

트리기반 멀티미디어 데이터의 의미적 검색을 위한 프레임워크 설계

안형근^o 고재진 양상석
울산대학교 컴퓨터·정보통신공학부
{hkahn^o, jkjh, wintli}@mail.ulsan.ac.kr

Design of the Framework for Tree-based Semantic Retrieval of Multimedia Data

Hyoung-Keun An^o, Jae-Jin Koh, Sang-Seok Yang
School of Computer Engineering & Information Technology, University of Ulsan

요 약

현재 웹을 사용하는 사람, 디지털 콘텐츠, 정보의 생성이 계속적으로 증가하여 오고 있다. 최근 웹 데이터는 멀티미디어 데이터(비디오, 오디오, 웹TV, 이미지 등)라고 말해도 무방할 것이다. 이런 방대한 멀티미디어 데이터의 검색은 사용자들에게는 아주 중요한 작업이 되었으며, 효과적인 멀티미디어 검색 시스템의 필요성이 대두되고 있다. 최근까지의 멀티미디어 데이터 검색은 내용-주석기반의 검색이 주를 이루고 있다. 하지만 내용-주석 기반 검색은 의미-추론적 검색에서는 부족함을 보여주고 있는 실정이다. 따라서 본 논문에서는 검색 방법 중에서도 주석 기반 검색에서 가지는 문제점을 살펴보고 그에 대한 해결 알고리즘과 프레임워크를 제안한다.

표준을 기초로 하기 때문에 그에 대한 기본 개요와 표현을 알아보고, 연구 동기가 된 주석 기반 검색 질의에 대한 문제점에 대해 먼저 살펴본다.

1. 서 론

최근 디지털 형식의 멀티미디어 데이터가 상당히 증가하고 있는 실정이다. 따라서 멀티미디어 데이터로부터 유용한 정보의 추출과 멀티미디어 검색 엔진 같은 실질적인 시스템의 정보 조작은 여전히 문제로 남아 있다. 가장 중요한 문제점이라고 할 수 있는 것이 멀티미디어 데이터의 이해력과 단순하고 유연한 표현이 부족하다는 것이다. 최근 MPEG-7 표준은 거의 모든 멀티미디어 정보를 포함하기 위해 기술들의 일반적인 라이브러리를 제공함으로써 문제점들을 해결하려고 하고 있다.[1]

최근 멀티미디어 모델은 MPEG-7 표준에 정의되어 있는 멀티미디어 기술 구조와 기술 정의의 언어인 XML 스키마를 사용하여 멀티미디어 데이터의 특성 및 계층적인 구조를 표현하고 있다. XML 스키마는 MPEG-7이 요구하는 구문, 구조, 관계, 타입, 제약을 표현하기에 적당하다.[5] 멀티미디어 검색에는 내용 기반(Content-based)과 주석 기반(Annotation-based) 검색이 있는데 그 중 주석 기반 검색은 멀티미디어 데이터에 대해 XML 기반으로 주석 처리 저장하고 검색하는 대표적인 예이다. 하지만 주석 기반 검색은 멀티미디어에 저장된 주석을 기초로 단순 검색이 이루어지는 단점을 보이고 있다.

따라서 본 논문에서는 MPEG-7에서 보이는 주석 기반 검색의 단점을 보완하기 위해서 멀티미디어 데이터의 의미적인 검색을 위한 해결에 초점을 맞추어 보았다. 따라서 MPEG-7 표준은 추론적인 지식 영역 표현을 위한 온톨로지로서 확장할 것이며, 확장된 데이터의 의미 검색을 위해 트리 삼입 기법을 사용, 이를 이용한 프레임워크를 제안하고자 한다.

나머지 논문의 구성은 다음과 같다. 2절에서는 MPEG-7 개요와 표현, 주석 기반 검색 문제점을 살펴본다. 3절에서는 프레임워크에 적용할 트리 삼입 기법에 대해서 알아보고, 4절에서 프레임워크 설계를 기술한다. 마지막 5절에서 결론 및 향후 과제로서 끝을 맺는다.

2. 관련 연구

본 논문의 사용자 질의 및 멀티미디어 데이터 모두 MPEG-7

2.1 MPEG-7 개요

MPEG-7은 공식적으로 "Multimedia Content Description Interface"로 불리며, 사용자가 요구하는 다양한 형태의 멀티미디어 데이터에 대한 빠르고 효율적인 검색을 해결할 목적으로 국제표준기구(ISO/IEC)에서 제정하였다.

MPEG-7은 멀티미디어의 구조와 색인의 부호화를 표준화 하는 것으로, MPEG-7에서는 이를 기술자(Ds : Descriptor), 기술체계(DSs : Description Scheme), 데이터정의언어(DDL : Data Description Language), 부호화 방법으로 나누어 표준을 정하고자 한다. 기술자(Ds)는 어떤 특징을 나타내는 표현 양식과 그 의미를 정의 한다.(예 : 질감, 컬러, 질감 주석, 매체 형식), 기술체계(DSs)는 하나 이상의 기술자와 기술체개로 구성되며, 기술자와 기술체계들의 관계를 명시하는 것이다. 기술정의 언어(DDL)은 기술자와 기술체계의 정의, 조작 등을 표현하는데 사용하는 컴퓨터 언어이다.

멀티미디어 데이터 기술은 아래 영역으로 구성이 된다.

- 저장 매체의 기술 : 파일, 코딩형식, 이미지크기
 - 작성과 생성 정보 : 작성 날짜, 위치, 제목, 장르
 - 내용의 의미 기술 : 사건, 개념, 객체
 - 내용 구조 기술 : 샷, 키 프레임, 질감, 동작의 특징
 - 기술에 대한 메타데이터 : 저자, 버전, 작성일자
- 위 모든 정보의 형태는 MPEG-7 기술 체계에 의해서 다를 수가 있다.[1]

2.2 MPEG-7 표현

MPEG-7 기술은 인스턴스의 기술체개로 XML 문서로서 표현한다. 여러 모델들이 XML 문서로 표현하기 위해 제안되었으며, W3C에서는 DOM(Document Object Model)이라는 일반적인 모델을 제안 하였다.[6] 이 모델에서는 XML 문서가 순서 레이블 트리로 표현되며, MPEG-7 기술에 또한 적용이 된다.

순서 레이블 트리는 $Tree(T) = (N, E, R(T))$ 로 표현할 수 있으며, 트리를 구성하는 세 개의 노드와 기타 요소들을 나열하면 다음과 같다.

- $Node(e)$: element nodes, 이름 노드
- $Node(t)$: text nodes, 값 노드
- $Node(a)$: attribute nodes, 이름과 값 노드를 가진
- $N = Node(e) \cup Node(t) \cup Node(a)$
- $E(edge)$: XML 문서에서 엘리먼트를 검색할 간선
- $R(T)$: root node

아래 그림 1은 MPEG-7 기술에서 비디오 데이터를 예로 할 경우의 MPEG-7 기술 트리다.

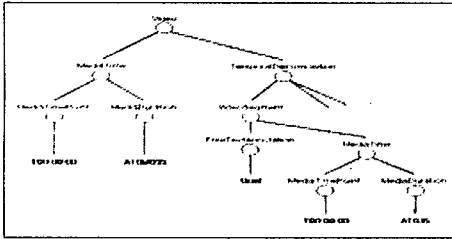


그림 1 MPEG-7 기술 트리

멀티미디어 기술체계와 마찬가지로 사용자 질의도 MPEG-7 기술에 의해서 해석되어지고 순서 레이블 트리의 질의 트리를 구성을 한다.

2.3 주석 기반 검색의 문제점

주석 기반 검색(Annotation-based Retrieval)은 멀티미디어 데이터에 대한 정보를 사용자가 결정하고 이를 주석으로 작성하여 저장한 후에 질의에 나타난 내용들을 주석과 비교하여 관련된 데이터를 찾는 방법이다. 이 방법은 사람이 인식할 수 있는 모든 내용을 주석으로 처리할 수 있기 때문에 여러 복잡한 개체간의 관계들로 이루어진 자료의 의미를 쉽게 모델링 할 수 있는 장점이 있다. 하지만 대용량의 데이터에 대하여 사람이 일일이 수동적으로 주석을 작성해야 하며 사용자의 관점에 따라 내용을 다르게 해석하거나 주석을 다르게 부여할 수 있으므로 데이터에 대한 일관성을 잃기 쉽다. 또 검색이 데이터에 부여된 주석에 의존하기 때문에 의미적인 검색이 어렵다는 단점을 가지고 있다. 다음은 주석 기반 검색의 한 예를 보여주고 있다. 그림 2와 같이 사용자 질의 "Run Animal"이 주어지고, 그림 3과 같은 멀티미디어 데이터베이스에서 주석 기반 검색의 결과가 어떻게 되는지를 고려해 보자.

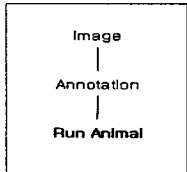


그림 2 Query

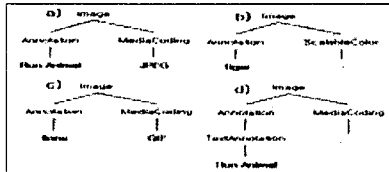


그림 3 멀티미디어 데이터베이스

위 결과 분명한 질의 결과는 a)이다. b), c)의 tiger나 lion은 분명히 달리는 동물이지만 이 질의의 결과로는 맞지가 않다. 여기에서 사용자 질의는 "Run Animal"이라는 주석 검색만을 요구했기 때문에 개념적인 달리는 동물들은 검색이 안 되는 문제점을 가지고 있다. 이 문제점의 해결을 위해서 개념적으로

유사한 데이터 추출을 위하여 추론적인 기법이 필요한데, 여기서 개념적인 동물도메인 온톨로지 구성으로 명확히 b)나 c)도 질의 답변으로 추출 될 수가 있을 것이다. 그림 4는 문제점 해결을 위한 동물도메인 온톨로지를 보여주고 있다.

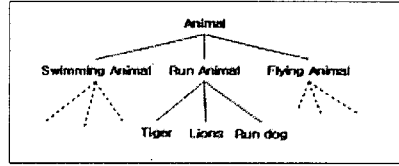


그림 4 동물도메인 온톨로지

3. 트리 삽입 전략(Tree Embedding strategy)

이미 트리나 그래프 삽입 기법은 이질적인 XML 문서 검색에서 많이 연구되었으며, 그 기법들 중 하나가 근사 트리 삽입 기법(Approximate Tree Embedding)이다.[2] MPEG-7 형식이 XML 기반에 의해 계층적 트리로 표현되므로, 표준화된 MPEG-7 기술도 트리의 특성을 가지는 계층적 구조 표현이 가능하다고 관련 연구에서 언급 하였고, 앞 절 예시에서 다루었던 주석 기반 검색에서는 의미적인 검색이 어렵다는 문제점을 기술하였다.

이 절에서는 4절에서 설명할 프레임워크의 검색 부분에 적용할 트리 삽입 기법을 살펴보고자 한다. 트리 삽입 기법에서는 몇 가지의 조건을 내포하고 있다. 트리의 노드 순행으로 후위 순행(postorder)을 따르며 방향성 방식으로 처리를 해야 한다. 또 노드들 간의 관계(예, 조상노드-자식노드)와 의미를 항상 유지하고 있어야 한다. 서로 간의 레이블이 동등하거나 포함관계가 있으면 그 관계도 보존을 하여야 한다.

그림 5는 운동경기에서 각 사람들의 역할에 대한 내용을 바탕으로 사용자 질의와 MPEG-7 데이터를 프레임워크의 처리 과정을 통해 생성된 질의 트리를 데이터 트리에 삽입하는 실행을 보여 주고 있다. 프레임워크에 관한 내용은 다음 절에서 다루는 것이다.

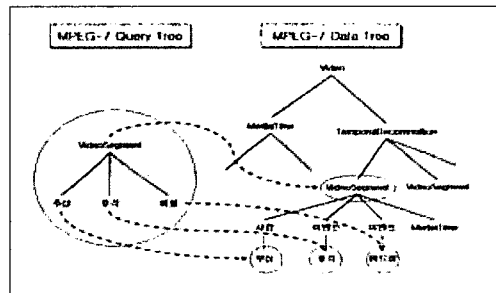


그림 5 데이터 트리에 질의 트리의 삽입

트리 삽입 알고리즘 그림 6과 같으며, 입력은 질의(Q)와 데이터(D) 노드이며, 질의 노드는 후위순행의 결과 노드이다. 출력은 불린 값으로 출력이 된다. 처리는 두 부분으로 나누어 처리된다. i) 잎 노드 질의는 동일한 레이블과 개념적인 포함관계 처리에 사용되며, ii) 중간 노드 질의는 그 중간 노드의 서브트리를 모두 데이터 노드의 서브트리로 삽입 구성한 다음 그 삽입된 데이터 노드에서 질의의 결과를 얻을 때 사용된다.

```

Algorithm (질의 의미의 데이터 트리 사이의 삼입 알고리즘)
Input : Q - 무위중행후 정렬된 질의 노드
       D - 데이터 노드
Output : Boolean (true, false)
Match = true
i = 1
while (Match = true and i ≤ n) do
    // n : 질의 노드의 수
    M(qi) = {}
    // 질의 질의 노드 처리
    for all d ∈ inverse(D) such (label(d) = label(qi) or label(d) ∈ label(qi)) do
        M(qi) = M(qi) ∪ {d}
    end for
    // 질의 내부 노드 처리
    for all d ∈ D such (label(d) = label(qi)) do
        S = {}
        for all qi ∈ children(qi) do
            Si = {}
            for all di ∈ M(qi) such (di is an ancestor of d) do
                Si = Si ∪ {di}
            end for
            S = S ∪ Si}
        end for
        if S ≠ {} then
            M(qi) = M(qi) ∪ {S ∪ {d}}
        end if
    end for
    if (M(qi) = {})) then
        Match = false
    else
        i = i + 1
    end if
end while
return Match;
    
```

그림 6 트리 삼입 알고리즘

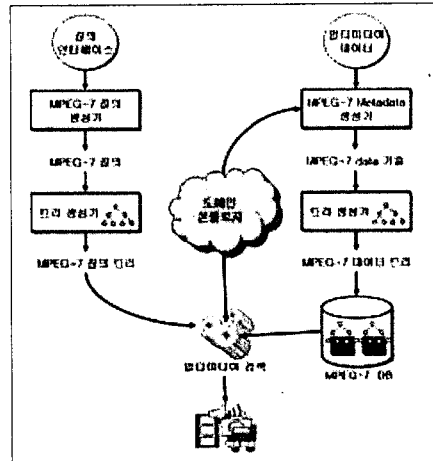


그림 7 프레임워크의 검색처리 과정

4. 프레임워크 구성

의미적인 멀티미디어 데이터 검색을 위한 트리 삼입 기반의 프레임워크를 제안한다. 이 프레임워크 처리과정 구조는 그림 7과 같으며, 프레임워크 내에 구체적인 도메인 온톨로지를 정의한다. 메타데이터 생성 과정에서 이 온톨로지를 통합한다. 메타데이터는 멀티미디어 내용의 색인을 위해 수동적인 처리부분이 되며, 이렇게 처리된 색인자들은 명확한 관계를 가지는 관계로 사용할 것이다. 물론 온톨로지도 색인을 포함한다.

프레임워크에는 서로 다른 네 개의 구성요소들이 있는데 이를 살펴보면 다음과 같다.

- MPEG-7 Metadata 생성기

도메인 온톨로지에 의해 유도되는 메타데이터 생성 처리 과정을 책임진다. 여기서 메타데이터는 멀티미디어 데이터와 텍스트 주석을 입력 받아서 산출되며, 자동적 또는 수동적으로 추출된다. 자동적인 경우는 색상, 파일 크기 등과 같은 낮은 레벨에서 적당하고, 객체나 그 객체의 속성들처럼 높은 레벨은 수동적 주석이 요구가 된다.
- MPEG-7 질의 생성기

사용자 질의를 MPEG-7에 맞는 양식으로 번역되며, 결과 질의는 MPEG-7 기술로써 트리 계층으로 표현된다.
- 트리 생성기

레이블 된 순서트리로 MPEG-7 기술을 모델화를 한다.
- 멀티미디어 검색

제한된 트리 삼입 기법 기반이다. 사용자 질의에 대한 해답을 검색하기 위해 MPEG-7 질의 트리는 MPEG-7 데이터베이스에 저장 되어 있는 MPEG-7 데이터 트리로 삼입이 되며, 사용된 온톨로지과 결합을 하여 원하는 결과 리스트를 얻는다.

5. 결론 및 향후과제

현 사용자들은 웹에서 보고 싶은 콘텐츠 혹은 콘텐츠의 일부분을 효과적이면서 효율적으로 접근할 수 있는 방법을 계속 요구하고, 심지어 지식 검색 수준까지 요구하고 있다. MPEG-7의 표준범위가 인터페이스, 자동검출 알고리즘, 검색엔진까지는 미치지 못하여 의미-추론적인 검색까지 생각 할 수가 없다.

본 논문에서는 MPEG-7의 기술에 도메인 온톨로지를 추가, MPEG-7 표준을 확장할 하고, 트리 삼입 기법을 사용, 검색에서의 확장된 의미 검색이 가능하도록 프레임워크 설계 제안 하였다. 향후 질의에 대한 검색 결과의 유사성과 질의와 유사한 결과가 얼마나 많이 검색 되었는지에 관한 유사성 및 검색 결과를 평가 할 수 있는 평가항수에 대한 연구를 지속적으로 할 계획이다.

[Acknowledgement]

본 연구는 정보통신부 및 정보통신연구진흥원의 대학 IT연구센터 육성·지원사업의 연구결과로 수행되었음.

참고 문헌

- [1] M. Lux, J. Becker, "XML and MPEG-7 for Interactive Annotation and Retrieval Using Semantic Metadata", In Proceedings of I-Know'02, volumn 8, page 965-984, J.UCS, 2002.
- [2] T. Schlieder and F. Nanumann, "Approximate tree embedding for querying XML data", In proceedings of the ACM SIGIR Workshop on XML and Information Retrieval, July 2000.
- [3] P. Liu and L. Sushu, "Queries of digital descriptions in MPEG-7 and MPEG-21 XML documents", XML Europe 2002, Barcelona, Spain, May 2002.
- [4] P. Kilplainen, "Tree Matching Problems with Applications to Structured Text Database", PhD thesis, Dept. of Computer Science, University of Helsinki, 1992.
- [5] W3C, "XML Schema Part 1:Structure Second Edition", <http://www.w3.org/TR/xmlschema-1/>, Oct. 28, 2004.
- [6] W3C, "Document Object Model(DOM)", <http://www.w3.org/DOM/>, 2004.