특집

MPEG-4 기반 오감 데이터 전송 및 재생 기술

양일식 · 김명규 · 손지연 · 이용희 · 박준석(한국전자통신연구원)

I. 서 론

HCI(Human-Computer Interaction)는 언제 어디서나 편리한 인터페이스로 생각의 범위를 제한하지 않는다. 사용자의 상황에 반응하여 적응적으로 인터페이스를 제공하는 상호작용 기술로서 사용자에게 불필요한 일을 요구하지 않으며 시스템의 존재를 느끼지 않고 원하는 작업에 전념하도록 하기 때문이다. 차세대 컴퓨팅 패러다임으로서 유비쿼터스 컴퓨팅 기술이 HCI를 중요한 이슈로 다루는 것은 사람과 컴퓨터 사이의 인터페이스가 컴퓨팅 분야에서 중요한 요소 기술임을 보여준다[1].

기존의 시청각 위주의 상호작용 기술은 촉각, 후각, 미각을 포함한 오감 서비스를 지원하는 방향으로 발전하고 있다. 또한 IT 인프라의 발전은 고정단말 뿐 아니라 이동단말의 연결성을 보장하면서 큰 대역폭을 지원하여 이동 중 멀티미디어 서비스를 제공할 것으로 예상된다. 결과적으로 멀티미디어에 촉각, 후각, 미각을 포함한 오감 콘텐츠가 필요하고 이러한 오감 콘텐츠를 사용자와 상호작용하면서 시공간의 제약 없이 서비스할 수 있는 시스템이 요구된다.

최근 HCI에서 상호작용 콘텐츠는 많은 연구가 진행되고 있으며 중요 토픽으로 다루어지고 있다[2]. 그러나 많은 연구가 네트워크를 고려하지 않고 있으며, 네트워크 기반 실시간 상호작용 미디어에 대한 연구는 미미한 실정이다. 특히 오감을 포함한 네트워크 기반 상호작용 콘텐츠는 개념화 단계라고 볼 수 있다.

본 고는 유비쿼터스 컴퓨팅 환경에서 중요한 이슈중 하나인 HCI기술에서 네트워크기반 오감 상호작용 콘텐츠와 이를 지원하기 위한 시스템을 구현한다. 오감 콘텐츠 모델은 MPEG-4의 데이터 포맷에 촉각, 후각데이터를 추가하여 제시하고 시스템 구조는 MPEG-4 표준 시스템에 기반 하여 구현한다. 이에 대해 MPEG-4의 전반적인 개요를 II절에서 알고보고 III절에서는 MPEG-4 기반 오감 데이터 전송 시스템에 대한 구현 내용을 기술한다. 그리고 IV절에서 MPEG-4 기반 오감 시스템에 대한 결론을 서술한다.

II. MPEG-4 개요

MPEG은 ISO/IEC내의 표준화기관으로서 1988년에 AV(Audio Visual)의 인코딩, 디코딩, 프로세싱을 표준화하기 위해 시작되었다[3]. MPEG-1, MPEG-2, MPEG-4, MPEG-7, MPEG-21 등의 AV 기술을 표준화하였으며 이 중 MPEG-4는 웹과 이동성을 고려한 것으로 고압축, 저전송률에 초점을 두고 제정되었다. 본 절에서는 MPEG-4의 주요 특징을 살펴보고 MPEG-4 구성 요소에 대해 간략히 소개한다.

1. MPEG-4 주요 특징

가. AV 데이터의 객체 기반 표현

MPEG-4는 객체 단위로 데이터를 처리하여 하나의 화면을 구성하고 장면들에서 여러 객체를 서로 통합하여 다룰 수 있다. 화면을 기술하고 객체사이의 관계를 구성하기위해 장면 기술 언어로 BIFS(Binary Format For Scene)를 사용한다.

나. AV 객체와 사용자 사이의 내용 기반 상호작용

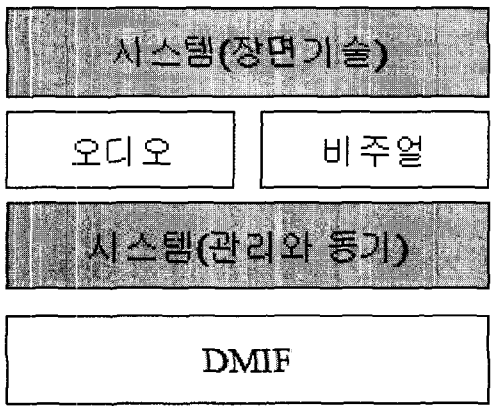
사용자가 BIFS언어로 표현된 장면의 각 객체와 상호작용하면서 장면을 동적으로 구성하게 한다.

다. 전송환경에 적응적이고 오류에 강인함

통신 상황에 따라 다양한 전송률을 지원하는 적응적인 코딩 기술을 사용하였으며 전송 중 오류가 발생하더라도 이미 수신된 데이터를 통해 예측할 수 있도록 하였다.

라. 다중화 및 동기화

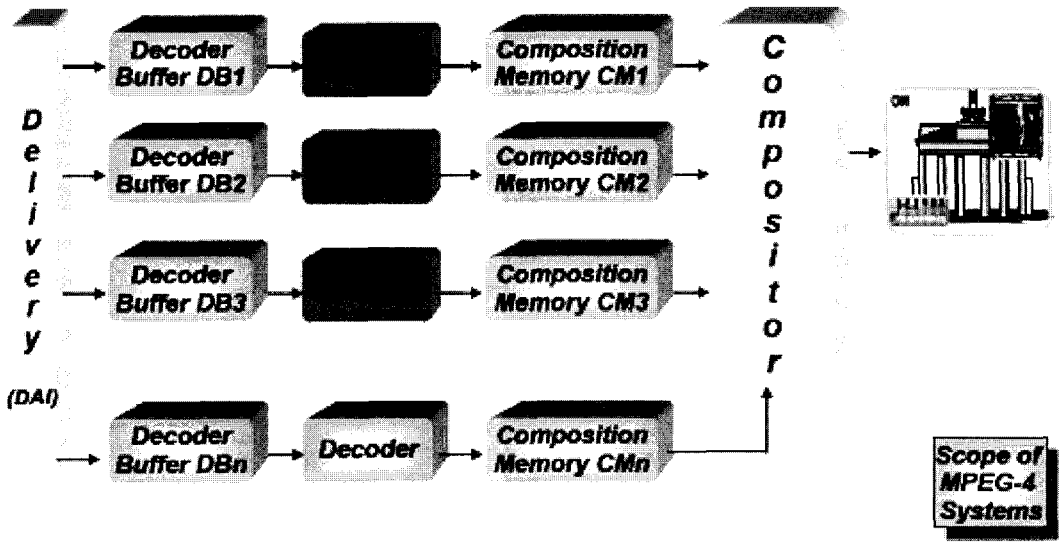
기본 스트림을 다중화하여 전송 효율을 고려했으며 동기화 계층을 두어 객체의 디코딩 시간과 화면에 구성되는 시간을 고려한 동기화를 지원하도록 하였다.



〈그림 1〉 MPEG-4 구조

2. MPEG-4 구성 요소

<그림 1>MPEG-4 구조는 시스템과 오디오, 비주얼, DMIF(Delivery Multimedia Integration Framework)와의 관계를 보여주고 있다. 비주얼과 오디오는 코딩 방법에 대한 표준이며 시스템은 비주얼과 오디오를 화면에 표현하는 측면과 미디어 스트림 동기화와 관리하는 방법을 다룬다. DMIF는 네트워크기반 전송 프레임워크를 제공한다.



〈그림 2〉 MPEG-4 디코더 모델[3]

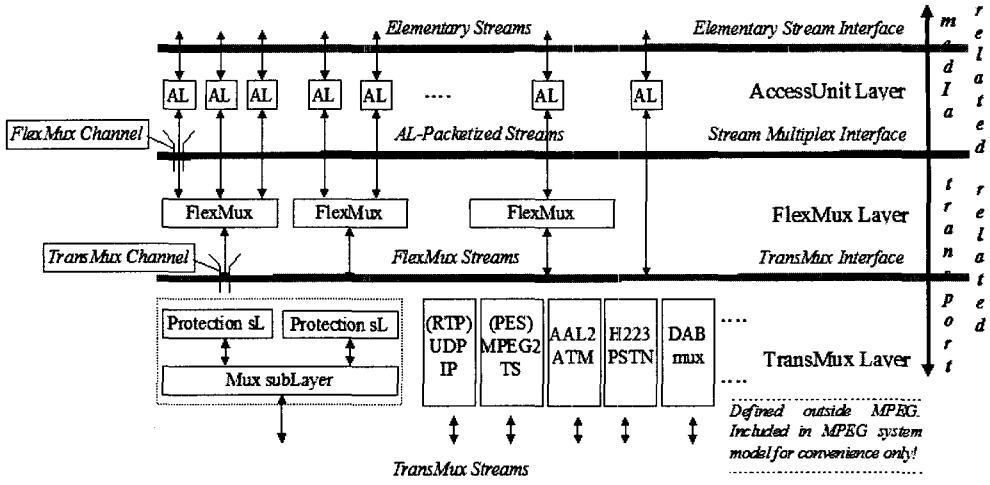
가. 비주얼/오디오

비주얼과 오디오분야 신호의 부호화 및 복호화 관련 규정을 기술하고 있다. 비주얼 기술은 객체 단위로 부호화하여 영상을 구성하며 효율성, 오류 내성, CG(Computer Graphic)영상 부호화, 법위성(Scalability)의 특징을 갖는다. 오디오 부호화 기술은 범용적인 부호화 방식을 목표로 개발되었다. 시간주파수 변환 부호화, 음성신호의 CELP(Code Excited Linear Prediction) 부호화, 저비트율/응악용의 매개변수 부호화, 계층 부호화 구조를 지원한다.

나. 시스템

<그림 1> MPEG-4 구조에서와 같이 시스템 부분은 멀티미디어 객체를 재현하기 위한 오디오/비주얼 정보를 감싸고 있는 구성요소로서 객체기반의 데이터 시스템 모델을 제시한다. 네트워크를 통해 전달된 기본 스트림과 화면과의 관계를 정의하기 위해 객체 디스크립터를 기술하였으며 미디어 객체의 시공간 관계를 기술하기 위한 포맷을 정의하였다. 객체 디스크립터와 객체 관계를 표현하는 포맷은 VRML(Virtual Reality Modeling Language)에 기반한 BIFS언어를 사용하여 기술된다.

MPEG-4 시스템은 스트림의 동기화를 위해 싱크레이어를 포함한 시스템 디코더 모델<그림 2>을 정의하였다. 기본 스트림은 엑세스 유닛으로 분할되며 복호/합성을 위한 시간 및 동기 처리 단위가 된다. Decoder Buffer는 엑세스 유닛을 저장하는 버퍼이고 Decoder는 엑세스 유닛을 입력으로 데이터의 기본단위인 부호화 데이터를 생성하여 Composition Memory에 저장된다. Compositor는 저장된 부호화 데이터를 화면에 구성한다. 엑세스 유닛 헤더는 시간정보를 포함하고 있어 디코딩과 융합 시간에 대한 동기화 정보를 제공한다.



〈그림 3〉 DMIF 구조[3]

다. DMIF

DMIF는 MPEG-4 표준의 6번째 파트로 콘텐츠 접근의 투명한 인터페이스를 위해 기술되었다. 사용자 데이터와 명령에 대한 별도의 포맷을 지원하며 세션 및 채널을 관리하는 메커니즘이 포함되어 있다. <그림 3>DMIF의 구조에서 DMIF구성요소는 ALFlexMux Layer와 TransMux Layer이다. 여러개의 엑세스 유닛 패킷을 다중화하여 전송효율을 높일 수 있는 FlexMux Layer를 두고 있으며, TransMux Layer는 다양한 통신 인터페이스를 서비스하기 위한 전송 기능을 수행한다.

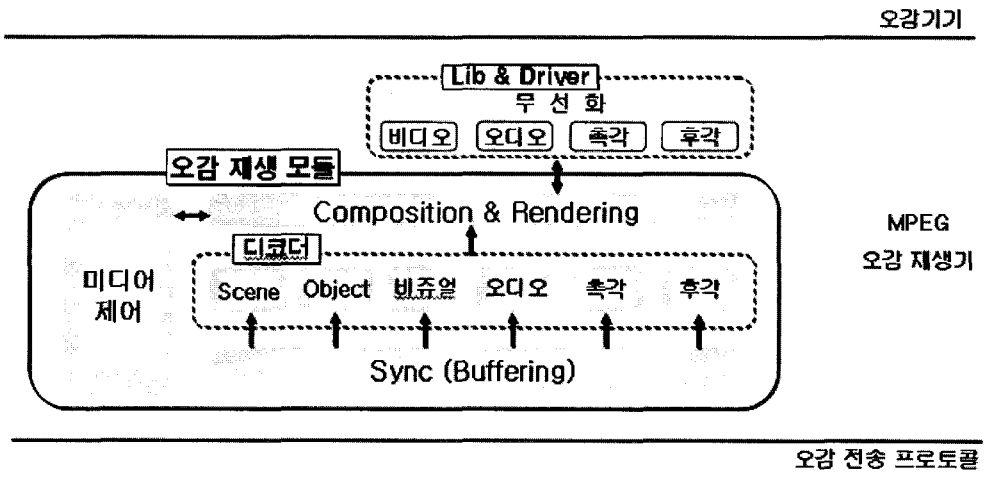
III. MPEG-4 기반 오감 데이터 전송 시스템

MPEG-4는 시청각 중심의 표준이지만 새로운 감각 데이터를 추가하기에 용이하도록 구성되어 있다. 각 스트림 요소를 객체로 처리하도록 하여 스트림 사이의 독립성을 지원하는 구조를 갖기 때문이다. 또한 객체사이의 관계를 기술할 수 있는 장면기술 언어인 BIFS를 사용하여 추가된 객체를 쉽게 융합, 표현할 수 구조를 지원한다.

본 절은 MPEG-4에 오감 데이터를 전송, 표현하기 위한 오감 전송 시스템 모델을 살펴보고 오감 데이터 재생을 위한 MPEG-4 오감 재생기 구조를 알아본다.

1. 시스템 개요

오감 전송 시스템은 클라이언트/서버 모델 기반으로 실시간 오감 상호작용 스트리밍 서비스를 위해 개발되었다. 클라이언트는 MPEG-4 표준 레퍼런스 소프트웨어[3]를 참조하여 개발하였고 UBC (University of British Columbia) MPEG-4 streaming 프로그램[4]을



〈그림 4〉 MPEG 오감 재생기 구조

수정하여 서버를 구성하였다. 오감 전송 프로토콜은 기존 DMIF와 호환을 유지하면서 각 감각의 특성을 고려한 전송 서비스를 제공하기 위해 설계되었다.

데이터 포맷은 MPEG-4 표준 파일 형식에 기반한 오감 MP4 파일구조를 갖는다. 촉각 및 후각의 부호화 데이터는 XML 형식으로 인코딩되어 MP4 저작툴을 이용하여 MP4 파일을 제작하였다. MP4 저작툴은 오픈 MPEG-4 저작툴인 GPAC 프로그램[5]에 촉각, 후각을 처리할 수 있는 모듈을 추가하여 개발하였다.

2. MPEG 오감 재생기 구조

<그림 4>는 MPEG-4 오감 전송 시스템의 클라이언트 프로그램으로서 MPEG 오감 재생기의 구조를 보여준다. MPEG 오감 재생기는 미디어 제어, Sync layer, 디코더, Composition & Rendering으로 구성된 재생 모듈과 비디오, 오디오, 촉각, 후각 관련 Lib & Driver로 구성된다. 미디어 제어 모듈은 주 프로그램으로 세션 및 데이터 송수신 제어를 담당하고 Composition & Rendering 모듈과의 인터페이스를 통해 사용자와의 상호작용 서비스를 제어한다.

Composition & Rendering 모듈은 디코딩된 스트림 객체를 화면에 구성하고 표현하여 사용자와의 인터페이스를 담당한다. 융합모듈은 Scene 디스크립터 정보를 통해 화면을 구성하는 역할로서 촉각 및 후각데이터에 대한 디스크립터를 추가하여 오감데이터를 처리하도록 한다. 표현 모듈은 기존 시청각에 후각에 대한 인터페이스를 정의하여 재현한다.

디코더 모듈은 Scene, Object, 비주얼, 오디오, 촉각, 후각으로 구성되는데 시청각 데이터에 촉각, 후각데이터를 디코딩 할 수 있는 XML 파서 프로그램이 추가된 구조이다. Sync(Buffering) 모듈은 MPEG-4의 Sync 레이어의 기능을 동일하게 사용하였다.

〈그림 5〉 MPEG-4 오감 재생기 실행 화면

3. 결과

<그림 5>은 MPEG 오감 재생기의 실행화면으로 대화형 오감 콘텐츠를 보여준다. 대화형 오감 콘텐츠는 MP4 파일형식으로 시청각 및 촉각, 후각데이터를 포함하고 있다. 사용자와의 상호작용을 통해 서버에 있는 데이터를 실시간으로 스트리밍하여 재생된다. 사용자는 MPEG-4 오감 재생기 실행 화면에서 <향기요청> 버튼을 클릭하여 발향서비스를 받을 수 있으며 직접 콘텐츠와 상호작용하며 다양한 서비스 시나리오를 전개할 수 있다.

Ⅳ. 결 론

오감정보 처리 기술에서 시청각은 안정화 단계라고 볼 수 있으나 촉 · 후 · 미각은 앞으로 많은 연구가 되어야 한다. 선진국에서도 연구가 미흡한 상황이어서 원천기술을 선점할 수 있는 좋은 시기라고 볼 수 있다. 특히 오감 융합측면에서의 연구는 시작 단계이다.

본 고에서는 MPEG-4 시스템을 활용하여 용이하게 개발할 수 있는 오감 융합 서비스를 제시하고 프로토타입을 구현하였다. BIFS 3D 관련 툴을 활용하면 3D환경으로 오감 융합 서비스 시스템을 확장할 수 있을 것으로 기대한다.

MPEG-4 시스템에 촉 · 후 · 미각 정보를 추가하여 오감 콘텐츠 전송 및 재생 시스템을 설

계 구현하는 방법은, 기존의 시청각 시스템과의 상호호환성을 유지하며 증강된 정보 서비스를 제공하는 자연스러운 시도이다. 본 고에서는 XML 기반 촉 · 후각 부호화를 MPEG 시스템에 적용하여 오감 전송 및 재생 시스템을 구현하였는데, 향후 촉 · 후각에 대한 부호화 표준을 수용할 수 있는 유연한 구조로 설계되었다.

〈감사의 글〉
본 연구는 정보통신부 및 정보통신연구 진흥원의 IT신성장 동력핵심기술개발사업의 일환으로 "07MH1810, 생체정보처리 기반 웨어러블 시스템 기술개발" 사업에 의해 지원되었습니다.

참고문헌

[1] Belkin, N. J., "Helping people find what they don' t know.", In Communications of the Association for Computing Machinery, 2000, Vol. 43, Issue: 8, pp. 58-61.

[2] Mithilesh, K. Akhilesh G. Sharad, S., "An Approach to Adaptive User Interfaces using Interactive Media Systems", Proceedings of the 11th international conference on Intelligent user interfaces, 2006, pp. 312-314.

[3] The MPEG Home Page, http://mpeg.org/

[4] UBC Lab for Advanced Networking, MPEG-4 Project, http://lan.ece.ubc.ca/apadana.html

[5] The GPAC Project, http://gpac.sourceforge.net/

저자소개

양 일 식

2002년 2월 전북대학교 컴퓨터과학과 학사
2004년 2월 전북대학교 컴퓨터정보공학과 석사
2004년-현재 한국전자통신연구원 디지털홈 연구단 연구원

주관심 분야 : 컴퓨터 네트워크, 이동 컴퓨팅, 임베디드 시스템, MPEG-4

김 명 규

1989년 2월 서울대학교 학사
1994년 12월 미국 매릴랜드 주립대학교 이학박사
1994년 10월-1997년 6월 서울대학교 자연과학종합연구소 연구원
1997년-현재 한국전자통신연구원 책임연구원

주관심 분야 : WPAN/WBAN, 오감 전송 프로토콜, 임베디드 시스템

손 지 연

1991년 2월 숙명여자대학교 학사
2001년 2월 한국정보통신대학원 석사
1991년 1월-현재 한국전자통신연구원 선임연구원

주관심 분야 : 무선기기자동인식및제어, 멀티미디어전송 프로토콜, WPAN(Wireless Personal Area Network)

저자소개

이 용 희

2002년 2월 충남대학교 컴퓨터과학과 학사
2004년 2월 충남대학교 컴퓨터과학과 석사
2004년-현재 한국전자통신연구원 디지털홈연구단 연구원

주관심 분야 : XML, MPEG-7, Intelligent User Interfaces, Digital Home System.

박 준 석

1984년 2월 인하대학교 전산학과 졸
1999년 8월 한국과학기술원 전산학과 석사
2006년 2월 인하대학교 컴퓨터정보공학과 박사
1987년 2월- 현재 한국전자통신연구원 차세대PC연구그룹, 스마트인터페이스연구팀장

주관심 분야 : 휴먼 컴퓨터 인터페이스, 멀티모달 인터페이스, 웨어러블컴퓨팅,Wireless Personal Area Network