

키워드 검색이 가능한 MPEG-4 저작도구 설계

임태진^o 이명원* 홍기천**

수원대학교 컴퓨터학과

*수원대학교 인터넷정보공학과

**수원대학교 정보통신공학과

{tjleem, mwlee, kchong}@suwon.ac.kr

Design of a MPEG-4 Authoring Tool with Keyword Retrieval

Tai Jin Leem^o, Myeong Won Lee*, Kicheon Hong**

Dept. of Computer Science, The U. of Suwon

*Dept. of Internet Information Engineering, The U. of Suwon

**Dept. of Information and Telecommunications Engineering, The U. of Suwon

요 약

최근 MPEG-4를 이용한 멀티미디어 응용이 활성화되고 있으나 아직까지는 단방향의 멀티미디어 스트리밍 전송이 대부분이다. 본 논문에서는 일반 사용자의 요구 사항을 서버 측에 전송할 수 있도록 하는 양방향의 멀티미디어 전송을 위한 시스템 설계에 대해 기술한다. 이 시스템은 서버측의 MPEG-4 콘텐츠를 클라이언트에서 키워드를 이용하여 원하는 부분의 내용을 검색하여 스트림 전송해서 받을 수 있도록 한다. 검색하고자 하는 키워드를 입력하면 해당 키워드의 장면트리구조에서 검색된 위치에서부터 MPEG-4 파일을 재생할 수 있는 것이다. 본 논문에서는 이러한 키워드 입력 인터페이스를 포함하는 MPEG-4 저작 도구의 구성에 대해 소개한다.

1. 서 론

MPEG-4는 방송의 영상매체 전달능력, 통신의 양방향성, 그리고 컴퓨터의 대화적인 사용자 인터페이스를 통합한 새로운 멀티미디어 부호화에 관한 국제 표준으로서 차세대 멀티미디어 서비스 분야의 핵심 기술이다[1][2]. MPEG-4 기술은 자연영상뿐만 아니라 그래픽의 합성 영상 등 다양한 멀티미디어 정보를 대화형으로 통신하고 사용자가 쉽게 접근, 편집, 처리할 수 있는 기능을 제공하기 때문에 현재 각광받고 있는 인터넷 정보 서비스를 포함한 멀티미디어 서비스에 있어서 필수적인 핵심 기술이다.

멀티미디어 정보 통신 서비스로는 영상 전화, 인터넷 방송, 원격 교육, 원격 게임/오락, 가상 쇼핑 및 VOD 등을 들 수 있으며 대부분의 서비스 내용에 있어서 미디어 스트림은 파일 단위로 순차적으로 플레이되는 것이 일반적이다. 이와 같은 응용 개발을 위해서는 일반사용자가 사용할 수 있는 저작도구가 필요하게 된다. 기존의 MPEG-4 저작도구로는 MPEG-4 저작 도구로는 프랑스의 ENST에서 개발한 MPEG-Pro[9], 그리스의 ITI에서 개발한 3D MPEG-4 저작도구, envivio의 EBS 및 2D MPEG-4 저작도구 등이 있다. 이러한 기존의 저작도구에서는 사용자의 키워드 입력에 의한 스트림 내용 검색의 기능은 제공하지 않고 있다.

본 논문에서는 미디어 스트림 중에 원하는 부분의 내용을 검색할 수 있도록 멀티미디어를 구성해주는 저작도구의 설계에 대해 기술한다. 이를 위해서는 미디어 스트림에 검색을 위한 정보를 포함시켜 전달하도록 하는 구

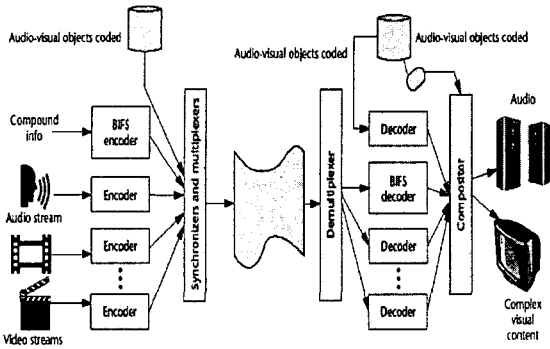
조가 필요하게 된다. 본 연구에서는 이러한 기능 제공을 위해서 MPEG-4를 사용하여 미디어 스트림과 검색 정보를 동기화시켜서 전송시키도록 구성하여 사용자의 요구에 따라 필요한 스트림을 서비스 받을 수 있도록 하였다.

본 논문에서의 연구 초점은 멀티미디어 서비스에 관한 국제 표준인 MPEG-4 기술을 이용하여 단순히 멀티미디어를 재생하여 보는데 그치는 것이 아니라 사용자의 선택에 따라 스트림 내의 검색된 위치에서부터 재생이 가능하도록 하는데 있다.

2. MPEG-4 시스템

멀티미디어 표준화 작업은 4.8~64kbps 정도의 비트율로 AV 데이터로 통신하고, AV 데이터베이스에 접근하며, AV 데이터를 다룰 수 있도록 하는 새로운 방법들을 지원하기 위한 부호화 표준이다. 현재 MPEG-4 표준의 산업화와 개선된 도구를 표준화하는 작업이 진행 중이다. MPEG-4는 예전의 H.261, MPEG-1, MPEG-2 기술과 같은 블록 중심의 부호화 방식과 차별하여 영상의 내용을 고려하는 부호화 방식을 택하였다[3].

MPEG-4는 다양한 객체들을 통합한 멀티미디어 부호화 표준으로 MPEG-4 시스템은 장면기술과 다중화의 2계층으로 구성되고, 다중화층 아래에 DMIF[4]라 불리는 기존의 전송 방식과의 인터페이스를 규정하는 층이 있다(그림 1). 이와 같이 MPEG-4 시스템은 여러 객체(오디오, 동영상, CG, 정지영상을 포함한 넓은 의미에서 미디어 객체)의 부호화 비트스트림의 다중화 및 동기뿐만 아니라, 장면기술에 따른 합성을 취급한다[5].



(그림 1) MPEG-4 시스템

MPEG-4 시스템의 BIFS(Binary Format for Scene)는 MPEG-4 를 이용하여 작성된 콘텐츠를 이용하여 오디오 비주얼 장면을 표현하기 위한 언어이다[6]. BIFS 는 VRML(Virtual Reality Modeling Language)을 기반으로 하여 개발되었으나 MPEG-4 응용프로그램을 개발하는데 용이하게 하기 위하여 VRML 노드의 일부를 삭제 또는 수정하였으며 스트리밍, 2D 장면을 지원하기 위한 노드 등을 추가하였다. 또한 간결한 이진형식으로 장면을 표현함으로써 장면을 기술하기 위한 메모리 양을 최소화하였다.

MPEG-4 에서는 BIFS 콘텐츠 저작 과정을 효과적으로 지원할 수 있도록 XMT(eXtensible MPEG-4 Textural format) 콘텐츠 저작 언어를 표준화하고 있다[7]. XMT 는 MPEG-4 장면 기술자인 BIFS 는 이진형식으로 정의되어 있어서 다루기에 불편함이 있는 제약을 보완하기 위해 개발되었으며 이를 이용하여 XML 기반의 MPEG-4 장면 기술과 시스템 특성을 기술할 수가 있다. XMT 는 XMT-0 와 XMT-A 두 가지 포맷이 있는데 XMT-0 는 상위 레벨의 장면기술자로서 사용자에게 보다 쉽게 장면과 시스템을 기술한 것이고, XMT-A 는 하위레벨에서 장면과 시스템을 기술하기 위한 것이다. 또한 이 두 포맷은 XMT-0 to A Mapper 모듈을 이용하여 상호 변환이 가능하다. XMT 의 이러한 통합성은 웹과 방송응용 사이의 상호 연동을 지원한다.

3. 키워드 검색이 가능한 MPEG-4 콘텐츠 생성 방법

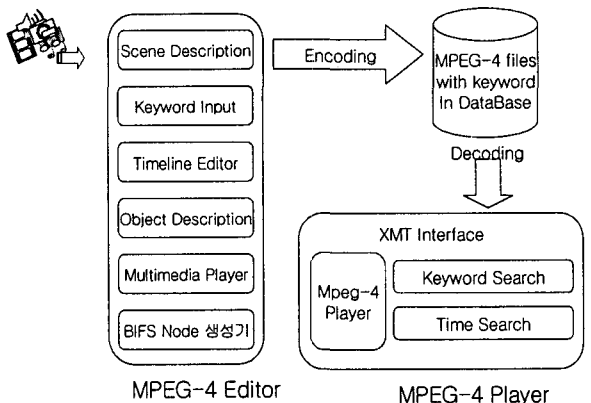
MPEG-4 콘텐츠가 키워드 검색이 가능하도록 구성하기 위해서는 두 가지 기능이 포함되어야 한다. 첫 번째는 콘텐츠를 구성하는 미디어들의 공간상의 동기화 정의와 관계되는 장면기술언어인 BIFS 재구성 방법이 필요하다. 장면을 기술하는 BIFS 에는 재생을 위한 멀티미디어 구성방법은 제공하지만 일부 스트림 검색을 위한 정보구성에 대해서는 정의되지 않는다. 이러한 기능을 제공하기 위해서는 BIFS 가 장면 자체의 정보 외에 장면을 검색하는데 필요한 부가적 정보를 포함하고 있어야 한다. 이를 위해 본 연구에서는 BIFS에 키워드 정보 구성을 위한 노드를 추가하도록 한다. 이 노드에는 키워드에 의해

스트림의 해당 위치 및 재생 관련 정보가 검색될 수 있도록 한다.

두 번째 기능으로는 BIFS에서는 시간 동기화 방법을 제공하지는 않으므로 미디어들의 시간 동기화 및 미디어 스트림 내 특정 부분의 재생을 위해 시간 처리 노드 추가를 필요로 한다. 이 노드 구성을 위해서는 객체간의 시간 동기화 기능과 함께 객체 내에서의 실제 시간 제어를 위한 기능이 포함된다. BIFS에는 각 미디어 객체를 제어하고 모니터링 하는 노드와 여러 객체들을 관리하고 객체의 시간적인 변화를 적용시키는 노드를 제공하고 있다. MediaContorl 노드는 재생, 반복, 소리 비활성 등의 객체 제어를 담당하는 노드이며 MediaSensor 노드는 재생 중에 있는 객체의 활성화 및 비활성을 제어한다. TemporalGroup 노드는 화면을 구성하는 객체들의 동시시작 및 동시종료에 대한 정보를 저장하며 Temporal Transform 노드는 재생중인 객체의 변화되는 정보들을 저장한다. 이러한 정보는 객체 단위로 지원되는 것이며 객체간이나 객체내에서의 시간 동기화는 정의되어 있지 않다. 본 논문에서는 위에서와 같이 객체간 시간 정보를 다루는 기능 외에 지정된 위치부터 주어진 시간 동안 일부 스트림의 재생을 지원하도록 한다.

4. 시스템 설계

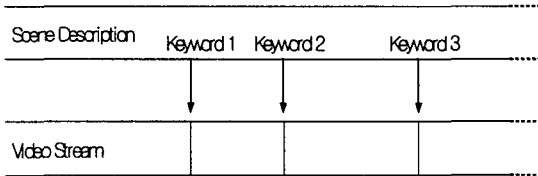
MPEG-4 스트림이 위의 기능을 갖기 위해서 MPEG-4 저작 시스템은 일반적인 미디어 객체를 포함한 장면 구성, 키워드 입력기, 타임라인 에디터, 객체 기술자, 멀티미디어 재생기, BIFS 노드 생성기를 포함하여야 한다. (그림 2)는 본 연구에서 제안하는 키워드 검색을 가능하게 하는 MPEG-4 저작도구 시스템 구성을 보여준다.



(그림 2) MPEG-4 저작도구 구성

키워드 검색을 하기 위해서는 저작도구에서 키워드 입력기를 이용하여 적절한 위치에 키워드를 입력해야 하는데, (그림 3)과 같이 비디오 스트림의 키워드를 입력할 해당위치에 장면 묘사로 키워드를 입력한다. 입력된 키워드는 BIFS 생성기에 의해 장면을 구성하며 MPEG-4 파

일과 함께 MPEG-4 파일 정보와 키워드 정보를 동시에 저장되도록 한다.



(그림 3) 키워드 입력

(그림 4)은 장면 트리에 대응하는 XMT 표현이다. 키워드를 300초에 삽입하고 1200초에 키워드를 삽입하는 경우를 보여준다. 나중에 검색을 할 경우 해당 값인 Chapter 1 과 Chapter 2를 검색할 경우 해당 위치의 검색이 가능하게 된다.

```

42 ...
43 <par begin="300s">
44 <insert atNode="keyword" atField="string" position="0" value="Chapter 1" />
45 </par>
46 <par begin="1200s">
47 <insert atNode="keyword" atField="string" position="0" value="Chapter 2" />
48 </par>
49 ...
    
```

(그림 4) 장면 검색을 위한 XMT 일부본

위와 같이 키워드 정보를 입력하여 검색을 하는 작업은 온라인 MPEG-4 재생기와 오프라인 MPEG-4 재생기에서 동일하게 적용될 수 있다. 온라인 MPEG-4 재생기란 서버로부터 MPEG-4가 스트림 전송되어 재생되는 경우를 말하며 오프라인 MPEG-4 재생기란 MPEG-4 파일로 독립적으로 재생되는 경우이다. MPEG-4 재생기에서 키워드 검색은 BIFS 정보에 접근하여 입력한 키워드를 검색한다. 검색 결과로는 해당 MPEG-4 파일이름과 해당 재생시간이 검색되며 원하는 키워드 검색 결과에서 선택을 하면 검색된 위치에서부터 재생이 가능하다.

본 시스템에서는 키워드 검색과 함께 시간 검색도 가능해서 원하는 시간을 입력하여 해당 시간부터 재생이 가능하도록 구성되었다. (그림 5)는 BIFS 에 추가되는 키워드 노드와 시간 노드의 정보 구성을 보여준다.

5. 결론

본 논문에서는 일반 사용자가 MPEG-4 응용을 개발할 수 있도록 하는 저작도구를 기술하였다. 이 저작도구에서는 키워드 입력을 포함한 멀티미디어 저작을 하고 이 결과를 MPEG-4 플레이어에서 재생하도록 해준다. 본 시스템의 장점은 사용자가 컨텐츠 저작시 대화형 인터페이스를 통해 키워드를 입력하고 멀티미디어 스트림을 구성할 수 있도록 해서, 원하는 위치를 검색하여 스트림을

	명 이름	데이터형식	길이	설명
키워드 노드 정보	idx	int	4	고유키 값
	file_name	varchar	50	파일 이름
	keyword	varchar	20	키워드
	descript	text		키워드에 대한 설명
	begin	int	4	시작 위치
시간 노드 정보	idx	int	4	고유키 값
	start_time	int	4	재생 시작 시간
	end_time	int	4	재생 종료시간
	duration	int	4	재생 실행시간
	offset	int	4	재생 지정시간
	current_time	int	4	현재 시간

(그림 5) 키워드 노드 및 시간 노드 정보

재생되도록 한 것이다. 본 연구의 결과는 원격 교육등의 인터넷 방송 응용 분야에서의 활용이 기대된다.

향후 연구계획은 윈도우 환경에서의 시스템 구현을 완성한 후 모바일 환경에서의 스트림 검색을 위한 임베디드 시스템 구현을 목표로 하고 있다.

참고문헌

- [1] Information technology-Coding of audio-visual objects Part 1: Systems, ISO/IEC 14496-1:2001.1
- [2] "Overview of the MPEG-4 Standard", <http://mpeg.telecomitalia.com/standards/mpeg-4/mpeg-4.htm>
- [3] 오승준, "ISO/IEC JTC1/SC29/WG11 MPEG-4 표준 동향 분석"
- [4] MPEG-4 DMIF Group, "Generic Coding of Audio-Visual Object : Part6 - DMIF", ISO/IEC JTC1/SC29/WG11 N2506 FDIS of ISO/IEC 14496-6, November 1998.
- [5] http://www.patentmap.or.kr/pm_inquiry/2002/2002-ee/2002-EE-07/EE07_html_chap_1/EE07-114_na.htm
- [6] "MPEG-4's Binary Format for Scene Description", http://ailab.chonbuk.ac.kr/~sjmun/mpeg4/tutorial/5-BIFS_paper.htm
- [7] ISO/IEC 14496-1, "Coding of Moving Pictures and Audio", ISO/IEC JTC 1/SC 29/WG11 N4699. Chapter 14-17, March 2002
- [8] 배수영, 김상욱, 마평수, "MPEG-4 저작 시스템에서 BIFS 생성 모듈", 정보과학회논문지, 제8권 제5호, p.520-p.529, 2002.10
- [9] Boughoufalah S., Dufourd J-C., Bouilhaguet F., "MPEG-Pro, an Authoring System for MPEG-4 with Temporal Constraints and Template Guided Editing", Proceedings of ICME2000, Vol. 1, pp.175-178, 2000.