Facial Expression Recognition Using AAMICPF

Jun-Sung Lee, Chi-Min Oh, and Chil-Woo Lee

School of Electronics and Computer Engineering, Chonnam National University, 500-757 Gwang-Ju, Korea
{aliesim,sapeyes}@image.chonnam.ac.kr, leecw@chonnam.ac.kr

Abstract. Recently, many interests have been focused on the facial expression recognition research because of its importance in many applications area. In the computer vision area, the object recognition and the state recognition are very important and critical. Variety of researches have been done and proposed but those are very difficult to solve. We propose, in this paper, to use Active Appearance Model (AAM) with Particle filter for facial expression recognition system. AAM is very sensitive about initial shape. So we improve accuracy using Particle filter which is defined by the initial state to particles. Our system recognizes the facial expression using each criteria expression vector. We find better result than using basic AAM and 10% improvement has been made with AAA-IC.

Keywords: Facial expression recognition, Active Appearance Model, Particle Filter.

1 Introduction

The face which is the best part on the human body represents inner psychological state except the language. Facial expression recognition is applied in many areas such as human computer interaction, multimedia information retrieval, medicine, and commercial product like digital camera and smart phone. So many research interests have been grown up in the facial expression analysis.

The processes for facial expression recognition are divided in two steps. The first, it extracts the features from the face expression. The main feature's areas are eyes, brows, nose, and mouth. The second, we design the recognizer for facial expression. Various techniques [4]-[8] have been proposed to detect and track facial features in face images. In general, two types of information are commonly utilized by these techniques. One is the image appearance of the facial features, which is referred as texture information; and the other is the spatial relationship among different facial features, which is referred as shape information.

In [7], a neural network is trained feature detector for each facial feature. Facial features will be located by searching the face image via the trained facial feature detectors. Similarly, Gabor wavelet networks are trained to locate the facial features in [3]. Since the shape information of the facial features is not modeled in both

J.A. Jacko (Ed.): Human-Computer Interaction, Part II, HCII 2011, LNCS 6762, pp. 268–274, 2011.

techniques, they are prone to image noise. Therefore, in [8], a statistical shape model is built to capture the spatial relationships among facial features, and multi-scale and multi-orientation Gaussian derivative filters are employed to model the texture of each the texture of facial feature. However, only the shape information is used when comparing two possible feature point configurations, ignoring the local measurement for each facial feature. Such a method may not be robust in the presence of image noise. Cootes [5] proposed Active Appearance Model that uses the shape and texture model. Shape and texture model are applied by Principle Component Analysis (PCA) from hand-marked sample data. Then the feature can be extracted and tracked by regression model on facial image. This approach robust in terms of the noise and it has high accuracy but the initial state is very important for result. If the initial state is wrong, the result is almost failed.

To solve this problem, we proposed to use AAM with Particle Filter (PF) for finding good initial state. We define the initial shape to particle and generate the particle distribution around the center of the face. Then it performs AAM fitting on each particle and selects the result which has a minimal error.

After extracting the facial features, there are two approaches to recognize the facial expression. Those are static and temporal approaches. The static classifiers such as the neural networks, the support vector machines and the linear discriminant analysis attempt to recognize the facial expression using one image. The temporal classifiers such as the Hidden Markov Model and the recurrent neural networks attempt the facial expression recognition using a sequence of images. In this paper, we use the criteria expression model using a sequence of images. We choose the 3 facial expressions (happy, angry, and neutral).

2 Extract the Facial Features

In this section, we introduce the AAM and AAM with Particle Filter. Then, we select the feature for facial expression recognition.

2.1 Active Appearance Model (AAM)

In recent year there has been a growing interest in face modeling. Active Appearance Model [1],[4] are generative, parametric models that of a certain visual phenomenon that show both shape and appearance variations. These variations are represented by a linear model such as PCA. The face model can be made up from the training data using AAM. And the face features extracting is achieved by fitting the trained model to a input sequence.

AAM is represented by a triangulated mesh with one vertex. The shape vector scan expression as $s=(x_1,y_1,\ldots,x_l,y_n)^T$ and shape variation is expressed by a linear combination of a mean shape s_0 and m shape basis vector s_i as

$$s = s_0 + \sum_{i=1}^{m} p_i s_i \tag{1}$$

Where p_i denote the i-th shape parameter, and $p=\{p_1,p_2,\ldots,p_m\}$ is the shape parameter vector of AAM for input face image. The mean shape s_0 and m shape basis vectors s_i

are normally obtained by applying PCA to the training data, where the vertices of image are marked by hand. The i-th shape basis vector s_i is the i-th eigenvector that corresponds to the i-th largest Eigen value.

The training images are warped to the mean shape s_0 using the piece-wise affine warp that is defined between the corresponding triangles in the landmarked shape of the training images and the mean shape. Then, we can define the appearance as a shape normalized image A(x) over the pixels x that belong to the inside of the s_0. The appearance variation is expressed by a linear combination of a mean appearance $A_0(x)$ and n appearance basis vectors $A_i(x)$ as

$$A(x) = A_0(x) + \sum_{i=1}^{n} \alpha_i A_i(x), \tag{2}$$

where α denote the i-th appearance parameter, and $\alpha=\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ is the appearance parameter vector of AAM for input face image. As with the shape model, the mean appearance $A_0(x)$ and n appearance basis vectors $A_i(x)$ are normally obtained by applying PCA to the training data. The i-th appearance basis vector (image) $A_i(x)$ is the i-th eigenvector that corresponds to the i-th largest Eigen value.

The Goal of AAM fitting is then to minimize the difference between the warped image and the appearance image. Thus, fitting AAM to a target image can be formulated as finding the model parameters of an AAM that minimize the following error as

$$E = \sum_{x \in s_0} [A(x) = A_0(x) + \sum_{i=1}^{n} \alpha_i A_i(x) - I(W(x;p,q)]^2, \tag{3}$$

where q is the global 2D pose parameter vector including the scale, rotation, and horizontal/vertical translation. A number of gradient-based fitting algorithms have been proposed, in which the Lucas-Kanade image matching algorithm [9], [10] is extended. In this paper, we selected the method that Ian Matthews and Simon Baker proposed Inverse compositional method.

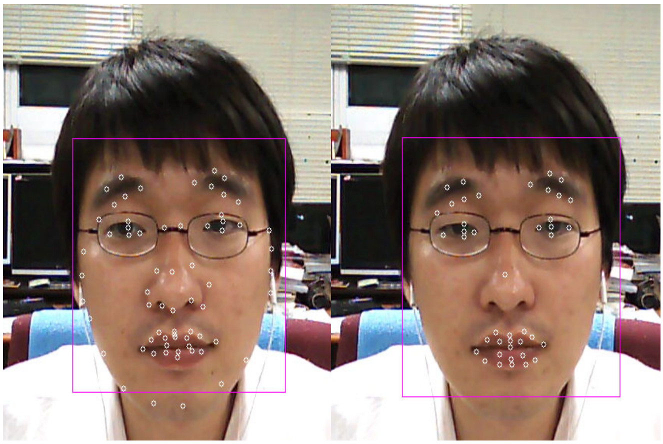
2.2 AAM Based on Inverse Compositional Using Particle Filter (AAMICPF)

The fitting step of tradition AAMs used the single sample area. So if this area is wrong, all results are almost failed. To solve this problem, we use the Particle Filter which can be constructed multi-sample area.

Particle filter is a sequential Monte Carlo methodology where the basic idea is the recursive computation of relevant probability distributions using the concepts of importance sampling and approximation of probability distributions with discrete random measures. The fundamental idea of Particle filter approximates the filtered posterior (next) distribution (density) by a set of random particles (sampling) with associated weights. It weights particles based on a likelihood score and then propagates these particles according to a motion model. Particle filtering assumes a Markov Model for system state estimation. Markov model states that past and future states are conditionally independent of a given current state. Thus, observations are dependent only on current state. In this paper, we define the initial shape to particle and distribute particles according to particle parameters (numbers, range, scale and rotation of particles). We perform AAM fitting algorithm in each particle location and the final result accomplished with the particle which has a minimal error.

2.3 Selecting Features for Facial Expression Recognition

We use the 68 vertices for constructing the shape model (Brows 12, eyes 10, nose 12, mouth 19, jaw 15). When the facial expression is changed, the several features change. Those parts are brows, eyes and mouth. So we just use the 42 vertex except nose and jaw vertex.



(1) (2)

Fig. 1. Result of extracting features (1) AAMICPF (2) feature for facial expression

3 Facial Expression Recognition

In this paper we recognize three facial expressions (happy, angry, and neutral). And we use the criteria vector for each facial expression. This method divides the two steps: training and recognition step. In training step, we extract the features using proposed AAMICPF and select the feature for facial expression. So we call this feature set F and denote the equation 4. We calculate the center of the feature and alignment and normalize using the equation 5, 6 and 7.

$$f = \{f_1, f_2, ..., f_i\} (1 \le i \le 42) \tag{4}$$

$$f_M = \frac{\sum_{1}^{42} f_i}{42} \tag{5}$$

$$F^{'} = F - f_M \tag{6}$$

$$\overline{F} = \frac{F^{'}}{|F|} \tag{7}$$

We calculate normalized distance histogram using iterating this step. Then, we normalize histogram of each expression using equation 8. This will be criteria vector.

$$\overline{F_M} = \frac{1}{N}\sum_{1}^{N}\overline{F_i} \tag{8}$$

Similarity, in recognition step, we extract the features using AAMICPF from input image and we calculate F_I using equation 4, 5, 6, 7 and 8. Then we calculate MSE using equation 9 between each expression criteria vector F_M and F_I.

$$MSE(F_I) = \sqrt{(\overline{F_I} - \overline{F_M})^2} \tag{9}$$

The final result is the minimum MSE value for facial expression recognition. Figure 2 shows the main working block diagram of facial expression recognition with AAMICPF in simplified form.

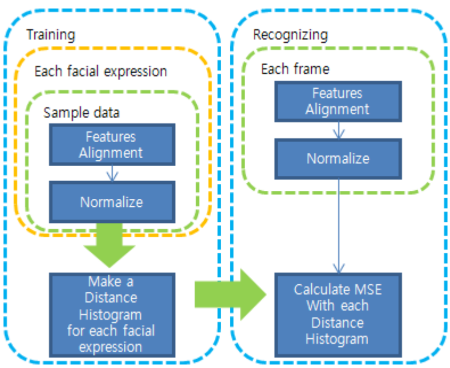


Fig. 2. Block diagram of facial expression recognition with AAMICPF algorithm

4 Experiments and Results

The proposed facial expression recognition technique is implemented using C++ on a PC with a i3 2.93GHz CPU and a 2GB RAM. The resolution of the captured images is 640x480 pixels, and the built system runs at approximately 25 fps.

We collected the database which included 10 Korean peoples (8 male and 2 female whose ages are between 24 and 30, and 5 peoples of them wear the glasses) and had three expressions (happy, angry and neutral) for testing and training. It consists of 150 frames for each expression and total 4500 frames. We use 50% of the database for training, the others is for testing. The Table 1 show that compare performance with traditional AAM, AAMIC and AAMICPF using facial expression recognition.

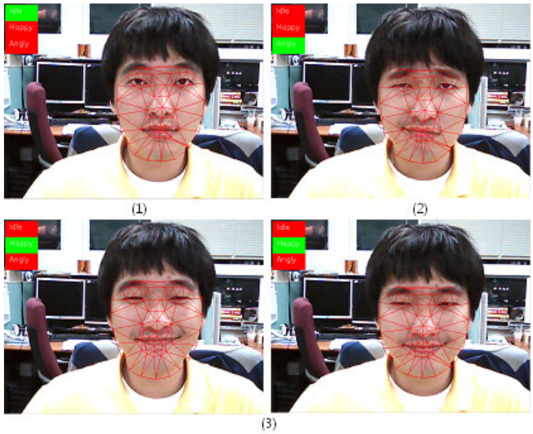


Fig. 3. Result of facial recognition (1) neutral (2) angry (3) happy

Table 1. Performance assessment with other algorithm (%)

	Neutral	Happy	Angry	Average
AAMICPF	**82.12**	**92.40**	**93.65**	**89.39**
AAMIC	70.38	79.60	80.25	76.74
Trad. AAM	68.23	75.00	76.43	73.22

5 Conclusion

In this paper, we propose to solve initial shape location problem using particle filter. Our System clearly improves the accuracy, and performs the facial expression recognition with proposed system and compares the result with traditional AAM and AAMIC.

In future work, we will improve more about our algorithm for extracting good features. Besides, our future research will concern more about the improvement of facial expression recognition and considering application of extracting facial features.

Acknowledgments. This research is supported by Ministry of Culture, Sports and Tourism (MCST) and Korea Culture Content Agency(KOCCA) in the Culture Technology(CT) Research & Development Program 2010.

References

1. Matthews, I., Baker, S.: Active appearance models revisited. International Journal of Computer Vision 60(2), 135–164 (2004)
2. Cohen, I., Sebe, N., Chen, L., Garg, A., Huang, T.: Facial expression recognition from video sequences: temporal and static modeling. Computer Vision and Image Understanding 91(1), 160–187 (2003)

3. Toyama, K., Feris, R.S., Gemmell, J., Kruger, V.: Hierarchial wavelet networks for facial feature localization. In: International Conference on AFGR (2002)
4. Edwards, G., Taylor, C., Cootes, T.: Active appearance models. IEEE Transaction on Pattern Analysis and Machine Intelligence 23(5), 681–685 (2001)
5. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. In: Burkhardt, H., Neumann, B. (eds.) ECCV 1998. LNCS, vol. 1407, pp. 484–498. Springer, Heidelberg (1998)
6. Wiskott, L., Fellous, J.M., Kruger, N., Malsburg, C.V.: Face recognition by elastic graph matching. IEEE Transactions on PAMI 19(7) (1997)
7. Reinders, M.J., Koch, R.W., Gerbrands, J.: Locating facial features in image sequences using neural networks. In: International Conference on AFGR (1996)
8. Burl, M., Leung, T., Perona, P.: Face localization via shape statistics. In: International Conference on AFGR (1995)
9. Baker, S., Matthews, I.: Lucas-Kanade 20 years on: a unifying framework: Part I. CMU Technical Report CMU-RI-TR-02-16 (2002)
10. Gross, R., Matthews, I., Baker, S.: Lucas-Kanade 20 years on: a unifying framework: Part 3. CMU Technical Report CMU-RI-TR-03-05 (2003)