MVA2011 IAPR Conference on Machine Vision Applications, June 13-15, 2011, Nara, JAPAN

Edge-Based Facial Feature Extraction Using Gabor Wavelet and Convolution Filters

Rosdiyana Samad
Dept. of Intelligent Mechanical System Eng.
Faculty of Engineering, Kagawa University,
2217-20 Hayashi-Cho, Takamatsu, Kagawa,
761-0396, Japan.
s09d501@stmail.eng.kagawa-u.ac.jp

Hideyuki Sawada
Dept. of Intelligent Mechanical System Eng.
Faculty of Engineering, Kagawa University,
2217-20 Hayashi-Cho, Takamatsu, Kagawa,
761-0396, Japan.
sawada@eng.kagawa-u.ac.jp

Abstract

Feature extraction is a crucial step for many systems of face detection and facial expression recognition. In this paper, we present edge-based feature extraction for recognizing six different expressions, which are angry, fear, happy, neutral, sadness and surprise. Edge detection is performed by using Gabor wavelet and convolution filters. In this paper we propose two convolution kernels that are specific for the edge detection of facial components in two orientations. In this study, Principal Component Analysis (PCA) is used to reduce the features dimension. To validate the performance of our proposed feature extraction, the generated features are classified using Support Vector Machine. 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.

1. Introduction

Facial feature extraction is an essential step in the face detection and facial expression recognition frameworks. The extracted features contain meaningful information of the face that describes the face behavior. 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 previous works, 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. 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. Edge detection can be done by using a convolution filter which applies a negative weight on one edge and a positive weight on the other edge. The classical edge detection filters are Laplacian, Sobel, Prewitt, Canny and Kirsch.

Edge detection is a straightforward technique and can be applied directly on a face image. However, some of the edge detectors need a fixed threshold value, which is difficult to set when the face image samples contain various lighting conditions. In general, edge detection has difficulty adapting to different situations. [1] had solved the manual thresholding problem that usually occurs in Canny detection method by using Adaptive Canny edge detection and combined it with AAM algorithm to extract the facial expression features. Some methods are based on edge detection like a Canny operator to extract at facial components such as eyes, eyebrows, lips, and nose.

Other method to detect edges in an image is a wavelet transform. It has been widely used in feature extraction, especially in pattern recognition research area. Haar wavelet and Gabor wavelet have been used actively in face detection. [2] and [3] 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.

Gabor wavelet is favored among many researchers because of its outstanding performance in the task of facial expression analysis [4]. Generally, Gabor filter bank which consists of filters with 5 frequencies and 8 orientations are used in many facial expression recognition systems. However, it has a limitation which is the processing time of Gabor feature extraction is very long and its dimension is prohibitively large [5]. In some previous proposed feature extraction used all 40 filters of Gabor filters while some of them used a few selected frequencies and orientations to reduce the processing time [5, 6, 7, 8].

In this paper, edge-based facial feature extraction is proposed. Our work is focusing on the extraction of specific facial components which are eyes, eyebrows, nose and mouth. Besides, the details or skin texture on face image like wrinkles are also extracted. In this proposed technique, our aim is to generate significant facial features that are able to be used in facial expression recognition system.

In this study, Gabor wavelet and convolution filters are used as edge detectors. We select specific frequency and orientation for Gabor wavelet and specific convolution kernels for convolution filters to extract the facial features. Each edge detector extracts specific part of facial features on face image. The output images of the edge detection are added together, so that the final output image presents a complete face. The number of generated features is large; therefore these features are compressed by using the Principal Component Analysis (PCA). To prove that our proposed feature extraction could generates significant facial features, we classify these features into 6 different facial expression using Support Vector Machine (SVM).

2. Edge-based Facial Feature Extraction

In this experimental study, we use images from 9 subjects that are selected from FEEDTUM [9] and our database. 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 120x120 pixels. Next, eyes detection is applied by using Haar-like features and AdaBoost algorithm then the eyes region is cropped automatically. To extract the facial features of different facial expressions, we apply edge detection on the both face and eyes region images. The explanation on the edge detection methods are explained in the next sub-section.

2.1 Filtering with Gabor Wavelet

The Gabor wavelet is a linear filter which impulse response is defined by a harmonic function multiplied by a Gaussian function. This filter can be used to detect line endings and edge borders over multiple scales and with different orientations. The Gabor wavelet can be defined as Eq.1 and Eq.2:

$$\psi(z) = \frac{p_{u,r}^2}{\sigma^2} \exp\left(-\frac{p_{u,r}^2 z^2}{2\sigma^2} \right) \left[\exp\left(i p_{u,r} z \right) \exp\left(-\frac{\sigma^2}{2} \right) \right] \quad (1)$$

where $z = (x,y)$, u and r define the orientation and scale of the Gabor wavelet, respectively. $p_{u,r}$ is defined as follows:

$$p_{u,r} = p_r e^{i\Phi_r} \quad (2)$$

where $p_r = p_{max}/f^{\,r}$ and $\Phi_u = \pi u / 8$. p_{max} is the maximum frequency, and f is the spacing factor between kernels in the frequency domain.

To detect the edges of the eyes, the parameters of Gabor wavelet are set as follows: $r=1$, $u=4$, $\sigma = 2\pi$, $p_{max} = \pi/2$ and $f = \sqrt{2}$. In this experiment, Gabor wavelet with horizontal orientation is selected because it produces discriminative Gabor features better than other orientation. The parameter r represents the scale of the filter. The greater value of r is set, the size of filter is becomes small. The parameter σ is a standard deviation of the Gaussian function along the x and y-axes. By using this parameter, the width of the filter can be changed and it can increase or decrease the thickness of the edges found in the image. For the eyes part, the eyebrows are included together since both parts give different shapes or patterns for different expressions. The extraction of eyes by employing Gabor wavelet from the face image of subject number 1 is shown in Fig.1. The selected parameters produce the large magnitude and brighter intensity of eyes edges. However, this filtering step does not give proper shapes of eyes and eyebrows but it produces unique patterns for different facial expressions.

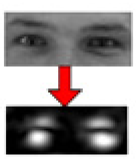


Figure1. Gabor wavelet on eyes image

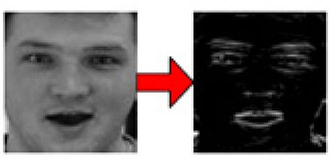


Figure 2. Edge detection on face image

The Gabor wavelet with different set of parameters is applied to the whole face image. The parameters are set as follows: $r=2$, $u=4$, $\sigma = \pi/2$, $p_{max} = \pi/2$ and $f = \sqrt{2}$. We change r and σ values to make the size and width of the filter smaller than the previous. This filter produces thinner edges but the detected edges are much more details. The result of this filtering step is shown in Fig. 2, where the edges of the eyes, eyebrows, nose and lips form a proper shape and each component is isolated. Besides, the edges of skin texture and wrinkles are detected. It gives complete edge detection on a face image. In this study we consider the skin texture and wrinkles as features because the detected edges have form specific patterns when the expression is changed. Edges of the wrinkles are significant at certain areas of the face such as forehead, between two eyebrows, both sides of mouth and sometimes at both sides of eyes. These features can help to differentiate facial expressions better.

The output image from this filtering technique needs some improvements so that the final features are really useful and can be categorized into different expressions easily. Edges that appear on this output image are quite blur and not strong enough. Therefore we propose another edge detection method to enhance the intensity of the edges.

2.2 Filtering with Convolution Filters

In this study, two convolution filters with sharpening kernels are used as edge detector. Generally, the amount of edges enhancement will depend on values of weight in the convolution kernel and the filter size. To emphasize edges, it is necessary for some weights in the kernel to be positive or negative values. The enhancement will depend on the sum of weight values of the kernel. The edge detector only detects the edges that visible in the image and removes the background information by having a sum value of 0. The classical edge detectors such as Laplacian, Sobel, Prewitt and Kirsch, are having similar sum values which is 0. In our case, we need a filter that can detects edges on eyebrows, eyes, nose and mouth without highlights on face details such as skin texture. The purpose of this edge detection is to maximize the edges of the facial component parts that resulted from the Gabor filtering.

To get the desire edges, two convolution kernels are used as shown in Fig. 3(a) and (b). These filters are suitable for the image samples that were taken in indoor environment, where the lighting condition was the same in all the time. The lighting source came from the room light, which made the facial image contain some shadows especially at the eyes and mouth parts. We construct filters which are sensitive to the edges at 90°

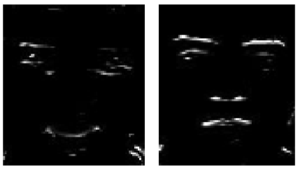
-3	-3	-3
1	0	1
1	1	1

(a)

1	1	1
1	0	1
-3	-3	-3

(b)

Figure 3. Convolution kernel image with (a) 90° (b) 270° orientations.



(a) (b)

Figure 4. Edge detection (a) 90° (b) 270°orientation.

Figure 5. Output image from the combination of 2 orientations of convolution filters.

Figure 6. Output image from the combination of Gabor and convolution filters.

and 270° orientations.

The selection of weight values for each kernel has been made by changing the kernel values that is based on sum values of the weights. We test a face image with the filter that has 0 to negative sum values and check its suitability with the image samples. From the observation, the edges can be detected but the magnitude of the edge is reduced or getting thinner when the sum value is changed from 0 to -4. The filter suppresses the background (details on face surface) and leaving only edges of lips, eyebrows and nose. This result is suitable with our aim, to extract only at facial component parts. In this case, the sum value for both filters that suitable with our image samples is -4. Fig. 4(a) and (b) show output images from the 90°and 270°orientation filters. Both output images are added together to get the complete shape of facial components as shown in Fig. 5. Finally all pixels of this image are added with the pixels of the output image from the Gabor filtering as shown in Fig. 6.

3. Dimension Reduction

At this stage, the sizes of face image and eyes 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. The Gaussian pyramid down sampling and image resizes operation are performed to produce the final size of the face image which is 30x30 while eyes image is 42x13 pixels. The number of generated features from this operation is 900 for face and 546 for eyes.

However, the number of features is still large, and again, the dimension of features is reduced by using PCA [10]. The number of features dimension is reduced to 60, 50, 40, 30, 20 and 10. 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.

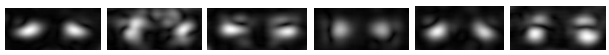
4. Classification of facial expression features

To evaluate the significance of the produced facial features, these features are classified by using SVM classifier into fear, angry, happy, neutral, surprise and sad classes. 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, Radial Basis Function (RBF) is used as kernel function while cross validation technique via grid search is used to find optimal parameter for the kernel function.

5. Results and Discussions

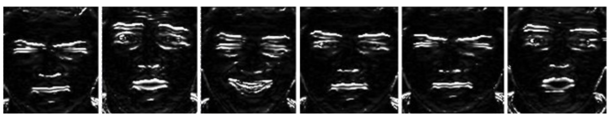
The feature extraction method is tested on face images of 9 subjects that contain angry, fear, happy, neutral, sad and surprise expressions. Here, the results from the edge detection method are discussed. Edge detection on eyes image using Gabor wavelet shows that the eyes for each facial expression has different patterns as shown in Fig.7. The resultant edges do not form a perfect shape of eyes. The edges of eyebrows are not appearing apparently, and it looks like disappear or the edges are merged with eyes edges. However, the edges have form unique shapes that can be differentiated from one expression to another expressions.

The results for the edge detection on the whole face taken from the samples of one subject number 1 are shown in Fig. 8. The edges of eyes, mouth, nose and eyebrows are thick and noticeable. Besides, small edges appearing around the face show the skin texture and wrinkles. The intensity of the wrinkles edges is less bright than other facial components, however it is clearly visible. The details of the wrinkles edges are shown in Fig. 9, where the edges usually appear around mouth, under the eyes and forehead. The resultant edges are obtained not only by the wrinkles but also by the shadows. It is usually happen at eyes as shown in Figure 9(b). We also plot the compressed features data from one of the subjects in our samples as shown in Fig. 10. It shows the features data are well separated and grouped into each type of facial expression.



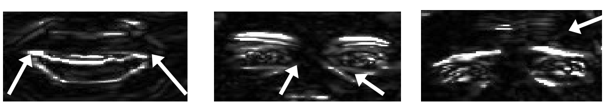
(a) (b) (c) (d) (e) (f)

Figure. 7 Edge detection using Gabor wavelet on eyes image for (a) angry, (b) fear, (c) happy, (d) neutral, (e) sad, (f) surprise.



(a) (b) (c) (d) (e) (f)

Figure.8 Edge detections from the combination of Gabor wavelet and convolution filters with 2 orientations for (a) angry, (b) fear, (c) happy, (d) neutral, (e) sad, (f) surprise expression.



(a) (b) (c)

Figure.9 Edges of wrinkles on (a) mouth, (b) eyes, (c) forehead.

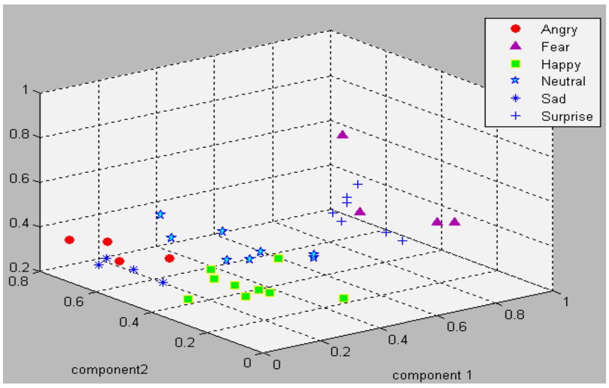


Figure.10. 3D Plots of facial expression data for Subject 1

Table 1. Facial expression recognition rate for 9 subjects

Subject	No. of feature vectors	Recognition rate (%)
S1	40	91.7
S2	50	84.6
S3	40	88.6
S4	20	88.2
S5	40	79.6
S6	40	82.1
S7	50	89.7
S8	40	80.5
S9	50	86.2

Table 2. Subject-dependent confusion matrix of facial expression recognition rate (%)

		Actual class					
		Ang	Fear	Hap	Neut	Sad	Surp
Predicted class	Ang	**86.8**	2.4	1.2	1.2	8.4	0.0
	Fear	6.6	**80.2**	7.6	0.9	0.0	4.7
	Hap	0.0	1.9	**97.2**	0.9	0.0	0.0
	Neut	1.1	3.4	8.0	**80.7**	1.1	5.7
	Sad	5.2	4.2	1.1	6.3	**83.2**	0.0
	Surp	2.0	12.7	2.0	0.0	0.0	**83.3**

We also performed the subject-dependent classification on the extracted features that have been compressed into 10, 20, 30, 40, 50 and 60 features. The recognition rates for subjects 1 to 9 are shown in Table 1. The result shows the number of feature vectors that were used to obtain the highest recognition rate for each subject. From the observation, 5 subjects used 40 features as input for the classification. The rest 3 subjects used 50 features and 1 subject used 20 features. It shows that the compressed features able to be classified even though the features dimension is small. Table 2 shows the classification results on each expression for all subjects. These results have proved that our edge detection method is successful to extract the facial features and these features can be classified into 6 expressions.

In the comparison with the conventional edge detection like Canny, our approach 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 our method is more suitable to use in the facial feature extraction.

6. Conclusion

Edge detection on eyes and face image using Gabor wavelet and convolution filters with 2 orientations is presented in this paper. The results show, the use of multiple edge detectors work slightly better compared to [1] and do not require manual thresholding. In this study, we found out that the extracted features from the proposed feature extraction method are able to be classified, since these features still contain significant information even though it have been compressed. It shows that these features can be used to represent each facial expression. Besides, our proposed approach 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.

References

[1] P. Z. Yi, Z. Y. Hui & Z. Yu, "Real-time Facial Expression Recognition Based on Adaptive Canny Operator Edge Detection", Proc. of 2nd IEEE Int. Conf. on MultiMedia and Information Technology, pp. 154-157,2010.

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

[3] 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.

[4] Y. B. Lee, S. B. Moon and Y.G. Kim, "Face and Facial Expression Recognition with an Embedded System for Human-Robot Interaction", in *Lect Notes Comput Sci*, vol. 3784, pp. 271-278, 2005.

[5] L.V. Praseeda and M. Sasikumar, "Facial Expression Recognition from Global and a Combination of Local Features", in *IETE Tech Rev*, vol 26, pp. 41-46, 2009.

[6] H.B.Denget. al,"A New Facial Expression Recognition Method Based on Local Gabor Filter Bank and PCA Plus LDA", in *Int. J. Inf. Technol*. vol.11 (11), pp. 86-96, 2005.

[7]M.P. Loh, Y.P. Wong and C.O. Wong, "Facial Expression Recognition for E-learning System Using Gabor Wavelet & Neural Network", Proc. of 6th IEEE Int. Conf. on Advance Learning Technology, pp. 523-525, 2006.

[8] D. Vukadinovic and M. Pantic, "Fully Automatic Facial Feature Point Detection Using Gabor Feature Based Boosted Classifier", Proc. of IEEE Int. Conference on Systems, Man and Cybernetics, vol. 2, pp. 1692-1698, 2005.

[9] F. Wallhoff, Facial Expressions & EmotioDatabase, Technische Universitat Munchen, web site: http://www.mmk.ei.tum.de/~waf/fgnet/feedtum.html

[10] L. I. Smith, "A Tutorial on Principle Component Analysis", Cornell University USA, pp.2-22, 2002.