_f = 180^o\}$
 8:             Compute five different fuzzy membership values such as $\{\mu_I(A_f, B_f, C_f)\}^s_{i,j,k}$, $\{\mu_R(A_f, B_f, C_f)\}^s_{i,j,k}$, $\{\mu_{IR}(A_f, B_f, C_f)\}^s_{i,j,k}$, $\{\mu_E(A_f, B_f, C_f)\}^s_{i,j,k}$, $\{\mu_O(A_f, B_f, C_f)\}^s_{i,j,k}$.
 9:             Store all membership values into the following vectors: $V_{I^s_{f,(i,j,k)}} = \{\mu_I(A_f, B_f, C_f)\}^s_{i,j,k}$, $V_{R^s_{f,(i,j,k)}} = \{\mu_R(A_f, B_f, C_f)\}^s_{i,j,k}$, $V_{IR^s_{f,(i,j,k)}} = \{\mu_{IR}(A_f, B_f, C_f)\}^s_{i,j,k}$, $V_{E^s_{f,(i,j,k)}} = \{\mu_E(A_f, B_f, C_f)\}^s_{i,j,k}$, $V_{O^s_{f,(i,j,k)}} = \{\mu_O(A_f, B_f, C_f)\}^s_{i,j,k}$.
10:         **end for**
11:         Concatenate all vectors in frame wise to obtain different fuzzy membership signatures such as: $FIS^s = V_{I^s_f} \cup V_{I^s_{f+1}}$, $FRS^s = V_{R^s_f} \cup V_{R^s_{f+1}}$, $FIRS^s = V_{IR^s_f} \cup V_{IR^s_{f+1}}$, $FES^s = V_{E^s_f} \cup V_{E^s_{f+1}}$, $OFS^s = V_{O^s_f} \cup V_{O^s_{f+1}}$
12:     **end for**
13:     Each signature is fed into MLP seperatly one by one
14:     Obtain emotional transition recognition
15: **end for**

---

4 Experimentation and result discussion

To ensure the performance consistency of our proposed recognition system, we orchestrate experimentations for five different proposed geometric signatures separately. Well-known video datasets of facial expressions viz. CK+ [19], MMI [36], and MUG [2] are used for our experimentations. The performance of each signature is evaluated on those three datasets separately. In this experimentation, each dataset is divided into three non-overlapping segments. The first segment includes training data (70%), the second segment includes validation data (15%) and the third segment includes test data (15%). Training data is utilized to initially learn the MLP network with the fitting of parameters such as learning rate, edge weights, etc. Mostly, the network does not give unbiased results with this training set. The network faces with overfitting problem during this training process due to the improper learning of network parameters. Thus subsequently the validation dataset is used to tune properly the parameters of the network. This set helps the network to find a starting point as well as an endpoint for the overfitting problem. With the help of such supervised learning, a best-fitted network model is obtained for our experimentation. Finally, the test dataset

is reserved from the main dataset comprising samples unused in training and validation. This dataset helps the network to gain recognition results from the best-fitted model as an unbiased outcome. The results are evaluated through the computation of confusion matrices from every dataset. For the detailed study on the influence of different signatures over those datasets, other measurement parameters such as False Acceptance Rate (FAR), False Rejection Rate (FRR), and Error rate (ERR) are also incorporated here. These parameters are defined by the following (25), (26) and (27).

$$FAR = \frac{FP}{FP + TN} \tag{25}$$

$$FRR = \frac{FN}{FN + TP} \tag{26}$$

$$ERR = \frac{FP + FN}{TP + TN + FP + FN} \tag{27}$$

Here, $FP \rightarrow$ False Positive, $FN \rightarrow$ False Negative, $TN \rightarrow$ True Negative and $TP \rightarrow$ True Positive.

– FAR: It measures the proportion that the classifier incorrectly recognizes an image sequence that is not available in the actual expressional class.
– FRR: It measures the proportion that classifiers incorrectly discard an image sequence to be recognized which is available in the actual expressional class.
– ERR: It measures the proportion by taking the ratio between all image sequences that are incorrectly recognized and discarded by the classifier and the total number of image sequences is available in the dataset.

Further, the effectiveness of all proposed signatures is justified by presenting the implementation of k-fold cross-validation in this article. Dataset-wise description and performance assessment follow subsection wise.

4.1 Discussion on extended Cohn-Kanade (CK+) dataset

The dataset stores facial expression profiles which are recorded from 210 different peoples. Among them, most belong to the female category (69%) and 81% of this female category are Euro-American, while the rest are from Afro-American (13%) and other cultures (6%). The ages of those people vary from 18 to 50 years. A total of 593 face sequences taken from 123 subjects are available within this dataset. Each sequence includes several image frames (vary from 6 to 60) which are captured and digitized into pixel arrays (640 × 490 or 640 × 480). Only 327 sequences are labeled with 7 different expressions: anger (AN), contempt (CON), disgust (DI), fear (FE), happiness (HA), sadness (SA), and surprise (SU). The dataset with those 327 labeled sequences is used in our proposed experimentation. The expression profile of happiness as a typical representative example is shown in the first row of Fig. 5 which displays the transition from a neutral expression to a happiness expression.

4.1.1 Result analysis on CK+ Dataset

In this section, the performances of all five fuzzy signatures on the CK+ dataset are described individually. Each signature is computed from this dataset to detect the transitional behavior of basic facial expressions. The associated number of various expressions are utilized from this dataset, 83 (AN), 18 (CON), 59 (DI), 25 (FE), 69 (HA), 28 (SA), and 83 (SU). Table 1 displays confusion matrices corresponding to those five signatures
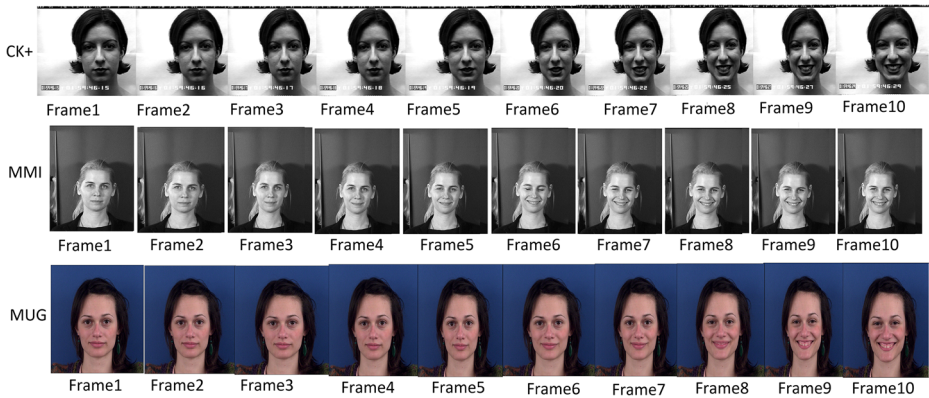
Fig. 5 Image sequence of CK+, MMI and MUG dataset

computed from the CK+ dataset resulted from the application of the MLP classifier. The corresponding analysis graphs of confusion matrices for CK+ dataset are displayed in Fig. 6.

Fuzzy Isosceles Triangle Signature (FIS) From Table 1, it is observed that the FIS signature is able to find out contempt, fear, happiness, and surprise with 100% recognition accuracy. 43 anger expressions are classified accurately but the remaining 2 expressions out of 45 anger expressions are misclassified with disgust. The signature classifies 55 disgust expressions and 26 sadness expressions perfectly. Among all disgust expressions, 3 are misclassified with anger and only 1 is mismatched with sadness expression. Out of the 28 sadness expressions, 1 is identified incorrectly with expressing anger, and 1 is with fear expression. Here, 97.55% overall accuracy is found on the CK+ dataset.

Fuzzy Right Triangle Signature (FRS) It is able to find the transition of expressions such as anger, contempt, disgust, happiness, and surprise properly without any error. In the case of sadness expression, 26 expressions are identified correctly and 2 expressions are wrongly identified as anger. The lower accuracy is found in the recognition of fear expression. Here 3 expressions are categorized into anger class and 1 is categorized into happiness class but 21 expressions are found appropriately classified. The overall recognition rate of 98.16% is found by applying this signature on the CK+ dataset.

Fuzzy Isosceles and Right Triangle Signature (FIRS) FIRS signature is able to identify the expressions (contempt, disgust, happiness, and surprise) perfectly with no misclassification. 44 anger transitions are detected properly and 1 is misclassified with sadness. Fear classifies 22 transitions correctly but it finds a total of 3 misclassifications distributed equally into contempt (1), sadness (1), and surprise (1). Among 28 transitions of sadness, only 1 transition is incorrectly identified as a surprise and the rest of the transitions (27) are properly classified into exact class. The signature could achieve overall accuracy of 98.47% on the CK+ dataset.

Fuzzy Equilateral Triangle Signature (FES) The transition of anger, happiness, and surprise are correctly recognized by this signature without any misclassification. 25 transitions of sadness, 24 transitions of fear, 58 transitions of disgust, and 16 transitions of contempt are

Table 1 Confusion Matrices for different signatures on CK+ Dataset

|	AN	CON	DI	FE	HA	SA	SU
Fuzzy Isosceles Triangle Signature							
AN	43	0	2	0	0	0	0
CON	0	18	0	0	0	0	0
DI	3	0	55	0	0	1	0
FE	0	0	0	25	0	0	0
HA	0	0	0	0	69	0	0
SA	1	0	0	1	0	26	0
SU	0	0	0	0	0	0	83
Fuzzy Right Triangle Signature							
AN	45	0	0	0	0	0	0
CON	0	18	0	0	0	0	0
DI	0	0	59	0	0	0	0
FE	3	0	0	21	1	0	0
HA	0	0	0	0	69	0	0
SA	2	0	0	0	0	26	0
SU	0	0	0	0	0	0	83
Fuzzy Isosceles and Right Triangle Signature							
AN	44	0	0	0	0	1	0
CON	0	18	0	0	0	0	0
DI	0	0	59	0	0	0	0
FE	0	1	0	22	0	1	1
HA	0	0	0	0	69	0	0
SA	0	0	0	0	0	27	1
SU	0	0	0	0	0	0	83
Fuzzy Equilateral Triangle Signature							
AN	45	0	0	0	0	0	0
CON	0	16	0	0	1	0	1
DI	0	0	58	0	0	1	0
FE	1	0	0	24	0	0	0
HA	0	0	0	0	69	0	0
SA	2	0	0	0	0	25	1
SU	0	0	0	0	0	0	83
Other Fuzzy Triangles Signature							
AN	43	0	1	0	0	1	0
CON	4	14	0	0	0	0	0
DI	2	0	57	0	0	0	0
FE	0	0	0	25	0	0	0
HA	0	0	0	0	69	0	0
SA	0	0	0	0	0	28	0
SU	0	1	0	0	0	0	82
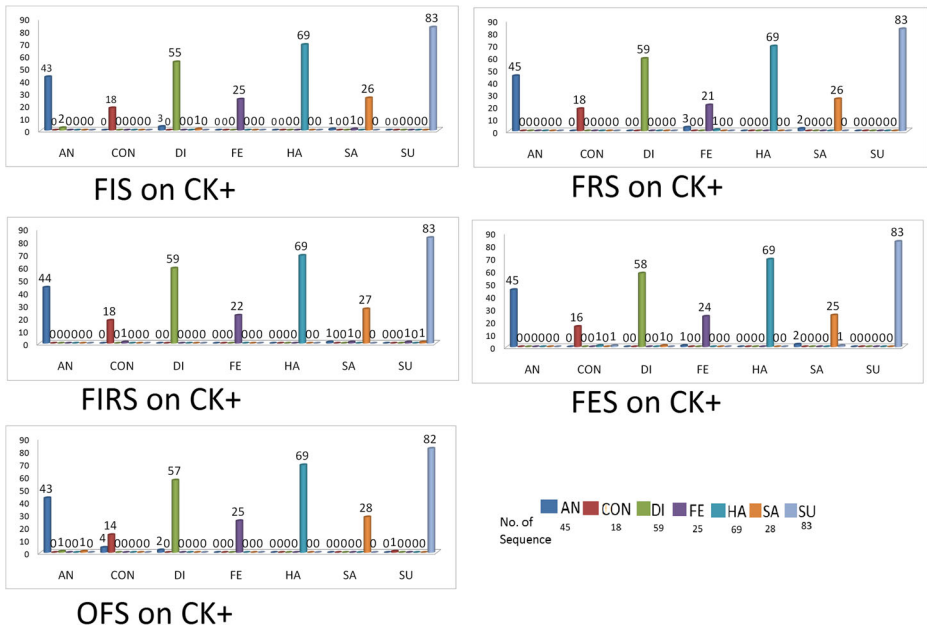
Fig. 6 Visualization graphs of CK+ confusion matrix

recognized perfectly. 1 transition of disgust is incorrectly recognized as sadness. 2 transitions of sadness and 1 transition of fear are incorrectly identified as anger. 1 transition of sadness and 1 transition of contempt are misclassified as a surprise. 1 transition of contempt is found confused with happiness. The overall recognition rate of 97.85% is achieved by this signature.

Other Fuzzy Triangles Signature (OFS) Fear, happiness, and sadness recognize all their transitions into exact classes accurately. 14 transitions of contempt, 43 transitions of anger, 57 transitions of disgust, and 82 transitions of surprise are recognized properly. 4 transitions of contempt and 2 transitions of disgust are misinterpreted as anger. 2 anger transitions are confused with disgust and happiness. 1 surprise transition is misclassified as contempt. The signature reports 97.24% overall accuracy on the CK+ dataset.

Other performance evaluation parameters: FAR, FRR and ERR are computed from CK+ dataset for proposed all different signatures figuring in Table 2. Figure 7 demonstrates FAR, FRR and ERR for CK+ dataset by showing graphs. After doing the discussion on performance, it is observed that transitions of happiness and surprise are recognized with 100% accuracy by all five different signatures which show that the signatures are able to find out the common pattern of changing behavior over the transitions of happiness and surprise emotions. On the other side, the fear transition is found more difficult to detect perfectly. Among five signatures, only two signatures (Fuzzy Isosceles Triangle Signature and Other Fuzzy Triangles Signature) are able to detect more information about triangle shape deformation in fear emotion than other signatures by showing 100% accuracy. Fuzzy Right Triangle Signature and Fuzzy Isosceles and Right Triangle Signature are showing higher performance than the other four signatures by achieving more than 98% overall accuracy on the CK+ dataset.

Table 2 FAR, FRR and ERR on CK+ dataset

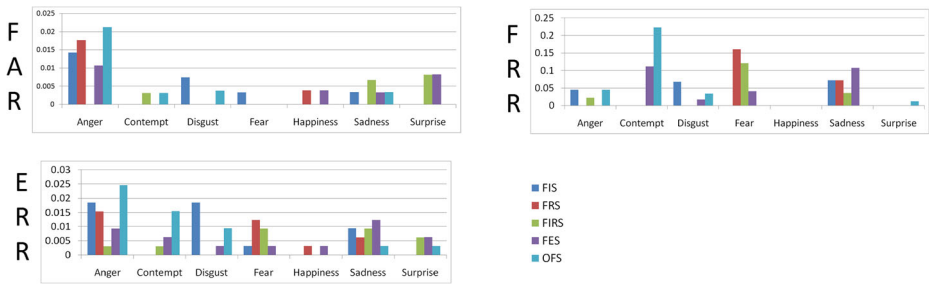|	FAR	FRR	ERR
Fuzzy Isosceles Triangle Signature			
AN	0.01428	0.0444	0.0184
CON	0	0	0
DI	0.0075	0.0677	0.0184
FE	0.0033	0	0.0031
HA	0	0	0
SA	0.0034	0.0714	0.0093
SU	0	0	0
Fuzzy Right Triangle Signature			
AN	0.0177	0	0.0153
CON	0	0	0
DI	0	0	0
FE	0	0.16	0.0123
HA	0.0039	0	0.0031
SA	0	0.0714	0.0061
SU	0	0	0
Fuzzy Isosceles and Right Triangle Signature			
AN	0	0.0222	0.003
CON	0.0032	0	0.003
DI	0	0	0
FE	0	0.12	0.0092
HA	0	0	0
SA	0.0067	0.0357	0.0092
SU	0.0082	0	0.0061
Fuzzy Equilateral Triangle Signature			
AN	0.0107	0	0.0092
CON	0	0.1111	0.0062
DI	0	0.0169	0.0031
FE	0	0.04	0.0031
HA	0.0039	0	0.0031
SA	0.0033	0.1071	0.0123
SU	0.0083	0	0.0062
Other Fuzzy Triangles Signature			
AN	0.0213	0.0444	0.0245
CON	0.0032	0.2222	0.0154
DI	0.0038	0.0338	0.0093
FE	0	0	0
HA	0	0	0
SA	0.0034	0	0.0031
SU	0	0.012	0.0031

Fig. 7 FAR, FRR, ERR on CK+ dataset

4.2 Discussion on M&M initiative (MMI) dataset

MMI has 236 expression profiles of emotion which are captured in video clips taking from different people of ages from 19 to 62 years. Most of the people belong to the female category from different cultures (European, Asian, or South American). Each expression profile contains the sequence of both frontal and side view face images. It is a very challenging task to map landmarks properly on such a dataset. We have collected transitions of frontal face expression in which facial expressions are found evolving from neutral expression to peak expression and turn back to a neutral expression. It is found that only 202 profiles of six expressions (anger, disgust, fear, happiness, sadness, and surprise) are labeled perfectly. The transition profile of happiness emotion is shown in the second row of Fig. 5.

4.2.1 Result analysis of MMI dataset

We demonstrated the recognition ability of five different signatures on the MMI dataset individually. The dataset consists of the following number of six expressions: 31 (AN), 32 (DI), 28 (FE), 42 (HA), 28 (SA), and 41 (SU). Different signature induced confusion matrices on the MMI dataset are presented in Table 3. The corresponding analysis graphs of confusion matrices for the MMI dataset are displayed in Fig. 8.

Fuzzy Isosceles Triangle Signature (FIS) This signature identifies the transition of surprise emotion without misclassification. 28 transitions of anger are perfectly identified but 2 transitions are found confused with disgust and 1 with fear. 31 transitions of disgust are properly recognized but only 1 transition gets misclassified with anger. Fear classifies 21 transitions perfectly while 4 transitions are misclassified with surprise and 3 transitions with anger. 27 transitions of sadness and 41 transitions of happiness are classified correctly but 1 transition of sadness and 1 transition of happiness are misinterpreted as anger emotion. The signature is able to achieve 93.56% overall accuracy on the MMI dataset.

Fuzzy Right Triangle Signature (FRS) All transitions of disgust emotion are recognized properly without any error. 30 anger transitions are identified but 1 transition is confused with sadness. 21 transitions of fear are correctly classified while 5 transitions are misclassified as anger and 2 transitions as a surprise. 41 happiness, 27 sadness, and 37 surprises are correctly recognized by the signature. 1 transition of happiness, 1 transition of sadness get confused with fear and anger respectively. 2 transitions of surprise are misclassified with

Table 3 Confusion Matrices for different signatures on MMI Dataset

	AN	DI	FE	HA	SA	SU
Fuzzy Isosceles Triangle Signature						
AN	28	2	1	0	0	0
DI	1	31	0	0	0	0
FE	3	0	21	0	0	4
HA	1	0	0	41	0	0
SA	1	0	0	0	27	0
SU	0	0	0	0	0	41
Fuzzy Right Triangle Signature						
AN	30	0	0	0	1	0
DI	0	32	0	0	0	0
FE	5	0	21	0	0	2
HA	0	0	1	41	0	0
SA	1	0	0	0	27	0
SU	1	1	2	0	0	37
Fuzzy Isosceles and Right Triangle Signature						
AN	30	0	1	0	0	0
DI	0	30	1	1	0	0
FE	1	0	24	0	0	3
HA	1	0	0	41	0	0
SA	2	0	0	0	25	1
SU	0	0	3	0	0	38
Fuzzy Equilateral Triangle Signature						
AN	30	1	0	0	0	0
DI	0	32	0	0	0	0
FE	4	0	21	0	1	2
HA	0	0	0	42	0	0
SA	2	0	1	0	25	0
SU	1	0	1	0	0	39
Other Fuzzy Triangles Signature						
AN	29	1	0	0	1	0
DI	1	31	0	0	0	0
FE	0	0	26	1	0	1
HA	1	0	0	40	1	0
SA	1	2	2	0	23	0
SU	2	0	0	1	0	38

fear, 1 transition of surprise is incorrectly identified as anger and 1 transition is misinterpreted as disgust. Overall accuracy 93.06% is found on the MMI dataset by using Fuzzy Right Triangle Signature.
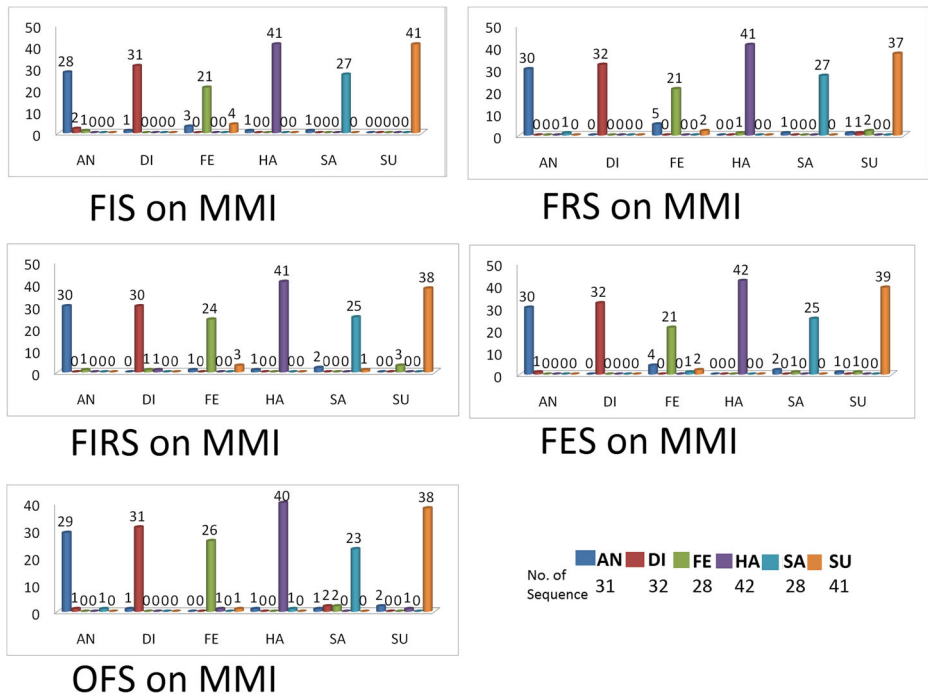
Fig. 8 Visualization graphs of MMI confusion matrix

Fuzzy Isosceles and Right Triangle Signature (FIRS) The signature classifies perfectly 30 transitions as anger, 30 transitions as disgust, 24 transitions as fear, 41 transitions as happiness, 25 transitions as sadness, and 38 transitions as a surprise. 1 fear, 1 happiness, 2 sadness are confused with anger emotion. 1 transition of anger, 1 transition of disgust, 3 transitions of surprise is misclassified as fear emotion. 1 transition of disgust is incorrectly identified as happiness. 3 transitions of fear and 1 transition of sadness are misinterpreted as surprise emotion. The signature achieves a 93.06% overall recognition rate on the MMI dataset.

Fuzzy Equilateral Triangle Signature (FES) Transitions of disgust and happiness are correctly identified by this signature without any error. Here, signature recognizes the following number of emotions correctly: AN (30), DI (32), FE (21), HA (42), SA (25), and SU (39). 4 transitions of fear, 2 transitions of sadness and 1 transition of surprise are confused with anger. 1 anger is misclassified as disgust. 1 sadness and 1 surprise are wrongly identified as fear. 1 transition of fear is misinterpreted as sadness while 2 transitions of fear as a surprise. 93.56% overall recognition rate is reported by this signature.

Other Fuzzy Triangles Signature (OFS) The signature recognizes the following number of emotions: AN (29), DI (31), FE (26), HA (40), SA (23), and SU (38). 1 transition of disgust, 1 transition of happiness, 1 transition of sadness, and 2 transitions of surprise mismatch with anger. 1 transition of anger and 2 transitions of sadness are confused with disgust. 2 transitions of sadness are misclassified as fear. 1 transition of fear, 1 transition of surprise are misclassified as happiness. 1 anger and 1 happiness are incorrectly identified as sadness.1

fear transition emotion is wrongly interpreted as a surprise. This fuzzy signature reports 92.57% overall recognition accuracy on the MMI dataset.

FAR, FRR, and ERR for all fuzzy signatures are described in Table 4. Figure 9 demonstrates FAR, FRR and ERR for MMI dataset by showing graphs. The dataset is found pretty challenging to recognize all basic emotions. Out of five fuzzy signatures, Fuzzy Isosceles Triangle Signature (FIS) and Fuzzy Equilateral Triangle Signature (FES) demonstrate

Table 4 FAR, FRR and ERR on MMI dataset

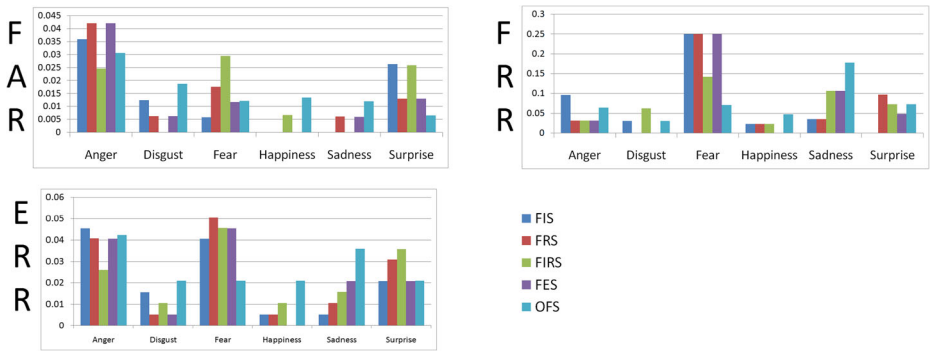|	FAR	FRR	ERR
Fuzzy Isosceles Triangle Signature			
AN	0.0359	0.0967	0.0454
DI	0.0125	0.0312	0.0156
FE	0.0059	0.25	0.0406
HA	0	0.0238	0.0052
SA	0	0.0357	0.0052
SU	0.0263	0	0.0207
Fuzzy Right Triangle Signature			
AN	0.0421	0.0322	0.0408
DI	0.0063	0	0.0052
FE	0.0176	0.25	0.0505
HA	0	0.0238	0.0052
SA	0.0061	0.0357	0.0105
SU	0.013	0.0975	0.0309
Fuzzy Isosceles and Right Triangle Signature			
AN	0.0246	0.0322	0.0259
DI	0	0.0625	0.0105
FE	0.0295	0.1428	0.0456
HA	0.0067	0.0238	0.0105
SA	0	0.1071	0.0157
SU	0.0259	0.0731	0.0358
Fuzzy Equilateral Triangle Signature			
AN	0.0421	0.0322	0.0406
DI	0.0063	0	0.0052
FE	0.0117	0.25	0.0454
HA	0	0	0
SA	0.006	0.1071	0.0207
SU	0.0131	0.0487	0.0207
Other Fuzzy Triangles Signature			
AN	0.0306	0.0645	0.0424
DI	0.0188	0.0312	0.0209
FE	0.0122	0.0714	0.0209
HA	0.0134	0.0476	0.0209
SA	0.012	0.1785	0.036
SU	0.0066	0.0731	0.0209

Fig. 9 FAR, FRR, ERR on MMI dataset

impressive performance on the MMI dataset compared to other fuzzy signatures by attaining 93.56% overall accuracy.

4.3 Discussion on multimedia understanding group (MUG) dataset

The dataset includes various expressional profiles of face, recorded from 86 people. Male (51) and female (35) both participate to perform their facial expression. All of them are of Caucasian origin and their age varies from 20 to 35 years. Each profile is taken with a frame rate of 19 f/s. Each frame is stored in a JPG image format with a minimum size of 240 KB and a maximum size of 340 KB. A total of 801 expression profiles are found with six emotions: anger, disgust, fear, happiness, sadness, and surprise. The third row of Fig. 5 describes the expression profile for happiness emotion taken from the MUG dataset. Here expression evolves from neutral face to the face with universal expression of happiness.

4.3.1 Result analysis on MUG dataset

We also use the MUG dataset for measuring the recognition capability of our proposed five fuzzy signatures separately. Each signature uses the following number of MUG emotions: AN (149), DI (117), FE (150), HA (107), SA (133), and SU (145). We computed confusion matrices on this dataset corresponding to every signature and these are displayed together in Table 5. The corresponding analysis graphs of confusion matrices for the MUG dataset are displayed in Fig. 10.

Fuzzy Isosceles Triangle Signature (FIS) From Table 5, it is found that the signature identifies the transition of disgust and fear with 100% accuracy. 145 anger transitions, 106 happiness transitions, 131 sadness transitions, and 142 surprise transitions are identified correctly. 3 transitions of anger, 1 transition of sadness, and 1 transition of surprise are found confused with the emotion happiness. 1 transition of sadness, 1 transition of happiness, 1 transition of surprise, and 1 transition of anger are misclassified as anger, disgust, disgust, and surprise respectively. The overall recognition rate of 98.87% is achieved by this signature on the MUG dataset.

Fuzzy Right Triangle Signature (FRS) 100% accuracy is achieved by this signature on the MUG dataset for the emotions: disgust, fear, sadness, and surprise. Anger is identified with

Table 5 Confusion Matrices for different signatures on MUG Dataset

	AN	DI	FE	HA	SA	SU
Fuzzy Isosceles Triangle Signature						
AN	145	0	0	3	0	1
DI	0	117	0	0	0	0
FE	0	0	150	0	0	0
HA	0	1	0	106	0	0
SA	1	0	0	1	131	0
SU	0	1	0	1	0	143
Fuzzy Right Triangle Signature						
AN	146	0	0	3	0	0
DI	0	117	0	0	0	0
FE	0	0	150	0	0	0
HA	3	0	0	104	0	0
SA	0	0	0	0	133	0
SU	0	0	0	0	0	145
Fuzzy Isosceles and Right Triangle Signature						
AN	149	0	0	0	0	0
DI	0	116	0	0	0	1
FE	1	0	148	0	0	1
HA	3	0	0	104	0	0
SA	0	0	0	1	132	0
SU	0	0	0	0	0	145
Fuzzy Equilateral Triangle Signature						
AN	145	0	0	3	0	1
DI	0	114	0	0	0	3
FE	0	0	150	0	0	0
HA	1	0	0	106	0	0
SA	0	0	0	0	132	1
SU	0	2	0	0	0	143
Other Fuzzy Triangles Signature						
AN	147	0	1	1	0	0
DI	0	116	0	0	0	1
FE	0	0	150	0	0	0
HA	2	0	0	105	0	0
SA	0	0	0	0	133	0
SU	1	1	0	0	2	141

the number of 146 while only 3 transitions of anger are wrongly identified as happiness. On the other hand, 104 transitions of happiness are correctly identified whereas 3 transitions of happiness are wrongly classified as anger. This signature reports a 99.25% overall recognition rate on the MUG dataset.
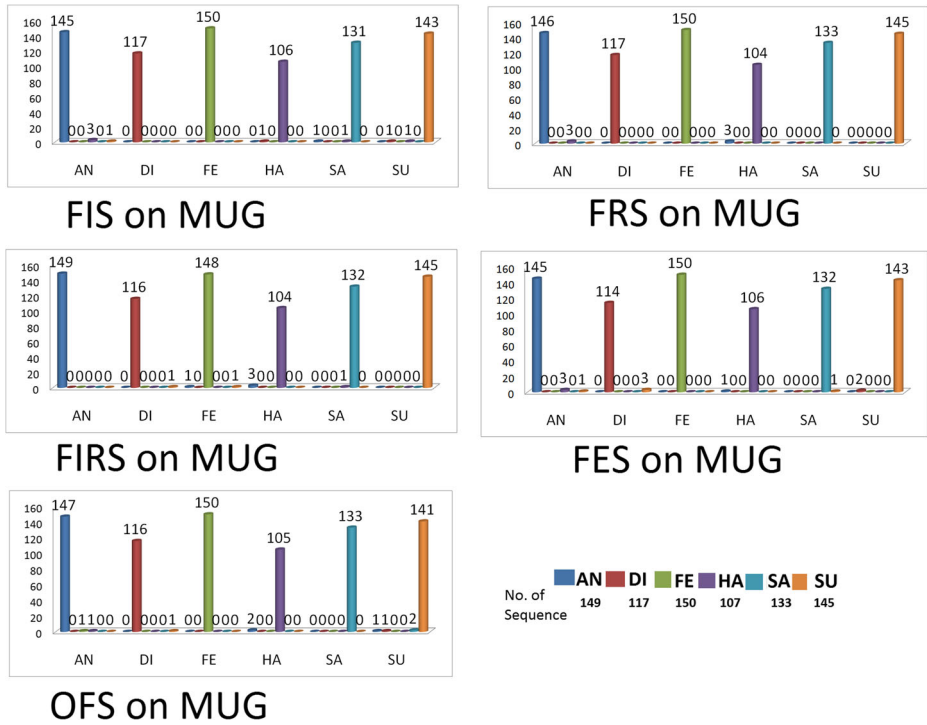
Fig. 10 Visualization graphs of MUG confusion matrix

Fuzzy Isosceles and Right Triangle Signature (FIRS) Here, anger and surprise are found attaining a 100% recognition rate by the signature. Out of 117 transitions of disgust emotion, only 1 transition gets confused with surprise while others are perfectly classified. 1 transition of fear is misclassified as anger, 1 transition of fear as surprise whereas 148 transitions of fear are identified perfectly. Out of 107 transitions of happiness, 3 mismatches with anger, and the remaining are classified correctly. 132 transitions of sadness are recognized accurately and only 1 transition of sadness is found misclassified with happiness. 99.12% overall recognition rate is obtained by Fuzzy Isosceles and Right Triangle Signature on the MUG dataset.

Fuzzy Equilateral Triangle Signature (FES) The signature is able to get 100% accuracy on fear emotion only. Out of 149 transitions of anger, 145 are appropriately-recognized, 3 are wrongly identified with happiness and 1 is found confused with surprise. 3 transitions of disgust are misclassified as a surprise while 114 transitions of disgust are correctly recognized. Out of 107 transitions of happiness, only 1 transition is incorrectly identified with anger and 106 are correctly classified. Similarly, only 1 transition of sadness is misinterpreted as surprise whereas 132 transitions of sadness are correctly classified. 143 transitions of surprise are recognized perfectly and only 2 transitions of surprise are misinterpreted as disgust emotion. The signature reports a 98.62% overall recognition rate on the MUG dataset.

Other Fuzzy Triangles Signature (OFS) Fear and sadness are classified with a 100% recognition rate by this signature. 1 transition of anger is confused as fear, 1 transition of anger

Table 6 FAR, FRR and ERR on MUG dataset

	FAR	FRR	ERR
Fuzzy Isosceles Triangle Signature			
AN	0.0015	0.0268	0.0062
DI	0.0029	0	0.0025
FE	0	0	0
HA	0.0072	0.0093	0.0075
SA	0	0.015	0.0025
SU	0.0015	0.0137	0.0037
Fuzzy Right Triangle Signature			
AN	0.0046	0.0201	0.0074
DI	0	0	0
FE	0	0	0
HA	0.0043	0.028	0.0074
SA	0	0	0
SU	0	0	0
Fuzzy Isosceles and Right Triangle Signature			
AN	0.0061	0	0.005
DI	0	0.0085	0.0012
FE	0	0.0133	0.0025
H	0.0014	0.028	0.005
SA	0	0.0075	0.0012
SU	0.003	0	0.0025
Fuzzy Equilateral Triangle Signature			
AN	0.0015	0.0268	0.0076
DI	0.0029	0.0256	0.0062
FE	0	0	0
HA	0.0043	0.0093	0.005
SA	0	0.0075	0.0012
SU	0.0076	0.0137	0.0087
Other Fuzzy Triangles Signature			
AN	0.0046	0.0134	0.0062
DI	0.0014	0.0085	0.0025
FE	0.0015	0	0.0012
HA	0.0014	0.0186	0.0037
SA	0.003	0	0.0025
SU	0.0015	0.0275	0.0062

mismatch with happiness, and 147 transitions of anger are properly recognized. Out of 117 transitions of disgust, only 1 is wrongly recognized as surprise whereas others are properly identified. 105 transitions of happiness, 141 transitions of surprise are correctly recognized. Only 2 transitions of happiness are misclassified as anger. 2 transitions of surprise are mis-interpreted as sadness, 1 transition of surprise mismatch with anger, and 1 transition of surprise as disgust. Here, 98.87% overall recognition rate is found on the MUG dataset.
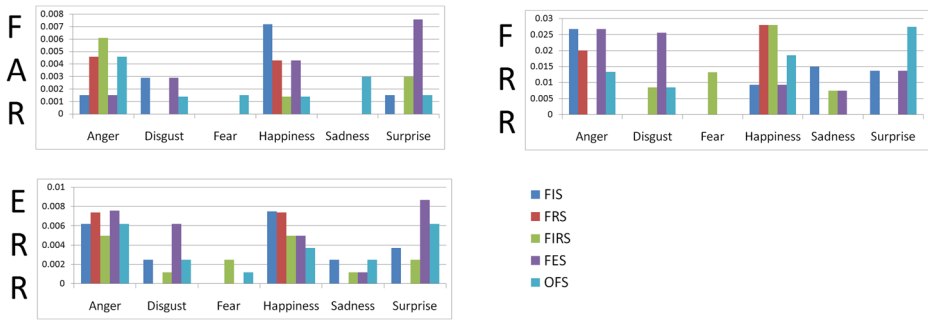
Fig. 11 FAR, FRR, ERR on MUG dataset

We consider other performance estimation parameters: FAR, FRR, and ERR for all five signatures on the MUG dataset, and these are presented in Table 6. Figure 11 demonstrates FAR, FRR and ERR for MUG dataset by showing graphs. By implementing our proposed approach on the MUG dataset, it is observed that the maximum shape variation information is collected from the transition of fear emotion through the application of all proposed signatures individually. On the other hand, the transition of happiness is found more difficult compared to other emotions available in the MUG dataset reporting 100% accuracy by our proposed five signatures. It is also seen that our all proposed signatures show superior performance on the MUG dataset and among them, two signatures (Fuzzy Right Triangle Signature, Fuzzy Isosceles, and Right Triangle Signature) ensure more impressive performance compared to others by confirming overall accuracy above 99%.

4.4 Performance comparison among five different signatures

For better adjudication of recognition performance of all different kinds of fuzzy signatures, we implemented 10 fold cross-validation techniques on three different datasets. Here, the dataset is split into 10 different folds. The validation part uses One fold and the training part reserves nine folds. Next, it calculates validation accuracy for each of 10 sets and takes the average of them. Table 7 displays accuracies of 10 fold cross-validation for all kind of fuzzy signatures in three different datasets. The performance comparison of five different kinds of fuzzy geometric signatures by using three benchmark datasets is demonstrated in Fig. 12. The overall recognition accuracy by using FIS, FRS, FIRS, FES and OFS signatures in CK+ dataset is 97.55%, 98.16%,98.47%, 97.85%, and 97.24%, in MMI dataset is 93.56%, 93.06%, 93.06%, 93.56%, and 92.57%, in MUG dataset is 98.87%, 99.25%, 99.12%, 98.62%, and 98.87% respectively. The performance of all five signatures on CK+ and MUG datasets is reportedly better compared to the MMI dataset. MMI shows lower performance because of its intrinsic complexity in data characteristics leading to likely improper mapping of the landmark points on MMI face images. Moreover, it is

	FIS	FRS	FIRS	FES	OFS
CK+	97.04%	96.94%	96.89%	97.18%	95.64%
MMI	89.38%	89.07%	89.78%	89.83%	89.22%
MUG	97.79%	97.88%	97.74%	97.68%	97.92%

Table 7 10 fold cross-validation accuracy for five different fuzy signatures on CK+, MMI and MUG Datasets
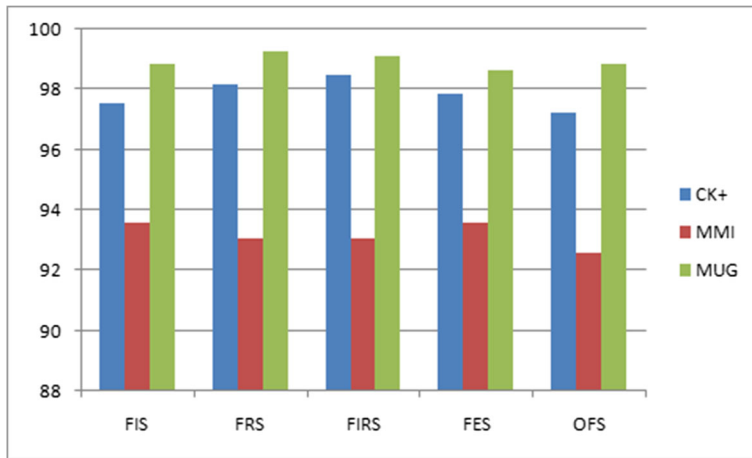
Fig. 12 Performance comparison for five fuzzy signatures on CK+, MMI, and MUG Datasets

also observed that the Fuzzy Right Triangle Signature (FRS) and Fuzzy Isosceles and Right Triangle Signature (FIRS) offer the best results on CK+ and MUG datasets.

5 Results comparison with existing state-of-the-art works

We endeavor to provide a comparison task for validating our system performance with other methods by using three datasets: CK+, MMI, and MUG. Details are described in subsequent sections.

5.1 Comparison on CK+ dataset

We compare recognition results obtained from five different fuzzy geometric signatures being on CK+ dataset with the results reported in literatures [13, 18, 23, 29, 39, 44]. Table 8 describes the comparison results of MLP recognizer. In [13], authors achieved an overall accuracy of 97.5% which is the highest rate compared to other literature methods described in Table 8. Out of the proposed five fuzzy signatures, three signatures (Fuzzy Right Triangle Signature, Fuzzy Isosceles, and Right Triangle Signature, Fuzzy Equilateral Triangle Signature) are found to perform better than that in article [13] by attaining overall accuracy 98.16%, 98.47%, and 97.85% respectively. Though the method in [13], could find an accuracy of 96.67% on sadness emotion which is higher than these three signatures, our Other Fuzzy Triangle Signature (OFS) is able to recognize sadness emotion with 100% accuracy. On the other hand, all our proposed signatures are able to obtain higher overall accuracy as well as a higher rate in individual emotion recognition than the approaches reported in [29] (83.01%), [18] (87.43%), [23] (83.9%). In [39, 44] authors could find higher accuracy 88% and 87% on fear emotion than reported by our Fuzzy Right Triangle Signature (84%). However, this signature is found to outperform those literature methods by ensuring a higher overall accuracy of 98.16% on the CK+ dataset.

Table 8 Comparison of results on CK+, MMI, MUG with other state-of-the-art works

Method	Database	AN	CON	DI	FE	HA	SA	SU	Avg. Acc.
FIS	CK+	95.55	100	93.22	100	100	92.85	100	**97.55**
FRS	CK+	100	100	100	84	100	92.85	100	**98.16**
FIRS	CK+	97.77	100	100	88	100	96.42	100	**98.47**
FES	CK+	100	88	98.3	96	100	89.28	100	**97.85**
OFS	CK+	95.55	77.77	96.61	100	100	100	98.79	**97.24**
[29]	CK+	80	NA	83	72	100	64.2	98.7	**83.01**
[44]	CK+	93.3	NA	100	88	100	85.7	96.4	**95.8**
[18]	CK+	66.67	NA	91.02	80	97.08	80	96.88	**87.43**
[13]	CK+	97.5	NA	96.67	96	100	96.67	100	**97.8**
[39]	CK+	91	NA	97	87	99	90	97	**90**
[23]	CK+	81	44	90	58	97	50	98	**83.9**
FIS	MMI	90.32	NA	96.87	75	97.61	96.42	100	**93.56**
FRS	MMI	96.77	NA	100	75	97.61	96.42	90.24	**93.06**
FIRS	MMI	96.77	NA	93.75	85.71	97.61	89.28	92.68	**93.06**
FES	MMI	96.77	NA	100	75	100	89.28	95.12	**93.56**
OFS	MMI	93.54	NA	96.87	92.85	95.23	82.14	92.68	**92.57**
[13]	MMI	70	NA	80	70	85	73.33	85	**77.22**
[42]	MMI	50	NA	66.67	75	66.67	100	100	**71.43**
[20]	MMI	98.3	NA	88.7	90.2	90.5	88.7	98.1	**93.53**
FIS	MUG	97.31	NA	100	100	99.06	98.49	98.62	**98.87**
FRS	MUG	97.98	NA	100	100	97.19	100	100	**99.25**
FIRS	MUG	100	NA	99.14	98.66	97.19	99.24	100	**99.12**
FES	MUG	97.31	NA	97.43	100	99.06	99.24	98.62	**98.62**
OFS	MUG	98.65	NA	99.14	100	98.13	100	97.24	**98.87**
[13]	MUG	100	NA	100	85	100	90	98	**95.5**
[28]	MUG	96.3	NA	93.83	88.89	95.06	80	78	**95.24**

5.2 Comparison on MMI dataset

We illustrate results on the MMI dataset to demonstrate the comparison task with other works. The comparative results of our proposed five fuzzy signatures are summarized in Table 8. Results reported in [13, 20, 42] are compared separately with the results obtained from five signatures on MMI dataset. In [20], the method could find overall accuracy of 93.53% on the MMI dataset which is the highest rate compared to other existing methods described in the Table 8. Even though authors in [20] are able to find higher accuracy on recognition of anger (98.3%) and fear (90.2%) transition compared to our proposed two fuzzy signatures (Fuzzy Isosceles Triangle Signature and Fuzzy Equilateral Triangle Signature) but our these two signatures could outperform the method [20] by achieving higher overall recognition rate 93.56% each. It is observed that our proposed signatures show performance domination over [13] by providing higher accuracy in individual emotion recognition as well as overall recognition. The study in [42] found 100% accuracy on

sadness and surprise emotions but average accuracy (71.43%) is found on the MMI dataset which is lower than our five signatures as a whole.

5.3 Comparison on MUG dataset

We also exhibited the comparison task in this section by considering the results obtained from the MUG dataset by using five fuzzy geometric signatures. Here, different recognition results computed by fuzzy signatures are compared with the results available in literature [13, 28]. From Table 8 it is observed that authors in the article [13], have reported an overall recognition rate of 95.5% on the MUG dataset with finding 100% accuracy on anger, disgust, and happiness emotions. Though the article [13] shows higher accuracy on anger and happiness compared to all our proposed signatures in view of overall performance evaluation, however, our proposed fuzzy signatures outperform it by ensuring an overall recognition rate of more than 95.5% each. It is also seen from Table 8 that all kinds of results on MUG reported in [28] report overall lower performance than our proposed signatures.

6 Conclusion

In this article, we exhibit a comprehensive study for understanding the changing behavior of human emotion. It is very important to know about the unsteady motions of various major face portions (eyes, eyebrows, nose, lips) while an expression evolves from a neutral face to a universal face. That is why our newly developed triangulation technique is involved in this study to form smaller triangle shapes on the face plane by taking a triplet of landmark points. Sometimes triangle shape does not match with a basic geometrical shape. Thus the shape is not exclusively sufficient to capture quantitative information for understanding the deficiency associated with the shape. We incorporate fuzzy membership features into the triangle shape to generate five fuzzy geometric signatures that describe the transition of basic emotion in five different ways. The dearth in transitional shape is well identified by utilizing our proposed fuzzy signatures into MLP classifier and superiority in classification results on CK+, MMI, and MUG datasets assert evidence for the good identification of this vagueness. The effectiveness of five fuzzy signatures is individually reviewed through the comparison of the results with state-of-the-art works as well as by performing 10 fold cross-validation on benchmark face sequence datasets CK+, MMI, and MUG. From our study, it is observed that the highest influence in emotional transition recognition on CK+, MMI, and MUG is shown by Fuzzy Isosceles and Right Triangle Signature (98.47%), Fuzzy Equilateral Triangle Signature (93.56%), and Fuzzy Right Triangle Signature (99.25%) respectively. On the other hand, We found a crucial benefit from our work that we used only 23 landmark points among 68 landmarks identified as a prominent feature in the context of highly deformable face portions. Thus it ensures to utilize less number of fuzzy triangles having maximum variations in shape over the frame in emotional transition that reduce the computational overhead of the system.

References

1. Ahn B, Han Y, Kweon IS (2012) Real-time facial landmarks tracking using active shape model and lk optical flow. In: 2012 9th international conference on ubiquitous robots and ambient intelligence (URAI). IEEE, pp 541–543
2. Aifanti N, Papachristou C, Delopoulos A (2010) The mug facial expression database. In: 11Th international workshop on image analysis for multimedia interactive services WIAMIS 10. IEEE, pp 1–4
3. Barman A (2020) Human Emotion Recognition from Face Images. Springer, Berlin
4. Barman A, Dutta P (2017) Facial expression recognition using distance and shape signature features. Pattern Recognit Lett
5. Barman A, Dutta P (2019) Facial expression recognition using distance and texture signature relevant features. Appl Soft Comput 77:88–105
6. Barman A, Dutta P (2019) Influence of shape and texture features on facial expression recognition. IET Image Process 13(8):1349–1363
7. Bastanfard A, Bastanfard O, Takahashi H, Nakajima M (2004) Toward anthropometrics simulation of face rejuvenation and skin cosmetic. Comput Animat Virt Worlds 15(3-4):347–352
8. Bastanfard A, Takahashi H, Nakajima M (2004) Toward e-appearance of human face and hair by age, expression and rejuvenation. In: 2004 International conference on cyberworlds. IEEE, pp 306–311
9. Choi H-C, Oh S-Y (2006) Realtime facial expression recognition using active appearance model and multilayer perceptron. In: 2006 SICE-ICASE International joint conference. IEEE, pp 5924–5927
10. Cootes TF, Taylor JC, Cooper DH, Graham J (1995) Active shape models-their training and application. Comput Vision Image Understand 61(1):38–59
11. Ekman P, Friesen WV (2003) Unmasking the face: A guide to recognizing emotions from facial clues. Ishk
12. Ghimire D, Lee J (2014) Extreme learning machine ensemble using bagging for facial expression recognition. JIPS 10(3):443–458
13. Ghimire D, Lee J, Li Z-N, Jeong S (2017) Recognition of facial expressions based on salient geometric features and support vector machines. Multimed Tools Appl 76(6):7921–7946
14. Gross R, Matthews I, Baker S (2004) Appearance-based face recognition and light-fields. IEEE Trans Pattern Anal Mach Intell 26(4):449–465
15. Happy SL, Routray A (2015) Robust facial expression classification using shape and appearance features. In: 2015 Eighth international conference on advances in pattern recognition (ICAPR). IEEE, pp 1–5
16. Kotsia I, Pitas I (2006) Facial expression recognition in image sequences using geometric deformation features and support vector machines. IEEE Trans Image Process 16(1):172–187
17. Kumari J, Rajesh R, Pooja KM (2015) Facial expression recognition: A survey. Procedia Comput Sci 58(1):486–491
18. Li Y, Wang S, Zhao Y, Ji Q (2013) Simultaneous facial feature tracking and facial expression recognition. IEEE Trans Image Process 22(7):2559–2573
19. Lucey P, Cohn JF, Kanade T, Saragih J, Ambadar Z, Matthews I (2010) The extended cohn-kanade dataset (ck+): a complete dataset for action unit and emotion-specified expression. In: 2010 Ieee computer society conference on computer vision and pattern recognition-workshops. IEEE, pp 94–101
20. Majumder A, Behera L, Subramanian VK (2014) Emotion recognition from geometric facial features using self-organizing map. Pattern Recogn 47(3):1282–1293
21. Meftah IT, Thanh NL, Amar CB (2012) Emotion recognition using knn classification for user modeling and sharing of affect states. In: International conference on neural information processing. Springer, pp 234–242
22. Mehrabian A, Russell JA (1974) An approach to environmental psychology. MIT Press, Cambridge
23. Mohammadian A, Aghaeinia H, Towhidkhah F (2015) Video-based facial expression recognition by removing the style variations. IET Image Process 9(7):596–603
24. Ojala T, Pietikäinen M., Harwood D (1996) A comparative study of texture measures with classification based on featured distributions. Pattern Recognit 29(1):51–59
25. Pantic M, Pentland A, Nijholt A, Huang TS (2007) Machine computing and understanding of human behavior Human A survey. In: Artifical intelligence for human computing. Springer, pp 47–71
26. Perikos I, Ziakopoulos E, Hatzilygeroudis I (2014) Recognizing emotions from facial expressions using neural network. In: IFIP International conference on artificial intelligence applications and innovations. Springer, pp 236–245

27. Rahul M, Kohli N, Agarwal R, Mishra S (2019) Facial expression recognition using geometric features and modified hidden markov model. Int J Grid Util Comput 10(5):488–496
28. Rahulamathavan Y, Phan RC-W, Chambers JA, Parish DJ (2012) Facial expression recognition in the encrypted domain based on local fisher discriminant analysis. IEEE Trans Affect Comput 4(1):83–92
29. Saeed A, Al-Hamadi A, Niese R, Elzobi M (2014) Frame-based facial expression recognition using geometrical features. Adv Human-Comput Interact 2014
30. Samadiani N, Huang G, Cai B, Luo W, Chi C-H, Xiang Y, He J (2019) A review on automatic facial expression recognition systems assisted by multimodal sensor data. Sensors 19(8):1863
31. Shan C, Gong S, McOwan PW (2009) Facial expression recognition based on local binary patterns: A comprehensive study. Image Vision Comput 27(6):803–816
32. Sharma G, Singh L, Gautam S (2019) Automatic facial expression recognition using combined geometric features. 3D Res 10(2):14
33. Terzopoulos D, Waters K (1993) Analysis and synthesis of facial image sequences using physical and anatomical models. IEEE Trans Pattern Anal Mach Intell 15(6):569–579
34. Tian Y-L, Kanade T, Cohn JF (2005) Facial expression analysis. In: Handbook of face recognition. Springer, pp 247–275
35. Tzimiropoulos G, Pantic M (2013) Optimization problems for fast aam fitting in-the-wild. In: Proceedings of the IEEE international conference on computer vision, pp 593–600
36. Valstar M, Pantic M (2010) Induced disgust, happiness and surprise: an addition to the mmi facial expression database. In: Proc. 3rd intern. Workshop on EMOTION (satellite of LREC): Corpora for research on emotion and affect, Paris, France, p 65
37. Vishnu Priya R (2019) Emotion recognition from geometric fuzzy membership functions. Multimed Tools Appl 78(13):17847–17878
38. Wolf L (2009) Face recognition, geometric vs. appearance-based. Encycloped Biomet 2
39. Yaddaden Y, Adda M, Bouzouane A, Gaboury S, Bouchard B (2017) Facial expression recognition from video using geometric features
40. Yeasin M, Bullot B, Sharma R (2006) Recognition of facial expressions and measurement of levels of interest from video. IEEE Trans Multimed 8(3):500–508
41. Zangeneh E, Moradi A (2018) Facial expression recognition by using differential geometric features. The Imaging Sci J 66(8):463–470
42. Zhang S, Pan X, Cui Y, Zhao X, Liu L (2019) Learning affective video features for facial expression recognition via hybrid deep learning. IEEE Access 7:32297–32304
43. Zhang S, Zhao X, Lei B (2012) Robust facial expression recognition via compressive sensing. Sensors 12(3):3747–3761
44. Zhao L, Wang Z, Zhang G (2017) Facial expression recognition from video sequences based on spatial-temporal motion local binary pattern and gabor multiorientation fusion histogram. Math Probl Eng 2017

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