

## 멀티미디어 파일에 대한 시맨틱 어노테이션에 관한 연구

황인문<sup>○</sup> 유남현 송길종 김원중  
순천대학교 컴퓨터학과  
kwj@sunchon.ac.kr

### A Study on Semantic Annotation System for Multimedia File

InMoon Hwang<sup>○</sup> NamHyun Yoo GilJong Song WonJung Kim  
Suncheon National University Dept. of Computer Science

#### 요 약

기존의 멀티미디어 자료들을 직접 분석하여 메타데이터를 부여하던 기존 연구들은 멀티미디어 파일에 부여된 키워드에 해당하는 온톨로지가 존재하지 않는 경우 해당 멀티미디어 파일에 대한 직접 검색을 실시하지 않거나, 잘못된 의미가 부여된 키워드에 의하여 검색 결과의 정확성과 재현율이 떨어지는 문제점을 가지고 있었다. 본 논문에서 제안한 SASM은 키워드에 부여된 다의성과 동의성의 개념을 이용하여 WordNet을 이용하여 확장한 후, 확장된 키워드들을 이용하여 온톨로지가 구축되어 있지 않더라도 해당 이미지에 대한 직접 분석을 실시할 수 있으며, 키워드에 대한 동의성과 다의성을 이용한 확장된 키워드들을 이용함으로써 검색 결과의 재현율과 정확성의 성능을 향상시켰다.

#### 1. 서 론

현재 시맨틱 웹을 구축하기 위해서는 인터넷 상에 존재하는 모든 정보들에 대하여 메타데이터를 부여하여야 한다. 부여된 메타데이터를 이용하여 추론 과정을 거치거나 정제 과정을 거친 후 온톨로지가 구축되고 구축된 온톨로지들을 이용하여 시맨틱 웹[1]을 구축할 수 있다. 최근 포털 사이트나 UCC(User Created Contents) 전문 사이트에서 요청하는 태그 정보도 역시 시맨틱 웹을 구현하기 위한 기반 기술이라고 할 수 있다 태그는 최초의 문서 작성 작성자가 해당 정보에 대하여 기초적인 설명인 메타데이터를 부여하는 것으로서 해당 정보에 대한 가장 정확한 메타데이터의 부여가 가능하며 사용자가 정보를 생성하는 마지막 단계에서 연관된 단어를 입력하기 때문에 사용자의 적극적인 참여만 보증된다면 매우 저렴하고도 효과적으로 해당 정보에 대한 메타데이터를 구축할 수 있는 장점이 있다 그러나 기 구축된 정보들이나 새롭게 생성된 정보 중에서 사용자가 관련 정보를 입력하지 않는 경우에는 이에 대한 대책이 필요하다 물론 텍스트 기반의 웹 문서의 경우 기존의 기술을 이용하여 만족할 만한 메타데이터를 구축할 수 있으나 멀티미디어 정보의 경우 해당 정보에 대한 메타데이터를 구축

하기 위해서는 해당 멀티미디어 정보를 최초로 작성한 작성자에 의하거나 전문적인 메타데이터 부여자에 의하여 해당 자료에 대한 메타 데이터를 할당하는 경우 이외에는 매우 어렵다. 이와 같은 문제를 해결하기 위하여 멀티미디어 정보의 내용을 직접 분석한 후 메타데이터를 부여하거나, 수동으로 메타데이터를 효과적으로 부여할 수 있도록 해주는 기존의 다양한 연구들이 존재해 왔다.

본 논문에서는 멀티미디어 정보를 직접 분석하여 메타데이터를 부여하는 기존의 연구들이 기 구축된 온톨로지가 정보가 없거나 잘못된 의미의 메타데이터를 내포함으로써 발생할 수 있는 문제점을 해결하기 위한 SASM(semantic Annotation System for Multimedia file) 제안하였다.

#### 2. 관련연구

##### 2.1 Semantic Web

웹의 창시자인 W3C의 Tim Berners-Lee가 1999년 제안하여 표준화가 진행되고 있는 시맨틱 웹 기술은 인간 지식의 처리를 기계가 수행할 수 있도록 지원하고

웹을 통해 확장, 공유할 수 있도록 하는 핵심 기술로 인정받고 있다. 이미 미국과 유럽에서는 시맨틱 웹 기술에 대한 많은 연구가 진행되고 있으며 최근에는 상업적인 시도가 이루어지고 있다 수년 안에 이러한 노력은 지식 기반 사회의 거의 모든 산업에 큰 영향을 끼칠 것이라고 예상하고 있다. 특히 우리나라에서도 정보 차원에서 IT839의 차세대 소프트 인프라로 시맨틱 웹의 연구 개발 육성에 힘을 쏟고 있다. 특히 유비쿼터스 기반의 정보 환경이 전개됨에 따라 휴대폰 서비스 등에 시맨틱 웹 기술을 응용하려는 연구 개발이 활발히 진행되고 있다

시맨틱 웹 기술은 온톨로지 엔지니어링 시맨틱 어노테이션, 시맨틱 정보 저장 관리 시스템 의미 정보 질의 및 추론 시스템, 응용 솔루션 및 서비스 시스템 등 다섯 가지로 나눌 수 있는데 본 논문에서 구현한 SASM의 경우 시맨틱 어노테이션과 시맨틱 정보 저장 관리 시스템과 연관이 있다.

## 2.2. MPEG-7

MPEG-7은 동영상 전문가 그룹인 MPEG(moving picture experts group)에서 개발한 ISO/IEC(international organization for standardization/international electro-technical commission) 표준으로 공식 명칭은 "멀티미디어 콘텐츠 기술을 위한 인터페이스(multimedia content description interface)"이다. 이전의 MPEG 표준들이 저장 매체에 동영상 및 오디오 정보를 저장하기 위한 MPEG-1, 디지털 텔레비전 방송을 MPEG-2, 인터넷 및 이동통신에서 사용하기 위한 멀티미디어 응용인 MPEG-4와 같이 동화상 또는 오디오 정보를 압축하거나 전송하기 위한 것이 목적이었다면 이에 비해 MPEG-7은 멀티미디어 데이터를 서술하기 위한 표준으로 제안되었다 1996년 10월 멀티미디어 내의 영상 데이터 콘텐츠를 서술하기 위한 표준을 만들기 위해 MPEG-7 그룹이 결성되었고, 이후 평가와 실험 모델을 거쳐 2001년 9월에 표준을 제안하였다. 기존 기술과 달리 MPEG-7은 콘텐츠 자체를 표현하는 것이 아니라 콘텐츠에 대한 정보 즉, 메타데이터를 표현한다는 점에서 기존의 표준들과 다르

며, 기존의 MPEG 표준들이 멀티미디어 콘텐츠를 효과적으로 이용할 수 있도록 지원 해주었다면 MPEG-7은 필요한 멀티미디어 콘텐츠의 검색을 가능하게 해준다[2].

## 3. SASM의 설계 및 구현

### 3.1 멀티미디어 파일에 메타데이터의 어노테이션

멀티미디어 파일에 메타데이터를 부여하기 위한 멀티미디어 파일을 직접 분석하거나, 멀티미디어 파일에 메타데이터를 부여해주는 연구들은 매우 다양하게 진행되어져 왔다. [3,4,5,6,7,8,9,10,11,12,13,14]. 기존의 다양한 연구들 중에는 수동 메타데이터 어노테이션 방식[3], 반자동 메타데이터 어노테이션 방식[4,5,6,7,8,9], 자동 메타데이터 어노테이션 방식[10,11,12,13,14]들로 구분될 수 있다. 자동 혹은 반자동 메타데이터 어노테이션 방식의 경우 멀티미디어 파일 자체의 저수준 기술자(low level descriptor)들을 산출하여 이미지에 대하여 추론을 진행한 후, 다시 전문가가 정제(refinement)할 있도록 제공하는 반자동 메타데이터 어노테이션 방식의 변형된 형태의 연구[6,7,8,9]도 있다. 멀티미디어 파일 자체를 직접 분석하는 경우에도 기본적으로 이미지와 관련된 키워드, 문장 등이 할당되어 있는 멀티미디어 파일들을 분석하는 경우와 직접 이미지 자체 내의 윤곽선(edge), 색상, 질감, 공간 관계 등을 분석하여 작업하는 등의 두 가지로 나눌 수 있다. 후자의 경우 해당 이미지에 대한 정확한 검색 결과에 의한 정확한 메타데이터 어노테이션 작업을 수행하기 위해서 소요되는 많은 비용에 비해 검색 결과의 질의 아직도 미미한 수준에 이르기 때문에 대부분의 연구들은 전자의 결과를 발전시키는데 중점을 두고 다양한 연구가 진행되고 있다

전자의 연구에서 가장 큰 문제점은 멀티미디어 파일에 키워드가 부여되어 있더라도 해당 키워드에 대한 온톨로지가 구축되어 있지 않는 경우에는 해당 멀티미디어 파일에 대한 직접 분석을 실시하지 않는다는 점이다 이에 본 논문에서는 WordNet을 이용하여 해당 멀티미디어 파일에 부여된 키워드에 해당하는 온톨로지 파일이 존재하

지 않더라도 키워드에 부여된 다의성과 동의성의 특징을 이용하여 멀티미디어 파일에 대한 직접 분석을 실시할 수 있는 SASM을 설계, 구현하였다.

### 3.2 SASM의 구조

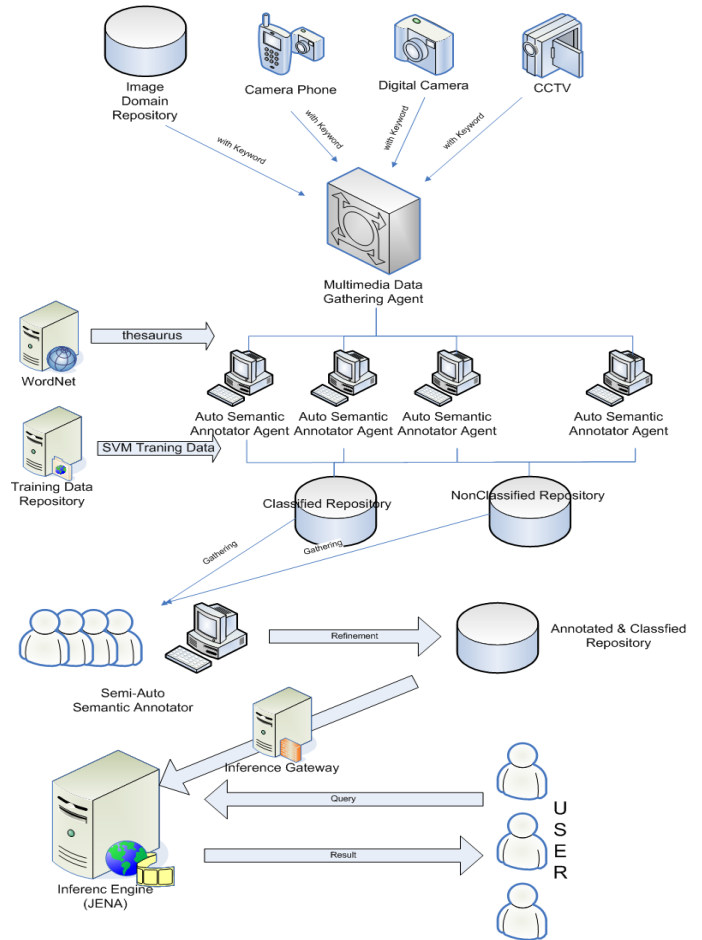
#### (1) SASM의 기본 개념

본 논문에서 설계, 구현한 SASM(Semantic Annotation System for Multimedia File)은 특정 키워드가 부여되어 있는 특정 도메인 중심의 이미지 그룹들이나 개인이 디지털 카메라나 모바일 폰의 디지털 카메라로 찍은 이미지, 캠코더 등에 의한 촬영된 동영상들을 자동으로 분석하여 각 이미지들에게 메타데이터를 어노테이션한 후 전문가나 개인 사용자가 다시 해당 분류에 대하여 수동으로 메타데이터를 정제 혹은 어노테이션 할 수 있도록 지원한다. 특히 SASM은 기존의 메타데이터 어노테이션 시스템들이 해당 이미지에 부여된 특정 키워드에 한정하여 구축된 온톨로지에 기반한 학습 데이터의 결과를 가지고 이미지에 대한 자동 분석하는 방법을 확장 및 개선하여, WordNet에서 제공하는 다의성과 동의성을 가지는 단어들의 연관성을 활용하여 도메인 혹은 해당 키워드와 연관된 온톨로지나 학습데이터가 존재하지 않더라도 해당 멀티미디어 파일에 대한 자동 분석 작업의 학습 범위를 확장한다. 이와 같이 확장된 학습 데이터의 활용으로 해당 이미지들에 대한 검색 결과에 대한 재현율(recall)과 정확율(precision)을 향상시켰다. 또한 자동 메타데이터 어노테이션 방식에 의하여 어노테이션된 메타데이터와 사용자가 직접 어노테이션한 메타 데이터를 JENA 추론 엔진에 RDF 형식으로 저장함으로써, 추후 사용자가 이미지를 검색하는 경우 사용자가 다양한 조건을 가지고 원하는 이미지를 검색할 수 있도록 지원할 수 있는 기반 환경을 제공한다.

#### (2) 멀티미디어 파일 수집 에이전트

본 논문에서 설계, 구현한 SASM은 구글, 네이버 등에서 제공하고 있는 키워드 중심의 이미지 도메인 저장소

에서 데이터를 수집할 수 있는 검색 엔진 에이전트일 반 사용자가 카메라 폰을 이용하여 찍은 사진을 온라인



[그림 1] SASUBI의 전체 구조

상으로 수집할 수 있는 모바일 에이전트 디지털 카메라, 캠코더 등과 같은 멀티미디어 전용기기를 이용해서 찍은 사진이나 영상 등을 사용자가 직접 업로드 하는 것을 수집할 수 있는 웹 에이전트들로 구성된다 이와 같은 다양한 멀티미디어 데이터 검색 에이전트들은 각종 멀티미디어 기기에서 자료를 획득할 경우에는 '제목'과 같이 해당 멀티미디어 데이터에 대하여 사용자가 부여한 키워드가 있어야 한다. 물론 사용자가 무의미하게 입력한 '제목'일지라도 꼭 해당 멀티미디어 파일에 대한 키워드가 포함되어 있어야 한다. 이와 같이 키워드가 필요한 이유는 [그림 1]의 자동 시맨틱 어노테이터(auto semantic annotator agent)가 멀티미디어 파일에 대한 직접 분석을 실시할 때 사용되는 학습 데이터를 WordNet을 이용하여 동의성과 다의성을 이용하여 활용하기 때문이다 만약

멀티미디어 파일 수집 에이전트(multimedia data gathering agent)가 수집한 멀티미디어 데이터에 키워드가 포함되어 있지 않는 경우에는 비분류 저장소(non-classified repository)로 바로 해당 멀티미디어 파일을 저장한다.

(3) 자동 시맨틱 어노테이터 에이전트

자동 시맨틱 어노테이터 에이전트(auto semantic annotator agent)는 멀티미디어 파일 수집 에이전트(multimedia data gathering agent)가 수집한 각종 멀티미디어 자료에 대한 직접 분석을 실시하여 얻은 메타 데이터를 어노테이션 하는 기능을 수행한다. 물론 수집된 멀티미디어 데이터에는 특정 키워드가 이미 부여되어 있어야 한다.

SASM에서 멀티미디어 데이터에 대한 직접 분석을 실시할 때, 기존 연구에서는 해당 키워드에 대한 온톨로지와 그와 연관된 학습 데이터만을 이용하지만 SASM은 해당 키워드에 대한 온톨로지나 그와 연관된 학습 데이터가 존재하지 않더라도 WordNet에서 해당 키워드에 대한 동의어 혹은 다의어 단어를 구한 후 기 부여된 키워드와 동의어 및 다의어에 대한 온톨로지 혹은 학습 데이터를 이용하여 멀티미디어 데이터에 대한 직접 분석을 실시하기 때문에 직접 분석한 결과에 대한 재현율과 정확율 성능을 향상 시킬 수 있다. 멀티미디어 파일에 대한 분석이 끝난 후, 메타데이터를 도출할 수 없는 경우에는 비분류 저장소(non-classified repository)에, 메타데이터를 도출한 경우에는 분류 저장소(classified repository)에 해당 키워드들과 함께 저장한다.

(3) 반자동 시맨틱 어노테이터

반자동 시맨틱 어노테이터(semi-auto semantic annotator)는 자동 시맨틱 어노테이터(auto semantic annotator agent)가 어노테이션한 멀티미디어 파일에 대한 메타데이터를 정제하거나 부여하는 작업을 수행한다. 본 논문에서는 정제 작업을 수행하는 사용자는 해당 자료에 대한 정확한 분석을 수행할 수 있는 숙련된 사용자라고

가정한다. 자동 시맨틱 어노테이터 에이전트(auto semantic annotator agent)가 분석한 결과는 해당 데이터에 대한 최종적인 정제 과정을 거친 후 메타데이터부여분류 저장소(annotated&classified repository)에 저장될 수 있다. 자동 시맨틱 어노테이터(auto semantic annotator agent)에서 수행한 결과가 옳은 경우에는 별도의 정제 과정을 거치지 않아도 되며, 수행한 결과가 옳지 않는 경우에 한해서만 정제의 과정을 거치므로 최종적인 메타데이터부여분류 저장소(annotated&classified repository)에 저장하기까지의 소요되는 시간을 줄일 수 있는 장점이 있다.

4. 성능분석

본 논문에서 멀티미디어 자료에 대한 직접 분석을 실시하기 위하여 UC Berkley Digital Library의 CalPhotos의 Animal을 사용하였으며, 기계 학습 도구로는 SVM(Support Vector Machine)을 이용하였다. 멀티미디어 자료에 대한 직접 분석 방법으로는 윤곽선 검출, 영상 이진화에 따른 패턴 추출, Dominant Color 등을 사용하였다. 멀티미디어 자료에 대한 직접 분석을 이용하여 CalPhotos의 일부 데이터를 추출하여 학습을 실시하였으며 학습을 실시할 때는 WordNet에서 제공하는 다의성과 동의성을 이용하여 확장될 수 있는 키워드까지 포함하여 학습을 실시하였다. 기존 시스템은 단순히 주어진 키워드에 의해서만 학습을 실시하고, 본 논문에서 제안된 시스템은 주어진 키워드 이외에도 다의성과 동의성을 이용하여 멀티미디어 파일에 대한 직접 분석을 실시하였다. [표 1]과 [표 2]는 기존 시스템과 키워드의 다의성과 동의성을 이용한 본 논문에서 제안한 방식으로서 본 논문에서 제안한 방식이 성능이 우수함을 알 수 있다.

[표 2] 다의성을 가지는 키워드가 부여된 이미지들에 대한 직접 분석 결과

구분	기존 시스템		SASUBI	
	개	소시지	개	소시지
인식율	43%	1%	54%	61%
거부율	57%	99%	46%	39%

[표 3] 동의성을 가지는 키워드가 부여된 이미지들에 대한 직접 분석 결과

구 분	기존 시스템	SASUBI
인식율	45%	51%
거부율	55%	49%
정확성	56%	70%

5. 결론

시맨틱 웹을 성공적으로 구현하기 위해서는 그에 따른 킬러 어플리케이션의 개발이 가장 시급하다. 킬러 어플리케이션의 개발하기 위해서는 시맨틱 웹이 구현되기 위한 기반 환경을 구축되어야 한다. 이러한 시맨틱 웹을 구현하기 위하여 기존의 생산된 정보에 효과적으로 적절한 메타데이터를 부여할 수 있는 연구가 필요하다. 본 논문에서는 설계, 구현한 SASM은 기존의 직접 분석에 의한 메타데이터 부여하는 방식이 해당 키워드에 대한 온톨로지가 기 구축되어 있지 않는 경우 분석을 거부하거나, 키워드에 부여된 잘못된 의미의 해석으로 발생할 수 있는 재현율과 정확성 저하 문제를 일정 부분 개선하였다.

참고 문헌

[1] Tim Berners-Lee, "Semantic Web Road Map," W3C September 1998, <http://www.w3.org/DesignIssues/Semantic.html>

[2] 임은주, "내용 기반 멀티미디어 검색을 위한 메타데이터 요소에 관한 연구, 중앙대학교 대학원 문헌정보학과 석사학위논문, 2005.12.

[3] Laura Hollink, et al., "Semantic Annotation of Image Collections," W3C Recommendation, <http://www.w3.org/TR/REC-xml>, 2000.

[4] Christian Halaschek-Wiener, et al., "A Flexible Approach for Managing Digital Images on the Semantic Web," MM: International Multimedia Conference, VOL.00 NO.00 pp.0483 - 0485 2001.

[5] Mischa M. Tuffield, et al., "Image Annotation with Photocopain," ACM Trans on Graphics, VOL.22 NO.03 pp.0409 - 0416 2003. 07

[6] Stephan Bloehdorn, et al., "Semantic Annotation of Images and Videos for Multimedia Analysis," In Proceedings of VLDB Conf., 1999.

[7] K. Petridis, et al., "Knowledge representation and semantic annotation of multimedia content," IEEE Data Engineering Bulletin 22(3), 2006.

[8] Kosmas Petridis, et al., "M-OntoMat-Annotizer: Image Annotation Linking Ontologies and Multimedia Low-Level Features," Workshop on Query processing for semistructured data and non-standard data formats, 2005.

[9] Adrian Matellanes, Alyson Evans, and Burack Erdal, "Creating an application for automatic annotation of images and video," In Proceedings of WebDB, 2000.

[10] G. Th. Papadopoulos, P. Panagi, S. and Dasiopoulou, "A Learning Approach to Semantic Image Analysis," In Proceedings of DBPL Conf., 2003.

[11] G. Th. Papadopoulos, P. Panagi, S. Dasiopoulou, "Semantic Image Analysis Using a Learning Approach and Spatial Context," In Proceedings of VLDB Conf., 2003.

[12] Evaggelos Spyrous, et al., "Fuzzy Support Vector Machines for Image Classification Fusing MPEG-7 Visual Descriptors," SIGMOD Record 26(3), 2005.

[13] King-Shy Goh, Edward Y. Chang, and Beita Li, "Using One-Class and Two-Class SVMs for Multiclass Image Annotation," IEEE Transactions on Knowledge and data engineering, Vol. 17, No. 10, 2005. 10.

[14] Edward Chang, et al., "CBSA: Content-Based Soft Annotation for Multimodal Image Retrieval Using Bayes Point Machines," IEEE Transactions on Circuits and Systems for Video Technology, Vol. 13, No. 1, 2003.01.

[15] 유남현, "유비쿼터스 환경에서 멀티미디어 자료 검색을 위한 시맨틱 어노테이션에 관한 연구, 순천대학교 대학원 컴퓨터학과 박사학위논문 2006.12.