FBML을 이용한 서로 다른 아바타 모델간의 제스처 통신

*이용후, *김상운, **青木 由直
*명지대학교 컴퓨터공학과, **북해도대학대학원공학연구과
e-mail : {yongwho, kimsw}@mju.ac.kr

Gesture Communications Between Different Avatar Models Using FBML

Yong-Hu Lee, *Sang-Woon Kim, **Yoshiki Arakawa
*Dept. of Computer Science & Engineering, Myongji University
**Graduate School of Engineering, Hokkaido University

Abstract

In order to overcome the limitation based on different avatar models, in this paper, we propose gesture communications between different avatar models using FBML (Facial Body Markup Language). The experimental results demonstrate a possibility that the proposed method could be used as an efficient means to overcome the problem.

Ⅰ. 서론

최근 인터넷 가상공간에서 서로 다른 언어를 구사하는 아바타들 사이에서 언어 장벽을 극복하기 위하여 지금까지 청각 장애자의 언어로만 인식되어온 수화를 이용하는 방법에 대한 연구가 진행되었다. 또한, 지적 통신방식과 3D 애니메이션 기술을 이용하여 수화 제스춰를 실시간으로 전송 및 재생하기 위한 연구 등이 수행되었다[1]. 그러나 가상공간에서 활동하는 아바타들은 서로 다른 구조를 가질 수 있다. 예를 들어 표정과 몸동작을 전달하기 위해서 일반적으로 FACS (Facial Action Coding System)의 AU (Action Unit)들과 관절각 정보를 사용하고 있다.

기존 연구[2]에서는 구조 혹은 모양이 다른 아바타 모델에 얼굴 표정을 생성하기 위해서 볼, 턱의 영역을 추출하는 방법을 제안하였다.

그렇지만 한 시스템에서는 FACS를 사용하고 있지만, 서로 다른 시스템상에서 아바타의 모양과 구조가 틀림으로 인해서, 서로 의사를 전달하는 것은 사실 불가능하다.

본 논문에서는 이러한 문제를 해결하기 위해서 서로 다른 아바타 모델에 전달되는 내용을 표준 문서로 만들어 FBML (Facial Body Markup Language)을 정의하였다. FBML에서 기본으로 사용하고 있는 시스템은 FACS와 관절각 정보를 이용하지만, 또한 다른 시스템을 사용할 수 있도록 하였다.

그림 1.1. 아바타들 간에 XML을 이용한 시스템

그림 1.1처럼, 서로 다른 모델간에 표정 전달을 가능하도록 하기 위해서, XML 데이터를 이용하여 전달할 수 있도록 하는 DTD(Document Type Definition)을 제안한다.

Ⅱ. 관련 연구

2.1 얼굴 관련 연구

Ekman과 Friesen[3]은 눈으로 식별 가능한 모든 얼굴의 변화를 표현할 수 있는 시스템을 만들었다. FACS(Facial Action Coding System)라고 하는 이 시스템은 얼굴 표정의 변화를 일으키는 기본적인 action unit(AU)들의 나열로 얼굴 표정을 기술한다. FACS에는 얼굴의 표정을 기술하는 46개의 action unit들과 head orientation과 gaze를 표현하는 12개의 AU들을 합쳐서 총 58개의 AU들이 정의되어 있다.

Essa와 Pentland은 기존의 FACS가 사람의 판단에

기반하는 heuristic 한 점을 지향하고자 얼굴의 근육구조와 피부의 물리학적 특정을 반영한 3차원 모델을 이용하여 muscle actuation을 직접 계산해 내고 이 값을 사용하는 새로운 표정 기술 시스템인 FACS+를 개발하고, 이를 이용해 표정을 분석하고 인식하였다.

2.2 지적 통신 방식에 대한 수화 제스처 통신

인터넷 가상공간에서의 지적 통신 방식에 의한 수화 제스처 통신 시스템이란 가상공간에서 활동하는 서로 다른 언어권의 아바타들 사이의 언어 벽을 극복하기 위하여 수화 제스춰를 통신 수단으로 이용하는 시스템으로서 중앙 집중적인 클라이언트-서버 구조를 가진다. 예를 들어 한글과 일본어를 구사하는 한-일 수화 통신 시스템의 구조는 다음과 같다. 클라이언트 'A'는 한글 수화(Korean Sign Language: KSL)를 구사하는 사이트이고, 클라이언트 'B'는 일본 수화 (Japanese Sign Language: JSL)를 쓰는 사이트이다. 여기서 실시간 통신을 위하여 실제의 수화 영상을 전송하는 대신에 수화 애니메이션을 생성하기 위한 수화 제스처 및 얼굴 표정 파라미터를 추출하여 전송하는 지적 통신 방식을 이용한다. 즉 클라이언트 'A'와 'B'에 각각 아바타 모델을 준비하여 놓고, 송신 측에서는 수화 및 표정 파라미터를 추출하여 보내면 수신 측에서 이들 파라미터를 받아서 해당하는 수화 및 표정 영상을 재생하는 통신방식이다.

따라서 서버와 클라이언트에는 수화 제스춰와 얼굴 표정으로부터 파라미터를 추출하고 해당하는 애니메이션을 재생하는 "영상 그래픽 모듈"과 이를 해당하는 언어권의 파라미터로 변환시킬 수 있는 "수화 단어 사전" 등이 필요하고, 또한 "통신 모듈"이 있어야 한다 [4].

Ⅲ. XML 데이터 생성을 위한 DTD 설계

DTD를 설계하기 위해서는 가장 일반적으로 사용되는 규칙중에 하나는 처음에 head와 body로 나누어야 한다. 그리고 body 부분에 전달하고자 하는 내용을 기술한다. 최상위 엘리먼트를 FBML(Facial Body Markup Language)라고 하며, 이것의 서브 엘리먼트는 <fbml_head>와 <fbml_body>로 나눈다. 그림 3.1은 <fbml_body> 부분의 서브 엘리먼트 구조이다.

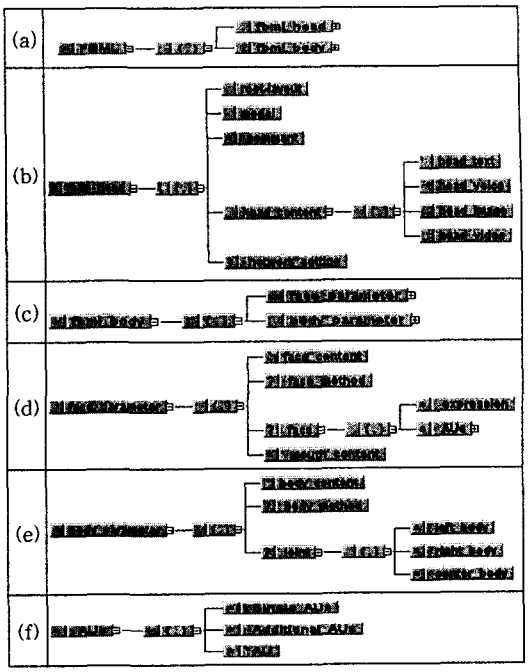

그림 3.1. FBML을 위한 DTD 구조
(a) FBML의 최상위 엘리먼트, (b) fbml_head, (c) fbml_body (d) face_parameter, (e) body_parameter (f) AUs의 서브 엘리먼트 구조

3.1 <FBML> 최상위 엘리먼트 정의

DTD 문서를 만드는 일반적인 규칙은 HEAD와 BODY로 나눈다. 기존의 HTML의 내용을 보면 <HTML>이 최상위 엘리먼트가 되며, 서브 엘리먼트로 <HEAD>와 <BODY>로 나누어지는 것과 같은 이치이다. 그러므로 그림 3.1(a)에서 FBML의 최상위 엘리먼트는 <fbml_head>와 <fbml_body>를 포함하고 있다. 그렇지만 여러번 반복해서 사용할 수 없다. 이러한 두개의 엘리먼트에는 각각의 기능들이 따로 있다.

3.2 <fbml_head> 엘리먼트 정의

그림 3.1(b)의 내용은 <FBML>의 첫 번째 서브 엘리먼트 내용들이다. <fbml_head>는 네 가지의 내용을 포함하고 있다. 첫 번째로 <root-layout>은 해상도와 관련된 정보를 포함하고 있으며, 두 번째인 <model>은 현재 사용하고 있는 모델의 종류를 기술하게 된다. 세 번째는 <head_content>이며 아바타 상호간에 의사전달에 가장 중요한 의미를 전달하기 위해서 만들어졌다. 마지막 <network_setting>이 있으며, 서버와 클라이언트에 상호간에 필요한 네트워트 정보들을 포함

하고 있으며, 서로 상호간에 의사전달을 하기위한 필요한 연결 세팅에 관련된 사항이다.

3.3 <fbml_body> 엘리먼트 정의

그림 3.1(c)는 <FBML>의 서브 엘리먼트인 <fbml_body>의 세부적인 사항들이다. <fbml_body>는 두 개의 엘리먼트를 가지고 있으며, <face_parameter>와 <body_parameter>이다. 전자의 세부사항은 그림 3.1(d)이며, 후자의 세부사항은 그림 3.1(e)이다. 그림 3.1(d)는 아바타 통신에서 얼굴에 대한 의사전달이 중요한 역할을 수행하기 때문에 <face_parameter>를 포함하였다. <face_parameter>에서 기본으로 포함된 시스템은 <facs>이며, 그 밖의 다른 방법들을 지원하기 위해서 <face_content>, <face_method>, 그리고 <mouth_content>를 포함하였다. <face_content>는 얼굴 움직임에 대한 내용을 텍스트 형태로 제공하기 위한 것이며, 마지막으로 <mouth_content>는 입술 움직임에 대한 정보를 포함하는 것이다. 그리고 <facs>는 각각의 AU들을 정의할 수 있도록 <AUs>를 제공하고 있으며, 자세한 내용은 그림 3.1(f)이다. 또한 AU들의 조합인 6가지 표정으로는 기쁨, 슬픔, 두려움, 역겨움, 화남, 그리고 놀람 표정이있다. 이를 전달할 수 있도록 지원하기 위해서 <expression>을 포함하고 있다. 마지막으로 그림 3.1(e)는 아바타 통신에서 몸에 동작에 대한 정보를 전달하는 것으로, 특히 수화 동작에서는 많은 영향을 끼치게 된다.

Ⅳ. 실험

4.1 실험 환경

본 논문에서 제안한 방법을 실험하기 위해 펜티엄 4 1.8G의 Windows XP환경에서 Visual C++ 6.0과 Open Inventer Library를 이용하여 구현하였다. 본 연구에 사용된 3차원 아바타 모델은 VRML에서 제시하고 있는 휴모노이드이며, 폴리곤이 3800개와 296개 폴리곤을 가진 얼굴 모델을 사용하였다. 2차원 모델은 전체 28개의 제어 점을 사용하였고 선을 위주로만 나타낼 수 있다. 기존 시스템과 제안된 시스템을 비교하기 위해서, 메세지 처리 형식을 먼저 살펴보고, 제안된 시스템의 처리 형식을 살펴본다. 또한 마지막으로 기존 시스템과 제안된 시스템에서 처리되는 시간을 측정하여 비교한다.

4.2 기존 방법에서 메세지를 사용한 동작

먼저, 기존 시스템에서 사용된 서버와 클라이언트간에 아바타 통신을 할 수 있는 파라메터는 표 4.1과 표 4.2이다. 전자는 클라이언트에서 서버로 보내는 메세지 내용이며, 후자는 서버에서 클라이언트로 보내는 메세지이다.

표 4.1 클라이언트에서 서버로 보내는 메세지

ID	코드	방번호	내용=(얼굴 표정:F)+(문장)
a	4 0	0	F0 나는 학교에 갑니다

표 4.2 서버에서 클라이언트로 보내는 메세지

ID	코드	내용=(얼굴 표정:F)+(입술:M)+(몸동작:P)
a	4	F0 M2 9 0 0 10 9 7 7 9 P1 0 0 0 0 0 0 0 0 0 0 ..생략..
a	4	F0 M3 10 0 0 10 10 6 2 10 10 5 5 10 P4 20 0 ..생략..
a	4	F0 M3 10 0 0 8 9 6 6 10 9 0 0 10 P2 0 0 0 0 ..생략..

예를 들어, 클라이언트에서 보낸 메세지의 구조(ID+코드+방번호+내용)를 바꿀 경우에는, 수신측인 서버에서 내용을 받아들일 수 있도록 구조를 바꾸어야 한다. 이러한 문제는 다른 아바타 모델을 연결하는 경우에 더욱 심각한 문제를 발생시킨다. 그러므로 이러한 문제를 해결하기 위해서 본 시스템에서는 메세지의 구조가 일부 변경된다해도 영향을 미치지 않는 XML 구조를 정의하였으며, 엘리먼트의 이름만으로 접근이 가능하도록 구현하였다.

4.3 3차원 모델에서 FBML을 사용한 동작

기존 시스템에서 제시한 문제점을 해결하는 방법으로 새로운 XML 문서를 제안하였다. 다양한 시스템에서는 아바타의 종류도 다양하며, 네트워크 설정도 일부 달라리게 된다. 현재 시스템에서 처리되는 형태는 클라이언트에서 아바타 동작에 필요한 텍스트를 서버에게 전달하게 되면, 받아들인 서버는 텍스트에 맞게 아바타를 동작 시킬 수 있는 정보들로 반환한다. 생성된 데이터는 입술의 움직임과 몸의 관절각 정보들을 전달하게 된다. 그림 4.1은 클라이언트에서 서버로 보내는 XML 문서이다.

반대로 서버에서는 클라이언트에서 보낸 XML 문서에 3가지 엘리먼트가 비어있다. 각각은 <face_content>, <mouth_content>, 그리고 <center_body> 이다. 이것은 서버에서 처리하여 내용을 XML 문서에 포함시켜야 한다. 먼저 얼굴 표정에 대한 정보는 <face_content>에 저장되고, 입술 동작에 대한 정보는 <mouth_content>에 포함시키며, 움직임에 대한 관절각 정보는 <center_body>에 포함시킨다. 특히 관절각 정보를 받은 클라이언트에서는 3D 아바타를 동작시킬 수 있게 된다

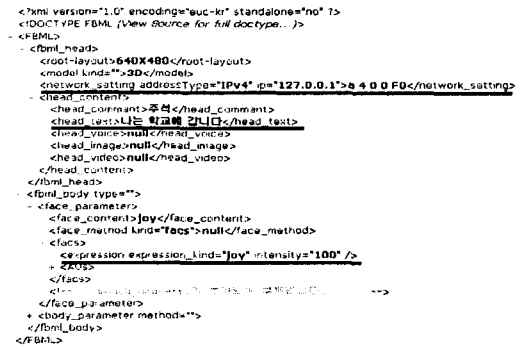

그림 4.1. 클라이언트에서 서버로 보낸 FBML

4.4 2차원 모델에서의 동작

클라이언트에서 보내는 메세지는 3차원에서 사용하는 모델과 동일한다. 즉, 클라이언트에서 보내는 메세지는 "나는 학교에 간다"에 대한 내용이 포함된 XML 문서가 된다. 반대로 서버에서 보내는 메세지는 3차원 모델과 다르다. 그 중에서도 가장 큰 차이점은 <mouth_content>와 <center_body>의 내용이다. 이는 2차원 모델에서 애니메이션을 시키기 위한 관절각 정보와 3차원 모델에서 사용하는 관절각 정보들이 서로 다르기 때문에 발생하는 문제점이다. 이러한 문제를 해결하면 3차원 모델과 같이 동작이 가능하다

4.5 제안된 시스템의 처리 시간 측정

표 4.3에서 2D 원본과 3D 원본을 측정하고, 제안된 시스템으로 2D에 XML을 포함한 것과, 3D에 XML을 포함한 시간을 측정하였다. 측정범위는 클라이언트에서 입력된 메세지를 처리하여 보낸 시간이 포함되며, 또한 서버로부터 받은 메세지를 처리하여 애니메이션 바로 이전까지의 시간을 측정한 것이다.

표 4.3 기존방법과 제안된 시스템간의 시간 분석 (ms)

	2D원본	2D-XML	3D원본	3D-XML
나는	15	139	5	168
학교에	17	142	6	175
가다	17	152	5	168
나는 학교에 가다	45	183	11	189

2D 원본에서는 최대 40밀리세컨트 까지 시간이 걸리지만, XML을 사용하는 경우에는 180밀리세컨트가 걸린다. 즉, 기존 시스템보다 4.5배 정도의 부하가 생기게되는데, 그렇지만 180밀리세컨트라도 사람이 느낄 수 없을 정도의 시간이 소비되므로, 실제 사용자가 느끼는 정도는 아주 미약하다. 기존 시스템에서 제시하지 못했던, 다양한 모델간의 통신을 XML 형식으로 사용함으로써 본 시스템의 타당성을 제시하였으며, 또한 네트워크의 많은 발전으로 인해, 시간측면에서도 성능 저하없이 처리되는 것을 보였다.

V. 결론

가상 공간에서 서로 다른 아바타들 간에 상호 데이터를 주고 받기 위해서 기존 시스템에서는 메세지 처리 형식을 사용하며, 이 형식은 서버-클라이언트 양단간에 규칙에 의해서 만들어진다. 따라서 다른 형식을 가지고 있는 단말기에서 다른 형식으로 메세지를 보낼 경우 기존 시스템들과의 통신은 이루어지기 어렵다.

이러한 문제점을 해결하기 위해서 본 논문에서는 XML 데이터를 만들어 사용하는 시스템을 제안하였다. 이를 위해 먼저 DTD를 정의하였으며, XML 문서를 이용함으로써 다양한 모델사이에 의사 전달이 가능하게 하였다. 또한 FBML에 포함된 내용중에서 유비쿼터스 컴퓨팅 처리를 하기위한 IPv6처리를 지원할 수 있도록 <network_setting> 엘리먼트 속성에 IPv6를 정의하였다. 실험에서는 2차원과 3차원 모델을 사용한 시스템에서 XML을 사용하였을 경우, 처리 가능하다는 점을 확인하였다. 그러나 아바타의 종류와 형태가 너무 다양하기 때문에, 앞으로 더 많은 실험을 필요로 한다.

참고문헌

[1] 金商雲, 李種雨, 青木由直, "3D 수화 애니메이션을 위한 팔과 손의 통합된 키 프레임 에디터," 대한전자공학회논문지-SP, vol. 37, no. 9, pp. 371-380,, 2000.

[2] 李容厚, 金商雲, 青木由直, "인터넷 가상공간에서 지적 아바타 통신을 위한 코믹한 얼굴 표정의 생성법", 대한전자공학회논문지-S, vol. 40, no. 1, pp. 59-73,, 2002.

[3] P. Ekman and W.V. Friesen, Unmasking the Face, New Jersey: Prentice Hall, 1975

[4] 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, 2000.