On the Application of Lip Features in Classifying Human Emotions

M. Karthigayan[1], M. Rizon[1,2], S. Yaacob[1], R. Nagarajan[1], M. Sugisaka[2,3], M. Rozailan[4] and D. Hazry[2]

[1]*School of Mechatronics Engineering, Kolej Universiti Kejuruteraan Utara Malaysia (KUKUM), Jejawi, Perlis, Malaysia*
[2]*Deparment of Electrical and Electronics Engineering, Oita University, 700 Oaza Dannoharu, Oita 870-1192, Japan*
[3]*Advance Research Institute for Science and Engineering, Waseda University, 2-2 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0135, Japan*
[4]*Terengganu Advanced Techical Institute,Kampus TATI, Jalan Panchor, Teluk Kalong, 24000 Kemaman, Terengganu, Malaysia*
Email: karthigayan@ieee.org

Abstract

In this paper, lip features are applied to classify the human emotion using a set of irregular ellipse fitting equations using Genetic algorithm. As South East Asian (SEA), is considered in this study. All six universally accepted emotions are considered for classifications. Lip is usually considered as one of the features for recognizing the emotion. In this work, three feature extraction methods are proposed and their respective performances are compared for determining the feature of the lips. The method which is fastest in extracting lip features is adopted in this study. Observation of various emotions of the subject lead to unique characteristic of lips. GA is adopted to optimize such irregular ellipse characteristics of the lip features in each emotion. That is, the top portion of lip configuration is a part of one ellipse and the bottom of different ellipse. Two ellipse based fitness equations are proposed for the lip configuration and relevant parameters that define the emotion. This has given reasonably successful emotion classifications.

Keywords: *Feature extraction, Irregular ellipse fitness function, Genetic algorithm, Face emotion recognition.*

1. Introduction

In recent years, there has been a growing interest in improving all aspects of interaction between humans and computers especially in the area of human emotion recognition by observing facial expression. Ekman and Friesen developed the most comprehensive system for synthesizing facial expression based on what they call as action units [1]. In the early 1990's the engineering community started to use these results to construct automatic methods of recognizing emotion from facial expression in still or video images [2]. Human being possesses an ability of communication through facial emotions in day to day interactions with others. Some emotions attracted most of the interest in human computer interaction environments. Double structured neural network has been applied in the methods of face detection and emotional extraction. In this, two methods are proposed and carried out; they are lip detection neural network and skin distinction neural network [3]. Facial action coding is given to every facial points. For example, code 23 is given for lip funnel, code 4 for eye brow lower, code 10 for chin raise etc. The cods are grouped for a specific facial emotion. In order to determine the category of emotion, 15 facial points in a face-profile sequence has been recommended. The algorithm performs both automatic segmentation of an input video images of facial expressions and recognition of 27 AUs occurring alone or in combinations in the input face-profile video. A recognition rate of 87% is reported [4]. The motion signatures produced are then classified using Support Vector Machines as either non-expressive or as one of the six basic emotions. The completed system is demonstrated in two simple but effective computing applications that respond in real-time to the facial expressions of the user, thereby providing the potential for improvements in the interaction between a computer user and technology [5]. The method uses edge counting and image correlation optical flow techniques to calculate the local motion vectors of facial feature [6]. Cauchy Naïve Bayes classifier is introduced in classifying the face emotion. The person dependent and Person-independent experiments shows that the Cauchy distribution assumption typically provides better results that the Gaussian distribution assumption [7]. The universally accepted categories of emotion, as applied in human computer interaction are: Sad, Anger, Joy, Fear, Disgust (or Dislike) and Surprise. In this paper, image preprocessing, filtering, edge detection methods that are suitable for feature extraction are presented and compared towards applying it to get lip features. A set of fitness function for the GA methods are also proposed as suitable for face emotion recognition. Such an approach of determining the emotions are highly suited for a personified face and cannot be common to all faces of all human being. A subject (South East Asian) as shown in Figure 1 has been used for emotion classification. The generalized process flow for the image processing, feature extraction and classifying the emotions is shown in Figure 2.

2. Face Image Processing

As the first step in image processing, the region of intesrest (ROI) has been selected in the acquired image. The ROI image is converted into grayscale image (0-256).

Figure 1. The Angry Emotion SEA

Before obtaining the filtered grayscale image, a histogram equalization method has been applied.Histogram equalization [8] improves contrast in the grayscale and the goal of histogram equalization is to obtain an uniform histogram. The histogram equalization method also helps the image to redistribute the intensity distributions as shown in Figure 3. New intensities will not be introduced into the image. Existing intensities values will be mapped to new values but the actual number of intensities in the resulting image will be equal or less than the original number of intensities. In the image sequence, the histogram equalized image is filtered using average and median filters in order to make the image smoother. Finally, Sobel edge detection method is applied to the filtered image. The cropped ROI lip region is shown in Figure 4.

Noises are added to the cropped ROI lip region. The salt and pepper noise are added to the image. This type of noise consists of random pixels being set to black or white. The application of the filter such as average filter and median filter to the noise added image is to remove the unwanted noise. This method creates a two-dimensional filter and returns with a correlation kernel. Median filtering [9] is that each output pixel is set to an average of the pixel values in the neighborhood of the corresponding input pixel. However, with median filtering, the value of an output pixel is determined by the median of the neighborhood pixels, rather than the mean. The median is much less sensitive than the mean to extreme values (called outliers). Median filtering is therefore better in the ability to remove these outliers without reducing the sharpness of the image. The median filter with various matrix sizes such as 3*3, 4*4, 5*5, 6*6, 7*7 and 8*8 are applied. The 5*5 size matrix has been found to be suitable in removing unwanted noise in the image.

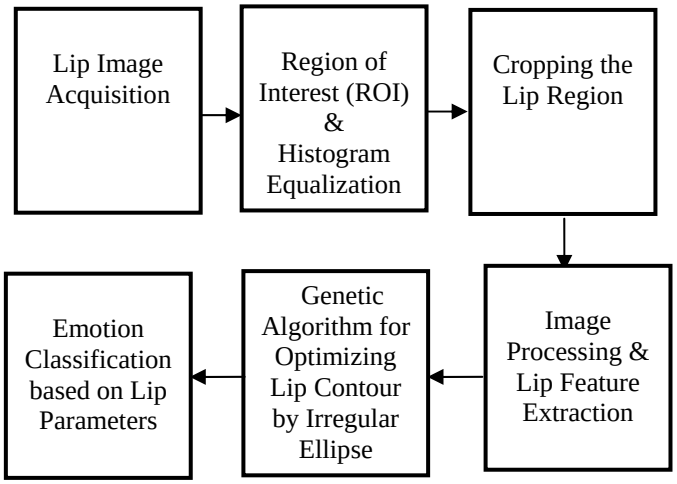

Figure 2. Process Flow of Image Processing

Figure 3. Histogram Equalization of Face ROI

Figure 4. Cropped ROI Lip

Threshold has been performed to the filtered image by selecting a suitable threshold value. Various edge detection methods such as Sobel, Prewitt, Canny, Roberts and Log have been applied to the image. A comparison has been made among the edge detection methods and it is found that the Sobel edge detection method [9] performed well compared to other four methods. The sobel edge detection region of lip region is shown in Figure 5

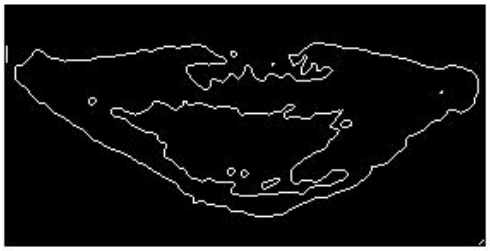

Figure 5. Sobel Edge Detected Lip Region

3. Feature Extraction

A feature extraction method is now to be applied to the edge detected image to extract features. Three feature extraction methods are considered and their capabilities are compared in order for adopting the one that is suitable for the proposed face emotion recognition problem. They are projection profile, contour profile and moments.

The performance of each of the above described feature extracting methods can be compared with respect to processing time using the edge detected image of the lips. The processing time includes the image reading, preprocessing, filtering, edge detection and feature extraction processes. Table 1 shows the processing time for obtaining the lip features of all three feature extraction methods. The projection profile is found to perform well in feature extraction with regards to the processing time and is adopted here. The projection profile has been also found to have performed well in earlier works [10,11].

Table 1. Processing Time for Lip Feature Extraction

Feature Extraction Method	Processing Time (Seconds)
Projection Profile	0.691 – 0.862
Moments	0.865 – 0.892
Contour Profile	23.07 – 33.31

4. Face Emotion Recognition using Genetic Algorithm

In the early 1970s, John Holland, one of the founders evolutionary computations, introduced the concept of genetic algorithm [12]. Genetic algorithm (GA), or GA is a heuristic method used to find approximate solutions to solve problems through application of the principles of evolutionary biology. GA adopts biologically-derived techniques such as inheritance, mutation, natural selection, and recombination (or crossover). GA is a particular class of evolutionary algorithms. A population containing a number of trial solutions each of which is evaluated (to yield fitness) and a new generation is created from the better of them. The process is continued through a number of generations with the aim that the population should evolve to contain an acceptable solution. GA is well known for optimization of nonlinear functions. It offers the best optimized value for any fitness or objective functions suitably selected for particular problems.

GA has been applied in various applications which include in image processing, control, design of aircraft, robot trajectory generation, air-injected hydro cyclone optimization, multiple fault diagnosis, a hybrid technique for engineering design optimization the traveling salesman, sequence scheduling and quality solutions where nonlinear optimization are required [13] Some aspects of vision system and image processing methodologies have been discussed towards approximating the face as a best ellipse using GA. In the feature extraction stage, the GA is applied to extract the facial features such as the eyes, nose and mouth, in a set of predefined sub regions. Some simulation has been carried out [9]. A method that extracts region of eyes out of facial image by genetic algorithm has been suggested recently [14].

The human lip shape is more of towards combination of two ellipse and we call this is as irregular ellipse. The word 'irregular' means that the ellipse has two different minor axes wherein a major axes remains the same. The preprocessed lip image is considered as a irregular ellipse. Lengths of minor axes of the lip feature for each emotion are computed. In Figure 6, the whitened area of edge detected lip image for a particular emotion of SEA is shown. The major axis "2a" (considered to be fixed) and two minor axes are "2b1" and "2b2" (to be computed). This is shown in Figure7. A general form of the regular ellipse is defined by

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \qquad (1)$$

Figure 6 Image Processed Lip for SEA

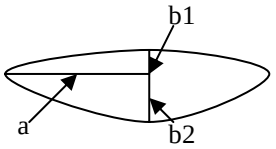

Figure 7 Irregular Ellipse with Minor and Major Axis

A fitness equation is suggested in this work in order to find the two minor axes so that the emotion changes can be recognized. A fitness function, to be discussed latter, for applying GA, is derived to optimally compute semi-minor axes, b1 and b2, for the top lip area and bottom lip area respectively; emotions can thus be related to the values of b1 and b2.

4.2 Fitness Function

A fitness function is a particular type of objective function that quantifies the optimality of a solution (that is, a chromosome) in a GA problem so that that particular chromosome may be ranked against all the other chromosomes. A fitness value reflecting the amount of overlapping between the regions covered by the overlaid boundaries is computed for each chromosome. A pair of individuals are selected with a probability proportional to their fitness and mated to reproduce their next generation. The process is repeatedly performed with the same number of individuals of the previous epoch. The fitness function equation, Equation (2) and Equation (3) with b1 and b2 are derived based on the general ellipse equation, Equation (1).

$$f(x)=\Big(\sum_{i}^{m}\sum_{j}^{n} col(j)-2\sqrt{X_1^2\Big(1-\frac{row(i)^2}{a^2}\Big)}\Big)^2 \qquad (2)$$

$if X_1 \geq 0$

$$\overline{f(x)}=\Big(\sum_{i}^{m}\sum_{j}^{n} col(j)-2\sqrt{X_2^2\Big(1-\frac{row(i)^2}{a^2}\Big)}\Big)^2 \qquad (3)$$

$if X_2 \leq 0$

In Equation (2) and Equation (3), X_1 and X_2 are variables which finally converged to optimized values as b1 and b2 respectively. White pixels of the preprocessed lip feature for each column are calculated. In Equation (2) and Equation (3), col(j) is sum of white pixels occupied by j^{th} column and row(i) is number of rows of white pixels in i^{th} row. Equation (2) and Equation (3) are fitness equations of top and bottom lip respectively.

4.3 Results and Discussion

In this study on a SEA subject, six emotions and one neutral have been considered. The lip features have been given as input to the genetic algorithm to find the optimized values. The selected values of GA parameters are indicated in the Table 2. The process of optimization has been carried out for 5 times for each emotion. This process of optimization is found to be giving favorable two minor axis value b1 and b2 for top and bottom areas of the lips. Table 3 indicates the manually measured values of b1 and b2 and the

corresponding optimized values of X1 and X2. The emotion based on minor axes of the lip feature can now be estimated for the subject. The experiment result shows that the two minor axis (b1 and b2) of the lip feature are different for each emotion there by distinctions are possible.

Table 2 Parameter Settings

Generation	250
Population size	20
Fitness scaling	Rank
Selection Function	Roulette
Mutation	Gaussian
Crossover	Scattered
Stall generation	50
Stall time	20

Table 3 Classification of Emotion (SEA)

Emotions	**Manually Computed Value**		**Optimized Mean Value by GA**	
	b1	**b2**	**X1**	**X2**
Neutral	44	56	49.9294	47.7074
Fear	38	57	40.8410	54.9664
Happy	31	178	22.5574	74.1890
Sad	44	16	45.5745	17.9480
Angry	19	61	12.6218	64.5633
Dislike	76	47	37.9206	41.9958
Surprise	60	75	52.1640	77.4762

5. Conclusion

In this paper, a set of suitable sequences in image processing and new fitness function for the GA application have been proposed for determining the top-lip and bottom-lip features. The face and lip are obtained under uneven lighting. In image processing task, the average and median filters are applied to smoothen the image. The Sobel edge detection is found to perform well for edge detection. The suggested sequence of image processing offers acceptable lip images compared to other methods suggested in this work. The lips are considered as an irregular ellipse. The GA is then applied to get the optimized values of the minor axes, b1 and b2, of the irregular ellipse by using a set of proposed fitness functions. These optimized values of minor axes indicate that the values b1 and b2 can be regarded as identifying the emotion. The new method has shown successful classifications emotion. The proposed method of emotion classification is a general approach to lips of any face. The parameters listed in Table 3 are for a personified face. Such a table look-up can be applied to a particular bed-ridden disabled dumb patient for identifying the patient's reaction.

References

1) Haibo Li, "Computer Recognition of Human Emotion", Proceedings of International Symposium on Intelligent Multimedia, Video and Speech Processing, Hong Kong, 2-4 May 2001, pp. 490-493.
2) Nicu Sebe, Michael S. Lew, Ira Cohen, Ashutosh Garg and Thomas S. Huang, "Emotion Recognition Using a Cauchy Naïve Bayes Classifier", Proc. of Sixteenth International Conference on Pattern Recognition, vol. 1, 11-15 August 2002, pp. 17-20.
3) Hironori Takimoto, Yasue Mitsukura, Minoru Fukumi and Norio Akamatsu, "Face Detection and Emotional Extraction System Using Double Structure Neural Network", *Proc. of the International Joint Conference on Neural Networks*, vol. 2, pp. 53-1257, 20-24 July 2003.
4) Maja Panti and Ioannis Patras, "Dynamics of Facial Expression: Recognition of Facial Actions and Their Temporal Segments from Face Profile Image Sequences, *IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics*, Vol. 36, No. 2, pp. 433-449, April 2006
5) Keith Anderson and Peter W. McOwan, "A Real-Time Automated System for the Recognition of Human Facial Expressions", *IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics*, Vol. 36, No. 1, pp. 96-105, February 2006.
6) Liyanage C De Silva and Suen Chun Hui, "Real-time Facial Feature Extraction and Emotion Recognition", Proceedings of Fourth Pacific RIM Conference on Multimedia, vol. 3, 15-18 Dec 2003, pp. 1310-1314.
7) Nicu Sebe, Miachel S. Lew, Ira Cohen, Ashutosh Garg and Thomas S. Huang , "Emotion Recogmnition Using a Cauchy Bayes Classsifier". Proceeding of 16th International Conference on pattern Recognition, vol. 1, 11-15 August 2002, pp. 17-20.
8) Rafael C Gonzalez and Richard E Woods, "Digital Image processing". Pearson Education, Inc, India, 2002.
9) Gary G. Yen and Nethrie Nithianandan, "Facial Feature Extraction Using Genetic Algorithm", Proceedings of Congress on Evolutionary computation, vol. 2, 12-17 May 2002, pp. 1895-1900.
10) M. Karthigayan, Mohammed Rizon, Sazali Yaacob and R. Nagarajan, "An Edge Detection and Feature Extraction Method Suitable for Face Emotion Detection under Uneven Lighting", The 2nd Regional Conference on Artificial Life and Robotics (AROB'06), July 14 - July 15, 2006, Hatyai, Thailand.
11) R. Nagarajan, Sazali Yaacob, Paulraj Pandiyan, M. Karthigayan, Marzuki Khalid and Shamsudin H Amin, "Real Time Marking Inspection Scheme for Semiconductor Industries", International Journal of Advance Manufacturing Technology (IJAMT), Springer Publisher, 2006. (Accepted for publication).
12) Michael Negnevitsky, "Artificial Intelligence", Addison Wesley, Pearson education Limited, England, 2002.
13) http://neo.lcc.uma.es/TutorialEA/semEC/cap03/cap_3.html, 1997.
14) Hideaki Tani, Kenji Terada, Shunichiro Oe and Junichi Yamaguchi, "Detecting of One's Eye from facial Image by Using Genetic Algorithm", The 27thAnnual Conference of the IEEE Industrial Electronics Society, 2001, pp. 1937-1940.