

비디오 내용기술을 위한 MPEG-7과 TV Anytime 메타데이터의 상호 변환

임화영⁰, 이창윤, 김혁만
 국민대학교 전산학과 멀티미디어 데이터베이스 연구실
 {juliet⁰, wof3, hmkim}@cs.kookmin.ac.kr

Metadata Transcoding between MPEG-7 and TV Anytime for Segmentation Information of Video

Hwayoung Lim⁰, Chang Yoon Lee, Hyeokman Kim
 Multimedia Database Lab., Dept. of Computer Science, Kookmin University

요 약

본 논문에서는 MPEG-7 스키마에 따라 비디오의 내용을 기술한 메타데이터를 TVA 스키마에 따른 메타데이터로 변환, 그리고 그 역방향으로의 변환 방법을 제안한다. 이를 위해 MPEG-7과 TVA 스키마를 분석하여 그들의 유사 점과 차이점을 밝히다. 또한 중첩 표현법과 참조 표현법으로의 상호 변환시 야기되는 id 처리 문제, 대표화면 정보의 처리, 위치 정보의 처리 등에 관한 방법을 기술한다.

1. 서론

비디오 콘텐츠는 다양한 포맷으로 상이한 분배 네트워크를 거쳐 사용자들에게 제공된다. 현재 비디오 분배 네트워크로 가장 큰 역할을 하는 것은 고품질 비디오를 제공하는 방송망과 저화질 비디오를 제공하는 인터넷, 그리고 오프라인에서의 테이프나 DVD를 들 수 있다. 이들 네트워크에서 유통되는 비디오는 같은 내용이라도 다양한 형태의 포맷으로 제공되고 있다. 한편 사용자들은 수많은 비디오로부터 자신이 원하는 비디오 혹은 비디오의 특정 장면을 찾기 위해 비디오의 내용을 브라우징하길 원하고 있다. 이런 브라우징을 위해서는 비디오의 내용을 계층구조로 표현한 메타데이터(metadata)가 있어야 한다.

이러한 네트워크 환경 및 사용자 요구를 바탕으로 하여, 국제 표준화 기구인 ISO에서는 MPEG-7이라는 멀티미디어 데이터의 내용 기술에 관한 국제 표준을 만들고 있다[1]. MPEG-7은 주로 웹 환경에서의 멀티미디어 정보 검색을 목적으로 개발되고 있으나, 그 적용 환경을 웹으로만 국한시키지는 않았다. 한편 PVR(Personal Video Record)의 등장으로 TV 방송을 하드디스크에 저장하여 원하는 시간대에 저장된 프로그램을 시청할 뿐만 아니라 기존에는 볼 수 없었던 다양한 서비스가 방송에서도 가능하게 되었다. 따라서 이런 변화된 방송 환경에 대처할 수 있는 메타데이터 표준이 TV Anytime Forum의 주도로 TVA (TV Anytime)란 명칭으로 개발되고 있다[2]. 두 표준 모두 스키마를 기술하는 도구로 XML Schema[3]를 사용하고 있으며, TVA는 MPEG-7에서 정의한 많은 데이터타입을 그대로 사용하고 있다.

두 표준에서 공통적으로 다루는 가장 중요한 문제중의 하나가 브라우징 및 검색을 위한 비디오의 계층구조 기술이다. 이미 두 표준에서 이에 대한 내용은 각자의 방식으로 정의되었다. 그러나 앞으로 많은 비디오의 내용이 MPEG-7이나 TVA 스키마로 기술될 경우, 이들 간의 상호 변환이 요구될 것이다. 예를 들어, 케이블 TV에서 보고싶은 영화를 자신의 PVR에 예약 녹화한 후, 웹을 통해 그 영화의 내용을 기술한 MPEG-7 메타데이터를 찾아낸 사용자라면 그 메타데이터를 TVA로 변환하여 PVR에 옮기고 녹화된 영화를 브라우징할 수 있다.

MPEG-7과 TVA 모두 비디오 콘텐츠의 내용과 관련 정보를

계층구조로 기술하고 있는데, 표현 방식 및 개념에 있어서 많은 유사성이 있으나 동시에 의미상의 차이도 존재한다. 본 논문에서는 이러한 점을 고려하여 MPEG-7 혹은 TVA 스키마에 따라 비디오의 내용을 기술한 메타데이터를 상대방의 스키마에 따라 변환시키는 방법을 제안한다.

2. MPEG-7과 TA Anytime의 계층 구조 표현

2.1 세그먼트와 그룹

MPEG-7과 TVA에서는 비디오를 세그먼트(segment)와 세그먼트 그룹(segment group)으로 구분하여 기술하고 있다. 여기서 세그먼트란 비디오의 연속된 특정 부분을 말하고, 그룹은 내용상 연관된 세그먼트 또는 그룹을 묶은 것으로 하나 이상의 세그먼트를 묶은 것, 하나 이상의 그룹을 묶은 것, 또는 하나 이상의 세그먼트와 그룹을 묶은 것이 될 수도 있다. 그림 1은 5개의 세그먼트로 이루어진 비디오에 대한 논리적 세그먼트 계층구조의 예를 보여주고 있다. 그림에서 단말노드인 S1, S2, S3, S4, S5는 세그먼트이고 S1, S2를 묶은 S6과, S4와 S5를 묶은 S7, 그리고 세그먼트 S5와 그룹 노드인 S6, S7을 묶은 S8은 그룹을 나타낸다.

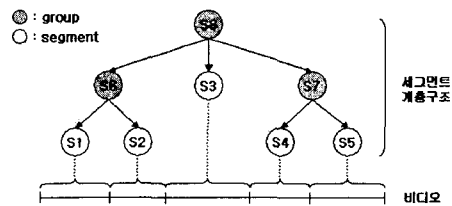


그림 1. 비디오의 논리적 계층 구조의 예

2.2 계층 구조의 표현

MPEG-7의 Summarization DS와 TVA의 Segmentation Information Table에서는 세그먼트와 그룹을 표 1과 같은 DS(Description Scheme)을 사용하여 기술한다. 그림 2는 표 1의 DS 간의 관계를 UML 다이어그램[4]으로 서술한 것이다.

표 1. MPEG-7과 TVA에서의 그룹과 세그먼트에 대한 DS

MPEG-7	그룹	SummarySegmentGroup
	세그먼트	SummarySegment
TVA	그룹	SegmentGroupInformation
	세그먼트	SegmentInformation

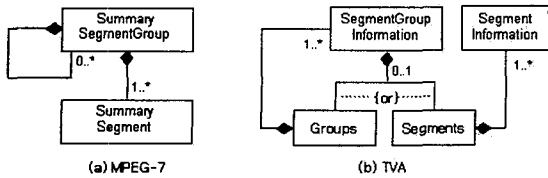
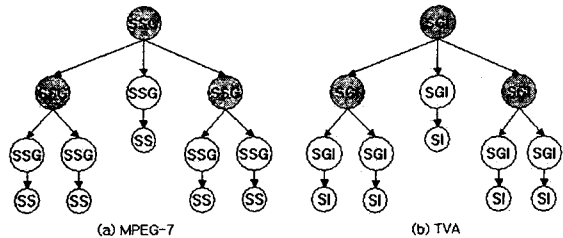


그림 2. 그룹과 세그먼트에 관한 UML 다이어그램

그림 1의 세그먼트 계층구조에서 그룹 S8의 자식 노드들을 서술하면 그룹 S6, 세그먼트 S3, 그룹 S7의 순으로 서술해야 한다. 하지만 MPEG-7에서는 SummarySegmentGroup 내에 속한 그룹과 세그먼트를 같은 형제노드로 기술이 가능하지만 그들의 순서를 유지할 수 없고, TVA의 SegmentGroupInformation에서는 세그먼트 혹은 그룹만 기술해야 한다. 즉, 세그먼트와 그룹의 혼합 서술이 불가능하다.

그림 1의 논리적 계층구조를 MPEG-7 혹은 TVA 스키마로 표현하려면, 모두의 스키마 정의에 위배되지 않으면서 메타데이터의 상호변환이 효율적으로 이루어지게 하는 표현이 필요하다. 본 논문에서는 이를 위해 그림 1의 논리적 계층구조에서 단말에 위치하는 세그먼트를 그룹을 기술하는 DS로 표현하고, 자식 노드로 세그먼트를 갖게 한다. 그림 3은 이 방법으로 그림 1의 논리적 계층구조를 MPEG-7과 TVA로 기술한 것이다. 결과적으로 계층구조의 레벨이 하나 증가하게 되나, 이렇게 함으로써 MPEG-7과 TVA 모두 같은 논리적 구조로 비디오의 계층구조를 서술할 수 있을 뿐만 아니라 자식 노드들의 순서도 유지가 가능하다. 또한 이 기술 방법은 MPEG-7 혹은 TVA용 비디오 브라우저 개발시 계층구조의 순회 논리가 매우 단순해지는 이점이 있다.



SSG : SummarySegmentGroup SGI : SegmentGroupInformation
 SS : SummarySegment SI : SegmentInformation
 그림 3. MPEG-7과 TVA 스키마에 의한 계층구조의 표현

3. 상호변환의 문제점 및 해결 방안

3.1 계층구조의 표현 방식

비디오 계층구조를 기술할 때 크게 두 가지의 표현방식이 있다. 각 그룹 노드의 자식 노드(그룹 혹은 세그먼트)를 그 노드의 내부에 중첩시켜 기술하는 중첩 표현법(nested description), 다른 곳에서 자식 노드를 기술하고 그룹 노드에서는 이미 기술된 자식노드를 참조하는 참조 표현법(reference description)으로 계층구조를 기술할 수 있다. MPEG-7은 이 두 가지 표현방식이 모두 가능하도록 스키마가 정의되어 있지만, TVA 스키마에서는 참조 표현법만 가능하다. 이러한 표현 방식의 차이로 MPEG-7과 TVA의 메타데이터 변환시 id의 처리가 중요하다. MPEG-7과 TVA의 참조 표현법에서 자식 노드의 참조는 자식 노드의 id를 통하도록 설계되었

다. 즉 참조하고자 하는 노드에 그 문서상에서 유일한 id값을 할당하고 그 id를 통해 참조하도록 한다. 따라서 중첩 표현법에 의한 MPEG-7 메타데이터를 TVA 메타데이터로 변환할 경우 id가 없는 그룹 또는 세그먼트에 대해서는 새로운 id를 할당해야 한다. 역으로 TVA 메타데이터를 참조 표현법을 사용한 MPEG-7으로 변환할 경우에는 id를 그대로 사용할 수 있다. MPEG-7 DS에서 id는 optional로 정의되어 있으므로 id를 그대로 사용하는 것은 변환시의 선택사항이다.

그림 4는 그림 1의 그룹 S6을 중첩 표현법을 사용한 MPEG-7 메타데이터를 참조 표현법의 TVA 메타데이터로 변환한 예이다. 그림에서 MPEG-7의 경우 그룹 노드에만 id가 할당되어 있고 중첩된 세그먼트 노드에는 할당되어 있지 않다. TVA로 변환시 이 세그먼트 노드가 그룹 노드에서 참조되어야 하므로 세그먼트 노드에 새로운 id인 S1_s과 S2_s라는 id를 할당하였음을 알 수 있다.

```
<SummarySegmentGroup id="S6">
  <SummarySegment>...</SummarySegment>
  <SummarySegmentGroup id="S6">
    <SummarySegment>...</SummarySegment>
  </SummarySegmentGroup>
  <SummarySegmentGroup id="S6">
    <SummarySegment>...</SummarySegment>
  </SummarySegmentGroup>
</SummarySegmentGroup>
```

(a) 중첩 표현법을 사용한 MPEG-7 메타데이터

```
<SegmentList>
  <SegmentInformation segmentId="S6_s1"/>
  <SegmentInformation segmentId="S6_s2"/>
</SegmentList>
<SegmentGroupList>
  <SegmentGroupInformation groupId="S6">
    <Groups refList="S6_s1 S6_s2"/>
  </SegmentGroupInformation>
  <SegmentGroupInformation groupId="S6">
    <Segments refList="S6_s1 S6_s2"/>
  </SegmentGroupInformation>
  <SegmentGroupInformation groupId="S6">
    <Segments refList="S6_s1 S6_s2"/>
  </SegmentGroupInformation>
</SegmentGroupList>
```

(b) 변환된 TVA 메타데이터 (참조 표현법)

그림 4. 그룹 메타데이터의 변환

3.2 대표화면의 표현

각 그룹 노드는 여러 개의 자식 노드 즉, 세그먼트 또는 그룹으로 구성되어 있다. 계층구조상의 그룹들을 재생하지 않고 대표화면(key frame)만으로 브라우저하기 위해 그룹 노드에 대표화면을 기술할 수 있다. TVA에서는 그룹 노드에 KeyFrameLocator DS가 존재하지만, MPEG-7에서는 대표화면을 포함한 대표 세그먼트의 개념으로 SummarySegment DS를 이용한다. 따라서 대표 세그먼트가 정의된 MPEG-7 메타데이터를 TVA로 변환하면, 대표 세그먼트의 구간 정보는 유실되고 대표화면만 표현할 수 있다. 반대로 대표화면이 정의된 TVA 메타데이터를 MPEG-7으로 변환하면, SummarySegment DS를 이용해 대표화면 정보를 표현할 수 있다.

그림 5는 대표화면 정보를 포함한 MPEG-7과 TVA 스키마의 계층구조를 나타낸 것이다. MPEG-7에서는 SummarySegment DS의 KeyFrame DS에서 대표화면을 기술하며, TVA에서는 KeyFrameLocator DS에서 기술한다. 참고로 두 스키마 모두 단말에 해당하는 세그먼트 노드에 그 세그먼트의 대표화면에 대한 정보가 저장될 수 있다. 따라서 그룹에 대한 대표화면 정보만 추가로 기술하면 전체 논리적 계층구조에 대해 대표화면만으로 브라

우장하는 것이 가능하다. 그림 6은 그림 4의 메타데이터에 대표 화면 정보를 추가하여 기술한 예를 보여주고 있다.

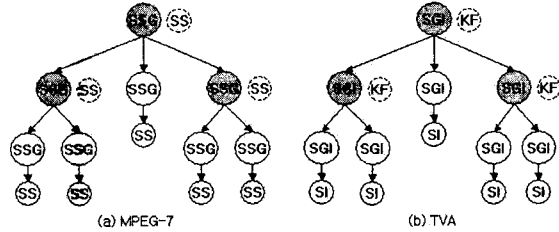


그림 5. 대표화면 정보를 포함한 트리 구조

```
<SummarySegmentGroup id="1" >
  <SummarySegment>
    <KeyFrame>
      <MediaTimePoint>T00:04:12</MediaTimePoint>
    </KeyFrame>
  </SummarySegment>
  <SummarySegmentGroup id="2" >
    <SummarySegment>...</SummarySegment>
  </SummarySegmentGroup>
  ...
</SummarySegmentGroup>
```

(a) 중첩 표현법을 사용한 MPEG-7 메타데이터

```
<SegmentList>...</SegmentList>
<SegmentGroupList>
  <SegmentGroupInformation groupId="1" >
    <Groups refList="1" >/>
    <KeyFrameLocator>
      <MediaTimePoint>T00:04:12</MediaTimePoint>
    </KeyFrameLocator>
  </SegmentGroupInformation>
  ...
</SegmentGroupList>
```

(b) 변환된 TVA 메타데이터 (참조표현법)
그림 6. 대표화면 정보의 기술

3.3 비디오의 위치 정보

두 표준은 비디오의 위치 정보 표현에 있어서 다른 접근 방법을 취하고 있다. MPEG-7은 URI를 사용하지만, TVA는 CRID(Content Referencing ID)를 이용해 기술하고 있다. CRID는 여러 방송 채널 혹은 set-top box의 하드디스크에 녹화된 비디오를 참조하기 위해 사용하는 논리적 식별자로서, CRID resolution 과정을 통해 실제 비디오의 위치를 가리키는 식별자로 전환된다[5]. URI로 표현된 위치정보를 CRID로 변환하거나, 그 역으로 변환하기 위해서는 CRID resolution 과정이 필요하다. 본 논문에서는 메타데이터 파싱시 발생할 validation 에러를 피하기 위해 단순히 URI와 CRID의 변환시 정의된 포맷에 맞게 변환하였다.

4. MPEG-7과 TV Anytime의 상호변환 알고리즘

4.1 MPEG-7 → TV Anytime

```
rootnode = root of SummarySegmentGroup;
Convert(rootnode);
Convert(node) {
  create SegmentGroupInformation, and add id to it
  append to SegmentGroupList;
  for childnodes of pnode {
    if(childnode.nodeName == "SummarySegmentGroup")
      Convert(childnode);
    else {
```

```
if(childnode == node to describe keyframe) {
  create KeyFrameLocator;
  add to SegmentGroupInformation;
} else { /* leaf node of SummarySegmentGroup */
  create SegmentInformation, and add id to it;
  append to SegmentList;
}
}
```

4.2 TV Anytime → MPEG-7

```
Select nested or reference description;
Find top level node(SegmentGroupInformation) in childnodes of
SegmentGroupList;
Convert(top level node);
Convert(node) {
  if( name of node == "SegmentGroupInformation") {
    create SummarySegmentGroup and add;
    create SummarySegment and add; /* to describe KeyFrame */
    idlist = ids of nodes referencing childnode
    for( i = 0; i < idlist.length; i++) {
      childnode = GetNode(idlist[i]);
      Convert(childnode);
    }
  } else { /* if name of node is "SegmentInformation" */
    create SummarSegment and add;
  }
}
```

5. 결론

MPEG-7은 웹 환경에서 비디오를 포함한 멀티미디어 콘텐츠의 내용을 기술하기 위해 개발되었고, TVA는 방송 환경에서 EPG(Electronic Program Guide)와 비디오 내용 기술을 위해 개발되고 있는 스키마이다. 웹과 방송 환경에서만 필요한 정보가 있으며 상호 변환시 모든 데이터를 일대일로 대응시켜 변환하기는 불가능하다. 따라서 변환 과정에서 많은 데이터가 누락될 수 있다. 하지만 비디오의 내용 기술 부분으로 국한하면 양쪽 스키마 모두 논리적으로는 유사한 정보를 기술하고 있으므로 공통적인 정보는 변환이 가능하다.

본 논문에서는 MPEG-7과 TVA의 스키마를 분석하여 그들의 유사점을 밝히고, 중첩 표현법과 참조 표현법으로의 상호 변환시 야기되는 id 처리 문제, 대표화면 정보의 처리, 위치 정보의 처리 등에 관한 방법을 제안하였다. 본 논문의 메타데이터 변환은 계층구조의 단말도 그룹으로 기술되었다는 가정하에 개발되었다. 이러한 제약이 없는 메타데이터에 대한 변환은 제안한 알고리즘을 부분적으로 수정하면 가능할 것이다. 앞으로 방송과 통신이 융합되는 환경에서 MPEG-7과 TVA 메타데이터가 보편화되고, 이로 인해 상호 변환 요구가 증대될 것이라 예상된다. 본 논문에서 제안한 메타데이터 상호 변환은 이와 같은 환경에서 유용하게 사용될 수 있을 것이다.

참고문헌

1. ISO/IEC 15938-5, "MPEG-7 Part 5: Multimedia description schemes (MDS)", Oct. 2001.
2. TV Anytime Forum, "TV-Anytime Specification Series S-3: on Metadata", Version 1.2, June 2002
3. XML Schema, <http://www.w3c.org/XML/Schema>
4. UML, <http://www.uml.org/>
5. TV Anytime Forum, "TV-Anytime Specification Series S-4: Content Referencing", Version 1.2, Feb. 2002