

## 2D 코믹 모델을 이용한 3D 얼굴 표정 에디터

○李 容厚\*, 金 商雲\*, 青木 由直\*\*

\* 明知대학교 컴퓨터工學部

\*\* 北海道大学大学院工學研究科

### A 3D Facial Emotion Editor Using a 2D Comic Model

○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

A 2D comic model, a comic-style line drawing model having only eyebrows, eyes, nose and mouth, is much easier to generate facial expressions with small number of points than that of 3D model. In this paper we propose a 3D emotional editor using a 2D comic model, where emotional expressions are represented by using action units(AU) of FACS. Experiments show a possibility that the proposed method could be used efficiently for intelligent sign-language communications between avatars of different languages in the Internet cyberspace.

#### 1. 서론

수화 제스처는 인터넷 사이버공간에서 개인을 상징하는 아바타(avatar)의 통신 보조 수단으로도 사용될 수 있다. 특히 얼굴 표정을 가진 수화 영상은 단지 팔과 손의 움직임만을 보여주는 영상에 비해 이해하기가 쉽다[1]. 또한 코믹 만화나 애니메이션 영화에서는 과장된 얼굴 표정이나 몸짓으로 사실적이지는 않지만 보다 확실하게 감정을 전달하고 있다. 따라서 문헌[2]에서는 2차원 및 3차원에서 코믹 모델과 특수기호를 이용한 감정표현의 효과를 고찰하였다. 그러나 3차원 얼굴 표정은 2차원에 비하여 감정의 전달 효과가 크게 개선되지 않았음을 보이고 있다[3].

한편, 실시간 수화 영상 통신을 위하여 수화 영상이나 압축 영상을 전송하는 대신에 수화 애니메이션 파라미터만을 전송하는 지적 통신방식을 이용한다. 즉

송신 측에서는 수화 영상을 수화 관절 각 파라미터로 변환하여 전송하면 수신 측에서는 이 파라미터를 이용하여 다시 수화 애니메이션을 재생하는 방법이다. 이때 애니메이션을 위한 관절 각 파라미터 값은 애니메이션의 경험에 의해 결정하였고, 얼굴 표정은 한 종류로 고정되는 등의 문제점이 있었다.

기존의 표정 에디터에는 MetaCreations사의 Poser, 3D Planet사의 3D Create 및 Blaxxun사의 Avatar Studio 등이 있다[4]. 본 논문에서는 비교적 감정 생성이 용이한 2차원 코믹 모델을 이용하여 표정을 편집한 후, 이를 이용하여 3차원 표정을 생성하는 방법을 제안한다. 즉 감정의 종류 및 강도를 조정하는 버튼을 이용하여 FACS (Facial Action Coding System)의 AU(Action Unit)들을 조합하여 2차원 코믹 모델 상에 원하는 감정을 표현 한 후, 동일한 AU를 이용하여 원하는 표정을 3차원으로 생성하였다. 또한 이 AU만을 전송하면 수신 측에서는 자기 모델에 이 파라미터를 적용하여 원하는 감정 표현을 재생할 수 있도록 한다.

이하 제 2장에서는 인터넷상의 가상공간에서 지적통신방식에 의한 수화 제스처 통신 시스템을 고찰하고, 제 3장에서는 과장된 표현이 가능한 코믹 모델 및 2차원 코믹 모델을 이용하는 3차원 표정 에디터를 제안한다. 그리고 제 4장에서는 실험 결과를 고찰한 후, 제 5장에서 결론을 맺는다.

#### 2. 지적 통신방식에 의한 수화 제스처 통신 시스템[5]

인터넷 가상공간에서의 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. 또한 서버는 각 클라이언트들간의 메시지의 송·수신을 관리하게 된다. 이러한 구조는 각 클라이언트들에게 돌아가는 파라미터가 동일하기 때문에 디스플레이되는 수화 동작을 통일할 수 있다는 점과 서버의 기능 확장만으로 새로운 서비스를 추가할 수 있다는 장점이 있다. 서로 다른 언어를 구사하는 한-일 수화 통신 시스템의 구조는 그림 1과 같다[5].

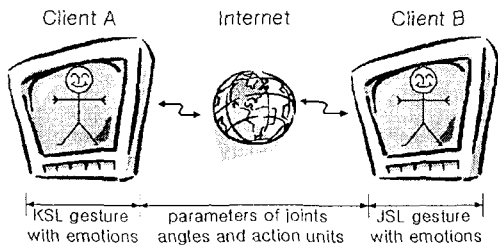


그림 1. 인터넷상에 구현된 한-일간의 지적 수화 통신 시스템.

그림 1에서 클라이언트 A는 한글 수화(Korean Sign Language: KSL)를 구사하는 사이트이고, 클라이언트 B는 일본 수화(Japanese Sign Language: JSL)를 쓰는 사이트이다. 여기서 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.

따라서 서버에는 수화문과 얼굴 표정을 파라미터로 변환시킬 수 있는 “수화 단어 사전”이 필요하다. 본 연구에서는 다양한 얼굴 표정 파라미터를 도출하기 위한 에디터에 대하여 검토한다.

### 3. 얼굴 모델 및 에디터 제작

얼굴 영상을 실시간으로 디스플레이하기 위해서는 가능한 적은 수의 제어 점으로 구성되는 모델 영상이 필요하다. 코믹한 모델이란 과장된 얼굴표정을 합성하기 위해 감정 표현에 꼭 필요한 눈썹, 눈, 눈동자, 입의 제어 점만을 사용한 모델이다. 즉 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 감정

표현을 위한 2차원 코믹 모델과 벡스 수가 각각 296과 452인 3차원 얼굴 모델이다.

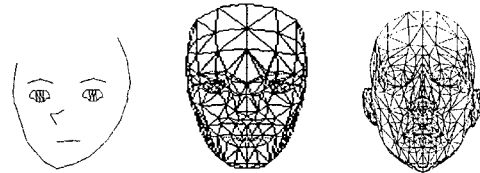


그림 2. 감정 표현을 위한 2차원 3차원 얼굴 모델. (a) 2차원 코믹모델, (b) 3차원 여성모델(296버텍스), (c) 3차원 남성모델(452버텍스).

본 연구에서는 2차원 코믹 모델에서는 감정 표현을 위해 전체 60개의 제어 점 중 28개를 사용하였고, 3차원 모델에서는 전체 295개의 제어 점 중 75개의 제어 점만을 사용하였다.

그림 3은 2차원과 3차원에서 동일한 AU를 사용하여 감정의 종류와 강도를 편집하는 에디터 화면이다.

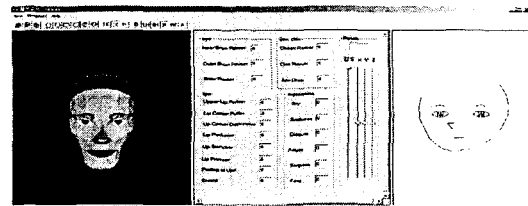


그림 3. 표정 에디터 화면의 예. 좌우 측은 표정 창이고, 중간은 AU값과 표정선택 및 강도 조절을 위한 제어 패널이다.

그림 3에서 우측 창의 2차원 코믹모델을 관측하면서 선택한 표정 및 강도를 조정한 후 제어 패널에 나타난 AU를 이용하여 우측 창에 3차원 표정을 생성한다. 3차원에서 바로 강도를 지정하는 경우에는 표현의 차이가 명확하지 않기 때문에 2D를 동시에 적용함으로써 표정 편집을 명확히 할 수 있다.

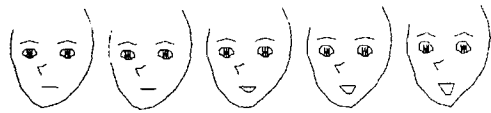
### 4. 실험

먼저 제안한 에디터를 이용하여 2D 코믹 모델과 3D 모델에서 감정을 표현하기 위하여 필요한 AU번호를 조사하였으며, 6가지 기본 감정에 대한 AU번호는 표 1과 같다. 즉 두 모델에서 기본적인 감정을 표현하기 위해 필요한 AU는 거의 같음을 알 수 있다.

표 1. 2D 코믹 모델과 3D 코믹 모델에서 6가지 기본 감정을 표현하기 위하여 필요한 AU의 비교.

Emotions	2D Comic Model	3D Model
JOY	12	12
SADNESS	1+15	1+15
DISGUST	1+48	1+48
ANGER	2+15	2+15
SURPRISE	1+2+5+26	1+2+5+26
FEAR	5+24	5+24

또한, 제안한 에디터를 이용하여 다양한 종류의 감정과 강도를 편집하는 실험을 하였다. 그림 4는 2차원 코믹모델과 3차원 모델에 "Surprise" 표정을 다섯 가지 강도로 생성한 예이다. 또한 그림 4의 표정을 생성하기 위한 AU의 종류 및 강도 값은 표 2와 같다.



(a)



(b)

그림 4. 얼굴 표정 "Surprise"를 다섯 가지 강도로 편집한 예. 여기서 (a)는 2차원 코믹모델의 표정이고 (b)는 3차원 모델의 표정이며, 좌측으로부터 표정 강도는 각각 0%, 25%, 50%, 75%, 100%이다.

표 2. 그림 2의 감정을 표현하기 위하여 필요한 AU의 비교.

표정	AU번호(강도 값)
SURPRISE-0	AU-1, 2, 5, 26 (0%)
SURPRISE-1	AU-1, 2, 5, 26 (25%)
SURPRISE-2	AU-1, 2, 5, 26 (50%)
SURPRISE-3	AU-1, 2, 5, 26 (75%)
SURPRISE-4	AU-1, 2, 5, 26 (100%)

한편, 인터넷 가상공간에서 지적 방식에 의한 수화 통신이나 아바타의 제스처 통신의 경우 각 클라이언트는 서로 다른 아바타 모델을 사용할 수 있으며, 또 모델을 구성하는 버텍스 수도 각각 다를 수 있다. 따라서 본 논문의 에디터를 이용하여 도출한 감정 표현 파라미터(AU번호 및 강도 값)를 서로 다른 얼굴 모델에 적용하여 비교하는 실험을 하였고, 또 같은 얼굴 모델에서 버텍스 수를 달리 하면서 감정을 표현하는 실험을 하였다.

그림 5는 그림 4의 감정을 생성하기 위한 AU 파라미터를 버텍스 수가 452개인 남자 모델(그림 4 모델의 버텍스 수는 296 이다)에 적용하였을 경우의 예이다. 또 그림 4(b)의 3차원 모델의 버텍스 제어점 수를 225개와 151개로 줄여 표 2의 AU 값을 그대로 적용하였을 경우의 예는 각각 그림 6, 7과 같다.



그림 5. "Surprise" 감정을 452 버텍스 해상도의 남자 모델에 표현한 예.



그림 6. "Surprise" 표정을 225 버텍스 해상도의 여자 모델에 표현한 예.



그림 7. "Surprise" 감정을 151 버텍스 해상도의 여자 모델에 표현한 예.

본 연구의 에디터를 이용하여 편집한 표정의 정확도를 측정하기 위하여 다음과 같이 표정의 인식률을 실험하였다. 실험에 참가한 학생들에게 편집한 얼굴 표정을 제시하고, 해당하는 감정의 종류와 강도를 맞히

도록 하였다. 네 명을 대상으로 한 인식 실험의 결과는 표 3과 같다.

표 3. 3차원 얼굴 표정에 대한 인식률(%)

표정종류 \ 표정 강도	25%	50%	75%	100%
JOY	100	100	100	100
SADNESS	25	75	50	50
DISGUST	50	50	25	25
ANGER	75	75	100	100
SURPRISE	50	75	100	75
FEAR	0	0	25	0

끝으로, 지적통신 방식으로 얼굴 표정을 전송하기 위한 데이터 량을 분석하였다. 클라이언트간 전송되는 제어 파라미터는 한 어절 단위로 이루어지며, 얼굴 표정을 나타내는 파라미터인 F, 강도를 나타내는 S, 수화 제스처를 나타내는 P, 입 모양을 나타내는 M으로 구성된다. 첫 번째 파라미터인 F는 얼굴 표정을 나타내는 코드로, 6가지 표정과 15가지 AU번호로 이루어진다. S는 표정강도로 0에서 100까지의 범위를 가진다. P는 제스처 동작을 구성하는 키 프레임 수를 나타내며, 한 동작 제스처는 54개 파라미터 열로 기술된다. 끝으로 입 모양 파라미터 M은 어절별로 초성에 1개 중성에 2개 종성에 1개로 총 4개가 존재한다. 예를 들어 "학교에"의 경우 총 239 (=3+4+2+54×4+2+4×3)개의 파라미터가 전송되며, 각 파라미터 당 2 바이트를 할당한다면 총 478 (=239×2)바이트가 된다.

한 프레임을 디스플레이 하는데 드는 시간은 0.1초라 하고, 각 키 프레임사이의 보간 프레임 수를 7이라 할 때 "나는 학교에 갑니다"라는 수화 문장에 대해 전송할 수화-표정 파라미터의 데이터 량과 데이터 속도는 다음과 같다. 먼저 "나는", "학교에", "갑니다"에 대한 데이터 량은 각각 73 (=3+4+2+54×1+2+4×2), 239 (=3+4+2+54×4+2+4×3), 131 (=3+4+2+54×2+2+4×3)로 전체 데이터 량은 443바이트가 된다. 그러므로, 자연스러운 수화 통신을 위해서는 1.45Kbps (=443×8×2)/0.7×(1+4+2)) 데이터 전송속도가 요구된다.

### 5. 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. 따라서 이 논문에서는 얼굴표정을 효과적으로 편집 생성할 수 있도록 하기 위하여, 2차원 코믹 모델을 이용하여 3차원 표정을 생성

하는 방법을 제안하였다. 비교적 제어가 용이한 2차원 코믹 모델 상에서 표정을 편집하여 AU종류 및 강도 값을 결정한 다음 이를 3차원 모델에 적용하는 방법을 이용하였다.

또한 가상공간에 참여하는 아바타들은 서로 다른 구조와 해상도를 가질 수 있다. 따라서 본 논문의 표정 에디터를 이용하여 도출한 얼굴 표정 파라미터를 구조와 해상도가 다른 캐릭터 모델에 적용하여 표정을 생성하는 실험을 하였다. 실험 결과 서로 다른 모델과 해상도가 떨어지는 모델에서도 원하는 얼굴 표정을 재현할 수 있었다. 따라서 제안 에디터를 이용하여 표정 및 수화 제스처 파라미터만을 추출하여 전송하는 방법으로 서로 다른 언어를 사용하는 아바타들 사이의 제스처 통신에 이용할 수 있을 것으로 사료되며, 앞으로의 과제이다.

### 참고 문헌

- [1] 金 商雲, 吳 芝英, 青木 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] 高 光訓, 金 商雲, 棚橋 眞, 青木 由直, "3D얼굴 모델의 코믹한 감정 표현", 대한전자공학회추계대회논문집, vol. 22, no. 2, pp.536 - 539, 1999. 11.
- [3] S.-W. Kim, Jong-Woo Lee, Ji-Yong Oh, and Yoshinao Aoki, "A comparative study on the sign-language communication system between Korea and Japan through 2D and 3D character models on the Internet," Proc. of 1999 IEEE International Conference on Image Processing, Kobe, Japan, 28AP1.9, Oct. 1999.
- [4] S. K. Wilcox, Web Developer.Com Guide to 3D Avatar, John Wiley & Sons, Inc., pp. 99 - 184, 1998.
- [5] 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.