

BiM 을 이용한 메타데이터의 효율적인 부호화 방법

양승준, 남제호, 김영태, 강경옥
한국전자통신연구원 방송미디어연구부
대전광역시 유성구 가정동 161

An efficient compression method of metadata using BiM

Yang Seung-Jun , Jeho Nam, Young-Tae Kim, Kyeong-Ok Kang
161 Gajeong-Dong, Yuseong-Gu, Taejon
sjyang@etri.re.kr

요약

ISO/IEC 15938-1(MPEG-7 Systems)에서는 멀티미디어 콘텐츠에 대한 메타데이터의 효율적인 전송과 저장에 대한 이진 표현 방법인 BiM(binary format for MPEG-7)을 제공한다. 멀티미디어 콘텐츠를 기술(description)하는 메타데이터의 텍스트 표현은 대체로 많은 저장 용량과 전송 리소스를 요구하기 때문에 효율적인 압축을 위해서는 이진 형식으로의 변환이 요구된다. 또한 텍스트 형식은 방송 환경과 같은 스트리밍 전송에는 적절하지 못한 단점이 있다.

BiM은 콘텐츠에 대한 기술을 전체 또는 2개 이상의 AU(access units) 단위로 분할하여 부호화 하는 방법을 지원함으로써 스트리밍 전송을 가능하게 한다. 이러한 구조는 이진 포맷 형태로 표현되는 헤더를 가지는 패킷 기반 형태이며, 융통성이 있는 전송 순서를 제공한다. 또한, 비트 스트림의 전체를 해석(parsing)하지 않고 랜덤 액세스 기능을 제공하는 장점이 있다.

BiM이 지닌 이러한 장점들로 인하여 현재 방송 산업계를 중심으로 메타데이터를 방송에 활용하기 위한 기술을 표준화하는 국제 민간 표준화 기구인 TV-Anytime 포럼에서는 방송 콘텐츠에 대한 메타데이터의 압축에 관한 요구사항을 만족하는 하나의 방법으로 BiM을 고려하고 있다. 본 논문에서는 이러한 MPEG-7 시스템의 BiM을 소개하고, 이를 이용하여 TV-Anytime 포럼의 메타데이터를 이진 포맷으로 부호화한 실험과 그 결과를 기술한다.

1. 서론

인터넷의 이용 증가, 디지털 방송 서비스의 도입에 따라 다양한 형태의 정보통신 서비스가 확대되고 디지털 멀티미디어 정보가 급속히 증가하고 있으며, 이에 따라 비디오, 오디오, 음성 및 정지 영상을 포함하는 다양한 형태의 멀티미디어 자료로부터 사용자가 원하는 멀티미디어 정보를 신속하고 효과적으로 찾아내고 그 정보를 사용자가 원하는 형태로 변형, 또는 전송할 수 있도록 하는 기술의 필요성이 대두되었다.

이러한 필요성에 따라 국제 표준화 기구인 ISO/IEC 산하의 MPEG 에서는 오디오비주얼 데이터를 중심으로 한 멀티미디어 데이터의 내용 기반 검색을 위한 MPEG-7 표준화를 진행하고 있다. 지금까지 표준화되었거나 현재 표준화가 진행되고 있는 MPEG-1, 2 및 4 에서는 오디오비주얼 데이터의 압축을 그 목표로 하였으나, MPEG-7 은 데이터 그 자체가 아닌 데이터의 내용에 대한 표현방법을 다루는 것이다. 즉, MPEG-7 은 오디오, 비주얼 그리고 멀티미디어에 대한 기술자(D: descriptor)와 기술 구조(DS: description scheme)의 표준 집합과 기술자 및 기술 구조를 정의하기 위한 정형적인 언어를 포함한다. MPEG-7 표준은 ISO/IEC 15938 로 언급되며, 시스템(Systems), 기술 정의 언어(DDL: Description Definition Language), 비주얼(Visual), 오디오(Audio), 멀티미디어 기술 구조(MDS: multimedia DSs), 참조 소프트웨어(Reference Software), 적합성(Conformance)의 7 개 부분에 대한 표준화가 진행 중이다. MPEG-7 시스템에서는 효율적인 전송/저장에 대한 MPEG-7 기술의 작성, 콘텐츠와 기술간의 동기화, 디코더의 적합성 개발과 같은 시스템 레벨의 기능들에 대해 기술한다. 기술은 텍스트 형식 또는 BiM 으로 불리우는 이진 형식으로 표현될 수 있다.[1][4]

개인용 대용량 저장 매체(local storage)를 갖는 사용자 환경에서 사용자 기반의 방송 서비스 제공을 위한 표준 개발을 목적으로 하는 민간 표준 기구인 TV-Anytime 포럼에서는 표준화 작업을 요구 규격과 기술 표준의 단계로 나누어 진행하고 있으며, 비즈니스 모델(Business Model), 시스템 기술(System Description), 메타데이터(Metadata), 콘텐츠 레퍼런싱(Content Referencing), 저작권 관리 및 보호(Rights Management & Protection)의 5개 작업 그룹(WG)으로 나누어 기술 표준화가 진행 중이다. 메타데이터 작업 그룹에서는 TV-Anytime 메타데이터에 관한 정의가 이루어지고 있다. 진행중인 표준안 1.1의 XML (eXtended Markup Language) 기반의 TV-Anytime 메타데이터 프레임워크에서는 대역폭과 시스템 성능을 최대화 하기 위해 이진화를 허용하고 있으며, MPEG-7 BiM 이 한 방법으로 논의되고 있다.[2][3]

본 논문에서는 TV-Anytime 포럼에서 메타데이터 부호화를 위한 하나의 방법으로 거론되고 있는 BiM의 소개와 TV-Anytime 포럼의 규격에 따른 메타데이터를 BiM을 이용하여 부호화하는 실험 및 결과에 대해 기술한다.

2. MPEG-7 Systems

2.1 MPEG-7 시스템의 기본 구조

멀티미디어 콘텐츠를 기술하는 메타데이터의 텍스트 표현(XML)은 인간이 읽고 처리하기 쉽게 되어있다. XML 기술자는 일반적인 툴(tool)을 사용하여 변환, 탐색, 해석하기 위한 구조화된 데이터를 허용한다. 이러한 일반적인 변환, 질의 인터페이스, 가용성 검사 도구들은 간단하고 안정적인 XML 문서의 조작을 허용한다. 이러한 특성으로 인해 XML은 구조화된 정보를 교환하는 어플리케이션 사이의 상호운영에 적합하며, 어플리케이션 개발을 용이하게 한다. 이러한 장점에 반하여 XML의 텍스트 구조는 몇 가지 단점이 있다. 텍스트 포맷은 이론적으로 필요한 것보다 더 많은 저장 공간과 전송 리소스를 요구한다. 따라서 텍스트 포맷의 이진으로의 변환을 위한 효율적인 압축 방법이 필요하게 된다. 추가로 텍스트 포맷은 스트리밍 어플리케이션에 적절하지 않다. 스트리밍 어플리케이션을 위해서는 엘리먼트의 전송 순서에 대한 고수준의 융통성이 요구된다. 손쉬운 랜덤 액세스는 전체 비트스트림의 해석없이 제공되어야 한다. 이러한 특성은 일반적으로 이진 포맷의 헤더 구조를 가지는 패킷 기반 구조를 요구한다.[5][7]

MPEG-7 시스템의 이진 포맷(BiM)은 그림 1과 같이 콘텐츠에 대한 기술의 요소화된 전송을 허용한다.

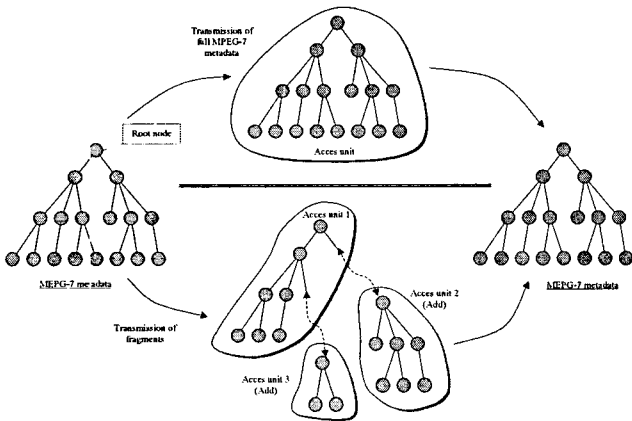


그림 1. MPEG-7 기술과 AU

MPEG-7의 비트스트림은 하나 또는 2개 이상의 연속적인 BiM Unit으로 구성된다. 각 BiM Unit의 구조는 그림 2와 같다.[1][7]

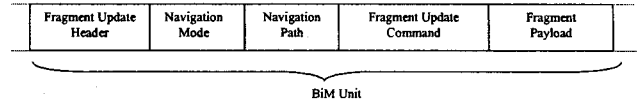


그림 2. BiM Unit의 구조

BiM Unit의 첫번째 부분은 스키마(schema)의 확장에 대한 정보를 포함한 프래그먼트 업데이트 헤더이다. 두번째 부분은 네비게이션 모드 부분으로 3비트의 고정길이를 가지고 절대 혹은 상대 주소 방식을 나타낸다. 세번째 부분은 네비게이션 패스 부분으로 프래그먼트의 절대 혹은 상대 주소를 나타내며, 노드와 자식 노드 사이의 TBC(tree-branch-codes)를 이용하여 주소가 계산된다. 네번째 부분은 프래그먼트 업데이트 커맨드로 갱신에 대한 종류를 지시하며, 추가, 대체, 삭제의 갱신 명령이 제공된다. 다섯번째 부분은 프래그먼트 payload 부분으로 이진 부호화된 XML 서브 트리를 가지며 프래그먼트 업데이트 커맨드에 의해 조작된다.[1][5][7]

2.2 MPEG-7 시스템 이진 포맷의 특성

BiM은 텍스트와 이진 XML 포맷 사이의 일대일 매핑을 제공한다. 부호화는 텍스트 XML 파일을 해석하며 유효성 검사를 하고 이진 파일을 생성한다. XML 전송과 관련된 일련의 단계에서, 기술은 원래의 텍스트 포맷으로 변환될 수 있고, XML 툴을 사용하여 처리될 수 있으며, 이진 포맷으로 부호화될 수 있다. BiM은 어느 특정 XML 언어에 종속되어 있지 않으며, 부호화와 복호화는 스키마 정의가 주어진 어떠한 XML 언어도 다룰 수 있다. 더욱이 참조 소프트웨어는 SAX(The Simple API for XML) API를 사용함으로써 XML 환경에 쉽게 적용될 수 있다. BiM은 구조화된 정보를 부호화하기에 매우 적합하다. 기술에 대해 고수준의 구조화가 이루어지고 스키마가 효율적으로 제한된 상황에서는 엘리먼트당 1비트 이하로 XML 문서의 구조를 부호화할 수 있다. 더욱이, BiM은 데이터를 부호화하기에 확장성 있는 프레임워크를 제공한다.[7]

BiM을 사용하면 전체 기술 트리를 하나의 AU(Access unit)로 전송하거나, 서브 트리로 나누어 2개 이상의 AU로 전송하는 것을 선택할 수 있다. 또한, 전송측에서 서브 트리의 전송 순서를 융통적으로 선택할 수 있다. 이러한 AU의 특성을 이용하면 유연한 갱신과 전송이 가능하다. 즉, 부호화기에서 기술 부분이 변화된 경우에 전체 기술의 재전송 없이 변화된 부분만 전송이 가능하다.[1]

또한, BiM은 표준에 따른 모든 버전이 호환할 수 있는 특성을 제공한다. 수신된 스키마가 가용한 경우, 그림 3과 같은 완전한 전방향/후방향 호환성을 제공하고, 네임스페이스(namespace)가 모든 엘리먼트와 속성에 적용될 수 있다. 이때 스키마 갱신으로 인해 수신된 스키마가 가용하지 않는 경우에 문제가 야기될 수 있다. 이러한 상태를 다루기 위해 BiM은 복호화기가 확장된 스트림을 부분적으로 복호화하는 것을 허용한다.[1][6][7]

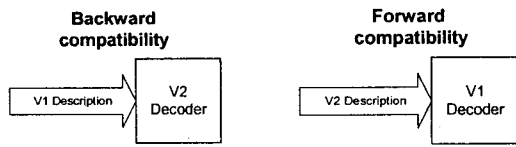


그림 3. 전방향/후방향 호환성

v: 버전

효율적인 압축률과 일대일 매핑의 측면에서 BiM은 추가적인 특성을 가지고 있다. BiM은 이진화된 XML 기술을 효과적으로 처리할 수 있는 API를 제공한다. 적절한 부호화가 이루어졌다면, BiM 복호화는 전체 기술 엘리먼트의 복호화 과정을 생략(skip)할 수 있다. 이러한 스킵핑(skiping) 처리는 엘리먼트 이름, 형(type) 또는 속성 값을 기초로 하여 선택적으로 수행될 수 있다. 스킵핑 특성은 요구되지 않는 기술 부분의 매우 빠른 복호화와 메타데이터 구조를 통한 브라우징과 탐색 속도를 향상 시키는데 이용될 수 있다. 더욱이 BiM 플래그먼트는 비트 단위의 패턴 비교를 통해 매우 빠르게 검출될 수 있다. 이러한 필터링 방법은 원하는 AU를 빠르게 선택할 수 있게 한다.[1][7]

결국, BiM은 텍스트 포맷에 비해 높은 압축효율과 XML 텍스트와 이진 포맷간의 양방향 매핑, XML 기반의 언어에 대한 적합, 융통성 있는 전송과 메타데이터의 갱신, 전송 순서의 랜덤성, 스키마의 확장성, 빠른 액세스와 필터링 능력, 비트스트림 레벨에서의 파싱과 필터링등의 기능을 제공함으로써 TV-Anytime 포럼에서 규정한 인코딩 기법, 탐색과 필터링, versioning, 갱신, 확장성등에 관한 요구사항을 만족시킨다. 결론적으로 BiM은 MPEG-7 기술의 부호화를 위해 설계되었지만 TV-Anytime 메타데이터에도 적합하다.

3. TV-Anytime metadata framework

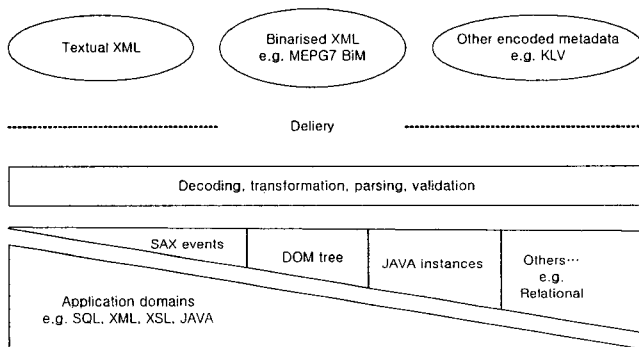


그림 4. XML 기반의 TV-Anytime 메타데이터 프레임워크

그림 4은 TV-Anytime 터미널에서 다양한 XML 포맷 또는 non-XML 포맷으로 인코딩된 메타데이터를 입력받는 구조를 보여준다.[2]

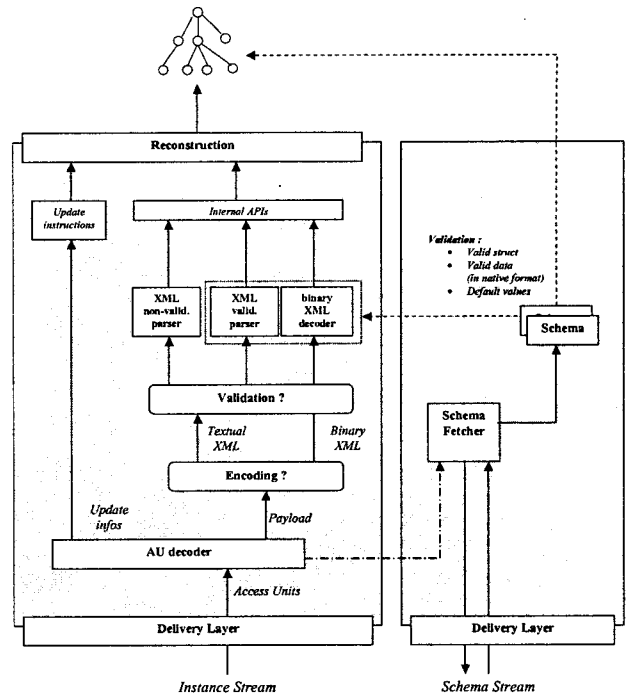


그림 5. 메타데이터 처리 구조

그림 5는 TV-Anytime 포럼의 메타데이터 처리 구조를 나타낸다. 인스턴스 메타데이터 과정은 전송 계층에서 메타데이터 인스턴스 스트림을 입력받아 먼저 AU를 AU 디코더로 출력하며, AU 디코더는 이를 갱신 정보와 payload 부분으로 분리한다. payload 부분은 인코딩 여부의 검사를 거쳐 텍스트 포맷인 경우 유효성(validation) 검사를 할 것인지를 결정한다. TV-Anytime 포럼에서 다루는 메타데이터가 매우 크고 고수준의 구조화된 메타데이터인 경우, 유용성 검사는 중요한 부분이다. 인터넷과 같은 불안정한 환경에서는 모든 기술에 대해서 완전한 유효성 검사를 거치는 것이 가장 좋은 방법이나, 이 경우 터미널 리소스의 소비가 크게 된다. 전송 환경이 안정적인 경우, 전송된 기술은 항상 유용할 것이며, 복호화는 단지 기술을 해석하고 엘리먼트와 속성에 기본 값과 형을 할당하면 된다. 이진 XML 디코더에서 출력된 결과는 내부 API를 통하여 문서에 접근될 수 있고, 재조립 단계를 거쳐 본래의 텍스트 형식의 메타데이터로 복원된다.[2]

4. 실험 및 결과

실험에 사용된 XML 파일은 TV-Anytime 포럼의 표준안 1.1. 버전에 따라 작성되었고, MPEG-7 시스템의 참조 소프트웨어를 이용하여 부호화 및 복호화를 수행하였다.

사용된 예제 파일의 ContentDes.xml은 TV-Anytime의 ContentDescription 스키마 구조를 기반으로 작성되었으며, UserPreference.xml에서는 UserDe-

scription 스키마 구조를 이용하여 작성하였다. ContentUser.xml 은 이러한 두가지 구조를 혼합한 형태이다. 각 파일의 크기는 파일의 크기에 따른 압축률 변화를 살펴보기 위해 약 6K, 10K, 14K 로 크기를 제작하여 실험하였다.

부호화 실험에 대한 결과는 표 1에서 확인할 수 있다. XML 파일의 압축 실험을 ZIP 과 비교한 결과, BiM 의 압축률이 우수하게 측정되었다. ZIP 은 어느 정도 일정 비율의 압축률을 가지는데 비해, BiM 은 압축률의 차이가 심하게 나타났다. 이러한 원인은 BiM 참조 소프트웨어가 부호화 과정에서 주석문을 제거시키며, 네임스페이스의 접두어를 간략한 기호로 변환시킴으로써 내용에 따라 부호화된 비율이 다르게 나타나는 것이다. 이에 따라 복호화된 파일의 크기가 원본에 비해 감소하게 된다.

표 2에서는 부호화 조건을 동일하게 하기 위해, 예제 파일내에 주석문을 모두 제거하고 실험을 반복하였다. 동일한 조건하에서의 실험에서도 BiM 이 ZIP 에 비해 우수한 성능을 나타내었다.

부호화 과정을 수행하고 복원된 XML 파일의 크기는 감소하였으나, 내용은 구조적으로 원본과 동일하며, 엘리먼트와 속성값도 모두 원본과 동일하게 보존되어 있음을 확인하였다. 따라서 복호화된 텍스트 포맷과 부호화 이전의 XML 원본 사이에 규범적인 동가성(canonical equivalence)이 유지됨을 확인할 수 있었다.

5. 결론

본 논문에서는 MPEG-7 시스템의 BiM 을 이용한 TV-Anytime 포럼의 메타데이터를 부호화하는 실험과 그 결과를 기술하였다. TV-Anytime 포럼의 스키마 규격에 따라 컨텐츠에 대한 메타데이터를 이진 포맷으로 부호화하였으며, 복호화된 텍스트 포맷(XML)이 부호화 이전의 XML 원본과 규범적인 동가성을 유지함을 확인할 수 있었다.

BiM 은 우수한 부호화 효율과 더불어 이진 상태에서 필터링과 갱신이 가능하게 설계 되어있기 때문에 XML 기반의 메타데이터의 부호화에 유용한 기법으로 활용될 것이 예상된다. 향후에는 이러한 이진 상태에서의 탐색과 갱신에 대한 연구가 진행할 예정이다.

참고 문헌

- [1] "Text of ISO/IEC FCD 15938-1 Information Technology- Multimedia content description interface-Part 1 Systems", ISO/IEC JTC1/SC29/WG11 N4001, Mar 2001.
- [2] "Provisional SP003V1.2", MD WG WD415, Jul 2001.
- [3] "Specification Series: S-3 on Metadata", MD WG, SP003v1.1, Aug 2001.
- [4] Shih-Fu, Thomas Sikora, Atul Puri, "Overview of the MPEG-7 Standard" IEEE Trans. Circuits and systems for video technology, vol. 11, pp. 688-695, Jun 2001.
- [5] Olivier Avaro, Philippe Salembier, "MPEG-7 System : Overview", IEEE Trans. Circuits and systems for video technology, vol. 11, pp. 760-764, Jun 2001.
- [6] Jane hunter, "An Overview of the MPEG-7 Description Definition Language", IEEE Trans. Circuits and systems for video technology, vol. 11, pp. 765-772, Jun 2001.
- [7] Cedric Thienot, Claude Seyrat, "A streamable XML binary encoding for TV Anytime metadata", MD WG, AN282, Jun 2001.

감사의 글

본 논문은 정보통신부의 "MPEG-7 기반 메타데이터 방송 기술 개발 사업" 과제의 일환으로 수행된 연구 결과이며, 도움을 주신 대화형방송연구팀원들께 감사드립니다.

표 1. 원본 파일의 실험 결과

(단위 : byte)

실험 파일		부호화된 크기		복호화된 크기	압축률	
파일	크기	ZIP	BiM	BiM	ZIP	BiM
ContentDes.xml	5,814	1,928	1,517	3,919	3.01	3.83
UserPreference.xml	9,310	1,856	1,057	4,758	5.01	8.80
ContentUser	13,291	2,697	1,972	7,748	4.92	6.73

표 2. 주석문을 제거한 실험 결과

실험 파일		부호화된 크기		복호화된크기	압축률	
파일	크기	ZIP	BiM	BiM	ZIP	BiM
ContentDes.xml	4,713	1,743	1,517	3,920	2.70	3.10
UserPreference.xml	5,960	1,399	1,057	4,759	4.26	5.64
ContentUser	9,653	2,226	1,972	7,749	4.34	4.90