

장면 기반 메타데이터 국제 표준화 동향

이상윤 한국전자통신연구원 공공안전지능화연구실 책임연구원

1. 머리말

5G 시대가 개막되면서 각종 동영상을 좀 더 빠르게 다운로드받아 시청할 수 있는 시대가 도래하였다. 예전에는 인기 있는 드라마를 보기 위해 본방 사수라는 말이 유행하기도 했다. 하지만 지금은 자신이 무엇인가를 보고 싶을 때 언제 어디서든 바로 동영상을 다운로드받아 보는 것이 일반화되었다. 그런데 최근 몇 년 전부터 동영상도 전체 영상보다는 자신이 보고 싶은 내용, 즉 클립 형태의 부분 동영상을 보는 소비 행태가 일반화되고 있다. 이에 발맞춰 유튜브, 방송사업자, 넷플릭스, 포털사업자들은 개인 맞춤형 동영상을 제공하기 위해 클립 형태의 동영상을 경쟁적으로 서비스하고 있다. 현재는 동영상 클립에 광고를 붙이는 것이 주요 수익원이지만, 광고 수익이 포화됨에 따라 새로운 수익을 창출할 수 있는 부가서비스 개발에 관심이 높아지고 있다. 특히, 최근에 컴퓨터 성능, AI 기술 등이 획기적으로 발전함에 따라 동영상에서 자동으로 메타데이터를 추출할 수 있다는 기대가 높아지고 있다.

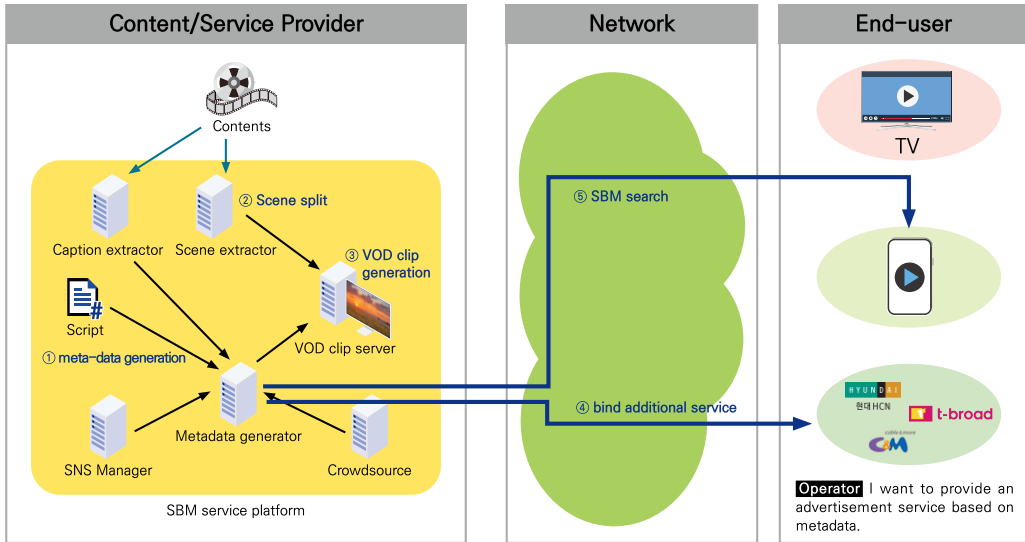
본고에서는 이러한 동영상 클립 기반의 부가서비스와 관련된 표준으로서 최근에 ITU-T에서 제정된 장면 기반 메타데이터 표준에 대해 소개하고자 한다.

2. 장면 기반 메타데이터 표준화 동향

2.1 서비스 시나리오

요즘 사람들은 TV에서 프로그램을 보는 동안 세컨드스크린에서 부가 정보를 검색하거나 찾아본다. 이때 방송 사업자는 장면 기반 메타데이터를 사용하여 제품 광고 및 매시업 서비스와 같은 부가 서비스를 제공할 수 있다. 장면 기반 메타데이터는 IPTV 서비스를 위해서 제정되었지만, 다른 방송 매체에서도 활용이 가능하다. [그림 1]은 장면 기반 부가 서비스의 한 서비스 시나리오이다.

- 1단계: 메타 데이터 생성기는 스크립트, 클라우드 소싱 데이터 및 장면 정보를 이용하여 장면 기반 메타데이터(SBM, Scene-based metadata)를 생성한다.
- 2단계: 장면 추출기는 방송 콘텐츠를 일련의 샷으로 분할한다. 그런 다음 자막 분석을 통해 의미 단위로 장면을 결합한다.
- 3단계: VOD(Video on Demand) 클립 서버는 분할된 장면을



※ 출처: ITU-T.H.753

[그림 1] 장면 기반 메타데이터 서비스 시나리오

- VOD 클립으로 인코딩하고 스토리 그래프로 표현된 비디오 스토리를 생성한다.
- 4단계: 사업자는 SBM을 기반으로 부가 서비스를 결합한다.
 - 5단계: 사용자는 세컨드스크린 장치에서 SBM 및 부가 서비스를 검색한다.

요소들은 부가 서비스와 결합할 경우, 다양한 매시업 서비스 창출이 가능한 것들을 우선적으로 고려하여 선정되었다.

2.2 장면 기반 메타데이터 구성 요소와 형식

장면 기반 메타데이터는 기본 요소를 비롯하여 위치, 음악, 객체, 통계 요소로 이루어져 있다. 이 구성

2.2.1 기본 요소 장면 기반 메타데이터

기본 요소 장면 기반 메타데이터는 동영상 클립에 대한 기본적인 콘텐츠 정보를 표현한 것이다. <표 1>은 장면 기반 메타데이터 중 기본 요소의 일부를 나타냈다.

<표 1> 기본 요소 장면 기반 메타데이터

엘리먼트명	설명	데이터 타입	사용 예
SceneID	장면 식별자	<IDREF>	"KBS1_T2017-02-01-10", (channel ID-program ID-episode ID-scene order)
SceneOrder	장면 번호	<integer>	"5", "24"
SceneStartTime	장면 시작 시간	<clock-value>	"00:53:38", "01:05:22"
SceneEndTime	장면 종료 시간	<clock-value>	"00:43:28", "01:13:45"
SceneDuration	장면 재생 시간	<integer>	"45", "234"
SceneKeyframeImageURL	장면을 표현하는 키프레임 이미지에 대한 URL	<URL>	http://52.198.166.6:8000/2016-017-KBS1_T2016-02-01-DescendentOfSun-302/SCENE/CLIP/KBS1_T2016-02-01-02-SCENE_KEYFRAME.jpg
EpisodeID	장면을 포함하고 있는 에피소드 식별자	<IDREF>	"KBS1_T2017-02-01", (channel ID-program ID-episode ID)
ProgramID	장면을 포함하고 있는 프로그램 식별자	<IDREF>	"KBS1_T2017-02", (channel ID-program ID)
ProgramName	장면을 포함하고 있는 프로그램명	<string>	"태양의 후예"

2.2.2 위치 요소 장면 기반 메타데이터

위치 요소 장면 기반 메타데이터는 동영상 클립에 포함된 위치 정보를 표현한 것이다. <표 2>는 장면 기반 메타데이터 중 위치 요소를 나타냈다.

2.2.3 음악 요소 장면 기반 메타데이터

음악 요소 장면 기반 메타데이터는 동영상 클립 재생 시 배경 음악에 대한 정보를 표현한 것이다. <표 3>은 장면 기반 메타데이터 중 음악 요소를 나타냈다.

〈표 2〉 위치 요소 장면 기반 메타데이터

엘리먼트명	설명	데이터 타입	사용 예
LocationName	장면에서 나타난 장소의 명칭	<string>	“서울”, “상해”, “도쿄”
LocationAddress	장면에서 나타난 장소의 주소	<string>	“대전시 유성구 가정동 한국전자통신연구원”
LocationURL	추가적인 정보를 제공하기 위해 현재의 장면과 연결된 위치의 URL	<URL>	https://www.whitehouse.gov/

〈표 3〉 음악 요소 장면 기반 메타데이터

엘리먼트명	설명	데이터 타입	사용 예
MusicTitles	장면에서 재생되고 있는 곡명	<string>	“You are my everything”, “넌 누구니?”
MusicGenre	장면에서 재생되고 있는 곡의 장르	<string>	“Blues”, “Jazz”, “Hip hop”
MusicSinger	장면에서 재생되고 있는 곡의 가수 이름	<string>	“조용필”
MusicURL	장면에서 재생되고 있는 곡의 URL	<URL>	“http://music.naver.com/album/index.nhn?albumId=623671”
MusicLyrics	장면에서 재생되고 있는 곡의 가사	<text>	
MusicVideoURL	장면에서 재생되고 있는 뮤직비디오의 URL	<URL>	“http://tv.naver.com/v/1943434/list/67096”

〈표 4〉 객체 요소 장면 기반 메타데이터

엘리먼트명	설명	데이터 타입	사용 예
ObjectID	객체식별자	<IDREF>	
ObjectName	객체명	<string>	“코카콜라”
ObjectCategory	객체분류	<string>	“가방”, “의류”, “액세서리”
ObjectNumber	객체일련번호	<string>	
ObjectColor	객체의 색상	<string>	“노랑”, “파랑”
ObjectPrice	객체의 가격	<string>	“12,000원”, “23,000원”
ObjectImageURL	객체 이미지의 URL	<URL>	“http://www.ssg.com/item/itemView.ssg?itemId=0000004553455&siteNo=6004&salestrNo=6005”
ObjectSteelCutImageURL	객체의 스틸컷 이미지	<URL>	섬네일 이미지
ObjectShoppingURL	쇼핑홈페이지 URL	<URL>	“www.11st.co.kr”
ObjectReleaseDate	상품 출시일	<date>	2017-05-25, 2016-09-14
ObjectCountry	객체 제조국명	<string>	“한국”, “대한민국”
ObjectSize	객체의 크기	<string>	“높이”, “넓이” (실제 크기)
WearOnActorName	상품을 착용한 배우명	<string>	“송준기”, “송해교”
WearOnActorRoleName	상품을 착용한 배우의 배역명	<string>	“강모연”, “류시진”

〈표 5〉 통계 요소 장면 기반 메타데이터

엘리먼트명	설명	데이터 타입	사용 예
ClickCount	장면을 클릭한 횟수	<integer>	34,498, 456,789
GoodRatingCount	장면에 대해 '좋음' 평가를 준 횟수	<integer>	792,345, 330,901
BadRatingCount	장면에 대해 '나쁨' 평가를 준 횟수	<integer>	832,316, 140,508

2.2.4 객체 요소 장면 기반 메타데이터

객체 요소 장면 기반 메타데이터는 동영상 클립에 등장하는 객체에 대한 정보를 표현한 것이다. 〈표 4〉는 장면 기반 메타데이터 중 객체 요소를 나타냈다.

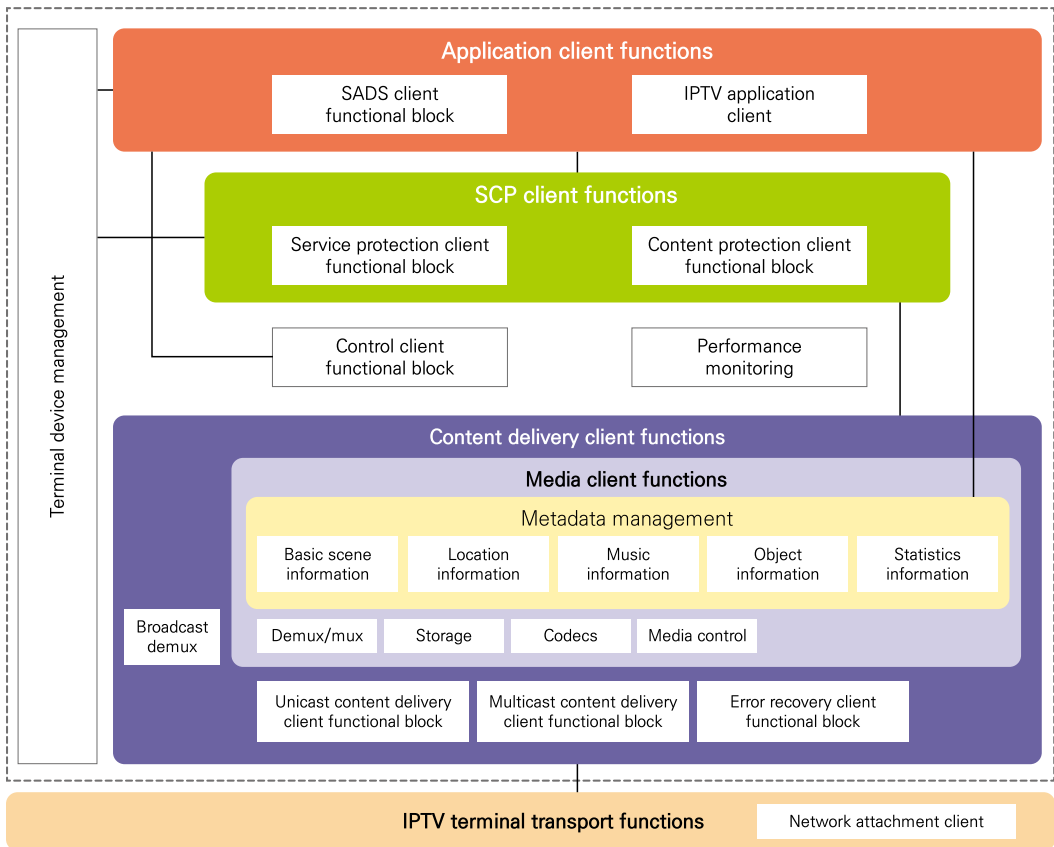
2.2.5 통계 요소 장면 기반 메타데이터

통계 요소 장면 기반 메타데이터는 동영상 클립을 대상으로 통계에 대한 정보를 표현한 것이다. 〈표 5〉는 장면 기반 메타데이터 중 통계 요소를 나타냈다.

2.3 장면 기반 메타데이터 관리 함수

2.3.1 장면 기반 메타데이터 관리 함수 정의

[그림 2]는 ITU-T H.753 규격에서 제정된 IPTV 터미널 디바이스 함수 아키텍처와 장면 전송 클라이언트 기능을 위한 갱신된 아키텍처를 보여준다. 미디어 클라이언트 함수 내에 메타데이터 관리 함수가 추가되었으며, 여기에는 기본 장면 정보, 위치 정보, 음악 정보, 객체 정보, 통계 정보를 전달할 수 있는 함수들이 정의되었다.



※출처: ITU-T H.753

[그림 2] 장면 기반 IPTV 서비스를 위한 IPTV 터미널 디바이스의 갱신된 함수 구조 아키텍처 블록 다이어그램

```

<link rel="get" href="/api/EpisodeList">
  <request>
    <http_method>GET</http_method>
    <url>
      <parameters_set>
        <parameter type="IDREF" required="true">
          <name>ProgramID</name>
        </parameter>
      </parameters_set>
    </url>
  </request>
  <responses>
    <response_set>
      <response type="IDREF">
        <name>EpisodeID</name>
      </response>
      <response type="date">
        <name>BroadcastingDate</name>
      </response>
      <response type="clock-value">
        <name>BroadcastingTime</name>
      </response>
      <response type="text">
        <name>storyline</name>
      </response>
    </response_set>
  </responses>
</link>


```

※ 출처: ITU-T H.753

[그림 3] 웹 기반 API 예제

2.3.2 장면 기반 메타데이터 서비스를 위한 웹기반 API

미디어 서비스를 위한 웹 기반 API는 클라이언트 측의 SBM 애플리케이션이 SBM 서비스 플랫폼에 접근하도록 허용한다. [그림 3]은 레스트풀 서비스 디스크립션 랭귀지(RSDL, Restful Service Description Language)로 표현된 한 예이다.

다. 최근에 이러한 부가서비스 개발을 위한 메타데이터 국제 표준이 제정된 것은 환영할 만하다. 이를 기반으로 다양한 메타데이터 기반의 부가 서비스가 제공되기를 기대하며 사업자와 소비자 모두 이 표준의 혜택을 받기를 바란다. 

3. 맺음말

방송 클립 기반의 콘텐츠 유통 서비스 시장이 점점 확대되고 있다. 하지만 조만간에 이 시장도 포화 상태에 다다를 것이고, 사업자들은 이에 대한 돌파구로 부가서비스 개발에 좀 더 투자할 것으로 전망된

※ 본 연구는 미래창조과학부 및 정보통신기술연구진흥센터의 정보통신·방송연구개발사업의 일환으로 수행하였음[2015-0-00219, 개방형 미디어 생태계 구축을 위한 시맨틱 클러스터 기반 시청상황 적응형 스마트방송 기술 개발].

참고문헌

[1] ITU-T H.753 Scene-based metadata for IPTV services

주요용어풀이

- VOD(Video on Demand): 주문형 비디오 서비스
- SBM(Scene-based Metadata): 동영상 클립의 내용을 표현하는 메타데이터
- RSDL(RESTful Service Description Language): HTTP 기반의 웹 애플리케이션을 기계와 인간이 읽을 수 있도록 XML로 표현한 것