

효율적인 비디오 브라우징 및 검색을 위한 통합 멀티미디어 응용 형식

조준호, 진성호, 양승지, 노용만

한국정보통신대학교 영상시스템 연구실

e-mail : chojunho@icu.ac.kr

Integrated Multimedia Application Format for Active Video Browsing and Retrieval

Jun Ho Cho, Sung Ho Jin, Seungji Yang, Yong Man Ro

Image and Video Systems Lab, Information and Communications University

요 약

본 논문에서는 MPEG의 멀티미디어 응용 표준인 MAF(Multimedia Application Format)를 기반으로, 효율적인 비디오 콘텐츠의 검색 및 활용을 위한 통합 미디어 구조, 즉 비디오 MAF를 제안한다. 제안하는 비디오 MAF는 ISO 미디어 포맷을 기반으로 하고 단일의 비주얼 스트림과 다중 음성을 지원하기 위한 다수의 오디오 스트림, 내용기반의 정보를 포함하는 메타데이터, 그리고 비디오 콘텐츠의 대표 이미지를 동시에 포함하는 구조이다. 제안하는 파일포맷의 유용성을 검증하기 위해 비디오 MAF로 생성 및 해석할 수 있는 부호기(encoder)와 복호기(decoder)를 설계하고 구현하여, 통합 미디어에 내재된 메타데이터를 이용한 효율적인 검색과 멀티트랙의 오디오 스트림을 활용한 다중 음성에 대한 지원이 가능함을 확인하였다. 또한 내재된 대표이미지는 비디오 콘텐츠에 대한 브라우징이 효과적으로 활용됨을 확인하였다.

1. 서론

최근 PMP(Portable Multimedia Player)와 같은 휴대용 멀티미디어 단말의 보급화와 저장할 수 있는 용량의 증대로 인해 사용자들은 다수의 미디어 콘텐츠를 시간과 장소에 구애 없이 소비 가능하게 되었다. 그러나 저장된 콘텐츠의 양의 증대와 함께 저장매체로부터 콘텐츠를 효율적으로 접근 및 관리하기 위한 연구의 필요성이 대두되고 있다. 이에 대한 해결책으로 비디오 콘텐츠와 내용기반의 정보들을 포함하는 메타데이터를 단일의 통합 미디어로서 활용하는 방법이 고려될 수 있다.

현재, MPEG-A(MPEG-Application: ISO/IEC 23000)에서는 이러한 통합 미디어에 대한 표준화 활동이 진행 중이다. 새롭게 표준화가 진행중인 멀티미디어 응용 형식(Multimedia Application Format, MAF)은 MPEG-1,2,4,7,21과 같은 MPEG의 표준들뿐만 아니라, JPEG과 같은 비 MPEG 표준들을 포함하여 상호 호환성 있도록 조합하는 방법을 제공한다. 표준화가 완성된

MAF를 활용하여 특정 어플리케이션이나 응용 서비스의 개발을 위해 필요한 표준 요소들을 선택하고, 적절하게 조합하는 노력을 줄임으로써 급변하는 산업의 요구에 빠르게 대응하고 MPEG 표준의 활용가치를 증가시키는 효과를 가져 올 수 있다[1].

본 논문에서는 효율적인 비디오 브라우징과 검색을 위해, 비주얼 스트림과 다중의 오디오 스트림, 비디오 콘텐츠의 대표 이미지, 내용기반의 정보인 MPEG-7 MDS 메타데이터를 동시에 포함하는 비디오 MAF의 구조를 제안하고, 이를 생성 및 해석할 수 있는 부호기 및 비디오 MAF 기반 검색 시스템을 개발한다.

본 논문의 구성은 다음과 같다. 2장에서 관련연구에 대해서 서술하고, 3장에서 비디오 MAF의 구성요소, 4장에서 부호화와 복호화 과정을 설명한다. 5장에서는 부호기와 복호기의 구현을 통해 통합 미디어의 유용성을 입증하고, 마지막으로 6장에서는 결론 및 향후 연구 과제를 서술한다.

2. 관련연구

MPEG-A 의 하부 그룹인 3on4(Music Player Application Format: ISO/IEC 23000-2)는 기존의 MP3 와 MPEG-7 MDS(Multimedia Description Scheme)의 내용기반 정보, 그리고 음악과 관련된 JPEG 영상을 추가하여 하나의 통합미디어를 구성한다. 이러한 통합미디어는 음악을 듣는 동시에, 관련된 정보를 획득하거나 앨범 재킷과 같은 이미지를 볼 수 있는 미디어로서 최근 일차적인 표준화가 완료되었다[2].

Jon4(Photo MAF Player: ISO/IEC 23000-3)는 내재된 메타데이터를 활용하여 사진들의 그룹화 또는 유사한 사진의 검색을 가능하게 하기 위해 JPEG 데이터에 MPEG-7 Visual, MDS 등의 메타데이터를 통합된 미디어로 구성하는 표준화로서 최근 작업표준안(Working Draft, WD)이 작성된 단계에 있다[3].

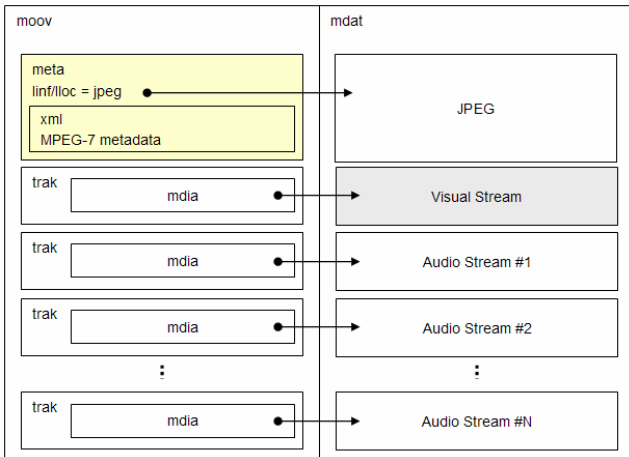
표준화가 시작 단계인 비디오 MAF 는 다양한 멀티미디어 단말간에 비디오 코딩 및 저작권 보호 관리와 관련된 구성요소들의 상호 호환성을 위한 통합미디어를 정의한다. 비디오 MAF 의 통합미디어는 콘텐츠에 대한 내용기반 정보인 메타데이터와 단일 비주얼 스트림에 대한 다중 언어를 지원하는 다수의 오디오 스트림, 그리고 비디오 콘텐츠를 대표하는 이미지 등의 구성요소들로 구성된다[4].

본 논문에서는 비디오 MAF 에 요구되는 구성요소들을 포함하는 통합 미디어의 구조를 제안한다.

3. 통합미디어의 구조 및 구성요소

3.1 통합미디어의 구조

그림 1 은 비디오 MAF 를 위한 ISO 미디어 파일포맷을 기반으로 하는 통합미디어의 구조를 나타낸다. ISO 미디어 파일포맷은 미디어 리소스를 저장하기 위한 미디어 데이터 아톰(Media Data Atom : mdat Atom)과 이를 해당 미디어 리소스의 위치와 이를 재생하기 위한 재생 정보 등을 저장하는 무브 아톰(moov Atom) 등의 객체 단위로 분리하여 저장하는 특징이 있으며 MPEG-4, MPEG-21, MJ2K 파일 포맷 등으로 확장될 수 있다[5][6].



(그림 1) 제안하는 통합 미디어 구조

본 논문에서는 ISO 미디어 파일 포맷의 확장으로서 MPEG-4 파일 포맷을 활용하였다. 비주얼 스트림을 위한 단일 트랙과 다수의 오디오 스트림을 위한 다중 트랙으로 구성되고, 각각의 트랙 내에는 미디어 아톰(Media Atom)이 존재하여 미디어 데이터 아톰내에 저장된 트랙이 가리키는 해당 비디오/오디오 스트림의 접근 및 재생을 위한 정보를 저장할 수 있다. 무브 아톰에 존재하는 메타 아톰(meta Atom)은 아이템 식별자(Item ID), 아이템의 오프셋/길이(offset/length) 정보를 포함하여 미디어 데이터 아톰에 저장된 JPEG 데이터에 접근 가능하게 한다. 또한, 비디오 콘텐츠의 내용기반 정보를 서술하기 위한 MPEG-7 메타데이터는 XML 텍스트 또는 바이너리 형식으로 xml Atom 내에 저장한다.

3.2 통합미디어의 구성요소

통합미디어의 구성요소로서 단일 비주얼 스트림과 각각 한국어와 영어로 구성된 다중 오디오 스트림을 포함한다. 또한, 비디오 콘텐츠의 대표 이미지로서 JPEG 데이터와 내용기반 정보로서 MPEG-7 MDS 메타데이터를 활용한다.

그림 2 는 MPEG-7 MDS 메타데이터[7]의 서술구조에 유효하게 저작한 XML 의 예이다. 미디어 정보 및 생성 서술구조(Media Information/Creation DS)는 비디오 콘텐츠에 대한 타이틀(Title), 장르(Genre), 생성자(Creator), 생성된 날짜(Date), 재생시간(Media Time), 통합미디어가 지원하는 음성 언어(Language)등의 정보를 포함하고, 내용기반 정보를 서술하기 위한 의미적 서술구조(Semantic DS)는 인물 정보 서술구조(Person DS)를 포함하여 등장 인물에 대한 배역이름, 실제이름, 인물에 대한 부가정보 등을 서술할 수 있도록 되어있다.

```
<?xml version="1.0" encoding="UTF-8" ?>
<Mpeg7 xmlns="urn:mpeg:mpeg7:schema:2001" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="urn:mpeg:mpeg7:schema:2001 ./mds-2001-1.xsd">
  <Description xsi:type="ContentEntityType">
    <MultimediaContent xsi:type="VideoType">
      <Video>
        <CreationInformation>
          <Creation>
            <Title>위기의 주부들</Title>
            <Title type="opusNumber">episode 1</Title>
            <Title type="episodeTitle">자살</Title>
          </Creation>
          <Abstract>
            <FreeTextAnnotation xml:lang="ko">4명의 주부들의 일상과 그에 대한 열망, 살인사건의 비밀을 풀
            여가는 마스터교역 드라마</FreeTextAnnotation>
          </Abstract>
          <Creator>ABC</Creator>
          <CreationCoordinates>
            <Date>
              <TimePoint>2004-08-18</TimePoint>
            </Date>
          </CreationCoordinates>
          <Classification>
            <Genre xsi:type="main">Drama</Genre>
            <Language xsi:type="original">English</Language>
            <Language xsi:type="dubbed">Korean</Language>
          </Classification>
          <CreationInformation>
            <TextAnnotation>
              <FreeTextAnnotation xml:lang="ko">황폐하고 부유한 교외 주택가, 평범한 주부미녀 매리 뿐이 권총 자살을
              하면서 그녀의 가족과 이웃의 친구들은 미움을 품기 시작한다... 미생물이 매력적인 성격 미혼녀 수진, 불
              감적으로서 총살이나 일상의 무류함에 17세 정원사에 배움을 보이는 가브리엘, 사회적 성공을 뒤로하고 네
              아들 뒤처닥거리며 허우허우 지저귀는 권석, 마사 스튜디오를 지향하는 양복주의 성격. 또한 가족들로부터 멀
              어지는 브리 등 네 주부와 그들의 가족, 애인, 이웃이 얽힌 이야기가 시작된다.</FreeTextAnnotation>
            </TextAnnotation>
            <MediaTime>
              <MediaTimePoint>T00:00:00</MediaTimePoint>
              <MediaDuration>PT50M30S</MediaDuration>
            </MediaTime>
          </Video>
        </MultimediaContent>
      </Description>
    </Description xsi:type="SemanticDescriptionType">
```

(a)

```

<Description xsi:type="SemanticDescriptionType">
  - <Semantics>
    + <Label>
      - <SemanticBase xsi:type="AgentObjectType">
        + <Label>
          - <Definition>
            <FreeTextAnnotation>Housewives</FreeTextAnnotation>
          </Definition>
          - <Object xsi:type="AgentObjectType" id="Person1">
            - <Label>
              <Name>Bree Van De Kamp</Name>
            </Label>
            - <Definition>
              <FreeTextAnnotation>워스타리아 가 주된 전체가 만장일치로 인정하는 완벽한 주부이자 엄마. 하지만 지
              니치가 끊임된 성격 때문에 경력 가족과는 불화가 심하다. 미혼을 요구하는 남편과 변동하지 못한 아들
              때문에 마음고생을 심하게 겪는다.</FreeTextAnnotation>
            </Definition>
          </Object>
          - <Agent xsi:type="PersonType">
            - <Name>
              <GivenName>Marcia</GivenName>
              <FamilyName>Cross</FamilyName>
            </Name>
            </Agent>
          </Object>
          + <Object xsi:type="AgentObjectType" id="Person2">
          + <Object xsi:type="AgentObjectType" id="Person3">
          + <Object xsi:type="AgentObjectType" id="Person4">
          + <Object xsi:type="AgentObjectType" id="Person5">
        </SemanticBase>
      </Semantics>
    </Description>
  </Mpeg7>
  
```

(b)

(그림 2) MPEG-7 MDS 메타데이터의 예

(a)미디어 정보 및 생성 서술구조(Media Information/Creation DS), (b)의미적 서술구조(Semantic DS)

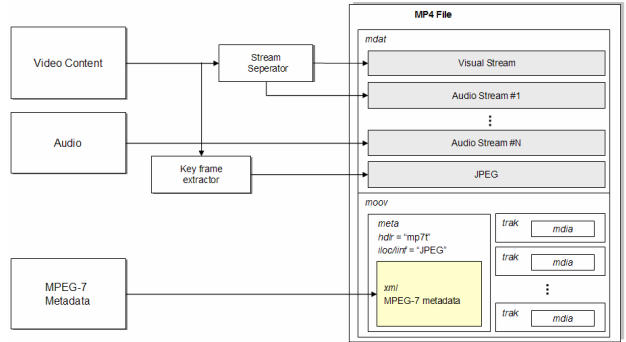
4. 통합미디어의 부호화 및 복호화

4.1 부호화(Encoding)

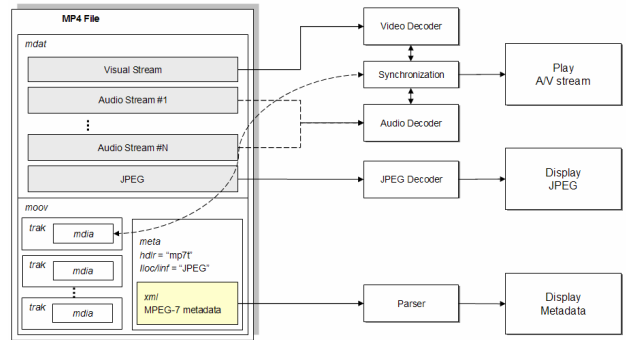
통합미디어를 생성하기 위한 부호화 과정은 그림 3 (a)와 같다. 입력되는 비디오 콘텐츠로부터 키프레임을 추출하여 이미지를 획득한다. 비디오 콘텐츠는 비주얼 스트림과 오디오 스트림으로 분리되어 미디어 데이터 아톰내에 저장되고 해당하는 스트림에 대한 재생정보는 트랙내에 미디어 아톰에 저장된다. 추출된 이미지는 미디어 데이터 아톰에 저장되고 이미지에 대한 재생 정보는 아이템 위치 및 아이템 정보 아톰 (*iloc/iinf Atom*)내에 오프셋/길이, 다수의 이미지를 저장할 때 식별가능 하도록 하기 위한 아이템 식별자, 해당 콘텐츠의 타입 정보(*image/jpeg*)등이 저장된다. XML 형태의 MPEG-7 MDS 메타데이터는 메타 아톰내에 *xml Atom*에 저장된다.

4.2 복호화(Decoding)

생성된 통합미디어는 그림 3 (b)와 같이 복호화 된다. 비주얼 스트림과 사용자에게 의해 선택되는 오디오 스트림은 각각의 트랙내에 존재하는 미디어 아톰내에 샘플 테이블 아톰(*Sample table Atom : stbl Atom*)으로부터 획득되는 정보에 의해 동기화가 이루어져서 재생된다. 아이템 정보 아톰으로부터 해당 JPEG 데이터의 아이템 식별자와 콘텐츠 타입 정보를 획득하고 아이템 위치정보 아톰으로부터 오프셋/길이 정보가 가리키는 미디어 데이터 아톰 내에 데이터를 추출하여 JPEG 디코더를 통해서 이미지를 브라우징한다. *xml Atom* 내에 존재하는 MPEG-7 MDS 메타데이터는 *hdlr Atom*로부터 MPEG-7 임을 가리키는 "mp7t" 정보를 획득한 후, XML 파서에 의해 내용기반의 정보들이 추출되어 사용자의 질의와 비교하여 검색 가능하게 한다.



(a)



(b)

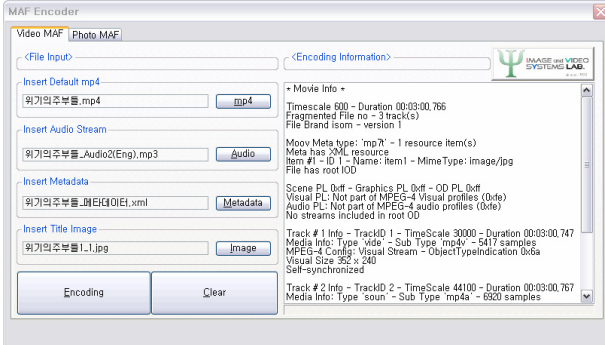
(그림 3) 통합미디어의 부호화 및 복호화 (a)부호화(Encoding), (b)복호화(Decoding)

5. 구현 및 고찰

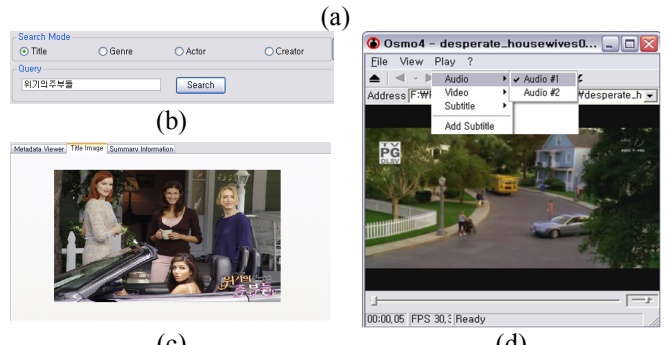
5.1 부호기

그림 4는 통합미디어를 생성하기 위해 구현된 비디오 MAF 부호기로부터 부호화된 결과를 나타낸다. 기본적으로 각각 단일의 비주얼 스트림과 오디오 스트림(한국어)으로 구성되어 있는 MPEG-4 파일과 추가하고자 하는 오디오 스트림(영어)과 XML 파일의 MPEG-7 MDS 메타데이터, 비디오 콘텐츠의 대표 이미지로서 JPEG 파일을 입력한다. 본 연구에서는 비디오 콘텐츠로부터 키프레임은 추출되었다고 가정하고 별도로 입력한다. 부호화된 후에는 우측의 화면으로부터 생성된 통합미디어의 부호화 정보를 통해 MPEG-4 파일 포맷에 유효한지의 여부를 확인할 수 있다.

통합미디어의 구성요소들을 입력하고 부호화한 결과, 각각의 구성요소들이 MPEG-4 파일 포맷에 유효하게 저장되었음을 확인하였고, 영어와 한국어로 이루어진 각각의 오디오 스트림은 미디어 아톰 내의 샘플 테이블 아톰 정보로부터 비주얼 스트림과의 동기화가 되고 있음을 확인할 수 있었다. XML 을 포함하고 JPEG 데이터의 위치정보를 저장하는 메타 아톰은 무브 아톰레벨에 존재하여 설계된 통합미디어의 구조에 따라 생성되었음을 확인하였다. 비디오 MAF 부호기는 GPAC Project 에서 제공하는 MPEG-4 Authoring tool(ver. GPAC v0.2.4)을 기반으로 구현되었다[8].



(그림 4) 비디오 MAF 부호기의 인터페이스



(그림 5) 비디오 MAF 검색 시스템

(a) 검색시스템 인터페이스, (b) 검색모드, (c) 통합미디어에 내재된 JPEG 데이터 디스플레이, (d) Osmo4 MPEG-4 플레이어

5.2 비디오 MAF 검색 시스템

그림 5 (a)는 생성된 통합미디어를 복호화하고 이를 활용하기 위해 구현된 비디오 MAF 기반 검색 시스템을 나타낸다. 좌측에는 저장된 통합미디어의 리스트가 표시되고 그림 5 (b)와 같이 비디오 콘텐츠의 타이틀 (Title), 장르(Genre), 등장인물(Actor), 생성자(Creator) 정보에 의해 저장된 통합미디어들에 대한 사용자의 질의가 가능하도록 설계되었다. 사용자가 원하는 질의를 입력하면 검색 시스템은 통합미디어에 내재되어 있는 메타데이터를 파싱하여 질의와의 일치 여부를 판단하고 검색 결과 리스트에 나타낸다. 검색 결과 리스트에 존재하는 통합미디어를 선택하면 우측 하단부에 내재된 메타데이터와 JPEG 데이터가 그림 5 (c)와 같이 디스플레이 되어 비디오 콘텐츠를 재생하지 않고도 충분한 해당하는 정보를 획득할 수 있다.

메타데이터와 JPEG 데이터를 활용하여 비디오 콘텐츠에 관한 사전적인 정보를 획득한 후 범용적인 MPEG-4 플레이어(Osmo4)를 통해서 통합미디어를 재생할 때, 통합미디어 내의 다수의 오디오 스트림을 활용할 수 있음을 그림 5(d)와 같이 확인하였다. 재생 도중 오디오 스트림을 변경하였을 때 이미 각각의 오디오 트랙의 미디어 아톰내에 샘플 테이블 아톰에 의해 비주얼 스트림과의 동기화 정보를 획득하여 변경 후의 오디오 스트림도 비주얼 스트림과 동기화 되어 재생되는 것을 확인할 수 있었다. 이는 비디오 콘텐츠에 대해서 다중음성 지원을 가능하게 하며 어학교육에 활용될 수 있을 것으로 예상된다.

6. 결론 및 향후 과제

본 논문에서는 비디오 MAF 에 대한 연구로서 비디오 스트림과 다중트랙의 오디오 스트림, 내용기반의 정보를 서술하는 메타데이터, 그리고 관련된 대표 이미지를 동시에 포함하는 통합 미디어의 구조를 제안하였다.

제안된 통합미디어를 생성하는 부호기와 이를 해석하는 복호기를 구현하여 효과적으로 통합미디어에 내재된 MPEG-7 MDS 메타데이터에 의한 효율적인 검색과 JPEG 데이터에 의한 비디오 콘텐츠에 대한 브라우징이 가능함을 확인하였다.

향후 과제로서 비디오 MAF 에 내재되는 콘텐츠는 음악 MAF 와 같이 콘텐츠에 대한 권리보호가 필요시 될 것이므로 현재 구성된 비디오 MAF 에 대한 추가 구성요소를 연구할 것이다. 또한, 다수의 비주얼 스트림을 포함하여 다수의 비디오 콘텐츠들을 의미 있는 정보에 의해 앨범화하여 통합미디어로 구성함으로써 휴대용 단말간의 공유 및 전송이 용이하게 할 수 있는 추가 연구를 수행할 것이다.

참고문헌

- [1] ISO/IEC JTC1/SC29/WG11, White Paper on MPEG-A(N7068), Busan, Korea, April 2005
- [2] ISO/IEC JTC1/SC29/WG11, MPEG Music Player Application Format FDIS(N7156), Busan, Korea, April 2005.
- [3] ISO/IEC JTC1/SC29/WG11, Photo Player Multimedia Application Format(N7093), Busan, Korea, April 2005.
- [4] ISO/IEC JTC1/SC29/WG11, MAFs under Consideration (N7080), Poznan, July 2005
- [5] ISO/IEC 14496-14, MPEG-4 File Format
- [6] ISO/IEC 14496-12, ISO File Format
- [7] ISO/IEC 15938 3:2002/Amd.1ISO/IEC TR 15938-8:2002 Information technology -Multimedia content description interface - Part 8: Extraction and use of MPEG-7 descriptions
- [8] GPAC Project – <http://gpac.sourceforge.net/index.php>

