2010 International Symposium on Nonlinear Theory and its Applications
NOLTA2010, Krakow, Poland, September 5-8, 2010

Facial Feature Animation and Its Artistic Representation

Yang Yang†‡, Yuanqi Su†, Yuehu Liu† and Yoshifumi Nishio‡

†Institute of Artificial Intelligence and Robotics
Xi'an Jiaotong University, Xi'an 710049, China
‡Dept. of Electrical and Electronic Engineering
Tokushima University, Tokushima 770-8506, Japan
Email: {yyang, yqsu, liuyh}@aiar.xjtu.edu.cn, nishio@ee.tokushima-u.ac.jp

Abstract—Facial animation and artistic representation are active research topics in the multimedia industry. This paper addresses facial object with geometric feature to generate expressions using the artistic representation. Firstly, we propose an animation method to transfer expressions from a source face to a target face automatically, which is effective when only a single face with neutral expression is available as the novel target. Then, we provide a facial artistic representation framework and exhibit its producing results by facial cartoon and sketches. Combining above work by building the corresponding map between the facial animation and representation features, we generate expressions for the target face vividly.

1. Introduction

Facial feature animation and its artistic representation has played an important role in the human's lives with increasing requirements in the entertainment industry and multimedia communication. Recently, the movie Avatar made a huge success by using the related techniques. It took us to a spectacular world by exhibiting real actors' performance onto fanciful characters in the movie. Meanwhile, more and more communication tools in the daily life, like the MSN and mobile phone, show their potential applications for the personalized interface. To generate the artistic facial animation will be welcomed by every user. In this paper, we firstly animate facial features by transferring expressions from a source performer's face to target faces, and then exhibit the animation results using the artistic representation.

The commonly used animation technique in the movie industry needs animators to repeatedly define facial animation parameters for every face. In addition, it often requires complex computing and manual adjusting because of the correlations among these parameters. Some blendshape methods [1]-[3] were proposed based on key examples of the source and target faces. They simplified the animation process just to estimate and transfer the blending weights. However, it is also a tedious work to prepare key examples for every face. In this paper, our animation method directly transfers and adjusts motion vectors of the source face to target faces, which can animate the novel target face effectively.

Some impressive works have been developed to produce facial artistic representations. PicToon [4] utilized a statistic model with non-parametric sampling to create cartoon from the given face image. It can generate the facial sketches and specific animations from templates. The MSN cartoon [5] was an online application published by Microsoft. Human's interactions are required by the system to select and adjust the personalized cartoon shape and accessories. A facial sketch generation method was proposed by Xu et al. [6], which gave a hierarchical facial representation for rendering facial sketches. In our early work about a face cartoon producer [7], a cartoon rendering method was proposed based on the automatic template feature fitting and adaptation. In this paper, we extend the producer to generate artistic animation results by the corresponding map built between the facial animation and representation features.

2. System Overview

The system overview is shown in Fig. 1. When input a target face image with neutral expression, a set of feature points are extracted by the ASM [8] method. We can render an artistic representation for this target face by the artistic face producer. For generating expression animation sequences, source expressions are transferred to target facial features firstly, and then the whole animation results are represented with the same style as the artistic target face. In the facial expression database, examples are labeled off-line on various expressions of different people. In the artistic face database, rendering templates are drawn by artists. In the following sections, we describe each function block in the facial feature animation and representation in more detail.

3. Expression Transfer for Facial Feature Animation

Human's faces share a common topological structure and exhibit a similar movement when making the same expression. When just one single face is available as the target input, Blanz et al. [9] gave an effective expression transfer method by the vector space operations. Their method is based on the assumption that the 3D displacements of surface points are the same for all individuals, which is just

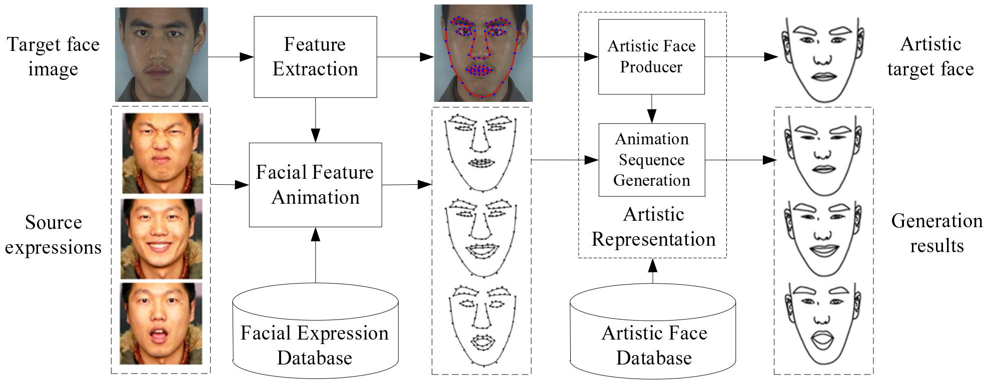

Figure 1: The system overview.

suited to very similar faces. We propose a two steps expression transfer method. Inspired by Blanz's work [9], we could firstly acquire a coarse animation result using the direct motion vector's transferring. Although that may generate an unreasonable face, it has mapped the source face's movement on the target face roughly. And then we adjust this coarse animation result according to the personality of the target face.

3.1. Global Movement Transferring

Facial feature is represented by a set of feature points' coordinates $C = (x_1, y_1, x_2, y_2, ..., x_N, y_N)$, where N is the points' number. We suppose that the first frame in source expression sequence is always with the neutral expression. Motion vectors are calculated by the source neutral face C_{SN} and expressional face C_{SE}, and transferred to the target facial feature C_{TN} directly as:

$$C_{TE} = C_{TN} + C_{SE} - C_{SN}. \quad (1)$$

Without the magnitude and direction adjusting, the transfer result C_{TE} is easily to cause an unreasonable face. To improve it, we project C_{TE} in the facial space and reconstruct it by principal component analysis (PCA) model:

$$C'_{TE} = \bar{C} + Pb, \quad (2)$$

where, $\bar{C}$ is the mean shape of training examples and P contains unit eigenvectors of the covariance matrix. b is the principal component for a given C_{TE}. The reconstructed C'_{TE} is the most similar reasonable face as C_{TE} for the reason of the PCA.

One problem is that the PCA model is learned based on various faces in the database, which is hard to maintain the personality especially from a coarse input. Moreover, the motion vectors of the source face also need to be adapted for the target face. In the following part, since a globally reasonable facial structure could be guaranteed by the PCA, we discuss the local feature adjustment for each facial component.

3.2. Local Personalized Adjustment

Xiong et al. [10] proposed a facial component model using neighbor reconstruction method for personalized expression synthesis. In the facial parameter space, they used neighbor examples to reconstruct target face, and transferred reconstruction weights for generating target expressional faces. However, every expression of the source face should be known in advance by the animator, since they needed re-preparing corresponding expression's examples. Here, we use the neighbor reconstruction as a constraint to formulate the personalized adjustment as an optimization problem.

We construct the facial expression database by a set of example pairs. Every example pair contains one neutral face and one expressional face belonged to the same person. All facial features are separated into seven local subsets according to facial components, including left and right eyes, left and right eyebrows, nose, mouth and face contour.

Firstly, each feature subset of target face finds the neighbor examples by the k-nearest neighbor algorithm from neutral faces in the example pairs.

And then, we calculate the reconstruct weights for local component based on two principles: one is the reconstructed similarity with target neutral face; the other is to match with the current expression. Using the global movement transferring result C'_{TE} as the current expressional face, the optimization function is built as follows:

$$\begin{aligned} &\min_{\alpha_i^c} (\|C'^c_{TE} - \sum_{i=1}^{K} \alpha_i^c S_{iE}^c\|^2 + \lambda \|C_{TN}^c - \sum_{i=1}^{K} \alpha_i^c S_{iN}^c\|^2) \\ &s.t.\ \ \alpha_i^c \geq 0\,, \sum_{i=1}^{K} \alpha_i^c = 1. \end{aligned} \quad (3)$$

We use $c \in (1, 2, ..., 7)$ as the index for seven facial components to represent facial feature subset. S_{iN}^c and S_{iE}^c are the neutral and expressional facial feature subset belonged to ith example pair in the K neighbors. Parameter λ balances two terms in Eq. (3).

This is a quadratic problem with linear constraints, where the objective function is positive semidefinite. We use fmincon function in Matlab to solve Eq. (3) and acquire a set of optimized weights α_i^{*c}.

Finally, the fine animation result C_{TE}^{*c} is acquired by reconstructing from neighbor expressional examples using weights α_i^{*c}:

$$C_{TE}^{*c} = \sum_{i=1}^{K} \alpha_i^{*c} S_{iE}^{c}. \quad (4)$$

4. Artistic Facial Representation

Our prior work [7] proposed a face cartoon producer for automatically rendering cartoon based on extracted features from static input face image. This producer can be easily extended to a multiple styles artistic face producer by just changing the rendering templates. Liu et al. [11] presented a three-layer framework for cartoon facial animation, but did not discuss the animation method for the sequence generation. In this section, we briefly introduce the framework of artistic face producer and then give the details about the animation sequence generation.

4.1. Artistic Face Producer

Given an input face image and extracted feature points, the process for generating artistic presentation has the following steps.

Template selection: A group of rendering templates are drawn by artists. All these templates are parameterized with rendering vector's information, rendering rules, facial geometric features, and so on. For the target face, the most similar template is selected by the facial features' comparison.

Model adjusting: The selected template is similar to the target face, but still needs to be adapted for exactly fitting the target facial feature by the geometric deformation algorithm.

Artistic Rendering: The producer combines the painting entities with the picked out rendering rules and spatial arrangements of the template, and outputs feature representation with artistic style in a vector way.

With above steps, facial features could be artistically presented correspondingly. Two classes of artistic face rendering results, facial cartoon and sketches, are shown in Fig. 2.

4.2. Animation Sequence Generation

The rendering style should be consistent to all the frames in the expression sequence for one target face. Facial features in the animation sequence have been acquired in the section 3. Here, we build the corresponding map between the facial animation and representation features. Then the rendering vectors of the artistic target face are warped to each animation result correspondingly.

Suppose the facial feature in artistic target face is represented as (u_i, v_i), and facial feature of an arbitrary frame in the animation sequence is (x_i, y_i). To build correspondence mapping between two set of features, we use thin-plate splines (TPS) [12] in the following form for 2D points.

$$\begin{aligned} f_u(x,y) &= d_1(x,y) + \sum_{i=1}^{N} \omega_{1i} U(\|(x,y)^T - (u_i,v_i)^T\|) \\ f_v(x,y) &= d_2(x,y) + \sum_{i=1}^{N} \omega_{2i} U(\|(x,y)^T - (u_i,v_i)^T\|). \end{aligned} \quad (5)$$

Here, $U(r) = r^2 \ln r$ is the kernel function in the TPS, d represents the affine transformation, and ω_i represents the non-affine deformation. All the coefficients are solved by a minimized energy function as defined in [12].

The reason for choosing TPS as the mapping function is its suitability in the facial deformation. There is a smoothness measure term $\beta \int\int [(\frac{\partial^2 f}{\partial x^2})^2 + 2(\frac{\partial^2 f}{\partial x \partial y})^2 + (\frac{\partial^2 f}{\partial y^2})^2] dxdy$ in the minimized energy function for the TPS fitting, which is in charged by the parameter β. We can control the smoothness of the transformation by different β flexibly. During people making expressions, rigid transformation dominates the eyebrow's motions, so that we choose a bigger β to animate eyebrow; but for the mouth, a smaller β is used for its variable shape.

At the end, we repeat the third step in the artistic face producer to render the entity in each frame for generating the final animation sequence. Some animation results are artistically presented in Fig. 3, in which the source expressions are collected from the Cohn-Kanade database [13].

5. Conclusion

In this paper, we presented a facial feature animation and artistic representation method, which can generate expression animation results with artistic style. Our method is suited for the novel input face image. No complex model parameters need to be adjusted, which saves a lot of manual work for the application. We extended our artistic face producer by the corresponding map built between the facial animation and representation features. The animation results were presented vividly.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant No. 60775017 and the National Basic Research Program of China under Grant No. 2007CB311005.

References

[1] I. Buck and F. Adam, "Performance-driven hand-drawn animation," *Proc. Int'l Symp. on Non-photorealistic Aniamtion and Rendering*, pp. 101-108, 2000.

[2] E. Chuang and C. Bregler, "Performance driven facial animation using blendshape interpolatio,"

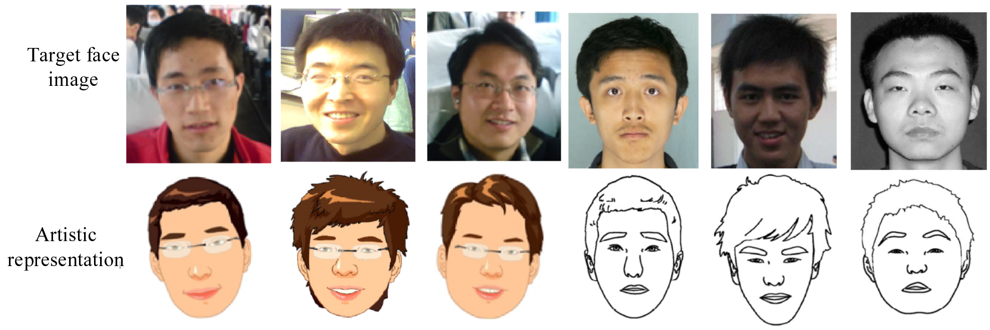

Figure 2: Facial artistic representation.

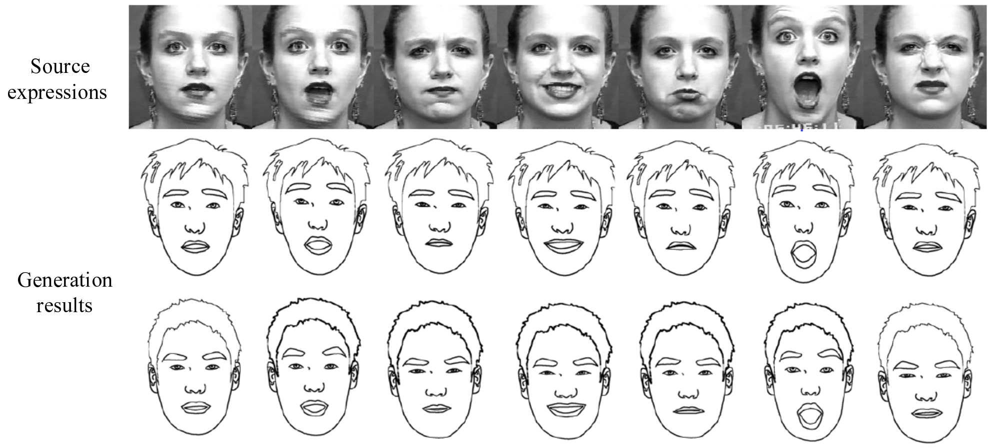

Figure 3: Animation results with facial sketches representation.

Standford University Computer Science Technical Report, CSTR-2002-02, 2002.

[3] X. Liu, T. Mao, et al., "Facial animation by optimized blendshapes from motion capture data," *Computer Animation and Virtual Worlds*, vol. 19, no. 3-4, pp. 235-245, 2008.

[4] H. Chen, Z. Liu, C. Rose, Y. Xu, H.Y. Shum, and D. Salesin, "Example-based composite sketching of human portraits," *Proc. Int'l Symp. on Non-photorealistic Aniamtion and Rendering*, pp. 95-101, 2004.

[5] MSN cartoon, http://cartoon.msn.com.cn.

[6] Z. Xu, H. Chen, S. Zhu, and J. Luo, "A hierarchical compositional model for face representation and sketching," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 30, no. 6, pp. 955-969, 2008.

[7] Y. Liu, Y. Su, Y. Shao, Z. Wu, and Y. Yang, "A face cartoon producer for digital content service," *Proc. Int'l Workshop of Mobile Multimedia Processing*, pp. 188-202, 2008.

[8] T.F. Cootes and C.J. Taylor, "Active shape models-their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38-59, 1995.

[9] V. Blanz, C. Basso, et al., "Reanimating faces in images and video," *Proc. EUROGRAPHICS*, pp. 641-650, 2003.

[10] L. Xiong, N. Zheng, et al., "Facial expression synthesis based on facial component model," *Int'l J. Pattern Recognition and Artificial Intelligence*, vol. 23, no. 3, pp. 637-657, 2009.

[11] Y. Liu, Y. Su, et al., "A facial sketch animation generation for mobile communication," *Proc. IAPR Conf. on Machine Vision Applications*, pp. 532-535, 2007.

[12] F.L. Bookstein, "Principal warps: thin-plate splines and the decomposition of deformations," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 11, no. 6, pp. 567-585, 1989.

[13] T. Kanade, J. F. Cohn, and Y. Tian, "Comprehensive database for facial expression analysis," *Proc. Automatic Face and Gesture Recognition*, pp. 46-53, 2000.