

텍스트 정보와 시각 정보를 이용한 결정 규칙 기반의 웹 이미지 자동 주석

박주현 최기석^o 낭중호

서강대학교 컴퓨터공학과

{parkjh, brix, jhnang}@sogang.ac.kr

Decision Rule based Web Image Annotation using Textual and Visual Information

Joohyoun Park Giseok Choe^o Jongho Nang

Dept. of Computer Science and Engineering, Sogang University

1 서론

텍스트 기반 웹 이미지 검색 시스템에 있어서 주석의 정확성은 검색 결과의 질을 좌우하는 매우 중요한 조건이다. 그러나 기존의 웹 이미지 자동 주석 방법들[1, 2, 3]은 모호한 주석 선택 기준과 이미지로부터 추출된 시각 정보를 사용하지 않거나 사용하더라도 개념간의 시각적 특성 차이를 고려치 않아 주석 선택의 정확성이 그리 높지 않다는 문제점을 가지고 있다. 본 논문에서는 이미지를 포함하고 있는 웹 문서의 텍스트 정보와 시각 정보를 함께 이용하여 웹 이미지에 대한 주석 선택의 정확성을 높일 수 있는 방법을 제안한다. 제안한 웹 이미지 자동 주석 방법은 웹 문서 구조 분석에 의한 주석 단계와 시각 정보에 의한 필터링 단계의 두 단계 과정으로 주석을 단다.

2 웹 이미지 자동 주석 규칙 생성과 시각 정보에 의한 필터링

이미지를 포함하고 있는 웹 문서는 이미지와 함께 포함되어 있는 임의의 단어들과 연관되어 있는지의 여부를 알려주는 많은 단서들을 포함하고 있다. 그 단서들은 웹 문서를 분석하여 얻을 수 있으며, 크게 태그로부터 얻을 수 있는 단서, 강조 태그로부터 얻을 수 있는 단서, 이미지와 단어간의 구조적인 정보 분석으로부터 얻는 단서, 그리고 웹 문서의 전체적인 특징으로부터 얻는 단서 등의 네 가지 카테고리로 나누어 생각해볼 수 있다. (표 1)은 이러한 네 가지 카테고리들을 기반으로 정의된 이미지와 단어간의 관계 속성을 보여준다. 제안한 16가

표 1: HTML 문서에 포함되어 있는 임의의 이미지 o 와 임의의 단어 c 간의 관계 속성들

ID	Attribute	Possible Values	비고
A1	$c \in \Psi_{ALT}$	{yes, no}	Ψ_{Tag} : 주어진 tag와 함께 나온 명사들의 집합
A2	$c \in \Psi_{Hx}$	{yes, no}	
A3	$c \in \Psi_{IMG}$	{yes, no}	
A4	$c \in \Psi_{TITLE}$	{yes, no}	
A5	$c \in \Psi_{URL}$	{yes, no}	
A6	$c \in \Psi_B \cup \dots, \Psi_{STRONG}$	{yes, no}	$B(\cdot)$: 주어진 이미지 또는 단어를 포함하고 있는 노드
A7	$B(o) \in P(B(c))$	{yes, no}	$P(\cdot)$: 노드의 Parent 노드
A8	$f_l(B(c), B(o))$	continuous	$f_l(\cdot, \cdot)$: 노드들간의 edge 개수
A9	$f_p(B(c), B(o))$	continuous	$f_p(\cdot, \cdot)$: 노드들간의 브라우저에서 보여지는 거리
A10	$f_s(B(c))$	continuous	$f_s(\cdot)$: 노드의 크기
A11	$f_s(B(o))$	continuous	
A12	Position of o	{center, outer edge}	
A13	File Type of o	{jpeg, other type}	
A14	Aspect Ratio of o	continuous	
A15	Frequency of c	continuous	
A16	Number of Image	continuous	

지 관계 속성에 근거하여 임의의 이미지 o 와 임의의 주석 c 쌍을 선택할 것인지의 여부를 판단하는 규칙(Rule)을 만들어야 한다. 본 논문에서는 이러한 규칙을 효과적으로 생성하기 위해 C4.5[4] 알고리즘을 이용하는 결정트리(Decision Tree)를 사용하였다. 결정 트리는 <A1, A2, A3, ..., A15, A16, SelectImage=yes or no >와 같이 표현되는 다수의 훈련용 샘플에 의해서 구축된다. A1 A15는 (표 1)에서 소개한 16개의 속성에 대한 값들의 나열이고 SelectImage는 대상이 되는 이미지와 주석의 선택 여부를 표현하는 목표 개념(target concept)을 의미한다. 한편,

각 이미지들이 어떤 규칙에 의해 후보 이미지로 선택되었는가는 중요한 의미를 갖는다. 정확성이 높은 규칙에 의해 선택된 이미지와 주석의 연관 정도가 정확성이 낮은 규칙에 의해 선택된 이미지와 주석의 연관 정도보다 더 가까울 것이라 예상할 수 있다. 본 논문에서는 90% 신뢰 구간 안에서 각 규칙의 정확성을 평가하였으며 이를 이미지와 선택된 주석간의 연관 정도 w_{oc} 로 정의하였다.

주석의 정확성을 높이기 위해 후보 이미지 집합은 시각정보에 의한 2차 필터링 과정을 거치게 된다. 본 논문에서는 이미지 클러스터링을 위한 시각정보로서 MPEG-7시각정보 기술자중 Dominant Color(DC), Color Layout(CL), Color Structure(CS), Edge Histogram(EH)을 사용하였다. 하지만 4개의 기술자 모두가 모든 개념에 대한 이미지들을 클러스터링하는데 좋은 성능을 보여주지 않기 때문에 효과적 이미지 클러스터링을 위해서는 각 개념에 따라 좋은 클러스터링 결과를 보여줄 수 있는 개념에 적합한 기술자를 선택해야만 한다. 본 논문에서는 각 개념에 적합한 좋은 기술자를 선택하기 위해서 클러스터링 결과 평가 척도인 CH Index[5]를 사용하였다. 클러스터링 시 이용할 임의의 두 이미지 o_i, o_j 간 거리 $d(o_i, o_j)$ 는 CH Index를 가중치로 한 4가지 기술자의 가중치 합으로 계산한다. 최종적으로 주어진 개념과 연관 정도가 낮은 클러스터들을 제외하여야 한다. 임의의 클러스터 C_i 와 주어진 개념 c 간의 연관 정도 $S_c(C_i)$ 는 선택된 규칙의 정확도와 클러스터에 포함된 이미지의 개수의 곱으로 계산한다.

3 실험 및 결론

제안한 방법의 성능을 평가하기 위해 15,185장의 이미지를 수집하여 실험하였고, 성능 평가 방법으로는 수집된 이미지들을 주요 개념 별로 정답 집합(Ground truth)를 만들어 정확률(Precision)을 측정하였다. (그림 1)에서처럼

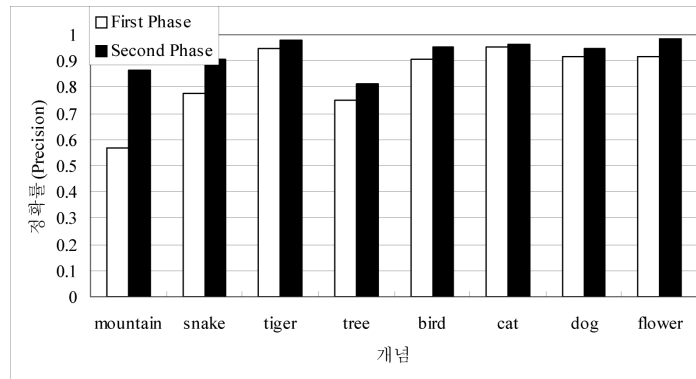


그림 1: 시각정보에 기반한 필터링을 한 경우와 그렇지 않은 경우의 정확률 비교

훈련에 의해 생성된 결정 규칙만을 이용하여 주석을 달 경우 평균 82%의 정확률을 보여주었다. 여기에, 시각 정보에 의한 필터링 과정을 추가할 경우 정확률을 평균 92%로 향상시킬 수 있었다. 제안한 웹 이미지 자동 주석 방법은 본 연구팀에서 개발한 웹 이미지 검색 시스템[6]에 적용하여 그 유용성을 검증하였다.

참고문헌

- [1] D. Cai, X. He, W. Ma, J. Wen, and H. Zhang, "Organizing www images based on the analysis of page layout and web link structure," in *Proceedings of IEEE International Conference on Multimedia Expo*, pp. 113–116, 2004.
- [2] X. Wang, W. Ma, and X. Li, "Data driven approach for bridging the cognitive gap in image retrieval," in *Proceedings of International Conference on Multimedia and Expo 2004*, pp. 2231–2234, 2004.
- [3] K. Yanai, "Generic image classification using visual knowledge on the web," in *Proceedings of ACM Multimedia 03*, pp. 167–176, 2003.
- [4] T. Mitchell, *Machine Learning*. McGRAW-HILL, 1997.
- [5] U. Maulik and S. Bandyopadhyay, "Performance evaluation of some clustering algorithms and validity indices," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 12, pp. 1650–1654, 2002.
- [6] J. Park and J. Nang, "Content based web image retrieval system using both mpeg-7 visual descriptors and textual information," in *Proceedings of ACM International Conference on Multimedia Modeling*, pp. 659–669, 2007.