

## 방송 환경에서 내용기반 검색 서비스에 관한 연구

추진호, 배태면, 진성호, 강경옥\*, 노용만  
한국정보통신대학원대학교 멀티미디어 그룹  
\*한국전자통신연구원 방송미디어 연구부

### A Study on Content Based Retrieval Service in Broadcasting Environment

Jin Ho Choo, Tae Meon Bae, Sung Ho Jin, Kyoungok Kang\*, Yong Man Ro  
Multimedia Group , Information and Communications University

E-mail: sselmo2000@icu.ac.kr

\*Electronics and Telecommunications Research Institute, Korea(ETRI)

#### Abstract

본 논문에서는 지능형 방송에서 내용 기반 검색 서비스에 관해 연구하였다. 이를 통해 MPEG-7 오디오/비주얼 서술자들을 이용하여 TVAF 메타데이터와 메타데이터 전송 서비스 스키마를 확장하였다. 그리고 MPEG-7 오디오 서술자를 이용한 내용 기반 검색 시나리에 제안된 스키마를 적용하여 효용성을 증명하였다.

#### I. 서론

디지털 방송 서비스의 시작으로 채널의 수가 증가하고 컨텐츠의 양이 방대해짐에 따라 사용자가 원하는 채널이나 컨텐츠를 빠르고, 효율적으로 검색하고 이용할 수 있도록 하는 서비스가 필요하게 되었다. 이러한 사용자의 요구를 충족시키기 위해 TV-Anytime Forum(TVAF)에서는 메타 데이터를 이용하여 관련된 기술과 애플리케이션에 대한 표준화 작업을 진행중이다. TVAF는 사용자가 다 채널 환경하에서 관련 메타데이터를 이용하여 모든 채널을 직접 검색하거나, 신문 등의 매체를 통해 채널 정보를 검색하지 않고 원하는 방송 프로그램을 빠르고 쉽게 획득할 수 있게 하는 것을

목적으로 한다. 그러나 TVAF에서는 컨텐츠 공급자에 의해 기본적인 프로그램 EPG 정보나 채널 등의 서비스 정보, 사용자 선호도 정보만이 주어지고 있어 공급자에 의해 사용자가 요구사항을 제한 받게 되는 단점을 가진다. 따라서 언제 어디서나 사용자의 취향과 요구에 따라 실감 있는 맞춤형 방송서비스 제공을 목표로 하는 SmarTV(Super-intelligent Multimedia Anytime-anywhere Realistic TV)에서는 내용기반 검색이 필수적인 요소가 되고 있다.[1] 이를 위해 내용 기반 검색을 제공하는 MPEG-7의 오디오/비주얼 서술자로 TVAF 메타데이터를 확장시킬 필요가 있다.

본 논문에서는 현재 개발 진행중인 지능형 방송 시스템인 SmarTV에서 내용기반 검색을 구현하기 위해 멀티미디어 컨텐츠의 검색 표준인 MPEG-7의 오디오/비주얼 서술자들을 사용할 수 있도록 TVAF 메타데이터를 확장하고 이를 검색에 사용하도록 메타데이터 전송 서비스를 확장한다. 그리고 제안된 스키마를 MPEG-7 오디오 서술자를 이용하는 내용 기반 검색 시나리오에 적용하여 효용성을 증명하였다. 본 논문의 구성은 다음과 같다. 2 절에서는 관련된 연구에 대해 설명하고, 3 절에서는 제안된 스키마들에 대해 설명한다. 4 절에서는 이를 이용한 서비스 시나리오를 설명하고, 5 절에서는 결론을 기술한다.

## II. 관련 연구

### 2.1. MPEG-7 오디오/비주얼 서술자

MPEG-7은 내용 기반 검색을 위해 컨텐츠의 오디오/비주얼 서술자들을 제공한다. 오디오 서술자는 컨텐츠의 오디오의 주파수, 하모닉스 등의 오디오 관련 특징 값들을 서술하고, 비주얼 서술자는 컨텐츠에 대해서 오브젝트의 모양, 크기, 질감, 컬러, 움직임, 카메라 움직임 등에 관한 특징 값을 서술하고 있다[3][4].

### 2.2. TV – Anytime Forum 메타데이터

TVAF의 메타데이터 시스템은 사용자가 enhanced broadcast나 interactive TV, Internet and local storage를 포함하는 내·외부 소스들로부터 컨텐츠의 검색, 네비게이션 그리고 관리할 수 있도록 하는데 그 목적이 있다[5].

#### 2.2.1 TVAF 메타데이터 종류

- 컨텐츠 서술 메타데이터(Content Description Metadata) 텔레비전 프로그램과 같은 단일 컨텐츠 혹은 프로그램 그룹에 관한 정보, 출연하는 연기자에 관한 정보를 서술하기 위한 메타데이터.

- 개체 서술 메타데이터(Instance Description Metadata) 컨텐츠의 위치, 서비스정보등 프로그램 개체에 관한 메타데이터.

- 사용자 메타데이터(Consumer Metadata)

사용자의 선호도나 프로파일, 사용 내역(Usage history)에 관한 정보를 기술하는 메타데이터이다.

- 세그먼테이션 메타데이터(Segmentation Metadata)

컨텐츠를 세그먼트 단위로 분할하여 재구성한 뒤 각 세그먼트에 대한 정보를 기술하는 메타데이터.

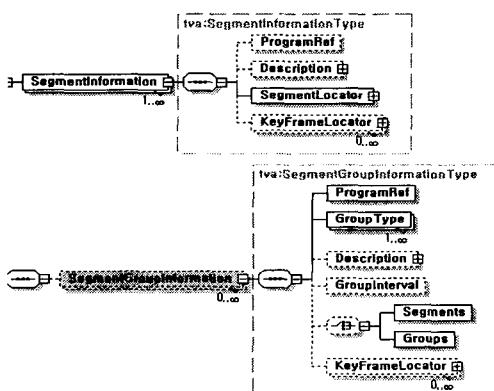


그림 1. TVAF SegmentInformationTable 구조

그림 1은 TVAF에서 표준화한 세그먼테이션 메타데이터의 SegmentInformation과 SegmentGroupInformation 구조를 보여주고 있다.

SegmentInformation는 세그먼트의 타이틀, 키워드 등을 기술한다. 특히 컨텐츠내 세그먼트의 위치를 나타내는 ‘SegmentLocator’ 요소와 세그먼트의 키 프레임을 나타내는 ‘KeyFrameLocator’ 요소가 포함된다.

SegmentGroupInformation은 여러 개의 세그먼트를 하나의 그룹으로 묶어 정보를 기술한다

#### 2.2.2 메타데이터 전송 서비스

TVAF는 양방향 방송환경에서 메타데이터의 전송 서비스를 위해 get\_Data Operation과 submit\_Data Operation의 두 종류 메타데이터 서비스 타입을 제공하고 있다[6].

- get\_Data Operation : 사용자가 서버에 메타데이터를 검색하는 기능. 쿼리 타입 지원.
- submit\_Data Operation : 사용자의 선호도, 사용 내역(Usage History) 메타데이터를 서버에 제공하는 기능

## III. 제안하는 TVAF 스키마

본 논문에서는 TVAF 세그먼테이션 메타데이터의 SegmentInformationTable 중 SegmentInformation과 SegmentGroupInformation을 MPEG-7 오디오/비주얼 서술자들을 사용하여 스키마를 확장하였다. 또한 추가된 메타데이터를 이용한 검색 서비스를 위해 메타데이터 전송 서비스 스키마도 확장하였다.

### 3.1. SegmentInformationTable 확장

단일 컨텐츠 정보를 기술하는 컨텐츠 서술 메타데이터나, 채널과 서비스 정보를 나타내는 개체 서술 메타데이터, 사용자 정보를 나타내는 사용자 메타데이터보다는 세그먼테이션 메타데이터의 각 키 프레임 정보에 키 프레임의 오디오/비주얼 서술자들을 추가하는 것이 가장 적합하다.

#### 3.1.1. SegmentInformation

컨텐츠 제공자는 컨텐츠를 각 세그먼트로 나눈 후, 세그먼트의 위치정보나 키프레임의 위치정보를 기술하게 된다. 이때 각 키 프레임마다 추출되는 MPEG-7 오디오/비주얼 서술자들을 키 프레임 정보 나타냄으로써 내용 기반 검색을 할 수 있다.

그림 2 는 *SegmentInformation* 의 확장된 스키마 구조를 보여준다. 그림 1 이 단순히 키 프레임의 위치만을 기술한 것에 반해, 그림 2 는 키 프레임에 대해서 요소들을 확장하고, 서술자 정보를 기술할 수 있도록 하였다. 이때 *AudioDescriptor* 는 MPEG-7 의 Part4 에서 지원하는 오디오 서술자를 가지고, *VisualDescriptor* 는 MPEG-7 의 Part3 에서 지원하는 비주얼 서술자를 가진다.

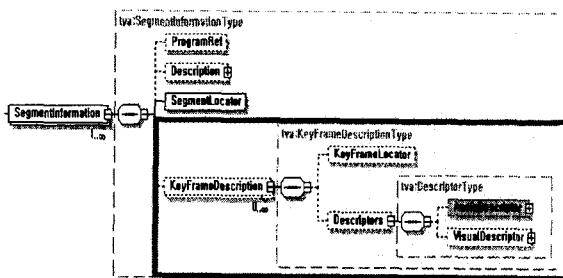


그림 2. 확장된 SegmentInformation

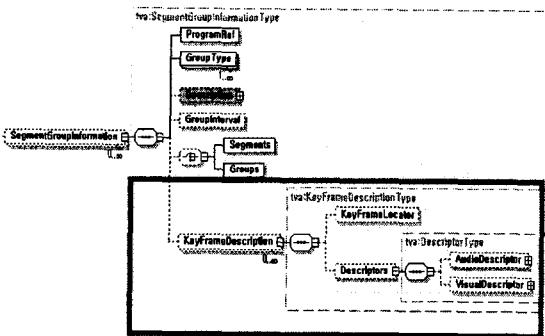


그림 3. 확장된 SegmentGroupInformation

### 3.1.2. SegmentGroupInformation

그림 3은 SegmentGroupInformation의 확장된 구조를 보여준다. 세그먼트 그룹에서도 존재하는 키 프레임에 대한 요소를 확장하여, 키 프레임 위치 정보와 함께 오디오/비주얼 서술자를 기술할 수 있도록 하였다.

### 3.2. 메타데이터 전송 서비스 타입 확장

본 논문에서는 사용자가 서버측에 쿼리를 전달할 수 있는 get\_Data Operation에 대해서 스키마를 확장하였다. submit\_Data Operation의 경우 사용자의 정보를 서버에 전달하는 기능을 하기 때문에 검색 기능과 관련이 없다.

`get_Data Operation` 쿼리 포맷의 경우 쿼리값들이 '`BinaryPredicate`' 요소에 속성값으로 지정되어 있다. 이 경우 타이틀이나, 키워드등 스트링 형태의 쿼리는 사용

가능 하지만, 새로 추가된 서술자 정보는 그 값 자체가 XML의 계층적 구조를 가지고 있어 한 요소의 속성값으로 대입하는 것 자체가 불가능하다. 따라서 새로운 타입의 쿼리 형태가 구성될 수 있다.

본 논문에서는 QueryConstraints 요소의 하위 계층으로 서술자 값들만을 서비스할 수 있는 QueryFeature 요소를 추가하였다. 이 요소는 세그먼테이션 서술자 확장에 쓰였던 Descriptortype을 가진다. 이 확장으로 인해 MPEG-7 서술자들을 서버에 쿼리로 보낼 수 있게 된다.

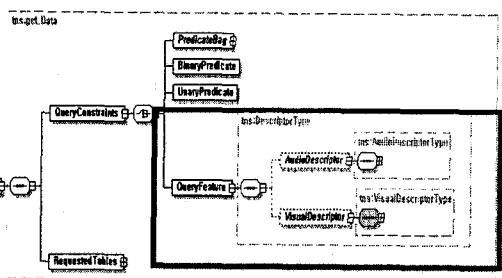


그림 4. 확장된 get\_Data Operation

#### IV. 서비스 시나리오

본 논문에서는 SmarTV에서 드라마의 배경음악 검색 서비스를 위하여 제안한 스키마를 적용하였다. 이 서비스는 사용자가 현재 방송중인 드라마의 배경음악에 대한 정보를 알지 못하는 경우, 오디오 특징값을 사용하여 배경 음악과 동일한 음악 파일을 자동으로 검색해 사용자에게 들려주는 시나리오이다.

이 시나리오의 겹색 과정은 다음과 같다.

- ① 사용자는 원하는 드라마의 배경음악을 일정구간 저장한 다음, 검색 버튼을 누른다. 이때 User Interface Agent는 저장된 데이터를 컨텐츠 관리 에이전트에 보내게 되고, 여기서 오디오 서술자들을 추출하여 그림 5와 같이 제안된 스키마를 사용한 쿼리 메타데이터를 만든다.
  - ② 쿼리 메타데이터는 각 에이전트들에 의해 서비스 제공자(SP)의 관리자 에이전트(MA)에 보내지게 된다.
  - ③ 관리자 에이전트는 서버 데이터 베이스에 저장된 컨텐츠의 메타데이터 내 오디오 서술자들을 이용하여 패턴매칭 기법을 이용하여 동일한 음악파일을 찾아 CRID's(Content Referencing ID)를 사용자에게 보낸다.
  - ④ CRID's를 사용자가 받게 되면 원하는 CRID를

선택하여 컨텐츠의 위치자(Locator)를 얻게 되고, 이것을 가지고 해당 컨텐츠를 획득하게 된다.

```
<get_Data xmlns="http://www.tv-anytime.org/2002/11/transport" xmlns:mpeg7=>
<QueryConstraints>
  <QueryFeature>
    <AudioDescriptor>
      - <mpeg7:HarmonicSpectralCentroid>
        <mpeg7:Scalar>811.099976</mpeg7:Scalar>
      </mpeg7:HarmonicSpectralCentroid>
      - <mpeg7:HarmonicSpectralDeviation>
        <mpeg7:Scalar>0.021578</mpeg7:Scalar>
      </mpeg7:HarmonicSpectralDeviation>
      - <mpeg7:HarmonicSpectralSpread>
        <mpeg7:Scalar>-0.585618</mpeg7:Scalar>
      </mpeg7:HarmonicSpectralSpread>
      - <mpeg7:HarmonicSpectralVariation>
        <mpeg7:Scalar>0.134047</mpeg7:Scalar>
      </mpeg7:HarmonicSpectralVariation>
      - <mpeg7:SpectralCentroid>
        <mpeg7:Scalar>987.950317</mpeg7:Scalar>
      </mpeg7:SpectralCentroid>
      - <mpeg7:AudioFundamentalFrequency lLimit="-25">
        <mpeg7:Scalar>188.770599</mpeg7:Scalar>
      </mpeg7:AudioFundamentalFrequency>
    </AudioDescriptor>
  </QueryFeature>
</QueryConstraints>
</get_Data>
```

그림 5. 내용 기반 검색을 위한 쿼리 메타데이터

```
<ProgramDescription>
  <ProgramInformationTable version="1">
    <ProgramInformation programId="crid://www.limbc.com/musickamp200100080" version="11">
      <BasicDescription>
        <Title>Hummingbird</Title>
        <Keywords>
          <mpeg7:Keywords>유카 구리 모드</mpeg7:Keywords>
        </Keywords>
      </BasicDescription>
      <ProgramInformationTable>
    </ProgramInformationTable>
  <SegmentInformationTable version="1">
    <SegmentsList>
      - <SegmentInformation segmentId="segid_humvn_1" refType="SegmentInformationType" version="15">
        <SegmentRef crid="crid://www.limbc.com/musickamp200100080"/>
        <SegmentStarts>
          <mpeg7:MediaTimePoint>1900-01-01T00:00:00</mpeg7:MediaTimePoint>
        <mpeg7:MediaDuration>PT30S</mpeg7:MediaDuration>
        <SegmentEnds>
        </SegmentEnds>
        <KeyframeDescription>
          - <KeyframeType>
            <AudioOpus>
              <mpeg7:HarmonicSpectralCentroids>
                <mpeg7:Scalar>861.783225</mpeg7:Scalar>
              </mpeg7:HarmonicSpectralCentroids>
              - <mpeg7:HarmonicSpectralDeviation>
                <mpeg7:Scalar>-0.022899</mpeg7:Scalar>
              </mpeg7:HarmonicSpectralDeviation>
              - <mpeg7:HarmonicSpectralSpread>
                <mpeg7:Scalar>0.589251</mpeg7:Scalar>
              </mpeg7:HarmonicSpectralSpread>
              - <mpeg7:HarmonicSpectralVariation>
                <mpeg7:Scalar>0.149569</mpeg7:Scalar>
              </mpeg7:HarmonicSpectralVariation>
              - <mpeg7:SpectralCentroid>
                <mpeg7:Scalar>987.950993</mpeg7:Scalar>
              </mpeg7:SpectralCentroid>
              - <mpeg7:AudioFundamentalFrequency lLimit="-25">
```

그림 6. 내용 기반 검색을 위한 Program Description

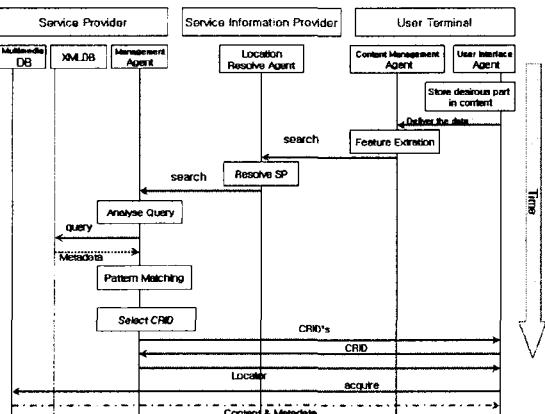


그림 7. SmarTV에서 내용 기반 검색 서비스 과정

그림 6은 서버에 저장되어 있는 Hummingbird라는 음악 파일에 대한 메타데이터의 예이다. ProgramInformation과 SegmentInformation 메타데이터에 대해 기술되어 있음을 알 수 있다. 여기서는 6개의

오디오 서술자가 사용되었다.

그림 7은 서비스 시나리오에 대한 시스템의 수행 과정을 나타낸 것이다.

## V. 결론

본격적인 사용자 중심의 지능형 방송 시스템에서 내용 기반 검색이 필수적인 요소이다. 본 논문에서 제시된 스키마와 이를 이용한 내용 기반 검색 시나리오는 TVAF 표준과 MPEG-7의 오디오/비주얼 서술자를 사용하여 사용자로 하여금 자기가 원하는 방법과 시간에 다양한 형태의 프로그램 및 관련된 멀티 미디어 컨텐츠들을 소비할 수 있음을 보여준다.

향후 연구에서는 방송용 환경에서 TVAF에서 제공하는 방송 컨텐츠 이외의 멀티 미디어 컨텐츠에 대한 메타데이터 정의와 이들을 효율적으로 획득할 수 있는 메커니즘에 대한 연구가 필요하다.

## Acknowledgment

본 논문은 정보통신부 지원 “지능형 방송 에이전트 및 메타데이터 관리 기술” 과제의 수행 결과임

## 참고문헌

- [1] Sung Ho Jin, Tae Meon Bae, Jin Ho Choo, “A MPEG-7 and TV-Anytime based Multi-Agents Platform for Interactive Broadcasting System,” Proceeding on SPIE Electronic Imaging, Jan. 2003
- [2] 진성호, 배태면, 노용만 “지능형 방송 서비스를 위한 위한 멀티에이전트 시스템 설계에 관한 연구,” 멀티미디어 추계 학술 발표 대회, V.5 No.2 p404-407, 2002
- [3] ISO/IEC FDIS 15938-4, “Part4:Audio,” MPEG-7, June 2001.
- [4] ISO/IEC FCD 15938-3 “Part3:Visual,” MPEG-7, March 2001
- [5] SP003v1.2, “Metadata,” TV-Anytime Forum, Dec 2002.
- [6] SP006v1.0, “Metadata Services over a Bi-directional Network,” TV-Anytime Forum, Jun 2002
- [7] SP002v1.3, “System Description,” TV-Anytime Forum, Feb 2003.
- [8] SP004v1.2, “Content Referencing,” TV-Anytime Forum, Jun 2002.