

# FACIAL EXPRESSION RECOGNITION USING DIGITALISED FACIAL FEATURES BASED ON ACTIVE SHAPE MODEL

Nan Sun<sup>1</sup>, Zheng Chen<sup>2</sup> and Richard Day<sup>3</sup>

Institute for Arts, Science & Technology  
Glyndwr University  
Wrexham, United Kingdom  
bruce.n.sun@gmail.com<sup>1</sup>  
z.chen@glyndwr.ac.uk<sup>2</sup>  
r.day@glyndwr.ac.uk<sup>3</sup>

## **ABSTRACT**

*Facial Expression Recognition is a hot topic in recent years. As artificial intelligent technology is growing rapidly, to communicate with machines, facial expression recognition is essential. The recent feature extraction methods for facial expression recognition are similar to face recognition, and those caused heavy load for calculation. In this paper, Digitalized Facial Features based on Active Shape Model method is used to reduce the computational complexity and extract the most useful information from the facial image. The result shows by using this method the computational complexity is dramatically reduced, and very good performance was obtained compared with other extraction methods.*

## **KEYWORDS**

*Facial Expression Recognition, Active Shape Model, Feature Digitalisation, Computational Complexity Reduction*

## **1. INTRODUCTION**

Computers are behaving more and more likely as human. They can talk and play with a human, but a deeper communication needs interactions of emotions. Thus Facial Expression Recognition (FER) function is important and highly required for future computers.

Since Paul Ekman introduced the 6 basic expressions [1], many FER methods have been developed. Local Binary Pattern (LBP) for texture analysis was introduced by T. Ojala [2, 3]. Then, T. Ahonen presented LBP-based methods for face detection and recognition [4, 5]. X. Tan solved the problem with difficult lighting conditions using LBP [6]. Some researchers believe that using Gabor wavelets for FER is a better way. M. Bartlett did his work utilizing the Gabor filter to recognize the facial expression [7]. The temporal extension of Gabor features in facial expression analysis work was done by L. Ma [8]. Both LBP and Gabor filters deal with the image pixels directly, the extracted features are therefore still related to pixel information. That means the facial images need to be aligned or treated by blocks. The disadvantage of using the above methods is that the complexity of calculation during training and classifying relies on the resolution of the image.

Active Shape Model (ASM) was developed by T. Cootes in 1992 [9], which detects the profile of the image. Unlike LBP and Gabor filters, the result of ASM for facial feature extraction is not related to pixels but the positions of facial landmarks. Many variations of the ASM method for FER have been introduced: Optimal Features ASM (OFASM) is high in accuracy but is more computationally expensive [10]. F. Sukno extended OFASM to allow application in more complex geometries [11]. However, the above methods do not consider the wrinkle features, which is quite important in real-world FER.

In this paper, we suggest a new method called Digitalized Facial Features based on Active Shape Model (DFFA) to extract the facial features. ASM is only used to get the landmarks and a part of the useful features. Wrinkle features and other useful facial features are extracted by edge detection and pixel analysis. Finally, all the features are digitalized to simple variables to represent their strength. Artificial Neural Network is used for the final facial expression classification.

The paper is organized as follows. Section 3 introduces basic concepts of ASM. Detailed digitalized facial feature extraction is described in Section 4. In Section 5 the comparisons and experiment results are discussed, and the conclusions are given in Section 6.

## **2. RELATED WORKS**

Z. Yu suggested to use ASM as facial landmarks locating method, with RS-SVM for selection and classification [12]. RS-SVM reduced the computational complexity. However, its attribute reduction stage was not specifically designed for facial features. Detailed local facial feature extraction methods are not discussed.

R. Shibib developed a whole set of methods for facial expression recognition [13]. The facial features were extracted directly from ASM, and were not processed. An example in Chapter 3 shows only using ASM may cause confusions.

C. Hsieh and M. Jiang focused on local facial features, some regions of interest (ROI) were discussed [14]. However, the final extracted data are still in pixel format which increases the computing load of the following classifier.

All the above research agree that ASM is essential for FER. More detailed locations of ROI and more simplified extracted data are needed for FER.

## **3. BASIC CONCEPT ABOUT ACTIVE SHAPE MODEL**

Active Shape Model (ASM) detects the profile of an object, it can be used to detect the shape of face and the parts of the face. To perform ASM to the face, a set of training samples of faces need to be labelled with landmark positions. Then apply Principal Component Analysis (PCA) to the data to find the eigenvalue and eigenvectors. Using different eigenvectors to form the new shapes which are limited by the eigenvalues.

The main idea of using ASM here is to locate the facial landmarks, so the locations of the other useful features can be found. With edge detection and pixel analysis, those useful features are accurately and easily extracted. Features such as the shapes of the nose and eyes will not be extracted because they are less related to facial expression, and this makes sure only the useful features are extracted.

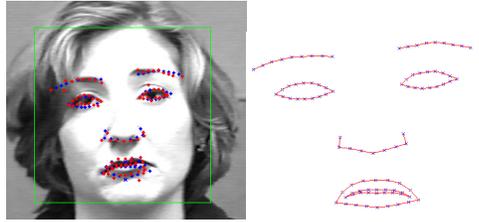


Figure 1. ASM Landmarks Locating

#### 4. DIGITALISED FACIAL FEATURE EXTRACTION

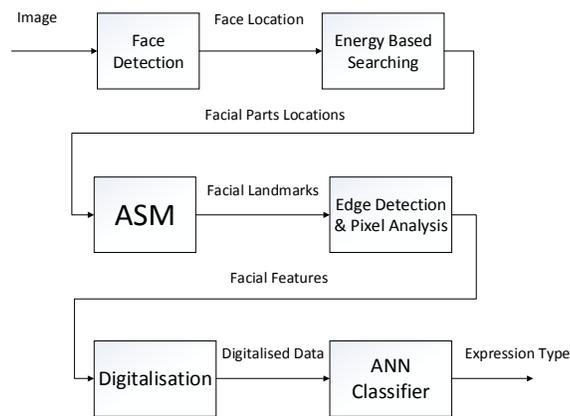


Figure 2. DFFA Flow Chart

The flow chart shows the whole progress of facial expression recognition using digitalized facial feature extraction. Face detection will be performed when an image is coming in, the rough face location will be found. Energy based searching will find more detailed facial parts locations. Then ASM will be applied to the facial image to search for the facial landmark. Edge detection and pixel analysis will extract the wrinkle features and eyeball feature. Then the useful features are digitalized. Passing through ANN classifier, the facial expression is recognized.

##### 4.1. Energy Based Searching

Pre-processing of the facial image is very necessary for landmarks locating. In Face Detection stage, Haar-like features are used for face detection, however, only knowing the rough location of the face is not enough for ASM searching, more details need to be given to make sure the searching will not be trapped into local optima. So Energy Based Searching is used here to find out the rough centers of the key facial parts.

The facial image will first pass to a Gaussian filter to make the image smooth. Then accumulate the pixels horizontally and vertically. At last, calculate the first derivative and second derivative to find the centers. Figure 3 shows the original facial image and the plot of first derivatives in horizontal and vertical for Energy Based Searching.

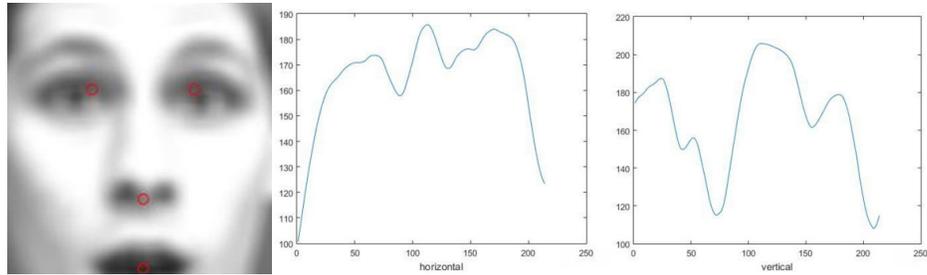


Figure 3. Energy Based Searching for Facial Parts Centers

By knowing the rough locations of the centers of eyes, nose and mouth, ASM searching is further limited. More accurate ASM results are assured.

#### 4.2. ASM landmarks locating

When the face is detected, and the centers of facial parts are located, ASM is applied to the facial image. By using ASM, the profile of the facial parts are found. However, the ASM result cannot be used directly for expression recognition, here is an example.

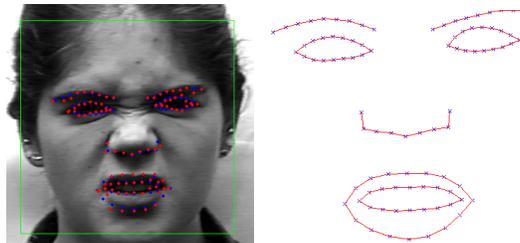


Figure 4. Confusing Expression with Only ASM Shapes without Wrinkles

Figure 4 shows a disgust expression, but if without wrinkles and using the ASM result only, it is difficult to tell the exact expression. So only using ASM is not capable of recognizing an expression. That means apart from ASM result, there is some other important information which are not extracted.

#### 4.3. Edge Detection and Pixel Analysis

From Figure 4, it can be seen not the whole face is sending expression message. Dealing with parts which send no information is a waste. By using ASM for locating, the useful locations can be found. They are the locations of the forehead, between eyes and beside the nose. Wrinkles in those locations are crucial for FER. Thanks to ASM landmarks, these locations are already found. Edge Detection with Sobel Operator can easily detect the wrinkles.

Another useful information for expression is the eyeball black/white ratio and the coverage of the upper and lower eyeball. ASM had already located the eye area, so the work here is to calculate the pixel inside the eye for analysis.

#### 4.4. Digitalization of the Features

Wrinkle features for different people vary, but sometimes different wrinkles are giving the same expression. Simply using pixels or LBP or Gabor filters for feature extraction will have very different results. If these wrinkle features can be just described in a variable with a scalar to show the strength of the wrinkle, the problem is solved. It is the further digitalization of the features. After the landmarks are found by ASM, five areas are the regions of interest (ROI). The forehead for wrinkle features (Area1), between eyes for wrinkle features (Area2), the eye feature (Area3), the mouth feature (Area4) and beside nose wrinkle features (Area5).

Features in the above 5 areas will be extracted and digitalized. In different areas, the features are selected differently. As the image may be stretched, the distance between centers of eyes and the distance between the end of nose to the point in the middle of eyes are used as a reference for horizontal and vertical.

To digitalize the wrinkles, the area needs to be defined. The digitalized wrinkle  $DW$  is expressed as

$$DW = \frac{\sum P_{an}}{N_P} \quad (1)$$

Where  $P_{an}$  is the magnitude of pixels which are above the threshold of Edge Detection,  $N_P$  is the total number of pixels in the area.

To digitalize the eye feature, a global average grey scale factor  $G_R$  for the face area need to be calculated. Eyeball black/white ratio  $De_r$  is expressed as

$$De_r = \frac{N_E - N_{Pw}}{N_{Pw}} \quad (2)$$

Where  $N_E$  is the total number of pixels in the eyes area,  $N_{Pw}$  is the number of pixels which are lighter than the average factor  $G_R$ .

The eyeball upper coverage  $De_u$  feature is expressed as

$$De_u = \frac{N_{Pub}}{N_{Eh}} \quad (3)$$

Where  $N_{Pub}$  is the number of pixels of the upper eyeball which are black.  $N_{Eh}$  is the number of pixels of half eyeball.

The eyeball lower coverage  $De_l$  feature is expressed as

$$De_l = \frac{N_{Plb}}{N_{Eh}} \quad (4)$$

Where  $N_{Plb}$  is the number of pixels of upper eyeball which are black.

Considering all the faces are highly symmetrical, only the right eye feature is used for digitalization.

To digitalize the mouth features  $D_m$ , four curvatures are used, the value of curvatures are from ASM result.

Finally, the facial features are expressed by

$$F = [Dw_f, Dw_{be}, Dw_{bn}, De_r, De_u, De_l, Dm_1, Dm_2, Dm_3, Dm_4] \quad (5)$$

## 5. EXPERIMENTS AND RESULTS

### 5.1. Experiment Conditions and Training

The facial expression database used for training is Cohn-Kanade (CK) database [15, 16]. Compared with other facial expression databases like Japanese Female Facial Expression (JAFPE) database which only has 213 faces from 7 women, CK database has better diversity and more images.



Figure 5: Example of Cohn-Kanade database

150 of images are chosen randomly from the CK database with six basic expressions and neutral expression equally, 101 of them are used for training, and 49 of them are for verification.

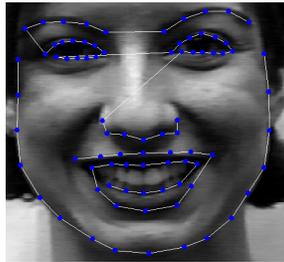


Figure 6. 82 points ASM landmarks

### 5.2. Comparison with Other Feature Extraction Methods

For DFFA, five key areas with ten parameters are extracted, the ASM searching uses 82 key landmarks. For Gabor filters, eight directions and five scales are used.

The first comparison is about the computational complexity of the extracted features for the following classifier between different feature extraction methods and different resolutions.

Table 1. Parameters after Extraction

Resolution	DFFA	Gabor Filters (8d5s)	LBP
64x64	10	1638400	4096
128x128	10	6553600	16384

Table 1 shows, after extraction, the computational complexity of DFFA is much lower, such scale of Gabor wavelet features are almost unable to be processed in real-time.

The next comparison is the recognition rate between the different feature extraction methods, the classifier is ANN.

Table 2. Feature Extraction Methods Comparison

Type of Method	Recognition Rate
Gabor Features	89.87%
LBP	71.4%
DFFA	85.7%

The result shows DFFA has acceptable recognition rate.

## 6. CONCLUSION

DFFA shows high performance in FER and is suitable for real-time FER. It reduces the computational complexity while keeping acceptable recognition rate. Thanks to the digitalization of facial features and edge detections, the huge facial feature data were extracted and only the most useful feature information is converted into small scaled data set. Future works will focus on a high-efficiency classifier design for the digitalized data.

## REFERENCES

- [1] P. Ekman, W. Friesen, Facial action coding system, Palo Alto CA, USA, Consulting Psychologist Press, 1978.
- [2] T. Ojala, M. Pietikinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distribution," *Pattern Recognition*, vol. 29, no. 1, 1996.
- [3] T. Ojala, M. Pietikinen, and T. Menp, "Multiresolution grayscale and rotation invariant texture classification with local binary patterns," *IEEE PAMI*, vol. 24, no. 7, July 2002.
- [4] T. Ahonen, A. Hadid, and M. Pietikinen, "Face recognition with local binary patterns," *ECCV*, 2004, pp. 469–481.
- [5] A. Hadid, M. Pietikinen, and T. Ahonen, "A discriminative feature space for detecting and recognizing faces," *IEEE CVPR*, June 2004, pp. 797–804.

- [6] X. Tan, B. Triggs, "Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions", *IEEE Trans. Image Processing*, Vol. 19, No. 6, pp. 1635-1650, June 2010.
- [7] M. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J. Movellan, "Recognizing facial expression: machine learning and application to spontaneous behavior," *Computer Vision and Pattern Recognition*, pp. 568-573, vol. 2, 2005.
- [8] L. Ma, D. Chelberg, and M. Celenk. "Spatio-temporal modeling of facial expressions using Gabor-wavelets and hierarchical hidden Markov models." *IEEE International Conference on Image Processing 2005*. Vol. 2, 2005.
- [9] T. F. Cootes and C. J. Taylor. "Active shape models". 3rd British Machine Vision Conference 1992, pages 266-275, 1992.
- [10] B. van Ginneken, A.F. Frangi, J.J. Staal, B.M. ter Har Romeny, and M.A. Viergever, "Active shape model segmentation with optimal features," *IEEE Trans. Medical Imaging*, vol. 21, no. 8, pp. 924-933, 2002.
- [11] F.M. Sunko, S. Ordaas, C. Butakoff, S. Cruz, and A.F. Frangi, "Active shape models with invariant optimal features: Application to Facial Analysis," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 29, no. 7, pp. 1105-1117, 2007.
- [12] Yu Zheng-Hong, Li Cong. "Research of Facial Expression Recognition Based on ASM Model and RS-SVM". *ISDEA*, pp. 772-777, 2014.
- [13] Shbib, Reda, and Shikun Zhou. "Facial expression analysis using active shape model." *Int. J. Signal Process. Image Process. Pattern Recognition*, vol. 8, no. 1, pp-9-22, 2015.
- [14] Hsieh, Chen-Chiung, and Meng-Kai Jiang. "A facial expression classification system based on active shape model and support vector machine." *Computer Science and Society (ISCCS), 2011 International Symposium on*. IEEE, pp. 311-314, 2011.
- [15] Kanade, T., Cohn, J. F., & Tian, Y. (2000). Comprehensive database for facial expression analysis. *Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG'00)*, Grenoble, France, 46-53.
- [16] Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010). The Extended Cohn-Kanade Dataset (CK+): A complete expression dataset for action unit and emotion-specified expression. *Proceedings of the Third International Workshop on CVPR for Human Communicative Behavior Analysis (CVPR4HB 2010)*, San Francisco, USA, 94-101.