

MPEG-7을 기반으로 한 뉴스 동영상 스키마와 요약 생성 방법의 설계 및 구현

심진선⁰ 정진국¹ 남종호² 김경수³ 하명환⁴ 정병희⁵
서강대학교 컴퓨터학과, *KBS 기술 연구소
iceshim@hanmail.net⁰

A Design and Implementation of Scheme and Summary Generation Mechanism for News Video based on MPEG-7 MDS

Jin-Sun Shim⁰ Jin-Guk Jeong¹ Jong-Ho Nang² Gyung-Su Kim³ Myung-Hwan Ha⁴ Byung-Hui Jung⁵
Department of Computer Science, Sogang University, *KBS Technical Research Institute

요 약

최근 디지털 동영상의 사용이 증가하면서 자동으로 구조를 분석하는 기술이 필요하게 되었다. 특히 뉴스 동영상은 다른 동영상에 비해 그 구조가 비교적 정형화 되어 있다는 특징 때문에 많은 연구에서 이용되어졌다. 뉴스 동영상을 이용하는 이전의 연구에서 문제점으로 제시될 수 있는 사항은 서로 다른 자료 구조와 시스템 구조로 인한 호환성의 부족을 들 수 있다. 본 논문에서는 이와 같은 호환성 부족을 해결하기 위해 멀티미디어 데이터를 기술하는 표준인 MPEG-7을 기반으로 한 뉴스 동영상 스키마를 제안하였다. 특히 전체 뉴스를 보지 않고 효율적으로 뉴스 동영상을 이해할 수 있도록 요약하는 방법을 제시하였으며 MPEG-7의 MDS를 이용하여 기술하였다. 본 논문에서 제안한 방법은 디지털 비디오 라이브러리와 같은 응용 분야에서 유용하게 이용될 수 있을 것이다.

1. 서 론

인터넷 및 컴퓨터 기반 기술이 발달함에 따라 멀티미디어를 이용한 다양한 서비스가 가능하게 되었다. 뉴스 아카이브 시스템 또한 이러한 이유로 인하여 최근 연구가 많이 되고 있는 분야이다. 뉴스 아카이브 시스템에서는 정형화된 뉴스의 특징을 이용하여 뉴스를 인덱싱하고 추상화 함으로써 사용자의 요구에 맞는 기사 검색을 가능하게 한다. 하지만 이러한 시스템들은 각종 자료들을 표현하고 저장하는 방법들이 다르기 때문에 서로 호환성이 없다는 문제점을 가지고 있다. 이와 같은 이유로 인해 나온 것이 동영상의 내용 정보들을 기술하기 위한 표준인 MPEG-7 이다.

본 논문에서는 MPEG-7 MDS[1]를 기반으로 하여 뉴스 동영상을 위한 스키마를 설계하고 이러한 스키마에 맞는 데이터를 생성하고 요약할 수 있는 도구를 구현한다. 이를 위하여 요구 사항을 분석한 결과 뉴스는 사용자들이 좀 더 다양한 질의를 통해 원하는 정보를 찾을 수 있도록 특정 단위로 인덱싱되고 요약되어야 함을 알 수 있었다. 이런 조건들을 만족하기 위해 본 논문에서는 MPEG-7 MDS의 VideoSegment, TextAnnotation, MediaInformation, Classification, SequentialSummary, HierarchicalSummary 등을 이용하여 스키마를 생성하고 요약방법을 제안하였다. 본

논문에서 제안한 방법은 디지털 비디오 라이브러리와 같은 응용 분야에서 유용하게 이용될 수 있을 것이다.

2. 뉴스 스키마 설계와 요약 생성 방법

2.1 요구 사항 분석 및 기본 설계 조건

뉴스는 특정 주제에 대해서 하나의 기사를 구성하게 되며 기사는 다시 샷으로 나뉘어진다. 기사는 기사가 다루고 있는 분야에 따라서 정치, 경제, 사회, 국제, 문화, 과학, 스포츠, 날씨로 나뉘고 샷은 보도의 유형에 따라서 앵커, 그래픽 기자, 인터뷰, 회견/연설, 일반적으로 분류[2]된다. 이러한 내용을 기반으로 하여 사용자에게 제공할 수 있는 요약 파일을 생성하게 되는데 사용자의 관심분야가 다양하기 때문에 모든 사용자를 만족시킬 수 있는 요약 파일을 미리 생성해 놓는 것은 불가능하다. 따라서 사용자의 요구사항을 반영할 수 있는 요약 시스템을 통해서 동적으로 요약을 생성할 수 있도록 하여야 할 것이다.

2.2 뉴스 동영상 스키마 설계

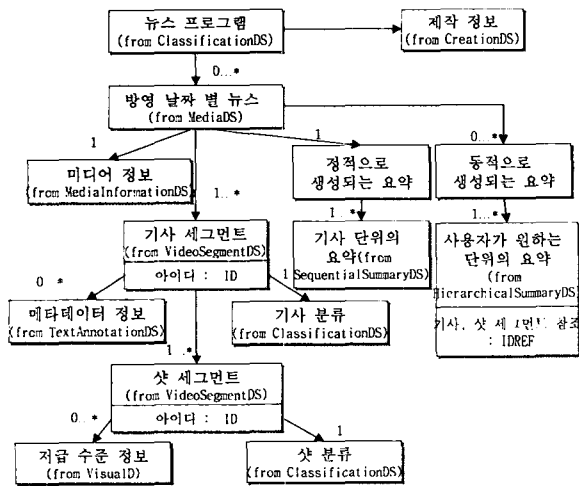
MPEG-7 MDS를 이용하여 동영상을 기술할 때 내용의 구조를 표현하기 위하여 사용되는 스키마는 SegmentDS이므로 본 논문에서 뉴스의 계층적인 구조를 표현하기 위해

Segment DS를 이용하였다. 세그먼트 단위인 기사와 샷은 시간을 기준(TemporalDecomposition)으로 구분하고 기사의 제목과 분류를 표현하기 위해 <그림 1>과 같이 FreeTextAnnotation 과 type 속성을 이용하였다. 또한 대표 프레임을 간단하게 기술하기 위해서 KeyFrame 타입을 세그먼트에 추가하였고 스크립트를 기술하기 위해서 Script 타입을 추가하였다. 뉴스와 세그먼트, 요약에 대한 전체적인 관계는 <그림 2>와 같다.

```

<VideoSegment id="Article1">
  <KeyFrame>
    <MediaUri>pkbs_20000309_1302.jpg</MediaUri>
  </KeyFrame>
  <TextAnnotation>
    <Script>
      ©김종진 앵커:정부의 대북정책기조가 달라집니다. ...
    </Script>
  </TextAnnotation>
  <TextAnnotation type="NewsClassificationCS:News:Section">
    <FreeTextAnnotation>정치</FreeTextAnnotation>
  </TextAnnotation>
  <!--중략-->
</VideoSegment>
    
```

<그림 1> 기사의 기술 예



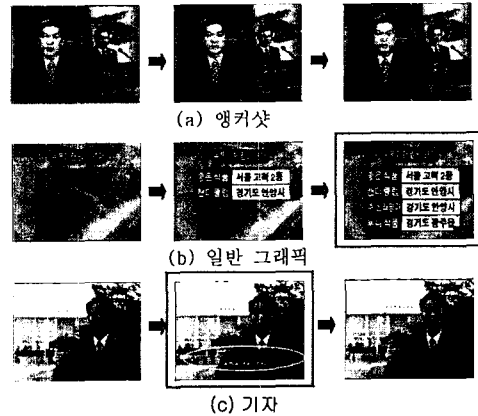
<그림 2> 뉴스 동영상 스키마 다이어그램

2.3 요약 생성 방법

2.3.1 샷의 대표 프레임 선택

뉴스는 샷의 타입별로 서로 다른 특징을 보이므로 본 논문에서는 샷의 타입을 이용하여 대표 프레임을 선택하였다.

앵커샷과 일반샷은 특별히 중요한 프레임을 찾을 수 없으나 처음 프레임과 마지막 프레임은 점진적인 변화로 인해 두 이미지가 겹칠 수 있으므로 제외하였다. 그래픽 샷은 <그림 3>의 (b)에서 보는 바와 같이 끝 부분의 프레임에 표현하고자 하는 모든 정보를 포함하므로 마지막 프레임을 선택하였다. 마지막으로 기자, 인터뷰, 회견/연설 샷은 <그림 3>의 (c)와 같이 중간 프레임에 부가정보를 자막으로 보여주므로 중간 프레임을 선택하였다.



<그림 3> 샷 타입별 프레임 변화

2.3.2 정적으로 생성되는 요약

기사에서 앵커의 말은 그 기사를 요약한 것이라 볼 수 있으므로 앵커 샷만을 연결해서 하나의 새로운 동영상을 만드는 것이 가장 쉬운 요약 방법이다. 그러나 앵커 샷은 거의 정지 영상에 가까워서 기사마다 거의 비슷하기 때문에 비디오 정보를 전달하는 데 부족하다. 본 논문에서는 비디오 정보와 오디오 정보를 순차적으로 모두 표현하기 위해서 SequentialSummary DS를 이용하였다. 먼저 동영상을 비디오와 오디오로 분리한 후 비디오에서는 비디오 클립이 아닌 대표 프레임들을 추출하고 오디오에서는 앵커 부분을 추출한 후 오디오가 재생되는 동안 대표 프레임들을 슬라이드 형태로 보여주게 된다. 기사의 대표 프레임들은 앵커 샷의 길이(초단위)를 1.5초로 나눈 개수 만큼 샷의 대표 프레임들로 선택하고 FrameProperty D를 이용하여 기술하였다. 앵커의 오디오 부분은 앵커 샷의 시간 정보를 이용하여 오디오에서의 위치를 추정하고 SoundProperty D를 이용하여 기술하였다. 또한 기사의 제목은 TextProperty D를 이용하여 기술한 후 요약에 포함시켰다. 이러한 요약은 각각의 기사에 대해서 개별적으로 만들어서 기사 단위로 요약을 볼 수 있도록 하였다. 이와 같은 방법으로 생성되는 요약은 <그림 4>와 같이 기술되게 된다.

```
<Summary id="기사1" xsi:type="SequentialSummaryType">
  <AudioSummaryLocator>
    <MediaUri>kbs_20000309.mp3</MediaUri>
    <BytePosition offset="191" length="6987"/>
  </AudioSummaryLocator>
  <FrameProperty>
    <ImageLocator>
      <MediaUri>kbs_20000309_90.jpg</MediaUri>
    </ImageLocator>
  </FrameProperty>
  <!-- 대표 프레임들 종락 -->
  <TextProperty>
    <FreeTextAnnotation>베를린</FreeTextAnnotation>
  </TextProperty>
</Summary>
<Summary id="기사2" xsi:type="SequentialSummaryType">
  <!-- 종락 -->
</Summary>
```

<그림 4> 정적으로 생성되는 요약의 기술 예

2.3.3 동적으로 생성되는 요약

MPEG-7 MDS의 HierarchicalSummary DS는 다양한 요약 시간과 조건에 맞는 계층적인 요약을 기술하기 위한 스키마[1] 이므로 동적으로 생성되는 요약을 위해서 본 논문에서는 HierarchicalSummary DS를 이용하였다. 우선 뉴스의 기사 타입과 샷 타입은 뉴스를 구분하는 요소이기 때문에 keyThemes로 사용하여 해당 세그먼트를 요약에 포함시킨다. 사용자는 주제 리스트에서 복수개의 항목을 선택할 수 있도록 했으며 사용자 인터페이스를 통해서 결과를 수정할 수 있도록 하였다. 또한 일반적으로 뉴스는 중요한 사건을 먼저 보도[3]하므로 요약 시간을 제한한 경우는 보도 순서대로 요약에 포함시켰다. <그림 5>는 그래픽을 선택하고 사용자가 수정을 한 후의 요약 정보를 기술한 예이다.

```
<Summary xsi:type="HierarchicalSummaryType"
  components="keyAVClips keyThemes" hierarchy="independent">
  <!-- 그래픽을 선택한 경우의 요약 -->
  <HighlightSummary id="Summary01" themelds="그래픽">
    <HighlightSegment>
      <keyAVClip>
        <MediaTime>
          <MediaTimePoint>126.292959</MediaTimePoint>
          <MediaDuration>7.640974</MediaDuration>
        </MediaTime>
      </keyAVClip>
    </HighlightSegment>
    <HighlightChild level="0">
      <!-- 세부 내용 종락 -->
    </HighlightChild>
    <!-- More HighlightSegment and HighlightChild -->
  </HighlightSummary>
</Summary>
```

<그림 5> 동적으로 생성되는 요약의 기술 예

3. 시스템 구현 및 분석

제안한 뉴스 동영상 스키마에 맞는 뉴스 아카이브 시스템을 구현하기 위해 본 논문에서는 XML Spy v4.0, C 언어 그리고 MSXML 4.0 파서를 이용하였다.

본 논문에서 제안한 정적으로 생성되는 요약은 일반적인 동영상 요약보다 비주얼한 내용 전달에 더 효율적이었고 데이터 양이 훨씬 적었다. 이러한 요약은 웹상이나 대역폭이 작은 모바일 환경에서 특히 더 효율적으로 사용될 수 있을 것이다. 동적으로 생성되는 요약은 기사의 섹션 타입과 샷의 보도 타입, 요약 시간에 의해서 생성이 되었다. 따라서 뉴스의 형식적인 측면에서는 사용자의 요구를 충분히 반영한다고 볼 수 있었다. 하지만 실제 고급 수준 내용 정보를 이용하지 않았기 때문에 내용적인 측면에서는 충분히 반영하지 못했다.

4. 결론 및 향후 연구 방향

본 논문에서는 보관과 호환성을 위해 MPEG-7을 기반으로 한 뉴스 동영상 스키마를 제안하였다. 또한 제안한 스키마에 맞는 문서를 생성하고 브라우징할 수 있는 도구를 구현하여 MPEG-7이 뉴스 동영상에 관한 내용을 풍부하게 기술할 수 있음을 살펴보았다. 이런 연구들은 향후에 서로 다른 시스템 간에서도 자료를 교환할 수 있고 내용적으로도 풍부한 기술 문서를 생성할 것이며 사용자의 요구를 더 충족할 수 있는 시스템 개발에 이용될 수 있을 것이다.

5. 참고문헌

[1] P. Beek, B. Ana, *Text of 15938-5 FCD Information Technology Multimedia Content Description Interface Part 5 Multimedia Description Schemes*, ISO/IEC/JTC1/SC29/WG11/ N3966, March 2001.
 [2] 한근주, *샷의 타입을 이용한 뉴스 아카이브 시스템의 설계 및 구현*, 서강대학교 석사 학위 논문, 1999.
 [3] B. Merialdo, K. Lee, et al, "Automatic Construction of Personalized TV News Programs," Proc. of the Seventh ACM International Conference on Multimedia, pp. 323-341, 1999.
 [4] H. Zhang, Y. Gong, S. Smoliar, "Automatic Parsing of News Video," *Proc. of Multimedia Computing and Systems*, pp. 45-54, 1994.
 [5] G. Ahanger, T. Little, "Data Semantics for Improving Retrieval Performance of Digital News Video Systems," *Proc. of IEEE Transactions on Knowledge and Data Engineering*, Vol. 13, pp. 352-360, 2001.