

온톨로지를 이용한 이미지의 고수준 의미 정보

자동 추출 기법

박경욱⁰ 이동호
한양대학교 컴퓨터 공학과
{kwpark⁰, dhlee}@cse.hanyang.ac.kr

Full-automatic high-level concept extraction for image using domain ontologies

Kyung-Wook Park⁰, Dong-Ho Lee
Knowledge and Data Engineering Laboratory
Dept. of Computer Science and Engineering, Hanyang University

요 약

최근 인터넷의 급속한 성장은 이미지와 같은 멀티미디어 정보의 급격한 증가를 가져왔다. 따라서 사용자 하여금 원하는 이미지를 검색하는데 있어서 좀 더 효율적이고 정확한 검색 방법의 필요성이 대두되어 왔다. 일반적으로 이미지 검색 방법에는 키워드 기반 방식과 내용 기반 방식이 존재한다. 그러나 위 두 방법은 지금의 대용량 이미지 데이터베이스 검색에 있어서 여러 문제점들을 가지고 있다. 특히, 키워드 기반 방식을 보완하기 위해서 제안되어진 내용 기반 방식의 경우, 사람이 인식할 수 있는 의미 정보가 아닌 시각 정보만을 이용하기 때문에 시맨틱 갭(semantic gap) 문제가 발생하게 된다. 본 논문에서는 이미지 객체의 시각 정보들에 대한 중간 의미값으로 구성된 시각 정보 온톨로지와 동물에 대한 분류 정보를 표현하고 있는 동물 온톨로지를 구축하고, 이를 이용하여 이미지로부터 고수준의 의미 정보를 완전 자동으로 추출하는 효율적인 방법을 제안한다.

1. 서 론

최근 인터넷 및 무선 통신 환경, 그리고 컴퓨터 장치의 급속한 발전으로 인하여 사용자가 대용량 멀티미디어 데이터베이스로부터 원하는 이미지를 검색하는데 있어서 좀 더 효율적이고 정확한 검색이 가능하도록 하는 기술의 필요성이 대두되고 있다. 그러나 기존의 전통적인 이미지 정보 검색 기술들이 최근의 급격한 환경 변화에 잘 대응하기에는 여러 문제점들을 가지고 있다. 따라서 본 논문에서는 기존의 전통적인 이미지 정보 검색 기술들이 가지고 있는 문제점들을 온톨로지를 이용하여 해결함으로써 사용자 하여금 좀 더 효율적이고 정확한 이미지 정보 검색이 가능하도록 하는데 있다.

기존의 이미지 정보 검색 방식에는 일반적으로 키워드 기반(keyword-based) 방식과 내용 기반(content-based) 방식이 존재한다. 키워드 기반 방식은 최초의 이미지 검색 방식으로서 이미지를 주제별로 분류하기 위해서 미리 정의된 단어들을 이용하여 사람이 직접 대상 이미지에다 키워드를 다는 방식이다. 그러나 키워드 기반 방식은 일반적으로 대상 이미지에 키워드를 다는데 있어서 너무 많은 시간을 소모하며 대상 이미지에 대해서 이미지가 지니는 색상(color)이나 무늬(texture) 등의 속성에 따라 사람마다 보는 관점이 서로 다를 수 있으므로 동일 이미지에 대해서 서로 다른 키워드를 달 수 있다는 문제점이 있다.

내용 기반 방식(CBIR)은 키워드 기반 방식의 문제점들을 해결하기 위하여 제안되어진 방식이다. 내용 기반 방식은 대상

이미지의 내용(즉, 시각 정보)만을 가지고 멀티미디어 데이터베이스에서 사용자가 원하는 이미지를 검색함으로써 키워드 기반 방식에 비해서 시간이 덜 소모되며 객관적일 수 있다는 장점이 있다. 그러나 내용 기반 방식도 몇 가지 문제점을 가진다. 첫째로, 이미지의 시각 정보(low-level feature)만을 이용하기 때문에 사람이 공통적으로 이해할 수 있는 좀 더 높은 수준의 개념(high-level concept)들로 해석될 수 없다. 두 번째로, 동일 객체들을 포함하는 이미지라도 빛이나 촬영 각도에 따라 이미지의 내용은 서로 다를 수 있다.

본 논문에서는 내용 기반 방식의 가장 큰 문제점 중 하나인 시맨틱 갭 문제를 온톨로지를 이용하여 해결함으로써 사용자 하여금 의미 질의가 가능하도록 하며 또한 이미지의 의미 정보를 자동으로 추출함으로써 키워드 기반 방식의 문제점을 해결하는데 초점을 맞추고 있다. 여기서 온톨로지는 높은 수준의 개념들을 구축하고 저장하기 위한 일종의 정형화된 언어이며 이미지의 시각 정보들과 고수준의 개념들 간에 맵핑(mapping)을 수행한다.

2. 관련 연구

일반적으로 이미지 안에 존재하는 객체는 해당 이미지를 판단하는데 있어서 매우 중요한 요소이다. 따라서 이미지 프로세싱 분야에서 이미지 안에 존재하는 객체를 식별하는 문제는 매우 커다란 이슈가 되어왔다. [1]의 연구에서 제안하는 시스템은

객체 온톨로지와 중간 계층 기술값(intermediate-level descriptor value)들을 기술함으로써 이미지의 시각 정보와 객체들 간에 맵핑을 도모한다. 여기서 객체 온톨로지는 그 객체를 최상위 개념으로 색상, 모양(shape)과 같은 시각 정보들을 하위 개념으로 가지는 온톨로지이다. 그리고 중간 계층값은 시각 정보의 벡터 값을 사람이 인식할 수 있는 의미값으로 변환한 것으로서 일정 범위를 주어 각각의 의미값을 부여하게 된다. 이 시스템은 온톨로지의 최상위 개념으로서 객체를 가지고 단순히 객체와 그 객체의 시각 정보들간의 is-a 관계(즉, 포함 관계)만을 표현함으로써 객체들 간의 관계는 표현할 수 없다는 문제점을 가지고 있다. 또한 각각의 객체에 대해서 이용하는 시각 정보들이 적다는 문제점이 있다. [2]의 연구에서는 중국의 고대 그림을 식별하기 위해서 온톨로지를 구축하였다. 이 온톨로지는 기본적으로 4 종류의 그림 작업(즉, 오일 페인팅, 전통 중국 화법, 예술 사진 그리고, 컴퓨터 그래픽 기법)을 포함한다. 그리고 사용자 하여금 여러 측면에서 질의할 수 있도록 다양한 종류의 개념들을 포함한다. 이미지는 이미지 처리 기법이나 패턴 분류 기술들을 통해 자동적으로 주석을 달게 된다. 그러나 이 시스템은 이용 가능한 이미지 처리 기법에 한계를 가지고 있다. 다시 말해서 이미지 처리 기법에 의존적인 문제점을 가진다.

3. 시스템 구조

본 장에서는 본 논문에서 제안하는 시스템에 대해서 자세하게 설명한다. 그리고 세 개의 MPEG-7 시각 정보 기술자(즉, 에지 히스토그램 기술자, 무늬 브라우징 기술자, 윤곽 기반 모양 기술자)를 통하여 추출된 시각 정보가 어떻게 중간 의미값으로 매칭되는지에 대해서 설명한다. 마지막으로 추출된 중간 의미값을 가지고 최종 객체를 추론하기 위하여 제안되어진 온톨로지들에 대해서 설명을 한다.

3.1 시스템 요약 설명

그림 1은 본 논문에서 제안하는 시스템의 전체적인 구조를 나타내고 있으며 수행 단계는 다음과 같다.

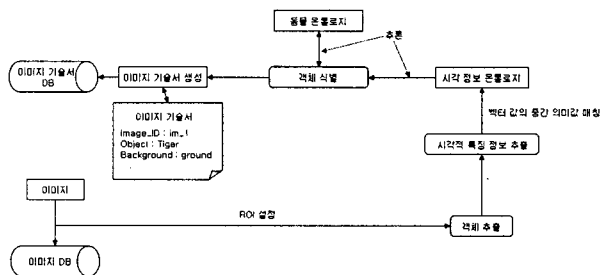


그림 1 전체적인 시스템 구조

1. 사용자는 원하는 이미지 객체에 대하여 ROI(region of interest)를 설정한다. 즉, 의미 정보를 추출하기를 원하는 이미지의 주요 객체에 대한 ROI를 설정한다.
2. ROI를 기반으로 객체가 추출되어지고 이 객체에 대하여 시

각 정보들이 추출된다.

3. 추출된 시각 정보는 중간 의미값으로 매칭되고 이 값들은 객체에 대한 시각 정보 온톨로지의 인스턴스들로서 저장된다.
4. 시각 정보 온톨로지에 규칙들을 적용해 동물 온톨로지로부터 객체를 식별한다.
5. 추론 결과 식별된 객체에 대한 정보(즉, 객체 종류, 시각 정보)와 객체 이외의 부가적인 정보(즉, 이미지 해상도, 이미지 생성 날짜 등)를 이용하여 이미지 기술서(image description)를 생성한다.
6. 최종적으로 이미지 기술서는 이미지 기술서 DB에 저장된다.

3.2 시각 정보 기술자의 중간 의미값 변환

본 논문에서는 이미지의 시각 정보들을 추출하기 위해서 MP EG-7 표준안의 시각 정보 기술자들을 이용한다. 특히, [3]의 연구에 따르면 이미지 객체를 식별하는데 있어서 하나의 시각 정보만을 이용하는 것보다 다양한 시각 정보를 이용하는 것이 검색 성능을 높일 수 있다. 따라서 이미지 객체의 무늬, 모양, 색상 등의 다양한 시각 정보들을 이용하기 위해서 에지 히스토그램 기술자, 윤곽 기반 모양 기술자, 컬러 구조 기술자를 이용한다. 또한 추출된 시각 정보 기술자의 벡터값들은 사람이 인식할 수 있는 중간 의미값으로 변환된다. 중간 의미값으로의 변환 이유는 위에서도 언급했듯이 벡터값만으로는 사람이 그 의미를 인식할 수 없으며 시각 정보 온톨로지를 이용한 객체 추론 과정에서 중간 의미값이 이용되기 때문이다. 여기서 중간 의미값이란, 특정 벡터값을 "low"나 "medium" 같이 하나의 텍스트로 표현한 것을 의미한다. 예를 들어, 이미지의 무늬 정보로서 이용되는 에지 히스토그램 기술자에 대한 대표 에지 E_1, E_2 를 구하는 공식은 다음과 같다.

A_{ij} 는 이미지 A의 i행 j열의 서브 이미지, 각각의 서브 이미지 A_{ij} 에 대한 히스토그램 $h_{ij} = \{e_0, e_1, e_2, e_3, e_4\}$ 일 때,

$$g_EH = \{E_0, E_1, E_2, E_3, E_4\}, \text{ 여기서 } E_k = \{0, 1, 2, 3, 4\} = \sum_{i=0}^3 \sum_{j=0}^3 e_k^{h_{ij}}$$

$$\therefore \text{대표 에지 } E_1^A = \max(g_EH), \quad E_2^A = \max(g_EH - E_1^A)$$

그 다음으로 추출된 대표 에지 E_1^A, E_2^A 는 다음과 같이 중간 의미값 IS_0, IS_1 으로 변환된다.

$$IS_k = \begin{cases} \text{Vertical,} & \text{if } E_k^A = e_0 \\ \text{Horizontal,} & \text{if } E_k^A = e_1 \\ 45\text{degree,} & \text{if } E_k^A = e_2 \\ 135\text{degree,} & \text{if } E_k^A = e_3 \\ \text{Nondirectional,} & \text{if } E_k^A = e_4 \end{cases}, k = \{0, 1\}$$

결과적으로 이미지 A는 에지 히스토그램 정보로서 두 개의 중간 의미값 IS_0, IS_1 을 가진다. 위와 비슷한 방법으로 나머지 두

개의 시각 정보에 대한 중간 의미값을 구하게 된다.

3.3 온톨로지 구축

Gruber[4]에 의하면 온톨로지는 특정 도메인 내의 지식(knowledge)을 개념화(conceptualization)하고 이를 명세화(specification)하는 것으로서 정의된다. 일반적으로 OWL(Ontology Web Language)은 RDF(Resource Description Framework)에 상-하위 관계 및 동등 관계, 제약 조건 등의 풍부한 표현이 가능하며 논리나 추론을 위한 정보를 표현할 수 있는 기반을 제공한다. 따라서 본 연구에서는 시각 정보 온톨로지와 동물 온톨로지를 구축하는데 있어서 OWL을 이용한다.

시각 정보 온톨로지는 중간 의미값들을 가지는 이미지 객체를 인스턴스로서 저장하기 위해서 구축되어진 간단한 온톨로지이다. 그림 2는 시각 정보 온톨로지를 보여준다.

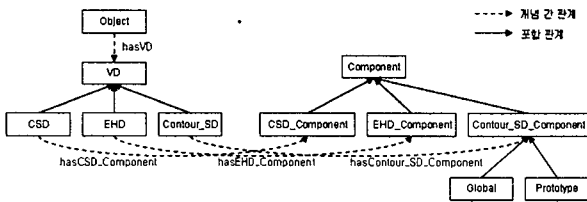


그림 2 시각 정보 온톨로지

그림 2에서 시각 정보 온톨로지는 몇 가지 클래스(예를 들어, Component, Object, VD 등)들과 클래스 간에 관계(예를 들어, hasVD, hasCSD_Component 등)를 기술한 OWL 온톨로지임을 알 수 있다. 임의의 이미지 객체의 시각 정보를 온톨로지로서 나타냄으로써 후에 설명할 'bossam' 추론 엔진을 통한 객체 추론이 가능해진다.

본 연구의 주요 도메인은 동물 이미지이다. 따라서 주요 식별 객체는 동물들이 되며 이를 위해서 동물 온톨로지를 구축하였다. 그리고 동물 온톨로지는 기본적으로 생태 분류학을 따른다. 또한 이것은 일종의 어휘 사전과 비슷하다. 따라서 이 온톨로지를 구성하는 클래스 간에는 어떤 관계 및 속성도 가지지 않는다.

3.4 추론

시각 정보 온톨로지로부터 객체를 식별하기 위해서 온톨로지 기반 추론 엔진인 "bossam" 추론 엔진[5]을 이용한다. 시각 정보 온톨로지와 동물 온톨로지로부터 임의의 이미지 객체 Object₁을 추론하는 과정은 다음과 같다.

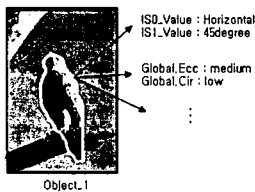


그림 3 Object₁ 객체

1. 규칙을 정의한다. 예를 들어 'Eagle' 객체를 식별하기 위한 규

칙은 다음과 같이 정의될 수 있다.

$$\begin{aligned} \text{rule_1 is } & \text{hasVD}(?x, ?y) \wedge \text{hasEHD_Component}(?y, ?z) \\ & \wedge \text{ISO_Value}(?z, \text{Horizontal}) \wedge \text{IS1_Value}(?z, 45\text{degree}) \\ \Rightarrow & \text{Eagle}(?x) \end{aligned}$$

여기서 rule₁은 임의의 이미지 객체가 대표 예지 ISO, IS1의 값으로서 'Horizontal', '45degree'를 가질 경우, 이 객체는 'Eagle'이라는 규칙이다.

2. 그림 3에서 임의의 이미지 객체 Object₁은 대표 예지 ISO, IS1의 값으로서 'Horizontal', '45degree'를 가지므로 이미지 객체 Object₁은 규칙 rule₁을 만족한다. 따라서 객체 Object₁은 'Eagle'이라는 것을 알 수 있다.

3. 아래 RDF 문서는 추론 결과 생성된 문서이다. 12~13번째 줄에서 Object₁의 rdf:type은 'Eagle'임을 알 수 있다.

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <!DOCTYPE rdf:RDF [
3 <ENTITY ns_1 "http://kde.hanyang.ac.kr/ontology/Animals#" >
4 <ENTITY ns_0 "http://kde.hanyang.ac.kr/ontology/VDO#" >
5 ]>
6 <rdf:RDF
7   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
8   xmlns:ns_1="http://kde.hanyang.ac.kr/ontology/Animals#"
9   xmlns:ns_0="http://kde.hanyang.ac.kr/ontology/VDO#"
10  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
11 >
12 <rdf:Description rdf:about="&ns_0:Object_1">
13 <rdf:type rdf:resource="&ns_1:Eagle"/>
14 </rdf:Description>
15 </rdf:RDF>
    
```

4. 추론 결과 생성된 RDF 문서는 표현 방식의 통일성을 위해서 OWL 문서로 변환된다.

4. 결론 및 향후 계획

본 논문에서는 임의의 이미지 객체에 대하여 추출된 시각 정보를 중간 의미값으로 변환하고 이 값들을 온톨로지로서 매칭함으로써 객체를 자동적으로 식별하는 방법에 대해서 제안하였다. 그러나 이미지의 의미 정보에는 객체 뿐만 아니라 배경 정보도 함께 포함되어야 한다. 따라서 현재 배경 정보도 포함하기 위한 방법에 관하여 연구 중이다.

5. 참고 문헌

- [1] Vasileios Mezaris, Ioannis Kompatsiaris, and Michael G. Strintz, "Region-based Image Retrieval using an Object Ontology and Relevance Feedback", EURASIP JASP, 2004
- [2] Shuqiang Jiang, Tiejun Huang, Wen Gao, "An Ontology-based Approach to Retrieve Digitized Art Images", Web Int-elligence, pp 131-137, 2004
- [3] Spyrou E., Le Borgne H., Mailis T., Cooke E., Avrithis Y. and O'Connor N, "Fusing MPEG-7 visual Descriptors for image classification", International Conference on Artificial Neural Networks, (ICANN 2005), pp 11-15, 2005
- [4] Thomas R. Gruber, "A Translation Approach to Portable Ontology Specifications." Knowledge Acquisition, 5(2), pp199-220, 1993
- [5] Minsu Jang, Joo-Chan Sohn, "Bossam: An Extended Rule Engine for OWL Inferencing", RuleML 2004, pp128-138. 2003