

MPEG-4 시스템 표준을 이용한 멀티미디어 시스템의 구현

임동근*, 호요성*, 조창식**, 마평수**

* 광주과학기술원 정보통신공학과
** 한국전자통신연구원 정보가전응용연구팀

Multimedia System Implementation using the MPEG-4 System Standard

Dong-Keun Lim*, Yo-Sung Ho*, Changsik Cho**, Pyeongsoo Mah**

* Kwangju Institute of Science and Technology (K-JIST)
** Electronics and Telecommunications Research Institute(ETRI)

요약

기존의 영상 압축 부호화 방식에서는 입력 영상에 담긴 내용과는 무관하게 영상의 화소 값만을 기본으로 데이터 압축을 수행하였다. 그러나, 이러한 압축 방법은 그 성능에 한계가 있기 때문에 다양한 멀티미디어 서비스의 응용 분야에서 요구되는 여러 가지 기능을 모두 충족하기 곤란하다. 따라서 영상의 내용에 대한 이해를 바탕으로 한 내용기반 부호화(Content-based Coding)에 중점을 둔 MPEG-4 표준이 연구되었다. 이 내용기반 부호화에서는 영상 내용을 각각의 객체(Object) 단위로 나누어 처리하여 전송하므로, 사용자의 의도에 따라 다양한 형태의 조작과 재현이 가능하다. 본 논문에서는 MPEG-4 표준의 시스템 부분에 따르는 멀티미디어 서비스 시스템의 구현에 대해 기술하였다. 최근 이 분야의 연구개발 동향은 주로 비디오와 오디오 신호의 개별적인 부호화에 치중하던 이전의 기술적 상태에서 벗어나 비디오와 오디오 신호를 묶어주는 시스템 기술의 중요성이 더욱 증가하고 있다. 따라서 본 연구에서 구현된 MPEG-4 시스템 기술은 다양한 멀티미디어 서비스 응용에 사용될 수 있는 기반 기술로서 중요성을 가진다.

1. 서론

최근 멀티미디어 서비스에 대한 요구가 크게 늘어남에 따라 멀티미디어 통신의 핵심 기술인 초저속 영상과 음성 신호 부호화 방식에 많은 관심이 모아지고 있다. 기존의 음성 신호 뿐만 아니라, 영상 신호까지 포함한 방대한 정보량을 가지는 멀티미디어 신호를 유선 통신망이나 무선 채널을 통해 효율적으로 전송하기 위해서는 디지털 신호의 압축 처리가 절대적으로 필요하다.

기존의 영상 부호화 표준들은 대개 블록 단위로 처리하기 때문에, 초저속 영상통신을 위해 데이터 압축 비율을 크게 높일 경우에, 블록화 현상 및 경계흐림 현상 등의 문제점이 발생한다. 따라서 MPEG-4 표준이나 H.26L 장기표준에서는 기존의 영상부호화 방식과 다르게 객체에 기반한 (Object-based) 새로운 부호화 기법을 도입하였다. 아울러, 기존의 부호화 방식에서 부호화 압축율과 재생 화질을 동시에 높이기 위한 여러 가지 연구들도 한창 진행되고 있다. 이와 같은 초저속 영상부호화 방식에

대한 연구 개발은 현재 빠르게 전개되고 있는 멀티미디어 통신 서비스의 기술 기반이 될 것이다.

ISO/IEC JTC1/SC29/WG11(MPEG) 그룹에서는 1993년부터 MPEG-4 표준화 작업을 준비하였으며, MPEG-4 표준에서 지원할 8개의 주요 기능(Functionality)을 정했다. 이에 따라 1995년 말부터 본격적인 MPEG-4 표준화 작업을 진행하였으며, 이미 MPEG-4 표준 1차 버전이 완성되어 구현되어 있고, 2차 버전도 거의 완성 단계에 이르렀다.

현재 국내외의 여러 연구 기관들과 회사들은 MPEG-1, MPEG-2 또는 MPEG-4 기술[1-6]을 이용한 멀티미디어 서비스 개발에 많은 노력을 기울이고 있지만, 이러한 기술들은 아직 그렇게 활발히 이용되지 못하고 있는 상태이다. 국내에서는 기술적으로는 약간 부족한 상태이지만, 외국 기술을 이용한 다양한 형태의 멀티미디어 서비스들이 이미 소개되고 있다. 특히, 현재 그 사용이 날로 급증하고 있는 인터넷이나 이동통신과 같은 제한된 용량의 전송 채널을 통한 멀티미디어 통신 서비스들이 왕성하게 도입되고 있다.

본 논문에서는, 이러한 기술들의 국산화를 위한 노력하의 일환으로, 멀티미디어 서비스 시스템을 구축하기 위한 MPEG-4 시스템 표준의 구현 방법을 설명한다. 본 작업에서는 MPEG-4 영상[1]과 음성 신호[2]의 압축/복원 기술을 연구 개발하였고, MPEG-4의 확장성과 호환성의 특징을 최대한 살리기 위해 MPEG-4 시스템 코덱[3]을 소프트웨어로 구현하였다.

현재 표준화 작업이 마무리되어 가고 있는 MPEG-4 부호화 기술은 인터넷이나 이동통신, 또는 초고속정보통신망을 통한 다양한 멀티미디어 통신 서비스에 사용될 핵심적인 기술이다. 따라서 구현된 MPEG-4 시스템을 이용하면, 초고속정보통신망을 사용하는 멀티미디어 서비스 사용자들에게 좋은 화질의 영상 서비스를 실시간으로 제공할 수 있을 것이다.

2. MPEG-4 시스템 표준

2.1. 내용

MPEG-4 표준에서는 대화형(interactive) 오디오-비주얼 화면으로 구성된 멀티미디어 정보를 교환하기 위한 목표 시스템을 설정하고 있는데, 다음과 같은 사항을 포함하고 있다.

- 실제 영상 혹은 인공영상, 오디오 혹은 비주얼 2차원 혹은 3차원 객체의 부호화된 표현
- 오디오-비주얼 객체의 상호관계와 시-공간적인 위치의 부호화된 표현
- 동기일치 (synchronization), 인식 (identification), 설명 (description)과 비트열의 조합 (association)과 같은 데이터 비트열의 관리와 관련된 정보의 부호화된 표현
- 데이터 비트열 전송계층(delivery layer) 함수와의 일반적인 인터페이스

오디오-비주얼 화면들로 구성된 멀티미디어 정보의 통신을 위한 시스템의 전체 동작은 다음과 같다.

전송단 측에서 오디오-비주얼 화면 정보는 각각의 부호화기를 이용하여 압축되고, 동기일치를 위한 시간 정보와 결합된다. 이 결합된 하나 이상의 부호화된 이진 비트열은 전송되거나 저장된다. 수신단 측에서 받은 이 비트열은 역다중화되고 압축이 풀어진다. 오디오-비주얼 객체들은 화면구성 정보(scene description)와 동기일치 정보에 따라 재구성되어 사용자에게 제시된다. 사용자는 이 표현된 것에 따라 대화형 처리를 위한 상호작용을 할 수 있다. 상호작용 정보는 수신단 측에서만 처리될 수도 있고, 전송단 측으로 전송될 수도 있다. MPEG-4 시스템 부분은 자세한 복호화 처리 방법 뿐만 아니라, 그러한 화면 정보를 운반하는 비트열들의 구문구조(syntax)와 의미(semantics)를 정의하고 있다.

2.2. 구조

MPEG-4 시스템 부분에 규정된 표현 정보는 부호화된 오디오-비주얼 정보와 해당되는 화면구성 정보(scene description)에 대한 상호 연관된 오디오-비주얼 화면을 생성하는 방법을 설명한다.

초기 설정 정보는 부호화된 콘텐츠의 표현 형태인 하나 이상의 기초비트열에 적용되도록 허용되어야 한다. 이러한 기초비트열들의 일부는 MPEG-4 시스템 표준[3]에 설명된 다중화 도구(tool)을 사용하여 묶인다. 기초비트열은 오디오 데이터 혹은 비주얼 데이터, 그리고 화면구성 정보 중의 하나에 대한 부호화된 표현물을 포함한다. 기초비트열은 비트열을 인식하고, 비트열 사이의 상호 의존성을 설명하며, 비트열의 콘텐츠와 관련된 정보를 설명하는 정보 자체일 수도 있다. 각 기초비트열은 하나의 데이터 형태만을 포함한다. 기초비트열들은 그들의 비트열과 관련된 복호화기를 사용하여 복호화된다. 오디오-비주얼 객체들은 화면구성 정보에 따라 편집되며, 터미널의 표현 장치에 출력된다. 이러한 동작들은 동기일치 계층(synchronization layer, sync layer)에서 제공되는 동기일치 정보를 이용하여 시스템 복호화기 모델(systems decoder model, SDM)에 따라 동기가 일치된다.

가. 비트열의 다중화: 전송 계층(delivery layer)

전송 계층은 MPEG-4 표준에 따르는 콘텐츠를 전송하거나 저장하도록 기존의 전송 프로토콜(transport protocol)의 일반적인 개념으로 사용된다. 이 계층의 기능은 MPEG-4 시스템 표준의 범위 안에 있지 않으며, 단지 이 계층의 인터페이스만이 고려된다. 이 인터페이스는 MPEG-4 표준 DMIF 부분[6]에 규정된 DAI(DMIF Application

Interface)이다.

다시 말하면, 시스템 부분은 DMIF로부터 FlexMux 패키지를 입력 받은 후, SL 패키지들을 패키지 헤더와 패키지 페이로드로 분리하고, 패키지 헤더에 있는 타이밍과 패키지화 정보를 이용하여 기초비트열을 분리해 낸다.

나. 비트열의 동기일치: 동기일치 계층(Sync Layer)

기초비트열들(elementary stream, ES)은 비트열 데이터에 대한 기본적인 대상이다. 기초비트열들은 DAI(DMIF Application Interface)에서 동기일치 계층으로 패키지화된 비트열(SL-Packetized) 형태로 전송된다. 이 패키지화된 표현물은 임의 접근(random access) 정보 뿐만 아니라, 타이밍과 동기일치 정보를 추가적으로 제공한다. 동기일치 계층(SL)은 동기일치된 복호화와 계속적으로 기초비트열 데이터의 컴포지션이 가능하도록 타이밍 정보를 추출한다.

다. 압축 계층(Compression Layer)

압축 계층은 데이터와 그것의 압축된 형식인 기초비트열(ES)을 수신하여 이 데이터의 복호화를 위해 필요로 하는 동작을 수행한다. 복호화된 정보는 터미널의 편집(composition), 렌더링(rendering), 표현(presentation)에 사용된다.

2.3. 기술어 (Descriptor)

기술어들은 MPEG-4 시스템 부분에 규정된 각 기초비트열들간의 인식, 조합, 동기일치 등의 동작을 하기 위하여 사용되는 메타데이터(metadata)이다. <그림 1>은 전체적인 구성을 보이며, <그림 2>는 그 동작을 설명하고 있다.

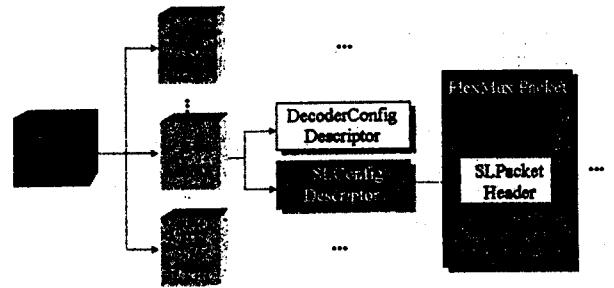


그림 1. 기술어(Descriptor)의 구성

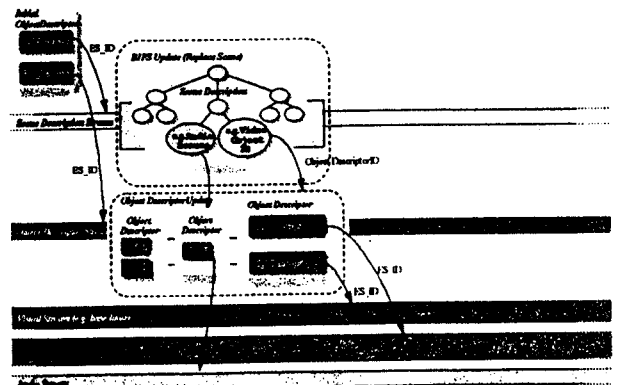


그림 2. 기술어(Descriptor)들의 동작

객체기술어 OD(Object Descriptor)는 개별적인 기초비트열과 그것의 성질을 설명하기 위해 그들 항목에 대한 다양한 추가적인 설명어(descriptor)들을 포함한다.

기초비트열 기술어 ESD(ES_Descriptor)는 특정한 기초비트열과 관련된 정보를 운반하며, 세 부분으로 이루어져 있다. 첫번째 부분은 ES_ID와 같이 ESD를 유일하게 식별하기 위한 표시이며, 부모 OD와의 의존성을 설명하는 매커니즘이다. 두번째 부분은 기초비트열이 요구하는 특성을 기술하는 해당 항목 기술어로 구성되어 있다. 세번째 부분은 미래의 확장이나 역방향 호환성(backward compatible)을 제공하기 위한 선택적인 기술어이다.

복호기설정 기술어 DecoderConfigDescriptor는 복호기 형과 해당 기초비트열이 요구하는 복호기 자원에 관한 정보를 제공한다. 이것은 수신단에서 해당 기초비트열을 복호할 수 있는지 여부를 결정하는데 필요하다.

동기일치설정 기술어 SLConfigDescriptor는 SL Packet Header에서 참조하는 값을 전송하는데, 해당 변수(parameter)가 부호화되는지 여부를 나타내는 플래그(flag)와 변수(parameter)를 몇 비트 길이로 부호화할 것인지의 정보를 가지며, <그림 3>과 같은 구조로 되어 있다.

SL 패킷은 헤더와 페이로드(payload)부분으로 구성되는데, 헤더에서는 페이로드로 전송되는 데이터의 시간적인 관계와 연결에 관련한 정보를 포함한다. SL 패킷헤더는 각 개별적인 기초비트열의 요구사항에 따라서 설정된다. 변수들은 타임스탬프들과 클럭 참조의 존재 유무, 해상도, 시간의 정확도에 관한 사항을 포함한다. <그림 4>는 SL 패킷의 구조를 설명하고 있다.

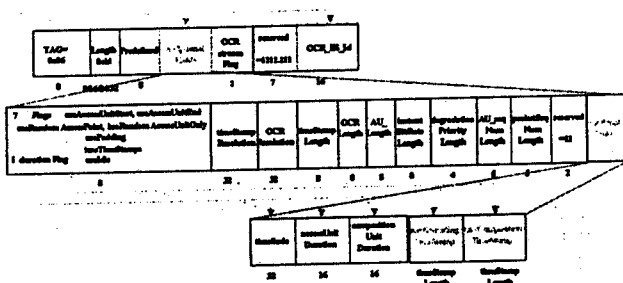


그림 3. SLConfigDescriptor의 구조

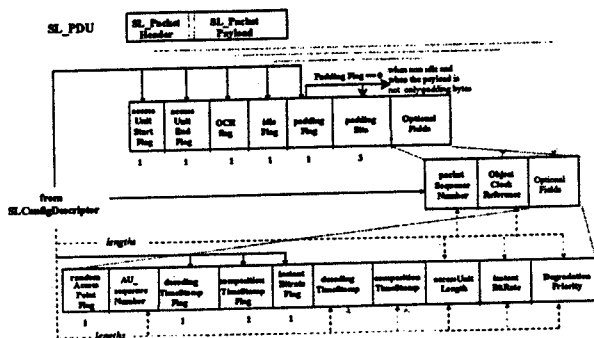


그림 4. SL Packet의 구조

3. MPEG-4 시스템의 구현 방법

3.1. 전체 시스템의 구조

본 작업에서 구현된 멀티미디어 서비스 시스템은 참여한 기관들이 개별적으로 부분을 구현한 각각을 결합시켜 이루어진다. 우리는 이 중에서 <그림 5>와 같이 MPEG-4 비주얼, 오디오, 시스템 부분을 구현하였는데, 본 논문에서는 시스템 부분의 구현 방법에 대해서 설명한다.

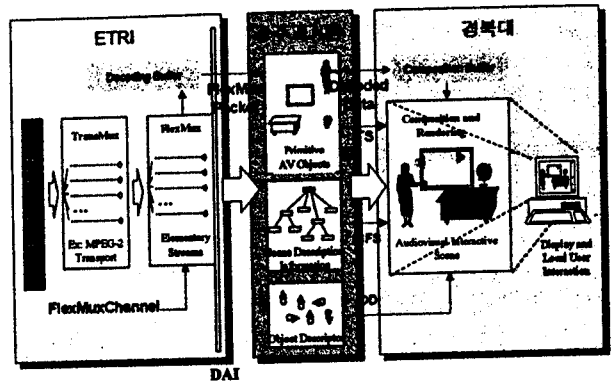


그림 5. 전체 멀티미디어 시스템

3.2. MPEG-4 시스템 구현 구조

<그림 6>은 복호기로 전송할 TransMux 비트열을 만드는데 사용되는 FlexMux 비트열을 만드는 MPEG-4 부호화기의 동작을 개략적으로 보여주고 있다.

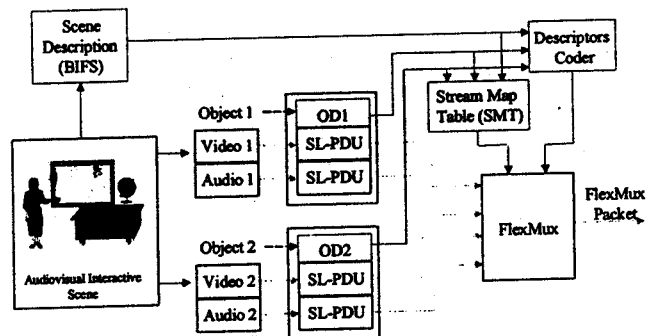


그림 6. MPEG-4 부호화기의 개략적 모델

<그림 6>에서 시스템 부분은 각 object를 구성하는 기초비트열들을 SL 패킷(SL-PDU)으로 만든다. 이 SL 패킷에는 AU(access unit)단위로 구성된 기초비트열 데이터, 타이밍 정보인 DTS(decoding time stamp)와 CTS(composition time stamp), 그리고 각 object를 복호화하는 시점을 나타내는 OCR(object clock reference) 정보가 포함되어 있다. 이 값들은 복호화기에서 필요로 하는 개별 복호화기 블록에 공급된다. OCR정보는 어느 객체(object)가 복호화를 시작할 것인지를 나타내는 참조 시점으로, 해당 패킷에서 추출된 DTS나 CTS는 이 값과 비교하여 그 사용 시점을 결정한다. 기초비트열들은 비디오, 오디오, 텍스트, 그리고 화면구성 정보(scene description, SD)가 될 수 있다.

MPEG-4 부호화기는 <그림 7>과 같이 구현되었다. 부호화하려는 여러 종류의 자료 즉, 비디오, 오디오, 텍스트 등의 기초비트열을 입력받아 조합할 때, 기술어(descriptor) 정보를 파일의 헤더 부분에 위치시키고, 서로간의 동기를 맞추기 위한 타이밍 정보는 SL-PDU에

부가정보로서 추가시킨다.

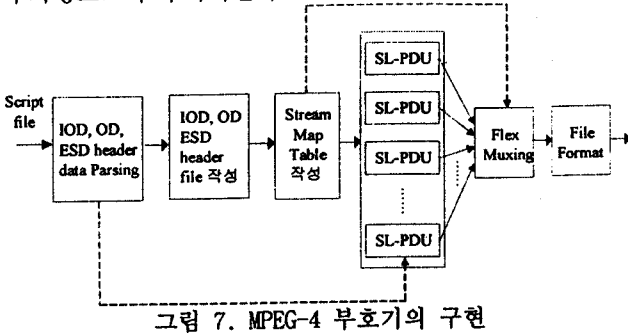


그림 7. MPEG-4 부호기의 구현

이때 해당되는 부가정보를 조합하기 위한 기본 입력 화일이 필요하며, 이것은 스크립트(script) 화일 형태를 가진다. 그 내용은 기술어에 삽입되는 내용에 대한 것과 SL 패킷을 만들 때 필요한 타이밍, 타임스탬프(DTS, CTS), 클록 참조(OCR) 등에 관한 사항을 가지고 있다.

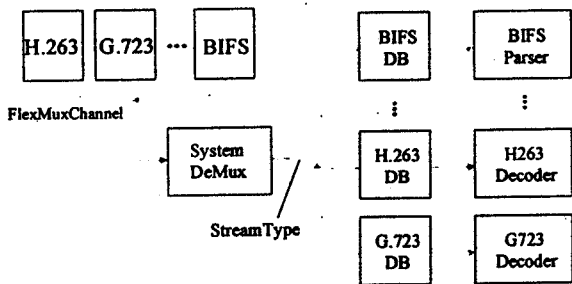


그림 8. MPEG-4 복호기의 구현

MPEG-4 복호화기는 <그림 8>과 같이 구현되는데, 내부에 비디오, 오디오, 텍스트 등을 복호하는 개별 복호화기를 내장하고 있다. 네트워크를 통해 MPEG-4 복호화기, 즉 멀티미디어 시스템으로 전송되어온 비트열 속의 다양한 형태의 비디오, 오디오, 텍스트 등의 멀티미디어 자료를 기초비트열(Elementary stream, ES)들로 구분하여 각각에 적합한 복호화기로 분배하여 원래의 비디오, 오디오, 텍스트 등으로 복원하는 역할을 한다. 그런데, H.263 혹은 G.723 등의 복호화기에서 필요로 하는 타이밍 정보와 AU(Access Unit)를 임시로 저장하기 위해서는 system DeMux 블록과 Decoder 블록 사이에 디코딩 버퍼(decoding buffer, DB)가 필요하다. 이 버퍼에 들어 있는 내용을 입력 받아 복호화기는 복호화를 시작하고, 그 결과를 결합 버퍼(Composition Buffer, CB)에 저장하고, 화면구성 정보(Scene Description, SD)를 사용하여 화면의 해당 위치에 출력한다.

4. 구현된 멀티미디어 시스템

본 작업에서 최종적으로 구현된 시스템의 화면 구조는 <그림 9>와 같다. 총 9 개의 객체가 사용되는데, 비디오와 오디오, 그리고 시간에 따라 동적으로 변화하는 자막 부분으로 구성되어 있다. SD나 OD와 같은 일부 객체는 화면이나 음성으로 출력되지 않고, 각 미디어 데이터들의 연결이나 동기 일치에 사용되었다.

이 멀티미디어 시스템은 또한 개발중인 더욱 더 큰 멀티미디어 시스템의 일부분(MPEG-4 Player)으로서 사용된다. 여러 가지 다양한 형태의 규격을 가지는 멀티미디어

정보를 다루기 위하여 시스템들 사이의 결합이 자연스럽게 이루어지고 있다.

MPEG-4 PLAYER

종료



그림 9. MPEG-4 Player 구현화면

5. 결론

본 논문에서 구현된 MPEG-4 표준 방식의 복호화기는 빠른 속도로 확산되고 있는 저가적이면서 고성능인 멀티미디어 통신 서비스를 고려하여 유동적으로 기존의 방식 및 새로운 기능들을 모두 지원할 수 있는 부호화 도구를 제공한다. 또한 MPEG-4 표준 자체의 특성인 양방향성, 높은 압축률 및 다채로운 접속을 가능케 하는 AV(Audio-Visual) 표준화 방식 뿐만 아니라, 급격히 발전하는 관련 기술들을 이용할 수 있도록 고도의 융통성과 확장성을 제공할 것이다. 또한 MPEG-4 표준 방식을 이용하여, 기존의 단방향 통신의 일방적인 정보 전달의 차원을 벗어나, 수신자의 요구와 선택에 따라 필요한 정보를 원하는 시간에 제공할 수 있는 양방향 서비스가 실현될 수 있을 것이다.

감사의 글

본 연구는 한국전자통신연구원(ETRI)의 조립식 실시간 OS 기술개발 과제, 광주과학기술원(K-JIST) 초고속광네트워크(UFON) 연구센터를 통한 한국과학재단 우수연구센터(ERC)와 교육부 두뇌한국 21(BK21) 정보기술사업단의 지원에 의한 것입니다.

참고문헌

- [1] ITU-T Recommendation H.263, Line Transmission of Non-telephone Signals, 1995.
- [2] ITU-T Recommendation G.723, Dual Rate Speech Coder For Multimedia Communications Transmitting At 5.3 & 6.3 kbits/s, 1995.
- [3] ISO/IEC 14496-1: Information Technology - Coding of audio-visual objects - Part 1: Systems, 1999.
- [4] ISO/IEC 14496-2: Information Technology - Coding of audio-visual objects - Part 2: Visual, 1999.
- [5] ISO/IEC 14496-3: Information Technology - Coding of audio-visual objects - Part 3: Audio, 1999.
- [6] ISO/IEC 14496-6: Information Technology - Coding of audio-visual objects - Part 6: Delivery Multimedia Integration Framework (DMIF), 1999.