



# Emotion Profiling: Ingredient for Rule based Emotion Recognition Engine

By Dr. Preeti Khanna & Dr. Sasi M Kumar

*SVKMs NMIMS - School of Business Management, India*

*Abstract-* Emotions are considered to be the reflection of human thinking and decision-making process which increase his/her performance by producing an intelligent outcome. Hence it is a challenging task to embed the emotional intelligence in machine as well so that it could respond appropriately. However, present human computer interfaces still don't fully utilize emotion feedback to create a more natural environment because the performance of the emotion recognition is still not very robust and reliable and far from real life experience. In this paper, we present an attempt in addressing this aspect and identifying the major challenges in the process. We introduce the concept of 'emotion profile' to evaluate an individual feature as each feature irrespective of the modality has different capability for differentiating among the various subsets of emotions. To capture the discrimination across target emotions w.r.t. each feature we propose a framework for emotion recognition built around if-then rules using certainty factors to represent uncertainty and unreliability of individual features.

*Keywords:* human computer interaction; emotion; rule based system; emotion profile; multimodalities.

*GJCST-F Classification :* I.7.5 , I.4.0



*Strictly as per the compliance and regulations of:*



# Emotion Profiling: Ingredient for Rule based Emotion Recognition Engine

Dr. Preeti Khanna <sup>α</sup> & Dr. Sasi M Kumar <sup>σ</sup>

**Abstract-** Emotions are considered to be the reflection of human thinking and decision-making process which increase his/her performance by producing an intelligent outcome. Hence it is a challenging task to embed the emotional intelligence in machine as well so that it could respond appropriately. However, present human computer interfaces still don't fully utilize emotion feedback to create a more natural environment because the performance of the emotion recognition is still not very robust and reliable and far from real life experience. In this paper, we present an attempt in addressing this aspect and identifying the major challenges in the process. We introduce the concept of 'emotion profile' to evaluate an individual feature as each feature irrespective of the modality has different capability for differentiating among the various subsets of emotions. To capture the discrimination across target emotions w.r.t. each feature we propose a framework for emotion recognition built around if-then rules using certainty factors to represent uncertainty and unreliability of individual features. This technique appears to be simple and effective for these kind of problems.

**Keywords:** human computer interaction; emotion; rule based system; emotion profile; multimodalities.

## I. INTRODUCTION

The difference between machine and human is not only that a human being is intelligent but also that he/she is emotional (Marinez-Miranda et al., 2005). The emotions enable human to interact intelligently and effectively with other humans. Same concept could be extended to human computer interaction (HCI). It deals with various procedures and methods through which humans interact with computer. According to Foley (1996) HCI is a socio-technological discipline whose goal is to bring computer and communication system to society and its people in such a way that both become accessible and hence useful in working, learning, communicating and recreational lives (Foley, 1996). The study related to HCI draws from supporting knowledge on both the machine and the human side. On the machine side computer graphics, operating systems, and programming languages are relevant while on the human side communication, social sciences, cognitive psychology, and human performance are relevant. As computers become more pervasive in culture, researchers are increasingly looking for new and innovative ways to design these interfaces more interactive and efficient. By embedding emotions in the

interaction of human with machine, machine would be in a position to sense the mood of the user and change its interaction accordingly. The system will be friendlier to the user and its responses will be more similar to human behavior. Motivations for emotional computing are manifold. From a scientific point of view, emotions play an essential role in decision making, as well as in perception and learning. Emotions influence various cognitive processes of people (Lisetti & Nasoz, 2005) including perception and organization of memory (Bower, 1981), categorization and preference (Zajonc, 1984), goal generation, evaluation, and decision-making (Damasio, 1994), strategic planning (Ledoux, 1992), focus and attention (Derryberry & Tucker, 1992), motivation and performance (Colquitt et al., 2000), intention (Frijda, 1986), communication (Birdwhistle, 1970; Ekman & Friesen, 1975; Chovil, 1991), and learning (Goleman, 1995). A common everyday task is driving, and yet research suggests that people emote while driving and their driving is affected by their emotions (James & Nahl, 2000). The inability to control one's emotions while driving is often identified as one of the major causes for accidents. Also by knowing the user's emotions, computer agents can become more effective in tutoring. A computer agent can learn the student's preferences and offer better interactions. Surveillance is another application domain in which the reading of emotions may lead to better performance in predicting the future actions of subjects. In this way, the emotion driven technology can enhance the existing systems for the identification and prevention of terrorist attacks in public places. Certainly not all computers need to pay attention to emotions, or have emotional abilities. Some machines are useful as rigid tools, and it is fine to keep them that way.

The paper begins by identifying the challenges in problem domain of emotion recognition. A complete framework of emotion recognition using rule based approach independent of any modalities (like speech or facial expressions) is then introduced. Core of this approach is feature analysis which has been explored using 'emotion profiling'. Finally the whole approach of rule based emotion recognition has been implemented using a running case scenario of facial expressions. Finally the performances of recognizing the target emotions have been reported. We conclude the paper by summarizing the results and consider some challenges facing the researchers in this area.

**Author α σ:** e-mails: preeti.khanna@nmims.edu, pkhanna0578@gmail.com, the.little.sasi@gmail.com

## II. EMOTION RECOGNITION IS CHALLENGING!

Research related to emotion recognitions is tough because understanding emotion is difficult. To address the problem of emotion recognition various modalities like speech (Khanna & Kumar, 2010; Khanna & Kumar, 2011), facial expression, gesture, keyboard interaction (Khanna & Kumar, 2010), etc. had been explored. Some of the major challenges in this domain are as follows.

### a) *Choice of Features*

The number of features used in the process of recognizing emotions from different modalities varies and depends on the application. Having a large number of features increases the complexity of the system, normally results in longer system training time and demand rich set of training data. Hence selection of features is a critical task.

### b) *Choice of Machine Learning Techniques*

Depending on the context and the type of data, the classification algorithms used for emotion recognition have been constantly evolving in course of time. Various recognition methods have been used in the literature. One major dimension for variability among algorithms is the nature of knowledge representation used by these algorithms.

### c) *Emotional Database Issues*

Data is of utmost importance. Having an appropriate database that is collected with a particular application and target user profile in mind can be expected to minimize the confusions that occur while organizing and labeling the emotional database. For example issues like emotion elicitation method (i.e., whether the elicited emotion displays are posed or spontaneous), size (the number of subjects), modality (audio, visual, etc.), emotion description (category or dimension), and labeling scheme is tedious job. Hence choice of database is another concern.

### d) *Choice of Emotions*

This requires identification of the emotional states which have a bearing on HCI. It is not important to track all variants of emotion as a principle. Literature defines various subsets of emotions based on desired granularity and other parameters. Researchers still do not agree on what an emotion is and many of them do not consider a specific subset of emotions as 'basic set'. Hence defining and identifying the emotional state is a challenge.

### e) *Choice of Modalities and Fusion*

The studies show multiple different modalities as the source for emotion like face, voice, gesture, etc. The accuracy of recognition from different sources may vary with time, for example, facial expression recognition is dependent on the system getting an adequate frontal

view under good illumination. But in reality, one uses a combination of all these and they do not exist independently. Indeed, at times, the signals from the different sources may conflict each other, indicating different emotional states. However, most of the time the different sources provide additional information reinforcing the estimates made using one source and thus in determining the states with better confidence. Given the difficulties in mapping emotional states to recognizable characteristics in the various individual modalities, it becomes important to use multiple sources together. Picard (1997) observes that affect recognition is most accurate when it combines multiple modalities, information about the user's context, situation, goal, and preferences. But too much information from different modalities simultaneously seems to be confusing for human judges (Picard, 1997). Does this pertain in HCI too, needs to be addressed?

Hence due to multimodalities, problems related to data fusion are common. Humans simultaneously employ modalities of sight and sound. Does this tight coupling persist when the modalities are used for human behavior analysis, as suggested by some researchers, or not, as suggested by others? Does this depend on the machine learning techniques employed? In literature, some attempts like (De Silva & Ng, 2000), (Sebe et al., 2006) and (Zeng et al., 2007) have considered the integration of information from facial expressions and speech. Kim and Andre (2006) concentrated on the integration of physiological signals and speech signals for emotion recognition based on short term observation (Kim & Andre, 2006). In general there are two broad approaches which combine the inputs from different sources- feature based fusion and decision based fusion. Feature based fusion involves simply merging the features of each modality into a single feature vector. Decision based fusion is based on the fusion of decisions from each modality where the input coming from each modality is processed independently and these results are combined at the end. Several works like (Corradini et al., 2003), (Liao, 2002), (Kettebekov & Sharma, 2000), and (Sharma, 1998) discussed many issues and techniques of multimodal fusion. Finding an optimal fusion type for a particular combination of modalities is not straightforward. Hybrid fusion attempts to combine the benefits of both feature level and decision level fusion method. This may be a good choice for some multimodal fusion problems. However, based on existing knowledge and methods, how to combine the information coming from different modalities for the target set of emotions is still an open problem. In this paper we propose a rule based approach to recognize target emotions. This approach remains independent of modalities (speech or facial expressions or others).

### III. RULE BASED APPROACH

A rule based system, in general, consists of if-then rules, a bunch of facts, and an interpreter controlling the application of the rules. One of the major strength of rule based representation is its ability to represent various uncertainties. Uncertainty is inherently part of most human decision making. This uncertainty could arise from various sources like incomplete data or domain knowledge used being unreliable. If – then rules is often represented like 'If A, B, C ----> then D, with certainty 'X', where X represents the degree of belief or confidence in the rule (Kumar et al., 2007). To handle uncertainties, there are two broad approaches, those representing uncertainty using numerical quantities and those using symbolic methods. For example, Bayesian reasoning (Shortliffe & Buchanan, 1975), Evidence theory (Gordon & Shortliffe, 1984) and Fuzzy set approaches (Negoita, 1985) are numerical models. On the other hand, symbolic characterization of uncertainty is mostly aimed at handling incomplete information, for example Assumption Based Reasoning (Doyle, 1979), Default Reasoning (Reiter, 1980) and Non-monotonic Logic (McDermott & Doyle, 1980). In our domain, the basic problem is that there are hardly any feature or feature combinations which can infer any emotion to complete certainty. Therefore, we concentrate on numerical approaches for handling the uncertainty. We have adopted the 'Confirmation Theory' as used in MYCIN approach (Shortliffe & Buchanan, 1975). This approach works well with rule based representation of domain knowledge.

Shortliffe and Buchanan, 1975 developed the Certainty Factor (CF) model in the mid-1970s for MYCIN, an expert system for the diagnosis and treatment of infections of the blood (Shortliffe & Buchanan, 1975). Since then, the CF model has been widely adopted for uncertainty management in many rule based systems. Each rule is assigned CF by domain experts. Higher CF indicates that the conclusion can be asserted with higher confidence when the conditions are true. CF denotes change in belief in a hypothesis given some evidence. A value of +1.0 indicates absolute belief and -1.0 indicates absolute disbelief. The method generally used to propagate the measure of uncertainty in the antecedents and the uncertainty attached to the rule to the conclusions being derived is briefly explained below. This propagation is done in two steps (Kumar, et al., 2007).

- The different antecedents in the rule, in general, have different values of uncertainty attached to them. As a first step, we aggregate these values into a single CF, using the option considering the strength of the weakest link in a chain as the strength of the chain. This is defined as:  

$$CF_{\text{antecedents}} = \{\text{minimum of CFs of all antecedents}\} \quad (1)$$

- Then this measure (uncertainty for the set of antecedents) is combined with the measure of uncertainty attached to the rule to give a measure of uncertainty for the conclusion of the rule.

$$CF \text{ of the conclusion from rule} = \{CF \text{ associated with rule } R1\} * \{CF_{\text{antecedents}}\}, \text{ provided } CF_{\text{antecedents}} \geq \text{threshold} \quad (2)$$

It can be seen that the CF obtained for a conclusion from a particular rule will always be less than or equal to the CF of the rule. This is consistent with the interpretation of the CF used by MYCIN, that is, the CF of a rule is the CF to be associated with the conclusion if all the antecedents are known to be true with full certainty. In a typical rule based system, there may be more than one rule in the rule base that is applicable for deriving a specific conclusion. Some of them will not contribute any belief to the conclusion, because CF of antecedents is less than the threshold. The contributions from all the other rules for the same conclusion have to be combined. For MYCIN model, initially CF of a conclusion is taken to be 0.0 (i.e. there is no evidence in favour or against) and then as different rules for the conclusion fires, the CF gets updated. MYCIN uses a method that incrementally updates the CF of the conclusion as more evidence for and against is obtained. Let  $CF_{\text{old}}$  be the CF of the conclusion so far, say, after rules  $R_1, R_2, \dots, R_m$  have been fired. Let  $CF_{\text{in}}$  be the CF obtained from firing of another rule  $R_n$ . The new CF of the conclusion (from rules  $R_1, R_2, \dots, R_m$  and  $R_n$ ),  $CF_{\text{new}}$ , is obtained using the formulae given below.

$$CF_{\text{new}} = CF_{\text{old}} + CF_{\text{in}} * (1 - CF_{\text{old}}) \text{ when } (CF_{\text{old}}, CF_{\text{in}} > 0) \quad (3)$$

$$CF_{\text{new}} = CF_{\text{old}} + CF_{\text{in}} * (1 + CF_{\text{old}}) \text{ when } (CF_{\text{old}}, CF_{\text{in}} < 0) \quad (4)$$

$$CF_{\text{new}} = (CF_{\text{old}} + CF_{\text{in}}) / (1 - \min(|CF_{\text{old}}|, |CF_{\text{in}}|)) \text{ otherwise} \quad (5)$$

We adopt this calculus in our model and explained later with a running example in section VI. Before that the concept emotion profiling and the complete framework of emotion recognition system has been introduced.

### IV. CONCEPT OF EMOTION PROFILING

Emotion profile (EP) is introduced for understanding the variation of each feature for different emotional states. This is the core domain of rule based system to classify the target emotions. We define the emotion profile as the degree by which a given feature could reasonably differentiate among target emotions. If E is denoted as set of emotions and  $2E$  is the set of subsets of emotions, then EP of feature ( $F_i$ ) is defined as

$$EP(F_i) = \{X_i | X_i \in 2E; i = 1, 2, \dots, N\}$$

such that all elements of E occurs once and only once in the emotion profile.

There could be two extreme scenarios mentioned below as 'worst scenario' and 'best scenario'.

$$EP(F_i) = \{\{E_1, E_2, E_3, E_4, E_5, E_6, \dots, E_N\}\}$$

represents 'worst scenario' because the feature  $F_i$  is not able to differentiate between any of the target emotions. This is normally due to the variation in the feature value being independent of the emotional state, and generally means the feature is not a useful one for this purpose.

$$EP(F_i) = \{\{E_1\}, \{E_2\}, \{E_3\}, \{E_4\}, \{E_5\}, \dots, \{E_N\}\}$$

represents 'best scenario' as the feature  $F_i$  is strong enough to differentiate between every individual emotions. For example, if feature  $f_1$  (distance between nose and lip) observed to differentiate the emotional states 'disgust' and 'happy' but not able to differentiate between 'fear' and 'sad' and 'anger' from 'neutral' (as their range of values are very close) then we represent emotion profile of the feature  $f_1$  as

$EP(F_1) = \{\{H\}, \{D\}, \{F, S\}, \{A, N\}\}$  where D, H, F, S, A and N stands for 'disgust', 'happy', 'fear', 'sad', 'neutral', respectively.

This is further validated by certain rules (illustrated in section VI). As all the features considered for our problems of emotion recognition are numeric in nature, we considered the average value as the final value to define the range and hence to understand the partition between emotions. This process is very much useful in finding the useful set of features. Relevant set of features acts as an ingredient for emotion classification problem. The next section will illustrate the complete process with a concrete example of emotion recognition using facial expressions.

## V. GENERAL FRAMEWORK FOR EMOTION RECOGNITION

The conceptual framework for emotion recognition includes preprocessing, feature extraction, feature analysis, selection of the features, formulation of rules and measuring performance to classify the target emotional states. This will be explained using facial expression as an input in next section.

### a) Preprocessing and Feature Extraction

The objective of preprocessing is to make the input data in a standard format and suitable for extracting the desired features. Feature extraction involves identifying relevant features and formulating algorithms to extract these features from their respective input data.

### b) Feature Analysis and Emotion Profiling

Once the basic feature set is ready, the next step is analysis of these features. The question, 'how does each of these features vary with the emotion'

needs to be answered here. Each feature has been analyzed carefully by looking its emotion profile respectively. Usually all features don't contribute to the same extent to recognize different emotional states.

### c) Formulation of Rules Using Features

Influential and useful features can be used to define rules, as follows:

- Emotion profile had been created for each feature to analyze its ability to distinguish among the target emotional states, and accordingly useful features were shortlisted.
- Rules are formed using each of these features for different target emotional states. A feature may yield one or more rules. Generally these rules have the form: if feature  $F_1$  has value less than  $T_1$  and feature  $F_1$  has value greater than  $T_2$  then conclude emotion =  $e_1$ . For each rule, the cut off points  $T_1$  and  $T_2$  for a given emotion class is taken to be the approximate average of the value of that emotion with its immediate emotion neighbor.
- To each rule, we associate CF values for each emotional class. These values of CFs are decided as per guidelines mentioned in Table 1.

There may be multiple rules associated with each feature. Multiple rules when fired simultaneously (based on values of different features) may saturate the values of CF associated with them. To minimize this possibility, we have chosen relatively lower range of CF values. Given our observation that most features do not provide a high degree of discrimination for any of the target emotion, a high value did not appear justified for any individual feature. The chosen range also allows the CF value to climb steadily to a high range, when there are many features supporting an emotion. The rules may point to a specific emotional state or a set of emotional states. If the distance of an emotion with its neighboring emotion is found to be less than 5% to 6% of the entire spread (overall range i.e. difference between upper value and lower value) for that features value, then these emotions are grouped as a subset. Allocation of these CF value to the target classes is done based on the three interclass rules (IR-1, IR-2 and IR-3). This is derived based on analysis of the emotion profile.

#### IR-1 (High Interclass Distance):

If the interclass distance of an emotional class (either singleton or non-singleton) with its neighbors (left side as well as right side) is more than 15% of the entire spread for that feature, then the chances of a confusion with the neighboring class is low and hence the CF value associated with this class for that feature is considered to be 0.3.

#### IR-2 (Medium Interclass Distance):

If the interclass distance of a emotional class (either singleton or non-singleton) with its neighbors (left side as well as right side) is in between 6% to 15% of the

entire spread for that feature, then the CF value associated with this class is considered to be 0.2.

#### IR-3 (Low Interclass Distance):

If the interclass distance of a emotional class (either singleton or non-singleton) with its neighbors (left side as well as right side) is less than 6% of the entire spread for that feature, then the CF value associated with this class is considered to be 0.1.

#### d) Recognizing Emotions a using Rules

Overall system's performance for recognizing emotions was measured with the final value of CF corresponding to all the emotional states for all images in the test set. The highest value of final CF is considered.

## VI. CASE STUDY FOR FACIAL EXPRESSION

The standard database, Cohn-Kanade (CK) (Kanade, et al., 2000) of the static images have been used, where individuals are constrained to look straight at the camera and they are photographed with single colored background and illumination conditions do not vary drastically. Therefore, preprocessing issues are not a concern here. Total of 184 images from 57 subjects (32 female and 25 male subjects) have been selected for the emotional states of neutral, anger, happy, fear, sad, and disgust.

Table 1: Defining Certainty Factor (CF) for Rules			
Range of the CF	CF Values	Belief and Disbelief	Indicated by
Greater than 0.2 and up to 0.4	0.3	High evidence	High Inter class distance
Greater than 0.1 and up to 0.2	0.2	Moderate evidence	Medium Inter class distance
Equal to 0.1	0.1	Low evidence	Low Inter class distance

#### a) Feature Extraction

The frontal view face model (Pantic & Rothkrantz, 2000b) is composed of many elements like mouth, nose, eyes and brows (Figure 1 and Table 2). By using a set of 18 points in the frontal view image, total of 21 features ( $f_3, f_4, f_5, f_6, f_7, f_8, f_9, f_{10}, f_{11}, f_{12}, f_{13}, f_{14}, f_{15}, f_{16}, f_{17}, f_{19}, f_{20}, f_{21}, f_{22}, f_{23}, f_{24}$ ) as shown in Figure 1, mostly in the form of inter-point distances had been extracted. For example, the feature  $f_3$  is the distance between left eye

outer corner, A to left eyebrow outer corner, E. Similarly feature  $f_4$  (symmetrical to  $f_3$ ) is the distance between right eye outer corners,  $A_1$  to right eyebrow outer corner,  $E_1$ . Each of these points has been extracted from the image. The distances are compiled and are used for further analysis. All these distances were obtained for different emotions including the neutral state for all subjects. Facial expressions are often characterized by variation of a feature from its value in the neutral state, rather than its absolute value in a given state.

Therefore, standardization of these features w.r.t their neutral value was done. These parameters were normalized in the following manner:

$$\text{Normalized Value} = (\text{Measured Value} - \text{Neutral State Value}) / \text{Neutral State Value} \quad (6)$$

Hence forth, in the remaining paper the reference made to use these normalized values as feature value as an input variable.

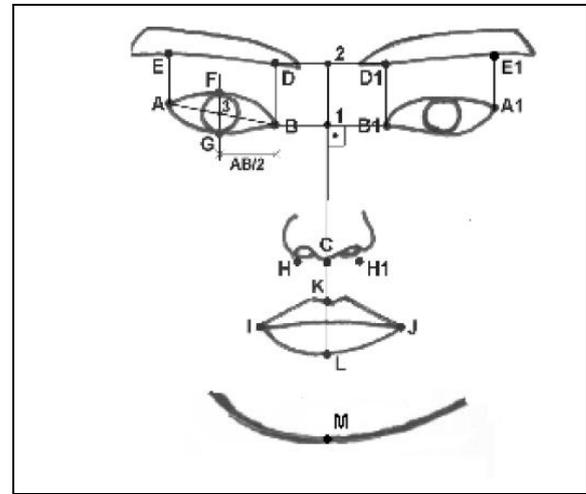


Figure 1 : Facial points (Pantic & Rothkrantz, 2000b)

#### b) Feature Analysis and Emotion Profiling

As discussed earlier all features might not be useful in forming the rules. Individually each of these has to be analyzed. For example, the feature, lip distance (horizontal distance-  $f_{16}$  and vertical distance-  $f_{17}$ ) could be seen as varying with emotions (Figure 2 and Figure 3).

The emotion profile of these feature (i.e.  $f_{16}$  and  $f_{17}$ ) are represented as

$$EP(f_{16}) = \{\{A, D\}, \{N\}, \{S\}, \{F\}, \{H\}\}$$

$$EP(f_{17}) = \{\{A\}, \{S\}, \{N, D\}, \{F\}, \{H\}\}$$

The lip movement (horizontal lip distance,  $f_{16}$  and vertical lip distance,  $f_{17}$ ) provides good separation between 'happy', 'sad', 'fear' w.r.t 'neutral' state (as per table-1). The emotions 'anger' and 'disgust' appear to be very close with each other for  $f_{16}$ . But the feature  $f_{17}$  is able to discriminate 'anger', 'sad', 'fear' and 'happy' but 'disgust' if found to be in the vicinity of 'neutral'. This

exercise is done for all features. Few observations are as follows:

Symmetrical pairs of features (like left eye vertical distance,  $f_9$  and right eye vertical distance,  $f_{10}$ ) do not always have the same emotion profile. For example,

$f_9$  clearly differentiates between 'disgust' and 'fear', but doesn't show a reasonable separation between other pairs of emotions e.g. {'anger', 'happy'} and {'neutral', 'fear'}.

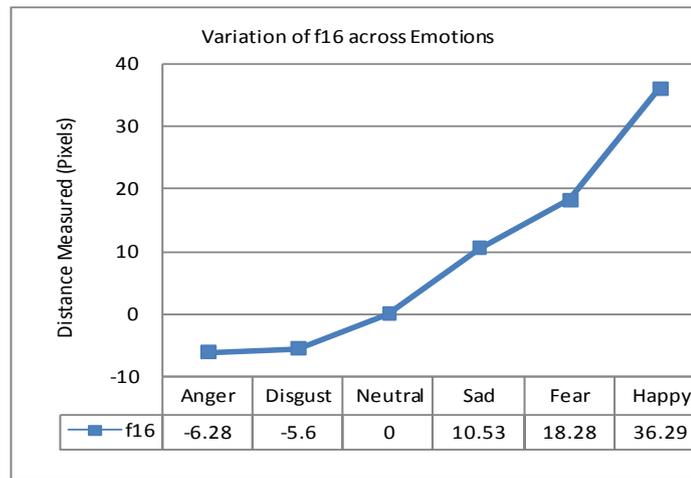


Figure 2 : Variation of feature  $f_{16}$  across emotions

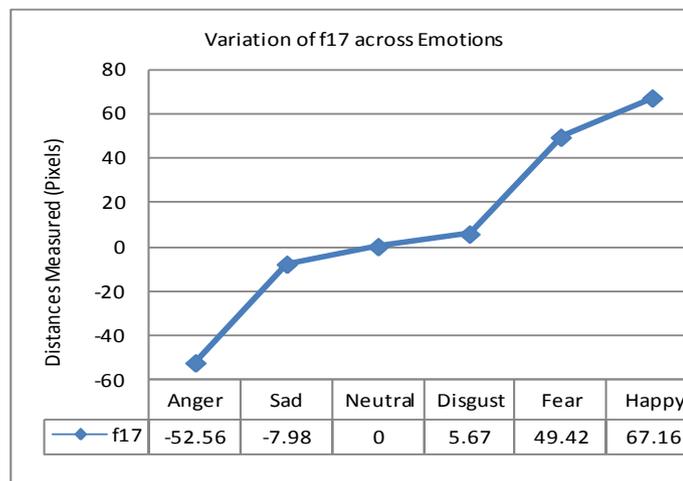


Figure 3 : Variation of feature  $f_{17}$  across emotions

The feature  $f_{10}$  differentiates reasonably well between all the target emotions. The same is true with  $f_{12}$  feature (distance between left lip and left eye). But  $f_{13}$  feature (distance between right lip and right eye) differentiate a cluster of emotion {'happy', 'disgust'} and {'sad', 'fear'}.

The symmetrical features,  $f_{12}$  and  $f_{13}$  show the same results for 'neutral' and 'anger' only. It is observed that total of eleven features (i.e.  $f_3, f_4, f_9, f_{10}, f_{11}, f_{12}, f_{13}, f_{14}, f_{15}, f_{16}$ , and  $f_{17}$ ) shows significant variation across the target emotional states among all twenty one features. Hence these 11 features will be most relevant and useful in designing rules further for recognizing emotions.

TABLE 2: Features of the facial points (Pantic & Rothkrantz, 2000b)

Features	Feature Description
$f_3$	Distance AE
$f_4$	Distance $A_1E_1$
$f_5$	Distance 3F, 3 is the centre of AB (See Figure 1)
$f_6$	Distance $4F_1$ , 4 is the centre of $A_1B_1$ (See Figure 1)

$f_7$	Distance 3G
$f_8$	Distance $4G_1$
$f_9$	Distance FG
$f_{10}$	Distance $F_1G_1$
$f_{11}$	Distance CK, C is $0.5HH_1$
$f_{12}$	Distance IB
$f_{13}$	Distance $JB_1$
$f_{14}$	Distance CI
$f_{15}$	Distance CJ
$f_{16}$	Distance IJ
$f_{17}$	Distance KL
$f_{19}$	Image intensity in circle ( $r(0.5BB_1)$ , C(2)) above line (D, $D_1$ )
$f_{20}$	Image intensity in circle ( $r(0.5BB_1)$ , C(2)) below line (D, $D_1$ )
$f_{21}$	Image intensity in circle ( $r(0.5AB)$ , C(A)) left from line (A, E)
$f_{22}$	Image intensity in circle ( $r(0.5A_1B_1)$ , C( $A_1$ )) right from line ( $A_1$ , $E_1$ )
$f_{23}$	Image intensity in the left half of the circle ( $r(0.5BB_1)$ , C(I))
$f_{24}$	Image intensity in the right half of the circle ( $r(0.5BB_1)$ , C(J))
Total	21 Features

### c) Formulation of Rules

From the trend of feature  $f_{16}$  (Figure 2), it is seen that the emotions 'neutral', 'sad', 'fear' and 'happy' are distinguishable individually, whereas the emotions, 'disgust' and 'anger' are found to be close together (as the distances with its neighbour are found to be in the range of 5% to 6% of the entire spread). Depending on the interclass distances of these classes CFs has been allocated (as per Table 1) and rules have been formed. For each rule (of the type if – then), the cutoff point (i.e., upper limit,  $T_2$  and lower limit,  $T_1$ ) belonging to the emotion class is taken to be the average of the value of that class with its immediate emotional class. From the figure 2, it is clear that for 'sad' emotion the cutoff points (i.e.  $T_1$  and  $T_2$ ) to be considered are 5 and 14, forming

the singleton class and due to high inter class distances the CF values is to be considered as 0.3 (see Table 1). Similarly, the feature  $f_{17}$  also varies across emotions (Figure 3). It is observed that 'neutral' along with 'disgust' is forming a non-singleton class while rest of the emotions is acting as singleton classes. It is observed that for 'sad' emotion the cutoff points (i.e.  $T_1$  and  $T_2$ ) to be considered are -30 and -3.

Depending on distances between these classes, CFs has been allocated and rules have been formed. We found a total of five conditions each for the feature  $f_{16}$  and feature  $f_{17}$  to classify emotions. Examples of rules (Rule 1 and Rule 2) are shown below.

*Example Rule 1:* Using  $\text{dist\_horizontal\_lip}$  ( $f_{16}$ ) for emotion identification

- (i) if ( $\text{dist\_horizontal\_lip} \leq -3$ )  
 $CF_{Dis}=0.2; CF_{Ang}=0.2;$
- (ii) if ( $(\text{dist\_horizontal\_lip} > -3) \ \&\& \ (\text{dist\_horizontal\_lip} \leq 5)$ )  
 $CF_{Neu}=0.3;$
- (iii) if ( $(\text{dist\_horizontal\_lip} > 5) \ \&\& \ (\text{dist\_horizontal\_lip} \leq 14)$ )  
 $CF_{Sad}=0.3;$
- (iv) if ( $(\text{dist\_horizontal\_lip} > 14) \ \&\& \ (\text{dist\_horizontal\_lip} \leq 27)$ )  
 $CF_{Fear}=0.3;$
- (v) if ( $\text{dist\_horizontal\_lip} > 27$ )  
 $CF_{Hap}=0.3;$

*Example Rule 2:* Using  $\text{dist\_vertical\_lip}$  ( $f_{17}$ ) for emotion identification

- (i) if ( $\text{dist\_vertical\_lip} < -30$ )  
 $CF_{Ang}=0.3;$
- (ii) if ( $(\text{dist\_vertical\_lip} < -3) \ \&\& \ (\text{dist\_vertical\_lip} \geq -30)$ )  
 $CFSad=0.2;$
- (iii) if ( $(\text{dist\_vertical\_lip} < 27) \ \&\& \ (\text{dist\_vertical\_lip} > -3)$ )  
 $CF_{Neu}=0.3; CF_{Dis}=0.3;$
- (iv) if ( $(\text{dist\_vertical\_lip} \geq 27) \ \&\& \ (\text{dist\_vertical\_lip} < 58)$ )  
 $CFFear=0.3;$
- (v) if ( $\text{dist\_vertical\_lip} \geq 58$ )  
 $CF_{Hap}=0.3;$

TABLE 3: Examples of computed values of CF using rules from face for female subject

Updated Value of CF computed using rules for respective emotion							
Subject	Actual Emotion	CF <sub>Sad</sub>	CF <sub>Neutral</sub>	CF <sub>Angry</sub>	CF <sub>Happy</sub>	CF <sub>Fear</sub>	CF <sub>Disgust</sub>
s <sub>1</sub>	S	0.83	0.56	0.30	0.36	0.72	0.00
s <sub>1</sub>	N	0.00	0.97	0.00	0.00	0.30	0.00
s <sub>1</sub>	A	0.10	0.37	0.91	0.20	0.50	0.51
s <sub>1</sub>	H	0.30	0.30	0.30	0.82	0.36	0.30
s <sub>1</sub>	F	0.78	0.37	0.00	0.20	0.84	0.00
s <sub>1</sub>	D	0.00	0.00	0.85	0.51	0.20	0.87

Such kind of exercise is done for each of the selected features. Symmetrical pair of features like ( $f_3$ ,  $f_4$ ), ( $f_9$ ,  $f_{10}$ ), ( $f_{12}$ ,  $f_{13}$ ) and ( $f_{14}$  and  $f_{15}$ ) do not vary in the same way across different emotions and hence the resulting rules may differ. Total of 11 rules have been formed for emotion identification using facial static images.

#### d) Recognizing Emotions using Rules

All these rules have been tested on the database and final value of CF has been computed corresponding to each of the 6 emotional states. The emotion with the highest value of final CF is considered and counted against the expected emotion class for each image for all the subjects. For example, Table 3 shows the computed values of CF corresponding to all the six emotions - sad (S), neutral (N), anger (A), happy (H), fear (F) and disgust (D).

A row in this table indicates an input image of an individual subject ( $s_1$ ) in a particular emotional state. Final outcome for the same is indicated in these CF values under the six columns labelled from CF<sub>Sad</sub> to CF<sub>Disgust</sub>. For example, row 3 corresponds to subject-1 (i.e.  $s_1$ ) in 'angry' state; the table shows the maximum value of CF under the emotion class of 'anger' (i.e. 0.91) showing correct identification. Similarly, the maximum value of CF for the subject-1 (i.e.  $s_1$ ; row 6) is 0.87 and is for the target emotion of disgust. Though the value belonging to 'anger' is coming close to this value, we are considering the highest value of CF to identify the target emotion associated with the input image. Hence, the computed emotion matches with the 'predicted emotion' which is 'disgust' in this case and 'anger' in the previous case. Similarly computed value of CF has been analyzed for each of the emotions. Table 4 shows the overall recognition results for each emotion class using confusion matrices.

The literature discusses results for various face based emotion classification systems (Azcarate et al., 2005; Zhao & Kearney, 1996; Fasel & Luetttin, 2003; Pantic & Rothkrantz, 2000a; Sebe et al., 2007; Pantic & Rothkrantz, 2000b, Kobayashi & Hara, 1992; Edwards et al., 1998; Lyons et al., 1999; Huang & Huang, 1997; Hong et al., 1998 and Kulkarni et al., 2009). The average expression recognition rate of all of these systems is around 82% (in the range of 64% to 100%). Some of these studies have used limited testing data for training and for testing.

In comparison, the overall correctness of recognizing emotions using our rule based approach from facial expression is found to be 86.43%. The recognition rates are found to be 80% and 88.89% for female and male subjects respectively. Recognition rate of 'anger' and 'fear' is high for male subjects as compared to female subjects. For example, it is observed that rate of recognition for 'anger' is coming out to be 100% for male and 69% for female.

This rule based approach could be extended to any other modalities easily as it is based on the set of rules which could be extracted from different modalities (e.g. facial expression, speech or others). The overall process remains same i.e. to design the rules all the relevant features needs to be studied in more detail in the similar fashion. Emotion Profiling of each feature acts as an important ingredient as it is useful to map the relevant feature set for target emotional states. Influential and useful features were selected for defining the rules. Performance of the system could be improved by modifying, adding and deleting rules.

**TABLE 4: Confusion Matrices for Gender Independent Case**

	S	Z	A	H	F	D
Sad (S)	8	1	1	0	2	1
Neutral (N)	0	57	0	0	0	0
Anger (A)	1	1	20	0	0	3
Happy (H)	0	1	1	36	2	2
Fear (F)	1	0	0	1	10	0
Disgust (D)	0	2	2	2	1	28

## VII. CONCLUSION AND FUTURE WORK

Emotion is assuming increasing importance in HCI, in general, with the growing feeling that emotion is central to human communication and intelligence. While various aspects of this problem have been addressed in the literature, the full problem has not received much attention so far. The primary concern in emotion recognition is inaccurate knowledge and data. There are hardly any features or feature combinations which can infer any emotion to complete the certainty. In general, there are no features that are universally effective for recognizing all emotions. There are some features which provide reasonable discrimination among various subsets of emotions. Hence the concept of 'emotion profile' is useful for extensive analysis and evaluation of individual features. We used the confirmation theory as used in MYCIN system where the values of CF are allocated to the emotional classes based on interclass distances. These have been derived based on the analysis of the emotion profile of individual features. Rule based systems have certain advantages. Because of the uniform syntax, each rule can be easily analyzed. The syntax is usually quite simple, so it is easy to understand the rules without an explicit translation. Rules could be considered as independent pieces of knowledge about the domain and this independence leads to a high degree of modularity. Performance of the system could be improved by modifying / adding / deleting rules. This rule based system is applicable for any modalities like speech, gesture, facial expressions, etc. if provided with set of features. To validate further a study was done on speech and keyboard usage modality using the above mention rule based system.

Given the vast scope of the work needed to build reliable emotion recognition system and use the same for enhancing the HCI, and the unavailability and difficulty in collecting reliable datasets for emotion

recognition, this work covers only a part of the journey. A number of aspects require further investigation and refinement. To mention a few limitations against the use of certainty factor is that they have no sound theoretical basis; though, they often work well in practice. We allocated the values of CF to the emotional classes based on heuristic rules as defined in section III. These have been derived based on the analysis of the individual features across different emotions. In this work, we have ignored the possibility of having more than one emotional state at a time. Also the investigation to alternative uncertainty models like the Dempster-Shafer Theory is still open. Dempster Shafer theory provides more flexibility in assigning belief to various subsets of emotions. The databases used for the expression analysis are all based on subjects who "performed" a series of different expressions. There is a significant difference between expressions of a spontaneous and of a deliberate nature. Without a database of spontaneous expressions, the expression analysis system cannot be robust enough. This database issue is common for all the modalities – may be speech, facial expressions, etc. The multimodal data fusion for emotion recognition remains an open challenge as several problems still persist, related to finding optimal features, integration and recognition. Completely automated multimodal emotion recognition system is still at the preliminary phase, shows very limited performance and is mostly restricted to the lab environment.

## REFERENCES RÉFÉRENCES REFERENCIAS

1. Marinez-Miranda, J. & Aldea, A. (2005). Emotions in Human and Artificial Intelligence, in *Computers in Human Behaviour Journal*, R. D. Tennyson (Ed.), 21(2), 323-341.
2. Foley, J.D. (1996). JTEC Panel Report on Human Computer Interaction Technologies in Japan, March 1996, Retrieved from <http://www.wtec.or/pdf/hci.pdf> / on January, 2013.
3. Lisetti, C.L. & Nasoz, F. (2005). Affective intelligent car interfaces with emotion recognition HCI 2005, 11th International Conference on Human Computer Interaction, July 22-27, Las Vegas, USA.
4. Bower, G. (1981). Mood and Memory, *American Psychologist* 36 (2), 129-148.
5. Zajonc, R. (1984). On the Primacy of Affect, *American Psychologist* 39, 117-124.
6. Damasio, A. (1994). *Descartes' Error*, Avon Books, New York, NY.
7. Ledoux, J. (1992). Brain Mechanisms of Emotion and Emotional Learning, *Current Opinion in Neurobiology*, 2, 191-197.
8. Derryberry, D. & Tucker, D. (1992). Neural Mechanisms of Emotion, *Journal of Consulting and Clinical Psychology* 60 (3), 329-337.

9. Colquitt, J.A., LePine, J.A. & Noe, R.A. (2000). Toward an integrative theory of training motivation: a meta-analytic path analysis of 20 years of research, *Journal of Applied Psychology* 85, 678–707.
10. Frijda, N.H. (1986). *The Emotions*, Cambridge University Press, New York.
11. Birdwhistle, R.L. (1970). *Kinesics and Context: Essays on Body Motion and Communication*, University of Pennsylvania Press.
12. Ekman, P. & Friesen, W.V. (1975). *Unmasking the Face: A Guide to Recognizing Emotions from Facial Expressions*, Prentice Hall, Inc, New Jersey.
13. Chovil, N. (1991). Discourse-Oriented Facial Displays in Conversation, *Research on Language and Social Interaction*, 25, 163-194.
14. Goleman, D. (1995). *Emotional Intelligence*, Bantam Books, New York.
15. James, L. & Nahl, D. (2000). *Road Rage and Aggressive Driving: Steering Clear of highway Warfare*, Prometheus Books, Amherst, NY.
16. Khanna, P. & Kumar, S. (2010). Recognizing Emotions from Human Speech, *Think Quest*, International Conference on Contours of Computing Technology in association with Springer Publications, Mumbai India.
17. Khanna, P. and Kumar, S. (2011). Application of Vector Quantization in Emotion Recognition from Human Speech, *Springe Series in Communications in Computer and Information Science (CCIS)*, ICISTM, 118-125.
18. Khanna, P. & Kumar, S. (2010). Recognizing Emotions from Keyboard Stroke Pattern, *International Journal of Computer Applications*, 11(9).
19. Picard, R.W. (1997). *Affective computing*, The MIT Press, Cambridge, MA.
20. De Silva & Ng (2000). Bimodal Emotion Recognition, *Automatic Face and Gesture Recognition*, in IEEE International Conference, 332 – 335.
21. Sebe, N., Cohen, I., Gevers, T., & Huang, T.S. (2006). Emotion Recognition Based, On Joint Visual and Audio Cues”, *Pattern Recognition, International Conference on*, 1, 1136–1139.
22. Zeng, Z., Jilin, Tu., Liu, Huang, Pianfetti, Roth & Levinson. (2007). Audio-Visual Affect Recognition, *IEEE Transactions on multimedia*, 9 (2), 424-428.
23. Kim, J., & André, E. (2006). Emotion recognition using physiological and speech signal in short-term observation, *Perception and Interactive Technologies: LNAI 4201*, 53-64. Springer-Verlag Berlin Heidelberg.
24. Corradini, A., Mehta, M., Bernsen, N. & Martin, J.C. (2003). Multimodal input fusion In Human computer interaction on the example of the on-going nice project, In *Proceedings of the NATO-ASI conference on Data Fusion for Situation Monitoring, Incident Detection, Alert and Response Management*, Yerevan (Armenia).
25. Liao, H. (2002). *Multimodal Fusion*, Master's thesis, University of Cambridge.
26. Kettebekov, S. & Sharma, R. (2000). Understanding Gestures in Multimodal Human Computer Interaction, *International Journal on Artificial Intelligence Tools*, 9(2), 205-223.
27. Sharma, R., Pavlovic, V., & Huang, T. (1998). Toward Multimodal Human Computer Interface, in *Proceedings of the IEEE*, 86(5), 853-860.
28. Kumar, S., Ramani, S., Raman, S.M., Anjaneyulu K.S.R. and Chandrasekar, R. (2007). *Rule Based Expert Systems – A Practical Introduction*, Narosa Publishers.
29. Shortliffe, E.H. & Buchanan, B.G. (1975). A Model of Inexact Reasoning in Medicine, *Mathematical Biosciences*, 23, 351-379.
30. Gordon, J. & Shortliffe, E.H. (1984). The Dempster-Shafer Theory of Evidence, 272-292 in text reference: (Buchanan and Shortliffe, 1984).
31. Negoita, C.V. (1985). *Expert Systems and Fuzzy Systems*, Benjamin/Cummings.
32. Doyle, J.A. (1979). Truth Maintenance System, *Artificial Intelligence*, 12, 231-272.
33. Reiter, R. (1980). A Logic for Default Reasoning, *Artificial Intelligence*, 13, 81-132.
34. McDermott, D. & Doyle, J. (1980). Non-monotonic Logic I, *Artificial Intelligence*, 13, 41-72.
35. Kanade, T., Cohn, J. & Tian, Y. (2000). Comprehensive Database for Facial Expression Analysis, in *Proceedings of the International Conference on Automatic Face and Gesture Recognition*, 46-53.
36. Azcarate, A., Hageloh, F., van de Sande, K. & Valenti, R. (2005). *Automatic facial Emotion recognition*, University of Amsterdam Technical Report.
37. Zhao, J. & Kearney, G. (1996). Classifying Facial Emotions by Back propagation Neural Networks with Fuzzy Inputs, *Proceedings of Conference on Neural Information Processing*, 454-457.
38. Fasel, B. & Luetin, J. (2003). Automatic Facial Analysis A Survey, *Pattern Recognition*, 36, 259-275.
39. Pantic, M. & Rothkrantz, L. (2000a). Automatic Analysis of Facial Expressions: The State of the Art, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 22, 1424-1445.
40. Sebe, N., Lew, M.S., Sun, Y., Cohen, L. Gevers, T., & Huang, T.S. (2007). Authentic Facial Expression Analysis, *Image and Vision Computing*, 25, 1856-1863.
41. Pantic, M. & Rothkrantz, L.J.M. (2000b). Expert System for Automatic Analysis of Facial Expression, *Journal of Image and Vision Computing* 18, 881-905.

42. Kobayashi, H. & Hara, F. (1992). Recognition of Six Basic Facial Expressions and Their Strength by Neural Network, Proceedings of International Workshop Robot and Human Communication, 381-386.
43. Edwards, G.J., Cootes, T.F. & Taylor, C.J. (1998). Face Recognition Using Active Appearance Models, Proceedings of European Conference on Computer Vision, 2, 581-695.
44. Lyons, M.J., Budynek, J. & Akamatsu, S. (1999). Automatic Classification of Single Facial Images, IEEE Transactions on Pattern Analysis and Machine Intelligence, 21, 357-362.
45. Huang, C.L. & Huang, Y.M. (1997). Facial Expression Recognition Using Model Based Feature Extraction and Action Parameters Classification, Journal of Visual Communication and Image Representation 8, 278-290.
46. Hong, H., Neven, H. & Malsburg, C von der. (1998). Online Facial Expression Recognition Based on Personalized Galleries, Proceedings of International Conference on Automatic Face and Gesture Recognition, 354-359.
47. Kulkarni, S.S., Reddy, N.P. & Hariharan, S.I. (2009). Facial expression (mood) recognition from facial images using committee neural networks, Bio Medical Engineering.

