Facial Emotion Recognition Using Fuzzy Systems

Austin Nicolai
Electrical & Computer Engineering
Mercer University
Macon, GA, USA
austin.l.nicolai@live.mercer.edu

Anthony Choi
Electrical & Computer Engineering
Mercer University
Macon, GA, USA
choi_ta@mercer.edu

***Abstract*—This paper presents a fuzzy logic based emotion recognition system. The system is comprised of an image processing stage followed by an emotion recognition stage. In the image processing stage, the subject's face and facial features (eyes, mouth, etc.) are extracted. Next, the relevant identifying points are extracted from each facial feature. In the the emotion recognition stage, the identifying points are used to fuzzify and determine the strength of different facial actions. These strengths are then used to determine the subject's displayed emotion. The Japanese Female Facial Expression database was used to evaluate the system's performance resulting in an overall successful detection rate of 78.8%.**

Keywords - Fuzzy system; Emotion recognition; JAFFE; Feature extraction

I. INTRODUCTION

Emotion recognition is a problem that is relevant in a wide variety of fields. The recognition of emotion from facial expressions has been a main focus of many psychological studies over the past decade [1]. It is no wonder that many studies have arisen from this topic. Applications of this topic include automated security systems, interactive robotic aids assistants, and diagnosing medical patients [2]. In a security system, accurate emotion recognition could help identify potential threats. Similarly, it would allow robotic aids to more effectively interact with and help a human counterpart. In each case, accurate and timely recognition plays an integral role in improving the performance of the system.

Today, there has been a growing interest in improving interaction between humans and computers. Some say that to get the best response, computers need to interact with the user just like humans interact with each other. Emotional responses can be seen through gestures and speech. Even though speech is the main method, gestures can portray a lot of vital information about a current state or emotion [3, 4].

Psychological theory states that human emotions can be classified into six archetypal types: surprise, fear, disgust, anger, happiness, and sadness. To display these archetypal types, the muscles in a person's face can change, the tone of voice can be altered, and the energy of a person's voice can increase or decrease. The shape and formation of the lips can also greatly contribute to understanding of speech in a noisy environment. All of these play a role in the process of communicating different feelings. Even if these signals are subtly displayed, humans can recognize these signals by processing information from the ears and eyes [5, 6]. It has been said that individuals focus more attention on projecting their own facial expressions and perceiving others' facial expressions than they do other nonverbal channels and often more than verbal communication as well [7].

When approaching emotion recognition, a variety of techniques can be employed. Approaches include analyzing body language, voice patterns, and facial features. In analyzing the face, Ekman and Friesen, in 1978, presented a now well-known facial expression model called the Facial Action Coding System (FACS). FACS breaks the face into 46 different Action Units (AU's). Each AU is defined as the contraction or relaxation of one or a group of muscles. Some, but not all, of these AU's contribute to the identification of six basic emotions identified by Ekman and Friesen: happiness, sadness, fear, anger, disgust, and surprise. The AU's pertaining to the eyes, eyebrows, and mouth are most associated with the expression of emotion [8-11].

The fuzzy system presented in this paper is based on using these AU's to recognize emotion. Using the Japanese Female Facial Expression (JAFFE) database [12], the system compares an image of interest to a neutral expression baseline image to determine the expressed emotion. This database (JAFFE) contains 213 images containing 7 different facial expressions posed by 10 different females. To achieve this comparison between a neutral and expressed emotion, the need for any time consuming training period is negated. First, the two images are processed to extract relevant facial features. The differences in position of the different feature points are then fuzzified to obtain the strength of exhibited AU's. These strengths are then fed through fuzzy rule sets and defuzzified to obtain the exhibited strength of the six basic emotions. The recognized emotion is chosen by a *winner-take-all* technique in which the highest exhibited strength is chosen.

II. RELATED WORK

In recent years, much research has been done regarding the problem of emotion recognition. Many different approaches utilize the JAFFE database to measure performance. Other systems use fuzzy logic based techniques.

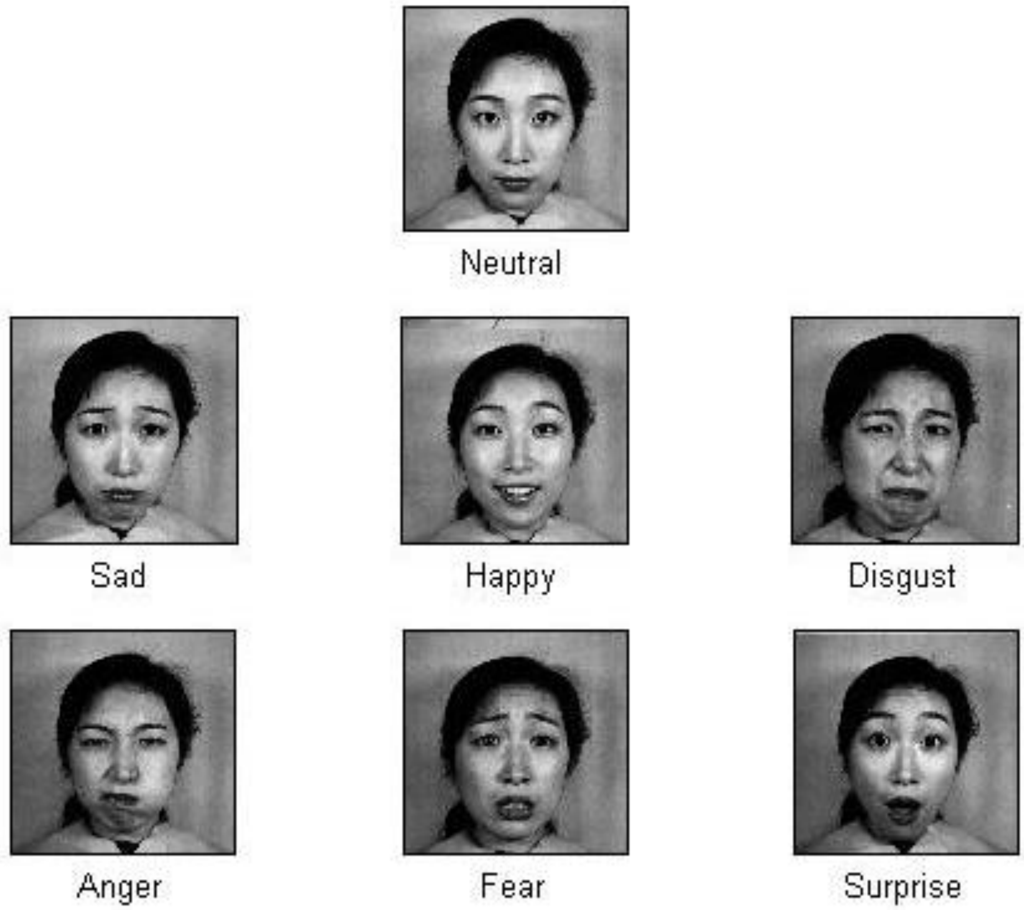

Figure 1. Seven emotion images for one subject

A study done on Alzheimer's disease shows that in the early stages, the most important structures dealing with emotional processing are damaged by the disease. These three areas that are damaged are the hippocampus, the amygdala, and the posterior association areas [13]. In similar cases, people with autism, a disorder characterized by profound difficulties with social interaction, also have trouble recognizing facial emotion recognition. This is due to abnormalities in brain structure or function. The main causes of autism are still under debate as to whether it is a genetic disorder or people are born susceptible to it [14]. The amygdala is often damaged in people with temporal lobe epilepsy as well. This area of the brain has been associated with emotional recognition in faces.

Since the beginning of a person's life, facial expressions are a key expression that humans analyze. These facial "features" can tell a person a lot about their environment or the emotional state of an individual while speaking [15]. From these studies, it can be seen that these three areas of the brain greatly influence a person's ability to recognize emotion.

Since this paper uses the JAFFE database to measure performance, it is useful to look at the results of other systems using the same database. One system using Gabor filters at different scales and orientations followed by Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) was able to successfully recognize the six basic emotions, as well as neutral, with a 92% success rate [16]. Another system using Gabor wavelet coefficients and geometric positioning was able to distinguish the six basic emotions, as well as neutral, with 90.1% success [17]. Using PCA for feature extraction and LDA as a classifier, another system was able to recognize emotion with an 87.6% success rate [18].

In addition to looking at systems using the JAFFE database, it is also important to look at the results of other systems utilizing fuzzy logic. A fuzzy rule based system with a Genetic Algorithm (GA) to improve performance was able recognize four emotions (happiness, sadness, surprise, and anger) with an accuracy rate of 88% using the Cohn-Kanade database [19]. A fuzzy rule based system operating on the angles between different facial feature points achieved an overall success rate of 72% in detecting four emotions (happiness, sadness, fear, and anger) in real-time [20]. It is worth noting that the detection rates for happiness and sadness (63% and 58%) were significantly lower than that of anger and fear (72% and 90%) for this system.

Figure 2. Multiple happiness expressions for one subject

Some applied recognition techniques have been used in many aspects ranging from the Xbox Kinect to treatment of Asperger's to even affective computing. This is done through continuous video monitoring. To analyze this input, features are extracted from the current time and classification is carried out using this single time point. To determine if the facial feature was of a certain type, differencing of Gaussians and histogram differencing were used. Even though this seems promising, there was about 60% accuracy on correct identification [21]. Another study has developed a scoring system to systemically categorize the physical expression of emotions. Ratliff and Patterson used FACS to analyze facial expressions. Results of this study showed that results varied by person anywhere from 60 to 100% accuracy. The average accuracy was between the 80% to 90% range [22]. The results of this study bring up another interesting characteristic, why is it more difficult for certain people to be recognized compared to others? The fact is that each person has a different facial structure. Therefore, a large training set of data would be a good idea to help accurately identify these facial expressions. Other ideas that have been used include hidden Markov models. These are great for modeling stationary signals. These models are; however, quite time consuming [23].

An important factor to consider when looking at these results is that most of these systems take advantage of complex and robust feature detection schemes while using more straightforward emotion categorization methods. The exception is the real-time system [20], which utilizes a time constrained feature detection scheme.

III. Database Overview

The image database used for measuring the performance of the system was the JAFFE database [12]. The database is comprised of ten Japanese females identified only by their initials (KA, KL, KM, KR, MK, NA, NM, TM, UY, and YM). For each subject, seven different expressions are presented. These expressions include neutral, happy, sad, angry, disgust, fear, and surprise. For each subject, two to four example images are present for each expression.

In total, there are 213 images in the database. The images are of size 256 x 256 and in TIFF format. Fig. 1 shows the seven expressions presented for one of the subjects. Fig. 2 shows multiple examples of the happiness expression for one subject.

IV. Facial Feature Extraction and Fuzzy Classification

A. Facial Feature Extraction

In order to detect the six basic emotions identified by Ekman and Friesen, the relevant facial feature points must first be extracted from the images. Since this system focuses on the fuzzy classification of emotions, only simple feature extraction is performed. A basic thresholding method is used. This is possible because there is sufficient change in contrast between the subjects' skin tone and that of the eyes, eyebrows, and mouth. This is accomplished using the MATLAB computer Vision and Image Processing toolboxes [24, 25].

The first step for the feature extraction is to find the facial features of interest (eyes, eyebrows, and mouth). First, the subject's face is identified. After cropping the original image, the subject's eyes, eyebrows, and mouth are located. These are identified by MATLAB with a bounding box in which the feature is contained, as shown in Fig. 3.

Next, the actual facial feature must be found within the region identified by MATLAB. To do this, the image is first thresholded. The actual feature is then identified by tracing the boundaries of objects found in the thresholded regions. Fig. 4 shows the traced boundaries for the identified eyes, eyebrows, and mouth for one subject.

Last, the identifying feature points are obtained from the feature boundaries. For the eyes and mouth, the top, bottom, left, and right corners are used. For the eyebrows, the left, right, and center points are used. Fig. 5 shows the identified

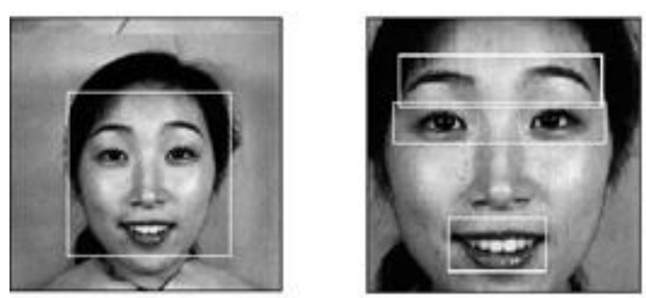

Figure 3. Facial features of interest

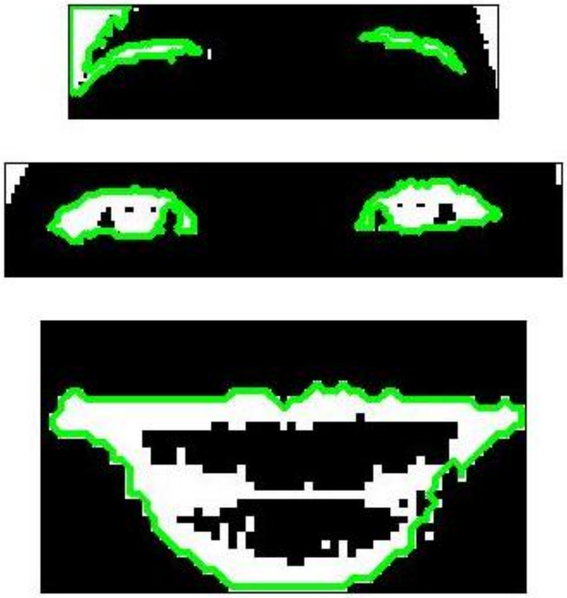

Figure 4. Traced boundaries

feature points overlaid on the original image of the subject.

B. Fuzzy Classification

With the relevant feature points identified, the data must now be fuzzified for classification [26]. Using the identified feature points, relevant information for each AU is calculated. This information includes parameters such as eyelid height, inner brow height, mouth width, etc. The AU information is calculated for both a neutral expression reference image as well as the image of interest. The values are then compared for each AU and recorded. The comparison is done as a ratio of the image of interest to the reference image. For example, if the inner brow height increases from 8 pixels in the neutral image to 12 pixels in the image of interest, the value recorded is 12:8, or 1.5.

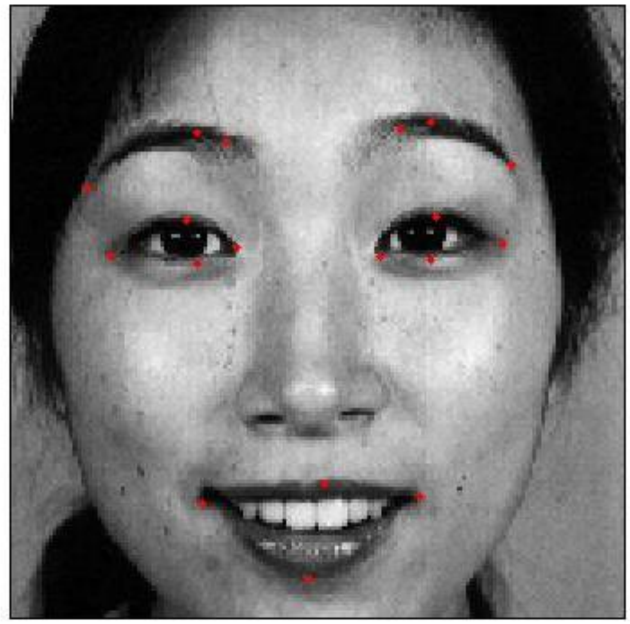

Figure 5. Overlaid facial feature points

When the ratios for all AU's have been recorded, the data is fuzzified. For each AU, the strength exhibited is denoted by the linguistic variables *none*, *weak*, *normal*, and *strong*. The membership functions used to map inputs to outputs are comprised of triangular and trapezoidal curves. For AU's that depend on multiple facial parameters, each parameter is represented by its own membership function. Fig. 6 shows an example membership function for AU 1 (inner brow raised).

For each emotion, there is a separate fuzzy rule system. Using a separate rule system for each emotion has a two-fold benefit. Since each emotion does not rely on every AU,

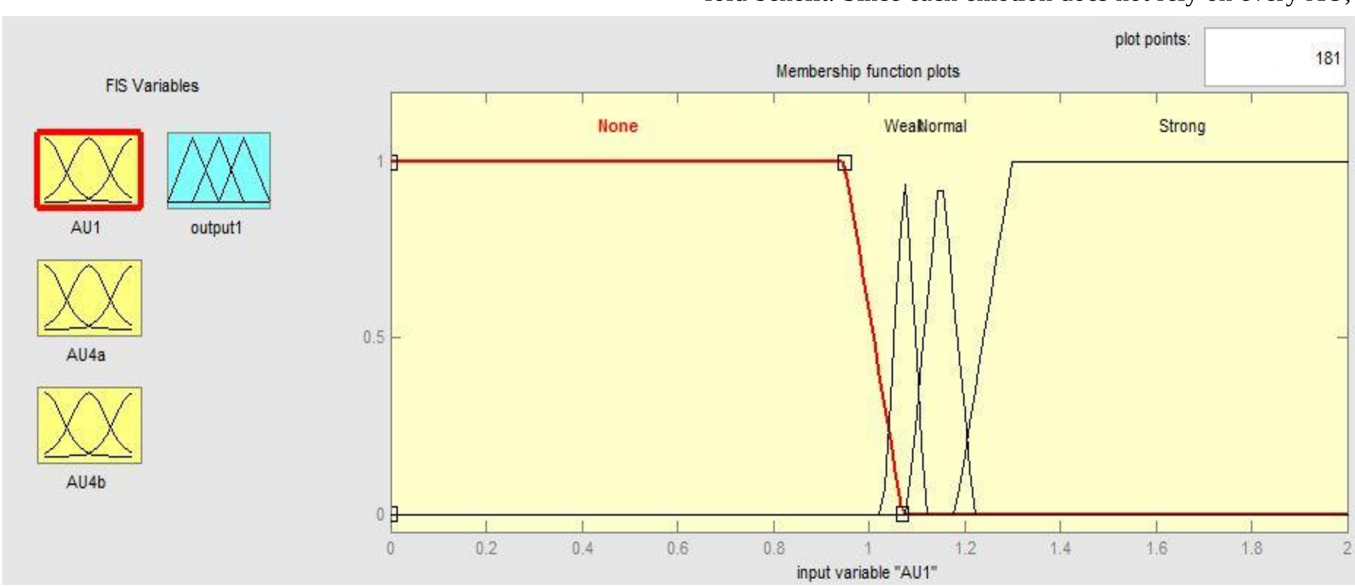

Figure 6. AU 1 membership function

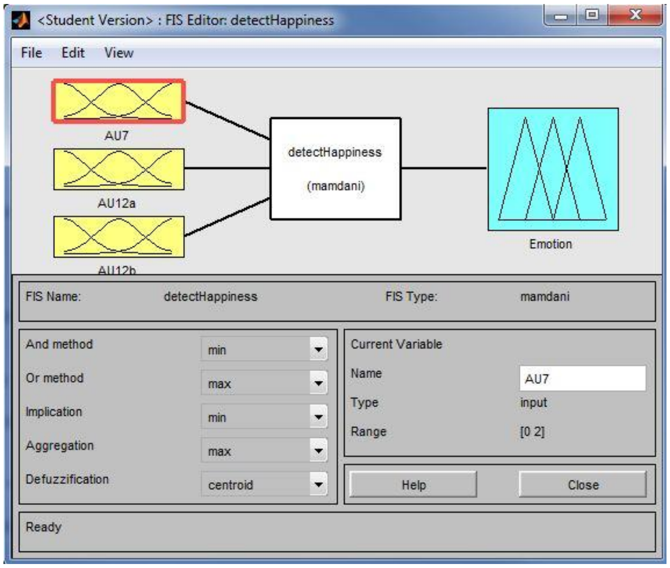

Figure 7. Happiness fuzzy system

using multiple rule systems drastically reduces the total number of rules required for the system. In reducing the number of rules required, the system processing time is reduced as well.

The output from each rule system is a value between 0 (not displayed) and 6 (very strongly displayed). This value represents the degree to which the image of interest displays the given emotion. Fig. 7 shows an example rule system. The fuzzified AU data is fed through each of the fuzzy rule systems and the outputs are compared. A *winner-take-all* method is used to determine the emotion exhibited by the image of interest.

TABLE I. EMOTION RECOGNITION SUCCESS RATES

	Success Rate					
	Happiness	***Sadness***	***Anger***	***Fear***	***Surprise***	***Disgust***
Proposed	80%	78%	90%	75%	87%	63%
Dubuisson [18]	88%	89%	91%	83%	90%	85%
Esau [20]	63%	58%	72%	90%	N/A	N/A

V. RESULTS

All of the non-neutral expression images in the JAFFE database were used for testing. A single neutral image was used as a reference for each subject. The results are presented in Table I. The average recognition rate was 78.8%. Happiness, anger, and surprise were detected with the highest success rate. Sadness and fear were detected reasonably well while disgust was detected with by far the lowest accuracy. Disregarding disgust, the average recognition rate rises to 82%.

Disgust was the most difficult emotion to detect. This is likely because detecting disgust largely relies on the nose region and surrounding area. The system presented here does not take the nose into consideration.

In some cases, visual analysis of expressions can be rather difficult due to how a person’s expression is portrayed. Visual inspection of the images in the JAFFE database reveals that several subjects perform some emotions incorrectly according to Ekman and Friesen's description; in fact, additional AU’s can be seen that are not typically shown for that emotion. Therefore, expanding the data set or excluding the incorrect portrayals will help achieve better results.

In other cases, the image processing performed poorly in identifying the desired feature points. This was most notable in the left/right corner points for the mouth as well as the upper/lower eyelid points. This is expected due to the simplicity of the feature extraction method used. Even with this simple method, the fuzzy system still performs well. Using more accurate feature extraction will likely increase system performance.

Another issue stems from the resolution of the test images. With a resolution of 256 x 256, certain parameters (eyelid height, inner brow height, etc.) are measured as a single digit pixel value. The smallest possible change (one pixel) represents a 10% or more increase/decrease. This level of granularity limits the amount of fuzzification the membership functions can perform.

Comparing the results to similar work with more complex feature detection, the system presented performs with comparable success rates for some emotions while performing worse for the rest [18]. The system presented performs with slightly increased success rates compared to that of the less complex real-time system [20].

The difference in level of complexity of the image processing between these other systems and this system likely explains the differences in success rates. Increasing the accuracy of input data to the fuzzy system should improve the successful recognition rate to the level achieved in these similar systems.

VI. ACKNOWLEDGEMENTS

Authors thank Roland Adams for his help with editing.

REFERENCES

[1] R. Adolphs, “Recognizing emotion from facial expressions: psychological and neurological mechanisms,” Behavorial and Cognitive Neuroscience Reviews, vol. 1, pp. 21-62, March 2002.

[2] C. Gagliardi, E. Frigerio, D. Buro, I. Cazzaniga, D. Pret and R. Borgatti, “Facial expression recognition in Williams syndrome,” Neuropsicologia 41, 2003, pp. 733-738.

[3] A. Azcarate, F. Hageloh, K. Sande, and R. Valenti, “Automatic facial emotion recognition,” Universiteit van Amsterdam, pp. 1-6, June 2005.

[4] A. Metallinou, C. Busso, S. Lee, and S. Narayanan, “Visual emotion recognition using compact facial representations and viseme information,” ICASSP 2010, pp. 2474-2477, 2010.

[5] C. Busso, Z. Deng, S. Yildirim, M. Bulut, C. Lee, A. Kazemzadeh, S. Lee, U. Neumann, and S. Narayanan, “Analysis of emotion recognition using facial expressions, speech, and multimodal information,” Proceedings of the 6th international conference on Multimodal interfaces, pp. 205-211, 2004.

[6] S. Ioannou, A. Raouzaiou, V. Tzouvaras, T. Mailis, K. Karpouzis, and S. Kollias, “Emotion recognition through facial expression

analysis based on a neurofuzzy network," Neural Networks, pp.423-435, 2005.

[7] H. Elfenbein, A. Marsh, and N. Ambady, "Emotional intelligence and the recognition of emotion from facial expressions," The wisdom of Feelings: Processes Underlying Emotional Intelligence, pp. 1-19, 1998.

[8] P. Ekman, "Emotions in the human face," New York, NY, Cambridge University Press, 1982.

[9] P. Ekman and R. J. Davidson, "Nature of emotion," New York, NY, Oxford University Press, 1994.

[10] P. Ekman and W. V. Friesen, "The facial action coding system," Palo Alto, CA, Consulting Psychologists Press, 1978.

[11] P. Ekman and K. R. Scherer, "Approaches to emotion," Hillsdale, NJ, Erlbaum Associates, 1984.

[12] M. J. Lyons, M. Kamachi, and J. Gyoba, "Japanese Female Facial Expression (JAFFE) Database of Digital Images," http://www.kasrl.org/jaffe.html, 1997.

[13] J. Drapeau, N. Gosselin, L. Gagon, I. Peretz, and D. Lorrain, "Emotional Recognition from face, voice, and music in dimentia of the alzheimer type," The Neurosciences and Music III, vol. 1169, pp. 342-345, 2009.

[14] M. Harms, A. Martin, and G. Wallace, "Facial emotion recognition in autism spectrum disorders: a review of behavorial and neuroimaging studies, vol. 20, pp.290-322, 2010.

[15] C. Cristinzio, D. Sander, and P. Vuilleumier, "Recognition of emotional face expressions and amygdala pathology," Epileptologie 2007, vol. 24, pp. 130-138, 2007.

[16] M. Lyons, J. Budynek and S. Akamatsu, "Automatic classification of single facial images," IEEE Trans. Patt. Anal. Mach. Intell. 21, 1999, pp. 27-38.

[17] Z. Zhang, M. Lyons, M. Schuster and S. Akamatsu, "Comparison between geometry-based and Gabor-wavelets-based facial expression recognition using multi-layer perceptron," Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition, 1998, pp. 454-459.

[18] S. Dubuisson, F. Davoine and M. Masson, "A solution for facial expression representation and recognition," Sign. Process.: Imag. Commun. 17, 2002, pp. 657-673.

[19] A. Jamshidnezhad, "A learning fuzzy model for emotion recognition," European Journal of Scientific Research 57(2), 2011, pp. 206-211

[20] N. Esau, E. Wetzel, L. Kleinjohann and B. Kleinjohann, "Real-time facial expression recognition using a fuzzy emotion model," IEEE Fuzzy Systems Conf., 2007, pp. 1-6

[21] A. Cruz, B. Bhanu, and N. Thakoor, "Facial emotion reconition in continuous video," 21st International Conference on Pattern Recognition, pp. 1880-1883, November 2012.

[22] M. Ratliff and E. Patterson, "Emotion recognition using facial expressions with active appearance models," HCI Proceedings of the Third IASTED International Conference on Human Computer Interaction, pp. 138-143, 2008.

[23] I. Cohen, A. Garg, and T. Huang, "Emotion recognition from facial expressions using multilevel HMM," pp. 1-7, 2000.

[24] "Computer Vision System Toolbox," *MathWorks*, 2012, Web.

[25] "Image Processing Toolbox," *MathWorks*, 2012, Web.

[26] "Fuzzy Logic Toolbox," *MathWorks*, 2012, Web.