

태그간 의미관계를 이용한 효율적인 태그 기반 이미지 검색 기법

홍현기*, 정진우**, 이동호***

한양대학교 컴퓨터공학과

e-mail:{route*, selphyr**, dhlee72***}@hanyang.ac.kr

An Efficient Technique for Tag-based Image Search using Semantic Relationship between Tags

Hyun-Ki Hong*, Jin-Woo Jeong**, Dong-Ho Lee***

Dept. of Computer Science and Engineering, Hanyang University

요 약

최근, 소셜 미디어 공유 시스템의 사용자-참여형 아키텍처를 구성하는 핵심요소인 폭소노미에 기반하여 이미지를 공유하고 검색하고자 하는 다양한 시도들이 진행되고 있다. 그러나 폭소노미에 기반한 현재의 이미지 공유 시스템에서는 태그의 문법적, 의미적 모호성과 이미지에 대한 태그들의 중요성 또는 상관관계를 고려하지 않아 태그 기반 이미지 검색시 정확성 및 신뢰성을 보장할 수 없다. 이러한 문제를 해결하기 위해, 본 논문에서는 태그간 의미관계를 이용한 이미지 태그 랭킹 기법을 활용하여 태그들을 이미지와의 관련정도에 따라 정렬하여 할당된 후, 이미지의 태그 순위를 고려하여 이미지와 질의어와의 관련성에 따라 효율적으로 이미지를 검색하기 위한 기법을 제안한다. 또한, 제안한 기법이 기존의 이미지 공유 시스템의 검색 결과보다 정확성을 높일 수 있음을 실험 예제를 통하여 확인하였다.

1. 서론

최근 휴대용 디지털 카메라의 보급과 더불어 Flickr[1]와 같은 소셜 미디어 공유 사이트의 등장에 따라, 콘텐츠의 양이 기하급수적으로 증가하고 있다. 이러한 소셜 미디어 공유 사이트를 통해 사용자들은 블로그, 기사, 이미지, 비디오 등과 같은 콘텐츠를 업로드하고 공유한다. 소셜 미디어 공유 시스템의 사용자-참여형 아키텍처를 구성하는 핵심요소로 인식되고 있는 폭소노미(Folksonomy)는 과거 택소노미(Taxonomy)와 같이 전문가에 의하여 구축되는 분류 체계가 아닌 사용자들이 협동적으로 태그(Tag)들을 만들고 관리하는 소셜 태깅(Social Tagging)에 의한 분류 시스템이다. 이러한 폭소노미에 기반한 소셜 미디어 공유 시스템들은 사용자들이 업로드한 콘텐츠에 대하여 사용자들의 적극적인 태깅을 유도하고 있다. 이처럼 다양한 사용자들의 태깅을 통하여 구축되는 폭소노미는 콘텐츠에 대한 풍부한 정보를 제공하여 콘텐츠의 분류와 검색을 용이하게 한다.

폭소노미를 구성하는 핵심적 요소인 태그는 콘텐츠를 기술하기 위하여 사용자에게 의하여 자유롭게 할당되는 키워드이다. 그러나 태그는 특별한 제약 없이 사용자 주관에 의해 만들어지기 때문에 오타, 약어, 은어 등 표준어가 아닌 표기로 인한 문법적 모호성과 동음이의어, 다의어와 같은 의미적 모호성 등을 수반하게 된다.



(그림 1) Flickr의 'banana' 질의어에 대한 검색 결과

현재, Flickr와 같은 이미지 공유 시스템에서는 태그에 대한 단순 키워드 매칭방법을 통하여 태그 기반 이미지 검색을 수행하고 있으며, 이에 따른 검색 결과에 질의어와는 관련이 적은 이미지들이 다수 포함되어 있어 사용자에게 원하는 정보를 정확하게 제공하지 못하는 한계점을 보이고 있다. 그림 1은 Flickr에서 질의어 'banana'로 이미지 검색했을 때의 결과로 사용자의 요구수준과는 큰 차이를 보이고 있다. 이는 태그의 문법적, 의미적 모호성과 이미지에 대한 태그들의 중요성 또는 상관정도를 고려한 태그 기반 이미지 검색 서비스를 제공하지 않고, 단순히 질의어가 각 이미지의 태그 리스트에 포함되어 있는 이미지들을 검색 결과로 보이기 때문이다.

이러한 문제를 해결하기 위해 본 논문에서는 저자들의 선행연구인 태그간 의미관계를 이용한 이미지 태그 랭킹

*** 교신저자

기법[2]을 활용하여 신뢰성 높은 이미지 검색 결과를 위한 효율적인 태그 기반 이미지 검색 기법을 제안한다.

본 논문의 구성은 다음과 같다. 먼저 2장에서 기존 연구들을 살펴보고 문제점들을 기술한다. 3장에서는 태그간 의미관계를 고려한 효율적인 태그 기반 이미지 검색 기법에 관하여 자세히 기술하고, 4장에서는 실험 예제를 보인 후에 5장에서 결론 및 향후 과제로 논문을 마무리한다.

2. 관련 연구

최근, 효율적이고 신뢰성 있는 태그 기반 이미지 검색을 위하여 태그 전달(태그 자동 부여 및 확장)과 태그 랭킹 기법에 관한 다양한 연구들이 수행되고 있다.

[3]에서는 시각적 콘텐츠와 폭소노미를 사용한 태그 전달 기법 기반의 자동 이미지 어노테이션 기법을 제안하였다. 이를 위해 이미지로부터 MPEG-7의 ColorLayout descriptor와 Orientation Histogram의 특징 정보를 추출한 후, 비선형 거리(non-linear distance)와 유클리드 거리(Euclidean distance)기법을 활용하여 k 개의 최근접 이미지를 획득한다. 그 후, 높은 출현빈도를 갖는 태그들 또는 상·하위관계에 있는 태그들을 선정하여 태깅되지 않는 이미지에 할당한다. 하지만 실험결과를 살펴보면, 최고 30%, 평균 19%의 정확성을 보여 전달된 태그들과 이미지와의 관련정도가 매우 적은 것을 알 수 있다. 그리고 전달된 태그들의 랭킹정보가 없어 여전히 태그의 중요성 및 이미지와의 관련성을 제공할 수 없다. 또한 이미지와 관련이 없는 스팸 태그(Spam Tag)나 부적절한 태그가 유사한 이미지들의 태그들에 많이 포함되어 있는 경우 전달되는 태그들의 정확성이 심각하게 낮아지는 단점이 존재한다.

[4]에서는 특정 이미지의 태그들을 이미지 콘텐츠에 관련된 정도에 따라 자동적으로 랭킹하는 시스템을 제안하였다. 이를 위해 특정 이미지에 태깅된 초기 태그들을 대상으로 비모수 확률 밀도 추정 기법 중 하나인 KDE(Kernel Density Estimation) 기법과 Random Walk-based Refinement 방법을 통해 이미지와의 관련정도를 계산하고, 그 결과에 따라 랭킹을 수행한다. 이 방법은 비교적 우수한 태그 랭킹 결과를 보이지만, 각 태그와 이미지와의 관련정도를 측정하는 과정에 있어 상당한 계산복잡성과 공간복잡성을 요구한다. Random Walk-based Refinement 방법에 사용되는 유사도 행렬(Similarity Matrix)을 획득하는 과정 역시 이미지 데이터베이스 전체를 대상으로 복잡한 계산을 수행해야 한다. 또한 태그와 이미지의 관련정도는 유사도 행렬을 Random Walk-based Refinement 방법으로 값이 수렴할 때까지 반복 적용하여 이에 따른 오버헤드가 발생하게 된다.

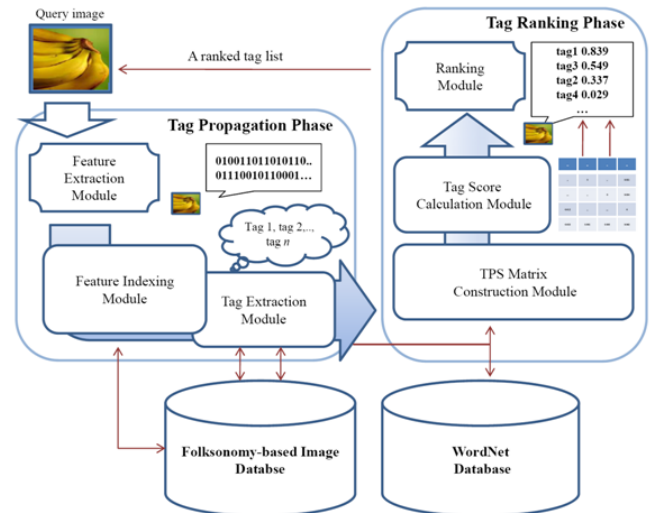
[5]에서는 이미지에 대해 임의의 사용자가 태깅한 특정 태그가 시각적으로 이웃한 이미지에 다른 사용자가 태깅한 태그와 동일한 경우, 특정 태그에 대해 투표한 것으로 간주하는 이웃 투표(Neighbor Voting)기법을 제안하였다. 먼저, kNN(k Nearest Neighbor) 알고리즘을 이용하여 시

각적으로 이웃한 이미지들을 수집한 후, 특정 이미지에 대하여 임의의 사용자가 기술한 모든 태그를 대상으로 이웃 투표 기법을 수행한다. 누적된 각 태그의 이웃 투표 결과를 이용하여 각 태그와 이미지의 관련정도를 측정하여 이에 기반한 태그 랭킹을 수행한다. 하지만, [5]의 기법은 계산복잡성이 낮은 반면 태그 랭킹결과가 kNN 알고리즘을 통한 유사한 이미지 획득의 정확성에 상당히 의존적이라는 단점을 가진다.

3. 태그간 의미관계를 고려한 효율적인 태그 기반 이미지 검색 시스템

본 장에서는 저자들의 선행연구인 태그간 의미관계를 이용한 이미지 태그 랭킹 기법[2]을 통해 이미지와의 관련정도에 따라 태그들을 랭킹하는 기법에 대해 알아보고, 이미지의 태그 순위를 고려하여 이미지와 질의어와의 관련성에 따라 효율적으로 이미지 검색을 수행하는 기법을 자세히 기술한다.

3.1 태그간 의미관계를 이용한 이미지 태그 랭킹 기법



(그림 2) 태그간 의미관계를 이용한 이미지 태그 랭킹 기법의 시스템 구조

이미지의 태그간 의미관계를 이용한 효율적인 이미지 태그 랭킹 기법의 전체적인 시스템 구조는 그림 2와 같다. 이미지에 보다 적합한 태그들을 전달하기 위해 유사한 이미지들의 태그들을 사용하는 1) 태그 전달 단계(Tag Propagation Phase)와 WordNet[6]을 활용하여 전달받은 태그들 간의 의미적 관련정도를 기반으로 랭킹을 수행하는 2) 태그 랭킹 단계(Tag Ranking Phase)로 구성되어 있다. 먼저, 태깅되지 않은 이미지가 주어질 경우, 특정 추출 모듈(Feature Extraction Module)을 이용하여 이미지로부터 색상, 무늬 정보와 같은 시각적 특징 정보들을 추출한 후, 이러한 특징 정보들을 기반으로 인덱싱(Indexing)되어 있는 이미지 데이터베이스로부터 k 개의 최근접 이미지를 획득한다. 그 이후, 유사 이미지 집합으로부터 태그 추출 모듈(Tag Extraction Module)을 이용하여 n 개의

적합한 후보 태그 군을 획득하여 태그 랭킹 단계에 전달한다. 태그 랭킹 단계의 태그 쌍 점수(Tag-Pair Score:TPS) 행렬 계산 모듈(TPS Matrix Construction Module)에서는 WordNet을 활용하여 전달받은 태그들 간의 유사도를 측정하고 다양한 가중치를 적용하여 태그 쌍 점수를 계산한다. 그 후, 태그 쌍 점수를 각 태그 점수로 환산하는 태그 점수 계산 모듈(Tag Score Calculation Module)을 통해 각 태그의 이미지와 관련정도를 측정하고, 랭킹 모듈(Ranking Module)을 통해 각 태그의 출현빈도를 함께 고려한 랭킹점수에 따라 태그 랭킹을 수행한다.

3.1.1 태그 전달 단계

먼저, FCTH(Fuzzy Color and Texture Histogram)[7]를 이용하여 이미지의 색상 및 무늬 정보를 추출한 후, 추출된 정보들을 통해 이미지 데이터베이스로부터 시각적으로 유사한 k 개의 최근접 이미지들을 획득한다. 최근접 이미지들의 태그 리스트에서 각각의 태그의 출현빈도를 계산한 후, 태그들을 출현빈도순으로 정렬하여 최상위 n 개의 태그들과 그 출현빈도를 태그 랭킹 단계에 전달한다.

3.1.2 태그 랭킹 단계

태그간 의미관계에 기반한 효율적인 태그 랭킹을 수행하기 위해, WordNet에서 개념간 유사도 측정방법을 사용하여 태그들 간의 유사도를 태그 쌍(Tag-Pair)별로 계산한 초기 태그 쌍 점수 TPS를 구한다. 그 후, 최소 공통 상위어의 깊이(lcs_depth), WordNet 계층구조에서 같은 가지에 속하는지 여부(Same Branchness:SB), 상대적인 상·하위어(Relative Hypernym/Hyponym:RHH) 관계를 고려한 각각의 가중치를 태그 쌍 점수에 적용하여 최종 태그 쌍 점수를 계산한다. 이 최종 태그 쌍 점수를 각 태그 별로 합산하여 태그 점수로 환산하고 각 태그의 태그 점수와 출현빈도를 함께 고려한 랭킹점수에 따라 태그 랭킹을 수행한다. 보다 상세한 내용은 [2]에 기술되어 있다.

3.2 태그 기반 이미지 검색 시스템

3.1절의 태그간 의미관계를 이용한 이미지 태그 랭킹 기법을 활용하여 이미지 데이터베이스의 저장된 이미지들에 대해서는 이미지의 태그들을 재정렬하고, 태그가 없는 새로운 이미지에 대해서는 자동으로 이미지와 관련성이 높은 순서로 태그들을 할당하여 이미지 데이터베이스에 저장한다. 그 후, 이미지의 태그 순위를 고려하여 이미지와 질의어와의 관련성에 따라 효율적으로 이미지 검색을 수행한다. 이를 위한 태그간 의미관계를 이용한 효율적인 태그 기반 이미지 검색 알고리즘 1과 같다.

알고리즘 1의 1-4번째 줄에서는 사용자로부터 질의어 q 를 입력받아 이미지 데이터베이스의 이미지 i 에 대하여 q 가 i 의 태그 리스트 T_i 에서 일치하는 태그의 순위를 p_i ($p_i = 1, 2, \dots, n$)에 저장한다. 5-6번째 줄에서는 p_i 의 값에 기반하여 q 와 i 의 관련점수 $r(q, i)$ 를 계산한다. 그 후, r 에 따라 이미지들을 정렬하여 반환한다. $r(q, i)$ 를 구하는 식(알고리즘 1의 6번째 줄)은 다음과 같은 특징을 보인다.

Algorithm 1 Image Search

input : query tag q

output : relevant images

t_j : j -th tag of tag list

T_i : ranked tag list of image i

p_i : the position of the tag that is matched to q in the tag list of image i

$sim(q, t_k)$: similarity score between q and tag t_k
($0 \leq sim_{lin}(q, t_k) < 1$)

$r(q, i)$: relevance score between q and image i

```

1 for each image do
2   for tag  $t_j \in T_i$  do
3     if  $q$  is matched to  $t_j$  then
4        $p_i = j$ ;
5       if  $p_i == 1$  then  $k = 2$  else  $k = 1$ ;
6        $r(q, i) = -p_i + sim_{lin}(q, t_k)$ ;
7     end if
8   end if
9 end for
10 end for
11 Sort images by the  $r$  ;
```

- $p_i < p_j$ 인 경우, 항상 $r(q, i) > r(q, j)$ 의 결과를 획득한다. 이는 i, j 의 각 태그 리스트에서 q 와 일치하는 태그의 순위가 더 높은 이미지 i 에 더 많은 관련점수를 주기 때문이다.
- $p_i = p_j$ 인 경우, i, j 의 각 태그 리스트에서 또 다른 태그 t_k ($t_k \neq q$)와 q 의 의미적 유사도 $sim_{lin}(q, t_k)$ 에 따라 결정된다. 이 때, t_k 는 이미지의 태그 리스트에서 q 와 일치하는 태그의 순위에 따라 다르게 계산된다. 먼저, 이미지 i 에 대해 p_i 가 1일 경우, 즉, q 가 i 의 태그 리스트에서 최상위에 나타난 경우, i 의 태그 리스트에서 2번째 태그(t_2)와, 그 외의 경우 최상위 태그(t_1)와 q 의 의미적 유사도 $sim_{lin}(q, t_k)$ 를 측정하여, 태그 순위와 함께 질의어와 이미지와의 관련점수를 계산한다. 이는 태그 리스트에서 q 와 일치하는 태그의 순위가 같은 이미지들에 대해 질의어와의 관련성 순위를 판별하기 위한 방법으로, t_k 와 q 의 의미적 유사도 $sim_{lin}(q, t_k)$ 가 큰 이미지일수록 q 와 더 높은 관련성을 갖는다.

태그간 의미적 유사도 $sim_{lin}(q, t_k)$ 는 WordNet에서 개념간 유사도 측정방법을 사용하여 획득한다. 이는 소셜 태그된 대부분의 태그는 명사형 키워드로 기술되며, 각각의 명사형 키워드는 어휘 데이터베이스인 WordNet에서 정의된 명사의 개념(Concept)에 해당되어 사용할 수 있기 때문이다. 본 논문에서는 WordNet에서 개념간 유사도 측정방법 중 lin [8]을 사용하며, lin 방법은 두 개념과 그들의 최소공통상위어(Lowest Common Superconcept:lcs)



(그림 3) 'apple'과 'tomato'의 질의어에 대한 이미지 검색 결과

의 정보량(Information Content: IC)[8]을 고려하여 두 개념의 유사도를 측정한다.

$$sim_{lin}(t_1, t_2) = \frac{2 \times IC(lcs(t_1, t_2))}{IC(t_1) + IC(t_2)} \quad (1)$$

식 1은 lin 방법을 이용한 두 태그 t_1, t_2 의 태그간 의미적 유사도를 나타내며, $lcs(t_1, t_2)$ 는 WordNet 계층구조에서 t_1, t_2 의 최소공통상위어를 표현한다.

4. 실험

본 연구에서는 실험을 위하여 Flickr Fruit&Veg pool로부터 10,000장 이상의 이미지를 수집하여 데이터베이스를 구축하였다. 그 후, WordNet 3.0, JWNL(Java WordNet Library)을 활용하여 태그간 유사도를 측정하였다. 이미지 데이터베이스에 대해 태그 기반 이미지 검색 시 기존의 이미지 공유 시스템의 검색 방법과 제안한 검색 기법을 이용하여 다양한 질의어로 실험을 수행하였다.

그림 3은 태그 기반 이미지 검색 시 기존의 검색 방법과 제안한 검색 기법을 통한 실험 예제들의 Top 10 검색 결과를 보여준다. (a)와 (c)는 현재 Flickr의 태그 기반 이미지 검색과 같이 단순 키워드 매칭기법을 통한 이미지 검색 시스템에서 'apple'과 'tomato'의 질의어에 대하여 검색을 수행한 결과를 보인다. 이에 반해, (b)와 (d)는 각 질의어에 대해 제안한 이미지 검색 기법에 기반한 검색 결과를 보인다. (a)와 (c)의 결과에서는 단순히 질의어가 이미지의 태그 리스트에 포함되어 있는지를 파악하여 해당 이미지들을 사용자에게 결과로 보여주기 때문에 빨간색 박스의 이미지들처럼 질의어와 관련성이 적은 이미지들이 다수 포함되었다. 하지만, (b)와 (d)에서는 이미지의 태그 순위를 고려하여 이미지와 질의어와의 관련성에 따라 효율적으로 이미지를 검색을 수행하여 질의어와 관련성이 높은 이미지들이 결과에 포함되었다. 실험 예제들을 통해 많은 사용자가 만족하고 공감할 수 있는 검색 결과를 보여 제안한 기법의 정확성을 확인할 수 있었다.

5. 결론 및 향후 연구

본 논문에서는 이미지의 태그간 의미관계를 이용한 이미지 태그 랭킹 기법을 활용하여 효율적인 태그 기반 이미지 검색 기법을 제안하였다. 실험 예제를 통해 제안한 기법이 정확하고 만족스러운 결과를 보이는 것을 확인할 수 있었다. 향후 본 논문에서 제안한 기법의 객관적인 성능 평가를 위하여 추가적인 실험을 수행할 예정이다. 또한 태그 기반 이미지 검색시 사용자 의도를 반영하기 위해 유의어, 동의어와 같은 모호한 질의어를 처리하기 위한 질의어 확장 기법에 대한 연구를 수행할 예정이다.

참고문헌

- [1] Flickr. <http://www.flickr.com>.
- [2] 홍현기, 허지욱, 정진우, 이동호, "태그간 의미적 관계를 이용한 효율적인 이미지 태그 랭킹 기법", 한국정보과학회 2010 한국컴퓨터종합학술대회, 데이터베이스, 2010.
- [3] Stefanie Lindstaedt et al., "Automatic Image Annotation using Visual Content and Folksonomies", MULTIMED TOOLS APPL, Vol. 42, Issue 1, pp. 97-113, 2009.
- [4] Dong Liu et al., "Tag Ranking", Proc. of the 18th Intl. Conf. on WWW, 2009.
- [5] Xirong Li et al., "Learning Tag Relevance by Neighbor Voting for Social Image Retrieval", Proc. of the 1st ACM Intl. Conf. on MIR, pp. 351-360, 2008.
- [6] Fellbaum, C., "WordNet: An Electronic Lexical Database", Cambridge, MA: MIT Press, 1998.
- [7] Savvs A. Chatzichristofis and Yiannis S. Boutalis, "FCTH: Fuzzy Color and Texture Histogram - A Low Level Feature for Accurate Image Retrieval", WIAMIS 2008, pp. 191-196, 2008.
- [8] Siddharth Patwardhan et al., "Using Measures of Semantic Relatedness for Word Sense Disambiguation", Proc. of the 4th Intl. Conf. on ITPCL, Vol 2588, pp. 241-257, 2003.