Classification of Facial Expression Using Transformed Features

M. Arfan Jaffar and Eisa Al Eisa

***Abstract*—Automatic identification of facial expressions structures the elementary nature of a variety of next generation computing devices together with sentimental computing expertise, intellectual tutoring methods, and patient sketch delicate wellness scrutinize methods etc. Therefore, we have proposed a facial expression recognition system that has the aptitude of incrementally learning and thus can learn all possible patterns of expressions that may be generated in feature. Proposed system consists of different phases including face detection, features extraction and classification. First of all, face detection has been performed by using Voila & Jones method which is robust and then transformed features has been extracted for classification using local window. Three types of features have been extracted using Discrete Cosine Transform, Haar Wavelet transform, and Gabor Wavelet. Then these features have been fused and used for classification. The results of proposed technique are compared using different quantitative measures with some of the existing techniques which show its performance.**

***Index Terms*—Facial expression, classification.**

I. Introduction

The area generation of computing; such as Computational Intelligence and omnipresent ambient temperature should be human-centered user interfaces that respond easily to multimodal natural human communication to develop. These interfaces need to be able to spot and recognize the intentions and emotions, which are expressed by the social and emotional indicators. Applications of facial expressions recognition have a wide range of human interaction, computer, robotics and security of the fusion of the face images for the conversion between the sexes and fusion [1], etc. Therefore, automatic face recognition different age groups of expressions is the essence of the various IT tools of the next generation, including affective computing technologies, intelligent tutoring systems and systems monitoring profiles patients personal welfare, etc.

Speaking from man to man, articulation and perception of facial expressions form a communication channel in addition to the voice, the essential information about the mental, emotional and even physical persons in conversation state bears. In its simplest form, facial expressions mean when a person happy or angry. In a more subtle vision expressions Feedback can be either planned or unplanned supply listener to the speaker to indicate understanding, sympathy, or even disbelief to what the speaker is saying. Recent research has shown that certain facial expressions can be also questioned whether a person is trying to deceive their interviewer [1], [2].

Generally established predictions in the computing will move to the background if they did not perform well. This goal is to achieve the computation of the next generation; as ubiquitous computing and ambient intelligence, are human -centered user interfaces that develop multimodal easily react to natural human communication [1]. These interfaces need to be able to spot and recognize the intentions and emotions, which are expressed by the social and emotional indicators. This vision of the future motivates the research for the automatic recognition of non-verbal language and action. Facial expression recognition has attracted the attention more and more in computer vision, pattern recognition and human-computer interaction research community. Automatic recognition of facial expressions is thus the essence of the various IT tools of the next generation, including affective computing technologies, intelligent tutoring methods, surveillance systems the benefit of patients personal profiles, etc.

Indulgent how people practice and distinguish facial expressions has been an exigent assignment in the field of computer vision for an elongated time. Various methods have been proposed to conjure up the human progression, in which, a variety of adaptive methods are established, such as, artificial neural networks, genetic programming and algorithms, and support vector machines [2]. However, eventual solution for this is still being pursued. One of the difficulties in the expressions recognition tasks is to enhance the robustness over the spatial and temporal variations of human facial expressions due to the growth (or aging) and the changes in lighting conditions, face directions, expressions, make-up, and so forth. Conventional expression recognition systems can achieve excellent performance when tested over a benchmark dataset [3]. However, the performance significantly may decreased if such systems FER operated in a practical environment. This is because all the pictures of the face is training inadequate or unsuitable for expression patterns that could be generated accordingly. Any changes that will occur in the future will not be taken into consideration in advance; Performance and high recognition can hardly be expected in concrete situations with only one set of static data. Another problem by the regulations for the classification of facial expressions during the preamble of the new mode of expression, or the amount on a variety of facial expressions in the confirmation innovative found occurred. Existing systems can not have the ability to such a large amount of data to record. This is a problem for practical speech recognition system. Consider the system of a

Manuscript received February 1, 2014; revised April 17, 2014.

The authors are with the Al-Imam Muhammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia (e-mail: aleisa@ccis.imamu.edu.sa, arfan.jaffar@ccis.imamu.edu.sa).

DOI: 10.7763/IJIEE.2014.V4.446

particular environment in which they take interest only a few of the conditions in the case of the bill said [1]. Later, when we use this system in a different position, which is used in a number of terms, while usually does not offer a solution the system want [4].

Fig. 1. Expression taken from dataset.

The paper is organized according to following outlined: related work will be discussed in Section II and proposed method in Section III; results has been discussed and presented in Section IV.

II. Related Work

Study of the expression dates back to 1640, when John Bulwer done studies on facial expressions by the biological point of view. In the 19th century, the expressions of people and animals examined from a psychological point of view [1]. Later, Darwin wrote a book called " The Expression of the Emotions in Man and Animals." Before the mid- 1970s, the analysis of facial expression has attracted the interest of many groups of computer vision. Several statistical techniques applied to select functions. Two major techniques of parametric representation of facial expressions, has been proposed. In 1978, defines a new scheme for describing facial movement [1], [6]. This is called the Facial Action Coding Scheme (FACS). FACS combines 64 units share (AU) and a combination of the AU represents the movement of facial muscles and said information about facial expressions. In 1998 [7], the facial animation parameters (FAP), which are part of the MPEG - 4 synthetic / natural hybrid coding (SNHC) have been proposed [1].

A number that have been developed for the method of analysis of the expression [8], [9], and analysis of surface structures, measuring facial movements of optical flow. Methods based on the geometry typically retrieve geometric characteristics such as the shape and location of the components of the object (such as the mouth, eyes, eyebrows and nose in the face objects) and represented by a vector characteristics to characterize the geometry of the object [10]-[12]. In general, different facial expressions are different representations of entities. Since it is very difficult to locate and extract the geometrical characteristics of the geometry -based methods in many practical applications, methods based on appearance - increasingly popular for facial expression recognition performance and better performance are based only the geometry in terms of detection accuracy and. Methods based on the overall appearance convert each face image into a feature vector and apply analysis subspace techniques to extract statistical characteristics of the representation of facial expression [13]-[16].

Image filters, such as principal component analysis (PCA) [17], linear discriminant analysis (LDA) [18] regularized discriminant analysis (RDA) [15] and Gabor wavelet analysis [14], can be applied to the entire surface or specific regions of the face, facial features to extract the changes. It should be noted that bend calculation expensive facial images with a set of Gabor filters to extract multi- scale and multi - orientation coefficients. Furthermore, in practice, the size of Gabor functions is so high that the conditions of memory and calculation are high. In recent years, an effective face descriptor called local binary pattern (LBP) [14], which was originally proposed for texture analysis [12], have attracted great interest in the representation of facial expression.

In 2011, [13], a new approach to learning subspace transfer is proposed, a feature space that results from the training data to test data, the detection performance against in - learn to improve the data transfer scenarios. According to this concept, they made four new methods for transferring learning space, ie, transferred PCA (TPCA), transfer (LDA TLDA) transfer LPP (TLPP) and transferred ONPP (TONPP) for recognition international data facial expressions. A comparison of our proposed with the techniques proposed in this paper system was presented in Section IV.

III. Proposed Method

A. Face Detection

In this case the problem to be addressed is the recognition of the facial expression. To detect the expression from a static image of a person in our area of interest is the face of this person. For example, we find that when a person smiles, changes in the area of only the surface observed. Smile has nothing to do with hair, ears and neck of a person in the picture. So we have to extract the known area of the surface of the entire image, the region of interest. Therefore, we used for ROI extraction Viola and Jones face [6] detection technique based on AdaBoost algorithm. The technology of face recognition in AdaBoost consists of three elements: the integral image a strong classifier from weak classifiers based on the AdaBoost learning algorithm and an architecture which consists of a cascade of classifiers number of solids. A cascade of classifiers enhanced layer 25 is formed to realize multi-view face. A whole area of the sample and not the face images (indicated as background) are used for training. AdaBoost face detection algorithm to detect faces in a fast and robust way [1].

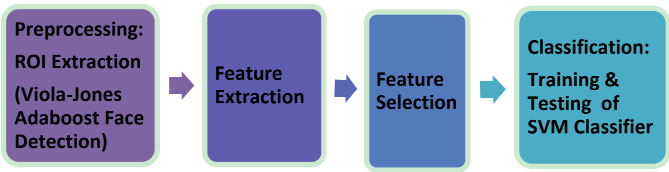

Fig. 2. Proposed architecture.

B. Hybrid Fused Features Extraction

We have extracted DCT based features, Haar wavelet transform based features and then Gabor wavelet features. After then we have fused these features and used for classification.

First, we have extracted features by using DCT. 2-Dimensional DCT of the input is defined by the following equations:

$$C(u,v) = \alpha(u)\alpha(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)cos\left[\frac{\pi(2x+1)u}{2N}\right]cos\left[\frac{\pi(2y+1)v}{2N}\right]$$

where $0 \leq u \leq N$, & $0 \leq v \leq N$, and

$$a(u) = \begin{cases} \sqrt{\frac{1}{N}} & for\ u = 0 \\ \sqrt{\frac{2}{N}} & for\ u \neq 0 \end{cases}$$

We input image into four non-overlapping windows divided horizontally and vertically. Then, each window was divided into four sub- windows know. In this way, there are 16 different windows do not overlap. Then DCT is applied to each window and separately extract the top sixteen functions zigzag scanning. In this way, we have a total of 256 functions in a single image. After the APC asked these functions to reduce dimensionality and selected the top ten functions. Thereafter, they can be reduced and are selected for classification used DCT characteristics.

Then we extract features using Haar wavelet transform [2]. The features are extracted from a Haar wavelet transform. Haar wavelet is in the recognition algorithm used texture. It can be used in order to analyze the texture, and to detect the edges of the face. For this purpose, the surface portion of the first image is recorded [16], [17]. Then, perform convolution on the detected face image with the filter of hair. Then convert the facial image in a folded bit pattern by swelling of the image. After extracted and save the model as binary traits of the process described above.

Gabor filtering is a process of take out texture information transform base. The utilization of Gabor filters by Gabor filtering is stimulated " strongly correlated with the human visual system. " [18]. Gabor filter is a Gaussian modulated Sinusal sin form and take out an image in the outward appearance of a reaction image by the appliance of diverse parameters. [18].

$$G(x,y) = e^{\left(\frac{x_\theta^2\ - \gamma^2 y_\theta^2}{\alpha^2} + \frac{2\pi x\theta^i}{\lambda}\right)} \tag{1}$$

where

$$x_\theta = x\cos\theta\ + y\sin\theta \tag{2}$$

$$y_\theta = -x\sin\theta + y\sin\theta \tag{3}$$

And σ is the standard deviation of the Gaussian function, λ is the wavelength of the harmonic function, θ is the orientation, and γ is the spatial aspect ratio that is constant at 0.5. The spatial frequency bandwidth is the ratio σ/λ and is detained constant and the same to.56. Hence there are 2 parameters that modify when figure out a Gabor filter θ and λ. Input image is segregated into 9×9 non-overlapping sections. The Gabor filter is then convolved with diverse parameters; and it will produce the response images. All parameters are defined in [18]. Michael [18] utilized simply the odd constituent of the Gabor filter that does not create imaginary output:

$$G_o(x,y) = \exp\left(\frac{x_\theta^2 - \gamma^2 y_\theta^2}{\alpha^2}\right)\sin\left(\frac{2\pi x\theta}{\lambda}\right) \tag{4}$$

After that, it will convolve the image with twelve Gabor filters corresponding to 4 directions (θ) and 3 frequencies ($1/\lambda$).

C. Classification

SVM support vector machine [4] was used for classification. After extracting features the second important task is a good classifier to have hasty and vigorous to a scrupulous problem. This segment illustrates the structural drawing, instruction and testing the classifier stock. There are numerous classifiers which can be employ for multi-classification dilemma, but there is a require for such a amalgamation of a classifier or classifiers that can efficiently consign expression with towering accuracy [6], [15]. We found that the combination of binary classifiers so that each classifier is trained to recognize a particular expression to achieved capable results and recovers the simplification presentation on individual classifiers significantly. So we premeditated a stock of SVM for the identification of facial expression. We get the results of the classifiers designed differently and make decisions, the result are superior. This is awfully comparable to our judgments in everyday life we search for numerous views ahead of building a judgment [10]. The reaction of every classifier stock classifier merged with the highest power. Thus, the grouping of twofold classifiers is employed for multi- classification of facial expressions [1].

IV. Results and Discussion

We have three diverse datasets to estimate the performance of proposed method for facial expression classification technique. The three datasets used in this case are JAFFE [12] MMI [13] and CAFE [14]. The details of these records will be omitted and provided in [19]. In the relevant documents No facial image of the sample from these data sets are shown in Fig. 1 and Fig. 4, the number of images that are each of the seven groups of expression about the same. We tested through a series of three experiments: 20 percent for training and 80 for testing, 50 percent for training and 50 percent for testing and 80 percent for training and 20 percent for testing. We have the power of the classifier in the calculation and analysis of sensitivity and specificity. We have tested the performance of classifiers by calculating and analyzing, sensitivity and specificity [19]. These are defined as follows: Accuracy: *Number of classified mass / Number of total mass* (TP + TN)/(TP + TN + FP + FN) ,where TP is known as True Positive, TN is known as True Negative, FP known as False Positive and FN is considered as False Negative [19].

TABLE I: FACIAL EXPRESSION ACCURACIES IN PERCENTAGE

DS / exp	MMI	MUG	JAFFE
ANG	61.70	68.26	83.69
DIS	64.70	56.49	82.17
FEA	71.20	66.42	85.39
HAP	72.50	69.72	87.70
SAD	69.36	66.37	86.30
SUR	75.27	72.45	87.60
NEU	68.47	66.50	84.40
AVG	**69.02**	**66.60**	**85.32**

A. Experiment I: 20% Training—80% Testing

During this experiment, we have partition the dataset into training and testing data ratios like twenty percent of the images of both classes to structure the training set, and the lingering eighty percent of the images employ to structure the test set. Table I and Fig. 4 demonstrate results by using this experiment via SVM.

B. Experiment II: 50 % for Training—50 % for Testing

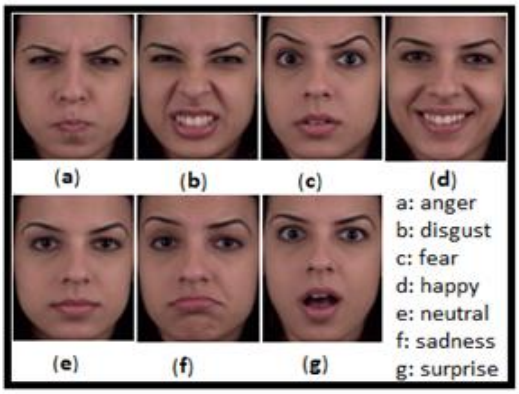

Fig. 3. Seven basic emotions (pictures courtesy MUG database).

TABLE II: FACIAL EXPRESSION ACCURACIES IN PERCENTAGE

DS / exp	MMI	MUG	JAFFE
ANG	74.30	74.71	93.12
DIS	76.50	69.53	92.15
FEA	72.30	73.42	85.33
HAP	74.58	74.53	87.15
SAD	72.26	78.37	96.20
SUR	76.59	72.56	87.50
NEU	73.26	71.40	84.30
AVG	**74.25**	**73.50**	**89.30**

During this experiment, we have partition the dataset into training and testing data ratios like fifty percent of the images of both classes to structure the training set, and the lingering fifty percent of the images employ to structure the test set. Table II demonstrates results by using this experiment via SVM.

C. Experiment III: 80 % for Training—20 % for Testing

During this experiment, we have partition the dataset into training and testing data ratios like eighty percent of the images of both classes to structure the training set, and the lingering twenty percent of the images employ to structure the test set. Table III demonstrates results by using this experiment via SVM.

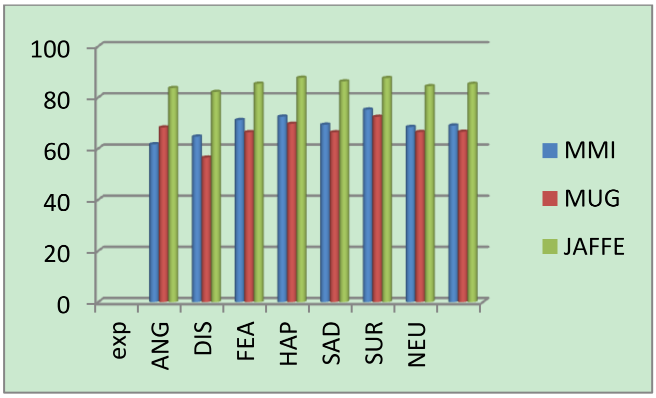

Fig. 4. Facial expression accuracies in percentage using experiment 1.

TABLE III: FACIAL EXPRESSION ACCURACIES IN PERCENTAGE

DS / exp	MMI	MUG	JAFFE
ANG	73.20	78.36	93.70
DIS	74.30	75.83	95.73
FEA	73.53	77.84	95.45
HAP	75.26	79.81	96.10
SAD	78.29	78.74	96.80
SUR	79.37	74.64	97.80
NEU	78.82	76.63	94.30
AVG	**76.11**	**77.40**	**95.69**

In this study, we observed different expressions for facial expression classification problem in this paper. These expressions include sadness, disgust, neutral, happiness, anger, surprise and fear. For The images in the database images only in grayscale and color in png, jpg and tiff file formats. These are frontal posture and eyes are approximately the similar location. It is evident from the results that execute every proposed method and clear for facial expression classification problems.

V. CONCLUSION

Automatic classification of facial expression method has been proposed in this paper. We discover and extract the region of interest (ROI) and in this problem the ROI is the facial area. Then, the functions of the window depending on the local features extracted DCT, wavelet, Haar transform, and Gabor wavelet -based functions. These functions are provided for a bank for SVM learning classification algorithm available. The pictures of expressions were of different classes. A detailed experimentation has been performed to show results.

REFERENCES

[1] S. B. Kazmi, Qurat-ul-Ain, and M. A. Jaffar, "Wavelets based Facial Expression Recognition using a Bank of Neural Networks," in *Proc. Fifth International Conference on Future Information Technology*, 2010, pp. 1-6.

[2] R. Polikar, L. Udpa, S. S. Udpa, and V. Honavar, "Learn++ An incremental learning algorithm for supervised neural networks," *IEEE Trans. Systems, Man. Cybern. C*, vol. 31, no. 4, pp. 497–508, 2001.

[3] M. A. Jaffar, A. Hussain, and A. M. Mirza, "Fuzzy entropy based optimization of clusters for the segmentation of lungs in CT scanned images," *Knowledge and Information Systems*, vol. 24, issue 1, pp. 91-111 July 2010.

[4] S. B. Kazmi, Qurat-ul-Ain, and M. A. Jaffar, "Wavelets based facial expression recognition using a bank of support vector machine," *Soft Computing*, pp. 369-379, vol. 16, issue 3, March 2012.

[5] C. Darwin, "The expression of the emotions in man and animal," *J. Murray*, London, 1872.

[6] P. Ekmanand and W. Friesen "A Technique for the Measurement of Facial Movement," *Facial Action Coding System*, Consulting Psychologists Press, 1978.

[7] MPEG Video and SNHC Text of ISO/IEC FDIS 14 496-3: Audio, in Atlantic City MPEG Mtg. Doc. ISO/MPEG N2503, 1998.

[8] M. Pantic and L. Rothkrantz, "Automatic analysis of facial expressions: "The state of the art," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1424–1445, 2000.

[9] J. F. Cohn and T. Kanade, "Recognizing action units for facial expression analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 97 – 115, 2001.

[10] I. Kotsia and I. Pitas, "Facial expression recognition in image sequences using geometric deformation features and support vector machines," *IEEE Trans. on Image Processing*, vol. 16, no. 1, 2007.

[11] Y. Cheon and D. Kim, "Natural facial expression recognition using differential-AAM and manifold learning," *Pattern Recognition 42*, pp. 1340-1350, 2009.

[12] L. Lin *et al.*, "A stochastic graph grammar for compositional object representation and recognition," *Pattern Recognition*, vol. 42, pp. 1297–130, 2009.

[13] R. Zhi and Q. Ruan, "Discriminant spectral analysis for facial expression recognition," in *Proc. IEEE Int. Conf. on Image Processing*, pp. 1924-1927, 2008.

[14 C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: a comprehensive study," *Image and Vision Computing*, vol. 27, pp. 803-816, 2009.

[15] G. Zhao and M. Pietikainen, "Boosted multi-resolution spatiotemporal descriptors for facial expression recognition," *Pattern Recognition Letters*, vol. 30, pp. 1117-1127,2009.

[16] L. Lin *et al*., "Representing and recognizing objects with massive local image patches," *Pattern Recognition*, vol. 45, pp. 231–240, 2012.

[17] Z. Niu and X. Qiu, "Facial expression recognition based on weighted principal component analysis and support vector machines," *Advanced Computer Theory and Engineering (ICACTE)*, pp. 174-178, 2010.

[18] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," *IEEETrans.Patt, Anal. Mach.Intell*., vol. 19, no. 7, pp. 711–720, 1997.

[19] M. S. Zia and M. A. Jaffar, "An adaptive training based on classification system for patterns in facial expressions using SURF descriptor templates," Multimedia Tools and Applications, December 2013 DOI: 10.1007/s11042-013-1803-3.

M. Arfan Jaffar received Ph.D. degree from the Department of Computer Science, NUCES, Islamabad, Pakistan in 2009. He has worked as a research professor in SIPL, GIST, Korea for two year. He is working as an assistant professor in College of Computer and Information Sciences at Imam University, Riyadh, Saudi Arabia. His research interests include computer vision, image processing and machine learning techniques.

Eisa Abdullah Aleisa received Ph.D. degree from the Department of Computer Science, Lehigh University, Bethlehem, PA, USA in 2000. He is working as the dean of College of Computer and Information Sciences at Imam University, Riyadh, Saudi Arabia. His research interests include heterogeneous distributed networks, image processing and machine learning techniques.