

SVM과 온톨로지를 이용한 이미지 의미 관계 자동 추출 기법

정진우^{0*} 주영도^{**} 이동호^{*}

한양대학교 컴퓨터공학과*, 강남대학교 컴퓨터미디어공학부**
{jwjeong⁰, dhlee72}@cse.hanyang.ac.kr, ydjoo@kangnam.ac.kr**

Automatic Extraction of Semantic Relationships from Images Using Ontologies and SVM Classifiers

Jin-Woo Jeong^{0*}, Joo, Young Do, ** Dong-Ho Lee^{*}

Dept. of Computer Science & Engineering, Hanyang University*
Division of Computer and Media Engineering, Kangnam University**

요 약

효과적인 이미지 검색을 위하여, 이미지의 저수준 시각 정보로부터 고수준 의미 정보를 추출하는 기술에 관한 많은 연구가 이루어지고 있다. 특히 최근에는 Support Vector Machine과 같은 기계 학습 기법을 이용한 이미지 어노테이션 시스템의 개발이 활발히 진행중이다. 그러나 기존의 연구들은 단편적인 이미지 정보만을 추출함에도 불구하고, 그 성능이 여전히 만족스럽지 못하다

본 논문에서는 Support Vector Machine과 온톨로지를 이용하여 이미지의 다양한 정보를 효과적으로 추출 및 기술할 수 있는 시스템을 제안한다. 특히 온톨로지는 특정 도메인의 상세한 지식 표현과 추론을 위한 지식베이스로서, 본 논문에서는 Support Vector Machine을 이용하여 이미지 안에 존재하는 객체들의 컨셉을 판별하고 이미지 어노테이션 온톨로지와 생태계 온톨로지를 이용하여 공간 관계, 천적 관계와 같은 객체간 의미 관계를 자동적으로 추출하는 방법을 제안한다.

1. 서 론

일반적으로, 사용자들은 원하는 이미지를 얻기 위하여 이미지 내에 포함되어 있는 개념들을 기반으로 이미지를 검색한다. 그러나 기존의 내용 기반 검색 기법[1,2]들은 ‘의미적 불일치’라 불리는 치명적인 문제점을 안고 있기 때문에, 고수준 개념을 이용한 이미지 검색 시, 만족스러운 성능을 발휘할 수 없다. ‘의미적 불일치 문제’는 이미지간의 시각적 유사성이 결코 의미적 유사성을 의미하지는 않는다는 점을 지적하고 있다.

따라서, 이러한 문제점을 해결하기 위하여, 기계 학습 기법을 이용한 다양한 연구들이 진행되고 있다[3,4]. 특히, 최근에는 Support Vector Machine (SVM) 기반의 자동 이미지 어노테이션 기법들이 제안되고 있다[5,6]. SVM은 감독 기계 학습 기법의 한 종류로써, 우수한 데이터의 분류 성능을 보인다.

그러나 기존의 SVM 기반의 기술들은 다양한 고차원 의미 정보들을 지원하지 못한다는 한계를 가진다. 예를 들어, 기존의 시스템들은 “호랑이의 왼쪽에 얼룩말이 서 있는 이미지를 찾아라” 라는 질의를 처리할 수 없다. 이러한 한계점들을 극복하기 위하여, 우리는 SVM과 온톨로지를 이용하여 이미지 내의 객체의 컨셉 뿐만 아니라 객체들간의 위치관계, 천적관계와 같은 고차원 의미 정보를 추출하고 기술할 수 있는 새로운 프레임워크를 제안한다.

특히 온톨로지는 특정 분야에 존재하는 컨셉들에 대한 상세한 정의와 컨셉들간의 관계에 대하여 상세히 정의를 제공한다. 특히, 이미지 검색의 측면에서는, 온톨로지를 이용하여 이미지 내에 존재하는 다양한 지식 정보들을 상세하게 기술함으로써, 사용자로 하여금 원하는 이미지들을 보다 쉽게 찾을 수 있도록 도움을 줄 수 있다.

본 논문에서 제안하는 시스템은 크게 이미지 분류 단계와 의미 처리 단계로 이루어진다. 첫 째, 이미지 분류 단계에서, 이미지의 저수준 시각 정보들이 추출되어진 후 색상/무늬 SVM 분류자에 의하여 ‘노랑’, ‘75각도 무늬’와 같은 중간수준의 컨셉으로 분류되어진다. 그 후에, 이러한 중간수준 컨셉의 집합들이 객체 SVM 분류자에 의하여 ‘호랑이’, ‘말’과 같은 객체컨셉으로 분류되어진다. 특히, 이 과정에서 SVM이 잘못된 분류결과를 도출 할 수도 있기 때문에, 이러한 오류를 수정할 수 있는 수정모듈을 제안한다. 두번째로, 의미 처리 모듈에서는 이미지 분류단계에서 생성된 각종 시각 정보들과 이미지 어노테이션 온톨로지, 생태계 온톨로지를 이용하여 이미지 안에 존재하는 객체들 간의 다양한 고수준 의미 관계를 추론한다. 예를 들어, 동물 객체간의 천적 관계 혹은 이미지 안에 존재하는 객체들 간의 절대적/상대적 공간 관계 등을 추론하게 된다.

본 논문의 구성은 다음과 같다. 제 2장에서는 기존

연구들의 문제점들을 기술한다. 제3장에서는 제안하는 시스템의 구조와 이미지의 컨셉, 고수준 의미 정보를 추출하는 방법에 관하여 자세히 기술하고, 제 4장에서 실험 결과에 대한 분석을 기술한다. 마지막으로, 제 5장에서 결론과 함께 향후 연구 방향에 대하여 기술한다.

2. 관련 연구

자동 이미지 어노테이션을 위하여 SVM을 기반으로한 다양한 기술들이 제안되어 왔다.

Jianping Fan [5]은 n개의 객체 컨셉을 식별하기 위하여 n 개의 컨셉 검출 함수를 생성하였다. 각각의 검출 함수는 이미지가 해당 컨셉에 적합한지 아닌지를 판단하는 분류자이다. 그러나 이러한 방법은 각각의 컨셉 검출 함수에 대한 최적의 파라미터를 찾아야만 한다는 문제점을 지닌다. 또한, 트레이닝 데이터와 객체 컨셉의 수가 증가함에 따라 높은 계산 복잡성을 보이게 된다.

Tsai [7]은 다수의 SVM 분류자를 이용한 2-단계 학습 프레임워크를 제안하였다. 먼저, 각각의 이미지들은 같은 크기를 가지는 5개의 부분이미지로 분할된다. 그 후에, 각각의 부분 이미지로부터 저수준 시각 정보들이 추출된다. 첫 번째 매핑 단계에서, 이러한 저수준 정보들은 ‘빨강’과 같은 적합한 용어로 분류되고, 두 번째 매핑 단계에서, SVM을 통하여 이러한 용어들의 집합을 ‘호랑이’와 같은 객체 컨셉으로 분류시킨다.

비록 이러한 방법들이 보다 효과적이고 정확한 이미지 어노테이션을 위한 시도를 하였지만, 다음과 같은 한계점들을 지니고 있다. 첫째, 위의 기술들은 컨셉 별 대표 특징요소에 관한 고려가 없기 때문에, SVM에 의한 분류 성능을 떨어뜨릴 수 있다. 일반적으로, 각 컨셉은 그들만의 대표 특징요소를 가질 수 있다. 예를 들어, 호랑이들은 주황/검정 줄무늬를 자신들의 대표 특징요소로 가진다. 따라서, 이러한 대표 특징 요소에 대한 고려는 유사한 시각 정보를 가지는 객체들 중에서 특정 객체를 보다 효과적으로 구별할 수 있도록 도움을 줄 수 있다. 둘째, SVM을 이용한 기존의 어노테이션 시스템들은 공간 관계나 천적 관계와 같은 고수준의 의미 정보들을 포함하지 않는다. 따라서 이미지 검색 시, 이미지 내의 객체 간 의미 관계를 이용한 질의를 사용할 수 없다. 예를 들어, “얼룩말과 천적 관계에 있는 객체를 포함하는 이미지를 찾아라”와 같은 질의를 처리할 수 없게 된다. 결론적으로, 최근의 SVM을 이용한 연구들은 이미지로부터 추출된 정보에 귀납/연역과 같은 어떠한 지식 가공을 하지 않기 때문에, 이미지의 단편적인 정보만을 추출한 것에 불과하게 된다.

따라서, 보다 풍부하고 고차원적인 의미 정보를

추출하기 위하여 이미지 정보에 대한 지식 가공절차가 반드시 요구되며, 본 논문에서는 온톨로지와 컨셉 별 대표 특징 요소에 대한 고려를 포함한 SVM 분류자를 이용하여 다양한 고수준 의미정보들을 자동으로 추출하는 2단계 프레임워크를 제안한다.

3. 시스템 구조

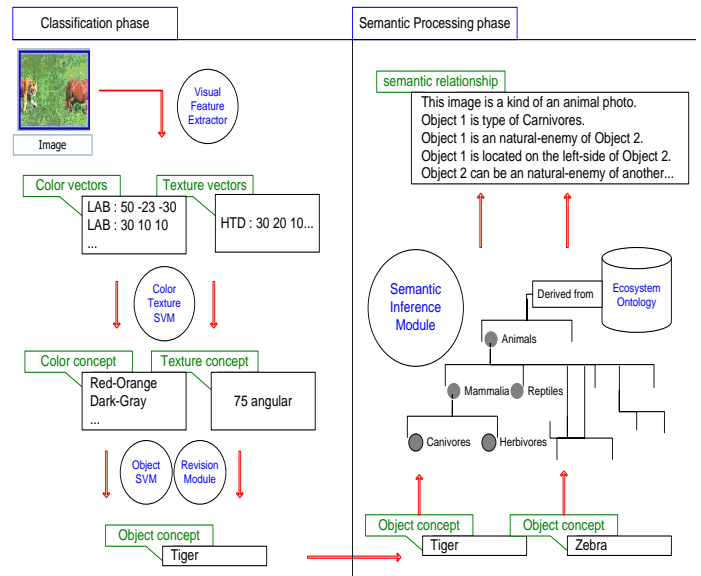


그림 1 프레임워크

제안하는 시스템의 전체적인 구조는 그림 1과 같다. 우리의 시스템은 크게 이미지 분류 단계와 의미 처리 단계로 구성되어 있다. 이미지 분류 단계에서는 이미지로부터 시각 정보들을 추출한 후, 적절한 객체 컨셉으로 분류시킨다. 특히, 이 과정에서 수정모듈을 통하여 SVM 분류자가 잘못 분류할 가능성이 있는 결과들을 올바르게 교정한다. 이미지 분류 단계에서 생성된 정보들은 의미 추론 모듈로 전달되고 이미지 어노테이션 온톨로지와 도메인 온톨로지를 이용하여 다양한 의미 정보들을 추론하게 된다.

3.1 이미지 판별 단계

언급한 바와 같이, 각 개체는 고유의 대표 특징 요소들을 가질 수 있다. 본 연구에서는, 개체의 컨셉을 식별하기 위한 데이터로서, CIE-LAB로 기술되는 4개의 대표 색상 정보와 MPEG-7 Homogeneous Texture Descriptor(HTD)로 기술되는 하나의 대표 무늬 정보를 사용한다. 또한, 수정 모듈을 위하여 각 개체의 색상/무늬 히스토그램도 추출된다.

트레이닝 이미지가 주어지면, 각 이미지는 먼저 8x8의 블록들로 분할되어 진다. 그 후, 각 블록으로부터 대표 색상/무늬 정보가 추출되어진다. 여기서, 대표 색상과 대표 무늬는 각각 Wu-양자화 알고리즘[8]과 MPEG-7 HTD를 통하여 추출된다. 그림 1에서 볼 수

있듯이, 추출되어진 저수준 시각 정보들은 먼저 색상/무늬 SVM 분류자를 통하여 ‘노랑’, ‘75각도 무늬’와 같은 중간 단계의 색상/무늬 컨셉으로 분류되고, 이러한 중간 단계 컨셉들이 하나의 집합을 이루어 적절한 객체 컨셉으로 분류된다. 예를 들어, 그림 1의 호랑이 이미지의 경우, 먼저 이미지로부터 4개의 대표 색상 벡터와 1개의 무늬 벡터가 추출된다. 그림 1에서 볼 수 있듯이, 추출된 저수준 시각 정보는 [50, -23, -30], [30 20 10 ... 30] 과 같은 벡터 형태로 표현되고, 색상/무늬 SVM 분류자를 거쳐 [주황, 선홍, 회색, 검정, 75 각도 무늬]와 같은 하나의 중간 단계 컨셉 집합을 생성하게 된다. 객체 SVM 분류자는 [주황, 선홍, 회색, 검정, 75 각도 무늬, (호랑이)]와 같은 중간단계 컨셉의 집합과 해당 객체 컨셉의 레이블로 구성되는 트레이닝 데이터들을 기반으로 저수준 시각 정보로부터 객체 컨셉을 식별 해낸다.

일반적으로 SVM분류자는 분류 결과로서 하나의 예상 컨셉과 각 컨셉에 해당되는 컨피던스 값을 생성한다. 컨피던스는 이미지와 이미지 DB안에 저장되어 있는 각각의 객체 컨셉 사이의 연관성을 나타내는 값이다. 예를 들어, 이미지 i 의 분류 결과로서 각 컨셉 {말, 늑대, 사자, 곰, 기린}에 대한 컨피던스 값 {0.7, 0.15, 0.05, 0.07, 0.03}이 생성되었다고 하자. 이 경우, SVM 분류자는 객체 컨셉 ‘말’을 이 이미지의 예상 컨셉으로 결정한다. 컨셉 ‘말’에 대한 컨피던스 값이 다른 값들에 비해 월등히 높기 때문에, 우리는 이 결정이 신뢰적이라고 말할 수 있다. 그러나 이미지 j 의 분류결과로서 {0.35, 0.30, 0.12, 0.27, 0.03}이 생성되고, 예상 컨셉으로 ‘말’이 결정된 경우를 생각해보자. 그러나 이 경우는 상위 3개의 컨피던스 값들의 차이가 거의 존재하지 않는다. 이것은 곧 이 이미지의 시각정보가 객체 컨셉을 뚜렷하게 식별할 만큼 특징적이지 못하다는 것을 의미하고, 이는 SVM 분류자가 충분히 잘못된 결과를 도출할 수도 있음을 의미한다.

따라서, 본 논문에서는 이러한 경우의 분류 결과를 보이는 상황에 대하여 각 객체 별 특징 정보에 대한 가중치를 부여, 객체 간 차이를 극대화시킴으로써 올바른 컨셉을 찾도록 유도한다.

```

Algorithm_1 ConceptDecision(p,conf[]) returns Object Concept
// 입력 값 p는 객체SVM에 의해 예상된 컨셉, conf[]는 객체SVM의
  분류 결과로 함께 생성되는 컨피던스값들을 의미한다
  candidates <- highly ranked 5 values of conf[]
  std_cand <- stddev(candidates)
  If(std_cand < Th1 && confidence(p) < Th[p])
  Then return revision(p); // 수정모듈을 수행한다
  Else return p;
    
```

그림 2 컨셉 결정 알고리즘

그림 2는 객체 SVM의 분류 결과로 인해 생성되는 예상 컨셉 p 와 컨피던스 값들로부터 최종적인 객체

컨셉을 결정하는 과정을 보여준다. 객체 SVM의 결과를 수정하기 위한 과정은 다음과 같다. 먼저, 분류 결과로 생성된 컨피던스 값들 중 상위 다섯 개의 표준편차를 계산한다. 이 표준편차 값이 특정 기준치 $Th1$ 보다 작게 되면, 앞서 설명한 두 번째 예제와 부합되기 때문에, 수정과정을 진행한다. 그러나 예상 컨셉 p 에 대항하는 컨피던스 값이 기준치 ThP 보다 크다면, 수정 과정은 진행되지 않는다. ThP 는 객체 SVM 분류자의 결과의 신뢰성을 보장하는 최소한의 컨피던스 값을 의미한다.

객체 컨셉의 수정은 식(1)과 같은 객체별 특징 정보요소에 대한 가중치를 고려한 수정된 유클리디안 거리 측정 함수를 사용, 가장 적은 차이를 보이는 컨셉을 객체의 컨셉으로서 결정함으로써 이루어진다.

$$\sum (W_x^c \cdot R_x^c - W_x^c \cdot I_x^c)^2 + (W_x^t \cdot R_x^t - W_x^t \cdot I_x^t)^2 \quad (1)$$

$$W_x^c = 1 - (bin_x^c / Area_c)^2, W_x^t = 1 - (bin_x^t / Area_t)^2 \quad (2)$$

여기서, $W_x^{c,t}$ 는 각각 색상/무늬 히스토그램에서의 특정 빈 x 의 가중치로서, 식(2)에서처럼 각각의 히스토그램에서 특정 빈 x 가 차지하는 비율을 계산함으로써 얻어진다. 식(1)에서 $R_x^{c,t}, I_x^{c,t}$ 는 각각 특정 컨셉을 가지는 이미지들의 평균 색상/무늬 히스토그램, 테스트 이미지의 색상/무늬 히스토그램 상에서의 특정 빈 x 를 나타낸다.

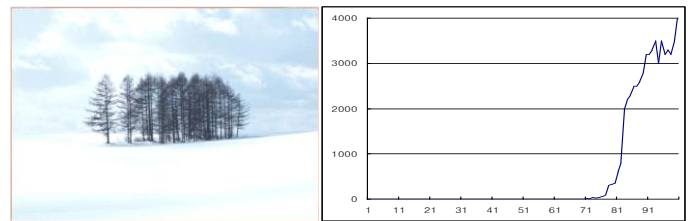


그림 3 설원 이미지와 그에 해당하는 조도 히스토그램

즉, 특정 컨셉을 가지는 이미지들의 히스토그램을 분석하여 각 히스토그램에서 많은 비율을 차지하는 빈들에 가중치를 두고, 이 가중치를 고려하여 분류하고자 하는 이미지와 각 컨셉들의 차이를 계산한다. 예를 들어, 그림 3과 같은 설원 이미지의 경우에는 CIE-LAB 색상 공간 중 조도에 해당하는 L 히스토그램이 특이성을 보이게 된다. 따라서 이러한 특이성을 반영할 수 있도록 가중치를 부여, 최종적으로 이미지 어노테이션 시스템의 정확도를 높이게 된다.

3.2. 의미 처리 모듈

이 절에서는 이미지 분류단계를 통하여 생성된 정보들과 이미지 어노테이션 온톨로지, 도메인 온톨로지를 이용하여 공간관계, 천적관계와 같은 고수준 정보를 추론하는 과정을 기술한다.

그림 1에서 볼 수 있듯이, 이미지 분류 단계에서 생성되어진 모든 정보들은 의미 처리 모듈로 전달되게 된다. 그 후, 의미 추론 모듈은 이미지 어노테이션 온톨로지와 생태계 온톨로지(도메인 온톨로지), 그리고 이미지 분류단계에서 얻어진 정보들을 기반으로 다양한 고수준 정보를 추론한다.

효율적인 정보 및 지식의 추론/관리를 위해서는 이를 위한 효율적인 지식 구조 / 지식 베이스가 필요하다. 따라서 우리는 그림 4와 같은 이미지 어노테이션 온톨로지를 제안한다. 이미지 어노테이션 온톨로지의 중요한 클래스에 대한 설명은 다음과 같다.

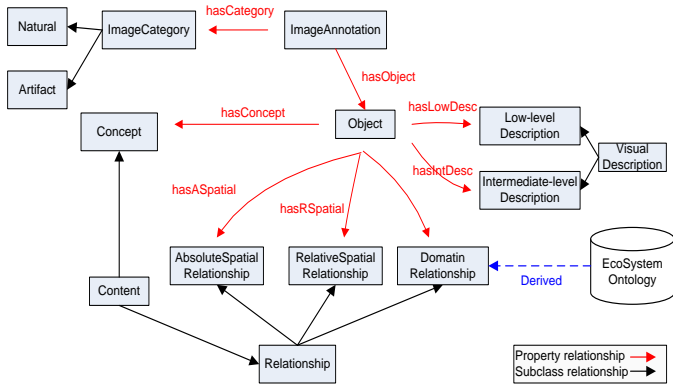


그림 4 이미지 어노테이션 온톨로지

- *ImageAnnotation* 클래스는 이미지를 기술하기 위한 기본 유닛이 되는 클래스로, 이미지 카테고리 기술하는 *ImageCategory* 클래스와 이미지 안에 존재하는 객체들에 대한 정보를 담고 있는 *Object* 클래스와 관계를 맺고 있다.

- *Object* 클래스는 해당 객체의 시각 정보들을 담고 있는 *VisualDescription* 클래스와 객체 개념에 대한 정보를 담고 있는 *Concept* 클래스, 그리고 각종 의미 관계들을 기술할 수 있는 *Relationship* 클래스와 연결되어 있다.

- *Concept* 클래스는 해당 개념의 정의, 유사어, 동의어, 계층적 레벨등을 기술하는 클래스로서, 이러한 정보들은 워드넷 2.1[9]로부터 파생된다.

- *Relationship* 클래스는 객체 간에 존재할 수 있는 관계들에 대하여 기술하는 클래스이다. 예를 들어, 하나의 이미지 안에 존재하는 객체들 간의 상대적인 공간관계, 친적 관계에 대하여 기술하게 된다. 특히, 친적 관계와 같은 도메인에 종속적인 관계성들은 도메인 온톨로지에 존재하는 관계성 및 인스턴스들을 상속받아 기술할 수 있다.

본 시스템은 이미지에 다양한 의미 추론 규칙들을 적용시킴으로써 상응하는 고수준 의미 관계들을 자동 추출한다. 이를 위하여 *부침개*라는 룰 언어를 제공하는 *보쌈* 룰 엔진[10]을 사용한다.

3.2.1 상대적 위치 관계 추론

이미지에 다수의 객체가 존재할 경우, 더욱 복잡한 질의를 이용한 검색이 가능하다. 예를 들어, “곰의 왼편에 호랑이가 위치하는 이미지를 찾아라.”와 같은 질의가 가능하다. 상대적 공간 관계는 *isLeftOf*, *isNorthOf*와 같은 키워드로 나타낼 수 있다. 이러한 상대적 공간 관계를 추론하기 위하여, 먼저 각 객체가 차지하는 공간을 MBR(Minimum Bounding Rectangle)로 표시한다. 그 후에, 공간 관계 추론을 위한 추론 규칙들을 객체들에 적용함으로써 그들간의 상대적 공간 관계를 추론한다. 두 객체 A, B가 그림 6의 1)과 같이 위치할 경우, 객체 A가 객체 B의 위에 위치한다는 사실은 다음과 같은 룰을 적용함으로써 추론할 수 있다.

```

Rule isNorthOf is:
  MBR(?A), MBR(?B)
  LowerRightY(?A, ?y1), UpperLeftY(?B, ?y2)
  IF isLessThan(?y1, ?y2)
  Then isNorthOf(?A, ?B)
    
```

그림 5 객체간 공간 관계 추론을 위한 추론규칙

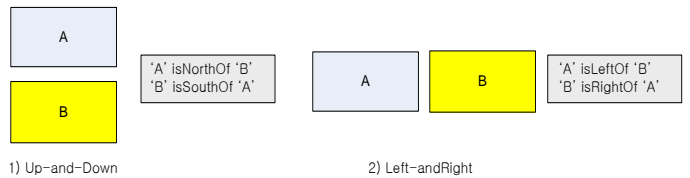


그림 6 객체 간 상대적 공간 관계

유사한 방법으로, *isSouthOf*나 *isLeftOf*와 같은 공간 관계도 쉽게 추론해낼 수 있다.

3.2.2 이미지 카테고리 추론

```

Rule isNaturalScene is:
  If hasObject(?x, ?y) and hasConcept(?y, ?z) and
  hasHierarchicalLevel(?z, ?w) and
  contains(?w, 'living thing') or
  contains(?w, 'NaturalPhoenomenon')
  Then isNaturalScene(?x)
    
```

그림 7 이미지 카테고리링을 위한 추론 규칙

이미지 카테고리링을 위한 추론 규칙들을 이미지들에 적용함으로써 각 이미지의 카테고리를 추론할 수 있다. 예를 들어, “풀밭 위에 앉아 있는 호랑이” 이미지는 *Natural-Scene* 이미지와 *Animal* 이미지로 분류될 수 있다. *Natural-Scene* 이미지에 대한 추론은 어떤 이미지 안에 존재하는 객체들이 자연적 현상/사건에 관련된 객체이거나 생물 객체일 경우, 그 이미지는 *Natural-Scene* 을 묘사한다는 사실을 기반으로 이루어진다. 그림 7은 *보쌈* 룰 엔진의 *부침개* 룰 언어를 이용하여 *Natural-Scene* 을 추론하기 위한 추론규칙을 보여준다.

3.2.3 천적 관계 추론

본 연구에서는, 그림 8과 같은 생태계 온톨로지를 이용하여 생태계와 관련된 고수준 의미 정보를 추론하는 과정에 대하여 기술한다. 생태계 온톨로지와 같은 도메인 종속적인 온톨로지는 직접 제작하여 사용하거나, 기존에 존재하는 온톨로지를 상속하여 사용할 수도 있다. 본 연구에서 사용하고자 하는 생태계 온톨로지에는 객체의 종(種), 포식자와 같은 정보가 상세히 기술되어 있기 때문에, 이미지 분류 단계에서 추출된 객체 컨셉을 기반으로 다양한 관계들을 추론해 낼 수 있다

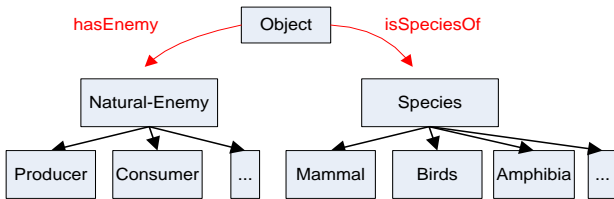


그림 8 생태계 온톨로지

그림 8은 생태계에 존재하는 개체들에 대한 상세한 정의와 각 개체들간의 관계를 기술하고 있는 온톨로지로서, 그림에서 볼 수 있듯이 각 개체의 생산자 / 소비자, 종(種)과 관련된 정보가 기술된다. 예를 들어, 이미지 내에 얼룩말 객체가 존재할 경우, 얼룩말의 천적은 호랑이, 사자와 같은 개체이며 얼룩말의 주식은 풀이라는 정보가 이미지 어노테이션 온톨로지에 추가된다.

```

1 <ImageAnnotation rdf:ID = "Image_001">
2   <hasImageCategory rdf:ID = "Animals"/>
3   <hasObject rdf:resource = "#Object_01"/>
4   <hasObject rdf:resource = "#Object_02"/>
5 <Object rdf:ID = "Object_01">
6   <hasConcept rdf:resource = "#Zebra"/>
7   <hasDomainRelship rdf:resource = "#ecosystem_zebra"/>
8   ...
9 </Object>
10 <DomainRelship rdf:ID = "ecosystem_zebra">
11   <isConsumerOf rdf:resource = "#Grass_01"/>
12   <isProducerOf rdf:resource = "#Tiger"/>
13   ...
14 </DomainRelship>
    
```

그림 9 이미지 디스크립션 OWL 문서

그림 9는 위의 이미지 분류 모듈과 의미 추론 모듈을 통하여 생성된 정보들이 이미지 어노테이션 온톨로지를 통하여 이미지에 대한 정보를 기술하는 모습을 보여준다. 라인 5-9에는 수정 모듈을 적용한 이미지 분류 모듈의 분류 결과로서 추론된 객체의 컨셉인 얼룩말에 대한 정보와 해당하는 관계성들이 기술되어 있다. 또한 라인 10-14를 통하여 얼룩말의 생태계적 관계에 놓여져 있는 호랑이, 풀 개체에 대한 정보 또한 기술되어 있음을 볼 수 있다. 이미지 디스크립션 OWL 문서에는 객체들의 컨셉과 같은 개념적인 정보 뿐만 아니라 객체들의 저수준 시각정보, 중간단계 컨셉정보, 그림의 크기나 해상도와 같은 정보들 또한 기술되어진다. 사용자들은 이러한

정보를 바탕으로 내용기반 질의(이미지의 색상이나 모양과 같은 시각적인 정보를 이용한 질의)를 이용한 이미지 검색을 시도할 수 있다. 이러한 이미지 디스크립션 OWL 문서에 기반하여 우리가 제안하는 시스템은 다양한 의미 질의에 대한 처리를 수행한다.

4. 성능 평가 및 분석

이 장에서는 제안한 이미지 분류 모듈의 성능을 실험/분석하고 의미 질의 처리 과정이 진행되는 모습을 간략하게 살펴본다. 성능 평가를 위한 실험은 20여개의 컨셉을 포함하는 6000개 이상의 컬러 이미지 데이터베이스를 기반으로 수행되었다.

4.1 이미지 분류 모듈 평가

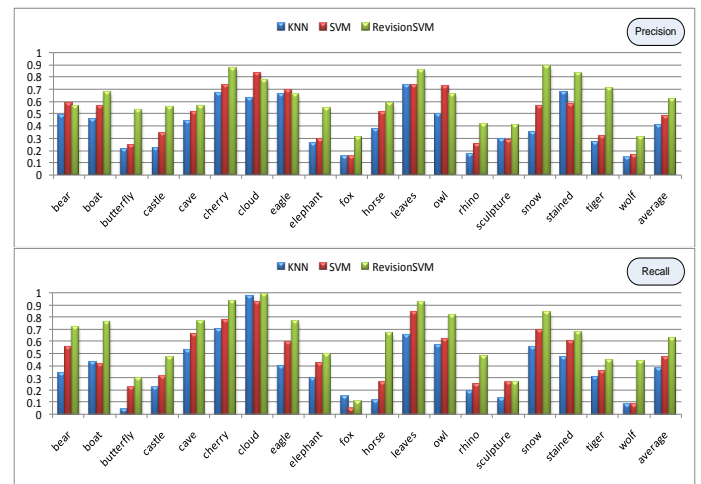


그림 10 이미지 분류 모듈의 성능평가

보다 객관적인 성능 평가를 위하여, 본 논문에서 제안하는 수정 모듈을 첨가한 SVM을 이용한 분류 모듈, 수정 모듈을 첨가하지 않는 SVM을 이용한 분류 모듈, 그리고 효율적인 데이터 분류 알고리즘으로 잘 알려져 있는 K-Nearest Neighbor (KNN) 알고리즘을 비교 분석하였다. 또한, 수정 모듈을 위한 기준치 $Th1$, ThP 는 각각 0.35와 0.25로 설정하였다. 이미지 분류 모듈의 성능은 이미지 어노테이션의 정확도 ρ 와 재현율 r 를 바탕으로 측정되었다.

$$\rho = \eta / (\eta + \varepsilon) , r = \eta / (\eta + \vartheta) \quad (3)$$

식(3)은 정확도와 재현율이 의미하는 바를 나타내는 공식이다. 식 (3)에서 η 은 검색된 이미지들 중 올바르게 어노테이션된 이미지들의 개수, ε 는 검색된 이미지들 중 올바르게 어노테이션 되지 않은 이미지들의 개수, ϑ 는 올바르게 어노테이션 되었으나 검색되어 지지 않은 이미지들의 개수를 나타낸다.

각 알고리즘들에 대한 성능 평가는 그림 10에 나타나 있다. 그림에서 볼 수 있듯이, 본 연구에서 제안한 수정 모듈을 포함하는 SVM모듈은 평균적인 정확도와 재현율 면에서 SVM 단독사용모듈에 비해서 30%의 성능 향상을, KNN 알고리즘에 비해서 50%의 성능 향상을 보임을 알 수 있다.

실험 결과에 따르면, 나비, 체리, 나뭇잎, 눈, 호랑이와 같은 객체는 기존 SVM에 의한 분류 성능을 탁월하게 향상시킴을 볼 수 있다. 이는, 위에 언급한 객체들의 경우, 객체 고유의 시각적인 정보가 매우 특징적이기 때문이며, 이러한 특징들에 대한 가중치를 고려한 수정 모듈을 통하여 효과적으로 분류 결과가 개선되었음을 의미한다. 그 밖에, 여우, 사슴, 조각상과 같이 상대적으로 덜 특징적인 시각정보를 가진 객체들에 대해서도 기존의 KNN 알고리즘과 SVM만을 단독으로 이용한 분류 결과보다는 우수한 분류 정확성을 보임을 알 수 있다.

4.2 의미 처리 모듈 과정

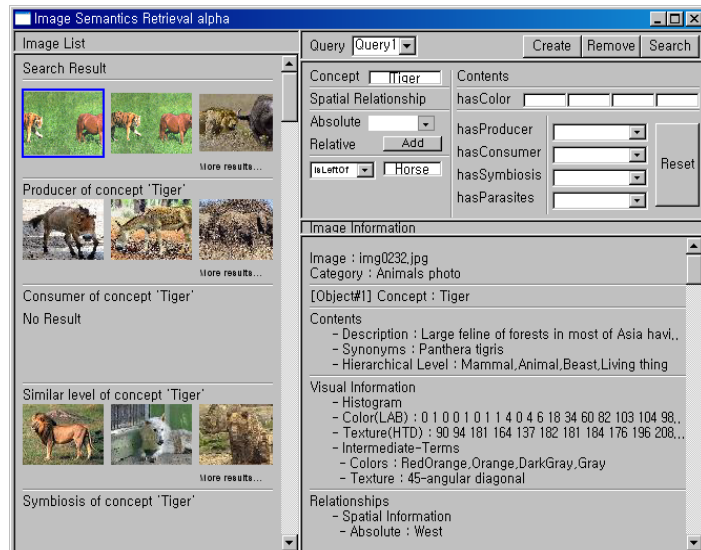


그림 11 의미 처리 과정의 예

그림 11은 본 시스템의 프로토타입으로서, 의미 질의에 대한 결과를 나타낸다. 사용자가 “말의 왼쪽에 호랑이가 위치하고 있는 이미지를 찾아라” 라는 질의를 이용하여 이미지를 검색하게 될 경우, 시스템은 이미지 데이터베이스에 존재하는 이미지들 중, 호랑이 객체를 가지고 있으며, 이 호랑이 객체의 *Relative Spatial Relationship* 클래스의 속성 중 *isLeftOf*의 값이 같은 이미지에 존재하는 말 객체인 이미지들을 검색하게 된다.

그림 11에서 볼 수 있듯이, 질의를 만족하는 이미지들의 리스트가 *Search Result* 패널에 나타나게 된다. 또한, 질의하고자 하였던 컨셉 호랑이와 관련있는 이미지들도 자동적으로 검색하여 결과 이미지들을 보여주게 된다. 예를 들어, 먹이 사슬관계에서 호랑이와 비슷한 레벨에 있는 사자, 곰과 같은 객체의 이미지들이 추가적으로 검색된다. 사용자는 검색된 이미지들 중 하나를 클

릭하여 이미지에 담긴 객체에 대한 상세한 정보를 *Image Information* 패널에서 확인할 수 있다.

5. 결론

본 논문에서, 우리는 풍부한 이미지 어노테이션을 위하여 SVM과 온톨로지를 이용한 새로운 프레임워크를 제안하였다. 본 연구에서는 SVM과 수정 모듈을 통하여 이미지 객체의 컨셉을 구분하고, 추론 규칙들을 이미지에 적용함으로써 고수준 의미 관계들을 추출하였다. 본 연구는 다음과 같은 장점을 가진다. 첫째, 본 시스템은 이미지 DB상의 다양한 의미 프로세싱을 지원한다. 둘째, 본 연구에서 제안한 수정 모듈을 기존의 SVM을 이용한 이미지 어노테이션의 정확성을 더욱 향상시킬 수 있다. 셋째, 이미지 어노테이션 온톨로지는 도메인 독립적이기 때문에, 다른 분야에도 쉽게 적용시킬 수 있다. 넷째, 트레이닝 데이터와 추론 규칙을 추가하는 것 외에는 모든 과정이 자동적으로 진행된다.

향후 과제로서, 더욱 방대한 데이터와 다양한 정보들을 가지는 비디오 데이터베이스를 위한 자동 어노테이션 기술들에 관하여 연구하고자 한다.

6. Acknowledgement

이 논문 또는 저서는 2006년 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임”(KRF-2006-521-D00457).

7. 참고 문헌

- [1] John R. Smith and Shih-Fu Chang: VisualSEEK: a fully automated content-based image query system. ACM Multimedia 96
- [2] Chad Carson, Megan Thomas, Serge Belongie, Joseph M. and Hellerstein: Blobworld: A System for Region-Based Image Indexing and Retrieval. Visual Information Systems (1999)
- [3] T.P. Minka and R.W. Picard: Interactive Learning Using a Society of Models. Pattern Recognition' 97
- [4] Q. Zhang, S.A. Goldman, W. Yu, and J.E. Fritts: Content-Based Image Retrieval Using Multiple-Instance Learning. Machine Learning'02
- [5] Jianping Fan, Yuli Gao, Hangzai Luo, and Guangyou Xu: Automatic Image Annotation by Using Concept-Sensitive Salient Objects for Image Content Representation. ACM SIGIR'04
- [6] .B.Li,K.-S.Goh,andE.Y.Chang: Confidence-based Dynamic Ensemble for Image Annotation and Semantics Discovery, ACM MM'03
- [7] C.-F. Tsai, Ken McGarry, and John Tait: Automatic Metadata Annotation of Images via a Two-Level Learning Framework, ACM SIGIR'04
- [8] Xiaolin Wu: Color quantization by Dynamic Programming and Principal Analysis,ACMTOG' 92
- [9] WordNet 2.1 <http://wordnet.princeton.edu/>
- [10] Minsu Jang and Joo-Chan Sohn: Bossam: An Extended Rule Engine for OWL Inferencing