Fuzzy emotion: a natural approach to automatic facial expression recognition from psychological perspective using fuzzy system

by Dewi Yanti Liliana

Cognitive Processing
https://doi.org/10.1007/s10339-019-00923-0

RESEARCH ARTICLE

Fuzzy emotion: a natural approach to automatic facial expression recognition from psychological perspective using fuzzy system

Dewi Yanti Liliana[1] · T. Basaruddin[1] · M. Rahmat Widyanto[1] · Imelda Ika Dian Oriza[2]

Received: 16 November 2018 / Accepted: 7 June 2019

Abstract
Many studies in automatic facial expression recognitions merely limit their focus on recognizing basic emotions, ignoring the fact that humans show various emotions in their daily life. Moreover, from psychological perspective humans express multiple emotions simultaneously. Up to now, researchers recognize two basic emotions at the same time, called mixed emotions. Nevertheless, the mixed emotion still does not reflect how humans express the emotion naturally. This paper advances the concept of mixed emotion into a generalized fuzzy emotion. Fuzzy emotion captures multiple emotions in a single image using fuzzy inference engine. We propose a fuzzy emotion framework which consists of processing system and knowledge system. The processing system extracts facial expression parameters, and the knowledge system employs a fuzzy knowledge-based engine, elicited from the psychologist knowledge to recognize facial expressions. Some advantages are offered: (1) no facial template comparison; (2) no training efforts needed; (3) moreover, fuzzy emotion can recognize ambiguous facial expressions adaptively. The experiment gives a recognition result with the highest accuracy rate of 0.90. A research agenda for future study of mixed emotion recognition is proposed.

Keywords Artificial intelligence · Affective computing · Emotion recognition · Facial expression · Fuzzy emotion · Fuzzy system

Introduction

To date, there are two main problems in automatic emotion recognition. The first is related to the categorization of emotion, and the second is related to the variety of facial expressions. In emotion theory, the classification of human emotion into fundamental emotion categories was carried out by psychologist Paul Ekman (1975) based on his long-term research on emotions across cultures (Ekman and Friesen 1975). He defined that there were six categories of human basic emotions—happiness, anger, sadness, disgust, fear, and surprise. This statement was then used as a main reference for many studies, including research in computer vision and machine learning which aimed to classify basic emotion classes from facial expression images. Various annotated facial expression datasets for basic emotion recognition were also developed, such as CK+, JAFFE, DISFA, and MUG. However, those studies neglected the combinatory nature of emotion which existed in human real life that formed a mixture class of basic emotion (or the so-called mixed emotion) by limiting the recognition into basic emotion classes. This explains the first main problem of emotion categorization.

Handling editor: Tina Iachini (University of Campania Luigi Vanvitelli).

Reviewers: Igor Aleksander (Imperial College London), Antonio Jesus Bandera Rubio (University of Malaga).

✉ Dewi Yanti Liliana
dewiyanti.liliana@tik.pnj.ac.id

T. Basaruddin
chan@cs.ui.ac.id

M. Rahmat Widyanto
widyanto@cs.ui.ac.id

Imelda Ika Dian Oriza
imelda.ika@ui.ac.id

1 Faculty of Computer Science, Universitas Indonesia, Kampus UI, Depok, Indonesia

2 Faculty of Psychology, Universitas Indonesia, Kampus UI, Depok, Indonesia

Published online: 17 June 2019

The second main problem is the variety of the individual facial expressions. Based on psychological analysis, there is a relationship between basic emotion and the variational characteristics of individual in expressing basic emotions (Izard 1992). It means that an expression may have ambiguous emotions interpretation, or someone may express the same emotion in different ways with others. As an example, anger emotion has more than 60 expression variations (Ekman 1992), whereas the existing research simplified and even ignored these facts by classifying emotions into six basic emotion classes in absolute terms.

In computer science, few researchers have put their works to solve mixed emotion recognition (Du et al. 2014; Aly and Tapus 2015; Liliana et al. 2017). They generalized mixed emotion as a combination of two basic emotions, either simultaneously or sequentially in a rapid transition. However, naturally mixed emotion is not limited to only two emotions, in line with Larsen (2014) who stated that mixed emotion represented multiple emotions simultaneously (Larsen and McGraw 2014). To overcome the gap in mixed emotion recognition, multiple emotions in a single image can be modeled into a generalized fuzzy emotion using fuzzy approach from psychological perspective. Fuzzy emotion approach can tolerate the difference in individual facial expression characteristics by utilizing the fuzzy knowledge-based engine and fuzzy inference engine; thus, the system can recognize mixed emotion in a face for a better human and machine interaction.

In this paper, we propose a fuzzy emotion recognition framework based on psychologist principle by observing the changes in facial components, ambiguity, and vagueness of facial expressions. Fuzzy emotion framework naturally recognizes mixed emotion in human facial expression along with the intensity of each emotion. It is consistent with the psychological nature of emotion recognition since it is imitating human activities in recognizing emotions by expressing emotions in their type and intensity, instead of firmly classifying it into several discrete types of emotion. Fuzzy emotion approach is also relevant to artificial intelligence which attempts to mimic human behavior. Fuzzy emotion framework is potentially applicable for the real human–machine interaction and natural emotion recognition.

The contributions of the proposed framework are as follows: (1) solving the limitation of the existing mixed emotion recognition by introducing fuzzy emotion; (2) designing a framework which is adaptive to recognize various facial expressions and multiple emotions; (3) processing robust feature extraction using high-level linguistic conditions of facial features based on pixel coordinates; and (4) developing Indonesian Mixed Emotion Dataset (IMED) to test the performance of the proposed framework as well as to enrich the diversity of available facial expression datasets. These contributions will improve the research in artificial intelligence, especially in social interaction where machine needs to understand not just a single emotion, but multiple emotions displayed by human through facial expressions. Section 2 discusses facial expression and emotion recognition. Next, Sect. 3 describes the fuzzy emotion framework. Section 4 presents results and findings from the experiment, while Sect. 5 discusses the results interpretation. Lastly, Sect. 6 provides concluding remarks.

Facial expression and emotion recognition

In their book, psychologist Ekman (1975) explained the face as a primary signal system to convey emotions (Ekman and Friesen 1975). Facial signals are shown by changes in facial components (eyebrows, eyelids, forehead, cheeks, nose, lips, and chin). Humans read emotions by observing the temporary changes in facial components. The face is a multi-signal message system which provides three types of signals: static, slow, and rapid. Static signals are skin color and face shape; slow signals are permanent wrinkle and skin texture; and rapid signals are the movement of facial muscles such as raising eyebrows and opening mouth. Ekman emphasized that emotion message was delivered by rapid signals.

Emotion is an important element in social interaction. The function is to show a response in a communication. Emotion in human can be recognized from nonverbal cues such as facial expression, gesture, eye gaze, and intonation (Vinciarelli et al. 2009). Among those cues, facial expression is the most dominant signal in conveying emotion where human can identify the related emotion by observing the facial expression. Generally, two types of emotion are basic emotion and non-basic emotion. Some terms used for non-basic emotion are blend, compound, or mixed emotion. According to a psychology theory of emotion, basic emotion is a separate discrete emotion that is significantly different from each other (Ekman 1992). Six categories of universal basic emotions are happiness, sadness, anger, surprise, disgust, and fear (Ekman and Friesen 1975). Meanwhile, mixed emotion is defined as the co-occurrence of two or more basic emotions at the same time (Larsen and McGraw 2014). In addition, facial expression is commonly used as feature descriptors for emotion recognition (Chakrabarti and Dutta 2013; Du et al. 2014; Siddiqi et al. 2015).

Human has many facial expression combinations and different styles to express emotion. Three factors that cause individual variations in facial expressions are: (1) static facial signals that distinguish individual expressions, such as the location of cheek bones, the shape of eyes and eyebrows; (2) personal experiences, which result in the difference of emotion expressions between individuals; and (3) personal display rule which is a characteristic of how people express their emotions (Ekman and Friesen 1975). Based on this explanation, individual facial expressions are ambiguous and irrelevant to be separated into absolute emotion classes. Figure 1 displays an example of ambiguous facial expressions from JAFFE dataset (Kamachi et al. 1997), where a face image which is labeled as sad in Fig. 1b has no significant difference with the neutral class image in Fig. 1a, except the inner eyebrow is slightly raised in the sad image.

The case in Fig. 1 is an example of the ambiguity between two different emotion classes which have nearly identical facial expressions. Even binary classifier using powerful machine learning methods failed to classify the ambiguous facial expressions because the classifier must separate two classes absolutely. Therefore, we propose a fuzzy approach to model the emotion recognition from a single emotion classification into fuzzy emotion which is more adaptive for multiple emotions recognition. Fuzzy logic can handle the case of partial truth where the truth value exists between true and false (Zadeh 1965). Thus, fuzzy emotion accommodates different human perceptions and tolerates non-fix facial expression values; this is what other classifiers cannot do. Machine learning methods process fixed value input, while fuzzy emotion can accept different emotion values.

Recently, several works on facial expression and emotion recognition have been proposed in the literature using fuzzy approaches (Ilbeygi and Shah-Hosseini 2012; Nicolai and Choi 2015; Sujono and Gunawan 2015; Liliana et al. 2016). Those approaches were actually solving basic emotion, not mixed emotion recognition. Table 1 summarizes the aforementioned studies of facial expression and emotion recognition using fuzzy approach. Note that we do not compare the recognition accuracy rate reported by each method since different datasets were used, so the results were incomparable. The last row of Table 1 is the proposed fuzzy emotion framework.

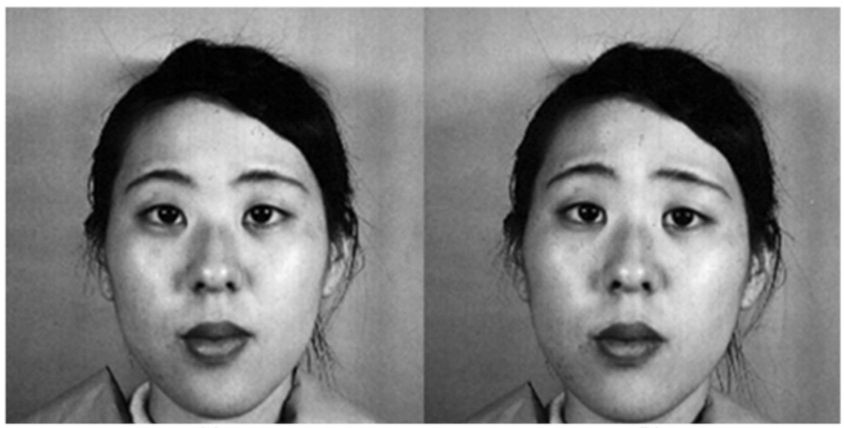

Fig. 1 Ambiguous facial expressions labeled as **a** neutral; **b** sad from JAFFE (Kamachi et al. 1997)

Table 1 Recent works on emotion recognition using fuzzy approach

First author	Feature extraction	Recognition methods	Recognized emotions
Chakraborty (2009)	Three facial features (mouth, eyes, and eyebrows) using fuzzy C-means clustering (FCM)	Fuzzy relational model from facial expression	Six basic emotions
Ilbeygi (2012)	Six facial features (eyebrows, eyes, nose, mouth, lip, and teeth) using Sobel operator	Mamdani fuzzy inference system (FIS) for partially occluded facial expression	Six basic emotions
Nicolai (2015)	Three facial features (mouth, eyes, and eyebrows) using simple thresholding method	Facial action unit (AU) as input and compared with neutral template using Mamdani FIS	Six basic emotions
Sujono (2015)	Three AU (lip depressor, lip raiser, and lip stretcher) using active appearance model (AAM)	Mamdani FIS	Three basic emotions (sadness, normal, and happiness)
Liliana (2016)	Facial points using AAM	FCM clustering	Six basic emotions, plus contempt, and neutral
Proposed framework	AAM and geometric facial features of ten facial components	Mamdani FIS for facial component and Sugeno FIS	Fuzzy emotion (mixed emotions)

Fuzzy emotion recognition framework

The general stages of every emotion recognition system are as follows: face detection; feature extraction; and emotion classification (Ilbeygi and Shah-Hosseini 2012). This proposed framework also follows the same principle of emotion recognition stages; those are (1) facial points detection; (2) geometric features extraction; and (3) fuzzy emotion recognition. Figure 2 depicts the architecture of fuzzy emotion recognition framework. The framework shows four subsystems work sequentially. At the first stage, the input image passes through the active appearance model (AAM) detection subsystem to produce facial points. At the second stage, two subsystems: geometric facial component feature extraction and fuzzy facial component inference system (FFCIS), generate high-level linguistic features of facial components based on FFCIS knowledge-based engine. At the last stage, the fuzzy emotion inference system (FEIS) acts as a classifier which produces fuzzy emotion output based on FEIS knowledge-based engine. The next subsection explains fuzzy emotion subsystems in more detail.

AAM facial points detection

AAM is a statistics-based template matching model for human face detection (Cootes et al. 2001). AAM combines the variability of shape and texture features to determine the location of facial landmarks and return a set of facial point coordinates of face image (Liliana et al. 2016). AAM works by creating a template model which is obtained from training sizeable amount of face images. This model becomes the template reference for face detection. In this study, AAM module developed by Tzimiropoulos and Pantic (2013) is employed for facial points detection. The AAM process steps are described as follows:

1. Initiate face template model fitting.
2. Read shape features and apply Procrustes procedure to remove scale–rotation–translation.
3. Read texture features and apply piecewise affine warping procedure
4. Combine shape and texture into a single feature vector and create the coordinates of AAM facial points.

Shape feature is a configuration of initial facial points. Shape is aligned to a face template model using Procrustes procedure. The alignment between template model and input image is to find parameters of scaling, rotation, and translation that match with the template model. Texture feature is pixel intensities in a face area. This is obtained by superimposing the template model to the image using piecewise affine warping procedure. Details of the AAM steps have been explained in the previous research (Liliana et al. 2016). The example of AAM process is displayed in Fig. 3.

AAM face detection yields 68 facial points coordinate on a cartesian plane. Those points are located on the important facial component areas such as eyebrows, eyes, nose, mouth, and jaw. The next stage after AAM facial point detection is geometric facial component feature extraction.

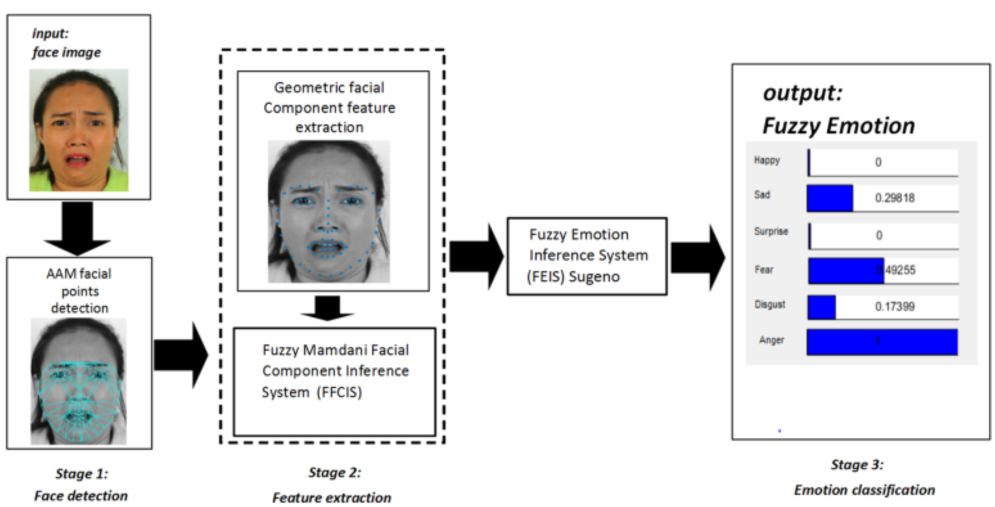

Fig. 2 Fuzzy emotion recognition framework

(a) (b)

Fig. 3 **a** AAM template matching process; **b** AAM facial points result

Geometric facial component feature extraction

Stage 2 in fuzzy emotion recognition is feature extraction. Facial feature extraction is claimed to be the most complicated and time-consuming stage in emotion recognition (Ilbeygi and Shah-Hosseini 2012), since the recognition accuracy fully depends on the precision of the feature representation. Therefore, the selection of feature extraction method is crucial in any emotion recognition system. The main problem in facial expression recognition is the variability of facial expressions that are diverse and almost similar to one another. Variability in facial expressions appears frequently in facial geometric components, while wrinkles and textures have little variation compared to geometric features (Sadeghi et al. 2013). Besides, the geometric-based feature is more reliable against changes in illumination (Loconsole et al. 2014). Here we design a geometric feature extraction method which works on local facial components and transforms them into a vector of facial component values. Geometric features utilize the shape of facial components such as eyebrows stretching, eyes opening, and mouth opening. The geometric feature extraction method is simple and fast since it works on a pixel coordinate basis. This is different from other extraction methods which work on the pixel intensity values.

Ten facial components are used: left eyebrow, right eyebrow, inner eyebrow, nose, upper lip, lower lip, inner mouth, and outer mouth. Geometric feature extraction results in vector $V = \{v_1, v_2, \ldots, v_i\}$, where v_i is the facial point coordinate and $i = 1, \ldots, 68$ is the index of facial point. The objective is to transform the vector V into a more meaningful descriptor to represent the facial expression values using geometric feature extraction methods. The description of facial component is available in Table 2.

Table 2 Description of geometric facial components

Code	Description	Number of facial points	Number of features
gf_1	Left eyebrow	5	2
gf_2	Right eyebrow	5	2
gf_3	Inner eyebrow	4	2
gf_4	Left eye	6	2
gf_5	Right eye	6	2
gf_6	Nose	9	2
gf_7	Upper lip	12	2
gf_8	Lower lip	12	2
gf_9	Inner mouth	8	2
gf_{10}	Outer mouth	12	2

Ten facial components $gf_1, \ldots, gf_{10}$ which are formed from 68 facial points are depicted in Fig. 4. Left and right sides of facial components are based on camera viewpoint.

We process a local feature of facial components through geometric feature extraction. Each facial component has more than one feature value. The result of geometric feature extraction is a set of geometric feature vector $GF = \{gf_{1j}, gf_{2j}, \ldots, gf_{ij}\}$, where gf_{ij} is the geometric feature of ith facial component, $i = 1, \ldots, 10$ is the index of facial component, and j is the number of geometric features in each facial component.

Two types of geometric features applied on facial components are eccentricity and distance ratio. Eccentricity is a property of an ellipse which is used to measure how elliptic an object is; the value ranges from zero to one (Loconsole et al. 2014). Circle is an ellipse with zero eccentricity value. While distance ratio is the distance comparison between two values, e.g., for mouth opening, the distance ratio is the comparison between mouth height and width. Distance ratio is also applied

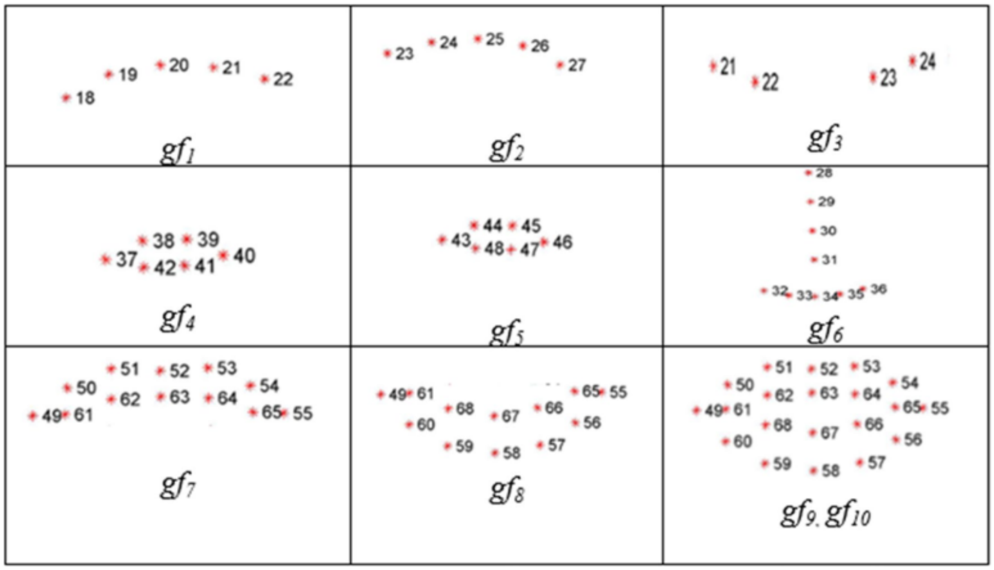

Fig. 4 Ten facial components

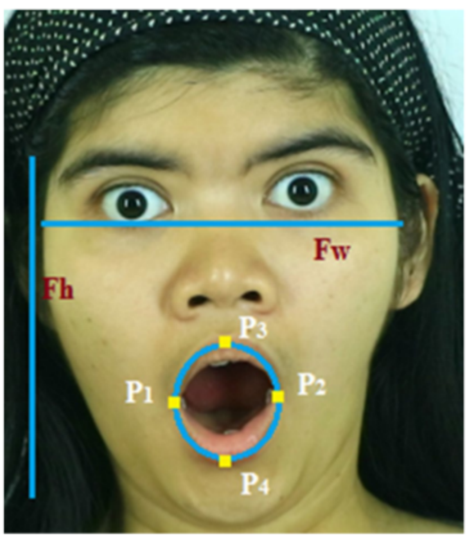

Fig. 5 Example of geometric feature points

to the face height (Fh) and the face width (Fw). Figure 5 gives an example of geometric facial feature points.

Using Fig. 5 as an example, the eccentricity feature is measured using formula (1).

$$e = \frac{\sqrt{a^2 - b^2}}{a} \tag{1}$$

where e is the eccentricity value, and a and b are

$$a = \frac{P_{2x} - P_{1x}}{2} \tag{2}$$

$$b = \frac{P_{3y} - P_{4y}}{2}, \tag{3}$$

where P_1, P_2, P_3, P_4 are facial points as illustrated in Fig. 5, and x and y are the value of x-axis and y-axis of facial point P_i. Meanwhile, distance ratio d is measured using the formula in Eq. (4),

$$d = \frac{P_{3y} - P_{4y}}{P_{2x} - P_{1x}} \tag{4}$$

Eccentricity feature is applied on facial components which have either full or half-elliptical shapes, such as eyebrows, eyes, lips, and mouth; eyebrows and lips are half-ellipses. Geometric feature extraction results in GF feature vector with length = 20. This vector is forwarded as an input to a fuzzy facial component inference system (FFCIS) which will generate the high-level linguistic condition of facial components. Next subsection describes the FFCIS subsystem as a high-level linguistic facial component feature extraction.

Fuzzy facial component inference system (FFCIS)

The purpose of FFCIS subsystem is to transform the geometric feature vector GF into a set of high-level features of linguistic conditions. The linguistic feature is naturally used by a psychologist to describe the facial components states (e.g., raising eyebrows, tightening lips, wrinkling nose). Each facial component has its own FFCIS; hence, there are ten FFCIS subsystems for all facial components. We use Mamdani-type fuzzy inference since the output is a fuzzy set (Mamdani 1974). For each FFCIS, *GF* parameter values become the input and the outputs are fuzzy linguistic conditions of the facial component. The input has three membership functions: low, medium, and high. Each FFCIS has different outputs which correspond to the linguistic condition of the facial component itself; e.g., eyebrows are lower, normal, and raised; eyes are narrow, normal, wide; nose is normal, wrinkled; lips are thin, normal, thick; and mouth is tight, normal, open. The FFCIS rules are enumerated based on psychologist knowledge. Fuzzy triangular function is used as membership function. Figure 6 gives the example of FFCIS input membership functions. A set of rules for each FFCIS is stored in the fuzzy knowledge-based engine. Ten FFCIS output are then forwarded into the fuzzy emotion inference system.

Fuzzy emotion inference system (FEIS)

FFCIS output or facial component linguistic condition values subsequently become input for FEIS subsystem. There are six separate FEIS subsystems, each of it for different classes of basic emotion. The six FEIS subsystems are independent of one another. Figure 7 depicts the FEIS process diagram. The reason behind the separation of emotion into six fuzzy emotion inference engines is that those emotions are not in the same dimension or interval. Thus, it cannot be classified into a single output dimension since it is irrelevant; each emotion class has its own dimension and interval value with different intensities.

FEIS has three main processes: input fuzzification, fuzzy emotion inference, and output defuzzification. Ten FFCIS values become input which passes through the six FEIS. Although each emotion inference subsystem is independent, the input for each subsystem is different from one another

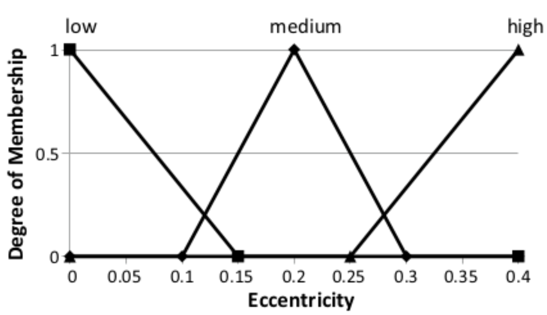

Fig. 6 FFCIS input membership functions

and depends on the facial components that dominantly define an emotion. The mapping between input and output is as follows:

1. Anger: inner eyebrow, left eye, right eye, inner mouth
2. Disgust: left eyebrow, right eyebrow, nose, inner mouth
3. Fear: left eyebrow, right eyebrow, left eye, right eye, upper lip, lower lip
4. Happiness: left eye, right eye, upper lip, lower lip, inner mouth
5. Sadness: left eyebrow, right eyebrow, inner eyebrow, upper lip, lower lip, inner mouth
6. Surprise: left eyebrow, right eyebrow, inner eyebrow, left eye, right eye, nose, upper lip, lower lip, inner mouth, outer mouth.

Each FEIS has a different facial component input, while the results are fuzzy emotion values of every emotion with the value range from 0 to 1. Fuzzification transforms input into fuzzy sets based on fuzzy membership functions. We use two types of fuzzy membership functions, triangular and trapezoidal, in the FEIS fuzzification process. Table 3 describes the facial component fuzzification properties. Each facial component (FC) possesses different linguistic variables based on its characteristic.

Emotion inference involves facial components input, emotion rules, and emotion output. Fuzzy Sugeno type-0 is performed with constant output for inferencing input into output based on the reason that the emotion output has a categorical value. The assigned output values are 1 = slight; 2 = moderate; and 3 = extreme. As an example, in anger FEIS we are mapping four inputs (gf_3, gf_4, gf_5, gf_9) into three categorical fuzzy output values: slight, moderate, and extreme. Figure 8 is an example of inner mouth membership functions.

The emotion rules are elicited from the psychologist knowledge which is extracted from the Ekman's theory (Ekman and Friesen 1975) as well as from the observation and the interview with psychologist experts. These rules are stored in a knowledge-based engine. Each FEIS owns different numbers of rules: 42 anger rules, 37 disgust rules, 68 fear rules, 73 happiness rules, 44 sadness rules, and 71 surprise rules. The number of rules is relevant to the number of FC inputs. Some examples of FEIS rules for each emotion class are as follows:

1. IF inner eyebrow is closer AND left eye is normal AND right eye is NORMAL and inner mouth is widely open THEN anger is extreme.
2. IF left eyebrow is raised AND right eyebrow is raised AND nose is wrinkled AND inner mouth is narrow THEN disgust is moderate.

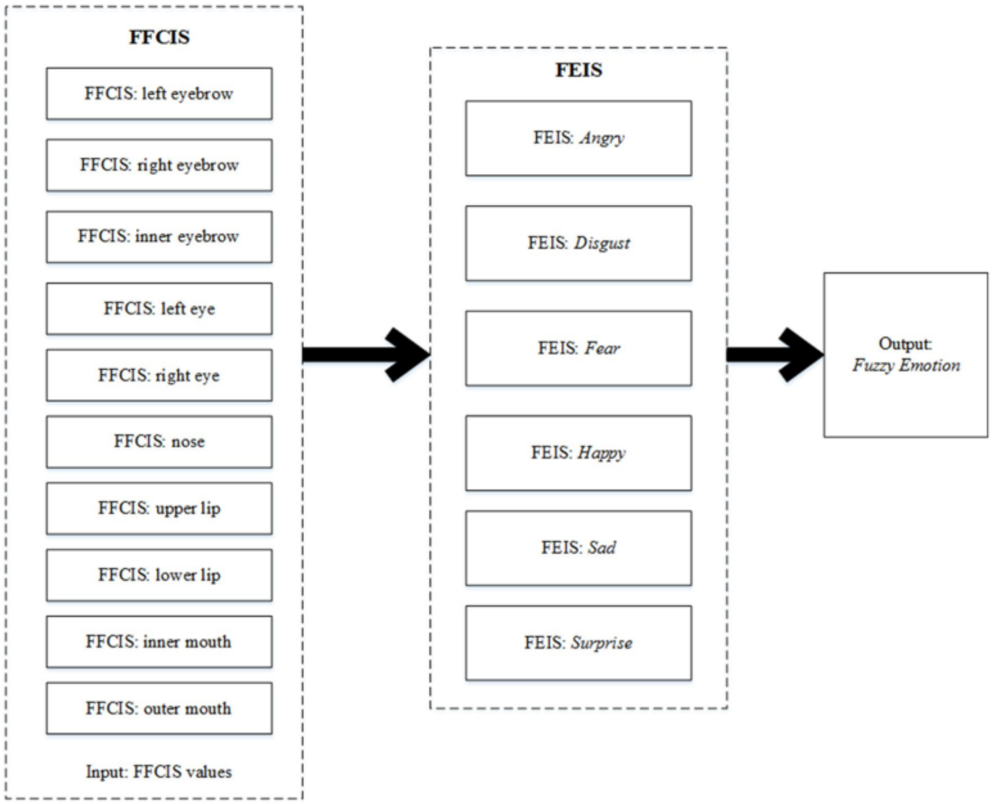

Fig. 7 Fuzzy emotion inference system process diagram

Table 3 Description of geometric facial components

FC code	FC input	Linguistic variables	Membership function
gf_1	Left eyebrow	Lower, normal, raised	Triangular
gf_2	Right eyebrow	Lower, normal, raised	Triangular
gf_3	Inner eyebrow	Closer, normal	Trapezoidal
gf_4	Left eye	Narrow, normal, wide	Triangular
gf_5	Right eye	Narrow, normal, wide	Triangular
gf_6	Nose	Normal, wrinkled	Trapezoidal
gf_7	Upper lip	Thin, normal, thick	Triangular
gf_8	Lower lip	Thin, normal, thick	Triangular
gf_9	Inner mouth	Close, narrow, normal, open, widely open	Triangular
gf_{10}	Outer mouth	Close, narrow, normal, open, widely open	Triangular

3. IF left eyebrow is normal AND right eyebrow is normal AND left eye is narrow AND right eye is narrow AND upper lip is thin AND lower lip is thin THEN fear is slight.
4. IF left eyebrow is lower AND right eyebrow is lower AND upper lip is thick AND lower lip is thin AND inner mouth is normal THEN happiness is moderate.
5. IF left eyebrow is raising AND right eyebrow is raised AND inner eyebrow is normal AND upper lip is normal AND lower lip is thin AND inner mouth is close THEN sadness is slight.
6. IF left eyebrow is raised AND right eyebrow is raised AND left eye is wide AND right eye is wide AND upper

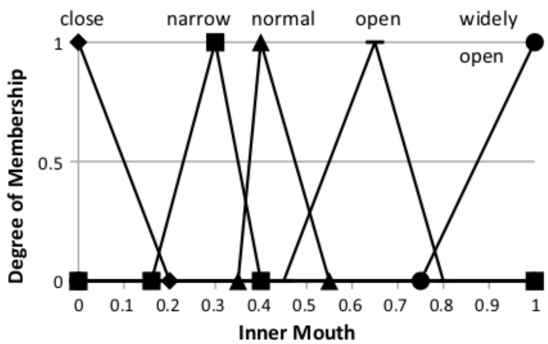

Fig. 8 FEIS inner mouth input membership functions

lip is normal AND lower lip is normal AND outer mouth is widely open THEN surprise is extreme.

Fuzzy emotion works by inferencing linguistic conditions from the knowledge-based engine to produce fuzzy emotion values. All firing rules are then composed and mapped into a solution in the output space. FEIS performs defuzzification to gain the normalized intensities of each emotion class. (Value ranges between 0 and 1.) Six FEIS subsystems are running concurrently. The results are six intensities of multiple emotions. The next section discusses experiment results and findings of the proposed system.

Results

The implementation of fuzzy emotion framework is in MATLAB 2014 program. The recognition performance is being observed through experiments. Two experiment scenarios are arranged to test the proposed system where each of it had different objectives: First, to measure the recognition accuracy in several facial expression datasets and, second, to compare the recognition performance of the proposed system with other classifiers. In the experiment, the most dominant emotion intensity determined the emotion class of a test image. The recognition result was correct if the system accurately captured the emotion which was cross-checked to the image label on the dataset and the expert rater validation. Thirty expert raters who had expertise in fuzzy approach and emotion recognition validated the results to avoid subjective claims.

The first testing scenario used four facial expression datasets. The first dataset was extended Cohn–Kanade or CK+ facial expression dataset using 304 face images (Lucey et al. 2010). The second was Japanese Female Facial Expression or JAFFE dataset using 183 face images (Kamachi et al. 1997). The third was Denver Intensity of Spontaneous Facial Action or DISFA dataset using 147 face images (Mavadati et al. 2013). The last one was the custom-made dataset: Indonesian Mixed Emotion Dataset (IMED) using 270 images. IMED consisted of 15 subjects from different ethnics, ages, and gender who posed mixed emotion which has been validated by psychologists (available at http://imed.cs.ui.ac.id).

Figure 9 gives examples of fuzzy emotion recognition using images from CK+.

Fuzzy emotion displayed mixed emotions in a facial expression through the intensity bar ranging from 0 to 1 with three classifications: slight, moderate, and extreme emotion intensities. As shown in Fig. 9a, the proposed system acknowledged the intensity of sadness was 0.521; it was a moderate sadness with slight intensity of disgust 0.02. Meanwhile in Fig. 9b, the image was recognized as the combination of moderate sadness, moderate anger, and slight disgust with intensities 0.5, 0.5, and 0.2, respectively. This is an important finding in understanding fuzzy emotion that produces multiple emotions recognition and also the intensity of each emotion in a facial expression image. We summarize the result of scenario one in Table 4 that shows the performance of fuzzy emotion recognition tested on four datasets.

From Table 4, we can see that the recognition performance using four datasets is varying from 0.80 to 0.95 (on a scale of 0–1). The average accuracy rate for all datasets is 0.88 with standard deviation 0.04. Surprise FEIS gained the highest average accuracy rate for all testing data (0.92), while the lowest are anger and sadness (0.85). Meanwhile, the highest average accuracy rate is CK+ (0.90), followed by DISFA (0.89) and IMED (0.87), while the lowest is JAFFE (0.82). In CK+, the highest recognition is happiness (0.95); in JAFFE it is surprise (0.93); in DISFA it is disgust (0.93); and in IMED it is sadness and surprise (0.89). Graphic in Fig. 10 visualizes the performance of fuzzy emotion on different datasets.

The horizontal and vertical axes in Fig. 10 are emotion classes and the accuracy rate, respectively. We can observe the graphic by focusing on each emotion category in the horizontal axes. Four different bold markers are the four datasets. The stacked markers indicate that the recognition performance has similar accuracy; for example, in anger class, DISFA and IMED are stacked; in disgust, CK+ and JAFFE are stacked; and in sadness, CK+ and DISFA are stacked. The experiment result in scenario 1 confirmed the finding about fuzzy emotion recognition performance using different datasets which was consistently high with average accuracy rate of 0.88. The main contribution was the multiple fuzzy emotions recognition with intensities for each input image. However, in line with the idea of discrete emotion categorization, the highest intensity class was being determined as the emotion label to validate the result with the ground truth label on the dataset to get the accuracy rate.

Fig. 9 Fuzzy emotion output: **a** sadness–disgust; **b** sadness–disgust–anger

Table 4 Fuzzy emotion recognition results

Dataset	# Data	FEIS accuracy rate						Average accuracy
		Anger	Disgust	Fear	Happiness	Sadness	Surprise	
CK+	304	0.89	0.83	0.87	0.95	0.86	0.94	0.90
JAFFE	183	0.80	0.83	0.84	0.84	0.81	0.93	0.82
DISFA	147	0.85	0.93	0.92	0.91	0.86	0.92	0.89
IMED	270	0.84	0.87	0.87	0.87	0.89	0.89	0.87
Average		0.85	0.87	0.87	0.89	0.85	0.92	

Average accuracy rate $=0.88\pm0.04$

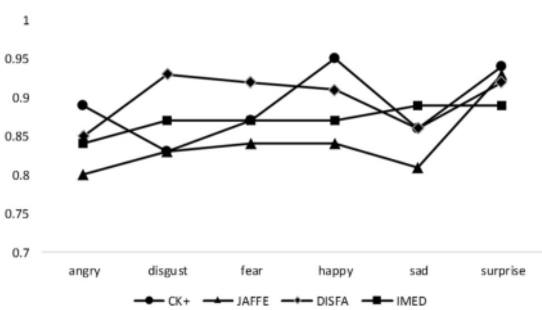

Fig. 10 Fuzzy emotion comparison using different datasets

In case of single facial expression, fuzzy emotion was also able to result in a single emotion intensity.

The second testing scenario compared the performance of FEIS with two other classifiers which performed basic emotion classification. Fuzzy C-means (FCM) and support vector machine (SVM) classifiers were tested under the same condition using CK+ dataset consisting of 304 images, and the accuracy rate of the proposed system was compared with those of other classifiers. The result is presented in Table 5.

Table 5 describes the recognition performance of three classifiers: FCM, SVM, and the proposed FEIS. SVM obtained the highest average accuracy rate 0.92 followed by

Table 5 Comparison with different classifiers using CK + dataset

Classifier	FEIS accuracy rate						Average accuracy
	Anger	Disgust	Fear	Happiness	Sadness	Surprise	
FCM	0.80	0.71	0.80	0.85	0.84	0.87	0.81
SVM	0.92	0.88	0.93	0.95	0.91	0.93	0.92
FEIS	0.89	0.83	0.87	0.95	0.86	0.94	0.90

FEIS 0.90, while FCM obtained the lowest average accuracy rate 0.81. Although FEIS obtained a slightly lower accuracy rate than SVM, it is important to highlight the fact that FEIS can perform multiple recognitions. Unlike traditional classifiers (SVM and FCM) which voted only a single emotion class and ignored the rest emotions occurrence in facial expression, fuzzy emotion preserved all the emotions contained in a facial expression by showing the intensity of each emotion class. It is interesting to note that in FEIS, high-level linguistic features of facial components determine the emotion classification result. Meanwhile, SVM and FCM used low-level image features; the facial point coordinate. Low-level feature is only understandable by machine, while high-level feature is close to human intelligence. This implies that FEIS has performed a human natural recognition through a set of facial component rules and has provided the high accuracy of recognition results.

Discussion

This section recaps findings and contributions made based on experiment results. By comparing results, we determine the main finding that the proposed fuzzy emotion recognition framework has successfully recognized single and multiple emotions from facial expression images. Moreover, fuzzy emotion has three types of emotion intensities: slight, moderate, and extreme. This idea comes from the psychology theory of emotion that was coined by Ekman who stated that the intensity of emotion was not absolute but did not consist of several degrees such as slight anger, moderate sadness, or extreme surprise (Ekman 1992). Two testing scenarios were used to measure the performance using different facial expression datasets and to compare the recognition accuracy with other methods. The experiment on four facial expression datasets, including the custom-made mixed emotion dataset (IMED), gave a high performance with average accuracy rate of 0.88. Experiments on different datasets proved that recognition performance is consistently high and uniform, not only in one dataset variant, with the accuracy range in 0.82–0.92.

When compared to other classifiers (SVM and FCM) using CK+ dataset, fuzzy emotion system obtained an accuracy rate of 0.90. It has a slight difference with SVM classifier (0.92) which performed single class classification, but far more accurate than FCM classifier (0.81). Furthermore, fuzzy emotion displays each emotion intensity based on facial expression and facial component rules. Fuzzy emotion has adaptively recognized mixed emotion and its intensities; this is more advanced compared to prior research on mixed emotion recognition conducted by Du et al. (2014) which only recognized compound emotion as a combination of two basic emotions.

Although the framework consists of many subsystems and seems complex, the system is simple yet robust in the implementation, unlike other artificial intelligence methods with data-driven nature (such as SVM and FCM), which require large amount of training data for learning process. The proposed system has a knowledge-driven nature, and this is the true nature of artificial intelligence which utilizes human expert knowledge. By defining emotion rules and storing it in knowledge-based engines, it allows the system to generalize the recognition result without any effort to train the model. Moreover, simple pixel-based calculation techniques for geometric facial feature extraction work fast by using pixel coordinates location, and they do not require any facial template to extract the geometric features. The limitation of the present study is in the fuzzy input–output parameter setting which was done based on the experiment. This parameters setting is very important in influencing the recognition result. Hence, parameter optimization technique is needed to improve the recognition accuracy.

Conclusion

This study is introducing fuzzy emotion as a new approach to determine mixed emotion in a single face image from psychological perspective. In general, the fuzzy emotion recognition framework consists of two main parts: the processing subsystem and the knowledge subsystem. The processing subsystems are as follows: facial point detection by AAM, geometric feature extraction, fuzzy facial component inference system (FFCIS), and fuzzy emotion inference system (FEIS). While knowledge subsystems are fuzzy facial component rule-based engine and fuzzy emotion rule-based engine, both rule-based engines are extracted from the psychologist knowledge. Fuzzy emotion recognition aims to classify multiple emotions from six basic emotions that coexist in a single facial expression image. The performance of the system on four facial expression datasets which is investigated by 30 expert raters has reached excellent results with an average accuracy rate of 0.88. The accuracy is comparable to a well-known SVM classifier for emotion recognition. Although existing researches on automatic emotion recognition mostly performed distinct emotion classification, the results reported here show otherwise. Fuzzy emotion advances the existing researches by capturing multiple emotions in a face image, as well as determining the intensity of mixed emotions. Fuzzy emotion recognition has a potential to be applied in social robotics, where robot can analyze human real emotions as well as synthesize fuzzy emotion as a natural response in an interaction. Moreover, fuzzy emotion recognition can help people with alexithymia disorder who are unable to recognize or describe their emotions. In the future, we will enhance the performance of the system by modifying image features, optimizing the fuzzy parameters, and improving the IMED metadata for a better utilization.

Acknowledgements The authors would like to thank all the contributors as well as participants in this research. The first author would also like to thank the Indonesia Endowment Fund for Education (LPDP) for the doctoral study sponsorship.

Funding This study was funded by the Indonesia Ministry of Research Technology and Higher Education through the Postgraduate Team Grant number 120/SP2H/PTNBH/DRPM/2018.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

Aly A, Tapus A (2015) An online fuzzy-based approach for human emotions detection: an overview on the human cognitive model of understanding and generating multimodal actions. Spring Tracts Adv Robot 106:185–212. https://doi.org/10.1007/978-3-319-12922-8_7

Chakrabarti D, Dutta D (2013) Facial expression recognition using eigenspaces. Proc Technol 10:755–761. https://doi.org/10.1016/j.protcy.2013.12.419

Chakraborty A, Konar A, Chakraborty UK, Chatterjee A (2009) Emotion recognition from facial expressions and its control using fuzzy logic. IEEE Trans Syst Man Cybern Part A Syst Humans 39(4):726–743. https://doi.org/10.1109/TSMCA.2009.2014645

Cootes TF, Edwards GJ, Taylor CJ (2001) Active appearance models. IEEE Trans Pattern Anal Mach Intell 23:681–685. https://doi.org/10.1109/34.927467

Du S, Tao Y, Martinez AM (2014) Compound facial expressions of emotion. Proc Natl Acad Sci 111:E1454–E1462. https://doi.org/10.1073/pnas.1322355111

Ekman P (1992) An argument for basic emotions. Cogn Emot 6:169–200. https://doi.org/10.1080/02699939208411068

Ekman P, Friesen WV (1975) Unmasking the face: a guide to recognizing emotions from facial clues. Prentice-Hall, Oxford

Ilbeygi M, Shah-Hosseini H (2012) A novel fuzzy facial expression recognition system based on facial feature extraction from color face images. Eng Appl Artif Intell 25:130–146. https://doi.org/10.1016/j.engappai.2011.07.004

Izard CE (1992) Basic emotions, relations among emotions, and emotion-cognition relations. Psychol Rev 99:561–565. https://doi.org/10.1037/0033-295X.99.3.561

Kamachi M, Lyons M, Gyoba J (1997) The Japanese female facial expression (jaffe) database. http://www.kasrl.org/jaffe.html. Accessed 3 Oct 2018

Larsen JT, McGraw A (2014) The case for mixed emotions. Soc Personal Psychol Compass 8:263–274. https://doi.org/10.1111/spc3.12108

Liliana DY, Widyanto MR, Basaruddin T (2016) Human emotion recognition based on active appearance model and semi-supervised fuzzy C-means. In: 2016 international conference on advanced computer science and information systems (ICACSIS), pp 439–445

Liliana DY, Basaruddin C, Widyanto MR (2017) Mix emotion recognition from facial expression using SVM-CRF sequence classifier. In: Proceedings of international conference on algorithms, computing and systems - ICACS'17, pp 27–31. https://doi.org/10.1145/3127942.3127958

Loconsole C, Miranda CR, Augusto G et al (2014) Real-time emotion recognition—novel method for geometrical facial features extraction. In: Proceedings of the 9th international conference on computer vision theory and applications, pp 378–385

Lucey P, Cohn JF, Kanade T et al (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, CVPRW, pp 94–101. https://doi.org/10.1109/cvprw.2010.5543262

Mamdani EH (1974) Application of fuzzy algorithms for control of simple dynamic plant. Proc Inst Electr Eng 121:1585–1588. https://doi.org/10.1049/piee.1974.0328

Mavadati S, Mahoor M, Bartlett K, Trinh P, Cohn J (2013) DISFA: A Spontaneous Facial Action Intensity Database. IEEE Trans Affect Comput 4(2):151–160. https://doi.org/10.1109/T-AFFC.2013.4

Nicolai A, Choi A (2015) Facial emotion recognition using fuzzy systems. In: 2015 IEEE international conference on systems, man, and cybernetics, pp 2216–2221. https://doi.org/10.1109/smc.2015.387

Sadeghi H, Raie A-A, Mohammadi M-R (2013) Facial expression recognition using geometric normalization and appearance representation. In: 2013 8th Iranian conference on machine vision and image processing (MVIP). IEEE, pp 159–163

Siddiqi MH, Ali R, Khan AM et al (2015) Human facial expression recognition using stepwise linear discriminant analysis and hidden conditional random fields. IEEE Trans Image Process 24:1386–1398. https://doi.org/10.1109/TIP.2015.2405346

Sujono, Gunawan AAS (2015) Face expression detection on Kinect using active appearance model and fuzzy logic. Proc Comput Sci 59:268–274. https://doi.org/10.1016/j.procs.2015.07.558

Tzimiropoulos G, Pantic M (2013) Optimization problems for fast AAM fitting in-the-wild. In: 2013 IEEE international conference on computer vision, pp 593–600

Vinciarelli A, Salamin H, Pantic M (2009) Social signal processing: understanding social actions through nonverbal behaviour analysis. In: 2009 IEEE computer society conference on computer vision and pattern recognition workshops, vol 3, pp 42–49. https://doi.org/10.1109/cvprw.2009.5204290

Zadeh LA (1965) Fuzzy sets. Inf Control 8:338–353. https://doi.org/10.1016/S0019-9958(65)90241-X

Fuzzy emotion: a natural approach to automatic facial expression recognition from psychological perspective using fuzzy system

ORIGINALITY REPORT

4%	2%	3%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Praveen Kulkarni, Rajesh T. M.. "A Multi-Model Framework for Grading of Human Emotion Using CNN and Computer Vision", International Journal of Computer Vision and Image Processing, 2022 Publication	1%
2	docs.autismresearchcentre.com Internet Source	1%
3	Paramartha Dutta, Asit Barman. "Human Emotion Recognition from Face Images", Springer Science and Business Media LLC, 2020 Publication	1%
4	boyuan.global-optimization.com Internet Source	1%

Exclude quotes	On	Exclude matches	< 1%

Exclude bibliography On

Fuzzy emotion: a natural approach to automatic facial expression recognition from psychological perspective using fuzzy system

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11

PAGE 12

PAGE 13