

# 의미적 멀티미디어 메타데이터 생성을 위한 MPEG-7 기술기반 주석도구의 개발

안 형 근<sup>†</sup> · 고 재 진<sup>††</sup>

## 요 약

최근 멀티미디어 데이터의 급격한 양적 팽창은 원하는 데이터를 빠르고 정확하게 검색해야 한다는 새로운 과제를 안겨주었다. 이러한 효율적 검색을 위해서 가장 중요한 기반이 되는 것이 바로 멀티미디어 데이터의 적절한 표현이다. 국제 표준으로 제정된 MPEG-7은 바로 이러한 이유로 멀티미디어 데이터의 표현에 대한 표준화를 다루고 있다.

본 논문에서 메타데이터 생성을 위한 새로운 접근법을 제안한다. 사용자는 주어진 멀티미디어 콘텐츠를 작은 단위들로 분해를 하고, 분해된 단위들에 시간, 위치 등과 같은 추가적인 기본정보뿐만 아니라 MPEG-7 표준을 따르는 사건, 관계 등과 같은 분류정보를 쉽게 주석할 수 있다. 이 주석의 목적은 자동적으로 의미기술을 만들기 위한 것이고, 이 의미기술에서 노드들은 사건들이고, 링크는 그들 사이의 관계인 하나의 그래프이다. 마지막으로 제안된 기법을 기반으로 의미기술을 위한 주석도구(SMAT)를 구현하였고, 실제 실험을 통하여 성능을 평가하였다. 최종적으로, 제안 도구는 재사용성과 확장성의 두 개의 중요한 타당의 특징이 있다고 말할 수 있다.

키워드 : MPEG-7 기술, 의미 주석, 메타데이터, SMAT

## Development of MPEG-7 Description-based Annotation Tool for Production of Semantic Multimedia Metadata

Hyounggeun An<sup>†</sup> · Jawjin Koh<sup>††</sup>

### ABSTRACT

Recently, an increasing in quantity of multimedia data have brought a new problem that expected data should be retrieved fast and exactly. The adequate representation for the multimedia data is the key element for efficient retrieval. For this reason, MPEG-7 standard was established for description of multimedia data.

In this paper, we propose a new approach to metadata production. The user can decompose a given content into units and easily annotate each unit by adding basic information such as time, place, etc. as well as classification information such as event, relationship, etc. according to the MPEG-7 standard. The objective is to build automatically a pure semantic description; the nodes are the events and the links are the graphs which describe the relationships among the events. Finally, we have implemented an annotation tool(SMAT) for semantic description based on proposed technique and assess some of the experiment results. In conclusion, we can say that the proposed annotation tool is characterized by two important proprieties : reusability and extendibility.

Key Words : MPEG-7 Description, Semantic Annotation, Metadata, SMAT

### 1. 서 론

최근 디지털 카메라 또는 개인용 미디어 단말기처럼, 멀티미디어를 생성할 수 있는 장비의 증가로 인하여 디지털 데이터의 양이 기하급수적으로 증가하고 있다. 대부분 데이터들은 오디오, 이미지, 비디오와 같은 멀티미디어 데이터들을 포함하고 있고[1, 2], 이러한 멀티미디어 데이터와 연관된 유용

한 정보를 빠르고 손쉽게 접근할 수 있는 새로운 검색 방법을 증명하려고 노력하고 있으며, 전 세계적으로 활발한 연구의 진행이 이루어지고 있는 상황이며, 효율적인 검색을 위하여 반드시 이루어져야 할 기반적 요소가 바로 데이터의 적절한 표현이라 하겠다. 이러한 멀티미디어의 다양한 정보를 조직적으로 표현하기 위한 정의의 언어와 스키마가 필요함에 따라 MPEG(Motion Picture Expert Group)은 멀티미디어 정보를 기술하고 풍부한 도구들을 제공하는 MPEG-7을 발표하였다[3]. MPEG-7은 그 정식 명칭인 “멀티미디어 콘텐츠 표현에 관한 인터페이스”에서 알 수 있듯이 검색을 위한 콘텐츠 표현의 규격을 만드는데 그 목적을 두고 있으며[4], 또한 상

※ 본 연구는 2005학년도 울산대학교 학술연구 지원으로 수행되었음.

† 정 회 원 : 울산대학교 컴퓨터정보통신공학부 박사과정 수료

†† 정 회 원 : 울산대학교 컴퓨터정보통신공학부 교수(교신저자)

논문접수 : 2006년 7월 2일, 심사완료 : 2007년 1월 8일

호호환성의 문제를 해결하기 위한 멀티미디어 데이터의 메타데이터 표현에 대한 표준이기도 하다[7, 8].

MPEG-7에 기반하여 수행된 수많은 연구에도 불구하고 저급수준(low level) 특징들은 완전히 자동적인 방법으로 주석도구들에 의해서 추출되어 주석되고 있는 반면에, 고급수준(high level) 특징으로 생각되는 의미적인 정보들을 생성하기 위하여 개발된 주석도구들은 그 수가 얼마 되지 않으며, 표현에 있어 비정형적 텍스트(free Text)로 제한되어 있다. 이러한 내용에서 볼 때 지금까지 고급수준의 특징들을 자유롭게 표현하여 생성할 수 있는 도구들이 없으며, 사용자가 의미정보를 쉽게 생성하기 위해 이용할 수 있는 대화식 의미도구 또한 현재로는 극소수라는 것이다. 이러한 고급수준인 의미정보 표현의 부족함으로 사용자에게 제공하는 정보와 기대하고 있는 정보 사이에는 차이가 발생하는데, 이러한 차이를 “의미적 차이(semantic gap)”이라고 하며 아직까지 멀티미디어 콘텐츠 검색에 있어 존재하고 있다[5]. MPEG-7에서는 이러한 문제 해결을 목적으로 MPEG-7 기술구조(Multimedia Description Scheme : MDS)를 정의하고 있다[6]. MPEG-7 기술구조는 멀티미디어 생성에 관한 정보(제목, 장르, 제작자 정보 등), 멀티미디어의 고급수준 정보(육하원칙에 따른 사건 설명), 저급수준 정보(색상, 질감, 움직임 등) 등을 포함하여 멀티미디어를 표현하는 문서의 구조를 정의하며, XML 형태의 문서로서 다른 기술구조를 이용해서 새로운 기술구조를 만들 수 있다.

따라서, 본 논문에서는 멀티미디어 데이터의 의미적 메타데이터 생성에 직접적인 관련이 있는 MPEG-7 기술구조의 의미기술(Semantic Description)을 중심으로 분석하여 정의하고, 이를 기반으로 의미적 멀티미디어 메타데이터 생성을 위한 MPEG-7 기반의 대화식 주석도구 SMAT(Semantic Metadata Annotation Tool)를 구현한다. SMAT 도구는 의미적인 멀티미디어 메타데이터 생성을 위하여 기존의 비정형적 텍스트 메타데이터와는 다른 의미 그래프 기반의 메타데이터 생성이다. 이 방법을 위하여 사용자는 주어진 멀티미디어 데이터가 가지는 의미정보를 구조기술(Structural Description)에 기반한 수동의 입력된 텍스트 값을 의미변형을 통하여 의미적 메타데이터 생성에 필요한 구성요소 별로 분류기술(classification description)하고, SMAT 도구의 구성요소로서 저장하게 된다. SMAT 도구는 이러한 과정에서 발생된 구성요소를 그래프의 노드(node)와 간선(edge)으로 표현하여 사용자에게 구성요소의 드로잉(drawing) 작업으로서 메타데이터를 쉽게 생성할 수 있는 사용자 인터페이스이다. 본 논문에서 제안하는 SMAT 도구는 기존의 메타데이터 생성도구와는 다른 재사용성(Reusability), 확장성(Extendibility)과 추가적으로 사용자 편의성(user convenience)을 고려하여 개발되었다.

이 후 본 논문의 구성은 다음과 같이 전개된다. 1장 서론에 이어 2장에서는 관련기술로 MPEG-7 표준 기술에 대한 개요와 관련연구로 기존 주석 도구들에 대하여 살펴본다. 3

기술의 결과를 위한 구조기술과 분류기술에 대하여 서술한다. 4장에서는 앞 장에서 제시한 정의와 분석을 기반으로 주석도구(SMAT)를 구현하고, 이미지와 비디오 데이터에 대한 적용 결과를 제시한다. 또한 기존의 도구와의 기능적인 측면으로 비교평가를 내린다. 마지막으로 5장에서는 결론을 통한 향후 연구방향을 살펴보고 글을 맺는다.

## 2. 관련 연구

이미지나 비디오와 같은 멀티미디어 데이터는 다양하고 방대한 양의 정보를 포함하고 있어 효율적인 검색을 위해서는 데이터를 기술하는 구조적이고 체계화된 형태의 메타데이터가 요구된다. 멀티미디어 데이터의 보편화로 빈번히 사용되는 관계로 많은 표준과 통합이 이루어졌다. 하지만 멀티미디어 데이터가 내포하고 있는 다양하고 복잡한 의미정보를 효율적으로 처리하는데 제약점이 많다. 본 장에서는 멀티미디어 메타데이터 생성과 관련된 기존 주석도구에 대하여 살펴보고, 그에 따른 관련기술에 대하여 설명을 한다.

### 2.1 MPEG-7 표준 기술의 개요

주어진 멀티미디어 데이터의 의미적 메타데이터 표현으로 널리 사용될 수 있는 표준으로 MPEG-7을 들 수 있다[7, 8]. MPEG-7은 다양한 응용을 위하여 제정된 표준으로써, 오디오 비주얼 데이터를 기술하는데 유연하고 확장 가능한 프레임워크를 제공한다. MPEG-7은 기술자(Descriptor : D), 기술구조(Description Scheme : DS), 기술정의언어(Description Definition Language : DDL) 및 기술을 인코딩하기 위한 방법을 표준화하며, 또한 MPEG-7 멀티미디어 기술구조(Multimedia Description Scheme : MDS)에서 제공하는 메타데이터를 카테고리별로 총 8개의 파트로 구성하고 있다. 따라서, MPEG-7은 멀티미디어 데이터와 관련된 모든 메타데이터를 구조적으로 기술할 수 있으며, 의미적인 정보를 기술하기 위한 규약과 형식을 제공한다.

위 정의를 기반으로 본 논문에서는 모든 의미적인 메타데이터를 MPEG-7 기술로써 표현되어질 수 있다. 많은 종류의 메타데이터는 주어진 멀티미디어 데이터에 연관시킬 수가 있으며, 마찬가지로 MPEG-7 기술의 많은 형식들을 사용한다. 이러한 많은 형식들 중에 의미적 메타데이터를 구성하는 기술은 세 가지 기술을 응용 활용한다[9]. 구조기술, 의미기술, 분류기술 등이 이에 속한다.

- 구조기술(Structural Description)

주어진 멀티미디어 데이터의 구조적 계층 정보를 말한다. 예를 들어 멀티미디어 데이터의 관련된 특징(feature)을 가지는 샷(shot)의 시간적인 분해나 객체들의 분해를 말한다.

- 분류기술(Classification Description)

멀티미디어 데이터를 적절하게 기술하기 위한 특정 도메인 영역에 있는 의미적 엔티티들의 집합을 말한다. Classification-Scheme(CS)으로 정의된다.

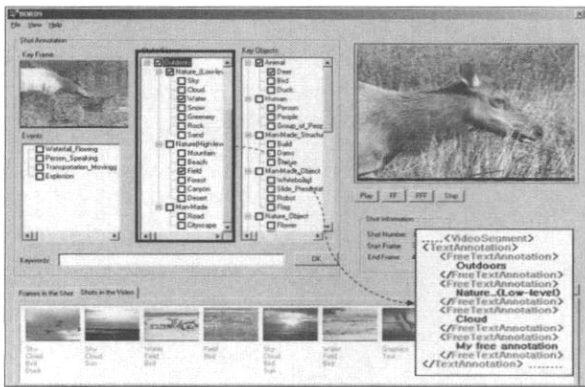
• 의미기술(Semantic Description)

의미적 엔티티들 서로 간의 관계 집합으로 구성 된 것을 말한다.

MPEG-7기반의 구조기술 및 분류기술은 의미기술을 생성하기 위하여 조합되어지며, 후에 의미적 메타데이터로 구성되고 주석되어진다. 또한 의미기술 생성을 위한 작업은 도구를 통하여 그래프 형태의 드로잉 작업으로 가능해진다.

2.2 기존 멀티미디어 주석도구

MPEG-7 기반의 의미주석을 위해 개발된 가장 인기 있는 도구의 하나가 IBM에서 개발한 VideoAnnEx(그림 1)이다 [11, 12]. 사용자는 입력되는 멀티미디어 데이터를 연속적으로 선택할 수 있으며 자유롭게 분해를 할 수 있고, 각 세그먼트에 주석을 달수가 있다[13-15]. 이 주석에는 세그먼트에 속하는 분류항목들의 선택주석과 비정형적 텍스트 주석으로 작성된 것들을 말한다. 분류항목들은 사용자가 각 세그먼트와 특별한 의미(“Outdoors”, “Indoors” 등)를 동일하게 하며, 이전 생성된 항목을 사용할 수도 있고 새로운 항목을 생성하여 사용할 수도 있다. 또한 항목들의 분류는 서로 다른 수준들로 나타내기 위해서 계층적인 구조의 카테고리 식으로 특징을 기술하고 있다. 예를 들어, 구름(cloud)을 포함하고 있는 하늘을 표현하는 세그먼트는 “Outdoors” 항목에 속하며, 저급 수준의 “Nature”의 “Cloud”를 선택하게 된다. 사용자가 각 세그먼트에 주석할 때에 기술의 모든 구성요소는 MPEG-7 표준을 따르며, MPEG-7 표준의 FreeTextAnnotations 태그를 이용하는 메타데이터의 XML 문서로 표현된다[9].



(그림 1) VideoAnnEx 도구와 XML 문서

이 도구는 최종적으로 이러한 분류정보를 포함하는 XML 기반의 메타데이터로 저장이 된다. 하지만 이러한 분류정보가 MPEG-7 표준에서 제공하는 모든 것을 이용하지 못하고 있다. 특히 분류항목들과 주석된 구성요소들의 기술 사이에는 어떠한 개념적인 관계연결도 할 수 없다. 또 하나의 문제점으로 도구에서 제공되는 분류정보가 인터페이스 필요조건만 만족을 하고 있다.

Ricoh Movie Tool[16]은 MPEG-7 표준을 따르는 비디오

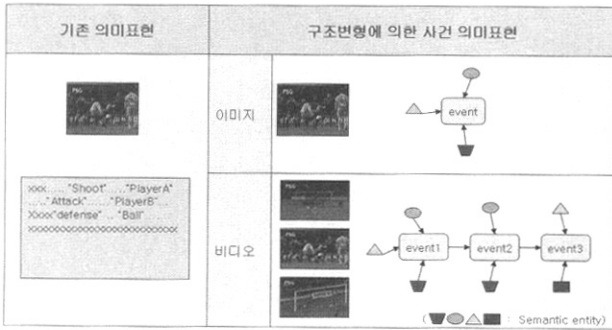
내용 기술들을 대화적으로 생성하기 위한 도구이다. 도구는 비디오를 로딩하는 동안 구조적 기반의 MPEG-7 기술들을 대화적으로 생성을 한다. 빌트인 편집 기능들은 MPEG-7 기술을 수정할 수 있다. 적당한 MPEG-7 태그들을 선택하기 위해, 비주얼 도표는 후보 태그와 결합하는데 대화적인 구조 편집동안 사용자들을 편리하게 한다. 비주얼 구조와 MPEG-7 기술의 관계는 시각화하게 된다. 모든 MPEG-7 기술은 정의 스키마와 일치시켜 확인할 수가 있다. 마지막으로 정의된 메타데이터는 비정형적 텍스트 주석에 의한 것이다. Ricoh Movie Tool의 문제점으로는 추상의 고급수준 메타데이터를 가진 의미기술의 표현은 가능하지가 않다는 것이다. Ricoh Movie Tool과 유사한 방법으로 MPEG-7 Annotation Tool[17]은 EPFL 연구소에서 개발되었다. 이 도구는 각 세그먼트에 사용자가 주석할 수 있는 계층적인 시간 분류를 제공하고 있으며, 비정형적 텍스트의 형태로 메타데이터 값이 저장된다.

위와 같이 모든 도구들은 전체적인 주석과 구조적 분류를 처리하는 동안 의미 엔티티와 그들 사이의 관계를 사용자에게 함축적으로 제공하지 못하고 있다. 실제로 사용자들은 세그먼트의 분해를 통해 알지 못하는 비정형적 텍스트 메타데이터 또는 구조적 분류 카테고리과 같은 정보를 제공받고 있다. 전체적인 내용으로 기존 주석도구에서 발생하는 문제점을 다음과 같이 정리할 수 있다.

- 편의성 : 메타데이터 생성을 위한 사용자 인터페이스를 사용하기에 어렵다.
- 제한성 : 구조적 텍스트 메타데이터 및 비정형적 텍스트 메타데이터는 의미정보 표현범위에 제한이 있기 때문에, 이후 검색에서 일치되는 의미질의 텍스트가 없으면 검색을 하지 못하는 문제점이 있다.
- 확장성 : 작업 중 필요시에 추가적인 의미적 엔티티 표현이 어렵다.
- 일관성 : 생성된 전체적인 메타데이터에 대한 의미주석 표현의 일관성이 떨어진다.
- 재사용성 : 주석과 검색에서 있어 기존 메타데이터 생성에 사용된 의미 엔티티를 제공 받지 못하여, 주석 때마다 새로운 텍스트의 키워드를 입력하여야 한다.
- 비디오 데이터의 경우 세그먼트들 서로 간에 어떠한 개념 관계를 성립시키지 못한다.

2.3 멀티미디어 데이터의 구조화

이미지 데이터의 경우 색상, 질감, 형태, 위치 등 시각적인 정보를 히스토그램이나 윤곽선 검출(edge detection) 알고리즘 등을 사용하거나 사람의 주관적인 판단에 따라 객체들의 특성을 비정형적 텍스트로 기술하는 방식으로 메타데이터를 구축하였다. 히스토그램이나 윤곽선 검출을 이용한 내용기반 검색은 비용이 많이 들고 정확도가 낮은 단점이 있으며, 텍스트로 기술하는 방식의 주석기반 검색은 데이터



(그림 2) 멀티미디어 데이터의 구조화

가 커지고 종류가 다양해짐에 따라 주석 처리에 소요되는 시간이 증가하고 사람의 주관적인 판단에 따라 기술됨으로써 기술의 객관성에 문제가 있다.

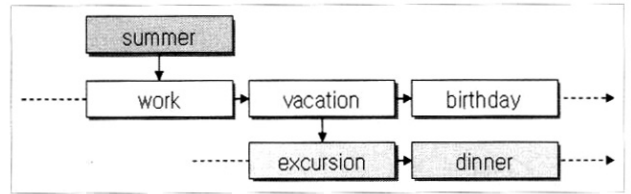
비디오 데이터는 프레임, 샷 장면 등의 계층적 구조로 표현하며, 이를 위해서 비디오 분석을 통해 그 경계 검출하는 것이 필요하다. 많은 연구에서 비디오 프레임에 대한 색상 히스토그램과 대표 색상을 기반으로 장면과 샷의 경계 검출이 이루어져 왔다. 하지만 비디오 분석은 여전히 어렵고 발전시켜야 할 과제로 남아 있다. 비디오 데이터에 대해 쉽고 빠르게 의미 정보를 해석하는 것은 사람이며, 사람에 의한 주석처리는 필수적이라 할 수 있다.

위 내용으로 기존의 멀티미디어 데이터 분석은 사용자가 멀티미디어 데이터를 직접 분석함으로써 적절한 내용을 텍스트 형태로 입력하는 방식이 대부분이다[10]. 이와 같은 방법은 좋은 성능을 가질 수 있는 시스템을 구성할 수 있지만, 사용자에게 많은 시간과 노력을 요구하며, 데이터에 대한 일관되고 정확한 표현이 어렵다. 따라서 이러한 주석기반이 멀티미디어 데이터에 대한 모든 정보를 제공할 수 없더라도 객체들 간의 관계를 일관성 있게 해석할 수 있는 모델을 정립하여 의미정보를 자유롭게 표현할 수 있는 도구가 필요하다.

이러한 도구를 위하여 멀티미디어 데이터를 작은 단위로 구조화하여 각 단위는 사건(event)으로 개념화 한다. 따라서 이미지 데이터는 자체가 하나의 데이터이기 때문에 하나의 사건으로, 비디오 데이터의 경우 여러 개의 세그먼트로 표현되어 다수 개의 사건들로 대응되도록 의미변형을 한다. 결론적으로 멀티미디어 데이터들은 사건의 집합체라고 할 수 있으며, 시스템 측면으로 이러한 사건을 의미 엔티티로 표현을 한다. 더불어 사건 의미 엔티티에는 또 다른 의미 엔티티 정보를 추가적으로 덧붙여 표현하게 된다, (그림 2)는 기존의 멀티미디어 데이터의 비정형적 텍스트 메타데이터와 구조화 변형 후의 사건 의미 엔티티의 표현을 보여주고 있다.

### 3. 의미 메타데이터를 위한 기술들의 생성

본 장의 기술적인 이해를 돕기 위하여 (그림 3)과 같이 “summer” 예시 멀티미디어 데이터를 제시한다. “summer” 멀티미디어 데이터는 시간적 세그먼트 분해에 의해서 “work”, “vacation”, “birthday”와 같이 세 개의 세그먼트가 생성되



(그림 3) “summer” 예시 멀티미디어 데이터 구조

며, 다시 “vacation”은 “excursion”, “dinner”와 같이 두 개의 세그먼트로 되는 것을 볼 수가 있다. 본 장에서는 제시된 “summer” 예시 멀티미디어 데이터를 통하여 MPEG-7 기술인 구조기술, 분류기술, 의미기술에 대해 정의하고자 한다.

#### 3.1 사건을 위한 구조적 분해

MPEG-7 기술은 n 개의 서로 다른 “의미” 수준을 이용하여 멀티미디어 데이터의 구조를 기술할 수 있다. (그림 3)과 같이 “summer” 멀티미디어 데이터를 두 수준으로 기술한다면, 첫 번째 수준에서 “work”, “vacation”, “birthday” 등의 시간적 분해로 세그먼트를 구성할 수 있다(단, 시간적으로 “work” 후에 “vacation”이 다음으로 “birthday”가 주어진다). 두 번째 수준은 이전의 수준에서 정의된 각 세그먼트를 더 분해 할 수가 있다. 예를 들어 “vacation”은 “excursion”과 “dinner” 등으로 나눌 수가 있다(“excursion”이 선행하고 후에 “dinner”가 온다). 이러한 분해는 함축적 의미를 각 세그먼트에 관련시킨 계층구조를 생성한다. 세그먼트는 시각적인 원본 정보를 보이며 의미정보는 포함하고 있지 않다. 따라서 각 세그먼트에 맞는 의미에 따라 해당 사건들과의 사이에 일대일의 관계를 성립시킨다.

생성된 구조적 표현을 위한 핵심으로는 개별 사용자 선택에 따른 각 사건의 분류이다. 이를 위해서 (그림 4)와 같이 MPEG-7기반의 이용할 수 있는 정의된 분류기술의 CS를 사용할 수가 있다. 예로서 “vacation”, “work”, “birthday” 등이 속한다. 또는 기존에 정의된 CS로는 모든 사건을 표현할 수가 없기 때문에 요구조건에 맞는 적당한 새로운 CS를 만들 수도 있다. 예로서 예시의 멀티미디어 사건정보로 보이지 않은 특별한 상황과 주관적 관점의 “leisure”와 같은 사건을 의미한다. 이런 방법에서 사용자는 함축적인 의미의 CS를 생성할 수 있고, 직접적으로 만든 것을 사용할 수도 있다. 도구로 생각하면 사용자가 CS 데이터베이스에서 선택 항목으로 이용할 수 있는 사건으로 분류할 수가 있다. 만일 CS 데이터베이스에서 하나의 항목이라도 이용할 수 없다면 사용자는 새로운 항목을 정의할 수 있고 이러한 새로운 항목 추가 작업으로 CS 데이터베이스의 항목들은 계속 증가하고 다양하게 된다.

각 사건은 단 하나의 세그먼트와 관련되며, (그림 5)에서와 같이 SemanticBase에 EventType 속성 값을 사용하는 세그먼트를 의미적 엔티티에 연관시킬 수 있다. 사건의 재사용을 위한 구성이다. 또한 사용자는 각 사건에 정보를 추가할 수 있다. 즉, MPEG-7 의미도구의 정의를 따르는 Time, Place 그리고 Agent 등의 정보를 (그림 6)과 같이 추가할 수 있다. 이

```
<ClassificationScheme uri="urn:mpeg:mpeg7:cs:event_lev0">
  <Term termID="birthday">
    <Name xml:lang="en">Birthday</Name>
  </Term>
  <Term termID="vacation">
    <Name xml:lang="en">Vacation</Name>
  </Term>
  <Term termID="work">
    <Name xml:lang="en">Work</Name>
  </Term>
</ClassificationScheme>
```

(그림 4) ClassificationScheme(CS) XML 문서

```
<VideoSegment id="Vacation2006">
  <SemanticBase xsi:type="EventType" id="E1_vac2006">
    <Label href="urn:mpeg:mpeg7:cs:event_lev0:vacation">
      <Name>Vacation</Name>
    </Label>
  </SemanticBase>
</VideoSegment>
```

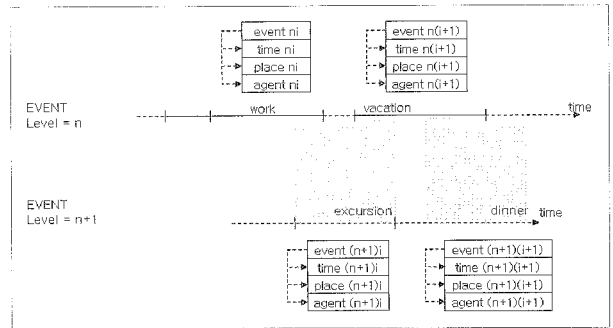
(그림 5) EventType XML 문서

```
<SemanticBase xsi:type="EventType" id="E1_vac2006">
  <Label href="urn:mpeg:mpeg7:cs:event_lev0:vacation">
    <Name>Vacation</Name>
  </Label>
  <Relation target="#ST1_vac2006"
    type="urn:mpeg:mpeg7:cs:SemanticRelationCS:2006:time" />
  <Relation target="#SP1_vac2006"
    type="urn:mpeg:mpeg7:cs:SemanticRelationCS:2006:place" />
  <Relation target="#AO1_vac2006"
    type="urn:mpeg:mpeg7:cs:SemanticRelationCS:2006:agent" />
</SemanticBase>
<SemanticBase xsi:type="SemanticPlaceType" id="SP1_vac2006">
  <Label href="urn:mpeg:mpeg7:cs:semantic_place_lev0:germany">
    <Name>Germany</Name>
  </Label>
</SemanticBase>
<SemanticBase xsi:type="SemanticTimeType" id="ST1_vac2006">
  <Label href="urn:mpeg:mpeg7:cs:semantic_time_lev0:2006">
    <Name>2006</Name>
  </Label>
</SemanticBase>
<SemanticBase xsi:type="AgentObjectType" id="AO1_vac2006">
  <Label href="urn:mpeg:mpeg7:cs:role-travellers">
    <Name>Travellers</Name>
  </Label>
  <Agent xsi:type="PersonGroupType">
    <Member>
      <Name>user1</Name>
      <GivenName>user1</GivenName>
      <Name>
      <Member>
    </Agent>
  </SemanticBase>
```

(그림 6) Event Relation XML 문서

리한 정보 또한 각각의 의미적 엔티티(SemanticBase)로 표시할 수가 있다. 각 타입을 보면 SemanticTimeType, SemanticPlaceType, AgentObjectType이다. 사용자가 추가적인 특징 정보가 필요하면 해당 사건에 정보를 관계시키는 작업이 이루어진다. 이런 작업을 위해서 SemanticBase의 EventType은 Relation 엘리먼트를 포함한다. 이미 정의된 CS인 SemanticRelationCS를 소유하고 있으면, Relation 타입을 이용하여 각 사건에 추가적인 정보를 관계시킬 수 있다. 시스템 구현에서는 간단한 관계집합으로 Time, Place, Agent와 같이 기술하였다.

이 결과 기술은 계층적 사건기반의 구조로 (그림 7)과 같이 추출된다. 게다가, 각 사건은 간단한 링크를 통하여 의미적 엔티티(Event, Time, Place, Agents, Objects)들에 의해서 기술된다. 이 구조기술은 사용자 인터페이스에 의해서 반자동적으로 얻어지며, 구조기술을 위한 사용자 인터페이스에서는 다양한 의미적 구성요소들을 반복적으로 선택할 수가 있다.



(그림 7) 사건으로의 구조적 분해

```
<mpeg7>
  <Description xsi:type="ClassificationSchemeDescriptionType">
    <ClassificationScheme uri="urn:mpeg:mpeg7:cs:event_lev0">
      <Term termID="vacation">
        .....
      </ClassificationScheme>
    </Description>
  </mpeg7>
  <mpeg7>
    <Description xsi:type="ContentEntityType">
      <MultimediaContent xsi:type="VideoType">
        <Video>
          <TemporalDecomposition gap="true" overlap="false">
            <VideoSegment id="Vacation2006">
              <Semantic>
                <Label/>
                <SemanticBase xsi:type="EventType" id="E1_vac2006">
                  .....
                </SemanticBase>
              </Semantic>
            </VideoSegment>
          </TemporalDecomposition>
        </Video>
      </MultimediaContent>
    </Description>
  </mpeg7>
```

(그림 8) 최종 구조기술의 XML 문서

MPEG-7 표준을 따르는 실행코드 구현을 위하여, 구조기술은 (그림 8)에서와 같이 모든 분류기술인 CS들은 Description의 ClassificationSchemeDescriptionType을 이용하여 시작하고, 다음으로 Description의 ContentEntityType을 이용하여 사건으로 분해된 각 기술들이 기술된다. 각 사건의 SemanticBase 목록은 각 사건과 부 사건을 기술하기 위해서 사용하는 VideoSegment의 지식요소인 Semantic 요소에 포함된다.

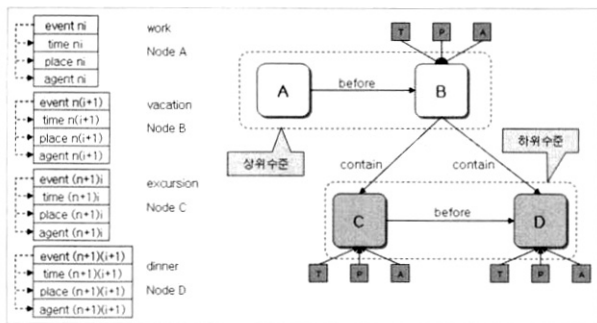
### 3.2 의미기술의 추출

이전 절에서, 구조기술과 분류기술이 어떻게 정의되는지를 살펴보았다. 이들 기술들은 다양한 의미기술을 추출하기 위해서 처리될 것이며, 이를 위해서 구조기술은 콘텐츠의 세그먼트로 기술되며, 사건과 같은 의미적 엔티티와 대응되며, 이러한 사건은 Time, Place, Agents, Objects와 같은 의미적 엔티티들을 포함 또는 관계 연결한다.

이 처리의 목적은 구조기술 및 분류기술에서 추출된 의미적 엔티티를 이용하여 자동적으로 의미기술을 그래프로 이용하여 생성하기 위함이다. 그래프의 구성을 보면 노드는 사건의 의미적 엔티티이며, 사건들 사이의 링크는 의미적 관계를 의미한다. MPEG-7 기반 의미기술은 Description의 SemanticDescriptionType 속성 값으로 시작한다. 그 다음의 필수요소로는 SemanticBagType의 Semantics 태그이다. 이 SemanticBagType의 의미는 Semantics가 의미적 엔티티

(SemanticBase)의 집합을 포함할 수 있어 그 만큼 많은 의미적 엔티티들을 가진다는 의미에서 가방의 “bag”을 표현한 것이다. 단순 연결된 구조기술에서 이용할 수 있는 모든 의미적 엔티티를 수집하기 위해서 반복적으로 위 속성들은 사용한다. 의미적 엔티티들은 사건과 이미 대응 관계가 되어 있지만 모든 사건들은 동일한 수준에 있다. 이러한 관계를 표현하기 위해서 본 논문에서는 의미적 그래프를 이용하여 멀티미디어 데이터의 구조적 분해를 통하여 사건들로 보이게 되며, 사건들 사이에 의미적 관계를 생성할 수 있다. 사건은 그래프의 노드로 표현되고, 사건들 사이의 관계는 링크로 표현된다.

위 내용을 기반으로 (그림 3)의 “summer” 예시 멀티미디어 데이터에 대한 구조기술을 고려해 보자. “summer”는 “work” 후에 “vacation”이 오며, “vacation” 기간 동안은 “excursion”을 하고, 이후 “dinner”가 나오는 시나리오이었다. 여기서 “summer”를 세그먼트로 분해를 한 후, 사건에 대응하여 다음과 같이 표현을 할 수가 있다. 상위 단계의 사건 “work(node A)”와 사건 “vacation(node B)”로 다시 사건 “vacation(node B)”는 하위단계의 부 사건으로 “excursion(node C)” 및 “dinner (node D)”로 나타 낼 수 있으며, 사건들 사이의 관계를 설명하면 다음과 같이 전개를 할 수가 있다.



(그림 9) 의미그래프를 이용한 의미기술 표현

```

<Mpeg7>
<Description xsi:type="SemanticDescriptionType">
  <Semantics>
    <AbstractionLevel dimension="0" />
    <Label />
    <SemanticBase xsi:type="EventType" id="E1_vac2006"> ...
    <SemanticBase xsi:type="SemanticPlaceType" id="SP1_vac2006"> ...
    <SemanticBase xsi:type="SemanticTimeType" id="ST1_vac2006"> ...
    <SemanticBase xsi:type="AgentObjectType" id="AO1_vac2006"> ...

    <SemanticBase xsi:type="EventType" id="E1_exc2006"> ...
    <SemanticBase xsi:type="SemanticPlaceType" id="SP1_exc2006"> ...
    <SemanticBase xsi:type="SemanticTimeType" id="ST1_exc2006"> ...
    <SemanticBase xsi:type="AgentObjectType" id="AO1_exc2006"> ...

    <SemanticBase xsi:type="EventType" id="E1_din2006"> ...
    <SemanticBase xsi:type="SemanticPlaceType" id="SP1_din2006"> ...
    <SemanticBase xsi:type="SemanticTimeType" id="ST1_din2006"> ...
    <SemanticBase xsi:type="AgentObjectType" id="AO1_din2006"> ...
    .....
  </Semantics>
  <Graph>
    <Node id="Vacation" href="#E1_vac2006"/>
    <Node id="Excursion" href="#E1_exc2006"/>
    <Node id="Dinner" href="#E1_din2006"/>
    <Relation source="#Vacation" target="#Excursion"
      Type="urn:mpeg:mpeg7:cs:TemporalRelationCS:2006:contain" />
    <Relation source="#Vacation" target="#Dinner"
      Type="urn:mpeg:mpeg7:cs:TemporalRelationCS:2006:contain" />
    <Relation source="#Excursion" target="#Dinner"
      Type="urn:mpeg:mpeg7:cs:TemporalRelationCS:2006:before" />
  </Graph>
</Description>
</Mpeg7>
  
```

(그림 10) Semantic Description XML 문서

- “vacation(node B)” 사건은 “excursion(node C)” 사건과 “dinner(node D)” 사건을 포함한다(contain).
- “excursion(node C)” 사건은 “dinner(node D)” 사건 이전에 발생되어야 한다(before).

위와 같이 전개된 사건들의 노드 표현과 사건과의 관계를 의미적 그래프로 표현하면 (그림 9)과 같이 나타나어 표현할 수가 있다.

MPEG-7 표준에 따라서, 위의 그래프를 표현하기 위해서 (그림 10)과 같이 DS의 Graph 태그를 사용한다(Graph 요소는 Semantics 요소의 자식 요소들 중 하나이다.). Node는 “vacation”, “work”, “dinner” 등과 같은 사건의 집합이 된다. “contain”, “before”와 같은 링크는 Relation이 된다. 이와 같은 관계의 경우는 CS의 TemporalRelationCS를 사용하여 표현을 한다.

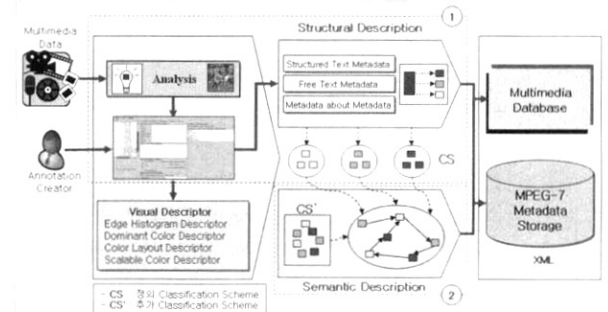
#### 4. 주석도구의 구현 및 비교평가

앞 장까지 기술한 정의를 이용하여 의미적 메타데이터를 생성하기 위한 MPEG-7 기반 주석도구인 SMAT를 소개하고, 이미지와 비디오 멀티미디어 데이터에 대한 적용 결과를 기술한다. 그리고 기존의 시스템들과 기능적인 측면에서 비교 평가한다.

##### 4.1 SMAT 도구의 구현

(그림 11)에서와 같이 본 도구 SMAT는 의미기술 생성을 위하여 멀티미디어 데이터를 입력받아 크게 두 부분의 처리 과정을 거치게 된다. 이러한 과정은 주어진 멀티미디어 데이터의 구조적 분해를 통해 의미적 엔티티인 사건으로 대응하고 또 다른 의미 엔티티들을 추가하는 구조기술 단계(그림 11:①)와 전 단계에서 생성된 의미 엔티티들 서로 간에 노드와 링크를 이용한 의미적 그래프 인터페이스 작업을 통하여 의미기술을 생성하는 의미기술 단계(그림 11:②)로 나누어진다.

분류기술은 모든 과정에서 생성된다. 특히 기본적인 분류기술은 구조기술 단계에서 생성(CS)되며 이후 과정에서의 분류기술은 상황에 따라 추가적인 생성(CS')이 가능하며,



(그림 11) SMAT 전체적인 작업흐름

항상 최신의 것으로 데이터베이스에 저장되어 갱신된다. 또한 본 논문에서는 제외되었지만 추가적으로 (그림 11)에서 보듯이 SMAT 도구는 구조기술 단계에서 저급수준의 비주얼 기술자 메타데이터(에지 히스토그램, 도미넌트 컬러 등), 구조적 텍스트 메타데이터, 비정형적 텍스트 메타데이터, 기술에 대한 메타데이터 등 다양한 메타데이터들의 생성도 지원하고 있다.

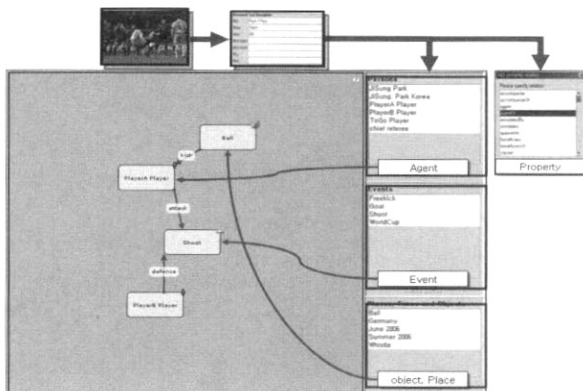
전체적인 작업 과정 후의 메타데이터는 MPEG-7 표준 정의를 따르는 XML 문서로서 주석 되고 저장되어진다.

#### 4.2 이미지 데이터의 적용

이미지 데이터는 비디오 데이터와는 달리 원본 그 자체가 하나의 사건과 동일하다. 따라서 세그먼트 분해의 과정은 필요가 없으며, 이미지 원본 그 자체의 의미정보만으로 메타데이터를 구성하여야 한다. 이미지 데이터를 SMAT 도구를 이용한 처리과정으로 다음과 같이 기술할 수 있다.

구조기술 단계에서는 이미지 원본 자체에서 의미 정보를 구조적 텍스트 메타데이터의 양식에 텍스트 입력함으로 의미적 엔티티들이 도구에서 생성이 된다. 구조적 텍스트 양식은 멀티미디어 데이터의 주석에 유용하게 사용되는 MPEG-7 기술의 6WIH(Who, Where, When, What Object, What Action, Why, How)로 구성되며, 입력된 6WIH의 텍스트 메타데이터 값은 의미적 엔티티 및 속성의 값으로 변형이 이루어진다. 의미기술 단계에서 이미지는 구조적인 수준 차이가 0인 멀티미디어 데이터이다. 사건의 의미적 엔티티는 다른 의미적 엔티티와 관계를 이룰 수가 있는데, 0 수준의 이미지 데이터는 수준과는 관계없으며 단일 사건의 의미적 엔티티를 가지고 의미적 그래프 작업을 하여 메타데이터를 생성한다.

위 과정의 이해를 돕기 위하여 (그림 12)는 축구 경기에서 “슛(Shoot)” 하는 장면의 이미지 데이터[18]이며, 의미정보를 기술하면 다음과 같다. 사건의 의미적 엔티티는 “Shoot”으로 정의하며 그에 추가적인 의미적 엔티티로는 “Ball”, “playerA”, “PlayerB”등으로 나타낼 수 있다. “Shoot” 사건과 “playerA”, “PlayerB” 선수와의 관계는 “attack”과 “defence”이다. 또한 “playerA”는 “Ball”을 골대를 향해서 차고 있다(kick). 이외



(그림 12) “Shoot” 이미지의 의미기술 그래프 작업화면

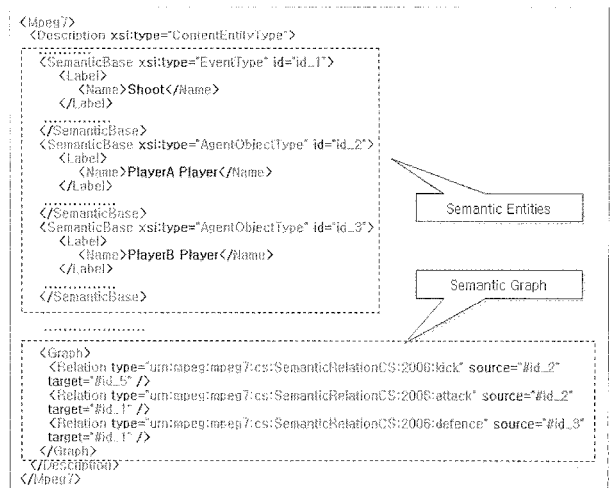
사용자가 생각을 하지 못하는 의미적 엔티티도 있을 수가 있지만, 필요시 마다 도구에서 추가 가능하다. 예시 이미지를 제외한 나머지 그림은 예시 이미지를 의미기술하기 위한 그래프 작업의 화면을 보여 주고 있다.

사건 “Shoot”에 대한 “PlayerA”, “PlayerB”는 “agent” 분류이며 관계에 대한 “attack”과 “defence”의 속성으로 링크된 것을 보여주고 있다. “PlayerA”와 “Ball”에 대한 관계로 “kick” 속성으로 링크된다. 이러한 의미 그래프 작업에 의해서 의미기술이 생성되며, 이 결과 MPEG-7 기반의 의미적 멀티미디어 메타데이터가 생성된다. 이러한 메타데이터는 XML 문서 형태로 저장이 된다. (그림 13)는 의미적 엔티티의 기술부분과 의미적 그래프 표현에 따른 의미기술 부분의 XML 문서를 보여주고 있다.

#### 4.3 비디오 데이터의 적용

앞에서 살펴본 것과 같이 비디오 데이터는 프레임 및 샷 장면 등의 계층적인 구조로 표현된다. 비디오 데이터는 주석을 위한 의미기술 메타데이터 생성 전에 세그먼트 분해를 하게 되며, 분해 된 세그먼트는 의미 사건으로 각각 대응되게 된다. 따라서 비디오 데이터를 여러 사건의 집합이라고도 한다. 이러한 사건들 간의 의미 관계를 위하여 SMAT 도구는 하나의 세그먼트에 대한 의미 메타데이터 처리 작업뿐만 아니라, 확장하여 세그먼트와 세그먼트 간에 의미 메타데이터를 생성할 수 있다. SMAT 도구를 이용한 처리과정을 보면 다음과 같이 기술할 수 있다.

구조기술 단계에서 비디오 데이터는 세그먼트 분해를 하게 된다 각 세그먼트의 대표 프레임에 대한 구조적 텍스트 메타데이터의 텍스트 값을 입력하게 되며, 세그먼트는 사건의 의미적 엔티티와 대응된다. 사건은 다시 또 다른 의미적 엔티티들과 관계를 이룰 수 있다. 의미기술 단계에서 각 사건에 대한 의미적 그래프 작업이 수행되고, 사건과 사건에 대한 의미 관계를 구성하여 의미기술 한다. 그 결과 비디오 데이터의 의미적 메타데이터를 생성하게 되는 것이다.



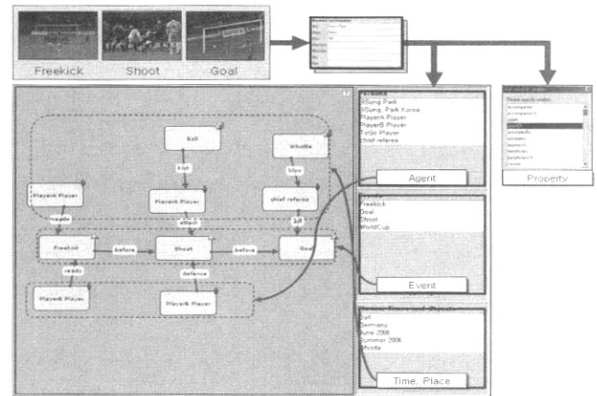
(그림 13) “Shoot” 이미지의 의미기술 XML 문서

SMAT 도구에 적용할 비디오 데이터는 (그림 14)와 같이 세 개의 세그먼트를 가지는 “축구” 비디오 클립이다[18]. 해당 비디오 세그먼트는 각각 “Freekick(node A)”, “Shoot(node B)”, “Goal(node C)”이라는 사건으로 대응된다. 각 사건에 대한 내용은 구조적 텍스트 메타데이터 양식을 통하여 의미 정보를 수동적으로 기록하게 된다.

(그림 14)에 참고하기 위한 각 사건들과 의미적 엔티티들 사이의 의미관계를 시나리오하면 다음과 같다.

- “Freekick(node A)” 사건은 “Shoot(node B)” 사건 전에 발생되어야 한다(*before*).
- “Shoot(node B)” 사건은 “Goal(node C)” 사건 전에 발생되어야 한다(*before*).
- “Freekick(node A)” 사건에서는 “PlayerA”, “PlayerB” 선수들이 프리킥을 준비를 하고 있다(*ready*).
- “Shoot(node B)” 사건에서는 “PlayerA” 선수가 공을 차고(*kick*), “PlayerB” 선수들은 수비(*defence*)를 하고 있다.
- “Goal(node C)” 사건에서는 골인 된 장면이며, 주심(chief referee)이 호각(Whistle)을 불고(*blow*) 있다.

위 작업과정에서 중앙부분에 세 개의 사건이 나열되어 사건 서로 간에 “before” 관계 링크를 하고 있다. 각 사건마다 의미적 엔티티들이 추가되어 의미정보를 확장 기술되고 있다. 이러한 모든 의미적 엔티티들의 연결로 의미기술이 생성된다.



(그림 14) “축구” 비디오의 의미기술 그래프 작업화면

작업과정에서 공(Ball), 주심(chief referee) 및 호각(whistle)의 의미적 엔티티를 볼 수 있는데, 작업과정에서 새로이 추가된 의미적 엔티티들이다. 다음은 위 작업에 의한 전체적인 비디오 메타데이터의 의미기술 부분은 (그림 15)와 같은 XML 문서의 결과로 생성된다.

#### 4.4 기존 주석도구와의 비교평가

본 논문에서는 관련연구로 가장 대표적인 기존의 주석도구로 VideoAnnEx[11-12], Ricoh Movie Tool[16], MPEG-7 Annotation Tool[17] 등을 소개하였다. 비교 평가는 주석도

```

<mpeg7>
  <Description xsi:type="ClassificationSchemeDescriptionType">
    <ClassificationScheme uri="urn:mpeg:mpeg7:cs:event_level0">
      <Term termId="freekick">
        <Name xml:lang="en">Freekick</Name>
      </Term>
      <Term termId="shoot">
        <Name xml:lang="en">Shoot</Name>
      </Term>
      <Term termId="goal">
        <Name xml:lang="en">Goal</Name>
      </Term>
    </ClassificationScheme>
  </Description>
  <Agent />
  <Agent />
  <Description xsi:type="ContentEntityType">
    <MultimediaContent xsi:type="VideoType">
      <Video>
        <TemporalDecomposition gap="true" overlap="false">
          <VideoSegment id="Freekick2006">
            <Semantic>
              <Label />
              <SemanticBase xsi:type="EventType" id="E1_fre2006">
                <Label href="urn:mpeg:mpeg7:cs:event_level0:freekick">
                  <Name>Freekick</Name>
                </Label>
                <Relation target="#ST1_fre2006"
                  type="urn:mpeg:mpeg7:cs:SemanticRelationCS:2006:time" />
                <Relation target="#SP1_fre2006"
                  type="urn:mpeg:mpeg7:cs:SemanticRelationCS:2006:location" />
                <Relation target="#AO1_fre2006"
                  type="urn:mpeg:mpeg7:cs:SemanticRelationCS:2006:agent" />
              </SemanticBase>
            </Semantic>
          </VideoSegment>
          <VideoSegment id="Freekick2006">
            <Semantic>
              <Label />
              <SemanticBase xsi:type="EventType" id="E1_sho2006">
                <Label href="urn:mpeg:mpeg7:cs:event_level0:shoot">
                  <Name>Shoot</Name>
                </Label>
                <Relation target="#ST1_sho2006"
                  type="urn:mpeg:mpeg7:cs:SemanticRelationCS:2006:time" />
                <Relation target="#SP1_sho2006"
                  type="urn:mpeg:mpeg7:cs:SemanticRelationCS:2006:location" />
                <Relation target="#AO1_sho2006"
                  type="urn:mpeg:mpeg7:cs:SemanticRelationCS:2006:agent" />
              </SemanticBase>
            </Semantic>
          </VideoSegment>
          <VideoSegment id="Goal2006">
            <Semantic>
              <Label />
              <SemanticBase xsi:type="EventType" id="E1_goa2006">
                <Label href="urn:mpeg:mpeg7:cs:event_level0:goal">
                  <Name>Goal</Name>
                </Label>
                <Relation target="#ST1_goa2006"
                  type="urn:mpeg:mpeg7:cs:SemanticRelationCS:2006:time" />
                <Relation target="#SP1_goa2006"
                  type="urn:mpeg:mpeg7:cs:SemanticRelationCS:2006:location" />
                <Relation target="#AO1_goa2006"
                  type="urn:mpeg:mpeg7:cs:SemanticRelationCS:2006:agent" />
              </SemanticBase>
            </Semantic>
          </VideoSegment>
        </TemporalDecomposition>
      </Video>
    </MultimediaContent>
  </Description>
  <Graph>
    <Semantic>
      <VideoSegment>
        <Semantic>
          <Label />
          <SemanticBase xsi:type="EventType" id="E1_fre2006">
            <Label href="urn:mpeg:mpeg7:cs:event_level0:freekick">
              <Name>Freekick</Name>
            </Label>
            <SemanticBase xsi:type="AgentObjectType" id="AO1_fre2006">
              <Label />
            </SemanticBase>
            <SemanticBase xsi:type="AgentObjectType" id="AO1_sho2006">
              <Label />
            </SemanticBase>
            <SemanticBase xsi:type="AgentObjectType" id="AO1_goa2006">
              <Label />
            </SemanticBase>
          </SemanticBase>
          <Graph>
            <Code id="Freekick" href="E1_fre2006" />
            <Code id="Shoot" href="E1_sho2006" />
            <Code id="Goal" href="E1_goa2006" />
            <Relation source="#Freekick" target="#Shoot"
              type="urn:mpeg:mpeg7:cs:TemporalRelationCS:2006:before" />
            <Relation source="#Shoot" target="#Goal"
              type="urn:mpeg:mpeg7:cs:TemporalRelationCS:2006:before" />
          </Graph>
        </Semantic>
      </VideoSegment>
    </Semantic>
  </Graph>

```

(그림 15) “축구” 비디오의 의미기술 결과 XML 문서



〈표 1〉 주석도구의 비교평가

구 분	VideoAnnEx	Ricoh Movie Tool	MPEG-7 Annotation Tool	SMAT
이미지 처리	×	×	×	○
비디오 처리	○	○	○	○
의미기반 주석	△	△	△	○
특징기반 주석	×	△	×	○
엔티티 재사용성	△	×	×	○
엔티티 확장성	△	×	×	○
세그먼트 간의 의미화	×	×	×	○
의미주석 방법	free Text	free Text	free Text	Semantic Graph

(범례 : ○ 가능 △ 부분적 가능 × 불가능)

구에서 제공하는 기능 측면이며, 비교 결과는 <표 1>과 같이 나타낼 수가 있다.

기존 도구들의 의미주석 방법은 대부분 도구들이 비정형적 텍스트 입력방식이다. 이러한 방법은 해당 멀티미디어 데이터의 의미정보를 사용자가 생각하는 주관적 면을 텍스트로 주석하는 것이며, 주석이 이루어질 때마다 항상 같은 의미의 정보도 상이할 수 있다는 문제점을 가지고 있다. VideoAnnEx 도구는 그러한 단점을 주석을 위한 구조적 카테고리를 지원하여 작업할 수 있게 하고 있다. 하지만 카테고리에 새로이 구성요소를 추가하는데 문제점이 있으며, 분류된 구성요소들이 멀티미디어 데이터의 의미정보를 지원하는 분류는 아닌 것이다. 제안된 SMAT 도구는 이러한 문제점을 의미적 엔티티 개념으로 해석하고 주석을 위한 의미정보 표현을 위한 구성요소들을 데이터베이스에 저장하고 주석 작업 때마다 저장된 구성요소를 재사용하여 의미 메타데이터를 생성할 수 있는 장점을 가지고 있으며, 구성요소들을 새로이 추가 또한 가능하다. 더불어 SMAT 도구는 기존 텍스트 입력이 아닌 의미적 그래프를 이용한 드로잉 작업방법을 채택하고 있어 사용자 편의성도 가지고 있다.

또한, 비디오 데이터만을 처리하는 기존 도구들은 세그먼트와 세그먼트 서로 간의 의미관계는 지원을 하지 않고 있다. 단지, 하나의 세그먼트에 비정형적 텍스트 입력으로 의미 주석하고 있는 반면에, SMAT 도구는 세그먼트들 서로 간의 의미관계를 개념화 할 수 있으며, 각각의 세그먼트에 추가적인 의미 정보를 표현할 수도 있다.

### 5. 결론 및 향후 연구

본 논문에서는 멀티미디어 사용자가 멀티미디어 데이터와 관련된 의미정보를 가지는 MPEG-7 기반의 기술(description)을 생성하기 위해 도움을 주는 대화식 주석도구(SMAT)를 소개하였다. 생성된 결과 기술은 계층적 구조의 사건기반이다. 이 사건들은 Time, Place, Agent, Object의 특징을 가지는 의미적 엔티티들로 관계하여 표현하였다. 이러한 목적으로 SMAT 도구는 사건들과 의미적 엔티티들과의 관계를 표

현하기 위하여 개인화된 분류구조(Classification Scheme : CS)를 사용자로 하여금 생성할 수 있게 제공하며, 생성된 모든 CS는 데이터베이스에 저장 관리된다. 따라서 새로운 주석 작업이 주어지면 기존의 CS를 이용하여 기술들을 표현할 수가 있어 일관성 있는 메타데이터 생성이 가능하였다. 또한 메타데이터 생성 중에 필요에 따라 CS를 추가 또는 삭제가 가능하다.

SMAT 도구는 두 부분의 처리과정을 통하여 사용자에게 의미기술을 제공하였다. 첫 번째 단계에서 도구는 멀티미디어 데이터를 분석하여 구조화 텍스트 메타데이터 작업을 통하여 구조기술을 추출하였다. 두 번째 단계는 추출된 구조기술과 분류기술을 기반으로 의미적 그래프 작업을 통하여 의미기술을 추출하였다. 이러한 작업에서 그래프의 노드들은 사건의 의미적 엔티티 및 다른 의미적 엔티티들을 말하며, 링크는 이러한 의미적 엔티티들 간의 관계 표현을 의미한다. 사용자들은 이러한 의미적 그래프 작업도구를 통하여 편리한 의미적 멀티미디어 메타데이터 작업을 할 수가 있었다.

결론적으로, 본 논문에서 개발 제안된 SMAT 도구는 기존의 주석도구와는 다른 재사용성(reusability)과 확장성(extendibility)의 강한 특징을 가지고 있으며, 추가적으로 의미적 메타데이터 생성 작업에서 기존의 키워드 생성이 아닌 의미 그래프 기반의 드로잉(drawing) 작업으로 사용자에게 편의성을 제공해 주고 있다.

향후 연구로 현재 SMAT 도구는 의미기술 추출을 위해 필요한 분류기술인 CS의 증가로 의미 엔티티들 간에 계층 관계 및 중복관계를 명확하게 해결을 하지 못하고 있다. 따라서 증가하는 CS 구성에 있어 비교 분석하는 알고리즘을 연구 중에 있으며, 의미적 그래프를 이용한 멀티미디어 데이터 검색 도구도 개발 진행 중에 있다.

### 참고 문헌

[1] Dongge Li, Ishwar K. Sethi, Nevenka, Thomas McGee, "Classification of general audio data for content-based retrieval," Pattern Recognition Letters, Vol.22, No.5,

pp.533-544, 2001.

[2] M. Flickner et al., "Query by Image and Video Content : The QBIC System," Computer, Vol.28, No.9, pp.23-32, 1995.

[3] Moving Picture Experts Group, "Overview of the MPEG-7 standard," ISO/IEC JTC1/SC29/WG11 N4509, Pattaya, Thailand, December, 2001.

[4] S. F. Chang, "Overview of the MPEG-7 Standard," IEEE Trans. On Circuits and Systems for Video Technology, Vol.11, No.6, pp.688-695, June, 2001.

[5] Tsuhan Chen. Low-Level Features to High-Level Semantics : Are We Bridging the Gap ? In EWIMT, London, UK, November, 2004.

[6] M. Martinez, "Overview of the MPEG-7 standard," ISO/IEC JTC1/SC29/WG11 N4031, March, 2001.

[7] M. Martinez. Overview of MPEG-7 Description Tools, Part 2. IEEE MultiMedia, pages 83-93, 2002.

[8] M. Martinez. MPEG-7 Overview. Multimedia Content Description Interface. <http://www.chiariglione.org/mpeg/standards/mpeg7/mpeg-7.html/>, 2003.

[9] ISO/IEC 15938-5. Information Technology - Multimedia Content Description Interface -- Part 5 : Multimedia Description Schemes, September, 2001.

[10] B. Y. Ricardo and R. N. Bethier, Modern Information Retrieval, ACM press, 1999.

[11] <http://www.alphaworks.ibm.com/tech/videoannex>

[12] <http://www.research.ibm.com/VideoAnnEx>

[13] 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.

[14] 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.

[15] 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.

[16] <http://www.ricoh.co.jp/src/multimedia/MovieTool/index.html>

[17] <http://tswwww.epfl.ch/~newuma/>

[18] soccer video clips, <http://ghs.pasco.k12.fl.us/clips/clips.html>

[19] 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.

[20] A. Graves and M. Lalmas. Video retrieval using an MPEG-7 based inference network. In the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pages 339-346. ACM Press, August 2002.

### 안 형 근



e-mail : [hkahn@mail.ulsan.ac.kr](mailto:hkahn@mail.ulsan.ac.kr)

2000년 한국방송통신대학교

컴퓨터과학과(이학사)

2003년 울산대학교 정보통신대학원

정보통신공학과(공학석사)

2006년 울산대학교 대학원

컴퓨터정보통신공학부(박사수료)

1997년 ~ 2004년 현대오토시스템 기술지원부

2004년 ~ 현재 (주)CFIC 기업부설연구소 연구소장

관심분야 : 멀티미디어DB, DB설계/분석, ERP, BPM, Workflow

### 고 재 진



e-mail : [jjkoh@mail.ulsan.ac.kr](mailto:jjkoh@mail.ulsan.ac.kr)

1972년 서울대학교 응용수학과(공학사)

1981년 서울대학교 대학원

계산통계학과(이학석사)

1990년 서울대학교 대학원

컴퓨터공학과(공학박사)

1975년 ~ 1979년 한국후지쯔(주) 기술개발부 사원

1979년 ~ 현재 울산대학교 컴퓨터정보통신공학부 교수

관심분야 : DB시스템, 전문가 시스템, DB설계, ERP