

VOD 시스템에서 내용기반 검색을 위한 데이터모델에 관한 연구

전근환*, 신성윤**, 강오형**, 이양원**, 류근호*

*충북대학교 전자계산학과

**군산대학교 컴퓨터정보과학과

e-mail:khjeon@kunjang.ac.kr

A Study on Data Model for Content-based Retrieval in VOD System

Keun-Hwan Jeon*, Seung-Yoon Shin**, Oh-Hyeong Kang**, Yang-Weon Rhee**, Keun-Ho Ryu*

*Dept of Computer Science, Chung-Buk University

**Dept of Computer Information Science, Kun-san University

요약

비디오 데이터의 내용기반 검색을 위해서는 비디오가 가지는 많은 객체와 이들간의 관계성에 관한 정보를 표현할 수 있는 데이터모델이 요구된다. 이 논문에서는 이러한 비디오 데이터의 내용정보를 표현할 수 있는 메타데이터 모델을 제안한다. 제안 모델은 특정 응용분야에 국한하지 않도록 물리적 정보와 논리적 정보 모듈을 포함하기 위하여 OMT기법을 이용하여 설계하였다. 이 논문에서는 제안한 모델을 이용하여 VOD 시스템에 적용할 수 있는 영화 비디오 표현을 제시하였다.

1. 서론

VOD(Video-on-Demand) 시스템은 컴퓨터 네트워크를 이용하여 데이터베이스와 멀티미디어 기술을 가능하게 한 대표적인 응용분야라 할 수 있다. 멀티미디어 데이터의 대표적 데이터인 비디오데이터는 데이터 자체가 시간적 구조(temporal structure)와 공간적 구조(spatial structure)로 이루어져 있기 때문에, 비디오 데이터에 대한 내용 기반 검색은 두 관계를 중심으로 이루어질 수 있다. 비디오 데이터는 많은 객체들을 포함하고 있으며, 여러 미디어가 혼합되어 있기 때문에, 이들 간의 관계성을 이용한 정보검색을 하려면, 시간동기화, 시간 전후 관계, 공간관계 등이 먼저 멀티미디어 데이터 모델로써 표현되어야 한다. 비디오가 다른 이미지 영상과 가장 다른 점은, 시간에 따라 끊임없이 변화한다는 것이다. 따라서 비디오 데이터를 의미적으로 추상화하고, 그들 간의 시간적 표현을 간략히 표현 가능하도록 하는 비디오 모델이 제공되어야 하고, 이를 기반으로 한 시간 질의어(temporal query), 시간 함수(temporal operator) 등이 지원되어야 한다. 또한 비디오 데이터의 색인 기능을 수행하기 위한 구조적이

고 확장성 있는 메타데이터 모델의 정립은 비디오 데이터에 대한 내용기반 검색을 보다 효율적으로 지원하기 위한 필수적 요인이다.[1,2].

이 논문에서는 비디오 데이터에 대한 내용기반 검색을 지원하기 위한 메타데이터 모델을 제안한다. 효과적인 모델제안을 위해 2장에서 관련연구를, 3장에서 비디오 데이터 모델에서 포함되어야 하는 속성 정의를, 4장에서 메타데이터 모델을 제안한 후, 5장에서 결론을 맺는다.

2. 관련연구

비디오 데이터는 시청각 정보뿐만 아니라 의미적 정보를 포함하고 있기 때문에 이에 대한 정보를 구하는 사용자의 질의를 모두 지원할 수 있어야 한다. 이를 위해서는 비디오 데이터에 대한 정보의 추출뿐만 아니라 추출된 정보를 체계적으로 구축하여 관리하기 위한 비디오의 응용분야에 관련된 사용자의 요구를 분석하고, 질의 유형을 파악하여 목적에 부응하는 메타데이터를 분류하고, 모델링하는 작업이 우선되어야 한다.

비디오 데이터에 대한 메타데이터 구축은 내용이

반 검색에 있어서 이정표와 같은 역할을 하므로 효율적인 메타데이터의 구축 여부는 비디오 데이터베이스의 성능을 높이고 가용성의 향상 여부를 좌우할 수 있다. 또한, 모든 비디오 데이터는 자신에 대한 정보를 메타데이터를 통하여 표현하고, 사용자는 어떠한 질의를 수행하더라도 메타데이터를 통하여 실제 데이터에 접근하게 됨으로써 메타데이터는 사용자와 비디오 데이터를 연결하는 매개자 역할을 수행한다. 따라서 비디오 데이터를 효율적으로 관리하고 제어할 수 있는 기술과 함께 이들을 손쉽게 검색할 수 있는 메타데이터의 구축과 관리가 성공적인 비디오 데이터베이스 시스템을 구축하는 초석이 될 수 있다.

미국 Berkeley 대학에서 개발한 Plateau는 교육용 비디오를 제공하는 VOD 시스템으로 수백 시간 분량의 교육용 비디오를 저장한 비디오 데이터베이스를 근거리 네트워크나 인터넷을 통해 사용자에게 제공하고 지역적으로 서로 다른 위치에 저장된 비디오를 지역 비디오 파일서버에 저장해 두었다가 재생시키는 분산 VOD 시스템이다. Plateau 시스템 개발에 앞서 사용자로부터 동영상에 대한 설문조사를 실시하여 시스템에서 필히 제공되어야 하는 질의 타입을 도출한 후 메타데이터를 비디오 일반정보, 계층정보, 그리고 키 프레임에 포함된 개체, 대사등을 포함하는 내용기반검색정보로 분류하였다[3].

ViMOD[4]는 미국 Michigan 대학에서 개발한 비디오 제작을 위한 메타데이터 모델로서 Xenomania [5]와 같은 시각정보 관리 시스템에서 참고로 사용하고 있다. ViMOD에서는 질의의 특성을 정확매칭(exact match)과 유사매칭(inexact match)으로 구분한 매칭정도(certainty), 시간(temporal)과 공간(spatial)으로 구분한 차원성(dimensionality), 그리고 데이터의 통계적 수치를 기반으로 한 의존성(dependence) 등 세 가지 관점에서 메타데이터를 분류하였다.

VideoSTAR[6]는 비디오 데이터와 메타데이터의 공유와 재사용을 지원하기 위해 노르웨이 공대에서 개발된 비디오 데이터베이스 시스템이다. VideoSTAR는 범용 비디오 데이터베이스 시스템을 지향한 프로토타입 시스템으로 응용 프로그램에 의존하는 질의 유형들의 분석은 제외하고, 메타데이터를 미디어 데이터, 미디어 데이터간의 관계를 정의한 데이터, 비디오 문서 버전, 전체정보, 구조정보, 주석정보, 특징정보, 줄거리, 기타 등 9가지로 확장

하였다.

3. 비디오 메타데이터 모델링

3.1 비디오 데이터의 특성 분석

2장에서 기술한 비디오 데이터의 특성을 기반으로 하여 이 연구에서 내용 기반 검색을 위한 메타데이터를 구축하기 위해 비디오 데이터의 특성을 다음과 같이 분류하였다.

- (1) 물리적 특성(Physical Feature)
- (2) 논리적 특성(Logical Feature)
- (4) 시각적 특성(Visual Feature)
- (5) 시간 특성(Temporal Feature)
- (6) 공간 특성(Spatial Feature)
- (7) 일반적 특성(General Feature)
- (8) 구체적 특성(Specific Feature)
- (9) 구조적 특성(Structural Feature)

위와 같은 비디오 데이터의 특성을 포함하기 위한 메타데이터를 크게 내용기반 정보와 내용 독립 정보로 나누어 다음 표 1과 같이 분류할 수 있다.

<표 1> 메타데이터의 속성 분류

분류	클래스	속성	
Content-based information	Topic annotation data	subject : 주제	
		keyword : 주제어	
		type : 주제가 관련된 분야	
		actor : 등장 인물	
	Event	main_actor	
		play_time	
	Low-level data	spatio-temporal data	object position
			frame duration
		visual feature data (샷 및 카메라의 특수효과)	shot angle
			shot distance
shot motion			
shot effect			
camera_pan			
camera tilt			
camera_height			
Content-independent information	General data	title : 제목	
		date : 제작일	
		producer : 제작사	
		staff&cast: 제작진과 배우	
		genre : 장르	
		language : 사용 언어	
		summary : 줄거리	
	running time(total frame#)		
	Structural data	shot	
		scene	
sequence			
		Video Document	

3.2 메타데이터를 이용한 시·공간 특성 모델링

이 절에서는 표 1의 속성을 기반으로 다음과 같은 조건을 만족할 수 있는 메타데이터를 모델링 한다.

- 첫째, 검색하고자 하는 내용에 관하여 체계적이고 구조적이어야 한다.
- 둘째, 어떤 종류의 비디오 데이터도 내용기반 검색이 가능하도록 공통적, 포괄적 내용을 포함해야 한다.
- 셋째, 내용을 근거로 하거나 내용에 무관한 내용, 그리고 물리적 내용도 다룰 수 있어야 하고, 비디오 데이터의 가장 대표적인 특징인 시·공간적인 특성을 다룰 수 있어야 한다.
- 넷째, 새로운 데이터의 추가가 용이하도록 융통성있는 구조를 가져야 하며, 비디오 데이터의 형식에 구애받지 않는 데이터 독립성을 제공해야 한다.

3.2.1 비디오 내용 독립정보를 위한 메타데이터

- 1) 비디오 일반 정보: 장르나 응용분야에 관계없이 모든 비디오가 가지고 있는 특성으로 비디오 형식과 제작에 관한 내용을 포함한다.
- 2) 비디오 구조 정보: 비디오 구조는 물리적 구조 측면인 시퀀스, 씬, 샷의 물리적 구성을 모델링한다.

3.2.2 비디오 내용기반 정보를 위한 메타데이터

- 1) 시퀀스 정보 : 시퀀스는 비디오 계층 구조에서 비디오 다음의 계층에 속하는 종류(Type)다. 시퀀스는 전체적인 내용의 흐름에 따라 기, 승, 전, 결 과 유사한 분류이다.
- 2) 씬 정보 : 비디오의 씬은 논리적인미를 갖는 단위로서 하나 이상의 씬들이 모여 주제를 구성한다.
- 3) 샷 정보 : 하나의 카메라가 멈추지 않고 연속으로 촬영한 영상을 의미하는 샷은 연속된 프레임으로 구성된다. 이러한 샷은 키 프레임으로 구분되며 중심객체에 대한 카메라의 특수효과가 부여되는 특징을 갖는다.
- 4) Object : 샷을 구성하는 객체에 대한 모델링이다. 객체에 대한 메타데이터는 객체 이름과 객체를 구성하는 컬러정보, 객체의 종류에 관한 정보를 갖는다.
- 5) Visual Feature : 시각적 특징은 샷이 갖는 물

리적 정보로서 편집정보나 카메라의 특수효과정보를 갖는다.

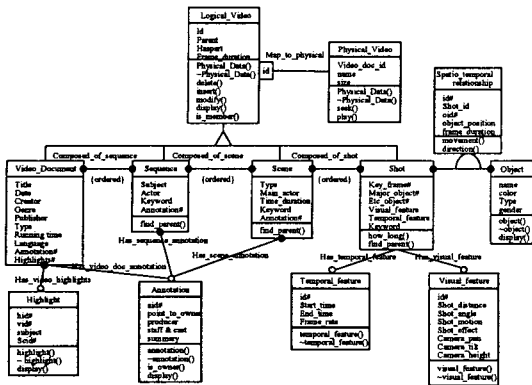
- 6) Temporal Feature : 샷이 갖는 시간적 특징으로서 샷이 갖는 시간간격에 대한 정보와 동기화를 위한 초당 프레임 속도 정보를 갖는다.
- 7) Spatio-Temporal relationship : 시·공간 관계정보는 샷에 포함된 객체가 시간의 흐름에 따라 변화되는 물리적 위치 정보로서 객체들 간의 상대적 위상, 방향 등의 공간관계를 추론할 수 있는 정보를 갖는다.
- 9) 하이라이트 정보 : 하이라이트는 비디오 계층 구조와는 독립적이며 비디오의 주요 장면들에 대한 클러스터 집합으로 구축된다.

4. 메타데이터 제안 및 적용

현재 개발되었거나 개발 중인 비디오 데이터베이스 시스템들에서 메타데이터의 추출 및 모델링 작업은 시스템 설계 단계에서 이루어지고 있으며 모델링 방법으로 ER(Entity-Relationship)모델, 객체지향모델 방법 등을 사용하고 있다. 2장에서 조사한 Plateau 는 ER모델로, VideoSTAR는 EER모델로 비디오 데이터를 모델링 하였다. ER모델이나 EER 모델들은 엔티티의 속성과 엔티티 간의 관계만 표현 가능하지만, 객체지향 모델은 엔티티를 클래스화하여 속성과 연산을 할 수 있고 다음과 같은 장점을 가진다.

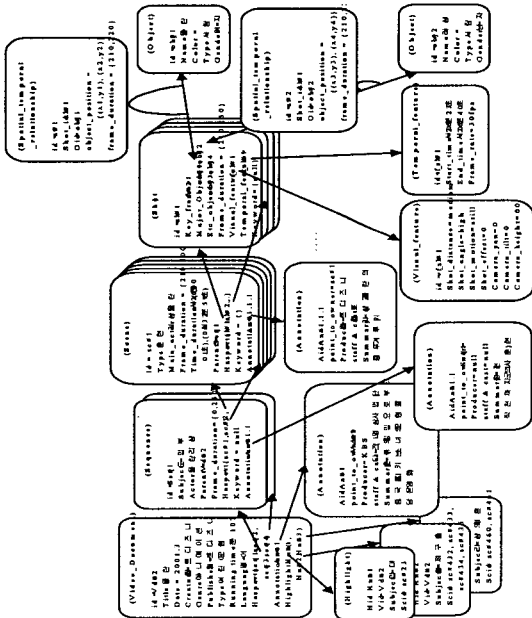
- 첫째, 클래스 계층 간의 정보의 상속으로 정보의 불필요한 중복을 줄여 모델이 간단하다.
- 둘째, 각 클래스간의 관계와 관계 사이의 속성들도 클래스화 할 수 있으므로 관계의 정의가 명확해진다.
- 셋째, 연산의 오버로딩(overloading)기능으로 데이터 관리에 확장성을 제공한다.
- 넷째, 모듈화가 가능하므로 모델에 확장성과 융통성을 가진다.

이에 이 연구에서는 Rumbaugh의 OMT 기법의 객체 모델 구성단계를 이용한 메타데이터 모델을 다음 그림 1과 같이 구성하였다.



(그림 1) 메타데이터 객체 다이어그램

본 장에서 제안한 모델은 범용성과 확장성을 목표로 비디오가 가질 수 있는 많은 정보를 표현할 수 있도록 일반적, 구체적, 시각적, 구조적, 시간적, 공간적 특성을 추출하여 구축하였다. 특히 시공간적인 특성 표현을 위하여 샷과 객체간의 관계를 정의한 Spatio_temporal_relationship 클래스를 이용하여 시공간 질의가 가능하며 세분화 정도에 따라 보다 상세한 비디오 질의가 가능하게 하였다. 그림 2는 영화 "물란"에서 주인공 "물란"과 "리상"이 군사 훈련 도중 쿡푸 겨루기를 하는 장면에 대해 메타데이터를 구축한 예이다.



(그림 2) 영화 "물란"에 대한 메타데이터 구축 예

5. 결론

멀티미디어 데이터의 대표적 데이터인 비디오 데이터는 데이터 자체가 시공간 구조로 이루어져 있기 때문에 비디오 데이터에 대한 내용기반 검색은 두관계를 중심으로 이루어질 수 있다. 이러한 시공간 관계를 기반으로 하는 내용기반 검색은 이들 특성을 표현할 수 있는 데이터 모델이 요구된다.

이 논문에서는 내용기반 검색이 가능할 수 있게 하기 위해 비디오에 포함된 많은 객체들과 그들간의 시공간 관계를 표현할 수 있는 비디오 데이터 모델을 제안하였다. 제안한 비디오 데이터 모델은 논리적 특성인 상위레벨 정보 뿐만 아니라 프레임적용된 카메라 및 샷의 특수효과 등의 물리적 특성인 하위레벨 정보까지를 포함하여 특정 응용분야에 국한되지 않는 범용성을 갖게 하였다.

참고문헌

- [1] Sheth, B., "A Learning Approach to Personalized Information Filtering", MIT Department of EECS MS Thesis, Feb. 1994
- [2] William I. Grosky, Ramesh J., Rajiv Mehrotra, "The Handbook of Multimedia Information Management", Prentice Hall PTR, 1997
- [3] L.A. Rowe, J.S. Boreczky, and C.A. Eads, "Indexes for User Access to Large Video Databases," Storage and Retrieval for Images and Video Databases II, IS&T/SPIE Symposium on Electronic Imaging Science & Technology, San Jose, CA, February 1994
- [4] R. Jain and A. Hampapur., "Metadata in Video Databases", ACM SIGMOD RECORD, 23(4):27-33, 1994
- [5] Bach, J., Paul, S. & Jain, R., "A visual information management system for interactive retrieval of faces", IEEE Transactions on Knowledge and Data Engineering, 5(4), 619-628, 1993
- [6] Rune Hjelsvold and Roger Midtstraum., "Modelling and querying video data", In Proceedings of the 20th International Conference on Very Large Databases (VLDB'94), Santiago, Chile, 1994