Interactive 3D Avatar Synthesis with a Photograph and its Facial Expressions on a Smartphone

[1]In-Ho Choi, [2]Hyeonjoon Moon, [3]Joo-Hyung Kim and [4]Yong-Guk Kim*

[1,2,4]Department of Computer Engineering, Sejong University, 98 Kunja-Dong, Kwangjin-Gu, Seoul, Korea
[3]Department of Electrical Engineering, Korea University, Anam Dong, SungBuk-Gu Seoul Korea

*Email Address: ykim@sejong.ac.kr

Abstract. We present a method by which a user can create his 3D avatar using a photograph. To model the face, we use an Active Appearance Model since it provides a reliable and efficient way to model a face including several facial features within it. The user can change the shape of the face as well as facial features interactively. In addition, this avatar is able to express diverse facial expressions. All software components are implemented as an application program running on a smartphone.

Keywords: 3D facial avatar, Active Appearance Model, Facial Expressions, Smartphone

1 INTRODUCTION

Humans have been evolved to express their emotions and to recognize other's emotions to communicate with each other. It is known that humans identify one another by their facial shape and color, and yet facial features are also important visual attributes. In perceiving human facial expressions, variations of facial features within the given face play a critical role. This makes the face modeling a challenging issue in computer graphics and computer vision. Well-constructed face models find many interesting applications, such as movies and 3D avatars controlled by facial expression recognition. In the cyber space, we like to express and recognize other's emotion for fun in the remote place. In such case, creating an avatar is necessary for replacing a real person. To make an avatar is not an easy task since it requires a lot of equipment and software. Although there are a few avatar application programs for smartphones, most of them are for 2D avatar. Moreover, it is difficult to find any applications for making a 3D avatar using a photograph of user himself.

In this study, we propose a new system by which a user can make a personal 3D avatar that can express facial expressions on a smartphone. Many people normally think that humans being could generate and express hundreds of different facial expressions. And why we deal with only six emotions in the present and other studies?

During last century, researchers in this area found that human being have a group of universal facial expressions across diverse ethnic groups, and moreover with these six basic facial expressions we could generate other subtle facial expressions as well.

2 FACE MODEL

2.1 Active Appearance Model (AAM)

When an image containing a face is given, the first task is to locate the position of the face, called it face detection. Locating the face is a time consuming task since one has to search it within the entire image. Among many, the AdaBoost algorithm, adopted in this study, is the most popular because of its high speed, though it requires a long training session. The second task is to detect several facial features concurrently within the detected face. The AAM was initially developed for modeling the face using the shape and appearance parameters. So it allows us model the whole face including facial features. Although it requires an extensive training session using the face DB, it operates fast and reliable when the modeling process is completed.

Now the main task of an operating AAM is to synthesize a face model for the given face image. Then it tries to minimize the error between the given face and the synthesized face through iterations. To synthesize a face, we need to find the model parameters using an image alignment algorithm. Such image alignment algorithm normally takes a lot of computing time. The Inverse Compositional Image Alignment (ICIA) method [2] is a recent efficient way to reduce the processing time, as it is possible to track a face even on a real-time basis.

The shape in the AAM is represented by meshes generated by connecting landmarks, and mathematically defined using a vector $S =$ $x_2, y_2, \ldots, x_{„}\ y_V I^T$. The shape vector allows us linear variation of the vector. Using such manipulation, we **can** define it as follow:

$$S = S_o + p\ _i s \qquad (1)$$

where coefficient pi is the parameter of the shape and si is the orthomormal vector that is obtained by Principal Components Analysis (PCA), and s_o is the mean shape.

The appearance (or color) in the AAM is defined by the pixel vector $x = (x, y)^T$ within the shape s, warp to the mean shape s_o. Similar to the shape case, the appearance can be defined by

$$A(x) = A_o(x) + 2,21, (x) \qquad (2)$$

where the coefficient Xi is the appearance parameter, and Ai is the orthomormal vector that is also obtained by the PCA method. A_o is the mean appearance.

$$M(W(x; p)) = A(x) \quad (3)$$

In (3), M represents the model instance. It can be generated by warping the pixel x within the shape s by updating the shape s using parameter p. Here, the parameter p is calculated using the ICIA method. Once an average model is ready, we could apply that average model (or a face template) to the input face image to model the face. In the present system, an avatar synthesis for a person consists of two stages. First, the AAM is applied to the input image. And then we separate three major facial features such as eyes, nose, and mouth. The user could adjust (or fit) each part using the touch-based interface. When the fitting tasks are completed, the system generates skin color of the model face. Secondly, the user can vary the shape of the face using the shape controller interface (see Fig. 2) depending on user's personal preference.

3 3D AVATAR SYNTHESIS

As an AAM is applied to the face in the given image, three facial features (Eyes, nose, mouth) are separated and warped to generated facial texture. The remaining area except the face can be textured using the colors from the face. In generating the texture of the face, there are options: one is to use the whole face like one to one mapping and the other is to utilize the symmetry of the face. In the latter, the can chooses either the left part of his face or the right part. As the last step, a mean filter is used between two texture parts. Empirically, it appears that the mean filter outperforms the Gaussian filter visually.

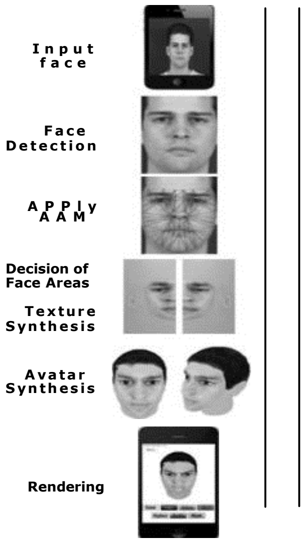


Fig. 1. The System Flow

3.1 Shape Synthesis

The ratio of the face is calculated using the coordinates of AAM's shape vector. The face shape has the index of the shape and the regional index of the face is shown in Table 1. In our system, the shape of an avatar can be changed according to the user's preference.

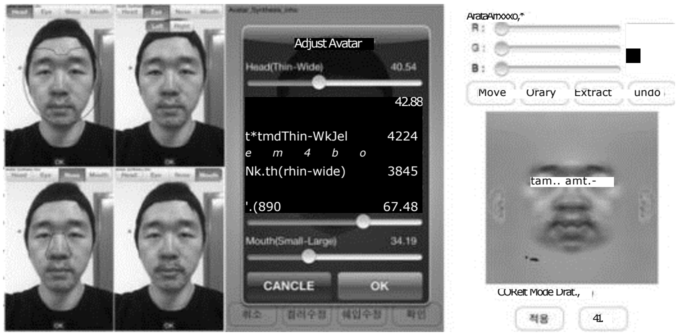
REGION	FACIAL AREA	X		Y		RATIO
1	HEAD (WIDE OR THIN)	3	4	1	2	WIDTH
2	JAW (WIDE OR THIN)	5	6	1	2	WIDTH
3	NOSE (WIDE OR THIN)	7	8	1	2	WIDTH
4	MOUTH (WIDE OR THIN)	9	10	1	2	WIDTH
5	EYE (LARGE OR SMALL)	11	12	1	2	WIDTH
6	MOUTH (LARGE OR SMALL)	14	15	2	13	HEIGHT

Table 1. Ratio Index for different facial regions

4 INTERACTIVE CONTROLLER

4.1 Controller for Fitting & Adjusting

We have developed an interactive controller by which the uses can adjust and fit the shape of the avatar within the mobile environment. The controller can improve the synthesis process and change the shape and color (or texture) of the avatar. The left of Fig. 2 shows the shape controller by which user can change the



shapes of the eyes, nose, and mouth, respectively

Fig. 2. Shape Controller (left) and Skin Texture Controller (right)

4.2 Skin Controller

The right of Fig. 2 shows the skin controller by which one can remove the scar, spot, unwanted part within the avatar. In addition, the color (or texture) can vary using the three color coordinate as well.

5 Facial Expression Controller

5.1 Diverse Facial Expressions

The main idea of the facial expression controller is similar to that of the AAM. To make a facial shape is to build a linear combination of the shape components, since the shape formation of the AAM is based on PCA (Principal Component Analysis). A facial expression E of an avatar is constructed using 3D coordinates (i.e. x, y, z). In the present study, the vertex number for a 3D face is 6293. The user can draw a sketch within a rainbow disk as shown in the right part of Fig. 3. The sketch trajectory indicates how the facial expressions of the 3D avatar evolves as time goes by. The facial expression can be also instructed as a point. The sketch facial expression scenario can be replayed by pressing a replay button. The time span of the given scenario can be chosen as shown in the menu.

5.2 A Fashion Avatar with Facial Expressions

We would like to apply the facial expression controller to a walking avatar who wears a 3D fashion. The fashion models often express their emotion to interact with audience. The right part of Fig. 3 shows different facial expressions of a fashion avatar. The emotion can be programmed using a facial expression controller. The trajectory of the facial expression corresponds to the evolution of a facial expression of the avatar.

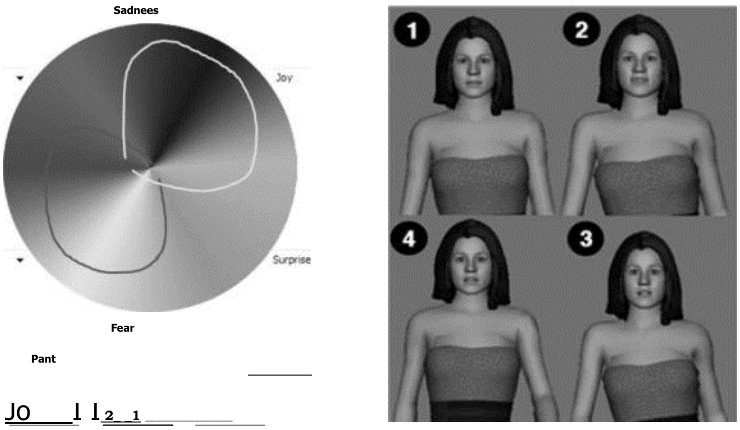


Fig. 3. Facial Expression Controller (left) and a Fashion Avatar with Different Facial Expressions (right)

6 Conclusion and Future Work

A facial expression controller, which is based on AAM, is proposed for the mobile applications. We present a new method by which a user can program facial expressions for an avatar either discrete or continuous facial expressions. The facial expression controller has been applied to the 3D fashion avatar walking on a runaway as an example. In the future, we plan to apply this method for the animation movies.

1. P. Ekman (1999). Basic Emotions. In T. Dalgleish and M. Power (Eds.). Handbook of Cognition and Emotion. Sussex, U.K.: John Wiley & Sons, Ltd.
2. I. Matthews, S. Baker (2004) Active Appearance Models Revisited, International Journal of Computer Visio, pages 135-164,
3. G. Edwards, C.J. Taylor, and T.F. Cootes. (1998) Interpreting Face Images Using Active Appearance Models. In Proc. International Conference on Automatic Face and Gesture Recognition, June (1998), 300-305.
4. P. Ekman. (1972) Universal and cultural differences in facial expressions of emotion. In J Cole (Ed) Nebraska symposium on motivation 1971 (pp 207-283) Lincoln: University of Nebraska Press
5. N. Stoiber et al. (2010). Facial animation retargeting and control based on a human appearance space. Computer Animation and Virtual Worlds, Vol. 21, No. 1, pp. 39-54.