

의미적 메타데이터를 이용한 멀티미디어 주석 및 검색

안형근^o 고재진

울산대학교 컴퓨터·정보통신공학부

{hkahn^o, jjkoh}@mail.ulsan.ac.kr

Multimedia Annotation and Retrieval using Semantic Metadata

HyoungKeun An^o JaeJin Koh

School of Computer Engineering & Information Technology, University of Ulsan

요 약

최근 멀티미디어의 이용과 멀티미디어 접근을 위한 기술이 많이 증가하고 있다. 그렇지만 멀티미디어 검색엔진과 같은 실용시스템에서 멀티미디어에 대한 유용한 정보 추출과 정보의 응용은 여전히 문제로 있다. 특히, 멀티미디어 이용자는 검색의 효율성을 위하여 저장소를 직관적인 구조로 생성을 하고 있다. 그 예로 "KISS 추계학술 대회 이미지"와 같은 데이터 폴더를 만들거나, 각 멀티미디어 데이터에 Free Text 기반의 주석을 하여 관리를 하였다. 하지만 이러한 검색들에도 한계점을 가지고 있으며, 또 다른 지능적인 의미 검색에 있어서도 인간이 바라는 검색의 정확도에 미치지 못하고 있다.

본 논문에서는 이러한 문제점을 해결하기 위한 새로운 접근법을 소개한다. 목적을 위하여 멀티미디어의 의미적인 작업을 위하여 콘텐츠 획득과 분류를 위한 새로운 사용자 도구를 소개하고자 한다. 도구를 이용하는 멀티미디어 사용자는 주어진 콘텐츠를 인간이 생각하고 콘텐츠가 내포하는 의미의 일정한 구조적 단위로 분해하고, 각 단위들에 MPEG-7 표준기반의 추가적인 기술 정보(Description Information)를 부여하여 새로운 의미적 메타데이터를 생성할 수 있다. 이러한 의미적 메타데이터는 멀티미디어 검색을 위해 사용자들에게 효율성을 줄 것이라 본다.

1. 서 론

디지털카메라 사용자들은 한 해 걸쳐 많은 멀티미디어 콘텐츠를 만들고 있으며, 그들의 개인용 컴퓨터들에 멀티미디어 데이터를 저장하고 있는 실정이다. 얼마의 시간이 지난 후에 이 많은 양의 멀티미디어 콘텐츠들은 전문적인 도구로서도 해결을 할 수 없는 한계점에 도달하리라 본다. 최근까지 증가하는 멀티미디어 콘텐츠와 연관된 유용한 정보를 빠르고 손쉽게 접근할 수 있는 새로운 방법을 증명하려고 노력을 하고 있다. 수많은 연구 성과로 오늘날 멀티미디어 콘텐츠의 하위수준(Low level) 특징들은 완전히 자동적인 방법으로 추출되고 있다. 이에 반해 몇 되지 않은 수에 제한된 능력을 가지는 주석 도구들은 상위수준(High level) 특징을 기술할 수 있는 의미 정보를 생성하기 위해서 개발되어 오고 있다.

상기 내용에서 두 가지의 핵심을 살펴보면 다음과 같이 기술할 수가 있다. 첫째로 현재 사용 중인 몇몇 도구들은 상위수준의 특징을 자세하게 추출할 수가 없다. 그리고 사용자가 의미정보를 쉽게 생성하기 위해 이용할 수 있는 대화식 의미 도구가 현재로서는 거의 없다는 것이다. 둘째로 사용자에게 제공할 수 있는 내용정보와 사용자들이 기대하고 있는 정보 사이에는 조그마의 차이가 존재한다. 다시 말하면, 하위수준 특징과 상위수준 특징 사이의 차이로 인하여 더 좋은 검색 수행에 있어 큰 장애가 되고 있다는 것이다. 이러한 차이를 "의미적 차이(Semantic Gap)"이라고 하며 아직까지 멀티미디어 콘텐츠 검색에 있어 얼마간의 애로점으로 존재하고 있다[1].

하지만 가장 큰 문제의 장벽은 확장성 및 지능적인 부분과 멀티미디어 어플리케이션을 공동 이용이 가능한 표준, 그리고 이해력, 멀티미디어의 유연한 표준의 부족이다. 따라서 멀티미디어 환경에서 사용자가 원하는 콘텐츠를 검색하고 선택할 수 있는 정보를 제공하는 콘텐츠의 메타데이터와 인덱싱에 대한 필요성이 대두되고 있다. 이에 MPEG-7은 멀티미디어 콘텐츠에 대한 구조 정보와 의미 정보를 기술할 수 있는 국제 표준 규격을 제정함으로써 이러한 메타데이터 정보를 이용하여 원하는 콘텐츠를 검색하고 선택하여 콘텐츠의 효율적인 접근과 사용에 있어서 상호 운용성을 제공하고 있다. 따라서 멀티미디어 콘텐츠에 대한 복잡한 메타데이터를 생성하기 위해서는 MPEG-7 표준을 준수하며 사용하기에 편리한 새로운 메타데이터 생성도구의 필요성이 절실히 요구되고 있는 실정이다. MPEG-7 기반 멀티미디어 기술 구조(MDS : Multimedia Description Schemes)는 멀티미디어 콘텐츠에 대한 정보 표현을 구조화함으로써 사용자가 원하는 멀티미디어 콘텐츠의 특정 부분의 접근을 용이하게 하는 가장 효과적인 도구이며, MPEG-7 오디오 및 비주얼 기술자는 오디오 비주얼 특징에 대한 효과적인 표현 방식을 제공함으로써 내용 기반 오디오 비주얼 색인 및 검색을 가능하게 한다[2][3].

따라서 본 논문에서는 멀티미디어 주석과 검색에 있어 효율성을 제공할 수 있는 의미적인 메타데이터를 생성하기 위한 새로운 방법론을 연구하는 것이 목적이며, 최소한 사용자가 쉽게 이용할 수 있는 새로운 대화식 도구를 개발하는 것이다. 제안하는 도구의 구성은 하위수준의 시각적 특징 값을 위한 MPEG-7 Descriptor 자동생성

부분과 의미적인 메타데이터의 식별을 위한 메타데이터를 위한 메타데이터 부분, 텍스트 기반의 주석(Free Text)을 포함하고 있는 구조기술 부분, 멀티미디어 콘텐츠의 상위수준을 기술할 수 있는 의미기술(Semantic Description) 부분으로 되어 있다. 이러한 모든 내용들은 MPEG-7 표준을 따르며, 결과 생성된 의미적 메타데이터를 이용하여 검색에 있어 효과적인 멀티미디어 검색이 될 수 있도록 도움을 줄 것이다.

본 논문의 구성은 다음과 같이 전개된다. 1장 서론에 이어, 2장에서는 논문의 기본적인 배경 지식으로 MPEG-7 Descriptor와 MDS(Multimedia Description Scheme)에 대하여 살펴본다. 3장에서는 기존 주석기반 검색의 문제점에 대하여 기술하고, 4장에서는 기존 검색에서 가지는 문제점을 보완할 수 있는 의미적인 메타데이터 생성에 대하여 기술한다. 5장에서는 제안된 의미적 메타데이터를 이용한 주석 및 검색에 관련된 시스템을 구현, 소개한다. 마지막으로 6장에서 결론 및 향후 과제로 끝을 맺는다.

2. 배경 지식

2.1 MPEG-7 Visual Descriptor

MPEG-7 비주얼 표준은 이미지나 비디오의 비주얼 특징 정보(Low-level features)를 표준화된 포맷으로 표현하기 위한 기술자에 대한 구문과 의미를 정의한다. 비주얼 특징 정보로는 컬러, 텍스트, 형태, 질감, 질감, 움직임 등을 포함하며 비주얼 각각의 해당 기술자를 이용하여 내용 기반 인덱싱이 가능하다. 이러한 내용 기반 비주얼 특징정보 색인은 자동 특징 추출 및 기술자 값으로의 표현이 자동으로 이루어질 수 있는 틀을 제공하고 있다 [2]. 아래 그림 1은 MPEG-7 표준에서 제공하고 있는 MPEG-7 비주얼 기술자에 대한 구성을 나타낸다.

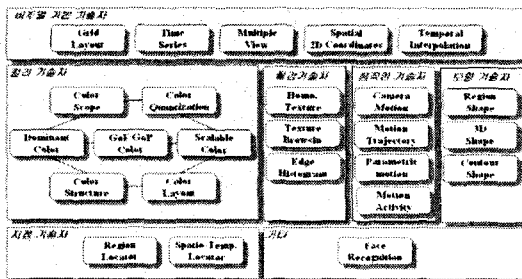


그림 1. MPEG-7 비주얼 기술자의 개요

본 논문에서는 컬러 정보를 색인하기 위한 기술자로서 스케일러블 컬러 기술자(Scalable color descriptor)와 영상의 에지 특성을 나타내기 위한 에지 히스토그램 기술자(Edge Histogram Descriptor)를 메타데이터 도구에 통합 구현하였으며, 스케일러블 컬러 기술자는 Haar Transform으로 변환된 HSV 컬러 공간의 컬러 분포 값을

히스토그램으로 표현하고 히스토그램의 각 빈값에 대한 표현 정확도(precision)를 여러 형태로 표현 할 수 있게 하였다. 에지 히스토그램 기술자는 영상 콘텐츠에 에지 성분에 대한 분포를 표현함으로써 영상의 전체 내용 특징을 전역적으로 표현하기에 적합하다.

2.2 MPEG-7 MDS(Multimedia Description Scheme)

MPEG-7 MDS는 MPEG-7 기술 정의 언어(DDL : Description Definition Language) 형식을 이용하여 멀티미디어 콘텐츠의 구조와 의미 정보를 기술하는 도구이며, 아래 그림 2와 같은 요소들로 구성되어 있다.

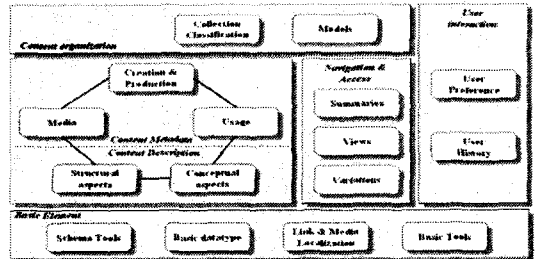


그림 2. MPEG-7 MDS 개요

MDS의 기본적인 바탕을 차지하고 있는 Basic Elements는 상위 수준의 Description tool을 위한 기초적인 블록 역할을 한다. Content Description은 멀티미디어 콘텐츠의 구조적, 의미적 특징 정보에 대하여 기술하며 Structure DS와 Semantics DS로 표현된다. Content Metadata 부분에서는 멀티미디어 콘텐츠와 관련된 정적 데이터인 콘텐츠의 제목, 장르 등의 제작과 포맷, 크기에 대한 미디어 형식, 사용 정보 등의 정보를 표현하며 Creation & Production DS, Media DS 그리고 Usage DS에 의해 기술된다. Navigation & Access tool은 멀티미디어 콘텐츠의 요약, 뷰, 변화에 대하여 정의하여 특정 부분으로의 접근을 용이하게 한다. Content Organization은 멀티미디어 콘텐츠의 컬렉션을 조직화하고 모델링하기 위한 기술 스키마이다. 사용자의 멀티미디어 콘텐츠의 소비에 관련한 사용 기록과 장르, 배우 등의 선호도를 확인하고 그에 적합한 콘텐츠를 선별하여 제공할 수 있는 정보는 User Interaction DS와 User Preference DS에 의해 기술된다[3].

3. 기존 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 개체 간의 관계들로 이루어진 자료의 의미를 쉽게 모델링할 수 있는 장점이 있

다. 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 쉬우며, 확장성이 부족하다고 할 수 있다. 또 검색이 데이터에 부여된 주석에 의존하기 때문에 의미적인 검색이 어렵다는 단점이 있다. 다음은 주석 기반 검색의 한 예를 보여주고 있다. 그림 3의 좌측과 같이 사용자 질의로 "Run Animal"이 주어지고, 우측에 멀티미디어 데이터베이스에서 주석 기반 검색의 결과가 어떻게 되는지를 고려해 보자.

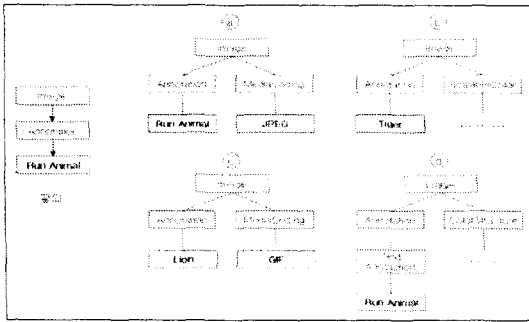


그림 3. 멀티미디어 주석 검색 예

위 결과 분명한 질의 결과는 ㉠이며, ㉡㉢의 Tiger나 Lion은 달리는 동물이지만 이 질의의 결과로는 맞지 않다. 여기에서 사용자 질의는 "Run Animal"이라는 주석 검색만을 요구하기 때문에 실제 달리는 동물들은 검색이 안 되는 문제점을 가지고 있다. 이러한 문제점의 해결을 위해서 유사한 데이터 추출을 위하여 추론적인 기법이 필요한데, 여기서 개념적인 동물영역의 온톨로지 구성으로 명확히 ㉡㉢도 질의 답변으로 추출 될 수가 있을 것이다. 그림 4는 위 문제점 해결을 위한 동물도메인 온톨로지를 보여주고 있다.

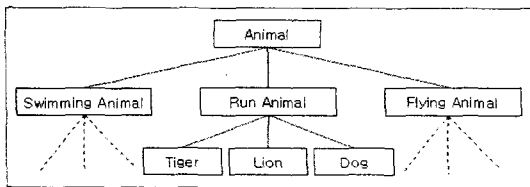


그림 4. 동물도메인 온톨로지

하지만 이러한 동물도메인 온톨로지의 구성에 의한 검색 또한 제한된 범위 내에서의 정확성은 가질 수는 있으나 의미 확장 검색에서는 단점을 보이고 있다. 실제로 "박지성선수와 토고 선수가 볼을 다루고 있다."라는 멀티미디어 데이터에서 콘텐츠를 보고 "박지성", "토고선수", "볼"과 같은 부분은 주석처리 가능하지만 개최지는 "독일", 개최년도는 "2006년"과 같은 내포하고 있는 의미 정보에 대해서는 주석에 대한 어려움이 있다[4].

4. 의미적 메타데이터의 생성

본 논문에서 제시하는 의미적 메타데이터는 MPEG-7 기반으로 이미지 또는 비디오 세그먼트와 같은 멀티미디어의 특정 인스턴스에 의미(Semantic)를 추가하여 주석할 수 있으며, 프레임워크는 멀티미디어 특징을 표현하는 기술자(Descriptor)와 기술체계(Description Scheme)로 구성된다. 보다 자세하게 멀티미디어 콘텐츠의 의미적 주석을 위해서 세 가지의 MPEG-7 기술(Description)로써 나타낼 수가 있으며 아래와 같다.

- 구조 기술(Structural Description)
주어진 멀티미디어 콘텐츠를 형식에 구조적으로 기술한 것을 말한다.(예로, 구조적 텍스트 기술, 시각적 기술자, 메타데이터의 메타데이터 등)
- 분류 기술(Classification Description)
주어진 어떠한 영역에서 멀티미디어 콘텐츠를 기술하기에 적당한 특징 키워드들의 집합을 의미한다.(예로, Object, Event, Person, Place 등)
- 의미 기술(Semantic Description)
주어진 추상적 수준에서 의미적 엔티티들 간의 관계 속성들로 구성된 집합을 의미한다.(예로, UserOf, LocationOf, agentOf 등)

아래 그림 5는 상기에서 기술한 기술들을 설명하기 위한 예시이다. 멀티미디어 콘텐츠에서 분해를 통한 의미적 엔티티(세그먼트)는 타원의 도형(node)이며, 엔티티들 서로 간의 관계는 선분(edge)으로 표현된다. 선분의 표현 값은 엔티티들 간의 관계가 가지는 의미 속성 값을 말한다. 엔티티와 의미 속성 값들은 분류 기술로 정의하여 관리하게 된다. 구조 기술에서 제시한 값들과 분류 기술 및 의미 기술에 의한 표현 값들의 조합으로 의미적인 메타데이터 주석이 생성되는 것이다.

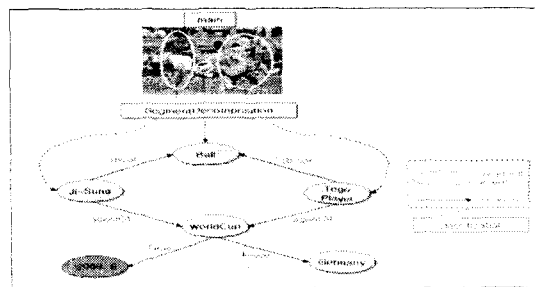


그림 5. 멀티미디어 콘텐츠 분해와 의미기술

그림 5는 '2006년 독일 월드컵에서 한국과 토고 축구 경기 중에 박지성(Ji-Sung) 선수와 토고 선수(Togo Player) 간에 볼을 다루고 있는 멀티미디어 콘텐츠를 구조 및 의미 엔티티들로 분해하여 서로 간의 의미 속성을 부여한 후 구성된 전체적인 의미적 메타데이터 도식화한 것을 보여주고 있다. 이러한 사람에게 의한 의미 기술은 수동적인 작업으로 이루어지며, 그래픽 작업으로 가

능하다. 그 결과 모든 내용은 MPEG-7 기반의 XML 문서로 산출된다.

5. 의미적 메타데이터를 이용한 주석 및 검색

본 논문에서 제안하는 시스템은 메타데이터를 이용한 의미적 주석 시스템(m-SAS)과 의미적 검색 시스템(m-SRS)으로 구성이 된다. 주석 시스템은 하위수준의 자동적 주석과 상위수준의 인간기반의 지능적 주석을 혼합하여 보다 검색에서 효율성을 가질 수 있도록 하였으며, 주석함에 있어 의미적인 주석 부분은 쉽게 활용 가능하도록 그래프 기반의 모델링 개념으로 설계를 하였다. 이 결과 생성된 의미적인 메타데이터 주석은 멀티미디어 검색에 있어 한층 더 자세한 검색 결과를 보이게 될 것이다.

5.1 의미적인 주석시스템(m-SAS)

m-SAS(metadata-Semantic Annotation System)의 전체적인 프레임워크를 살펴보면 그림 6과 같이 도식화할 수 있다. 보는 바와 같이 크게 두 단계로 나누어서 설명을 할 수가 있다. 사용자 인터페이스를 통한 생성 결과인 구조기술과 분류 기술과 혼합된 의미기술이다. 분류 기술은 첫 번째 단계에서 생성이 되며, 사용자가 작업에 따라 새로이 추가하여 최신의 것으로 갱신할 수 있다.

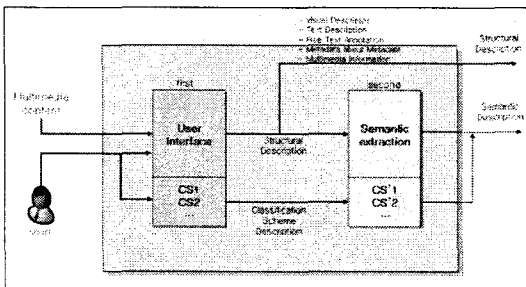


그림 6. m-SAS 전체적인 프레임워크

• 구조 기술 처리

구조기술에서는 Visual Descriptor, Text Description, Free Text Annotation, Metadata about Metadata, Multimedia Information 등으로 구성이 되며, Visual Descriptor에는 Edge-Histogram, DominantColor, ColorLayout, ScalableColor 등으로 구성된다. 아래 그림 7은 주석 시스템에서 Visual Descriptor 부분을 보이고 있다.

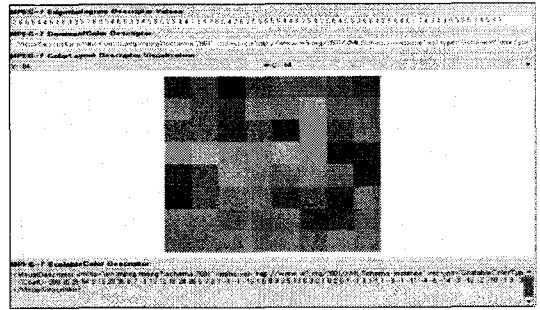


그림 7. m-SAS의 Visual Descriptor 부분

Visual Descriptor의 값을 XML 문서로 나타내면 아래 그림 8과 같다.

```
<VisualDescriptor xsi:type="ColorLayoutType">
  <VDCoeff>30</VDCoeff>
  <CbDCCoeff>15</CbDCCoeff>
  <CrDCCoeff>35</CrDCCoeff>
  <YACCoeff63>17 15 14 15 13 17 16 14 13 17 15 18 19 16 15 16 17 15 17 15 13 18 14 15 18 16 16 14 16 16 16
  <CbACCoeff63>18 22 16 17 19 12 15 12 20 16 17 12 15 18 17 17 20 17 15 17 15 17 16 14 15 16 16 14
  <CrACCoeff63>18 21 11 17 16 12 17 15 13 14 15 15 13 14 15 17 14 13 15 17 15 16 16 15 16 17 16 15 16
</VisualDescriptor>
<VisualDescriptor xsi:type="ScalableColorType" numOfBitPlanesDiscarded="0" numOfCoeff="255">
  <Coeff>208 70 25 54 3 13 29 35 6 7 -1 13 12 16 24 26 0 7 3 7 -1 -1 -10 1 5 8 3 35 13 8 3 0 1 0 2 0 1 -1 1
</VisualDescriptor>
<VisualDescriptor xsi:type="EdgeHistogramType">
  <BinCounts>2 6 5 4 4 5 4 6 3 3 5 7 6 3 5 4 6 5 3 3 4 5 6 6 2 5 4 4 7 1 4 7 6 6 4 2 6 7 5 5 5 5 5 4 4 4 7 5 4
</VisualDescriptor>
<VisualDescriptor xsi:type="DominantColorType">
  <SpatialCoherency>D</SpatialCoherency>
</VisualDescriptor>
```

그림 8. 추출된 Visual Descriptor의 XML 문서

• 분류기술과 의미기술 처리

구조기술에서 추출된 엔티티와 의미적인 엔티티로 분류하여 이들 엔티티 사이의 관계 설정과 의미적인 속성 값을 부여함으로써 멀티미디어의 의미적인 주석이 이루어진다. 이러한 엔티티와 의미적 관계 속성 값은 분류기술로서 정리된다. 그림 9는 멀티미디어의 의미적인 메타 작업을 위한 작업 화면을 보여주고 있으며, 그림 10은 작업에 의한 저장된 결과 XML 문서이다.

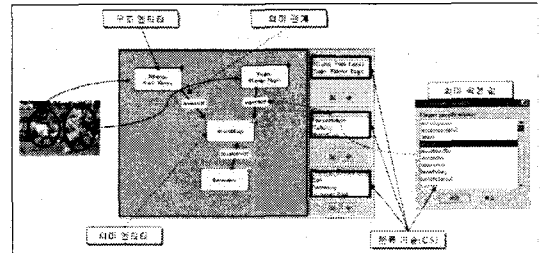


그림 9. 의미적인 메타데이터 생성 작업 화면

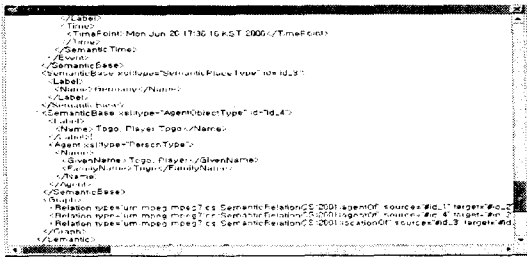


그림 10. 의미적인 주석의 XML 결과 문서

6. 결론 및 향후 연구

최근까지 멀티미디어의 이용과 멀티미디어 접근을 위한 기술이 많이 증가하고 있지만 검색엔진과 같은 실용 시스템에서 멀티미디어에 대한 유용한 정보 추출과 정보의 응용은 여전히 문제로 있다. 가장 중요한 장벽은 확장성과 지능적, 멀티미디어 어플리케이션을 공동 이용이 가능한 표준, 이해력, 멀티미디어의 유연한 표현의 부족이다. 이러한 문제점을 해결하기 위해서 의미적인 주석을 위한 메타데이터 생성하는 연구가 진행되고 있으며 확장성을 가지는 사용자 인터페이스를 고려하고 있는 실정이다. 본 논문에서는 멀티미디어 데이터의 의미적인 주석에 도움을 줄 수 있는 의미적인 메타데이터를 생성하는 확장 가능한 새로운 방법을 제시하였다.

본 논문에서 제안하는 의미적인 주석은 기존의 Free Text, 온도로지, 의미 분류 항목을 이용한 주석에서 표현을 하지 못하는 부분을 확장하여 의미적인 표현을 할 수가 있다. 이 목적을 위하여 자료구조의 그래프 표현으로 볼 때에 의미 메타데이터를 위한 의미 엔티티와 구조 엔티티는 노드로써 엔티티들 간의 관계를 위한 간선으로 연결을 하고 관계 값으로 의미 속성 값을 항목으로 설정하여 의미적인 관계 값을 추가를 하였다. 이렇게 생성되는 의미적인 메타데이터 값과 하위수준의 자동 추출 값과 통합되어 MPEG-7 기반의 표준을 따르는 XML 문서로서 저장되고, 이러한 의미적인 주석 값은 다양한 검색 방법으로 멀티미디어 데이터를 인간 측면에 가까운 정도의 검색 결과를 이룰 수가 있을 것이다. 이러한 멀티미디어 주석과 검색을 위한 사용자 인터페이스로 m-SAS와 m-SRS를 소개하였다.

향후 과제로 주석의 사용자 인터페이스에 의해 제공되는 세 개의 기술로 다른 멀티미디어 정보를 추정할 수 있는지 조사 연구할 계획이다. 또한 세 개의 기술로서 의미적인 메타데이터를 생성하는 기존의 과정을 축소하고, 검색에 있어 질의에 이용할 엔티티들의 항목 리스트 효율적으로 관리하는 방법 또한 연구 중에 있다. 마지막으로 이러한 그래프를 이용한 의미 메타데이터 생성 모델링은 다른 응용분야 연구에도 다양하게 적용될 것으로 기대가 된다.

참고 문헌

- [1] Tshuan Chen. Low-Level Features to High-Level Semantics : Are We Bridging the Gap ? In EWIMT, London, UK, November 2004.
- [2] ISO/IEC 15938-3, Information Technology Multimedia Content Description Interface : Visual.
- [3] ISO/IEC 15938-5, Information Technology Multimedia Content Description Interface : Conformance Test.
- [4] 안형근 외 2명, "트리기반 멀티미디어 데이터의 의미적 검색을 위한 프레임워크 설계", 한국정보과학회 한국컴퓨터종합학술대회, Vol.32, No.1,

5.2 의미적인 검색 시스템(m-SRS)

m-SRS(metadata-Semantic Retrieval System)의 전체적인 프레임워크를 살펴보면 그림 11과 같이 도식화하여 나타낼 수가 있다. m-SRS 시스템에서는 검색 방법으로 키워드 질의, 유사도 질의, 의미 질의, XPath, 그래프 질의 등을 제공하고 있다. 특히 의미 질의 및 그래프 질의 부분은 m-SAS에서 분류기술로 정의된 항목을 재사용하여 검색에 활용을 할 수가 있다. 이러한 재사용은 검색에 있어 질의의 표준화를 암시한다.

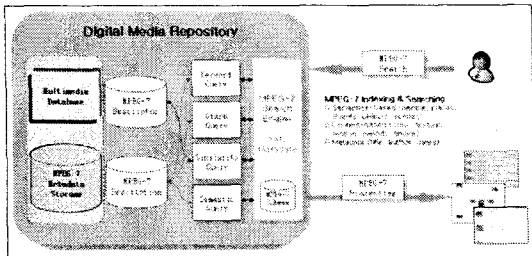


그림 11. m-SRS 전체적인 프레임워크

멀티미디어 검색을 위하여 질의를 월드컴의 개최자가 독일이고 선수들 중 박지성 선수 포함하는 멀티미디어 데이터를 검색하는 경우의 예를 보도록 하겠다. 본 논문에서는 m-SAS 시스템에서 정의된 분류기술 항목을 사용하여 그래프를 활용한 질의를 예로 들것이다. 아래 그림 12는 그래프를 이용한 질의검색과 결과를 보여주고 있다.

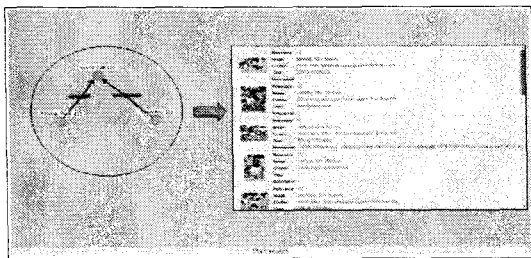


그림 12. 그래프를 이용한 질의검색과 결과

- pp.133-135, 2005. 7.
- [5] M. Naphade, C.-Y. Lin, J. R. Smith, B. L. Tseng, and S. Basu. Learning to annotation video database. In SPIE Electronic Imaging 2002 - Storage and Retrieval for Media Database, San Jose, CA, USA, January 2002.
 - [6] B. L. Tseng, C.-Y. Lin, and J. R. Smith. Video personalization and summarization system. In SPIE Photonics East 2002 - Internet Multimedia Management System, Boston, MA, USA, August 2002.
 - [7] B. L. Tseng, C.-Y. Lin, and J. R. Smith. Video summarization and personalization for pervasive mobile devices. In SPIE Electronic Imaging 2002 - Storage and Retrieval for Media Database, San Jose, CA, USA, January 2002.
 - [8] J.M. Martinez. Overview of MPEG-7 Description Tools, Part 2. IEEE MultiMedia, pages 83-93, 2002.
 - [9] A. Benitez, H. Rising, C. Jorgensen, R. Leonardi, A. Bugatti, K. Hasida, R. Mehrotra, A. M. Tekalp, A. Ekin, and T. Walker. Semantic of multimedia in MPEG-7. In Processing (ICIP-2002), Rochester, New York, USA, September 2002.
 - [10] Eidenberger, Horst, Statistical analysis of MPEG-7 image description, ACM Multimedia System journal, Springer, 2/2004.