

개념 기반 이미지 검색 시스템을 위한 WordNet 적용 방안

조미영⁰, 최준호, 김판구
조선대학교 전자계산학과
(irune⁰, spica, pkkim)⁰@kiss.or.kr

Applying Method WordNet for Concept based Image Retrieval system

Mi-Young Cho⁰, Jun-Ho Choi, Pan-Koo Kim
Dept. of Computer Science, Chosun University

요 약

기존의 키워드 기반 이미지 검색에서는 의미적 내용 인식을 위해 일반적으로 어휘적 정보나 텍스트 정보를 인간이 주석 형태로 달아주었다. 그러나 이런 텍스트 정보 기반 이미지 검색은 개념적 매칭이 아닌 스트링 매칭이므로 주석을 달아놓은 단어와 정확한 매칭이 없다면 찾을 수가 없다. 이러한 문제를 해결하기 위해 본 논문에서는 개념 기반 이미지 검색 시스템을 위한 WordNet의 적용 방안에 대해 연구했다.

WordNet은 단어형이 아닌 단어의 의미 즉 synset이 구성 요소라는 특징을 이용해 각각의 이미지에 텍스트 정보 대신 적합한 개념의 synset번호를 저장한다. 그리고 검색시 개념간의 유사성 측정을 이용해 검색어와 개념적으로 유사한 모든 이미지를 검색하도록 한다.

1. 서 론

이미지의 내용 인식 문제는 시각적으로 직접 보이는 것 외에 그 이미지에 포함된 의미적 내용인식 메커니즘이 필수적이다. 내용인식의 일반적인 방법으로는 주석자가 어휘적 정보나 텍스트 정보를 주석 형태로 달아주어 그 내용 이해를 위해 사용한다. 이 모델은 기존의 텍스트 정보 검색 기술을 그대로 사용할 수 있기 때문에 구현이 쉽고 비교적 높은 검색 성능을 나타낸다는 장점을 가지고 있지만 주석 처리된 단어와 정확한 매칭이 없다면 찾을 수가 없다는 단점을 가지고 있다.

이를 해결하기 위해 전통적인 방법인 키워드 기반 검색 시스템에서와 마찬가지로 Thesaurus와 같은 용어사전을 사용해 용어 재쓰기(rewriting) 기법, 질의어 확장(query expanding), Relevance Feedback과 같은 기법을 써서 문제를 해결해 왔다. 그러나 이런 방법은 시스템의 정확도에 상당한 문제가 있으며, 여전히 개념적 매칭이 아닌 패턴(스트링) 매칭의 문제를 안고 있다고 볼 수 있다.

따라서, 본 논문에서는 이미지 주석의 패턴 매칭이 아닌 개념적 매칭을 위해 WordNet을 적용하는 방안에 대해 연구했다.

본 논문의 구성은 다음과 같다. 2장에서는 기존의 이미지의 개념 정보 생성의 문제점을 지적하고 3장에서는 노드 기반과 에지 기반 유사성 측정법을 비교한 후 새로운 유사성 측정법을 제안한다. 그리고 4장에서는 실질적으로 개념 기반 이미지 검색에 적용하는 방안에 대해 살펴보고 마지막 5장에서 결론 및 향후 연구 방향을 제시한다.

2. 이미지의 개념 정보 생성

기존의 이미지의 개념 정보 생성은 크게 이미지의 저차원 특징들을 자동적으로 추출하여 의미적 데이터를 찾는 방법과 이미지에 의미를 부여하기 위해 Text주석을 다는 방법으로 나눌 수 있다.

첫 번째 접근 방법은 기존의 Color, Texture, Region 정보, Spatial Color Distribution등의 저차원(low-level) 특징 정보를 이용하여 객체를 추출하고 의미적 특징의 시간적, 공간적, 객체 정보 등과 같은 내용을 표현하는 방법이다. 하지만 이미지의 의미적 요소는 단순한 객체 인식 그 이상이므로 이 방법은 이미지 의미 파악에 한계가 있다고 볼 수 있다.

두 번째 접근 방법은 시각적으로 직접 보이는 것 외에 보이지 않는 이미지의 의미적 내용을 이해하도록 주석자가 텍스트 주석 처리하는 방법이다. 이 방법에서는 부여된 키워드를 가지고 Thesaurus와 같은 어휘 사전을 이용하여 개념 기반 정보 검색을 수행하고 있지만, 이런 경우 기존의 키워드 기반 텍스트 정보검색의 한계를 벗어나지 못하는 문제가 있다.

Ontology의 일종으로 간주되기도 하는 WordNet은 인간의 어휘지식에 대한 심리언어학 연구의 성과를 토대로 1985년도 중반부터 프린스턴대학 인지 과학연구실이 구축해온 영어어휘 데이터베이스다.[6] WordNet은 인간의 어휘지식을 모방한 만큼 다의성과 동의의 관계를 이용하여 의미를 최대한 정확히 표현하고 있고 개념간의 관계 표현 등을 통해 개념을 계층적으로 표현하고 있다.[1] 이러한 특징을 이용해 본 논문에서는 기존의 텍스트 주석에서는 표현할 수 없었던 정확한 의미 표현과 단어의 개

념적 접근을 시도하고자 한다. 먼저 개념적 접근을 위한 개념간의 유사성 측정 방법을 제안한 후 개념 기반 이미지 검색시 이 측정 방법의 적용 방안에 대해 논하겠다.

3. 개념간의 유사성 측정

이 장에서는 WordNet의 계층적 구조를 이용한 단어의 개념적 접근을 위한 개념간 유사성을 측정 방법에 대해 논하겠다.

3.1 노드(node) 기반 유사성 측정

노드 기반 유사성 측정은 노드의 확률에 기반한 정보량(엔트로피) 측정에 의한 접근방식이다.[2] 먼저 정보량의 측정을 위한 노드의 확률을 구하면 다음과 같다.

$$P(c) = \frac{freq(c)}{N} \quad (1)$$

수식 1에서 N 은 개념의 총 수를 의미하고 $P(c)$ 는 개념 c 와 마주칠 확률로 계층적 구조의 경우 $freq(c)$ 는 개념 c 에 포함된 모든 하위 개념들의 합을 의미한다. 만약 c_1 IS-A c_2 라면 $P(c_1) \leq P(c_2)$ 가 된다. 그러므로 WordNet에서 유일한 top node ϕ 의 확률은 1이 된다.

정보 이론에 따르면 개념 c 의 정보량은 $-\log P(c)$ 로 확률이 증가하면 정보량은 감소하므로 더 추상적인 상위 개념은 낮은 정보량을 가진다. 즉, ϕ 의 정보량은 0이 된다.

$$H(c) = -\log P(c) \quad (2)$$

측정하고자 하는 두 개념이 공유한 정보는 WordNet에서 두 개념들을 포함하고 있는 개념의 정보량으로 표현할 수 있다. 따라서 두 개념간의 노드 기반 유사성 측정은 다음과 같이 나타낼 수 있다.

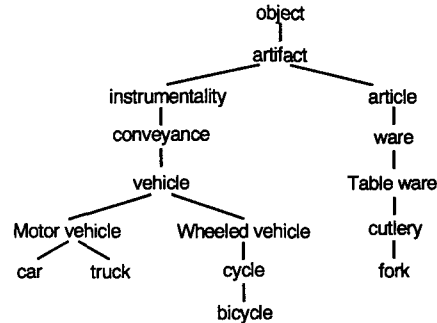
$$\text{sim}(c_1, c_2) = \max_{c \in S(c_1, c_2)} [-\log P(c)] \quad (3)$$

수식 3에서 $S(c_1, c_2)$ 는 c_1 과 c_2 두 개념을 포함하고 있는 상위 개념 c 를 의미한다.

두 개념이 공유하고 있는 상위 개념이 많을수록 두 개념은 더 유사하다고 할 수 있다. 따라서 개념간 유사성은 c_1 과 c_2 를 포함하는 상위 개념 중 가장 큰 H 를 가지는 개념의 H 이다. 예를 들어 [그림 1]에서 car와 truck 그리고 car와 bicycle간의 유사성을 측정한다고 하면 car와 truck이 공유한 개념은 car와 bicycle이 공유한 개념 보다 더 많으므로 더 유사하다고 할 수 있다. 그러므로 $\text{sim}(car, truck)$ 은 $H(\text{motor vehicle})$ 이고 $\text{sim}(car, bicycle)$ 은 $H(\text{vehicle})$ 이다.

2.2 에지(edge)기반 유사성 측정

에지 기반 측정 방식은 개념간의 거리를 측정하므로 의미적 유사성을 평가하는데 더 직접적인 방식이라고 할 수 있다. 이 측정 방식에서는 밀도(부모 노드에서 자식노드로 미치는 링크의 총수), 계층구조에서 노드의 깊이, 링크 타입 그리고 마지막으로 가장 중요한 에지 링크의



[그림 1] car와 truck, bicycle간의 관련트리

길이 등의 특성이 고려된다.

Sussna는 두 개념간의 패스(pass) 길이의 weight로 개념간 거리를 정의했다.[3] 개념 c_i 와 c_j 사이의 weight는 다음과 같다.

$$w(c_i, c_j) = \frac{w(c_i \rightarrow_r c_j) + w(c_j \rightarrow_r c_i)}{2d} \quad (4)$$

$$w(c_x \rightarrow_r c_y) = \max_r \frac{\max_r - \min_r}{n_r(x)} \quad (5)$$

\rightarrow_r 는 개념 c_i 와 c_j 의 관계(relationship) 타입 r 을 의미하고, \rightarrow_r 는 역관계 타입이다. d 는 두 개념중 더 깊은 개념의 깊이를 의미하고, \max 와 \min 은 각각 관계 타입 r 에 대해 최대, 최소 weight들이다. 마지막으로 $n_r(x)$ 는 개념 x 에서 하위 개념들로 미치는 관계의 총 수를 의미한다. 두 개념간의 거리는 최소 패스 길이로 값이 클수록 두 개념은 더 유사하다.

2.3 제안한 개념간 유사성 측정 방법

앞서 살펴본 두 가지 방법은 다른 각도에서 의미적 유사성에 접근한 만큼 각각 다른 특성을 가진다. 에지 기반 측정은 더 직관적인 방법으로 미리 정의된 계층적 단어 네트워크의 구조에 많이 의존한다. 반면 노드 기반 측정은 단어 네트워크의 자세한 구조는 필요하지 않지만 다양한 링크 타입의 문제는 간과되기 쉽다. 그리고 단어의 경우 의미가 아닌 오직 단어 데이터의 빈도만을 고려하므로 실제 정보량보다 더 큰 값으로 계산 될 것이다.

본 논문에서는 각각의 측정법의 단점을 보완하기 위해 노드 기반과 에지 기반 측정법을 결합한 새로운 방법을 제안한다. 먼저 인접한 두 노드(두 개념간 최단 경로가 하나의 에지인 경우)에 유사성 측정은 다음과 같다.

$$S_{Adj}(c_i, c_j^{-1}) = d(L_{i \rightarrow j}) \cdot [H(c_i) - H(c_j^{-1})] \cdot f(d) \quad (6)$$

수식 6에서 $d(L_{i \rightarrow j})$ 는 i 와 j 간 링크 타입에 기반한 밀집도 함수를 나타내고 $H(c_i)$ 는 c_i 의 정보량을 나타낸다. 그리고 마지막으로 $f(d)$ 는 깊이 요소를 반환한다.

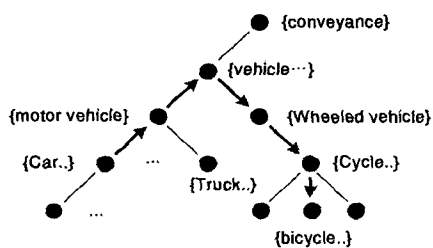
수식 6이 두 개념 사이의 최단 경로가 하나 이상의 에지인 경우로 확장될 경우 개념간 최단 경로의 집합을 P

라고 하면 다음과 같이 나타낼 수 있다.

$P = \{(t_0, c_0, c_1), (t_1, c_1, c_2) \dots (t_{n-1}, c_{n-1}, c_n)\}$ (7)
 따라서 두 개념간의 거리는 최단 경로의 합이라고 할 수 있으며 c_i 와 c_j 사이의 유사성 거리 측정은 다음과 같은 식으로 표현된다.

$$S_{edge}(c_i, c_j) = \sum_{k=0}^n w(t_k) \cdot S_{ADJ}(c_k, c_{k+1}) - (8)$$

$w(t)$ 는 step 함수로 IS_A관계이면 1을 반환하고 다른 관계는 1보다 작은 값을 반환하도록 한다. [그림 2]는 car에서 bicycle간의 최단 거리를 위한 경로를 나타낸다.

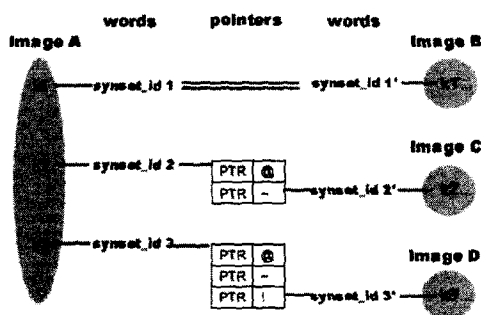


[그림 2] car와 bicycle의 최단 거리

수식 8은 링크 타입과 에지의 수만을 고려하고 있으므로 반의어의 경우 음수값이 리턴될 것이다. 예를 들어 "man"과 "woman"의 경우 WordNet상에서 반의어 관계로 하나의 에지로 연결되어 있다. 그러나 공유하고 있는 즉 "entity", "life form", "person"같은 상위 개념들을 유사성 측정시 고려한다면 두 개념은 유사하게 표현 될 것이다. 즉, 개념간 유사성은 공유한 개념의 정보량에 비례한다. 이 요인을 고려하면 수식 8은 다음과 같이 된다.

$$S(c_i, c_j) = \frac{\sum_{k=0}^n w(t_k) \cdot S_{ADJ}(c_k, c_{k+1})}{\max_{c \in S(c_i, c_j)} [-\log p(c)]} - (9)$$

4. 개념 기반 이미지 검색 시스템 적용



[그림 3] 개념 기반 이미지 검색에 WordNet 적용

[그림 3]은 개념 기반 이미지 검색에 WordNet의 적용을 설명한 그림이다. k1, k2, k3는 이미지 A의 키워드를 의미하고 다른 이미지의 키워드로 저장된 k1', k2', k3'는 각각 이미지 A의 키워드와 유사한 개념의 키워드이거나 동일한 개념의 키워드이다. 주석자는 이미지의 개념 정보를 생성하기 위해 키워드의 가장 적합한 의미를 표현하는 단어의 synset 번호를 이미지와 함께 저장한다. 그리고 검색시 사용자가 적합한 의미의 검색어를 입력하면 개념간의 유사성 측정을 통해 이미지 A의 키워드와 유사한 개념의 k1', k2', k3' 키워드를 포함한 이미지 B, C, D를 검색한다.

5. 결론 및 향후 연구

본 논문에서는 WordNet상에서 개념간의 유사성 측정을 이용한 개념 기반 이미지 검색 모델을 제안했다. 특히, 이 모델은 정확한 개념에 의한 검색을 지원함으로써 단순 패턴(스트링)매칭에 의한 텍스트 기반 검색의 단점을 보완할 수 있을 뿐만 아니라 검색어의 개념적 접근을 통해 개념적으로 가장 유사한 이미지를 검색할 수 있다.

향후 연구 과제로는 내용 기반 이미지 검색 시스템들이 제시하는 저차원의 특징 추출 기술을 적용하여 "빨간 열매가 열린 나무" 같은 자연어 형태의 사용자 질의를 처리할 수 있도록 한다.

참고 문헌

[1] George A. Miller "Introduction to WordNet: An On-line Lexical Database" 1993
 [2] Philip Resnik "Using Information Content to Evaluate Semantic Similarity in a Taxonomy" IJCAI-95
 [3] Sussna "WordSense Disambiguation for Free-text Indexing Using a Massive Semantic Network" CIKM'93
 [4] Jay J Jiang, David W. Conrath "Semantic Similarity Based Corpus Statistics and Lexical Taxonomy" 1997
 [5] Y.C. Park, P.K. Kim, F Golshani, S Panchanathan "Concept-based visual information management with large lexical corpus" 2001
 [6] <http://www.cogsci.princeton.edu/~wn/>
 [7] 최재훈, 박성희, 박수준, 강희범 "에지 디스크립션 템플릿을 이용한 개념 기반 이미지 검색" 2002 정보과학회 추계 학술 발표