# Estimating Face Emotion Using Genetic Algorithm

**M. Karthigayan, Mohamed Rizon, Sazali Yaacob and R. Nagarajan**

**Status: Published**

**The International Conference in Man-Machine System (KUKUM)**
**15-16 September 2006, Langkawi**

# Estimating Face Emotions using Genetic Algorithm

**M. Karthigayan, Mohamed Rizon, Sazali Yaacob, R. Nagarajan**

*School of Mechatronics Engineering*
*Northern Malaysia University College of Engineering (KUKUM), 02600, Jejawi, Perlis, Malaysia*
*Email: karthigayan@ieee.org*

## Abstract
*Recognition of emotion through face features (Face Emotion) is a recent concept undertaken by several researchers. Face features have to be extracted from face images before applying the emotion recognition techniques. This paper contemplates these aspects in two parts. The first part describes various processing stages in order to obtain a suitable processed image for applying techniques of face emotion recognition. Several stages of processing lead to segmentation of image. Three methods are proposed in extracting the features and their capabilities are compared. The second part discusses a Genetic Algorithm methodology of estimating the emotions from eye feature alone. Observation of various emotions lead to an unique characteristic of eye, that is, the eye exhibits ellipses of different parameters in each emotion. Genetic Algorithm is adopted to optimize the ellipse characteristics of the eye features. A new form of fitness function is proposed for the Genetic Algorithm. It is ensured through several experiments that the optimized parameters of ellipse reveal various emotional characteristics.*

## Keywords
*Feature extraction, Ellipse fitness function, Genetic algorithm, 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 through 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 beings possess an ability of expressing their emotion through eyes in day to day interactions with others. One set of category of emotions has attracted most of the interest in human computer interaction environments. The universally accepted category of emotions, as applied in human computer interaction is: Sadness, Anger, Joy, Fear, Disgust or Dislike and Surprise. This paper consists of mainly two parts. The first part of this paper describes various stages in image processing towards applying the techniques of emotion determination. The processing stages include preprocessing, filtering and edge detection. Three methods of extracting face features (eyes and lips) are proposed and their characteristics towards determining the emotions are compared. The second part of this paper discusses a Genetic Algorithm (GA) based approach for estimating the emotion through eye feature alone. A new form of fitness function adopting the characteristics of ellipse is suggested. It is ensured through several experimentation that the eye feature is well suited in finding the emotions through GA.

## 2. Face Image Processing
Face features are unique to one face thus vary from one person to another. Any emotion determining parameters derived from face features cannot be the same within any two persons. For example, the eye features of a Chinese are totally different from those of ASEAn. Hence, the determination of emotion is personalized and hence suitable to one person only. Figure 1 indicates an ASEAn face expressing the emotion 'anger'. Figure 2 describes the block schematic of image processing and emotion determination stages that is to be undertaken in this paper.

In this work, a digital camera with 3.2 Mega pixels has been used to acquire face image. The region of interest (ROI) has been selected in the acquired image. The ROI is then converted into grayscale image (0-256 range) as the first stage of image processing.

*Figure 1. The Angry Emotion*

Before applying the filter to the grayscale image, a histogram equalization method is applied. Histogram equalization [3] improves contrast and achieves the goal of this equalization method which is to obtain an uniform histogram. The histogram equalization method also helps the image to reorganise the intensity distributions as shown in Figure 3. New intensities will not be introduced into the image. Existing 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.

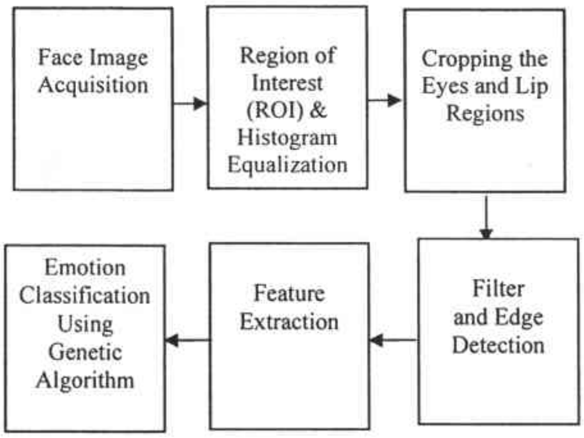


*Figure 2.A General Process Flow of face emotion recognition*

*Figure 3. Histogram Equalization of Face ROI*

The histogram equalized image is filtered using average and median filter to make the image smoother. Finally, Sobel edge detection method is applied to the filtered image. Due to the problem of light intensity variations, the segmentation process applied to entire image was not successful. In the edge detected image of the whole face, the eyes are properly segmented where as the lip segmentation is poor as shown in Figure 4. So the histogram equalized image is split into eyes ROI ( region of interest) and lip ROI regions. These ROI regions are shown in Figure 5 and Figure 6.

*Figure 4. Sobel Edge Detection for ROI Face*

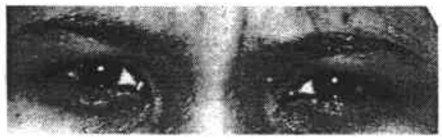

*Figure 5. Cropped ROI Eyes Region*

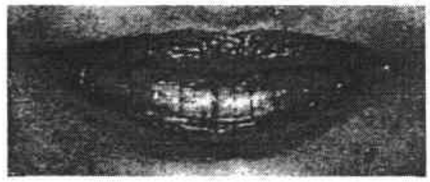

*Figure 6. Cropped ROI Lip Region*

Noises are added to the cropped ROI eyes and lip regions. 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. The noise added images are shown in Figure 7 and Figure 8. The Average filter thus creates a two-dimensional filtered image and returns with a correlation kernel. Median filtering [4] makes each output pixel 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 its 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 to the image.

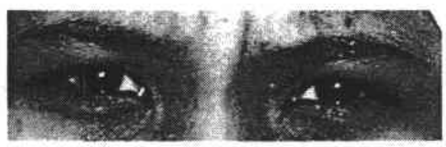

*Figure 7. Noise Added Eyes Region*

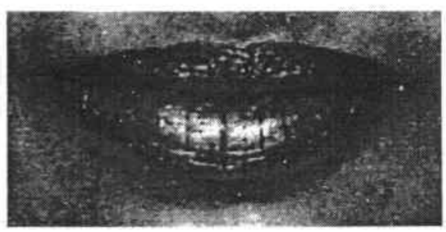

*Figure 8. Noise Added Lip Region*

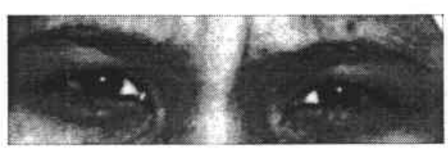

*Figure 9. Smoothened Eyes Region Image by Median Filter*

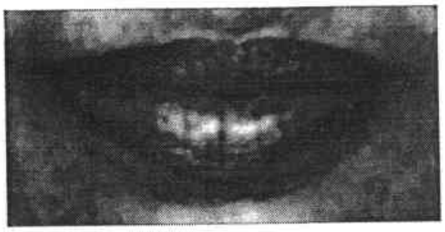

*Figure 10. Smoothened Lip Region image by Median Filter*

Thresholding is then performed to the filtered image by selecting a suitable threshold value. Then, 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 [4] performed well compared to other four methods. The sobel edge detection applied to eyes region and lip region are as shown in Figure 11 and Figure 12 respectively.

*Figure 11. Sobel Edge Detected Eyes Region*

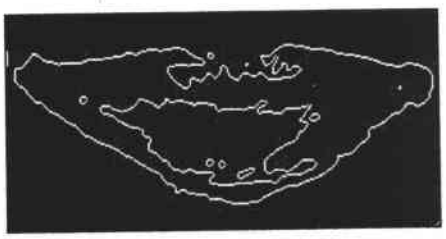

*Figure 12. Sobel Edge Detected Lip Region*

The processing time for all these methods such as preprocessing, filtering and edge detection of the face image is in the range of 1.03 to 1.23 seconds. The processing time is based on a PC with Pentium Mobile processor 1400 MHz.

## 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 for adopting the one that is suitable for the proposed face emotion recognition problem. They are projection profile, contour profile and moments.

### 3.1 Projection Profile

This feature extraction method is connected with the row sum and column sum of white pixels [6]. The pattern of row-sum ($P_h$) and the pattern of column-sum ($P_v$) of white pixels define as the feature of each region. The set { $P_h$ , $P_v$ } is known as projection profile and this set is the extracted feature. These patterns are extracted for each region of image. Let S(m,n) represent a binary image of m rows and n columns. Then, the vertical profile is defined as the sum of white pixels of each column perpendicular to the x-axis; this is represented by the vector $P_v$ of size n as defined by

$$p_{v[j]} = \sum_{i=1}^{m} s[i,j] \qquad j = 1,2,3,\ldots,n \tag{1}$$

The horizontal profile is the sum of white pixels of each row perpendicular to the y-axis; this is represented by the vector $P_h$ of size m, where

$$p_{h[i]} = \sum_{j=1}^{n} s[i,j] \qquad i = 1,2,3,\ldots,m \tag{2}$$

### 3.2 Moments

The moments have been widely used in pattern recognition [5]. Several desirable properties that can be derived from moments are also applicable to shape analysis. The processing time of Central moments is faster than Zernike moments and the moments invariant. Central moments of binary image orders can be obtained for each column of the image. The moment orders can be of 2 or 3. In the moment of order 1, the moment values are zero. On the other hand, orders more than 3 produce smaller and smaller moment values that cannot be generally useful for feature extraction.

Let f(x,y) be an image. Then, the 2D continuous function of the moment of order (p+q), $M_{pq}$, is defined as [5].

$$M_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q f(x,y)\,dx\,dy \tag{3}$$

The central moment, $\mu_{pq}$, of f(x,y) is defined as [5].

$$\mu_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\bar{x})^p (y-\bar{y})^q f(x,y)\,dx\,dy \tag{4}$$

where $\bar{x} = \frac{M_{10}}{M_{00}}$ and $\bar{y} = \frac{M_{01}}{M_{00}}$.

If f(x,y) is a digital image then Equation (4) becomes

$$\mu_{pq} = \sum_{x}\sum_{y} (x-\bar{x})^p (y-\bar{y})^q f(x,y) \tag{5}$$

where p and q are of nonnegative integer values. The moment values can be considered as extracted features.

### 3.3 Contour Profile

This is one of the fundamental techniques used for object identification in the field of pattern recognition [6]. The outer vertical and horizontal profiles of black pixels in white background are computed. Contour profile can be used here as the feature of the lip and eyes region. The contours for eyes and lips region are shown as in Figure 13 and Figure 14 respectively.

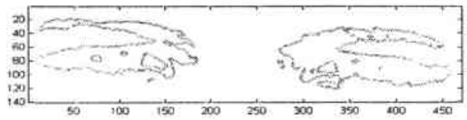


*Figure 13. Contour for Eyes Region*

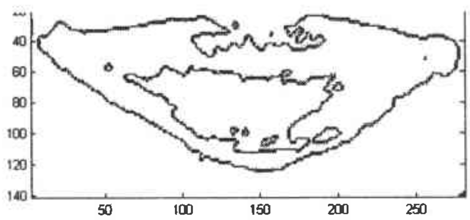


*Figure 14. Contour for Lip Region*

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 eyes and lip. The processing time includes the image reading, preprocessing, filtering, edge detection and feature extraction processes. Table 1 shows the processing time (including eyes and 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.

*Table 1. Processing Time for Feature Extraction*

| Feature Extraction Method | Processing Time (Seconds) |
|---|---|
| Projection Profile | 1.06 - 1.23 |
| Moments | 1.24 – 1.28 |
| Contour Profile | 50.43 – 68.90 |

## 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 [7]. Genetic algorithm (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 as an optimizing method. It offers the best optimized value for any fitness and objective function. It can be used in minimization or maximization problems.

GA has been applied in various applications such as in image processing, control, design of aircraft, robot trajectory, multiple fault diagnosis, engineering design optimization, the traveling salesman, sequence scheduling and so on [8]. The GA is applied to extract the facial features such as the eyes, nose and mouth, in a set of predefined sub regions. Some simulations have been carried out using GA [4]. A method has been suggested to extract regions of eyes out of facial image by GA in order to detect one's eye [9].

The human eye shape is of more towards ellipse. The preprocessed eye image can be considered as an ellipse. The major axis of this ellipse is more or less fixed for the eye of a particular person Then, the minor axis of the eye feature can be related to the emotion.. The whitened area of edge detected eye image for a particular emotion is shown in Figure 15. The ellipse can be parameterised by its minor and major axes. The major and the minor axes are "2a" ( more or less fixed) and '2b" (to be computed) respectively. This is shown in Figure 16. The ellipse is defined by its equation as

Figure 15 Image Processed Eye

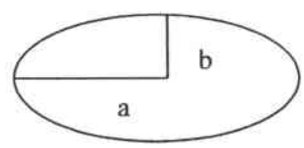


Figure 16 Ellipse with Minor and Major Axis

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \qquad (5)$$

A fitness function, to be described latter, for applying GA is derived to optimally compute minor axis, b, such that the emotions can be easily related to the values of b.

### 4.1. The Algorithm

GA represents an iterative process; each iteration is known as generation. The algorithm is illustrated as follows [7]:

**Step 1**. Represent the problem variable domain as chromosome of a fixed length, chromosome population, the cross over probability and mutation probability
**Step 2**. Define a fitness function to measure the performance or fitness of an individual chromosome in the problem domain
**Step 3**. Randomly generate an initial population of chromosome of size N
**Step 4**. Calculate the fitness of each individual chromosome
**Step 5**. Select a pair of chromosomes for matting from the current population. Parent chromosomes are selected with a probability related to their fitness. Highly fit chromosomes have a higher probability of being selected for mating than the less fit chromosomes.
**Step 6**. Create a pair of offspring chromosomes by applying the genetic operators – the crossover and the mutation
**Step 7**. Place the created offspring chromosomes in the new population
**Step 8**. Repeat from step 5 until the size of new chromosome population becomes equal to the size of the initial population
**Stpe 9**. Replace the initial chromosome population with the new population
**Step 10**. Go to step 4, and repeat the process until the termination criterion is satisfied.

### 4.2 The Fitness Function

A fitness function is a particular type of objective function that quantifies the optimality of a solution (that is, a chromosome) in the GA. This 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 performed in a repetitive manner with the same number of individuals of the previous epoch is formed. The fitness function, Equation (2), as shown below, is developed based on the general ellipse Equation (1). The proposed fitness function is to find the minor axis of the eye feature using GA. The fitness function is derived as

$$f(x) = \left( \sum_{i}^{m} \sum_{j}^{n} col(j) - 2\sqrt{X^2\left(1 - \frac{row(i)^2}{a^2}\right.} \;\right)^2 \quad (2)$$

where col(j) is the sum of white pixels occupied in each column, row(i) is number of rows and "X" is expected optimized value of "b". The constant "a" is a measure of major axis. The values of m and n represent the number of rows and the number of columns respectively.

### 4.3 Results and Discussion

In this analysis, a subject (South East Asian) expressing 6 different emotions and the neutral face is considered. The eye feature is given as input to the GA to compute the optimized value of b. The optimization is performed for more than 5 times for each emotion in reaching consistent minor axis value of b. Table 2 indicates one set of parameters used in the process of obtaining the optimized values of b. Table 3 illustrates the values of b for each emotion. The experiment results demonstrate that the minor axis of the eye feature is different for each emotion. This is the advantage of using GA to get the optimized values of b from unpredicted shapes of processed images towards classifying the emotions. Table 3 concludes that each emotion has unique minor axis values. Table 3 also indicates the number of rows (m, as a measure of minor axis ) and the number of columns (n, as a measure of major axis ).

## 5. Conclusion

In this paper, a set of suitable preprocessing, filtering, edge detection and feature extraction methods for applying Genetic Algorithm have been proposed. Three different feature extraction methods are suggested and compared for their individual performances. Projection profile method of feature extraction has been adopted for its speed of response and for its requirement of less complex computations. All these image processing stages work favorably well even when the face image is obtained under uneven lighting. Genetic Algorithm is then applied for estimating the emotions. Eye is used as the emotion determining feature. The developed fitness function for applying the Genetic Algorithm exhibits its capability in identifying the emotions. The fitness function is derived from the theory of ellipse in which a measure of minor axis, b, is related to the emotions. It is to be indicated here that a mean value of b for an emotion is obtained through several experiments. However, variations of b do occur from one experiment to another. These variations can give information on real time changes of emotions from one to another. Work is progressing to use fuzzy logic in considering these variations towards determining the transition of one emotion to another.

*Table 2 Parameter Settings*

| | |
|---|---|
| Generation | 200 |
| Population size | 20 |
| Fitness scaling | Rank |
| Selection Function | Roulette |
| Mutation | Gaussian |
| Crossover | Scattered |
| Stall generation | 50 |
| Stall time | 20 |

*Table 3 Classification of Emotion*

| Emotions | n | m | Optimized Mean Value of "b" | Duration of Emotion recognition by GA (sec) |
|---|---|---|---|---|
| Neutral | 202 | 72 | 33.7993 | 142 |
| Fear | 184 | 70 | 37.8598 | 162 |
| Happy | 176 | 60 | 27.6636 | 118 |
| Sad | 186 | 66 | 33.8305 | 149 |
| Angry | 192 | 50 | 22.4746 | 99 |
| Dislike | 162 | 50 | 25.2514 | 113 |
| Surprise | 210 | 78 | 39.9059 | 172 |