

# 지적 아바타 통신을 위한 코믹한 얼굴 표정의 생성법

○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 Generation Method of Comic Facial Expressions for Intelligent Avatar Communications

○Yung-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

The sign-language can be used as an auxiliary communication means between avatars of different languages in cyberspace. At that time, an intelligent communication method can also be utilized to achieve real-time communication, where intelligently coded data (joint angles for arm gestures and action units for facial emotions) are transmitted instead of real pictures.

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, a comic-style facial model having only eyebrows, eyes, nose, and mouth is employed. Then generation of facial emotion animation with the parameters is also investigated. Experimental results show a possibility that the method could be used for the intelligent avatar communications between Korean and Japanese.

### 1. 서 론

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 송신 측에서는 수화 영상과 얼굴영상을 생성할 수 있는 수화 관절 각 파라미터와 표정 AU만을 전송하면 수신 측에서는 이 파라미터를 이용하여 다시 수화 애니메이션을 재생하는 방법이다. 이 때 애니메이션을 위한 관절 각 파라미터 값은 애니메이션의 경험에 의해 결정하였고, 얼굴 표정은 한 종류로 고정되는 등의 문제점이 있었다.

기존의 표정 에디터는 MetaCreations사의 Poser, 3D Planet사의 3D Create 및 Blaxxun사의 Avatar Studio 등이 있다[3]. 이들 시스템에서는 아바타의 표정을 생성하기 위해 AU를 미리 정의해서 표정을 생성하기 때문에 아바타 모델에 AU가 정의되어 있지 않으면, 표정을 생성할 수 없다. 따라서 표정을 생성할 수 있는 아바타 모델의 종류를 제한하고 있다. 그러나 가상공간상에서 활동하는 아바타는 크기, 폴리곤 수, 구조 등에서 다양한 종류가 있다. 따라서 임의의 아바타 모델에 표정을 생성하기 위해서는 필요한 AU를 찾아내는 작업이 필요하다.

본 논문에서는 모양과 구조가 다른 아바타 모델에서 표정을 생성하기 위한 AU를 찾아내는 방법을 제안한다. AU를 찾기 위하여 폴리곤 정보나 색상 정보, 또는 이 두 정보를 모두 이용한다. 또한 지적 통신방식에서 찾아낸 AU만을 전송하면 수신 측에서는 자기 모델에 이 AU를 적용하여 원하는 감정 표현을 재생할 수 있도록 한다. 또한, 사람마다 표정이 다양하듯이 아바타 종류에 따라서 표정의 형태가 다르게 하며, 표정을 나타낼 수 있는 범위는 아바타의 얼굴 크기에 따라서 다르게 생성할 수 있도록 한다.

이하 2장에서는 인터넷상 가상 공간상에서 구현한 지적 아바타 수화 통신 시스템을 고찰하고, 제 3장에서는 얼굴 모델을 설명한다. 제 4장에서는 다양한 아바타 모델에서 코믹한 얼굴표정을 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. 여기서 서버의 역할은 클라이언트로부터 오는 수화 문장을 각 클라이언트가 요구하는 언어권의 수화 파라미터로 번역하여 내보내 주는 것이다. 그림 1은 인터넷상에 구현된 한-일간의 지적 수화 통신 시스템 구조이다[4].

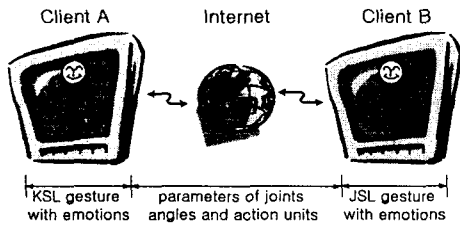


그림 1. 인터넷상에 구현된 한-일간의 지적 수화 통신 시스템.

여기서, 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 송신 측에서는 보낸 수화 및 표정 파라미터를 이용하여 수신 측에서 수화 영상을 재생한다.

## 3. 얼굴 모델

얼굴 영상을 실시간으로 디스플레이하기 위해서는 가능한 적은 수의 제어 점으로 구성되는 모델 영상이 필요하다. 반면에 보다 사실적인 얼굴 표정을 재현하기 위해서는 충분한 수의 폴리곤 모델이 필요하다. 따라서 본 연구에서는 적은 수의 폴리곤으로 감정을 효율적으로 표현할 수 있는 코믹모델을 이용한다. 코믹한 모델이란 과장된 얼굴표정을 합성하기 위해 감정 표현에 꼭 필요한 눈썹, 눈꺼풀, 눈동자, 입의 제어 점만을 사용한 모델이다.

기존 방식에서는 사진과 같은 사실적인 감정 표현을 위해 AU6 (뺨을 올린다)와 같이 얼굴 근육의 미세한 움직임을 표현하기 때문에 많은 AU의 조합이 필요하였다. 그러나 코믹 모델에서는 얼굴 근육의 필요한 부분만을 강조하여 과장된 표현을 하기 때문에 보다 적은 수의 AU만으로도 효과적인 감정 표현이 가능하게 된다.

그림 2는 폴리곤 수가 3,800, 1,900, 그리고 296개인 3차원 wireframe 얼굴 모델이다

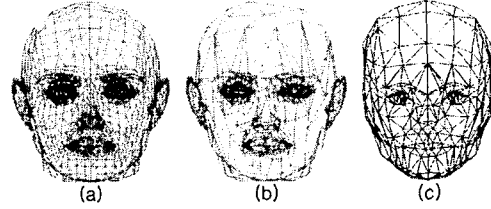


그림 2. 감정 표현을 위한 3차원 wireframe 얼굴 모델 (a) 3,800, (b) 1,900, (c) 296 폴리곤 모델.

여기서 (a)와 (b)는 동일한 구조의 모델로써 폴리곤의 수를 다르게 한 모델이고, 마지막으로 (c)는 구조와 폴리곤 수를 모두 다르게 한 모델이다.

## 4. 얼굴 표정 재생 방법

그림 3은 3차원 모델에 포함되어 있는 폴리곤 정보와 색상 정보를 이용하여 AU와 AU 주변을 추출하고, 6가지 표정을 생성하는 절차를 나타낸 것이다. 여기서 3D Model에는 3차원 아바타 모델을 생성하기 필요한 정보들이 포함되어 있다. 즉, 폴리곤, 벡터스 및 색상 정보를 가지고 있다. 단계 ①, ②, ③에서는 벡터스, 폴리곤, 색상 정보를 참조하며, 단계 ④, ⑤에서는 AU를 찾을 수 있도록 위치적 정보를 제공한다. 그리고 단계 ⑥에서는 폴리곤과 색상정보를 같이 사용하여 상호보완적으로 참조하여 그룹핑을 수행한다. 또한 단계 ⑦, ⑧은 폴리곤 정보와 색상정보를 이용하여 그룹핑하며, 단계 ⑨, ⑩은 AU 추출과 AU 주변 추출을 한다.

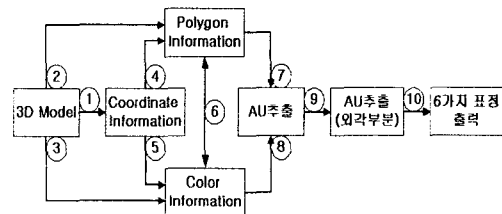


그림 3. 얼굴 표정 생성을 위한 AU 추출 절차.

여기서 그룹핑이란 폴리곤들 간에 혹은 색상들 간에 떨어져 있으면 다른 그룹으로 분류하는 것이며, AU 주변(외각부분)이란 AU가 움직이므로써 영향을 미칠 수 있는 주변이 되는 영역을 의미한다.

표 1은 AU를 찾아내기 위한 3가지 방법이며, 그림 3의 3D Model에서 제공되는 정보들 중에서 어떠한 부분을 참조하는가에 따라서 3가지로 나눌 수 있다. 즉, 폴리곤 정보 이용하여 AU를 찾지 못하면, 색상 정보를 이용하여 AU를 찾는다. 만약 색상 정보를 이용하여 AU를 찾지 못하면, 폴리곤 정보와 색상정보를 같이 이용하여 AU를 찾는다.

표 1. AU를 찾아내기 위한 3가지 방법

AU를 찾는 방법	AU 찾는 순서
폴리곤을 이용하는 방법	① ② ④ ⑦ ⑨ ⑩
색상 정보를 이용하는 방법	① ③ ⑤ ⑧ ⑨ ⑩
폴리곤과 색상 정보를 모두 이용하는 방법	① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩

그림 4은 코의 꼭지점과 그룹핑된 그룹들 사이에 있는 상관정보를 이용하여 중요한 얼굴 구성요소를 찾는 과정을 나타낸다.

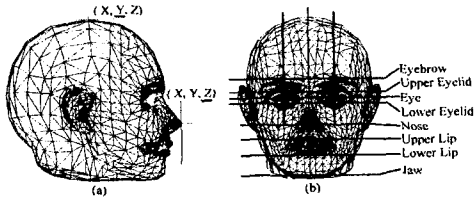


그림 4. 얼굴 구성요소의 추출. (a)는 코의 꼭지점 위치를 추출, (b)는 코의 꼭지점 기준과 그룹들의 위치 비교.

그림 5은 AU 주변을 추출하기 위해서는 그림2의 단계 ⑦, ⑧에서 찾아낸 AU의 위치를 기준으로 사각형의 크기를 생성하는 방법을 나타낸다. 여기서 사각형은 주변을 포함하는 크기를 의미하며, 사각형의 크기는 눈썹과 눈꺼풀, 그리고 입술의 평균위치에 따라서 크기가 다르다. 눈썹과 눈꺼풀 주변의 크기는 눈썹과 눈꺼풀의 높이 차이이며, 입술 주변의 크기는 윗입술과 아랫입술의 높이 차이로 나타낼 수 있다. 사각형은 큰 사각형과 작은 사각형으로 분리하며, AU 움직임이 100%인 경우, 작은 사각형은 66% 움직이며, 큰 사각형은 33% 움직인다.

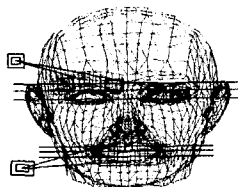


그림 5. AU 주변을 추출하는 사각형 AU 주변을 추출하기 위해서는 AU의 위치를 기준으로 한다.

턱을 추출하는 방식은 인접한 폴리곤을 찾아 레벨을 지정하고 Normal Vector 계산을 사용한다.

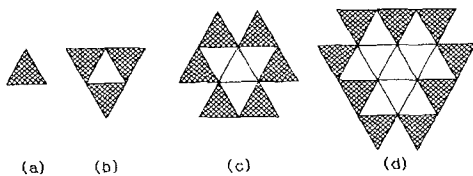


그림 6. (a) 레벨0, (b) 레벨1, (c) 레벨2, (d) 레벨3 아래 입술의 가정 근접한 폴리곤을 찾아서 레벨을 0

으로 처리하며, 레벨0 폴리곤에 멀어질수록 가중치를 부여한다. 레벨0을 기준으로 Normal Vector를 사용하여 비교한다. 즉, 레벨0과 레벨1을 Normal Vector(N)를 비교하며, 레벨1과 레벨2, 다음으로 레벨2와 레벨3을 비교하는 순서로 하고. 하나에 폴리곤에 3개의 Vertex(V)를 각각  $V_1 = (X_1, Y_1, Z_1)$ ,  $V_2 = (X_2, Y_2, Z_2)$ ,  $V_3 = (X_3, Y_3, Z_3)$  라고 할 때, 이 벡터들 사이에 외적을 취하면

$$N = (V_2 - V_1) \times (V_3 - V_1)$$

이 된다. 여기서 X, Y, Z중에서 Z축만을 기준으로 턱을 추출하기 때문에

$$N = (Z_2 - Z_1) \times (Z_3 - Z_1)$$

이 된다. 여기서

$$0 < |N| < 0.0005$$

를 만족할 경우 이 벡터를 턱을 구성하는 벡터로 판정하였다. 이 때 범위는 모델에 따라서 다를 수 있으며, 실험적으로 구한 값이다.

### 5. 실험

다음 그림 7은 그림 4와 그림 5에서 찾아낸 AU와 AU 주변을 사용하여 3800, 1900, 그리고 296 벡터를 가진 3차원 wireframe 모델과 비교한 것이다. 좌측부터 눈썹, 눈꺼풀, 입술, 턱을 추출하는 화면이다.

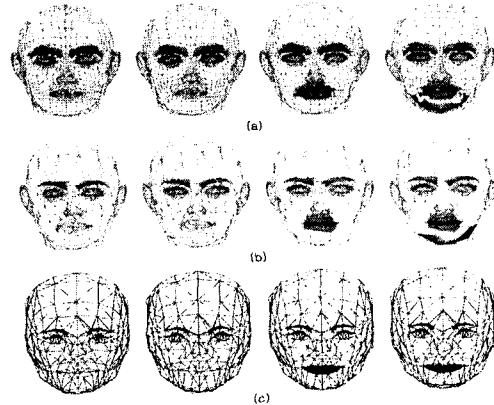


그림 7. AU와 AU 주변을 찾아내는 과정.

(a) 3800개, (b) 1900, (c) 296개 벡터를 가진 모델

그림 8은 그림 3에서 ⑩번 과정에서 표정을 생성한 화면이며, 좌측부터 기쁨, 슬픔, 억저음, 화남, 놀람, 그리고 두려움을 나타내는 6가지 표정 화면이다.

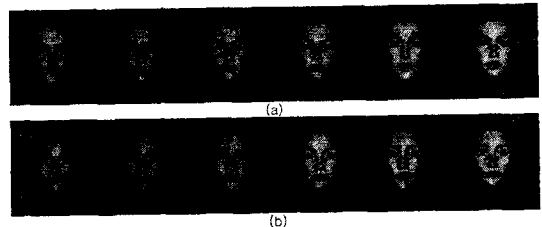


그림 8. 1900개 모델에 6가지 표정을 생성. (a) 3800개, (b) 1900개 벡터를 가진 모델

표정을 나타낼 때 사용할 수 있는 6가지 표정은 내적상태 기술단어의 차원값 분포[5]에서 도표로 제시할 수 있다. 이 도표에 의하면 화남과 혐오(역겨움)은 분포가 비슷하며, 두려움(부서움)과 슬픔(울적한)도 비슷한 위치에 있다. 표정을 나타낼 때 동일한 AU를 사용하는 것과 다른 AU를 사용하는 것을 분류하여, 다른 AU를 사용하는 것을 과장된 표현함으로써 6가지 표정들을 분리시킬 수 있다.

표 2는 6가지 표정에서 나타낼 수 있는 표정의 종류를 나열하고, 가장 잘 어울리는 표정을 음영으로 처리하였다. 여기서 AU(2, 4)란 화남과 혐오에서 공통적으로 사용하는 AU이고, AU(1, 4)는 두려움과 슬픔에서 같이 사용하는 AU이다. 또한 AU(1, 2)는 기쁨과 놀람에서 같이 사용하는 AU를 의미한다.

6가지 표정은 AU공통과 표현 방법의 종류를 결합하여 표정을 생성할 수 있다.

표 2. 표정을 나타낼 때 사용할 수 있는 AU의 종류

	AU 공통	얼굴 표정의 표현 방법						
		1	2	3	4	5	6	7
화남	2,4	5	15	20	5,15	5,20	15,20	5,15,20
혐오	2,4	10	18	48	10,18	10,48	18,48	10,18,48
두려움	1,4	20	25	26	20,25	20,26	25,26	20,25,26
슬픔	1,4	15						
기쁨	1,2	12	25	12,25				
놀람	1,2	5	26	5,26				

표 3은 296개의 벡터를 가지고 있는 모델은 폴리곤과 색상을 각각 따로 이용한 경우에 AU를 찾아낼 수 없으며, 표 1에서 "폴리곤과 색상 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.

표 3. 폴리곤과 색상을 이용한 그룹핑 생성(296개 벡터를 가진 모델)

그룹번호	방식	폴리곤과 색상 정보를 모두 이용하는 방법	
		폴리곤(첫 번째)	색상(두 번째)
1		머리카락	머리카락
2		얼굴	얼굴
3			왼쪽눈썹
4			오른쪽 눈썹
5		왼쪽 눈	왼쪽 눈꺼풀
6			왼쪽 홍채
7			왼쪽 흰자
8		오른쪽 눈	오른쪽 눈꺼풀
9			오른쪽 홍채
10			오른쪽 흰자
11		윗입술	윗입술
12		아랫입술	아랫입술
그룹 합계		6개	12개

그림 9는 표 2에서 음영으로 처리된 표정과 표 3에서 폴리곤과 색상을 이용한 그룹핑 된 내용을 가지고 6가지 표정을 생성한 것이다.

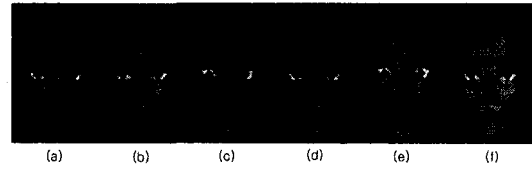


그림 9. 296개 모델에 과장된 표현으로 생성한 6가지 표정, (a) 기쁨, (b) 슬픔, (c) 역겨움, (d) 화남, (e) 놀람, (f) 두려움

## 6. 결론

인터넷 가상공간에서 이루어지는 지적 수화 제스처 통신에서는 수화 영상을 전송하는 대신에 수화 제스처 및 얼굴표정 애니메이션을 생성할 수 있는 관절 각 및 AU 파라미터만을 전송한다.

가상 공간에서 다양한 아바타 모델들이 있으며, 아바타 모델마다 AU를 미리 지정하여 하는 것은 매우 어렵다. 논문에서는 다양한 아바타 모델에 AU를 미리 지정하지 않아도 표정을 생성하기 위해서 아바타 모델에 AU를 찾아내기 위한 3가지 방법과 표정 생성하는 과정을 제안하였다.

그리고 6가지 표정 중에 화남과 혐오, 그리고 두려움과 슬픔에서 분포가 비슷한 위치에 있으며, 표정을 나타내는데 표정들 간에 다른 AU를 사용하는 부분을 과장된 표현으로 함으로써 6가지 표정을 분리시켰다.

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 商雲, 吳 芝英, 靑木 山直, "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] 高 光訓, 金 商雲, 棚橋 眞, 靑木 山直, "효율적인 수화 통신을 위한 코믹한 감정 표현", 대한전자공학회계중합학술대회 논문집, vol. 22, no. 1, pp. 651-654, 1999. 6.
- [3] S. K. Wilcox, Web Developer.Com Guide to 3D Avatar, John Wiley & Sons, Inc., pp.99 - 184, 1998.
- [4] Sang-Woon Kim, Ji-Yong Oh, Shin Tanahashi, and Yoshinao Aoki, "Preliminary study on a sign-language chatting system between Korea and Japan for avatar communication on Internet." IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, vol. E83-A, no. 2, pp. 386 - 389, Feb. 2000
- [5] [http://psylab.yonsei.ac.kr/~vision/cschung/표정\\_국문/index.htm](http://psylab.yonsei.ac.kr/~vision/cschung/표정_국문/index.htm)