

Intelligent web inference model based facial micro emotion detection for media store playback using artificial feed forward neural network and facial coordinate matrix.

Ashok Krishna EM^{1*}, Dinakaran K²

¹Department of Computer Science and Engineering, Karpagam University, Coimbatore, Tamil Nadu, India

²PMR Engineering College, Chennai, Tamil Nadu, India

Abstract

The problem of facial emotion detection has been approached by various methods in several articles but suffers from the problem of poor accuracy in detecting the facial emotion. The psychological condition of any human can be obtained by facial emotions or facial reactions, which helps to solve various problems. To solve the issue of facial emotion detection, in this paper an intelligent fuzzification model is described which uses the artificial feed forward neural network and facial coordinate matrix. More than our previous methods, the proposed method considers the facial coordination matrix which stores the coordinate point of different facial components extracted at different facial emotions considered. The method extracts the facial features and computes various measures of facial features like eye size, lip sizes, nose size, and chin size. With the above-mentioned features, the method extracts the coordinate points of skin at different emotional situations. The extracted features are used to construct the neural network and the feed forward model computes various measures by each neuron and forwards them to the next layer. At each layer, the neuron computes multi-variant emotion support for each facial emotion and forwards them to the next layer. Finally, a single emotion is selected according to the computed measure based on fuzzy approach. Identified emotion is used to play songs and videos from the media store available on the server. The proposed fuzzification based method improves the performance of facial emotion detection in micro level with more accuracy.

Keywords: Facial emotion detection, Facial coordinate matrix, Artificial neural network (ANN), Feed forward model, WIM.

Accepted on October 21, 2016

Introduction

The psychological condition of human can be identified from the facial expression and emotions. Whatever the condition of human mind can be expressed on the face and by monitoring the changes in the facial components the human emotion can be identified. The face has number of components namely eye, mouth, lips, nose and chin. The change in psychology of any human has noticeable change in the human facial components. So the psychological conditions can be identified using the facial emotions and the facial features.

The face emotion recognition is an on-going research based on biological features and signals now a day. There are various researchers working on this to support different fields of personal identification and monitoring systems. Facial emotion recognition will become vitally important in future visual communication systems, for emotion translation between cultures, which may be considered analogous to speech translation. However so far the recognition of facial emotions is mainly addressed by computer vision researchers, based on

facial display. Facial emotion recognition is very essential in various human-to-computer communications method. Even though so many examinations are made, it is difficult to locate applications in the real time system due to its undervaluing about human being's emotion recognition in any form. Various other features also under usage for the person identification like hand, eye, fingerprint and etc. To satisfy customers the automobile industries working on playing audios and videos based on reading the mind of the driver. On focusing this, we have to identify the emotions of the human and the bio-signal and facial features are the key to perform this task.

The facial coordinate matrix is one which represents the edge points of each feature from the nose, mouth, lips, chin, and skin. By identifying the edge points of the facial components, the facial coordinate matrix can be generated and used to perform facial emotion detection. Because, whenever the psychological emotion has been changed then the edge points of the facial components also gets changed. By monitoring and maintaining coordinate matrix at different situations, the

problem of emotion detection can be performed in an efficient manner.

The artificial neural network is a collection of neurons formed in different layers where each neuron is capable of performing specified task with given input and produces an output to the next level neurons. In this paper, a neuron is careful as an independent component which takes facial features and facial coordinate matrix as input and computes emotional support measure at different inputs. Based on generated values of facial support measures of same emotion, a fuzzy value are produced using which the face emotion is detected. Every single neuron computes the facial emotion support measure and based on which the neuron detects the emotion at final output. The feed forward model is one which takes the input and computes measures to produce output to the next layers.

The human can show any emotion in his/her face like happy, surprised, angry, sadness and etc. The emotions can be used for person identification and can be applied with many factors. We apply such emotions in playing media store where the song will be played according to the emotions of the user to satisfy his thoughts. The web inference model is an optimization method where the input data will be processed on the web container. The input data will be submitted to the web process through the web interfaces available, and the processed output will be given to the user through the interfaces available. In our case, the input image which is captured will be submitted to the web portal through the browser and will be processed to generate output.

Related Works

There are various methodologies been carried out to identify the facial emotions but whatever the emotions could be identified through the features of the face. In the facial image, emotions are most widely represented with eye and mouth expressions. We explore few of the methods discussed in this case.

Real-time facial feature extraction and emotion recognition [1], propose a method which uses edge counting and image-correlation optical flow techniques to calculate the local motion vectors of facial feature. Thereafter the determination of the emotional state of a subject using a neural network is discussed. The main objective of the paper is the real-time implementation of a facial emotion recognition system. The techniques used in conventional off-line calculations are modified to make real-time implementation possible. A variety of improved techniques to the general algorithms such as edge focusing, global motion cancellation are proposed and implemented.

Human-computer interaction using emotion recognition from facial expression [2], describes emotion recognition system based on facial expression. A fully automatic facial expression recognition system is based on three steps: face detection, facial characteristic extraction, and facial expression classification. We have developed an anthropometric model to detect facial feature points combined to Shi et al.'s method.

The variations of 21 distances which describe the facial features deformations from the neutral face were used to coding the facial expression. Classification step is based on SVM method (Support Vector Machine). Experimental consequences demonstrate that the planned approach is an effective method to recognize emotions through facial expression with an emotion recognition rate more than 90% in real time. This approach is used to control music player based on the variation of the emotional state of the computer user.

Facial emotion recognition using multi-modal information [3], propose a method of facial emotion detection by using a hybrid approach, which uses multi-modal information for facial emotion recognition.

Analysis of emotion recognition using facial expressions, speech and multimodal information [4], examines the strengths and the limitations of organizations based only on facial expressions or acoustic gen. It also discusses two methods used to fuse these two modalities: decision level and mouth level integration. Using a database chronicled from an actress, four emotions were classified: sadness, anger, happiness, and neutral state. By the use of markers on her face, thorough facial motions were captured with gesture capture, in combination with simultaneous speech soundtracks. The results disclose that the system based on facemask expression gave a better presentation than the system based happening just acoustic information for the emotions considered. Results also show the complementarily of the two modalities and that when these two modalities are fused, the presentation and the robustness of the feeling recognition system improve noticeably.

Meta-analysis of the first facemask expression recognition task [5], presents a meta-analysis of the first such challenge in automatic recognition of facial expressions, held during the IEEE conference on face and gesture recognition 2011. It details the challenge data, evaluation protocol, and the results attained in two sub-challenges: AU detection and classification of facial expression imagery in terms of a number of discrete emotion categories. We also summarize the lessons learned and reflect on the future of the field of facial expression recognition in general and on possible future challenges in particular.

In sign-judgment approaches [6], a widely used method for manual labelling of facial actions is Facial Action Coding System (FACS). FACS associates facial expression changes with actions of the muscles that produce them. It defines 9 different AUs in the upper face, 18 in the lower face, and 5 AUs that cannot be classified as belonging to either the upper or the lower face. In addition, it defines the-called action descriptors, 11 for head position, 9 for eye position, and 14 additional descriptors for miscellaneous actions. AUs are considered to be the smallest visually discernible facial movements. AU intensity scoring is defined on a five level ordinal scale by FACS. It also defines the makeup of AUs' temporal segments (onset, apex, and offset) but goes short of defining rules on how to code them in a face video or what rules governing the transitions between the temporal segments. Using FACS, humanoid coders can physically code nearly an anatomically possible facial expression, decomposing it into

the specific AUs and their temporal segments that produced the expression.

A real-time facial expression recognition system based on active appearance models using gray images and edge images present an approach for facial expression classification, based on active appearance models [7,8]. To be able to work in real-world, we applied the AAM framework on edge images, instead of gray images. This yields to more robustness against varying lighting conditions. Additionally, three different facial expression classifiers (AAM classifier set, MLP and SVM) are compared with each other. An essential advantage of the developed system is that it is able to work in real-time-a prerequisite for the envisaged implementation on an interactive social robot. The real-time capability was achieved by a two-stage hierarchical AAM tracker and a very efficient implementation.

Facial emotion gratitude using context based multimodal approach [9], present a method for the emotion acknowledgement from a facial expression, hand, and physique posture. Our model uses multimodal emotion recognition system in which we use two different models for facial expression credit and for hand gesture recognition and then uniting the result of both classifiers using a third classifier which gives the resulting sentiment. Multimodal organization bounces more accurate consequence than a signal or bimodal system.

The methods discussed here have used different features for emotion detection problem; we propose a rule-based approach which works based on real time web based inference model.

Proposed Method

The proposed fuzzification based micro-emotion detection using ANN feed forward model with the facial coordinate matrix, first pre-process the captured facial image to remove noise introduced by the electronic device. At the second stage, the method extracts various features of nose, eyes, lips, mouth and generates the facial coordinate matrix. The generated features and matrix are used to generate the artificial neural network with feed forward model. There are a number of neurons present in each layer where each layer is considered as micro-emotion, and each neuron performs the micro-emotion support measure with the input features and the pre-computed values. Based on the values computed a single micro-emotion is identified as a result. The entire process can be split into a number of stages where each has their own dedicated service

Table 1. Reactions and features for different emotions.

Emotion reaction	Feature representatives
Normal	None
Anger	Lowered brows/drawn together/line between brows/lower lid tense/ may be raised/upper lid tense/lowered due to brows' action/lips are pressed together with corners straight down or open
Surprise	Brows raised/skin below brow stretched/not wrinkled/horizontal wrinkles across forehead/eyelids opened/jaw drops open or stretching of the mouth

namely pre-processing, feature extraction, image capturing, facial coordinate matrix generation, ANN generation, micro emotion detection. This section will explain each functional stage in detail.

Image capturing

We have designed an interface to access the camera attached to the computer through which the face image of the user could be accessed. We capture the facial image of the user through the web cam attached, and captured image is transferred to the web portal through an http request.

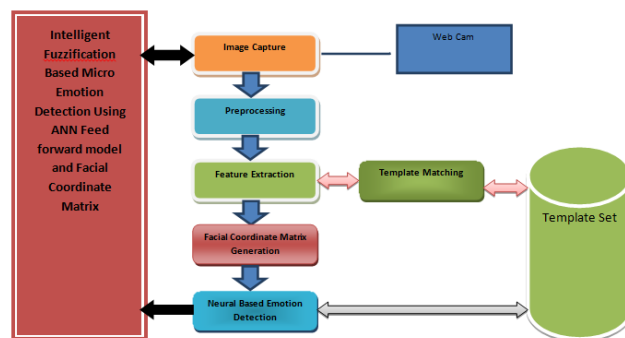


Figure 1. Proposed method architecture.

The Figure 1 shows the architecture of proposed ANN Feed forward model for micro-emotion detection approach and its functional components.

Pre-processing

The pre-processing is to prepare the image for feature extraction. The input image is applied with histogram equalization technique to improve the image quality. For each pixel from the image, red, green and blue values are normalized to generate the equalized image. Skin matching technique is applied to separate the skin feature from other regions. The most commonly used technique to determine the regions of interest ROI) is skin color detection. A previously created probabilistic model of skin-color is used to calculate the probability of each pixel to represent some skin. Thresholding then leads to the coarse regions of interest. Some further analysis could, for example, involve the size or perimeter of the located regions in order to exclude regions such as the face. The pre-processing makes the image as completely ready for feature extraction.

Fear	Brows raised and drawn together forehead/wrinkles drew to the center/upper eyelid is raised, and lower eyelid is drawn up/mouth is open/lips are slightly tense or stretched and drawn back
Happiness	Corners of lips are drawn back and up mouth parted/not with teeth exposed/not, cheeks are raised, lower eyelid shows wrinkles below it, wrinkles around the outer corners of the eyes
Disgust	The upper lip is raised, the lower lip is raised and pushed up to the upper lip or it is lowered, nose is wrinkled, cheeks are raised brows are lowered
Sadness	Inner corners of eyebrows are drawn up the upper lid, inner corner is raised, corners of the lips are drawn downwards

Feature extraction

From the pre-processed image, we extract the facial features like eye, lips, and nose. We apply template matching technique to extract the necessary features from the pre-processed image. The proposed method maintains set of templates for each case of an emotion considered in this paper according to Table 1. With the template matching technique, region growing methods are used to extract the features.

Facial coordinate matrix generation

The facial components are subject to change their shape, and their coordinates or boundary points are also subject to change according to the micro-emotions. When the user smiles the skin and chin shapes, and their coordinates will change and increase and this will identify the micro-emotion in an efficient manner. At this stage, the method computes the coordinate points of facial features extracted at the feature extraction phase and for each facial feature being identified, the method computes the coordinate points and selects the most deviating points to maintain in the matrix.

Algorithm

Input: Feature set Fs

Output: Coordinate matrix Cm.

Start

Extract eye feature E.

Compute coordinate points Cp of left eye,

CPL= $\int \sum$ Edge points ϵ LE

For each point Pi from CPL

Identify distanced suitable curvature points.

$$DSCP = \int_{i=1}^{size(CPL)} \int_{CPL(i)}^{size(CPL)} Dist(CPL(i), CPL(j)) \in Maxx(CPL) //$$

This function identifies the set of points which covers remaining points in the edges.

End

Add to matrix CM.

CM (1, :)=DSCP

Compute coordinate points Cp of right eye,

CPL= $\int \sum$ Edge points ϵ RE

For each point Pi from CPL

Identify distanced suitable curvature points.

$$DSCP = \int_{i=1}^{size(CPL)} \int_{CPL(i)}^{size(CPL)} Dist(CPL(i), CPL(j)) \in Maxx(CPL) //$$

This function identifies the set of points which covers remaining points in the edges.

End

Add to matrix CM.

CM (2, :)=DSCP

Compute coordinate points Cp of nose

CPL= $\int \sum$ Edge points ϵ Nose

For each point Pi from CPL

Identify distanced suitable curvature points.

$$DSCP = \int_{i=1}^{size(CPL)} \int_{CPL(i)}^{size(CPL)} Dist(CPL(i), CPL(j)) \in Maxx(CPL) //$$

This function identifies the set of points which covers remaining points in the edges.

End

Add to matrix CM.

CM (3, :)=DSCP

Compute coordinate points Cp of mouth

CPL= $\int \sum$ Edge points ϵ Mouth

For each point Pi from CPL

Identify distanced suitable curvature points.

$$DSCP = \int_{i=1}^{size(CPL)} \int_{CPL(i)}^{size(CPL)} Dist(CPL(i), CPL(j)) \in Maxx(CPL) //$$

This function identifies the set of points which covers remaining points in the edges.

End

Add to matrix CM.

CM (4, :)=DSCP

Stop.

The above-discussed algorithm computes the set of all edge points of each facial feature or facial component which will be used to compute the emotional support.

Multi-variant emotional support estimation

For each extracted facial feature, boundary points are identified and extracted. Using the extracted coordinates, the method computes various measures of the facial feature are computed. Then for each facial feature different metrics are computed, i.e. in the case of the eye, the distance from pupil to upper eye brow, the distance between eye brows, the presence of the line in between eye brows are computed. With the mouth feature, the distance between lips, the tightness of lips, height and width of lips, stretchiness is computed. Then for each feature set trained and with the coordinate matrix, the method computes the micro facial emotion support measure using the entire feature set trained under each emotion class.

Algorithm

Input: Feature set Fs, Coordinate Matrix CM.

Output: Micro Emotional Support Measure MESM.

Step 1: Start

Step 2: Read Feature set Fs.

Step 3: Initialize feature values set FV.

Step 4: For each feature Fi from Fs

Eye feature E (Fi)=Fi (Eye).

Identify pupil center pc= $\Omega \times (E)$.

Identify upper eye brow boundary Eub=gray pixel at the top of E.

Compute pupil to upper eye brow PB=Eub-pc.

Compute inner lower brow position lb.

Compute inner upper brow position ub.

Compute distance between brows Db=ub-lb.

Identify end of left eye brow lb.

Identify starting of right brow rb.

Compute distance between brows dib = lb-rb.

For each pixel between lb and rb

Check for the presence of gray pixel.

If found then

Line present

Else

Line not present.

End

End.

//Estimating Lips features

Compute width and height of lips.

L_{height}=Location of a first gray pixel in column wise location of last gray pixel in column wise.

L_{width}=Location of first gray pixel in row wise location of last gray pixel in row wise.

Compute the distance between lips L_{Dis}.

L_{Dis}=Distance between the boundary of both lips.

Store feature points in FV (i)={PB, Db, Dib, line present, L_{height}, L_{width}, L_{Dis}}.

End

For each feature vector Fi from FVSet

$$\text{Compute pupil similarity } ps = \frac{\sum_{i=1}^{size(FVset)} Dist(pb, Fvset(i), pb)}{size(FVset)}$$

$$\text{Compute Eye Brow similarity } Ebs = \frac{\sum_{i=1}^{size(FVset)} Dist(Db, Fvset(i), Db)}{size(FVset)}$$

$$\text{Compute Lips similarity } Ls = \frac{\sum_{i=1}^{size(FVset)} Dist(Dlib, Fvset(i), Dlib)}{size(FVset)}$$

$$\text{Compute Shrink Line Similarity } SLS = \frac{\sum_{i=1}^{size(FVset)} Dist(Lp, Fvset(i), Lp)}{size(FVset)}$$

$$\text{Compute Coordinate Feature Similarity } CFS = \frac{\sum_{i=1}^{size(Fvset)} \sum_{j=1}^4 \frac{Dist(Fvset(1,j), CM(j))}{size(Fvset)}}{size(Fvset)}$$

Compute Micro Emotional Support Measure MESM.

MESM=ps+Ebs+LS+SLS/CFS

End

Step 5: stop.

The above-discussed algorithm computes the micro facial emotion support measure by computing various measures using the facial features and the coordinate matrix. The computed emotional support measure will be used to identify the emotion of the user.

ANN based emotion detection

In this stage, the method constructs the neural network with a number of layers where each layer has a number of neurons. The neurons are assigned with the facial features extracted and the volumetric measures computed. For each emotion considered there will be a layer of neurons and has trained features. At each layer, the neuron computes similarity measure with the rule set available with the feature given as input to the neuron. The output will be given to the next layer to proceed with the emotion detection. Finally, the neurons at the last layer estimate the similarity measures according to the measures completed in the previous layers and select a single emotion to produce the result. Based on selected emotion, a song from the media store will be play backed. The audio songs are stored according to the emotion which will be used for playing according to the emotion identified.

Algorithm

Step 1: Start

Read number of emotions considered.

Noe= \sum Emotions ϵ Emotion set

Create Neural network Nn.

$$Nn = \int_{i=1}^{\text{size(Emotion)}} \int_{j=1}^{\text{size(Feature set)}} \text{Create-Neuron(Feature)}$$

Step 3: Initialize Emotion set E.

Step 4: Read feature set Fv.

Step 5: For each layer Li

Compute Multi variant Emotional support MESM.

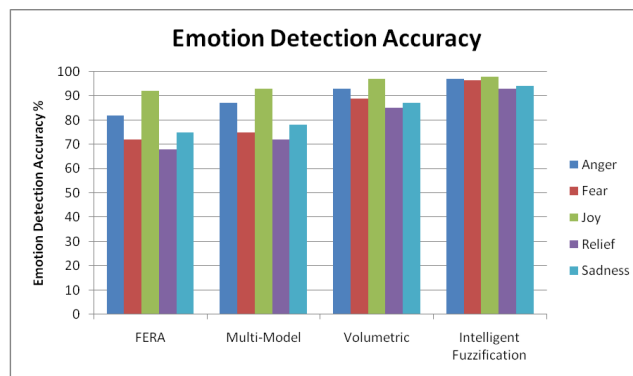
Forward Similarity measure with features to next layer.

End

Step 6: Choose the emotion with more similarity measure.

Step 7: Playback the song according to selected emotion.

Step 8: Stop.



Graph 1. The accuracy of emotion detection of different methods.

The Graph 1 shows the comparison of emotion detection accuracy produced by different methods and the results shows that the proposed method has produced a more efficient result than other methods [10-17].

Conclusion

We proposed an intelligent fuzzification approach for web-based facial emotion detection in a micro level using artificial feed forward neural networks which detect the emotion with the support of facial coordinate matrix. We have used skin matching technique to remove unwanted skin features from the image then we have used template matching technique which has few set of templates for each of the feature. The template matching is used to identify the location of the facial feature, and the feature is extracted to compute various measures like the size of features. The method computes the coordinate matrix of extracted features and generates artificial feed forward neural network with a number of neurons at each layer of emotions. Each neuron computes similarity on each feature and finally computes the facial micro-emotion support measure at each level. Computed measures will be passed through the next layers, and finally, a single emotion is selected based on fuzzy values where the micro-emotion can be identified at the micro level deviation of the similarity measure. Based on identified final emotion a related audio will be played for the user. The proposed method has reduced the time complexity and increased the accuracy of emotion detection than other previous methods [10-17].

References

1. Real-time facial feature extraction and emotion recognition. IEEE Conf Inform Commun Sig Proc 2003; 3: 1310-1314.
2. Human-computer interaction using emotion recognition from facial expression. UK Sim European Symp Comp Model Simul 2011; 196-201.
3. De S. Facial emotion recognition using multi-modal information. International Conf Inform Commun Sig Proc 1997; 1: 397-401.
4. Carlos B, Zhigang D, Serdar Y, Murtaza B, Chul ML, Abe K, Sungbok L, Ulrich N, Shrikanth N. Analysis of emotion recognition using facial expressions, speech and

The above-presented algorithm computes the emotional support measure using the computed values the method selects an emotion accordingly.

Results and Discussion

The proposed Fuzzification based web inference model for micro-emotion detection using coordinate matrix and artificial neural network based feed-forward model has produced good results. Feature extraction accuracy as high as 95.3% for best combination facial feature extraction was achieved. Emotion recognition accuracy as high as 87% was achieved based on the AM-FED and FERET face database. Speed improvement achieved over the previous method is in excess of 20 times. The proposed method has replaced the offline emotion detection methods by providing an online model. FERA is the baseline method we have used to evaluate our proposed method.

Table 2. Size of data set used for evaluation.

Emotion	Train	Test	Total
Anger	42	28	70
Fear	36	16	52
Joy	48	16	74
Relief	38	12	50
Sadness	29	17	46

Table 2 shows size of data set used for evaluation.

We have used different size of data for training and testing phase, and the proposed method has evaluated at all the cases for the accuracy of detection.

- multimodal information. Proc Int Conf Multimod Interface 2004; 13-15.
5. Valstar MF, Mehu M, Bihan Jiang, Pantic M, Scherer K. Meta-analysis of the first facial expression recognition challenge. IEEE Trans Syst Man Cybern B Cybern 2012; 42: 966-979.
 6. Cohn J, Ekman P. Measuring facial action by manual coding, facial emg, and automatic facial image analysis. Non-verbal Behav Res Meth Affect Sci Oxford Univ Press 2005; 9-64.
 7. Muid M, Assia K. Fuzzy rule based facial expression recognition. Int Conf Comp Intel Model Contr Autom Int Conv Intel Agents IEEE 2006.
 8. Martin C. A real-time facial expression recognition system based on active appearance models using gray images and edge images. IEEE Conf Autom Face Gest Recogn 2008; 1-6.
 9. Priya M. Facial emotion recognition using context based multimodal approach. Int J Emerg Sci 2012; 2: 171-182.
 10. Hatice G, Massimo P. A bimodal face and body gesture database for automatic analysis of human nonverbal affective behaviour computer vision research group. IEEE Int Conf Patt Recogn 2006.
 11. Loic K, Ginevra C, George C. Multimodal emotion recognition in speech-based interaction using facial expression. Body Gest Acoust Anal 2009.
 12. Ming-Hsuan Y. Detecting faces in images. Surveyieee Trans Patt Anal Mach Intel 2002; 24.
 13. Littlewort G, Whitehill J, Wu T, Butko N, Ruvolo P, Movellan J, Bartlett M. The motion in emotion-a cert based approach to the fera emotion challenge. Proc IEEE Int Conf Autom Face Gesture Anal 2011; 897-902.
 14. Littlewort G, Whitehill J, Wu T, Fasel I, Frank M, Movellan J, Bartlett M. The computer expression recognition toolbox (cert). Proc IEEE Int Conf Autom Face Gesture Recog 2011; 298-305.
 15. Liu C, Yuen J, Torralba A. Siftow: Dense correspondence across scenes and its applications. IEEE Trans Pattern Anal Mach Intell 2010; 33: 978-994.
 16. Baltrusaitis T, McDuff D, Banda N, Mahmoud M, El Kaliouby R, Robinson P, Picard R. Real-time inference of mental states from facial expressions and upper body gestures. Proc IEEE Int Conf Autom Face Gesture Anal 2011; 909-914.

***Correspondence to**

Ashok Krishna EM

Department of Computer Science and Engineering

Karpagam University

Tamil Nadu

India