The Improvement of Parameterized Face Model of Candide Based on MPEG-4 and FACS

Baoying Ma
Changji University, Changji, Xinjiang, China
Email: 377313144@qq.com

Junfeng Yao
Software School of Xiamen University, Xiamen, China
Email: yao0010@xmu.edu.cn

Ruohe Yan
Cognitive Science Department of Xiamen University, Xiamen, China
Email: 190129510@qq.com

Binbin Zhang
Virhum Information Technology Co., Ltd., Xiamen, China
Email: zbb922@gmail.com

***Abstract*—The research of face has wide range of applications. Face expression simulation is a new application among them. In this paper, we have improved and perfected parameterized face model of Candide through comparing Candide with MPEG-4 and FACS. After the model has improved, the experiment results indicate that the head is more complete, the face and the head contour are smoother, and the expression is more rich, natural and true. We have achieved good facial expression simulation results.**

***Index Terms*—expression simulation, MPEG-4, FACS, parameterized face model of Candide**

I. Introduction

According to different viewpoints, there are two methods of facial motion simulation: based on geometry and based on physics. The former is from geometry, and the latter is from dynamic theory. The former is more intuitive and less computation, but it cannot guarantee that the motion is correct and reasonable. The latter is more realistic and difficult to produce unreasonable facial motion, but modeling is complex and more computation [1] and [2]. In this paper, we describe the motion features of facial expression by improving parameterized model of Candide, and we control facial motion by MPEG-4 face standard based on facial geometry model and Facial Action Coding System (FACS) based on facial muscle model. This approach is more eclectic, since it takes account of the merits of motion simulation method based on geometry and physics and makes up for their shortcomings.

Manuscript received July 31, 2013; revised October 24, 2013.

II. Candide Model

Candide model is a parameterized model based on Facial Action Coding System (FACS) [3]. It is only defined by a small amount of triangle (about 100), so it is easy to calculate and reconstruct quickly. The shape of the Candide model is controlled by some global and local Action Unit (AU).The global AUs control rotation around three axes, and the local AUs control a variety of facial underlying behavior to generate different expressions. The initial Candide model contains 75 vertices, 100 triangles; Candide-1 model, the first officially published version, contains 79 vertices and 108 triangles; then Bill Welsh's [4] Candide-2 model added feature points of hair, teeth and shoulder based on the former, contains 160 vertices and 238 triangles. For representing the eyes and the mouth in detail and considering the application background of video feature tracking, Ahlberg gave Candide-3 model. The model simplified some unnecessary details contains 113 vertices and 168 triangles (Fig. 1). Candide-3 model can directly establish various path of 2D facial template.

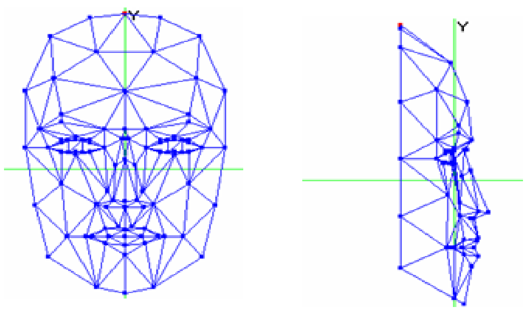

Figure 1. Original candide-3 face model.

doi: 10.12720/ijoee.2.2.99-102

Candide-3 model has its unique advantages. It is compatible with both MPEG-4 facial standard and Facial Action Coding System (FACS). It defines almost all FAPs of MPEG-4 and part AU of FACS (They can be collectively referred to as expression motion parameters within the model).Therefore, facial expression simulation animation generated by FAPs of MPEG-4 facial expression animation standard can also be achieved by expression motion parameters of Candide-3 model.

III. The Contrast between Candide Model and MPEG-4 and FACS

A. Candide Model and MPEG-4 Facial Standard

MPEG-4, a new generation of multimedia compression standard, is object-based coding and allows the object in the scene to be encoded independently [5]-[7], thus it is convenient to encode and decode.MPEG-4 takes face as a special object, and establishes a series of standards for facial animation.MPEG-4 uses two parameter sets to define the face: FDP and FAP.FDP is used to define geometry information and texture information of the face, which describes appearance and feature of face. In FDP, MPEG-4 defines 84 feature points which can control the deformation of face. MPEG-4 totally defines 68 FAPs which is divided into 10 groups. FAP represents a complete set of basic facial motion. A complete facial expression and motion can be composed of multiple FAPs. Different combinations of FAP can form a variety of facial expressions and motions. Most of the facial animation parameters (FAP) of MPEG-4 have been achieved in Candide-3 model and most of facial definition parameters (FDP) of MPEG-4 facial animation standard have corresponding vertices, but few feature points have no corresponding vertices in Candide -3, following that:

- Candide-3 model does not have the tongue, so sixth FDPs of MPEG-4 facial animation standard have no corresponding vertices.
- Candide-3 model does not represent the whole head model, so the top of the head (FDPs7.1), the back of the head (FDPs11.6) of MPEG-4 facial animation standard have no corresponding vertices.
- Candide-3 model does not have the teeth, so FDPs9.9, FDPs9.11 of MPEG-4 facial animation standard have no corresponding vertices.
- Candide-3 model does not have the ear, so FDPs10.1, FDPs10.6 of MPEG-4 facial animation standard have no corresponding vertices.
- Candide-3 model does not have the hair, so the hair of MPEG-4 facial animation standard has no corresponding vertices.
- FAP 43-FAP 47(tongue),FAP 65-FAP 68(ear) of MPEG-4 facial animation standard have not been achieved in Candide-3 model; moreover FAP 48-FAP 50 which represent head deflection in Candide-3 model are no longer regarded as animation, but replaced with global parameters.

B. Candied Model and FACS

Facial Action Coding System (FACS), an integrated system which distinguishes facial motion as much as possible, established by Elman etc, is based on in-depth study of facial muscles motion and the control of expression. It is a description method of facial motion based on facial action anatomy. It is by far the most detailed, the most delicate facial motion measuring technique, and it can measure and record all facial behavior which can be observed.

TABLE I. Part of Motion Units in FACS.

AU	Description	Facial muscle
1	Inner Brow Raiser	Frontalis, pars medialis
2	Outer Brow Raiser	Frontalis, pars lateralis
4	Brow Lowerer	Corrugator supercilii, Depressor supercilii
5	Upper Lid Raiser	Levator palpebrae superioris
6	Cheek Raiser	Orbicularis oculi, pars orbitalis
9	Nose Wrinkler	Levator labii superioris alaquae nasi
15	Lip Corner Depressor	Depressor anguli oris (a.k.a. Triangularis)
20	Lip Stretcher	Risorius w/ platysma
23	Lip Tightener	Orbicularis oris
26	Jaw Drop	Masseter, relaxed Temporalis and internal Pterygoid
27	Mouth Stretch	Pterygoids, Digastric
28	Lip Suck	Orbicularis oris
41	Lid Droop	Relaxation of Levator palpebrae superioris
42	Slit	Orbicularis oculi

TABLE II. Combination of the Basic Expression in FACS.

Basic Expression	Involved Action Units
Surprise	AU1, 2, 5, 15, 16, 20, 26
Fear	AU1, 2, 4, 5, 15, 20, 26
Digust	AU2, 4, 9, 15, 17
Anger	AU2, 4, 7, 9, 10, 20, 26
Happiness	AU1, 6, 12, 14
Sadness	AU1, 4, 15, 23

FACS divides face into three parts: forehead--eyebrow, eye-eyelid, nose, cheek--mouth, lip. According to facial muscle anatomy, they use photographic and video to record reaction by stimulating pieces of muscular tissue. They match the photos of each muscular activity when 6 kinds of emotions--happiness, surprise, disgust, anger, fear, sadness occurs with each emotion, then they identify that which change of each part of face is caused by which muscular movement for determining facial muscular motion combination standard of each emotion. Ekman totally find 14 kinds of single muscular activity units and 19 kinds of compound muscular activity units, and then he details these muscular activities and records them by video. Although the system has many defects, the method is still widely used to denote facial expression in the perception field of human and machine. Table I lists part of action units in FACS.

FACS contains 44 action units (AUs), combination of different action units produce different facial expressions. For example, combination of action unit AU1 (Inner Brow Raiser), AU4 (Brow Lower), AU15 (Lip Corner Depressor), AU23 (Lip Tightener) can create a sad expression. Table II contains basic expressions combined by part of action units.

FACS uses the movement of the feature points to combine and analyze expression. The method combines expression by using expression parameters to deform the shape of wire frame model. The question is how to define expression parameters. The human expression is formed by the stretching of the facial muscles, however, the stretching of the facial muscles cannot be observed from the outside, so direct analysis of such parameters is almost impossible. Hence, we can only define facial expression parameters by facial feature points which can be seen from the outside.

Since definition of feature points of Candide-3 model based on FACS, and part of action units of FACS (denoted by AUV in Candide model) have been achieved in Candide-3, as shown in Table III.

TABLE III. DEFINED AU NUMBERS IN CANDIDE-3

Number	Name	AU(set)	Effect
AUV0	Upper lip raiser	AU10	Raise upper lip
AUV11	Jaw drop	AU26, 27	Drop jaw
AUV2	Lip stretcher	AU20	Stretch lip
AUV3	Brow lowerer	AU4	Drop brow
AUV14	Lip corner depressor	AU13, 15	Depress lip corner
AUV5	Outer brow raiser	AU2	Raise outer brow
AUV6	Eyes closed	AU42, 43, 44, 45	Close eyes
AUV7	Lid tightener	AU7	Tight lid
AUV8	Nose wrinkler	AU9	Wrinkle nose
AUV9	Lip presser	AU23, 24	Press lip
AUV10	Upper lid raiser	AU5	Raise upper lid

Since many AUs have not been achieved, and AU of FACS is similar to FAP of MPEG-4 facial animation standard, we need to supplement and improve other AUs which have not been achieved in order to achieve real facial expression motion.

IV. Improvement of the Model

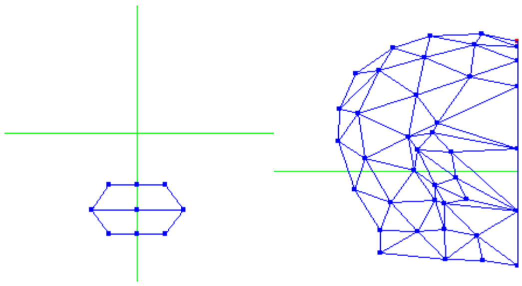

Figure 2. Grid of teeth and the back of the head.

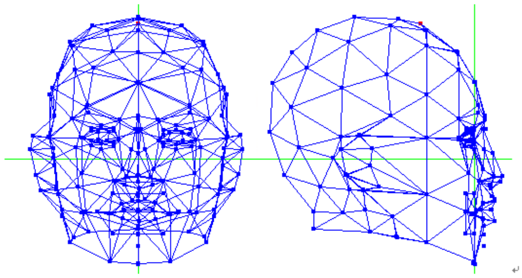

Figure 3. Improved Candide facial model.

In order to be able to represent the characteristics of facial expression motion more accurately, Candide-3 model has been improved by increasing the part of the back of the head and teeth (Fig. 2). New model contains 182 vertices, 316 triangles, as shown in Fig. 3.

In addition, in order to meet the need of expression simulation, we modify the subtle expression motion parameters (AU and FAP) defined in Candide model.

V. Experimental Results

Through the experiment, the simulation results of the head model, teeth and facial expression of Candide before and after improved are compared, as shown in Fig. 4-Fig. 7.

Figure 4. Laugh expression simulation sequence of Candide before improved.

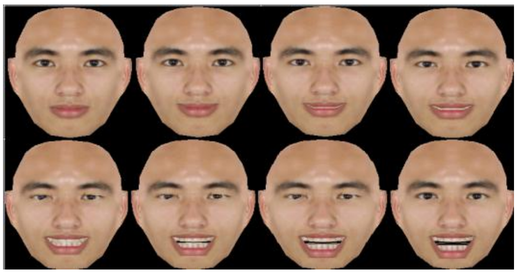

Figure 5. Laugh expression simulation sequence of Candide after improved.

Figure 6. Painful expression simulation sequence of Candide before improved.

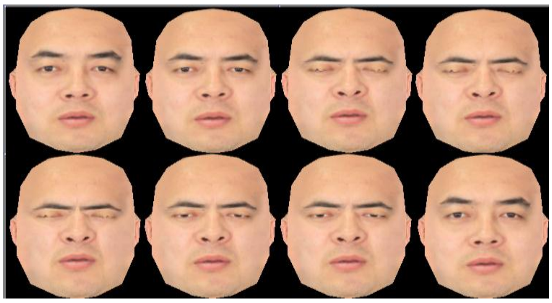

Figure 7. Painful expression simulation sequence of Candide after improved.

VI. Summary and Scope

In this paper, through comparing Candide with MPEG-4 and FACS, we increase part of the back of the head and the teeth based on the original Candide. We modify the subtle expression movement parameters(AU and FAP), and we increase feature points, motion vector, 13 action units of the region involved in chin opener, lip corner depressor, nostrils riser, left and right nose riser, left and right eye opener, upper teeth raiser, lower teeth lowerer. Through improvement of parameterized Candide model and combination control of facial geometry model and facial muscle model FACS, we do experiment for facial expression motion characterization. The experiment results indicate that the model before and after improved have a big difference. The model after improved has the back of the head and the teeth, the model before improved (in Fig. 4) does not have the teeth, and the model after improved (in Fig. 5) has the teeth and the head. After improved, the shape of face is more complete, the face and the head profile are more smooth, the motion expression of chin, mouth corner, nostril, nose, eye when face laugh and pain are more rich ,nature, true, we achieved good facial expression simulation results. However, the model is a triangular mesh model, feature points increase too much influence calculation and reconstruction speed directly, so the facial contour curves are still not perfect. Next, we will further optimize and improve facial expression parameterized model.

Acknowledgment

This work supported by the National Natural Science Foundation of China (Grant No. 61174161, the second collaborator) and the 2011 Fujian Provincial Science and Technology Foundation (Grant No. 2011H0031).

References

[1] S. Morishima, “Real-time talking head driven by voice and its application to communication and entertainment,” in *Proc. International Conference on Auditory-Visual Speech Processing*, vol. 13, 1998, pp. 64-68.

[2] T. Capin, I. Pandzic, N. M. Thalmann, and D. Thalmann, *Avatars in Networked Virtual Environments*, John Wiley &Sons, 1999, pp. 47-48.

[3] Q. S. Zhang and Z. C. Liu, “Geometry-driven photorealistic facial expression synthesis proc ACM symp,” *Computer Animation*, San Diego, CA, vol. 28, no. 9, pp. 426-436, July 2003.

[4] K. Kahler, J. Haber, and H. P. Seidel, “Reanimating the dead: Reconstruction of expreesive faces form skull data,” in *Proc. ACM Transactions on Graphics, SIGGRAPH Conference Proceedings*, vol. 22, 2003, pp. 554-561.

[5] C. Kuo, “Reconstruction techniques and applications based on the facial regional control model,” *Journal of Zhejiang University*, 2006.

[6] Q. S. Zhang and G. L. Chen, “A realistic 3d facial animation,” *Journal of Software*, vol. 22, pp. 643-650, 2003.

[7] F. I. Parke, “A parametrized model for facial animation,” *IEEE Computer Graphics and Applications*, vol. 2, no. 9, pp. 61-70, February 1982.

Baoying Ma was born in Xinjiang, China, in 1965. She is currently an Associate Professor with Changji College, Xinjiang, China. She has devoted her study life to 3D Facial Simulation for several years.

Junfeng Yao was born in Shanxi, China, in 1973. He received the B.E., M.E. and Ph.D. degrees in data modeling engineering from Central South University, Hunan, China, in 1995, 1998, and 2001, respectively, and the Post Degree in Simulation and control engineering from Tsinghua University, Beijing, China, in 2003. From 2004 to 2005 he was a Visiting Researcher with University of SUSSEX, Brighton, UK, and between 2009 and 2010 he was a Visiting Researcher with Southern Polytechnic State University, Ga, USA. He is currently a Professor with Xiamen University, Xiamen, China, and is engaged in research on virtual reality, simulation and optimization of complex industry proces.

Ruohe Yan is a PhD student in Cognitive Science Department of Xiamen University. He has devoted his study life to 3D Facial Simulation for 6 years.

Binbin Zhang was born in Gansu, China, in 1986. She received the B.E., M.E. of Software from Xiamen University. She has devoted his study life to the area of Virtual Reality for 6 years.