

폐쇄자막을 기반한 자막기반 동영상 검색 시스템

김효진, 황인정, 이은주, 이용혁*, 민홍기

한국산업기술대학교 전자공학과*, 인천대학교 공과대학 정보통신공학과

Video retrieval system based on closed caption

Hyo Jin Kim, Ein Jung Hwang, Eun Joo Lee, Tae Sung Hur*, Hong Ki Min

Dept of Computer and Information System , In-ha technical college* ,

Dept of Information and Telecommunication Engineering , University of Incheon,

요약

동영상 데이터와 같은 영상자료는 많은 활용분야를 가지고 있지만, 비정형성과 구조의 복잡성으로 인하여 손쉽게 검색을 하거나 재사용하는 데 있어서 많은 어려움을 가지고 있다. 본 논문에서는 동영상과 같은 멀티미디어 데이터를 구조적이고 체계화된 형태로 기술한 언어인 SMIL 과 SAMI를 사용하여 동영상과 동기화 되어있는 폐쇄자막을 기반으로 동영상을 검색하는 시스템을 구현하였다. 사용자의 검색어를 입력받아 자막파일에서 검색어를 가지고 있는 문자열에서 시간 요소를 추출하여 해당 하는 동영상 화면을 보여주는 구조를 가지고 있다.

abstract

Even if the video data is utilized for a lot of field, its very difficult to reuse and search easily because of its atypical(unfixed form) and complicated structure. In this study, We presented the video retrieval system which is based on the synchronized closed caption and video, SMIL and SAMI languages which are described to structured and systematic form like multimedia data These have next structure; At first, a key word is inputted by user, then time stamp would be sampling from the string which has a key word in the caption file. To the result, the screen shows an appropriate video frame.

서론

멀티미디어와 인터넷이 생활의 한 부분이 되어 가면서 일반적인 평면데이터 보다는 화상, 소리, 동영상등의 멀티미디어 정보, 특히 동영상 정보에 대한 수요가 급증함에 따라 사용자들에게는 멀티미디어 데이터를 효과적으로 운용할 수 있는 기술의 필요성이 증가하게 되었다. 방대한 양의 분산된 멀티미디어 데이터를 처리 할 수 있는 색인 및 검색도구의 요구가 점차 늘고 있다.

특히, 동영상 데이터를 사용하는 VOD, 원격

교육, 전자도서관등에서의 동영상 데이터 검색 기술은 필수적인 요소기술이다. 이러한, 동영상검색 기법에는 주석기반 검색방법과 내용기반 검색의 2가지 방법이 있다. 주석기반 검색은 동영상에 대한 의미 정보를 미리 주석으로 작성하여 저장한 후에 사용자가 원하는 동영상을 기술한 문자 질의와의 비교를 통하여 유사한 동영상을 찾는 방법이다.

내용기반 검색방법은 동영상 검색을 위해서 가장 많이 연구되고 있고, 이는 동영상으로부터 특징적인 내용을 표현하는 특정 데이터를 자동으로 추출하여 저장한 후에, 사용자가 원하는 비디오에 포함되는 이미지나 비디오 클립등의 질의로 입력하면 그로부터 추출된 특징 데이터와의 비교를 통하여 유사한 비디오를 찾는 검색이다.

그러나, 주석기반의 검색은 정확한 검색을 위해서 수작업을 통하여 정보를 입력으로 입력자의 주관에 따르기 때문에 객관적인 판단이 힘들다. 내용기반 검색은 클립등을 통한 키프레임에 한정되어 사용자가 원하는 장면을 입력하는데에 어려움이다. 있다. 본 논문에서는 폐쇄자막에 기반한 검색방법을 사용하여 검색효율을 증가시켰다.

자막 방송이란 본 방송을 내보내면서 동시에 화면 하단 혹은 상단에 자막을 내보내는 방송을 말한다. 이미 공중파에선 한글 자막방송을 실시하고 있고, 인터넷 방송 및 멀티미디어 데이터에도 많은 부분 채택이 되고 있다. 특히, 마이크로소프트의 SAMI(Synchronized Accessible Media Interchange)를 이용하여 폐쇄자막을 표현한 동영상들이 많이 사용되어지고 있다. 이뿐만 아니

라 멀티미디어 동기화 언어 표준으로 제정된 SMIL (Synchronized Multimedia Integration Language)을 사용하여 동영상과 동기화가 이루어진 문서를 쉽게 구현할 수 있다. SAMI(Synchronized Accessible Media Interchange)로 구현된 폐쇄자막(Closed caption)을 토대로 내용기반 검색을 하여, 사용자가 원하는 장면을 찾아주는 검색 시스템을 구현하였다.

I. 멀티미디어 동기화 언어

1. 폐쇄자막을 지원하는 언어의 종류

Microsoft사에서 제공하는 SAMI는 단일문서에 동영상 데이터 표현과 동기화된 자막편집에 최적화 되어 있는 형식의 언어이다. HTML과 비슷한 마크업 언어의 형식을 취하고 있으며, 윈도우 미디어 플레이어로 재생이 가능하고, 공중파에서 제공하는 폐쇄자막의 표준을 수용하고 있다. 특히, 인터넷의 대부분의 영상파일의 SAMI 형식을 사용하고 있으며 점차 사용이 늘어가고 있는 추세이다.

폐쇄자막을 표현할 수 있는 또 다른 언어는 SMIL이다. SMIL은 XML (eXtensibel Markup Language)의 형태를 그대로 상속하여 어떻게 데이터가 보일 것인지 보다는 어떻게 데이터를 구조화시킬 것인지에 초점을 맞추고 있다. SMIL을 이용하여 오디오, 비디오, 이미지, 텍스트 등의 멀티미디어 객체를 동기화 하는데 필요한 시간적, 공간적 관계를 기술하는데 필요한 마크업 태그를 만들 수 있다. SMIL은 동영상과 자막파일 뿐만 아니라 여러 가지 형태의 멀티미디어 데이터를 모두 동기화 시켜 표현 할 수 있다.

2. 멀티미디어 동기화 언어의 구조

멀티미디어 동기화 언어인 SMIL과 SAMI는 모두 HTML과 같은 마크업 언어형식의 언어이다. SMIL과 SAMI 파일은 동영상 데이터와 동일한 이름을 가진 텍스트 파일의 형태로 존재한다.

SAMI 파일의 경우에는 자막데이터를 자기 자신이 가지고 있지만, SMIL 파일의 경우에는 또 다른 자막 파일을 로드 하는 형태를 가지고 있다.

이들 동기화 언어들은 동영상 데이터와 시간축을 기준으로 재생되어진 시간을 확인하여 동영상

데이터와 함께 자막을 제공한다. 이를 위한 멀티미디어 동기화 언어의 구조는 다음과 같다.

첫 번째로, 공중파 폐쇄자막의 표준을 수용하고 있는 SAMI는 각각의 동영상 파일과 자막파일(*.smi,*.sami)로 구성되어져 있다. 이는 동기화를 위한 텍스트 파일을 따로 사용하지 않고, 텍스트로 된 자막파일 내에 동기화 태그가 포함된 구조를 가지고 있다. HTML형식을 대부분 수용하고 있으며, 폐쇄자막에서 보여주는 여러 가지의 보여주기 형식을 가지고 있다. SAMI는 [그림 1]과 같이 <SYNC> 태그의 "strat"라는 속성을 사용하여 현재 미디어 파일의 시작시간을 ms단위로 나타낸다.

```
<SAMI>
<HEAD>
<Title> CC Video Template</Title>
<STYLE TYPE="text/css">
</HEAD>
<BODY>
<SYNC Start=3000><P Class=KRCC> little lamb.
<SYNC Start=6000><P Class=KRCC> little lamb.
<SYNC Start=9000><P Class=KRCC>Mary had a little lamb.
<SYNC Start=12000><P Class=KRCC>whose fleece was white as snow.
</BODY>
</SAMI>
```

[그림 1] SAMI file

두 번째로, SMIL은 리얼플레이어사의 동영상 데이터(*.rm,*.ra)와 자막 데이터(*.rt) 그리고 이들을 동기화한 파일(*.smi,*.smil)로 구성되어 있다. 그러나, 리얼플레이어사의 한정된 지원만을 제공하는 것은 아니다. 특히, rt 파일은 리얼플레이어사의 자체 표준으로 지원하는 text file로 다양한 효과를 지원하지만, 본 논문에서는 일반 텍스트를 사용한 자막파일을 기준으로 다룬다.

시간적인 동기화 정보의 표현은 <seq>와 <par>블록의 사용을 통해 간편한 표현기법을 사용하고 있다. <seq> 태그와 <par>태그를 사용한 미디어간의 동기화 표현을 나타낸다. <seq> 태그는 순차적인 순서로 미디어를 표현하고, <par> 태그는 동시에 2가지 미디어가 재생되는 것을 표현한다.

자막파일과 동기화시키는 SMIL 파일을 살펴보면 [그림 2]와 같다. "src"라는 HTML과 같은 형식의 프로퍼티를 사용하여 각기 다른 미디어 파일을 지칭하고 있다. <layout>태그에 지정한 각각의 레이어에 동영상 데이터와 텍스트 데이터를 로드하여 각각의 파일에 명시된 시간과 SMIL파일에 있는 <seq>,<par> 태그를 기준으로 동기화 하여 화면에 출력한다.

```

<smil>
<head>
<meta name="title" content="CC Video Template"/>
<meta name="author" content="RealNetworks, Inc."/>
<meta name="copyright" content="(c) 1998"/>
<layout>
<root-layout width="176" height="176"/>
<region id="video_region" width="176" height="144" left="0" top="0"/>
<region id="text_region" width="176" height="32" left="0" top="144" />
</layout>
</head>
<body>
<par>
<textstream src="text/text.rt" region="text_region" fill="freeze"/>
<video src="video/video.rm" region="video_region" fill="freeze"/>
</par>
</body>
</smil>
    
```

[그림 2] SMIL 파일

자막을 나타내는 rt 파일에서는 동영상 데이터와 자막 데이터를 동기화하기 위해서 시간축을 기준으로 동영상 데이터가 진행된 시간을 나타낸다. <time> 태그의 속성 'begin'을 통하여 진행 시간을 나타낸다. begin은 시작시간을 나타내고 이외에도 끝나는 시간을 표시하는 'end', 지속시간을 표시하는 'dur' 속성도 가지고 있다.

II. 검색시스템의 설계 및 구현

1. 시스템의 설계

구현된 SMIL을 이용한 동영상 검색 및 재생 도구는 다양한 종류의 데이터들을 재생하고 검색할 수 있다. SMIL과 MS에서 지원하는 SAMI 파일을 지원하며 특히, 동영상 정보에서 자막을 기반으로 한 검색기법을 통하여 사용자가 찾고자 하는 장면을 쉽게 검출할 수 있도록 한다.

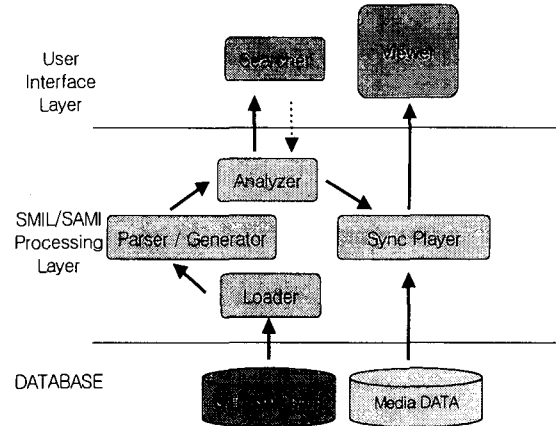
이 시스템은 [그림 3]과 같이 사용자 인터페이스 계층과 SMIL/ SAMI 처리 계층, 데이터베이스 계층의 3계층 구조를 갖는다.

사용자 인터페이스 계층은 쉽게 원하는 내용을 검색할 수 있게 검색 창을 제공하여 폐쇄자막을 기반으로 내용기반 검색을 할 수 있도록 한다.

SMIL처리 계층은 SMIL문서를 파싱하여 동영상 파일과 동기화를 시켜 보여줄 수 있는 형태의 공통된 데이터형태로 변환하여 재생할 수 있도록 한다. 또한 사용자의 검색한 부분을 재생할 수 있도록 하여 검색어만으로 쉽게 재생위치까지 가는 기능을 수행한다.

데이터베이스계층은 미디어 파일과 SMIL/SAMI 파일을 저장하여 데이터베이스 전체를 검

색할 수 있는 환경을 제공한다.



[그림 3] 시스템의 구조

2. 시스템의 구현

본 논문에서 구현한 시스템은 검색된 내용들을 색인하여 사용자에게 보여주고 재생을 하는 등의 작업을 효율적으로 수행하기 위해서 로더, 파싱/생성기, 분석기, 동기화 재생기등의 총 4개의 모듈로 구성되어 있다.

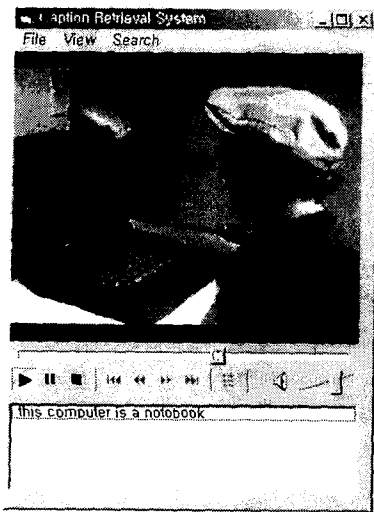
로더는 SMIL이나 SAMI로 작성된 문서를 읽어들이기 위한 모듈이다. 파싱/생성기는 SMIL 문서를 파싱하고 파싱한 문서들을 재생할 수 있는 형태로 문서를 재생성하는 역할을 한다. 분석기는 파싱/생성기로부터 넘겨받은 데이터를 동영상 파일과 동기화 시킬 때 필요한 요소별로 분류를 하며 검색할 수 있는 메타데이터를 생성한다. 또 사용자의 검색을 받아 필요한 데이터를 찾을 수 있도록 데이터들을 분류한다. 동기화 재생기는 동기화가 가능한 형태로 변환된 데이터를 미디어 파일의 시간원소를 기준으로 동기화하여 별도로 구역이 설정된 곳에 자막을 재생할 수 있도록 하는 기능을 한다.

3. 사용자 인터페이스 계층의 구현

재생을 하는 일반적인 플레이어와 비슷한 인터페이스로 구현하였다. 기존의 미디어 플레이어나, 리얼플레이어와 비슷한 환경에서 재생할 수 있다. 검색창을 통해 저장된 미디어들을 원하는 영상을 검색하여 재생할 수 있는 구조를 가지고 있다. 특히, 검색 부분은 자막의 내용뿐만 아니라 제목 및 기타 정보를 통해서도 검색을 할 수 있도록 하여 사용자에게 보다 검색이 편리한 환경을 제공한다.

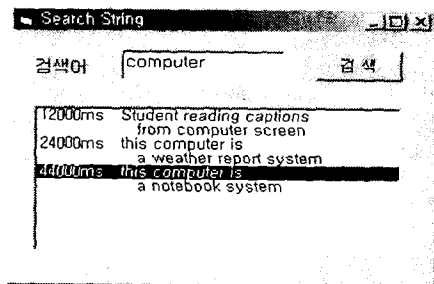
III. 구현결과 및 고찰

본 연구를 통해 구현된 재생기는 SMIL1.0 / SAMI 파일로 되어진 문서를 일반재생 및 검색 재생을 할 수 있도록 하였다. 검색을 통한 재생 시에는 검색된 목록을 제공하여 원하는 자막이 있는 부분을 재생하는 구조를 가지고 있다.



[그림 4] 검색 결과 재생화면

시스템의 실행화면은 [그림 4]와 같다. 일반적인 멀티미디어 플레이어에서 제공하는 하는 기능들을 모두 지원하고, 자막을 표현하는 부분은 동영상 재생창의 하단에 위치한다. 자막파일 이름 따로 지정하지 않아도 기본적으로 동영상 파일과 동일한 이름을 가지고 있는 "*.smi" 파일을 로드한다. 존재하지 않을 경우에는 기존의 멀티미디어 플레이어와 같은 작동을 하고 검색은 지원하지 않는다. 자막 파일은 자막파일에서 정의한 형식에 따라 화면 하단의 자막 출력 창에 출력한다.



[그림 5] 자막검색 창

시스템의 'Search' 메뉴를 클릭하면 검색을 할 수 있는 검색창이 화면상에 나타난다. 검색어를 입력하면 검색어가 있는 자막문장을 재생시간과 함께 리스트 박스에 [그림 5]와 같이 리스트를 출력한다.

[그림 5]의 검색 결과에서 검색결과를 클릭하면 이에 해당하는 장면부터 미디어파일이 재생된다.

결론

비디오 데이터는 최근 멀티미디어 서비스용으로 이용가치가 한층 부각되고 있는 대표적인 멀티미디어 데이터이다. 비디오 데이터는 기존의 텍스트 데이터에 비해 방대한 종류와 양으로 인해 기존의 데이터베이스 관리 시스템에서는 효율적인 관리가 어렵다. 이러한 비디오 데이터를 멀티미디어 서비스에 이용하기 위해서는 데이터를 활용할 서비스와 분야의 목적에 따라 분류하여 필수적인 데이터만 추출할 필요가 있다. 정확한 검색 결과를 얻기 위해서는 무엇보다도 제공되는 디지털 콘텐츠를 많이 확보할 수 있는 연구가 더 필요하다.

(본 연구는 인천대학교 멀티미디어 연구센터의 일부 지원으로 수행되었습니다.)

[참고 문헌]

- [1] W3C, Extensible Markup Language(XML) 1.0, <http://www.w3.org/TR/1998/REC-xml-19980210,1998>
- [2] W3C, Synchronized Multimedia Integration Language(SMIL)1.0Specification, <http://www.w3c.org/TR/1998/REC-smil-19980615,1998>
- [3] 김기병, 김형주, "내용 기반 검색 및 주석 기반 검색을 통합하는 비디오 데이터 모델의 설계 및 구현", 한국정보과학회논문지(C), 3(2), pp115-126, 1997
- [4] 김기욱, 김형주. "비디오 주석 시스템의 설계 및 구현", 한국정보과학회논문지, 24(6), June 1997.
- [5] Microsoft, "SAMI Adding Closed Captions to Multimedia", October 1999
- [6] Realaudio, Technical Blueprint and Whitepaper, February, 1997