

MPEG-4 저작 시스템에서 다양한 멀티미디어 언어로의 변환기법

임영순*, 김희선**, 김상욱*
경북대학교 컴퓨터과학과*, 위덕대학교 컴퓨터멀티미디어공학부**

Conversion Mechanism of variable Multimedia Languages in MPEG-4 Authoring System

Y. Lim, H. Kim, S. Kim
Dept. of Computer Science, Kyungpook National Univ.*
Dept. of Division of Computer & Multimedia Engineering, Uiduk Univ.**

요 약

XMT 는 텍스트 형식의 MPEG-4 씬 기술 언어로 방송용 오디오/비디오 편집 및 사용자 중심의 미디어 콘텐츠 개발에 활용될 수 있다. 본 논문에서는 다양한 재생 환경에서 콘텐츠의 상호 교환을 지원하는 XMT 저작 시스템을 제안한다. XMT 저작 시스템은 XMT 의 두 가지 파일 포맷인 α 와 Ω 를 생성한다. 두 파일 포맷은 같은 객체를 표현하는 방법이 다르므로, α 와 Ω 를 위한 추상화된 인터페이스를 제공한다. 또한, 두 개의 파일 포맷을 지원하는 내부 자료 구조를 정의하고, XMT- α 를 BIFS 로 변환하는 기능과 XMT- Ω 를 SMIL 과 XMT- α 로 변환하는 기능을 제공하여 XMT 의 다양한 환경에서 멀티미디어의 상호 교환성을 제공한다.

1. 서론

MPEG-4 시스템은 다양한 멀티미디어 데이터로 이루어지는 상호작용 가능한 시청각 씬의 부호화된 표현에 대한 표준이다. MPEG-4 시스템은 BIFS(Binary Format for Scene)외에 BIFS 스트림 표현의 텍스트 포맷 인 XMT(eXtensible MPEG-4 Textual format)를 정의하여 여러 멀티미디어 언어 간의 상호 교환성을 제공한다[1-3].

XMT 는 XML 포맷으로 작성되어 있으므로 씬 구성 정보에 대한 가독성이 높으며, XSL(XML Stylesheet Language)[4]을 이용하여 다른 언어로의 변환이 용이하므로 MPEG-4 씬을 XMT 로 저작하면 용도에 따라 다른 언어로 변환할 수 있으므로, SMIL, VRML, MPEG-4 재생기를 사이에서 상호 교환될 수 있다. XMT 는 2 개의 추상화된 언어를 제공하는데 그 중 하나는 BIFS 의 이진 선택스가 XML 로 일대일 대응되는 구조를 가진 XMT- α 이다. XMT- α 는 X3D[5]와 유사하고, 씬의 세부적인 영역까지 표현하는 하위 레벨 디스크립션 언어이다. 다른

본 연구는 한국과학재단 목적기초연구(R04-2002- 000-20026-0) 자원으로 수행되었음.

하나는 XMT-Ω라 불리는 상위 레벨 디스크립션 언어로 SMIL[6] 기반으로 작성되어 있다. 따라서 XMT 로 작성된 콘텐츠 정보는 각각 BIFS, VRML(Virtual Reality Modeling Language)[7], SMIL(Synchronized Multimedia Integration Language) 등의 언어로 변환이 용이하다.

본 논문의 제 2 절에서는 다양한 멀티미디어 언어들에서 대해서 설명하고 제 3 절에서는 MPEG-4 저작 시스템의 구조를 설명하고, 각 구조별 특징에 대해서 설명한다. 그리고 제 4 절에서는 평가를 보이고 끝으로 제 5 절에서는 결론을 내린다.

2. 다양한 멀티미디어 언어

멀티미디어 프라젠테이션은 여러 미디어 요소들(비디오, 오디오, 이미지, 텍스트, 2D, 3D 그래픽 등)로 구성되어 있다. 멀티미디어 디스크립션 언어에 따라 이들 미디어 요소들은 시간적 순서와 공간 배치 등을 기술하는 방법이 다르다. 예를 들어 MPEG-4 는 BIFS 라는 이진 멀티미디어 디스크립션 포맷을 제공한다. 표 1 은 각 멀티미디어 파일들의 특징을 기술 한 것이다[8].

표 1 멀티미디어 언어 비교

	VRML	BIFS	SMIL
공간구성	3D	2D, 3D	2D 만 가능
압축	Zip 압축	압축이진포맷	지원 안 함
스트리밍	지원 안 함	스트리밍 지원	지원 안 함
동적구성	지원 안 함	업데이트메커니즘에 의해 지원	지원 안 함
다른 환경과의 통합	Web 환경에 제한	Web 환경에 MPEG-4 매핑 가능	Web 어플리케이션에 제한
2D 그래픽	지원 안 함	지원	SVG 사용
3D 그래픽	지원	지원	지원 안 함
애니메이션	VRML 과 유사	SMIL BIFS-Anim 메커니즘	SMIL 애니메이션 모듈

BIFS 기반의 MPEG-4 는 기본적으로 2, 3 차원의 공간구성을 모두 제공하고 있다. MPEG-4 의 텍스트 포맷의 하나인 XMT 는 2, 3 차원 객체의 디스크립션 정의가 가능한 반면 VRML, SMIL 은 각각 3 차원, 2 차원으로 제한되어 있다. SMIL 은 2 차원

미디어들의 시공간 요소의 결합으로 이루어진다. XMT-Ω는 SMIL 을 기반으로 만들어졌기 때문에 유사한 요소들이 많다. 그러나 XMT-Ω는 MPEG-4 의 특징을 텍스트 형태로 표현한 언어이기 때문에 SMIL 이 갖지 않는 특성을 가진다. SMIL 에서 표현되는 미디어 객체는 audio, video, still picture, still text, text stream, animation 등이고, XMT-Ω에서 기술하는 객체는 audio, video, img, string, rectangle, circle, lines 등으로 SMIL 에서 포함하지 않는 기하객체들을 포함하고 있다. 기하객체를 표현하기 위한 노드들은 SMIL 의 미디어 요소로 존재하지 않으므로 2 차원 기하객체를 표현하기 위하여 2 차원 벡터 그래픽을 표현하는 언어인 SVG[9]를 이용하여 SMIL+SVG 스펙으로 나타낸다.

3. 저작 시스템 구조

본 논문에서 제안하는 XMT 저작 시스템은 XMT 에 대한 시각적 저작 환경을 제공하여, 사용자가 쉽고 직관적으로 XMT 노드를 저작할 수 있도록 돕고, 생성된 XMT 포맷은 다양한 재생 환경을 지원하기 위하여 다른 멀티미디어 포맷으로 변환 가능하다.

3.1 시스템 구조

XMT 저작 시스템은 크게 유저 인터페이스와 씬 디스크립션 트리 관리기, XMT 파서 / 파일 생성기와 멀티미디어 파일 변환기로 구성된다. 그림 1 은 MPEG-4 저작 시스템의 구조이다.

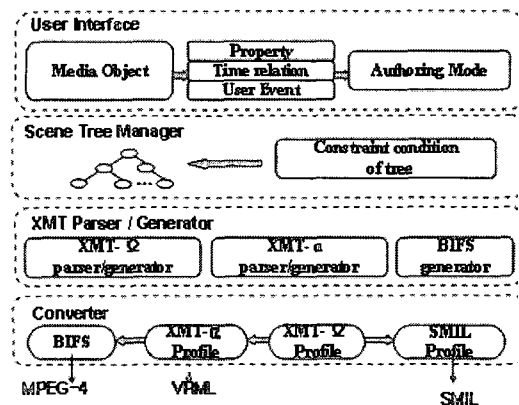


그림 1 MPEG-4 저작 시스템 구조

사용자 인터페이스는 저작을 쉽게 하기 위한 시각적인 환경을 지원한다. 하나의 장면을 구성하는 2 차원 기하객체, 이미지, 비디오, 오디오, 텍스트, 애니메이션 등의 미디어 객체들을 저작공간에 쉽게 나타내기 위하여 아이콘을 제공하며, 객체의 속성과 썸을 이루는 객체들 사이의 시간관계에 대한 속성과 이벤트 속성을 지원한다. 또 XMT 파서/파일 생성 모듈은 XMT 표준의 2 가지 파일 포맷의 DTD 혹은 스키마 정의에 의해 XMT- α 와 XMT- Ω 로 각각의 생성된다. 사용자 인터페이스에서 설정한 저작모드에 의해 파일 포맷을 구분하고, 설정 포맷에 따라 썸 디스크립션 트리를 탐색하여 필요한 객체들의 정보를 기재하게 된다. 생성된 각각의 포맷은 멀티미디어 변환기를 거쳐 사용자의 목적에 따라 MPEG-4, SMIL 혹은 VRML로 생성할 수 있다.

3.2 XMT 파일 생성

저작도구의 중간 산출물은 XMT- Ω 와 XMT- α 파일이다. 사용자 인터페이스에서 생성한 장면에 대한 내부적인 자료구조인 썸 디스크립션 트리는 XMT의 두 개의 프로파일을 모두 지원할 수 있도록 구성되어 있다. 파일 생성기는 썸 디스크립션 트리를 깊이 우선 탐색하여 XMT 텍스트 형태로 출력하고, 각각의 프로파일은 XMT- Ω 와 XMT- α 의 DTD 혹은 스키마에 따른다. 또한 썸 디스크립션 트리로부터 시청각 객체의 속성 정보 및 라우트 정보, 시간 정보 등을 추출하여 XMT 파일을 생성한다. XMT 파일은 XML을 기반으로 정의되어 있기 때문에 제한하는 저작도구에서 XMT의 파싱은 XML 파서를 이용하고 파싱의 결과로 생성된 DOM 트리를 다시 내부 자료구조인 썸 디스크립션 트리로 변환하는 기능을 가진다.

3.3 멀티미디어 변환기의 구조

XMT 파일 파서/생성기를 통해 생성된 XMT 파일을 다양한 멀티미디어 파일로 변환하기 위하여 본 저작도구에서는 XSLT 엔진을 사용한다. XMT- Ω 로 생성된 썸 정보는 변환을 목적으로 미리 정의한 XSL 파일과 XSLT 엔진에 의하여 SMIL 파일로 생성된다.

변환기의 입력은 저작 시스템에서 생성한 XMT 파일과 변환을 적용하기 위한 XSL 파일이다. 이 두 파일은 MSXML4에 의해 각각 DOM 트리로

파싱된다. 파싱한 DOM 트리를 변환엔진에 적용하면 SMIL 혹은 VRML 파일이 생성된다. 작성된 XSL 파일은 XMT- Ω 파일을 SMIL로 변환하는 방법, XMT- α 를 VRML로 변환하는 방법, 혹은 XMT- Ω 를 XMT- α 로 변환하여 BIFS로 변환하는 방법을 기술한 언어이다. 실제로 변환 규칙이 XSL 파일에 의해 기술된다. XMT의 SMIL로의 변환의 경우에 XSL 파일은 크게 SMIL의 Head를 기술하는 부분과 Body를 기술하는 부분으로 나눌 수 있다. head 부분은 썸의 크기, 배경 색등의 정보와 각 객체 단위로 출력될 위치를 기술한 layout 정보, transition 정보, 메타 정보, switch에 관한 정보 처리방법에 대하여 중점적으로 기술하고, body 부분은 출력될 객체들의 정보에 관한 변환 규칙을 기술한다. 작성된 XSL 파일은 여러 개의 템플릿으로 구성되어 있다. 루트 템플릿으로부터 각 미디어 별로 해당되는 템플릿을 적용함으로써 객체 및 모듈 별로 변환이 가능하다. 실제 XMT DOM 트리의 정보를 SMIL의 DOM 트리로 변환 구조를 그림 2에 나타내었으며, XMT의 프로파일의 트리 구조는 XSL 파일에 의하여 트리 구조의 변경이나 새로운 노드가 생성 또는 삭제된다.

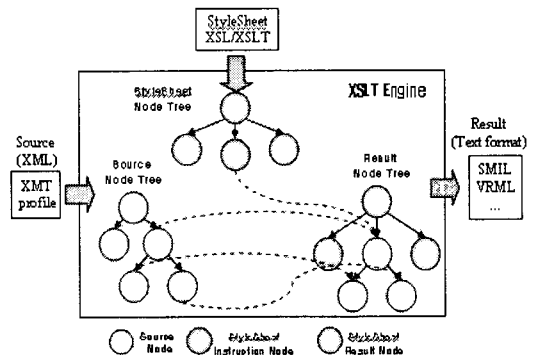


그림 2 XMT 파일의 변환

4. 평가

제안하는 저작 시스템은 멀티미디어 재생을 위한 썸 구성을 MPEG-4의 썸 기술 언어 중 하나인 XMT로 저장하고, XMT의 두 가지 포맷을 이용하여 SMIL, MPEG-4 등 다양한 멀티미디어 포맷으로 출력한다. 그림 3은 제안하는 저작도구에서 3 차원

노드를 제외한 XMT-Ω 태그의 SMIL 태그로의 변환 정도를 나타낸다. 아무런 변환 기법을 사용하지 않고 태그만 비교 하였을 경우 54%의 태그만이 두 멀티미디어 언어에서 일치함을 알 수 있다. 이는 대부분의 미디어 모듈이 SMIL 에서 지원하는 않은 경우가 많기 때문이다. 이러한 문제점을 해결하기 위하여 XMT-Ω의 태그인 <audioClip>, <sound>, <string>은 SMIL 의 <audio>나 <text> 태그의 이름을 변경하여 그 역할을 할 수 있도록 하고, <fontStyle>,<material>, <transformation>, <outline>의 XMT 태그는 SMIL 의 미디어 속성으로 변경하여 재생된다. 또한 SMIL 에서 지원하지 않는 2 차원 기하객체인 <circle>, <curve>, <lines>, <points>, <polygons>, <rectangle> 태그는 SMIL+SVG 표준을 따르도록 SVG 로 변환하고, 일부 노드들은 변환규칙을 적용한 XSL 파일을 통하여 변환한다. 저작도구 내부에서 제공하는 파일 변환기를 거쳐 나온 SMIL 파일은 XMT 에서 제공하는 태그의 90% 가량을 변경할 수 있다.

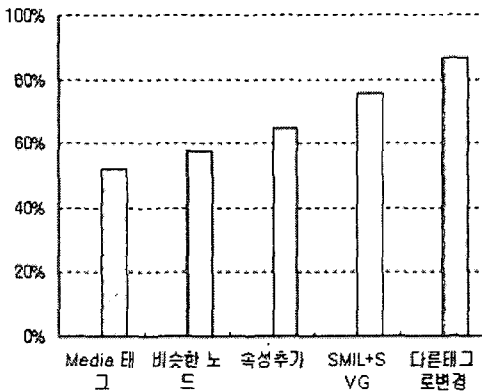


그림 3 태그변환정도

5. 결론

본 논문에서는 XMT 를 생성하는 MPEG-4 저작 시스템을 개발하고, 생성된 XMT 파일 중 XMT-Ω를 SMIL 로 변환하는 방법을 제시하였다. 제안하는 저작도구는 썬의 정보를 XMT 형태로 저장할 수 있을 뿐 아니라 SMIL 로도 출력 가능하므로 XMT 의 본래 목적인 다양한 재생 환경을 지원한다. XMT-Ω는 SMIL 기반으로 정의되어 있어 SMIL 과 노드 구성 방식이 유사하다. 그러나 XMT-Ω를 구성하는

객체 모두가 SMIL 로 표현 가능한 것은 아니다. 제안하는 저작 도구는 저작 기능과 변환 기능이 저작도구 내에 있어 저작과 변환을 따로 했던 기존에 다른 도구들에 비해 효율적이다. 제안하는 저작도구와 관련하여 현재 XMT 의 VRML 로의 변환에 관한 연구가 진행 중에 있으며 XMT 의 2 차원 노드를 VRML 의 3 차원 노드로 변환하는 방법과 노드 매핑 간의 연관성에 대하여 초점을 두고 있다.

[참고문헌]

- [1] K. Cha, H. Kim and S. Kim, "The Design and Development of MPEG-4 Contents Authoring System," Journal of The KISS, Vol.7, No.4, pp.309-315, August 2001
- [2] ISO/IEC 14496-1:2000 MPEG-4 Systems October 2000
- [3] ISO/IEC JTC1/SC29/WG11/N4091, Study of ISO/IEC 14496-1:2001/PDAM2:Extensible MPEG-4 Textual Format(XMT), March 2001
- [4] XSLT 2nd Edition, Programmer's Reference by Michael Kay
- [5] ISO/IEC 19776 X3D February, 2002
- [6] World Wide Web Consortium, "Synchronized Media Integration Language (SMIL 2.0)", <http://www.w3.org/TR/smil20/>
- [7] VRML 97, ISO/IEC DIS 14772-1, 1997
- [8] C.Concolato, J.-C. Dufourd "Comparision of MPEG-4 BIFS and some other Multimedia description languages" Workshop and Exhibition on MPEG-4
- [9] Scalable Vector Graphics (SVG) 1.1 Specification, W3C Recommendation, <http://www.w3.org/TR/SVG11/> REC - smil, 2003