

## **Lip-Sync in Human Face Animation based on Video Analysis and Spline Models**

Author

Tang, SS, Liew, AWC, Yan, H

Published

2004

Conference Title

10TH INTERNATIONAL MULTIMEDIA MODELLING CONFERENCE, PROCEEDINGS

DOI

[10.1109/MULMM.2004.1264973](https://doi.org/10.1109/MULMM.2004.1264973)

Rights statement

© 2004 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Downloaded from

<http://hdl.handle.net/10072/22676>

Griffith Research Online

<https://research-repository.griffith.edu.au>

# Lip-Sync in Human Face Animation Based on Video Analysis and Spline Models

Sy-sen Tang<sup>1</sup>, Alan Wee-Chung Liew<sup>1</sup> and Hong Yan<sup>1,2</sup>

<sup>1</sup>*Department of Computer Engineering and Information Technology  
City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong.*

<sup>2</sup>*School of Electrical and Information Engineering  
University of Sydney, NSW 2006, Australia.*

*Email: {itjohn, itwcliew, ityan}@cityu.edu.hk*

## Abstract

*Human facial animation is an interesting and difficult problem in computer graphics. In this paper, a novel B-spline (NURBS) muscle system is proposed to simulate a 3D facial expression and talking animation. The system gets the lip shape parameters from the video, which captures a real person's lip movement, to control the proper muscles to form different phonemes. The muscles are constructed by the non-uniform rational B-spline curves, which are based on anatomical knowledge. By using different number of control points on the muscles, more detailed facial expression and mouth shapes can be simulated. We demonstrate the flexibility of our model by simulating different emotions and lip-sync to a video with a talking head using the automatically extracted lip parameters.*

## 1. Introduction

With the fast increase in computing power, 3D modeling is very common nowadays. The 3D facial animation technique has become more and more important and has found many applications in different areas such as entertainment and teleconferencing [1, 2]. There are mainly five approaches in facial modeling and animation by geometric manipulations [3], i.e., interpolation-based, parameterization-based, pseudo-muscle modeling, performance-based and physics-based modeling. Interpolation techniques use in-between method to generate the frames between two key-frames [4]. It is fast and relatively easy to generate primitive facial animation but difficult to create a wide range of realistic facial configurations. Parameterization techniques use a set of independent parameter values to

specify any possible facial expressions [5, 6]. However, there might be conflict between parameters when they affect the same vertices. The pseudo-muscle modeling based method applies geometric deformation like free form deformation (FFD) animation [2]. As it does not follow the actual facial anatomy, it is not suitable for facial animation that requires high fidelity. The performance-based [16] method captures real human facial movements and uses that data for animation. The physics-based muscle modeling method tries to simulate real human facial muscles for animation. Since it is based on human anatomy, it is the closest to realistic human facial animation.

The physics-based muscle approach can be further divided into three categories [3]: mass spring systems, vector muscles and layered spring meshes. The mass spring method [7] uses the spring network to simulate forces applied on face. The layered spring mesh [8] extends it into three connected mesh layers. The vector muscle method [12, 15] simply defines a vector field to act as muscle to attract the mesh vertex. It consumes less computation power than the mass spring and layered mass spring systems. However, it only considers the muscle effect on skin and cannot simulate the fatty tissue. Recently, Huang and Yan [9] have presented a NURBS curves based method which separates the human face into five facial units and uses the NURBS curves to control them. It uses fuzzy sets to associate the vertices with the NURBS curves. However, it is very hard to locate a control polygon on the mesh model as the control polygons are not based on anatomical knowledge but are just used to roughly separate the face into several parts.

In this paper, a realistic facial expression animation system is developed by using a NURBS-based vector muscle system. The proposed system can simulate linear muscle, sphincter muscle and the non-linear part of the

face like fatty tissue. The system allows more control on linear muscles to simulate a particular expression more realistically by modifying the weights of different control points. As it uses NURBS to simulate muscles, the control points can be put on the surface of face mesh based on the facial anatomy. This makes it easier to locate facial muscles. Through the control points, the curve can be formed under the mesh like the muscle under the face. By changing the weights of the control points, the knots will form a motion vector to control the movement of the mesh vertex within certain region. The number of control points can be determined by the complexity of different parts of face. To animate the mouth, a lip contour extraction technique by Liew et al. [14] is employed to extract the lip shape parameters from the video. These parameters will be used to control virtual model to form different phonemes. In Section 2, a brief review of the lip contour extraction technique and NURBS curves are given. Section 3 describes the proposed NURBS-based system for facial expression animation. Simulation results are presented in Section 4, followed by conclusions in Section 5.

## 2. Brief Review

This section provides a brief introduction of the lip contour extraction technique and the definition of the NURBS for the NUBRS muscle animation approach.

### 2.1 Lip Contour Extraction Technique

In Liew's lip contour extraction system, it uses color video images to analyze lip movement and a geometric deformable model to describe the lip shape (see Fig.1). This model is formed by two curves. The shape of the model is pre-constrained to the expected lip shape based on prior knowledge. The lip equations are defined as follows,

$$y_1 = h_1 \left( \left( \frac{x - sy_1}{w} \right)^2 \right)^{1+\delta^2} - h_1 \quad (1)$$

$$y_2 = \frac{-h_2}{(w - x_{off})^2} (x - sy_2 - x_{off})^2 + h_2 \quad (2)$$

for  $x \in [-w, w]$  with the origin at (0,0). The parameter  $s$  describes the skewness of the lip shape and exponent  $\delta$  describes the deviation of the curve from a quadratic. Liew's system aims to find an optimum partition of a given lip image into lip and non-lip regions based on the

pixel's intensity and color. Since this model is defined by a small set of parameters with clear physical interpretation, it is very suitable for controlling the NURBS muscles. A NUBRS muscle is controlled by a few control points. The weighting of the control point can be made to correspond to these parameters.

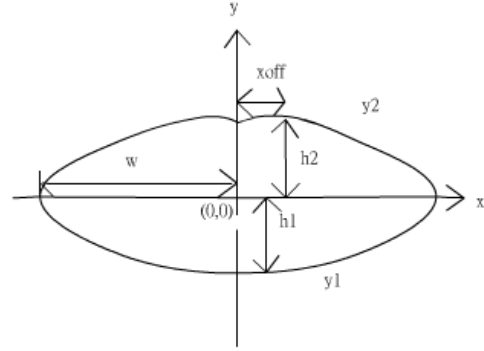


Figure 1. A geometric lip model

### 2.2 Definition of NURBS

A degree  $n$  NURBS curve [10] is defined as

$$e(u) = \frac{\sum_{i=0}^n B_{i,n}(u) \omega_i P_i}{\sum_{i=0}^n B_{i,n}(u) \omega_i} \quad 0 \leq u \leq 1 \quad (3)$$

where  $e(u)$  is the knot of the curve. In our approach, it will be used for calculating the motion vector as described in 3.1.  $\omega_i$  is the weight,  $P_i$  is the control point and  $B_{i,n}(u)$  is the blending function defined as follow,

$$B_{i,n}(u) = \frac{n!}{i!(n-i)!} u^i (1-u)^{n-i} \quad (4)$$

One important property of the NURBS curve is the convex hull property, i.e.

$$\sum_{i=0}^n B_{i,n}(u) \omega_i = 1 \quad (5)$$

This property ensures that the polynomial can smoothly follows the control points without erratic oscillations. Another important property is endpoint interpolation, which means that the curve always passes through the first and last control points [10, 11].

### 3. NURBS Muscle System

In this system, three to five control points will form a NUBRS muscle. By changing the weight of the control points, the NUBRS muscle contracts like real muscle. Modification of control point's weight can force the knots of NURBS curve to move, which will then influences the nodes on the 3D mesh. Since NURBS curve can have several control points, it can be used to control the face skin and tissue to move to any desired position.

#### 3.1 NURBS Linear Muscle and NURBS Sphincter Muscle

The control points of the NURBS model are classified into two groups. They are reference control point and current control point (see Fig.2). Reference control point is used to relate the knot vector and the node of the mesh inside the influence region. Current control point is the control point whose weight is currently being modified. For the NBURS linear muscle, the current and reference control points are the same as illustrated in Fig.3.

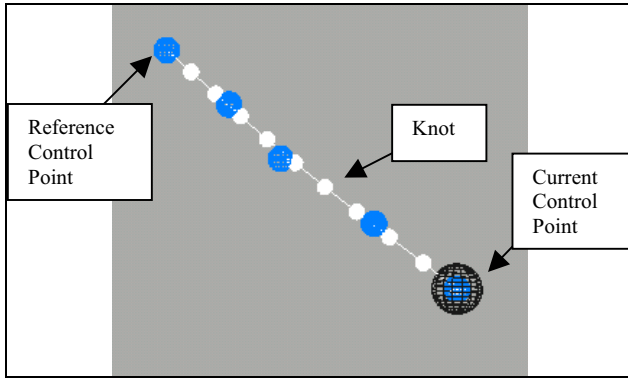


Figure 2. A NURBS curve

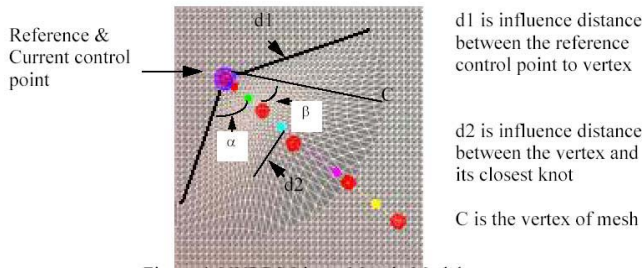


Figure 3. NURBS linear muscle model

Fig. 4 shows the relationship between the reference point, the knot and the vertex.  $B$  is the average of the knots' position before movement and  $B'$  is average of the

knots' position after movement.  $\overrightarrow{B'B}$  is a vector formed by the knots movement and can be obtained by

$$\overrightarrow{B'B} = \left( \sum_{i=0}^n (e_i(u)' - e_i(u)) \right) / (n+1) \quad (6)$$

where  $e_i(u)$  is the node before movement,  $e_i(u)'$  is the node after the movement.  $C$  is the vertex of the mesh which is within the influence region. A virtual vector is used to control the vertex direction. Assume  $C$  is repositioned to  $C'$ , this will form a vector  $\overrightarrow{C'C}$ . We can use vector  $\overrightarrow{B'B}$  to find out vector  $\overrightarrow{C'C}$  by the following rules,

$$\angle BAC = \angle B'AC' \quad (7)$$

$$\angle ABC = \angle AB'C' \quad (8)$$

$\overrightarrow{C'C}$  is the vector which will be used to calculate the new position of vertex  $C$ . Finally, the displacement of the vertex, i.e.,  $C''$ , can be calculated by adding a displacement vector [8],

$$C'' = C + D \cdot R \cdot \overrightarrow{C'C} \quad (9)$$

where  $K$  is contraction factor,  $R$  is an angular displacement factor,

$$R = \frac{\cos \beta}{\cos \alpha} - 1 \quad (10)$$

and  $D$  is a radial displacement factor given by

$$D = \left( 1 - \left( \frac{\|AC\|}{d1} \right) \right) \cdot \sin \left( \frac{\|AC\|}{d1} \cdot \frac{\pi}{2} \right) \quad (11)$$

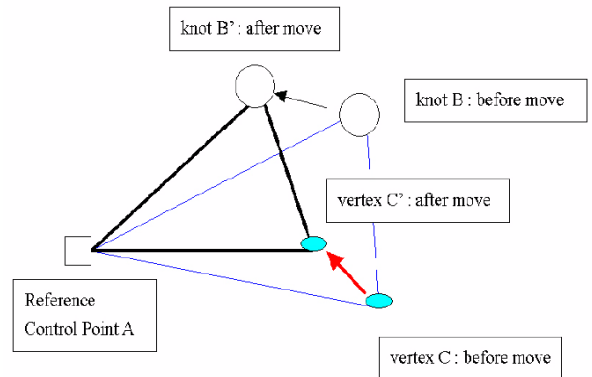


Figure 4. Relationship between control point, knot and vertex

As the NUBRS muscle can perform non-linear movement, the motion vectors are calculated independent for each mesh node. This ensures the smooth movement of the mesh. Moreover, it increases the controllability of the muscle system.

Since the sphincter muscle does not have angular displacement, it can be formed by multiplying two radial displacement parameters as shown in Fig. 5,

$$D2 = \cos\left(\left(\frac{a}{d2}\right) \cdot \frac{\pi}{2}\right) \times \cos\left(\left(\frac{b}{d1}\right) \cdot \frac{\pi}{2}\right) \quad (12)$$

where  $a$  is the length between knot and vertex,  $b$  is the length between control point and vertex.

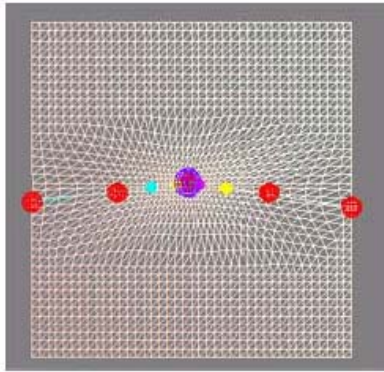


Figure 5. NURBS sphincter muscle model

### 3.2 Usage of NURBS Muscle

The linear NURBS muscles are mainly responsible for facial expression. The sphincter muscle can be used for the deformation of the mouth region. However, not all phoneme shapes can be formed. To control the mouth to form more phoneme shapes, two more Linear NURBS curves have to be added. Fig. 6 illustrates how the NURBS curves can control the mouth shape. The reference control point on the curve is fixed to ensure that the mesh would not form some unexpected mouth shape. The current control can be any one of the control points. With such arrangement, the muscle can control the lip to form a desired phoneme shape.

Since human face has fatty tissue under the skin, when we use muscle to control mesh skin we have to consider the effect on fatty tissue. To create a more realistic expression, our system provides a way to simulate the fatty tissue on a face as shown in Fig. 7. This is achieved by adding additional control points between two end control points. The control points will drag the

mesh slightly up, simulating the fatty tissue. As the control point is placed on the surface of a face, the curve will follow the shape of the face.

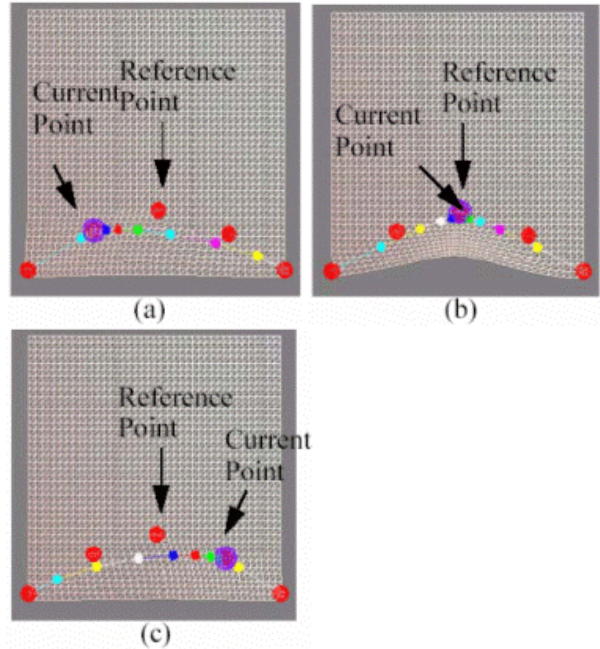


Figure 6. NURBS muscles with same reference point but different control points

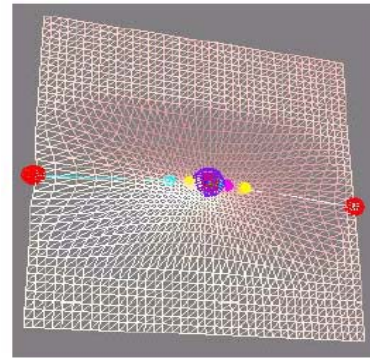


Figure 7. Fatty tissue simulation

### 3.3 Control the Lip Shape by Lip Contour Parameters

In Liew's [14] lip model, five parameters are used to control the lip contour. They are given by

$$p = \{w, h_1, h_2, x_{off}, \delta\}.$$

Using  $p$  to control the corresponding NURBS muscles can result in realistic lip movement. In our system, one

parameter could drive more than one muscle and one muscle could be driven by more than one parameter. Following are the correspondences between muscles and parameters (see Fig.9 for the muscle set we used):

Zygomatic Major and Angular Depressor :  $w$

Mentalis and Orbicularis Oris :  $h1$

Levator Labii and Orbicularis Oris :  $h2$

Orbicularis Oris :  $x_{off}$

Mentalis :  $\delta$

The parameters control the weights of different muscle's control points to activate the muscle to form the desire lip shape. The lip shape parameters are linearly linking to the muscle's weight,

$$\text{Weight of NURBS curves} = S * p \quad (13)$$

where  $S$  is a scale parameter. The scale  $S$  is determined experimentally and is different in different part of the face model. For example, the width parameter  $w$  directly controls the Zygomatic Major and Angular Depressor muscles since the function of these two muscles is to stretch the lip to make it wider. In practice, these two muscles are controlled by the same weight using equation (13). It ensures that the mouth is only stretched wider but not to other direction.

Fig.8. compares the video lip extraction result to the 3D simulation result of uttering the digit "four" in Cantonese. The results from lip extraction control the virtual model. The lip tracking system gives meaningful parameters to control the NUBRS muscle to create realistic lip shape.

#### 4. Results and Discussions

Twenty-two muscles are implemented based on the facial action coding system (FACS) [13]. FACS use Action Unit (AU) to describe the basic actions of human face. The actions are based on the facial anatomy and each action represents a single muscle. Through different combinations of AU, we can create different facial expressions. Fig. 9 illustrates the mesh with NURBS muscles.

A NURBS facial modeling system is implemented to test the described method. The modeling system is developed by VC++/OpenGL. It runs on a P4 1.9GHz PC with 3D graphics interface. We create a variety of expressions and mouth shapes using NURBS-based linear and sphincter muscles. Each expression simulation is according to the Action Unit of the FACS. Fig. 10 shows the six basic facial expressions, in which only linear muscles are used in the expressions.

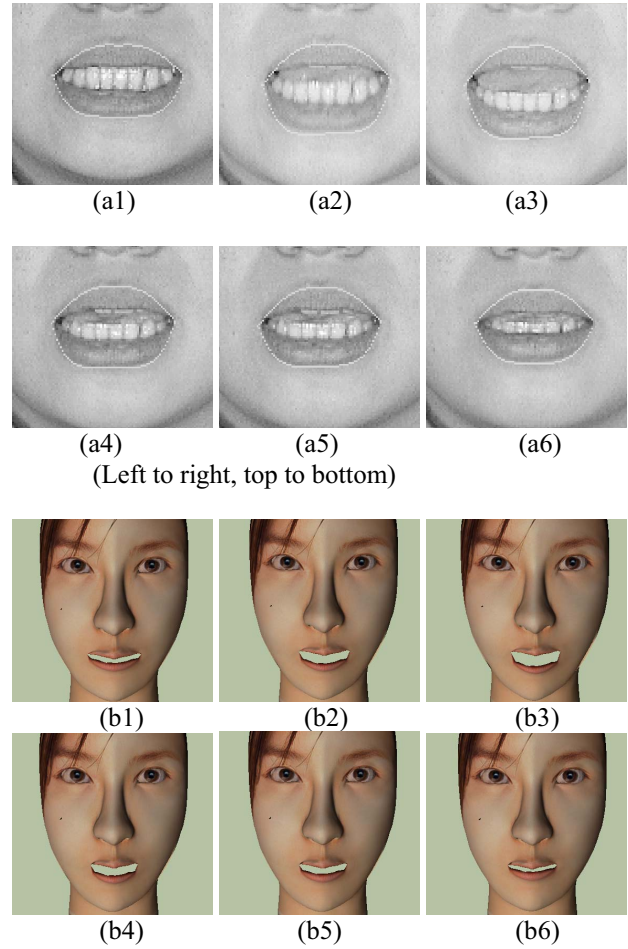


Figure 8. Comparing real video lip shape to the animated lip sequence when uttering the digit "four" in Cantonese (Left to right, top to bottom).

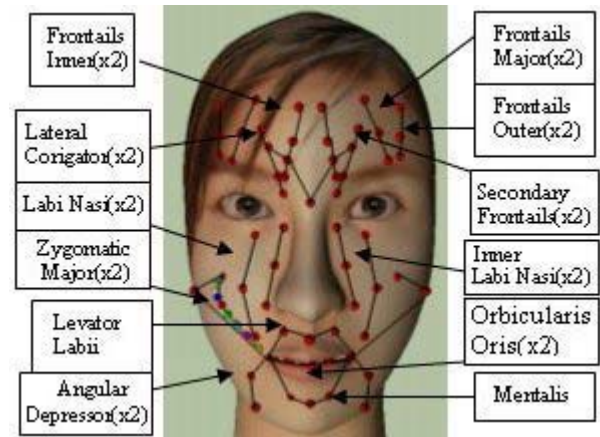


Figure 9. A set of muscles

The muscles around the mouth are responsible for changing the mouth shapes. The fatty tissues surrounding the cheek and the mouth may make the expression look more realistic especially when animating the model. Fig. 11 shows a more realistic simulation of linear muscle with fatty tissues. Fig.12 shows two more expressions.

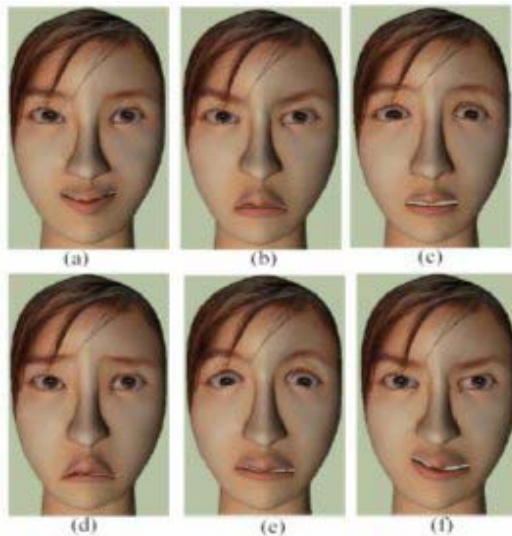


Figure 10. The six basic expressions: (a) happiness, (b) anger, (c) surprise, (d) sadness, (e) fear, (f) disgust

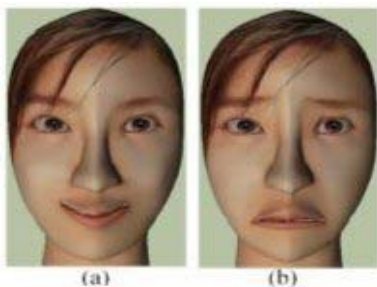


Figure 11. (a) happiness, (b) sadness, with fatty tissue modeling

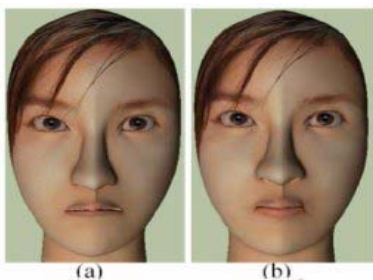


Figure 12. (a) bite lip, (b) close mouth

## 5. Conclusion

In this paper, we have presented a novel method for facial expression animation, which uses NURBS curves to create linear muscle, sphincter muscle, and control the mouth shape. Our NURBS-based muscle model can simulate more realistic facial expressions by incorporating fatty tissue simulation. The flexibility of NURBS shape makes the face more controllable. As the control points are attached to the surface of the mesh, the muscles position can be easily located. The flexible facial muscles also allow realistic simulation and animation of talking mouth when animated with a set of physically meaningful lip shape parameters extracted from video of a talking person. When combine with the automatic lip tracking technique we developed, the system can perform accurate real time lip-sync.

## Acknowledgment

This work reuses part of code source by Keith Waters ([waters@crl.dec.com](mailto:waters@crl.dec.com)). The project is supported by the Center for Media Technology at City University of Hong Kong.

## References

- [1] G. Breton, C. Bouville and D. Pele, "FaceEngine: A 3D Facial Animation Engine for Real Time Applications", Proceedings of the 2001 Conference on 3D Technologies for World Wide Web, February 2001.
- [2] S. Coquillart, "Extended Free-Form Deformation: A Sculpturing Tool for 3D Geometric Modeling", Computer Graphics, 1990, vol. 24, pp. 187-193.
- [3] J. Noh and U. Neumann, "A Survey of Facial Modeling and Animation Techniques", USC Technical Report, 1998, pp. 99-105.
- [4] F. Pighin, J. Hecker, D. Lishchinski, R. Szeliski and D. H. Salesin, "Synthesizing Realistic Facial Expression from Photographs", SigGraph proceedings, 1998, pp. 75-84.
- [5] F. I. Parke, "Parameterized Models for Facial Animation", IEEE Computer Graphics and Applications, 1982, vol. 2(9), pp.61-68.
- [6] J. Noh, D. Fidaleo and U. Neumann, "Animated Deformations with Radial Basis Functions", ACM Symposium on Virtual Reality Software and Technology (VRST 2000), October 2000, pp. 166-174.
- [7] Y. Zhang, Edmond C. Prakash and Eric Suang, "Real-time Physically-based Facial Expression Animation Using Mass-spring System", Computer Graphics International 2001. Proceedings, 2001, pp. 347-350.
- [8] D. Terzopoulos and K. Waters, "Physically-based facial modeling, analysis, and animation." J. of Visualization and Computer Animation, March, 1990, vol. 1(4), pp. 73-80.
- [9] D. Huang and H. Yan, "Modeling and animation of human expression using NURBS curves based on facial anatomy", Signal Processing Image Communication, 2002, pp. 457-465.

- [10] L. Piegl and W. Tiller, *The NURBS Book*, 2nd Ed., Springer, Berlin, 1997.
- [11] D. Hearn and M. Pauline Baker, *Computer Graphics C version*, 2nd Ed., Prentice-Hall International, Inc., 1998.
- [12] K. Waters, "A Muscle Model for Animation Three-dimensional Facial Expression", ACM SIGGRAPH Computer Graphics, Proceedings of the 14th Annual Conference on Computer Graphics and Interactive techniques, August 1987, vol. 21, Issue 4.
- [13] P. Ekman and W. V. Friesen, *Facial Action Coding System*. Consulting Psychologists Press, Palo Alto, CA, 1978.
- [14] A.W.C. Liew, S.H. Leung and W.H. Lau, "Lip Contour Extraction from Color Images Using a Deformable Model", Pattern Recognition, vol. 35, Dec 2002, pp. 2949-2962.
- [15] G. Breton, C. Bouville and D. Pele, "FaceEngine a 3D facial animation engine for real time applications", International Conference on 3D Web Technology, Feb, 2001, ACM SIGGRAPH.
- [16] I. Buck, A. Finkelstein, C. Jacobs, A. Klein, D. H. Salesin, J. Seims, R. Szeliski and K. Toyama, "Performance-driven hand-drawn animation", Symposium on Non-photorealistic animation and rendering, June 2000. ACM SIGGRAPH