

# 내용 기반 검색 및 주석 기반 검색을 통합하는 비디오 데이터 모델의 설계 및 구현

## (Design and Implementation of a Video Data Model Integrating Content-Based Retrieval and Annotation-Based Retrieval)

김기병<sup>†</sup>   김형주<sup>\*\*</sup>  
(Ki-Byoung Kim) (Hyoung-Joo Kim)

**요약** 본 논문에서는 내용 기반 검색과 주석 기반 검색을 일관성있게 통합하는 비디오 데이터 모델(MuVi)을 제안하고 비디오 검색에 필수적인 연산들을 제시하였다. MuVi 모델은 비디오 스트림 계층, 특징 계층, 특징 시퀀스 계층, 부 개념 계층 및 개념 계층으로 이루어져 있다. 특히 MuVi 모델은 부개념 계층(semi-conceptual layer)을 이용하여 무의미 데이터(non-semantic data)와 의미 데이터(semantic data)의 관계를 표현할 수 있으며, 내용 기반 검색과 주석 기반 검색을 일관성있게 지원할 수 있다.

MuVi 모델의 계층화된 특성은 물리적 비디오 데이터와 특징 데이터, 특징 데이터와 개념적 데이터의 분리 및 이에 따른 질의 처리의 분리 계층을 제공하므로 이러한 데이터의 독립성에 의해 각 계층 간의 종속성이 최소화된다. 이러한 독립성을 이용하여 내용 기반 검색과 주석 기반 검색을 지원하는 비디오 검색 시스템을 쉽게 모듈화할 수 있으며, 다른 계층에 무관하게 개선되거나 수정될 수 있다.

**Abstract** This paper presents Multi-Layered Video Model(MuVi) which integrates content-based retrieval and annotation-based retrieval. We also identify essential operations to implement a video retrieval system on MuVi model. MuVi consists of 5 layers such as video stream layer, feature layer, feature sequence layer, semi-conceptual layer, and conceptual layer. Especially, MuVi can represent the relationship between non-semantic data and semantic data using semi-conceptual layer, thus, content-based retrieval and annotation-based retrieval are supported within a unique framework.

Using these 5 layers, MuVi provides two kinds of separation such as the separation of physical video data and feature data, the separation of feature data and conceptual data. By these separation, dependencies between each layers are minimized. Thus, we can implement each modules of video retrieval system without the effect to other layers.

### 1. 서론

멀티미디어 기술의 급속한 발전으로 방대한 비디오 자료를 접하기가 어렵지 않다. 이에 따라, 비디오 자료의 검색에 대한 필요성은 방송이나 교육용 비디오를 매체로 이용하여 정보를 전달하는 부분으로부터 점차로 증대되

고 있다. 그러나, 비디오를 기존의 텍스트 데이터처럼 검색하기는 매우 어렵다. 이에 따라 비디오가 담고 있는 내용을 표현하고 검색하기 위한 방법으로 두가지 접근법이 주로 사용되어져왔다. 내용 기반 검색[8, 9]과 주석 기반 검색[8, 9, 11, 14]이 이러한 방법들의 예이다.

내용 기반 검색은 비디오가 가지고 있는 의미를 이미 지 분석 기법 등을 이용하여 추출한 후 이를 특징 벡터로 표현하고, 이를 검색에 이용하는 방법이다. 예를 들면, 대상 객체의 궤적(trajjectory)을 추적하여 대상물의 움직임을 알아내거나[4], 대상 객체의 특징을 추출하여 대상 객체의 동작을 모델링하는 등의 기법이 제안되어 있으며[10] 이러한 방법들은 특정 정의역에 대해서는 좋

· 본 연구는 정보통신부의 초고속통신망 응용 개발 과제 No. 96-194, "초고속통신망에서의 비디오 교육질의시스템의 개발" 및 과학재단 목적기초 No 95-2022-A, "다중매체 요구형 시스템기술에 관한 연구"의 부분적인 지원에 의한 것임.

† 종신회원 : 서울대학교 컴퓨터공학과  
 \*\* 종신회원 : 서울대학교 컴퓨터공학과 교수  
 논문접수 : 1996년 3월 19일  
 심사완료 : 1996년 12월 18일

은 검색 효과와 성능을 낸다. 그러나 비디오가 담고 있는 일반적인 의미를 추출하는 것이 매우 어려우므로 일반적인 비디오 검색 시스템으로 사용하기에는 문제가 있다.

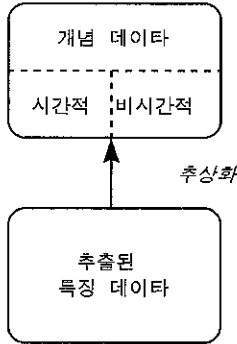


그림 1 기존 추상화 계층

주석 기반 검색은 비디오가 담고 있는 의미를 사람이 먼저 파악한 후 이를 자연어를 이용하여 표현하고, 검색에는 미리 부여된 주석을 이용하는 방법이다. 주석 기반 검색 방법을 이용하면, 사람이 인식할 수 있는 모든 내용을 주석으로 처리할 수 있기 때문에, 일반적인 비디오 검색 시스템으로 확장이 가능하다. 이는 자동화된 방법으로 알아내기 힘든 비디오 데이터의 여러가지 의미를 손쉽게 모델링할 수 있고 검색에 응용할 수 있다는 장점이 있다. 그러나, 사용자의 관점에 따라 주석을 다르게 부여하거나 다르게 해석할 수 있으므로, 비디오의 의미의 일관성을 잃기 쉬우며, 동일한 의미에 대해서 주석에 사용되는 언어의 표현의 다양성으로 인하여 일관성 있는 주석시스템을 만들기가 어렵다. 또한, 비디오가 담고 있는 의미를 사용자가 일일이 부여하여야 한다는 문제점들이 있다.

이러한 두가지 검색 방법은 비디오 검색 시스템을 구현함에 있어서 서로 보완적인 역할을 한다. 비디오 검색 시스템의 궁극적인 목적은 사용자가 원하는 조건을 만족하는 비디오를 찾아내는 데 있다. 비디오를 검색하는 여러가지 방법은 이러한 목적을 위해 적용되어지는 것이다. 그러므로 내용 기반 검색 기법을 적용 가능한 부분에 대해서는 이를 적용하고, 특징을 추출하기 어려운 의미에 대해서는 사용자가 직접 주석을 부여하여 검색할 수 있게 하는 것처럼, 보다 다양한 검색 기법을 동시에 이용할 수 있는 비디오 검색 시스템을 구현한다면, 그 검색 성능을 향상할 수 있다.

그러나 내용 기반 검색에 사용되는 특징 데이터와 주석 기반 검색에 사용되는 개념 데이터 사이에는 의미상의 차이가 존재하므로, 특징 데이터를 개념 데이터에 직접 대응시키고 검색에 사용하는 것은 의미상의 차이에 따른 문제점을 내포한다. 예를 들어, 특징 데이터는 비디오 데이터로부터 직접 추출한 데이터이므로 인스턴스의 속성이 강한 반면 개념 데이터는 추상화되어 있으므로 타입의 속성이 강하다. 그러므로, 우리는 특징 데이터와 개념 데이터 간의 관계를 기술하기 위하여 부개념 계층을 도입하고 이 계층에 추출된 특징 데이터를 그룹화하고, 사용자가 부여한 개념 데이터와 instance-of 관계를 부여하여 유지하였다. 이를 통하여 무개념 데이터와 개념 데이터 간의 의미상의 차이를 줄이면서 연관 관계를 유지하게 된다. 부개념 계층을 이용하여 특징 추출을 위한 비디오 데이터 분석 계층과 주석 계층에 대해서 독립적인 데이터의 조작성 허용함과 동시에, 질의 처리 시에는 다양한 비디오 질의 처리 기법을 적용할 수 있게 된다.

본 논문에서는 이를 위한 다중 계층 비디오 모델 (Multi-Layered Video Model, MuVi)을 제안하고 그 구현 예를 보였다. MuVi 모델은 내용 기반 비디오 검색과 주석 기반 비디오 검색을 지원하는 비디오 검색 시스템의 기본 모델로 사용될 수 있을 뿐 아니라 이 두가지를 통합하여 검색에 사용하는 비디오 검색 시스템의 기본 모델로 사용되어질 수 있다.

데이터베이스의 계층화된 설계의 목적은 데이터의 직접적인 조작으로부터 사용자에게 추상화된 관점을 제공하여 데이터의 독립성을 획득하고, 이를 통하여 물리적 데이터의 성질에 무관하게 데이터를 일관성있게 조작할 수 있다[12]. 본 논문에서는 이러한 계층 구조를 비디오 모델에 적용하였다.

본 모델은 이미 녹화된 비디오 데이터의 집합을 검색 대상으로 가정하였다. 그러므로 이러한 비디오 데이터들은 처음 녹화된 형태로만 존재한다는 특징과 여기에 포함된 의미의 형태는 각 프레임이 가지는 의미와 프레임들이 나열된 시퀀스에서의 의미로 분류할 수 있다. 여기에서는 이러한 특징을 표현하고 검색하고자 하였으며, 데이터를 합하거나, 편집하여 새로운 비디오 장면의 합성하거나, 이러한 비디오들 간의 동기화에 대해서는 본 논문의 범위를 넘으므로 크게 다루지 않았으나, 본 모델의 계층화된 성질을 이용하면 이를 지원하도록 쉽게 확장할 수 있다.

본 논문의 구성은 다음과 같다. 2절에서는 비디오 검색과 관련한 기존의 모델 및 검색 기법에 대해서 설명하

고, 확장된 모델의 필요성에 대해 기술하였다. 3절에서는 내용 기반 검색과 주석 기반 검색을 통합 지원하기 위한 비디오 데이터 모델인 다중 계층 비디오 모델(MuVi)을 소개한다. 4절에서는 MuVi 모델에 근간한 비디오 질의의 종류를 보인다. 5절에서는 MuVi 모델에서 질의 처리 과정을 보이고 MuVi 모델에 근거하여 비디오 검색시스템을 구현한 예를 보인다. 6절에서는 본 논문에서 제시한 MuVi 모델의 특징을 요약하고, 앞으로의 연구 방향에 대해 이야기한다.

## 2. 관련 연구

최근 10여년 동안, 비디오 검색에 관한 연구들이 Video-On-Demand, 비디오 데이터베이스 등의 분야에서 많이 행해져 왔다. 특히 비디오 검색에 관한 연구는 대용량의 비디오 데이터가 점차로 널리 보편화되고, 초고속 통신망이 확산되면서 많은 관심의 대상이 되고 있다. 비디오 검색에 관한 연구는 비디오가 가지는 의미를 추출하고 표현하는 방법에 관한 연구와, 검색의 방법으로 내용 기반 검색과 주석 기반 검색에 관한 연구들이 있다.

Oomoto와 Tanaka는 [14]에서 비디오 검색을 위한 객체지향 모델 OVID를 제안하였다. 그들은 의미있는 개념과 이러한 개념을 합성하는 방법을 제안하였으나, 시간적 속성과 개념적 속성을 분리하지 못함으로써 개념적인 데이터를 재사용하거나 공유하기에는 문제점을 가지고 있다.

Hjelsvold와 Midtstraum는 [8, 9]에서 비디오 데이터 모델을 제안하고 질의어를 확장하였다. 여기서는 성층화(stratification)방법을 이용하여 원 비디오 데이터를 전체 비디오 문서와 개별 프레임의 두단계로 추상화하고, 비디오의 물리적 구조를 따라 비디오를 검색할 수 있는 다섯가지의 비디오 질의를 제안하였다.

Day, Dagtas 및 Iino는 [3]에서 비디오 데이터를 위한 객체지향 개념 모델을 제안하였다. 여기서는 방향 그래프(directed graph) 및 개념 시간 객체(conceptual temporal object), 개념 공간 객체(conceptual spatial object), 물리 객체(physical object)를 이용하여, 원 비디오 데이터에 포함되어 있는 공간적 속성과 시간적 속성을 분리하여 추상화하여 물리적 독립성과 개념적 독립성을 가지는 모델을 제시하였다.

Chua와 Ruan은 [2]에서 shot 계층과 scene 계층을 가지는 2-계층 개념 모델을 제안하였다. Shot 계층은 원 비디오 데이터로부터 기본적인 개념을 표현하기 위해 몇개의 일련의 비디오 프레임을 모은 것이며, scene

계층에서는 보다 포괄적인 의미나 주석을 위한 계층으로 제공된다. 이 모델에서는 두 계층 모두 의미를 표현하기 위한 계층이며 의미의 크기에 따른 계층으로서, 하나의 개념적 계층을 분할한 형태이다.

여기서 제시된 비디오 모델들은 비디오에 포함된 인스턴스를 개념에 직접 대응시켜서 표현하고 있다. 그러므로 개념 데이터에 의해 표현되는 타입의 성질과 인스턴스의 불일치를 거론하고 있지 못하다.

Weiss와 Duda는 [17]에서 비디오 대수를 제안하였다. 여기서는 주로 원 비디오 데이터에 대해 생성, 합성 및 출력, 설명 기술등의 연산을 제공하며, 중첩 논리 구조(nested logical structure)를 이용하여 비디오 데이터를 추상화한다. 이 모델은 그러나 비디오 표현(presentation)에는 적합한 연산을 제공하나, 검색을 위한 효율적인 개념적 구조의 관리에 대해서는 제시하고 있지 않다.

Dimitrova와 Golshani는 [4]에서 객체의 궤적을 추적하여 궤적에 의한 내용 기반 검색을 지원하는 Rx 시스템을 제안하였다. 여기서는 비디오 데이터 모델로서 의미 다중 해상도 계층 구조(semantic multiresolution hierarchy)를 제안하였다. 이를 이용하면, 원 비디오 데이터로부터 대상물의 궤적의 추적등으로 얻어지는 의미를 분리할 수 있다. 그러나 여기서는 객체의 움직임에 대한 정보만을 표현할 수 있으며, 객체나 동작에 대한 설명을 사용하고는 있으나 개념과 특징 데이터의 차이에 대해서 고려하고 있지는 않다.

Smoliar와 Zhang은 [16]에서 내용 기반 비디오 인덱싱을 제안하였다. 여기에서는 비디오 데이터를 검색하기 위해 내용을 분석하는 과정과, 분석된 내용을 효율적으로 표현하기 위한 내용 모델이 중요함을 이야기하였다. 그러나, 원 비디오로부터 추출된 의미적 데이터를 원데이터와 같이 효율적으로 관리하거나 검색하기 위한 구조의 제시는 미비하다.

이외에도 IBM Almaden 연구소에서는 [5]에서 내용 기반 검색 기법을 이용하여 이미지나 비디오를 검색하는 QBIC 시스템을 구현하였다. 여기에서는 내용 기반 검색을 위해 컴퓨터가 인식하기 좋은 형태로 질의를 기술하므로, 사용자가 의도한 내용과 다른 검색 결과가 얻어질 수도 있다.

Gibbs 등은 [6]에서 시간에 종속적인 데이터들을 시간적 스트림(timed stream)을 기본적인 추상화 단위로 하여 모델링하였다. 여기서는 시간적 스트림만을 이용하여, 비디오가 가지는 개념적 데이터를 모델링하게 되므로, 비디오 데이터에 포함된 개념적 데이터들이 시간적 스트림에 종속되어, 스트림과 개념의 분리가 명확히

이루어지지 않게 된다. 이 모델에서는 스트림을 다시 개념적인 단위로 보기 위해 번역(interpretation) 및 유도(derivation), 합성(composition) 계층을 제공하여, 스트림에 대한 의미 부여와 스트림의 편집이나 통합을 지원한다.

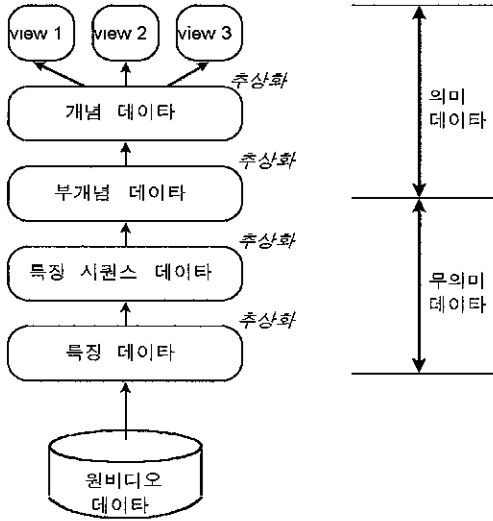


그림 2 MuVi 모델에서의 비디오 계층 구조

기존의 많은 비디오 검색 시스템이나 모델들이 계층화된 비디오 모델의 필요성과 유용성에 대해 동의하고 있다[2, 3, 4, 8, 9, 15, 17]. 그러나 기존의 비디오 검색 시스템에서 제안된 모델은 그 추상화 단계가 그림 1과 같이 물리적 데이터와 개념적 데이터만으로 분리하여져 있는 경우가 대부분이다.

그러나 내용 기반 검색을 지원하는 비디오 검색 시스템에서는 비디오 데이터의 검색에 사용되는 데이터는 이미지 분석 및 주석의 부여 등과 같이 다단계의 처리를 거치게 되며, 이에 따라 물리적인 데이터로부터 무의미 데이터(non-semantic data), 의미 데이터(semantic data)로 근접해가게 된다.

이러한 단계를 표현하고 검색에 활용하는 것은 비디오 데이터내에 포함된 의미를 명확하게 나타내고 검색을 효율적으로 행할 수 있는 기반이 된다. 이를 위해 본 논문에서 제안한 모델에서는 물리적 데이터로부터 의미 데이터로 추상화하는 과정에서 물리적 데이터나 의미 데이터로 보기 어려운 중간 데이터를 그림 2와 같이 별도의 독립적인 계층으로 모델링하고 검색에 활용할 수 있게 하였다.

본 논문에서 제안한 비디오 모델은 의미 데이터와 무의미 데이터간의 관계를 질의의 작성이나 검색 시에도 추출된 특징 벡터에 의한 검색과 더불어 개념에 의하여서도 내용 기반 검색을 수행하거나 통합된 질의 처리를 유일한 프레임 워크 안에서 할 수 있다는 장점이 있다

### 3. MuVi: 다중 계층 비디오 모델

3절에서는 대용량의 비디오 저장소로부터 효율적인 비디오 검색을 지원하기 위한 다중 계층 비디오 모델을 보다 형식적으로 정의하였다. 이 모델을 통해 그림 3에서 보여지는 바와 같이 우리는 물리적 비디오 데이터와 중간 단계로서의 추출된 특징 데이터, 주석에 의한 개념 데이터와 질의에 의해 제시되는 개념데이터들이 서로 독립성을 유지하면서, 각 계층별로는 고유의 연산을 지원하고, 계층 간에는 추상화 연산을 지원하여 각 계층에 대한 질의 연산이나 이들이 결합된 질의 연산에 대해 일관성 있는 검색을 지원할 수 있도록 하였다.

MuVi 모델은 각 계층 별로 독립성을 유지하므로 비디오 질의 처리를 위해 내용 기반 검색이나 사용자 주석 기반 검색의 특정 방법에 종속되지 않는 질의 처리를 구현할 수 있게 된다. 각 계층 별로 다른 단계에 영향을 받지 않고 구현되거나, 수정, 개선될 수 있다.

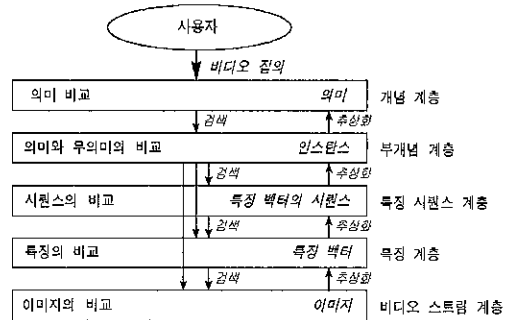


그림 3 MuVi 모델의 각 계층 및 역할

#### 3.1 계층 1: 비디오 스트림 계층

비디오 스트림은 이미지의 시퀀스로서 정의되어진다. 비디오 스트림 계층은 원 비디오 데이터를 이미지의 집합으로 표현한다. 이를 이용하면 비디오가 포함하고 있는 이미지들을 모델링하고, 원하는 프레임의 이미지를 얻을 수 있으며 이미지 데이터베이스의 원 데이터로서 제공되어질 수 있다. 비디오 스트림 계층은 원비디오 데이터가 가지고 있는 속성이나 비디오 데이터의 포맷에

무관하게 데이터에 접근할 수 있도록 물리적인 독립성을 제공하는 역할을 한다.

비디오 스트림은 이미지의 시퀀스로서 표현되는데, 이는 다음과 같이 정의한다.

정의 3.1 이미지의 시퀀스  $I_S$ 는  $\langle v_i \rangle$ 로 정의된다. 단,  $v_i$ 는 하나의 이미지를 나타내는 특징 벡터이며,  $v_i$ 는  $v_1, v_2, v_3, \dots, v_n$ 의 시퀀스를 나타낸다. 여기서 시퀀스를 표현하기 위하여  $v_i$ 와 같은 기호를 이용하였다.

정의 3.2 비디오 스트림  $Q$ 는 연속된 이미지의 시퀀스로서 순서쌍  $\langle vid, \langle v_i \rangle, f \rangle$ 로 정의된다. 단,

- $vid$ : 비디오 조각의 식별자로서 하나의 비디오 세그먼트를 나타낸다.
- $\langle v_i \rangle$ : 비디오 이미지의 시퀀스.
- $f$ : 샘플링 속도. 초당  $f$  개의 이미지가 재생되어야 함을 나타낸다.

함수	설명
$search(R) \rightarrow E$	대상 비디오에 대해 구간 $R$ 을 포함하는 비디오 구간을 검색

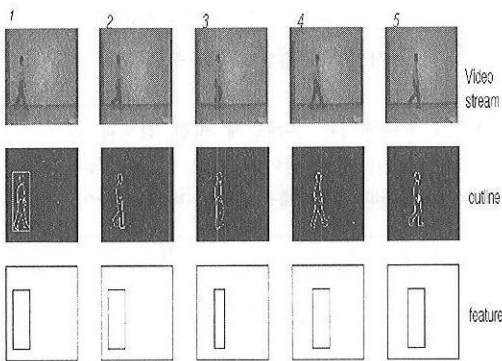


그림 4 걷는 동작에 대한 특징 시퀀스의 예

계층 1에는 비디오 스트림 상에서 임의의 프레임에 접근하기 위한 연산이 정의된다. 이를 이용하여, 비디오 데이터로부터 특징 추출을 위해 임의의 프레임을 얻는데 사용한다. 다음과 같은 연산들이 계층 1에 정의된다.  $R$ 은 비디오 구간,  $E$ 는 의미에 따라 분할된 비디오 조각들의 집합을 나타낸다. 주어진 비디오 스트림은 비디오 이미지의 급격한 변화나 추출된 특징 벡터의 급격한 변

화에 의해 찾아진 경계를 기준으로 비디오 조각으로 분할되거나, 사용자가 지정한 프레임을 기준으로 분할된다. 자동화된 방법에 의한 분할은 컬러 히스토그램의 기준치 이상의 변화나 대상 객체의 특징 벡터의 기준치 이상의 변화에 의해 결정된다. 이렇게 주어진 비디오 스트림은 그 분할 방법이나 분할 기준에 따라 여러번 분할될 수 있으며, 일부 구간이 다른 조각과 겹칠 수도 있다.

### 3.2 계층 2: 특징 계층

비디오 스트림을 구성하는 프레임들은 각각으로 보면 이미지로 구성된다. 특징 계층은 계층 1에서 선택된 임의의 이미지 프레임으로부터 이미지 분석 기법을 이용하여 특징을 추출한 후 이를 유지하기 위한 계층이다. 이미지 데이터로부터 특징 추출은 비디오의 내용을 잘 표현할 수 있는 프레임을 먼저 선정된 후 행하게 되는데, 비디오의 대상물에 관심이 있을 경우, 동일한 대상물이 존재하는 구간에 대해 하나의 프레임이 선정되며, 동작이나 시간적 데이터에 관심이 있을 경우에는 구간의 비디오 데이터에 대해, 일정한 간격으로 이미지 프레임을 추출하게 된다. 그림 4에서 Video stream 부분은 원 비디오로부터 얻어진 비디오 장면들이며, 외곽선이나 특징 부분이 특징 계층에서 관리되는 특징 데이터의 예이다.

특징 벡터로 표현되는 이미지의 특징 데이터를 검색하기 위해서는 각 이미지의 특징 벡터를 비교 메트릭으로 비교하여 검색을 수행할 수 있다. 즉 이미지로부터 추출된 특징 벡터는 무의미 데이터므로 이미지가 가지는 의미를 검색하는 것이 아니고, 추출된 특성을 기반으로 동일성 검색이나 유사성 검색 등의 기법을 적용하여 검색을 수행한다.

추출된 특성에 대한 검색을 위해 여러가지 기법이 사용된다. 예를 들면, 이미지 데이터로부터 외곽선이나, 볼륨, 모멘트 등을 얻어 이로부터 원 프레임 데이터가 가지는 무의미 데이터의 정보를 표현하는  $n$ 개의 원소를 가지는 특징 벡터를 추출할 수 있다. 특징 계층은 이렇게 추출된 특징들을 표현하기 위한 계층이다.

정의 3.3 특징 벡터  $F_V$ 는 순서쌍  $\langle vid, i, j, \langle f_i \rangle, t_i \rangle$ 로 정의된다. 단,

- $vid$ : 비디오 조각의 식별자.
- $pos$ : 비디오 스트림에서 하나의 프레임의 위치.
- $v$ :  $n$ 차원의 특징 벡터.
- $t_i$ : 이미지 분석 기법 식별자.

특징 계층에서는 특징 벡터들을 비교하고 검색하는

연산들이 제공된다.  $Re$ 은 실수의 집합,  $E$ 는 비디오 구간의 집합이며,  $F_V$ 는 특징 벡터의 집합을 나타낸다.

함수	설명
$search(v_i) \rightarrow E$	주어진 이미지를 포함하는 비디오 구간을 검색
$search(F_V) \rightarrow E$	주어진 특징 벡터를 포함하는 비디오 구간을 검색
$distance(v_i, v_j) \rightarrow Re$	두 이미지로부터 특징 벡터의 차이를 구하여 반환.
$distance(F_V, F_V) \rightarrow Re$	두개의 특징 벡터의 차이를 계산하여 반환

**3.3 계층 3: 특징 시퀀스 계층**

특징 시퀀스 계층에서는 비디오 데이터에 존재하는 시간에 종속적인 데이터를 모델링하기 위한 계층으로, 특징 계층에서 추출이 된 특징 벡터들을 시퀀스로 구성한 후, 이로부터 시간에 종속적인 무의미 데이터로 모델링한다. 특징 시퀀스 계층의 각 이미지 프레임들은 반드시 동일한 비디오 스트림으로부터 유도되어야하며 유도된 특징벡터의 시퀀스를 이용하여 비디오 스트림의 일정한 구간에 대해 포함된 정보를 표현하게 된다.

계층 3은 이미지 데이터에 대한 단순한 특징 추출만으로 표현하기 어려운 시간적 속성을 특징 시퀀스 계층의 구성 요소인 시퀀스를 이용하여 표현하게 되는데, 이를 통해 동작과 같은 시간적 요소가 표현된다. 시퀀스를 구성하는 방법은 단순히 특징 벡터를 늘어놓는 방법이나 특징 벡터의 시퀀스에 DFT(Discrete Fourier Transform)와 같은 변환을 적용하는 방법 등이 사용된다.

정의 3.4 특징 시퀀스  $F_S$ 는 순서쌍  $(vid, i, j, \langle f \rangle, t_s)$ 로 정의된다. 단,

- $vid$  : 비디오 스트림의 식별자.
- $i, j$  : 비디오 구간의 시작 및 끝 위치.
- $\langle f \rangle$  :  $f_i$ 는 특징 벡터의 시퀀스. 단,  $f_i \in F_V$ 이며 동일한  $vid, t_s$ 를 가진다.
- $t_s$  : 시퀀스 분석 방법 식별자.

특징 시퀀스 계층에서는 시퀀스의 비교 연산이 필요하다. 특징들이 추상화되어 표현되는 모든 의미는 시퀀스를 이용하여 표현되어진다.  $Bool$ 은 논리형 데이터,  $E$ 는 비디오 구간,  $DID$ 는 설명 식별자,  $Re$ 는 실수의 집합을 나타낸다.

함수	설명
$search(F_S) \rightarrow E$	특징 시퀀스를 포함하는 비디오 구간을 검색
$search(DID) \rightarrow E$	개념데이터의 식별자를 담고 있는 비디오 구간을 검색
$distance(F_S, F_S) \rightarrow Re$	주어진 두개의 특징 시퀀스의 차를 반환.
$temporal\_op: E, E \rightarrow Bool$	두개의 구간이 시간 관계 $temporal\_op$ 를 만족하는지 확인

$temporal\_op$ 는 [7, 9]등에서 거론되어진  $equal, before, meet, overlap, contain, start, finish$  연산이다.

**3.4 계층 4: 부개념(semi-conceptual) 계층**

부개념(semi-conceptual) 계층은 특징 계층이나 특징 시퀀스 계층의 데이터가 무의미 데이터로서 의미 데이터인 개념 계층의 데이터와 의미 상의 차이를 줄이고, 의미의 중복을 막기 위해 제안된 계층이다. 부개념은 특징 데이터나 특징 시퀀스 데이터 중 그 차이가 임계값보다 작을 때, 같은 그룹으로 분류한 후 이를 하나의 인스턴스화 한 것이다. 이를 통하여 추출된 특징 데이터에 의미를 부여할 때, 중복된 의미의 부여를 줄이고, 개념 계층과의 관계를  $instance-of$ 를 이용하여 표현해준다. 이와 같이 특징 계층과 개념 계층의 연관 관계를 표현해 줌으로써, 질의 시 사용자가 표현하기 어려운 특징 계층이나 특징 시퀀스 계층의 무의미(non-semantic) 데이터를 의미 데이터를 빌어서 표현할 수 있게 되며, 특징 데이터나 특징 시퀀스 데이터의 추상적인 의미를 일반화시켜 검색에서 공유할 수 있도록 하여준다.

정의 3.5 부개념 데이터(semi-conceptual data)  $C_S$ 는  $(sid, V, S)$ 로 정의된다. 단,

- $sid$  : 부개념(non-semantic) 데이터의 유사 객체를 추상화한 그룹의 식별자.
- $V$  : 특징 벡터의 집합  $F_V$ 의 부분집합.
- $S$  : 특징 시퀀스의 집합  $F_S$ 의 부분집합.

부개념 계층에서는 부개념에 의한 검색과 비교 연산을 제공한다. 이 때 검색이나 비교 연산은 특징 계층으로의 비교 연산과 개념 계층으로의 비교 연산이 있다.

함수	설명
$search(C_S) \rightarrow E$	개념스키마의 인스턴스인 부개념을 포함하는 비디오 구간을 검색
$distance(C_S, C_S) \rightarrow Re$	두개의 부개념 인스턴스의 차를 반환.
$distance(C_S, T) \rightarrow Re$	두개의 개념데이터와 부개념 데이터의 인스턴스의 차를 반환

**3.5 계층 5: 개념 계층**

개념 계층에는 비디오 데이터에 대한 의미 정보를 유지하는 계층이다. 이 계층은 의미 데이터의 클래스와 이들간의  $is\_a$  관계에 의해 구성된다. 이 계층에서는 개념 데이터의 클래스를 추가하거나 삭제하고,  $is\_a$  관계를 부여하는 연산들이 제공된다. 이 계층은 객체지향 DBMS나 객체지향 모델에서의 객체지향 스키마 구조와 유사하다. 이 계층에서 제공되는 모든 연산은 객체지향

DBMS를 이용할 경우 스키마 연산을 이용하여 쉽게 구현된다. 이 계층의 질의는 비 시간적 개념에 대한 검색이다.

정의 3.6 개념 스키마  $C_C$ 는 순서쌍  $(T, is\_a, \tau, C, instance\_of)$ 로 정의된다. 단,

- $(T, is\_a, \tau)$ :  $\tau$ 이 최상위 노드인 유향 그래프로서 타입  $T$ 의 원소들에 대한 일반화 계층 구조를 나타낸다.
- $T$ : DBMS 내에 존재하는 모든 타입의 집합의 부분 집합.
- $is\_a$ : 두 개의 타입 간의 일반화 관계를 나타낸다.
- $\tau$ : 시스템이 정의한 메타 식별자 타입으로 모든 타입의 상위 타입.
- $C$ : 부개념 데이터  $C_S$ 의 부분 집합.
- $instance\_of$  부개념 데이터와 개념 데이터의 관계.

개념 스키마 계층에서는 스키마 조작 연산과 검색 연산이 제공되어진다. 다음과 같은 연산들이 계층 5에서 제시되었다.

함수	설명
Add/drop a type	개념 스키마에 타입을 추가/삭제
Add/drop an instance	개념 스키마에 인스턴스를 추가/삭제.
Add/drop is-a relationship	타입간의 is-a 관계를 추가/삭제
Search( $T$ ) $\rightarrow E$	개념 데이터를 포함하는 비디오 구간을 검색.
distance( $T, T$ ) $\rightarrow Re$	개념 데이터간의 정량적 차이를 반환
distance( $C_S, T$ ) $\rightarrow Re$	부개념 데이터와 개념 데이터의 정량적 차이를 반환

여기서 제공하는 개념 스키마에는 개념 데이터와 이의 식별자가 포함되어 있다. 여기서의 검색 기법이 개체 지향 DBMS의 질의와 다른 점은 일단 검색된 개념 데이터에 대해 해당 개념의 인스턴스에 해당하는 개념 식별자들을 얻어내게 된다. 그리고 얻어진 식별자를 포함하는 비디오 구간들을 비디오 데이터베이스로부터 검색해낸다. 즉 개념 데이터는 데이터 베이스 내에 유일하게 저장되고 개념 데이터의 식별자는 시간의 추이에 따라 중복되어 저장될 수 있다. 시간에 따른 데이터의 중복은 개념 식별자에 대해서만 발생한다. 개념 식별자는 인스턴스로 간주되며 특징 데이터나 특징 시퀀스 데이터처럼 무개념 데이터와 같이 취급된다. 특징 데이터인 경우에는 개념적인 데이터와 인스턴스 데이터를 분리하여 처리할 수 있도록 개념 식별자와 개념을 분리하였다.

계층 4 및 5에서는 개념 검색이 계층 1, 2에서는 시

간적 속성이 제거된 비시간적 속성 만이 유지되며, 시간적 검색은 개념 식별자가 계층 3으로 전달되어 검색을 수행하게 된다.

### 3.6 추상화 연산

*MuVi* 모델에는 다섯가지 계층이 제공된다. 각 계층간의 데이터의 추상화를 위해 다음과 같은 추상화 연산이 지원된다.

함수	설명
Abstract $L_1 \rightarrow L_2$ $I \rightarrow F_V$	프레임을 분석하여 특징 벡터를 추출.
Abstract $L_2 \rightarrow L_3$ $F_V \rightarrow F_S$	특징 벡터의 집합을 분석하여 시퀀스를 추출
Abstract $L_3 \rightarrow L_4$ $F_S \rightarrow C_S$	특징 벡터의 시퀀스를 의미 데이터의 인스턴스로 받음
Abstract $L_2 \rightarrow L_4$ $F_V \rightarrow C_S$	특징 벡터를 대해 의미데이터의 인스턴스로 받음.
Abstract $L_1 \rightarrow L_4$ $I \rightarrow C_S$	이미지 프레임을 의미데이터의 인스턴스로 받음.
Abstract $L_4 \rightarrow L_5$ $C_S \rightarrow T$	하나의 인스턴스를 의미데이터의 인스턴스로 받음.

여기서 제시된 추상화 연산은 각 계층 연산의 구현과 서로 독립적이다. 즉 각 계층의 연산은 함수로 제공된다. 이에 따라, 각 함수는 이의 인자에만 영향을 받게 되는데, 함수의 내부를 서로 독립적으로 구현하게 되면, 구현상의 종속성은 없다.

Abstract  $L_1 \rightarrow L_2$ 는 디지털 이미지 프로세싱 기법을 이용하여 주어진 이미지로부터 특징 벡터를 추출하는 함수로 구현된다. Abstract  $L_2 \rightarrow L_3$ 는 DFT등과 같은 시퀀스 프로세싱 기법을 통하여 구현된다. Abstract  $L_3 \rightarrow L_4$ 는 거리 매트릭을 이용하여 유사한 특징 데이터 들끼리 그룹화하여 구현된다. Abstract  $L_4 \rightarrow L_5$ 는 무의미 데이터를 의미데이터화하는 과정이다. 그룹화된 데이터의 의미 부여는 사용자의 판단에 의해 부여된다. 컴퓨터를 이용한 방법은 현재의 기술로서는 응용에 부적절함이 이미 알려져 있다[5]. 그러므로 여기서는 이를 사용자가 검색을 위한 개념을 부여하는 단계로 가정하였다. 무개념 데이터에 직접 개념 데이터를 부여하는 방법에 비해 무개념 데이터를 부개념 데이터(semi-conceptual data)로 추상화한 후 개념 데이터와의 관계를 부여하는 방법은 특징 데이터와 개념 데이터와의 관계를 수치적으로 줄일 수 있을 뿐 아니라 매번 인스턴스가 등장할 때마다 반복되는 개념 데이터의 중복을 현저하게 줄일 수 있다.

부개념 데이터와 개념 데이터 간의 관계는  $n$ 개의 부개념 데이터와  $m$ 개의 개념 데이터 간의  $n \cdot m$  관계를 유지할 수 있는 상관 행렬(correlation matrix)을 이용하고 여기에 사용자가 상관도를 부여할 수 있도록 하였다. 기존의 주석 방법은 여기에서는 상관 관계가 1 : 1 이면

서 상관도가 1인 것에 해당한다. 검색 시 사용자가 주석을 하나의 비디오 인스턴스에 대해 직접 부여하는 것에 비해, 주석 부여와 검색의 효율을 높일 수 있다. 그림 5는 무개념 데이터와 개념 데이터의 상관 관계와 계층 구조와의 관계를 보여주고 있다.

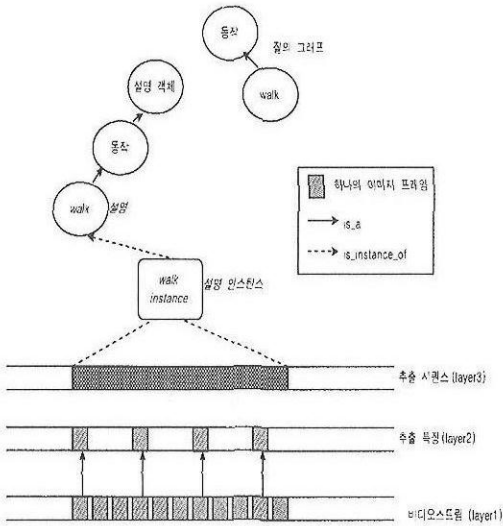


그림 5 다중 계층 비디오 모델에서의 비디오 표현 및 동작 인스턴스

4. 비디오 검색

4.1 질의 종류

Hjelsvold와 Midtstraum은 [8, 9]에서 비디오의 물리적인 구조에 근거하여 비디오 질의의 종류를 제시하였으며, Dimitrova와 Golshani는 [4]에서 비디오 데이터에 포함된 대상물, 동작, 비디오 조각의 세가지 요소 중 주어진 요소를 이용하여 나머지 요소를 찾는 OMV(Object, Motion, Video) 모델을 제시하였다. Sibel 등은 [1]에서 비디오 데이터를 대상물, 동작에 대한 추상화된 개념 및 인스턴스 개념으로 분류한 후 질의를 제시하였다. 본 논문에서는 MuVi 모델의 계층 구조에 근거하여, 각 계층 별로 제시된 연산에 근거한 질의와 개념 계층에서 지원되는 질의 및 이러한 것들이 복합적으로 나타나는 질의로 구별하여 분류하였다. 이러한 분류는 계층 구조에 의한 데이터의 독립성을 활용한 분류이다. 제안된 질의의 종류는 다음과 같다.

•Type-1 질의: 프레임 기반 질의

- Type-2 질의: 무의미 특징 질의  
추출된 특징에 대해 이미지의 내용 기반 검색
- Type-3 질의: 무의미 시퀀스 질의  
추출된 특징 시퀀스나 시간적 속성에 대한 검색
- Type-4 질의: 개념 질의  
사용자가 부여한 주석이나, 내용 기반 검색에서 추상화된 개념 단위의 검색
- Type-5 질의: 복합 질의  
위의 모든 질의 또는 둘이상의 질의가 복합적으로 표현되어 있는 검색

프레임 기반 질의는 임의의 프레임을 포함하는 비디오 조각들을 검색하는 것으로서 단순히 비디오의 위치로 비디오를 검색하기 위한 방법이다.

무의미 특징 질의는 검색의 조건으로서 하나의 프레임이 가질 수 있는 대상물을 기술하는 질의이다. 프레임이 가질 수 있는 조건의 예로는 사람의 얼굴, 장소, 등장하는 대상물의 지정 등이 될 수 있으며 이는 이미지 검색 시스템에서의 공간 질의(spatial query)에 해당한다. 프레임 기반 질의나 무의미 특징 질의는 OMV 모델에서의 객체 대상 질의나 Sibel 등이 제시한 비디오 정보 시스템의 객체 질의가 내용 기반 검색으로 확장된 형태가 된다.

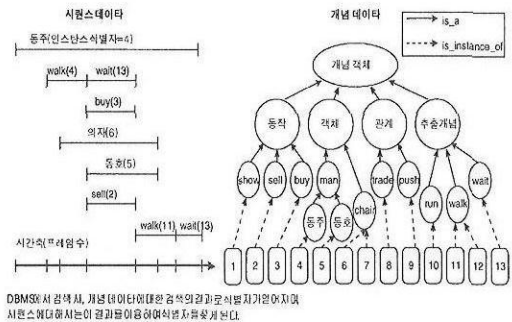


그림 6 특징 시퀀스, 개념 스키마 및 개념 식별자

무의미 시퀀스 질의는 OMV 모델에서의 동작 질의에 해당된다. 무의미 시퀀스 질의는 무의미 특징 질의가 공간 질의인 것에 비해 시간 질의(temporal query)에 해당한다. 질의에 표현된 검색의 조건이 시간의 추이에 따라 연결된 특징들의 시퀀스이거나 시간에 종속적인 개념 등이 된다.

개념 데이터에 대한 질의는 주어진 개념 데이터에 민



저 추상화된 개념 데이터에 대한 검색을 행한 후 검색의 결과로 선택된 개념 데이터의 인스턴스가 존재하는 비디오 조각을 찾아내는 검색이다. 이러한 검색을 위한 자료 구조가 그림 6에 나와 있다.

복합 질의는 주석에 의한 내용 기반 검색과 특징에 의한 주석 기반 검색을 동시에 포함한다. 질의에 포함된 조건의 내용이 특징이나 특징 시퀀스 등과 연관을 가지고 있다면 이를 계층 4, 5간에 설정되어 있는 상관 관계를 통하여 상관도의 범위 내에서 검색에 참여시킬 수 있게 된다. 검색 시 내용 기반 검색이 이용될 경우, 내용 검색은 그 검색의 결과가 사용자가 의도한 것과 다를 수 있으므로 이를 상관도에 대한 임계값을 설정하여 오류의 가능성에 대해 대비한다.

4.2 비디오 질의 처리

본 절에서는 비디오 스트림으로부터 생성한 비디오 조각의 집합에 대해 비디오 검색이 수행되는 과정을 보았다. 주어진 비디오 질의는 먼저 단순 조건으로 분할되어 각각의 조건에 대해 type 1, 2, 3, 4의 질의가 수행된다.

특정 데이터나 특정 시퀀스를 질의에 사용하기 위해서는 비디오 데이터가 미리 분석되어야하는데, 비디오나 프레임 이미지의 제시된 추상화 함수들에 의해 행해진다. 비디오 및 이미지 분석과정은 그 수행시간이 길기 때문에 off-line으로 미리 수행하여 검색에 적합한 형태로 저장된다. 데이터베이스 내에서 가공된 특징 데이터 특징 시퀀스데이터 및 개념데이터는 그림 7과 같이 저장되고 관리된다.

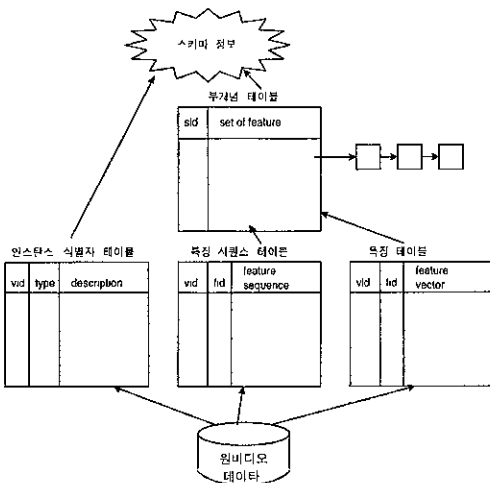


그림 7 MuVi 비디오 검색 시스템의 자료 구조

사용자가 제시한 질의는 다음과 같은 단계를 거쳐서 처리된다. 여기서는 질의 처리 과정을 보다 구체적으로 제시하였다.

알고리즘 1 SimpleVideoSearch(condition)

```

1 result = ϕ /* result를 초기화한다. */
2 DID = ϕ /* 검색의 결과인 식별자의 집합을 초기화한다. */
3 if condition has a range of a video stream
4     result = result ∪ search(R) /* 검색의 조건이 비디오의 범위로 주어질 경우 */
5 else if condition has a feature
6     result = result ∪ search(Fv) /* 검색의 조건이 특징 벡터로 주어질 경우 */
7 else if condition has a feature sequence
8     result = result ∪ search(Fs) /* 검색의 조건이 특징 시퀀스로 주어질 경우 */
9 else
10 /* 검색의 조건이 의미 데이터로 주어질 경우 */
11     DID = DID ∪ search(T)
12 end if
13 /* 의미데이터의 검색 및 그 결과인 식별자의 집합을 이용한 특정 데이터, */
14 /* 특정 시퀀스 데이터의 검색을 수행하고 */
15 /* 주어진 식별자와 관계있는 모든 인스턴스를 찾는다.*/
16 for j in DID
17     result = result ∪ search(Cs)
18 end for j
19 return result
    
```

이제 MuVi에서 제공하는 질의의 타입에 따른 질의의 예와 그 처리 과정을 설명하였다. 질의에서 사용한 개념 스키마와 특징 시퀀스는 그림 6에 나와있다.

•Type-3 질의: Retrieve video segment where action="178"<sup>1)</sup>

1. 특징 시퀀스 "178"과 관련 있는 특징 시퀀스를 검색한다<sup>2)</sup>.
2. 단계 1에서 추출된 특징 시퀀스가 포함된 비디오 구간을 검색한다.

•Type-4 질의: Retrieve video segment where person="DongJu".

1. 설명 "DongJu"를 개념 스키마에서 검색한

1) '178'은 'walk'의 단순화된 특징 시퀀스를 나타내는 것으로서 그림 4에 그 예가 나와있다.  
 2) 여기서는 비디오의 동작들이 [33]등에 제시된 방법으로 분석되어 있음을 가정한다.

- 다.
- 2. 단계 1의 식별자를 찾는다
- 3. 단계 2의 식별자를 이용하여 비디오 구간을 검색한다.

·Type-5 질의: Retrieve video where action="walk".

- 1. 설명 "walk"과 관련된 개념데이터를 찾는다.
- 2. 단계 1의 개념 데이터의 부개념데이터를 찾는다.
- 3. 단계 2의 부개념 데이터를 이용하여 특정 데이터를 검색한다.
- 4. 단계 3의 특정 데이터를 포함한 비디오 구간을 검색한다.

5. 원형(prototype) 구현

우리는 *MuVi* 모델에 기반한 비디오 검색 시스템의 원형(prototype)을 구현해보았다. 각 계층은 그림 8에서와 같이 비디오 검색 시스템의 주요 부분으로서 구현되었다. 계층 1은 원비디오 데이터(raw video data)를 재생하거나 임의의 이미지 프레임을 추출하기 위한 비디오 재생기를 포함한다. 원 비디오 데이터는 데이터베이스 내에서 대형 객체(large object) 형태로 저장된다. 계층 2에는 특징 벡터를 비교하여 일치 또는 유사한 이미지 프레임을 찾을 수 있는 벡터 질의와 공간 질의가 포함되어 있다. 계층 3은 비디오의 시간적 속성을 검색하는 부분으로서 특징으로부터 추출된 정보를 표현하기 위해 시퀀스를 표현하고 조작하기 위한 함수들로 이루어져 있다. 계층 4는 의미 데이터와 무의미 데이터의 의미 상의 차이를 줄이기 위해 존재하는 것으로서 유사한 특징 벡터 또는 특징 시퀀스를 가지는 무의미 데이터를 그룹화하여 이를 무의미 데이터의 추상화된 개념으로 제공할 수 있도록 한다. 그룹화를 통한 추상 데이터는 부개념 데이터로 생성하여 계층 4에서 관리된다. 부개념 데이터는 계층 5에서 관리되는 개념 데이터와 연관 관계를 가지게 되며 이때 개념 데이터와의 연관 정도를 나타내기 위해  $n \times m$  상관 행렬(correlation matrix)을 이용하였다. 상관 행렬에서 개념 데이터와 부개념 데이터의 상관도는 DB를 생성하는 사용자가 부여한다. 계층 5에서는 비디오 데이터가 포함하고 있는 모든 개념적 데이터를 관리하고 검색하기 위한 계층이다. 이 계층의 모든 데이터는 데이터베이스를 만들 때 주석 기반 검색에서의 주석 부여와 같은 형태로 이루어진다. *MuVi* 모델에서는 개념 데이터를 효율적으로 관리하기 위해서 개념데이터를 추상 데이터 및 인스턴스 데이터와 같이 구분하고 추

상 데이터에는 중복을 허용하지 않게 하고, 인스턴스 데이터는 추상 데이터의 인스턴스와 하여 비디오 스트림 내에서 중복을 허용하였다. 이 때 각각의 인스턴스 데이터들은 내용 기반 검색의 특징 데이터 및 특징 시퀀스 데이터와 대칭을 이루며, 추상 데이터는 내용 기반 검색의 부개념 데이터와 대칭을 이루게 된다. 이에 따라 특징 데이터는 부개념 데이터의 인스턴스로 간주한다. 이를 위한 자료 구조는 그림 7과 같다.

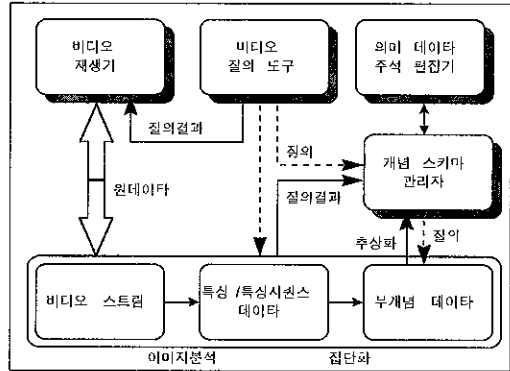


그림 8 *MuVi* 모델에 기반한 비디오 검색 시스템의 구조

본 비디오 검색 시스템의 원형은 Sparc 20, GNU C++ 및 X/Motif, Tcl/tk를 이용하여 구현되었으며, 비디오 데이터를 저장하기 위한 저장시스템으로는 SNU Object-Oriented DBMS Platform(SOP)을 이용하였다. 비디오 데이터는 카메라를 고정하고 일정한 거리만큼 떨어져서 13가지 동작을 수행하게 한 후 이를 SunVideo 장비를 이용하여 MPEG1으로 변환하였다. MPEG 재생기로부터 얻은 이미지와 이미지의 시퀀스에 대해 sobel 알고리즘을 적용하여 외곽선을 추출하였으며 추출한 외곽선으로부터 MBR(minimum bounding rectangle)에 기반하여 대상을 단순화한 특징 벡터를 얻었다. 본 구현에서는 대상물 및 대상물의 동작에 한하여 특징을 추출하고자 하였으며, 이에 따라 대상물의 위치나 크기에 대해 특징 벡터가 영향을 받지 않도록 얻어진 특징 벡터를 이동(transform) 및 크기변환(scaling)을 적용하여 크기나 위치의 차이에 따른 값이 없어지도록 정규화시킨 후, 이를 동작 시퀀스로 생성하였다.

비디오 검색 시스템을 통하여 비디오 검색에 사용될 조건을 개념 데이터나 무 개념 데이터를 이용하여 기술할 수 있으며, 그 결과를 재생하여 볼 수 있다. 그림 9는

구현된 비디오 검색 시스템의 검색 화면으로서 검색 조건을 특징 또는 의미로 기술할 수 있다.

### 6. 요약 및 향후 연구

본 논문에서는 이미 존재하는 비디오 데이터의 효율적인 검색을 위해, MuVi 모델을 제안하고, 이를 이용하여 이미지 분석 기법을 이용하여 추출되거나 시퀀스 분석 기법을 이용하여 추출된 특징 데이터와 같은 무개념 데이터와 사용자의 주석에 의해 부여되는 개념 데이터의 질의를 일관성있게 처리할 수 있는 모델을 제시하였다. 또한 무개념 계층(semi-conceptual layer)을 도입하고 이를 이용하여 개념 데이터와 무개념 데이터와의 관계를 모델링할 수 있다.

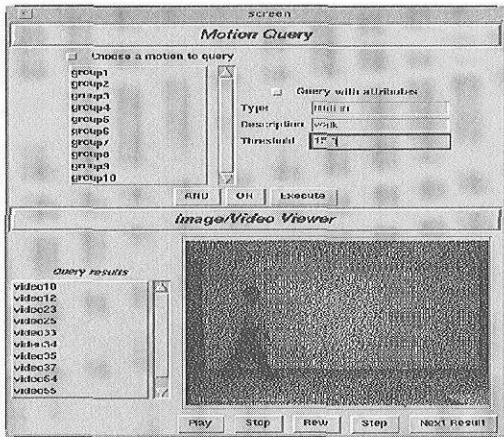


그림 9 의미 및 특징 데이터 조건에 의한 비디오 질의 윈도우

계층의 분리를 통하여 가공된 비디오 데이터의 계층 간의 물리적 독립성과 논리적 독립성을 제공하였으며, 각 계층에서는 계층에 적합한 연산을 제시하고 계층 간에는 데이터의 일반화를 지원하기 위한 추상화 연산들을 제시하였다. 이를 통하여 내용 기반 검색 및 주석 기반 검색이 효과적으로 수행될 수 있으며 이들이 동일한 프레임워크로 통합될 수 있음을 보였다.

각 계층이 독립적으로 이루어져 있으므로, 각 계층의 연산의 개선이나 추가 등이 질의 계층에 영향을 미치지 않는다. 이에 따라, 각 계층의 연산들은 여러가지 기법을 바꾸어가며 시험해 볼 수 있는 테스트 시스템의 기능을 제공할 수 있다. 이에따라 다양한 기법의 내용 기반 검색이나, 시퀀스 검색 기법의 적용이 가능하다. 본 논문에

포함되지는 않았지만 검색에서 지원이 거의 되지 않는 오디오 데이터에 대한 검색의 지원도 MuVi 모델의 계층 3을 이용하면 쉽게 지원될 수 있다.

추후 연구로는 본 모델을 이용하여 다양한 이미지 검색 기법과 시퀀스 검색 기법을 지원하는 비디오 검색 시스템으로의 확장이 필요하며, 의미의 차이가 큰 특징 데이터와 의미 데이터를 간의 관계를 보다 효율적으로 표현할 수 있는 메트릭에 대한 연구등이 필요하다. 질의 처리에서는 사용자가 질의를 계층 간의 추상화 연산과 계층간의 유사성 메트릭등을 이용하여 개념 계층의 조건만을 이용해서 내용 기반 검색이나 주석 기반 검색을 처리할 수 있도록 하는 유일한 검색 인터페이스로의 확장에 관한 연구가 필요하다.

### 참고 문헌

- [1] Sibel adali, K. Selcuk Candan, Su-Shing Chen, Kutluhan Erol, V. S. Subrahmanian, "The Advanced Video Information System: data structures and query processing". *Multimedia Systems*, 4, 1996.
- [2] Tat-Seng Chua, Li-Qun Ruan, "A Video Retrieval and Sequencing System". *ACM Transactions on Information Systems*, 13(4), 1995.
- [3] Young Francis Day, Serhan Dagtas, Mitsutoshi Iino, "Object-Oriented Conceptual Modeling of Video Data". In *Proceedings of International Conference on Data Engineering*, pages 401-408, 1995.
- [4] Nevenka Dimitrova, Forouzan Golshani "Rx for Semantic Video Database Retrieval". In *Proceedings of Second ACM International Conference on Multimedia*, pages 219-226, October 1994.
- [5] Myron Flickner, et. al., "Query by Image and Video Content: The QBIC System". *IEEE Computer*, 28(9), 1995.
- [6] Simon Gibbs, Christian Breiteneder, Dennis Tsichritzis, "Data Modelling of Time-Based Media". In *Proceedings of the 1994 ACM SIGMOD*, pages 91-102, 1994.
- [7] C. L. Hamblin, "Instants and Intervals". In *Proceedings of 1st Conf. Int. Soc. for the Study of Time*, pages 324-331, Springer Verlag, 1972.
- [8] Rune Hjelsvold, "Video Information Contents and Architecture". In *Proceedings of The 4th International Conference on Extending Database Technology*, March 1994.
- [9] Rune Hjelsvold, Roger Midtstraum, "Modelling and Querying Video Data". In *Proceedings of 20th International Conference on VLDB*, pages 686-694, October 1994.
- [10] 김기병, 김형주, "비디오 데이터베이스에서 동작의 빠른 검색"

- 색 기법”. 정보과학회 논문지 심사중, 1996.
- [11] 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.
- [12] Henry F. Korth, Abraham Silberschatz, *Database System Concept*, McGraw Hill Inc., 2nd edition, 1991.
- [13] M. K. Leung, Y. -H. Yang, “First Sight: A Human Body Outline Labeling System”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(4):359-377, April 1995.
- [14] Eitetsu Oomoto, Katsumi Tanaka, “OVID: Design and Implementation of a Video-Object Database System”. *IEEE Transactions on Knowledge and Data Engineering*, 5(4), August 1993.
- [15] Gerhard A. Schloss, Michael J. Wynblatt, “Providing definition and temporal structure for multimedia data”. *Multimedia Systems*, 3, 1995.
- [16] Stephen W. Smoliar, HongJiang Zhang, “Content-Based Video Indexing and Retrieval”, *IEEE Multimedia*, 1(2):62-72, 1994.
- [17] Ron Weiss, Andrzej Duda, “Composition and Searching with a Video Algebra”. *IEEE Multimedia*, 2(1), 1995.



김 기 병

1990년 서울대학교 계산통계학과 졸업.  
1992년 서울대학교 계산통계학과 전산과  
학전공 석사. 1991년 ~ 1992년 서울대학  
교 중앙교육연구전산원 조교. 1992년 서  
울대학교 컴퓨터공학과 박사과정. 관심분  
야는 멀티미디어 데이터베이스, 비디오

모델링, 비디오 검색



김 형 주

1982년 서울대학교 컴퓨터공학과 졸업.  
1985년 Univ. of Texas at Austin 전산  
학 석사. 1988년 Univ of Texas at  
Austin 전산학 박사. 1988년 Univ. of  
Texas at Austin Post-Doc. 1988년 ~  
1989년 Georgia Institute of

Technology 조교수. 1991년 ~ 1996년 서울대학교 컴퓨터  
공학과 부교수. 관심분야는 객체지향시스템, 사용자인터페이  
스, 데이터베이스.