

TV-Anytime Presentation Engine을 통합한 저작 도구의 개발

이형진, 이선임, 이종설*, 조위덕*, 문영식
한양대학교 컴퓨터공학과, * 전자부품연구원
전화 : 031-407-8991 , 팩스 : 031-419-1162

Implementation of authoring tool integrating TV-Anytime Presentation Engine

Hyung Jin Lee, Sun Im Lee, Jong Sul Lee, We Duke Cho, Young Shik Moon
Dept. of Computer Science Engineering, Hanyang University
E-mail : hjlee@cse.hanyang.ac.kr

Abstract

In this paper, we describe an implementation of indexing and presentation system conforming to TV-Anytime standards. This system has a metadata indexing scheme for multimedia contents, a video indexing algorithm for news and soccer videos, and an MPEG-2 TS player for nonlinear browsing. We also implement an authoring tool for the algorithms mentioned above.

I. 서론

2001년 하반기부터 시작된 국내 디지털 지상파 TV 본 방송으로 이제 본격적인 디지털 방송 시대가 열리게 되었다. 디지털 방송의 등장과 함께 지상파, 케이블, 위성 매체를 통한 다 매체 다 채널 방송 시청환경으로의 변화는 폭발적인 방송 프로그램의 증가를 가져오고 있으며, 시청자에게 다양한 방송 프로그램의 선택과 시청자 중심의 맞춤형 방송 서비스의 기회를 제공하는 새로운 패러다임인 방송 서비스 환경의 시대가 열릴 것으로 예상된다. 이러한 맞춤형 방송 서비스는 사용자가 원하는 방송프로그램을 자동으로 저장해서 사용자가 원하는 시간에 볼 수 있게 하는 서비스를 말하며, 사용자 취향에 기반한 방송 프로그램들의 선택적 저장, 자동 추천, 대화형 광고 및 방대한 콘텐츠에서의 비 선형 브라우징, 효과적인 네비게이션 등의 기능을 필요로 하고 있다. 또한, 로컬 저장매체를 갖는 PDR(Personal Digital Recorder)과 같은 셋탑 박스와, 사용자의 취향(preference) 및 요구(needs)에 맞는 다양한 휴대형 미디어 단말기들을 통한 방송 서비스 및 방송 프로그램의 Segmentation에 대한 메타데이터를

수신함으로써, 녹화된 방송 프로그램 중에서 사용자가 보고 싶은 부분만을 비 선형적으로 시청 할 수 있게 하는 메타데이터 색인 기술 등을 필요로 하고 있다.

현재 TV-Anytime Forum은 “상이한 방송환경에서의 원활한 상호운용성(Interoperability)”이라는 주제로, 그 표준화를 위해 유럽, 미국, 일본 등 선진 국가들의 방송사, 콘텐츠 제작사, 가전사, 통신사 등을 중심으로 활발한 준비를 하고 있으며, 표준화 활동에 국내 기업 등이 참여하여 집진적 활동의 폭을 늘리고 있다.

본 논문에서는 이러한 맞춤형 방송서비스의 제공을 위한 멀티미디어 콘텐츠들의 메타데이터 색인 기법, 뉴스 및 축구 동영상에서의 키 프레임 추출 기법, 그리고 MPEG-2 TS 동영상의 비 선형 브라우징이 가능한 플레이어 및 이들을 통합한 TV-Anytime 1.3 Specification을 지원하는 저작 도구(Authoring Tool)에 대하여 기술한다.

II. 메타데이터 저작 도구

본 논문에서는 축구와 뉴스 동영상과 같은 멀티미디어 콘텐츠들에 대해서 TV-Anytime 규격에 의거한 메타데이터를 정의하고, 이를 구현하기 위해서 TV-Anytime 1.3 규격의 네임스페이스를 사용하였다. 저작 도구에는 장면 전환 검출기와 뉴스 및 축구 동영상에서의 키 프레임 과 텍스트 프레임 추출기, 그리고 MPEG-2 TS를 지원하는 구간 플레이어와 메타데이터 생성기를 통합하여 구현하였다.

그림 1에서 보는 바와 같이 저작 도구에 대한 전체적인 구조는 동영상 파일이 프레임 추출기를 통하여 뉴스에서의 앵커 프레임이나 텍스트 프레임 같은 각 프레임 영상을 추출하고, 추출된 프레임에 대한 정보

로써 영상이름 (예, KeyFrame0001.bmp), 시작 프레임 번호, 그리고 마지막 프레임번호를 저장하게 된다. 그런 다음 메타데이터 생성기는 자동으로 입력되는 DB 내의 영상정보와 수동으로 입력하는 기본정보, 부가정보를 이용하여 TV-Anytime 1.3 규격에 의거한 메타데이터를 생성하게 된다.

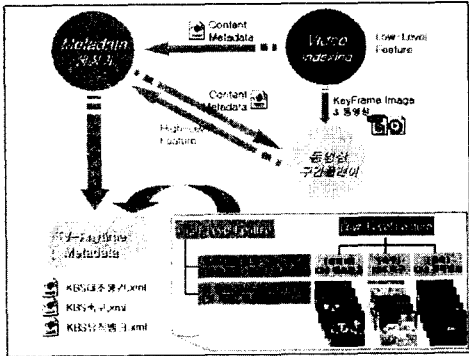


그림 1. 전반적인 메타데이터 구성도

또한, 메타데이터로 작성된 동영상 정보에 대해서 메타데이터 디코딩을 통한 비 선형 브라우징이 가능하도록 MPEG-2 TS 플레이어라 구성되어 있다. 그림 2는 메타데이터 저작 도구의 구성 화면이다.

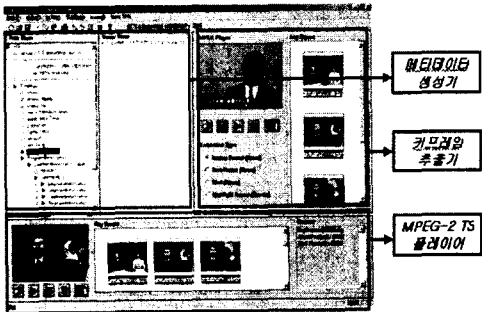


그림 2. 메타데이터 저작 도구의 구성화면

2.1 뉴스 동영상에서의 키 프레임 추출

폭발적인 멀티미디어 정보의 증가와 컴퓨터 성능의 향상, 네트워크 속도의 향상으로 이미 일반 사용자들은 원격지에 있는 동영상을 볼 수 있는 환경이 구축되었고, 이와 함께 특정 방송 프로그램의 특정 구간만을 시청할 수 있는 비 선형 브라우징에 대한 요구를 받게 되었다. 뉴스 동영상의 경우 특정한 사건, 사고 소식만

을 선택적으로 시청할 수 있기를 원하며, 이러한 서비스를 제공하기 위해서 뉴스 동영상을 뉴스 단위로 구조화하는 작업이 필요하다.[1]

일반적으로 뉴스는 시간 순서로 앵커가 말하는 부분과 뒤이어서 나오는 비 앵커 구간인 기자가 설명하는 보도자료 화면으로 규칙적인 구성을 가지고 있다.

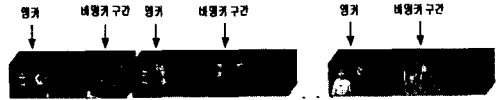


그림 3. 뉴스 동영상의 일반적 구성

이와 같이 하나의 뉴스 프로그램은 여러 조각의 뉴스 영상으로 나누어 놓을 수 있으며, 그림 3은 이러한 뉴스 동영상의 구조를 나타낸다. 따라서, 본 논문에서는 이러한 뉴스 비디오의 구조를 잘 구성하기 위한 효율적인 앵커 구간 추출 알고리즘(그림 4)을 제안한다.

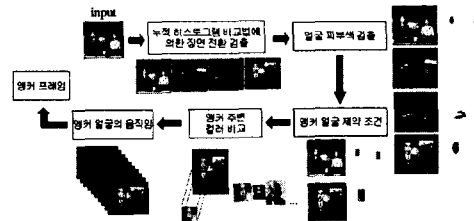


그림 4. 앵커 구간 추출 알고리즘

먼저, 입력받은 뉴스 동영상은 누적 히스토그램을 이용하여 장면 전환을 검출하게 된다[2]. 검출된 장면 전환 영상들은 앵커 후보 영상의 추출을 위하여 얼굴 피부색 검출과 얼굴 제약 조건을 사용하여 얼굴 영역을 검출하게 되며, 추출된 앵커 후보 영상들에 대해서 얼굴 주변 컬러와 앵커의 움직임은 아주 미세하다는 특징을 이용하여 앵커 구간을 추출한다[1][3].

2.2 축구 동영상에서의 관심장면 추출

축구 동영상에서 사람들의 관심 대상이 되는 슛 장면, 골인 장면 등을 관심장면(highlight)이라고 정의하면, 그림 5에서처럼 축구 동영상의 하이라이트 구간은 관심장면이 나온 후 연속되는 프레임에서 플레이어를 확대하고 있다는 것을 알 수 있다. 또한, 축구 경기 동영상에서 주 카메라는 대개 좌우로 이동하며 전체적인 경기 흐름을 찍고, 보조 카메라들은 부분적인 경기 요소들을 찍는다면, 하이라이트 구간은 골대 장면 후 연속적인 프레임에서 플레이어를 확대하므로 주 카메라



그림 5. 하이라이트 장면 후 연속 프레임

와 보조 카메라에서 찍은 프레임의 연속이라고 볼 수 있다[4]. 이런 하이라이트의 고유 특성을 이용하여 그림 6과 같은 하이라이트의 구조를 분석하는 알고리즘을 제안한다.

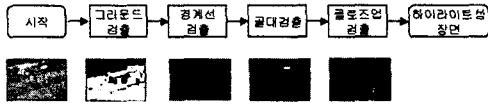


그림 6. 전체적인 알고리즘 순서도

먼저, 축구 동영상을 입력받아 RGB 색상 모델을 HSV 색상 모델로 변환한다. HSV 색상 정보를 이용하여 그라운드 영역과 비 그라운드 영역으로 이진화 영상을 구성하고, 이 영상에 대하여 모폴로지 필터를 적용하여 잡음을 제거한다. 잡음이 제거된 영상에서 그라운드와 비 그라운드 사이의 경계선에 대한 기술을 검출하고, 추출된 기술기 값이 일정한 구간 값을 가지면 후보 영상으로 추출하고 골대의 유무를 판단하여 골대의 크기가 일정한 문턱치 이상의 크기를 만족하면 연속되는 프레임에서 선수의 클로즈업 유무를 판단하여 하이라이트성 장면으로 추출한다. 골대의 유무는 입력 영상으로부터 골대가 가질 수 있는 색상(흰색)을 가진 화소를 추출하여 이들 화소들에 대해서 y축 방향으로 투영(projection)함으로써 문턱치 이상이 되는 두 개의 최대치(peak)가 나타나는 거리로 판단하고, 선수 클로즈업 장면을 검출하기 위해서는 CCV(Color Coherence Vectors)를 이용한다.

2.3 메타데이터 생성기

본 논문에서 구현한 메타데이터 생성기는 TV-Anytime 1.3 규격에 의거하여 작성되었다[6]. TV-Anytime 메타데이터는 구문(syntax), 타입들(types), 그리고 기본적인 값들(default values)을 정의하기 위해서 XML 기반의 MPEG-7 DDL를 이용한다[5]. 그러므로, TV-Anytime 메타데이터는 XML 문서로 표현되어 질 수 있으며, "TVAMain" 이라는 단 하나의 루트 요소(root element)를 포함하고, "TVAMain" 요소 안에는 많은 TV-Anytime 관련 타입들이 설명되어진다. 그림 7은 키프레임 추출기 등을 통해 자동적으로 추출되어지는 영상의 Feature 값들에 대해 메타데이터

를 작성하기 위한 요소와 타입들이다. 동영상에서 키프레임을 기준으로 한 색인은 그림의 <SegmentList> 요소 안에 여러 개의 <Segment>요소들로 구성되어지고, 그 하위 요소들로 <Title>, <Synopsis>, <Keyword>, <MediaLocator> 등을 갖는다.

따라서, 메타데이터 생성기는 키프레임 추출기를 이용하여 뉴스와 축구 동영상에서 앵커 프레임 구간과 관심장면 구간에 대한 정보 즉, 앵커 구간과 축구 관

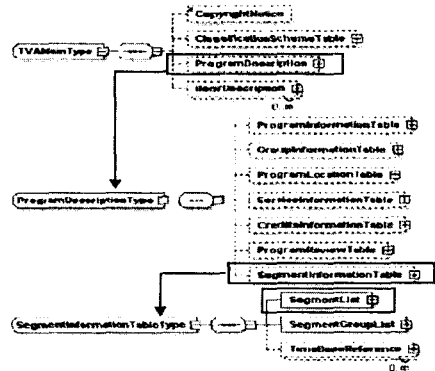


그림 7. TV-Anytime 메타데이터 요소와 타입들

심 장면 구간을 대표하는 영상들과 그 구간을 나타내는 시작 프레임과 끝 프레임의 프레임번호를 입력받아, 이 정보들을 기반으로 하여 서비스 제공자가 부가 정보를 추가하면 프로그램에 대한 메타데이터를 작성할 수 있도록 구현하였다. 또한, XML문서의 파싱 및 검색을 위하여 마이크로소프트사에서 제공하는 MSXML DOM Parser 4.0을 사용하여 메타데이터를 XML 문서로 표현한다.

메타데이터 생성기의 구성은 그림 8과 같다

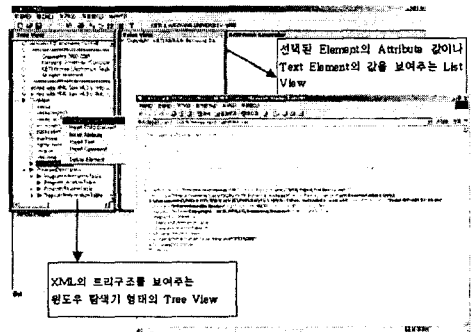


그림 8. 메타데이터 생성기의 구성 모습

이런 과정들이 모두 끝나게 되면 뉴스 및 축구 동영상

에 대해서 요약한 XML 기반의 메타데이터를 생성하게 된다.

2.4 MPEG-2 TS 플레이어

MPEG-1의 확장인 MPEG-2는 5~10Mbps 이상의 더욱 높은 비트율에서 고화질을 실현하기 위한 동영상과 음향압축의 표준이며, 또한 HDTV(High Definition TV)의 동영상 압축 표준이기도 하다. MPEG-2는 뛰어난 성능과 유연성에 따라 디지털 위성방송, HDTV, 디지털비디오디스크(DVD), 주문형 비디오(VOD) 등 많은 분야에서 채용이 결정되어 멀티미디어 혁명을 주도하는 원동력이 되고 있다. 본 연구에서 사용되는 MPEG-2 TS(Transport Stream) 플레이어는 키 프레임 추출기와 메타데이터 생성기에서 만들어진 메타데이터에 대해서 메타데이터 디코딩을 통한 동영상 구간 검색을 제공하기 위한 재생기이다. 또한, 사용자가 원하는 부분만을 볼 수 있는 비선형 브라우저 서비스를 제공한다. 플레이어의 구성은 아래 그림 9와 같다.

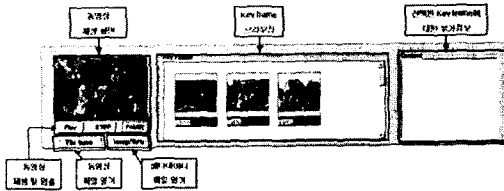


그림 9. MPEG-2 TS 플레이어의 구성 모습

MPEG-2 TS 플레이어는 마이크로소프트사의 DirectX 9.0 SDK에서 제공하고 있는 DirectShow 필터를 기본적으로 사용하고 있으며, MPEG-2 TS를 재생하기 위해서 Moonlight사에서 제공하고 있는 MPEG-2 디코더를 사용하고 있다.

III. 결론 및 향후과제

본 논문에서는 TV-Anytime 규격에 의거한 검색 및 Presentation 기법 개발을 목표로 하여 다음과 같은 알고리즘 개발과 시스템을 구축하게 되었다. 뉴스나 축구와 같은 멀티미디어 콘텐츠들에 대해서 여러 가지 특징 값들을 추출하여 효과적으로 검색 및 Presentation 할 수 있는 메타데이터 생성기를 개발하였으며, 이 메타데이터 생성기는 TV-Anytime 1.3 규격을 완벽하게 지원하고 있다. 또한 메타데이터를 구성하기 위해서 여러 가지 특징 값들을 추출할 수 있는 알고리

즘들을 개발하였고, 장면 전환 검출 및 키 프레임 추출할 수 있는 알고리즘을 개발하였다. 축구 동영상에 대해서는 사용자가 관심 있는 장면(Highlight)을 추출할 수 있는 알고리즘과, 뉴스 동영상에서 중요한 내용을 포함하고 있는 자막을 추출할 수 있는 알고리즘을 개발하였다. 또한 사용자가 MPEG-2 TS 동영상에 대해서 비 선형 시청이 가능하도록 구간 플레이어를 개발하였으며, 마지막으로 메타데이터 생성기와 키프레임 추출기, 그리고 MPEG-2 TS 플레이어를 통합한 TV-Anytime Presentation Engine 저작 도구 (Authoring Tool)를 개발하였다.

향후 시스템의 효율성과 영상에 대한 검색 기능을 지원하기 위하여 MPEG-7 표준에 정의되어 있는 기술자들을 이용하여 색인 할 수 있는 연구가 진행되어야 하며, 현재 개발된 알고리즘 및 시스템에 대한 속도 향상과 성능 향상을 위해 좀더 효과적인 색인 기법 개발과 같은 추가적인 연구가 필요할 것으로 생각된다.

참고문헌

- [1] 진승철, 박성환, "뉴스 정보의 단위 구조화를 위한 효율적인 앵커구간 추출 알고리즘," 한국방송공학 회 논문지, 제6권, 제3호, pp.260-269, 2001
- [2] 황두선, 이종설, 조위덕, 문영식, "누적 히스토그램과 에지 정보를 이용한 장면 전환 검출," 대한전자 공학회 하계종합학술대회, 2002.6
- [3] WeiQiang Wan, Wen Gao, "A Fast Anchor Shot Detection Algorithm on Compressed Video," Proc. Second IEEE Pacific Rim Conference on Multimedia, pp.873-878, 2001
- [4] 우성형, 진승철, 박성환, "축구경기 동영상에서의 효율적인 골 영역 검출 방법," 대한전자공학회 추계 종합학술대회 논문집, 제23권, 제2호, pp.81-84, 2000
- [5] P. Beek, B. Ana, Text of 15938-5 FCD information Technology Multimedia Content Description Interface Part 5 Multimedia Descriptor Sechemes, ISO/IEC/JTC/SC29/WG11/N3966, 2001
- [6] The TV-Anytime Forum Specification Series: S-3 On: Metadata Final Specification, Version 1.3, December 15, 2002