Emotion Recognition from Facial Expressions and Its Control

A Thesis

submitted in partial fulfilment of the requirements for

the Degree of

Master of Technology in Intelligent Automation and Robotics

in the Faculty of Engineering & Technology

Jadavpur University

May 2024

By

Suchismita Mondal Dey

Registration No: 160235 of 2021-2022

Examination Roll No: M6IAR24017

Under the Guidance of

Prof. Amit Konar

Department of Electronics & Telecommunication Engineering

Jadavpur University

Kolkata-700032

India

FACULTY OF ENGINEERING AND TECHNOLOGY

JADAVPUR UNIVERSITY

CERTIFICATE

This is to certify that the dissertation entitled "**Emotion Recognition from Facial Expressions and its Control**" has been carried out by **Suchismita Mondal Dey** (University Registration No: 160235 of 2021-2022, Examination Roll No: M6IAR24017) under my guidance and supervision and accepted in partial fulfilment of the requirement for the Degree of Master of Technology in Intelligent Automation and Robotics. The research results presented in the thesis have not been included in any other paper submitted for the award of any degree to any other University or institute.

_____

Prof. Amit Konar (Supervisor)
Dept. of Electronics & Telecommunication Engineering
Jadavpur University

_____

Prof. Manotosh Biswas
Head of Department
Electronics & Telecommunication Engineering

_____

Prof. Dipak Laha
Dean, Faculty Council of Engineering and Technology
Jadavpur University

FACULTY OF ENGINEERING AND TECHNOLOGY

JADAVPUR UNIVERSITY

CERTIFICATE OF APPROVAL

The forgoing thesis is hereby approved as a creditable study of a Master of Technology subject and presented in a manner satisfactory to warrant acceptance as prerequisite to the degree for which it has been submitted. It is understood that by this approval the undersigned do not necessarily endorse or approve any statement made, opinion expressed or conclusion drawn there in but approve the thesis only for which it is submitted.
Final examination for evaluation of the thesis.

_____

Signature of Examiner

_____

Signature of Supervisor

FACULTY OF ENGINEERING AND TECHNOLOGY

JADAVPUR UNIVERSITY

DECLARATION OF ORIGINALITY AND COMPILANCE
OF
ACADEMIC THESIS

I hereby declare that this thesis entitled "**Emotion Recognition from Facial Expressions and its Control**" contains literature survey and original research work by the undersigned candidate, as part of her Degree of Master of Technology in Intelligent Automation & Robotics.

All information have been obtained and presented in accordance with academic rules and ethical conduct.

I also declare that, as required by these rules and conduct, I have fully cited and referenced all materials and results that are not original to this work.

Name: SUCHISMITA MONDAL DEY

Examination Roll No:

Thesis Title: Emotion Recognition from Facial Expressions

_____

Signature of the candidate

ACKNOWLEDGEMENTS

First and foremost, I would like to express my earnest gratitude and heartfelt indebtedness to my advisor, Prof. Amit Konar, Department of Electronics and Telecommunication Engineering, for the privilege and the pleasure, of allowing me to work under him towards my Degree of Master of Technology in Intelligent Automation and Robotics. This work would not have been materialized, without his whole-hearted help and support. Working under him has been a great experience. I sincerely thank my supervisor, particularly for all the faith he had in me.

I also show my gratitude to my group head Mr. Asok Bandyopadhyay, Associate Director, ICT&Service Group, CDAC – Kolkata and my section head Mrs Barnali Pal, Joint Director, ICT&Service Group, CDAC – Kolkata at my office for giving me the permission to do my Degree of Master of Technology in Intelligent Automation and Robotics. I would also like to thank my all colleagues for their support.

I am indebted to my classmates and friends for their constant support and good wishes who supported me in a number of ways.

Lastly, I would like to thank my Family members for their love, support and guidance through the course work.

_____

SUCHISMITA MONDAL DEY

Examination Roll No:

Date: Jadavpur University

Place: Kolkata - 700032.

List of Content:

CERTIFICATE i
CERTIFICATE OF APPROVAL ii
DECLARATION OF ORIGINALITY AND COMPILANCE OF ACADEMIC THESIS iii
ACKNOWLEDGEMENTS iv
LIST OF CONTENT v
LIST OF FIGURES vi
LIST OF TABLES viii
ABSTRACT x
INTRODUCTION & OBJECTIVE 1
RELATED WORK 7
LITERATURE SURVEY 11
CHAPTER 1 15
Emotion Recognition from Facial Expressions and Its Control Using Fuzzy Logic 15
FILTERING, SEGMENTATION, AND LOCALIZATION OF FACIAL COMPONENTS 15
DETERMINATION OF FACIAL ATTRIBUTES 18
FUZZY RELATIONAL MODEL FOR EMOTION DETECTION 21
EXPERIMENTS AND RESULTS 24
VALIDATION OF THE SYSTEM PERFORMANCE 32
EMOTION TRANSITION AND ITS CONTROL 33
CHAPTER 2 45
General and Interval Type-2 Fuzzy Face-Space Approach to Emotion Recognition 45
PRELIMINARIES ON T2 FUZZY SETS 45
UNCERTAINTY MANAGEMENT IN FUZZY FACE SPACE FOR EMOTION RECOGNITION 46
FILTERING UNWANTED DATA POINTS IN FEATURE SPACE USING INTERVAL APPROACH 60
EXPERIMENTS DETAILS 63
PERFORMANCE ANALYSIS 72
CHAPTER 3 77
Sample run of Programs 77
CHAPTER 4 84
Conclusion & Future Scope 84
REFERENCE 87

LIST OF FIGURES:

LITERATUTE SURVEY

Boolean Logic vs Fuzzy Logic 12

CHAPTER 1

Fig. 1.	**Original face image**	**18**
Fig. 2.	**Median-filtered image.**	**18**
Fig. 3.	**Image after applying FCM clustering.**	**18**
Fig. 4.	**Measurement of MO from the dips in average intensity plot.**	**18**
Fig. 5.	**Synthetic eye template**	**19**
Fig. 6.	**Determination of the EO.**	**20**
Fig. 7.	**Determination of EBC in the selected rectangular patch, identified by image segmentation**	**21**
Fig. 8.	**Structure of a fuzzy relational matrix.**	**24**
Fig. 9.	**Movie clips containing four frames (row-wise) used to excite anxiety, disgust, fear, happiness, and sadness.**	**28**
Fig. 10.	**Arousal of anxiety, disgust, fear, happiness, and sadness using the stimulator in Fig. 9 (female subject).**	**29**
Fig. 11.	**Arousal of anxiety, disgust, fear, happiness, and sadness using the stimulator in Fig. 9 (male subject).**	**30**
Fig. 12.	**Validation of the proposed system by tuning the parameters a, b, m, and σ of the encoders.**	**34**
Fig. 13.	**Proposed emotion transition graph**	**36**
Fig. 14.	**Emotion transition dynamics.**	**38**
Fig. 15.	**Complete scheme of Mamdani-type emotion control.**	**43**
Fig. 16.	**Fig. 16. Submission of audiovisual stimulus of strength 28, 16, and 8 for controlling the emotion of a subject from the disgusted state (leftmost) to a final happy state through sad and anxious states (in that order).**	**45**

CHAPTER 2

Fig. 1.	**Experimental FOU for feature f_i = Mouth-Opening.**	**48**
Fig. 2.	**The IT2FSS-based emotion recognition.**	**53**
Fig. 3.	**General type-2 fuzzy face-space construction for *m* features, *k* emotion classes, and *n* subjects.**	**55**
Fig. 4.	**Computing support of the general type-2 fuzzy FS for emotion class i.**	**58**
Fig. 5.	**GT2FFS-based emotion classification**	**59**
Fig. 6.	**(a) The primary membership function for a given feature and (b) its corresponding secondary membership function obtained by minimizing J_i.**	**62**
Fig. 7.	**(a) Localized eye search region, and (b) detection of eye features.**	**64**
Fig. 8.	**(a) Mouth search area, (b) lip cluster, and (c) graph of average intensity over each row against the row position.**	**65**
Fig. 9.	**Experiment done on different databases: (a) JAFFE, (b) Indian women database (Jadavpur University), (c) Cohn-Kanade database.**	**67**
Fig. 10.	**Membership distributions for emotion anger and feature EOL**	**68**

Fig. 11.	**(a) Secondary membership curve of subject 1. (b) Secondary membership curve of subject 2 for emotion anger.**	**68**
Fig. 12.	**Facial image of an unknown subject**	**69**
Fig. 13.	**Graphical selection of FOU by testing that the point (m_l,m_r) = (0.0755,0.1257) plotted in the figure lies under the triangular zone obtained by satisfying inequalities in (28).**	**71**

LIST OF TABLES:

RELATED WORK

TABLE I	**Comparison of existing surveys on emotion recognition**	**10**

CHAPTER 1

TABLE I	**ASSESSMENT OF THE AROUSAL POTENTIAL OF SELECTED AUDIOVISUAL MOVIE CLIPS IN EXCITING DIFFERENT EMOTIONS**	**26**
TABLE II	**THEME OF THE MOVIE CLIPS CAUSING STIMULATION OF SPECIFIC EMOTIONS**	**28**
TABLE III	**SUMMARY OF RESULTS OF MEMBERSHIP EVALUATION FOR THE EMOTIONS IN FIG. 10**	**31**
TABLE IV	**RESULTS OF PERCENTAGE CLASSIFICATION OF AROUSED EMOTIONS IN SIX CLASSES FOR ADULT MALES**	**31**
TABLE V	**RESULTS OF PERCENTAGE CLASSIFICATION OF AROUSED EMOTIONS IN SIX CLASSES FOR ADULT FEMALES**	**31**
TABLE VI	**RESULTS OF PERCENTAGE CLASSIFICATION OF AROUSED EMOTIONS IN SIX CLASSES FOR CHILDREN IN AGE GROUP 8–12 YEARS**	**32**
TABLE VII	**INDEXING OF EMOTIONAL STATES**	**39**
TABLE VIII	**MEMBERSHIP FUNCTIONS OF MAGNITUDE OF ERROR AND UNSIGNED STRENGTH OF POSITIVE/NEGATIVE INFLUENCE**	**40**
TABLE IX	**MEMBERSHIP FUNCTIONS OF SIGN OF ERROR**	**40**
TABLE X	**SELECTED PARAMETERS OF THE MEMBERSHIP FUNCTIONS**	**41**

CHAPTER 2

TABLE I	**TRAINING AND TEST DATA FOR THREE DATABASES**	**67**
TABLE II	**EXTRACTED FEATURES OF FIG. 13**	**69**
TABLE III	**CALCULATED TYPE-2 PRIMARY MEMBERSHIP VALUES FOR THE FEATURE: EOL UNDER EMOTION: DISGUST**	**69**
TABLE IV	**CALCULATED RANGES OF PRIMARY MEMBERSHIPS AND CENTRE VALUE FOR EACH EMOTION**	**69**
TABLE V	**RESULTS OF EXECUTION OF IA ON FEATURE EOL DATA SET FOR EMOTION: ANGER**	**70**
TABLE VI	**CALCULATED TYPE-2 MEMBERSHIP VALUES FOR THE FEATURE: EOL UNDER EMOTION: DISGUST**	**72**
TABLE VII	**CALCULATED RANGES OF PRIMARY MEMBERSHIP CENTRE VALUE FOR EACH EMOTION**	**72**

TABLE VIII	**PERCENTAGE ACCURACY OF OUR PROPOSED METHODS OVER THREE DATABASES**	**73**
TABLE IX	**STATISTICAL COMPARISON OF PERFORMANCE USING MC NEMAR'S TEST WITH THREE DATABASES**	**75**
TABLE X	**AVERAGE RANKING OF CLASSIFICATION ALGORITHMS BY FRIEDMAN**	**75**

Abstract

Facial expressions of a person representing similar emotion are not always unique. Naturally, the facial features of a subject taken from different instances of the same emotion have wide variations. In the presence of two or more facial features, the variation of the attributes together makes the emotion recognition problem more complicated. This variation is the main source of uncertainty in the emotion recognition problem, which has been addressed here in two steps using type-2 fuzzy sets. First a type-2 fuzzy face space is constructed with the background knowledge of facial features of different subjects for different emotions. Second, the emotion of an unknown facial expression is determined based on the consensus of the measured facial features with the fuzzy face space. Both interval and general type-2 fuzzy sets (GT2FS) have been used separately to model the fuzzy face space. The interval type-2 fuzzy set (IT2FS) involves primary membership functions for m facial features obtained from n-subjects, each having l-in- stances of facial expressions for a given emotion. The GT2FS in addition to employing the primary membership functions mentioned above also involves the secondary memberships for individual primary membership curve, which has been obtained here by formulating and solving an optimization problem. The optimization problem here attempts to minimize the difference between two decoded signals: the first one being the type-1 defuzzification of the average primary membership functions obtained from the n-subjects, while the second one refers to the type-2 defuzzified signal for a given primary membership function with secondary memberships as unknown. The uncertainty management policy adopted using GT2FS has resulted in a classification accuracy of 98.333% in comparison to 91.667% obtained by its interval type-2 counterpart. A small improvement (approximately 2.5%) in classification accuracy by IT2FS has been attained by pre-processing measurements using the well-known interval approach.

This thesis presents a fuzzy relational approach to human emotion recognition from facial expressions and its control. The proposed scheme uses external stimulus to excite specific emotions in human subjects whose facial expressions are analysed by segmenting and localizing the individual frames into regions of interest. Selected facial features such as eye opening, mouth opening, and the length of eyebrow constriction are extracted from the localized regions, fuzzified, and mapped onto an emotion space by employing Mamdani-type relational models. A scheme for the validation of the system parameters is also presented. This paper also provides a fuzzy scheme for controlling the transition of emotion dynamics toward a desired state. Experimental results and computer simulations indicate that the proposed scheme for emotion recognition and control is simple and robust, with good accuracy.

INTRODUCTION & OBJECTIVE

HUMANS often use nonverbal cues such as hand gestures, facial expressions, and tone of the voice to express feelings in interpersonal communications. Unfortunately, currently available human–computer interfaces do not take complete advantage of these valuable communicative media and thus are unable to provide the full benefits of natural interaction to the users. Human–computer interactions could significantly be improved if computers could recognize the emotion of the users from their facial expressions and hand gestures, and react in a friendly manner according to the users' needs and preferences [4].

The phrase *affective computing* [30] is currently gaining popularity in the literature of human–computer interfaces [35], [46]. The primary role of affective computing is to monitor the *affective states* of people engaged in critical/accident-prone environments to provide assistance in terms of appropriate alerts to prevent accidents. Li and Ji [25] proposed a probabilistic framework to dynamically model and recognize the users' affective states so as to provide them with corrective assistance in a timely and efficient manner. Picard *et al.* [31] stressed the significance of human emotions on their affective psychological states. Rani *et al.* [32] presented a novel scheme for the fusion of multiple psychological indices for real-time detection of a specific affective state (anxiety) of people using fuzzy logic and regression trees, and compared the relative merits of the two schemes. Among the other interesting applications in affective computing, the works of Scheirer *et al.* [35], Conati [7], Kramer *et al.* [22], and Rani *et al.* [32], [33] deserve special mention.

Apart from human–computer interfaces, emotion recognition by computers has interesting applications in computerized psychological counselling and therapy, and in the detection of criminal and antisocial motives. The identification of human emotions from facial expressions by a machine is a complex problem for the following reasons. First, identification of the exact facial expression from a blurred facial image is not an easy task. Second, segmentation of a facial image into regions of interest is difficult, particularly when the regions do not have significant differences in their imaging attributes. Third, unlike humans, machines usually do not have visual perception to map facial expressions into emotions.

Very few works on human emotion detection have so far been reported in the current literature on machine intelligence. Ekman and Friesen [9] proposed a

scheme for the recognition of facial expressions from the movements of cheek, chin, and wrinkles. They have reported that there exist many basic movements of human eyes, eyebrows, and mouth that have direct correlation with facial expressions. Kobayashi and Hara [18]–

[20] designed a scheme for the recognition of human facial expressions using the well-known backpropagation neural algorithms [16], [39]–[41]. Their scheme is capable of recognizing six common facial expressions depicting happiness, sadness, fear, anger, surprise, and disgust. Among the well-known methods of determining human emotions, Fourier descriptor [40], template matching [2], neural network models [11], [34], [40], and fuzzy integral [15] techniques deserve special mention. Yamada [45] proposed a new method of recognizing emotionsthrough the classification of visual information. Fernandez Dols *et al.* proposed a scheme for decoding emotions from facial expression and content [12]. Kawakami *et al.* [16] analysed in detail the scope of emotion modelling from facial expressions. Busso and Narayanan compared the scope of facial expressions, speech, and multimodal information in emotion recognition [4]. Cohen *et al.* [5], [6] considered temporal variations in facial expressions, which are displayed in live video to recognize emotions. She proposed a new architecture of hidden Markov models to automatically segment and recognize facial expressions. Gao *et al.* [13] presented a methodology for facial expression recognition from a single facial image using line- based caricatures. Lanitis *et al.* [23] proposed a novel technique for the automatic interpretation and coding of face images using flexible models. Examples of other important works on the recognition of facial expression for conveying emotions include [3], [8], [10], [24], [26], [34], [37], [38], and [44].

This thesis provides an alternative scheme for human emotion recognition from facial images, and its control, using fuzzy logic. Audio visual stimulus is used to excite the emotions of subjects, and their facial expressions are recorded as video movie clips. The individual video frames are analysed to segment the facial images into regions of interest. Fuzzy C-means (FCM) clustering [1] is used for the segmentation of the facial images into three important regions containing mouth, eyes, and eyebrows. Next, a fuzzy reasoning algorithm is invoked to map fuzzified attributes of the facial expressions into fuzzy emotions. The exact emotion is extracted from fuzzified emotions by a denormalization procedure similar to defuzzification (fuzzy decoding). The proposed scheme is both robust and insensitive to noise because of the nonlinear mapping of image attributes to emotions in the fuzzy domain. Experimental results show that the detection

accuracies of emotions for adult male, adult female, and children of 8–12 years are as high as 88%, 92%, and 96%, respectively, outperforming the percentage accuracies of the existing techniques [26], [43]. This work alsoproposes a scheme for controlling emotion [36] by judiciously selecting appropriate audio visual stimulus for presentation be-fore the subject. The selection of the audio visual stimulus is undertaken using fuzzy logic. Experimental results show that the proposed control scheme has good experimental accuracy and repeatability.

EMOTION recognition is currently gaining importance for its increasing scope of applications in human–computer interactive systems. Several modalities of emotion recogni- tion, including facial expression, voice, gesture, and posture have been studied in the literature. However, irrespective of the modality, emotion recognition comprises two fundamental steps involving feature extraction and classification [78]. Fea- ture extraction refers to determining a set of features/attributes, preferably independent, which together represents a given emo- tional expression. Classification aims at mapping emotional features into one of several emotion classes.

Performance of an emotion recognition system greatly de- pends on feature selection and classifier design. A good clas- sification algorithm sometimes cannot yield high classification accuracy for poorly selected features. On the other hand, even using a large set of features, describing an emotion, we oc- casionally fail to recognize the emotion correctly because of a poor classifier. Most commonly used techniques for feature selection in the emotion recognition problem include principal component analysis (PCA) [101], independent component anal- ysis [102], rough sets [84], [103], Gabor filter [62], and Fourier descriptors [67]. Among the popularly used techniques for emotion classification, neural net-based mapping [45], [46], [60], fuzzy relational approach [56], linear discriminate analysis [60], support vector machine (SVM) [50], and hidden Markov model [101], [104] need special mention. A brief overview of the existing research on emotion recognition is given next.

Ekman and Friesen took an early attempt to recognize facial expression from the movements of cheek, chin, and wrinkles [66]. Their experiments confirmed the existence of a good correlation between basic movements of the facial action units [55], [61] and facial expressions [43], [44], {47], [49], [52],[61]–[64]. Kobayashi and Hara [57]–[59] designed a scheme for the recognition of human facial expressions using the well-known back-propagation neural networks [80], [85]. Their scheme is capable of recognizing six common facial

expressions depicting happiness, sadness, fear, anger, surprise, and disgust. Yamada proposed an alternative method of emotion recognition through classification of visual information [91].

Fernandez-Dols *et al.* proposed a scheme for decoding emotions from facial expressions and content [92]. Kawakami *et al.* [85] designed a method for the construction of emotion space using neural networks. Busso and Narayanan [93] analyzed the scope of facial expressions, speech, and multi-modal information in emotion recognition. Metallinou *et al.* [113] employed content-sensitive learning for audio-visual emotion recognition. In [115], Metallinou *et al.* proposed a novel approach to visual emotion recognition using a compact representation of face and viseme information. In [116], Metallinou *et al.* presented an approach to decision level fusion for handling multi-modal information in emotion recognition. Lee *et al.* [117] employed a hierarchical binary tree for emotion recognition. Mower *et al.* designed an interesting scheme about human perception of audio-visual synthetic emotion character in the presence of conflicting information [118]. Cohen *et al.* [94] developed a scheme for emotion recognition from the temporal variations in facial expressions obtained from the live video sequence of the subjects. They used hidden Markov model to automatically segment and recognize facial expression. Gao *et al.* presented a scheme for facial expression recognition from a single facial image using line based caricatures [95]. Among other significant contributions in emotion recognition, the works presented in [48], [50], [51], [53], [54], [57]–[59], [65]–[70], [72]–[74], [77],

[82], [88], [98], [99], [102], [112], [114], [119]–[122] need special mention. For a more complete literature survey, which cannot be given here for space restriction, readers may refer to two outstanding works by Pantic*etal.* [99], [109].

Emotional features greatly depend on the psychological states of the subjects. For example, facial expressions of a subject, while experiencing the same emotion, have wider variations, resulting in significant changes in individual feature. Further, different subjects experiencing the same emotion have differences in their facial features. Repeated experiments with a large number of subjects, each having multiple instances of similar emotional experience, reveal that apparently there exists a small but random variation of facial features around specific fixed points [107]. The variation between different instances of facial expression for similar emotive experience of an individual can be regarded

as an *intra-personal level uncertainty* [83]. On the other hand, the variation in facial expression of individuals for similar emotional experience can be treated as *inter- personal level uncertainty* [83].

The variations in features can be modeled with fuzzy sets. Classical (type-1 (T1)) fuzzy sets, pioneered by Zadeh [108], have widely been used over the last five decades for modeling uncertainty of ill-defined systems. T1 fuzzy sets employ a single membership function to represent the degree of uncertainty in measurements of a given feature. Hence, it can capture the variation in measurements of a given feature for different instances of a specific emotion experienced by a subject. In [56], the authors have considered a fixed membership function to model the uncertainty involved in a feature for a given emotion, disregarding the possibility of variation in the membership curves for different subjects.

This thesis, however, models the above form of inter-personal level uncertainty by interval type-2 (T2) fuzzy sets (IT2FS). IT2FS employs an upper and a lower membership function (UMF and LMF) to capture the uncertainty involved in a given measurement of a feature within the bounds of its two membership curves at the point of the measurement. However, the degree of correct assignment of membership for each membership curve embedded between the UMF and LMF in IT2FS is treated as unity, which is not always appropriate.

General T2 fuzzy set (GT2FS) can overcome the above problem by considering a secondary membership grade that represents the correctness in (primary) membership assignment at each measurement points. Naturally, GT2FS is expected to give us better results in emotion classification for its representational advantage over IT2FS.

One fundamental problem in GT2FS that limits its application in classification problems, perhaps, is due to users' in- ability to correctly specify the secondary memberships. In this work, we determine the secondary memberships by extracting certain knowledge from the individual primary assignments for each feature of a given emotion for a subject. The knowledge extracted is encoded as an optimization problem with secondary memberships as unknown. The solution to the optimization problem carried out offline provides the secondary grades. The secondary grades are later aggregated with the primary memberships of individual feature for all subjects at the given measurement point to obtain modified primary memberships.

The work provides two alternative approaches to emotion recognition from an unknown facial expression, when the emotion class of individual facial expression of a large number of experimental subjects is available. The first approach deals with IT2FS to construct a fuzzy face space based on the measurements of a set of features from a given set of facial expressions carrying different emotions. An unknown facial expression is classified into one of several emotion classes by determining the maximum support of individual emotion classes to a given set of measurements of a facial expression. The class having the maximum support is declared as the emotion of the unknown facial expression. In spirit, this is similar to how a fuzzy rule- based system for classification works.

The second approach employs GT2FS to construct a fuzzy face space, comprising both primary and secondary member- ship functions, obtained from known facial expressions of several subjects containing multiple instances of the same emotion for each subject. The emotion class of an unknown facial expression is determined by computing the support of each class to the given facial expression. The class with the maximum support is the winner. The maximum support evaluation here employs both primary and secondary memberships, and thus is slightly different than the IT2FS-based classification.

Experiments reveal that the classification accuracy of emotion of an unknown person by the GT2FS-based scheme is as high as 98%. When secondary memberships are ignored, and classification is performed with IT2FS, the classification accuracy falls by a margin of 7%. The additional 7% classification accuracy obtained by GT2FS, however, has to pay a price for additional complexity of ($m \times n \times k$) multiplications, where m, n, and k denote the number of features, number of subjects, and number of emotion classes, respectively. A 2.5% improvement in classification accuracy by IT2FS has been attained by pre-processing measurements and selecting membership functions using the well-known interval approach (IA) [110].

RELATED WORK

Emotion recognition is an active research challenge in the domain of Social Signal Processing (SSP) [124]. Many studies proposed different methods as well as different nonverbal cues as a social signal to recognize emotion; from facial expression, eye gaze, vocal intonation, and gesture [125]. Amongst all, facial expression is a nonverbal cue which gives more information about human emotion [126]. Many existing researches use facial expressions to recognize basic emotion [127]–[130]. Basic emotion recognition becomes popular in the community based on assumption that basic emotion classes have clear and distinguished characteristics. Meanwhile, less effort has been made on another type of emotion, or mixed emotion class. There are several existing studies in the literature which utilized various methods for basic emotion recognition, such as multilayer neural network [131], multi-kernel SVM [132], and convolutional neural network (CNN) [133], [134]. Those researches were reporting a very good performance. The problem is, human real emotions are not limited to basic emotion types only. This remains an open challenge, to recognize human real and natural emotions, in terms of mixed emotion. Mixed emotions are defined as an affective experiences which activating two or more basic emotions, usually in opposite valences such as happy and sad [135]. Another name for this emotion class is complex emotion, multiple emotion, blended emotion, compound emotion, and so on. Compound emotion recognition as a combination of two different basic emotions in a static image was proposed by Du (2014) using Support Vector Machine classifier [136]. Another work for mixed emotion recognition in a temporal image was done by Liliana (2017) using Conditional Random Fields [137]. Unfortunately, the limitation of those works are in capturing real emotions which were limited only to two basic emotions in a single occurrence in Du [136], and a single basic emotion in each image frame which formed a sequence of mixed emotion in Liliana [137]. The novelty of this proposed research is to overcome the limitation of the existing studies in recognizing human natural emotion as a combination of multiple basic emotions. Fuzzy emotion framework recognizes the occurrence of multiple emotions in a facial expression image as well as describes the intensity of each emotion involved. To the best of our known, this approach has not been proposed in anywhere. The existing fuzzy approaches were using fuzzy logic as a basic emotion classifier [138]–[140], but not in recognizing a real fuzzy emotion. This is different from the proposed research, where fuzzy knowledge-based is used for fuzzy emotion recognition to show the fuzzy emotions and the

fuzzy degrees of each emotion contained in a face image. This is conceptually distinguishing between existing research and the proposed research. Feature extraction has an important role in classification process. Facial feature selection is a mechanism to provide a meaningful representation of a facial expression image. Several feature extraction techniques were used for facial expression recognition, most of them were based on facial shape [141], appearance [142], or hybrid features which combined shape and appearance features [128], [130]. Active Appearance Model (AAM) is a very well-known template-based method that combines shape and appearance features to produce facial landmarks as an important facial features [143]. This research utilizes AAM to detect facial important points in an input image which results a set of facial points in a Cartesian coordinates [144]. The next section describes our proposed work, the fuzzy emotion recognition framework.

Several works have investigated methods for achieving high accuracy of emotion recognition. Older survey papers published in 2014: Wu et al. (2014) and Kołakowska et al. (2014) reviewed recent research on emotion recognition but did not adopt a comprehensive approach. The most relevant current emotion recognition research survey studies, to our knowledge, are Sharma & Dhall (2021), Imani & Montazer (2019) and Jam et al. (2021). Sharma & Dhall (2021) clearly pointed out the benefits of emotion recognition on visual, speech, text and EEG. The two fusion methods are thoroughly explained, and many datasets are investigated. But only unimodal methods are considered in the survey. In Imani & Montazer (2019), the authors evaluated numerous emotion recognition technologies used in e-learning systems and listed their benefits and drawbacks. Yet tools used for emotion acquisition are not mentioned. Jam et al. (2021) mainly focused on identifying social and emotional feature of human-robot interactions using deep learning methods. More research into the correlation of detailed descriptions of emotions to emojis is needed in the survey.

Alswaidan & Menai (2020) presented a detailed taxonomy of text emotion recognition methods, which also provided an in-depth explanation of main features and different available approaches. Various explicit and implicit emotion recognition techniques in text are explained and the different approaches found in the literature, their main features, their advantages and limitations, and comparisons are illustrated in their survey. On the other hand, the survey does not discuss tools and datasets. Abdullah et al. (2021) presented

different multimodal emotion recognition methods that combine modalities, such as images and text, and facial expression combined with body's physiological responses. The findings in this survey indicated that combining multiple modalities in emotion recognition improves the recognition accuracy of emotions. Baltrušaitis et al. (2018) provided a five-tiered taxonomy based on the challenges that are encountered in every multimodal emotion recognition method. These five challenges include: fusion, alignment, representation, co-learning, and translation. Their paper identified co-learning as an area to be explored. However, this overview report mostly concentrated on multimodal research from the last ten years.

Lim et al. (2020) mainly focused on eye tracking methods and classifiers used for this data. This study underlines the necessity for combining physiological signals with eye-tracking modality. However, the results of this survey were not properly discussed or compared between inter-subject and intra-subject classification performance rates. Malla et al. (2020) specified a taxonomy for speech emotion recognition methods. The three main components of their work: speech data, feature extraction, and DFC taxonomy (classification), were classified using a CNN-based speech emotion recognition. This increases the likelihood of correctly identifying the emotion in the speech. The survey is exhaustive in its domain but is also limited as it only focuses on speech emotion resulting in a very low number of modalities.

Table 1. Comparison of existing surveys on emotion recognition.

Authors	**Taxonomy followed**	**Method**
Alswaidan & Menai (2020)	Text emotion recognition methods	Rule-based, classical learning, deep learning, and hybrid.
Abdullah et al. (2021)	Multimodal emotion recognition methods	Neural network architecture and deep learning technique.
Wu et al. (2014)	audiovisual emotion recognition techniques	Deep learning, datasets (GEMEP,RML and VAM) are explained.

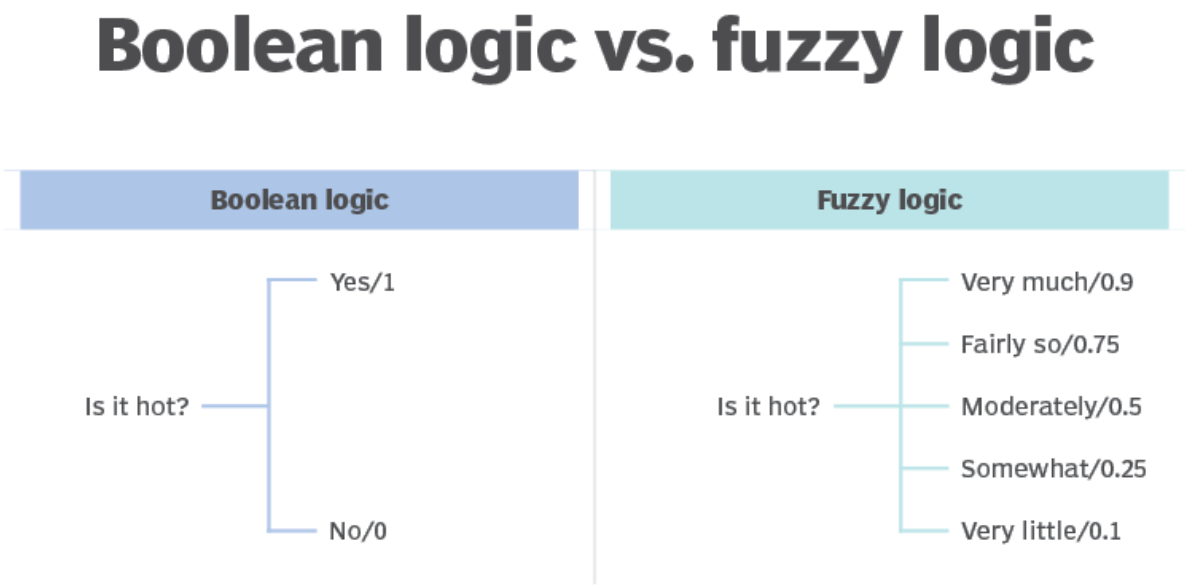
Authors	**Taxonomy followed**	**Method**
Baltrušaitis et al. (2018)	Technical challenges faced by multimodal researchers	Different co-learning methods.
Jam et al. (2021)	Social signals and emotional expressions in real-world human-robot interactions.	Deep learning methods.
Lim et al. (2020)	Taxonomy involves Eye tracking methods	machine learning algorithms.
Malla et al. (2020)	Speech emotion recognition methods	MFCC, STFT, ECC and Classification is done by combining CNN and LSTM.
Kołakowska et al. (2014)	sensory data detection available on modern smartphones.	machine learning approaches implemented to recognize emotional states.
Imani & Montazer (2019)	Various emotions recognition methods have been represented for online learning platforms.	Classification based on feature and machine or deep learning methods.
Sharma & Dhall (2021)	emotion recognition on visual, speech, text, EEG.	Deep learning classifiers and feature level and decision level fusion.

LITERATURE SURVEY

Facial Emotion Recognition is a technology used for analysing sentiments by different sources, such as pictures and videos. It belongs to the family of technologies often referred to as 'affective computing', a multidisciplinary field of research on computer's capabilities to recognise and interpret human emotions and affective states and it often builds on Artificial Intelligence technologies.

Fuzzy logic is an approach to computing based on "degrees of truth" rather than the usual "true or false" (1 or 0) Boolean logic on which the modern computer is based.

The idea of fuzzy logic was first advanced by Lotfi Zadeh of the University of California at Berkeley in the 1960s. Zadeh was working on the problem of computer understanding of natural language. Natural language -- like most other activities in life and indeed the universe -- is not easily translated into the absolute terms of 0 and 1. Whether everything is ultimately describable in binary terms is a philosophical question worth pursuing, but in practice, much data we might want to feed a computer is in some state in between and so, frequently, are the results of computing. It may help to see fuzzy logic as the way reasoning really works and binary, or Boolean, logic is simply a special case of it.



What is Fuzzy Control?

- It is a technique to embody human-like thinkings into a control system.
- It may not be designed to give accurate reasoning but it is designed to give acceptable reasoning.
- It can emulate human deductive thinking, that is, the process people use to infer conclusions from what they know.
- Any uncertainties can be easily dealt with the help of fuzzy logic.

Advantages of Fuzzy Logic System

- This system can work with any type of inputs whether it is imprecise, distorted or noisy input information.
- The construction of Fuzzy Logic Systems is easy and understandable.
- Fuzzy logic comes with mathematical concepts of set theory and the reasoning of that is quite simple.
- It provides a very efficient solution to complex problems in all fields of life as it resembles human reasoning and decision-making.
- The algorithms can be described with little data, so little memory is required.

Disadvantages of Fuzzy Logic Systems

- Many researchers proposed different ways to solve a given problem through fuzzy logic which leads to ambiguity. There is no systematic approach to solve a given problem through fuzzy logic.
- Proof of its characteristics is difficult or impossible in most cases because every time we do not get a mathematical description of our approach.
- As fuzzy logic works on precise as well as imprecise data so most of the time accuracy is compromised.

Application

- It is used in the aerospace field for altitude control of spacecraft and satellites.
- It has been used in the automotive system for speed control, traffic control.
- It is used for decision-making support systems and personal evaluation in the large company business.
- It has application in the chemical industry for controlling the pH, drying, chemical distillation process.
- Fuzzy logic is used in Natural language processing and various intensive applications in Artificial Intelligence.
- Fuzzy logic is extensively used in modern control systems such as expert systems.
- Fuzzy Logic is used with Neural Networks as it mimics how a person would make decisions, only much faster. It is done by Aggregation of data and changing it into more meaningful data by forming partial truths as Fuzzy sets.

In artificial intelligence (AI) systems, fuzzy logic is used to imitate human reasoning and cognition. Rather than strictly binary cases of truth, fuzzy logic

includes 0 and 1 as extreme cases of truth but with various intermediate degrees of truth.

As a result, fuzzy logic is well-suited for the following:

- Engineering for decisions without clear certainties and uncertainties, or with imprecise data -- such as with natural language processing technologies; and regulating and controlling machine outputs, according to multiple inputs/input variables -- such as with temperature control systems.

Various types of AI systems and technologies use fuzzy logic. This includes vehicle intelligence, consumer electronics, medicine, software, chemicals and aerospace.

- In automobiles, fuzzy logic is used for gear selection and is based on factors such as engine load, road conditions and style of driving.
- In dishwashers, fuzzy logic is used to determine the washing strategy and power needed, which is based on factors such as the number of dishes and the level of food residue on the dishes.
- In copy machines, fuzzy logic is used to adjust drum voltage based on factors such as humidity, picture density and temperature.
- In aerospace, fuzzy logic is used to manage altitude control for satellites and spacecrafts based on environmental factors.
- In medicine, fuzzy logic is used for computer-aided diagnoses, based on factors such as symptoms and medical history.
- In chemical distillation, fuzzy logic is used to control pH and temperature variables.
- In natural language processing, fuzzy logic is used to determine semantic relations between concepts represented by words and other linguistic variables.

- In environmental control systems, such as air conditioners and heaters, fuzzy logic determines output based on factors such as current temperature and target temperature.
- In a business rules engine, fuzzy logic may be used to streamline decision-making according to predetermined criteria.

CHAPTER 1

Emotion Recognition from Facial Expressions and Its Control Using Fuzzy Logic

FILTERING, SEGMENTATION, AND LOCALIZATION OF FACIAL COMPONENTS

The identification of facial expressions by pixel-wise analysis of images is both tedious and time consuming. This chapter attempts to extract significant components of facial expressions through segmentation of the image. Because of the differences in the regional profiles on an image, simple segmentation algorithms, such as histogram-based thresholding techniques, do not always yield good results. After conducting several experiments, we concluded that for the segmentation of the mouth region, a color-sensitive segmentation algorithm is most appropriate. Further, because of apparent nonuniformity in the lip color profile, a fuzzy segmentation algorithm is preferred. A color-sensitive FCM clustering algorithm [9] has, therefore, been selected for the segmentation of the mouth region.

Segmentation of the eye regions, however, in most images has successfully been performed by the traditional thresholding method. The hair region in human face can also easily be segmented by the thresholding technique. Segmentation of the mouth and eye regions is required for the subsequent determination of MO and EO, respectively. Segmentation of the eyebrow region is equally useful in determining the length of EBC. The details of the segmentation techniques of different regions are presented below.

A. *Segmentation of the Mouth Region*

Before segmenting the mouth region, we first represent the image in the $L^* a^* b^*$ space from its conventional red–green–blue (RGB) space. The $L^* a^* b^*$ system has the additional benefit of representing a perceptually uniform color space. It defines a uniform matrix space representation of color so that a perceptual color difference is represented by the Euclidean distance. The color information, however, is not adequate to identify the lip region. The position information of pixels together with their color would be a good feature to segment the lip region from the face. The FCM clustering algorithm that we

employ to detect the lip region is supplied with both color and pixel-position information of the image.

The FCM clustering algorithm is a well-known technique for unsupervised pattern recognition. However, its use in image segmentation in general and lip region segmentation in par- ticular is a novel area of research. A description of the FCM clustering algorithm can be found in books on fuzzy pattern recognition (see, e.g., [1], [17], and [45]). In this work, we just demonstrate how to use FCM clustering in the present application.

A pixel in this work is described by five attributes: three attributes of color information (L a b) and two attributes of position information (x, y). The objective of the clustering algorithm is to classify the set of 5-D data points into two classes/partitions—the lip region and the nonlip region. Initial membership values are assigned to each 5-D pixel, such that the sum of the memberships in the two regions is equal to one. That is, for the kth pixel x_k, we have

$$L(x_k) + NL(x_k) = 1 \qquad (1)$$

where $L(x_k)$ and $NL(x_k)$ denote the membership of x_k to fall in the lip and nonlip regions, respectively.

Given the initial membership values of $L(x_k)$ and $NL(x_k)$ for $k = 1$ to n^2 (assuming that the image is of size $n \times n$), we use the FCM algorithm to determine the cluster centers V_L and V_{NL} of the lip and the nonlip regions as

$$V_L = \sum_{k=1}^{n^2} [L(x_K)]^m x_K / \sum_{k=1}^{n^2} [L(x_k)]^m \qquad (2)$$

$$V_{NL} = \sum_{k=1}^{n^2} [NL(x_K)]^m x_K / \sum_{k=1}^{n^2} [NL(x_k)]^m . \qquad (3)$$

where V_j denotes the jth cluster center for $j \in \{L, NL\}$.

Determination of the cluster centers [by (2) and (3)] and membership evaluation [by (4) and (5)] are repeated several times following the FCM algorithm until the positions of the cluster centers do not further change.

Fig. 1 presents a section of a facial image with a large MO. This image is passed through a median filter, and the resulting image is shown in Fig. 2. Application of the FCM algorithm to the image in Fig. 2 yields the image in Fig. 3.

Fig. 4demonstrates the computation of MO (Section III).



Fig. 1. Original face image.



Fig. 2. Median-filtered image.



Fig. 3. Image after applying FCM clustering.

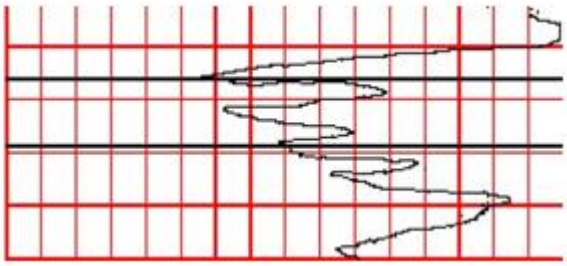


Fig. 4. Measurement of MO from the dips in average intensity plot.

B. *Segmentation of the Eye Region*

The eye region in a monochrome image has a sharp contrast to the rest of the face. Consequently, the thresholding method can be employed to segment the eye region from the image. Images grabbed at poor illumination conditions have a very low average intensity value. Segmentation of the eye region in these cases is difficult because of the presence of dark eyebrows in the neighborhood of the eye region. To overcome this problem, we consider

images grabbed under good illuminating conditions. After segmentation of the image, we need to localize the left and right eyes on the image. In this work, we use a template matching scheme to localize the eyes. The eye template we used is similar to the template shown in Fig. 5. The template matching scheme, taken from our previous works [2], [21], attempts to minimize the Euclidean distance between a fuzzy descriptor of the template and the fuzzy descriptor of the part of the image where the template is located. Even when the template is not a part of the image, the nearest matched location of the template in the image can be traced.

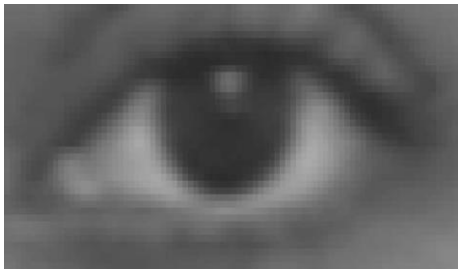


Fig. 5. Synthetic eye template.

C. *Localization of EBC Region*

In a facial image, eyebrows are the second darkest region after the hair region. The hair region is easily segmented by setting a very low threshold in the histogram-based thresholding algorithm. The eye regions are also segmented by thresholding. A search for a dark narrow template can easily localize the eyebrows. Note that the localization of the eyebrow is essential for determining its length. This will be undertaken in the next section.

DETERMINATION OF FACIAL ATTRIBUTES

In this section, we present a scheme for the measurements offacial extracts such as MO, EO, and the length of EBC.

A. *Determination of MO*

After segmentation of the mouth region, we plot the average intensity profile against the MO. The dark region in the segmented image represents the lip profile, whereas the white regions embedded in the dark region indicate the teeth. Noisy images, however, may include false white patches. Fig. 4, for instance, includes a white patch on the lip region.

Determination of MO in a black and white image is easier because of the

presence of the white teeth. A plot of the average intensity profile against the MO reveals that the curve has several minima, out of which the first and third correspond to the inner region of the top lip and the inner region of the bottom lip, respectively. The difference between the preceding two measurements along the Y-axis gives a measure of the MO. An experimental instance of MO is shown in Fig. 4, where the pixel count between the thick horizontal lines gives a measure of MO. When no white band is detected in the mouth region, MO is set to zero. When only two minima are observed in the plot of average intensity, the gap between the two minima is the measure of MO.

B. Determination of EO

After the localization of the eyes, the count of dark pixels (intensity < 30) plus the count of white pixels (intensity > 225) is plotted against the x-position. If the peak of this plot occurs at x = a, then the ordinate at x = a provides a measure of the EO (Fig. 6).

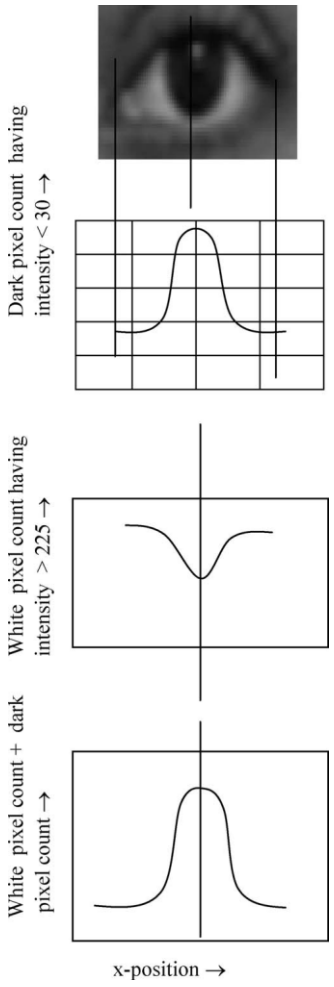


Fig. 6. Determination of the EO.

C. Determination of the Length of EBC

Constriction in the forehead region can be explained as a collection of white and dark patches called hilly and valley regions, respectively. The valley regions are usually darker than the hilly regions. Usually, the width of the patches is around 10–15 pixels for a given facial image of 512 x 512 pixels.

Let I_{av} be the average intensity in a selected rectangular profile on the forehead, and let I_{ij} be the intensity of pixel(i, j). To determine the length of EBC on the forehead region, we scan for variation in intensity along the *x*-axis of the selected rectangular region. The maximum x-width that includes variation in intensity is defined as the length of EBC. The length of the EBC has been measured in Fig. 7 by using the preceding principle. An algorithm for EBC is presented as follows.

1) Take a narrow strip over the eyebrow region with thick- ness two-thirds of the width of the forehead, which is determined by the maximum count of pixels along the length of projections from the hairline edge to the top edges of the eyebrows.
2) The length *l* of the strip is determined by identifying its intersection with the hair regions at both ends. Determine the center of the strip, and select a window of *x*-length 2*l*/3 symmetric with respect to the center.
3) For *x*-positions central to window-right-end, do the following.

   a) Select nine vertical lines in the window and compute the average intensity on each line.

   b) Calculate the variance of the nine average intensity values.

   c) If the variance is below a threshold, stop. Else, shift one pixel right.

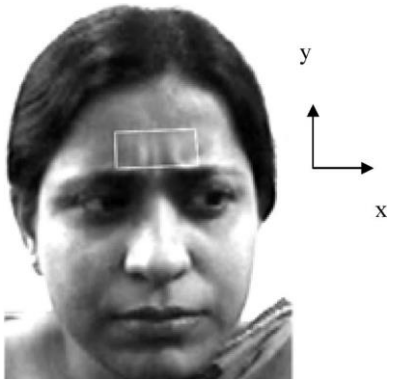


Fig. 7. Determination of EBC in the selected rectangular patch, identified by image segmentation.

4) Determine the total right shift.
5) Following a procedure similar to step 3, determine thetotal left shift.
6) Compute length of EBC = total left shift +total right shift.

FUZZY RELATIONAL MODEL FOR EMOTION DETECTION

Fuzzy logic is a mathematical approach that can be used to make predictions about human emotions and facial expressions. This technique is based on the idea that there is a lot of uncertainty and imprecision in the way people feel and express themselves. By using fuzzy logic, it is possible to make more accurate predictions about what someone is feeling by considering all the different variables that can affect their emotions. This approach has been shown to be particularly effective in predicting the sentiment of social media posts and comments.

Fuzzy logic has been used in many different fields, including sentiment analysis. To predict human emotions, facial expressions are analysed and categorized into six basic emotions: happiness, sadness, anger, fear, surprise, and disgust. Each of these emotions can be further divided into more specific sub-categories. For example, happiness can be subdivided into joy, contentment, love, etc. Once the facial expressions have been categorized, they are then represented as a set of fuzzy rules. These rules are then used to make predictions about future emotions based on current and past facial expressions. Fuzzy logic is a very effective tool for sentiment analysis because it considers the subjective nature of human emotions. Unlike other methods which rely on objective criteria, such as the tone of voice or the choice of words used, fuzzy logic can Capture the nuances and subtleties of human emotion that are often missed by other methods. This makes it an invaluable tool for research into human emotions and their triggers.

In this section, the encoding of facial attributes and their mapping to the emotion space are described using Mamdani- type implication relations.

A. *Fuzzification (Encoding) of Facial Attributes*

The measurements we obtain on MO, EO, and EBC are encoded into three distinct fuzzy sets: HIGH, LOW, and MODERATE. The typical membership

functions [29] that we have used in our simulation are presented below. For any continuous feature *x*, we have

$$\text{HIGH}(x) = 1 - \exp(-ax), \quad a > 0$$
$$\text{LOW}(x) = \exp(-bx), \quad b > 0$$
$$\text{MODERATE}(x) = \exp - (x - x_{mean}) / 2\sigma_2^2$$

where x_{mean} and σ^2 are the mean and variance of the parameter x, respectively.

For the best performance, we need to determine the optimal values of a, b, and σ. Details of these are discussed in Section VI.

B. *Fuzzy Relational Model for Emotion Detection*

Examination of a large facial database reveals that the degree of a specific human emotion, such as happiness or anger, greatly depends on the degree of MO, EO, and length of EBC. The following two sample rules describe the problem of mapping from the fuzzified measurement space of facial extracts to the fuzzified emotion space.

Rule 1: IF (*eye-opening* is MODERATE) & (*mouth-opening* is SMALL) & (*eyebrow-constriction* is LARGE)

THEN emotion is VERY-MUCH-DISGUSTED.

Rule 2: IF (*eye-opening* is LARGE) &

(*Mouth-opening* is SMALL / MODERATE) & (*eyebrow-constriction* is SMALL)

THEN emotion is VERY-MUCH-HAPPY.

Since each rule contains antecedent clauses of three fuzzy variables, their conjunctive effect is taken into account to determine the fuzzy relational matrix. The general formulation ofa production rule with an antecedent clause of three linguistic variables and one consequent clause of a single linguistic vari-able is discussed below. Consider, for instance, the following fuzzy rule.

If x isA and y isB andz isCthen w is D.

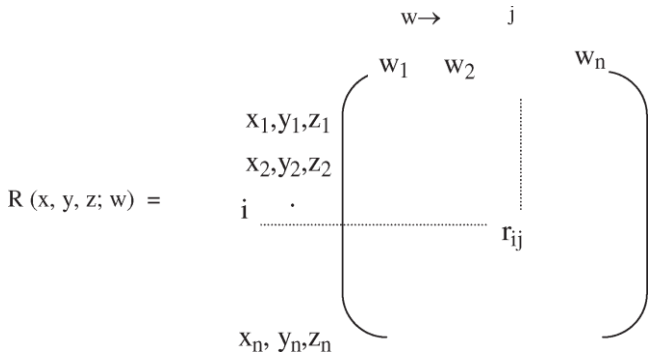


Fig. 8. Structure of a fuzzy relational matrix.

Let A(x), B(y), C(z), and D(w) be the membership distributions of linguistic variables x, y, z, and w belonging to A, B, C, and D, respectively. Then, the membership distribution of the clause "x is A and y is B and z is C" is given by t[A(x), B(y), C(z)], where t denotes the fuzzy t-norm operator [29]. Using Mamdani-type implication operator, the relation between the antecedent and consequent clauses for the given rule is described by

$$R(x, y, z; w) = \text{Min}\,[t\,(A(x), B(y), C(z))\,, D(w)]\,. \qquad (6)$$

Taking Min as the t-norm, we can rewrite the preceding expression as

$$R(x, y, z; w) = \text{Min}\,[\text{Min}\,(A(x), B(y), C(z))\,, D(w)]$$
$$= \text{Min}\,[A(x), B(y), C(z), D(w)]\,. \qquad (7)$$

Now, given an unknown distribution of $(A'(x), B'(y), C'(z))$, where $A' \neq A$, $B' \neq B$, and $C' \sim C$, we can evaluate $D'(w)$ by the following fuzzy relational equation:

$$D'(w) = \text{Min}\;\; (A'(x), B'(y), C'(z) \;\; o\; R(x, y, z; w). \qquad (8)$$

For discrete systems, the relation R (x, y, z; w) is represented by a matrix (Fig. 8), where x_i, y_i, z_i, and w_i denote specific arguments (corresponding to the variables x, y, z, and w, respectively).

In our proposed application, the row index of the relational matrix is represented by conjunctive sets of values of MO, EO, and EBC. The column index of the relational matrix denotes the possible values of six emotions: anxiety, disgust, fear, happiness, sadness, and surprise.

For determining the emotion of a person, we define two vectors: fuzzy descriptor vector F and emotion vector M. The structural forms of these two vectors are given as

F = [S(eo) M(eo) L(eo) S(mo) M(mo) L(mo)

×S(ebc) M(ebc) L(ebc)] . (9)

where S, M, and L stand for SMALL, MEDIUM, and LARGE, respectively. We also have

M = [VA(emotion) MA(emotion) N-SoA(emotion) VD(emotion)MD(emotion)N-SoD(emotion) VAf(emotion) MAf(emotion) N-SoAf(emotion) VH(emotion) MH(emotion) N-SoH(emotion) VS(emotion) MS(mod) N-SoS(emotion) VSr(emotion) MSr(emotion) N-SoSr(emotion)] (10)

where V, M, N-SO denote VERY, MODERATELY, and NOT-SO, and A, D, Af, H, S, and Sr denote ANXIOUS, DISGUSTED, AFRAID, HAPPY, SAD, and SURPRISED, respectively.

The relational equation used for the proposed system is given by

$M = F' \circ R_{FM}$ (11),

where RFM is the fuzzy relational matrix with the row and column indices as described above, and the ith component of the F' vector is given by Min*{*A'(xi)*,*B'(yi)*,*C'(zi)*}*, where the variables xi, yi, zi *∈{*eo*,* mo*,* ebc*}*, and the fuzzy sets A', B', C' *∈{*S*,*M*,* L*}* are determined by the premise of the ith fuzzy rule.

Given an F' vector and the relational matrix RFM, we can easily compute the encoded emotion vector M using the preceding relational equation. Finally, to determine the membership of the emotions from their fuzzy memberships, we need to employ a decoding scheme (12), where w1, w2, and w3 denote the weights of the respective graded memberships, which in the present context have (arbitrarily) been set to 0.33 each.

EXPERIMENTS AND RESULTS

The experiment is conducted in a laboratory environment, where illumination, sounds, and temperature are controlled to maintain uniformity in experimental conditions. Most of the subjects of the experiments are students, young faculty

members, and family members of the faculties. The experiment includes two sessions: a presentation session followed by a face-monitoring session. In the presentation session, audiovisual clips from commercial films are projected on a screen in front of individual subjects as a stimulus to excite their brain for arousal of emotion. A computer-controlled pan-tilt-type high resolution camera is used for the online monitoring of facial expressions of the subjects in the next phase. The grabbed images of facial expressions are stored in a computer for feature analysis in the subsequent phase.

Experiments were undertaken over a period of two years to identify the appropriate audiovisual movie clips that cause arousal of six different emotions: anxiety, disgust, fear, happiness, sadness, and surprise. A questionnaire was prepared to determine the consensus of the observers about the arousal of the first five emotions using a given set of audiovisual clips. It includes questions to a given observer on the percentage level of excitation of different emotions by a set of 60 audiovisual movie clips. The independent responses of 50 observers were collected, and the results are summarized in Table I. Clearly, the percentages in each row total 100.

The arousal of surprise, however, requires a subject to possess prior background information about an object or a scene, and arousal starts when the object/scene significantly differs from the general expectation. The main difficulty to get someone surprised with a movie clip is that the clip should be long enough to prepare the background knowledge before a “strange” scene is presented. To eliminate possible errors in the selection of the stimulus due to background differences among the subjects, it is reasonable to employ alternative means, rather than audiovisual movie clips, to cause arousal of surprise. Our experiments showed that an attempt at recognizing lost friends (usually schoolmates) from their current photographs causes arousal of surprise.

TABLE I

ASSESSMENT OF THE AROUSAL POTENTIAL OF SELECTED AUDIOVISUAL MOVIE CLIPS IN EXCITING DIFFERENT EMOTIONS

Subjects used to assess the emotion aroused by the audio-visual clips	Title of audio-visual clip	Percentage of arousal of different emotions by a clip					
		Anger	Anxiety	Happiness	Sadness	Fear	Relax-ation
Subject 1	Clip 1	0	0	80	0	0	20
Subject 2	Clip1	0	0	75	0	0	25
……	…….						
Subject 50	Clip1	0	0	78	0	0	22
…….							
Subject 1	Clip 2	0	82	0	9	9	0
Subject 2	Clip 2	0	80	0	12	8	0
…….	…….						
Subject 50	Clip 2	0	84	0	10	6	0
Subject 1	Clip 60	78	10	0	0	12	0
Subject 2	Clip 60	80	16	0	0	4	0
……	…….						
Subject 50	Clip 60	84	8	0	0	8	0

To identify the right movie clips capable of exciting specific emotions, we need to define a few parameters that would help indicate a consensus of the observers about the arousal of an emotion.

We have the following.

Oji,*k* Percentage level of excitation of emotion j by an observer

K using audiovisual clip i.

Eji Average percentage score of excitation assigned to emotion

j by n observers using clip i.

*σ*ji Standard deviation of the percentage score assigned to emotion j by all the subjects using clip i.

n Total number of observers.

Eji and *σ*ji are evaluated using the following expressions:

$$E_{ji} = \sum_{k=1}^{n} O_{ji,k}/n \qquad (13)$$

$$\sigma_{ji} = \sqrt{\left[\sum_{k=1}^{n} (O_{ji,k} - E_{ji})^2/n\right]}. \qquad (14)$$

The emotion w for which Ewi = max*j{*Eji*}* is the most likely aroused emotion due to excitation by audiovisual clip i.

The Ewi values are obtained for i = 1 to 60.

We next select six audiovisual clips from the pool of 60 movie samples such that the selected movies best excite six specific emotions. The selection was made by using the average-to-standard-deviation ratio for competitive audiovisual clips employed to excite the same emotion. The audiovisual clip for which the average-to-standard-deviation ratio is the largest is considered to be the *most significant sample* to excite a desired emotion.



Fig. 9. Movie clips containing four frames (row-wise) used to excite anxiety, disgust, fear, happiness, and sadness.

TABLE II
THEME OF THE MOVIE CLIPS CAUSING STIMULATION OF SPECIFIC EMOTIONS

Movie title	Emotion stimulated by the movie clip	Theme of the movie clip
Mr. and Mrs. Iyer (Indian, subtitled in English)	Anxiety	A scene of possible murder
Razz (Indian, Hindi)	Disgust	A sudden high-pitch shout of a lady in the jungle
Volcano (American)	Fear	A girl surrounded by lava erupted from a volcano
Chhadmabeshi (Indian, Bengali)	Happiness	Funny actions and words
Sarfarosh (Indian, Hindi)	Sadness	Sudden killing of tourists

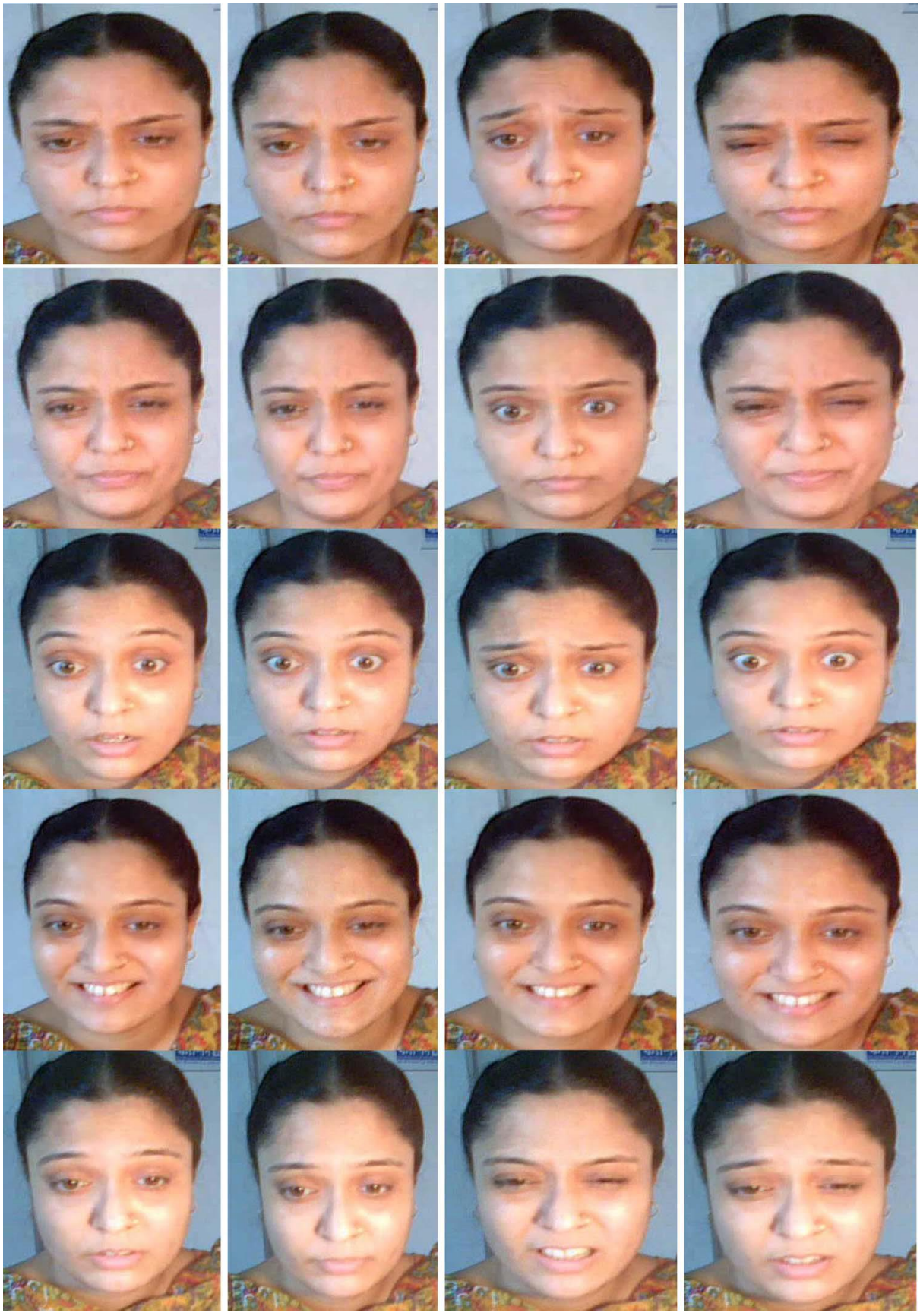


Fig. 10. Arousal of anxiety, disgust, fear, happiness, and sadness using the stimulator in Fig. 9 (female subject).



Fig. 11. Arousal of anxiety, disgust, fear, happiness, and sadness using the stimulator in Fig. 9 (male subject).

TABLE III

SUMMARY OF RESULTS OF MEMBERSHIP EVALUATION FOR THE EMOTIONS IN FIG. 10

Audio visual clip title	Emotion expected to be aroused	Feature extracted (in pixels)			Memb. for eo			Memb. for mo			Memb. for ebc		
		eo	mo	ebc	Low	Mod	Hi	Low	Mod	Hi	Low	Mod	Hi
Clip 1	Anxiety	6	2	20	0.8	0.6	0.2	0.9	0.6	0.1	0.1	0.5	0.7
Clip 2	Disgust	5	1	26	0.9	0.5	0.1	0.95	0.5	0.8	0.1	0.4	0.9
Clip 3	Fear	10	6	5	0.1	0.3	0.9	0.2	0.8	0.6	0.8	0.2	0.1
Clip 4	Happiness	8	11	1	0.1	0.4	0.8	0.1	0.7	0.9	0.9	0.3	0.09
Clip 5	Sadness	3	4	2	0.9	0.4	0.1	0.7	0.8	0.2	0.9	0.4	0.09

TABLE IV

RESULTS OF PERCENTAGE CLASSIFICATION OF AROUSED EMOTIONS IN SIX CLASSES FOR ADULT MALES

Desired Emotion	Aroused emotion obtained from facial expression					
	Anger	Anxiety	Fear	Surprise	Happiness	Sadness
Anger	88.5	5.5	4.0	0.0	0.0	2.0
Anxiety	4.0	88.2	3.8	0.0	0.0	4.0
Fear	0.0	1.0	88.4	0.0	0.0	10.6
Surprise	0.0	0.0	0.0	91.2	8.8	0.0
Happiness	0.0	0.0	0.0	9.8	90.2	0.0
Sadness	0.0	3.8	8.0	0.0	0.0	88.2

TABLE V

RESULTS OF PERCENTAGE CLASSIFICATION OF AROUSED EMOTIONS IN SIX CLASSES FOR ADULT FEMALES

Desired Emotion	Aroused emotion obtained from facial expression					
	Anger	Anxiety	Fear	Surprise	Happiness	Sadness
Anger	92.2	3.8	2.0	0.0	0.0	2.0
Anxiety	1.4	92.4	4.2	0.0	0.0	2.0
Fear	0.0	2.0	92.6	0.0	0.0	5.4
Surprise	0.0	0.0	0.0	92.2	7.8	0.0
Happiness	0.0	0.0	0.0	7.2	92.4	0.0
Sadness	0.0	2.0	5.4	0.0	0.0	92.6

TABLE VI

RESULTS OF PERCENTAGE CLASSIFICATION OF AROUSED EMOTIONS IN SIX CLASSES FOR CHILDREN IN AGE GROUP 8–12 YEARS

Desired Emotion	Aroused emotion obtained from facial expression					
	Anger	Anxiety	Fear	Surprise	Happiness	Sadness
Anger	96.0	1.2	2.2	0.0	0.0	0.6
Anxiety	0.6	96.2	2.2	0.0	0.0	1.0
Fear	0.0	1.6	96.4	0.0	0.0	2.0
Surprise	0.0	0.0	0.0	96.5	3.5	0.0
Happiness	0.0	0.0	0.0	3.4	96.6	0.0
Sadness	0.0	1.8	2.2	0.0	0.0	96.0

Fig. 9 presents the five most significant audiovisual movie clips selected from the pool of 60 movies, where each clipwas the most successful to excite one of the five emotions. Table II summarizes the theme of the selected movies. The selected clips are presented before 300 people, and their facial expressions are recorded. Figs. 10 and 11 show two representative examples of one female and one male in the age group 22–25 years.

Image segmentation is used to segment the mouth region, eye region, and eyebrow region of individual frames for each clip. MO, EO, and length of EBC are then determined for the individual frames of each clip. The averages of EO, MO, and EBC over all the frames under a recorded emotion clip are then evaluated. The membership of EO, MO, and EBC in the three fuzzy sets (Low, Medium, and High) are then evaluated using the membership functions given in Section IV. The results of membership evaluation for the emotion clips given in Fig. 10 are presented in Table III.

After evaluation of the memberships, we determine the emotion vector M by using (11), and then employ decoding rule [see (12)] to determine the membership of different emotions for thefive clips. The emotion that comes up with the highest value isregarded as the emotion of the individual clips. The preceding analysis is repeated for 300 people including 100 children, 100 adult males, and 100 adult females, and the results of emotion classification are presented in Tables IV–VI. Eachof these tables shows six "aroused" emotions, whereas the "desired" emotion refers to the emotion tag of the *most significant* audiovisual sample.

The experimental results obtained from Tables IV–VI revealthat the accuracy in the classification of emotion for adult male, adult female, and children (8–12

years) are 88.2%, 92.2%, and 96%, respectively. The classification accuracies obtained in this work are better than the existing results on accuracy reported elsewhere [19], [20], [26], [43].

VALIDATION OF THE SYSTEM PERFORMANCE

After a prototype design of an intelligent system is complete, we need to validate its performance. The term validation here refers to building the right system that truly resembles the system intended to be built. In other words, validation refers to the relative performance of the system that has been designed, and suggests reformulation of the problem characteristics and concepts based on the deviation of its performance from that of the desired (ideal) system [18]. It has experimentally been observed that the performance of the proposed system greatly depends on the parameters of the fuzzy encoders [28]. To determine optimal settings of the parameters, a scheme for the validation of the system's performance is proposed in Fig. 12.

In Fig. 12, we tune the parameters a, b, xmean (or m), and σ of the fuzzy encoders by a supervised learning algorithm, so as to generate the desired emotion from given measurements of the facial extract. The backpropagation algorithm has been employed to experimentally determine the parameters a, b, m, and σ. The feedforward neural network used for the realization of the backpropagation algorithm has three layers with 26 neurons in the hidden layer. The number of neurons in the input and output layers are determined by the dimensions of F/ vector and M vector, respectively. The root mean square error accuracy of the algorithm was set to 0.001. For a sample space of approximately 100 known emotions of different persons, the experiment is conducted, and the following values of the parameters are found to yield the best results: a = 2.2, b = 1.9, xmean = 2.0, and σ = 0.17.

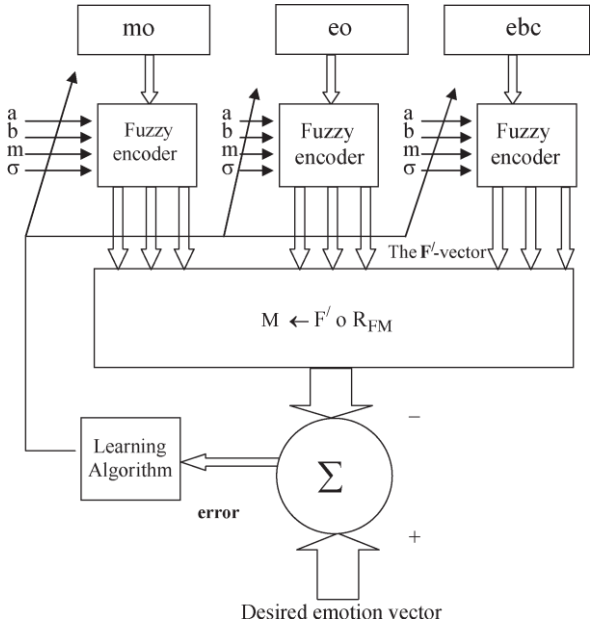


Fig. 12. Validation of the proposed system by tuning the parameters a, b, m, and σ of the encoders.

EMOTION TRANSITION AND ITS CONTROL

The emotion of a person at a given time is determined by the current state of his/her mind. The current state of one's mind is mainly controlled by the positive and negative influences of input sensory data, including voice, video clips, ultrasonic signals, and music. We propose a model of emotion transition dynamics where the strength of an emotion at time t + 1 depends on the strength of all emotions of a subject at time t, and the positive/negative influences that have been applied as stimulation at time t.

A. Model

We have the following.

$m_i(t)$ Positive unnormalized singleton membership of the ith emotion at time t.

$[w_{ij}]$ Weight matrix of dimension n n, where w_{ij} denotes a cognitive(memory-based) degree of transitionfrom the ith to the jth emotional state,and is a signed finite real number.

POS-IN(strength$_k$, t) Fuzzy membership distribution of an input stimulus with strength k to act as a positive influence at time t.

NEG-IN(input$_l$, t) Fuzzy membership distribution of an input stimulus with strength *l* to act as a negative influence at time t.

b_{ik} Weight representing the influence of aninput with strength k on the ith emotional state.

c_{il} Weight representing the influence of aninput with strength *l* on the ith emotional state.

The unnormalized membership value of an emotional state i at time t + 1 can be expressed as a function of the unnormalized membership values of all possible emotional states j, and the membership distribution of the input positive and negative influences at time t, i.e.,

$$m_i(t+1) = \sum_{\forall j} w_{i,j} \cdot m_j(t) + \sum_{\forall k} b_{i,k} \cdot \text{POS-IN}(\text{strength}_k, t) - \sum_{\forall l} c_{i,l} \cdot \text{NEG-IN}(\text{strength}_l, t). \quad (15)$$

The first term in the right-hand side of (15) accounts for the cognitive transition of emotional states, which concerns human memory, whereas the second and third terms indicate the effect of external influence on the membership of the ith emotional state. The weights wij of the cognitive memory are considered time invariant. Since wij's are time invariant, controlling the transition of emotional states can be accomplished by POS-IN(strength*k,* t) and NEG-IN(strength*l,* t). The positive terms on the right side of (15) indicate that with a growth in mj and POS-IN(strength*k,* t), mi(t + 1) also increases. The negative sign of the third term signifies that with a growth in NEG-IN(strength*l,* t), mi(t + 1) decreases.

A person's emotional state changes from happy to anxious on a negative input, e.g., when he/she runs the risk of losing something. An anxious person becomes sad when he suffers a loss. A sad person becomes disgusted when he realizes that he is not responsible for his loss/failure. In other words, with increasing negative influence (neg), the human emotion undergoes a transition in the following order:

$$\text{disgusted} \xleftarrow{\text{neg}} \text{sad} \xleftarrow{\text{neg}} \text{anxious} \xleftarrow{\text{neg}} \text{happy}.$$

Alternatively, with increasing positive influence (pos), the human emotion has a gradual transition from the disgusted state to the happy state, as follows:

$$\text{disgusted} \xrightarrow{\text{pos}} \text{sad} \xrightarrow{\text{pos}} \text{anxious} \xrightarrow{\text{pos}} \text{happy}.$$

Combining the preceding two state transition schemes, we can represent emotion transitions by a graph (see Fig. 13).

We have the following.

M = [mi] Unnormalized membership vector of dimension n × 1, whose ith element denotes the unnormalized singleton membership of emotion i at time t.

$\boldsymbol{\mu}$ = [POS-IN(strength*k,* t)] Positive influence membership vector of dimension m×1 whose kth component denotes the fuzzy membership of strength k of the input stimulus.

$\boldsymbol{\mu}'$ = [NEG-IN(input*l,* t)] Negative influence membership vector of dimension (m × 1), whose *l*th component denotes the fuzzy membership of strength *l* of the input stimulus.

B = [bij] n × m companion matrix to $\boldsymbol{\mu}$ vector.

C = [cij] n × m companion matrix to $\boldsymbol{\mu}'$ vector.

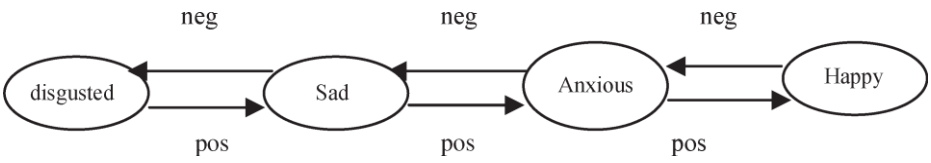


Fig. 13. Proposed emotion transition graph.

Considering the emotion transition dynamics [see (15)] for i = 1 to n, we can represent the complete system of emotion transition in vector–matrix form as

$$\mathbf{M}(t+1) = \mathbf{W}.\mathbf{M}(t) + \mathbf{B}\boldsymbol{\mu} - \mathbf{C}\boldsymbol{\mu}'. \quad (16a)$$

The weight matrix W in (16a) is given by

$$W = \begin{array}{c|cccc} \text{To} \backslash \text{From} & \text{Happy} & \text{Anxious} & \text{Sad} & \text{Angry} \\ \hline \text{Happy} & W_{1,1} & W_{1,2} & W_{1,3} & W_{1,4} \\ \text{Anxious} & W_{2,1} & W_{2,2} & W_{2,3} & W_{2,4} \\ \text{Sad} & W_{3,1} & W_{3,2} & W_{3,3} & W_{3,4} \\ \text{Angry} & W_{4,1} & W_{4,2} & W_{4,3} & W_{4,4} \end{array}$$

To keep the membership vector **M** normalized, we use the following scaling

operation:

$$\mathbf{M}^{S}(t+1) = \mathbf{M}(t+1) / \max_{i=1}^{n} \{m_i(t+1)\}. \qquad (16b)$$

Normalization of memberships in [0, 1] is needed for convenience of interpretation, but is not directly related to the emotion control problem undertaken here.

A control-theoretic representation of emotion transition dynamics [(16a) and (16b)] is given in Fig. 14, where the system has an inherent delayed positive feedback along with provision for control with external input vectors $\boldsymbol{\mu}$ and $\boldsymbol{\mu}'$.

Properties of the Model

In an autonomous system, the system states change with- out application of any control inputs. The emotion transition dynamics can be compared with an autonomous system witha setting of $\boldsymbol{\mu} = \boldsymbol{\mu}' = 0$. The limit cyclic behavior of an autonomous emotion transition dynamics is given in the followingtheorem.

Theorem 1: The vector $\boldsymbol{M}(t+1)$ in an autonomous emotion transition dynamics with $\boldsymbol{\mu} = \boldsymbol{\mu}' = \mathbf{0}$ exhibits limit cyclic behavior after every k-iterations if $\boldsymbol{W}^k = \boldsymbol{I}$.

Proof: Since $\boldsymbol{\mu} = \boldsymbol{\mu}' = \mathbf{0}$, we can rewrite (16a) as

$$\mathbf{M}(t+1) = \mathbf{W}.\mathbf{M}(t).$$

Iterating t = 0 to (k −1), we have

$$\mathbf{M}(1) = \mathbf{W}.M(0)$$

$$\mathbf{M}(2) = \mathbf{W}.\mathbf{M}(1) = \mathbf{W}^2\mathbf{M}(0)$$

:

:

$$\mathbf{M}(k) = \mathbf{W}^k\mathbf{M}(0).$$

Since the membership vector exhibits limit cyclic behavior after every k-iterations, we have

$$\mathbf{M}(k) = \mathbf{M}(0)$$

Which in turn requires $\mathbf{W}^k = \mathbf{I}$.

Theorem 1 indicates that without external perturbation, the cognitive memory helps in maintaining a recurrent relation in emotional states under a restrictive selection of memory weights, satisfying $\mathbf{W}^K = \mathbf{I}$.

For steering the system states in Fig. 13 toward the happy state, we need to provide positive influences in the state diagram at any state. Similarly, for

controlling the state transitions toward the disgusted state from any states, we submit a negative influence at states. This, however, demands a prerequisite of controllability of the membership state vector **M**. The control- lability of a given state vector **M** to a desired state is examined by Theorem 2.

Theorem 2: The necessary and sufficient condition for the state transition system to be controllable is that the controllability matrices should have rank equal to n.

$$\mathbf{P} = (\mathbf{B}\ \mathbf{WB}\ \mathbf{W}^2\mathbf{B}, \ldots, \mathbf{W}^{n-1}\mathbf{B})$$
$$\mathbf{Q} = (\mathbf{C}\ \mathbf{WC}\ \mathbf{W}^2\mathbf{C}, \ldots, \mathbf{W}^{n-1}\mathbf{C})$$

Proof: The proof follows from the test criterion of controllability of linear systems [27].

Emotion Control by Mamdani's Model

In a closed-loop process control system [21], *error* is defined as the difference of the set point (reference input) and the process response, and the task of a controller is to gradually reduce the error toward zero. When the emotional state transition is regarded as a process, we define error as the qualitative difference between the desired and the current emotional states. To quantitatively represent error, we attach an index to the individual emotional states in Fig. 13, such that when the error is positive (negative), we can guide the error toward zero by applying a positive (negative) influence. One possible indexing scheme that satisfies the above principle is given in Table VII.

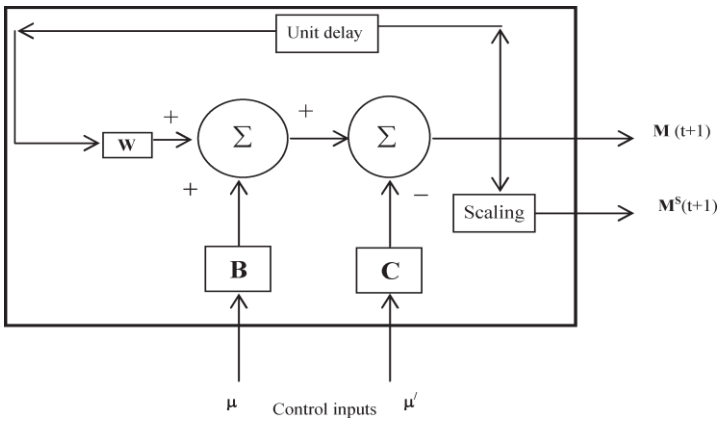


Fig. 14. Emotion transition dynamics.

TABLE VII
INDEXING OF EMOTIONAL STATES

Emotion	Index
Happy	4
Anxious	3
Sad	2
Disgusted	1

We can now quantify error as the difference between the desired emotional index (DEI) and the current emotional index (CEI). For example, if the desired emotion is happiness, and the current emotion is sadness, then the error is e = DEI − CEI = 4 − 2 = 2. Similarly, when the desired emotion is disgust, and the current emotion is happiness, error e = DEI − CEI = 1− 4 = −3.

Naturally, for the generation of control signals, we need to consider both sign and magnitude of the error. To eliminate the effect of noise in controlling emotion, instead of directly using the signed errors, we fuzzify the magnitude of errors into four fuzzy sets (i.e., SMALL, MODERATE, LARGE, and VERY LARGE) and the sign of errors into two fuzzy sets (i.e., POSITIVE and NEGATIVE) using nonlinear (Gaussian-type) membership functions. The nonlinearity of the membership functions eliminates the small Gaussian noise (with zero mean and small variance) over the actual measurement of error.

Further, to generate fuzzy control signals $\boldsymbol{\mu}$ and $\boldsymbol{\mu}/$, we need to represent the strength of positive and negative influences in four fuzzy sets (i.e., SMALL, MODERATE, LARGE, and VERY LARGE) as well. Tables VIII and IX provide a list of membership functions used for fuzzy control.

In this work, parameter selection of the membership functions has been achieved by trial and error. To attain the best performance of the controller, we considered 50 single-step control instances and found that the settings in Table X gave the best performance for all the 50 instances.

Let AMPLITUDE and SIGN be two fuzzy universes of *error*, and let STRENGTH be a fuzzy universe of positive/negative influences. Here, SMALL, MODERATE, LARGE, and VERY LARGE are fuzzy subsets of AMPLITUDE and STRENGTH, whereas POSITIVE and NEGATIVE are fuzzy subsets of the universe SIGN. Let x, y, and z be fuzzy linguistic variables, with x and y denoting error and z denoting positive/negative influence. Let Ai and Ci be any singleton fuzzy subset of *{*SMALL*,*MODERATE*,*LARGE*,*VERY-LARGE*}*, and let Bi be a singleton subset of *{*POSITIVE*,*NEGATIVE*}*. The general form of the ith fuzzy control rule is as follows.

Rule Ri: If x is Ai and y is Bi Then z is Ci.

Typical examples of the ith rule are given below.

Rule1: If *error* is SMALL and *error* is POSITIVE

Then apply *positive influence* of SMALL strength.

Rule2: If *error* is MODERATE and *error* is NEGATIVE

Then apply *negative influence* of MODERATE strength.

The Mamdani-type implication relation for Rulei is now given by

Ri(x, y; z) = Min [Min (Ai(x),Bi(y)) ,Ci(z)] . (17)

Now, for a given distribution of A/(x) and B'(y), where A' ≈ Ai, and B' ≈ Bi, we can evaluate

C'*i*(z) = MiA/(x),B/(y) o Ri(x, y; z). (18)

To take the aggregation of all the n rules, we determine

$$C'(z) = \max_{i=1}^{n} \left[C_i'(z) \right] . \qquad (19)$$

Example 1: In this example, we illustrate the construction of Ri(x, y; z) and the evaluation of C'i (z). Given

POSITIVE(error) ={1/0.2, 2/0.5, 3/0.6}

SMALL(error) ={0/0.9, 1/0.1, 2/0.01, 3/0.005}

SMALL(pos-in) ={2/0.9, 10/0.6, 20/0.21, 30/0.01}.

TABLE VIII

MEMBERSHIP FUNCTIONS OF MAGNITUDE OF ERROR AND UNSIGNED STRENGTH OF POSITIVE/NEGATIVE INFLUENCE

Signal	Membership Functions			
	SMALL	MODERATE	LARGE	VERY-LARGE
Magnitude of error (|e|)	$\exp(-k_1.\|e\|)$	$\exp(-k_2.(\|e\| - 1.5)^2)$	$1- \exp(-k_3.\|e\|)$	$1-\exp(-k_4.\|e\|)$
Unsigned strength of influence (|s|)	$\exp(-k_1'.\|s\|)$	$\exp(-k_2'.(\|s\| -1.5)^2)$	$1-\exp(-k_3'.\|s\|)$	$1- \exp(-k_4'.\|s\|)$

TABLE IX

MEMBERSHIP FUNCTIONS OF SIGN OF ERROR

Signal	Membership Functions	
	POSITIVE	NEGATIVE
Error (e)	$1 - \exp(-k_5.\ e),\ 0<e<3$	$1 - \exp(-k_5.\ \|e\|),\ -3<e<0$

TABLE X

SELECTED PARAMETERS OF THE MEMBERSHIP FUNCTIONS

Parameter	k_1	k_2	k_3	k_4	k_1'	k_2'	k_3'	k_4'	k_5
Selected value	0.15	0.12	0.16	0.008	0.16	0.14	0.18	0.008	0.10

The relational matrix R (error, error; pos-in) can now be evaluated by Mamdani's implication function as follows:

Min [POSITIVE(error), SMALL(error)]

= {(1, 0)/0.2, (1, 1)/0.1, (1, 2)/0.01, (1, 3)/0.005 ;

(2, 0)/0.5, (2, 1)/0.1, (2, 2)/0.01, (2, 3)/0.005 ;

(3, 0)/0.6, (3, 1)/0.1, (3, 2)/0.01, (3, 3)/0.005}.

Using (17), the R (error, error; pos-in) matrix is now obtained as

$$
\begin{array}{cc}
 & \text{pos-influence} \\
e,e & \begin{array}{cccc} 02 & 10 & 20 & 30 \end{array} \\
\left(\begin{array}{cc}
1,0 \\ 1,1 \\ 1,2 \\ 1,3 \\ 2,0 \\ 2,1 \\ 2,2 \\ 2,3 \\ 3,0 \\ 3,1 \\ 3,2 \\ 3,3
\end{array}\right. &
\left.\begin{array}{cccc}
0.2 & 0.2 & 0.2 & 0.01 \\
0.1 & 0.1 & 0.1 & 0.01 \\
0.01 & 0.01 & 0.01 & 0.01 \\
0.005 & 0.005 & 0.005 & 0.005 \\
0.5 & 0.5 & 0.2 & 0.01 \\
0.1 & 0.1 & 0.1 & 0.01 \\
0.01 & 0.01 & 0.01 & 0.01 \\
0.005 & 0.005 & 0.005 & 0.005 \\
0.6 & 0.6 & 0.2 & 0.01 \\
0.1 & 0.1 & 0.1 & 0.01 \\
0.01 & 0.01 & 0.01 & 0.01 \\
0.005 & 0.005 & 0.005 & 0.005
\end{array}\right)
\end{array}.
$$

Let us now consider the observed membership distribution of error to be positive and small as follows:

POSITIVE′(error) ={1/0.1, 2/0.5, 3/0.7}

SMALL′(error) ={0/0.2, l/0.1, 2/0.4, 3/0.5}.

We can evaluate

SMALL′(pos-in)

= Min POSITIVE′(error), SMALL′(error)

- R(error, error; pos-in)

= [0.1 0.1 0.1 0.1 0.2 0.1 0.4 0.5 0.2 0.1 0.4 0.5]

- R(error, error; pos-in)

= [02/0.2 10/0.2 20/0.2 30/0.01].

To determine the strength of audiovisual movies to be selected for presentation to the subject, we need to defuzzify (decode) the control signal C_i(pos-in) or C_i(neg-in). Let

$$C(\text{pos-in}) = (x_1/C(x_1),\ x_2/C(x_2),\ldots,\ x_n/C(x_n)\}.$$

Then, centroidal-type defuzzification (decoding) [21] yields the value of the control signal

$$x_{\text{deffzy}} = \sum_{i=1}^{n} C(x_i).x_i / \sum_{i=1}^{n} C(x_i).$$

Example 2: Let

C(pos-in) = [1/0.2, 3/0.2, 6/0.2, 9/0.1].

The defuzzification of the control signal by the centre of gravity method is obtained as

$$x_{\text{deffzy}} = \frac{1 \times 0.2 + 3 \times 0.2 + 6 \times 0.2 + 9 \times 0.1}{0.2 + 0.2 + 0.2 + 0.1}$$
$$= 4.14.$$

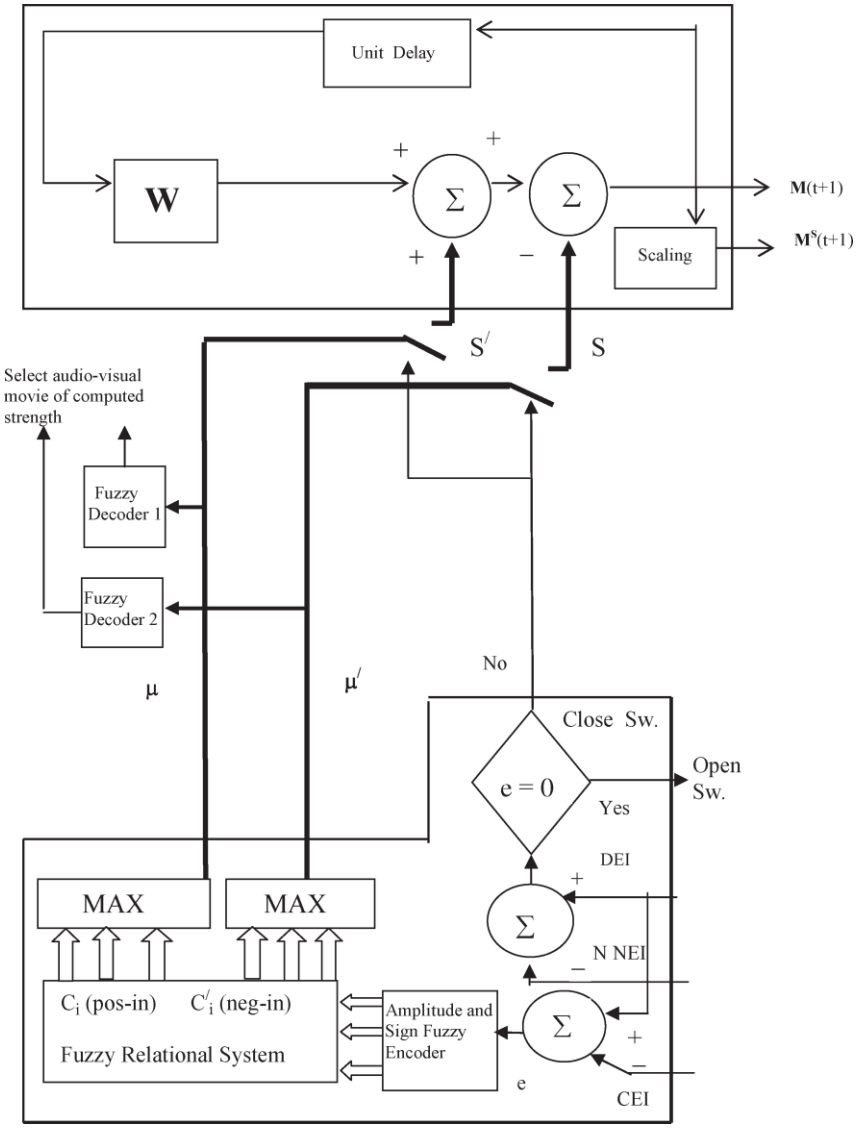


Fig. 15. Complete scheme of Mamdani-type emotion control.

Architecture of the Proposed Emotion Control Scheme

A complete scheme of the Mamdani-type emotion control is presented in Fig. 15. The scheme includes an emotion transition dynamics with provision for control inputs $\boldsymbol{\mu}$ and $\boldsymbol{\mu}'$, and a fuzzy controller to generate the necessary control signals. The fuzzy controller compares DEI and CEI, and their

difference is passed through the AMPLITUDE and SIGN fuzzy encoder. A fuzzy relational approach to the automatic reasoning introduced earlier in this work is then employed to determine the membership distribution of C_i(pos-in) and C_i'(neg-in) using the ith fired rule. The MAX units determine the maximum of C_i(pos-in) and C_i'(neg-in) that corresponds to all the fired rules, resulting in μ and μ', which are supplied as control inputs to the emotion transition dynamics. The generation of control signals μ and μ' are continued until the next emotional index (NEI) is equal to the DEI. When NEI is equal to DEI, the control switches S and S', which were closed since startup, are opened, as further transition of emotion is no longer required.

The decoding units are required to defuzzify the control signals μ and μ'. The decoding process yields the absolute strength of audiovisual movies in the range [a, +b], where integers a and b are determined by system resources, as discussed in Section V. The decoding process thus selects the audiovisual movie of appropriate strength for presentation to the subjects. Note that there are two fuzzy decoders in Fig. 15. When the error is positive, only fuzzy decoder1 is used. On the other hand, when the error is negative, fuzzy decoder 2 is used.

Experiments and Result

The proposed architecture (Fig. 15) was studied on 300 individuals at Jadavpur University. The experiment began with 100 audiovisual movies labeled with a positive/negative integer in [a, +b], which represents the strength of positive/ negative external stimulus. The labeling was done by a group of 50 volunteers who assigned a score in [a, +b]. The average of the scores assigned to an audiovisual movie stimulus is rounded-off to the nearest integer and used as its label. When the variance of the assigned scores for a movie is above a selected small threshold (1.8), we drop it from the list. In this manner, we selected only 60 movies, dropping 40 movie samples. This signifies that the scores obtained from 50 volun teers are close enough for the selected 60 audiovisual movies, which indicates good accuracy of the stimulus.

The current emotion of a subject is detected by the emotion recognition scheme presented in the previous sections. The desired emotion is randomly selected for a given subject. The control scheme outlined in Section IV is then invoked. When fuzzy decoder 1 or 2 (Fig. 15) generates a signed score, the nearest available average-scored audiovisual movie is selected for presentation to the subject. Note that when error e is equal to m (m <= 3), we need to select a sequence of at least m audiovisual movies until NEI becomes equal to DEI. An

experimental instance of emotion control is shown in Fig. 16, where the current emotion of the subject at time t = 0 is disgust, and the desired emotion is happiness. This requires three state transitions in emotions, which can be undertaken by presenting three audiovisual movies of strength +28 units, +16 units, and +08 units, in succession, to the subject. Here, DEI = 3, and CEI = 0; therefore, error *e* = DEI CEI = 3 >0. Since the error is positive, we apply positive instances of suitable strength as decided by the fuzzy controller. If the error were negative, then the fuzzy controller would have selected audiovisual stimuli of negative strength.

Two interesting points of the experiment include 1) good experimental accuracy and 2) repeatability. Experimental ac- curacy ensures that we could always control the error to zero. Repeatability ensures that for the same subject and the same pair of current and desired emotional states, the selected set of audiovisual movies is unique. Robustness of the control algorithm is thus established. In Fig. 16, the widths of the control pulses are 3, 2, and 2 s. At time t = 0, the error is large and positive. Therefore, the control signal generated has a positive strength of long duration. Then, with a gradual decrease in error, the strength of the control signal and its duration decrease.

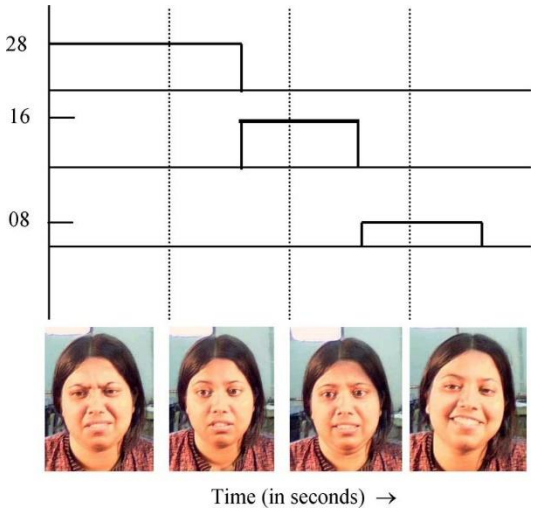


Fig. 16. Submission of audiovisual stimulus of strength 28, 16, and 8 for controlling the emotion of a subject from the disgusted state (leftmost) to a final happy state through sad and anxious states (in that order).

CHAPTER 2

General and Interval Type-2 Fuzzy Face-Space Approach to Emotion Recognition

PRELIMINARIES ON T2 FUZZY SETS

In this section, we define some terminologies related to T1 and T2 fuzzy sets. These definitions will be used throughout the work.

Definition 1: Given a universe of discourse *X*, a conventional *T1 fuzzy set A* defined on *X*, is given by a 2-D membership function, also called T1 membership function. The (primary) *membership function*, denoted by $\mu A(x)$, is a crisp number in [0, 1] for a generic element $x \in X$. Usually, the fuzzy set *A* is expressed as a two tuple [78], given by

$$A = \{(x, \mu A(x)) \mid \forall x \in X\} . \qquad (1)$$

An alternative representation of the fuzzy set *A* is also found in the literature as given in (2).

$$A = \int_{x \in X} \mu_A(x)|x \qquad (2)$$

Where $\int$ denotes union of all admissible x. *Definition 2:* A *T2 fuzzy set* $A\tilde{}$ is characterized by a 3-D membership function, also called T2 membership function, which itself is fuzzy. The *T2 membership function* is usually denoted by $\mu A\tilde{}(x, u)$, where $x \in X$, and $u \in Jx \subseteq [0, 1]$ [81].

Usually, the fuzzy set $A\tilde{}$ is expressed as a two tuple:

$A\tilde{} = \{((x, u), \mu\tilde{} A(x, u)) \mid x \in X, u \in Jx \subseteq [0, 1]\}$ (3) , where $\mu A\tilde{}(x, u) \in [0, 1]$. An alternative form of representation of the T2 fuzzy set is given in (4)

$$\tilde{A} = \int_{x \in X} \int_{u \in J_x} \mu_{\tilde{A}}(x, u)|(x, u), J_x \subseteq [0, 1] \qquad (4)$$

$$= \int_{x \in X} \left[\frac{\int_{u \in J_x} f_x(u)}{u} \right] / x, J_x \subseteq [0, 1] \qquad (5)$$

Where $fx(u) = \mu A\tilde{}(x, u) \in [0, 1]$. The $\int\int$ denotes union over all admissible x and u [81].

Definition 3: At each point of *x*, say $x = x'$, the 2-D plane containing axes u and $\mu(x', u)$ is called the *vertical slice* of $\mu A\tilde{}(x, u)$. A *secondary membership function*

is a vertical slice of $\mu\tilde{A}(x, u)$. Symbolically, it is given by $\mu\tilde{A}(x, u)$ at $x = x/$ for $x/ \in X$ and $\forall u \in Jx/ \subseteq [0, 1]$

$$\mu_{\tilde{A}}(x = x', u) = \int_{u \in J_{x/}} f_{x/}(u)|u, J_{x/} \subseteq [0, 1] \qquad (6)$$

Where $0 \leq fx'(u) \leq 1$. The amplitude of a secondary membership function is called secondary grade (of membership). In (6) Jx' is the primary membership of x'.

Definition 4: Uncertainty in the primary membership of a T2 fuzzy set $\tilde{A}$ is represented by a bounded region, called *footprint of uncertainty* (FOU) [81], which is the defined as the union of all primary memberships, i.e.,

$$FOU(\tilde{A}) = \bigcup_{x \in U} J_x. \qquad (7)$$

If all the secondary grades of a T2 fuzzy set $\tilde{A}$ are equal to 1, i.e.,

$$\mu_{\tilde{A}}(x, u) = 1 \forall x \in X, \forall u \in J_x \subseteq [0, 1] \qquad (8)$$

then $\tilde{A}$ is called *IT2FS*. The FOU is bounded by two curves, called the *Lower* and the *Upper Membership functions*, denoted by $\mu\tilde{A}(x)$ and $\mu\tilde{A}(x)$, respectively, where $\mu\tilde{A}(x)$ and $\mu\tilde{A}(x)$ at all x, respectively, take up the minimum and the maximum of the membership functions of the embedded T1 fuzzy sets [80] in the FOU.

UNCERTAINTY MANAGEMENT IN FUZZY FACE SPACE FOR EMOTION RECOGNITION

This section provides a general overview of the proposed scheme for emotion recognition using T2 fuzzy sets. Here, the emotion recognition problem is considered as uncertainty management in fuzzy space after encoding the measured facial attributes by T2 fuzzy sets.

Let $F = \{f1, f2, \ldots, fm\}$ be the set of m facial features. Let $\mu\tilde{A}(fi)$ be the primary membership in [0,1] of the feature fi to be a member of set $\tilde{A}$, and $\mu(fi, \mu\tilde{A}(fi))$ be the secondary membership of the measured variable fi in [0,1]. A primary and secondary membership function corresponds to a particular emotion class c, are denoted by $\mu\tilde{A}c(fi)$ and $\mu(fi, \mu\tilde{A}c(fi))$, respectively. If the measurement of a facial feature, fi, is performed p times on the same subject experiencing the same emotion, and the measurements are quantized into q intervals of equal size, we can evaluate the frequency of occurrence of the measured variable fi in q quantized intervals. The interval containing the highest frequency of

occurrence then can be identified, and its center, *mi*, approximately represents the mode of the measurement variable *fi*. The second moment, *σi*, around *mi* is determined and a bell-shaped (Gaussian) membership function centered at *mi* and with a spread *σi* is used to represent the membership function of the random variable *fi*. This function represents the membership of *fi* to be CLOSE-TO the central value, *mi*. It may be noted that a bell-shaped (Gaussian-like) membership curve would have a peak at the center with a membership value one, indicating that membership at this point is the largest for an obvious reason of having the highest frequency of *fi* at the center.

On repetition of the above experiment for variable *fi* on *n* subjects, each experiencing the same emotion, we obtain *n* such membership functions, each one for one individual subject.

Naturally, the measurement variable *fi* now has both intraand inter-personal level uncertainty. The intra-level uncertainty occurs due to the pre-assumption of a specific (Gaussian) primary membership function, and the inter-level uncertainty occurs due to multiplicity of the membership functions for *n* subjects. Thus, a new measurement for an unknown facial expression can be encoded using all the n-membership curves, giving *n* possible membership values, thereby giving rise to uncertainty in the fuzzy space.

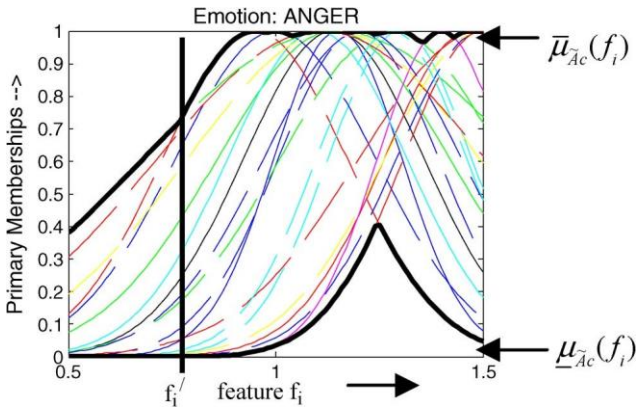


Fig. 1 Experimental FOU for feature f_i = Mouth-Opening.

The uncertainty involved in the present problem has been addressed here by three distinctive approaches: 1) IT2FS, 2) IA-IT2FS, and 3) GT2FS. The first approach is simple, but more error prone as it ignores the intra-level uncertainty. The second and the third approaches are robust as they are capable to take care of both the uncertainties. However, the modality of uncertainty management by the second and the third approaches is significantly different. The second approach models each subject's interval using a uniform probability distribution, and thus the mean and variance of each interval are mapped into

an embedded T1 fuzzy set. The third approach handles intra- and inter-personal level uncertainty compositely by fusing the primary and the secondary membership functions into an embedded interval T2 membership function. All three approaches have many common steps. Hence, we first present the steps involved in IT2FS and then explain the two techniques without repeating the common steps further.

Principles Used in the IT2FS Approach

A. Principles Used in the IT2FS Approach The primary membership functions for a given feature value *fi* corresponding to a particular emotion *c* taken from *n*-subjects together forms a IT2FS *A˜c*, whose FOU is bounded by a lower and an upper membership curves $\overline{\mu}_{\tilde{A}c}(f_i)$ and $\underline{\mu}_{\tilde{A}c}(f_i)$ respectively, where

$$\underline{\mu}_{\tilde{A}c}(f_i) = Min\left\{\mu^1_{\tilde{A}c}(f_i), \mu^2_{\tilde{A}c}(f_i), \ldots, \mu^n_{\tilde{A}c}(f_i)\right\}, \qquad (9)$$

$$\overline{\mu}_{\tilde{A}c}(f_i) = Max\left\{\mu^1_{\tilde{A}c}(f_i), \mu^2_{\tilde{A}c}(f_i), \ldots, \mu^n_{\tilde{A}c}(f_i)\right\} \qquad (10)$$

are evaluated for all *fi*, and $\mu^j_{\tilde{A}c}(f_i), 1 \leq j \leq n.$ denotes the primary membership function of feature *fi* for subject *j* in IT2FS *A˜c*. Fig. 1 provides the FOU for a given feature *fi*. Now, for a given measurement *f/ i* , we obtain an interval $[\underline{\mu}_{\tilde{A}c}(f_i^/), \overline{\mu}_{\tilde{A}c}(f_i^/)]$, representing the entire span of uncertainty of the measurement variable *f ‘i* in the fuzzy space, induced by *n* primary membership distributions:

The interval $[\underline{\mu}_{\tilde{A}c}(f_i^/), \overline{\mu}_{\tilde{A}c}(f_i^/)]$ is evaluated by replacing *fi* by *f/* in (9) and (10), respectively. If there exist *m* different facial features, then for each feature, we would have such an interval, and consequently we obtain m such intervals given by

$$\left[\underline{\mu}_{\tilde{A}c}\left(f_1^/\right), \overline{\mu}_{\tilde{A}c}\left(f_1^/\right)\right], \left[\underline{\mu}_{\tilde{A}c}\left(f_2^/\right), \overline{\mu}_{\tilde{A}c}\left(f_2^/\right)\right], \ldots\ldots\ldots, \left[\underline{\mu}_{\tilde{A}c}\left(f_m^/\right), \overline{\mu}_{\tilde{A}c}\left(f_m^/\right)\right].$$

The proposed IT2FS reasoning system employs a particular format of rules, commonly used in fuzzy classification problems [89]. Consider for instance a fuzzy rule, given by *R*c: if *f*1 is *A˜*1 AND *f*2 is *A˜*2 . . .. AND *fm* is *A˜m* then emotion class is *c*. Here, *fi* for *i* = 1 *to m* are m-measurements (feature values) and *A˜*1,*A˜*2, . . . ,*A˜m* are IT2FS on the respective domains

$$\tilde{A}_i = \left[\underline{\mu}_{\tilde{A}c}(f_i), \overline{\mu}_{\tilde{A}c}(f_i)\right], \forall i. \qquad (11)$$

Since an emotion is characterized by all of these m features, to find the overall support of the m features (m measurements made for the unknown subject) to the emotion class *c* represented by the *n* primary memberships, we use the fuzzy meet operation

$$S_c^{\min} = Min\left\{\underline{\mu}_{\tilde{A}c}\left(f_1^{/}\right), \underline{\mu}_{\tilde{A}c}\left(f_2^{/}\right) \ldots, \underline{\mu}_{\tilde{A}c}\left(f_m^{/}\right)\right\} \quad (12)$$

$$S_c^{\max} = Min\left\{\overline{\mu}_{\tilde{A}c}\left(f_1^{/}\right), \overline{\mu}_{\tilde{A}c}\left(f_2^{/}\right), \ldots, \overline{\mu}_{\tilde{A}c}\left(f_m^{/}\right)\right\}. \quad (13)$$

Thus, we can say that the unknown subject is experiencing the emotion class *c* at least to the extent *smin c* , and at most to the extent *smax c* .

To reduce the nonspecificity associated with the interval *sc–i* = [*smin c , smax c*], different approaches can be taken. For example, the most conservative approach would be to use lower bound, while the most liberal view would be to use the upper bound of the interval as the support for the class *c*. In the absence of any additional information, a balanced approach would be to use center of the interval as the support for the class *c* by the *n* primary memberships to the unknown subject. This idea is supported byMendel [84] and Lee [90].We compute the center *Sc* of the interval *Sc–i*

$$S_c = \frac{\left(s_c^{min} + s_c^{max}\right)}{2}. \quad (14)$$

expression is in emotion class *c*. Now, to predict the emotion of a person from his facial expression, we determine *S*c for each emotion class. Presuming that there exist *k* emotion classes, let us denote the degree by which the emotion classes 1*,* 2*, . . . , k* support the unknown facial expression be *S*1*, S*2*, ...,Sk*, respectively.

Since a given facial expression may convey different emotions with different degrees, we resolve the conflict by ranking the *Si* for *i* = 1 *to k*, and thus determine the emotion class *r*, for which *Sr* >= *S*i for all *i*.

The principle of selection of the emotion class *r* from a set of competitive emotions, satisfying the above inequality holds, since the joint occurrence of the fuzzy memberships, induced by (54)–(56), for all the features of the given facial expression for emotion r is the greatest among the same values for all other emotions.

Principles Used in the GT2FS Approach

The previous approach employs a reasoning mechanism to compute the degree of support of *k* emotion classes induced by *m* features for each class to an unknown facial expression using a set of *k* × *m* IT2FS. The GT2FS-based reasoning realized with measurements taken from n-subjects, however, requires *k* × *m* × *n* GT2FSs to determine the emotion class of an unknown facial expression. The current approach tunes the primary membership values for the given measurements using the secondary memberships of the same measurement, and thus reduces the degree of intra-level uncertainty of the primary distributions. The reduction in the degree of uncertainty helps in improving the classification accuracy of emotion at the cost of additional complexity required to evaluate T2 secondary distributions and also to reason with *k* × *m* × *n* fuzzy sets.

Let f_i be the measurement of the *i*th feature for a subject with an unknown emotion class. Now, by consulting the *n* primary membership functions that were generated from n-subjects in the training data for a given emotion class, *c*, we obtain *n* primary membership values corresponding to f_i for emotion class *c* as given by $\mu^1_{\tilde{A}c}(f_i), \mu^2_{\tilde{A}c}(f_i), \ldots, \mu^n_{\tilde{A}c}(f_i)$. Let the secondary membership values for each primary membership value, respectively, be $\mu(f_i, \mu^1_{\tilde{A}c}(f_i)), \mu(f_i, \mu^2_{\tilde{A}c}(f_i)), \ldots, \mu(f_i, \mu^n_{\tilde{A}c}(f_i))$.

Note that, these secondary membership values correspond to emotion class *c*. unless clarity demands, we have avoided (here and elsewhere) use of a subscript to represent the emotion class. We now fuse (aggregate) the evidences provided by the primary and secondary membership values to obtain the modified primary membership supports. A plausible way of fusing would be to use a T-norm. Here, we use the product. The product always lies within the FOU and thus satisfies MendelJohn Representation Theorem [39]. Further higher is the secondary membership, higher is the product representing new embedded fuzzy membership. Since the secondary membership represents the degree of correctness in primary membership, the product helps in reduction of intra-level uncertainty. Thus, for subject *j* of the training data representing emotion class *c*, we obtain

$$^{\text{mod}}\mu^j_{\tilde{A}c}(f_i) = \mu^j_{\tilde{A}c}(f_i) \times \mu\left(f_i, \mu^j_{\tilde{A}c}(f_i)\right) \forall j = 1, \ldots, n \quad (15)$$

where $^{\text{mod}}\mu^j_{\tilde{A}c}(f_i)$ denotes the modified primary membership value for *j*th training subject for *c*th emotion class. The secondary membership values used in the

above product function are evaluated using their primary memberships obtained by a procedure discussed $^{\text{mod}}\mu_{\tilde{A}}^{j}(f_i^{/})$ in Section IV.

The next step is to determine the range of $[^{\text{mod}}\underline{\mu}_{\tilde{A}}(f_i^{/}), ^{\text{mod}}\overline{\mu}_{\tilde{A}}(f_i^{/})]$, for $j = 1$ *to* n, comprising the minimum and the maximum given by

$$^{\text{mod}}\underline{\mu}_{\tilde{A}}\left(f_i^{/}\right) = Min\left\{^{\text{mod}}\mu_{\tilde{A}}^{1}\left(f_i^{/}\right), ^{\text{mod}}\mu_{\tilde{A}}^{2}\left(f_i^{/}\right), \ldots, ^{\text{mod}}\mu_{\tilde{A}}^{n}\left(f_i^{/}\right)\right\} \quad (16)$$

$$^{\text{mod}}\overline{\mu}_{\tilde{A}}\left(f_i^{/}\right) = Max\left\{^{\text{mod}}\mu_{\tilde{A}}^{1}\left(f_i^{/}\right), ^{\text{mod}}\mu_{\tilde{A}}^{2}\left(f_i^{/}\right), \ldots, ^{\text{mod}}\mu_{\tilde{A}}^{n}\left(f_i^{/}\right)\right\}. \quad (17)$$

Now, for m features, the rule-based T2 classification is performed in a similar manner as in the previous section with $\underline{\mu}_{\tilde{A}}(f_i^{/})$ and $\overline{\mu}_{\tilde{A}}(f_i^{/})$ by $^{\text{mod}}\underline{\mu}_{\tilde{A}}(f_i^{/})$ and $^{\text{mod}}\overline{\mu}_{\tilde{A}}(f_i^{/})$, Respectively.

Methodology

We briefly discuss the main steps involved in fuzzy facespace construction based on the measurements of m facial features for n-subjects, each having *l* instances of facial expression for a particular emotion. We need to classify a facial expression of an unknown person into one of *k* emotion classes. *IT2FS-Based Emotion Recognition:*

1) We extract m facial features for *n* subjects, each having *l* (*l* could be different for different emotion classes) instances of facial expression for a particular emotion. The above features are extracted for *k*-emotion classes.

2) We construct a fuzzy face space for each emotion class separately. The fuzzy face space for an emotion class comprises a set of *n* primary membership functions for each feature. Thus, we have m groups (denoted by m rows of blocks in Fig. 2) of n-primary membership functions (containing *n* blocks under each row of Fig. 2). Each primary membership curve is constructed from *l*-facial instances of a subject attempted to exhibit a particular emotion in her facial expression by acting.

3) For a given set of features $f_1', f_2', \ldots, f_m'$ obtained from an unknown facial expression, we determine the range of membership for feature f_i', given by $[\mu_{\tilde{A}}(f_i'), \mu_{\tilde{A}}(f_i')]$, where $\tilde{A}$ is an IT2FS with a primary membership function defined as CLOSE-TO-center-value-m of the respective membership function.

4)

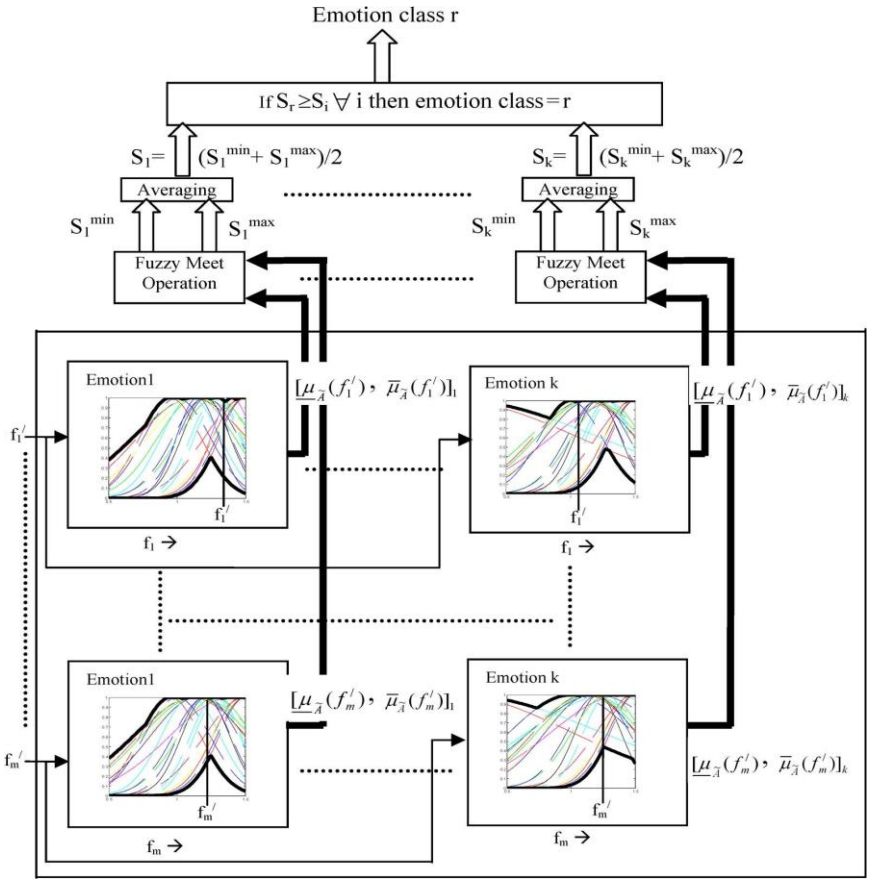


Fig. 2.The IT2FSS-based emotion recognition.

5) Now, for an emotion class *j*, we take fuzzy meet operation over the ranges for each feature to evaluate the range of uncertainty for individual emotion class. The meet operation here is computed by taking cumulative t-norm (here we use *min*) of $\mu\tilde{A}(f_i')$ and $\mu_{\tilde{A}}(f_i')$ separately for $i = 1$ *to* m, and thus obtaining S_j^{min} and S_j^{max}, respectively (see top of Fig. 2).

6) The support of the *j*-th emotion class to the measurements is evaluated by computing the average S_j of S_j^{min} and Sjmax.

7) Now, we determine the maximum support offered by all the *k* emotion classes, and declare the unknown facial expression to have emotion r, if $S_r \geq$

S_i for all emotion class i = 1 to k. The suffix j in $[\mu^{\min}{}_A(f_i'), \mu^{\max}{}_A(f_i')]_j$ refers to the range in that interval for emotion j.

GT2FS-Based Emotion Recognition:

1) This step is same as the step 1 of IT2FS-based emotion recognition.
2) The construction of the primary membership functions here follows the same procedure as given in step 2 of IT2FS-based recognition scheme. In addition, we need to construct secondary membership functions for individual primary membership curves. The procedure for construction of secondary membership functions will be discussed in Section IV. The complete scheme of construction of T2FFS, considering all k emotion classes, is given in Fig. 3.
3) For a given feature f_i', we consult each primary and secondary membership curve under a given emotion class, and take the product of primary and secondary membership at $f_i = f_i'$. The resulting membership value obtained for the membership curves for the subject w in the training data is given by

$$^{\text{mod}}\mu_{\tilde{A}}^{w}\left(f_i'\right) = \mu_{\tilde{A}}^{w}\left(f_i'\right) \times \mu\left(f_i', \mu_{\tilde{A}}^{w}\left(f_i'\right)\right) \quad (18)$$

where the notations have their usual meaning. Now, for w = 1 *to* n, we evaluate $^{\text{mod}}\mu_{\tilde{A}}^{w}(f_i')$, and thus obtain the minimum and the maximum values of $^{\text{mod}}\mu_{\tilde{A}}^{w}(f_i')$, to obtain a range of uncertainty $[^{\text{mod}}\mu A\tilde{}(f_i'), ^{\text{mod}}\ \mu_A\tilde{}(f_i')]$. This is repeated for all features under each emotion class. In Fig. 4 we, unlike conventional approaches, present secondary membership functions against feature f_i', for i = 1 *to* m. Such representation is required to demonstrate the computation of $^{\text{mod}}\mu A\tilde{}(f_i')$.
4) Step 4 is the same as that in IT2FS-based recognition scheme with the replacement of $\mu\ \tilde{}(f_i')$ and $\mu_A\tilde{}(f_i')$,

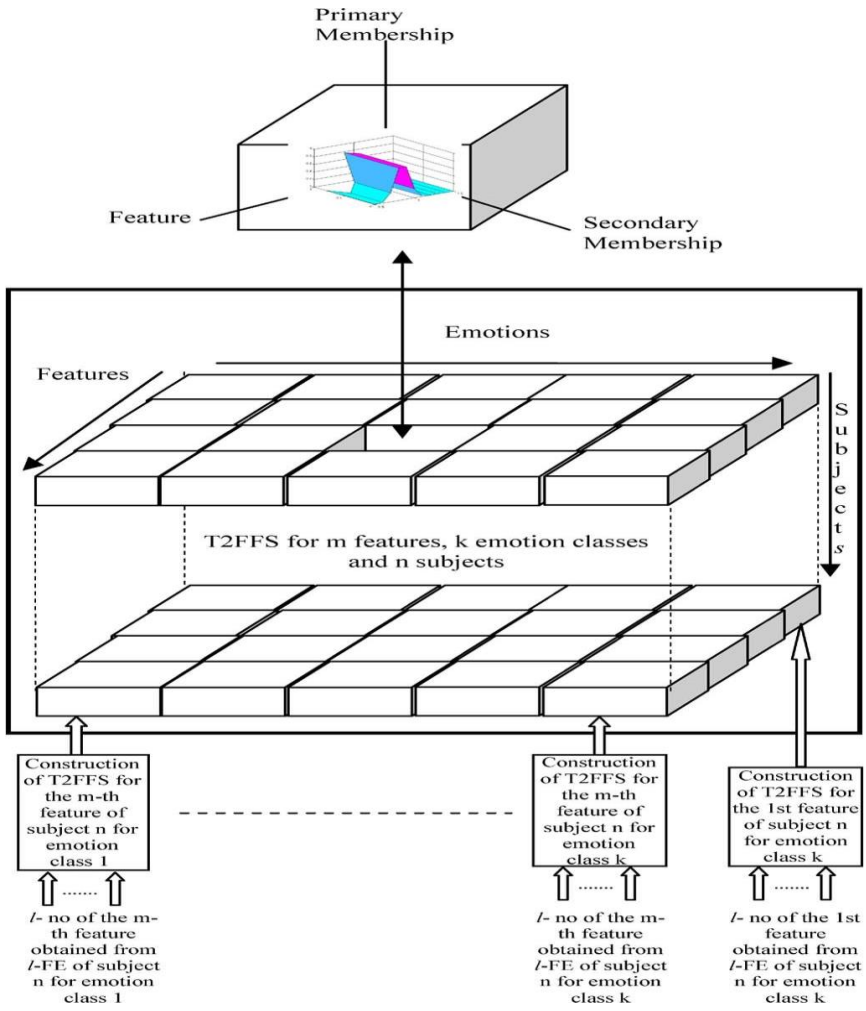


Fig. 3. General type-2 fuzzy face-space construction for *m* features, *k* emotion classes, and *n* subjects.

respectively, by ${}^{\text{mod}}\mu_{\tilde{A}}(f_i')$ and $\overline{{}^{\text{mod}}\mu_{\tilde{A}}}(f_i')$. Steps 5 and 6 are exactly similar to those in IT2FS-based recognition scheme. A complete scheme for GT2FS-based emotion recognition, considering support of k-emotion classes is given in Fig. 5.

. FUZZY T2 MEMBERSHIP EVALUATION

In this, we discuss T2 membership evaluation [89]–[81]. Although theoretically very sound, T2 fuzzy set has limitedly been used over the last two decades because of the users' inadequate knowledge to correctly assign the secondary memberships. This work, however, overcomes this problem by extracting T2 membership function from its T1 counterpart by an evolutionary algorithm. A brief outline to the construction of secondary membership function is given in this section.

Intuitively, when an expert assigns a grade of membership, she is relatively more certain to determine the location of the peaks and the minima of the function, but may not have enough background knowledge to correctly assign the membership values at other points. Presuming that the (secondary) membership values at the peak and the minima are close to 1, we attempt to compute secondary memberships at the remaining part of the secondary membership function. The following assumptions are used to construct an objective function, which is minimized to obtain the solution of the problem.

1) Let $x = x_p$ and $x = x_q$ be two successive optima (peak/minimum) on the primary membership function $\mu_A(x)$. Then, at any point x lying between x_p and x_q, the secondary membership $\mu(x,\mu_A(x))$ will be smaller than both $\mu(x_p,\mu_A(x_p))$ and $\mu(x_q,\mu_A(x_q))$.

2) The fall-off in secondary membership at a point x away from its value at a peak/minimum $\mu(x_p,\mu_A(x_p))$ is exponential, given by

$$\mu(x,\mu_A(x)) = \mu(x_p,\mu_A(x_p)).\exp(-|x - x_p|). \quad (19)$$

3) The secondary membership at any point x between two consecutive optima at $x = x_p$ and $x = x_q$ in the primary membership is selected from the range $[\alpha,\beta]$, where

$$\left.\begin{aligned}\alpha &= \mu\left(x_p, \mu_A(x_p)\right).\exp\left(-|x - x_p|\right)\\ \beta &= \mu\left(x_q, \mu_A(x_q)\right).\exp\left(-|x - x_q|\right)\end{aligned}\right\}. \quad (20)$$

T1 defuzzification over the average of n primary membership functions should return the same value as obtained by T2 defuzzification for a given primary membership function for any given source. This assumption holds because the two modalities of defuzzification, representing the same real-world parameter, should return close values, ignoring the average inter-personal level of uncertainty while taking the average of n-primary membership functions.

4) The unknown secondary membership at two values of x separated by a small positive δ should have a small difference. This is required to avoid sharp changes in the secondary grade.

Let the primary membership functions for feature $f_i = x$ from n sources be $\mu_{\tilde{A}}^1(x), \mu_{\tilde{A}}^2(x), \ldots, \mu_{\tilde{A}}^n(x)$. Then, the average membership function which represents a special form of fuzzy aggregation is given by

$$\mu_{\bar{\mathrm{A}}}(x) = \frac{\sum_{\mathrm{i}=1}^{\mathrm{n}} \mu_{\bar{\mathrm{A}}}^{\mathrm{i}}(x)}{n}, \forall x \qquad (21)$$

i.e., at each position of $x = x_j$, the above membership aggregation is employed to evaluate a new composite membership profile $\mu_A\tilde{}(x)$. The defuzzified signal obtained by the centroid method [36] from the averaged primary membership function is given by

$$c = \frac{\sum_{\forall} x.\mu_{\mathrm{A}}}{\sum} = \frac{\mathrm{x}}{\mu\tilde{}(x)} \qquad .(22)$$

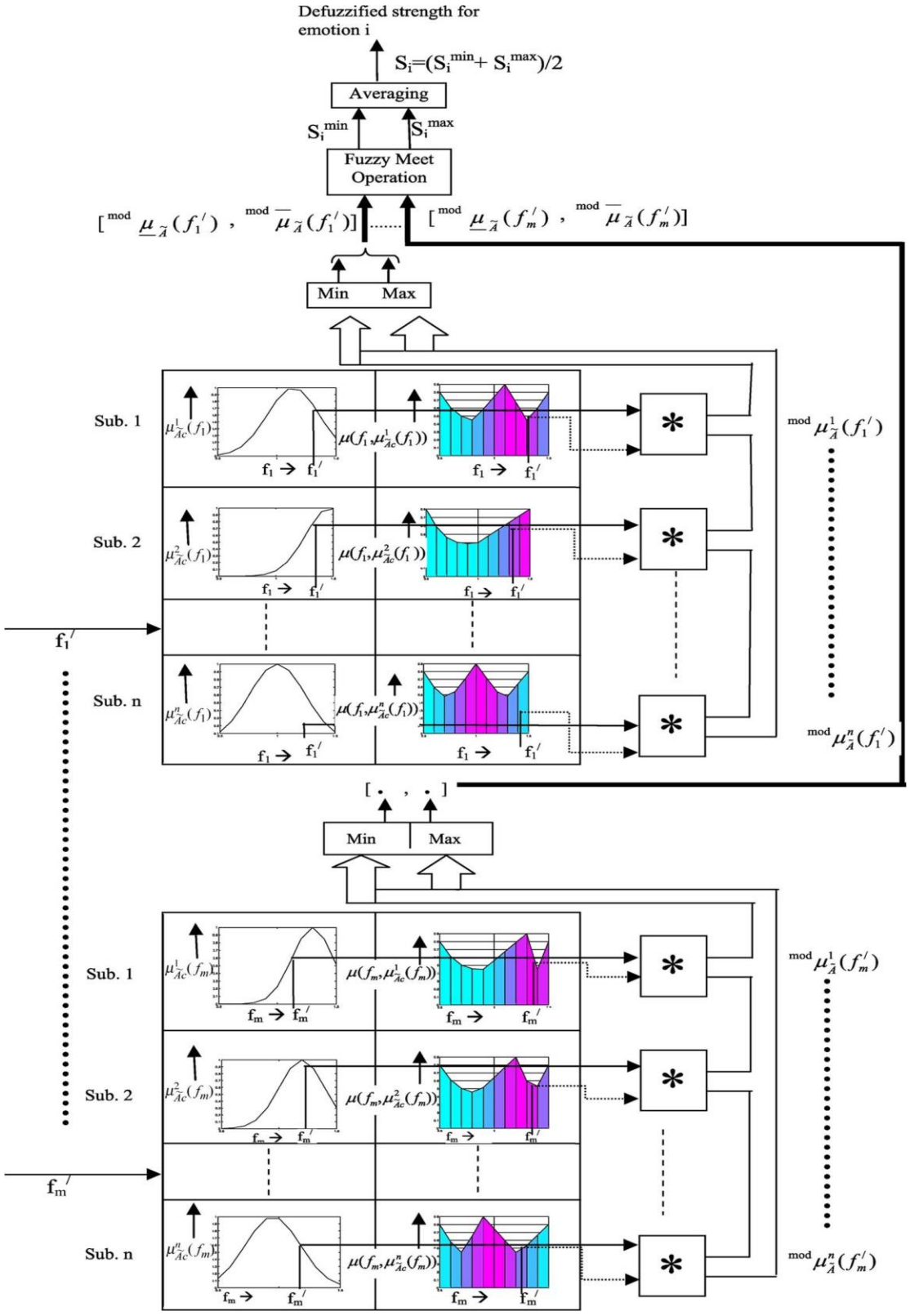


Fig. 4. Computing support of the general type-2 fuzzy FS for emotion class *i*.

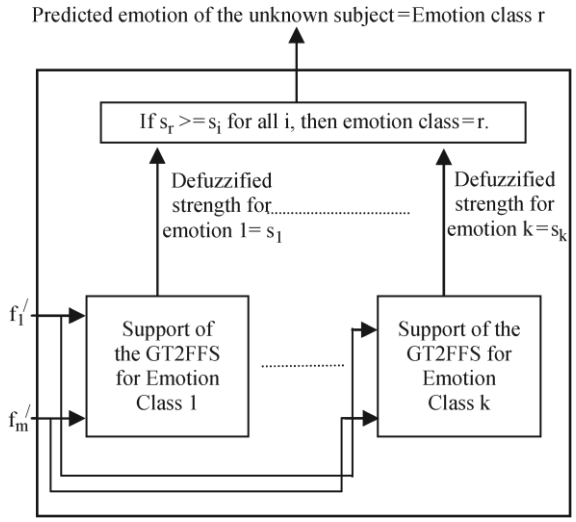


Fig. 5.GT2FFS-based emotion classification.

Further, the T2 centroidal defuzzified signal obtained from the ith primary and secondary membership functions here is defined as

$$\overline{c_i} = \frac{\sum_{\forall x} x.\mu_{\tilde{A}}^{i}(x).\mu\left(x, \mu_{\tilde{A}}^{i}(x)\right)}{\sum_{\forall x} \mu_{\tilde{A}}^{i}(x).\mu\left(x, \mu_{\tilde{A}}^{i}(x)\right)}. \qquad (23)$$

The products of primary and secondary memberships are used in (23) to refine the primary memberships by the degree of certainty of the corresponding secondary values.

Using assumptions 3 and 4, we construct a performance index J_i to compute secondary membership for the ith subject for a given emotion

$$J_i = (\overline{c_i} - \overline{\overline{c}})^2 + \sum_{x=x_1}^{x_{R-1}} \left\{\mu\left((x+\delta), \mu_{\tilde{A}}^{i}(x+\delta)\right) - \mu\left(x, \mu_{\tilde{A}}^{i}(x)\right)\right\}^2 \qquad (24).$$

The second term in (24) acts as a regularizing term to prevent abrupt changes in the membership function. In (24),x_1 and x_R are the smallest and the largest values of a given feature considered over R sampled points of$\mu_{\tilde{A}}^{i}(x)$. In (24), $\delta = (x_R - x_1)/(R - 1)$ and $x_k = x_1 + (k - 1)\delta$ for $k = 1,....,R$. The secondary membership evaluation problem now transforms to minimization of J_i by selecting $\mu(x, \mu_{\tilde{A}}^{i}(x))$ from a given range $[\alpha,\beta]$, where α and β are the secondary memberships at the two optima in secondary membership around the point x. Expressions (20) are used to compute α and β for each x separately. Note that, for each subject

carrying individual emotion, we have to define (23) and (24) and find the optimal secondary membership functions.

Any derivative-free optimization algorithm can be used to minimize J_i with respect to secondary memberships, and obtain $\mu(x, \mu^i_{\tilde{A}}(x))$ at each x except the optima on the secondary membership. Differential evolution (DE) [76] is one such derivative-free optimization algorithm, which has fewer control parameters, and has outperformed the well-known binary coded genetic algorithm [96] and particle swarm optimization algorithms [97] with respect to standard benchmark functions [87]. Further, DE is simple and involves only a few lines code, which motivated us to employ it to solve the above optimization problem.

An outline to basic DE [76] is given in the Appendix. An algorithm to compute the secondary membership function of a T2 fuzzy set from its primary counterpart using DE is given below.

1) Obtain the averaged primary membership function $\mu_{\tilde{A}}(x)$ from the primary membership functions obtained from n sources, i.e., $i = 1,....,n$. Evaluate c, and also c_i for a selected primary membership distribution $\mu^i_{\tilde{A}}(x)$ using (22) and (23), respectively.
2) Find the optima on $\mu^j_{\tilde{A}}(x)$ for a given j. Let the set of x corresponding to the optima be S. Set the secondary membership $\mu(x, \mu^j_{\tilde{A}}(x))$ to 0.99 (close to one) for all $x \in S$.
3) For each $x \in X$, where $x \notin S$, identify the optima closest around x from S. Let they be located at $x = x_p$ and $x = x_q$, where $x_p < x < x_q$. Determine α and β for each x, given by (20).
4) For each x, where $\mu(x,\mu^j\tilde{A}(x))$ lies in $[\alpha,\beta]$, minimize J_j by DE.
5) Obtain $\mu(x,\mu^j\tilde{A}(x))$ for all x after the DE converges. 6) Repeat step 2 onwards for all j.

For a Gaussian primary membership function, the minimum occurs at infinity, but the minimum value is practically zero when x is $m \pm 4\sigma$, where m and σ are mean and standard deviation of x. In Step 2, the minimum is taken as $m \pm 4\sigma$,

and we obtain x by dividing the range $[m - 4\sigma, m + 4\sigma]$ into equal intervals of same length (here 20 intervals).

An illustrative plot of secondary membership function for a given primary is given in Fig. 6.

FILTERING UNWANTED DATA POINTS IN FEATURE SPACE USING INTERVAL APPROACH

The IT2FS-based scheme for emotion recognition given in Section III is computationally efficient with good classification accuracy. However, its performance depends greatly on the measurements obtained from facial expressions of the experimental subjects. In order to reduce the effect of outliers, we here present a scheme of data pre-processing/filtering and selection of membership functions following the well-known IA [110].

The important steps of IA used in the present context are re-structured for the present application as outlined below. Let $[a^{(i)}, b^{(i)}]$ be the end-point interval of measurements of a given facial feature for the ith subject obtained from l instances of her facial expressions for a specific emotion.

Step 1) (Outlier processing): This step divides the two sets of lower and upper data end-points: $a^{(i)}$ and $b^{(i)}$, respectively, for i = 1 to n subjects in quartiles, and tests the acceptability of each data end-point by satisfying the following criteria:

$$\left.\begin{array}{l} a^{(i)} \in [Q_a(0.25) - 1.5IQR_a, Q_a(0.75) + 1.5IQR_a] \\ b^{(i)} \in [Q_b(0.25) - 1.5IQR_b, Q_b(0.75) + 1.5IQR_b] \\ L^{(i)} \in [Q_L(0.25) - 1.5IQR_L, Q_L(0.75) + 1.5IQR_L] \end{array}\right\} \quad (25)$$

where $Q_j(x)$ denotes the quartile ranges containing the first x% of the data points in the *i-th* data set. Here, $j \in \{a,b,L\}$ and a,b denote lower, upper end points of intervals, and L is the length of an interval. *IQR* denotes intra-quartile range and is defined by $Q(0.75)$ minus $Q(0.25)$. The suffixes a,b and L in *IQR* denote the *IQR* for left, right end points and interval length, respectively. $L^{(i)}$ is defined as the length of data interval = $b^{(i)} - a^{(i)}$, for i = 1 *to* n. The reduced set of data end-points after outlier processing is n'.

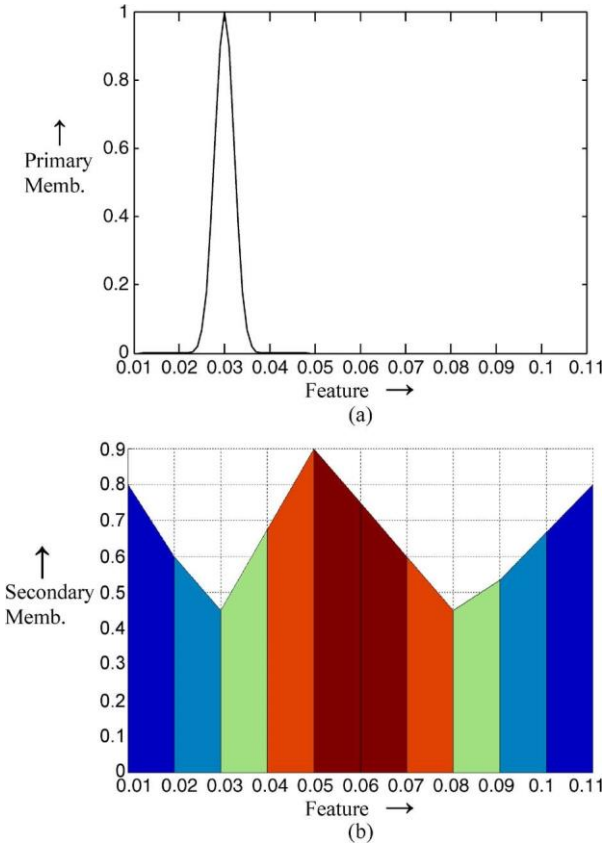


Fig. 6. (a) The primary membership function for a given feature and (b) its corresponding secondary membership function obtained by minimizing J_i.

Step 2) (Tolerance limit processing): This step deals with tolerance limit processing by presuming the data distributions to be Gaussian, and testing whether lower/upper data end-points: $a^{(i)}$,$b^{(i)}$ and interval length $L^{(i)}$ lie within mean plus/minus k(= 2.752) times the standard deviation of the data points. The number 2.752 appears in the scenario for statistical validation with 20 data end-point intervals for 20 subjects [110].

If a data interval $[a^{(i)},b^{(i)}]$ and its length $L^{(i)}$ satisfy (26), the interval is accepted, otherwise rejected:

$$\left.\begin{array}{l} a^{(i)} \in [m_l - ks_l, m_l + ks_l] \\ b^{(i)} \in [m_r - ks_r, m_r + ks_r] \\ L^{(i)} \in [m_L - ks_L, m_L + ks_L] \end{array}\right\} \quad (26)$$

where, m_j and s_j denotes sample mean and standard deviation for $j \in \{l,r,L\}$, for the n' set of data points/intervals. After tolerance processing, the reduced set of data end-points is n''.

Step 3) (*Reasonable-interval test*): This step checks whether intervals are reasonable, i.e., they are overlapped. This has been performed by

computing ξ^*, given in (27) and then by testing whether lower bounds of each interval $a^{(i)} < \xi^*$ and upper bound $b^{(i)} > \xi^*$, where $\xi^*$is one of the possible values of

$$\xi^* = \frac{\left(m_r\sigma_l^2 - m_l\sigma_r^2\right) \pm \sigma_l\sigma_r\left[(m_l - m_r)^2 + 2\left(\sigma_l^2 - \sigma_r^2\right)\ln\left(\frac{\sigma_l}{\sigma_r}\right)\right]^{\frac{1}{2}}}{\sigma_l^2 - \sigma_r^2} \quad (27)$$

where m_l and σ_l are sample mean and variance of the n''' left endpoints and m_r and σ_r are sample mean and variance of the n''' right endpoints. If $m_l <= \xi^* <= m_r$ is satisfied, then the data intervals are retained and dropped otherwise. The remaining number of data points after the drop of some intervals is called n'''.

Step 4) (FOU selection): This step is used for the selection of the right FOU among triangle, left shoulder, and right shoulder. For each FOU, the criteria can be found in [110]. We here reproduce the results for triangular FOU only, as our results to be given in Section VI yields triangular FOU. For triangular FOU, the conditions are

$$\left.\begin{aligned} m_r &\leq 5.831m_l - 1.328\frac{s_c}{\sqrt{n'''}} \\ m_r &\leq 0.171m_l + 8.29 - 1.328\frac{s_d}{\sqrt{n'''}} \\ m_r &\geq m_l \end{aligned}\right\} \quad (28)$$

where s_c = standard deviation of $[b^{(i)} - 5.831a^{(i)}]$ for $i = 1$ *to* n''', s_d = standard deviation of $[b^{(i)} - 0.17a^{(i)} - 8.29]$ for $i = 1$ *to* n'''.

Step 5) (FOU parameter evaluation): This step deals with parameter evaluation of the triangular membership functions for the existing data intervals $[a^{(i)}, b^{(i)}]$. For each interval $[a^{(i)}, b^{(i)}]$, we obtain the parameters $a_{MF}^{(i)}$ and $b_{MF}^{(i)}$ representing the end-points of the x-coordinates of the base for a symmetric triangular membership function as reproduced below [68]:

$$\left.\begin{aligned} a_{MF}^{(i)} &= \tfrac{1}{2}\left[\left(a^{(i)} + b^{(i)}\right) - \sqrt{2}\left(b^{(i)} - a^{(i)}\right)\right] \\ b_{MF}^{(i)} &= \tfrac{1}{2}\left[\left(a^{(i)} + b^{(i)}\right) + \sqrt{2}\left(b^{(i)} - a^{(i)}\right)\right] \end{aligned}\right\} \quad (29)$$

We use these membership functions in place of Gaussian membership functions in our IT2FS approach and call this approach as IT-IT2FS

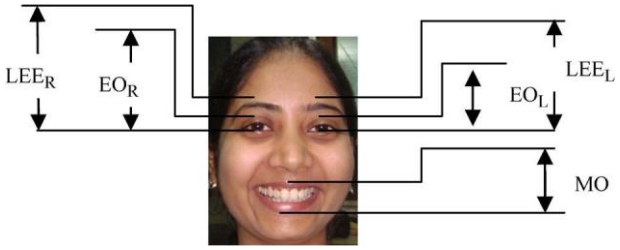


Fig. 7. Facial features.

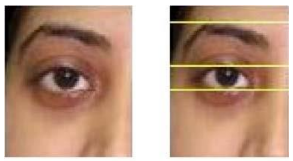


Fig. 7.(a) Localized eye search region, and (b) detection of eye features.

EXPERIMENTS DETAILS

Here, we consider the following *k*(= 5) emotion classes: anger, fear, disgust, happiness, and relaxation. The experiment is conducted with two sets of subjects: 1) the first set of *n*(= 20) subjects is considered for designing the fuzzy face space and 2) the other set of 40 facial expressions taken from six unknown subjects is considered to validate the result of the proposed emotion classification scheme. Five facial features (i.e., *m* = 5) have been used here to design the T2 fuzzy face space.

We now briefly overview the main steps of feature extraction followed by fuzzy face-space construction and emotion recognition of an unknown subject using the pre-constructed face space.

Feature Extraction

Feature extraction is a fundamental step in emotion recognition. This work considers extraction of features from emotionally rich facial expressions synthesized by the subjects by acting. Existing research results [56], [70] reveal that the most important facial regions responsible for the manifestation of emotion are the eyes and the lips. This motivated us to select the following features: Left eye opening (EO_L), right eye opening (EO_R), Distance between the Lower Eyelid to the Eyebrow for the Left Eye (LEE_L), distance between the lower eyelid to eyebrow for the right eye (LEE_R), and the maximum mouth opening (MO) including the lower and the upper lips. Fig. 7 explains the above facial features on a selected facial image.

For extraction of any of the features mentioned above, the first step that needs to be carried out is to separate out the skin and the non-skin regions of the image.

Estimation of Eye Features (EO_L, LEE_L, EO_R, and LEE_R): To compute the eye features, we first localize the eye region as shown in Fig. 8(a). The image in Fig. 8(a) is now transformed to gray scale, and average intensity over each row of pixels is evaluated. Now, we identify the row with the maximum dip in average intensity, while scanning the image from top. This row indicates the first dark region from top, i.e., the eyebrow region (Fig. 8(b)). Similarly, we detect the lower eyelid by identifying the row with the maximum dip in intensity in the gray scale image, while scanning the face up from the bottommost row. The location of the top eyelid region is identified by scanning the face up from the marked lower eyelid until the maximum dip occurs in the gray scale image.

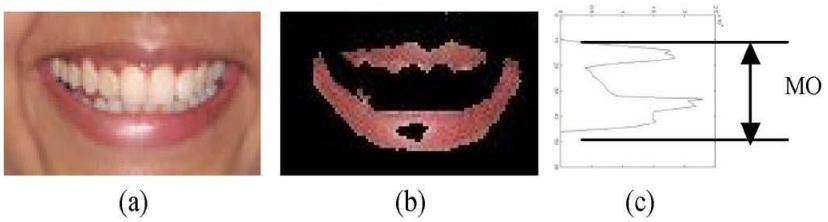


Fig. 8. (a) Mouth search area, (b) lip cluster, and (c) graph of average intensity over each row against the row position.

Estimation of MO: In order to estimate the MO, we first localize the mouth region as shown in Fig. 9(a). Then, a conversion from R-G-B to perceptually uniform $L^* - a^* - b^*$ color space is undertaken in order to represent the perceptual difference in color by Euclidean distance [111]. The k-means clustering algorithm is applied next on this image to segment the image into three clusters, namely skin, lip, and teeth regions. Each cluster is now transformed to gray scale, and the one with the highest average gradient of the boundary points (in intensity) is declared as the lip region. Now, to obtain the MO, we plot the average intensity over each row of Fig. 9(b) against the row number. The width of the zero-crossing zone in the plot (Fig. 9(c)) provides a measure of MO.

Experiments are undertaken both on colored image database such as the Indian Women (Jadavpur University) database, and gray scale images including Japanese Female Facial Expression (JAFFE) and Cohn-Kanade databases. The principles of feature extraction introduced above are equally applicable in both color and gray scale images. However, for color images, we need a conversion

to gray scale to determine the features of eye and mouth of the subject. In addition, for the gray scale facial images, segmentation of lip-, skin-, and teeth-regions is performed with intensity data only, unlike the case in color images, where we use the 3-D data points (L^*, a^*, b^*) as the input to the k-means algorithm for segmentation.

Selective images from three facial expression databases are given in Fig. 10. Training and test image data partition for three experimental databases is given in Table I. The training data in Table I include *l* instances for *n* subjects for *k* distinct emotions.

The following explanation in this section is given with respect to Indian Woman Database (Jadavpur University).

Creating the T2 Fuzzy Face Space

The interval T2 fuzzy face space contains only the primary membership distributions for each facial feature. Since we have five facial features, and the experiment includes five distinct emotions of 20 subjects, we obtain $20 \times 5 \times 5 = 500$ primary membership curves. To compute primary memberships, ten instances of a given emotion are used. These 500 membership curves are grouped into 25 heads, each containing 20 membership curves of 20 subjects for a specific feature for a given emotion. Fig. 11 gives an illustration of one such group of 20 membership functions for the feature EO_L for the emotion: Anger.



Fig. 9. Experiment done on different databases: (a) JAFFE, (b) Indian women database (Jadavpur University), (c) Cohn-Kanade database.

TABLE I

TRAINING AND TEST DATA FOR THREE DATABASES

Databases used	Training Images *(n×l×k)*	Test Images selected at random
JAFFE	5×3×5	40
Indian Woman(J.U)	20×10×5	40
Cohn-Kanade	10×5×5	40

For each primary membership function, we have a corresponding secondary membership function. Thus, we obtain 500 secondary membership functions. Two illustrative T2 secondary memberships for subjects 1 and 2 for the feature EO_L for the emotion anger are given in Fig. 12. The axes in the figure represent feature (EO_L), primary and secondary membership values as indicated.

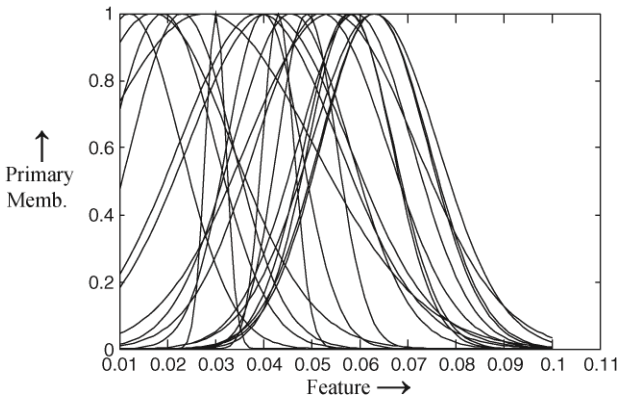


Fig. 10.Membership distributions for emotion anger and feature EO_L.

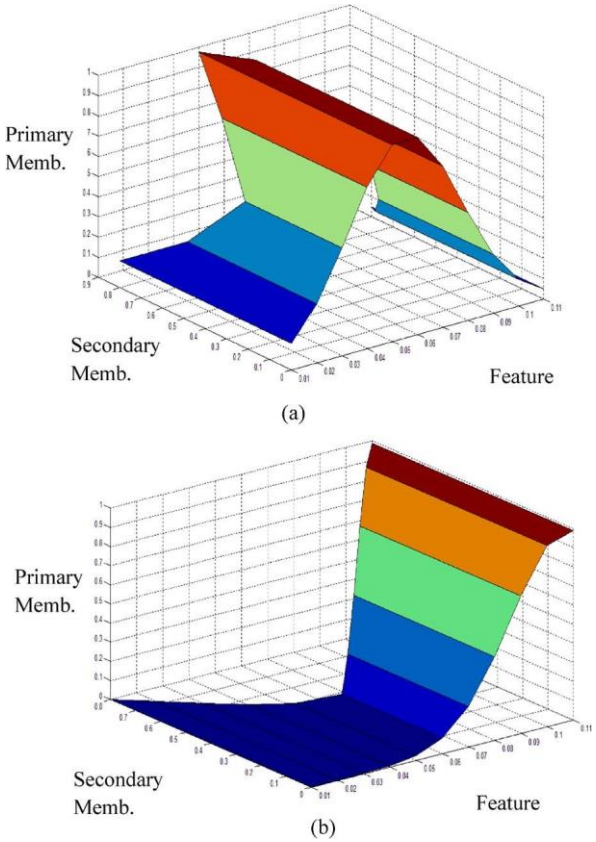


Fig. 11. (a) Secondary membership curve of subject 1. (b) Secondary membership curve of subject 2 for emotion anger.

Emotion Recognition of an Unknown Facial Expression

The emotion recognition problem addressed here attempts to determine the emotion of an unknown person from her facial expression. To keep the measurements in an emotional expression normalized and free from distance variation from the camera focal plane, we construct a bounding box, covering only the face region, and the reciprocal of the diagonal of the bounding box is used as a scale factor for normalization of the measurements. The normalized features obtained from Fig. 13 are listed in Table II. We now briefly explain the

experimental results obtained by two alternative reasoning methodologies incorporating IT2FS and GT2FS.



Fig. 12.Facial image of an unknown subject.

TABLE II

EXTRACTED FEATURES OF FIG. 13

EO_L	EO_R	MO	LEE_L	LEE_R
0.026	0.026	0.135	0.115	0.115

TABLE III

CALCULATED TYPE-2 PRIMARY MEMBERSHIP VALUES FOR THE FEATURE: EO_L UNDER EMOTION: DISGUST

Feature: $EO_L(pri)$									
Primary Memberships (μ_{pri})									
0.65	0.10	0.15	0.45	0.18	0.55	0.06	0.41	0.16	0.12
0.38	0.45	0.09	0.19	0.67	0.68	0.52	0.44	0.37	0.55
Range (min{ μ_{pri}}, max{ μ_{pri}})= [0.06, 0.68]									

TABLE IV

CALCULATED RANGES OF PRIMARY MEMBERSHIPS AND CENTRE VALUE FOR EACH EMOTION

Emotion	Range of Primary Membership for Features					Range S_c^j after fuzzy Meet operation (**centre**)
	EO_L	EO_R	MO	LEE_L	LEE_R	
Anger	0-0.18	0-0.24	0.076 -0.764	0-0.215	0.001-0.234	0-0.18 (0.09)
Disgust	0.06-0.68	0.064-0.65	0-0.52	0-0.58	0-0.78	0-0.52 (0.26)
Fear	0 - 0.067	0-0.071	0.194-0.914	0.042-0.74	0.038-0.729	0-0.067 (0.0335)
Happiness	0 - 0.687	0-0.694	0.12-0.897	0.57-0.85	0.64-0.89	0-0.687 (**0.3435**)
Relaxed	0 - 0.384	0-0.393	0-0.052	0.076-0.89	0.081-0.92	0-0.052 (0.026)

IT2FS-Based Recognition: The IT2FS-based recognition scheme considers a fuzzy face space of five sets of 20 primary membership functions as in Fig. 11, where each set refers to one particular feature obtained from 20 sources for an individual emotion. Consequently, for five distinct emotions, we have 25 such sets of primary membership functions. Table III provides the evaluation of T2 primary membership values for the feature, EO_L, consulting 20 primary functions obtained from 20 subjects, representing the facial expression for disgust. The range of these memberships is given in the last row of Table III. For each feature, we obtain five tables like Table III, each one for a given emotion. Thus, for five features, we would have altogether 25 such tables.

Table IV provides the results of individual range in primary membership for each feature experimented under different emotional conditions. For example, the entry (0–0.18) corresponding to the row anger and column EO_L gives an idea about the extent of the EO_L for the unknown subject matches with known subjects from the emotion class anger. The results of computing fuzzy meet operation over the range of individual features taken from facial expressions of the subjects under the same emotional condition are given in Table IV. The average of the ranges along with its center value is also given in Table IV. It is observed that the center has the largest value (= 0.3435) for the emotion: happiness.

TABLE V

RESULTS OF EXECUTION OF IA ON FEATURE EO_L DATA SET FOR EMOTION: ANGER

Data Preprocessing
Data points taken: 20 pairs of $a^{(i)}$, $b^{(i)}$ for i=1 to 20 subjects
Step-1: Outlier Processing Result: deleted point is [0.021, 0.113]
Step- 2: Tolerance Limit Processing Result: no deletion
Step- 3: Reasonable- interval Rest Result: no deletion
FOU Selection:
Step- 4: Computed values for: S_c= 0.0934; S_d= 0.0179; Test Condition (given in Fig. 14) Result: FOU = Triangle as (m_l, m_r) = (0.0755, 0.12257) lies in the interior FOU (triangle) obtained by satisfying (28)
Triangle Parameter Evaluation
Step- 5: a_{MF}, b_{MF} evaluated from (29) Result: Given in Fig. 15.

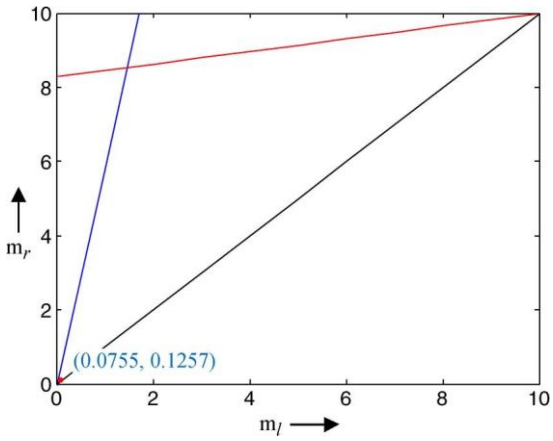


Fig. 13. Graphical selection of FOU by testing that the point $(m_l, m_r) = (0.0755, 0.1257)$ plotted in the figure lies under the triangular zone obtained by satisfying inequalities in (28).

IT2FS-Based Recognition With Pre-Processing of Features Using the Interval Approach (Hereafter IA-IT2FS): The IA introduced in Section V has two fundamental merits. It eliminates noisy data points obtained from facial data of the subjects. It also helps in identifying the primary membership functions for each feature of a facial expression representing a specific emotion by a statistically meaningful approach. The results of execution of adapted IA algorithm given in the last section for the feature EO_L for the emotion anger are given in Table V for convenience. After similar tables for all features of all possible emotions are determined, we use the statistically significant FOU for each feature of each emotion. In Fig. 14, we provide an illustrative experimental FOU for the feature EO_L for emotion anger by performing step 4 of Section V. The parameters of the FOU, here triangles, are evaluated by step 5 of Section V. Now, for an unknown facial expression, we follow the steps of IT2FSbased approach to recognize the emotion exhibited in the facial expression. Our experiments reveal that the pre-processing steps by IA help in improving the recognition accuracy of the IT2FS scheme by 2.5% (Fig. 15).

GT2FS-Based Recognition: We now briefly illustrate the GT2FS-based reasoning for emotion classification. Here, the secondary membership function corresponding to the individual primary membership function of five features obtained from facial expressions carrying five distinct emotions for 20 different subjects are determined using membership functions like Fig. 12.

Table VI provides the summary of the primary and secondary memberships obtained for EO_L for the emotion: disgust. The range computation for the feature EO_L is also shown in the last column of Table VI. The same computations are repeated for all emotions, and the range evaluated in the last column of Table

VII indicates that the center of this range here too has the largest value (= 0.301) for the emotion: happiness.

TABLE VI

CALCULATED TYPE-2 MEMBERSHIP VALUES FOR THE FEATURE: EOL UNDER EMOTION: DISGUST

Feature	Primary Memberships (μ_{pri})	Secondary memberships (μ_{sec})	$\mu^{mod}=\mu_{pri}\times\mu_{sec}$	Range ($\min\{\mu^{mod}\}$, $\max\{\mu^{mod}\}$)
EO_L	0.65	0.06	0.039	0.039-0.4355
	0.1	0.87	0.087	
	0.15	0.85	0.1275	
	0.45	0.53	0.2385	
	0.18	0.74	0.1332	
	0.55	0.52	0.286	
	0.08	0.88	0.0704	
	0.41	0.53	0.2173	
	0.16	0.78	0.1248	
	0.12	0.81	0.0972	
	0.38	0.67	0.2546	
	0.45	0.58	0.261	
	0.09	0.89	0.0801	
	0.19	0.72	0.1368	
	0.67	0.65	0.4355	
	0.68	0.58	0.3944	
	0.52	0.55	0.286	
	0.44	0.67	0.2948	
	0.37	0.78	0.2886	
	0.55	0.53	0.2915	

TABLE VII

CALCULATED RANGES OF PRIMARY MEMBERSHIP CENTRE VALUE FOR EACH EMOTION

Emotion	Range of Secondary Membership for Features					Range S_c^j after fuzzy Meet operation (**centre**)
	EO_L	EO_R	MO	LEE_L	LEE_R	
Anger	0-0.21	0 - 0.27	0.26 - 0.983	0.0006 - 0.763	0.0006- 0.790	0-0.21 (0.105)
Disgust	0.039- 0.4355	0.031- 0.433	0-0	0-0.15	0.-0.13	0-0 (0)
Fear	0 - 0.312	0-0.295	0.04- 0.713	0.044- 0.564	0.038- 0.571	0-0.295 (0.1475)
Happiness	0 - 0.602	0-0.606	0.273- 0.98	0.06- 0.93	0.064- 0.97	0-0.602 (**0.301**)
Relaxed	0 - 0.425	0-0.421	0-0	0.001- 0.758	0.001- 0.742	0-0 (0)

TABLE VIII

PERCENTAGE ACCURACY OF OUR PROPOSED METHODS OVER THREE DATABASES

	JAFFE	Indian Women (Jadavpur University)	Cohn-Kanade	Average Accuracy (of last 3 columns)
IT2FS	90%	92.5%	92.5%	91.667%
IA-IT2FS	92.5%	95%	95%	94.167%
GT2FS	97.5%	100%	97.5%	98.333%

PERFORMANCE ANALYSIS

Performance analysis for emotion recognition itself is an open-ended research problem, as there is a dearth of literature on this topic. This work, compares the relative performance of the proposed GT2FS algorithms with five traditional emotion recognition algorithms/techniques and the IA-IT2FS and IT2FS-based schemes introduced here, considering a common framework in terms of their features and databases. The algorithms used for comparison include linear SVM classifier [70], (T1) fuzzy relational approach [56], PCA [75], multilayer perceptron (MLP) [43], [71], radial basis function network (RBFN) [43], [71], IT2FS, and IA-IT2FS [110].

Table VIII shows the classification accuracy of our proposed three algorithms using three facial image databases, i.e., JAFFE, Indian Women Face Database (Jadavpur University), and Cohn-Kanade database. Experimental classification accuracy obtained for different other algorithms mentioned above using the three databases is given in Table X.

Two statistical tests called McNemar's test [100] and Friedman test [101], and one new test, called root mean square error test are undertaken to analyze the relative performance of the proposed algorithms over existing ones.

McNemar's Test

Let f_A and f_B be two classifiers obtained by algorithms A and B, when both the algorithms have a common training set R.

Let n_{01} be the number of examples misclassified by f_A but not by f_B, and n_{10} be the number of examples misclassified by f_B but not by f_A. Then, under the null hypothesis that

both algorithms have the same error rate, the statistic Z in (30) follows a χ^2 with degree of freedom equals to 1 [101]:

$$Z = \frac{\left(|n_{01} - n_{10}| - 1\right)^2}{n_{01} + n_{10}} . \qquad (30)$$

Let A be the proposed GT2FS algorithm and B is one of the other seven algorithms. We thus evaluate $Z = Z_1$ through Z_7, where Z_j denotes the comparator statistic of misclassification between the GT2FS (Algorithm: A) and the *j*th of the seven algorithms (Algorithm: B), where the suffix *j* refers to the algorithm in row number *j* of Table IX.

Table IX is evaluated to obtain Z_1 through Z_7 and the hypothesis has been rejected, if $Z_j > \chi_1^2, 0.95 = 3.84$, where $\chi_1^2, 0.95 = 3.84$ is the value of the chi square distribution for 1 degree of freedom at probability of 0.05 [123].

The last inequality indicates that if the null hypothesis is true, then the probability of χ^2 to be more than 3.84 is less than 0.05. If the hypothesis is not rejected, we consider its acceptance. The decision about acceptance or rejection is also included in Table IX.

It is evident from Table IX that McNemar's test cannot distinguish the performance of the five classification algorithms: IT2FS, IA-IT2FS, SVM, fuzzy relational approach, and PCA that support the hypothesis. Hence, next we use the Friedman test for ranking the algorithms.

TABLE IX

STATISTICAL COMPARISON OF PERFORMANCE USING MC NEMAR'S TEST WITH THREE DATABASES

Reference Algorithm A=GT2FS						
Classifier Algorithm B used for comparison	JAFFE Database		Indian Database (Jadavpur University)		Cohn-Kanade Database	
	Z_j	Comments on acceptance/ rejection of hypothesis	Z_j	Comments on acceptance/ rejection of hypothesis	Z_j	Comments on acceptance/ rejection of hypothesis
IT2FS	1.333	Accept	1.333	Accept	0.5	Accept
IA-IT2FS	0.5	Accept	0.5	Accept	0	Accept
SVM	1.333	Accept	0	Accept	1.333	Accept
Fuzzy Relational Approach	2.25	Accept	1.333	Accept	1.333	Accept
PCA	0	Accept	2.25	Accept	2.25	Accept
MLP	8.1	Reject	8.1	Reject	8.1	Reject
RBFN	11.077	Reject	10.083	Reject	10.083	Reject

Friedman Test

The Friedman test [100] ranks the algorithms for each data sets separately. The best performing algorithm gets rank 1. In case of ties, average ranks are used.

TABLE X

AVERAGE RANKING OF CLASSIFICATION ALGORITHMS BY FRIEDMAN

TEST, WHERE, CA = Classifier Algorithm, A = GT2FS, B_1 = SVM, B_2 = IT2FS, B_3 = IA-IT2FS, B_4 = Fuzzy Relational Approach, B_5 = PCA, B_6 = MLP, B_7 = RBFN

CA	Classification Accuracy tested by databases			Ranks obtained through experiments with databases			Average Rank (R_j)
	JAFFE	Indian	Cohn-Kanade	JAFFE	Indian	Cohn-Kanade	
A	97.5	100	97.5	1	1	1	1
B_1	90.33	97.57	88.11	4	2	5	3.667
B_2	90	92.5	92.5	5	4	3	4
B_3	92.5	95	95	3	3	2	2.667
B_4	87.5	92	90	6	5	4	5
B_5	95	87.5	87.5	2	6	6	4.667
B_6	72.5	75	72.5	7	7	7	7
B_7	65	67.5	67.5	8	8	8	8

Let r_i^j be the rank of *j*th algorithm on the ith data set. The average rank of algorithm *j* then is evaluated by

$$R_j = \frac{1}{N}\sum_{\forall i} r_i^j. \qquad (31)$$

The null hypothesis here states that all the algorithms are equivalent, so their individual ranks R_j should be equal. Under the null hypothesis, for large enough *N* and *k*, the Friedman statistic χ^2_F in (32) is distributed as a χ^2 with k-1 degrees of freedom. Here, *k* = 8 and *N* = 3. A larger *N* of course is desirable; however, emotion databases being fewer, finding large *N* is not feasible. Here, we consider percentage accuracy of classification as the basis of rank. Table X provides the percentage accuracy of classification with respect to three databases, JAFFE, Indian Woman (Jadavpur University), and Cohn-Kanade and the respective ranks of the algorithm

$$\chi_F^2 = \frac{12N}{k(k+1)}\left[\sum_j R_j^2 - \frac{k(k+1)^2}{4}\right]. \qquad (32)$$

Now, using *N* = 3, *k* = 8, and the ranks in Table X, we obtain $\chi_F^2 = 17.67 > \chi^2_{7,0.95} = 14.067$ [123], where $\chi^2_7, 0.95 =$ 14.067 is the value of the chi square distribution for 7° of freedom at probability of 0.05 [123]

$$\chi_F^2 = \frac{12N}{k(k+1)}\left[\sum_j R_j^2 - \frac{k(k+1)^2}{4}\right]$$
$$= 17.67 > \chi^2_{7,0.95}(14.067).$$

Thus, the hypothesis that the algorithms are equivalent is rejected. Therefore, the performances of the algorithms are determined by their ranks only. The order of ranking of the algorithm is apparent from their average ranks. The smaller the average rank, the better is the algorithm. Let "*>*" be a comparator of relative ranks where *x > y* means the algorithm *x* is better in rank than algorithm *y*. Table X indicates that the relative order of ranking of the algorithm by Friedman test as,

GT2FS>IA−IT2FS > SVM > IT2FS > PCA >Fuzzy Relational Approach > MLP > RBFN. It is clear from Table X that the average rank of GT2FS is 1 and average rank of IT2FS and IA-IT2FS are 4 and 2, respectively, claiming GT2FS outperforms all the algorithms by Friedman test.

CHAPTER 3

Sample run of Programs

We have to perform median filtering over an input bitmap image file: pic.bmp and generates an output file: med_file.bmp. The output image should be (103 * 77) pixels and removes noise from the source file to an appreciable extent.

After successfully running the program, again we have to read the image file 'med_file.bmp' (which was already created) and writes the pixel features (gray scale or any suitable transform of gray scale) in the file 'pix_dat.txt'. The Initial random seeds to be used as initial membership values for FCM classification are further generated and written in the file 'init_file.txt'.

After successfully running the above two programs, we have to perform image segmentation by fuzzy c-means clustering algorithm. It requires two input files: 'pix_dat.txt' & 'init_file.txt' (Already generated) and generates an output file cluster.bmp. The file 'pix_dat.txt' contains pixel information such as l', a', b' values and 'init_file.txt' contains initial membership of pixels to fall in the lip and non-lip class. The 'cluster.bmp' is the output file containing the segmented mouth region.

Finally we have to determine the average intensity in each row of the segmented mouth region and plots it against the height of the image. (Basically this program calculates the peak mouth opening. The input image file is 'cluster.bmp' and the output image file is 'intensity.bmp'

In this work, user's emotion using its facial expressions will be detected. These expressions can be derived from the live feed via system's camera or any pre-exisiting image available in the memory. Emotions possessed by humans can be recognized and has a vast scope of study in the computer vision industry upon which several researches have already been done. The work has been implemented using Python (2.7, Open Source Computer Vision Library (OpenCV) and NumPy. The scanned image(testing dataset) is being compared to the training dataset and thus emotion is predicted. The objective of this paper is to develop a system which can analyze the image and predict the expression of the person. The study proves that this procedure is workable and produces valid results.

OpenCV is Open Computer Vision Library [4]. It is a free for all extensive library which consists of more than 2500 algorithms specifically designed to carry out Computer Vision and Machine Learning related projects. These algorithms can be put to use to carry out different tasks such as Face Recognition, Object Identification, Camera Movement Tracking, Scenery Recognition etc.

For successful implementation of this project, the following packages of Python 2.7 have to be downloaded and installed: Python 2.7.x, NumPy, Glob and Random. Python will be installed in the default location, C drive in this case. Open Python IDLE, import all the packages and start working.

NumPy is one of the libraries of Python which is used for complex technical evaluation. It is used for implementation of multidimensional arrays which consists of various mathematical formulas to process. [9]. The array declared in a program has a dimension which is called as axis.

Random Module picks or chooses a random number or an element from a given list of elements. This module supports those functions which provide access to such operations.

Code:

```
import numpy as np
import cv2
import skfuzzy as fuzz
import matplotlib.pyplot as plt

def preprocess_image(input_image):
    img = cv2.imread(input_image)
    if img is None:
        raise ValueError("Image not found.")
    # Convert to grayscale for edge detection and to Lab for color
segmentation
    gray_img = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
    lab_img = cv2.cvtColor(img, cv2.COLOR_BGR2Lab)
    # Save the grayscale image resized to 103x77
    resized_gray_img = cv2.resize(gray_img, (103, 77))
    cv2.imwrite("med_file.bmp", resized_gray_img)
    return img, gray_img, lab_img[:, :, 1]

def extract_features_and_seeds(lab_img, feature_file, seed_file):
    L, a, b = cv2.split(lab_img)  # Split the Lab image into channels
    features = np.stack([L.flatten(), a.flatten(), b.flatten()], axis=1)
    # Save features for clustering
    np.savetxt(feature_file, features, fmt='%d %d %d')
    # Generate seeds for fuzzy c-means
    seeds = np.random.rand(features.shape[0], 2)

                        # Create two random membership values for each
pixel
    seeds[:, 1] = 1 - seeds[:, 0]
    np.savetxt(seed_file, seeds, fmt='%f %f')

def fuzzy_c_means(feature_file, num_clusters):
    pixel_data = np.loadtxt(feature_file)
    data = np.transpose(pixel_data)
    n_samples, n_features = data.shape
    # Randomly initialize centers
    np.random.seed(42)
    initial_centers = np.random.rand(num_clusters, n_features)
    cntr, u, _, _, _, _, _ = fuzz.cluster.cmeans(
        data, num_clusters, 2.0, error=0.005, maxiter=1000,
init=initial_centers)
    return cntr, u

def segment_mouth(img, gray_img, a_channel, output_image):
    red_mask = cv2.adaptiveThreshold(a_channel, 255,
cv2.ADAPTIVE_THRESH_MEAN_C, cv2.THRESH_BINARY_INV, 15, 10)
    edges = cv2.Canny(gray_img, 50, 150)
    combined_mask = cv2.bitwise_or(red_mask, edges)
    kernel = cv2.getStructuringElement(cv2.MORPH_ELLIPSE, (5, 5))
    morphed_mask = cv2.morphologyEx(combined_mask, cv2.MORPH_CLOSE,
kernel)
    morphed_mask = cv2.morphologyEx(morphed_mask, cv2.MORPH_OPEN, kernel)
```

```
    mouth_interior = cv2.bitwise_and(gray_img, gray_img,
mask=morphed_mask)
    _, interior_mask = cv2.threshold(mouth_interior, 30, 255,
cv2.THRESH_BINARY)
    output = np.full(img.shape, 255, dtype=np.uint8)
    output[interior_mask == 255] = [50, 50, 50]  # Gray
    cv2.imwrite(output_image, output)

def analyze_and_plot_intensity(output_image):
    img = cv2.imread(output_image, cv2.IMREAD_GRAYSCALE)
    average_intensity = np.mean(img, axis=1)
    plt.figure()
    plt.plot(average_intensity, range(len(average_intensity)))
    plt.gca().invert_yaxis()
    plt.title('Average Intensity vs. Image Height')
    plt.xlabel('Average Intensity')
    plt.ylabel('Image Height')
    plt.savefig('intensity_plot.png')
    plt.show()

def main():
    input_image = r"C:\Users\CDAC\Desktop\Suchismita\JU\PIC.bmp"
    feature_file = r"C:\Users\CDAC\Desktop\Suchismita\JU\pix_dat.txt"
    seed_file = r"C:\Users\CDAC\Desktop\Suchismita\JU\init_file.txt"
    output_image = r"C:\Users\CDAC\Desktop\Suchismita\JU\cluster.bmp"
    num_clusters = 2

    img, gray_img, a_channel = preprocess_image(input_image)
    lab_img = cv2.cvtColor(img, cv2.COLOR_BGR2Lab)
    extract_features_and_seeds(lab_img, feature_file, seed_file)
    cntr, u = fuzzy_c_means(feature_file, num_clusters)
    segment_mouth(img, gray_img, a_channel, output_image)
    analyze_and_plot_intensity(output_image)
print("Picture Coversion successful.")
print("The input file is PIC.bmp.")
print("The output file after median filtering is med_file.bmp.")
print("Height of the image is 77.")
print("Width of the image is 103.")
print("Two files named pix_dat.txt and init_file.txt generated
successfully.")
print("cluster.bmp is created successfully after applying fuzzy-C-means
clustering.")
print("Output file is intensity.bmp.")
print("The intensity data of the picture has been extracted.")

if __name__ == "__main__":
    main()
```

Sample Run:

C:\Python37\python.exe C:/Users/CDAC/Desktop/Suchismita/JU/All.py

Picture Coversion successful.

The input file is PIC.bmp.

The output file after median filtering is med_file.bmp.

Height of the image is 77.

Width of the image is 103.

Two files named pix_dat.txt and init_file.txt generated successfully.

cluster.bmp is created successfully after applying fuzzy-C-means clustering.

Output file is intensity.bmp.

The intensity data of the picture has been extracted.

Process finished with exit code 0

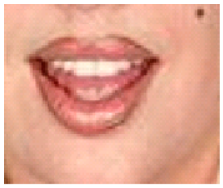


Fig. 1. Input image (PIC.bmp)



Fig. 2. Output image after median filtering (med_file.bmp)

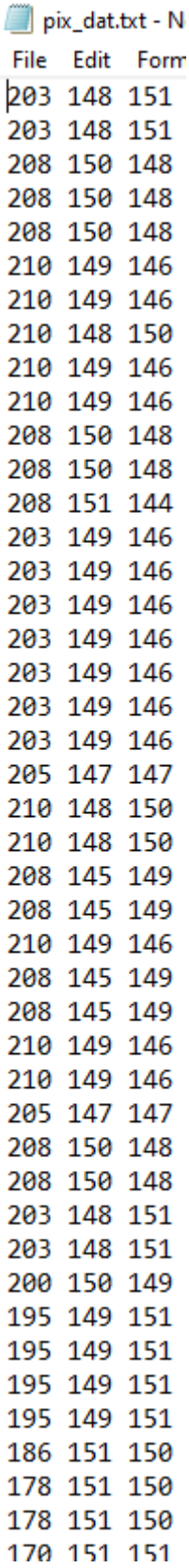


Fig. 3. Content of pix_dat.txt

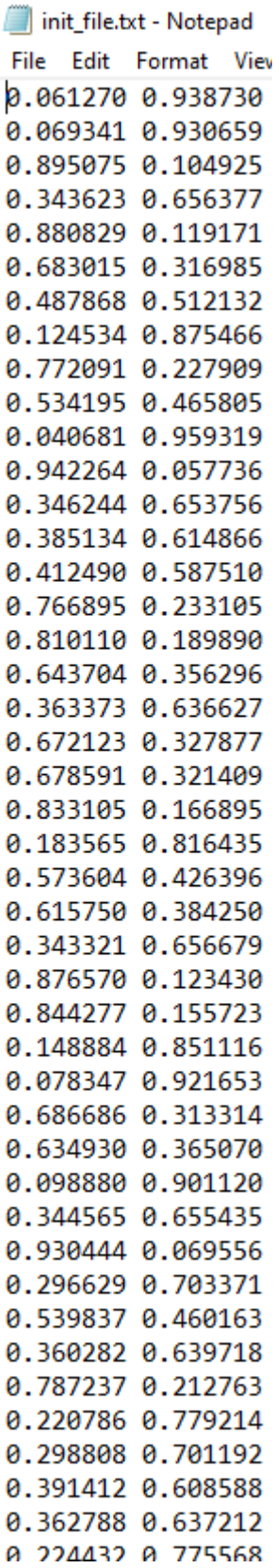


Fig. 4. Content of init_file.txt



Fig. 5. cluster.bmp

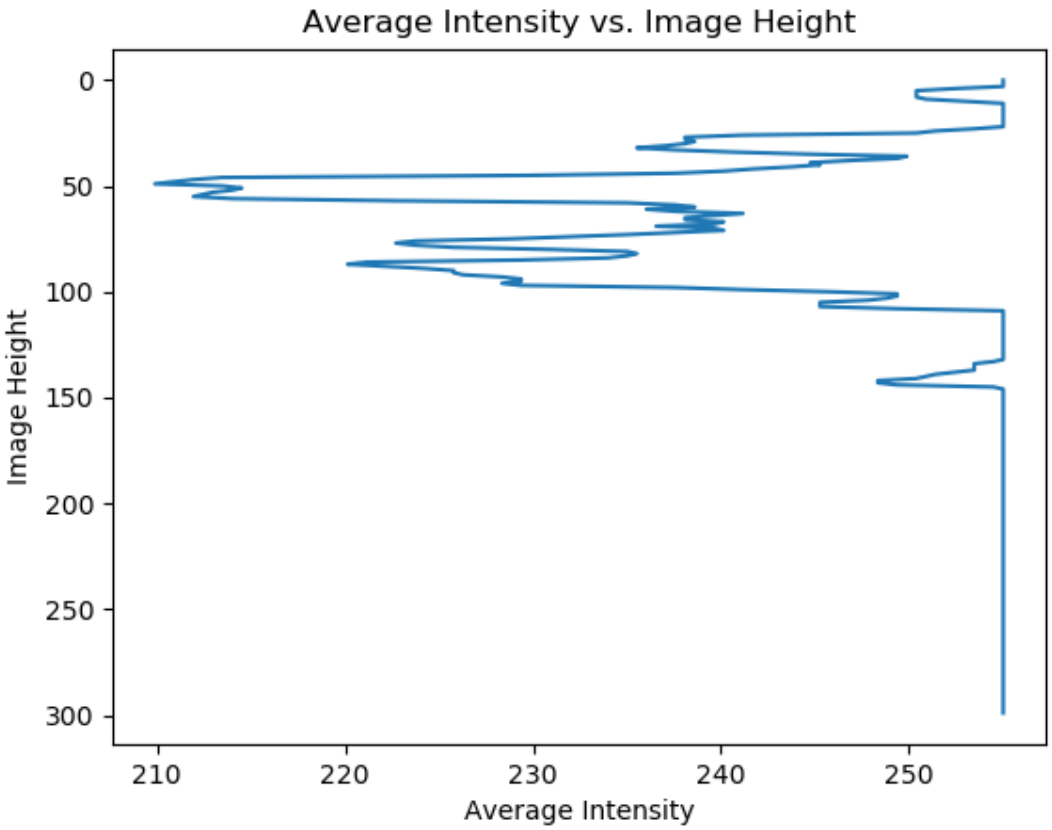


Fig. 6. Average Intensity VS. Image Height

CHAPTER 4

Conclusion & Future Scope

The merits of the proposed scheme for emotion detection lie in the segmentation of the mouth region by FCM clustering and determination of the MO from the minima in average- pixel intensity plot of the mouth region. The proposed EO determination also adds value to the emotion detection system. The fuzzy relational approach to emotion detection from facial feature space to emotion space has a high classification accuracy of around 90%, which is better than other reported results. Because of its good classification accuracy, the proposed emotion detection scheme is expected to have applications in next-generation human–machine interactive systems.

The ranking of audiovisual stimulus considered in this project also provides a new approach to determining the best movies to excite specific emotions.

The existing emotion detection methods hardly consider the effect of near-past stimulation on the current arousal of emotion. To overcome this problem, we submitted an audiovisual stimulus of relaxation before submission of any other movie clip to excite emotions. A state transition in emotion by an audiovisual movie thus always occurs at a relaxed state of mind, giving full effect of the current stimulus on the excitatory sub-system of the brain and causing arousal of the desired emotion with its full manifestation on the facial expression. Feature extraction from the face becomes easy when the manifestation of facial expression truly resembles the aroused emotion.

An important aspect of this work is the design of an emotion control scheme. The accuracy of the control scheme ensures convergence of the control algorithm with a zero error, and repeatability ensures the right selection of audiovisual stimulus. The proposed scheme of emotion recognition and control can be applied in system design for two different problem domains. First, it can serve as an intelligent layer in the next-generation human–machine interactive system. Such a system would have extensive applications in the frontier technology of pervasive and ubiquitous computing [42]. Second, the emotion monitoring and control scheme would be useful for psycho-logical counseling and therapeutic applications. The pioneering works on the “structure of emotion” by Gordon [14] and the “emotional control of cognition” by Simon [36] would find a new direction with the proposed automation for emotion recognition and control.

The project presents three automatic emotion recognition systems based on IT2FS, IA-IT2FS, and GT2FS. In order to classify an unknown facial expression, these systems make use of the background knowledge about a large face database with known emotion classes. The GT2FS-based recognition scheme requires T2 secondary membership functions, which are obtained using an innovative evolutionary approach that is also proposed in this work. All the schemes first construct a fuzzy face space, and then infer the emotion class of the unknown facial expression by determining the maximum support of the individual emotion classes using the pre-constructed fuzzy face space. The class with the highest support is assigned as the emotion of the unknown facial expression.

The IT2FS-based recognition scheme takes care of the inter subject level uncertainty in computing the maximum support of individual emotion class. The GT2FS-based recognition scheme, however, takes care of both the inter- and intra-subject level uncertainty, and thus offers higher classification accuracy for the same set of features. Using three data sets, the classification accuracy obtained by employing GT2FS is 98.333%, by IT2FS is 91.667%, and by IA-IT2FS is 94.167%.

The more the number of subjects used for constructing the fuzzy face space, the better would be the fuzzy face space, and thus better would be the classification accuracy. Since the fuzzy face space is created offline, the online computational load to recognize emotion is insignificantly small in IT2FS. The computational load in GT2FS, however, is large as it includes an optimization procedure to determine the secondary membership for each emotion and for each subject. However, this additional complexity in GT2FS, offers approximately 7% improvement in classification accuracy in comparison to that by IT2FS. The IA-IT2FS has around 2.5% gain in classification accuracy with no more additional computational complexity than IT2FS. It may be noted that the necessary computations in IA-IT2FS for data filtering and membership function selection is performed offline. The statistical tests employed clearly indicate that GT2FS outperforms the seven selected algorithms.

The problems that may be taken up as future research are briefly outlined below. First, new alternative strategies are to be designed to determine secondary memberships without using optimization techniques. Second, a more formal and systematic approach to fuse secondary and primary memberships to reduce uncertainty in the fuzzy face space is to be developed. Last, we would try

to explore the power of fuzzy logic to determine emotion classes in absence of sufficient (or even no) measurements. Facial features, for example MO, may be directly encoded into fuzzy features with fuzzy sets, such as “a little,” “more,” and “not so large,” and then an IT2FS-based model may be adopted to recognize emotion of unknown subjects.

Classification accuracy under this circumstance could be poor, but a more human-like interpretation of emotion can be given in the absence of precise measurements.

The perception of an ‘emotion’ is still the major lacunae of the FER systems. The manifestation of emotions differs across cultures, ethnic groups, and geographical locations and sometimes even between individuals. The observations in this work highlight the significance of different configurations of fuzzy inference systems, aiding the design of an apt FER system for understanding a perception of emotion. The future scope of this work is to evaluate the performance of a fuzzy-based FER system with a combination of heterogeneous membership functions for more datasets and diverse feature subsets to obtain robust recognition systems.

REFERENCE

[1] J. C. Bezdek, "Fuzzy mathematics in pattern classification," Ph.D. dissertation, Appl. Math. Center, Cornell Univ., Ithaca, NY, 1973.

[2] B. Biswas, A. K. Mukherjee, and A. Konar, "Matching of digital images using fuzzy logic," *AMSE Publication*, vol. 35, no. 2, pp. 7–11, 1995.

[3] M. T. Black and Y. Yacoob, "Recognizing facial expressions in image sequences using local parameterized models of image motion," *Int. J. Comput. Vis.*, vol. 25, no. 1, pp. 23–48, Oct. 1997.

[4] C. Busso and S. Narayanan, "Interaction between speech and facial gestures in emotional utterances: A single subject study," *IEEE Trans. Audio, Speech Language Process.*, vol. 15, no. 8, pp. 2331–2347, Nov. 2007.

[1] I. Cohen, "Facial expression recognition from video sequences," M.S. thesis, Univ. Illinois Urbana-Champaign, Dept. Elect. Eng., Urbana, IL, 2000.

[2] I. Cohen, N. Sebe, A. Garg, L. S. Chen, and T. S. Huang, "Facial expression recognition from video sequences: Temporal and static mod- eling," *Comput. Vis. Image Underst.*, vol. 91, no. 1/2, pp. 160–187, Jul. 2003.

[3] C. Conati, "Probabilistic assessment of user's emotions in educational games," *J. Appl. Artif. Intell., Special Issue Merging Cognition Affect HCT*, vol. 16, no. 7/8, pp. 555–575, Aug. 2002.

[4] G. Donato, M. S. Bartlett, J. C. Hager, P. Ekman, and T. J. Sejnowski, "Classifying facial actions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 10, pp. 974–989, Oct. 1999.

[5] P. Ekman and W. V. Friesen, *Unmasking the Face: A Guide to Recognizing Emotions From Facial Clues*. Englewood Cliffs, NJ: Prentice-Hall, 1975.

[6] I. A. Essa and A. P. Pentland, "Coding, analysis, interpretation and recognition of facial expressions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 757–763, Jul. 1997.

[7] W. A. Fellenz, J. G. Taylor, R. Cowie, E. Douglas-Cowie, F. Piat,

S. Kollias, C. Orovas, and B. Apolloni, "On emotion recognition of faces and of speech using neural networks, fuzzy logic and the ASSESS systems," in *Proc. IEEE -INNS-ENNS Int. Joint Conf. Neural Netw.*, 2000, pp. 93–98.

[8] J. M. Fernandez-Dols, H. Wallbotl, and F. Sanchez, "Emotion cate- gory accessibility and the decoding of emotion from facial expression and context," *J. Nonverbal Behav.*, vol. 15, no. 2, pp. 107–123, Jun. 1991.

[9] Y. Gao, M. K. H. Leung, S. C. Hui, and M. W. Tananda, "Facial expression recognition from line-based caricatures," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 33, no. 3, pp. 407–412, May 2003.

[10] R. N. Gordon, *The Structure of Emotions: Investigations in Cognitive Philosophy*, ser. Cambridge Studies in Philosophy. Cambridge, U.K.: Cambridge Univ. Press, 1990.

[11] K. Izumitani, T. Mikami, and K. Inoue, "A model of expression grade for face graphs using fuzzy integral," *Syst. Control*, vol. 28, no. 10, pp. 590–596, 1984.

[12] F. Kawakami, S. Morishima, H. Yamada, and H. Harashima, "Construction of 3-D emotion space using neural network," in *Proc. 3rd Int. Conf. Fuzzy Logic, Neural Nets Soft Comput.*, Iizuka, Japan, 1994, pp. 309–310.

[13] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[14] H. Kobayashi and F. Hara, "The recognition of basic facial expressions by neural network," *Trans. Soc. Instrum. Contr. Eng.*, vol. 29, no. 1, pp. 112–118, 1993.

[15] H. Kobayashi and F. Hara, "Measurement of the strength of six basic facial expressions by neural network," *Trans. Jpn. Soc. Mech. Eng. (C)*, vol. 59, no. 567, pp. 177–183, 1993.

[16] H. Kobayashi and F. Hara, "Recognition of mixed facial expressions by neural network," *Trans. Jpn. Soc. Mech. Eng. (C)*, vol. 59, no. 567, pp. 184–189, 1993.

[17] A. Konar, *Computational Intelligence: Principles, Techniques and Applications*. Heidelberg, Germany: SpringerVerlag, 2005.

[18] A. F. Kramer, E. J. Sirevaag, and R. Braune, "A psycho-physiological assessment of operator workload during simulated flight missions," *Hum. Factors*, vol. 29, no. 2, pp. 145–160, Apr. 1987.

[19] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Automatic interpretation and coding of face images using flexible models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 743–756, Jul. 1997.

[20] H. Li, P. Roivainen, and R. Forchheimer, "3D motion estimation in model-based facial image coding," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 6, pp. 545–555, Jun. 1993.

[21] X. Li and Q. Ji, "Active affective state detection and user assistance with dynamic Bayesian networks," *IEEE Trans. Syst., Man, Cybern. A, Syst.,*

Humans, vol. 35, no. 1, pp. 93–105, Jan. 2005.

[22] K. Mase, "Recognition of facial expression from optical flow," *Proc. IEICE Trans., Special Issue Comput. Vis. Appl.*, vol. 74, no. 10, pp. 3474– 3483, 1991.

[23] K. Ogata, *Modern Control Engineering*. Englewood Cliffs, NJ: Prentice- Hall, 1990.

[24] W. Pedrycz and J. Valente de Oliveira, "A development of fuzzy encoding and decoding through fuzzy clustering," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 4, pp. 829–837, Apr. 2008.

[25] W. Pedrycz and F. Gomide, *An Introduction to Fuzzy Sets: Analysis and Design*. Cambridge, MA: MIT Press, 1998.

[26] R. Picard, *Affective Computing*. Cambridge, MA: MIT Press, 1997.

[27] R. W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: Analysis of affective psychological states," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 10, pp. 1175–1191, Oct. 2001.

[28] P. Rani, N. Sarkar, and J. Adams, "Anxiety-based affective communication for implicit human–machine interaction," *Adv. Eng. Inf.*, vol. 21, no. 3, pp. 323–334, Jul. 2007.

[29] P. Rani, N. Sarkar, C. Smith, and L. Kirby, "Anxiety detecting robotic systems—Towards implicit human–robot collaboration," *Robotica*, vol. 22, no. 1, pp. 83–93, 2004.

[30] M. Rosenblum, Y. Yacoob, and L. Davis, "Human expression recognition from motion using a radial basis function network architecture," *IEEE Trans. Neural Netw.*, vol. 7, no. 5, pp. 1121–1138, Sep. 1996.

[31] J. Scheirer, R. Fernadez, J. Klein, and R. Picard, "Frustrating the user on purpose: A step toward building an affective computer," *Interact. Comput.*, vol. 14, no. 2, pp. 93–118, Feb. 2002.

[32] H. Simon, *Motivational and Emotional Control of Cognition, Models of Thought*. New Haven, CT: Yale Univ. Press, 1979, pp. 29–38.

[33] D. Terzopoulos and K. Waters, "Analysis and synthesis of facial image sequences using physical and anatomical models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 6, pp. 569–579, Jun. 1993.

[34] Y. Tian, T. Kanade, and J. Cohn, "Recognizing action units for facial expression analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 2, pp. 97–115, Feb. 2001.

[35] N. Ueki, S. Morishima, and H. Harashima, "Expression analysis/synthesis system based on emotion space constructed by multilayered neural network," *Syst. Comput. Jpn.*, vol. 25, no. 13, pp. 95–103, 1994.

[36] O. A. Uwechue and S. A. Pandya, *Human Face Recognition Using Third-Order Synthetic Neural Networks*. Boston, MA: Kluwer, 1997.

[37] P. Vanger, R. Honlinger, and H. Haykin, "Applications of synergetic in decoding facial expressions of emotions," in *Proc. Int. Workshop Autom. Face Gesture Recog.*, Zurich, Switzerland, 1995, pp. 24–29.

[38] A. Vasilakos and W. Pedrycz, *Ambient Intelligence, Wireless Net- working and Ubiquitous Computing*. Norwood, MA: Artech House, Jun. 2006.

[39] Y. Yacoob and L. Davis, "Computing spatio-temporal representations of human faces," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, Jun. 1994, pp. 70–75.

[40] Y. Yacoob and L. Davis, "Recognizing human facial expression from long image sequences using optical flow," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 6, pp. 636–642, Jun. 1996.

[41] H. Yamada, "Visual information for categorizing facial expression of emotion," *Appl. Cogn. Psychol.*, vol. 7, no. 3, pp. 257–270, 1993.

[42] Z. Zeng, Y. Fu, G. I. Roisman, Z. Wen, Y. Hu, and T. S. Huang, "Spontaneous emotional facial expression detection," *J. Multimedia*, vol. 1, no. 5, pp. 1–8, Aug. 2006……

[43] D.-J. Kim, S.-W. Lee, and Z. Bien, "Facial emotional expression recognition with soft computing techniques," in *Proc. 14th IEEE Int. Conf. Fuzzy Syst.*, 2005, pp. 737–742.

[44] R. Cowie, E. Douglas-Cowie, J. G. Taylor, S. Ioannou, M. Wallace, and S. Kollias, "An intelligent system for facial emotion recognition," in *Proc. IEEE ICME*, Amsterdam, The Netherlands, Jul. 2005, pp. 904–907.

[45] A. Bhavsar and H. M. Patel, "Facial expression recognition using neural classifier and fuzzy mapping," in *Proc. IEEE INDICON*, Chennai, India, 2005, pp. 383–387.

[46] Y. Guo and H. Gao, "Emotion recognition system in images based on fuzzy neural network and HMM," in *Proc. 5th IEEE ICCI*, 2006, pp. 73–78.

[47] M. H. Bindu, P. Gupta, and U. S. Tiwary, "Cognitive model based emotion recognition from facial expressions for live human computer interaction," in *Proc. IEEE Symp.CIISP*, 2007, pp. 351–356.

[48] M. Rizon, M. Karthigayan, S. Yaacob, and R. Nagarajan, "Japanese face emotions classification using lip features," in *Proc. GMAI*, 2007, pp. 140–144.

[49] N. Esau, E. Wetzel, L. Kleinjohann, and B. Kleinjohann, "Real-time facial expression recognition using a fuzzy emotion model," in *Proc. IEEE Int. Fuzzy Syst. Conf.*, 2007, pp. 1–6.

[50] H. Zhao, Z. Wang, and J. Men, "Facial complex expression recognition based on fuzzy kernel clustering and support vector machines," in *Proc. 3rd ICNC*, 2007, pp. 562–566.

[51] Y. Lee, C. W. Han, and J. Shim, "Fuzzy neural networks and fuzzy integral approach to curvature-based component range facial recognition," in *Proc. Int. Conf. Converg. Inf. Technol.*, 2007, pp. 1334–1339.

[52] A. Khanam, M. Z. Shafiq, and M. U. Akram, "Fuzzy based facial expression recognition," in *Proc. Congr. Image Signal Process.*, 2008, pp. 598–602.

[53] G. U. Kharat and S. V. Dudul, "Neural network classifier for human emotion recognition from facial expressions using discrete cosine transform," in *Proc. 1st Int. Conf. Emerging Trends Eng. Technol.*, 2008, pp. 653–658.

[54] J. M. Sun, X. S. Pei, and S. S. Zhou, "Facial emotion recognition in modern distant system using SVM," in *Proc. 7th Int. Conf. Mach. Learn. Cybern.*, Kunming, China, Jul. 2008, pp. 3545–3548.

[55] H. Zhao and Z. Wang, "Facial action units recognition based on fuzzy kernel clustering," in *Proc. 5th Int. Conf. Fuzzy Syst. Knowl. Discov.*, 2008, pp. 168–172.

[56] A. Chakraborty, A. Konar, U. K. Chakraborty, and A. Chatterjee, "Emotion recognition from facial expressions and its control using fuzzy logic," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 39, no. 4, pp. 726– 743, Jul. 2009.

[57] H. Kobayashi and F. Hara, "Measurement of the strength of six basic facial expressions by neural network," *Trans. Jpn. Soc. Mech. Eng. (C)*, vol. 59, no. 567, pp. 177–183, 1993.

[58] H. Kobayashi and F. Hara, "Recognition of mixed facial expressions by neural network," *Trans. Jpn. Soc. Mech. Eng. (C)*, vol. 59, no. 567, pp. 184–189, 1993.

[59] H. Kobayashi and F. Hara, "The recognition of basic facial expressions by neural network," *Trans. Soc. Instrum. Contr. Eng*, vol. 29, no. 1, pp. 112– 118, 1993.

[60] Y. Sun, Z. Li, and C. Tang, "An evolving neural network for authentic emotion classification," in *Proc. 5th Int. Conf. Natural Comput.*, 2009, pp. 109–113.

[61] C. Chang, Y. Huang, and C. Yang, "Personalized facial expression recognition in color image," in *Proc. 4th Int. Conf. Innov. Comput., Inf. Contr.*, 2009, pp. 1164–1167.

[62] M. Fu, Y. Kuo, and K. Lee, "Fusing remote control usages and facial expression for emotion recognition," in *Proc. 4th Int. Conf. Innov. Comput., Inf. Contr.*, 2009, pp. 132–135.

[63] C. Huang and Y. Huang, "Facial expression recognition using modelbased feature extraction and action parameters classification," *J. Vis.*

Commun. Image Represent., vol. 8, no. 3, pp. 278–290, Sep. 1997.

[64] K. Huang, S. Huang, and Y. Kuo, "Emotion recognition based on a novel triangular facial feature extraction method," in *Proc. IJCNN*, 2010, pp. 1–6.

[65] H. Tsai, Y. Lai, and Y. Zhang, "Using SVM to design facial expression recognition for shape and texture features," in *Proc. 9th Int. Conf. Mach. Learn. Cybern.*, Qingdao, China, Jul. 11–14, 2010, pp. 2697–2704.

[66] P. Ekman and W. V. Friesen, *Unmasking the Face: A Guide to Recognizing Emotions From Facial Clues*. Englewood Cliffs, NJ: Prentice-Hall, 1975.

[67] O. A. Uwechue and S. A. Pandya, *Human Face Recognition Using ThirdOrder Synthetic Neural Networks*. Boston, MA: Kluwer, 1997.

[68] B. Biswas, A. K. Mukherjee, and A. Konar, "Matching of digital images using fuzzy logic," *AMSE Publ.*, vol. 35, no. 2, pp. 7–11, 1995.

[69] M. Paleari, R. Chellali, and B. Huet, "Features for multimodal emotion recognition: an extensive study," in *Proc. CIS*, 2010, pp. 90–95.

[70] S. Das, A. Halder, P. Bhowmik, A. Chakraborty, A. Konar, and

A. K. Nagar, "Voice and facial expression based classification of emotion using linear support vector," in *Proc. 2nd Int. Conf. Develop. Syst. Eng.*, 2009, pp. 377–384.

[71] M. Gargesha and P. Kuchi, "Facial expression recognition using a neural network," *Artif. Neural Comput. Syst.*, pp. 1–6, 2002.

[72]J. M. Mendel, "Fuzzy sets for words: A new beginning," in *Proc. IEEE Int. Conf. Fuzzy Syst.*, 2003, pp. 37–42.

[73]Y. Zhang, Q. J. I., "Active and dynamic information fusion for facial expression understanding from image sequences," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 5, pp. 699–714, May 2005.

[74]C. Liu and H. Wechsler, "Probabilistic reasoning models for face recognition," in *Proc. Comput. Vis. Pattern Recog. Conf.*, 1998, pp. 827–832.

[75]G. U. Kharat and S. V. Dudul, "Emotion recognition from facial expression using neural networks," in *Proc. Human Comput. Sys. Interact.*, 2009, pp. 207–219.

[76]K. V. Price, R. M. Storn, and J. A. Lampinen, Eds., *Differential Evolution: A Practical Approach*. New York: Springer-Verlag, 2005.

[77]A. Konar and A. Chakraborty, *Emotional Intelligence: A Cybernetic Approach*. Heidelberg, Germany: Springer-Verlag, 2009.

[78]H. J. Zimmermann, *Fuzzy Set Theory and its Applications*. Berlin, Germany: Springer-Verlag, 2001.

[79]J. M. Mendel and H. Wu, "Type-2 Fuzzistics for symmetric interval type2 fuzzy sets: Part 1, forward problems," *IEEE Trans. Fuzzy Syst.*, vol. 14, no. 6, pp. 718–792, Dec. 2006.

[80]J. M. Mendel and D. Wu, *Perceptual Computing*. Piscataway, NJ: IEEE Press, 2010.

[81]J. M. Mendel and R. I. John, "Type-2 fuzzy sets made simple," *IEEE Trans. Fuzzy Syst.*, vol. 10, no. 2, pp. 117–127, Apr. 2002.

[82]A. Konar and A. Chakraborty, Eds., *Advances in Emotion Recognition*. New York: Wiley, 2011.

[83]J. M. Mendel, R. I. John, and F. Liu, "Interval type-2 fuzzy logic systems made simple," *IEEE Trans. Fuzzy Syst.*, vol. 14, no. 6, pp. 808–821, Dec. 2006.

[84]J. M. Mendel, "On the importance of interval sets in type-2 fuzzy logic systems," in *Proc. Joint 9th IFSA World Congr. 20th NAFIPS Int. Conf.*, Vancouver, BC, Canada, Jul. 25–28, 2001, pp. 1647–1652.

[85]F. Kawakami, S. Morishima, H. Yamada, and H. Harashima, "Construction of 3-D emotion space using neural network," in *Proc. 3rd Int. Conf. Fuzzy Logic, Neural Nets Soft Comput.*, Iizuka, Japan, 1994, pp. 309–310.

[86] H. Wu, Y. Wu, and J. Luo, "An interval type-2 fuzzy rough set model for attribute reduction," *IEEE Trans. Fuzzy Syst.*, vol. 17, no. 2, pp. 301–315, Apr. 2009.

[87] P. N. Suganthan, N. Hansen, J. J. Liang, K. Deb, Y.-P. Chen, A. Auger, and S. Tiwari, "Problem definitions and evaluation criteria for the CEC 2005 special session on real-parameter optimization," Nanyang Technol. Univ., Singapore, 2005, Tech. Rep..

[88] J. Romero Agüero and A. Vargas, "Calculating functions of interval type-2 fuzzy numbers for fault current analysis," *IEEE Trans. Fuzzy Syst.*, vol. 15, no. 1, pp. 31–40, Feb. 2007.

[89] O. Cordón, M. J. del Jesus, and F. Herrera, "A proposal on reasoning methods in fuzzy rule-based classification systems," *Int. J. Approx. Reasoning*, vol. 20, no. 1, pp. 21–45, Jan. 1999.

[90] C.-H. Lee, J.-L. Hong, Y.-C. Lin, and W.-Y. Lai, "Type-2 fuzzy neural network systems and learning," *Int. J. Comput. Cogn.*, vol. 1, no. 4, pp. 79–90, Dec. 2003.

[91] H. Yamada, "Visual information for categorizing facial expression of emotion," *Appl. Cogn. Psychol.*, vol. 7, no. 3, pp. 257–270, Jun. 1993.

[92] J. M. Fernandez-Dols, H. Wallbotl, and F. Sanchez, "Emotion category accessibility and the decoding of emotion from facial expression and context," *J. Nonverbal Behav.*, vol. 15, no. 2, pp. 107–123, Jun. 1991.

[93] C. Busso and S. Narayanan, "Interaction between speech and facial gestures in emotional utterances: A single subject study," *IEEE Trans. Audio, Speech Language Process.*, vol. 15, no. 8, pp. 2331–2347, Nov. 2007.

[94] I. Cohen, N. Sebe, A. Garg, L. S. Chen, and T. S. Huang, "Facial expression recognition from video sequences: Temporal and static modeling," *Comput. Vis. Image Understand.*, vol. 91, no. 1/2, pp. 160–187, Jul. 2003.

[95] Y. Gao, M. K. H. Leung, S. C. Hui, and M. W. Tananda, "Facial expression recognition from line-based caricatures," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 33, no. 3, pp. 407–412, May 2003.

[96] Z. Michalewicz, *Genetic Algorithms + Data Structure = Evolution Programs*. Berlin, Germany: Springer-Verlag, 1992.

[97] P. Engelbrecht, *Fundamentals of Computational Swarm Intelligence*. New York: Wiley, 2005.

[98] R. W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 10, pp. 1175–1191, Oct. 2001.

[99] M. Pantic and I. Patras, "Dynamics of facial expression: Recognition of facial actions and their temporal segments from face profile image sequences," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 36, no. 2, pp. 433–449, Apr. 2006.

[100] J. Demsar, "Statistical comparisons of classifiers over multiple data sets," *J. Mach. Learn. Res.*, vol. 7, pp. 1–30, Dec. 2006.

[101] T. G. Dietterich, "Approximate statistical tests for comparing supervised classification learning algorithms," *Neural Comput.*, vol. 10, no. 7, pp. 1895–1923, Oct. 1998.

[102] A. Halder, P. Rakshit, S. Chakraborty, A. Konar, A. Chakraborty, E. Kim, and A. K. Nagar, "Reducing uncertainty in interval type-2 fuzzy sets for qualitative improvement in emotion recognition from facial expressions," in *Proc. IEEE FUZZ*, Brisbane, Australia, 2012, to be published.

[103] D. Datcu and L. J. M. Rothkrantz, "Emotion recognition using bimodal data fusion," in *Proc. 12th Int. Conf. Comput. Syst. Technol.*, 2011, pp. 122–128.

[104] M. Zia Uddin, J. J. Lee, and T.-S. Kim, "Independent shape componentbased human activity recognition via hidden Markov model," *Appl. Intell.*, vol. 33, no. 2, pp. 193–206, Oct. 2010.

[105] Y. Yang, Z. Chen, Z. Liang, and G. Wang, "Attribute reduction for massive data based on rough set theory and mapReduce," in *Proc. RSKT*, 2010, pp. 672–678.

[106] C.-Y. Chang, S.-C. Li, P.-C. Chung, J.-Y. Kuo, and Y.-C. Tu, "Automatic facial skin defect detection system," in *Proc. BWCCA*, 2010, pp. 527–532.

[107] H. Zhang, D. Liu, and Z. Wang, *Controlling Chaos: Suppression, Synchronization and Chaotification*. New York: Springer-Verlag, 2009.

[108] L. A. Zadeh, "Fuzzy sets," *Inf. Contr.*, vol. 8, no. 3, pp. 338–353, Jun. 1965.

[109] M. Pantic and L. J. M. Rothkrantz, "Automatic analysis of facial expressions: The state of the art," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 12, pp. 1424–1445, Dec. 2000.

[110] F. Liu and J. M. Mendel, "Encoding words into interval type-2 fuzzy sets using an interval approach," *IEEE Trans. Fuzzy Syst.*, vol. 16, no. 6, pp. 1503–1521, Dec. 2008.

[111] L. Guan, S.-Y. Kung, and J. Larsen, *Multimedia Image and Video Processing*. Boca Raton, FL: CRC Press, 2000.

[112] A. Metallinou, A. Katsamanis, and S. Narayanan, "A hierarchical framework for modeling multimodality and emotional evolution in affective dialogs," in *Proc. ICASSP*, Kyoto, Japan, Mar. 2012, pp. 2401–2404.

[113] A. Metallinou, M. Wollmer, A. Katsamanis, F. Eyben, B. Schuller, and S. Narayanan, "Context-sensitive learning for enhanced audiovisual emotion classification," *IEEE Trans. Affect. Comput.*, vol. 3, no. 2, pp. 184– 198, 2012.

[114] A. Metallinou, A. Katsamanis, Y. Wang, and S. Narayanan, "Tracking changes in continuous emotion states using body language and prosodic cues," in *Proc. ICASSP*, 2011, pp. 2288–2291.

[115] A. Metallinou, C. Busso, S. Lee, and S. S. Narayanan, "Visual emotion recognition using compact facial representations and viseme information," in *Proc. ICASSP*, 2010, pp. 2474–2477.

[116] A. Metallinou, S. Lee, and S. S. Narayanan, "Decision level combination of multiple modalities for recognition and analysis of emotional expression," in *Proc. ICASSP*, 2010, pp. 2462–2465.

[117] C.-C. Lee, E. Mower, C. Busso, S. Lee, and S. Narayanan, "Emotion recognition using a hierarchical binary decision tree approach," *Speech Commun.*, vol. 53, no. 9/10, pp. 1162–1171, Nov. 2011.

[118] E. Mower, M. Mataric, and S. Narayanan, "Human perception of audiovisual synthetic character emotion expression in the presence of ambiguous and conflicting information," *IEEE Trans. Multimedia*, vol. 11, no. 5, pp. 843–855, Aug. 2009.

[119] M. Grimm, E. Mower, K. Kroschel, and S. Narayanan, "Primitives based estimation and evaluation of emotions in speech," *Speech Commun.*, vol. 49, no. 10/11, pp. 787–800, Oct. 2007.

[120] C. M. Lee and S. Narayanan, "Emotion recognition using a data-driven fuzzy inference system," in *Proc. EUROSPEECH*, Geneva, Switzerland, 2003, pp. 157–160.

[121] C. Busso, M. Bulut, C.-C. Lee, A. Kazemzadeh, E. Mower, S. Kim, J. Chang, S. Lee, and S. Narayanan, "Interactive emotional dyadic motion capture

database," *J. Language Resourc. Eval.*, vol. 42, no. 4, pp. 335– 359, Dec. 2008.

[122] A. Metallinou, C.-C. Lee, C. Busso, S. Carnicke, and S. Narayanan, "The USC Creative IT database: A multimodal database of theatrical improvisation," in *Proc. MMC*, Valletta, Malta, May 2010.

[123] J. N. Kapur and H. C. Saxena, *Mathematical Statistics*. New Delhi, India: S. Chand & Company Ltd., 1980.

[124] M. Pantic, R. Cowie, F. D'Errico, D. Heylen, M. Mehu, C. Pelachaud, I. Poggi, M. Schroeder, and A. Vinciarelli, "Social Signal Processing: The Research Agenda," Vis. Anal. Humans, pp. 511–538, 2011.

[125]] A.Vinciarelli, H. Salamin, and M. Pantic, "Social Signal Processing: Understanding Social Actions through nonverbal behaviour analysis," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work. CVPR Work. 2009, vol. 3, no. 231287, pp. 42–49, 2009.

[126] Vinciarelli, M. Pantic, H. Bourlard, and A. Pentland, "Social signals, their function, and automatic analysis: A survey," Proc. 10th Int. Conf. Multimodal interfaces - IMCI '08, pp. 61–68, 2008.

[127]] D. Y. Liliana, M. R. Widyanto, and T. Basaruddin, "Human emotion recognition based on active appearance model and semisupervised fuzzy C-means," in 2016 International Conference on Advanced Computer Science and Information Systems, ICACSIS 2016, 2017, pp. 439–445.

[128] L. Wang, R. Li, and K. Wang, "A Novel Automatic Facial Expression Recognition Method Based on AAM," Journal of Computers, vol. 9, no. 3, pp. 608–617, 2014.

[129] X. Li, Q. Ruan, Y. Jin, G. An, and R. Zhao, "Fully automatic 3D facial expression recognition using polytypic multi-block local binary patterns," Signal Processing, vol. 108, pp. 297–308, 2015.

[130] M. H. Siddiqi, R. Ali, A. M. Khan, Y. T. Park, and S. Lee, "Human Facial Expression Recognition Using Stepwise Linear Discriminant Analysis and Hidden Conditional Random Fields," IEEE Trans. Image Processing, vol. 24, no. 4, pp. 1386–1398, 2015.

[131] K. Anderson and P. W. McOwan, "A real-time automated system for the recognition of human facial expressions.," IEEE Trans. Syst. Man. Cybern. B. Cybern., vol. 36, no. 1, pp. 96–105, 2006.

[132] Z. Ming, J. Rouas, T. Shochi, Z. Ming, J. Rouas, T. Shochi, "Facial Action Units Intensity Estimation by the Fusion of Features with Multi-kernel Support Vector Machine", 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), 2015.

[133] P. Khorrami, T. Le Paine, K. Brady, C. Dagli, and T. S. Huang, "How Deep Neural Networks Can Improve Emotion Recognition on Video Data," IEEE International Conference on Image Processing (ICIP), pp. 1–5, 2016.

[134] D. A. Pitaloka, A. Wulandari, T. Basaruddin, and D. Y. Liliana, "Enhancing CNN with Preprocessing Stage in Automatic Emotion Recognition," Procedia Computer Science, vol. 116, pp. 523–529, 2017.

[135] C. E. Izard, "Basic emotions, relations among emotions, and emotion-cognition relations.," Psychological Review, vol. 99, no. 3, pp. 561–5, 1992.

[136] S. Du, Y. Tao, and A. M. Martinez, "Compound facial expressions of emotion," Proceedings of the National Academy of Sciences, vol. 111, no. 15, pp. E1454–E1462, Apr. 2014.

[137] D. Y. Liliana, C. Basaruddin, and M. R. Widyanto, "Mix Emotion Recognition from Facial Expression using SVM-CRF Sequence Classifier," Proc. Int. Conf. Algorithms, Comput. Syst. - ICACS '17, pp. 27–31, 2017.

[138] A. Chakraborty, A. Konar, U. K. Chakraborty, and A. Chatterjee, "Emotion Recognition From Facial Expressions and Its Control Using Fuzzy Logic," IEEE Trans. Syst. Man, Cybern. - Part A Syst. Humans, vol. 39, no. 4, pp. 726–743, 2009.

[139] A. Khanum, M. Mufti, M. Y. Javed, and M. Z. Shafiq, "Fuzzy case-based reasoning for facial expression recognition," Fuzzy Sets and System, vol. 160, pp. 231–250, 2009.

[140] M. Ilbeygi and H. Shah-Hosseini, "A novel fuzzy facial expression recognition system based on facial feature extraction from color face images," Engineering Applications of Artificial Intelligence, vol. 25, no. 1, pp. 130–146, 2012.

[141] K. Lekdioui, R. Messoussi, Y. Ruichek, Y. Chaabi, and R. Touahni, "Facial decomposition for expression recognition using texture/shape descriptors and SVM classifier," Signal Processing and Image Communication, vol. 58, no. July, pp. 300–312, 2017.

[142] J. Chen, D. Chen, Y. Gong, M. Yu, K. Zhang, and L. Wang, "Facial expression recognition using geometric and appearance features," Proc. 4th Int.

Conference on. Internet Multimedia and Computer Service - ICIMCS ’12, p. 29, 2012.

[143] T. F. Cootes, G. J. Edwards, and C. J. Taylor, “Active appearance models,” IEEE Trans. Pattern Analysis and. Machine Intelligence, vol. 23, no. 6, pp. 681–685, Jun. 2001.

[144] G. Tzimiropoulos and M. Pantic, “Optimization Problems for Fast AAM Fitting in-the-Wild,” in 2013 IEEE International Conference on Computer Vision, 2013, pp. 593–600.

[145] J. Kumari, R. Rajesh, and K. M. Pooja, “Facial Expression Recognition: A Survey,” Procedia Computer Science, vol. 58, pp. 486–491, 2015.

[146] Black M., Kim, S. & Simeral, J. (2008). Neural control of computer cursor velocity by decoding motor cortical spiking activity, Journal of Neural Engineering, 5, 455–476.

[147] Chakraborty, A. & Konar, A. (2009). Emotion recognition from facial expressions and its control using fuzzy logic, IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 39(4), 726–743.

[148] Contreras R., Starostenko, O. & AlarconAquino,V. (2009). A Knowledge-based Framework for Analysis of Facial Expressions Using FACS and MPEG-4 Standards, 10th International Conferenceon Pattern Recognition and Information Processing, Minsk, Belarus, 251–256.

[149] Ekman, P. & Friesen. W.V. (1978). Facial Action Coding System (FACS). CA, USA: Consulting Psychologists Press.

[150] Esau N., Wetzel, E., Kleinjohann, L. & Kleinjohann, B. (2007). Real-Time Facial Expression Recognition Using a Fuzzy Emotion Model. IEEE International Fuzzy Systems Conference, London, England, 1–6.

[151] Gomathi, V. & Ramar, K. (2009). Human Facial Expression Recognition using MANFIS Model. World Academy of Science, Engineering and Technology, 50.

[152] Information technology – Coding of audiovisual objects. Part 2: Visual, ISO/IEC14496-2:2001(E), Second edition.

[153] Kanade, T., Cohn, J.F. & Yingli, T. (2000). Comprehensive database for facial expression analysis. 4th IEEE Conference on Automatic. Face and Gesture Recognition, Grenoble, France, 46–53.

[154] Khanum, A., Mufti, M. & Javed, M.Y. (2009). Fuzzy case-based reasoning for facial expression recognition. Journal of Fuzzy Sets and Systems, 160(2), 231–250.

[155] Kharat, G.U. & Dudul, S.V. (2008a). Human Emotion Recognition System Using Optimal SVM. WSEAS Journal Transactions on Computers, 7 (6), 650–659.

[156] Kharat G.U. & Dudul S.V. (2008b). Neural Network Classifier for Human Emotion Recognition. Conference on Emerging Trends in Engineering and Technology, Iran, 1–6.

[157] Kyoung, S.C., Yong-Guk, K. & Yang-Bok, L. (2006). Real-Time Expression Recognition System Using Active Appearance Model and EFM. Computational Intelligence and Security Conference, Guangzhou, China, 1–6.

[158] Lin, D.T. (2006). Facial Expression Classification Using PCA and Radial Basis Function Network. Journal of Information Science and Engineering, 22 (5), 1033–1046.

[159] López, J.M., Gil, R. & Cearreta, R. (2008). Towards an Ontology for Describing Emotions. Lecture Notes in Artificial Intelligence, 5288, 96– 104.

[160] Maglogiannis, I. et al. (2009). Face detection and recognition of human emotion using Markov random fields. Ubiquitous Computing Journal, 13 (1), 95– 101.

[161] Mufti M., & Khanam, A. (2006). Fuzzy Rule Based Facial Expression Recognition, International conference on Computational Intelligence for Modeling, Control and Automation, Sydney Australia, 57.

[162] Muthukaruppan, K., et al. (2007). Development of a Personified Face Emotion Recognition Technique Using Fitness Function. Japan: Springer.

[163] Pantic, M. & Rothkrantz, L.J. (2004). Facial Action Recognition for Facial Expression Analysis from Static Face Images. IEEE Transaction on Systems, Man, and Cybernetics, 34 (3), 1449–1461.

[164] Pantic M., Valstar M.F. & Rademaker R. (2005). Web-based Database for Facial Expression Analysis, IEEE Conference and Expo, Netherlands, 1–6.

[165] Plutchik R. (2001). The nature of emotions. American Scientist, 89 (4), 344–350.