

논문 2004-41CI-5-6

FBML을 이용한 서로 다른 아바타 모델간의 제스처 통신

(Gesture Communications Between Different Avatar Models Using A FBML)

이 용 후*, 김 상 운**, 아오끼 요시나오***

(Yong-Who Lee, Sang-Woon Kim, and Yoshinao Aoki)

요 약

인터넷 가상공간에서 서로 다른 언어 사이의 장벽을 극복하기 위한 방법으로 수화 통신 시스템이 연구되고 있다. 그러나 지금까지의 연구는 같은 유형의 아바타만을 대상으로 하였기 때문에 모델 구조가 서로 다를 경우에는 적용할 수 없다는 문제가 있다. 따라서, 본 연구에서는 FBML (Facial Body Markup Language)을 이용하여 서로 다른 아바타들 간의 수화 제스처 통신을 구현하는 방법을 제안하였다. 즉, FBML을 정의하여 제스처를 위한 관절각 및 얼굴 표정 등 서로 다른 아바타 모델들 사이에 전달할 내용을 표준 문서로 만들었다. 본 방법은 윈도우 플랫폼에서 Visual C++와 Open Inventor 라이브러리를 이용하여 구현하였으며, 실험 결과 언어의 장벽을 넘을 수 있는 비언어 통신수단으로 이용될 수 있는 가능성을 확인하였다.

Abstract

As a means of overcoming the linguistic barrier between different languages in the Internet cyberspace, a sign-language communication system has been proposed. However, the system supports avatars having the same model structure so that it is difficult to communicate between different avatar models. Therefore, in this paper, we propose a new gesture communication system in which different avatars models can communicate with each other by using a FBML (Facial Body Markup Language). Using the FBML, we define a standard document format that contains the messages to be transferred between models, where the document includes the action units of facial expression and the joint angles of gesture animation. The proposed system is implemented with Visual C++ and Open Inventor on Windows platforms. The experimental results demonstrate a possibility that the method could be used as an efficient means to overcome the linguistic problem.

Keywords : 아바타, 제스처 통신, FBML (Facial Body Markup Language), XML (eXtensible Markup Language)

I. 서 론

최근 인터넷 가상공간에서 서로 다른 언어를 구사하는 아바타들 사이에서 언어 장벽을 극복하기 위한 노력의 일환으로, 지금까지 청각 장애자의 언어로만 인식되어 온 수화를 이용하는 방법에 대한 연구가 진행되고 있다^{[1][2][3]}. 김상운 등^[2]은 수화 제스처를 효율적으로 생성

하기 위해 역 운동학 기법을 이용하여 팔과 손의 관절 각을 자동으로 추출할 수 있는 통합된 키 프레임 에디터를 연구하였다. 또한 김상운 등^[3]은 지적 통신방식과 3D 애니메이션 기술을 이용하여 수화 제스처를 실시간으로 전송 및 재생할 수 있는 방법을 연구하였다. 그 밖에도 얼굴 표정을 가진 수화 영상은 단지 팔과 손의 움직임만을 보여주는 영상에 비해 의사 전달기능이 뛰어나다는 점에 주목하여 효율적인 얼굴표정의 생성법에 대한 연구^[4]와 사실적인 감정 표현보다는 만화 애니메이션에서와 같이 코믹한 표현이 더 효과적이라는 연구 보고도 있다^[5].

가상공간에는 다양한 서로 다른 아바타 모델들이 있다. 그러나 지금까지의 기존 수화 제스처 통신 연구에서

* 학생회원, ** 정회원, 명지대학교 컴퓨터공학과
(Department of Computer Science and Engineering,
Myongji University)

*** 홋카이도대학 전자정보공학연구과
(Graduate School of Electronics and Information
Engineering, Hokkaido University)
접수일자: 2003년12월24일, 수정완료일: 2004년8월30일

는 동일한 구조를 가진 아바타 모델들만을 대상으로 하였다. 즉, 기존 시스템에서는 메시지 처리 형식을 사용하였고, 메시지 형식은 특정한 아바타의 구조에 맞게 설계되었다. 따라서 메시지 형식이 다른 아바타 모델들 간에는 통신할 수 없다는 문제가 있다. 또한 아바타 시스템 환경은 웹과 웹이 아닌 환경으로 나누어지며, 이들 환경들 간에 통신이 문제가 된다. 이는 웹과 웹이 아닌 환경 간에 서로 주고받을 수 있는 통일된 문서가 없기 때문이다.

이러한 문제점을 해결하기 위해서 본 논문에서는 FBML (Facial Body Markup Language)을 정의하여 사용하는 시스템을 제안한다. 이를 위해 먼저 XML (eXtensible Markup Language)^{[6][7]}에서 제공하는 DTD (Document Type Definition)를 이용하여 FBML의 엘리먼트와 속성을 정의한다. FBML은 웹 서비스를 제공하기 때문에 웹과 웹이 아닌 환경에서 서비스할 수 있다는 장점이 있다. 특히 FBML 문서는 유비쿼터스 컴퓨팅 처리에서 사용하는 IPv6^[8]를 지원할 수 있다.

일반적으로 표정 및 몸동작은 전달하기 위해서 FACS (Facial Action Coding System)의 AU (Action Unit)[4]들과 관절각 정보를 사용하고 있다. 본 논문에서는 서로 다른 구조를 갖는 아바타들 간의 통신문제를 해결하기 위해서, 전달할 내용을 FBML을 이용하여 표준 문서로 만든다. FBML은 FACS와 관절각 정보를 포함하며, 다른 시스템을 사용할 수 있도록 확장성을 고려하여 설계하였다.

본 연구에서는 서로다른 모델의 아바타들 사이에 통신이 가능함을 실험하기 위해 가상환경 서버를 구축하였고, 가상환경 서버와 클라이언트들 사이에 FBML을 이용하여 통신하도록 하였다. 이하, 제 II장에서는 XML 관련 연구를 개괄한 후, 제 III장에서는 XML 데이터를 생성하기 위한 DTD를 설명한다. 제 IV장에서는 실험 결과를 고찰한 후, 제 V장에서 결론을 맺는다.

II. XML 관련 연구

최근 인터넷의 폭발적 사용 증가는 웹이라 불리는 WWW (World Wide Web)의 발전으로 인한 것이다. 이는 웹을 통하여 원하는 정보에 쉽게 접근할 수 있는 정보 접근의 용이성과 비전문가들도 쉽게 정보를 만들고 공유할 수 있는 HTML의 보편성에서 유래되었다. 그러나 HTML은 동적인 콘텐츠를 제공하거나, 문서 구조를 표현하지 못하는 한계를 가지고 있다.

이와 같은 제한점을 해소하고 사용자의 다양한 요구를 수용하기 위하여 1996년 W3C (World Wide Web Consortium)에서는 웹의 새로운 표준으로 XML을 채택하였다. XML은 문서에 대한 구조를 기술하는 메타언어이다. 웹상에서 구조화된 문서를 제공하는 것이 가능하도록 설계된 표준화된 텍스트 형식으로써, 문서의 내용과 연관된 태그를 사용자가 직접 정의하여 사용한다. 이는 체계적이고, 일관적인 접근 방식을 정의하고 많은 기술을 포괄할 수 있는 기본 프레임워크를 제공한다.

일반적으로 문서는 구조, 내용, 표현의 세 가지 구성요소로 이루어지는데, 서로 다른 응용 사이에서 사용하는 문서의 표현 방식이 상이한 경우에는 문서 교환이 어려워지게 된다. XML은 문서의 구성 요소 중 응용에 상관없이 공통적으로 사용되는 요소인 구조와 내용만을 기술하는 언어이므로, XML 포맷의 문서는 데이터 교환에 용이하다. 또한 구조를 기술하는 것은 문서를 파일 단위가 아닌 구조를 이루고 있는 각각의 태그 단위로 관리가 가능하므로, 문서 관리의 효율성을 높여주는 역할을 한다.

기존의 HTML에 비하여 XML이 가지는 장점은 다음과 같다. 첫째, XML은 HTML과는 달리 문서의 내용과 표현방식을 분리할 수 있다. 둘째, DTD와 XML 스키마를 이용하여 구조화된 문서 정의를 지원한다. 셋째, 확장된 링크 기능이 있다. 마지막으로, 문서 처리를 위한 표준 API를 제공한다.

XML 문서는 트리 형태의 데이터 구조를 가지고 있다. 이와 같은 형태의 자료 구조를 접근하기 위해서는 특별한 방법이 필요하다. DOM (Document Object Model)^[7]은 XML 문서의 각 요소들을 하나의 객체로 표현하고, 그 요소들 사이의 관계를 통해 각 요소에 접근할 수 있는 방법을 제공해주는 API이다.

다양한 XML응용 프로그램들은 대부분 DOM API를 XML 문서 조작의 표준 인터페이스로 사용하고 있다. DOM 인터페이스는 트리 조작에 필요한 함수들을 제공하기 때문에, 응용 프로그램이 DOM 인터페이스를 통해서 XML 문서를 조작하게 되면, 질의처리 시 노드 탐색을 쉽게 수행할 수 있을 뿐만 아니라, XML 문서의 트리 구조를 쉽게 변경할 수 있다. 이외에도 다양한 트리 조작 기능이 가능하다.

DOM은 저장소에 저장되어 있는 XML 문서들을 파싱해서, 그 결과로 생성되는 객체들을 메모리 상에 상주시킨다. 다양한 응용 프로그램들은 메모리 상에 있는 객체들을 DOM 표준 인터페이스를 통해서 참조하는데, 이런

방법을 통해서 응용 프로그램들은 XML 문서를 접근하고 조작할 수 있다. DOM은 XML 문서를 접근하기 위한 인터페이스의 표준으로 채택된 만큼 이용 가능성이 높다.

III. FBML을 위한 DTD 설계

사이버 공간에서 활동하는 아바타에는 간단한 사용자 별명에서부터 텍스트, 이미지, 3차원 모델 등 많은 종류들이 있다. 그림 3.1은 본 연구에서 고려한 일부 모델이다.

본 연구에서 정의한 FBML은 계층구조를 이루고 있다. 최상위 엘리먼트는 <FBML>이고, 하부구조는 <fbml_head>와 <fbml_body>로 나뉜다. 이는 HTML에서 <HTML>이 최상위 엘리먼트가 되며, 서브 엘리먼트로 <HEAD>와 <BODY>로 나누어지는 것과 같다. <fbml_head>에는 모델 종류와 네트워크 관련 내용을 기술하고, <fbml_body>에는 전달하고자 하는 내용을 기술한다. 다음은 DTD로 정의한 엘리먼트의 부분 예이다.

```
<!ELEMENT FBML (fbml_head , fbml_body)>
<!ELEMENT fbml_head (root-layout , model , commant* ,
    head_content , network_setting*)>
<!ELEMENT head_content (head_text , head_voice ,
    head_image , head_video)>
<!ELEMENT fbml_body (face_parameter ,
    body_parameter)>
<!ELEMENT face_parameter (face_content , face_method? ,
    facs? , mouth_content*)>
<!ELEMENT facs (expression* , AUs*)>
<!ELEMENT expression ANY>
```

```
<!ELEMENT AUs (Single_AUs* , Additional_AUs* , AU*)>
<!ELEMENT body_parameter (body_content , body_method? ,
    joint?)>
<!ELEMENT joint (left_body* , right_body* ,
    center_body*)>
... 생략 ...
```

1. <fbml_head> 엘리먼트 정의

엘리먼트 <fbml_head>는 다섯 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. <commant>는 참고 자료를 기술한다. 내용을 기술하지 않아도 상관없지만, 부가적으로 필요한 내용들을 포함시킨다. 네 번째는 <head_content>이며 아바타 상호간에 의사전달에 가장 중요한 의미를 전달하기 위해서 만들어 졌다. 즉, 아바타들 간에 의사전달은 텍스트, 음성, 이미지, 그리고 비디오 영상으로 구분 지었다. 아바타 통신에서 필요한 내용을 채워서 전달하면 된다. 마지막 <network_setting>은 서버와 클라이언트에 상호간에 필요한 네트워크 정보들을 포함하고 있으며, 서로 상호간에 의사전달을 위해 필요한 연결 설정에 관련된 사항이다. 포함되는 내용은 간단한 네트워크 연결부터 유비쿼터스에서 사용하는 IPv6에 까지 다양한 관련된 정보들도 포함할 수 있다.

2. <fbml_body> 엘리먼트 정의

엘리먼트 <fbml_body>는 <face_parameter>와 <body_parameter>를 가지고 있다. 아바타 통신에서 얼굴 표정이 중요한 역할을 하기 때문에 <face_parameter>를

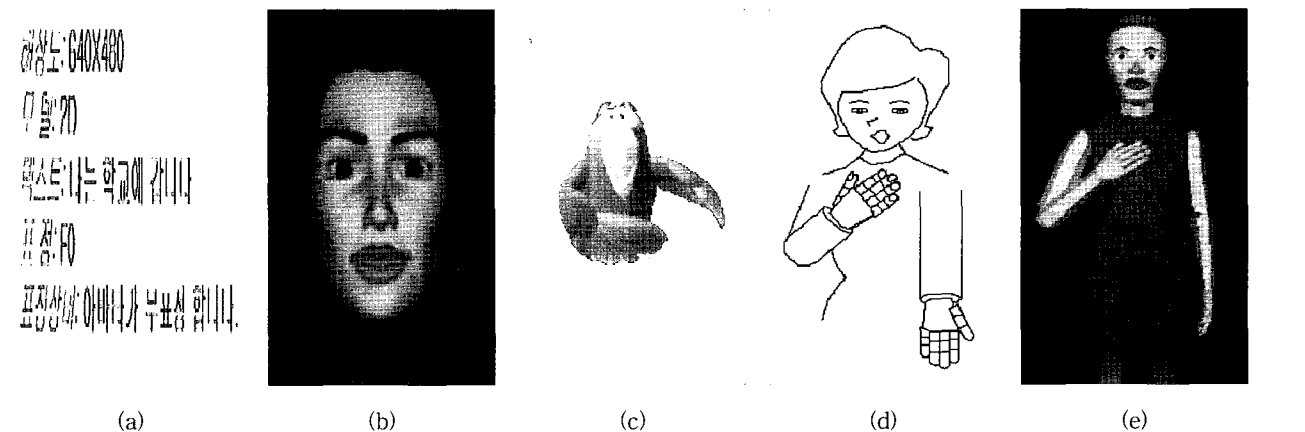


그림 3.1. 서로 다른 구조의 아바타 모델들. (a) 텍스트 정보, (b) 일반 표정 이미지, (c) 웹에서 동작하는 3D 모델 (MS-Agent), (d) 라인으로 구성된 2D 모델, (e) VRML에서 제공한 Nancy 모델
 Fig. 3.1. Avatar models having different structures. (a) Text Information, (b) General Expression Image, (c) 3D Model (MS-Agent) in Web, (d) 2D Model organized by Line, (e) Nancy Model using VRML.

포함하였다. <face_parameter>에서 기본으로 포함된 시스템은 <facs>이며, 그 밖의 다른 방법들을 지원하기 위해서 <face_content>, <face_method>, 그리고 <mouth_content>를 포함시켰다. <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 있다. 또한 AU들의 조합인 6가지 표정으로는 기쁨, 슬픔, 두려움, 역겨움, 화남, 그리고 놀람 표정이 있다. 이를 표정 정보를 전달할 수 있도록 <expression>을 포함하고 있다. 마지막으로 <body_parameter>는 아바타 통신에서 몸동작에 대한 정보를 전달하는 것으로, 특히 수화 동작에서 많은 영향을 끼치게 된다. <body_parameter>에서 기본으로 사용하는 관절각에 대한 정보로 <joint>을 이용한다. 관절각은 왼쪽, 오른쪽, 중심 기준의 세 부분으로 나누었다.

3. FBML에서 엘리먼트의 속성들

각각의 엘리먼트 내에는 엘리먼트의 이름과 구조, 그리고 반복 횟수를 포함하고 있다. 또한 각 엘리먼트는 한 개 이상의 속성을 가지며, <!ATTLIST>들로 정의한다. 예를 들어, 엘리먼트 <model>, <network_setting>, <fbml_body>, <face_method>, <expression>의 속성에 대한 DTD정의를 다음과 같다.

```
<!ATTLIST model kind (2D | 3D) "3D">
<!ATTLIST network_setting addressType (IPv4 | IPv6) "" ip
  CDATA #IMPLIED>
<!ATTLIST fbml_body type (MPEG4 | H-Anim) "">
<!ATTLIST expression expression_kind (joy | surprise |
  anger | sadness | fear | disgust) "" intensity CDATA
  #IMPLIED>
<!ATTLIST AU number (1 | 2 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
  11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22
  | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 |
  34 | 35 | 36 | 37 | 38 | 39 | 41 | 42 | 43 | 44 | 45 | 46)
  ""
intensity CDATA #REQUIRED>
<!ATTLIST mouth_content noname ENTITY #REQUIRED
  model (2D | 3D) "">
<!ATTLIST center_body number CDATA "" angle CDATA
  #REQUIRED model (2D | 3D) "">
... 생략 ...
```

여기서, 엘리먼트 <model>의 속성은 kind로 2D 또는 3D 두개의 기술이 가능하며, "3D"는 디폴트가 3D임을 의미한다. 또한, <network_setting>의 속성은 addressType으로, 어드레스 종류는 IPv4와 IPv6 중 선택이 가능하며 디폴트는 없다. 그리고 후반부에 기술된 'ip CDATA #IMPLIED'는 IPv4 (또는 IPv6)에 해당하는 네트워크 주소를 ip 속성의 문자로 기술할 수 있고, 반듯이 입력하지 않아도 되는 선택사항임을 나타낸다. 또한 엘리먼트 <fbml_body>의 속성은 type으로 VRML에서 제공한 H-Anim과 MPEG4 중에서 선택할 수 있음을 나타낸다. 같은 방법으로, 엘리먼트 <expression>의 속성은 expression_kind로 6가지 얼굴 표정을 기술할 수 있다.

그 밖에, 엘리먼트 <left_body>와 <right_body>를 이용하여 MPEG4에서 제공하는 관절각 정보들 까지 전달할 수 있도록 정의하였고, FAP(Facial Animation Parameter)를 지원하기 위하여 엘리먼트 <Single_AUs>에 33가지의 속성을 정의하였다. 또한 엘리먼트 <Additional_AUs>의 number속성에는 11가지의 속성 값 중에서 하나를 선택할 수 있도록 하였다.

IV. 실험결과

1. 실험 환경

본 논문에서 제안한 방법을 실험하기 위한 시스템은 펜티엄-4 1.8G을 사용하며, Windows XP 환경에서 Visual C++ 6.0, Open Inventor Library, MS-Agent (TTS, 음성인식), IIS (Internet Information Server) 웹 서버를 이용하여 구현하였다. 또한 XML을 사용하기 위해서 MSXML 4.0 DOM Parser를 사용하였다. 본 연구에 사용된 3D 아바타 모델은 VRML에서 제시하고 있는 사람과 비슷한 휴모노이드이다. 또한, 2D 모델은 전체 28개의 선으로 구성된 모델을 사용하였다.

2. FBML을 사용한 통신시스템의 전체구조

그림 4.1은 2D와 3D 모델을 같이 사용할 수 있도록 하는 전체 통신 시스템 구조이다. 기존 방식은 파라미터 방식을 사용하기 때문에, 2D와 3D 모델을 사용하는 각각의 서버와 클라이언트가 존재해야 한다. 통합 서버에서는 이러한 문제를 해결하기 위해서 FBML를 서버와 클라이언트 사이에 주고받을 수 있다.

그림 4.1의 통신 구조에서는 실제 웹상에 서비스를 제공하는 것에 문제가 거의 발생하지 않는다. 왜냐하면, XML이 웹을 바탕으로 한 서비스를 기본적으로 제공하

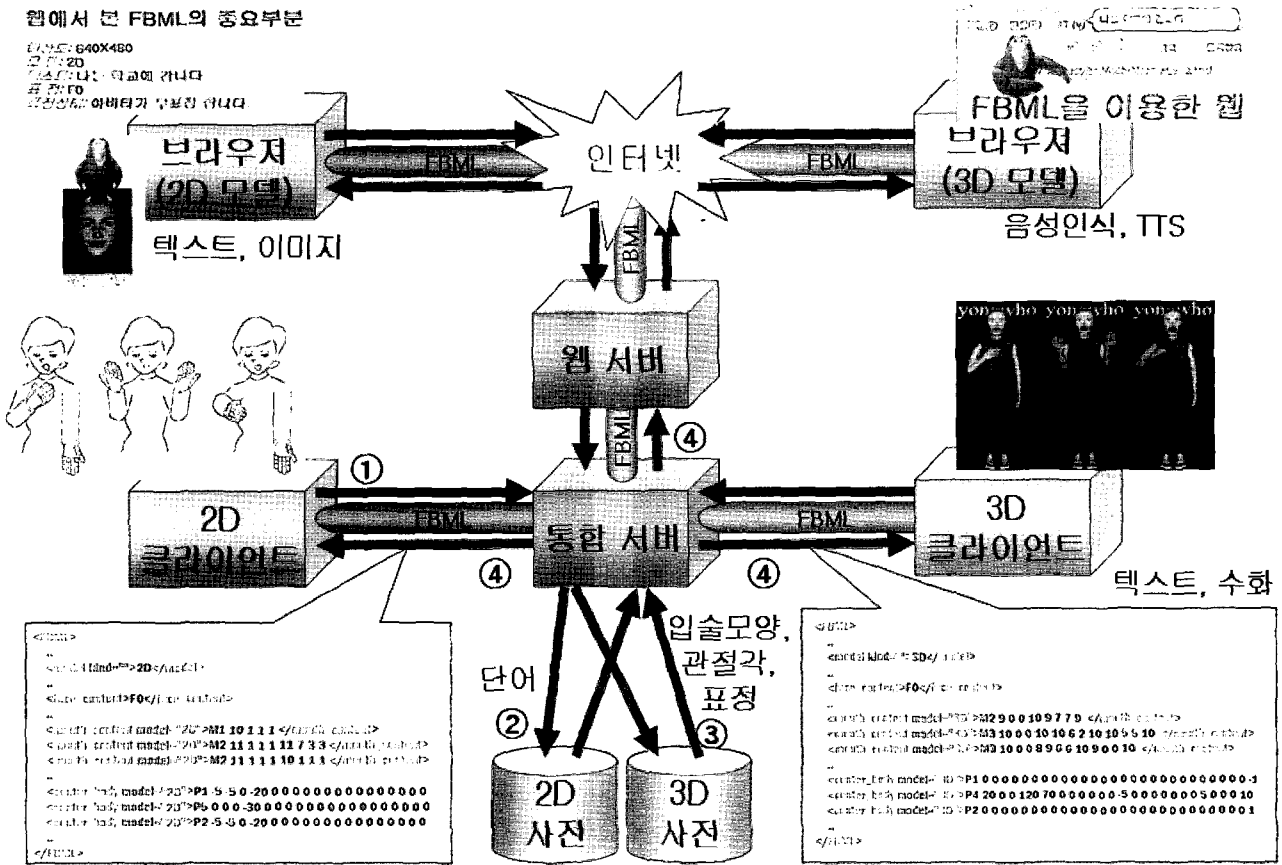


그림 4.1. 2D와 3D서버를 통합한 시스템 구조.
 Fig. 4.1. A system structure of combining 2D and 3D servers.

기 때문이다. 또한 웹에서 XML 문서를 보는 방법은 XSLT, DOM, 그리고 HTML 등 다양한 방법들을 이용하여 개발이 가능하기 때문에, 실제 확장 범위는 상당히 넓다.

그림 4.1의 통신 구조에서 통신과정은 네 가지 단계로, ① 클라이언트들은 서버에게 모델 정보와 텍스트 정보들을 FBML을 사용하여 전달하게 된다. ② 통합 서버는 모델에 관절각 정보를 얻기 위해서, 클라이언트에서 받은 문장 내용을 2D와 3D 사전에 보낸다. ③ 사전에서는 텍스트 정보를 이용하여 2D와 3D에 맞는 관절각 정보를 추출해서, 통합 서버에게 보낸다. ④ 통합 서버는 사전에서 가져온 관절각 정보들을 FBML에 포함시키고, 2D와 3D 클라이언트에게 각각 보내게 된다.

3. 클라이언트와 서버의 동작

그림 4.1의 시스템에서 처리되는 형태는 클라이언트에서 동작에 필요한 텍스트를 서버에게 전달하게 되면, 서버는 받아들인 정보를 텍스트형태에 맞게 아바타를 동작시킬 수 있는 정보들로 변환한다. 생성된 데이터는 입술

의 움직임과 몸의 관절각 정보들을 포함하게 된다. 또한 부가적인 요청이 있을 때 필요한 사항으로, 네트워크 연결에 대한 정보들을 포함시켜야 한다. 또한, 현재 시스템에서는 <network_setting>의 속성인 addressType에 IPv4가 사용되었고, 자신의 IP 주소를 ip 속성에 포함시켜 보내게 된다. 현재는 자신의 주소를 "127.0.0.1"로 하였다. 그림 4.2는 클라이언트에서 서버로 보낸 FBML의 예이다. 또한, 그림 4.3은 서버에서 클라이언트에게 보내는 FBML의 예이다.

그림 4.3에서처럼 2D와 3D 모델에서는 네 가지 정보를 필요로 한다. 먼저 ID와 코드는 <network_setting>에 저장되며, 얼굴 표정 정보는 <face_contents>에, 입술 정보는 <mouth_contents>에 있다. 그리고 마지막으로 몸 동작에 대한 정보는 <center_body>에 저장된다. 이들 네 개의 엘리먼트를 보면, 기존 시스템에서 사용된 파라미터 정보들이 모두 포함되어 있음을 알 수 있다.

이러한 정보를 받은 클라이언트 쪽에서는 관절각 정보를 이용하여 3D 아바타를 동작시킬 수 있게 된다. 물론, 2D 모델도 3D와 마찬가지로 동작이 가능하다는 것을 그

```
<?xml version="1.0" encoding="euc-kr" standalone="no" ?>
- <FBML>
  - <fbml_head>
    <root-layout>640X480</root-layout>
    <model kind="">2D</model>
    <network_setting>yongwho 4 0 0</network_setting>
    <head_content>
      <head_comment>주석</head_comment>
      <head_text>나는 학교에 갑니다</head_text>
      <head_voice>null</head_voice>
      <head_image>null</head_image>
      <head_video>null</head_video>
    </head_content>
  </fbml_head>
  - <fbml_body>
    <face_parameter>
      <face_content>F1</face_content>
      <face_method kind="facs">null</face_method>
    </face_parameter>
    <facs>
      <expression expression_kind="joy" intensity="100" />
    </facs>
    <AUs>
      <Single_AU number="" intensity="">null</Single_AU>
      <Additional_AU number="" intensity="">null</Additional_AU>
      <!-- <AU number="1" intensity="100"/> -->
    </AUs>
    </face_parameter>
    <!-- <mouth_content model="">입술 모양</mouth_content> -->
  </fbml_body>
  - <body_parameter method="">
    <body_content>null</body_content>
    <body_method>null</body_method>
  </body_parameter>
  </fbml_body>
</FBML>
```

2D 모델에서 참조

3D 모델에서 참조

그림 4.2. 클라이언트에서 서버로 보낸 FBML.

Fig. 4.2. A FBML message sent from Client to Server.

```
<?xml version="1.0" standalone="yes" ?>
<!DOCTYPE FBML (View Source for full doctype...)>
- <FBML>
  - <fbml_head>
    <root-lyout>640X480</root-lyout>
    <model kind="">3D</model>
    <network_setting>yongwho 4</network_setting>
    <head_content>
      <head_comment>주석</head_comment>
      <head_text>나는 학교에 갑니다</head_text>
      <head_voice>null</head_voice>
      <head_image>null</head_image>
      <head_video>null</head_video>
    </head_content>
  </fbml_head>
  - <fbml_body>
    <face_parameter>
      <face_content>F0</face_content>
      <face_method kind="facs">null</face_method>
    </face_parameter>
    <facs>
      <expression expression_kind="joy" intensity="100" />
    </facs>
    <AUs>
      <Single_AU number="" intensity="">null</Single_AU>
      <Additional_AU number="" intensity="">null</Additional_AU>
      <!-- <AU number="1" intensity="100"/> -->
    </AUs>
    </face_parameter>
    <!-- <mouth_content model="">입술 모양</mouth_content> -->
    <!-- <mouth_content model="2D">M1 10 1 1 1 1</mouth_content> -->
    <!-- <mouth_content model="2D">M2 11 1 1 1 1 7 3 3</mouth_content> -->
    <!-- <mouth_content model="2D">M2 11 1 1 1 1 10 1 1 1</mouth_content> -->
    <!-- <mouth_content model="3D">M2 9 0 0 10 9 7 7 9</mouth_content> -->
    <!-- <mouth_content model="3D">M3 10 0 0 10 10 6 2 10 10 5 5 10</mouth_content> -->
    <!-- <mouth_content model="3D">M3 10 0 0 8 9 6 6 10 9 0 0 10</mouth_content> -->
  </fbml_body>
  - <body_parameter method="">
    <body_content>null</body_content>
    <body_method>null</body_method>
  </body_parameter>
  </fbml_body>
  </FBML>
```

2D 모델에서 참조

3D 모델에서 참조

그림 4.3. 서버에서 클라이언트에 보낸 FBML.

Fig. 4.3. A FBML message sent from Server to Client.

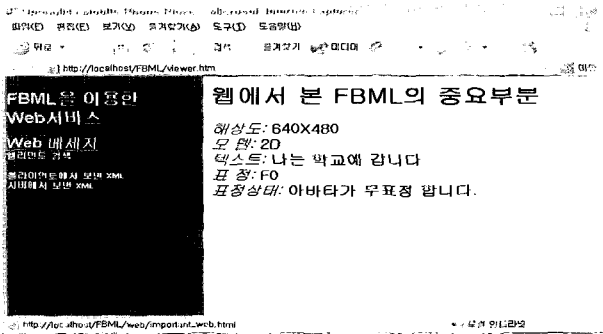


그림 4.4. 웹브라우저의 텍스트 동작 예.
Fig. 4.4. Text information on Web browser.

림 4.2와 그림 4.3을 보면 알 수 있다.

4. 웹 브라우저의 텍스트 동작

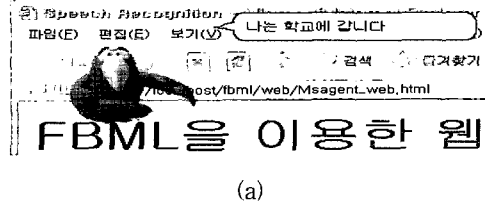
제안된 시스템에서 생성된 FBML 내용은 웹 서버와 연결되어 있으며, 서버에 어떠한 동작이 이루어지는 경우에 웹 서버는 변경된 내용을 감지하여 클라이언트에게 보내준다. 웹 서버로부터 받은 FBML 내용을 받아서 웹상에 보여주는 방법은 다양하다. 주로 XSLT, DOM, 그리고 HTML이 사용된다. 그림 4.4는 FBML에서 해상도, 모델 종류, 텍스트 정보, 그리고 표정 정보들을 추출하여 웹 브라우저에서 텍스트 형태로 보여주는 화면 예이다.

5. 웹 브라우저의 3D 동작

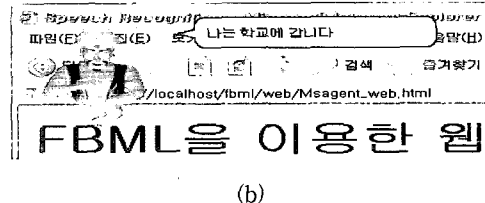
웹에서는 HTML을 이용하여 텍스트 정보뿐만 아니라, 3D 모델을 지원하고 있다. 본 실험에서는 MS-Agent^[9]라고 하는 3D 모델을 이용하여 표정 및 의사를 전달하는 방법을 검토하였다. MS-Agent는 3D 모델에 세 가지 주요 컴포넌트를 제공하고 있으며, 이들 컴포넌트를 사용하기 위해서 웹 브라우저에서는 <OBJECT> 태그를 이용하여 설정을 해주어야 한다. 먼저, 첫 번째 컴포넌트는 MS-Agent이며, 3D 캐릭터 정보에 대한 정보를 가진다. 그림 4.5에서는 FBML로부터 구성한 네 종류 캐릭터 예를 보이고 있다. 여기서 그림 4.5(a)는 "Peedy"라고 하는 앵무새 캐릭터이다.

두 번째 컴포넌트는 TTS(Text-To-Speech)로, FBML에서 받아들이는 텍스트 정보를 음성으로 변환한다. 영문과 한글을 포함하여 모두 11개 언어를 지원한다. 마지막 세 번째 컴포넌트는 사용자 음성을 인식할 수 있는 음성인식 기능을 지원한다.

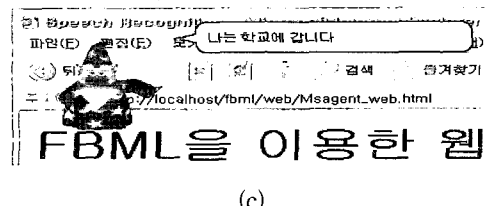
그림 4.5는 통합서버에서 보낸 FBML 문서를 이용하여 Peedy를 작동시키는 예이다. Peedy는 미리 정의된 90가지의 동작만 가능하며 수화동작은 할 수 없지만, TTS를



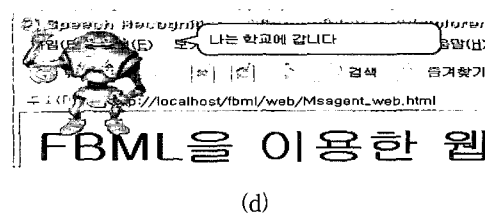
(a)



(b)



(c)



(d)

그림 4.5. FBML로부터 구성한 네 종류 캐릭터 예. (a) Peedy, (b) Merlin, (c) Robby, (d) Genie.

Fig. 4.5. Four characters generated by using FBML. (a) Peedy, (b) Merlin, (c) Robby, and (d) Genie.

지원한다. 이러한 컴포넌트를 이용함으로써, FBML에서 추출한 텍스트 정보를 음성으로 변환할 수 있다. 따라서 수화 통신 시스템이 가지고 있는 시각적인 정보들 외에 다른 정보를 결합된 멀티미디어 서비스를 제공할 수도 있다. 예를 들면, 제스처 이외에 표정 및 음성 정보를 추가한 응용을 고려할 수 있다.

6. 처리시간의 비교

기존 시스템과 제안된 시스템의 차이점은 통신 메시지의 형식이 다르다는 점이다. 기존 시스템에서는 특정 송수신 시스템에 맞도록 최적화된 형식으로 전송한다. 따라서 이 방법은 해당하는 두 시스템에는 가장 효율적일 지라도 다른 시스템에서는 사용할 수 없게 된다. 반면에 본 논문의 FBML은 웹과 웹이 아닌 시스템에서도 사용할 수 있도록, XML을 기반으로 작성되었다. 또한 그 밖에 다양한 시스템에서 사용할 수 있도록 FACS나 관절각 정보를 포함시켰다.

표 4.1. 처리시간 비교(단위: 밀리초).

Table 4.1. A comparison of processing CPU-times (unit: msec).

| 2D | 2D-XML | 3D | 3D-XML |
|-------|--------|-------|--------|
| 45.10 | 183.43 | 10.51 | 188.76 |

표 4.1은 수화 문장 “나는 학교에 갑니다”에 대한 처리시간을 비교한 결과이다. 여기서 첫 번째와 세 번째 열에 있는 “2D”와 “3D”는 기존시스템의 처리시간이며, 두 번째와 네 번째의 “2D-XML”과 “3D-XML”는 제안된 시스템에서 채택한 FBML을 사용한 결과이다. 여기서 측정된 시간은 단순한 전송시간으로, 클라이언트에서 서버로 보낸 처리와 서버에서 클라이언트에게 보낸 시간을 합한 것이다. 즉, 시작 시간은 클라이언트 쪽에서 보내기 버튼을 클릭한 상태이며, 마친 시간은 클라이언트에게 FBML이 도착한 상태이다. 단위는 밀리 세컨드 (ms)이다.

기존 시스템인 메시지 형식은 각 양단간에 최적화된 형식을 갖추어 전송한다. 그러므로 처리되는 시간은 다른 시스템 보다 빠를 수밖에 없다. 그러므로 기존의 방법 (2D)에서는 최대 45ms까지 시간이 걸리지만, FBML을 사용하는 경우 (2D-XML)에는 183ms, 4.5배 정도의 부하가 생겼다. 그러나 180ms도 사람이 느낄 수 없는 실시간 범주이다. 또한 기존 시스템에서 사용한 메시지 형식은 특정한 양단간 시스템들 간에 처리시간이 단축되지만, 그 밖에 다른 서비스를 제공하기 위해서는 다양한 메시지 형식을 고려해야 한다. 이러한 다양한 메시지 형식을 지원 할수록 시스템의 부하는 더욱 커지게 된다. 반면에 FBML에서는 다양한 시스템을 지원하면서도 시스템의 부하는 증가하지 않는다. 이점은 다음 절에서 고찰한다.

7. 전송할 파라미터의 비교

표 4.2는 그림 3.1에 나열되어 있는 다섯 개 모델이 지원할 수 있는 파라미터와 지원하지 못하는 파라미터를 비교하였다. 먼저, 파라미터를 아홉 가지 종류로 나눈 후, 각 모델에서 몇 개의 파라미터를 지원하는지 분석하여, 각각을 기술하기 위해서 필요한 정보량을 조사하였다.

표 4.2에서 각각의 파라미터 정보 단위는 바이트이다. 여기서 정의한 일정한 크기란 엘리먼트가 가지는 공간의 크기를 의미한다. 또한 효율성을 측정하기 위해서 사용율을 정의한다. 사용율이란 FBML을 사용한 바이트 수를 각각의 모델이 사용한 바이트 수의 합으로 나눈 것이다. 여기서 측정된 바이트 수란 실제 주고받은 데이터가 아니라 미리 정의해 놓은 저장공간이다. 즉, 실제 전송 데이터

표 4.2. 그림 3.1의 모델들이 지원하는 파라미터들.

Table 4.2. Parameters employed in the models of Fig. 3.1.

| 모델 | 모델 정보 | 네트 워크 정보 | 내용 정보 | | | | 얼굴 동작 | | 몸 동작 |
|-----|-------|----------|-------|-----|-----|-----|-------|----|------|
| | | | 텍스트 | 이미지 | 오디오 | 비디오 | 입술 | 표정 | |
| (a) | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 6 | 0 |
| (b) | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 6 | 0 |
| (c) | 4 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 90 |
| (d) | 1 | 1 | 1 | 0 | 0 | 0 | 11 | 20 | 57 |
| (e) | - | 1 | 1 | 0 | 0 | 0 | 11 | 20 | 57 |

가 아니라 미리 정의해 놓은 태그 공간을 의미한다.

표 4.2에서 첫 번째 (a) 텍스트 (웹기반)는 네트워크, 텍스트, 얼굴 표정의 세 가지 서브 엘리먼트만을 가진다. 따라서 총 사용된 바이트 수는 295(= 네트워크 40 + 텍스트 52 + 표정 203)개가 된다. 같은 방법으로, (b) 이미지 2D의 경우 322, (c) 웹 기반 MS-Agent 3D 모델은 730, (d) 라인 2D 모델은 730, (e) VRML 3D 모델은 730 바이트를 사용한다. 한편, FBML을 정의하기 위해 필요한 전체 바이트 수는 대략 826 (= 모델 27+네트워크 40+텍스트 52+이미지 32+오디오 32+비디오 32+입술 56+표정 203+동작 352)이다. 이는 FBML이 가지고 있는 모든 특성 공간을 나열하여 합한 것이다.

따라서, 다섯 모델을 각각 정의하였을 경우에, 중복된 파라미터의 정보를 제거함으로써 생성되는 이용율은 0.294 = 826/(295+322+730+730+730)이다. 즉, 모델의 종류가 많을수록 이용율은 감소하여 결국 네트워크의 부하를 줄일 수 있게 된다. 또한 다양한 아바타 모델들을 사용하면 할수록 데이터의 중복으로 많은 부분을 제거할 수 있게 되어, 전송할 데이터는 아바타 수에 따라 선형으로 증가되지 않게 됨을 알 수 있다.

V. 결 론

본 논문에서는 서로 다른 구조를 갖는 아바타 모델들 간의 통신을 지원하기 위해 FBML (Facial Body Markup Language)의 표준 문서를 정의하여 사용하는 방안을 제안하였다. FBML은 웹 서비스를 기본으로 하기 때문에 웹과 웹이 아닌 환경을 모두 서비스할 수 있다. 또한 MPEG과 VRML 모델에서 기본으로 채택한 6가지 표정을 사용할 수 있다.

간단한 실험으로, FBML 문서를 이용하여 서로 다른

환경에서 다양한 아바타 모델 간에 통신할 수 있음을 확인하였다. 또한 제안된 시스템을 처리시간과 데이터 량 측면에서 기존 시스템과 비교하였으며, 데이터 양은 아바타 수가 많아 질수록 선형적으로 증가하지 않음을 알 수 있었다. 또한 FBML을 이용한 TTS 등 다양한 서비스들의 가능성을 검토하였다. 그러나, 아바타의 구조와 모양이 너무 다양하다. 따라서 모든 아바타의 구조를 지원할 수는 없다는 한계는 여전히 남아 있다.

참 고 문 헌

[1] 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.

[2] 金商雲, 李種雨, 青木由直, "3D 수화 애니메이션을 위한 팔과 손의 통합된 키 프레임 에디터," 대한전자공학회는문지, 제37권 SP편, 제9호, 371-380쪽, 2000년 9월.

[3] 金商雲, 吳芝英, 青木由直, "인터넷상에 3차원 모델을 이용한 한-일 실시간 수화 통신 시스템의 구축을 위한 기초적인 검토," 대한전자공학회는문지, 제36권 S편, 제7호, 811- 820쪽, 1999년 7월.

[4] S.-W. Kim, Y.-W. Lee, Y. Aoki, and Y. Arakawa, "A comic expression method of universal emotions for intelligent avatar communications using principal action units" Proc. of ITC-CSCC, Tokushima, Japan, pp. 240-243, Jul 2001.

[5] 李容厚, 金商雲, 青木由直, "인터넷 가상공간에서 지적 아바타 통신을 위한 코믹한 얼굴 표정의 생성법," 대한전자공학회는문지, 제40권 CI편, 제1호, 59-73쪽, 2003년 1월.

[6] International Organization for Standardization, "ISO 8879: Standard Generalized Markup Language (SGML)", 1986.

[7] W3C, "Document Object Model (DOM) Level 2 Core Specification", Nov 2000, <http://www.w3.org/TR/DOM-Level-2-Core/>.

[8] <http://www.research.microsoft.com/easyliving>

[9] <http://www.microsoft.com/msagent/default.asp>

저 자 소 개



이 용 후(학생회원)
2000년 2월 울산대학교 컴퓨터 공학과 (공학사).
2002년 2월 명지대학교 컴퓨터 공학과 대학원(공학석사)
2002년 3월~현재 명지대학교 컴퓨터공학과 대학원 박사과정

<주관심분야: 패턴인식, HCI>



김 상 운(정회원)
제40권 CI편 제1호 참조



아오끼 요시나오
(전자정보통신학회, 정회원)
1964년 일본 홋카이도대학 졸업.
1966년 동대학원 석사과정 졸업.
1969~1971년 캐나다 라발대유학.
1966년~현재 홋카이도대 대학원 전자정보공학연구과 교수

<주관심분야: 신호화상처리, 컴퓨터그래픽스>
北海道新聞文化賞(2000년). 北海道文化賞 (2001년)
北海道科學技術賞 (2003년) .
주요저서: 「オペレータ法デジタル信號處理」 (コロナ社), 「コンピュータグラフィックス講義」 (コロナ社), 「情報リテラシとプレゼンテーション」 (コロナ社)、情報處理學會펠로우, 電子情報通信學會펠로우.