Facial Image Morphing by Self-Organizing Feature Maps

Mu-Chun Su and I-Chen Liu
Department of Electrical Engineering, Tamkang University, Taiwan, R.O.C.
E-mail : muchun@ee.tku.edu.tw

Abstract

In this paper, we propose a new facial image morphing algorithm based on the Kohonen self-organizing feature map (SOM) algorithm to generate a smooth 2D transformation that reflects anchor point correspondences. Using only a 2D face image and a small number of anchor points, we show that the proposed morphing algorithm provides a powerful mechanism for processing facial expressions.

I. Introduction

In recent years, facial morphing algorithms have been widely used in various applications, such as face recognition, criminal identification, animation, virtual reality, low-bandwidth video conference image transmission, intelligent man-machine interface, and entertainment industry etc. Most image morphing algorithms can be classified into three classes. The first one is mesh morphing which requires a control mesh on an image whereas its feature may have an arbitrary structure [1]-[3]. The second class is called field morphing which is specified by defining a weighted average of the influence fields around each of the features of the image to be warped [4]. The last one is point-based morphing where each feature is distinguished by a set of called anchor points in the image [5]-[9]. A warping can be derived by constructing the surfaces that interpolates these scattered anchor points. The radial basis functions [6] and the thin plate splines [7] are the two most popular warping functions.

The human face is an elastic object that consists of organs, numerous muscles, and bones. Muscles are generally fix by one side to the skeleton and by the other side to the skin. When a muscle contracts, the transformation of the skin area attached to the muscle may result in a certain type of visual effect. Both the contraction strength and the tissue elasticity account for the skin deformation area. These deformations have the feature of locality. That is, an expression is weighted by the contraction strength of the different muscles involves in the considered expression instead of all facial muscles. For example, a happy expression mainly results form the contraction of the chin square and the superciliary.

Recently, numerous technical reports have been written about successful applications of the self-organizing feature map (SOM). These applications widely range from vector quantization, speech recognition, to face recognition and cluster analysis. Extensive and good overview of the applications of the SOM can be found in [10]. In this paper, we try to explore the possibility of applying the SOM to facial image morphing.

II. Review of the SOM algorithm

The principal goal of self-organizing feature maps is to transform patterns of arbitrary dimensionality into the responses of one- or two-dimensional arrays of neurons, and to perform this transform adaptively in a topological ordered fashion. The essential constituents of SOMs are as follows [11]

- an array of neurons that compute simple output functions of incoming inputs of arbitrary dimensionality,
- a mechanism for selecting the neuron with the largest output, and
- an adaptive mechanism that updates the weights of the selected neuron and its neighbors.

Note that each component of the input pattern $\underline{x} = (x_1, \cdots, x_n)^T$ is simultaneously connected to each of an N×N array of neurons. The output of the *jth* neuron is defined as

$$Out_j^{(t+1)}(\underline{x}(t+1)) = f\Big(\sum_{i=1}^{n} w_{ji} x_i(t+1) + \sum_{l \in L_j} w'_{jl} Out_l^{(t)}(\underline{x}(t))\Big) \qquad (1)$$

where w_{ji} denotes the connection weight from the *i*th input component to the *j*th neuron, w'_{jl} denotes the lateral feedback weight from neuron l to neuron j, L_j denotes the subset that contains the neurons having lateral feedback weights connected to neuron j, $f(\cdot)$ is a suitable activation

0-7803-5529-6/99/$10.00 ©1999 IEEE

function, t denotes a discrete time index, and $\underline{x}(t)=[x_1(t),\cdots,x_n(t)]^T$ represents the tth input pattern. The training algorithm for forming a SOM is summarized as follows:

Step 1: Initialization: Choose random values for the initial weights $\underline{w}_j(0)$.

Step 2: Winner Finding: Find the winning neuron j^* at time t, using the minimum-distance Euclidean criterion:

$$j^* = \arg \min_j \|\underline{x}-\underline{w}_j\| \quad ,j=1,\cdots,N^2 \quad (2)$$

where $\|\cdot\|$ indicates the Euclidean norm.

Step 3: Updating: Adjust the weights of the winner and its neighbors, using the following rule:

$$\underline{w}_j(t+1)=\underline{w}_j(t)+\eta(t)h_{j,j^*}(\underline{x}(t)-\underline{w}_j(t)) \quad (3)$$

where $\eta(t)$ is a positive constant and h_{j,j^*} is the topological neighborhood function of the winner neuron j^* at time t.

A typical choice of h_{j,j^*} is the Gaussian function

$$h_{j,j^*} = \exp\left(-\frac{d^2_{j,j^*}}{2\sigma^2(t)}\right) \quad (4)$$

where d_{j,j^*} denotes the lateral distance between winning neuron j^* and excited neuron j. The Euclidean distance is usually adopted to measure the lateral distance, that is, $d_{j,j^*}=\|\underline{r}_j-\underline{r}_{j^*}\|$ where the vector $\underline{r}_j$ and $\underline{r}_{j^*}$ denote the positions of excited neuron j and winning neuron j^* in the discrete output space, respectively. The parameter $\sigma(n)$ is the "effective width" of the topological neighborhood function. It measures the degree to which excited neurons in the vicinity of the winning neuron participate in the learning process.

III. The Proposed Facial Image Morphing Algorithm

We let $\underline{s}_i$ and $\underline{d}_i$ with $i=1,2,\cdots I$, be respectively the two-dimensional coordinate vectors for the original and the destination anchor points in the image. The value of I is predetermined by the user. These anchor points are obtained manually by mouse clicks on the image. They are selected as they have greater potential in revealing the changes in visual effects.

The neural network we consider is of the same size as the image. Each neuron has a 2-D dimensional synptic weight vector denoted as $\underline{w}_j$ with $j=1,\cdots,N\times M$ where $N\times M$ represents the size of the image. These $N\times M$ neurons are initially situated in the positions of the corresponding $N\times M$ pixels. That is, these neurons are arranged in a $N\times M$ matrix. It should be emphasized that I of the $N\times M$ neurons have their weight vectors coincide with the I original anchor points $\underline{s}_i$ with $i=1,2,\cdots I$ in the image. The I destination anchor points are regarded as the input patterns to the network to be trained.

When the stimulus vector $\underline{d}_i$ is presented to the neural network, a winner neuron is determined as the one with its weight vector initialized to be $\underline{s}_i$. Let j^* be the index of this neuron and $\underline{w}_{j^*}$ be its weight vector. The updating rule of the weights is given by

$$\begin{aligned}\underline{w}_j(t+1) &= \underline{w}_j(t)+\Delta\underline{w}_j(t)\\ &= \underline{w}_j(t)+\eta(t)\Lambda_{j,j^*}(\underline{d}_i-\underline{w}_{j^*})\\ &= \underline{w}_j(t)+\eta(t)\Lambda_{j,j^*}(\underline{d}_i-\underline{s}_i)\end{aligned} \quad (5)$$

where $\eta(n)=\frac{n}{K}$, K denotes the maximum number of iterations or K-1 equals the number of intermediate images, and Λ_{j,j^*} is the neighborhood function of the winner neuron j^*. Note that the weight increment given by the second term on the right-hand side of Eq. (5) is not the usual one in the conventional Kohonen self-organizing feature map algorithm as shown in Eq.(3). The term $(\underline{d}_i-\underline{s}_i)$ denotes the displacement that should be taken when the original anchor point $\underline{s}_i$ moves toward to its corresponding destination anchor point $\underline{d}_i$.In order to get realistic warped images, we should carefully deal with the determination of the neighborhood function Λ_{j,j^*}.

We know that the human face is a complex elastic object where muscles and bones interact with each other. Every facial movement is the result of muscular action. The realistic animation of a face is extremely difficult to simulate by a computer. Presently, five types of animation models might be identified.

- The muscle-based method consists in precise and physical simulation of the shapes of facial muscles [12]-[14].
- The geometrical method uses geometric deformations to simulate the visual results of muscle contractions but not the muscles themselves [15].
- The performance-driven method is to transfer a human actor's facial movements to face model on the computer [16].

- The interpolation method is where key-frames, or basic expressions, are located at different moments in the animated sequence. Intermediate expressions are then simply interpolated between two successive basic expressions [17].
- The parametric method uses a set of parameters to control facial expressions [18].

A detailed discussion on the facial animations given in [19]. Although these 3D models can exhibit impressive results they usually are time consuming. Since we focus on 2D facial image morphing, our aim is to build a simplified 2D facial muscle model so that we know how to appropriately determine the shapes of the neighborhood functions. This model only takes into account the visual results of some muscle contractions but not the actual facial tissue dynamics. The human face contains numerous muscles, as shown in Fig. 1(a). Anatomy distinguishes two kinds of muscles. The first one are long-limbed muscles connected to the bone and the skin. The second one are the circular or elliptical muscles, generally located around an aperture such as the eyes, and the lips. Different kinds of muscles result in different types of skin deformation. For example, the oblique muscles contract in an angular direction and the circular muscles form ellipical rings. Based on these observations, the determinations of the shapes of neighborhood functions should depend on the locations of the anchor points. The simplified 2D facial muscle model shown in Fig. 1(b) illustrates the relationships between the locations of the anchor points and the corresponding neighborhood functions. A general form of the neighborhood functions of interest is

$$\Lambda_{j,j^*} = e^{-\left\|\underline{w}_j(0)-\underline{w}_{j^*}(0)\right\|_\Sigma^2/2\sigma^2} = e^{-\left\|\underline{w}_j(0)-\underline{s}_i\right\|_\Sigma^2/2\sigma^2} \tag{6}$$

where $\|\cdot\|_\Sigma$ denotes the Mahalanobis distance between $\underline{w}_j(0)$ and $\underline{w}_{j^*}(0)$ (or $\underline{s}_i$). The covariance matrix Σ is determined according to Fig. 1(b). The parameter σ determines the influence zone of the corresponding muscle.

IV. Experimental Results

We used a 480×640 man's color facial image to demonstrate the proposed facial image morphing algorithm. Presently, we simulate two muscles groups, namely the eyes muscles and the mouth muscles because they provide more visual effect than other muscles group.

Experiment I. Eye

Fig. 2(a) and 2(b) show the images where the original and the destination anchor points are marked by small white boxes. Fig. 2(c) then illustrates the morphing result of closing the right eye.

Experiment II. Mouth

The object is to warp the normal expression of the man to the sadness expression. Fig. 3(a) and 3(b) are the source images of a normal expression with anchor points located in the regions of the mouth. The warped image is shown in Fig. 3(c). We see that the normal expression has been warped to be a sadness expression.

V. Conclusion

An efficient facial image morphing algorithm based on the self-organizing feature map algorithm is developed here. A simplified facial muscle model is also proposed to provide us the information of how to determine the neighborhood functions so as to achieve a more natural look of the warped images. The value of the parameter σ in the neighborhood function plays an important role of local influence zone of reference points. The selection of σ should depend on the size of facial image and the location of the anchor point.

Reference

[1] T. Nishita and E. Nakamae, "Metamorphosis using Bezier Clipping," Proceeding of the 1st Pacific Conference in computer graphics and application, Korea, 1993, pp. 62-173.

[2] G. Wolberg, Digital Image Warping, IEEE Computer Society Press, 1990.

[3] G. Wolberg, "Recent Advances in Image Morphing," Computer Graphics International, Korea, pp. 24-28, 1996.

[4] T. Beier and S. Neely, "Feature based image metamorphosis," SIGRPH'92, pp. 25-42, 1992.

[5] N. Arad, N. Dyn, and D. Reisfeld, "Image warping by radial basis functions: Application to facial expressions," Computer Vision, Graphics, and Image Processing, vol. 56, no. 3, Mar. 1994, pp. 161-172.

[6] N. Arad and c. Reisfeld, "Image warping using few anchor points and radial functions," Computer Graphics forum, 14(1):35-46, 1995.

[7] F. L. Bookstein, "Principal warps: Thin plate splines and the decomposition of deformations," IEEE Transactions on Pattern Analysis Machine Intelligence, 11(6):567-585, 1989.

[8] H. Aboul-Ella nad M. Nakajima, "Image morphing of facial images transformation based on navier elastic body splines," Computer Animation 98. Proceedings, 119-125, 1998.

[9] A. Goshtasby, "Piecewise linear mapping functions for image registration," Pattern Recognition, 20(5):525-533, 1987.

[10] T. Kohonen, E. Oja, O. Simula, A. Visa, and J. Kangas, "Engineering application of the self-organizing map," Proceedings of the IEEE, Vol. 84, No. 10, pp. 1358-1383, 1996.

[11] T. Kohonen, Self-Organization and Associate Memory, Springer-Varlag, London, 1984.

[12] K. Waters, "A muscle model for animating three-dimensional facial expression," Computer Graphics ACMSIGGRAPH, 1987, vol. 24(4), pp. 17-24, 1987.

[13] D. Terzopoulos et K. Waters, "Physically based facial modeling, analysis and animation," Visualization and Computer Animation, vol. 1(2), pp. 73-80, 1990.

[14] S. M. Platt and N. I. Badler, "Animating facial expressions," Computer Graphics, Vol. 15, No. 3, pp. 245-252, 1981.
[15] P. Kalra, A. Mangili, N.M. Thalmann et D. Thalmann, "Simulation of Facial Muscle Actions Based on Rational Free Form Deformations," Eurographics 92, vol. 11(3), pp. 59-69, 1992.
[16] I. Essa, A. Pentland, "Facial expression recognition using a dynamic model and motion energy," Proc. of Int. Conf. On Computer Vision, pp. 360-367, Los Alamitos, CA, 1995.
[17] F. I. Parke, Computer Generated Animation of Faces, Master's Thesis, University of Utah, 1972.
[18] F. I. Parke, "Parameterized models for facial animation," IEEE Computer Graphics and Applications, Vol. 2, No. 9, pp. 61-68, 1982.
[19] F. I. Parke and K. Waters, Computer Facial Animation, A. K. Peters, 1996.

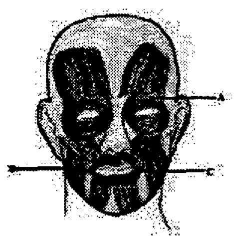


Fig. 1(a): Human facial muscles where A is the orbicularis oculi muscle, B is the masseter muscle and C is the orbicularis oris muscle.

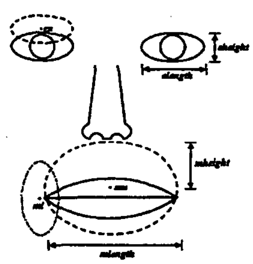
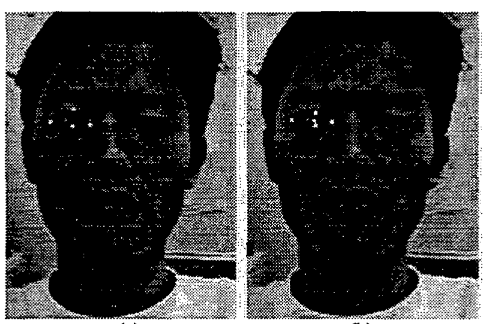


Fig. 1(b): The simplified 2D facial muscle model.



(a) (b)

Fig. 2: (a) The image with the original anchor points. (b) The image with the destination anchor points.

Fig. 2(c) The morphing result of closing the right eye.

(a) (b)

Fig. 3: 3(a) The image with the original anchor points. (b) The image with the destination anchor points.

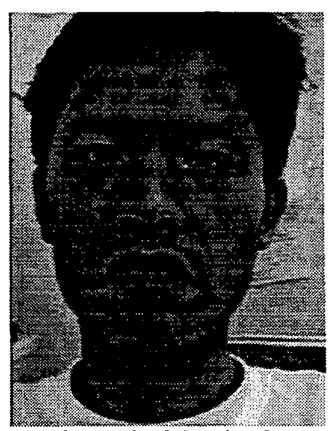


Fig. 3(c) The morphing result of changing the normal expression to the sadness expression.