

XML 기반의 동영상콘텐츠 검색 시스템 설계 및 구현

A Study on Implementation of XML-Based Information Retrieval System for Video Contents

김 용(Yong Kim)*

소민호(Min-Ho So)**

초 록

일반적으로 동영상콘텐츠에 대한 탐색과정에서 이용자는 해당 기관 또는 사이트에서 제공하는 간단하게 요약한 요약 동영상과 텍스트 정보를 이용하여 원하는 동영상을 선택한다. 이러한 이용자의 동영상콘텐츠의 탐색과정에서의 정확하고 신속한 동영상콘텐츠의 검색을 위하여 본 연구에서는 동영상을 구성하는 자막과 이미지 정보를 이용하여 동영상을 내용에 따라 여러 개의 클립으로 구분하고, 이용자의 필요에 따라 원하는 클립을 빠르게 검색하여 제공하는 시스템을 제안한다. 동영상콘텐츠 및 관련 메타데이터의 관리 및 제어를 위해서 본 연구에서는 XML 계층구조를 생성하여 저장한다. 한편, 이용자의 검색 요청시 XML을 기반으로 동영상콘텐츠의 계층구조를 반영하는 Xpath를 생성하여 질의처리를 수행함으로써 이용자가 원하는 동영상콘텐츠에 대한 검색결과를 제공한다. 제안된 방법을 기반으로 본 연구에서는 XML 계층구조를 이용한 동영상 검색 시스템을 설계하고 구현하였다.

ABSTRACT

Generally, a user uses briefly summarized video data and text information to search video contents. To provide fast and accurate search tool for video contents in the process of searching video contents, this study proposes a method to search video clips which was partitioned from video contents. To manage and control video contents and metadata, the proposed method creates structural information based on XML on a video and metadata, and saves the information into XML database. With the saved information, when a user try to search video contents, the results of user's query to retrieve video contents would be provided through creating Xpath which has class structure information. Based on the proposed method, an information retrieval system for video clips was designed and implemented.

키워드: 동영상, 검색, 구조화동영상 검색, 정보검색

XML, MPEG-7, video retrieval, information retrieval

* 전북대학교 문헌정보학과 조교수(yk9118@jbnu.ac.kr) (제1저자)

** 한국과학기술원 학술정보처 팀장(minhoso@kaist.ac.kr) (공동저자)

■ 논문접수일자: 2009년 11월 12일 ■ 최초심사일자: 2009년 11월 16일 ■ 게재확정일자: 2009년 11월 24일
■ 정보관리학회지, 26(4): 113-128, 2009. [DOI:10.3743/KOSIM.2009.26.4.113]

1. 서론

초고속 네트워크 및 정보의 전달을 위한 백본망의 전송속도가 빨라지고 각종 서비스 서버의 처리속도 및 성능이 증가함에 따라 인터넷을 이용한 다양한 디지털 콘텐츠에 대한 요구 또한 크게 증가하였다. 특히, 전통적인 텍스트 기반의 정보와 함께, 폭발적으로 증가하는 동영상 정보요구에 따른 효과적인 동영상콘텐츠의 분류 및 검색에 대한 필요성은 이용자의 효과적인 동영상콘텐츠의 이용, 동영상콘텐츠의 생산 및 유통에 있어서 절대적인 영향을 미칠 수 있다. 최근의 Web 2.0 개념을 수용하고 있는 서비스로서 각광받고 있는 YOUTUBE와 같이 이용자가 직접 제작 및 생산한 동영상을 제공하는 서비스 사이트의 성장은 이러한 동영상 검색에 대한 요구를 더욱 증대 시키고 있다. 기존의 동영상 검색은 크게 공간 영역에서의 동영상 검색(Farag and Wahab 2003)과 MPEG (Moving Picture Experts Group)과 같은 동영상 부호화 표준을 이용한 압축 영역에서의 동영상 검색(Kim and Chua 2005; Chen, Mark and Kin 2005; Ardizzone et al 1999)으로 분류할 수 있다. 공간 영역에서의 동영상 검색은 다양한 영상 처리 기법을 통해 비교적 정확한 검색이 가능하다. 그러나 대부분의 동영상은 MPEG과 같은 동영상 부호화 표준에 의해 압축된 비트 스트림(bitstream)의 형태로 저장 및 전송되고 있다. 따라서 공간 영역에서 동영상을 검색하기 위해서는 압축된 동영상 전체를 복호해야하는 단점이 있다. 이러한 단점을 극복하기 위해 압축된 비트 스트림 내에서 필요한 정보만을 추출해 동영상 검색에 이용하는

압축 영역에서의 내용기반 동영상 검색 방법이 주로 사용되고 있다. 이와 같은 검색기법을 기반으로 기존의 동영상 서비스에서는 간단하게 요약된 동영상이나 텍스트 정보를 이용하여 원하는 동영상콘텐츠를 검색한다. 이러한 서비스에서는 검색된 동영상이 이용자가 원하는 동영상이 아닐 수도 있으며, 동영상의 일부를 보기 위해서 동영상을 처음부터 브라우징해야 하는 경우도 생긴다. 그러나 전체 동영상이 아닌 한편의 동영상을 여러 개의 클립으로 구분하고 클립을 대표하는 메타정보를 제공하면 이용자가 필요로 하는 클립을 검색하여 제공함으로써 전체 동영상을 검색하는데 따른 시간 및 검색 속도에 있어서 보다 효율적인 동영상 검색 서비스의 제공이 가능할 수 있다. 즉, 스트리밍 서버를 이용하여 검색된 클립만을 별도로 제공하거나 또는 분할된 동영상 정보를 통하여 보다 정확한 검색서비스가 제공될 수 있으며 이용자는 편리하게 동영상에서 필요한 부분만을 검색하고 시청할 수 있다. 한편, 동영상콘텐츠를 포함한 멀티미디어 자료를 검색하고 선택할 수 있는 정보를 제공하는 메타데이터와 색인의 필요성에 따라 국제 표준활동이 활발하게 전개되고 있다. 이와 같이 동영상콘텐츠의 검색에 메타데이터와 색인을 활용하고자 하는 주된 이유는 동영상콘텐츠는 주로 색상, 형태 등의 요소로서 검색이 수행되는데 이와 같은 요소로서 검색률을 높이는데 있어서의 기술적인 어려움 때문이라고 할 수 있다. 동영상콘텐츠에 대한 기술정보 및 구조와 관련된 대표적인 표준화 활동 기구로서 MPEG-7에서는 멀티미디어 자료에 대한 구조정보와 의미정보를 기술할 수 있는 국제 표준 규격을 제정함으로써 메타데이

터를 이용하여 원하는 자료를 검색하고 선택하여 자료에 대한 효율적인 접근과 이용에 있어서 상호운용성을 제공하고 있다. 해당 표준에서는 멀티미디어 자료에 대한 정보 표현을 위하여 XML을 사용하고 XML 스키마를 기반으로 하여 MPEG-7 서술도구의 구조를 정의하였으며 질감, 색상 등의 하위레벨 특징정보와 의미정보가 포함된 상위레벨로 동영상콘텐츠의 정보를 표현한다(남윤영, 황인준 2002). 이러한 MPEG-7 표준을 기반으로 하는 멀티미디어 기술구조(MDS: Multimedia Description Schemes)는 동영상콘텐츠에 대한 의미 및 내용을 구조화함으로써 이용자가 원하는 멀티미디어 자료의 특정 부분에 대한 접근성을 용이하게 하는 가장 효과적인 도구이다(안형근, 고재진 2006). 이러한 국제표준의 활성화는 동영상과 같은 멀티미디어 자료의 생산 및 이용의 활성화를 더욱 촉진 시킬 수 있는 촉매역할을 수행할 수 있을 것이라고 기대되고 있다. 그러나 MPEG-7에서는 동영상콘텐츠에 대한 구조 및 의미정보에 대한 내용을 정의하고 있으며 해당 자료에 대한 검색과 추출에 대한 표준을 정하지 않아 MPEG-7 자료를 처리하기 위한 방법이 요구된다(김택곤, 김우생 2003). 한편, XML을 기반으로 문서의 구조화 정보를 이용한 검색방법은 전통적인 전문검색에 비하여 많은 연산과정과 복잡성을 필요로 하지만 높은 검색 성능을 보여준다(윤소영, 문성빈 2006). 따라서 검색과정에서의 연산 및 복잡성을 줄이면서 높은 검색성능을 보여주기 위한 다양한 방법들에 대한 연구가 진행되고 있다. 이와 같은 연구들에 있어서 주요한 분야로서 XML 기반의 동영상콘텐츠의 내용정보를 활용한 연구

가 시도되고 있다. 따라서 본 연구에서는 동영상콘텐츠의 분할된 하위 요소인 클립을 활용하면서 해당 동영상콘텐츠의 내용정보를 XML 기반으로 메타데이터를 활용한 검색기법에 대한 제안을 통하여 해당 검색 시스템의 구현을 목표로 하고 있다.

2. 연구 범위 및 목표

검색 시스템들은 데이터의 제어 및 관리를 효과적으로 하기 위해 여러 가지의 데이터 제어 및 관리 방식을 사용하는데, 특히 동영상콘텐츠와 같은 대용량의 데이터에 대한 검색에 있어서 효과적인 데이터 제어 및 관리 방식이 더욱 요구된다. 기존의 동영상콘텐츠 검색 시스템들은 검색에 필요한 메타데이터와 동영상 데이터를 데이터베이스(Database)에 저장하고, 이용자의 검색 요청이 있는 경우에 있어서 기존의 저장된 메타데이터를 참조하여 원하는 비디오 데이터를 데이터베이스에서 찾아서 보여주는 방식을 채택하거나, 비디오 데이터를 데이터베이스에 저장하지 않고 파일 시스템의 형태로 저장하여 검색하는 방식을 채택하였다. 그러나 파일시스템의 형태로 데이터를 저장하는 경우에 있어서 대용량의 동영상콘텐츠들을 하나의 디렉토리 내에 저장하게 되면 데이터를 검색하는데 소요되는 시간이 길어지게 되므로 상기 데이터들을 디렉토리(Directory)화 하는 방식을 사용하였다. 이때, 동영상콘텐츠에 대한 식별을 위해 각각의 동영상콘텐츠 고유의 식별 번호(Identification Number)를 부여하여 관리하는 방법이 이용되거나, 동영상콘텐츠에 대

한 파일명을 해싱(Hashing)하여 디렉토리 시스템에서 원하는 동영상콘텐츠의 위치를 찾아내는 방법이 이용되었다. 그러나 개별 시스템들에 사용되는 데이터 관리 방식의 상이성과 시스템의 변경, 데이터양의 증가 또는 검색 필드의 변경 등에 따라 관리 방식도 변경이 요구됨으로써 데이터에 대한 관리가 용이하지 않았다. 특히, 동영상콘텐츠와 같이 데이터의 구성이 반복적인 특징을 갖는 경우, 즉 하나의 동영상콘텐츠안에 여러 세그먼트들이 포함되며 각각의 세그먼트 안에 또다시 여러 개의 세그먼트들이나 클립들이 구성된 경우에 있어서 이들 각각의 세그먼트들이나 클립들에 대해 검색을 하기 위한 기존의 검색방법에서는 데이터들에 효과적으로 접근하여 검색하기가 쉽지 않았으며, 데이터 관리에 있어서도 복잡성이 증대되는 문제점이 있었다. 즉, 이용자의 검색 요청시 원하는 데이터에 대해 그 위치를 찾아내기 위한 시간이 길어지는 문제점이 있었으며, 데이터 관리에 있어서는 데이터에 고유의 식별번호를 부여하여 유지하는 경우에 데이터가 삭제되면 식별번호 역시 같이 삭제해야 하고, 새로운 데이터가 추가되는 경우에도 추가되는 데이터에 고유의 식별번호를 부여해야 하므로 미디어 자체에 대한 관리뿐만 아니라 식별번호 자체에 대한 관리 모듈이 추가되어야 하는 문제점이 있었다.

따라서 본 연구에서는 기존의 동영상콘텐츠 검색시스템의 문제점을 해결하고 동영상콘텐츠의 효과적인 검색과 이용을 위하여 동영상콘텐츠의 상위레벨 정보를 포함하는 의미정보를 기반으로 하는 검색 시스템에 대한 설계 및 구현을 목표로 한다. 특히, XML을 활용하여 동

영상콘텐츠에 대한 의미정보를 활용한 계층구조 기반의 동영상 검색 시스템을 구현하고자 한다. 제안하고 있는 동영상 검색 시스템을 구현하기 위하여 본 연구에서는 하위 요소로서 동영상콘텐츠에 대한 정보를 입력하는 입력 모듈, 동영상콘텐츠의 의미정보를 XML로 표현할 수 있는 XML 계층구조 반영 모듈, 개별적인 모듈들에서 생성된 정보를 기반으로 실제 검색이 수행되는 검색 모듈 및 이용자 인터페이스에 대한 설계와 구현을 수행하고자 한다. 세부적으로 개별 하위 요소의 기능을 알아보면 다음과 같다. 제안하고 있는 동영상 검색 시스템을 구현하기 위해서 먼저 전체 동영상을 의미적으로 연속성을 가지는 부분으로 분할된 클립으로 분할하고 동영상의 주요 구성요소인 이미지와 텍스트 정보를 별도로 분석하여 클립 구분 및 검출에 통합하여 이용할 수 있는 모듈이 요구된다. 즉, 해당 모듈에서 동영상콘텐츠를 클립단위의 하위요소로 분할하고 해당 정보는 XML 기반의 데이터베이스에 저장한다. 위의 과정을 통하여 XML 데이터베이스에 저장된 동영상콘텐츠에 대한 정보를 통하여 동영상에 대한 검색을 수행할 수 있는 검색모듈은 본 연구에서 제안하고 있는 동영상 검색 시스템의 핵심요소라고 할 수 있다. 이를 위하여 본 연구에서는 동영상정보의 효과적인 검색과 체계적인 데이터 관리를 위하여 동영상 클립에 대한 키워드를 기반으로 메타데이터와 XML의 계층구조를 활용하였다. 이와 같은 방법은 동영상콘텐츠에 대한 기술의 단순성 및 시각정보의 효과적인 표현방법으로 인하여 매우 유용한 방법이라고 할 수 있다(김진아, 김태수 2001). 한편, 동영상콘텐츠의 검색을 위하여 동영상 검

색시스템은 동영상콘텐츠에 대한 데이터 제어를 위하여 데이터 입력, 데이터 처리, 데이터 조작을 할 수 있는 기능을 제공해야 한다. 예를 들어, 새로운 데이터가 입력되거나 기존의 입력된 불필요한 데이터에 대해서 삭제해야 할 필요가 있을 경우 데이터를 갱신하여 일관되게 데이터를 유지해야 한다. 또한 이용자의 질의를 입력 받을 수 있는 이용자 인터페이스는 이용자의 검색 편의성을 위하여 매우 중요한 하위요소라고 할 수 있다. 이와 같은 연구의 목표를 위하여 본 연구에서는 첫째, 이미지 특성을 이용하여 동영상을 의미를 가지는 클립으로 분할하는 기법과 자막정보를 이용하여 클립을 검출하는 기법을 제안하고 둘째, 모듈별 개발 환경 및 XML 계층구조를 이용한 동영상 검색 시스템의 구성 및 개별 기능들을 수행하는 모듈에 대한 설계 및 구현을 수행하고자 하며, 마지막으로 본 연구의 결론과 향후 연구과제에 대해서 알아보려고 한다.

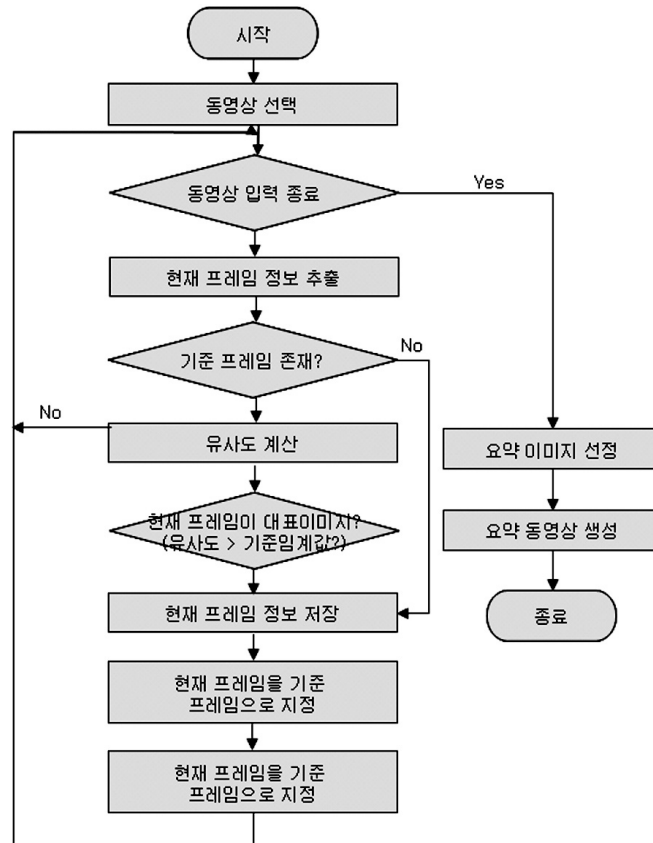
3. 클립 분할 및 검출

3.1 동영상클립 생성 과정

본 연구에서 제안하고 있는 동영상콘텐츠의 효과적인 검색을 위해서는 전처리 과정에서 먼저 동영상콘텐츠를 하위 요소인 클립형태로의 분할이 요구된다. 이와 같은 과정을 통하여 분할된 동영상콘텐츠에 대한 키워드 및 자막을 기반으로 생성되는 해당 동영상콘텐츠 및 동영상 클립에 대한 메타데이터를 생성한다. 이를 위하여 기존의 동영상콘텐츠를 분할된 동영상

클립의 형태로 생성하기 위해서는 분할을 위한 기준이 요구된다. 이와 같은 기준을 기반으로 동영상콘텐츠의 하위요소로서 클립에 대한 분할이 이루어진다.

〈그림 1〉에서는 이와 같은 동영상클립 생성 과정을 도식화하고 있다. 이를 간략히 알아보면 먼저 동영상신호나 동영상화일을 선택하거나 입력으로 받아들여 입력된 동영상을 선택하고 선택된 동영상신호나 동영상화일의 입력이 종료되었는지를 검사한다. 동영상콘텐츠의 입력이 종료되지 않았으면 동영상의 현재 프레임 정보를 추출하고 현재 프레임이 동영상 클립을 구성하는 새로운 클립의 대표 이미지 여부를 확인하여 현재 프레임이 대표이미지가 아니면 동영상 입력이 종료되었는지의 여부를 검사하는 과정으로 넘어가게 된다. 기준 프레임이 존재하는지를 검사하여 기준 프레임이 존재하지 않으면 현재 프레임을 대표이미지로 결정한 다음에 이를 다시 이전 과정으로 전달한다. 한편, 기준 프레임이 존재하면 기준 프레임과 현재 프레임간의 유사도를 계산한 후에 계산한 유사도가 대표이미지가 되기 위한 기준값에 비하여 높은지에 대한 여부를 검사한다. 검사 결과에 따라 유사도가 기준값보다 크면 현재 프레임이 대표이미지가 아닌 것으로 판단하며 유사도가 기준값보다 작으면 현재 프레임을 대표이미지로 결정하게 된다. 이와 같은 유사도와 기준값을 검사하는 과정에서 현재 프레임의 DCT(Discrete Cosine Transform) 상수중의 DC(Direct Current) 성분 히스토그램이나 DC 성분 차이 또는 두 가지 요소를 동시에 이용하여 유사도를 검사하는 방법을 사용할 수 있다. 이를 이용하면 비디오를 구성하는 프레임들의 DCT 성분 정



〈그림 1〉 동영상클립 생성 흐름도

보를 보유하고 있는 MPEG 등의 비디오화일을 입력으로 할 경우에는 매우 빠르게 대표 이미지를 추출할 수 있는 장점이 있다.

현재 프레임이 대표이미지인 경우에는 추출된 대표이미지 정보를 대표이미지 정보 테이블에 저장하고 현재 프레임을 새로운 이미지 선정을 위한 기준 프레임으로 지정한 후에, 동영상 입력의 종료여부를 검사하는 과정으로 넘어간다. 한편, 동영상 입력이 종료되었으면 저장된 대표이미지 정보 테이블의 정보를 이용하여 동영상클립 구성에 필요한 요약 이미지의 수만큼 대표이미지로부터 요약이미지를 선정한 후

에 선정된 요약이미지를 이용하여 입력된 동영상신호나 동영상화일로부터 이미지를 추출하여 동영상클립을 생성한 후에 종료한다.

3.2 이미지를 이용한 클립 분할

동영상콘텐츠는 텍스트, 오디오 및 동영상 이미지의 세 가지 요소로 구성되어 있다. 동영상콘텐츠는 프레임이라 불리는 이미지들을 연속적으로 제공하여 정보를 표현한다. 동영상콘텐츠에서 물체의 자연스러운 움직임은 시각의 잔상효과를 이용하여 이루어진다. 즉, 동영상콘

텐츠의 내용이 사람의 시각에 보여주는 원리에 있어서 인간의 시각은 각각의 프레임이 연속적으로 연결된 영상에 대한 빠른 움직임들을 독립적으로 구분하지 못하고 망막에 잔상을 남기고 이러한 잔상들이 연결되어 자연스러운 움직임을 만든다. 따라서 연속된 이미지는 시각적으로 매우 유사하며 장면의 전환이 이루어지는 시점에서 연속되는 두 이미지의 차이가 커진다. 이러한 성질을 이용하여 시각적으로 차이를 보이는 경계점을 추출하며, 이러한 경계점 사이에 있는 이미지들을 컷(Cut)이라 정의한다. 컷의 경계점을 검출하기 위해서는 각 프레임 이미지의 세 가지 특징인 색상 히스토그램, 이미지 레이아웃, 텍스트를 추출한다. 컷 경계의 기준이 되는 특징의 차이인 임계값(Threshold)은 시뮬레이션을 통해 조절하여 적절하게 컷이 분할되도록 지정한다. 이때 색상 히스토그램은 전통적인 방식을 사용하여 추출한다(Ballard 1991). 이미지 레이아웃과 텍스트는 웨이블릿 변환(Wavelet transform)을 사용함으로써 이미지 특징(Image features)을 추출하기 위해 요구되는 연산시간을 최대한 줄여 대용량 이미지 처리에 적합하도록 하였다(Leong and Selvanathan 1997; Scheunders et al 1997). 이미지의 세 가지 특징은 정확한 컷 검출과 빠른 검출을 위해 적절하게 조합하여 사용할 수 있도록 하였다.

대부분의 동영상은 컷이 여러 개 모여서 내용적으로 유사한 하나의 클립을 만든다. 시각적인 차이를 이용하여 자동으로 경계를 결정하는 컷과는 달리 클립은 경계 선정의 기준이 명확하지 않으며, 동영상의 장르나 상황에 따라 경계를 달리하고 컷을 통합하여 클립을 만드는 것은 매우 어려운 일이다. 본 연구에서는 동영

상콘텐츠의 검색을 위한 최소단위로서 동영상을 분할하는데 여기서의 최소단위는 클립(Clip)으로 정의한다. 클립은 의미상으로 연속성을 갖는 컷들을 모은 것으로서 메타데이터를 부여할 수 있는 최소단위로서 물리적 연속성을 갖는 컷의 집합이라 정의할 수 있다. 이러한 클립을 생성하기 위한 방법으로서 일반적으로 문자정보의 연속성을 이용, 특정 이미지로 시작되는 영역을 분할 및 음성 정보를 이용하는 방법이 있으며 전문적인 지식이나 기술을 가진 전문가에 의한 수작업에 의한 방법도 적용할 수 있다. MPEG-7이 표준화되어 동영상콘텐츠의 저작자가 클립의 경계를 직접 기술할 때 까지는 다양한 기법이 클립 검출에 이용될 것이지만 다양한 분야에 모두 잘 적용되는 기법의 개발은 빠른 시일 내에 이루어지지는 않을 것이다. 본 연구에서는 키워드 기반 클립 경계 검출 기법과 더불어 클립의 경계 이미지를 제공함으로써 클립의 경계검출을 수행하였다. 제안하고 있는 기법은 컷의 시작이미지와 미리 정의된 몇 장의 기준 이미지를 비교하여 컷의 시작 이미지와 기준이미지 중에서 하나와 정의된 임계값을 비교하여 두 값이 유사하면 해당 컷을 새로운 클립의 시작으로 결정하는 방법이다. 이 기법은 뉴스와 같이 항상 앵커나 로고와 같은 특정한 이미지로 새로운 뉴스를 제공하는 경우에 매우 유용하다.

3.3 자막정보를 이용한 클립 검출

효율적인 동영상콘텐츠의 검색을 위해서는 동영상이 제공하여 주는 한글 또는 영문의 자막정보의 이용이 필수적이다. 실질적으로 동영

상 콘텐츠에 대한 내용을 기반으로 검색을 수행하는 방법에 있어서는 아직은 기술적인 한계 점이 존재한다. 그러나 동영상콘텐츠의 내용을 기술하고 있는 텍스트를 기반으로 검색을 하는 방법은 기존의 텍스트 기반의 검색기법을 적용할 수 있기 때문에 보다 높은 검색효율을 기대할 수 있다. 동영상콘텐츠에 대한 내용을 보여주는 텍스트정보로서 자막정보는 영상의 분할 및 분할된 영상의 요약 등에 있어서 매우 효과적인 정보로 이용될 수 있다. 현재 자막을 동영상에 넣어 주는 방법은 영문 자막 방식과 한글 자막 방식이 있다. 자막 디코더(Decoder)를 사용하면 문자 정보를 화면상에서는 볼 수 있으나 자막 자체를 추출 할 수는 없다. 따라서 자막 정보를 추출 하기 위해서는 별도의 자막 추출 기술 개발이 필요하고 본 시스템은 영문과 한글의 자막을 추출하는 별도의 모듈을 포함하고 있다. 이 모듈은 오스프레이 캡처 보드를 이용하여 구성되었으며 자막은 문자 단위와 라인 단위로 추출할 수 있다. 추출된 자막은 동영상과의 동기화를 위하여 시간 태그를 부착하였다. 또한 동영상의 스트리밍서비스에 이용할 수 있도록 시간정보와 자막정보를 포함하는 SAMI (Synchronized Accessible Media Interchange 형식¹⁾)으로의 출력 기능도 구현하였다.

추출된 자막정보는 이미지를 이용한 경계검출 과정에서 추출된 컷의 경계시간을 이용하여

자막의 경계검출을 수행한다. 즉, 전체 자막정보를 컷 경계의 시간정보를 이용하여 각 컷에 대응하는 자막으로 분리하고, 분리된 자막정보는 키워드 추출방법을 통하여 해당 컷에 대한 키워드를 생성한다. 본 연구에서는 사전기반 키워드 추출방법을 사용하며 자막으로부터 추출하고자 하는 키워드를 사전에 추가함으로써 키워드 추출을 가능하게 하였다.

개별적인 컷들에 대한 키워드는 동영상의 클립 분할에 이용된다. 즉, 연속된 컷의 키워드의 유사도를 비교하여 기준 임계값 이상의 유사도를 가지면 두 컷은 동일한 클립을 형성하는 것으로 간주한다. 제안된 시스템에서는 이미지를 이용한 클립검출과 키워드를 이용한 클립검출을 선택적으로 이용하거나 동시에 적용할 수 있도록 하였다. 또한 하나의 클립에 포함된 컷에 대응하는 키워드들은 융합되어 해당 클립을 표현하는 키워드가 되어 클립의 검색에 이용된다.

4. 시스템 구성 및 기능

4.1 구현 환경

본 연구에서는 효과적인 동영상콘텐츠의 검색을 위하여 동영상 검색 시스템에 대한 설계

1) Synchronized Accessible Media Interchange의 약자이며 Microsoft에서 제정한 규격이다. 이것의 개념은 캡션(자막) TV와 유사하며 캡션 TV와 비슷하게 미디어 콘텐츠에 선택적으로 자막을 넣는 것이다. 이전에는 멀티미디어 콘텐츠에 자막을 넣기 위해서는 상당히 귀찮고 복잡한 과정을 거쳐야 했다. VHS 테이프라면 자막기를 활용하여 자막을 넣어야 했고 동영상 화일은 Adobe Premiere와 같은 편집 툴을 이용하여 수동으로 자막을 넣지 않으면 안됐다. 이런 복잡한 문제로 인해 동영상 콘텐츠를 제공하는 기업, 교육 기관에서는 상당한 노력을 들여 콘텐츠를 제작하거나 동영상 콘텐츠 제작을 포기해야 했다. 이런 문제를 해결하기 위해 동영상과 자막 화일을 분리하여 동영상은 동영상대로 읽어 들이고 자막은 자막대로 읽어 들여 이 둘을 합쳐 재생하는 방식이 탄생했다. 이러한 연구 결과로서 SAMI형식이 대표적이라고 할 수 있다.

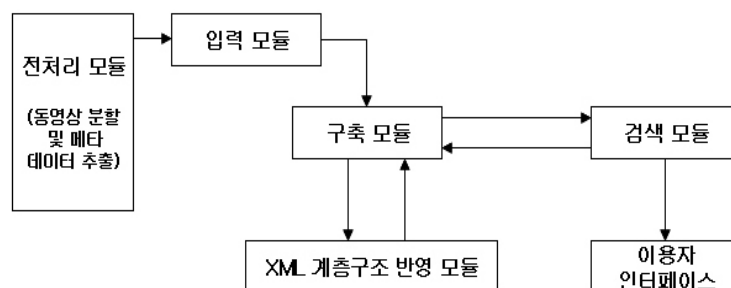
및 구현을 수행하였다. 이와 같은 동영상 검색 시스템을 구현하기 위하여 동영상콘텐츠를 분할하고 분할된 하위요소인 클립에 대한 정보 및 추출된 정보를 계층화된 형식으로 구현하기 위하여 기능모듈들이 요구된다. 이와 같은 기능을 수행하는 하위 요소로서 동영상콘텐츠에 대한 정보를 입력하는 입력모듈, 검색모듈, XML 계층구조를 반영하는 계층구조 반영모듈 및 사용자 인터페이스로 구성되었다. 본 연구에서 제안하고 있는 동영상 검색 방법을 기반으로 구현된 동영상 검색 시스템은 Java JSP와 XML DBMS인 엑셀론을 이용하여 구현하였으며 동영상콘텐츠에 대한 정보 입력 및 구축을 위해서 Java와 관련 Dom API를 사용하였다. 웹상에서의 사용자 인터페이스 및 사용자 질의 처리를 위해서 JSP, Java를 이용하였다. 한편, 웹기반의 사용자 환경을 구현하기 위하여 마이크로소프트사의 IIS 웹서버와 서블렛 엔진인 resin을 적용하였다.

4.2 동영상 검색시스템 구성

동영상콘텐츠의 검색을 위한 주요모듈의 구현과 함께, 각각의 모듈을 기반으로 하는 동영

상 검색시스템의 구성도는 <그림 2>와 같다. 구현된 동영상 검색시스템에서 동영상콘텐츠의 검색을 위하여 먼저 동영상콘텐츠의 클립으로의 분할 및 자막정보를 이용한 동영상 클립의 검출이 수행된다. 이와 같은 과정은 XML 기반의 계층정보 및 메타데이터를 활용한 동영상 검색을 위한 전단계로서 수행되었다. 이와 같은 전처리과정을 통하여 추출된 동영상콘텐츠에 대한 자막정보, 즉 추출된 키워드는 해당 동영상콘텐츠에 대한 메타데이터로서 적용된다. 따라서 실질적인 검색을 위하여 전처리 과정으로서 동영상콘텐츠에 대한 의미정보에 대한 XML기반의 구조정보를 생성하기 위한 메타데이터의 생성이 이루어진다.

<그림 2>에서는 본 연구에서 제안하고 있는 XML 계층구조를 이용한 동영상 검색시스템의 구성도이다. 시스템을 구성하는 개별 모듈에 대한 기능을 간략히 살펴보면 다음과 같다. 먼저 전처리 모듈에서는 전체 동영상에 대하여 이미지를 이용한 동영상 클립으로의 분할 및 개별 클립에 대한 자막정보를 기반으로 하는 키워드를 추출하고 이를 개별 클립에 대한 메타데이터로 추출한다. 입력 및 구축모듈에서는 동영상 검색을 위한 동영상화일과 관련 메타데



<그림 2> 동영상 검색시스템 구성도

이터를 입력받아 XML화일을 스캔하면서 필요한 ASX화일을 생성하고 이를 디렉토리에 동영상 별로 저장한다. XML 계층구조 반영모듈에서는 실제 동영상, 세그먼트, 대표프레임의 검색을 위하여 생성된 가상 디렉토리에 대한 상세정보 및 계층정보를 XML 구조에 반영하는 기능을 수행한다. 이와 같이 이용자 질의를 통하여 검색이 수행되는 일련의 과정은 최종적으로 실제 검색을 수행하는 검색모듈이 Xpath를 생성하고, 이용자 인터페이스를 통해 생성된 결과를 보여주는 것으로 마무리된다.

4.3 입력 모듈

입력모듈에서는 동영상콘텐츠와 이에 따른 XML의 계층화된 구조로 이루어지는 동영상 콘텐츠에 관련된 메타데이터를 입력받는 모듈로서 개별적인 동영상콘텐츠에 대하여 검색에서 활용될 수 있는 정보를 입력받는 기능을 수행한다. 즉, 동영상콘텐츠에 대한 정보를 포함하여 전처리 과정을 통하여 생성된 동영상콘텐츠에 대해서 상세정보들이 입력모듈을 통하여 입력된다고 할 수 있다. 입력모듈을 통하여 입력되는 정보들은 먼저 동영상 제목, 주제, 주석, 세그먼트수, 클립수, 개별 세그먼트에 대한 인식번호(Identification Number), 세그먼트와 관련된 키워드, 주석, 대표 프레임화일, 클립별 인식번호, 클립키워드 및 주석 등의 정보가 포함된다. 한편, 본 동영상 검색시스템에서는 동영상 전체와 동영상을 세그먼트로 나눈 세그먼트별 검색 그리고 세그먼트내의 클립들에 대한 검색

을 가능하게 하기 위해서 <그림 3>과 같은 형태로 개별적인 동영상콘텐츠에 대한 계층구조를 포함하고 있는 XML 화일을 입력받는다.

```

<mpeg7>
<VideoRoot>
<VideoRootMetaData> * </VideoRootMetaData>
<Video Id="Video1">
<VideoMetaData> ** </VideoMetaData>
</VideoRoot>
<SegmentMetaData> * </SegmentMetaData>
<Segment Id=Segment1>
<VideoRoot>
<ClipMetaData> * </ClipMetaData>
<Clip Id="Clip1"> ** </Clip>
<Clip Id="Clip2"> ** </Clip>
<Clip Id="Clipn"> ** </Clip>
</VideoRoot>
<Segment>
<Segment Id=Segment2> * </Segment>
<Segment Id=Segment3> * </Segment>
<Segment Id=Segmentn> * </Segment>
</Segment>
<VideoRoot>
<Video Id="Video2"> ** </Video>
<Video Id="Video3"> ** </Video>
<Video Id="Video4"> ** </Video>
*
<Video Id="Video5"> ** </Video>
</VideoRoot>
</mpeg7>
    
```

<그림 3> 동영상 데이터에 대한 XML 계층 구조의 예

본 연구에서 적용하고 있는 동영상 화일의 형식은 ASF 화일의 형식으로서 하나의 동영상 화일에 대해서 각각의 ASF화일²⁾과 XML화일

2) ASF(active streaming format): ASF는 오디오, 동영상, 슬라이드 쇼, 그리고 동기화된 이벤트 등을 지원하는 마이크로소프트의 스트리밍 미디어 형식이다. ASF는 인터넷을 통해 오디오, 동영상 및 생방송을 수신하는 유틸

을 입력데이터로서 입력을 받는다. 입력받은 데이터들에 대해서 XML 화일을 분류(Parsing) 하기 위하여 실제 필요한 정보들을 추출하여 세그먼트를 위한 ASX 화일을 생성해 낸다. 즉, 하나의 ASF 화일에 대해 여러 ASX 화일을 생성하여 디렉토리 형태로 관리한다. <그림 4> 는 세그먼트를 보이여 주기 위한 ASX 화일의 구조를 도식화 하여 표현하고 있다.

```
<ASX VERSION="3.0"> <ENTRY>
<TITLE>demo1.asf_Seg4.asx</TITLE>
<REF HREF="mms://nf3000/demo1.asf" />
<STARTTIME VALUE="00:05:59.62" />
<DURATION VALUE="00:03:18.87" />
</ENTRY> </ASX>
```

<그림 4> 세그먼트별로 생성되는 ASX 화일

입력받은 XML 화일들은 XML DBMS인 엑셀론에 저장시키며 동영상 파일(ASF 화일)과 세그먼트화일(ASX 화일) 그리고 대표프레임 이미지화일은 화일시스템에 디렉토리 형태로 저장시킨다.

4.4 XML 계층구조 반영 모듈

입력된 XML 계층구조에는 동영상콘텐츠화일, 동영상콘텐츠별로 구분된 세그먼트화일 및 대표프레임 화일에 대해서 설정된 가상 디렉토리에 대한 경로를 XML 화일에 반영해 주어야 한다. 특히, 세그먼트화일을 실행하거나 대표

프레임을 웹상에서 보여주기 위해서 관련 정보에 대한 가상 디렉토리를 XML 계층구조에 반영해 주어야 한다.

입력데이터와 함께, 실제 입력데이터를 어떠한 형태의 XML 계층구조로 변환하여 주느냐에 따라 Xpath를 통한 데이터의 위치경로가 다르게 형성되지만 동영상 XML 형태로 변환된 데이터 중에서 실제 데이터인 동영상 세그먼트 및 동영상 클립들에 대해서 XML 계층구조에 실제 원시 데이터에 대한 위치정보를 포함하고 있어야 한다. 이와 같은 방법은 이용자의 질의가 있을 경우에 해당 동영상 세그먼트 또는 동영상 클립에 대한 위치를 계산하지 않고 세그먼트의 XML 계층구조에서 텍스트 정보를 찾듯이 바로 위치를 확인할 수 있는 장점이 있다. 즉, 이용자의 질의 요청이 있는 경우에 전처리 과정을 통하여 생성된 XML 계층구조 정보를 활용하여 해당 질의에 적합한 동영상콘텐츠에 대한 위치를 Xpath 경로를 통하여 검색함으로써 검색속도를 향상시킬 수 있다.

XML 계층구조 반영 모듈은 검색 이전에 XML 기반의 동영상 시스템에서 검색을 하기 위해서 필요한 데이터의 경로를 XML 화일에 반영해야 한다. <그림 5>는 XML 화일에 표현되는 계층구조중에서 세그먼트와 관련된 정보를 나타낸 것이다. 개별 세그먼트에 대해 ASX와 대표 프레임에 대한 가상 디렉토리가 설정된다.

리틴인 마이크로소프트의 NetShow에서 사용된다. 여기에는 이 화일 형식과 관련된 두 가지 화일 형태가 있다. 확장자가 .asx인 화일은 웹브라우저에게 윈도우 미디어 플레이어를 호출하고, 스트리밍 콘텐츠가 담겨있는 .asf 화일을 로드하도록 신호를 보내는데 사용된다.

```

<Segment id="Seg1">
<TextAnnotation> 오늘의 뉴스 요약 </TextAnnotation>
<MediaTime>
<MediaRelTimePoint>00:00:00,00</MediaRelTimePoint>
<MediaDuration>00:01:03,50</MediaDuration>
</MediaTime>
<Keyword> 대통령 순방 </Keyword>
<Abstract> 대통령의 동남아 순방 </Abstract>
<Annotation> 아시아국가들과의 관계를 계선
</Annotation>
<Mosaic>
<MediaLocator>
<MediaUri>/Upload/demo.asf/Seg1.asx</MediaUri>
<MediaImageUri>/Upload/demo.asf/test.gif
</MediaImageUri>
</MediaLocator>
</Mosaic>
</Segment>
    
```

〈그림 5〉 세그먼트 계층구조

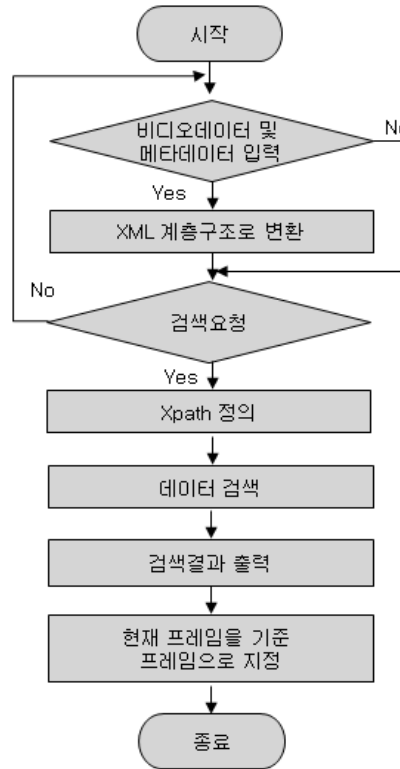
4.5 검색 모듈

동영상콘텐츠의 하위요소인 클립 검색을 포함한 동영상 검색 시스템에서 가장 중요한 기반 기술은 대용량의 동영상콘텐츠에서 원하는 정보를 고속으로 찾아줄 수 있도록 색인과정을 수행하는 방법이라고 할 수 있다. 따라서 본 연구에서는 대용량 동영상콘텐츠의 효율적인 색인을 위하여 동영상콘텐츠를 하위요소인 클립 단위로 분할하고 이에 대한 메타데이터 정보를 활용할 수 있는 방법을 활용하였다. 제안된 시스템에서 색인은 크게 두 단계로 이루어진다. 먼저 첫 번째 단계에서는 모든 동영상콘텐츠의 분할을 통하여 얻어진 클립을 포함하는 동영상콘텐츠를 하위요소로 분할하고 이에 대한 트리구조를 생성하며 두 번째 단계에서는 동영상 클립에 대한 내용정보를 표현하는 키워드들에 대한 트리구조를 생성하는 단계로 구성되었다.

키워드 트리구조에 포함된 모든 키워드들은 해당 키워드가 표현하고 있는 모든 동영상콘텐츠의 하위요소로서 클립을 포함할 수 있도록 한다. 이렇게 생성된 키워드는 해당 동영상에 대한 메타데이터로서 XML 구문으로 표현되어 XML 데이터베이스에 저장된다.

검색 모듈은 제목, 주제, 키워드, 요약, 동영상 생성일자 등의 메타데이터 정보를 기반으로 동영상콘텐츠 자체에 대한 검색 뿐만 아니라 동영상 내의 세그먼트, 세그먼트 내의 클립에 대해서 검색이 가능하다. 이를 위해서 XML 계층구조를 이용하여 구성된 XML 화일들을 XML 데이터베이스에 저장한 후 이용자의 검색 요구가 발생하면 해당 질의에 대하여 Xpath를 생성하여 이용자 질의를 처리한다. 이와 같은 과정을 통하여 추출된 검색 결과는 XSLT를 이용하여 웹상에 표현된다. 〈그림 6〉은 동영상콘텐츠에 대한 정보입력을 시작으로 클립으로의 분할 과정 및 이용자 질의에 대한 처리과정을 단계별로 표현하여 주고 있다. 한편, 동영상 검색은 동영상화일 자체에 대한 검색, 동영상의 세그먼트별로의 검색, 세그먼트내의 클립별로의 검색으로 나눌 수 있다. 결국 동영상들이 있고 동영상내에 세그먼트들이 있고, 세그먼트 내에 클립들로 구성된다. 이와 같은 동영상콘텐츠의 구조를 XML의 계층구조 형태로 나타낸다. 동영상에 대한 입력정보로는 동영상 제목, 동영상 장르, 키워드, 주석, 동영상 생성일자, 요약정보 등이 있으며 원시 동영상화일 이름도 포함된다.

입력된 데이터에 대해 XML 계층구조가 만들어지고, 검색 요청이 오면, 검색모듈에서는 Xpath를 통한 질의에 대한 결과를 이용자 인



<그림 6> 동영상 검색 단계

<표 1> XML 계층구조에 대한 Xpath 적용 예

동영상에 대한 메타데이터에 대한 Xpath	/mpeg7/VideoRoot/VideoRootMetaData
세그먼트 데이터에 대한 Xpath	/mpeg7/ VideoRoot/Video/VideoRoot/Segment
세그먼트의 메타 데이터에 대한 Xpath	/mpeg7/ VideoRoot/Video/VideoRoot/SegmentMetaData
클립의 메타데이터에 대한 Xpath	/mpeg7/VideoRoot/Video/VideoRoot/Segment/VideoRoot/ClipMetaData
클립 데이터에 대한 Xpath	/mpeg7/VideoRoot/Video/VideoRoot/Segment/VideoRoot/Clip

터페이스를 통해서 보여주게 된다.

<표 1>은 검색 모듈에서 동영상콘텐츠에 대한 이용자의 검색 요청시 생성될 수 있는 XML 계층구조에 대한 Xpath를 적용한 사례를 보여주고 있다. 즉, 동영상 검색을 위해 만들어진 XML 계층구조에 대한 정보를 기반으로 이용자의 질의에 의하여 요청되는 요소에 대한 Xpath를 생성

할 수 있다.

이와 같은 방법을 통하여 검색된 결과내역은 클립의 시작시간, 종료시간, 해당 클립의 키워드 리스트 및 대표이미지를 제공한다. 검색결과로 제공하는 클립의 대표이미지를 클릭하면 Microsoft의 windows media player를 통해 동영상의 해당 클립을 볼 수 있다. 이를 위해 본

연구에서는 하나의 원본 동영상을 보유하고 해당 클립을 스트리밍 서비스를 제공할 수 있는 ASX 파일을 만들어 제공한다. 즉, 대표이미지를 클릭하면 대표이미지와 연결되어 있는 ASX 파일이 활성화되어 스트리밍서버는 ASX 파일에 기술되어 있는 대로 이용자가 요구한 클립에 해당하는 원본 동영상의 일부분을 전송한다.

4.6 사용자 인터페이스

<그림 7>에서는 본 연구에서 구현된 동영상 검색 시스템의 사용자 인터페이스로서 제안된 동영상 검색시스템을 이용하여 “대통령”이라는 키워드를 포함하고 있는 클립을 검색한 결

과이다. 특히, 사용자 인터페이스모듈은 검색 모듈부터 전달된 결과 데이터를 이용자가 알기 쉽도록 보여 주는 역할과 이용자의 동영상콘텐츠에 대한 검색 질의를 검색 모듈에 직접 전달하는 역할을 수행한다. 한편, XSLT³⁾를 활용하여 XML 파일을 검색하기 위해서는 이전과정에서 생성된 Xpath를 통하여 검색모듈에서 이용자의 질의어에 해당하는 검색결과를 이용자 인터페이스에 보여지게 된다.

또한, 같은 데이터 결과에 대해서 다른 XSL⁴⁾를 적용함으로써 검색결과를 이용자가 원하는 형식으로 보여줄 수 있다. 즉, <그림 7>에서는 “대통령”이라는 이용자 질의어에 대한 검색결과에 대한 대표이미지를 검색결과로서 제공하



<그림 7> 검색 결과의 예

- 3) XSLT는 한 XML 문서 구조를 다른 구조를 갖는 XML 문서로 변환하는 방법을 기술하기 위한 표준적인 방법으로, W3C의 권고사항 중 하나이다. XSLT는 XSL의 확장판이라고도 생각할 수 있다. XSL은 XML 문서, 예를 들어 XML 문서 내에 기술되어 있는 데이터가 어떤 방식으로 웹페이지 내에 표현되어야 하는가를 보여주는 문서를 형식화하기 위한 언어이다. XSLT는 XML 문서가 다른 데이터 구조로 어떻게 재구성되어야 하는지를 보여준다.
- 4) XSL은 XML을 사용하고 있는 웹을 통해 보내어지는 데이터가 이용자에게 어떻게 보일 것인지를 설명한 스타일 시트를 만드는데 사용되는 언어이다. 예를 들면, 보험회사의 자동차에 대해 하나 이상의 특성을 묘사하고 있는 XML 페이지에서, “<automfg>”를 열고 닫는 한 쌍의 태그는 자동차 제작회사의 이름을 포함하고 있을 수 있다. XSL을 사용하면, 표시되어야 할 자동차 제작회사의 이름과, 페이지 상에 표시되어야 할 위치, 그리고 그것이 진한 글꼴로 표시되어야 한다는 사실 등을 웹브라우저에게 전해줄 수 있을 것이다. XSL은 DSSSL 및 CSS 표준에 기반을 두고 확장되었다.

는 방법을 보여주고 있다. 즉, 검색된 결과는 “대통령”이라는 질의어를 메타데이터로서 포함하고 있는 동영상 클립을 검색하고 이에 대한 대표이미지를 최종 결과물로서 사용자 인터페이스에 표현하고 있다. 한편, 본 연구에서는 검색결과를 표현하는 방법으로서 <그림 7>과 같이 대표이미지를 표현하는 방법과 검색된 결과에 대한 요약정보를 테이블의 형식으로 표현하여 보여줄 수 있는 기능을 제공하고 있기 때문에 이용자는 이용자의 편의에 따라 검색결과에 대한 표현을 선택할 수 있도록 하였다.

5. 결론

본 연구에서는 동영상콘텐츠에 대한 검색효율성을 높이기 위하여 해당 콘텐츠에 대한 이미지 정보와 자막정보를 이용하여 동영상콘텐츠를 여러 개의 클립으로 분리하고, 키워드를 이용하여 원하는 클립을 검색하는 동영상콘텐츠 검색시스템을 제안하였다. 이를 위하여 동영상콘텐츠에 대응되는 자막추출을 위하여 자막자체와 해당자막이 표현되는 시간을 검출하는 모듈을 개발하였다. 이를 기반으로 개별 동영상콘텐츠의 하위요소인 클립에 대한 키워드를 추출하고 이를 해당 클립에 대한 메타데이터로 활용하였다. 동영상콘텐츠의 검색성능의 향상에 대한 고려와 함께, 본 연구에서 제안하고 있는 동영상 검색시스템에서는 이용자의 검색 및 결과에 대한 편의성을 위하여 사용자 인터페이스에 있어서 이용자가 자신이 원하는 동영상콘텐츠에 대한 키워드를 입력하면 즉시 해당하는 클립에 대한 정보를 볼 수 있고, 대표이

미지를 클릭하면 선택한 클립을 실시간으로 스트리밍기법을 통하여 제공한다. 이를 위하여 각각의 동영상콘텐츠를 클립이라는 하위요소로 분할하여 검색 및 검색결과를 표현할 수 있도록 하고 있기 때문에 검색된 동영상콘텐츠에 대하여 이용자가 원하는 부분만을 빠르게 검색할 수 있다. 본 연구에서는 XML 계층구조를 이용한 동영상 검색 시스템에 대하여 전체적인 구성과 하위 모듈에 대한 기능을 살펴보았다. 특히, 동영상 데이터와 관련 메타데이터에 대한 정보를 포함하고 있는 XML 화일을 생성하고 이에 대해 데이터를 관리하고 이용자의 검색 요청시에 Xpath 경로를 구성하여 XML DBMS를 이용하여 검색하고, 해당 결과를 XSLT를 이용하여 보여주도록 하였다. 이를 통하여 XML 기반의 데이터 변환 및 데이터 관리를 수행함으로써 원본 동영상콘텐츠와 같이 동영상, 세그먼트, 클립 등과 같은 반복적으로 구성된 데이터에 대해 효과적인 데이터관리 및 검색이 가능하다. 한편, 본 연구에서 제안하고 있는 검색방법을 보다 효과적으로 만들기 위해서는 이미지 컷을 클립형식으로 통합하는 기법을 뉴스 이외의 다양한 응용분야에서도 이용할 수 있도록 발전시켜야 한다. 이를 위해서는 이미지의 특성과 키워드를 이용한 클립 검출 분야에 대한 보다 세밀한 검색기술을 강화시키고 동영상의 주요 구성요소 중의 하나인 오디오 정보를 활용하는 기술을 개발하여야 한다. 즉, 동영상을 구성하는 오디오의 에너지나 패턴과 같은 오디오 고유의 특성을 추출하고 이를 컷이나 클립의 검출에 이용하면 보다 정확한 클립분할이 가능하여 검색 성능을 높일 수 있을 것이다. 또한, 본 연구를 기반으로 추가적인 연구분야

로서 동영상콘텐츠의 효율적인 관리와 검색을 위하여 실제 동영상콘텐츠의 저작 시스템과의 연동을 위한 연구가 요구되며 한편, 보다 높은

검색 성능을 향상하기 위해서는 분야별 동영상 콘텐츠에 대한 체계화된 메타데이터에 대한 연구가 요구된다.

참 고 문 헌

- 김진아, 김태수. 2001. 동영상 정보의 메타데이터 구축에 관한 연구. 『한국정보관리학회지』, 18(4): 15-35.
- 김택근, 김우생. 2004. MPEG-7에 기반한 동영상 검색 시스템을 위한 계층형 인덱스 기법. 『한국통신학회논문지』, 29(10C): 1444-1450.
- 남윤영, 황인준. 2002. XMARS: XML 기반 멀티미디어 주석 및 검색시스템. 『한국정보처리학회 논문지』, 9-B(5): 541-548.
- 안형근, 고재진. 2006. 의미적 메타데이터를 이용한 멀티미디어 주석 및 검색. 『한국정보과학회 2006 가을 학술발표논문집』, 33(2C): 199-204.
- 윤소영, 문성빈. 2006. 엘리먼트 기반 XML 문서검색의 성능에 관한 실험적 연구. 『한국정보관리학회지』, 23(1): 201-219.
- Farag, W. E. and H. A. Wahab. 2003. "A human-based technique for measuring video data similarity." *Proc. of IEEE on International Symposium on Computers and Communication*, 2: 768-774.
- Kim, Y. T. and T. S. Chua. 2005. "Retrieval of news video using video sequence matching." *Proc. of IEEE Conf. on Multimedia Modeling*, 68-75.
- Chen, J. F., H. Y. Mark Liao, and C. W. Lin. 2005. "Fast video retrieval via the statistics of motion." *Proc. of IEEE Conf. on Acoustics, Speech and Signal Processing*, 2: 437-440.
- Ardizzone, E., M. La Cascia, A. Avanzato, and A. Bruna. 1999. "Video indexing using MPEG motion compensation vectors." in *IEEE international Conference on Multimedia Computing and Systems*, 2: 725-729.
- Swain, M. and D. Ballard. 1991. "Color indexing." *International Journal of Computer Vision*, 7(1): 11-32.
- Leong, Mun Chak and N. Selvanathan, "Image database query by content using wavelet coefficient signature." *Proceeding of AIRTC'97 - IFAC International Symposium on Artificial Intelligence & Real-time Control*.
- Scheunders, P., S. Livens, G. Van de Wouwer, P. Vautrot, and D. VanDyck, 1998. "Wavelet-based texture analysis." *International Journal Computer Science and Information management*, 1(2): 22-34.