

# MPEG-4의 3차원 재생

이현주 김남영<sup>0</sup> 이동훈 김상욱

경북대학교 컴퓨터과학과

{hyunju, nykim<sup>0</sup>, dhlee, swkim}@woorisol.knu.ac.kr

## 3D Playback of MPEG-4

Hyun-Ju Lee Nam-Young Kim<sup>0</sup> Dong-Hoon Lee Sang-Wook Kim

Dept. of Computer Science, Kyungpook National University

### 요 약

MPEG-4 시스템에서는 씬 디스크립션 언어로 BIFS와 XMT를 정의하고 있다. 씬 디스크립션은 MPEG-4 씬을 구성하는 객체들의 시공간적 구성관계를 기술하는 것이다. 본 논문에서는 MPEG-4에 정의된 다양한 데이터 포맷과 3차원 MPEG-4 객체의 효율적인 재생 지원을 위해 XMT에 노드를 기술하고 재생하는 시스템을 제안한다. 제안하는 시스템은 XMT 파일을 XMT 파서를 이용하여 파싱하고, 그 결과로 생성되는 DOM 트리에서 DOM APIs를 이용하여 재생에 필요한 노드 정보를 얻는다. 사용자 인터페이스상에서는 DOM 트리에서 얻은 노드를 기준으로 그래픽 라이브러리를 호출하여 화면상에 재생하는 구조를 가진다. 또한 재생이 진행되는 도중에 발생하는 사용자 이벤트도 감지 및 처리하며, 필요한 정보에 따라서 화면과 중간 자료구조를 재구성한다.

### 1. 서 론

MPEG-4는 오디오, 비디오 데이터를 포함한 객체 단위로 구성된 다양한 형태의 멀티미디어 데이터를 정의하고 있다[1]. MPEG-4 씬은 씬 내의 미디어 객체들간의 표시방법과 특성을 포함한 시공간적인 관계를 기술하고 이들간의 상호작용을 정의하는 씬 디스크립션에 의해서 구성된다. 따라서 MPEG-4 씬을 생성하기 위해서는 씬 구성을 기술하는 씬 디스크립션이 필요하다.

MPEG-4 파일 형식중 씬 디스크립션 정보를 저장하는 BIFS(Binary Format for Scenes)형식은 씬 디스크립션을 이진 부호화한 것이다. BIFS는 VRML을 기반으로 하고 있는데, 이를 확장하여 정의하고 있는 MPEG-4 씬 디스크립션이 XMT(eXtensible MPEG-4 Texture Format)이다[2, 3]. XMT로 저장된 MPEG-4 콘텐츠는 XML에 기반을 두고 있기 때문에 멀티미디어 콘텐츠를 인터넷상에서도 활용할 뿐만 아니라 기존 언어의 확장성 및 이식성을 용이하게 하는 등의 특징을 가질 수 있다.

따라서 본 논문에서는 각 객체들을 기술한 XMT 파일을 XMT 파서가 파싱하여 정보를 구조화한 후 DOM 트리를 생성한다[4]. 생성된 DOM 트리에서 DOM APIs를 이용하여 노드 정보를 얻어 재생에 필요한 그래픽 라이브러리를 호출하여 화면에 재생하는 시스템을 소개한다. XMT로 씬을 정의함으로써 기존 BIFS에 의한 씬 디스크립션을 확장하여 노드를 기술할 수 있다. 또한 사용자

제어에 의한 추가적인 씬을 재구성함으로써 기존의 재생 기로는 노드 기술 및 구현이 제한적이었던 MPEG-4에서 정의된 다양한 멀티미디어 객체로 구성된 씬들을 노드로 기술하여 XMT 파일을 생성하고 재생하는 시스템을 구현하였다.

본 논문의 제 2절에서는 MPEG-4 재생기의 전체적인 구조에 대해서 설명하고, 제 3절에서는 XMT를 이용한 3차원 재생 방법을 설명하고, 제 4절에서는 구현결과 및 개발환경을 보이고, 제 5절에서는 결론 및 향후 연구방향을 제시한다.

### 2. MPEG-4 재생기 구조

로컬의 MPEG-4 파일은 MPEG-4 재생기 시스템에서 디멀티플렉서에 의해 디멀싱(demuxing)이 수행된다. 디멀싱된 결과로 생성되는 씬 디스크립션 정보와 각 미디어 데이터들은 디코딩과 컴포지션 과정을 거친다. MPEG-4 재생기는 씬 디스크립션 정보에 따라 각 미디어 렌더링 모듈에 의해 사용자 인터페이스 상에서 재생이 되는 구조이다. 재생중에 발생하는 이벤트처리는 사용자 이벤트 매니저에 의해 관리 및 처리된다.

본 논문에서 제안하는 재생기의 전체적인 구조는 그림 1과 같다.

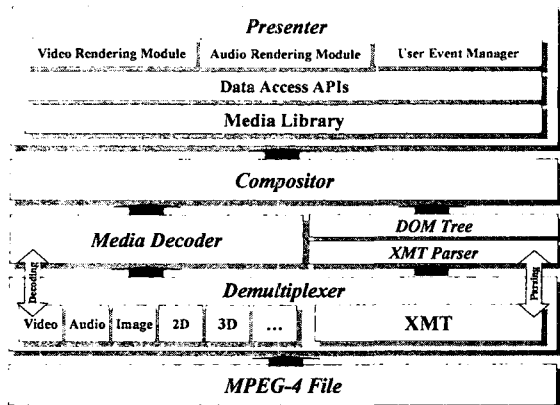


그림 1. 재생기 구조

### 3. XMT를 이용한 3차원 재생방법

#### 3.1 XMT 파일

MPEG-4에서는 콘텐츠 전체의 구성을 나타내는 장면 기술(Scene Description) 정보와 이 장면에 포함된 개개 콘텐츠 객체의 종류 및 비트스트림의 특성 정보를 나타내는 객체 기술(Object Description) 정보를 통해 콘텐츠의 전반적인 구성과 특성을 표현하게 된다. 그와 관련되어 MPEG-4에서는 장면을 기술하는 씬 디스크립션 언어로 BIFS와 XMT를 표준화하고 있다. XMT는 XML에 기반을 두는 것으로 기존의 BIFS가 가졌던 객체의 재사용이나 확장에 대한 한계성을 극복하고, 파싱기능과 중간 자료구조에 대한 정의를 간단히 함으로써 사용자 제어 측면에서 기존의 BIFS보다는 효과적이라고 볼 수 있다. 이러한 특징은 이미 생성된 기존의 다른 MPEG-4파일의 XMT에도 적용하여 추가적으로 사용자가 삽입이나 삭제 작업등을 할 수 있으므로 그 효율성이 높다고 할 수 있다.

본 논문에서는 사용자에 의한 씬 제어가 쉬운 XMT 형식에서 오디오 및 비디오 노드, 2D 노드 및 3차원 재생을 위한 3D 노드 등을 기술하고, 사용자 인터페이스 상에서 각 객체 렌더링을 위한 그래픽 모듈 라이브러리를 호출함으로써 MPEG-4를 재생 할 수 있는 시스템을 소개하고 구현하였다.

#### 3.2 3D 노드 디스크립션

MPEG-4 씬을 기술하게 되는 노드는 다양한 노드들의 집합으로 표현될 수 있는데, 노드는 MPEG-4에서 규격화된 비디오와 오디오를 지원하기 위한 노드, 2D 씬을 나타내기 위한 노드, 3차원 씬 재생을 위한 3D 노드 등으로 구성될 수 있다. 씬 디스크립션은 또한 2D 씬의 일부에 3D 씬을 조합시켜 복잡한 씬을 기술할 수 있을 뿐

만 아니라, XMT 형식의 노드 기술에서는 사용자 조작과 제어 작업을 직접 함으로써 씬을 자유롭게 생성될 수 있다.

본 논문에서 제안하는 XMT의 3D 객체의 노드 기술은 객체 렌더링에 필요한 노드 정의뿐만 아니라, 3D 노드가 가진 특성이 기술되는 속성노드와 노드의 값이 기술된다. 그리고, 각 노드는 그룹노드와 자식노드로 직접 지정이 가능하며, 또한 MPEG-4 각 객체는 하나의 노드를 생성하고, 생성된 노드에는 여러 특성들에 해당하는 속성노드를 생성한다. 그 노드는 객체노드의 자식노드로 불게 되고 그러한 객체노드들은 그룹을 형성하여 그룹노드의 자식노드가 되는 구조로 여러 단계의 계층구조를 가질 수 있다. XMT에서의 노드 기술은 노드명 자체를 기술하게 되는데, 일반 XML 문서에서처럼 요소(Element)들을 가지고 있으며, 요소들은 다시 자신들의 특성(Attribute)을 가지며, 이러한 요소와 특성들은 여러 단계의 계층 구조를 가질 수 있는 형태를 가지고 있다. 그러한 노드는 하나나 그 이상의 자식노드를 가질 수 있다. 그림 2는 3D 노드로 구성된 XMT파일의 예를 나타낸다.

```

<?xml version="1.0" encoding="UTF-8" ?>
<XMT xmlns="http://www.nonspace.com/schemas/xmt_0.xsd">
  <node>
    <meta NAME="GENERATOR" CONTENT="XMT Authoring Tool for Multimedia Content" />
    <meta NAME="AUTHOR" CONTENT="Computer Language/Media Lab. In KML" />
    <ns:sceneDescription />
  </node>
  <ns:scene>
    <ns:group DEF="Group0">
      <ns:node>
        <switch DEF="Switch001" whichChoice="1">
          <ns:group DEF="Group001">
            <ns:node>
              <switch DEF="Switch002" whichChoice="1">
                <ns:group DEF="Group002">
                  <ns:node>
                    <ns:material DEF="Material001" ambient="0.2" diffuseColor="0.8 0.8 0.8" emissiveColor="0.0 0.0 0.0" shininess="0.2" specularColor="0.0 0.0 0.0" transparency="0"><Material>
                      <ns:appearance>
                        <ns:box DEF="Box0" size="10 10 10"><Box>

```

그림 2. XMT파일의 3D 노드 구성 예

시스템에서 정의되는 3D 노드로는 Sphere, Cone, Cylinder, Box등이 있으며, 각 노드는 노드의 특성에 맞게 geometry, position, color, texture등과 같은 속성들을 가지게 된다. Cone노드인 경우 Shape노드를 부모노드로 가지며, bottomRadius, height, side, bottom같은 속성노드를 가진다. 노드들로 구성된 객체의 씬 디스크립션은 텍스트 기반의 특징으로 사용자 이벤트같은 추가적인 씬 제어로 노드의 확장이 가능하다. 또한 이러한 씬 디스크립션 방법은 기존의 BIFS에 의한 씬 디스크립션 방법을 포함하는 것이기도 하다.

### 3.3 XMT 파서

XMT 파일은 MPEG-4의 씬이 기술된 노드들의 집합으로 XMT 파서에 의해서 파싱이 이루어진다. 이것은 MPEG-4를 이루는 노드와 필드는 XMT의 요소(Element)와 특성(Attribute)으로 전환될 수 있기 때문에 가능한 것이다.

XMT 파서는 XMT 문서의 내부 자료구조 생성 및 분석을 위한 것이며, 파싱된 결과로 생성되는 DOM은 XML DOM에 기반을 두는 것으로, 미디어 정보에 대한 씬 디스크립션, 객체 디스크립션, 라우트 정보, 인코딩 정보등을 위한 노드의 다양한 트리구조를 가지게 된다.

DOM 트리는 XMT 문서안의 정보에 대한 깊이 우선 탐색으로 계층적 접근이 가능하며, 또한 노드는 2D, 3D 씬 디스크립션을 위해 사용되어 질 수 있다. 특히 XMT는 XML을 기반으로 정의된 언어이기 때문에 기존의 XML 파서를 이용한 파싱이 가능하며, 그 유효성을 검증할 수 있는 DTD나 스키마를 제공한다. XML을 기반으로 정의되어 있는 언어들의 일반적인 특성에 따라 XMT도 XML 기본 문법을 준수하고 언어마다 정의되어 있는 DTD나 스키마로 정의된 문법을 준수하는 유효한 문서이기만 하면 XML 파서를 이용한 해석이 가능하다.

따라서 본 논문에서 제안하는 MPEG-4 재생 시스템은 텍스트 형태의 XMT 파일을 XML 파서를 이용하여 파싱하며, 그 문서의 유효성 검증은 스키마를 통해 이루어진다. XMT를 위한 스키마에는 XMT노드에 대한 모든 구성요소와 특성이 정의되어 있다. 파싱된 결과로 생성되는 DOM 트리를 통하여 노드에 대한 정보를 얻고, 재생에 필요한 정보를 관리 및 처리하게 된다.

### 3.4 그래픽 모듈

그래픽 라이브러리는 사용자 인터페이스상에서 MPEG-4에 정의된 각 객체들의 재생을 위한 것으로 비디오, 오디오, 이미지, 2D, 3D 등의 다양한 미디어 객체들 데이터 액세스 APIs를 통해 호출하며, 기존의 MPEG-4 재생기에서 지원하지 못했던 MPEG-4에서 정의된 객체들의 다양한 데이터 포맷을 지원한다. 또한 이 라이브러리는 재생 시스템에서 독립적으로 관리되어질 수 있고, 필요시 호출될 수 있다.

### 4. 구현결과

본 논문에서 제안하는 MPEG-4 재생 시스템은 Windows XP 환경에서 Visual C++를 이용해서 구현하였다. 그림 3은 로컬의 MPEG-4 파일이 실제 구현되는 화면이다. 재생시에 발생하는 사용자 이벤트도 씬 디스크립션 정보안에 기술이 되어 있으며, 이벤트에 대한 감지 및 처리도 병행적으로 수행한다.

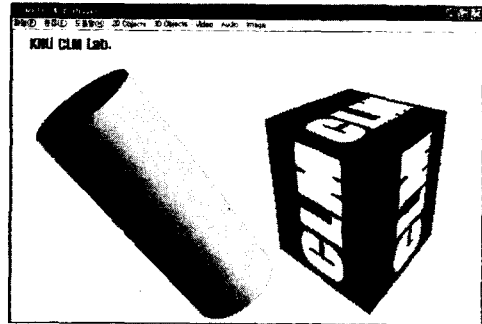


그림 3. MPEG-4 재생 화면

### 5. 결론

본 논문에서는 XML 기반의 MPEG-4 씬 정보를 기술하는 XMT를 이용하여 노드를 기술하고 속성을 추출하여 화면에 재생하는 시스템을 제안하고 구현하였다. 제안하는 시스템은 XMT 파일을 XMT 파서로 파싱하는데 XMT 파일의 유효성 검증을 위해서 스키마를 적용한다. 파싱된 후 문서 정보를 구조화하는 DOM 트리가 생성되면 DOM에서 제공하는 표준 DOM APIs를 이용해서 재생에 필요한 그래픽 라이브러리 호출을 위한 정보를 추출한다. 추출된 정보를 통해서 각 정보의 속성에 맞는 그래픽 라이브러리를 그래픽 모듈에서 호출하여 사용자 인터페이스에서 재생하게 된다.

제안하는 시스템을 통해서 사용자는 기존의 XMT 파일에도 추가적인 노드 확장 및 수정이 용이하며, 기존의 재생기에서의 구현이 제한적이었던 3D의 경우도 사용자는 자유롭게 필요에 따라서 텍스트 삽입과 그래픽 라이브러리 호출로서 구현의 한계를 극복할 수 있다.

향후 연구 과제로는 모바일 환경에서의 재생구현과 MPEG-4에 정의된 다양한 데이터 포맷을 XMT에 적용하는 것이다.

### 6. 참고문헌

- [1] R. Koenen, "MPEG-4 Overview - (V.18 - Singapore Version)," *ISO/IEC JTC1/SC29/WG11 N4030*, March 2001.
- [2] M. Kim, S. Wood and L.T. Cheok, "Extensible MPEG-4 textual format(XMT)," in Proc. ACM multimedia workshops., Los Angeles, California, United States, 2000, pp.71-74.
- [3] K. H. Kim, W. S. Cheong, H. C. Kim, J. W. Kim, "Interactive Broadcasting Contents Authoring System", *Telecommunications Review*, vol.11, No.5, Sep. 2001, pp. 688-698.
- [4] W3C recommendation(Ver1.0), Document Object Model(DOM) Level2 Core Specification, <http://www.w3.org/TR/DOM-Level-2-Core/>.