Double Spline Muscle Models for Human Face Animation based on Video Analysis

Author
Tang, Sy-sen, Liew, Alan Wee-Chung, Yan, Hong

Published
2004

Conference Title
Proceedings of the 2004 International Symposium on Intelligent Multimedia, Video & Speech Processing

DOI
https://doi.org/10.1109/ISIMP.2004.1434046

Copyright Statement
Copyright 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/10072/24864

Link to published version
http://ieeexplore.ieee.org/servlet/opac?punumber=9799
Double Spline Muscle Models for Human Face Animation
Based on Video Analysis
Sy-seu Tang, Alan Wei-Chung Liew and Hong Yan

1Department of Computer Engineering and Information Technology
City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong
2School of Electrical and Information Engineering
University of Sydney, NSW 2006, Australia
Email: {tijohn, ttwcliew, ityan}@cityu.edu.hk

ABSTRACT
Facial animation has been widely used in different areas such as entertainment and teleconferencing. Using video data to control facial animation is an interesting and difficult problem. In this paper, a novel double B-spline (NURBS) muscle system is proposed to simulate a 3D facial expression and talking animation based on video image. This system can simulate many different expressions of elastic human face. Each muscle is constructed by two non-uniform rational B-spline curves, which better approximate the real muscle system. There are 14 features points to represent facial expressions and we apply a lip contour extraction technique to determine lip shapes.

1. INTRODUCTION
There are mainly five approaches in facial modeling and animation by geometric manipulations [2], i.e., interpolation-based, parameterization-based, pseudo-muscle modeling, performance-based and physics-based modeling. The physics-based muscle modeling method tries to simulate real human facial muscles for animation. Since it is based on human anatomy, it is closest to realistic human facial animation.

The vector muscle model [1, 6] is a physics-based method. It defines a vector field to act as muscle to attract the mesh vertex and a NURBS curve basis to simulate muscles [10,11,12]. However, the vector muscle model cannot simulate the elastic effect of the face well.

In this paper, we present double spline muscle system to perform a realistic facial expression animation. It can simulate elastic human face tissues. Our system is a vision-based system, which does not require costly motion capture devices. It tracks facial expressions and lip shape parameters from image sequences, which control the contraction of the spline muscle in facial animation.

The paper is organized as follows. Section 2 describes the double spline muscle system. Section 3 shows the video tracking process and the usage of the input parameters. Simulation results are presented in Section 4, followed by conclusions in Section 5.

2. DOUBLE SPLINE MUSCLE MODEL
In our model, each muscle is formed by two splines. This is much closer to real human muscles, which control the face skin and fat tissue. One spline will be placed on the model surface which simulates the muscle control on the skin. The other spline is placed under the model surface to simulate the muscle control on other facial tissues, such as the fat tissue.

2.1 Definition of NURBS
A degree $n$ NURBS curve $[4]$ is defined as

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

(1)

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. $w_i$ is the weight, $P_i$ is the control point and $B_{i,n}(u)$ is the blending function defined as follow,

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

(2)

2.2. Spline Muscle
Spline muscle system uses movement of the NURBS' nodes to form a motion vector. Changing the weights of the control points can make the nodes move to a desired direction. This makes the system more flexible. Fig. 1 is a single NURBS curve muscle model.
Let \( B \) be the average of the knots' position before movement and \( B' \) be average of the knots' position after movement. Then \( \overrightarrow{B'B} \) is a vector formed by the knot movement and can be obtained by

\[
\overrightarrow{B'B} = \left( \sum_{i=0}^{n} (\varepsilon_{i}(u) - \varepsilon_{i}(u')) \right) / (n+1)
\]  

(3)

where \( \varepsilon_{i}(u) \) is the node before movement and \( \varepsilon_{i}(u') \) is the node after the movement. Let \( C \) be the vertex of the mesh which is within the influence region. If \( C \) is repositioned to \( C' \), it 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,

\[
\angle BAC = \angle B'AC'
\]

(4)

\[
\angle ABC = \angle AB'C'
\]

(5)

where \( A \) is the reference control point.

Fig 2 shows the final combined muscle model. Fig 2a represents the spline muscle in normal state. Fig 2b is the spline muscle in contraction state. The black arrows show the contraction direction. The muscle under the skin mainly controls the elastic effect. A very detailed facial expression can thus be simulated.

3. TRACKING PROCESS

In our method, the animation parameters obtained from the video sequence images are mapped to the appropriate spline muscles to drive the facial animation. In our system, facial expressions are described by 14 control parameters and the lip shape is tracked by a contour with 5 parameters.

3.1. Facial Expression Extraction

In our system, we track the features of the human face and transform these data into animation parameters based on the anthropometric knowledge. We use a face segmentation [5] technique to locate the face position by the human skin color. Fig 3a shows the original video capture image and Fig 3b shows the segmented face in red color. YCrCb color space is employed in the system. The skin range is set to Cr=[133 173] and Cb=[77 127]. After color segmentation, a density regularization and luminance regularization process are applied to filter the noise and refine the segmented region. Locating the face region constrains the tracking points on face.

The 14 tracking points for facial expression extraction are initialized by the user. As our system is muscle-based, we need to map the tracking data to the appropriate muscles to drive the model. Each part of the face can be tracked independently. The assignment of the tracking points is as follows:

Eyebrows: 4; Eye: 8; Nose: 2.

We use an 8x8 window to find correlation between the consecutive frames in the video. A block matching [8] technique is applied to get the cross-correlation between two images to track the feature points.

3.2. Lip Shape Extraction

We use a lip contour extraction method described in [7]. This model is formed by two curves. The lip equations are defined as follows,

\[
y_1 = h_1 \left( \frac{x - sy_1}{w} \right)^{14} - h_1
\]

(6)

\[
y_2 = \frac{-h_2}{(w - x_{off}^2)} \left( x - sy_2 \right) + h_2
\]

(7)

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.
3.2. Control Parameters

Twenty-two muscles are implemented based on the facial action coding system (FACS) [13]. Each muscle is constructed by two splines with same number of control points. Fig. 5 illustrates the mesh with NURBS muscles. As the face is considered symmetrical, we only need to consider 7 tracking parameters. The 7 parameters show in Fig.6 are: eyebrow control points 1 and 2, eye control points 1, 2, 3 and 4, and the nose control parameter 1.

Eye (parameter 2 and parameter 3): Frontails Major and Lab Nasi
Nose (parameter 1): Inner Labi Nasi
One control parameter may control more than one muscle and one muscle may be controlled by more than one parameter. Moreover, the parameters can control different control points in the same muscle. For example, the eye and nose parameters both control the Lab Nasi muscle. There are 3 control points in this muscle. The eye parameter will modify the weights of the control points 1 and 2. Nose parameter will modify the weights of control points 2 and 3. Modifying the weights of control points 1 and 3 will make that part contracted to the eye or nose. Control point 2 will make that part contract to the center. The equation is:

\[ \text{Expression Weight} = SI \cdot E \]  

(9)

where \( SI \) is a scale parameter determined experimentally and may be different for different muscles.

The lip control parameter is similar to the expression parameters. There are five parameters used to control the lip contour. They are given by

\[ p = \{ w, h_1, h_2, x_{off}, \delta \} \]

The following are the correspondences between muscles and parameters:

Zygomatic Major and Angular Depressor : w
Mentalis and Orbicularis Oris : h1
Levator Labii and Orbicularis Oris : h2
Orbicularis Oris : \( x_{off} \)
Mentalis : \( \delta \)

The lip shape parameters are linearly linking to the muscle's weight.

Weight of NURBS curves = \( S \cdot p \)  

(10)

where \( S \) is a scale parameter.

4. EXPERIMENTAL RESULTS

The double NURBS facial modeling system is developed using VC++/OpenGL. It runs on a P4 1.9GHz PC with 3D graphics interface. Based on the input video sequence, we create a variety of expressions and mouth shapes using NURBS-based muscles. Fig. 7 shows the video driven expressions. Fig. 8 compares the video lip extraction result to the 3D simulation result of uttering the digit “four” in Cantonese.
Figure 7: Example of video tracking results.

(a1) (a2) (a3)

(Left to right, top to bottom)

(a4) (a5) (a6)

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

5. CONCLUSION

In this paper, we have presented a novel method for facial expression animation, which uses double NURBS curves to simulate human face muscle, and control the mouth shape. One spline muscle control the skin movement and other one control the fat tissue effect. In order to extract the detailed lip shape, we use different approaches to track facial expressions and lip shapes. 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. Combining the video input data with this muscle model, we can simulate realistic facial expressions.

Acknowledgement: This project is supported by the Media Technology Center and a SRG Grant (Project 7001556) of CityU.

6. REFERENCE


