Facial Expression Mimicking System

Ryuichi Fukui Kouichi Katsurada Yurie Iribe Tsuneo Nitta
Toyohashi University of Technology
{r_fukui, iribe}@vox.tutkie.tut.ac.jp, {katurada, nitta}@tutkie.tut.ac.jp

Abstract

We propose a facial expression mimicking system that copies the facial expression of one person on the image of another. The system uses the active appearance model (AAM), a commonly used model in the field of facial expression processing. AAM compositionally comprises some parameters representing facial shape, brightness, and illumination environment. Therefore, in addition to the facial expression elements, the model parameters express other elements, such as individuality and direction of the face. In order to extract the facial expression elements from compositional parameters of AAM, we applied principal component analysis (PCA) to the AAM parameter values, collected with each change in facial expression. The obtained facial expression model is applied to the facial expression mimicking system and the experiment shows its effectiveness for mimicking.

1. Introduction

Numerous softwares that enable to synthesize facial expressions on computers were developed for video games, VR systems, and movies [1][2][3]. These softwares generally put a captured facial expression on a wireframe facial model and synthesize a facial image with computer graphics (CG) technologies. In this paper, we propose an appearance-based facial expression mimicking system that duplicates the facial expression of one person on the facial image of another without using wireframe models and CG.

Facial mimicking consists of two steps: recognizing a facial expression and then synthesizing a new facial image with that expression. In most of previous studies, some landmarks, such as eyes, brows, and mouth, are detected and a preliminary wireframe facial model is modified according to the location of the landmarks [4][5]. However, this type of system cannot mimic characteristics of the facial surface, such as skin roughness and furrowing; therefore, we introduced an active appearance model (AAM) [6] to mimic detailed facial expressions, including surface appearance.

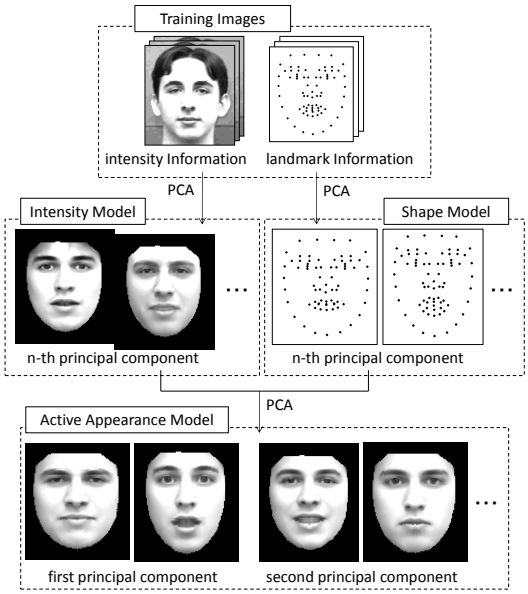

Figure 1. The procedure of AAM construction

AAM is an image appearance model that contains both shape and intensity information. In facial image processing, these factors indicate the individuality of faces, facial expressions, direction of faces, and illumination conditions. By changing the AAM parameters, we can synthesize many types of facial images. Our system first extracts the facial expression elements from the AAM parameters, and then reflects them on the image of another person. This mimicking process makes it possible to duplicate the surface appearance of the face in detail.

Figure 2. Fitting operation

2. Facial expression recognition and synthesis

2.1. Construction of AAM

AAM is a facial image synthesis model obtained from two models: a shape model and a gray-level intensity model. These two models are acquired from the landmark information and gray-level intensity of training images, respectively. Figure 1 outlines the procedure of AAM construction. Several dozen landmarks are put on training images preliminarily. After normalizing their size and rotation, principal component analysis (PCA) is applied to construct the shape model. The intensity model is obtained from the gray-level intensity data in a similar manner. Then, PCA is applied again to the parameters of these two models, and we obtain AAM.

2.2. Fitting operation

A facial expression is identified by searching parameters that synthesize a facial image similar to that of the user. This is called the fitting operation. This operation uses some parameter optimization methods, such as the steepest descent method. Similarities between a user's facial image and synthesized facial images are usually measured by the average difference of intensity per pixel. Figure 2 shows the sample images during a fitting operation. As shown in this figure, a Gaussian filter is applied in first few steps of the fitting operation to avoid a local solution caused by the roughness of intensity.

Although a Gaussian filter works well for avoiding local solutions, there are many cases where the fitting operation converged into a local solution because the steepest descent method is a type of gradient method. To solve this problem, simulated annealing-based method is proposed [7]. Simulated annealing is a probabilistic search method that permits deteriorative parameter update. It is a desirable feature for getting out from a local solution. The drawback of simulated annealing is that it takes a long time to find the optimum parameters. Therefore, we used both methods (steepest descent and simulated annealing) in different parts of our system.

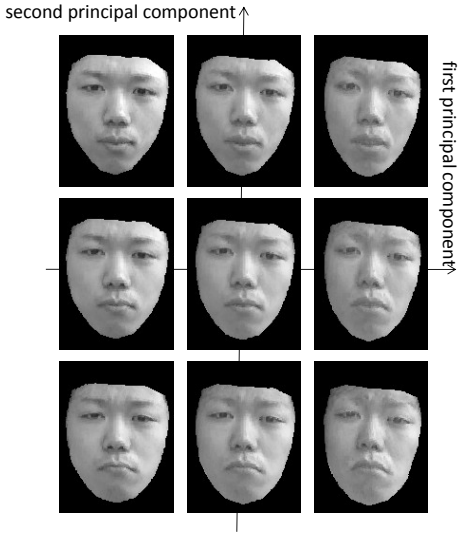

Figure 3. Facial expressions obtained by changing the facial expression parameters

2.3. Construction of facial expression model and facial expression mimicking

In addition to facial expression elements, AAM contains individuality of faces, direction of faces, and illumination conditions. Therefore, modifying parameters changes not only the facial expression but also some other components of the face. This is not desirable for mimicking facial expression. Here, we will show how facial expression elements are extracted from AAM.

The extraction procedure is as follows. First, the AAM parameters that synthesize the neutral expression of a person are detected by the fitting operation. Then AAM parameters for expressive faces are collected and differences from the neutral expression are calculated. These differences represent the movement of expressions. By applying PCA, we obtain a model that expresses the facial expression of the person—the facial expression model. Figure 3 shows the facial expressions obtained by changing the parameters (the facial expression parameters) of the facial expression model.

When an expressive facial image of a person is given, the facial expression parameters of the image can be detected as follows.

1) Prepare the AAM parameters of a neutral expression of the user.
2) Synthesize a facial image by changing the facial expression parameters from the parameters obtained in 1).

Figure 4 shows the process of searching for the facial expression parameters. The facial expression can be duplicated on the image of the face of another

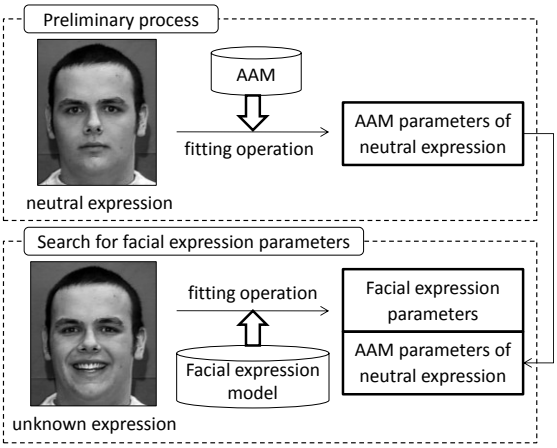

Figure 4. Searching for expression parameters

person by adding the facial expression parameters to the AAM parameters of their neutral expression.

3. A facial expression mimicking system

3.1. Algorithm of facial expression mimicking

The facial expression mimicking system duplicates the facial expression captured by a webcam onto a prepared facial image. The procedure of facial expression mimicking is shown in Figure 5. The following details the procedure.

3.1.1. Step 1: Detection of a face. The location of a user's face is detected by the open computer vision library (OpenCV) [8] developed by Intel Corporation. If a face is not found, the process is canceled; the detected location serves as the initial position of the face in the following steps.

3.1.2. Step 2: Searching for AAM parameters of a neutral expression. AAM parameters of a neutral expression of the user are searched for by the fitting operation. This search is carried out just one time. Since the accuracy of these parameters affects that of the subsequent steps, we employed simulated annealing for this fitting. If the average difference of gray-level intensity between the user's face and the synthesized face is greater than a certain threshold (e.g. 11 of 256 levels), this step is iterated until the average difference becomes smaller than the threshold.

3.1.3. Step 3: Searching for facial expression parameters. The user's facial expression parameters are searched for by the fitting operation. We employed the steepest descent method for this operation to maximize process speed. This step (and the following step) is iterated while the system is running. The initial

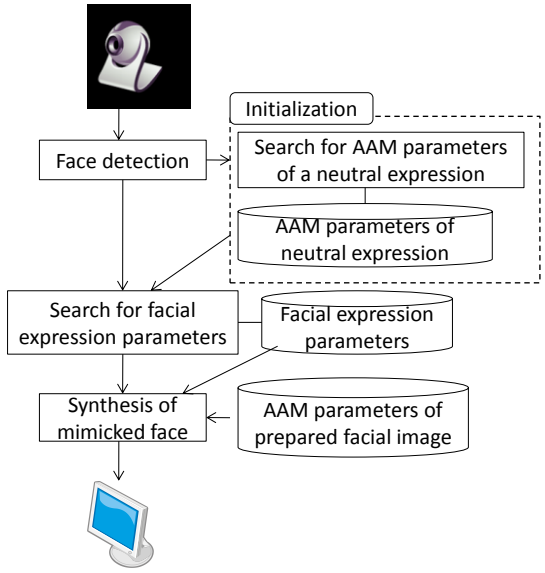

Figure 5. The process of mimicking expression

facial expression parameters for fitting are given by the converged expression parameters of the former frame to accelerate the fitting.

3.1.4. Step 4: Synthesis of mimicked face. The mimicked face is synthesized by adding the expression parameters to the AAM parameters of the prepared facial image. The generated facial image is displayed in a window.

3.2. Practical examples of the system

Figure 6 shows a sequence of images ranging from neutral to surprised expressions; user images, mimicking images of persons from Japan's history, and images of an ex-prime minister of Japan are shown on the first, second and third lines, respectively. These results show that the facial expression model and its parameters work well for duplicating the user's face on these two facial images.

4. Experiments

To investigate the correctness of the facial expression model, we compared the synthesized expressive facial image with the original facial image. We used two databases in the experiment: (A) 300 Japanese neutral facial images from every age group, published by Softpia Japan [9], and (B) 330 Japanese expressive facial images of 11 males in the laboratory. The image size is 640 × 320 pixels and all images are grayscale with 256 levels. First, database (A) is used to construct AAM. Then, 300 expressive facial images (10 males) in database (B) are used for making a facial

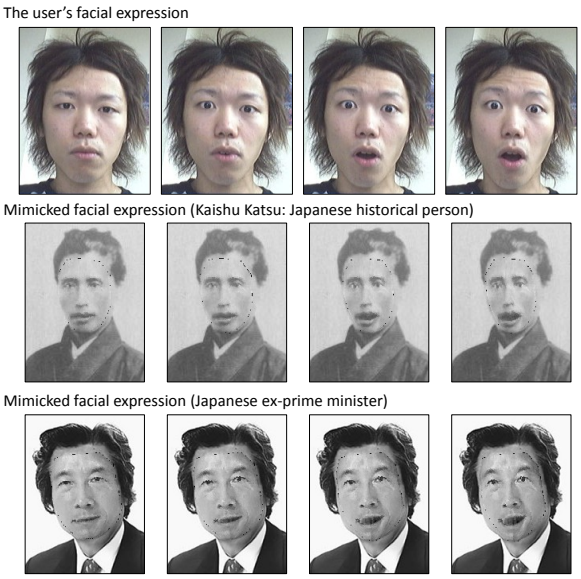

Figure 6. A screen shot of the system

expression model, and 30 expressive facial images (1 male) are used as test data.

We compared the mean square error of gray-level intensity and landmark position between the synthesized and original images. As shown in Table 1, the mean square error of the gray-level intensity is 8.25 (of 256 levels) and that of the landmarks is 5.38 pixels. Figure 7 shows two examples of synthesized images. Sample 1 is a successful example in which the expression parameters are correctly searched for (almost average of results). In this case, the expression is duplicated on the prepared images very well. However, sample 2 failed to search for the facial expression parameters and cannot duplicate expressions satisfactorily. This is because the gray-scale level in a mouth is largely dependent on the alignment of teeth, which differs from person to person.

5. Conclusions and future work

We proposed a facial expression mimicking system based on AAM in which a facial expression model is introduced to identify a facial expression. The experimental results show that the facial expression model and fitting operation gives precise detection of the facial expression parameters and well duplicated the expression on prepared images. Since the system requires only a facial image for the duplication, it can be utilized for various applications, such as an interaction system using the faces of famous people. Future work includes automatic construction of AAM, detection of the neutral expression of a user, and investigation of 3D models.

Table 1. Mean square error of gray-scale intensity and landmark position

	average (std deviation)
MSE: gray-sale intensity	8.25 (1.83)
MSE: landmark position	5.38 (1.54)

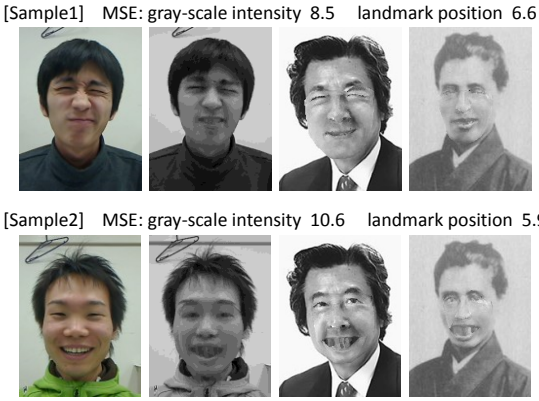

Figure 7. Extraction of expression parameters

6. Acknowledgements

This work was supported by the Grant-in-Aid for Young Scientists (B) 20700156 from MEXT - Japan.

References

[1] Pendulumn's AlterEgo, http://www.studiopendulum.com/alterego/index.html
[2] Mova contour, http://www.mova.com/
[3] Image Metrics | Superior Facial Animation., http://www.image-metrics.com/
[4] Breazeal, C., Buchsbaum, D., Gray, J., Blumberg, B.: Learning From and about Others: Towards Using Imitation to Bootstrap the Social Competence of Robots, Article Life, vol. 11(2005).
[5] Ghys, C., Taron, M., Paragios, N., Bascle, B., Komodakis, N.: Expression Mimicking: from 2D Monocular Sequences to 3D Animations, Advances in Visual Computing : Third International Symposium, ISVC 2007, Proceedings, pp. 621-630(2008).
[6] Cootes, T.F., Edwards, G.J., and Taylor, C.J.: Active Appearance models, IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(6), pp. 681-685(2001).
[7] Saad, A.A., El-Bialy A, Kandil, A.H., Sayed, A.A.: Automatic cephalometric analysis using active appearance model and simulated annealing, ICGST Int J on Graphics, Vision and Image Processing, Special Issue on Image Retrieval and Representation, 6, pp. 51-67(2006).
[8] OpenCV, http://sourceforge.net/projects/opencvlibrary/
[9] Softpia Japan facial image database http://www.softpia.or.jp/en/rd/facedb.htm