Edge-Based Facial Feature Extraction using Adaptive Canny Operator Edge Detection

Darli Myint Aung
University of Computer Studies, Mandalay
darlimyintaung@gmail.com

Abstract

Facial feature extraction is an essential step in the face detection and facial expression recognition frameworks. To develop a better facial expression recognition system, a good feature extraction method is needed. In this paper, an efficient Facial Feature Extraction method for recognizing four different expressions such as neutral, happy, surprise and sad is presented. In this study, adaptive canny operator edge detection method is used to reduce the computational complexity and improved the accuracy of feature point location. To validate the performance of the proposed feature extraction, the generated features are classified using Maximum Correlation Classifier (MCC). The experimental results demonstrated that the proposed feature extraction method could generate significant facial features and these features are able to be classified into each expression. Our results also showed that the proposed feature extraction method is more efficient than Gabor wavelet edge detection method.

Keywords: Adaptive Canny Operator Edge Detection, Facial Feature Extraction, Maximum Correlation Classifier, Gabar Wavelet edge detection method

1. Introduction

Facial expression is one of the most powerful, natural, and immediate means for human beings to communicate their emotions and intentions. Facial expression carries crucial information about the mental, emotional and even physical states of the conversation. It is a desirable feature of the next generation human-computer interfaces.

Facial feature extraction is a vital step in the face detection and facial expression recognition frameworks. The meaningful information of the face can be achieved from the extracted features. To develop a better facial expression recognition system, a good feature extraction method is needed so that the system can successfully recognize different facial expressions. In this research area, feature extraction is the most difficult and challenging task. Many researchers have proposed variety of techniques for feature extraction, and have tried to solve the problems that exist in this stage.

In prior works, there are numerous feature extraction techniques had been proposed such as Active Appearance Model (AAM), Active Shape Models (ASM), optical flows, eigenfaces and edge detection. Recently, some researchers combined several techniques in their algorithms to increase the recognition rate. Other method to detect edges in an image is a wavelet transform. Haar wavelet and Gabor wavelet have been used actively in face detection. [6] and [7] proposed Haar-like-features that have a similarity with Haar wavelet in their face detection system and then the system was successful to detect face in simulation and real-time application.

However these techniques are relatively complex and difficult to construct. Among those techniques, edge detection is the common technique and easy to implement in face detection and facial expression recognition. In general, edge detection has difficulty adapting to different situations.

The motivation of this paper is the effective facial feature extraction. This paper focuses on the extraction of specific facial components which are eyes and mouth. The contribution of this system is that the geometric features from a facial image are obtained by applying the adaptive canny operator edge detection. The main aim is to generate significant facial features that are able to be used in facial expression recognition system. In this study, adaptive canny edge detection method is used as edge detector.

The number of generated features is large; therefore these features are compressed by using the Principal Component Analysis (PCA). To prove that the proposed feature extraction could generate significant facial features, we classify these features into different face using maximum correlation classifier (MCC).

The paper is arranged as follows. Section 1 gives introduction, section 2 shows Related Work, section 3 describe system overview of facial expression recognition , section 4 give adaptive canny edge detection method, section 5 present Gabor Wavelet, section 6 explain results and discussion, and section 7 presents conclusions and future work.

2. Related Work

In recent years, the research of developing automatic facial expression recognition systems has attracted a lot of attention from many different fields. A more recent and complete overview can be found in [5, 2]. The approaches to facial expression recognition can be, roughly divided into two classes: geometrical feature-based approaches and appearance-based approaches [9].

The number of methods that have been developed for representing faces for identity recognition may also be powerful for expression analysis. These approaches are also included in the present comparison. These include Gabor wavelets [9], [4], linear discriminant analysis [5] and local feature analysis [2].

Facial expression recognition using still images often use feature based methods [1] for recognition and thus have fairly fast performance but the challenge in this approach is to develop a feature extraction method that works well regardless of variations in human subjects and environmental conditions.

Fasel fulfills the recognition of facial action units, i.e., the subtle change of facial expressions, and emotion-specified expressions. The optimum facial feature extraction algorithm, Canny Edge Detector, is applied to localize face images, and a hierarchical clustering-based scheme reinforces the search region of extracted highly textured facial clusters [3].

3. System Overview of the Facial Expression Recognition

The flow chart of facial expression recognition system is shown in figure 1. On the pre-processing stage, the face image is cropped manually and then transform to a grey scale representation and also histogram equalization. To get a uniform size of all images in the dataset, the face images are resized into 200x200 pixels. To extract the facial features of different facial expressions, we apply edge detection on the face region images.

At feature extraction stage, the mouth and eyes feature is extracted by using the training positions that train from the train images. And then, adaptive canny edge detection method is performed. The sizes of face image and eyes and mouth image are large, which produces a very large number of feature vectors. Since that, we need to reduce the image sizes so that the number of features would be smaller. However, the number of features is still large, and again, the dimension of features is reduced by using PCA [6]. The reduced features with different dimensions are examined to find the better features that can describe the similarities and differences of the facial expression data.

At last, to evaluate the significance of the produced facial features, these features are classified by using Maximum Correlation classifier (MCC) into facial expressions. The portion of data are made by choosing about 65% training data and 35% test data from the total number of data in each set of subject. During the training phase, Euclidean distance is used to find optimal parameter for facial expressions.

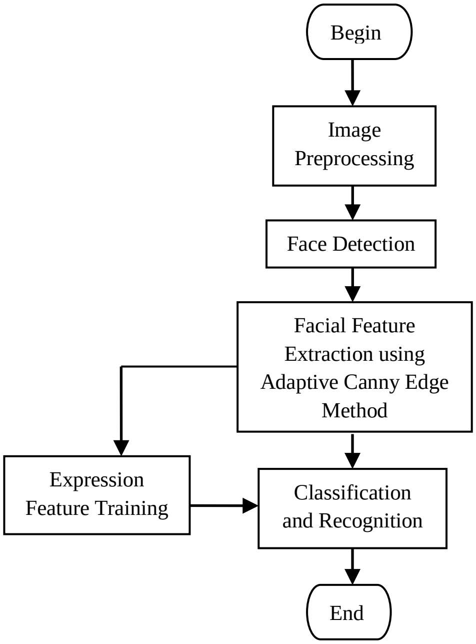


Figure 1. System Flow of Facial expression Recognition

4. Adaptive Canny edge detection method

The steps of Canny edge detection algorithm are as follows:

(1) The Gaussian filter is used to smooth the image. The chosen Gaussian function is as follows:

$$G(x) = \frac{1}{\sqrt{2\prod\sigma}} e^{-\frac{r^2}{2\sigma^2}} \quad (1)$$

where σ is the standard deviation of Gaussian curve, controlling the degree of smoothness.

(2) Use the finite difference of 2×2 neighborhood first-order partial derivatives to calculate the gradient magnitude and gradient direction of the smoothed data array I(x,y).

$$|G| = \sqrt{G_x^2 + G_y^2} \quad (2)$$

$$\theta = \arctan(\frac{G_y}{G_x}) \quad (3)$$

(3) Use non-maxima suppression method for the gradient magnitude. Non-maxima suppression is expressed as follows:

$$N[i, j] = NMS(M[i, j], \xi[i, j] \quad (4)$$

(4) Use double-threshold algorithm for the edge detection and connection.

First, we improve the calculation method of the gradient magnitude of the image. The method determine the pixel gradient amplitude by calculating a finite difference of first-order partial derivatives of the x direction, y direction, 45° direction, 135° direction in the neighborhood of 8 pixels. The calculation formula is (5):

$$M[i, j] = \sqrt{P_x[i, j]^2 + P_y[i, j]^2 + P_{135}[i, j]^2 + P_{45}[i, j]^2} \quad (5)$$

where: $P_x[i, j] = I[i+1, j] - I[i-1, j]$ (6)

$$Py[i, j] = I[i, j+1] - I[i, j-1] \quad (7)$$

$$P_{135}[i, j] = I[i+1, j+1] - I[i-1, j-1] \quad (8)$$

$$P_{45}[i, j] = I[i-1, j+1] - I[i+1, j-1] \quad (9)$$

Gradient direction is:

$$\theta[i, j] = \arctan(P_y[i, j] / P_x[i, j]) \quad (10)$$

Then in the calculating of the dynamic threshold value, the entire image is segmented into a number of sub-images and the sub-image regions can overlap, in which the scale parameter of the overlap region accounting for the sub-image is ρ. Then according to the results after the non-maxima suppression, set the high and low thresholds of the sub-images adaptively.

According to the edge gradient information of each sub-image, generate dynamic threshold adaptively by considering the feature information of the global edge gradient. Generally, the proportion of non-edge is much larger than the proportion of the edge in images. In the gradient histogram, the gradient value with a maximum number of pixels is called $H_{\max}$. The variance calculated by the gradient of all the pixels within the sub-image related to $H_{\max}$ is called $\sigma_{\max}$.

$$\sigma_{\max} = \sqrt{\sum_{i=0}^{k} (H_i - H_{\max})^2 / N} \quad (11)$$

where: k is the gradient maximum number of non-0 pixels. N is the total number of pixels.

The high threshold value τ_h of each sub-image must be selected outside the non-edge region in the gradient histogram. Otherwise, it will bring a lot of pseudo-edge noise to the final result. The formula calculating τ_h and τ_l is:

$$\tau_h = H_{\max} + \sigma_{\max} \tag{12}$$

$$\tau_l = 0.4 * \tau_h \tag{13}$$

Suppose τ_h and τ_l are respectively the global high and low threshold values of the whole image. The method accessing to them is in the same way as above. Eventually split high and low threshold on the sub-image as follows:

$$\tau_{High} = (1-\beta)\tau_H + \beta\tau_h \tag{14}$$

$$\tau_{Low} = (1-\beta)\tau_l + \beta\tau_l \tag{15}$$

where: 0< β <1. β is the threshold adjustment rate.

In this paper, in order to save time and improve efficiency, the proposed system completed the facial positioning work while the human face detection is in the process. At this time, this system only needs to carry out facial feature point location and feature extraction.

5. Gabor Wavelet

Gabor is a filter that transfers the image to wavelets in order to obtain and extract the desired feature. Gabor wavelets depend on spatial frequency, spatial locality, and orientation selectivity for extracting features. The features extracted from the Gabor filtered images might be very robust against the changing of illumination and facial expressions. In addition, the facial expression recognition system that is based on Gabor wavelet representation shows a large degree of correlation with the human semantic ratings. Gabor wavelets are represented by two-dimension plant waves in the spatial domain. The Gabor feature for Image I(x,y) is defined by the following:

$$G_{\mu,\nu}(x,y) = (x,y) \times \Psi_{\mu,\nu}(x,y) \tag{16}$$

where $G_{\mu,\nu}(x,y)$ is a convolution of image corresponding to Gabor filter μ orientation and ν scale $\Psi_{\mu,\nu}(x,y)$ is a Gabor Kernel derived from a Gaussian envelope and plane wave.

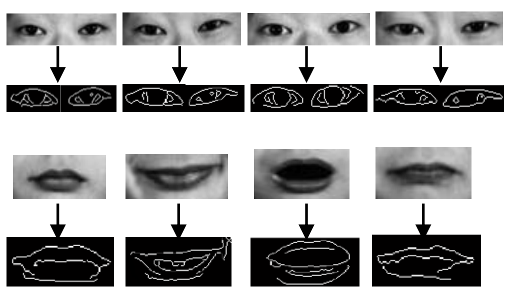
6. Results and Discussion

The feature extraction method is tested on face images from 8 subjects that are selected from Japanese Females Facial Expression (JAFFE) and our database. The Facial feature extraction method examines in four different expressions are Neutral, Happy, Surprise and Sad. Figure 2 shows the some sample image of face region by applying the step of the face detection.

Edge detection on eyes image using adaptive canny edge detection method and Gabor wavelet shows that the eyes and mouth for each facial expression has different patterns as shown in figure 3 and 4. According to the resultant edges of eyes and mouth, we can see that these features are clearly extracted rather than Gabor wavelet. The edges of eye and mouth are noticeable from the proposed method. Edges that appear on this Gabor output image are quite blur and not strong enough.

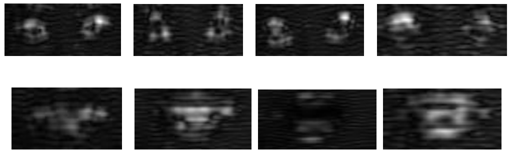
(a) (b) (c) (d)

Figure 2. Face region (a) neutral, (b) happy, (c) surprise, (d) sad



(a) (b) (c) (d)

Figure 3. Edge detection using the proposed method on eyes and mouth image for (a) neutral, (b) happy, (c) surprise, (d) sad



(a) (b) (c) (d)

Figure 4. Edge detection using Gabor wavelet on eyes and mouth image for (a) neutral, (b) happy, (c) surprise, (d) sad

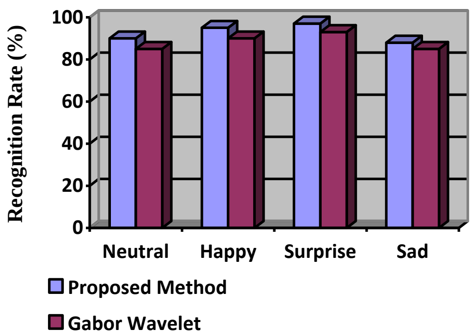


Figure 5. Recognition Rate of Facial Expressions using proposed method and Gabor wavelet of JAFFE Database

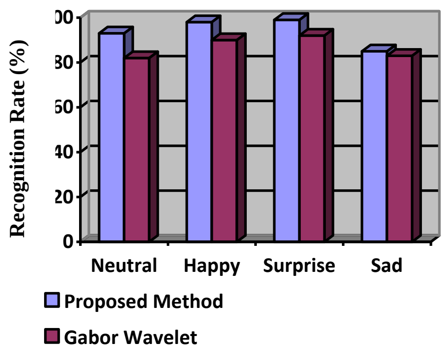


Figure 6. Recognition Rate of Facial Expressions using proposed method and Gabor wavelet of our Database

The Recognition Rate of various Facial Expressions tested on JAFFE and our database is compared with Gabor Wavelet (see Figure 5, 6).

We have experimented on 31 test images of different facial expressions from JAFFE and our database. The proposed system can extract features from the special part of the face in our database rather than JAFFE database. Therefore, we can see that the recognition rate showed a slight increase between our database and JAFFE database. These test images in testing dataset which are compared with 50 train images in training dataset to recognize facial expressions.

Computational burden is an important factor in practical applications, where the amount of required memory and speed of the processor have direct bearing on the final cost. Hence in the second experiment the proposed system compared the average time taken to process one face image by the feature extraction techniques. Results are showed in figure 7.

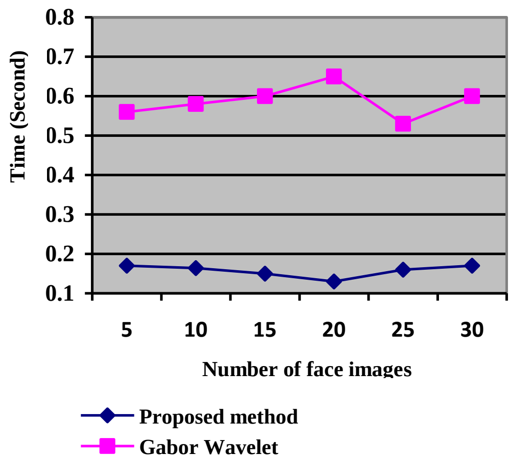


Figure 7. Comparison of Average time taken per face image (Pentium 4, 2.00 GHz)

In the comparison with the conventional edge detection like traditional Canny, the proposed method does not require manual threshold selection. Besides, the extracted edges are not very thick like a Kirsch filter which produces very strong edges and sometimes generates spotty edges on a face surface that look likes noises. Sobel filter also produces thick edges, but some parts of the mouth and nose are not complete detected. This comparison shows that the proposed method is more suitable to use in the facial feature extraction.

7. Conclusions and Future Work

This paper proposes an efficient feature extraction method for human face detection and recognition. A contribution of this paper is the use of adaptive canny edge detection method in order to extract the facial features. We can see that Gabor features are more computationally expensive to calculate taking about 0.17 seconds rather than the proposed method. The proposed system is simpler and easy to implement in the real-time system.

The future plans include further improvement of the robustness of the method and development of the real time facial expression system. These conclusions will be tested on larger data sets using various classification algorithms in the near future.

References

[1] Bartlett, M.S., Littlewort, G., Fasel, I., Movellan, J.R., "Real time face detection and facial expression recognition: development and applications to human computer interaction," International Conference on Computer Vision and Pattern Recognition Workshop, vol. 5, 2003.

[2] Fasel.B and J. Luettin. "Automatic facial expression analysis: A survey," Pattern Recognition, vol. 36, pp. 259-275, 2003.

[3] Jain, A.K., Duin R.P.W., Mao J., "Statistical Pattern Recognition: A Review", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 22, No. 1, pp. 4-37, 2000.

[4] Levine, M. D. and Yingfeng Yu. (2006) "Face recognition subject to variations in facial expression, illumination and pose using correlation filters", Journal of Computer Vision and Image Understanding, Vol 104, pp. 1-15.

[5] Pantic, M. and Rothkrantz, L., (2000) "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.

[6] P. Viola and M. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features", in C*omputer Visionand Pattern Recognition,* vol.1, pp.511-518, 2001.

[7] R. Lienhart and J. Maydt, "An Extended Set of Haar- Like Features for Rapid Object Detection", Proc. of IEEE Int. Conf. on Image Processing, pp.900-903, 2002.

[8] Tian.Y.L, T. Kanade, and J. F. Cohn. "Evaluation of Gabor-Wavelet-Based Facial Action Unit Recognition in Image Sequences of Increasing Complexity," in Proc. of the Fifth IEEE Int. Conf. on AFGR, pp. 229-234, 2002.

[9] Viola, P., and Jones, M., (2004) "Robust real-time object detection", International Journal of Computer Vision, Vol.57, No. 2, pp.137–154.