

개념간 유사성 측정을 이용한 개념 기반 이미지 검색

조미영⁰, 최준호, 신주현, 김판구

조선대학교 전자계산학과

{irune80⁰, spica, jhshin, pkkim}@mina.chosun.ac.kr

Concept based Image Retrieval Using Similarity Measurement Between Concepts

Mi-Young Cho⁰, Jun-Ho Choi, Ju-Hyun Shin, Pan-Koo Kim

Dept. of Computer Science, Chosun University

요 약

기존의 개념 기반 이미지 검색에서는 이미지의 의미적 내용 인식을 위해 일반적으로 어휘적 정보나 텍스트 정보를 이용했다. 이러한 텍스트 정보 기반 이미지 검색은 전통적인 검색 방법인 키워드 검색 기술을 그대로 사용하여 쉽게 구현할 수 있으나 텍스트의 개념적 매칭이 아닌 스트링 매칭이므로 주석처리된 단어와 정확한 매칭이 없다면 찾을 수가 없었다. 이에 본 논문에서는 ontology의 일종인 WordNet을 이용하여 깊이, 정보량, 링크 타입, 밀도 등을 고려한 개념간 유사성 측정으로 패턴 매칭의 문제를 해결하고자 했다. 또한 키워드로 주석처리 되어 있는 Microsofts Design Gallery Live의 이미지를 이용하여 개념간 유사성 측정법을 실질적으로 개념 기반 이미지 검색에 적용해 보았다.

1. 서 론

멀티미디어 정보 시스템들의 발전으로 대량의 이미지가 온라인상에 존재하게 되었으며, 이를 저장, 처리 및 검색하고자 하는 요구가 증가하고 있다. 이런 요구를 만족시키기 위한 노력으로 시각 정보 시스템에 있어서 많은 진전이 있었다. 지난 10여년동안 내용기반 시각 정보 검색 분야는 아주 유망한 연구분야로 대두되었으며, 여전히 기술개발에 박차를 가하고 있다. 하지만 이런 노력에도 불구하고, 여전히 내용기반 이미지 검색에 있어서 완전한 내용인식에는 그 범위가 미치지 못하고 있는 실정이다. 현재까지는 주로 컬러, 크기, 질감과 윤곽선 등에 기초를 두고 있다. 이는 시각적 특징을 벗어나 의미적 (semantic) 내용 인식까지는 아직 상당히 차이가 있다고 볼 수 있다.

이미지의 내용 인식 문제는 시각적으로 직접 보이는 것 외에 그 이미지에 포함된 의미적 내용인식 메커니즘이 필수적이다. 내용인식의 일반적인 방법은 주석자가 어휘적 정보나 텍스트 정보를 주석 형태로 달아주어 그 내용을 이해하는 것이다. 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.

본 논문에서는 주석 처리된 이미지 정보의 검색을 위한 정성적 측정에 대한 해결 방안을 찾고자 했다. Ontology의 일종으로 일반적으로 사용되는 WordNet을 이용하여 개념 기반 유사도 탐색을 위한 개념간 유사성 계산 수식을 유도했다. 또한 이를 이용해 개념 기반 이미지 정보 색인 및 검색이 되도록 했다.

본 논문의 구성은 다음과 같다. 2장에서는 텍스트 정보를 이용한 기존의 이미지 의미 인식 방안에 대해 살펴보고, 3장에서는 개념간 유사성 측정법에 대해 설명한다. 그리고 4장에서는 실질적으로 개념 기반 이미지 검색에 적용해보고 마지막 5장에서 결론 및 향후 연구 방향을 제시한다.

2. 관련 연구

Singapore 대학교의 Kian-Lee Tan의 연구[6]에서는 웹에 존재하는 이미지에 대해 의미를 부여하는 방법을 제시하고 있다. 주로 이미지와 함께 존재하는 주변의 텍스트를 보고 이미지에 의미를 부여하는 방법이다. 여기서는 Weight Chain Net을 제안하고 있는데, 대표적인 어휘들의 연결된 형태를 취하고 있다. 검색시 질의어와 관련된 어휘들의 Chain들과의 의미적 유사도 측정을 위한 척도를 제시하고 있다. 이 연구 또한 이미지 시각 데이터에 의미를 부여하여 의미기반 이미지 정보검색이 이루어지도록 하는 방법을 제안하고 있다고 볼 수 있다.

Naval Postgraduate School의 N. Rowe와 B.Frew의 연구[5]에서는 분류된 이미지에 주석을 달아서 얻는 장점을 제시하고 있다. 먼저 이미지 분류를 위해 자연 사진을 담은 데이터베이스에서 신경망과 Case기반 기법을 적용했으며, Shape template와 같은 이미지 인식 기법의 잘못을 지적하고 있다. 이 연구는 이미지 데이터에 대한 시각적 특성(Size, Center of mass, Surrounding color discontinuity)을 고려하고 있으며, 25가지의 미리 분류된 클래스로 나누었다. 또한 의미기반 검색을 위하여 WordNet을 사용해 주석 처리된 언어적 특성 또한 이용한다. 이와 같이 Rowe와 Frew도 의미 기반 이미지 정보 검색을 위해 WordNet을 사용하려는 시도가 있었다.

3. 개념간 유사성 측정

3.1. Link strength

단어의 개념적 접근을 위한 개념간 유사성 측정에 대해 논하기 전에 측정시 고려될 수 있는 특징들에 대해 간략히 살펴보겠다. 크게 다섯가지로 분류되는 각각의 특징들의 정의와 유사성 측정시 미치는 영향은 다음과 같다.

(1)깊이(depth)-WordNet상에서 노드의 위치로 깊을수록 더 세부적이고 구체적인 개념이다. 즉, 측정하고자 하는 개념을 포함한 노드의 깊이가 더 깊을수록 유사성 측정시 더 큰 값을 반환한다.

(2)정보량(information content)-노드의 확률에 기반한 값으로 정보 이론에 따르면 개념 c 의 정보량은 $-\log P(c)$ 다. 확률이 증가하면 정보량은 감소하므로 더 추상적인 상위 개념은 상대적으로 낮은 정보량을 가진다.

$$H(c)=-\log P(c) \text{ ----- (1), } P(c)=\frac{freq(c)}{N} \text{ ----- (2)}$$

여기서, 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 수를 의미한다. 만약 c_1 IS_A c_2 라면 $p(c_1) \leq p(c_2)$ 가 된다. 따라서 WordNet에서 유일한 top node \emptyset 의 확률은 1이 되며 정보량은 0이 된다.

(3)밀도(density)-부모 노드의 정보량과 자식 노드의 정보량의 차로 차이가 작을수록 두 노드간에 더 밀접하다.

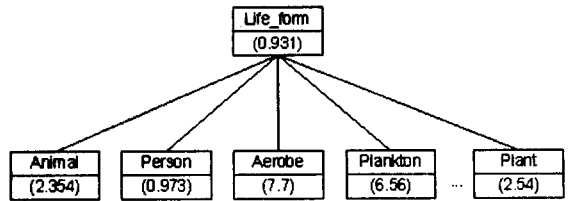
(4)거리(distance)-두 노드간의 거리로 노드간 에지 수를 의미하며 일반적으로 거리가 짧을수록 두 개념은 더 유사하다.

(5)링크 타입(link type)-인접한 두 노드 사이의 관계로 WordNet상에는 상/하의어(IS_A) 관계를 비롯하여 전체/부분, 반의 등과 같은 여러 관계들이 존재한다. 유사성 측정시 이들 각 관계에 대해 다른 가중치 고려하여 계산한다.

위의 5가지 사항들 중 먼저 깊이와 밀도를 고려하여 인접한 두 노드(두 개념간 최단 경로가 하나의 에지인 경우)간의 유사성 즉, link strength는 다음과 같이 정의한다.

$$S_{ADJ}(c_i, c_j^{-1})=d(c_{i-j}) \cdot f(d) \text{ ----- (3)}$$

위 수식에서 $d(c_{i-j})$ 는 노드 i 와 j 간 링크에 기반한 밀도 함수를 나타내고 $f(d)$ 는 깊이 요소를 반환한다. 밀도 함수의 경우 인접한 두 노드간의 정보량의 차로 값이 작을수록 더 유사하다. [그림 1]에서 Animal의 경우 Plankton보다 더 많은 자식 노드를 가지고 있으므로 상대적으로 낮은 정보량을 가진다. 부모 노드인 Life_form과 link strength 계산시 Plankton과 만날 확률보다 Animal이 만날 확률이 더 크기 때문에 더 밀접하게 연관되어 있다. 이를 반영하기 위해 밀도 함수에서 두 노드간의 정보량의 차를 구한 후 역수를 취한다.



[그림 1] Life_form과 그 자식 노드들의 정보량

3.2. 개념간 유사성 측정

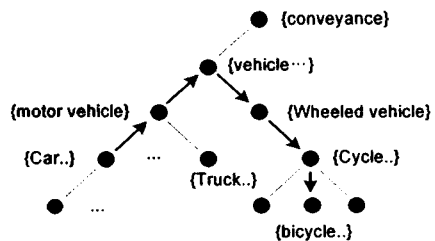
두 개념간 거리가 하나 이상의 에지로 확장될 경우 개념간 최단 경로는 링크 타입을 고려하여 다음과 같이 집합 P 로 나타낼 수 있다.

$$P=((t_0, c_0, c_1), (t_1, c_1, c_2) \dots (t_{n-1}, c_{n-1}, c_n)) \text{ ---- (4)}$$

두 개념간의 최단 거리는 최단 경로의 합 즉, link strength의 합이라고 할 수 있다. c_i 와 c_j 사이의 유사성은 $S_{edge}(c_i, c_j)$ 로 정의하며 다음과 같이 표현된다.

$$D(L_{j-i}) \cdot \sum_{k=0}^n w(t_k) \cdot S_{ADJ}(c_k, c_{k+1}) \text{ ----- (5)}$$

$D(L_{j-i})$ 는 두 노드간의 거리로 에지의 수이다. $w(t)$ 는 링크 타입에 따른 가중치 step 함수로 IS_A관계이면 1을 반환하고 다른 관계는 1보다 작은 값을 반환하도록 한다. [그림 2]는 car에서 bicycle간의 최단 거리경로를 나타낸다.



[그림 2] car와 bicycle간의 최단 경로

수식 (5)는 두 개념간 최단 경로 노드들의 링크 타입과 에지의 수만을 고려했을뿐 두 개념이 공유한 개념이 미치는 영향을 고려하지 않고 있다. 예를 들어, [그림 2]에서 $S(\text{bicycle}, \text{wheeled vehicle})$ 와 $S(\text{cycle}, \text{vehicle})$ 경우 최단 거리 에지의 수와 링크 타입은 IS_A로 같지만 개념간 유사성의 정도도 같다고 볼 수는 없다.

유사성 측정시 두 개념이 공유하고 있는 정보는 WordNet상에서 두 개념을 포함하고 있는 상위 개념의 정보량으로 표현할 수 있다. 즉, 두 개념이 공유하고 있는 상위 개념이 많을수록 두 개념은 더 유사하다고 할 수 있다. 정보량은 확률에 반비례하므로 c_1 과 c_2 를 포함하는 상위 개념 중 가장 큰 정보량을 가지는 최하위 개념의 정보량으로 공유 정보를 표현한다.

$$\text{sim}(c_1, c_2) = \max_{c \in S(c_1, c_2)} [-\log p(c)] \text{ ----- (6)}$$

여기서 $S(c_1, c_2)$ 는 c_1 과 c_2 두 개념을 포함하고 있

는 상위 개념의 집합을 의미한다. 수식 (5)에 상위 개념의 정보량을 고려하면 다음과 같이 나타낼 수 있다.

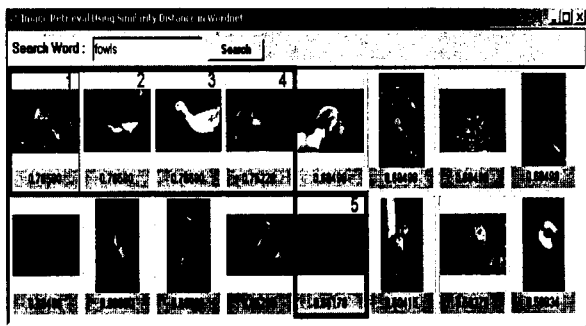
$$S(c_i, c_j) = S_{edge} \cdot \max[H(c)] \text{ ----- (7)}$$

여기서 c 는 c_i 와 c_j 가 공유하고 있는 상위 개념 중 가장 하위에 있는 개념이다.

4. 개념 기반 이미지 검색 적용

실험을 위해 본 논문에서는 약 60,000 개의 명사 synset으로 구성된 WordNet 1.6 버전을 이용하였으며, 분야별로 카테고리화되어 있는 Microsofts Design Gallery Live(<http://dgl.microsoft.com>)의 약 2,300개 가량의 동·식물 이미지를 이용했다. 이미지의 대부분은 일정하지 않은 배경의 컬러 이미지로 각 이미지당 3~5개 정도의 키워드로 주석 처리되어 있다.

[그림 3]은 "fowls" 검색어와 키워드간 개념간 유사성 측정을 통한 이미지 검색 결과이다.



[그림 3] "fowls"의 검색 결과

각 이미지의 하단부의 값은 이미지의 키워드들과 검색어간의 개념간 유사성 측정 후 그 측정치들의 평균치로 이를 기준으로 정렬하였다. 굵은 선으로 표시된 이미지는 "fowls"이라는 키워드가 포함된 이미지로 다른 이미지들 보다 좀더 상위 그룹에 나타났다. 그러나 5번 이미지의 경우 "fowls"이라는 키워드가 존재하나 동떨어진 다른 키워드들에 의해 순위가 낮아졌다. 또한 스트링 매칭이 아닌 개념 매칭이므로 fowl과 유사한 개념의 단어인 goose, hen 등의 키워드를 가진 이미지들도 검색되었다.

5. 결론 및 향후 연구 방향

이미지의 의미적인 인식을 위해 가장 흔히 쓰이는 특징은 키워드로 이 키워드를 이용한 이미지 검색은 기존의 텍스트 정보 검색 기술을 그대로 사용할 수 있기 때문에 구현이 쉽다는 장점을 가지고 있지만 주석 처리된 단어와 정확한 매칭이 없다면 찾을 수가 없다는 단점을 가지고 있다.

본 논문에서는 WordNet상에서 개념간의 유사성 측정을 이용한 개념 기반 이미지 검색을 제안했다. 단순 패턴(스트링)매칭에 의한 텍스트 기반 검색의 단점을 보완

할 수 있었을 뿐만 아니라 검색어의 개념적 접근을 통해 개념적으로 가장 유사한 이미지를 검색할 수 있었다.

향후 연구 과제로는 키워드뿐만 아니라 컬러, 질감 등 다른 저차원 특징과 결합한 이미지의 의미적 내용인식으로 의미 기반 이미지 검색이 되도록 하는 것이다.

참고문헌

- [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] N. Rowe, B. Frew, Automatic Classification of Objects in Captioned Depictive Photographs for Retrieval, Intelligent Multimedia Information Retrieval, The MIT Press, 1997.
- [6] Robert Tansley, et al, Automating the Linking of Content and Concept, Proceedings of the 8th ACM International Conference on Multimedia, Nov. 2000.
- [7] Y.C. Park, P.K. Kim, F Golshani, S Panchanathan "Concept-based visual information management with large lexical corpus" 2001
- [8] <http://www.cogsci.princeton.edu/~wn/>
- [9] 최재훈, 박성희, 박수준, 강희범 "에지 디스크립션 템플릿을 이용한 개념 기반 이미지 검색" 2002 정보과학회 추계 학술 발표