

# 지적 아바타 통신에서 볼과 턱 움직임을 사용한 코믹한 얼굴 표정

○李容厚\*, 金商雲\*, 青木由直\*\*

\* 明知大学校 컴퓨터工学科

\*\* 北海道大学大学院工学研究科

## A Comic Facial Expression Using Cheeks and Jaws Movements for Intelligent Avatar Communications

○Yong-Who Lee\*, Sang-Woon Kim\*, Yoshinao Aoki\*\*

\* Div. of Computer Science & Engineering, Myongji University

\*\* Graduate School of Engineering, Hokkaido University

yungwho@ce.myongji.ac.kr

### Abstract

In this paper, a method of generating the facial gesture CG animation on different avatar models is provided. At first, to edit emotional expressions efficiently, regeneration of the comic expression on different polygonal mesh models is carried out, where the movements of the cheeks and numerical methods. Experimental results show a possibility that the method could be used for intelligent avatar communications between Korea and Japan.

### 1. 서론

수화 제스처는 인터넷 사이버공간에서 개인을 상징하는 아바타(avatar)의 통신 보조 수단으로도 사용될 수 있다[1]. 특히 얼굴 표정을 가진 수화 영상은 단지 팔과 손의 움직임만을 보여주는 영상에 비해 이해하기가 쉽다. 또한 코믹 만화나 애니메이션 영화에서는 과장된 얼굴 표정이나 몸짓으로 사실적이지는 않지만 보다 확실하게 감정을 전달하고 있다[2]. 이때, 얼굴 구성 요소들에 대한 움직임을 표현하기 위해 FACS(Facial Action Coding System)의 AU(Action Unit)들의 조합을 이용한다.

한편, 실시간 수화 영상 통신을 위하여 수화 영상이나 압축 영상을 전송하는 대신에 수화 애니메이션 파라미터만을 전송하는 지적 통신방식을 이용한다. 즉 송신측에서는 수화 영상과 얼굴 영상을 생성할 수 있는 수화 관절 각 파라미터와 표정 AU만을 전송하면 수신측에서는 이 파라미터를 이용하여 다시 수화 애니메이션을 재생하는 방법이다.

얼굴 표정 생성에서 특히 AU의 6번인 “뺨을 올림”, AU의 26번인 “턱을 내림”에 대한 움직임은 자동화가

어려운 조작이지만 아바타 통신에서 얼굴 표정을 전달하기 위해서 중요한 역할을 수행한다. 따라서 본 논문에서는 구조 혹은 모양이 다른 아바타 모델에 얼굴 표정을 생성하기 위해서 코, 볼, 턱의 영역을 추출하는 방법을 제안하고, 이렇게 찾아낸 볼, 턱을 가지고 포물형 편미분식(Parabolic Partial Differential Equations)과 Runge-Kutta법을 이용하여 보다 자연스러운 표정을 생성할 수 있는 방법을 제안한다. 그리고 3차원 모델에 표정의 한계값을 지정하여 다양한 아바타 모델마다 자동적으로 표정을 생성할 수 있게 한다.

이하 2장에서는 인터넷 가상공간상에서 구현한 지적 아바타 수화 통신 시스템을 고찰한다. 제 3장에서는 얼굴 모델과 다양한 아바타 모델에서 코믹한 얼굴 표정을 위한 AU를 찾아내는 방법을 제안한다. 그리고 제 5장에서는 실험 결과를 고찰한 후, 제 6장에서 결론을 맺는다.

### 2. 지적 아바타 통신 시스템

인터넷 가상공간에서 지적 통신 방식에 의한 수화 제스처 통신 시스템은 중앙 집중적인 클라이언트-서버 구조를 가진다. 한-일간의 지적 수화 통신 시스템은 인터넷상에 구현된 구조이다[3].

예를 들어, 클라이언트 A는 한글 수화 (Korean Sign Language: KSL)를 구사하는 사이트이고, 클라이언트 B는 일본 수화 (Japanese Sign Language: JSL)를 쓰는 사이트이다. 이 때 실시간 통신을 위하여 수화 영상 대신에 수화 애니메이션을 위한 수화 제스처 및 얼굴 표정 파라미터를 전송하는 지적 통신 방식을 이용한다. 즉 클라이언트 A, B에 각각 캐릭터 모델을 준비하여 놓고 송신측에서 보낸 수화 및 표정 파라미터를 이용하여 수신측에서는 수화 영상을 재생한다.

### 3. 얼굴 표정 재생 방법

#### 3.1 얼굴 표정 생성 절차

얼굴 영상을 실시간으로 디스플레이하기 위해서는 가능한 적은 수의 제어 점으로 구성되는 모델 영상이 필요하다. 기존의 표정 생성 방법에서는 사진과 같은 사실적인 감정 표현을 위해 AU6(뺨을 올린다)와 같이 얼굴 근육의 미세한 움직임을 표현하기 위해 많은 AU의 조합이 필요하였다. 그러므로, AU추출과 표정생성을 위한 절차가 나타낸다.

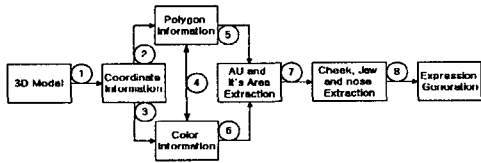


그림 1. 얼굴 표정 생성을 위한 AU 추출 절차

그림 1에서 ①~⑥까지는 3차원 모델에 포함되어 있는 정보 폴리곤과 색상 정보를 이용하여 AU 그리고 AU 주변을 추출하고, ⑦번에서 턱, 코, 볼에 대한 영역을 추출한다. 여기서 턱은 입술의 외각에 밑에서부터 기준하고, 코와 볼은 코의 끝점을 기준으로 하여 Normal Vector를 이용하며, 폴리곤들 사이에 Normal Vector를 비교하기 위해 레이블 작업을 수행한다. 마지막 ⑧번에서는 6가지 표정 중에서 턱과 볼에 대한 표정 생성에 중요한 역할을 담당한다. 때문에 각각의 움직임은 Runge-Kutta법과 포물형 편미분식을 사용한

#### 3.2 얼굴 영역 추출

296개의 버텍스를 가지고 있는 모델에서 폴리곤과 색상을 각각 따로 이용한 경우에 AU를 찾아낼 수 없으며, "폴리곤, 색상, 좌표 정보를 참조"하여 AU를 찾아낸다. 첫 번째 단계에서는 폴리곤을 이용하여 그룹핑 시켜서 6개의 그룹으로 분류한 후, 두 번째 단계에서 색상으로 다시 그룹핑하여 12개의 그룹을 생성한다. 세 번째 단계에서 좌표정보와 Normal Vector를 이용하여 턱, 코, 볼의 위치를 추출하여 총 15개의 그룹을 생성한다.

먼저 코 영역을 추출하는 방법은 얼굴 정면에서 가장 나온된 부분을 기준으로 Normal Vector의 z축값을 가지는 NVz의 값을 결정한다.( $0 \leq NVz < 0.00005$ )  
 다음으로 볼 영역을 추출하는 두 가지 절차가 있다. 첫 번째는 윗입술과 아랫입술 사이의 평균값과 그리고 위 눈꺼풀과 아래 눈꺼풀 사이의 평균값 경계를 벗어 나지 못한다. 두 번째는 코와 입술 주변(윗입술 주변, 아랫입술 주변, 턱 부분을 제외한 부분에서 코를 기준으로 NVz를 결정한다.( $0.00005 \leq NVz \leq 0.005$ )  
 마지막으로 턱 영역을 추출하는 두 가지 절차가 있다. 첫 번째는 윗입술과 아랫입술 사이의 경계를 기준으로 위의 영역을 포함하지 않는다. 두 번째는 외각 아랫입술 영역에 바로 밑에 있는 버텍스를 기준으로 NVz의 값을 결정한다.( $0 \leq NVz \leq 0.005$ )

#### 3.3 볼에 대한 처리

이전의 작업은 볼의 영역을 찾아낸다. 이후에 3차원 모델에 찾아낸 볼 영역의 움직임을 위해서 포물형 편미분식을 사용하였다. 그래서 볼의 움직임이 가능하도록 하였다.

포물형 편미분식은 다음과 같다.[4]

$$u_i^{j+1} = u_i^j + \alpha \frac{(\Delta t)}{(\Delta x)^2} (u_{i+1}^j - 2u_i^j + u_{i-1}^j)$$

여기서 입력으로  $\alpha, i, j, \Delta x$  그리고  $\Delta t$ 가 필요하다. 여기서  $\alpha$ 는 불룩한 정도를 나타내며, 값이 작을 수록 불룩한 정도가 날카로운 모양을 나타내고, 0.1에서 10까지의 실험치로 1을 지정하였다.  $i$ 는 거리  $x$ 에 대한 절점 번호를 나타내고,  $j$ 는 시간  $t$ 에 대한 반복횟수번호를 의미한다.  $\Delta x$ 는 거리구간이며,  $\Delta t$ 는 시간구간이다. 여기서  $\Delta x$ 과  $\Delta t$  식은

$$\Delta t \leq \frac{(\Delta x)^2}{2\alpha}$$

$$\Delta x = length / (n - 1) \quad \text{이다.}$$

여기서 입력으로  $length$ 는 볼의 반지름이며, 길이의 분할 횟수  $n$ 을 나타낸다.

다음으로 시간  $t=0$  ( $j=1$ )에서의  $(u_i^j, u_{i-1}^j, u_{i+1}^j)$  값으로부터  $u_i^j$ 는 분할된 영역의 정점이고,  $u_{i+1}^j$ 는 이전의 정점이며,  $u_{i-1}^j$ 는 다음의 정점이다. 마지막으로 시간  $(t + \Delta t)$ 에서  $u_i^{j+1}$  값을 계산한다 ( $i=2 \dots n$ ).

그러나 볼의 움직임을 나타내는 동작은 4가지 있으며, ①움직임은 "입술 주변 이동", ②움직임은 "볼의 중심에서부터 불룩하게 이동", ③움직임은 "볼의 중심에서부터 x와 y축으로 이동", 마지막으로 ④움직임은 "입술 주변에서부터 x와 y축으로 이동"이다. 이들 움직임 중에서 ①과 ③은 포물형 편미분식에서 생성된 정보를 참조한다.

그림 2는 4가지 동작들을 볼의 움직임에 영향을 끼쳐 변화하는 모습을 나타낸다.

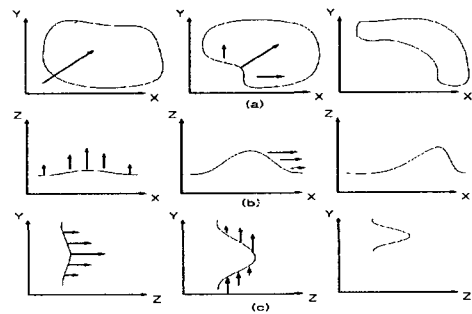


그림 2. 볼의 이동을 보기 위한 XY, XZ, YZ축에서 볼 때의 결과

그림 2(a)는 ①과 ④움직임을 얼굴의 X축과 Y축을 기준으로 표시하고, (b)는 ②와 ③움직임을 얼굴의 X축과 Z축을 기준으로 표시하고, 마지막 그림으로 (c)는

## 지적 아바타 통신에서 불과 턱 움직임을 사용한 코믹한 얼굴 표정

㉑와 ㉒ 그리고 ㉓ 움직임을 얼굴의 Y축과 Z축을 기준으로 표시한 결과이다.

### 3.4 턱에 대한 처리

AU가 움직이는 범위와 각도가 고정되어 있으므로, 다양한 모델에 AU의 움직임 범위를 제한하기 위해 입술의 한계 값은 그림 3(a)에서 정한다. 턱의 한계 값은 코와 턱의 기준축을 0°으로 하고 그림 3(b)에서 턱의 기준 축으로 Runge-Kutta법을 사용하며, 실험치로 75°을 초과하지 못하게 하였다.

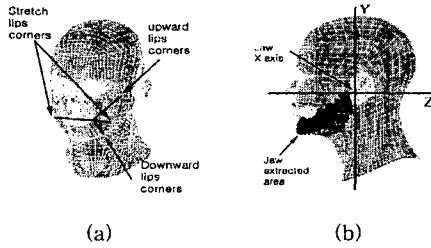


그림 3. (a) AU에 대한 움직임의 한계 값과 (b) 턱의 영역과 기준 축

그림 3(b)에서 턱의 기준축 (X, Y, Z)에서 X와 Y는 코를 중심으로 하고, Z는 얼굴의 앞면과 뒷면의 중앙을 기준으로 선정하였다.

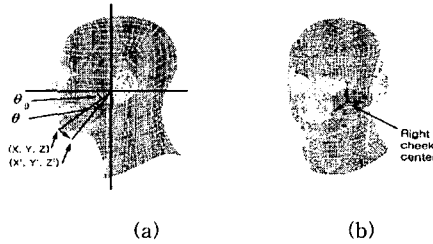


그림 4. 턱과 볼의 좌표이동 (a)은 Runge-Kutta법, (b)은 오른쪽은 포물형 편미분식을 사용

턱의 움직임을 기준축 기준으로 움직임과 함께 자연스러운 움직임을 지정할 수 있도록 턱의 움직임 사이에 보간을 할 수 있는 Runge-Kutta법을 사용하였다. 턱의 새로운 좌표 (X', Y', Z') 위치를 계산한다.

$$\begin{aligned} X' &= Xv \\ Y' &= d_n * \cos(t_n) \\ Z' &= d_n * \sin(t_n) * (\sin(ph)) \end{aligned}$$

초기각도 (θ₀)를 라디안 각도로 변환한 t<sub>n</sub>과 턱과 기준축 사이의 거리값인 d<sub>n</sub>을 사용한다. 여기서 턱이 옆으로 이동하는 것을 고려하지 않기 때문에 X'는 기존의 좌표와 동일하다. 또한 t<sub>n+1</sub>을 생성하기 위해서 Runge-Kutta법 식을 사용한다[4].

$$t_{n+1} = t_n + \frac{k_1}{6} + \frac{k_2}{3} + \frac{k_3}{3} + \frac{k_4}{6}$$

여기서, k<sub>1</sub>, k<sub>2</sub>, k<sub>3</sub> 그리고 k<sub>4</sub>는 다음과 같다.

$$\begin{aligned} k_1 &= h * f(\Omega_n, t_n) \\ k_2 &= h * f(\Omega_n + \frac{k_1}{2}, t_n + \frac{h}{2}) \\ k_3 &= h * f(\Omega_n + \frac{k_2}{2}, t_n + \frac{h}{2}) \\ k_4 &= h * f(\Omega_n + k_3, t_n + h) \\ \Omega_{n+1} &= \Omega_n + h \end{aligned}$$

여기서 사용된 입력은 t<sub>n</sub>, Ω<sub>n</sub>, h 그리고 N를 입력이 필요하며, θ<sub>0</sub>를 이용하여 t<sub>n</sub>을 산출한다. Ω<sub>n</sub>는 0으로 대입하고, h와 N은 각각 0.2 그리고 1로 대입한다. 즉, h가 누적되어 N이 될 때까지 반복하게 된다. 마지막으로 턱의 기준축과 턱의 벡터 Yv와 Zv의 차이로 각각 cy와 cz를 산출하고, 이를 이용하여 턱과 기준축 사이의 거리 d<sub>n</sub> = √(cy<sup>2</sup> + cz<sup>2</sup>)를 알아낸 이후에 θ<sub>0</sub>를 산출한다. 결과는 실험에서 관찰한다.

## 4. 실험

기존 연구에 표정생성 결과는 그림 5이고, 이 논문의 연구에서 제안된 표정 생성 결과는 그림 7, 또한 그림 6에서 MAX 툴을 이용하여 표정을 자동적으로 생성하기 어렵다. 따라서 수작업으로 표정을 생성하였다.

그림 5에서 "JOY" 표정이 명확하게 생성되기 어렵다. 다시 말해, 입술 주변의 범위를 초과하지 못한다.

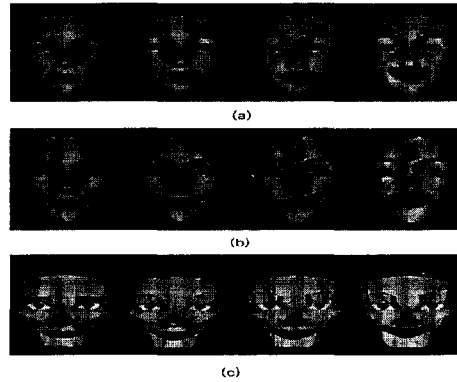


그림 5. 3800, 1900, 296개 개 벡터를 가진 모델에 기존의 연구에서 구현된 "JOY" 표정 생성 결과(왼쪽부터 강도 0, 30, 60 그리고 100)

그림 6에서 강도에 따라서 "JOY" 표정을 생성한 결과이며, 볼의 영역이 거의 소멸된 것을 볼 수 있다. 다시 말해, "JOY"에 대한 의미전달은 명확하지만, 자연스러운 표정 생성이 결여된다.

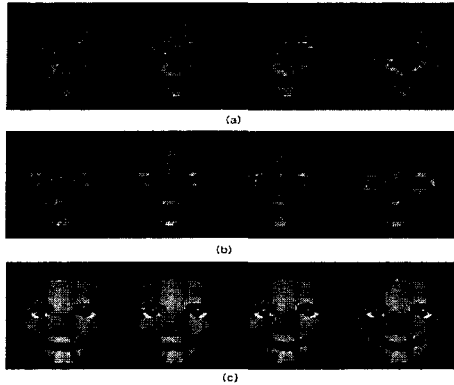


그림 6. 3800, 1900, 296개 개 버텍스를 가진 모델에 3D MAX로 구현한 "JOY" 표정 생성 결과

MAX를 이용하여 표정을 생성하는 방법으로 두 가지 방법을 사용하였다. 첫째로 FFD(Free Form Deformation modifier) 박스를 이용한 입술과 볼에 대한 움직임을 제어하는 영역을 선정하고, 둘째로 Affect Region을 이용한 선택된 영역에서 볼과 입술에 대한 움직임을 제어하였다.

그림 7은 본 논문에서 제안된 방법으로 생성된 "JOY" 표정을 나타낸 것이며, 표정에 대한 의미전달과 자연스러운 표정생성이 가능하였다.

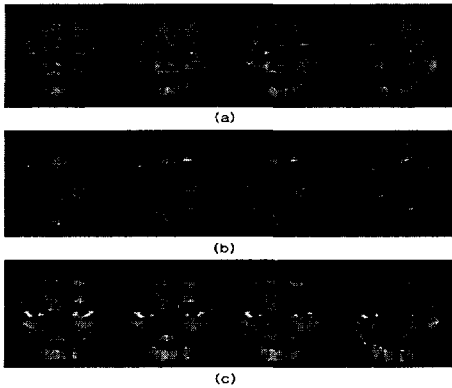


그림 7. 3800, 1900, 296개 버텍스를 가진 모델에 포물형 미분 방정식을 이용한 "JOY" 표정 생성

그림 8은 기존의 턱의 움직임을 수작업으로 이동시킨 결과이다. 그렇지만 그림 9에서 Runge-Kutta법을 이용하여 표정을 생성하였다. 따라서, 턱의 움직임이 자동으로 생성되었고, 보다 자연스러운 움직임이 생성되었다.



그림 8. 3800개 버텍스를 가진 모델에서 기존의 표정 생성(왼쪽부터 강도 0, 30, 60 그리고 100)

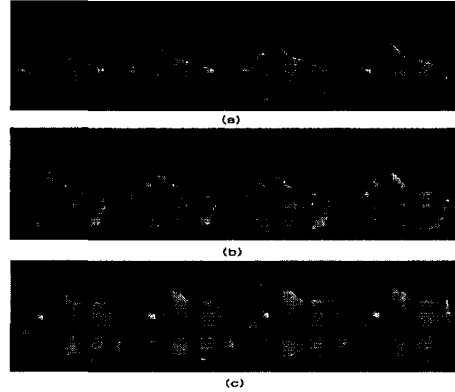


그림 9. 3800, 1900, 296개 버텍스를 가진 모델에서 Runge-Kutta법을 사용하여 "Surprise" 표정 생성에서 턱의 움직임을 표현한다

## 5. 결론

기존 방법에서는 AU의 6번인 "뺨을 올림"은 다른 AU(눈꺼풀, 눈, 입술 등)에 영향을 많이 줌으로써 뺨의 움직임에 대한 제한이 발생하였으나, 제안 방법을 이용함으로써  $\alpha$ 값을 조정하여 자연스러운 볼 모양을 생성할 수 있었다. 또한 AU의 26번인 "턱을 내림"은 턱의 기준 축에서 멀어질수록 움직임의 크기를 달리함으로써 자연스러운 턱의 움직임을 생성할 수 있을 뿐만 아니라 자동으로 표정을 생성한다.

그러나, 특별한 구조를 가지고 있는 모델에는 본 방법을 적용하기 어렵다. 또한 현재 제시되고 있는 아바타 모델은 Texture를 지원하지 않으므로 실제 사진과 같은 표정을 나타내기 어렵다.

## 참고 문헌

- [1] 金 尚雲, 吳 芝英, 靑木 由直, "인터넷상에 3차원 모델을 이용한 한-일 실시간 수화 통신 시스템의 구축을 위한 기초적인 검토", 대한전자공학회논문지-S, vol. 36, no. 7, pp. 811 - 820, 1999. 07.
- [2] 李 容厚, 金 尚雲, 靑木 由直, "지적 아바타 통신을 위한 코믹한 얼굴 표정의 생성법", 대한전자공학회추계종합학술대회 논문집, vol. 23, no. 2, pp. 227-230, 2000. 11.
- [3] S.-W. Kim, J.-Y. Oh, S. Tanahashi, and Y. Aoki, "Preliminary study on a sign-language chatting system between Korea and Japan for avatar communication on Internet." IEICE Trans., vol. E83-A, no. 2, pp. 386 - 389, Feb. 2000
- [4] Erwin Kreyszig. ADVANCED ENGINEERING MATHEMATICS, JOHN WILEY & SONS, pp.1040-1044, 1070-1075, 1993