

# 통합된 비디오 인덱싱 방법을 이용한 내용기반 비디오 데이터베이스의 설계 및 구현

(Design and Implementation of Content-based Video  
Database using an Integrated Video Indexing Method)

이 태 동 \* 김 민 구 \*\*

(Tae-Dong Lee) (Min-Koo Kim)

**요 약** 오늘날 멀티미디어 정보의 양이 매우 빠른 속도로 증가함에 따라 비디오 데이터베이스에 대한 효율적인 관리는 더욱 중요한 의미를 가지게 되었다. 그리고 초고속 정보통신망과 디지털 기술의 발전은 비디오 데이터 통신 및 컴퓨터와 결합하여 새로운 멀티미디어로 발전하고 있으며, 인터넷 방송, 주문형 비디오(VOD) 등에 크게 활용하고 있다. 비디오는 대용량적인 특성과 비정형적인 특성을 가지고 있으므로 신속하고 효율적으로 비디오를 검색하기 위해 비디오의 정확한 특징정보를 추출하여 비디오 데이터베이스를 구축하여야 한다. 비디오 데이터베이스는 텍스트 기반의 전통 데이터베이스와는 다른 모델링 방법과 검색방법을 사용한다. 따라서, 비디오 데이터베이스에서의 검색속도와 정확도를 향상시키기 위해서는 새로운 비디오 데이터베이스 생성기법과 효율적인 검색기법이 필요하다. 본 논문에서는 비디오의 의미적 구조와 사전 제작지식정보를 구조적으로 축적할 수 있는 내용기반 비디오 데이터베이스의 구축 방안과 생성기법을 제시하였다. 그리고 제안된 비디오 데이터베이스의 구축 방안과 생성기법을 사용하여 새로운 인터넷 방송 프로그램 콘텐츠 제작에 활용할 수 있는 비디오 데이터베이스를 구현하였다. 이를 위해 비디오 분할과 대표키 프레임 추출 시 비디오의 의미적 구조와 사전 제작지식정보의 상호관계관 기반으로 하여 비디오 데이터의 특징정보를 추출하고, 검색할 수 있도록 주석기반 검색과 내용기반 검색을 통합한 비디오 인덱싱 방법을 제시하였다. 통합된 비디오 인덱싱 방법은 비디오의 하위 레벨에 표현된 내용기반 메타데이터 유형과 비디오의 특징정보 추론이 어려운 상위 레벨에 표현된 주석기반 메타데이터 유형을 동시에 이용하므로 콘텐츠 검색의 성능을 향상시킬 수 있다. 마지막으로 본 논문에서 제시한 비디오 데이터베이스는 비디오의 의미적 구조와 사전 제작지식정보를 구조적으로 축적하여 데이터베이스를 구축하므로 정확한 인터넷 방송 콘텐츠 정보의 축적관리와 구축작업의 효율화가 가능하며, 또한 인터넷 방송 콘텐츠 제작 시 정보공유 및 재이용이 가능하므로 새로운 콘텐츠 제작의 효율성을 높일 수 있다.

**Abstract** There is a rapid increase in the use of digital video information in recent years, it becomes more important to manage video databases efficiently. The development of high speed data network and digital techniques has emerged new multimedia applications such as internet broadcasting, Video On Demand(VOD) combined with video data processing and computer. Video database should be construct for searching fast, efficient video be extract the accurate feature information of video with more massive and more complex characteristics. Video database are essential differences between video databases and traditional databases. These differences lead to interesting new issues in searching of video, data modeling. So, cause us to consider new generation method of database, efficient retrieval method of video. In this paper, We propose the construction and generation method of the video database based on contents which is able to accumulate the meaningful structure of video and the prior production information. And by the proposed the construction and generation method of the video database implemented the video database which can produce the new contents for the internet broadcasting centralized on the video database. For this production, We proposed the video

\* 정 회 원 : 경운대학 컴퓨터응용과 교수

LTD@kmc.ac.kr

\*\* 통신회원 : 아주대학교 정보및컴퓨터공학부 교수

minkoo@madang.ajou.ac.kr

논문접수 : 2001년 4월 20일

심사완료 : 2001년 8월 24일

indexing method which integrates the annotation-based retrieval and the content-based retrieval in order to extract and retrieval the feature information of the video data using the relationship between the meaningful structure and the prior production information on the process of the video parsing and extracting the representative key frame. We can improve the performance of the video contents retrieval, because the integrated video indexing method is using the content-based metadata type represented in the low level of video and the annotation-based metadata type impressed in the high level which is difficult to extract the feature information of the video at the same time. As a conclusion, the video database proposed in this paper can improve the efficiency of the construction process and the accurate management of the accumulated information for the internet broadcasting contents because the proposed method do construct the database by accumulating the meaningful structure and the prior production information systematically, and can increase the efficiency of the new contents production because the information sharing and reuse during the production process of the internet broadcasting contents is possible.

## 1. 서 론

### 1.1 연구의 배경 및 목적

오늘날 통신 및 컴퓨터 기술의 발달로 미디어의 새로운 혁명인 멀티미디어 시대에 도래하게 되었다. 최근 인터넷의 폭발적인 보급과 웹방송의 출현에 힘입어 인터넷 방송은 간단한 제작 도구와 압축 방식, 그리고 전송 상 초고속 전용회선만 있으면 가능하다는 점에서 빠른 성장 추세를 보이고 있다. 인터넷 방송의 주문형 비디오에 대한 서비스 예로는 TV중계, 원격수업, 뉴스전송, 원격진료 등과 같은 다양한 응용서비스들을 들 수 있다. 인터넷 방송은 다양한 멀티미디어 내용을 시간과 공간의 제약 없이 양방향의 대화형 데이터 방송에 대한 가능성을 부각시켰으며, 사용자가 원하는 콘텐츠에 대한 실시간 서비스를 할 수 있는 길을 열었다[1]. 인터넷 방송 콘텐츠 표현 기술의 대표적 모델은 현재 비디오를 표현하는 웹 콘텐츠가 될 것이다[2]. 인터넷 방송 콘텐츠 서비스는 전적으로 방송국에 의존하기보다는 이용자 측에도 상당 부분 권한이 있기 때문에 이들에게 다양한 채널과 내용을 자유롭게 선택할 권한이 있다. 따라서 인터넷 방송은 멀티미디어 형식으로 구성된 다양한 콘텐츠를 데이터베이스화하여 사용자에게 효율적으로 서비스하기 위해 반드시 필요하다. 멀티미디어 형식으로 구성된 다양한 콘텐츠 정보를 데이터베이스화하여 정보를 공유하고 재이용하므로 인터넷 방송 콘텐츠 제작의 효율성을 높일 수 있다. 또한 인터넷 방송 콘텐츠 제작 시 현재 충분히 축적, 관리되고 있지 않은 비디오의 의미적 구조와 제작과정의 사전 제작지식정보를 구조적으로 축적하여 데이터베이스화하므로 다양한 콘텐츠 서비스를 사용자에게 효율적으로 부가할 수 있다. 이러한 데이터베이스화는 비디오의 의미적 구조와 제작과정의 사전

제작지식정보(문자를 중심으로 이루어지는 정보와 비디오의 조작성을 중심으로 이루어지는 정보)를 기반으로 실현할 수 있다.

본 논문에서는 인터넷 방송 콘텐츠 제작과정에서 현재 충분히 축적, 관리되고 있지 않은 비디오의 의미적 구조와 제작지식정보를 구조적으로 축적할 수 있는 내용기반 비디오 데이터베이스의 구축 방안과 생성기법을 제시한다. 그리고 제안된 비디오 데이터베이스의 구축 방안과 생성기법으로 비디오 데이터베이스를 구축하여 새로운 인터넷 방송 프로그램의 콘텐츠 제작에 활용할 수 있도록 한다. 또한 비디오의 컷 검출과 분할, 그리고 대표키 프레임 추출 시 비디오의 의미적 구조와 사전 제작지식정보의 상호관계를 기반으로 비디오 데이터의 특징정보를 추출하고, 검색할 수 있도록 주석기반 비디오 검색과 내용기반 비디오 검색을 통합한 비디오 인덱싱 방법을 제시한다. 이를 위해 각 콘텐츠 제작과정의 정보공유에 의해 비디오의 의미적 구조와 제작과정의 사전 제작지식정보를 메타데이터로 정의하여 사용한다. 통합 비디오 인덱싱 방법은 하위 레벨에 표현된 내용기반 비디오 메타데이터 유형과 비디오의 특징정보 추출이 어려운 상위 레벨에 표현된 주석기반 비디오 메타데이터 유형을 함께 비디오 인덱스로 구성하여 제공하므로 비디오 검색의 성능을 향상시킬 수 있다. 본 논문에서 제시하는 비디오 데이터베이스는 비디오의 의미적 구조와 제작과정의 사전 제작지식정보를 구조적으로 축적하여 데이터베이스를 구축하므로 정확한 인터넷 방송 콘텐츠 정보의 추적관리와 구축작업의 효율화가 가능하며, 또한 인터넷 방송 콘텐츠 제작 시 정보공유 및 재이용이 가능하므로 새로운 콘텐츠 제작의 효율성을 높일 수 있다. 본 논문의 연구 목적은 비디오의 의미적 구조와 제작과정의 사전 제작지식정보를 구조적으로 축적하

어 인터넷 방송 프로그램의 새로운 콘텐츠 제작을 지원할 수 있도록 비디오 데이터베이스 시스템을 설계하고 생성기법을 연구하는데 그 목적이 있다.

### 1.2 논문의 구성

본 논문의 구성은 다음과 같다. 2장에서는 관련연구를 통하여 먼저 비디오 제작과정의 정보흐름을 살펴본 후 비디오 제작 과정에서 수집된 제작과정의 사전 제작지식정보를 구조적으로 축적할 수 있는 방법을 기술한다. 그리고 비디오 데이터베이스 시스템의 사례 연구에 대해 살펴보고 비디오 데이터베이스에 필요한 비디오 인덱싱 기법에 관하여 기술한다. 3장에서는 인터넷 방송 콘텐츠 제작 시 현재 충분하게 축적, 관리되고 있지 않은 비디오의 의미적 구조와 제작과정의 사전 제작지식정보를 구조적으로 축적하여 인터넷 방송의 새로운 콘텐츠 제작을 지원하는 내용기반 비디오 데이터베이스의 구축 방안과 생성기법을 제시한다. 그리고, 4장에서는 본 논문에서 제시하여 구현한 비디오 데이터베이스 시스템과 검색 시스템에 대한 실험결과를 기존의 사례연구와 비교 분석하고, 마지막으로 5장에서는 결론 및 향후 연구과제를 제시한다.

## 2. 관련 연구

최근 신속하고 효율적인 비디오 데이터베이스화에 관한 연구가 활발히 진행되고 있다. 이 절에서는 관련 연구를 통하여 비디오 제작과정의 정보흐름을 살펴본 후 비디오 제작과정에서 수집된 사전 제작지식정보를 구조적으로 축적할 수 있는 방법을 기술한다. 그리고 비디오 데이터베이스 시스템의 사례 연구에 대해 살펴보고 비디오 데이터베이스에 필요한 비디오 인덱싱 기법에 관하여 기술한다.

### 2.1 비디오 제작정보에 관한 연구

인터넷 방송 콘텐츠 제작과정에서 사용되고 있는 사전 제작지식정보를 통합하여 구조적으로 축적하려면 비디오 영상 제작 시 어떠한 작업이 이루어지며, 어떻게 정보가 취급되고 있는가를 분석하여야 한다. 일반적으로 비디오 영상 제작과정은 크게 'Pre-Production 단계', 'Production 단계', 'Post-Production 단계'의 총 3단계로 구분된다 [3]. 비디오 영상을 제작할 때 순서나 절차의 결정은 제작의 효율과 작품의 질을 결정하는 중요한 요소가 된다 [4]. 비디오 영상 제작과정의 Pre-Production 단계에서는 어떤 콘텐츠를 제작할 것인가를 기획하는 일, 콘텐츠 테마에 관한 정보수집과 취재하는 일, 콘텐츠를 구성하는 일, 대본과 콘티(Continuity) 등을 작성하는 단계이다. Pre-Production 단계에서 콘텐츠 제작 준비가 끝나

면 비디오 촬영에 들어가게 된다. 비디오 촬영은 작성된 콘티정보에 의해 카메라로 비디오를 수록하는 단계이다. 그리고 마지막 단계인 Post-Production 단계에서는 촬영된 필름이나 테이프를 가지고 편집 및 녹음 등의 과정을 거친 후 최종적으로 콘텐츠를 완성하게 된다.

이와 같이 콘텐츠 제작정보의 흐름절차에서 크게 두 가지의 정보 영역으로 구분할 수 있다는 것을 알 수 있다. 하나는 콘텐츠 제작과정에서 문자를 중심으로 이루어지는 정보영역이고, 또 다른 하나는 비디오의 조작용 중심으로 이루어지는 정보영역이다. 문자를 중심으로 이루어지는 정보는 제작할 콘텐츠의 의도와 그 의도를 비디오로 구성하기 위해 필요한 소재 리스트, 취재자료, 대본, 콘티, 구성표, 출연자 정보, 광고자료, 문헌, 뉴스, 서류 등의 정보이다. 그리고 비디오의 조작용 중심으로 이루어지는 영상소재 자료, 비디오 정보 등은 디지털 콘텐츠 제작 기기 또는 제작 소프트웨어로 작업이 이루어지는 정보이다.

### 2.2 비디오 인덱싱 기법에 관한 연구

비디오 정보검색 기술은 비디오 색인을 위한 비디오 분할(Video Parsing) 기술, 방대한 비디오 데이터를 저장하기 위한 비디오 데이터 압축 및 저장 기술, 비디오 검색 시 사용자 인터페이스를 위한 브라우징 방법 등을 핵심 요소로 하고 있다. 비디오 데이터는 영상 및 음성 정보뿐 아니라 문자 정보를 비롯한 여러 가지의 의미정보들을 담고 있으며, 또한 시각적 관계, 시간적 관계, 공간적 관계, 복합적 관계, 비정형적이고 대용량적인 관계의 특징 정보들을 동시에 가지고 있으므로 비디오 데이터는 기존의 문자 데이터 방식처럼 검색하기가 매우 어렵다. 이와 같은 비디오 데이터를 검색하기 위한 방법으로는 주석기반 검색 방법(Annotation-based Retrieval Method) [5,6,7,8,9,10,11,12]과 내용기반 검색 방법(Content-based Retrieval Method) [10,11,12,13,14,15,16]이 있다. 기존의 개발된 비디오 데이터베이스 시스템의 사례로는 다음과 같다. 1993년 일본 고베 대학에서 비디오 객체(Video Object)라는 개념을 도입하여 개발한 OVID [5], 1995년 노르웨이 공대에서 비디오 데이터와 메타데이터의 공유와 재사용을 지원하도록 개발한 VideoSTAR (Video Storage and Retrieval) [6], 1995년 미국 메릴랜드 대학에서 비디오의 상위 레벨에 표현된 객체(Object), 행위(Activity), 사건(Event) 등의 의미정보를 주석기반 메타데이터 유형으로 직접 부여하여 검색하도록 개발한 AVIS (Advanced Video Information System) [7], 1994년 미국 MIT 대학에서 비디오 정보를 모델링하여 주석기반 비디오 검색을 지원하기 위해 개발한 Algebraic

Video[8], IBM Almaden 연구소에서 개발한 QBIC (Query by Image and Video Content)[10,11], 미국의 콜롬비아 대학에서 개발한 VideoQ[12], 1996년 대만 청화 대학에서 개발한 Venus[13], 미국 카네기멜론 대학에서 개발한 Informedia[14] 등이 있다.

비디오 데이터를 표현하고 검색하기 위한 첫 번째 방식은 주석기반 검색 방법으로 각각의 비디오 데이터에 자동 인식이 어려운 의미정보를 수동으로 먼저 파악한 후, 이를 사용자가 직접 문자로 주석을 부여하여 저장한 후 검색 시 미리 부여된 주석을 이용하여 검색하는 방법이다[5,6,7,8,9]. 이 방법은 사용자가 직접 비디오를 보면서 내용을 주석으로 처리할 수 있기 때문에 비디오의 내용을 정확하게 표현하고 검색할 수 있다는 장점이 있다. 그러나, 각각의 비디오에 대해 문자로 사용자가 일일이 주석을 부여하여야 하므로 많은 시간과 노력을 필요로 하며 주석의 양이 방대하게 증가 할 수 있다는 단점이 있다. 또한 비디오 내용을 사용자의 관점에 따라 주관적으로 주석을 부여할 수 있으므로 검색 시 정확한 결과를 생성하기 어렵다는 단점이 있다. 그리고, 비디오 데이터의 시각적인 의미정보를 정확하게 문자로 표현하기가 어렵다는 단점이 있다.

두 번째 방식은 내용기반 비디오 검색 방법으로 비디오 데이터 자체로부터 그 데이터가 가지고 있는 특징정보들을 자동 또는 반자동적으로 추출하여 검색하는 방법이다[13,14,15,16]. 내용기반 검색 방법은 비디오 자체의 시각적 특징, 즉 비디오 특징정보를 자동 또는 반자동적으로 추출해 검색함으로써 기존의 주석기반 검색 방법의 단점을 해결할 수 있다. 그러나, 비디오의 특징정보를 정확하게 추출하기 위한 알고리즘 개발과 정확한 분류가 어렵다는 단점이 있다.

세 번째 방식은 주석기반 비디오 검색 방법과 내용기반 비디오 검색 방법을 통합하여 사용하는 방법이다[10,11,12]. 즉, 비디오 프레임의 각 객체들 사이의 시간 관계, 공간 관계 등의 하위 레벨 특징정보들은 내용기반 비디오 검색 방법을 이용하여 자동 또는 반자동적으로 특징정보를 추출해 내고, 그리고 자동 또는 반자동으로 추출이 불가능한 사건, 장소 등의 상위 레벨 의미정보들은 주석기반 비디오 검색 방법을 이용하여 사용자가 직접 의미정보를 기술해 주는 방법이다. 이 검색 방법은 비디오 데이터베이스를 구현함에 있어서 서로 보완적인 역할을 한다.

#### 2.2.1 기존의 제안된 비디오 컷 검출 기법

지금까지 비디오 구조의 컷 검출 기법들은 다양하게 연구되어 왔다. 컷 검출(Cut Detection) 또는 장면전환

검출(Scene Change Detection)은 샷의 경계를 구분해 내는 비디오 분할의 기본 단위로 보통 카메라 샷에 의해 처리된다. 비디오의 장면전환은 점진적 전환(Gradual Transition)과 순간적 전환(Abrupt Transition)의 두 가지로 분류된다. 점진적 전환은 비디오의 다양한 편집 효과(Fade-In, Fade-Out, Dissolve 등)에 의해 시작과 끝이 분명하게 구분되지 않는 장면전환을 의미한다. 점진적 전환은 샷 또는 장면 사이의 경계가 불분명하므로 검출하기가 어려우며, 특히 카메라의 촬영기법인 패닝(Panning)이나 주밍(Zooming)과 같은 특수효과 장면은 경계가 매우 불분명하므로 컷 검출에 많은 어려움이 따른다. 순간적 전환은 하나의 샷이 끝난 직후 바로 다음 샷을 이어 붙인 컷을 의미한다. 순간적 전환은 컷의 경계가 분명하므로 인접한 프레임들 사이의 픽셀 값 차이를 이용하거나 히스토그램(Histogram)의 차이를 이용하는 등의 통계적인 방법만으로도 경계 검출이 가능하다. 기존의 제안된 컷 검출 기법으로는 화소단위 비교법[17, 18,19], 유사율 측정법[20], 히스토그램 비교법[17,21,22, 23], 이중 비교법[22], 움직임 벡터 비교법[24], 에지 변화 비교법[25], 그리고 압축 상관 계수 비교법[26,27] 등에 기초를 두고 있다.

#### 2.2.2 기존의 제안된 비디오의 대표키 프레임 및 특징 정보 추출 기법

일반적으로 대표키 프레임 추출 기법은 컷으로부터 키 프레임들을 추출한 다음 내용이 유사한 키 프레임들을 그룹화하거나 삭제함으로써 대표키 프레임들을 추출할 수 있다. 기존의 대표키 프레임 추출 기법에 관한 몇 가지 연구사례를 살펴보면 다음과 같다.

일반적으로 대표키 프레임 추출 방법 중 가장 간단한 방법은 비디오 시퀀스의 내용에 관계없이 각 컷마다 임의의 한 프레임을 추출하여 이를 대표키 프레임으로 이용하는 방법이다. 이러한 방법으로 비디오 시퀀스의 내용에 관계없이 샷 내의 임의의 한 프레임을 추출하는 방식과 1초에 한 프레임씩 대표키 프레임을 추출하는 방식을 [28]에서 제안하였다.

[29]에서는 샷 경계의 검출 없이 비디오 영상의 정보만으로 대표키 프레임들을 추출하는 방법을 제안하였다. 이 방법은 먼저 일정한 단위로 프레임들을 나눈 후 각 단위의 특징정보 차이가 큰 그룹과 작은 그룹으로 나눈 다음, 특징정보 차이가 작은 그룹의 프레임들 중 첫 번째 프레임과 제일 마지막 프레임만 남기고 나머지 프레임들을 삭제한 후 다시 작은 그룹과 큰 그룹을 병합하여 이들을 다시 그룹화하고 삭제 작업을 반복하여 사용자가 원하는 수만큼의 대표키 프레임을 추출하도록 하였다.

[30]에서는 비디오 색인을 위한 분할 과정에서 프레임들 간의 특징정보 차이를 계승하여 대표키 프레임뿐만 아니라 관련된 키 프레임들도 동시에 추출하는 방법을 제안하였다. 이 기법은 n개의 대표키 프레임을 추출하기 위하여 먼저 비디오 클립을 균등하게 분할한 후 m(m>n)개의 프레임을 추출하여 추출된 프레임들 중 특징정보 차이가 가장 작은 연속된 두 개의 프레임을 선택하고, 이 두 프레임을 해당 프레임 그룹의 중간 프레임으로 대치하여 사용자가 요청한 n개의 대표키 프레임이 남을 때까지 반복적인 형태로 수행하여 대표키 프레임을 추출하도록 하였다.

### 3. 비디오 데이터베이스 시스템

이 장에서는 인터넷 방송 콘텐츠 제작과정에서 현재 충분하게 추적, 관리되고 있지 않은 비디오의 의미적 구조와 제작지식정보를 구조적으로 추적할 수 있는 내용기반 비디오 데이터베이스의 구축 방안과 생성기법을 기술한다.

#### 3.1 비디오 데이터베이스 시스템의 전체적인 구조

인터넷 방송 콘텐츠 제작을 위한 비디오 데이터베이스의 전체적인 시스템 구조는 그림 1과 같다. 그림 1과 같이 시스템의 전체적인 구조는 (1) Video Data Capture & Non-Linear Video Edition 기능 (2) Video Indexing Processing 기능 (3) Video Content Management 기능 (4) Video Encoding 기능 (5) User Interface 기능(① Image Processing ②Video Browsing ③Video Query Processing(주석기반 질의, 내용기반 질의, 주석기반 및 내용기반 통합 질의의 세 가지 방식) ④Video Playing) (6) Data Storage 기능 (7) Internet Broadcasting 기능으로 구분한다.

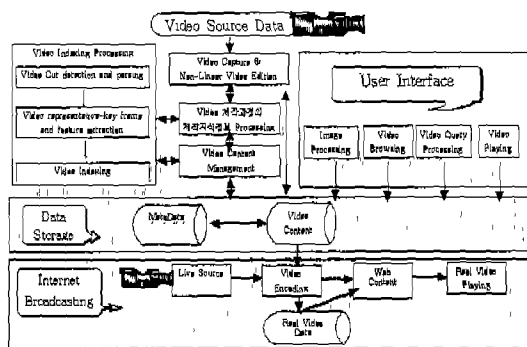


그림 1 비디오 데이터베이스 시스템의 전체적인 구조

#### 3.2 비디오 데이터 모델

인터넷 방송 콘텐츠 제작은 제작과정의 콘티(Continuity) 정보를 이용하여 제작할 콘텐츠의 의도와 그 의도를 비디오로 구성한다. 콘티정보는 촬영을 위한 구체적인 카메라 사이즈와 앵글, 카메라 움직임 그리고 배우의 동선 등을 컷별로 기록한다. 방송 콘텐츠 제작 시 이들 콘티정보에 의해 비디오 영상이 촬영되며, 촬영된 비디오는 편집과 녹음 과정을 거쳐 최종적으로 콘텐츠 제작이 완성된다. 콘티정보는 촬영대본 정보이며, 촬영을 위한 설계도이다. 촬영에 대한 모든 구상이 콘티정보에 있으므로 다른 어떤 정보들보다 더 중요하다는 것은 명백하다 [3]. 특히 여러 명의 제작팀이 함께 작업을 행하는 콘텐츠 제작에서는 콘텐츠의 제작목적과 비디오 촬영 등의 정보를 제작자간에 공유되어야 하기 때문에 콘티정보는 반드시 필요하다. 콘티정보를 기록하는 양식으로는 가로형과 세로형이 있으며, 그림 2는 가로형 콘티정보기록 양식의 예이다.

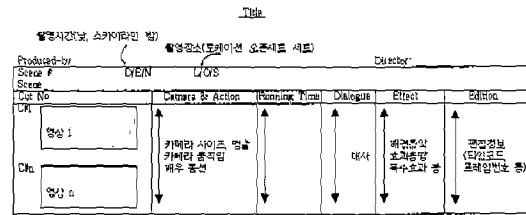


그림 2 가로형 콘티정보기록 양식 예

그림 2와 같이 비디오 제작과정의 콘티정보는 컷별로 제작할 콘텐츠의 의도와 그 의도를 비디오로 구성하기 위해 장면번호(Scene #), 장면(Scene), 컷 번호(Cut #), 컷(Cut), 촬영시간(D/E/N), 촬영장소(L/O/S), 카메라 사이즈와 앵글(Camera Size and Angle), 카메라 움직임(Camera Work), Produced-by, Director, Action & Dialogue, Running Time, Effect 등의 촬영 및 편집정보요소들로 구성된다. 이들 콘티정보요소 중 촬영에 필요한 정보요소들을 구체적으로 살펴보면 다음과 같다. (1) 카메라 사이즈 정보요소는 B.C.U(Big Close Up), C.U(Close Up), C.S(Close Shot), B.S(Bust Shot), W.S(Waist Shot), M.S(Medium Shot), F.S(Full Shot), E.L.S(Extreme Long Shot), One Shot, Two Shot, Three Shot, Group Shot 등으로 구분된다. (2) 카메라 앵글 정보요소는 촬영 시 카메라와 피사체 사이의 수직적, 수평적 각도를 나타낸 카메라 앵글 정보요소로 카메라

라 앵글의 수직적 관계에서 수평 앵글(Normal Angle 또는 Eye Level), 부각(High Angle), 양각(Low Angle)으로 구분된다. (3) 카메라 움직임 정보요소는 Panning Left & Right(콘티상 : Pan←, Pan→ 라고 표시), Panning Up & Down 또는 Tilt Up & Down(콘티상 : Pan↑, Pan↓ 또는 Tilt Up & Down으로 표시), Tracking(Truck Up & Back, Truck Left & Right, Follow, Crane Up & Down, Zoom In & Out) 등으로 구분된다.

본 논문에서 제시하는 비디오 데이터 모델은 비디오 데이터베이스 구축 시 신속하고 효율적으로 비디오 데이터를 접근하도록 하기 위해 응용 분야의 특성에 맞도록 설계된 비디오 데이터에 대한 표현 방법으로 비디오 데이터를 구조적으로 체계화하여 메타데이터를 정의한 비디오 데이터베이스의 스키마이다. 대용량의 비디오 데이터를 효율적으로 데이터베이스화하기 위해서는 비디오 데이터를 구조적으로 체계화한 메타데이터(Metadata)가 반드시 필요하기 때문에 비디오 데이터를 구조적으로 체계화하기 위해 비디오의 의미적 구조를 분석하여야 한다. 일반적으로 비디오의 의미적 구조를 계층적으로 분류하면 다음과 같다. 하나의 비디오 시퀀스(Video Sequence)는 비디오 클립(Video Clip), 장면(Scene), 샷(Shot), 프레임(Frame) 등의 요소로 구성된다. 그리고, 하나의 비디오 클립은 하나 이상의 비디오 장면으로 구성되고, 비디오의 각 장면은 하나의 주제를 갖는 연속된 일련의 비디오 샷으로 구성된다. 즉, 비디오의 각 장면은 비슷한 의미를 내포하고 있는 특정 대상을 연속 촬영한 일련의 샷들의 그룹으로 하는 비디오의 단위이다. 비디오의 각 샷은 캠코더에 의해 촬영된 연속적인 영상들을 표현한 비디오의 기본 단위이다. 또한 비디오

의 각 샷은 끊어지지 않은 여러 개의 프레임들로 구성된다. 그림 3은 비디오의 계층적 구조를 그림으로 나타낸 것이다.

인터넷 방송의 비디오 콘텐츠는 제작과정의 제작지식 정보인 콘티정보에 의해 제작하고자 하는 비디오 콘텐츠의 의미적 구조와 영상내용을 사전에 미리 알 수 있다. 본 논문에서는 비디오 콘텐츠 제작과정의 콘티정보에 의해 비디오의 의미적 구조정보와 사전 제작지식정보를 기반으로 비디오 데이터베이스의 메타데이터를 표현한다. 이 방법은 먼저 비디오 데이터를 구조적으로 체계화하기 위하여 비디오 콘텐츠 제작과정의 콘티정보에 의해 비디오의 의미적 구조정보와 사전 제작지식정보를 비디오 시퀀스의 내용별로 분류하여 비디오 데이터베이스의 메타데이터로 모델링한다.

본 논문에서 메타데이터는 비디오 데이터의 내용을 구조적으로 체계화한 데이터이다. 본 논문에서 제시하는 비디오 데이터베이스의 메타데이터는 비디오 콘텐츠 제작과정의 콘티정보에 의해 비디오의 의미적 구조정보와 사전 제작지식정보(문자를 기반으로 이루어지는 정보와 비디오의 조작용을 기반으로 이루어지는 정보)를 기반으로 메타데이터를 분류할 수 있다. 따라서, 본 논문에서는 인터넷 방송 콘텐츠 제작과정의 비디오 데이터들 구조적으로 체계화하기 위하여 제작과정의 콘티정보요소와 제작과정의 문자정보 및 비디오 조작용정보들을 비디오 데이터베이스의 메타데이터로 표현하여 비디오 데이터베이스를 생성하는 방법을 제안한다. 표 1은 비디오 콘텐츠 제작과정의 콘티정보에 의해 비디오의 의미적 구조정보와 사전 제작지식정보를 기반으로 비디오 데이터베이스 시스템의 메타데이터를 모델링하여 분류한 것이다.

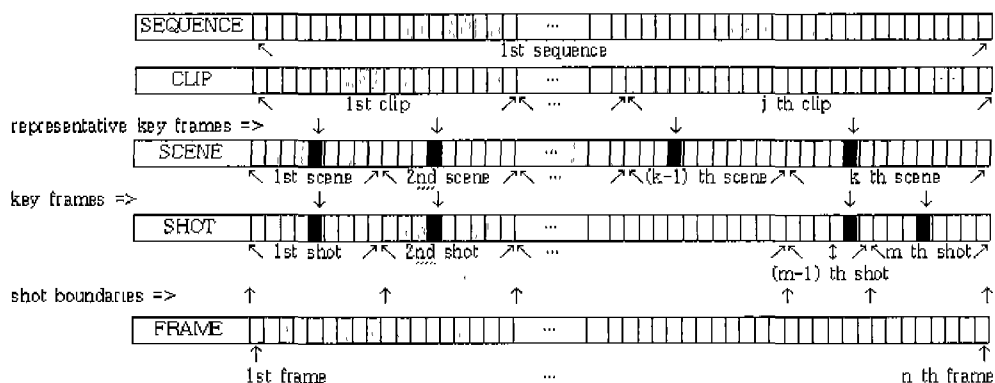


그림 3 비디오 시퀀스의 계층적 구조(Hierarchical Structure of Video Sequence)

표 1 비디오 데이터베이스의 메타데이터 모델

Metadata 분류	Metadata	내용
General Data	Identifier, Title, Subject, Director & Staff, 제작자, 제작일 등	비디오의 형식과 제작과정의 일반적인 정보 (General Feature)
Structural Data	Video Sequence Stream, Video Clip Stream, Video Cut Stream, Frame, Video Width, Video Height 등	비디오의 구조적인 특징정보(Structural Feature)
Temporal-Spatial & Sound and Visual Data	Begin, End, Duration, Time Code, Video Effect 등	비디오의 시공간적인 특징정보(Temporal-Spatial Feature)와 음향 및 시각적인 특징정보(Sound and Visual Feature)
Annotation & Content Data	Keyword, Video Sequence, Video Clip, Video Cut, Action & Dialogue, Camera Size, Camera Angle, Camera Work, Actor, Object Name, Object Type, Object Position, Object Feature 등	비디오의 주석기반 특징정보 및 내용기반 특징정보
Streaming Video Data	Real Video Stream 등	스트리밍(Streaming) 비디오 파일인 리얼 형식이나 도우 미디어 형식의 특징정보

표 2 비디오 데이터베이스의 메타데이터 클래스 및 속성 데이터 모델

Class	Attribute	Content
Video Sequence	Id, Title, Subject, Keyword, Duration, Time Code, Width, Height, Content, Director and Staff, Produced by, Date 등	Video Sequence Class의 Unique Identifier, Title, Subject, Keyword, Total Duration, End Feature Time Code, Height, Content, Director and Staff, 제작회사, 제작일자 등
Video Clip	Id, Pointer, Subject, Keyword, Content, Duration, Begin and End Time Code 등	Video Clip Class의 Unique Identifier, Pointer, Subject, Keyword, Content, Total Duration, Begin Frame and End Frame Time Code 등
Video Cut	Id, Pointer, Subject, Keyword, Content, Dialogue, Effect, Camera Size, Camera Angle, Camera Work, Actor, Duration, Begin and End Time Code 등	Video Cut Class의 Unique Identifier, Pointer, Subject, Keyword, Content, Dialogue, Effect(Music Effect, Sound Effect, Edition Effect etc.), Camera Size, Camera Angle, Camera Work, Actor, Total Duration, Begin Frame and End Frame Time Code 등
Video Object	Id, Pointer, Name, Type, Position, Event, Content, Feature, Time Code 등	Video Object Class의 Unique Identifier, Pointer, Object Name, Object Type, Object Position, Object Event, Object Content, Object Feature, Object Frame Time Code 등
Video Sequence Stream	Pointer, Video Sequence Stream, Real Video Sequence Stream 등	Video Sequence Class의 Pointer, Video Sequence의 Stream, Real Video Sequence의 Stream 등
Video Clip Stream	Pointer, Video Clip Stream, Real Video Clip Stream 등	Video Clip Class의 Pointer, Video Clip의 Stream, Real Video Clip의 Stream 등
Video Cut Stream	Pointer, Video Cut Stream, Real Video Cut Stream 등	Video Cut Class의 Pointer, Video Cut의 Stream, Real Video Cut의 Stream 등
Video Key frame	Pointer, Video Key Frame 등	Video Object Class의 Pointer, Representative-Key Frame 등

표 1의 비디오 데이터베이스 메타데이터는 비디오 데이터베이스의 메타데이터 클래스 및 속성 데이터로 활용하게 된다. 표 1에서와 같이 제작과정의 사전 제작지식 정보와 비디오의 의미적 구조에 의해 일반적인 특징정보 (General Feature), 구조적인 특징정보(Structural Feature), 시공간적인 특징정보(Temporal-Spatial Feature)와 음향·시각적인 특징정보(Sound and Visual Feature), 비디오의 주석기반 특징정보 및 내용기반 특징정보, 그리고 Streaming Video File 정보 등으로 분류하여 메타데이터를 분류하였다.

본 논문에서 제시하는 비디오 데이터베이스의 모든 비디오 콘텐츠 정보는 메타데이터를 통하여 표현되고, 메타데이터를 통하여 실제 데이터에 접근할 수 있다. 본 논문에서 메타데이터는 사용자와 비디오 데이터를 상호 연결하는 매개자 역할을 수행한다. 따라서 비디오 데이터의 효율적인 메타데이터 구축은 비디오 데이터베이스의 성능을 높이고 가용성을 향상시킨다. 비디오 데이터베이스의 메타데이터 유형은 비디오 데이터베이스 시스템의 응용 서비스에 따라 다르게 정의될 수 있다. 표 2는 표 1에서 분류한 메타데이터를 이용하여 비디오 데

이타베이스의 메타데이터 클래스와 속성(Attribute) 메타데이터를 모델링한 것이다.

비디오 데이터베이스의 메타데이터 클래스와 속성 메타데이터는 정보를 체계적으로 구축하여 관리하기 위하여 메타데이터 클래스간의 관계(Relationship)를 모델링하여야 한다. 본 논문에서는 OMT(Object Modeling Technique)[31] 객체 지향 모델링 기법을 사용하여 메타데이터 클래스들간의 관계를 모델링하였다. 그림 4는 OMT 기법을 이용하여 메타데이터의 각 클래스들간의 관계를 모델링한 시스템의 정적 구조이다.

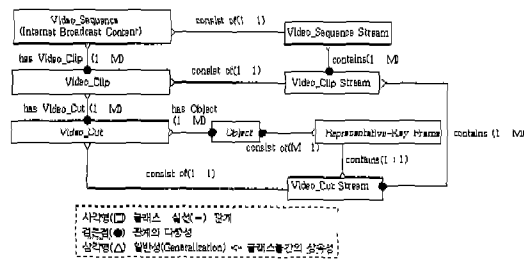


그림 4 메타데이터의 각 클래스들간의 관계 모델

그림 4에서 비디오 시퀀스는 하나 이상의 비디오 클립들로 구성되며, 비디오 클립은 하나 이상의 비디오 컷으로 구성된다. 비디오 시퀀스 및 비디오 클립의 내용은 비디오 제작과정의 사전 제작지식정보와 비디오 영상소재를 조합시켜 추출한 정보에 의해 별도의 주석으로 정의된다. 비디오 컷은 비디오 장면전환을 분할한 것이며, 이들 비디오 컷에는 하나의 비디오 장면과 하나 이상의 샷으로 구성된다. 비디오 컷에 관한 내용은 비디오 컷 검출과 비디오 분할 시 비디오 제작과정의 사전 제작지식정보와 비디오 영상소재를 조합시켜 추출한 컷 검출요소와 비디오의 특징정보에 의해 별도의 주석으로 정의된다. 그리고 비디오 컷은 장면과 샷을 구성하는 연속된 프레임들 중에서 자신을 상징적으로 나타내는 객체들과 객체들의 특징정보들을 추출하여 이들 객체가 포함된 프레임들 대표키 프레임으로 정의한다. 대표키 프레임은 하나의 정지 영상으로 취급되며, 비디오 컷의 객체에 의해 대표키 프레임을 추출한다. 비디오의 계층 구조 단위인 비디오 시퀀스, 비디오 클립, 비디오 장면, 비디오 샷들은 비디오를 통하여 프레임의 구간을 정의하고, 비디오에 연계되는 주석을 이용하여 포함된 내용을 메타데이터 클래스의 속성 데이터로 표현한다.

3.3 비디오 컷 검출 및 분할(Video Cut Detection and Parsing)

비디오 데이터베이스 검색의 기본이 되는 과정은 비디오 분할이다. 방대한 양의 콘텐츠를 효율적인 방법으로 저장하고 비디오의 의미적 구조를 분석하기 위해서는 적당한 방법에 의해 세그먼트 단위로 비디오 분할이 필요하며, 또한 효율적이고 체계적으로 비디오를 검색하고 브라우징 하기 위해 반드시 비디오 분할이 필요하다. 비디오 검색에서 컷 경계 검출과 키 프레임 추출은 비디오 분석에서 중요한 요소로 이용된다. 현재 비디오 인덱싱 방법은 일반적으로 계층적 구조를 가지는 비디오 시퀀스의 컷 검출 기법에 의한 방법이 주를 이루고 있다. 부가적으로 컷은 샷과 샷 또는 장면과 장면의 경계를 의미한다. 기존의 컷 검출의 방법은 비디오 시퀀스의 계층적 구조를 가지는 인접한 두 프레임간의 특징정보 차이가 일정한 임계값보다 클 경우 장면전환이 이루어졌다고 본다. 이들 컷 검출 방법은 시간적인 측면이나 자동 검출이라는 측면에서 여러 가지 장점이 있으나 사용자가 원하는 비디오 컷을 정확하게 검출하지 못한다는 단점을 가지고 있다. 특히, 점진적 장면전환과 같은 경우에는 비디오의 다양한 편집 효과에 의해 샷 또는 장면 사이의 경계가 불분명하므로 검출하기가 매우 어려우며, 또한 카메라의 촬영기법인 패닝이나 줌과 같은 특수효과 장면은 경계가 매우 불분명하므로 컷 검출에 많은 어려움이 따른다.

이를 극복하기 위하여 본 절에서는 콘텐츠 제작과정의 사전 제작지식정보와 비디오 영상소재를 조합시켜 장면전환에서 나타나는 분석적 표현의도의 특징정보들을 관찰하여 비디오의 의미적 구조를 정확하게 표현할 수 있도록 비디오 시퀀스의 타임 코드(Time Code)를 모델링 한 후, 이를 비디오 컷 검출과 비디오 분할에 이용하는 방법을 제시한다. 여기서 비디오 타임 코드는 비디오 편집 점을 정확하고 빠르게 찾아 편집하도록 하는 비디오의 절대위치와 시간정보를 화면 하나에 초당 30 프레임으로 신호를 붙인 것을 말하며, 타임 코드의 값은 00시 00분 00초 00 프레임부터 23시 59분 59초 29 프레임까지로 되어 있으며, 23시 59분 59초 29 프레임의 다음은 00시 00분 00초 00 프레임으로 돌아오는 24 시간제 시계와 같이 카운트되는 순환체제로 되어 있다. 비디오 시퀀스의 타임 코드는 제작할 콘텐츠의 의도와 그 의도를 비디오로 구성할 때 촬영한 비디오 영상소재 중에서 중요한 비디오 컷 검출과 분할하는데 이용되며, 또한 비디오 내용을 제작할 콘텐츠의 의도와 같도록 배열하고, 시간적 순서와 길이 등을 맞추어 영상과 음성을 정교하게 비디오를 구성하는데 이용된다. 이 방법은 콘텐츠 제작과정에서 비디오 시퀀스의 특징정보 변화패턴



을 가진 장면전환 프레임의 타임 코드를 모델링 한 후 이들 정보를 재이용하여 사용자가 원하는 비디오 컷을 정확하게 검출하고 분할할 수 있다.

3.3.1 비디오 컷 특징정보 추출 원리

비디오의 영상 표현에 있어서 컷은 필연적이다. 컷의 역할은 피사체의 공간적 배치 변화, 피사체 자체의 시간적 환경의 변화, 등장인물의 환경변화, 인물이 존재하는 시간의 변화 등이다. 비디오는 제작과정에서 샷의 시간적 순서를 결정된 후 서로 다른 샷들의 편집과정을 거치며, 이 편집과정의 마지막 컷에 원래의 샷의 프레임뿐만 아니라 편집 프레임을 추가하게 된다. 비디오 영상은 한 컷에 다의적인 뜻이 내포되어 있으므로 샷을 중복하여 맞추는 방법에 의해 의미가 확실하게 된다. 이와 같이 복수의 샷을 연결하면 샷과 샷간의 특정한 의미가 나타난다. 이들의 특정한 의미를 비디오 컷 검출에 이용할 수 있다.

3.3.2 비디오 컷 분할

컷의 검출은 비디오 제작과정의 분석적 표현의도에 따라 비디오 영상에 대한 의미상의 분할과 표현상의 분할에 의해 추출할 수 있다. 컷의 분할은 비디오 제작과정의 분석적 표현의도에 따른 카메라 사이즈와 앵글, 편집효과(Fade-In, Fade-Out, Dissolve 등), 카메라 워크, 와이드 샷, 샷의 변경 등의 검출요소에 의해 정확하게 컷을 분할 할 수가 있으며, 비디오 컷 분할 시 인터넷 방송 콘텐츠의 몽타주(구성 또는 편집)를 고려하여 분할하여야 한다. 본 절에서는 제작과정의 분석적 표현의도

에 의한 검출요소를 고려하여 비디오 컷을 분할하는 과정을 알고리즘으로 나타낸다. 알고리즘 1은 콘텐츠 제작과정의 사전 제작지식정보와 비디오 영상소재를 조합시켜 비디오 제작과정의 분석적 표현의도에 의해 컷을 분할하는 과정이다.

3.4 대표키 프레임 추출

비디오 시퀀스의 계층적 구조에서 샷의 내용을 대표하는 프레임을 비디오의 키 프레임(Key Frame)이라고 말하며, 장면의 내용을 대표하는 프레임을 대표키 프레임(Representative-Key Frame)이라고 말한다. 대표키 프레임은 비디오 시퀀스의 구조적인 인덱싱과 브라우징을 위해 반드시 필요하다. 각 비디오 컷 및 대표키 프레임은 데이터베이스 구축 시 비디오 데이터로 저장되며, 이 컷의 대표키 프레임은 정지 영상 형태로 저장된다. 비디오 시퀀스의 대표키 프레임 추출 방법은 계층적 구조를 가지는 비디오 클립의 순차적인 특성을 기반으로 하나의 비디오 클립을 장면단위로 분할한 후, 그 장면의 내용을 대표할 수 있는 프레임을 그림 5와 같이 선정한다.

이와 같이 대표키 프레임은 비디오 컷의 연속된 여러 프레임들 중 장면 내용의 주제를 이용하여 대표 프레임을 추출하는 것이므로, 의미적으로 중요한 프레임을 추출할 수 있는 기법이 필요하다. 본 절에서는 분할된 컷 영상 내의 객체와 객체의 유형, 그리고 객체간의 공간상 위치관계와 행위에 대한 특징정보를 이용하여 대표키 프레임을 추출하는 방법을 기술한다. 이 방법은 분할된 컷 영상 내의 객체와 객체의 유형, 그리고 객체간의 공

알고리즘 1 컷 분할 과정

<p>Step. 1 비디오 콘텐츠 제작과정의 사전 제작지식정보와 비디오 영상소재를 조합시켜 피사체의 공간적 배치 변화, 피사체 자체의 시간적 환경의 변화, 등장인물의 환경변화, 인물이 존재하는 시간의 변화 등의 컷 역할에 관한 분석요소를 고려하여 몇 개의 컷으로 구성되어 있는가를 추출한다.</p> <p>Step. 2 비디오 콘텐츠 제작과정의 사전 제작지식정보와 비디오 영상소재를 조합시켜 Step 1에서 추출한 비디오 컷 구성에 대한 비디오의 형식과 제작과정의 일반적인 특징정보(General Feature)를 추출하여 모델링한다.</p> <p>Step. 3 비디오 콘텐츠 제작과정의 사전 제작지식정보와 비디오 영상소재를 조합시켜 비디오 콘텐츠 제작과정의 분석적 표현의도에 따른 카메라 사이즈와 앵글, 카메라 워크, 와이드 샷, 샷 변경, 편집효과(Fade-In, Fade-Out, Dissolve 등) 등의 컷 검출요소 특징정보를 추출하여 모델링한다.</p> <p>Step. 4 비디오 콘텐츠 제작과정의 사전 제작지식정보와 비디오 영상소재를 조합시켜 Step 3에서 검출한 각 컷의 검출요소 특징정보를 이용하여 비디오의 시공간적인 특징정보인 각 컷의 개시 점과 종료 점에 대한 비디오 프레임의 타임 코드를 검출한다.</p> <p>Step. 5 비디오 콘텐츠 제작과정의 사전 제작지식정보와 비디오 영상소재를 조합시켜 비디오 제작과정의 분석적 표현의도에 Step 1에서 4단계까지 추출된 비디오의 특징정보를 이용하여 컷을 분할한다.</p> <p>If: 검출요소 = (의미상의 분석적 표현의도 &amp; 컷 분할 타임 코드)                  Then: 비디오의 의미상 분석적 표현의도에 의해 컷을 분할한다.                  즉, 시간축의 대폭적인 변화와 표현장소의 변화에 의해 컷을 분할한다.(왜냐하면, 비디오 촬영 및 편집 시에 시공간적인 변화는 별도의 컷으로 표현되기 때문에 컷을 분할할 수가 있다.)</p> <p>Else If: 검출요소 = (표현상의 분석적 표현의도 &amp; 컷 분할 타임 코드)                  Then: 비디오의 표현상 분석적 표현의도에 의해 컷을 분할한다.                  즉, 비디오 표현의 효과에 의해 컷을 분할한다.(왜냐하면, 비디오 촬영 및 편집 시에 비디오의 효과변화는 별도의 컷으로 표현되기 때문에 컷을 분할할 수가 있다.)</p>
---

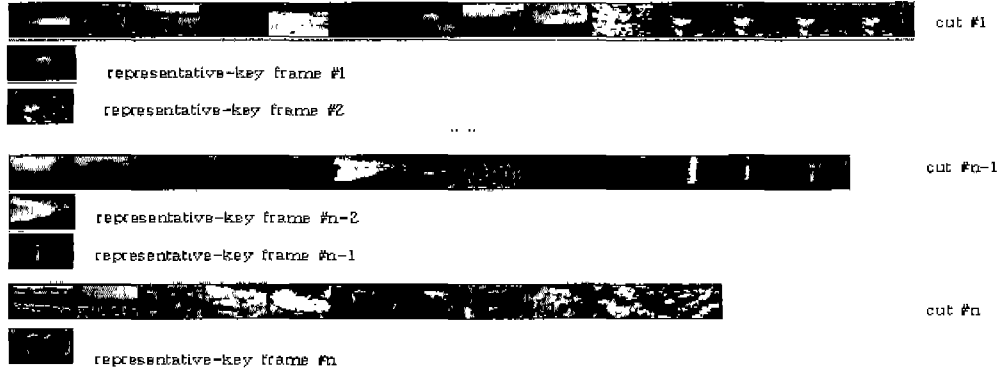


그림 5 대표키 프레임 선정(Choice of Representative-Key Frame)

간상 위치관계와 행위에 대한 특징정보를 5W 1H 분석 기준에 의해 콘텐츠 제작과정의 사전 제작지식정보와 비디오 영상소재를 조합시켜 각 샷의 피사체를 구체적으로 비교 분석하여 추출하므로 정확하고 쉽게 대표키 프레임을 추출할 수가 있다. 표 3은 비디오 콘텐츠 제작 과정의 사전 제작지식정보와 비디오 영상소재를 조합시켜 5W 1H 원칙에 의한 분할된 컷 영상의 분석 기준 예이다.

알고리즘 2는 비디오 콘텐츠 제작과정의 사전 제작지

식정보와 비디오 영상소재를 조합시켜 분할된 컷 영상 내의 객체와 객체의 유형, 그리고 객체간의 공간상 위치 관계와 행위에 대한 특징정보와 5W 1H 분석 기준에 의해 대표키 프레임을 추출하는 과정이다.

3.5 주석기반 및 내용기반 비디오 인덱싱

비디오 컷에서 추출된 대표키 프레임은 정지 영상 형태로 데이터베이스에 저장되며, 대표키 프레임의 특징정보는 비디오 인덱싱의 영상정보로 활용된다. 일반적으로 비디오 데이터베이스 구축 시 빠르고 정확하게 비디오

표 3 5W 1H 원칙에 의한 분할된 컷 영상의 분석 기준 예

구분	Video(영상)	Audio(음성)
WHEN	시대, 계절, 시각의 묘사 중속, 식물(계절) 등	효과음악(M.E), 효과음향(S.E) 등
WHERE	상황 설명의 로케이션·세트(L·S) 장소, 건물을 명시한 간판 문자 등의 설명소재 병용	현장은 상징하는 S.E 음성에 의한 설명 등
WHO	등장인물, 아나운서, 리포터 등 영상정보를 전하는 주체	언어 및 주인공의 역할을 명시한 S.E 등
WHAT	행동의 묘사, 완전 묘사가 아닌 구체적인 피사체 등	행동은 동반한 S.E 현장의 리얼리티를 높이는 S.E, 언어에 의한 설명 등
HOW	행동, 심리, 결과의 묘사설명의 방법 등	행동, 심리, 결과의 묘사언어의 S.E, M.E의 상승에 의한 설명묘사 등
WHY	언어를 중심으로 표정, 동작 등의 부가적 설명	언어 설명의 가장 중요한 요소 등

알고리즘 2 대표키 프레임 추출과정

Step. 1 비디오 콘텐츠 제작과정의 사전 제작지식정보와 비디오 영상소재를 조합시켜 분할된 컷 영상내의 객체의 변화, 객체유형의 변화, 객체간의 공간상 위치관계의 변화, 객체의 행위 등의 샷 역할에 관한 분석요소를 고려하여 몇 개의 샷으로 구성되어 있는가를 추출한다.
Step. 2 비디오 콘텐츠 제작과정의 사전 제작지식정보와 비디오 영상소재를 조합시켜 Step 1에서 추출한 비디오 샷의 연속 프레임 중 상징적인 객체 프레임을 5W 1H 분석 기준에 의해 특징정보를 추출하여 모델링한다.
Step. 3 비디오 콘텐츠 제작과정의 사전 제작지식정보와 비디오 영상소재를 조합시켜 Step 2에서 추출한 객체 프레임의 특징정보를 이용하여 객체 프레임의 타임 코드를 검출한다.
Step. 4 Step 3에서 추출한 객체 프레임의 타임 코드에 의해 대표키 프레임을 검출한다.
Step 4.1 If: 검출요소 = (추출한 객체 프레임의 타임 코드)
Then: 대표키 프레임은 검출

정보를 검색하고, 비디오 데이터의 효율적인 조작을 위해 비디오 데이터의 특징정보를 고려한 적절한 인덱싱 기법이 필요하다[32]. 기존의 비디오 검색 방법들은 비디오 데이터의 내용정보를 어떻게 추출해 내는가에 주로 초점을 맞추고 있으며 비디오 제작 시 현재 충분히 축적, 관리되고 있지 않은 비디오의 의미적 구조와 제작과정의 사전 제작지식정보를 기반으로 한 상호관계를 어떻게 표현해 줄 것인가에 대해서는 고려하지 않았다. 따라서 제작과정의 사전 제작지식정보간에 상호관계가 없어지므로 이들 상호관계정보를 비디오 검색에 이용할 수 없다. 비디오 검색의 궁극적인 목적은 사용자가 원하는 조건을 만족하는 비디오를 신속하고 정확하게 찾아내는 데 있다.

본 논문에서는 비디오 컷 검출과 분할, 그리고 대표키 프레임 추출 시 비디오의 의미적 구조와 제작과정의 사전 제작지식정보를 비디오 특징 데이터에 상호 관련지어 추출한 비디오의 하위 레벨에 표현된 내용기반 메타데이터 유형과 비디오의 특징정보 추출이 어려운 상위 레벨에 표현된 주석기반 메타데이터 유형을 함께 비디오의 색인으로 구성하는 방법을 제시한다. 그림 6은 콘텐츠 제작과정의 사전 제작지식정보와 비디오 영상소재를 조합시켜 비디오를 인덱싱하는 과정이다. 그림 6에서와 같이 비디오 콘텐츠 제작과정의 사전 제작지식정보와 비디오 영상소재를 조합시켜 비디오 몽타쥬(구성 또는 편집)에 의해 추출된 주석기반 영상 특징정보와 비디오의 영상소재에 의해 추출된 내용기반 영상 특징정보, 그리고 촬영기법에 의해 추출된 영상 특징정보의 메타데이터를 활용하여 비디오를 인덱싱할 수 있다. 영상의 특징정보는 패턴의 변화, 색상(Color), 질감(Texture), 영상에서의 개체 모양(Shape), 영상 내 개체간의 공간상 위치관계 등을 말한다. 영상으로부터 추출 가능한 특징정보는 영상의 모양, 색상, 질감 등과 같은 기본적인 특징정보(Primitive Feature)와 영상의 다양한 의미 정보

(Semantic Information)를 문자 형태로 표현하는 논리적인 특징정보(Logical Feature)로 나눌 수 있다[33].

이들 특징정보 중에서 색상정보를 가장 보편적으로 많이 사용되고 있다[34]. 그리고 색상정보를 이용한 질의 검색에는 평균 RGB를 이용한 질의 검색이 있다[35]. 평균 RGB를 이용한 질의 검색은 평균 색상으로 대표키 프레임 영상을 검색하므로 대표키 프레임 영상이 아님에도 검색이 될 수 있다는 단점이 있다. 이러한 내용기반 비디오 검색방법은 단일 특징만을 이용할 경우 정확한 검색이 이루어질 수 없으며 대용량의 데이터베이스에 적용하면 검색효율이 떨어질 수 있다. 이들의 단점을 극복하기 위해서는 검색방법들간의 상호 보완적인 역할이 필요하다. 따라서 검색방법들간의 상호 보완적인 역할을 갖도록 하기 위해서는 비디오 데이터베이스 시스템을 구현함에 있어서 다양한 검색방법을 제공하여야만 상호 보완적인 역할이 가능하다. 이와 같이 다양한 검색방법을 제공하기 위해서는 대표키 프레임 영상의 하위 레벨에 표현된 내용기반 특징정보와 상위레벨에 표현된 주석기반 특징정보를 반드시 추출하여야 한다.

본 논문에서는 비디오 데이터베이스 시스템을 구현함에 있어서 대표키 프레임 영상의 주석기반 특징정보와 내용기반 특징정보의 메타데이터 유형을 비디오의 색인으로 구성하여 주석기반질의, 내용기반질의, 그리고 주석기반 및 내용기반 통합질의 기능을 제공하는 방법을 제시한다. 이를 위해 비디오의 하위 레벨에 표현된 대표키 프레임 영상에 내포하고 있는 객체의 에지 특징정보를 추출하여 비디오 데이터베이스의 내용기반 비디오 인덱싱으로 구성하는 방법을 제시한다. 대표키 프레임의 에지 특징정보를 추출하는 방법은 라플라시안 마스크(Laplacian Mask) 계수 -0, -1, -0, -1, 4, -1, 0, -1, 0 과 엠보싱 효과의 컨벌루션 마스크(Convolution Mask) 계수 1, 1, 0, 1, 0, -1, 0, -1, -1을 대표키 프레임의 영상 처리에 적용하여 객체의 에지 특징정보를 추출한다.

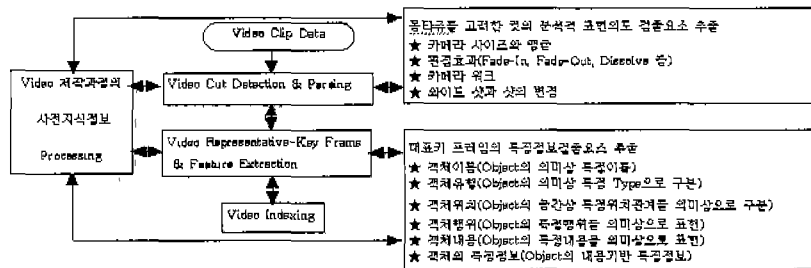


그림 6 비디오 인덱싱 과정

이러한 과정에 의해 추출된 객체의 에지 특징정보 값을 비디오 데이터베이스의 메타데이터로 제공하여 내용기반 비디오 인덱싱으로 구성한다.

### 3.6.1 대표키 프레임 영상의 내용기반 특징정보 추출

본 논문에서 제시하는 비디오 데이터베이스 시스템은 대표키 프레임에 있는 기본 색상특징정보, 객체의 에지 특징정보, 그리고 엠보싱 효과에 의한 객체의 에지 특징정보를 추출하여 비디오 데이터베이스 시스템의 비디오 인덱싱으로 구성한다. 이들 중 기본 색상특징정보를 비디오 데이터베이스의 인덱싱으로 구성하는 방법은 대표키 프레임 영상에서 추출한 각 채널별 RGB의 평균값을 메타데이터로 제공하여 비디오 데이터베이스의 인덱싱으로 구성한다. 그리고 엠보싱 효과에 의한 객체의 에지 특징정보를 비디오 데이터베이스의 인덱싱으로 구성하는 방법은 엠보싱 효과의 컨벌루션 마스크 계수를 대표키 프레임 영상에 적용하여 대표키 프레임 영상의 에지 외곽선 데이터를 추출한 후 추출한 에지 외곽선 데이터의 각 채널별 RGB 평균값을 메타데이터로 제공하여 비디오 데이터베이스의 인덱싱으로 구성한다. 이 방법은 라플라시안 마스크 계수에 의한 객체의 에지 특징정보 추출 방법과 추출하는 방법이 동일하므로 본 절에서는 라플라시안 마스크 계수에 의한 객체의 에지 특징정보 추출 방법만 기술한다. 대표키 프레임 영상의 에지 특징정보를 추출하는 방법은 다음과 같다.

대표키 프레임 영상에 있는 객체의 에지 특징정보를 추출하는 방법은 먼저 대표키 프레임 영상의 에지 외곽선 데이터를 추출한다. 본 논문에서는 다른 기법들보다 상대적으로 해당 에지 검출 성향이 높고 모든 방향의 에지 검출이 가능한 대표적인 2차 미분 연산자인 라플라시안 마스크 계수를 [36,37] 대표키 프레임 영상에 적용하여 에지 외곽선 데이터를 추출한다. 에지 외곽선 데이터 추출은 대표키 프레임 영상과 라플라시안 마스크 계수를 컨벌루션하여 식 (1)의 공간 컨벌루션 처리 수식에 의해 대표키 프레임 영상의 모든 픽셀을 대상으로 에지 외곽선 데이터 출력 값  $E_v(x,y)$ 를 구한다.

$$E_v(x,y) = (I_0(x-1,y-1) \times M_0) + (I_1(x,y-1) \times M_1) + (I_2(x+1,y-1) \times M_2) + (I_3(x-1,y) \times M_3) + (I_4(x,y) \times M_4) + (I_5(x+1,y) \times M_5) + (I_6(x-1,y+1) \times M_6) + (I_7(x,y+1) \times M_7) + (I_8(x+1,y+1) \times M_8) \quad (1)$$

$$E_v(x,y) = \begin{cases} 255, & \text{if } |E_v(x,y)| > 255 \\ 0, & \text{if } |E_v(x,y)| < 0 \end{cases} \quad (2)$$

그리고, 본 연구에서는 대표키 프레임 영상과 라플라시안 마스크 계수를 컨벌루션하여 에지 외곽선 데이터를 추출할 때 식 (2)와 같이 에지 외곽선 데이터 출력

값  $E_v(x,y)$ 의 크기를 비교하여 만약에  $E_v(x,y)$ 의 값이 음수일 경우에는  $E_v(x,y)$ 의 값을 0으로 치환하고, 그리고  $E_v(x,y)$ 의 값이 255이상일 경우에는  $E_v(x,y)$ 의 값을 255로 치환하여 영교차 처리한다. 에지 외곽선 데이터 출력 값  $E_v(x,y)$ 의 크기가 0인 것은 무시해도 될 만큼의 작은 변화를 가진 경우로서 보다 선명한 에지를 추출할 수 있다. 그리고 컨벌루션 함수를 구현할 때 대표키 프레임 영상의 가장자리 부분에 대한 처리는 여러 방법들이 있으나 본 연구에서는 컨벌루션 윈도우가 대표키 프레임 영상과 중첩되지 않는 (1,1) 위치의 픽셀에서 컨벌루션이 시작되도록 하여 (719,479) 위치의 픽셀에서 컨벌루션이 끝나도록 대표키 프레임 영상의 가장자리 부분을 처리한다. 그리고 출력 대표키 프레임 영상에서 컨벌루션 된 영상의 경계 부분 픽셀들은 입력 대표키 프레임 영상과 같은 해상도를 가진 영상으로 생성하기 위해 가장자리 부분을 복사하여 처리한다.

그리고, 다음 단계에서 대표키 프레임 영상의 에지 외곽선 데이터를 추출한 다음, 추출한 에지 외곽선 데이터에 대하여 세 채널을 갖는 완전 컬러 영상의 각 채널별 RGB 픽셀 값을 추출한다. 에지 외곽선 데이터에 대하여 RGB 픽셀 값을 추출하기 위한 raw 데이터로 변환해야 한다. raw 데이터는 영상 정보와 RGB 데이터 부분으로 구분된다. raw 데이터의 정보 부분은 영상의 폭, 영상의 높이, 영상의 색상 정보를 Space(0x20)나 Line Feed(0x0a)로 구분하여 나타낸다. 나머지 RGB 데이터 부분은 각 채널별 3byte씩 RGB 값을 나타낸다. 이러한 변환 결과에 의해 생성된 에지 외곽선 raw 데이터로부터 추출한 각 채널별 RGB 픽셀 값을 식 (3)과 같이 평균 RGB 방법을 이용하여 각각 더한 후 에지 외곽선 데이터의 전체 픽셀 값으로 나누어 각 채널별 RGB의 평균값을 구한다.

이러한 결과로 구해진 각 채널별 RGB의 평균값을 최종적으로 비디오 데이터베이스의 메타데이터로 제공하여 내용기반 비디오의 인덱싱으로 구성한다.

$$\begin{aligned} R_{avg} &= \left(\frac{1}{N}\right) \sum_{p=1}^N R(p) \\ G_{avg} &= \left(\frac{1}{N}\right) \sum_{p=1}^N G(p) \\ B_{avg} &= \left(\frac{1}{N}\right) \sum_{p=1}^N B(p) \end{aligned} \quad \begin{array}{l} p \text{ 픽셀, } N \text{ 전체 픽셀수} \end{array} \quad (3)$$

### 3.6.2 대표키 프레임 영상의 주석기반 특징정보 추출

대표키 프레임 영상의 내용기반 에지 특징정보 추출 시 추출한 에지 외곽선 데이터는 에지에 대한 특징정보를 표현하고 있으며, 또한 에지에 대한 시각적이고 공간적인 특징정보를 나타내고 있다. 따라서, 본 연구에서는

표 4 실험에 사용한 비디오 데이터

비디오 구분	File Size	Frame Size	Pixel Depth	Frame Rate	Total Frame	Total Duration
자료영상	60.43MB	720×480	24 Bit	30 fps	14514 Frame	484.32 Sec
볼륨이영상	25.73MB	720×480	24 Bit	30 fps	6179 Frame	206.21 Sec
자막영상	67.0MB	720×480	24 Bit	30 fps	10932 Frame	364.80 Sec

대표키 프레임 영상의 주석기반 예지 특징정보를 추출할 때 이들의 예지 외곽선 데이터에 대한 예지의 시각적인 특징정보를 이용하여 객체, 객체 유형, 객체 위치, 객체 이벤트, 객체 내용 등의 주석기반 예지 특징정보를 추출하는 방법을 제시한다. 대표키 프레임 영상의 주석기반 예지 특징정보를 추출하는 방법은 먼저 대표키 프레임 영상의 내용기반 예지 특징정보 추출 인터페이스를 통하여 대표키 프레임 영상에 대한 예지 외곽선 데이터의 시각적인 특징정보를 추출한다. 이러한 결과에 의해 추출된 예지 외곽선 데이터의 시각적인 특징정보를 대표키 프레임 영상과 직접 비교하여 대표키 프레임 영상내의 예지 외곽선 모양에 해당하는 객체에 한정하여 사용자가 직접 주석기반 예지 특징정보를 수동으로 추출하여 모델링한다. 이러한 모델링 결과에 의해 추출된 객체, 객체 유형, 객체 위치, 객체 이벤트, 객체 내용의 주석기반 예지 의미 정보를 메타데이터로 제공하여 비디오 데이터베이스의 인덱싱으로 구성한다. 이와 같이 본 논문에서 제시하는 대표키 프레임 영상의 주석기반 예지 특징정보 추출 방법은 추출된 예지 외곽선 데이터에 대한 시각적인 특징정보를 대표키 프레임 영상과 비교하여 주석기반 예지 특징정보를 직접 사용자가 추출하기 때문에 주석기반 예지 특징정보를 쉽게 추출할 수 있으며, 또한 주석기반 예지 특징정보를 정확하게 모델링할 수 있다.

4. 실험 및 결과

본 논문의 구현은 Windows 98 환경에서 Microsoft Visual C++ 6.0, Microsoft Access 2000 DBMS, Firebird 1.12, Adobe Premiere 5.1, Microsoft Visual InterDev, Active Server Page(ASP)를 사용하여 구현하였다. 웹 서버로는 Personal Web Server를 사용하였으며, 웹 클라이언트로는 Microsoft Internet Explorer를 사용하였다. 실험에 사용한 데이터들은 일반용 캠코더로 촬영한 비디오 콘텐츠 데이터를 Firebird 1.12를 사용하여 AVI 데이터 형식의 비압축 형태로 캡처하여 Adobe Premiere 5.1로 편집한 비디오 데이터를 사용하였다. 인터넷 방송 콘텐츠 데이터로는 편집 제작된 AVI

비디오 콘텐츠를 Firebird 1.12를 이용하여 스트리밍(Streaming) 비디오 파일인 윈도우 미디어 형식으로 변환하여 사용하였다. 본 실험에 사용한 AVI 파일은 Windows에서 멀티미디어 데이터 타입을 지원하기 위하여 만들어졌으며, 실험에 사용한 비디오 데이터는 표 4와 같다.

4.1 구현한 비디오 데이터베이스 시스템의 메인 구조

본 논문에서 구현한 비디오 데이터베이스 시스템의 메인 구조는 그림 7과 같다.

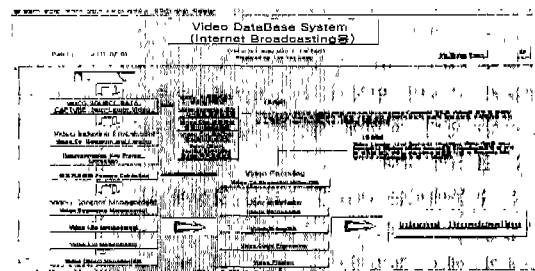


그림 7 비디오 데이터베이스 시스템의 메인 구조

그림 7에서와 같이 비디오 데이터베이스 시스템의 주요 기능은 (1) 일반 캠코더로 촬영한 비디오 소스 데이터를 캡처하여 인터넷 방송 콘텐츠에 맞도록 제작 및 편집하기 위한 Video Data Capture & Non-Linear Video Edition 기능 (2) 비디오 분할과 대표키 프레임 영상의 특징정보추출에 의해 비디오를 인덱싱하기 위한 Video Indexing Processing 기능 (3) 인터넷 방송 콘텐츠 제작과정의 사전 제작지식정보와 인터넷 방송 콘텐츠를 관리하기 위한 Video Content Management 기능 (4) 비디오 콘텐츠와 Live Source를 스트리밍(Streaming) 비디오 파일인 리얼 형식이나 윈도우 미디어 형식으로 변환하기 위한 Video Encoding 기능 (5) 사용자 인터페이스 기능으로 ① 영상 데이터와 대표키 프레임 영상을 처리하기 위한 Image Processing 기능 ② 각 비디오 데이터를 Clip별→Cut별→대표키 프레임 영상별로 특정 부분을 메타데이터와 대표키 프레임 영상 내용 브라우징(Content Browsing)하기 위한

표 5 실험에 사용한 대표키 프레임 영상 데이터

비디오 구분	Video Clip 수	Video Cut 수	대표키 프레임 수	Image Size	Pixel Depth	Color Mode
자료영상	1	30	30	720×480	24 Bit	True Color
불꽃놀이영상	1	2	10	720×480	24 Bit	True Color
차막영상	1	4	12	720×480	24 Bit	True Color

Video Browsing 기능 ③ 비디오 데이터를 주석기반 질의, 내용기반 질의, 주석기반 및 내용기반 통합 질의의 세 가지 방식으로 질의하여 비디오를 검색하기 위한 비디오 질의 처리 기능 ④ 비디오 콘텐츠(AVI 파일 형식)의 동영상을 재생하기 위한 Video Playing 기능 (6) 비디오 데이터베이스에서 관리되고 있는 인터넷 방송용 비디오 콘텐츠와 Live Source를 인터넷으로 방송하기 위한 Internet Broadcasting 기능으로 구현하였다.

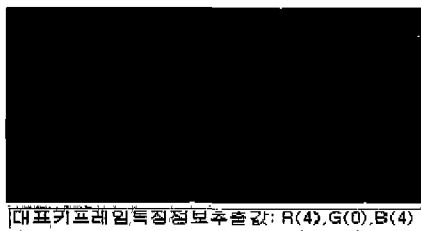
4.2 대표키 프레임 영상의 특징정보추출

본 논문에서는 인터넷 방송 콘텐츠 장면의 내용을 대표하는 대표키 프레임 영상을 그래픽 파일의 BMP, JPEG, TIFF, GIF 데이터 형식의 정지 영상 형태에서 대표키 프레임 영상의 특징정보를 추출하여 비디오를 인덱싱하였다. 실험에 사용한 대표키 프레임 영상 데이터는 표 5와 같다.

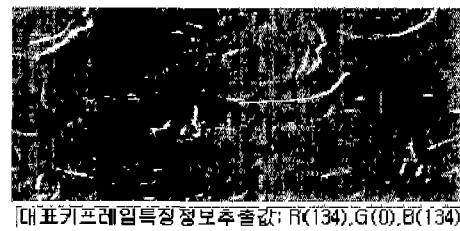
본 논문에서는 대표키 프레임 영상에 있는 기본 색상 특징정보, 객체의 에지 특징정보, 그리고 엠보싱에 의한 객체의 에지 특징정보를 내용기반 비디오 특징정보로

추출하여 비디오 데이터베이스 시스템의 비디오 인덱싱으로 구성하였다. 비디오 객체의 내용기반 에지 특징정보 추출은 라플라시안 마스크 계수 -0, -1, 0, -1, 4, -1, 0, -1, 0를 대표키 프레임 영상에 적용하여 추출하였으며, 엠보싱 효과의 에지 특징정보 추출은 컨벌루션 마스크 계수 1, 1, 0, 1, 0, -1, 0, -1, -1를 대표키 프레임 영상에 적용하여 추출하였다. 이러한 과정에 의해 추출한 대표키 프레임 영상의 특징정보 값을 비디오 데이터베이스의 메타데이터로 제공하여 내용기반 비디오 데이터의 인덱싱으로 구성하였다. 그림 8은 본 논문에서 구현하여 실험한 대표키 프레임 영상의 특징정보를 추출한 결과를 보여준 것이다. 그림 8의 (a)는 대표키 프레임 영상의 색상 특징정보를 추출한 결과이며, (b)는 대표키 프레임 영상의 에지 특징정보를 추출한 결과이다. 그리고, (c)는 엠보싱 효과에 의한 대표키 프레임 영상의 에지 특징정보를 추출한 결과이다. 본 논문에서는 추출한 대표키 프레임 영상의 주석기반 에지 특징정보와 내용기반 에지 특징정보를 메타데이터로 제공하여

(a) 대표키 프레임 영상의 색상  
특징정보추출결과 : R(119), G(0), B(124)



(b) 대표키 프레임 영상의 에지  
특징정보추출결과 : R(4), G(0), B(4)



(c) 엠보싱 효과에 의한 대표키 프레임 영상의 에지  
특징정보추출결과 : R(134), G(0), B(134)

그림 8 대표키 프레임 영상의 특징정보 추출 결과 예

비디오를 인덱싱하였다.

### 4.3 시스템 사용자 인터페이스

본 논문에서 구현한 비디오 데이터베이스 시스템의 사용자 인터페이스 기능으로는 ① 영상 데이터와 대표키 프레임 영상을 처리하기 위한 Image Processing 기능 ② 각 비디오 데이터를 Clip별→Cut별→대표키 프레임 영상별로 특정 부분을 메타데이터와 대표키 프레임 영상을 내용 브라우징하기 위한 Video Browsing 기능 ③ 주석기반 질의, 내용기반 질의, 그리고 주석기반 및 내용기반 통합질의 세 가지 방식의 비디오 질의 처리 기능 ④ 비디오 콘텐츠(AVI 파일 형식)의 동영상 재생하기 위한 Video Playing 기능으로 구분하여 구현하였다.

#### 4.3.1 비디오 브라우저

본 절에서는 비디오 데이터베이스 시스템의 사용자 인터페이스 내용 중 비디오 검색을 지원하기 위한 방법으로 구현된 비디오 브라우징 기능에 관하여 설명하였다. 비디오 브라우징의 목적은 전체 콘텐츠의 내용을 사용자가 빠르고 정확하게 검색할 수 있도록 지원해 주는 것이 주된 목적이다. 대표적으로 사용되는 비디오 브라우저의 종류로는 장면전환이 위치한 첫 화면을 단순히 대표 영상으로 지정하여 보여주는 샘플링 플래쉬(Sampling Flash) 브라우저와 비디오의 장면전환 또는 특정한 부분을 동영상으로 보여주는 러쉬(Rush) 브라우저, 그리고 비디오를 시간적으로 분할하여 여러 개의 아이콘을 차례로 나열하여 보여주는 스트로보스코픽(Stroboscopic) 브라우저 등이 있다.

본 논문의 비디오 브라우저는 비디오 데이터를 스트리브드별로 원하는 데이터를 브라우저에서 자체적으로 검색할 수 있도록 하는 기능과 비디오의 특정 부분을

내용 브라우징 할 수 있도록 하는 기능을 함께 제시하여 구현하였다. 그림 9는 대표키 프레임 영상 중 객체가 그룹인 영상을 자체 내용 브라우징하여 검색 실험한 결과 예이다.

그림 9와 같이 본 논문에서 제시하는 비디오 데이터베이스 시스템의 비디오 브라우저는 각 비디오에 대한 구조를 쉽게 파악할 수 있도록 트리 구조로 표현된 비디오 구조 정보를 Clip별→Cut별→대표키 프레임 영상별로 특정 부분을 메타데이터와 대표키 프레임 영상을 내용 브라우징 할 수 있도록 구현하였다. 그리고, Clip별→Cut별→대표키 프레임 영상별 세부 기능으로는 각자 별도의 검색 기능과 재생기능, 그리고 브라우징 기능을 갖도록 구현하였다.

#### 4.3.2 비디오 질의 처리

본 절에서는 비디오 데이터베이스 시스템의 사용자 인터페이스 내용 중 비디오 검색을 지원하기 위한 방법으로 구현된 비디오 질의 처리 기능에 관하여 설명하였다. 본 논문에서 구현한 비디오 데이터베이스 시스템의 비디오 질의 처리 기능은 비디오 데이터베이스에 저장된 데이터와 간접적인 유형의 데이터를 질의 조건 값으로 받아들여 질의를 수행하는 간접 질의(Indirect Query) 방식으로 이용하였다. 비디오 질의 처리 방식은 비디오 데이터베이스 시스템의 메타데이터와 대표키 프레임 영상의 상호관계를 기반으로 사용자가 비디오 질의 처리 인터페이스 화면을 통하여 비디오 데이터의 특징정보를 추출하고, 검색할 수 있도록 하였다. 비디오 질의 방법은 검색방법들간의 상호 보완적 역할을 갖도록 하여 신속하고 효율적으로 원하는 비디오를 질의할 수 있도록 주석기반질의, 내용기반질의, 그리고 주석기반 및 내용기반 통합질의 방식의 세 가지 질의방식을 제공하였다. 이 방법의 장점은 비디오의 하위 레벨에 표현된 내용기반 메타데이터 유형(기본 색상특징정보, 객체의 에지 특징정보와 임보싱 효과에 의한 에지 특징정보)과 비디오의 특징정보 추출이 어려운 상위 레벨에 표현된 주석기반 메타데이터 유형(Object, Object Type, Object Position, Object Event, Object Content)을 비디오의 색인으로 제공하여 질의하므로 비디오 콘텐츠 검색의 성능을 향상시킬 수 있다. 그리고, 사용자가 질의어에 대한 지식이 없더라도 쉽고 정확하게 비디오 데이터를 질의하고 검색할 수 있도록 사용자 인터페이스에 비디오 질의 연산방법을 제공하여 질의 및 검색에 편리성을 제공하였다. 비디오 질의 처리 사용자 인터페이스에 제공하는 비디오 질의 연산방법은 표 6과 같다.

본 절에서는 본 논문에서 구현한 비디오 질의 처리

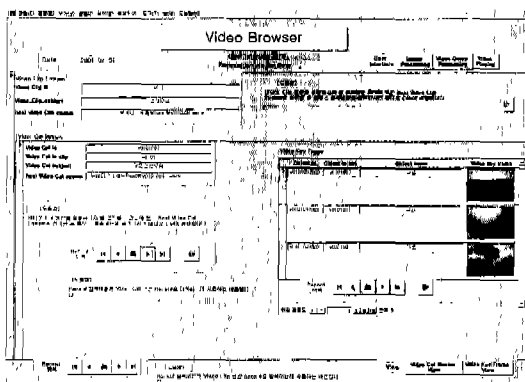


그림 9 비디오 내용 브라우징(Video Content Browsing)

표 6 비디오 질의의 연산방법

질의 방법	표현식	연산방법
주석기반질의	[객체명]	단일항목연산
	[객체명] OR [객체타입   객체위치   객체이벤트   객체내용]	OR 연산
	[객체명] AND [객체타입   객체위치   객체이벤트   객체내용]	AND 연산
내용기반질의	[기본색상특징추출정보   에지특징추출정보   엠보싱효과에지특징추출정보]	단일항목연산
통합질의 (주석기반 및 내용기반)	[기본색상특징추출정보   에지특징추출정보   엠보싱효과에지특징추출정보] AND [객체명]	복합연산
	[기본색상특징추출정보   에지특징추출정보   엠보싱효과에지특징추출정보] OR [객체명] OR [객체타입   객체위치   객체이벤트   객체내용]	OR 연산
	[기본색상특징추출정보   에지특징추출정보   엠보싱효과에지특징추출정보] AND [객체명] AND [객체타입   객체위치   객체이벤트   객체내용]	AND 연산

기능의 주석기반 질의, 내용기반 질의, 그리고 주석기반 및 내용기반 통합질의 기능인 세 가지의 비디오 질의 방법을 표 6에서 제시하는 비디오 질의 연산방법을 이용하여 실험하였다. 그리고 이 실험을 통하여 비디오 검색방법들간의 상호 보완적인 역할을 보여준다.

그림 10은 표 6에서 제시하는 비디오 질의 연산방법을 이용하여 주석기반 질의, 내용기반 질의, 그리고 주석기반 및 내용기반 통합질의 방식으로 비디오를 질의한 예이다.

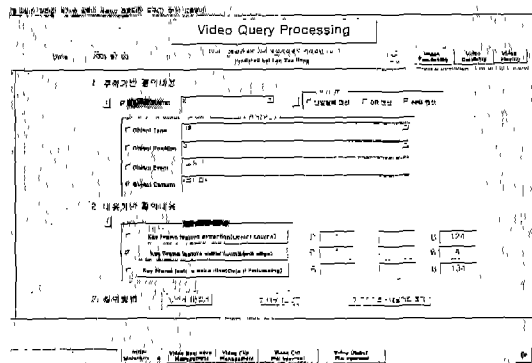


그림 10 비디오 질의 처리 예

그림 10에서 비디오를 질의한 데이터 중 비디오의 주석기반 에지 특징정보는 Object="꽃", Object Type="식물", Object Position="전체", Object Event="\*노랑색\*", Object Content="\*튤립\*" 값으로 질의된 메타데이터들이다. 그리고 비디오의 내용기반 에지 특징정보는 Object Source=(R:119, G:0, B:124), Object Edge=(R:4, G:0, B:4), Object Embossing=(R:134, G:0, B:134) 값으

로 질의된 메타데이터들이다. 그림 11은 그림 10에서 비디오의 주석기반 에지 특징정보 추출 값이 튤립 꽃인 비디오 대표키 프레임 영상을 주석기반 질의방식으로 질의하여 실험한 비디오 검색 결과를 보여준 예이다.

- (a) 비디오 질의 방법: 주석기반 질의
- (b) 질의 연산 방법: AND 연산
- (c) 표현식: [객체명] AND [객체타입 | 객체위치 | 객체이벤트 | 객체내용]
- (d) 주석기반 에지 특징정보: [Object: "꽃", Object Content: "\*튤립\*"]
- (e) 질의 결과:

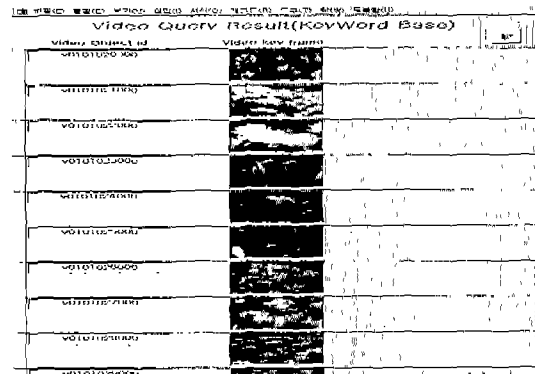


그림 11 주석기반 비디오 질의 검색 결과 예

그리고, 그림 12는 그림 10에서 에지 특징정보 추출 값이 R:4, G:0, B:4 인 대표키 프레임 영상을 내용기반 질의방식으로 질의하여 실험한 비디오 검색 결과를 보여준 예이다.



- (a) 비디오 질의 방법: 내용기반 질의
- (b) 질의 연산 방법: 단일항목연산
- (c) 표현식: [기본색상특징추출정보 | 에지특징추출정보 | 엠보싱효과에지특징추출정보]
- (d) 내용기반 에지 특징정보: [R:4, G:0, B:4]
- (e) 질의 결과:

Video Object Id	Video Key frame
v0101013000	
v0101019000	
v0101021000	
v0101022000	
v0101023000	
v0101027000	

그림 12 내용기반 비디오 질의 검색 결과 예

그리고, 그림 13은 그림 10에서 내용기반 에지 특징 정보 추출 값이 R:4, G:0, B:4 이고 주석기반 에지 특징 정보 추출 값이 노랑색 쫘인 대표키 프레임 영상을 주석기반 및 내용기반 통합질의방식으로 질의하여 실험한 비디오 검색 결과를 보여준 예이다.

- (a) 비디오 질의 방법: 주석기반 및 내용기반 통합질의
- (b) 질의 연산 방법: AND 연산
- (c) 표현식: [기본색상특징추출정보 | 에지특징추출정보 | 엠보싱효과에지특징추출정보] AND [객체명] AND [객체타입 | 객체위치 | 객체이벤트 | 객체내용]
- (d) 내용기반 에지 특징정보: [R:4, G:0, B:4]
- (e) 주석기반 에지 특징정보: [Object : “꽃”, Object Event : “\*노랑색\*”]
- (f) 질의 결과:

Video Object Id	Video key frame
v0101022000	

그림 13 주석기반 및 내용기반 비디오 통합질의 검색 결과 예

위 실험에서와 같이 비디오 데이터의 하위레벨에 표현된 내용기반 에지 특징정보와 상위레벨에 표현된 주

석기반 에지 특징정보를 이용하여 다양한 검색방법으로 질의함으로써 검색방법들간의 상호 보완적인 역할에 의해 신속하고 효율적으로 원하는 비디오 데이터를 정확하게 검색할 수 있다는 것을 알 수 있다.

#### 4.4 Internet Broadcasting 기능

인터넷 방송은 일반적인 방송 서비스와는 많은 차이가 있다. 인터넷 방송 콘텐츠 서비스는 전적으로 방송국에 의존하기보다는 이용자 측에도 상당 부분 권한이 있기 때문에 이들에게 다양한 채널과 내용을 자유롭게 선택할 권한이 있다. 따라서 인터넷 방송은 멀티미디어 형식으로 구성된 다양한 콘텐츠를 데이터베이스화하여 사용자에게 효율적으로 서비스할 수 있어야 한다. 그림 14는 본 논문에서 구현한 비디오 데이터베이스 시스템에서 관리되고 있는 인터넷 방송용 비디오 콘텐츠와 Live Source를 인터넷으로 방송하기 위한 인터넷 방송의 웹 콘텐츠를 실험한 예이다.

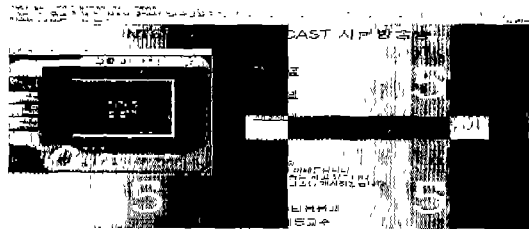


그림 14 인터넷 방송의 웹 콘텐츠를 실험한 예

본 논문에서 구현한 비디오 데이터베이스는 비디오의 의미적 구조와 사전 제작지식정보를 구조적으로 축적하여 데이터베이스를 구축하므로 정확한 인터넷 방송 콘텐츠 정보의 축적관리와 구축작업의 효율화가 가능하며, 또한 인터넷 방송 콘텐츠 제작 시 정보공유 및 재이용이 가능하므로 인터넷 방송의 TV중계, 원격수업, 뉴스 전송, 원격진료 등과 같은 다양한 응용서비스의 새로운 콘텐츠 제작에 효율성을 높일 수 있다. 그리고 인터넷 방송의 다양한 멀티미디어 내용을 시간과 공간의 제약 없이 양방향의 대화형 데이터 방송에 이용될 수 있으며, 웹 콘텐츠에 의해 사용자가 원하는 콘텐츠에 대한 실시간 서비스를 할 수 있다.

#### 4.5 실험결과 분석

본 논문에서 제안한 비디오 데이터베이스 시스템의 검색 적합성(Effectiveness)[38]을 알아보기 위하여 정확률(Precision)에 의한 성능평가 척도를 이용하여 실험을 하였다. 비디오 검색 시스템의 완전 일치 검색을 기반으로 하는 검색의 적합성은 재현율(Recall)과 정확률

(Precision)로 검색의 성능을 측정하고 있다[39,40]. 반면에 완전 일치 검색이 아닌 유사 검색(Similarity Retrieval) 시스템은 표준 정확율과 재현율(Normalized Precision and Recall)을 이용하여 평가한다[41]. 본 논문에서 구현한 비디오 데이터베이스 시스템은 완전 일치 검색을 기반으로 비디오 데이터를 검색하는 시스템이므로 정확율을 이용하여 비디오 검색의 적합성을 실험하였다. 비디오 데이터베이스 시스템의 실험은 총 52개의 비디오 대표키 프레임 영상 데이터를 3가지의 대표키 프레임 영상 종류별로 구분하여 각 대표키 프레임 영상에 대한 질의 검색결과를 실험해 보았다. 본 논문에서 제안한 비디오 데이터베이스 시스템의 검색 적합성에 대한 성능비교를 위해 식 (4) 재현율과 정확율의 성능평가 척도를 사용하였다.

$$Recall(\%) = \frac{R_r}{T} \times 100 \quad (4)$$

$$Precision(\%) = \frac{R_r}{T_r} \times 100$$

표 7은 각 대표키 프레임 영상에 대한 질의 검색의 적합성을 실험한 결과이다. 표 7에서 T는 비디오 데이터 질의에 적합한 전체 비디오 대표키 프레임 영상 수이며, R<sub>r</sub>은 비디오 데이터 질의에 의해 적합하게 검색된 비디오 대표키 프레임 영상 수이다. 그리고 T<sub>r</sub>은 비디오 데이터 질의에 의해 전체 검색된 비디오 대표키 프레임 영상 수이며, R<sub>r</sub>는 비디오 데이터 질의에 의해 부적합하게 검색된 비디오 대표키 프레임 영상 수이다.

비디오 데이터베이스 시스템의 비디오 질의 검색은 비디오 데이터베이스에 저장된 대표키 프레임 영상 데이터에 따라 재현율과 정확율은 다소 차이가 있었으나 본 실험에서는 표 7과 같이 대체로 높은 정확율을 얻을 수가 있었다.

본 논문에서 제안하여 구현한 비디오 데이터베이스 시스템의 비디오 검색 시간에 대한 실험결과 분석은 4.3.1절과 4.3.2절에서 실험한 결과와 같이 간접 질의 처리 방식에 의한 비디오 데이터의 메타데이터와 대표키 프레임 영상의 상호관계를 기반으로 사용자 비디오 질

의 처리 인터페이스 화면을 통하여 직접 비디오 데이터의 에지 특징정보를 추출하고, 검색방법들간의 상호 보완적 역할을 갖도록 하기 위해 비디오 질의 연산방법과 주석기반 및 내용기반 메타데이터의 비디오 인덱싱 구성에 의해 주석기반 질의, 내용기반 질의, 그리고 주석기반 및 내용기반 통합질의 방식으로 질의를 하므로 신속하고 효율적으로 원하는 비디오 데이터를 정확하게 검색할 수 있다. 4.3.1절과 4.3.2절의 실험에서와 같이 비디오 데이터의 하위레벨에 표현된 내용기반 에지 특징정보와 상위레벨에 표현된 주석기반 에지 특징정보를 이용하여 다양한 검색방법으로 질의함으로써 검색방법들간의 상호 보완적인 역할에 의해 신속하고 효율적으로 원하는 비디오 데이터를 정확하게 검색할 수 있다.

그리고, 비디오 데이터는 대용량적인 특성을 가지고 있으므로 비디오 데이터베이스 구축 시 비디오의 데이터에 따라 데이터베이스의 크기가 매우 크게 증가할 수 있다. 본 논문에서 제안하여 구현한 비디오 데이터베이스의 메모리 크기에 대한 실험결과 분석은 표 4에 나타난 실험 비디오 데이터에 대한 비디오 데이터베이스의 메모리 크기가 41.5Mb이었다. 본 논문에서 제안하여 구현한 비디오 데이터베이스 시스템은 표 2의 메타데이터를 기준으로 텍스트 기반의 전통 데이터베이스처럼 비디오 데이터베이스를 생성하고 구축할 수 있기 때문에 실험결과와 분석과 같이 비디오 데이터베이스의 메모리 크기를 줄일 수 있으며, 또한 비디오 콘텐츠의 축적관리와 비디오 데이터베이스의 메모리를 효율적으로 사용할 수 있다. 표 2의 비디오 데이터베이스의 메타데이터 클래스 및 속성 데이터는 비디오의 계층적 구조별로 비디오의 하위 레벨에 표현된 내용기반 메타데이터 유형과 비디오의 특징정보 추출이 어려운 상위 레벨에 표현된 주석기반 메타데이터 유형의 메타데이터이다.

#### 4.6 기존 시스템들과의 비교 분석

본 절에서는 제안된 비디오 데이터베이스의 데이터 모델과 비디오 질의 처리 기능, 그리고 비디오 브라우징 기능 등을 2장의 관련 연구에서 소개된 비디오 데이터

표 7 비디오 데이터베이스의 질의에 대한 검색 적합성 실험결과

비디오 구분	비디오 클립 수	비디오 컷 수	전체 프레임 수	대표키 프레임 수 (T)	검출 적합수 (R <sub>r</sub> )	검출 부적합수 (R <sub>t</sub> )	전체 검출수 (T <sub>r</sub> )	Precision (%)
자료영상	1	30	14514	30	30	4	34	88
꽃놀이영상	1	2	6179	10	10	1	11	91
자막영상	1	4	10932	12	12	0	12	100
계	3	36	31625	52	52	5	57	91

베이스 시스템들과 비교 분석하였다.

OVID는 비디오 객체 모델에 데이터베이스 스키마를 사용하지 않기 때문에 필요한 비디오 프레임 시퀀스를 비디오 객체로 정의할 수 있도록 허용한 반면에 비디오 데이터베이스의 구축 및 유지가 어렵다는 단점이 있다. OVID 모델은 비디오 객체를 주석 중심으로 구성되므로 내용기반 검색을 지원하지 못한다. 그리고 각 비디오 데이터 구조가 일반화되어 있지 않기 때문에 질의 형식을 정의할 수 없으므로 질의 모델이나 질의어가 빈약하다는 단점을 가지고 있다.

VideoSTAR는 범용 비디오 데이터베이스 시스템으로 비디오 데이터와 메타데이터의 공유와 재사용이 가능하다는 장점이 있다. 이 시스템은 관계형 데이터베이스 모델을 기반으로 비디오 데이터를 구조화하였으며, 전체적인 비디오의 계층적인 구조를 나타낼 수 있는 구조 브라우징 기능과 메타데이터를 보여줄 수 있는 내용 브라우징 기능을 지원한다. 반면에 데이터 모델이 복잡하고, 질의어가 별도로 제공되지 않으므로 질의 모델이나 질의어가 빈약하다는 단점이 있다. 그리고, 질의 처리는 먼저 메타데이터의 내용을 먼저 검색한 다음 메타데이터와 연관된 비디오 스트림을 대응시키는 단계적 접근 방법을 취하고 있다. 그러나 비디오 데이터 검색 시 해당 비디오 데이터와 메타데이터를 관련시켜야 하는 복잡성이 존재한다.

AVIS는 비디오의 상위 레벨에 표현된 객체, 행위, 사건 등의 의미정보를 주석기반 메타데이터 유형으로 직접 부여하여 검색하도록 되어 있다. 이 시스템은 비디오에 대한 데이터 모델 없이 사건과 객체 위주의 데이터 구조를 정의하여 데이터베이스를 구축하였다. 이 시스템의 단점은 비디오의 상위 레벨에 표현된 의미정보들만 주석기반 메타데이터 유형으로 부여하므로 하위 레벨의 비디오 의미정보들에 대한 메타데이터가 없다는 단점이 있다.

Algebraic Video는 비디오 정보를 모델링하여 주석기반 비디오 검색만 지원하며, 데이터 모델은 주석에 스키마를 정의하지 않기 때문에 필요한 속성을 사용자가 정의하여 사용할 수 있다. 그리고 사용자는 비디오의 의미정보를 비디오 세그먼트와 연계할 수 있으며, 동일한 비디오 세그먼트는 다양한 관점의 주석을 부여할 수 있다는 장점을 가지고 있는 반면에 특별한 질의어가 제공되지 않기 때문에 질의 인터페이스에서 주석에 정의된 속성과 속성 값만을 지정하여 질의를 구성하므로 질의 모델이나 질의어가 빈약하다는 단점을 가지고 있다.

Venus는 비디오 프레임에 나타나는 각 객체들과 객체들 사이의 시간 관계, 공간 관계들을 메타데이터로 한

내용기반 비디오 검색 시스템으로 비디오의 객체와 객체들간의 시간 및 공간 관계를 질의하는데 CVQL(Content-based Video Query Language)인 내용기반 비디오 질의어를 제공하였다. 이 시스템은 프레임 단위로 비디오의 특징정보들을 추출하기 때문에 데이터베이스 생성 시에 많은 양의 메모리가 필요하며, 또한 추출된 비디오의 특징정보들만을 기반으로 하여 메타데이터를 구성하므로 사건, 장소 등과 같은 상위 레벨의 비디오 의미정보들은 메타데이터로 구성하지 못하기 때문에 한정적인 비디오 검색에만 지원하게 되는 문제점을 가지고 있다.

본 논문에서 제안하여 구현한 비디오 데이터베이스 시스템은 비디오 콘텐츠 제작과정의 문자 정보와 비디오 정보를 비디오 데이터베이스에 생성할 수 있도록 비디오의 의미적 구조와 사전 제작지식정보를 기반으로 비디오 시퀀스의 내용별 특징정보를 조합하여 분류한 주석기반 및 내용기반 메타데이터 유형을 통합한 비디오 데이터 모델을 제시하였다. 그리고 비디오의 질의 검색 기능과 내용 브라우징 검색 기능의 사용자 인터페이스를 지원하여 쉽고 정확하게 비디오 데이터를 질의하고 검색할 수 있도록 질의 및 검색에 편리성을 제공하였다. 비디오 데이터 모델에 대한 비교는 표 8에 요약하였다.

그리고 본 논문에서 구현한 비디오 질의 처리는 비디오 데이터베이스에 저장된 데이터와 간접적인 유형의 데이터를 질의 조건 값으로 받아들여 질의를 수행하는 간접 질의 방식으로 처리하였다. 질의 방법은 사용자가 비디오 질의 처리 인터페이스 화면을 통하여 비디오 데이터의 주석기반 특징정보와 내용기반 특징정보를 직접 추출하여 질의 및 검색을 할 수 있도록 하였다. 그리고, 사용자가 질의어에 대한 지식이 없더라도 쉽고 정확하게 비디오 데이터를 질의하고 검색할 수 있도록 비디오 질의 처리 사용자 인터페이스에 비디오 질의 연산방법을 제공하여 질의 및 검색에 편리성을 제공하였다. 또한 비디오 질의 처리 사용자 인터페이스의 비디오 질의 연산방법을 이용하여 주석기반 질의, 내용기반 질의, 주석기반 및 내용기반 통합질의 세 가지 방식의 질의가 가능하도록 하였다. 그리고 비디오 브라우저는 각 비디오에 대한 구조를 쉽게 파악할 수 있도록 트리 구조로 표현된 비디오 구조 정보를 내용 브라우징 할 수 있도록 하였으며, 비디오 구조별 세부기능으로는 각자 별도의 검색 기능과 생성기능, 그리고 브라우징 기능을 갖도록 지원하였다. 비디오 질의 처리 및 브라우저에 대한 비교는 표 9에 요약하였다.

표 8 비디오 데이터 모델 비교

항목 시스템	데이터 모델링	메타데이터 유형	기반	구조 정보	제사용	특 징
OVID	No	Video Object	Annotation	No	메타데이터	스키마 및 Video 구조 정보가 없음 Raw Video 제사용 불가 메타데이터 유형이 단순하고, 객체지향 설계
VideoSTAR	Yes	Content Annotation 등 9가지 유형	Annotation	Yes	메타데이터, Raw Video	물리적 비디오의 공유와 메타데이터의 제사용 가능 공간관계 및 영상 특성 사용 불가 데이터 모델이 복잡
AVIS	No	Video Segment Tree 등 4가지 유형	Annotation	No	메타데이터	Video Segment Tree를 이용한 인덱싱 Video의 내용물 이벤트 위주로 처리 Raw Video 제사용 불가 Video 구조 정보가 없음 공간관계 및 영상 특성 사용 불가
Algebraic Video	No	Algebraic Video	Annotation	No	메타데이터	생성, 구성, 브라우징 등 다양한 연산 정의 Raw Video 제사용 불가 Video 구조 정보가 없음 메타데이터 유형이 단순
Venus	No	Object 등 5가지 유형	Feature	No	없음	프레임 기반특징추출정보를 메타데이터로 구성 프레임 기반특징추출에 의한 대용량의 메모리 필요 Raw Video 제사용 불가 Video 구조 정보가 없음
제안된 모델	Yes	Video Object 등 8가지 유형	Annotation, Feature	Yes	메타데이터, Raw Video	비디오 스키마 및 Video 구조정보 제공 Raw Video 제사용 가능 대표키 프레임 영상 기반의 예시 특징추출정보를 메타데이터로 구성 물리적 비디오 데이터 및 메타데이터 공유 가능(인터넷 방송 컨텐츠 제 작 시 정보공유 및 제사용 가능) 비디오 데이터 생성, 구성, 브라우징 등의 다양한 연산 정의 추석기반 및 내용기반 통합 비디오 데이터 모델 제공 비디오 데이터 모델이 단순 비디오의 하위레벨 특징정보 및 상위레벨 특징정보 추출 용이 대표키 프레임 영상 기반 예시 특징정보 추출 시 사용자 인터페이스 제공 비디오 메타데이터 유형이 단순

표 9 비디오 질의 처리 및 브라우저 비교

항목 시스템	질의언어 또는 Syntax	추석 기반	내용 기반	유사 질의	공간 질의	시간 질의	브라 우징	비주얼질의 인터페이스	특 징
OVID	VideoSQL	Yes	No	No	No	No	No	Yes	스키마 없이 추석기반 질의만 허용 시공간질의 불가 및 질의언어가 단순
VideoSTAR	없음	Yes	No	No	No	Yes	Yes	No	추석기반 질의만 허용 질의언어 지원 없음
AVIS	Find From Where	Yes	No	Yes	No	No	No	No	사전위주의 질의유형 정의 시공간질의 불가 질의언어가 단순
Algebraic Video	없음	Yes	No	No	No	Yes	No	No	질의연산 없이 객체의 손실에 시간관계를 값으로 사용 공간질의 불가 및 질의언어가 단순
Venus	CVQL	No	Yes	Yes	Yes	Yes	No	Yes	Macro 정의에 의해 유사질의 지원 객체의 움직임에 대한 연산정의 비디오의 상위레벨 내용 질의 불가
제안된 모델	상용 SQL, visual script	Yes	Yes	Yes	Yes	Yes	Yes	Yes	질의 검색기능과 브라우징 검색기능의 사용자 인터페이스 제공 대표키 프레임 영상 기반 예시 특징정보 추출 시 사용자 인터 페이스 제공 다양한 비디오 질의 방법 제공(추석기반질의, 내용기반질의, 추석기반 및 내용기반 질의) 다양한 비디오 질의 연산방법 제공 비디오 질의 연산방법에 의해 질의어 없이 질의가 가능토록 사용자 인터페이스 제공 트리 구조로 표현된 비디오 구조 정보를 Clip별→Cur별→대표 키 프레임 영상별로 내용 브라우징 기능 제공 비디오 Clip별→Cur별→대표키 프레임 영상별 각자 별도의 검 색 기능과 재생기능, 브라우징 기능 제공

마지막으로 본 논문에서 제시한 비디오 데이터베이스 모델은 비디오의 의미적 구조와 사전 제작지식정보를 구조적으로 축적하여 데이터베이스를 구축하므로 정확한 인터넷 방송 콘텐츠 정보의 축적관리와 구축작업의 효율화가 가능하며, 또한 인터넷 방송 콘텐츠 제작 시 정보공유 및 재이용이 가능하므로 새로운 콘텐츠 제작의 효율성을 높일 수 있다.

### 5. 결론 및 향후 연구과제

본 논문에서는 효율적인 인터넷 방송 콘텐츠 제작 정보관리와 활용을 목적으로, 인터넷 방송 콘텐츠 제작과정에서 현재 충분하게 축적, 관리되고 있지 않은 비디오의 의미적 구조와 제작지식정보를 구조적으로 축적할 수 있는 내용기반 비디오 데이터베이스의 구축 방안과 생성기법을 제시하여 비디오 데이터베이스를 구현하였다. 이를 위해 본 논문에서는 비디오 콘텐츠 제작과정의 문자 정보와 비디오 정보를 비디오 데이터베이스에 생성할 수 있도록 비디오의 의미적 구조와 사전 제작지식정보를 기반으로 비디오 시퀀스의 내용별 특징정보를 조합하여 분류한 주석기반 및 내용기반 메타데이터 유형을 통합한 비디오 데이터 모델을 제시하였다. 그리고 대표키 프레임 영상의 주석기반 특징정보와 내용기반 특징정보의 메타데이터를 비디오의 색인으로 구성하기 위해, 비디오의 하위 레벨에 표현된 대표키 프레임 영상의 예시 특징정보를 추출하는 방법과 비디오 데이터베이스의 주석기반 및 내용기반 메타데이터 유형을 비디오의 인덱싱으로 구성 방법을 제시하였다. 본 논문에서 제안하여 구현한 비디오 데이터베이스 시스템은 사용자가 질의어에 대한 지식이 없더라도 쉽고 정확하게 비디오 데이터를 질의하고 검색할 수 있도록 주석기반 질의, 내용기반 질의, 주석기반 및 내용기반 통합질의 기능을 가진 비디오 질의 처리 인터페이스와 비디오 브라우징 인터페이스를 제공하여 질의 및 검색에 편리성을 제공하였다. 또한 비디오 브라우저는 각 비디오에 대한 구조를 쉽게 파악할 수 있도록 트리 구조로 표현된 비디오 구조 정보를 내용 브라우징 할 수 있도록 하였으며, 비디오 구조별 세부기능으로는 각자 별도의 검색 기능과 재생기능, 그리고 내용 브라우징 기능을 갖도록 지원하였다. 마지막으로 본 논문에서 제시하여 구현한 비디오 데이터베이스 시스템은 비디오의 의미적 구조와 사전 제작지식정보를 구조적으로 축적하여 데이터베이스를 구축하므로 정확한 인터넷 방송 콘텐츠 정보의 축적관리와 구축작업의 효율화가 가능하며, 또한 인터넷 방송 콘텐츠 제작 시 정보공유 및 재이용이 가능하므로 인터넷

방송의 TV중계, 원격수업, 뉴스전송, 원격진료 등과 같은 다양한 응용서비스의 새로운 콘텐츠 제작에 효율성을 높일 수 있다. 그리고 인터넷 방송의 다양한 멀티미디어 내용을 시간과 공간의 제약 없이 양방향의 대화형 메타 방송에 이용될 수 있으며, 웹 콘텐츠에 의해 사용자가 원하는 콘텐츠에 대한 실시간 서비스를 할 수 있다.

향후 연구과제로는 본 연구에서 제안하여 구현한 내용기반 비디오 데이터베이스 시스템을 다양한 멀티미디어 콘텐츠에 적용될 수 있도록 확장하는 것이다. 그리고, 향후 인터넷 방송은 보다 고도의 멀티미디어 콘텐츠 제작에 의해 방송될 것으로 예상된다. 따라서, 복수의 스토리를 선택적으로 시청하기 위한 멀티스토리 방송 콘텐츠의 제작 등 구조화된 복잡한 정보를 정확하게 표현하고 조작할 수 있도록 멀티미디어 인터넷 방송 데이터베이스에 관한 연구가 필요하다. 그리고, 멀티미디어 인터넷 방송 데이터베이스의 에이전트 기술 등에 의한 콘텐츠 등록과 검색이 가능하도록 정보 환경의 구축이 기대되고 있다. 인터넷 방송 콘텐츠 서비스는 이용자 측에도 상당 부분 권한이 있기 때문에 이들에게 다양한 채널과 내용을 자유롭게 선택할 수 있도록 멀티미디어 형식으로 구성된 다양한 콘텐츠를 데이터베이스화하여 사용자에게 효율적으로 서비스할 수 있는 연구가 진행되어야 한다.

### 참고 문헌

- [1] 김우식, 방송제작론, p.297, 良書閣, 서울, 2000.
- [2] 권오형, 김용석, 김진용, 안치득, "국내 데이터 방송 표준화 현황", 정보과학회지: 제18권, 제10호, pp.4-11, 2000.10.
- [3] 양영철, 영화·TV제작입문, p.269, 학문사, 서울, 2000.
- [4] 황인선, 한인규, 영상제작기법, p.382, 圖書出版 技多利, 서울, 1997.
- [5] Eitetsu Oomoto, Katsumi Tanaka, "OVID: Design and Implementation of a Video Object Database System," IEEE Transactions on Knowledge and Data Engineering, Vol.5, No.4, pp. 629-643, Aug 1993.
- [6] R. Hjelmsvold, "VideoSTAR-A Database for Video Information Sharing," Ph.D. Thesis, Norwegian Institute of Technology, Nov. 1995.
- [7] Sibel Adali, et. al., "The Advanced Video Information System: data structures and query processing," Multimedia System, pp.172-186, 1996.
- [8] R. Weiss, A. Duda, and David K. Gifford, "Composition and Search with a Video Algebra," IEEE Multimedia, Spring 1995.

- [9] Ki-Wook Kim, Ki-Byoung Kim, Hyoung-Joo Kim, "VIRON: An Annotation-Based Video Information Retrieval System," In Proc. The Twentieth Annual International Computer Software and Applications Conference, 1996.
- [10] Myron Flickner and et. al, "Query by Image and Video Content: The QBIC system," IEEE Computer, 28(9), pp.23-32, 1995.
- [11] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Pekovic, P. Yanker, C. Faloutsos and G. Taubin, "The QBIC project: Querying images by Content using color, texture, and shape," Proc. of SPIE Storage and Retrieval for Image and Video Databases, pp.173-187, Feb 1993.
- [12] Shih-Fu Chang, William Chen, Horace meng, Hari Sundaram, and Di Zhong, "An Automated Content-Based Video Search System Using Visual Cues," ACM Multimedia, 1997 <http://www.ctr.columbia.edu/videoq>.
- [13] Tony C.T.Kuo and Arbee L.P.Chen, "A Content Based Query Language for Video Databases," In Proc. of IEEE MULTIMEDIA '96, pp.209-214, 1996.
- [14] A. Hauptmann and M. Witbrock, "Informedia News-On-Demand: Using Speech Recognition to create a Digital Video Library," May 1997, <http://www.informedia.cs.cmu.edu/>.
- [15] N. Dimitrova and M. A. Mottalab, "Content-based video retrieval by example video clip," SPIE Storage and Retrieval for Image and Video Database, Vol. 3022, pp.59-70, 1996.
- [16] H. J. Zhang, C. Y. Low, S. W. Smoliar and J. H. Wu, "Video Parsing Retrieval and Browsing: An Integrated and Content-Based Solution," Proc. ACM Multimedia '95, pp.15-24, 1995.
- [17] A. Nakasaka and Y. Tanaka, "Automatic Video Indexing and Full Video Search for Object Appearances," in Proc. IFIP TC2/WG2.6 second Working Conf. On Visual Database systems., pp.113-133, Oct 1991.
- [18] H. Zang, A. Kankanhalli, and S. Smoliar, "Automatic Partitioning of Full-Motion Video," Proc. ACM Multimedia System, ACM Press, New York, Vol. 1, pp.10-23, 1993.
- [19] R. Zabih, J. Miller, and K. Mai, "A Feature-Based Algorithm for Detecting and Classifying Scene Breaks," Proc. of ACM Multimedia '95, San Francisco, CA, pp.189-200, 1995.
- [20] R. Kasturi, R. Jain, "Dynamic vision," In Computer vision Principles, IEEE Computer Society Press, Washington, 1991.
- [21] H. Ueda, T. Miyatake, and S. Yoshizawa, "IMPACT: An Interactive Natural-motion-picture Dedicated Multimedia Authoring System," in proceedings of CHI, New York, pp.343-350, 1991.
- [22] Gulrukh Ahanger, and Thomas D. C. Little, "A Survey of Technolgies for Parsing and Indexing Video," Journal of Visual Communication and Image Representation, San Jose, February 1994.
- [23] J. S. Boreczky and L. A. Rowe, "A Comparison of Video Shot Boundary Detection Techniques," Journal of Electronic Imaging, 5(2), pp.122-128, Apr 1996.
- [24] D. Swanberg, C.F. Shu and R. Jain, "Knowledge Guided Parsing in Video Database," Proc. of SPIE-Storage and Retrieval for Image and Video Databases, San Jose, CA, Vol. 1908, pp.13-24, 1993.
- [25] Ramin Zabih, Justin Miller, Kevin Mai, "A feature-based algorithm for detection and classifying scene breaks," Proc. of ACM Multimedia '95, San Francisco, CA, pp.189-200, 1995.
- [26] F. Ahanger, A. Hsu, and M. Y. Chiu, "Image Processing on Compressed Data for Large Video Databases," Proc. 1st ACM Intl. Conf. on Multimedia, Anaheim CA, pp. 267-272, Aug 1993.
- [27] F. Idris and S. Panchanathan, "Storage and Retrieval of Compressed Image," IEEE Transactions on Consumer Electronics, Vol. 41, pp.937-941, August 1995.
- [28] Di Zhong, HongJiang Zhang, Shih-Fu Chang, "Clustering Methods for Video Browsing and Annotation," Storage and Retrieval for Image and Video Databases (SPIE), pp.239-246, 1996.
- [29] X. Sun, M. S. Kankanhalli, Y. Zhu, and J. Wu, "Content-Based Representative Frame Extraction for Digital Video," Intl. Conf. on Multimedia Computing and Systems, Austin Texas, pp.237-240, Jun 1998.
- [30] 김우성, "계승된 특징 차이를 이용한 효율적인 대표키 프레임과 관련 키 프레임 추론 기법", 정보과학회 논문지(B), 제26권 제5호, pp.657-664, 1999.5.
- [31] J. Rumbaugh et al., "Object-oriented modeling and design," Prentice-Hall International, Inc., 1991.
- [32] H.J Zhang et al, "Automatic Parsing and Indexing of News Video," ACM Multimedia Systems, ACM-Springer, pp.256-266, 1996.
- [33] V. N. Gudivada and V. V. Raghavan, "Content-Based Image Retrieval Systems," IEEE Computer, 28(9), 1995.
- [34] V. V. Vinod, H. Murase and C. Hashizume, "Focussed Color Intersection with Efficient Searching for Object Detection and Image Retrieval," Proceedings of the IEEE International Conference on Multimedia Computing and System,

- pp.229-233, 1996.
- [35] 홍성용, 나연목, "혼합형 이미지 메타데이터를 이용한 지능적 이미지 검색 시스템 설계 및 구현", 멀티미디어학회논문지, 3권 3호, 한국멀티미디어학회, pp.209-223, 2000.6.
- [36] M. Sonka, V. Hlavac, and R. Boyle, "Image Processing Analysis, and Machine Vision," Brooks/Cole Publishing, Second Edition, 1999.
- [37] Randy Crane, "A Simplified approach to Image Processing," Prentice-Hall, 1997.
- [38] D. Shasha and T-L Wang, "New techniques for best-match retrieval," ACM TOIS, 8(2) pp.140-158, April 1990.
- [39] 양명섭, "통합된 인덱싱 방법을 사용한 내용기반 뉴스 비디오 검색 시스템의 설계 및 구현", 전북대학교 박사학위논문, 1999.
- [40] 김수경, 정성환, "웨이블릿 변환영역에서 칼라 히스토그램과 에너지 벡터를 이용한 컷 검출", 한국정보과학회, 추계학술발표회, 제27회, CD-paper #21, 2000.10.
- [41] G. Salton and M. J. McGill, "Introduction to Modern Information Retrieval," McGraw-Hill, 1983.

이 태 동

정보과학회논문지 : 컴퓨팅의 실제  
제 7 권 제 5 호 참조

김 빈 구

정보과학회논문지 : 컴퓨팅의 실제  
제 7 권 제 5 호 참조