

질의완화를 기반으로 한 지능적인 비디오 검색 시스템

윤미희*, 조동욱**

*충북과학대학 컴퓨터정보과, **충북과학대학 정보통신과
e-mail:mihee@ctech.ac.kr

A Study on Intelligent Video Retrieval System based on query relaxation

Mi-Hee Yoon*, Dong-Uk Cho**

*Dept of Computer Information, Chungbuk Provincial college

**Dept of Information Technology, Chungbuk Provincial college

요 약

최근 하드웨어와 압축기술의 발달 및 보편화로 인해 사용자들의 비디오 데이터에 대한 요구가 증가하였다. 비디오 데이터는 비정형, 대용량의 특징을 가지고 있으므로 사용자의 다양한 요구를 만족시키기 위해서는 단순한 텍스트 형식의 데이터가 아닌 비디오 데이터에 대한 다양한 검색기법이 요구된다. 효율적인 비디오의 검색을 위해서는 사용자의 불완전한 질의에도 근사한 질의결과의 제시가 필요하다. 본 논문에서는 비디오데이터에 대한 효율적인 의미검색을 위해 주석기반과 특징기반을 혼합한 내용기반 검색을 지원하며 특히 사용자의 불완전한 질의에도 근접한 질의결과를 제시할 수 있는 지능적인 비디오 검색 시스템을 제안한다.

1. 서론

하드웨어와 압축기술의 발달로 인해 멀티미디어를 저장하고 관리하기 위한 다양한 요구가 최근 인터넷 및 초고속통신망의 구축으로 인해 사용자들의 멀티미디어에 정보에 대한 다양한 요구가 증가하였고, 이러한 다양한 요구에 대해 효율적으로 서비스를 제공하기 위해서는 단지 텍스트 데이터뿐만 아니라 이미지, 오디오, 비디오 데이터를 처리할 수 있는 멀티미디어 데이터베이스 시스템을 구축하는 것이 필요하다.[1,2] 다양하고 방대한 비디오데이터를 효율적으로 저장, 검색하는 기능은 비디오 데이터베이스에서 가장 중요한 요소이다.

비디오 정보는 방대할 뿐만 아니라, 텍스트나 수식 데이터와 같이 데이터 자체의 속성을 이용하여 처리하기 어렵기 때문에 비디오 데이터베이스에 저장된 객체들의 의미를 기반으로 한 내용기반 검색이 필요하다. 비디오 데이터베이스의 내용기반 검색을 제공하기 위해서는 비디오 데이터의 내용을 구조적

으로 체계화한 데이터가 필요하다. 이러한 데이터를 메타데이터라고 하며, 메타데이터는 내용기반 검색의 가장 필수적인 데이터이다. 그러므로 효율적인 내용기반 검색을 위해서는 메타데이터에 대한 모델링이 필요하다[5,6]. 또한 사용자의 불완전한 질의에도 근접한 질의 결과를 제시할 수 있는 기법이 요구된다.

본 논문에서는 특징기반 검색과 주석기반 검색을 통합하여 다양한 사용자의 의미검색을 지원하고, 유사성 질의와 내용기반 검색시 제시된 질의에 대한 정확한 검색 결과를 보여 주지 못하는 경우 질의 완화를 통한 근접 질의 결과를 제시할 수 있는 지능적인 비디오 검색 시스템(Intelligent Video Retrieval System)을 제안한다. IVRS는 기존의 선행 연구[7]를 바탕으로 원하는 비디오데이터의 검색을 위해 세 개의 계층을 가진 통합 객체지향 데이터 모델을 제안한다. 제안한 메타데이터 모델링을 기반으로 비디오 데이터의 내용기반 검색과, 장면과 객체의 유사

성 검색이 가능하다. 특히 사용자의 질의에 대해 만족할 만한 결과를 얻지 못하는 경우에 근접한 질의 결과를 제시하기 위해 속성을 기반으로 질의 완화를 수행하여 효과적인 질의결과를 제시한다.

본 논문의 구성은 다음과 같다. 2장에서는 관련연구로 기존의 비디오데이터베이스 시스템에 대해 알아보고, 3장에서는 비디오 데이터의 내용기반 검색을 위한 메타데이터 모델링과 비디오 질의언어[7]를 설명한다. 4장에서는 메타데이터와 비디오 질의언어를 기반으로 한 지능적인 비디오 검색시스템을 설명한다. 마지막으로 5장에서는 결론 및 향후 연구과제에 대해 설명한다.

2. 관련연구

본 장에서는 기존의 비디오 데이터베이스 시스템에 대해 설명한다.

AVIS(Advanced Video Information System)[4]는 미국 메릴랜드 대학에서 개발된 비디오데이터베이스 시스템으로 비디오 내용에 나타나는 객체, 사건 행위 유형에 대한 메타데이터를 정의하고, 이들을 비디오세그먼트와 연계시킴으로써 효율적인 검색 방법을 제안했다. 그러나 AVIS는 상위 레벨의 내용을 중심으로 하는 주석기반의 메타데이터만을 제공하고, 하위레벨을 위한 메타데이터가 없다는 단점이 있다.

Video-Star[1]는 비디오 데이터와 메타데이터의 공유와 재사용을 지원하도록 만든 범용 비디오 데이터베이스 시스템으로, basic context, secondary context 개념을 이용하여 데이터를 공유했다.

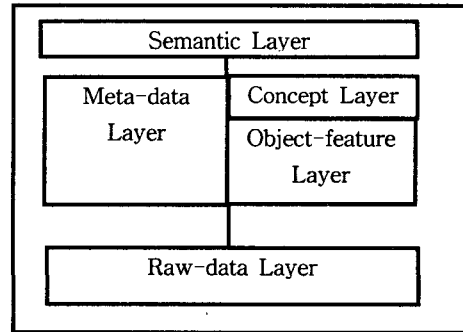
Jiawei는 지능적인 질의 처리 기능을 향상시키기 위해 일반화 규칙들과 개념계층[8][9], 지식발견 도구등을 사용하여 지능적인 질의응답을 제공하는 프레임워크를 제안했다. Jiawei의 연구는 다양한 방법을 이용한 지식발견을 통해 지능적인 질의 응답을 지원함으로써 질의처리기능을 향상시켰으나 질의 유형에 따른 처리과정이 비효율적이다.

3. 메타데이터 모델과 비디오 질의언어

본 논문에서 제안하는 세 개의 계층을 가진 통합 객체지향 메타데이터 모델(Three layered Hybrid Object-oriented Metadata Model : THOMM)은 원시데이터 계층(raw-data layer)과 비디오의 의미검색을 위해 특징과 주석을 통합하는 메타데이터 계층

(meta-data layer)과 질의 재형성을 위한 의미계층으로 구성되며 모델링을 위해 UML을 사용하였다.

비디오를 위한 THOMM은 (그림 1)과 같다. 각 계층에 대한 정의는 다음과 같다.



(그림 1) 세 개의 계층을 가진 통합 객체지향 메타데이터 모델 (THOMM)

(정의 1) 원시데이터 계층(raw data layer)

물리적인 비디오 스트림을 표현하는 계층

(정의 2) 메타데이터 계층(metadata layer)

비디오에 대한 정보와 비디오 스트림에 포함된 의미를 기술하기 위한 계층으로 개념 계층(concept layer)과 객체특징 계층(object feature layer)의 부계층(sublayer)으로 구성된다.

(정의 2-1) 개념 계층(concept layer)

비디오 정보와 비디오 스트림에 나타나는 추상적인 의미를 표현하기 위한 계층

(정의 2-2) 객체특징 계층(object feature layer)

장면에 나타나는 객체들의 특징을 기술하기 위한 계층

(정의 3) 의미 계층(semantic layer)

메타데이터 계층에 나타나는 의미들 사이의 연관관계를 표현하는 계층으로 IF - THEN규칙의 형태로 표현된다. 본 논문에서는 IF -THEN 규칙을 지식이라 칭하고 지식베이스를 구성한다.

본 논문에서 제안하는 지능적인 비디오검색 시스템에서 사용되는 비디오 질의언어의 구문은 다음과 같다[7].

```

    FIND <result>
    FROM <meta-db>
    WHERE <predicate_clause>
  
```

FIND절은 사용자가 검색하고자 하는 대상으로 비디오다큐먼트, 시퀀스, 장면, 객체가 될 수 있다.

FROM절은 질의의 탐색영역을 정의하며 WHERE절은 검색조건을 정의한다. 특히 장면에 대한 질의 시주어진 장면과 유사한 장면을 검색하기 위한 유사성구문과, 주어진 장면을 기반으로 객체의 추가, 삭제를 통해 검색의 범위를 축소하고 확대하면서 사용자가 원하는 좀 더 정확한 검색을 지원하는 예제기반 장면 질의 (SQBE)가 가능하다.

4. 지능적인 비디오 검색 시스템

(Intelligent Video Retrieval System : IVRS)

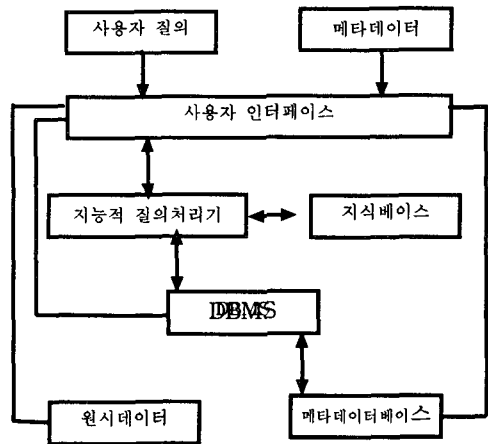
본 장에서는 지능적인 비디오 검색 시스템을 제안한다. IVRS의 구조를 설명하고 질의의 결과가 존재하지 않을 경우 지식을 기반으로 사용자의 질의를 완화하는 과정을 통해 질의가 재형성되는 방법을 제시한다.

4.1 IVRS의 구조

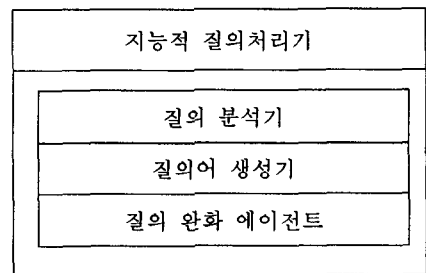
IVRS의 구조는 (그림 2)와 같다. IVRS는 사용자 인터페이스, 지능적 질의처리기(Intelligent Query Processor : IQP), 상용의 DBMS, 메타데이터베이스, 원시데이터 지식베이스로 구성된다. 사용자는 사용자 인터페이스를 통해 데이터의 입력과 질의를 수행한다. 메타데이터 입력은 동영상 플레이어에 의해 플레이되는 비디오를 보면서 해당 속성에 속성값을 부여한다. 원시데이터는 프레임으로 구성된 디지털 비디오이고, DBMS는 상용의 데이터베이스 관리 자로써 메타데이터베이스를 검색한다.

지능적 질의처리기는 사용자로부터 질의를 입력 받아 질의를 분석한 후 상용의 DBMS에서 사용되는 SQL문을 생성한다. 상용의 DBMS는 해당하는 질의의 결과를 찾기 위해 메타데이터베이스를 탐색하고 탐색결과를 플레이한다. 만일 질의의 결과가 존재하지 않는다면 질의 재형성 에이전트가 지식베이스를 이용하여 자동적으로 질의 완화과정을 수행하여 질의를 재형성한 후 검색을 수행한다. 지능적 질의처리기는 질의 분석기, 질의어 생성기와 질의완화 에이전트(query relaxation agent)로 구성되어 있다. 지능적 질의 처리기의 구조는 [그림 3]과 같다.

지능적 질의처리기는 질의어를 생성하여 상용의 DBMS에 의해 검색을 수행한다. 만약 검색결과가 존재한다면 검색결과가 제시되고 원하는 결과를 선택하면 동영상 플레이어가 실행되어 결과 영상을 플레이한다.



(그림 2) IVRS의 구조



(그림 3) 지능적 질의처리기의 구조

결과가 존재하지 않는다면 질의 완화 에이전트는 사용자에게 완화 불가능한 속성을 입력받을 수 있고 이런 속성이 입력되면 완화 불가능한 속성과 사건의 속성을 기준으로 뷰를 생성한다. 질의 완화 에이전트는 완화가 가능한 속성이 선택되었는지를 검사하고, 완화 가능한 속성이 선택되었다면 그 속성을 기반으로 지식베이스의 규칙에 의해 질의 재형성 과정을 수행하여 질의 문장을 재형성한다. 질의 문장이 재형성되면 만들어진 뷰에 의해 질의 재형성 과정에서 생성된 질의조건을 사용하여 결과를 생성한다. 만일 사용자로부터 질의완화가 불가능한 속성이 입력되지 않았다면 사건의 속성만으로 뷰를 생성한다.

4.2. 지식베이스

THOMM의 의미계층은 사용자의 정확한 질의뿐만 아니라 부정확한 질의에도 결과를 제시하기 위해 사용되는 계층이며 메타데이터 계층의 클래스의 속성을 연관시켜 만든 규칙(지식)으로 질의 재형성을 수행하기 위한 계층이다. 질의 재형성을 위한 의미

기반 속성 연관트리(Semantic-based Attribute Association Tree)는 [그림 4]와 같다. 이런 규칙의 모임으로 지식베이스는 생성되고 공유가 가능하다.

지식베이스를 생성하기 위해서는 SAAT를 기반으로 한 속성에서 그와 연관된 다른 속성으로의 변환 과정이 필요하다. 속성을 변화하기 위한 속성변환기법(Attribute Transfer Method : ATM)은 다음과 같다.

(정의 4) ATM

$ATM(X, Att_value) \rightarrow REP(Y, Att_Value1)$
 (단, $(X, Y \in \text{속성}), (Att_Value, Att_Value1 \in \text{속성값})$)

즉, X의 속성에 Att_value라는 값이 입력되었을 경우 원하는 결과를 찾을 수 없다면 질의 완화 에이전트에 의해 질의어가 재형성된다. 질의 재형성을 위해 Att_value의 속성값에 따라 X라는 속성은 Y라는 속성으로 변화되고 속성값인 Att_value도 Y속성에 해당하는 속성값인 Att_value1으로 변환된다. t이런 속성과 속성값의 변화과정을 통해 질의어는 재형성된다.

SAAT를 이용한 의미계층의 지식은 질의 일반화 규칙(query generalization rule)과 질의 재형성 규칙(query reformulation rule)으로 구성된다. 질의 일반화 규칙은 일반적인 의미를 가진 질의어를 찾기 위한 규칙으로 위쪽의 노드는 일반화된 의미를 가진 노드이고, 아래쪽 노드들은 세분화된 의미를 가진 노드이다. 질의 재형성 규칙은 주어진 질의어를 이용하여 질의를 재형성하기 위한 규칙으로 일반화된 질의어를 조건으로 하고 일반화 의미와 일반화 의미 노드의 자식노드인 상세화 의미들의 결과로 하는 규칙이다. 질의 재형성 규칙에 의해 일반화 의미와 일반화 의미와 관련된 모든 상세화 의미가 질의조건으로 확장된다.

4.3 질의 처리

IVRS는 먼저 사용자가 관심 있는 장르와 분야, 주제를 입력받아 지식베이스에 사실로 저장한다. 이 지식은 질의 결과의 수가 많을 경우 질의 처리시 필터링을 위해 사용된다. 사용자의 질의는 주석기반 질의와 특징기반 질의로 구성될 수 있으며, 이 질의어를 입력받아 질의 분석을 수행하고 질의어를 생성한 후 검색을 수행한다. 만일 검색결과가 존재한다면 검색결과를 플레이하고 존재하지 않는다면 질의어를 재형성한다. 질의어 재형성은 질의 완화 에이전

트(Query Relaxation Agent : QRA)에 의해 수행되며 QRA는 질의완화가 불가능한 속성을 제외한 나머지 속성들 중 SAAT를 기반으로 질의를 재형성한다.

7. 결론 및 향후 연구과제

본 논문에서는 메타데이터 모델링을 기반으로 비디오 데이터의 내용기반 검색과, 장면과 객체의 유사성 검색이 가능하며, 질의의 결과가 존재하지 않을 경우 질의 완화 과정을 통해 자동적으로 질의를 재형성하여 근접한 질의 결과를 제시한다.

향후 연구과제로는 효율적인 검색을 위한 인덱싱 기법과 음성 인식 및 자연어 이해 기술을 통합하여 메타데이터를 생성하고 질의할 수 있는 기법의 연구가 필요하며 이렇게 생성된 메타데이터를 기반으로 실시간으로 원하는 질의결과를 얻을 수 있는 방법에 대한 연구가 수행되어야 한다.

참고문헌

[1] R. Hjelmsvold, "VideoSTAR - A database for Video Information Sharing," Ph. D. Thesis, Norwegian Institute of Technology, Nov. 1995.
 [2] John Z. Li, M. Tamer Ozsu, and Duane Szafron, "Querying Languages in Multimedia Database Systems," TR95-12, The university of Alberta, Canada, 1995.
 [3] E. Oomoto and K. Tanaka, "OVID: Design and Implementation of a Video-Object Database System," IEEE Trans. on Knowledge and Data Engineering, Vol.5, No.4, 1993.
 [4] Sibel Adali, et. al., "The Advanced Video Information System : data structures and query processing," Multimedia Systems, 1996.
 [5] S. Smoliar and H. Zhang, "Content-Based video Indexing and Retrieval," IEEE Multimedia, 1994.
 [6] W. Klas and A. Sheth, "Metadata for Digital Media : Introduction to the special issue," SIGMOD Record, Vol.23, No.4, Dec. 1994.
 [7] 윤미희외, "비디오의 의미검색과 유사성 검색을 위한 통합비디오정보시스템", 한국정보처리학회 논문지 제 6권 8호, Aug. 1999.
 [8] Jiawei HAN, "Intelligent Query Answering by Knowledge Discovery Technique", Vol.8, No.3, June 1996.