

MPEG-7 Description 을 이용한 객체 기반 MPEG-4 콘텐츠 검색 시스템

정예선, 김규현
한국전자통신연구원 전파방송연구소 방송미디어부
E-mail: ysjoung, kkim@etri.re.kr

The Object oriented MPEG-4 Contents Retrieval System using MPEG-7 over MPEG-4 Scheme

Yesun Joung , Kyuheon Kim
Broadcasting Media Technology Dept., Radio & Broadcasting Laboratory
E-mail: ysjoung, kkim@etri.re.kr

요약

최근, 컴퓨터 환경의 발달로 각 분야에서 AV(Audio-Visual) 콘텐츠의 사용량이 기하급수적으로 증가되었다. 또한 컴퓨터, 방송 및 네트워크의 융합으로 이전 보다 다양한 종류의 어플리케이션에서 다양한 형태로 AV 콘텐츠가 사용되고 있다. 이에 따라, 이런 다양한 형태의 AV 콘텐츠를 표현하는 방법 및 검색하여 사용하는 방법에 대한 많은 연구들이 이루어지고 있다. 이런 연구들의 한 형태로 본 논문에서는 AV 콘텐츠를 표현하는 방법으로 MPEG-4 콘텐츠를 XML 에 기반하여 표현한 XMT(eXtensible MPEG-4 Textual format)와 AV 콘텐츠를 검색하는 방법으로 MPEG-7 표준을 사용하여 객체기반 MPEG-4 콘텐츠 검색 시스템을 제안하였다. 다양한 포맷의 미디어 데이터 및 그래픽 객체들을 각각의 하나의 객체로 간주하여 AV 콘텐츠를 구성하는 MPEG-4 시스템을 XML 을 기반으로 표현한 XMT 포맷으로는 MPEG-4 콘텐츠를 구성하는 미디어 데이터를 바탕으로 검색하는데 한계가 있다. 따라서 미디어 데이터 검색을 위한 메타데이터의 표준으로 사용되고 있는 MPEG-7 을 사용하여 객체 기반 MPEG-4 콘텐츠 검색이 효과적으로 이루어질 수 있도록 확장된 XMT 포맷을 제안하고, 이를 사용하는 시험 시스템을 구성하였다. 이렇게 함으로써 MPEG-4 로 표현된 객체 기반 AV 콘텐츠 검색 시 MPEG-7 메타데이터를 사용하여 보다 사용자가 원하는 콘텐츠를 효율적으로 검색하여 제공할 수 있을 것으로 사려된다.

1. Introduction

최근, 컴퓨터 환경의 발달로 각 분야에서 AV (Audio-Visual) 콘텐츠의 사용량이 기하급수적으로 증가되었다. 또한 컴퓨터, 방송 및 네트워크의 융합으로 보다 다양한 종류의 어플리케이션에서 다양한 형태로 AV 콘텐츠가 사용되고 있다. 이에 따라, AV 콘텐츠를 표현하는 방법 및 AV 콘텐츠를 검색하여 사용하는 방법에 대한 많은 연구들이 이루어지고 있다. 예를 들어, 요사이 AV 콘텐츠를 표현하는 방법으로는 XML에 기반한 SMIL, X3D, SVG, XMT 등이 있으며, AV 콘텐츠를 검색하는 방법으로는 텍스트 기반 검색, 내용기반 검색, MPEG-7, TV-Anytime 등이 있다. 이들 중에서 AV 콘텐츠를 표현하는 방법으로 XML에 기반하여 MPEG-4 시스템을 표현하는 XMT(eXtensible MPEG-4 Textual Format)는 SML(Synchronized Multimedia Integration Language), X3D(eXtensible 3D), SVG(Scalable Vector Graphics)등의 다른 AV 콘텐츠 표현 언어들에 바탕으로 설계되어 약간의 프로세서 과정을 거치면 각각의 표현 방식으로 상호작용(interoperability)이 가능하다. 또한 차세대 디지털 방송 콘텐츠의 가장 중요한 요구사항인 콘텐츠의 대화성(interactivity)을 지원하는 콘텐츠를 작성 할 수 있게 한다. 이런 XMT 포맷을 사용하여 표현된 MPEG-4 콘텐츠를 검색 하고자 할 때, MPEG-4 콘텐츠 특성상 다양한 포맷의 미디어 데이터 및 그래픽 객체들을 각

각 하나의 객체로 간주하여 하나의 콘텐츠를 구성하는데 이렇게 구성된 미디어 객체들을 바탕으로 사용자가 MPEG-4 콘텐츠를 검색하고자 할 때는 한계가 있다. 예를 들어, MPEG-4 콘텐츠를 구성하는 고호의 해바라기라는 그림을 바탕으로 검색하고자 한다면 AV 콘텐츠의 시공간적 위치 및 객체에 대한 정보만으로 구성된 XMT 만으로는 검색이 불가능하다.

따라서, 각각의 미디어 데이터 검색을 위해서는 미디어 데이터에 대한 메타데이터가 필요하다. 이런 미디어 데이터에 대한 메타데이터의 표준으로 MPEG-7 이 사용되고 있다. MPEG-7 표준은 종래의 텍스트 검색 방법과 내용 기반 검색 방법의 단점들을 보완하여 효과적으로 미디어 데이터를 검색할 수 있는 방법을 제공하고 있다. 기존의 텍스트 검색 방법은 AV 콘텐츠 내용을 키워드 중심의 텍스트 색인을 구성하는 것으로 사람의 주관적인 요소가 첨가 될 뿐 아니라, 텍스트가 지니는 한계에 의해 상당량의 잘못된 AV 콘텐츠 검색 결과를 가져오는 문제점이 있다. 반면에 내용 기반 검색 방법은 AV 콘텐츠에서 특징들을 자동적으로 추출하여 이를 검색하는데 사용하는 방법으로, 콘텐츠로부터 특징들을 추출하는데 시간이 많이 걸리며, 매번 검색 할 때마다 AV 콘텐츠의 특징을 추출해야 하는 문제점이 있다. 이런 문제점들을 극복하고 좀 더 효율적인 검색을 위하여 MPEG-7 표준을 제정하였다. MPEG-7 의 경우 AV 콘텐츠의 특징을 표현하는 기술자와 이런 기술자의 조합으로 구성되는 기술 구성, 기술 구성을 표현할 기술 언어를 표준화 하고 있다.

즉 텍스트 기반 메타데이터를 작성할 때 작성자의 주관적인 데이터가 적게 들어 갈 수 있도록 기술 할 수 있는 메타데이터 항목들을 표준으로 정하고, 미디어 데이터에 대한 고유의 특징인 color, texture, motion 과 같은 visual 특징들을 추출할 수 있는 표준적인 알고리즘을 정하고 이들을 표준으로 정한 기술언어로 표현할 수 있도록 하여 텍스트 검색과 내용 기반 검색 방법의 단점들을 적절히 보완하여 효과적으로 미디어 데이터를 검색할 수 있는 방법을 제공하고 있다.

본 논문에서는 위에서 언급한 바와 같이 MPEG-4 콘텐츠를 효과적으로 검색하기 위하여 MPEG-7 description을 적절히 사용하여 MPEG-4 콘텐츠 표현 방법을 구성하고, 이 표현 방법을 사용하여 MPEG-4 콘텐츠를 저작, 검색하는 시스템을 제안 하였다.

본 논문의 구성은 다음과 같이, 2 장에서 시험 시스템을 제한하게 된 동기 및 배경을 설명하고, 3 장에서 제안한 시험 시스템의 구현에 대해 설명할 것이다. 4 장에서 구현한 시험 시스템에서의 검색 결과에 대해 설명하고, 마지막으로 5 장에서는 결론을 맺고자 한다.

2. Background and Motivations

1 장에서 언급한 바와 같이, MPEG-4 콘텐츠를 구성하는 미디어 데이터를 중심으로 검색하고자 할 때, 미디어 데이터의 시공간적 위치 및 객체에 대한 간단한 정보로만 구성된 MPEG-4 텍스트 포맷, XMT 만으로는 검색에 한계가 있다. 따라서 미디어 데이터 검색을 위한 메타데이터의 표준인 MPEG-7 을 XMT 에 적절히 사용하여 이런 문제점을 해결하고자 한다. 이와 비슷한 시도를 한 두 가지 방법이 존재 한다.

첫 번째 방법[6]은 MPEG-7 을 이용하여 질의/브라우징/검색을 하고, 전송할 비디오 콘텐츠를 MPEG-4 를 사용하여 압축하는 시도이다. 이 방법의 경우 MPEG-7 의 color, motion, shape descriptor 들을 비디오 콘텐츠의 검색을 위한 메타데이터로 사용하고, 비디오 콘텐츠의 전송 압축 방식으로 MPEG-4 를 사용하고 있다. 하지만, 이 방법은 비디오 콘텐츠를 검색할 경우 각각의 비디오 콘텐츠에 대한 MPEG-7 메타데이터를 별도로 관리/전송 하여야 하며, 또한 color, motion, shape 과 같은 visual 특징들만 이용하므로 검색 질의어 구성에 한계가 존재 한다.

두 번째 방법[7]은 동기화하는 AV 콘텐츠를 MPEG-4 텍스트 포맷 보다 비교적 간단한 MPEG-7 메타데이터를 이용하여 저작 한 후, MPEG-4 XMT 포맷으로 전환하는 시도이다. 이 경우 역시 MPEG-4 콘텐츠를 검색할 때 별도의 MPEG-7 메타데이터를 이용하여야 하며, MPEG-4 interactivity 같은 기능은 MPEG-7 메타데이터로 표현하기 어렵다.

위의 두 가지 방법 모두 MPEG-4 콘텐츠의 미디어 데이터를 검색하고자 하고자 할 때 별도의 MPEG-7 메타데이터 파일을 사용해야 한다. 그러나 본 논문에서 제안한 방법은 MPEG-7 메타데이터를 MPEG-4 XMT 포맷 내부에 위치 시키기 때문에 별도의 MPEG-7 메타데이터 파일이 필요하지 않다. 이렇게 하기 위해서 본 논문에서는 두 가지 방법의 포맷을 제안한다. 그 첫 번째 방법은 그림 1 과 같다. MPEG-4 콘텐츠를 구성하

는 비디오, 오디오, 이미지와 같은 미디어 데이터에 대한 정보를 기술하는 Object Descriptor (OD)가 존재 하고, 이 OD 내부에 실제 미디어 데이터에 대한 위치 및 속성을 가리키는 Elementary Stream(ES)이 있다. MPEG-7 메타데이터를 Object Content Information(OCI) 혹은 IPMP(Intellectual Property and Management Protection) 스트림과 같이 미디어 데이터 OD 내부의 하나의 ES 로 간주 하여 포함시키는 방법이다. 이 경우는 MPEG-7 메타데이터의 크기가 큰 경우 유용하게 사용될 수 있다.

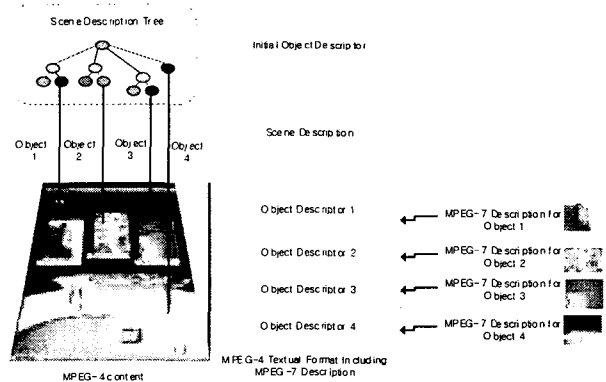


그림 1 MPEG-7 over MPEG-4 의 첫 번째 방법

두 번째 방법은 그림 2 와 같다. MPEG-7 메타데이터가 visual 특징들만으로 구성된 경우와 같이 MPEG-7 메타데이터의 크기가 작은 경우 두 번째 방법이 유용하게 사용될 수 있다. 이 방법은 ES 를 기술하기 위한 ES_Descriptor 내에 존재 하는 User defined Descriptor 의 내부에 MPEG-7 메타데이터를 포함시키는 방법이다. 그림 2 는 ES_Descriptor 의 XMT Schema 를 표현하는 다이어그램으로 DecoderConfigDescriptor 내부에 User defined Descriptor 가 존재 한다.

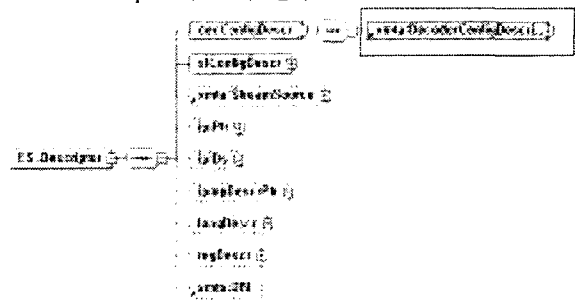


그림 2 MPEG-7 over MPEG-4 의 두 번째 방법

본 논문에서 제시한 방법은 MPEG-4 콘텐츠만으로 다양한 검색 기능을 제공 할 수 있으며, MPEG-4 콘텐츠의 가장 중요한 기능인 Interactivity 도 지원 할 수 있다.

3. Implementation

앞장에서 제시한 두 가지 방법 중에서 시험 시스템을 구성하기 위하여 두 번째 방법인 ES_Descriptor 내의 User defined Descriptor 를 사용하여 개발하였다.

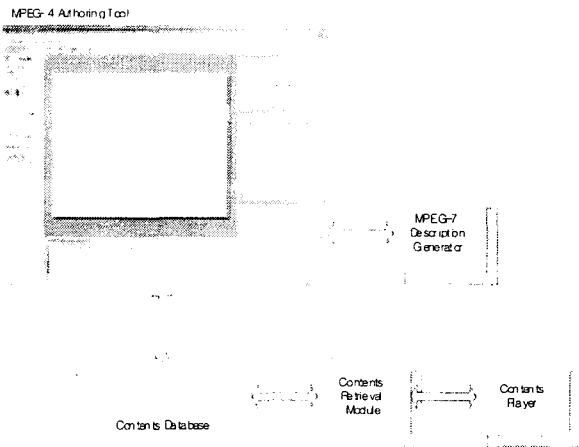


그림 3 MPEG-7 Description 을 이용한 객체 기반 MPEG-4 콘텐츠 검색 시스템 구조도

위의 그림 3 에서 보듯이 개발한 시험 시스템에는 제안한 MPEG-7 Description 을 이용한 MPEG-4 XMT 포맷을 사용하여 MPEG-4 콘텐츠를 저작 할 수 있는 MPEG-4 Authoring Tool, 미디어 데이터에 대한 MPEG-7 메타데이터를 생성할 수 있게 하는 MPEG-7 Description Generator, 저작된 MPEG-4 콘텐츠를 검색 할 수 있게 엔진 역할을 하는 Contents Retrieval Module, 검색을 위한 인터페이스를 제공하거나 검색된 MPEG-4 콘텐츠를 재생할 수 있게 하는 Contents Player 로 구성되어 있다. 자세한 모듈 설명은 다음과 같다.

3.1 MPEG-4 Authoring Tool

이전에 MPEG-4 콘텐츠를 저작하기 위하여 개발한 Interactive Rich-media Authoring Tool[8]을 바탕으로 제안한 MPEG-7 메타데이터를 활용한 MPEG-4 콘텐츠를 저작할 수 있도록 확장하였다. 이 저작 도구는 사용자가 쉽게 MPEG-4 콘텐츠를 저작할 수 있도록 강력한 그래픽 사용자 인터페이스를 제공하며, MPEG-4 시스템 표준에 적합한 XMT 및 MP4 포맷을 생성한다.

3.2 MPEG-7 Description Generator

현재의 구현에서 여러 가지 MPEG-7 메타데이터 중에서 visual 특징인 scalable color 및 homogeneous texture 에 대한 메타데이터를 사용한다. 사용자가 MPEG-4 Authoring Tool 을 사용하여 MPEG-4 콘텐츠를 저작하는 과정에서 미디어 데이터를 사용하면 이에 대한 위의 두 가지 메타데이터를 MPEG-7 Description Generator 를 통해 생성하고, 생성된 MPEG-7 메타데이터를 XMT 로 생성할 수 있도록 저작도구의 해당모듈로 전달하는 역할을 한다.

3.3 Contents Retrieval Module

현재 Contents Retrieval Module 의 경우 Contents Player 의 한 부분으로 설계되어 있다. Contents Player 의 질의 인터페이스를 통해 입력된 사용자 질의를 입력으로 하여 MPEG-4 콘텐츠를 검색하는 역할을 수행한다. Contents Retrieval Module 에는 MPEG-4 콘텐츠 포맷인 MP4 파일에서 해당하는 MPEG-7 메타데이터를 추출해 내는 모듈이 존재하며, 입력된 사용자 질의와의 유사도를 계산하는 모듈도 존재 한다. 이런 모듈들을 통해 Contents Retrieval Module 은 MPEG-4 콘텐츠와 계산된 유사도를 Contents Player 의 해당 모듈에 전달한다. 참고로 0.0 에 가까운 MPEG-4 콘텐츠 일수록 사용자 질의와 가까운 미디어 데이터를 포함하고 있는 것이다.

3.4 Contents Player

Contents Player 는 MPEG-4 콘텐츠를 재생하기 위한 IM-1 2D player 를 바탕으로 구현되었다. MPEG-4 콘텐츠를 검색하기 위한 사용자 인터페이스 및 Contents Retrieval Module 을 통해 검색된 결과를 위한 사용자 인터페이스를 기존의 player 에 첨가 하였다.

4. Result

MPEG-7 메타데이터를 사용하여 MPEG-4 콘텐츠를 검색하기 위해서는 그림 4에 나타난 순서도에 의해 이루어진다.

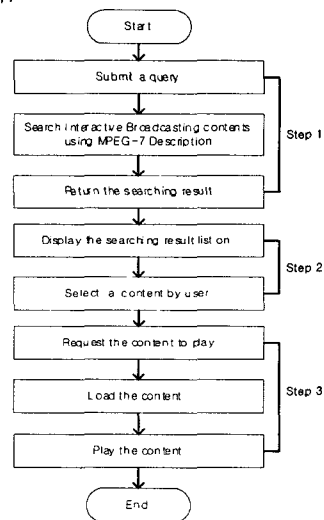


그림 4 객체 기반 MPEG-4 콘텐츠 검색 흐름도

Step 1: Retrieving step

Contents Player 의 사용자 질의 인터페이스를 통해 사용자가 찾기를 원하는 질의를 입력 받는다. 현재 구현에서는 Image 데이터를 입력 받는 query by example 형태를 가지고 있다. 그림 5 에서 보듯이 상단의 사용자 질의 인터페이스에서는 질의로 사용될 Image 데이터를 사용자가 지정할 수 있도록 구성되어 있으며 아래에서 해당 질의에 해당하는 정보를 사용자 에게 보여 준다.

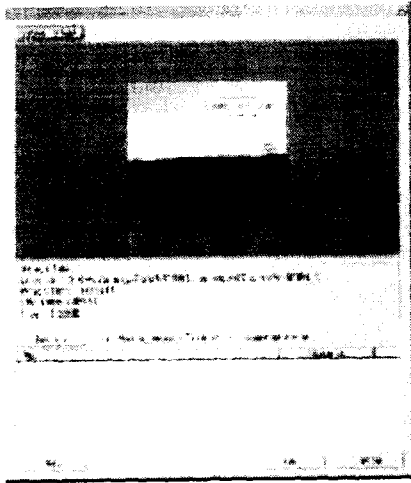


그림 5 사용자 질의 인터페이스

Step 2: Browsing step

그림 6 과 같이 Retrieving step 의 결과가 사용자 질의 인터페이스 하단에 보여진다. 결과는 MPEG-4 콘텐츠와 사용자 질의와의 유사도의 리스트로 보여지는데, 가장 유사한 순서대로 정렬된다.

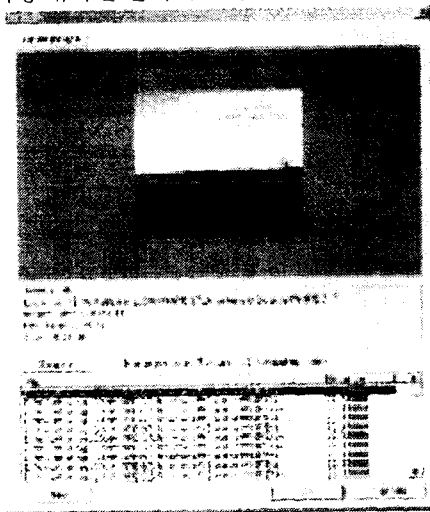


그림 6 검색 결과 인터페이스

Step 3: Playing step

그림 7 과 같이 검색 결과 인터페이스를 통해 사용자가 선택한 MPEG-4 콘텐츠가 Contents Player 를 통해 재생된다.

5. Conclusion

앞에서 언급한 바와 같이 디지털 방송의 인프라 개발이 점점 더 중요해짐에 따라, 디지털 방송 콘텐츠 개발 및 관리 또한 점점 더 중요한 이슈가 되고 있다. 디지털 방송 콘텐츠의 가장 중요한 요구사항인 양방향성을 잘 표현할 수 있는 MPEG-4 가 차세대 디지털 방송의 콘텐츠 표현 방법으로 여겨지고 있다.

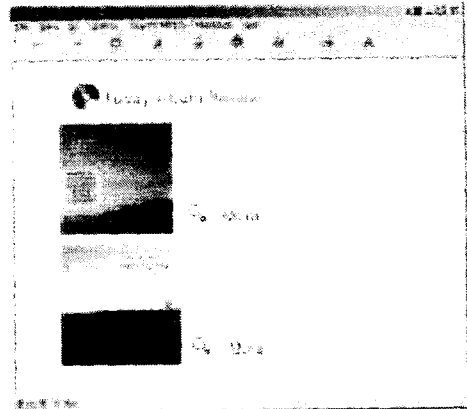


그림 7 콘텐츠 플레이어

이런 MPEG-4 콘텐츠가 널리 사용되면 디지털 방송 서버등과 같은 어플리케이션에서는 MPEG-4 검색과 같은 기능이 필요하게 될 것이다. 이런 경우 본 논문에서 제안한 MPEG-7 메타데이터를 활용한 MPEG-4 콘텐츠 표현 방법 및 검색 시스템을 사용한다면 미디어 데이터 검색에 대해 제약점을 벗어나 다양한 검색 기능을 제공할 수 있을 뿐만 아니라 높은 검색 효율을 제공할 것이다.

Acknowledgement

본 연구는 정보통신부의 “지능형 통합정보방송 (SMARTV) 기술 개발” 사업의 일환으로 수행된 연구 결과입니다.

참고 문헌

- [1] Patrick D. Griffiths, "Interactive TV in a digital era: A standards perspective", Int Conf, Broadcast Asia 2001, Session 3: Interactivity & Enhanced Television, Singapore, June 2001.
- [2] MPEG-4 Systems Group, "ISO/IEC 14496-1:2002 WD 2.0", ISO/IEC JTC1/SC29/WG11 N4848, March 2002.
- [3] MPEG-7 Systems Group, "ISO/IEC 15938-1:2001 FDIS", ISO/IEC JTC1/SC29/WG11, September 2001.
- [4] Aaron E. Walsh, and Mikael Bourges-Sevener, *MPEG-4 Jump-Start*, Prentice Hall PTR, 2002.
- [5] MPEG-7 Systems Group, "MPEG-7 Overview (version 8)", ISO/IEC JTC1/SC29/WG11 N4980, July 2002.
- [6] Miroslaw Bober, Kohtarō Asai, and Ajay Divakaran, "A MPEG-4/7 Internet Video and Still Image Browsing system", Proceedings of SPIE Vol. 4209(2001), pp33-38.
- [7] José M. Martínez, Luis F. Rubio, Francisco Moran, "Authoring 744: first results", ACM Multimedia, 2002.
- [8] Kyuheon Kim, Won-Sik Cheong, Hyun-Cheol Kim, and Jinwoong Kim, "Interactive Broadcasting Contents Authoring System", SK Telecommunication Review Journal, Vol. 11-5, pp688-698, 2001.
- [9] Michelle Kim, Steve Wood, and Lai-Tee Cheok, Extensible MPEG-4 Textual Format (XMT)", ACM Multimedia, 2000.
- [10] YeSun Joung, Kyuheon Kim, "An XMT API for generation of the MPEG-4 scene description", ACM Multimedia, 2002.
- [11] Yong man Ro, Munchul Kim, Ho Kyung Kang, B.S. Manjunath, and Jinwoong Kim, "MPEG-7 Homogeneous Textual Descriptor", ETRI Journal, Volume 23, Number 2, June 2001.
- [12] MPEG-7 Video and MDS subgroups, "Text of ISO/IEC 15938-8 DTR (Extraction and Use of MPEG-7 Descriptions)", ISO/IEC JTC 1/SC 29/WG 11 N4579, March 2002.