13-25

MVA2007 IAPR Conference on Machine Vision Applications, May 16-18, 2007, Tokyo, JAPAN

A Facial Sketch Animation Generator for Mobile Communication

Yuehu Liu, Yuanqi Su, Yang Yang, Fengjuan Wang, Maojun Yuan

Institute of Artificial Intelligence & Robotics,
Xi'an Jiaotong University, Xi'an, China, 710049
liuyh@mail.xjtu.edu.cn, yqsu@aiar.xjtu.edu.cn

Zhen Ren

France Telecom R & D Beijing Co., Ltd., Beijing, China, 100080
zhen.ren@francetelecom.com

Abstract

The ability to analyze and reproduce specific-person facial expression is an important multimedia content service for mobile communication. In this paper, a novel three-layer facial animation model is presented for generating the personalized facial sketch animation, and the three layers are motion sequence layer, the reference shape layer, and the painting shape layer. Based on the model, a facial sketch animation generator was implemented, allowing mobile subscribers to communicate using personalized facial sketch animations created from face snapshot taken by their phone's camera. The sender takes a snapshot of someone's face and transmits it to a multimedia content server, which then generates the specific-person facial sketch image or animation sequence. The server sends the image or animation to the receiver in multimedia message (MMS). The implementation of the generator architecture is low-cost and feasible for multimedia content service in mobile communication.

1. Introduction

Other than making phone calls, mobile phones now have numerous features: text messaging, Internet access, GPS, electronic money, and mobile movies, which have had a great influence on communication technology. There are four hundred million mobile phone subscribers in China, and that number will increase to six hundred million in 2008.

In pace with the fast growth of mobile subscribers, content service for mobile phones has a great market. For example, MMS has brought in extra revenue for telecom carriers by transmitting mobile phone snapshots. Various value-added multimedia content services have been developed for mobile communication and experience -sharing for geographically separate people, and ceaselessly adopted for satisfying the demands of cell phone subscribers. Personalized facial sketch and animation generation is one of the interesting new content services: a facial portrait taken by the phone camera is used to generate personalized multimedia content, and sent to subscribers or their friends as an MMS. As shown in shown in Figure 1, a subscriber sends a face snapshot to a multimedia content server, a personalized facial sketch image or animation sequence is generated in the server, and it is then sent to the phones of other subscribers. Therefore, the key problem is personalized facial sketch and animation generation.

To recreate 2D facial expressions in a sketch, Chen proposed a face-cartoon generation method: learning the drawing style of a group of facial sketches by a specific artist to generate a non-photorealistic rendering of the face[1]. Bruce Gooch resorted to combing the separately thresholded luminance and brightness images for generating human facial illustration. By exaggerating the difference between the particular face and the average face, a so-called caricature could be generated[2]. Miran introduced his system called LiveMail, which allows mobile subscribers to communicate using personalized 3D face models created from images taken by their phone cameras[3].

The pseudo-muscle model was used to implement 2D facial sketch animation[4]. A group of pseudo-muscles were predefined for controlling the motion of the painting path. To generate a specific expression on the facial sketch, empirical values were assigned to each pseudo-muscle as the facial motion vector. By tuning the facial motion of nearby regions around pseudo-muscles, the painting path of the facial sketch could be transformed to reproduce some different facial expressions.

In this paper, we present a facial sketch animation generator and its architecture, which generates personalized facial sketch and animation content as MMS for mobile communication. The generator is low-cost and feasible for multimedia content service in mobile communication.

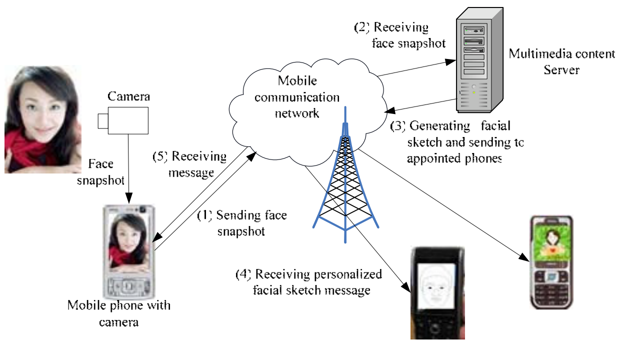

Figure 1. Multimedia contents service use-case scenario for mobile communication

2. System Architecture for Facial Sketch Animation Generator

The main functions of the facial sketch generator are the creation of personalized facial sketches, the generation of facial expression animations, and the transmission of such multimedia messages to their subscribers. As shown in Fig.2, the facial sketch generator is composed of three modules.

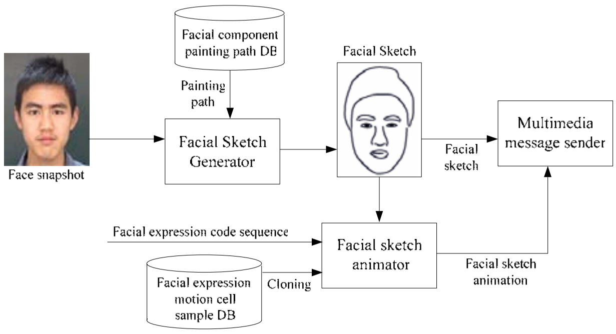

Figure 2. Architecture of personalized facial sketch generator

Facial sketch generation module: After receiving a face snapshot from a phone camera, model adaptation is introduced for generating the painting path of specific-person face components, and a personalized facial sketch image is generated.

Facial sketch animation module: Based on the previously created facial sketch, an expression motion sequence is used to drive the painting path to generate a personalized facial sketch animation sequence.

Multimedia message sending module: The generated facial sketch image or animation sequence can be sent to the user or his friends.

3. Personalized Facial Sketch Generation

To generate a personalized facial sketch from a person-specific face snapshot, the facial contour is necessarily extracted for constructing the painting path, and then the path is used to decide the layout of the strokes. The AdaBoost algorithm[5] can be used for detecting the regions of the face, and 87 predefined facial feature points are extracted using the active shape model (ASM) method[6]. By connecting the extracted feature points, a facial painting path can be generated for coarsely representing the contour of the original face. However, extra strategies should be sought for generating a personalized facial sketch.

3.1 Painting path adaptation for person-specific faces

The painting paths of the facial components should be modeled separately because of their different characteristics. The painting path of eyes and eyebrows is directly decided by connecting their feature points, and the contour of the eyeball is modeled as circle. The profile of hair can be got by segmenting its texture region with matting[7], or selecting a suitable hairstyle from a database. The contours of the nose and mouth are used as reference shapes for generating a personalized painting path using model adaptation.

Model adaptation[8] is based on the fact that generic facial component prototypes can be transformed into the specific face by comparison of their feature points. A facial component painting path database has been pre-built as a prototype. Given a specific face, a proper prototype is selected from the database for representing each component of the specific face. By comparing the feature points of the specific face with those of the prototype, the generic prototype is adapted to fit the specific facial component. This process is shown in figure 3(b)(c), and the proposed model adaptation can be divided into two phases: global adaptation and local adaptation.

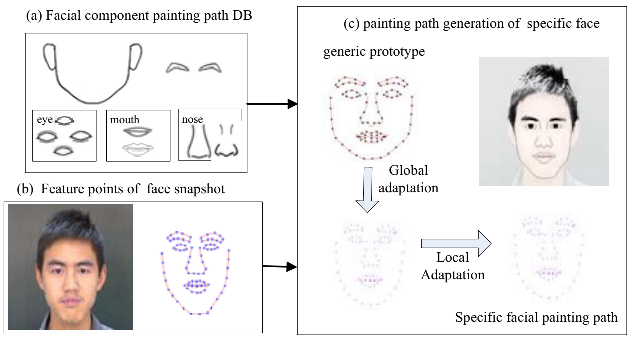

Figure 3. The flowchart of painting path generation based on model adaptation. A generic prototype is selected from the facial component sketch database, which is then transformed to fit specific facial components through global adaptation and local adaptation.

Global adaptation is responsible for tuning the facial component prototype, and this leads to a coarse fit of the specific face. An optimized pair of $\boldsymbol{R}$ and $\boldsymbol{c}$ is used to make two point sets fit each other, where $\boldsymbol{R}$ and $\boldsymbol{c}$ are the transformation matrix and displacement vector of the affine transformation respectively as formula (1) describes.

$$\min_{R,\mathbf{c}}\left\{f(R,\mathbf{c})=\sum_{i=1}^{N}\left\|R\mathbf{s}_i+\mathbf{c}-\mathbf{t}_i\right\|^2=\sum_{i=1}^{N}\left(R\mathbf{s}_i+\mathbf{c}-\mathbf{t}_i\right)'\left(R\mathbf{s}_i+\mathbf{c}-\mathbf{t}_i\right)\right\} \quad (1)$$

Where, $\boldsymbol{T}=\{\mathbf{t}_i=(x'_i,y'_i)\mid i=1,\cdots,N\}$ is the feature point set of the specific facial component, and $\boldsymbol{S}=\left\{\boldsymbol{s}_i=(x_i,y_i)'\mid i=1,\cdots,N\right\}$ is the feature point set of the generic prototype.

Local adaptation makes the generic prototype fit the specific face using geometric deformation in the local manner. We selecte a warping method based on the radial basis function (RBF) for its simplicity, and expressed it as formula (2).

$$\boldsymbol{B}(\mathbf{X})=\sum_{i=1}^{N}\boldsymbol{\alpha}_i\boldsymbol{g}\left(\left\|\mathbf{X}-\mathbf{X}_i\right\|\right)+\boldsymbol{A}(\mathbf{X}) \quad (2)$$

where $\{\mathbf{X}_i=(x_i,y_i)\mid i=1,\cdots,N\}$ denotes feature points of the new shape, $\boldsymbol{g}(\cdot)$ is the pure radial basis function, $\boldsymbol{\alpha}_i$ is the coefficient of the corresponding feature point, and $\boldsymbol{A}(\mathbf{X})$ is the affine component of the deformation function.

3.2 Facial Sketch Rendering

Facial sketch rendering is the process for combining the separately dealt painting path and line template. Hsu proposed a model called "skeletal strokes" for combining the painting path and arbitrarily selected pictures[9,10]. Just as with skeletal strokes, the generated line templates here are deformed along the synthesized painting path, forming a stylistic facial sketch rendering. Versatile facial sketches can be generated by changing the styles of the painting path and line template.

A group of face snapshot images are used to generate the facial sketches with different hairstyles shown in figure 4. You can see that the correspondence of most faces can be recognized.

(a) face snapshot Images (b) Personalized facial sketch with different hair styles

Figure 4. Some person-specific facial sketch rendering results with different hair styles and glass wearing

4. Personalized Facial Sketch Animation

Human facial expression reproduction is attractive for mobile multimedia content service. The facial sketch animator generates facial sketch animation sequences and reproduces specific-person facial expressions. Animating the facial sketch includes two parts: the expression motion control sequence and the facial animation model. The motion control sequence is used to drive the facial animation model for generating sketch animation sequences.

In the proposed facial sketch animation generator, dense motion features are extracted from facial expression video sample to obtain expression motion cells, and the motion cells synthesize the motion control sequence for specific expression.

To implement the facial sketch animation, the three-layer facial animation model is defined: the motion sequence layer, the reference shape layer, and the painting shape layer. The motion sequence layer models the facial expression motion; Simple geometric shapes are used to construct the facial reference shape, facial expression motion is implemented on the reference shape layer; and the personalized facial sketch animation sequence is generated by the painting shape layer.

As shown in figure 5, the reference shape constituted of simple geometric curves such as circle, parabola and so on, which can be easily animated with limited parameters.

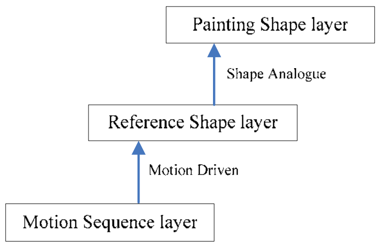

Figure 5. The three-layer facial animation model

Suppose that the referenced shape of eye is denoted by $y = f(x, \Theta^0)$, where Θ^0 represents the controlling parameters corresponded with current state of eye. The motion sequence is $\{(x_i^{(j)}, y_i^{(j)}), i = 1, \cdots, I; j = 1, \cdots, J\}$, I and J denotes the number of the landmarks and frames of the eye's reference shape respectively. Motion driven defines the mapping between Θ^0 and $(x_i^{(j)}, y_i^{(j)})$, and derives the referenced shape to generate sequence of eye.

The geometric deformation algorithms can be used to generate the sequence of painting shape, as you known, the painting shape is similar to the referenced shape in both size and shape, so small deformation is only need for driving the painting path. By the process called shape analogue, the motion control sequence drives the facial animation model to generate a personalized facial sketch animation sequence.

5. Conclusion and Further Works

We have introduced a facial sketch animation generator which can be used as a pure multimedia content server for mobile communication. Users deliver attractive face snapshots using their phone cameras, and can create their own multimedia messages, and send to their friends.

An active model adaptation algorithm is used to generate specific-person facial sketches. It includes two phases: the global adaptation and the local adaptation. The global adaptation tunes a generic facial component prototype to a coarse fit for specific-person facial painting path, while the local adaptation makes the generic prototype closely fit the specific-person face by geometric deformation in the local manner. The personalized facial sketch can be generated by combining the specific-person face painting path and the line template.

To generate a facial sketch animation sequence, the three-layer facial animation model is proposed. The expression motion sequence was used to drive the facial reference shape, and the painting shape was then adapted for generating personalized facial sketch animation.

A facial sketch animation generator has been implemented for multimedia content server, the sender takes a snapshot of someone's face, and transmit it to multimedia content server, then the generator generate some specific-person facial sketch images or animation sequences, and send to the receiver in MMS. The implementation of the generator architecture is low-cost and feasible for multimedia content service in mobile communication.

Further works will focus on rendering the facial sketch animation in advanced artistic style using texture information and exploring a time-efficient deforming algorithm for expression cloning.

Acknowledgments This work was supported by the NSF of China (No.60021302, No.90412010) and sponsored by SRF for ROCS, SEM. The authors would like to thank Dr. Gerard Eude and Dr. Haila Wang of France Telecom R&D, for contributing to the ideas in this paper. Zhengwang Wu, Yu Shao and Daitao Jia completed program of the generator. Dr. Kuizhi Mei and Mr. Luca gave very useful suggestion.

References

[1] Hong Chen, Ying-Qing Xu, Heung-Yeung Shum, Song-Chun Zhu, Nan-Ning Zheng. Example-based facial sketch generation with non-parametric sampling. ICCV'01. Jul 1998: pp.453-460.

[2] Bruce Gooch, Erik Reinhard, Amy Gooch. Human facial illustrations: Creation and psychophysical evaluation, Vol.23(1) , Jan 2004: pp. 27-44.

[3] Miran M. LiveMail: Personalized Avatars for Mobile Entertainment, MobiSys2005, pp.15-23.

[4] K.Waters. A Muscle Model for Animating Three-dimensional Facial Expression. ACM SIGGRAPH`87, Vol. 21(4), 1987: pp. 17-24

[5] G. Ratsch, T. Onoda and K.R.Muller. Soft Margin for AdaBoost, Machine Learning,Vol.42(3), 2001: pp.287-320

[6] T. F. Cootes, C. J. Taylor, D.H.Cooper, and J.Graham. Active Shape Models–Their Traing and Application. Computer Vision and Image Understanding, Vol.61(1), 1995:pp.38-59

[7] Jian Sun, Nanning Zheng, Tao Hai, and Heung-Yeung Shum. Image Hallucination with Primal Sketch Priors. IEEE CVPR2003.

[8] Yuehu Liu, Yunfeng Zhu, Yuanqi Su, Zejian Yuan. Image based Active Model Adaptation Method for Face Reconstruction and Sketch Generation[C]. Edutainment 2006, LNCS 3942, 2006: pp. 928-933.

[9] S. C. Hsu, I. H. H. Lee, N. E. Wiseman, Skeletal Strokes[C]. Proceedings of the 6th annual ACM symposium on User interface software and technology, Dec 1993: pp. 197-206.

[10]S. C. Hsu, Irene H. H. Lee, Drawing and animation using skeletal strokes[C], Proceedings of the 21st annual conference on Computer Graphics and Interactive Technique, July 1994: pp. 109-118.