

MPEG-4 비디오의 실시간 부호화기 구현

정 영 민, 김 중 호, 정 제 창
한양대학교 전자통신공학과

Implementation of a Real-time MPEG-4 Video Encoder System

Youngmin Jung, Jongho Kim, and Jechang Jeong
Department of Electronic Communication Engineering
E - Mail : dragon@ece.hanyang.ac.kr

요 약

멀티미디어 기술의 총아라고 불리는 차세대 디지털 방송의 핵심적인 기술인 MPEG-2 표준은 주로 방송용에 적합한 고화질 부호화가 목표였다. 1.5Mbps 이하의 저비트율, 즉 보다 고압축률의 부호화 방식에는 대응하고 있지 않았다. MPEG-4 표준 기술은 방송, 통신, 및 컴퓨터 분야에서 앞으로 차기 허용 대역폭이 증가하는 하겠지만, 현재의 네트워크 대역폭이 비디오 데이터를 전송하기에는 저대역폭인 상황에서 전송을 위해 새롭게 사용될 멀티미디어 부호화 기술로서 자연 영상뿐만 아니라 그래픽 합성 영상 등 다양한 멀티미디어 정보를 사용자가 대화형으로 쉽게 접근, 편집 처리할 수 있는 기능을 제공한다. 이에 본 논문에서는 향후 다양한 응용 분야에 적용될 MPEG-4 비디오 부호화기를 직접 설계 및 구현함으로써 응용의 가능성을 검증할 수 있는 시스템을 제안한다. 실험결과는 제안한 부호화기의 구조가 국제표준에 근거하여 다양한 응용 제품의 개발에 적합함을 보여준다.

I. 서 론

디지털 방송을 위한 MPEG-2 표준 규격에 이어 인

※ 본 논문은 ETRI 컴퓨터소프트웨어 기술연구팀의 위탁 과제 결과의 일부임.

터넷 기술 및 무선 단말 기술의 급속한 발전에 힘입어 멀티미디어 데이터에 대한 수요가 폭발적으로 증가하는 추세에 있다. 이러한 상황에서 다양한 멀티미디어 서비스를 위한 MPEG-4 표준 규격이 제정되어 이의 응용 제품의 개발이 관심의 초점이 되고 있다. MPEG-4 표준 기술은 이를 이용한 응용 제품과 관련하여 막대한 시장 창출은 물론이고, 방송, 통신, 컴퓨터와 관련한 타 산업으로의 파급효과도 매우 클 것으로 예상되기 때문에 당분간 지속적인 연구 개발이 이루어져야 할 분야이다. 이에 본 논문에서는 MPEG-4 시스템과 관련된 규격을 기반으로 하여 멀티미디어 정보를 네트워크를 통해 전송할 수 있는 시스템의 설계를 제안한다. 객체 기반의 부호화로 인한 대화형 기능, 보다 향상된 고압축 기능, 유무선 통신망을 통한 광범위한 접근 기능 등을 제공하는 MPEG-4의 특징에 대한 구현과 이를 검증하는 데에 목표를 두었다. MPEG-4 규격에서 사용되고 있는 도구는 MPEG-1, MPEG-2 및 H.263에서 사용되고 있는 도구를 기본으로 하고 있다. 이들 규격과의 차이점은 객체를 하나의 장면을 구성하는 단위로 도입했다는 점과 각 객체의 영상 데이터가 영상 정보를 포함하고 있다는 것이다. 이에 반해 종래의 휘도 및 색차 정보는 텍스처 정보라 불린다. 영상 정보의 부호화는 종래에 없는 기능이며, 이에 대한 각종 고능률 부호화 도구가 제안/채택되었다. 또, MPEG-4에서는 종래의 자연 영상 및 오디오 신호에 추가하여, 컴퓨터 그래픽스

(CG) 데이터 및 합성 음성 등의 인공 신호도 취급할 수 있다. 또한 시스템 규격에서는[1] 객체 단위의 영상과 음성에 접근할 수 있도록 여러 계층의 데이터 구조를 정하고 장면 기술을 위한 기술자를 규정하고 있다. 전송 규격에서는 기존의 다양한 전송 포맷을 사용할 수 있도록 시스템 스트림과 전송 프로토콜간을 연결해주는 인터페이스를 규정하고 있는데 이를 DMIF(Delivery Multimedia Integration Framework)라 한다.[5]

본 논문에서는 위에서 언급한 MPEG-4의 여러 특징을 반영하여 MPEG-4 객체 부호화 방식을 기본으로 한 여러 객체의 압축 비트 스트림을 부호화하는 인코더 시스템을 구현하고, 각 비트 스트림을 다중화하여 서버 저장용 데이터를 제작하였다. 이를 네트워크를 통해 전송한 후 복호하여 결과를 검증하고, 성능을 평가하였다. 본 논문의 구성은 다음과 같다. 2장에서는 MPEG-4 인코더 시스템의 설계 및 제작을 위한 규격들에 대해서 서술하고, 3장에서는 실제 구현 방법을 기술한다. 4장에서 실험결과를 보이고 이에 따른 결론과 향후 과제에 대해서 언급한다.

II. MPEG-4 기반의 부호화기 시스템을 위한 표준 규격

2.1 MPEG-4 비주얼 부호화

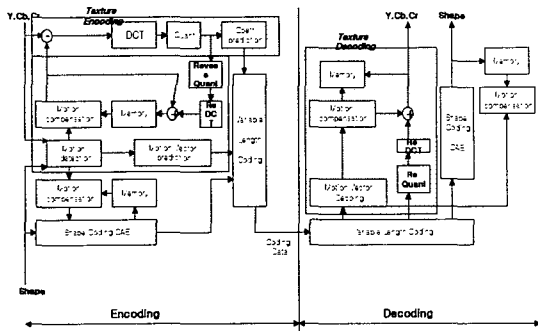


그림 1 비디오 객체 부호화·복호화

MPEG-4 비주얼 규격[2]은 영상 객체의 복호화에 필요한 데이터 선택스와 그 의미를 규정하고 있다. MPEG-4 표준 규격에서 대상으로 하는 영상 객체로는 자연 영상 뿐 아니라 컴퓨터 그래픽스와 같은 합성 영상을 포함하고 있다. MPEG-4에서는 한 장면을 구성하기 위해서 복수개의 객체 스트림을 다루기 때문에 계층

구조의 선택스를 가지고 있다. 각 비디오 객체는 특정 시간에 들어온 복수의 비디오 객체 평면(Video Object Plane : VOP)으로 구성된다. 이와 같이 복수의 객체로 구성된 한 평면(VOP)이 MPEG-4에서 취급하는 영상 데이터의 기본 단위이다. 구형 형상의 경우, VOP는 MPEG-1과 MPEG-2의 프레임 또는 필드에 해당한다. 그림 1은 비디오 객체 부호화/복호화의 구조도이다. 비디오 객체는 형상과 텍스처로 나누어 부호화된다. 부호화 알고리즘은 움직임 보상 및 예측 부호화를 이용한다. 본 논문에서 제안하는 MPEG-4 인코더 시스템은 단일 객체의 구형 영상을 대상으로 설계했다.

2.2 비주얼 규격의 프로파일과 레벨 정의

MPEG-4 비주얼 규격에서 정의하고 있는 프로파일은 여러 틀셋으로 정의된다. 일련의 의미 있는 처리 단위로 모은 도구의 조합을 객체라 부르고, 이 객체를 단독 또는 복수로 조합한 것을 프로파일이라고 부른다(그림 2).

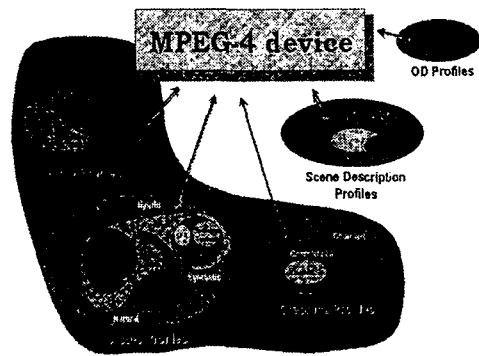


그림 2 MPEG-4 프로파일과 레벨 정의

각 프로파일에 요구되는 도구의 성능 기준, 구속 조건은 레벨에서 규정된다. 비주얼 프로파일에서는 자연 영상 및 합성 영상에 관한 총 9개의 프로파일이 정의되어 있다. 본문에서 제안한 시스템에서는 MPEG-4 심플 프로파일 레벨 1 (SP@L1)으로 하였다. 이 영상 압축 과정에서는 확장성을 고려하여, 향후 전송 속도 및 기술의 발전에 따라 profile 및 level은 확대 될 수 있으며, 이 경우에도 최소한 SP@L1은 지원되도록 하였다. 프로파일은 사용되는 요소 기술을 정의하고, 거기에서 사

용하는 영상의 크기 및 비트율 등은 레벨로서 정의된다. 이 논문에서 사용된 SP@L1의 규격은 다음과 같다.

- 인트라 부호화 모드(I-VOP)
- 인터 예측 모드(P-VOP)
- AC/DC 예측, 가역 VLC, 데이터 분할
- 슬라이스 재동기(resynchronization)
- 4 움직임 벡터
- 무제한 움직임 벡터

2.3 MPEG-4 시스템 규격

MPEG-4 표준은 지금까지의 방식과 달리 객체 단위의 압축 방식을 채택해 응용 범위를 확대하였다. 따라서 시스템 부분에서는 오디오/비디오 비트 스트림간의 다중화/동기화뿐만 아니라 각 객체 단위의 합성을 위한 장면 기술의 필요성이 제기된다.

MPEG-4 시스템 규격은 장면 기술을 위한 도구와 동기를 위한 패킷화 계층으로 크게 이루어져 있다(그림 3). 또한 MPEG-4 규격에서는 시스템과 전송 프로토콜 사이에 DMIF라고 하는 계층을 정의하여 단말이 특정 네트워크에 독립적으로 작동하는 것을 보장한다[5].

MPEG-4 시스템 규격의 독창성을 보여 주는 객체 및 장면 기술자는 MPEG-4의 객체 단위 부호화의 인해 생겨난 도구로 다음과 같은 작용을 한다.

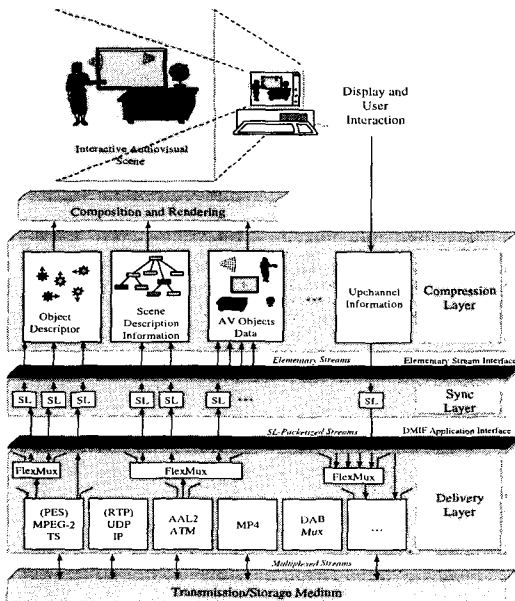


그림 3 MPEG-4 시스템 단말 구조

• 객체 기술자 :

각 미디어 객체의 비트 스트림은 객체 기술자에 기술되어 있는 미디어의 종류에 의하여, 지정된 복호화 알고리즘으로 복호된다. 이 객체 기술자는 비디오, 오디오 데이터의 기술 뿐만 아니라 MPEG-4 세션에서 최초로 전송되는 데이터이기도 하다. 초기 객체 기술자를 수신 단말에서 얻게 된 후 장면 기술자를 위한 객체 기술자를 얻고 그 후에 그 장면을 구성하는 실제 비디오, 오디오 데이터의 객체 기술자를 통해 사용자의 화면을 구성하게 된다.

• 장면 기술자 :

여러 미디어 객체의 시간적, 공간적 관계는 BIFS(Binary Format for Scene)에 의하여 행해진다. 미디어 객체는 BIFS 의하여 단말 화면상에 합성되어 배치되며 장면기술은 동적으로 갱신된다.

III. MPEG-4 비디오 부호화기의 구현

3.1 시스템의 구조

그림 4와 5에서 구현된 MPEG-4 부호화기 시스템의 전체 구성을 나타내었다. 또, 전체적인 시스템을 구성하는데 필요한 소프트웨어는 MPEG-4 표준화 기구에서 배포한 참조 S/W(reference S/W)를 활용하였다.

3.2 각 모듈의 기능과 구성

입력부를 통해서 분리된 각각의 오디오, 비디오 및 스크립트 정보를 부호화부의 입력으로 하며 영상 부호화부는 MPEG-4 SP@L1으로 하였다. 압축된 콘텐츠는 전송 대역폭을 고려하여 비디오와 오디오를 합하여 64Kbps 이내, 또는 384Kbps 이내의 비트율을 모두 포함한다. 음성 부호화부는 G.723.1을 사용하여 5.3Kbps의 비트율로 할 수 있도록 설계하였다. MPEG-4 시스템 규격에서는 객체의 표시 방법과 특성을 지정하기 위한 장면기술 언어로서 BIFS(Binary Format for Scene)를 규정하고 있다. BIFS 인코더 모듈에서는 BIFS로 기술된 장면과 객체에 관한 기술 정보를 가변장 부호화하여 스크립트를 바이너리로 부호화한다[7].

합성부는 각 부호화부를 거쳐 나온 출력 데이터를 합성하는 역할을 한다. 이들 데이터의 합성에 따라 MPEG-4 시스템 규격에서 규정하고 있는 MP4파일[6]

과 그 중간버전의 파일 포맷인 TRIF파일로 저장될 수 있다. 그림 6에서 구현된 시스템의 부호화부에 따른 모듈 구성에 대한 간략도를 보이고 그림 7에서 실제 구현 결과장면을 보였다.

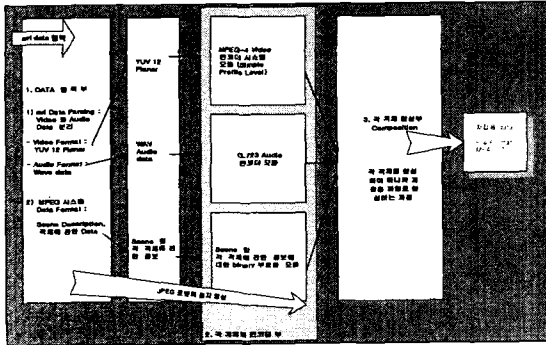


그림 4 인코딩 시스템의 구조

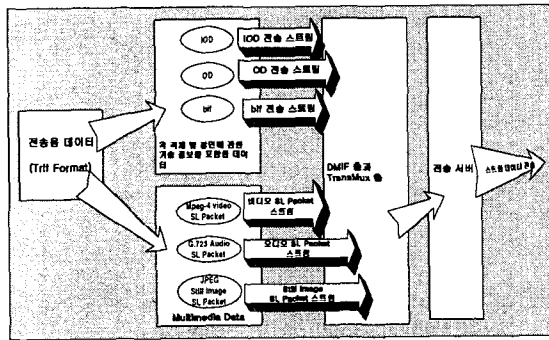


그림 5 전송용 스트림 출력 구조

IV. 실험 결과 및 결론

본 논문에서는 MPEG-4 시스템에서 영상 정보와 오디오 정보에 국한시켜 인터넷이나 무선망에서의 전송 대역폭에 맞추어 인코딩 시스템을 설계, 구현하였다. 이후에 MPEG-4 시스템과 관련된 규격을 기반으로 하여 멀티미디어 정보를 네트워크를 통해 전송할 수 있는 서버와 클라이언트로 구성되는 시스템의 설계와 그 응용 분야에 대해 연구 개발해야 할 문제는 여전히 남아 있다. MPEG-4 표준은 대역폭이 한정되어 있는 인터넷 망에서 사용자에게 기존의 코덱보다 더 좋은 화질을 제공하는 동영상을 제공할 수 있다. 또한 BIFS 장면 기술자를 사용하여 한 장면에서 다양한 객체들을 사용자에게 보여주거나, 사용자와의 대화형 기능을 부가할 수

있다는 점은 MPEG-4가 가지는 큰 매력이 될 수 있다.

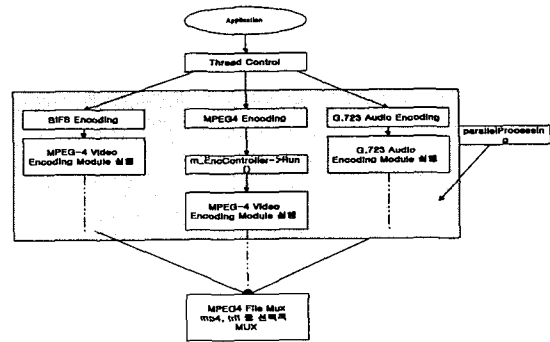


그림 6 멀티미디어 부호화부의 구조

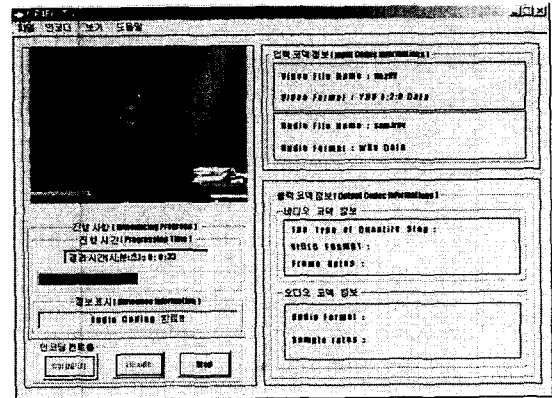


그림 7 MPEG-4 비디오 부호화기

참 고 문 헌

- [1] ISO/IEC 14496-1, "Information technology - Coding of audio-visual objects: Systems," International Standard, 1999
- [2] ISO/IEC 14496-2, "Information technology - Coding of audio-visual objects: Visual," International Standard, 1999
- [3] 정제창 역, "그림으로 보는 최신 MPEG," 교보문고, 1995
- [4] 정제창 역, "그림으로 보는 응용 MPEG," 교보문고, 1997
- [5] ISO/IEC 14496-6, "Information technology - Coding of audio-visual objects: DMIF," International Standard, 1999
- [6] ISO/IEC 14496-1/PDAM1 (MPEG-4 version 2 Intermedia Format -MP4)
- [7] ISO/IEC 14772-1:1997 "VRML97," International Standard, 1997