

# 적응형 VOD 시스템을 위한 비디오 메타 데이터 모델

전근환, 신예호

군장대학 컴퓨터응용학부, 극동대학교 정보통신학부

## Video Meta-data model for Adaptive Video-on-Demand System

Keun Hwan Jeon, Ye Ho Shin

School of Computer Application in Kunjang College, School of Communication  
Infomation in Far East University

### Abstract

The data models which express all types of video information physically and logically, and the definition of spatiotemporal relationship of video data objects. In This paper, we classifies meta-model for efficient management on spatiotemporal relationship between two objects in video image data, suggests meta-models based on Rambaugh's OMT technique, and expanded user model to apply the adaptive model, established from hyper-media or web agent to VOD. The proposed meta-model uses data's special physical feature: the effects of camera's and editing effects of shot, and 17 spatial relations on Allen's 13 temporal relations, topology and direction to include logical presentation of spatiotemporal relation for possible spatiotemporal reference and having unspecified applied mediocrity.

### 1. 서론

VOD(Video-on-Demand) 시스템은 컴퓨터 네트워크를 이용하여 데이터베이스와 멀티미디어 기

술을 가능하게 한 대표적인 응용분야라 할 수 있다. 멀티미디어 데이터의 대표적 데이터인 비디오 데이터는 데이터 자체가 포함된 객체들의 공간적 구조(spatial structure)와 초당 재생되는 프레임 수로 인해 생성되는 시간적 구조(temporal structure)로 이루어져 있기 때문에, 비디오 데이터에 대한 내용 기반 검색은 두 관계를 중심으로 이루어져야 한다. 비디오 속에서 시공간적 정보는 장면 속에 포함된 객체의 이동을 의미한다. 시공간적 특징을 기반으로 하는 검색은 범인 추적, 영화 편집, 비디오 콘텐츠 제작, 비디오를 이용한 모니터링 등 많은 응용분야에서 중요하다.

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. 그 중에서 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 구조적이고 확장성 있는 메타 데이터 모델의 정립은

비디오 데이터에 대한 내용기반 검색을 보다 효율적으로 지원하기 위한 필수적 요인이다[1,2].

메타 데이터는 비디오 데이터의 내용, 구조 그리고 프레임에 존재하는 객체들의 각종 정보를 다룬다. 메타 데이터의 정보는 응용 분야에서 필요한 매체 타입과 요구되는 정보의 형태에 의존한다. 비디오 데이터는 비디오 처리 기술에 의존하여 카메라 샷(shot), 샷 안의 성격, 샷의 배경 등을 식별해야 한다[3,4]. 식별된 정보들은 내용 기반 검색을 지원하기 위하여 사용자 질의를 지원할 수 있도록 적절히 구조화해야 한다[5]. 이 논문에서는 비디오 데이터도 일반 데이터와 같이 취급될 수 있도록 데이터베이스에 저장하고 관리되며, 데이터베이스 안에서 해석될 수 있는 비디오 데이터 모델을 제안하고자 한다.

## 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 메타 데이터를 분류하고, 모델링하는 작업이 우선되어야 한다.

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

구축하는 초석이 될 수 있다[8].

이 절에서는 비디오 데이터베이스 시스템으로 개발된 Plateau[6], ViMOD[4], VideoSTAR[7] 등 3가지 프로토타입 시스템을 중심으로 이 연구의 기반 모델인 메타 데이터모델 제안을 위한 기존 연구를 조사하고 기존모델들의 문제점을 살펴본다.

비디오 데이터의 특성중 이미지 처리 과정에서 자동으로 추출될 수 있는 비디오의 특성을 하위 레벨 특성이라 하고, 주로 사람에 의해 인지되고 입력되어야 하는 비디오의 특성을 상위 레벨 특성이라고 한다. Plateau[6]는 주로 상위 레벨 특성의 메타 데이터를 다루었기 때문에 내용에 대한 질의는 쉽게 처리 가능한 반면 장면 묘사 등 하위 레벨의 특성을 묻는 구체적 질의를 처리하기 용이하지 않다. ViMOD[4]의 메타 데이터는 뉴스 제작이라는 특수한 응용분야에 맞추어 모델링되었기 때문에 상위 레벨보다 하위 레벨의 특성을 상세히 분류하여 다루고 있다. 따라서 프레임 단위나 샷 단위의 상세한 하위레벨의 질의에는 유용하지만 비디오의 내용과 관계있는 특성을 다루는 메타 특성만으로는 영화나 스포츠 비디오내의 주제 및 이벤트 내용을 질의하기에는 부족하다. 반면 VideoSTAR[7]에서는 범용성을 지향하고 있기 때문에 Plateau나 ViMOD에 비해 상위레벨과 하위레벨 특성을 비교적 균형있게 다루고 있으나 범용성을 지향한 메타 데이터로서 상위 레벨에 비해 하위레벨의 특성들이 부족하다.

Plateau와 ViMOD는 각각 고유한 응용분야를 위해 개발된 시스템들로서 메타 데이터의 분류를 위해 먼저 질의어의 특성을 조사하여 정의하고 이에 적합한 메타 데이터를 분류하고 모델링하였기 때문에 각 응용분야에서는 전문성을 띠고 있으나 메타 데이터 모델링에 있어서는 범용성이 부족하다.

VideoSTAR는 범용 비디오 데이터베이스 시스템을 위해 개발된 프로토타입 시스템으로 일반적인 메타 데이터를 도출하여 메타 데이터 모델링 분야에서는 범용성이 돋보이나 모델이 복잡하고 응용분야 의존적인 부분에 대해서는 고려되지 않았기 때문에 이 모델을 이용하고자 하는 시스템에서는 응용에 의존적인 부분을 다시 고려해야하는

단점이 있다.

다음 표 1은 메타 데이터 모델들의 속성을 분석한 결과이다.

표 2. 기존 메타 데이터 종합 분석

시스템	속성분류	분류속성	분석	적용분야
Plat <u>e</u> au	사용자설문조사	Bibliographic Data	상위레벨 정보	VOD
		Structural Data	정의 우수	
		Content Data	하위레벨 정보	
		keyword	정의 없음	
		Representational	체계적인 분류가 취약	
ViMO <u>D</u>	매칭차원성존성의성	Video Q-frame	하위레벨 정보 정의 우수	뉴스제작
		Video R-frame		
		Image Q-frame	상위레벨 정보 분류 미흡	
		Image R-frame		
		Meta-feature		
Vid <u>e</u> o- <u>S</u> T <u>A</u> R	미분류	media-specific data	비디오 데이터의 공유와 재사용을 고려한 데이터 모델링 제공	범용
		compositional data		
		version data		
		bibliographic data		
		structural data		
		content annotations		
		content feature data		
		topic annotations		
		supplementary information		
	특정 응용분야에서는 재정의 또는 속성 추가 요구			

### 3. 메타데이터를 이용한 시-공간 특성 모델링

이 절에서는 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.

### 3.1 비디오 내용 독립정보를 위한 메타데이터

#### 1) 비디오 일반 정보

장르나 응용분야에 관계없이 모든 비디오가 가지고 있는 특성으로 비디오 형식과 제작에 관한 내용을 포함한다. 비디오 일반정보는 제작 정보와 형식정보로 구분하여 표 2와 같이 모델링한다.

표 3 비디오의 일반정보 메타데이터

속성	내용	비고
vid#	비디오 식별번호	비디오 객체 관리번호
Title	제목	제작정보
Date	제작일	
Creator	생산자	
Publisher	출판자	
Genre	장르	형식정보
Type	비디오형식	
Running Time	상영시간	
Language	언어	
Haspart	하위 요소들에 대한 id	구성요소정보
Annotation#	비디오 주석 번호	주석정보

#### 2) 비디오 구조 정보

비디오 구조는 물리적 구조 측면인 시퀀스, 씬, 샷의 물리적 구성을 표 3과 같이 모델링한다.

표 4. 비디오의 물리적 구조정보  
메타데이터

속성	내용
Video_doc_id	vid#
name	비디오 파일 이름
size	비디오 데이터의 물리적 크기
sequence	비디오를 구성하는 시퀀스 개수
scene	전체 씬 구성 개수
shot	전체 샷 구성개수

표 6. 씬에 대한 메타데이터

속성	내용
scid#	씬 식별번호
Type	씬의 주제
Main_actor	씬에서의 주연
Time_duration	씬의 상영시간(간격)
Parent	상위 Video_document 번호
Haspart	하위 구성요소(샷)들에 대한 식별번호
Keyword	키워드
Frame_duration	씬의 전체 길이 (프레임 간격)
Annotation#	씬에 관한 주석 번호

### 3.2 비디오 내용기반정보를 위한 메타 데이터

#### 1) 시퀀스 정보

시퀀스는 비디오 계층 구조에서 비디오 다음의 계층에 속하는 데이터이다. 시퀀스는 전체적인 내용의 흐름에 따라 기, 승, 전, 결 과 유사한 분류이다. 표 4에서는 이들 시퀀스들의 속성정보를 갖는 메타데이터의 속성요소와 각 속성이 갖는 내용에 대해서 모델링한 항목들이다.

표 5. 시퀀스에 대한 메타데이터

속성	내용
seqid#	시퀀스 식별번호
Subject	시퀀스의 주제
Actor	주연배우
Parent	상위 Video_document 번호
Haspart	하위 구성요소(씬)들에 대한 식별번호
Keyword	키워드
Frame_duration	시퀀스 전체 길이 (프레임 간격)
Annotation#	시퀀스에 관한 주석 번호

#### 2) 씬 정보

비디오의 씬은 논리적의미를 갖는 단위로서 하나 이상의 씬들이 모여 주제를 구성한다. 표 5에서는 이들 씬들의 속성정보를 모델링한 메타데이터의 속성요소와 각 속성이 갖는 내용이다.

#### 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. 다음 표 6에서는 샷을 메타데이터로 모델링한 항목들과 각 항목이 갖는 내용을 나타내고 있다.

표 7. 샷에 대한 메타데이터

속성	내용
shid#	샷 식별번호
Key_frame#	키 프레임 번호
Major_object#	중심객체 식별번호
Etc_object#	기타 구성 객체 식별번호(집합)
Visual_feature	샷 및 카메라 특수효과 정보 번호
Temporal_feature	시간 속성정보 번호
Frame_duration	시작 프레임과 끝 프레임 번호
Keyword	키워드

#### 4) Object

샷을 구성하는 객체에 대한 모델링이다. 객체에 대한 메타데이터는 객체 이름과 객체를 구성하는 컬러정보, 객체의 종류에 관한 정보를 갖는다. 표 7은 객체에 관한 메타데이터 속성들이다.

표 8. 객체에 대한 메타데이터

속성	내용
oid#	객체 식별번호
Name	객체명
Color	구성 컬러
Type	객체 속성
Gender	사람 일 경우 성별

### 5) Visual Feature

시각적 특징은 샷이 갖는 물리적 정보로서 편집 정보나 카메라의 특수효과정보를 갖는다.

표 9. 영상 특징에 대한 메타데이터

속성	내용
vfid#	식별번호
Shot angle	샷 각도 정보
Shot distance	샷 거리 정보
Shot motion	샷 모션 정보
Shot effect	샷 특수효과 정보
Camera_pan	카메라 상하/좌우 회전 값(정수)
Camera tilt	카메라 기울기 값(정수)
Camera_height	카메라 높이 값(정수)

### 6) Temporal Feature

샷이 갖는 시간적 특징은 다음 표 9와 같이 모델링 한다. 샷이 갖는 시간간격에 대한 정보와 동기화를 위한 초당 프레임 속도 정보를 갖는다.

표 10. 시간 특징에 대한 메타데이터

속성	내용
tfid#	식별번호
Start_time	시작시간
End_time	종료시간
Frame_rate	초당 프레임 수

### 7) Spatio-Temporal relationship

시-공간 관계정보는 샷에 포함된 객체가 시간의 흐름에 따라 변화되는 물리적 위치 정보로서 객체들 간의 상대적 위상, 방향 등의 공간관계를 추론

할 수 있는 정보를 갖는다.

표 11. 공간 특징에 대한 메타데이터

속성	내용
stid#	식별번호
oid#	객체 식별번호
Object_position	객체 위치
Frame_duration	프레임 간격

### 8) 하이라이트 정보

하이라이트는 비디오 계층 구조와는 독립적이며 비디오의 주요 장면들에 대한 클러스터 집합으로 구축된다. 표 11에서는 하이라이트들의 속성정보를 갖는 메타데이터의 속성요소와 각 속성이 갖는 내용에 대해서 기술하였다.

표 12. 하이라이트에 대한 메타데이터

속성	내용
hid#	하이라이트 식별번호
vid#	비디오 번호
Subject	하이라이트 주제
Actor	해당 하이라이트에서의 배우, 주인공
Frame_duration	하이라이트의 길이(프레임 간격)

## 3.3 메타데이터의 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 EER(Enhanced ER)모델로 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 오버로딩기능으로 데이터 관리에 확장성을 제공한다.  
넷째, 모듈화가 가능하므로 모델에 확장성과 융통성을 가진다.

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

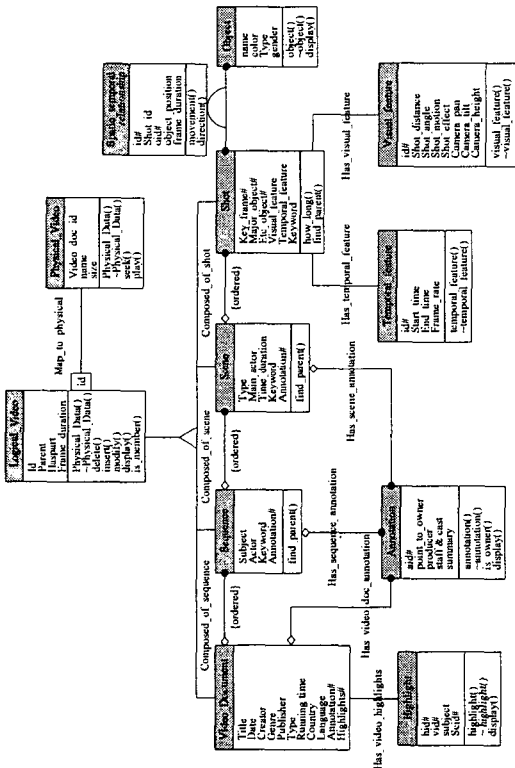


그림 161. 메타데이터 객체 다이어그램

이논문에서 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 클래스

를 이용하여 시공간 질의가 가능하며 세분화 정도에 따라 보다 상세한 비디오 질의가 가능하게 하였다. 다음 그림 2는 제시한 메타데이터 모델을 기반으로 영화 “물란”에서 주인공 “물란”과 “리상”이 군사 훈련도중 콩푸 겨루기를 하는 장면에 대해 메타데이터를 구축한 예이다.

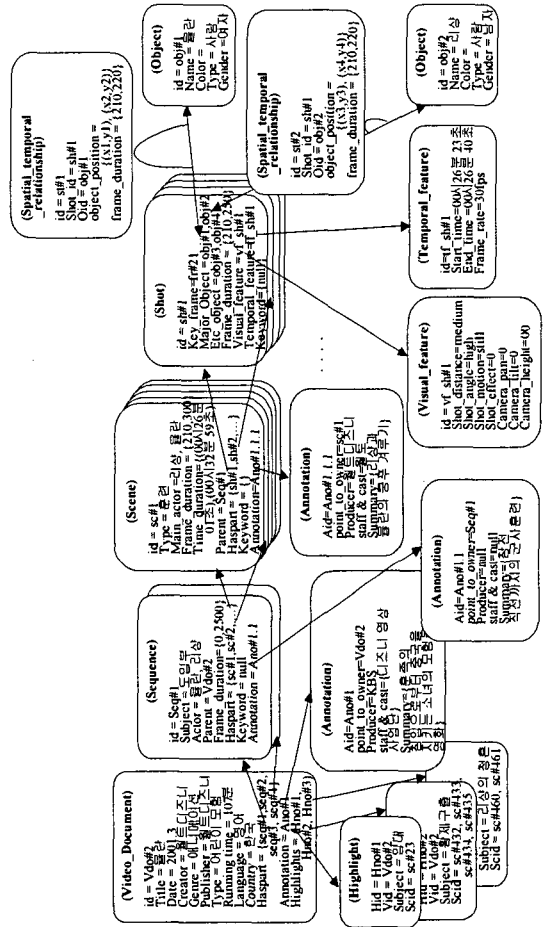


그림 162. 영화 “물란”에 대해 구축된 메타데이터의 예

#### 4. 결론

네트워크 기술과 멀티미디어 기술의 발전으로 인터넷을 통한 비디오의 상영이 가능해지게 되었다. VOD 시스템은 컴퓨터 네트워크를 이용하여

데이터베이스에 멀티미디어 기술을 통합시킨 대표적인 멀티미디어 응용분야로서 비디오 데이터베이스를 기반으로 사용자의 요구에 대해 그들이 원하는 장소에서 원하는 시간에 원하는 영화를 웹을 통하여 제공해주는 시스템이다. 비디오 데이터는 텍스트 데이터에 비해 정보에 대한 해석이 애매 모호할 뿐만 아니라 시공간의 다차원으로 인해 복잡한 내용 관계의 특성을 가지고 있다. 이 논문에서는 VOD 시스템에서도 비디오 데이터베이스 분야에서와 같이 비디오 데이터가 갖는 시공간 특성을 고려한 내용기반 검색을 지원하기 위해 비디오 데이터의 시공간 특성을 표현할 수 있는 메타데이터 모델을 내용기반 검색을 필요로 하는 다양한 응용분야에 적용이 가능하도록 특정 응용분야에 국한되지 않으면서 필수적인 특징뿐만 아니라 시각적, 시-공간적 특징을 포함하도록 메타데이터의 속성을 분류하였다.

또한 객체 지향 기법을 사용하여 메타데이터의 특성들을 추상화하여 클래스로 정의하고 각 클래스마다 메소드를 정의하여 응용분야에 의존적인 부분에 대한 확장이 용이하고 데이터 관리에 융통성을 갖을 수 있는 메타데이터를 기존 모델을 기반으로 확장하였다.

셋째, 메타데이터 모델에서 샷을 기반으로 하는 논리적 계층구조를 갖게하여 비디오 구성요소인 샷, 씬, 시퀀스, 비디오 도큐먼트들은 자신의 단계에서 필수 속성만 가지고도 하위 단계 컴포넌트의 속성을 모두 공유할 수 있으며, 검색에서 샷-기반(shot-based), 씬-기반(scene-based), 시퀀스-기반(sequence-based), 비디오 도큐먼트-기반(video document-based) 등으로 다단계 구조적 접근이 가능하게 하였다.

## 참고문헌

- (가) [1] Stephen W. Smoliar, HogJiang Zhang, "Content-Based Video Indexing and Retrieval", IEEE Multimedia, pp.66-72, Summer 1994
- (나) [2] William I. Grosky, Ramesh J., Rajiv Mehrotra, "The Handbook of Multimedia Information Management", Prentice Hall PTR, 1997
- [3] S.W. Smoliar and H.J. Zhang, "Content-Based Video Indexing and Retrieval", IEEE Multimedia, 1(2), pp. 62-72, 1994.
- [4] R. Jain and A.Hampapur, "Metadata in Video Databases", No. 4, ACM SIGMOD RECORD, pp. 27-33, December 1994,
- [5] T.D.C.Little, G.Ahanger, R.J.Folz, et al., "A Digital On-Demand Video Service Supporting Content-Based Queries", In Proceedings of ACM Multimedia'93 Anaheim, U.S.A, August 1993.
- (다) [6] 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
- (라) [7] 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
- (마) [8] 진성일, "동영상 메타정보 구조화 및 검색방법에 관한 연구", 한국전자통신연구소 부설 시스템공학연구소 최종연구보고서, 1997