Emotion Recognition System Using Open Web Platform

Tejeswi Vallabhaneni,
CSE Department, R.M.K Engineering College
Thiresh hasan,
CSE Department, R.M.K Engineering College
S.d. Lalitha,
AP, CSE Department, R.M.K Engineering College

----------ABSTRACT----------

This paper proposes a model for recognizing emotions through movement of facial muscles inspired by FACS (Facial Action Coding System) and FACSAID (Facial Action Coding System Affect Interpretation Dictionary). The computational implementation of the proposed model, here called WeBSER (Web-Based System for Emotion Recognition), was produced in Open Web Platform and is able to infer the user's emotional state in real time. The images of the user's face are captured using a webcam and emotions are classified using a Computer Vision system that uses the Web as a platform. Given the sequences of images acquired in real time via webcam, the WeBSER performs the following steps: Face detection and segmentation (eyes with eyebrows, nose and mouth); Entering reading points; Classification of emotions based on the movement of the reading points. For face detection and segmentation of face regions such as eyes, nose and mouth, the Viola-Jones method was used. Given the face image and the location of the segmented regions, 20 reading points were identified in image. The movement of each reading point is analyzed relatively to the other points. The direction of the movement of reading points is classified in bands of 45 degrees; Thus, each point can assume one of eight directions or remain stationary. Finally, the classification of emotions is made based on the movement of the reading points. This proposed model has a mean accuracy of 76,6% for determining exact emotions, and 84.4% to indicate uncomfortable states of persons suggesting suspicious behaviors.

Keywords - **Human-Machine Interaction, Computer Vision, Affective Computing, Image Processing, Emotion Recognition, Open Web Platform, Canvas, HTML 5.**

I. INTRODUCTION

As non-verbal communication has been crucial in human evolution, the use of nonverbal information is essential for the evolution of computer systems. In this context, new forms of data entries representing nonverbal information expressed by users can provide relevant contributions to man / machine interactions. In the interaction between humans, face provides a wide range of information both about the individual such as sex and age, as about his emotional state, for example if the person is happy or sad, interested or bored. Studies have shown that basic emotions such as happiness, disgust, surprise, anger, sadness and fear are innate and universal, making possible their use in computer systems worldwide, regardless of culture. The existing systems used in "smart security", try to support the task made for the security men helping them to identify dangerous people or situations that cannot be detected with naked eye. Face recognition technique is becoming one of the most important and interesting aspects in terms of the intelligent security by identification of images in several areas like banking, police records, biometric, etc. One of the main problem on today existing surveillance systems is the human factor. People who work observing an important number of screen monitors during a long period of time, certainly miss details that could be a potential threat. Due this, an intelligent surveillance system should help the operator with an automatic prediction warning of suspects behavior in order to prevent criminal events rather than identifying people after happened attacks [1]. Other challenging result is got on interviews and interrogations, where is necessary to use nonintrusive technologies to quickly measure the credibility of the statements made for some person. [2]. Those intelligent surveillance systems attempt to perceive tension, lying and psychological changes indicated by involuntary expressions appearing in human face when somebody is trying to hide emotions [3], [4]. Studies shows that methodologies usingemotion recognition could be used in intelligent Security Systems because they can do an objective analyzes to avoid the mistakes that traditional surveillance systems have, due the people efforts to hide facial expressions or to camouflage emotions. [5], [6], [7]. Expressed emotions can be recognized in the voice, gestures, and mainly through facial expressions [8]. The study to classify emotions through facial expressions can provide information about the affective state of the individual, including both momentary emotions like fear, anger, joy, surprise, sadness, disgust, and most enduring emotions such as humor, euphoria and irritability. In addition to information on the cognitive activity (bewilderment, boredom, and concentration), temperament and personality (sociability, hostility or timidity), truthfulness (lanes when the information provided in the words on the plans or actions is false) and psychopathology (depression, mania, schizophrenia and other disorders) [9].

II. FACIAL SYSTEMS SIGNS

According [10] face conveys information through four classes of signals: The first ones represent static facial

features of the face relatively permanent, such as bone and soft tissue masses, which contribute to form the appearance of an individual. The soft or mild facial signals represent changes in the appearance of the face that occur gradually over time, such as the appearance of wrinkles and permanent changes in the skin texture. The signs represent fake facial features of the face artificially determined like eyeglasses and cosmetics. Finally, the rapid facial signals represent immediate or rapid changes in physical activity, for example the neuromuscular that produce detectable changes in facial appearance. Just the rapid facial signals are relevant for studies on emotions and cognitive state of an individual, while the other three classes represent background noise in such studies. A facial signal system should measure the rapid facial signals based mainly on muscle actions; and the best way to build a measurement system is to understand the computational techniques for measuring facial behaviors, including those that humans use in their day to day [10].

III. IMAGE PROCESSING TECHNIQUES FOR RECOGNITION OF EMOTIONS

To detect human emotions by analyzing the facial expression in digital images it is necessary to apply image processing techniques to perform the location of the face (FaD – Face Detection) and relevant characteristics (FeD – Feature Detection) inherent to the application domain. In this study, Viola-Jones algorithm [11] was used to detect the face, the Sobel filter [12] was applied to the face image to highlight the edges and enable insertion of reading points of the proposed model which is based on FACS [11]. The LucasKanade algorithm [13], [14] was implemented to analyze the movement of read points that were classified using decision structures.

3.1 Face Detection

In the detection method proposed for [11] a classifier is introduced for object detection, in which the main goal is the detection of faces. The Viola-Jones detection technique is based on the appearance of the objects and it tries to "learn" characteristics from a set of training images [15]. This process occurs on static images and can be applied to real-time applications. The Viola-Jones method is based on features, being conducted in three modules: the creation of the full image [16], Adaboost algorithm [17] for classification using haar-like features (Fig. 1) [18] and creation of a tree structure called cascade of classifiers (Fig. 2). Fig. 1.Haar-Like characteristics.Adapted from [18] Fig. 2.Detector structure in cascade. Adapted from [25]

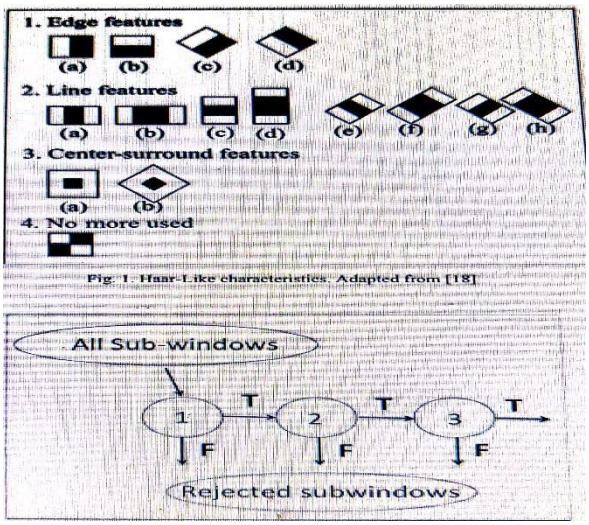

Fig 1 :Facial Action Coding System

3.2 Facial Coding System

The FACS (Facial Action Coding System) is a coding system for facial expressions developed by Ekman and Friesen in 1970 to determine how facial muscle contraction - alone and in combination with other muscles - changes the appearance of the face. Through the analysis of several shots of facial expressions were identified specific changes that occurred with muscular contractions and defined a way to differentiate them. The aim of the studies was Frisen Ekman and create a way to categorize the behavior indices through human facial. The FACS Manual was first published in 1978 [10], [19], [20].

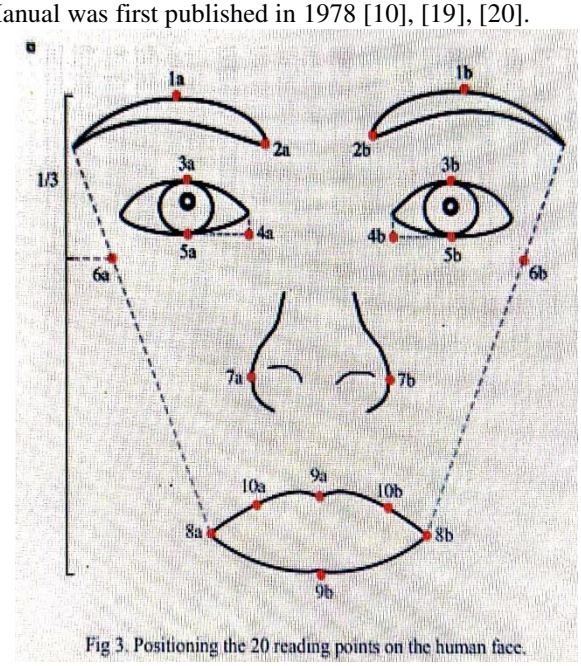

Fig 3. Positioning the 20 reading points on the human face.

The FACS consists of 46 AUs (Action Units) which are divided into two facial regions: upper, containing eyes, eyebrows and forehead, and bottom, which are considered the cheeks, chin, nose and mouth. Beyond 46 AUs, FACS has other codes that assist and complement these actions. These codes are organized in relation to head movement, eye movement, visibility of facial features and behaviors.

3.3 Feature Tracking Approach

The Lukas-Kanade method allows the analysis of moving regions in a video by testing optical flow. The pyramidal version of that method used in this work, create a Gaussian pyramid image having at the top the slightest part of the image (with less detail) and the base with the crucial image pixels (greater detail). Optical flow is calculated for points at the top of the pyramid and the results of the calculation of this level serve as a starting point for the calculation of the next level. This allows capture the largest movements using local windows which make method more efficient and robust because violates the initial assumption that the tracked pixel and its neighbors belong to the same closed surface [21].

IV. WEB-BASED SYSTEM FOR EMOTION RECOGNITION

The proposed system for recognizing emotions is based on web standards, allowing its use in any type of device having a rendering engine for Web - used by Web browsers. Moreover, the browser itself can be used. The current implementation performs the classification of the next emotional states: happiness, disgust, surprise, anger, sadness and fear. For that task are used 20 reading points which can move in 8 different direction divided in bands of 45 degrees. The emotion classification is done by a decision structure.

4.1 Implementation

The WeBSER was implemented in Open Web Platform. To capture images via webcam, an API (Application Programming Interface) was used: GetUserMedia [22]. For manipulation of captured images, 2D and 3D contexts of HTML 5 canvas element [23] were used, as well as the ECMAScriptlanguage[24].

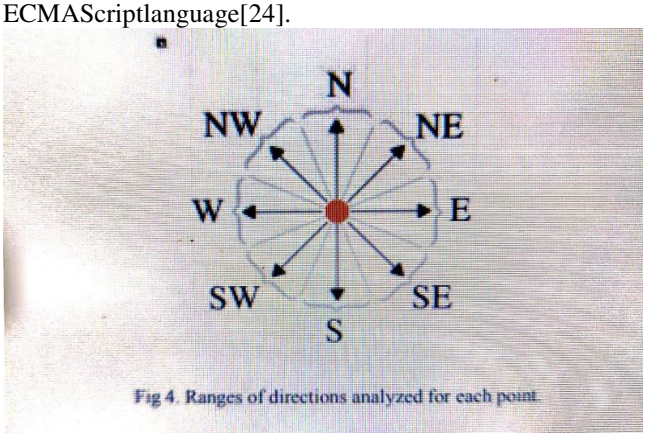

Fig 4. Ranges of directions analyzed for each point.

From this interface, the openCV library [25] ported for ECMAScript was used.

4.2 Reading Points

Given the face image and the location of the segmented regions of eyes, nose and mouth, 20 reading points were inserted as shown in Fig. 3. Fig 3.Positioning the 20 reading points on the human face. The Sobel filter was applied to enhance edges. Then, a threshold was applied to binarize the image. The image was scanned using masks with eyes of medium size (including eyebrows), nose and mouth in various scales for localization of regions where the points should be inserted. The movement of each reading point is analyzed relatively to the other points. The direction of the movement of reading points is classified in bands of 45 degrees. Thus, each point can assume one of eight directions or remain stationary. Fig. 4 illustrates the ranges of directions analyzed for each point.

4.3 Reading Points versus Emotions

The Table I presents the structure of decision given the twenty points of reading (1a to 10b), considering the eight possible directions of movement of each point to the emotions happiness (H), disgust (Di), surprise (Su), anger (A), sadness (Sa) and fear (F).

V. EXPERIMENTS AND RESULTS

The WeBSER system was tested by sexteen volunteers (nine female and seven male) aged between ten and sixty years, of three different nationalities (Brazilian, Thai and Chinese). Two experiments were conducted. In the first experiment, the volunteers expressed in a nonspontaneous way emotions such as happiness (H), disgust (Di), surprise (Su), anger (A), sadness (Sa) and fear (F) for two consecutive times. Table II shows the confusion matrix [26] of tests, where rows indicate the real emotion and columns represent the emotion interpreted by the system. The mean accuracy was 76,6%.

TABLE I. READ POINTS VERSUS EMOTIONS

Pts	Directions of Movement Read Points versus Emotions							
	N	NE	E	SE	S	SW	W	NW
1a	F/Su	F/Su						F/Su
1b	F/Su	F/Su						F/Su
2a	F/Su/Sa	F		A				
2b	F/Su/Sa					A	Su	F/Sa
3a	A/F/Su	A/F/Su		A/H/Sa	A/H/Sa	A/H/Sa		A/F/Su
3b	A/F/Su	A/F/Su		A/H/Sa	A/H/Sa	A/H/Sa		A/F/Su
4a	A	A	A					
4b								A
5a	A/H/Sa/D							
5b	A/H/Sa/D							
6a	H	A						
6b	H							A
7a	D						A	
7b	D		A					
8a	Sa		A/Su			Sa	F/Su	H
8b	Sa	H	F	Sa			A	
9a	H/D/Su	H/D/Su		A/F	A/F	A/F		H/D/Su
9b	A/F	A/F		H/Su	H/Su	H/Su		A/F
10a	D/Su				F	F	F	D
10b	D/Su	D			F	F	F	

VI. CONCLUSION

Either a person or computer cannot identify the exact feeling or emotion in other person, because people sometimes cannot even recognize their own emotions [27]. However, this study demonstrated that it is possible to recognize the user's basic emotions through a Computer Vision system. The tests showed that the tools available in Open Web Platform enable the use of the regular web systems in developing this type of programs. The results demonstrate that the proposed model was able to satisfactorily sort basic emotions simulated by volunteers. In addition, the second experiment showed that although the system does not accurately identify the emotion expressed by a person, it is possible to determine that the individual is in a state of discomfort, indicating a suspicious situation. The WeBSER enables applications in several areas, among them security, because it has a mean accuracy of 76,6% for determining exact emotions, and 84.4% to indicate uncomfortable states of persons suggesting suspicious behaviors.

REFERENCES

[1] R. Wang, B. Fang, "Affective Computing and Biometrics Based HCI Surveillance System". *Information Science and Engineering, 2008.ISISE '08.International Symposium, vol*. 1, pp. 192-195, 2008.

[2] Ryan, J. F. Cohn, S. Lucey, J. Saragih, P. Lucey, F. De la Torre, A. Rossi, "Automated Facial Expression Recognition System". Security Technology, 2009. 43rd Annual *2009 International Carnahan Conference, pp.* 172-177, 2009.

[3] M. Owayjan, A. Kashour, N. Al Haddad, M. Fadel, G. Al Souki, "*The design and development of a Lie Detection System using facial microexpressions". Advances in Computational Tools for Engineering Applications (ACTEA), 2012 2nd International Conference,* pp. 33-38, 2012.

[4] M. Rojas, C. M. Travieso, J.B. Alonso, M. A. Ferrer, "*Automatic lip identification applied under soft facial emotion conditions". Security Technology (ICCST), 2012 IEEE International Carnahan Conference*, pp. 218- 223, 2012.

[5] X. Chao ; F. Zhiyong, "Facial Expression Recognition and Synthesis on Affective Emotions Composition". Future *BioMedical Information Engineering, 2008.FBIE '08.International Seminar*, pp. 144-147, 2008.

[6] R. Srivastava, S. Roy, "3D facial expression recognition using residues".TENCON 2009 - 2009 *IEEE Region 10 Conference, pp. 1-5, 2009. [7] D. B. Megherbi, Y. Miao, "A distributed efficient face image screening and retrieval with affine transformations, disguised, and varying facial expressions".* Technologies for Homeland Security, 2009.HST '09. IEEE Conference, pp. 425-432, 2009.

[7] P. EKMAN, "The Argument and Evidence About Universals in Facial Expression of Emotion". *Handbook of Social Psychophysiology, Chichester: John Wiley, Ltd.,* pp. 143-164, 1989. [9] P. EKMAN et al., "NSF Report - Facial Expression Understanding", 1992.

[8] P. EKMAN, "*Facial signs: Facts, fantasies, and possibilities". In: SIGHT, Sound and Sense. Bloomington: T. Sebeok,* pp. 125-156, 1978.

[9] P. Viola, M. Jones, "Rapid Object Detection Using aBoosted Cascade of Simple Features". *Conference on Computer Vision and Pattern Recognition,* 2001.

[10] R. C. Gonzalez, R. E, Woods, "Processamento de ImagensDigitais", EdgardBlücher, São Paulo, 2000.

[11] B.D. Lucas, T. Kanade. "*An iterative image registration technique with an application to stereo vision".Journal of Japanese Society for Artificial Intelligence*, pp. 121–130, 1981.

[12] J. Y. Bouguet. "*Pyramidal Implementation of the Lucas Kanade Feature Tracker Description of the algorithm"*.Intel Corporation Microprocessor Research Labs, 2000.

[13] P. Viola, M. Jones, D. Snow, "*Detecting Pedestrians Using Patterns of Motion and Appearance". Proceedings of the Ninth IEEE International Conference on Computer Vision*, 2003.

[14] R. Lienhart, J .Maydt*, "An Extended Set of Haar-like Features for Rapid Object Detection"*. ICIP, pages 900–903, 2002.

[15] Y. Freund, R. E. Schapire, *"A decision-theoretic generalization of online learning and an application to boosting". 2th European Conference on Computational Learning Theory,* 1995.

[16] R. Lienhart, J. Maydt, *"An Extended Set of Haar-like Features for Rapid Object Detection", in Conference on Imaging Processing,* pp. 900–903, 2002.

[17] J. C. HAGER, P. EKMAN, "*The Inner and Outer Meanings of Facial Expressions". In: SOCIAL Psychophysiology. New York: Gilford Press,* pp. 287-306, 1983.

[18] J.C. HAGER, "DataFace: Facial Expressions, Emotion Expressions, Nonverbal Communication". Disponívelem: Acessoem: 03 out., 2011.