(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2009-0002133
(43) 공개일자 2009년01월09일

(51) Int. Cl.
G06Q 50/00 (2008.03) *G06F 17/40* (2006.01)
(21) 출원번호 10-2007-0059676
(22) 출원일자 2007년06월19일
심사청구일자 2007년06월19일

(71) 출원인
서강대학교산학협력단
서울 마포구 신수동 1-1 서강대학교
(72) 발명자
낭종호
서울특별시 강남구 논현동 105 동현아파트 1동 206호
박주현
경기 수원시 영통구 영통동 청명마을 325-604
(74) 대리인
이지연, 양부현

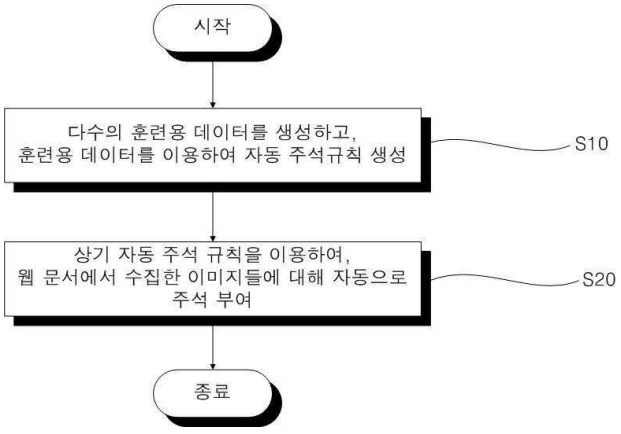
전체 청구항 수 : 총 6 항

(54) 결정 규칙에 기반한 웹 이미지 자동 주석 방법

(57) 요 약

본 발명은 웹 이미지에 대한 자동 주석 추출 방법에 관한 것이다. 상기 자동 주석 추출 방법은, (a) 다수의 훈련용 데이터를 이용해 자동 주석 규칙을 생성하는 단계, (b) 상기 자동 주석 규칙을 이용하여 웹에서 수집된 이미지에 자동으로 주석을 다는 단계를 구비한다. 상기 (a) 단계는 (a1) 다수의 웹 문서에서 임의의 이미지와 임의의 단어 간의 관계 속성을 추출하여 훈련용 데이터들을 준비하는 단계, 및 (a2) 상기 훈련용 데이터들을 이용하여 결정 트리 기반의 자동 주석 규칙들을 생성하는 단계로 이루어진다. 상기 (b) 단계는, (b1) 수집하고자 하는 이미지가 포함되어 있는 웹 문서에서 복수개의 후보 주석들을 선택하여 수집하고자 하는 이미지와 후보 주석들 간에 관계 속성들을 추출하는 단계, 및 (b2) 상기 자동 주석 규칙을 이용하여 (b1) 단계에서 추출한 이미지와 후보 주석 간의 관계 속성을 평가하여 최종 주석을 결정하는 단계로 이루어진다. 상기 훈련용 데이터는 웹 문서의 <IMG> 태그 및 강조 태그로부터 얻는 정보들, 이미지와 단어간의 구조적인 정보 분석으로부터 얻는 정보, 웹 문서의 전체적인 특징으로부터 얻는 정보 들 중 적어도 하나 이상을 선택적으로 포함한다.

대 표 도 - 도1



특허청구의 범위

청구항 1

텍스트 질의 이미지 검색을 위한 웹 이미지 데이터베이스 구축 시, 웹에서 수집된 이미지에 대하여 자동으로 주석을 부여하는 주석 추출 방법에 있어서,

(a) 다수의 훈련용 데이터를 이용해 자동 주석 규칙을 생성하는 단계;

(b) 상기 (a) 단계에서 생성된 자동 주석 규칙을 이용하여 웹에서 수집된 이미지에 자동으로 주석을 다는 단계;

를 포함하는 것을 특징으로 하는 웹 이미지에 대한 자동 주석 추출 방법.

청구항 2

제1항에 있어서, 상기 (a) 단계는

(a1) 자동 주석 규칙을 생성하기 위해 다수의 웹 문서에서 임의의 이미지와 임의의 단어 간의 관계 속성을 추출하여 훈련용 데이터들을 준비하는 단계, 및

(a2) 상기 훈련용 데이터들을 이용하여 결정 트리 기반의 자동 주석 규칙들을 생성하는 단계로 이루어지는 것을 특징으로 하는 웹 이미지에 대한 자동 주석 추출 방법.

청구항 3

제1항에 있어서, 상기 (b) 단계는,

(b1) 수집하고자 하는 이미지가 포함되어 있는 웹 문서에서 복수개의 후보 주석들을 선택하여 수집하고자 하는 이미지와 후보 주석들 간에 관계 속성들을 추출하는 단계; 및

(b2) 상기 자동 주석 규칙을 이용하여 (b1) 단계에서 추출한 이미지와 후보 주석 간의 관계 속성을 평가하여 최종 주석을 결정하는 단계;

를 포함하는 것을 특징으로 하는 웹 이미지에 대한 자동 주석 추출 방법.

청구항 4

제1항에 있어서, 상기 훈련용 데이터는 웹 문서의 <IMG> 태그 및 강조 태그로부터 얻는 정보들 중 적어도 하나 이상 포함하는 것을 특징으로 하는 웹 이미지에 대한 자동 주석 추출 방법.

청구항 5

제1항에 있어서, 상기 훈련용 데이터는 이미지와 단어간의 구조적인 정보 분석으로부터 얻는 정보를 포함하는 것을 특징으로 하는 웹 이미지에 대한 자동 주석 추출 방법.

청구항 6

제1항에 있어서, 상기 훈련용 데이터는 웹 문서의 전체적인 특징으로부터 얻는 정보를 포함하며, 상기 전체적인 특징은 단어의 출현 빈도수, 아이콘 이미지를 제외한 이미지의 갯수 중 적어도 하나 이상을 포함하는 것을 특징으로 하는 웹 이미지에 대한 자동 주석 추출 방법.

명 세 서

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

<5> 본 발명은 웹 이미지에 대한 자동 주석 추출 방법에 관한 것으로, 더 상세하게는 텍스트 질의 기반의 웹 이미지 검색을 위한 웹 이미지 데이터베이스 구축 시 웹으로부터 수집된 이미지에 대해 텍스트 정보와 시각 정보를 이용한 결정 규칙을 기반으로 하여 자동으로 주석을 추출하는 웹 이미지에 대한 자동 주석 추출 방법에 관한 것이

다.

<6> 웹 기술의 발달과 디지털 카메라와 같은 디지털 이미지 생산 장비의 대중화는 WWW 상에서 이미지 데이터의 양을 폭발적으로 증가시키는 결과를 가져왔다. 이러한 이미지들을 효과적으로 이용하기 위해서는, 웹에서 다양한 종류의 이미지를 자동으로 마이닝하여 데이터베이스화하고, 검색하여 보여줄 수 있는 웹 이미지 검색 시스템을 필요로 하게 된다. 이미 Google이나 Yahoo 와 같은 대형 웹 검색 업체들은 텍스트기반의 웹 이미지 검색 서비스를 제공하고 있다. 이러한 텍스트 기반 웹 이미지 검색은 텍스트로 주어진 질의와 이미지의 주석 정보를 비교하여 검색 결과를 만들기 때문에 주석의 정확성은 검색 결과의 질을 좌우하게 된다.

<7> 웹 이미지는 이미지로부터 추출할 수 있는 저급 수준 정보에 더해 텍스트 정보를 얻을 수 있어 기존의 이미지 자동 주석 방법과는 다른 접근 방법이 필요하다. 많은 웹 이미지 주석 및 검색에 관한 연구들에서는 이미지가 포함되어 있는 웹 문서의 타이틀이나 주변 텍스트들을 분석하여 이미지의 주석을 단다.

<8> 예컨대, ImageRover는 이미지가 포함되어 있는 웹 문서 안에 각 단어들이 위치한 HTML 태그에 따라 서로 다른 가중치를 적용한다. 즉, ALT 텍스트, 페이지 타이틀, 헤딩 및 볼드, 이탤릭 태그 등과 같이 이미지나 웹 문서의 내용을 직접적으로 표현하고 있거나 시각적으로 강조된 텍스트 등에 대해서는 높은 가중치를 부여하며 이미지로부터 떨어진 거리를 고려하여 가까운 곳에 위치한 단어에 좀 더 높은 가중치를 부여한다.

<9> WebSeer는 HTML 태그로부터 키워드를 추출하며 이미지 파일명, ALT 텍스트, 페이지 타이틀, URL 등을 이용한다. 웹 문서를 파싱하여 각 태그로부터 단어를 추출하고, ImageRover에서와 마찬가지로 각 단어들의 추출 태그에 따라 서로 다른 가중치를 부여한다.

<10> WebSeek은 웹 문서에 포함된 URL로부터 키워드를 추출한다. 웹 문서로부터 이미지 파일의 URL과ALT 텍스트, Anchor 태그 등을 추출한 후, Non-Alphabet 문자를 제거하여 얻어진 단어들을 색인어로 사용한다.

<11> 이러한 방식으로 추출된 웹 이미지 주석은 시각 정보만을 이용한 일반적인 이미지 주석보다는 높은 정확율을 보여 준다. 이러한 방법들에서는 이미지와 함께 웹 문서에 포함된 모든 단어는 이미지의 주석으로 사용할 수 있다. 물론 검색시 우선 순위를 결정하기 위하여 함께 쓰인 태그의 중요성이나 이미지와의 거리 등을 기준으로 가중치가 부여된다. 하지만, 가중치를 부여하는 기준이 시스템마다 다르고 주석 선택의 임계값이 제시되지 않아 색인 작업을 어렵게 하고 검색 결과의 질을 저하시키는 원인이 된다.

<12> 한편, 최근의 웹 이미지 검색에 대한 연구들에서는 키워드 추출의 정확성을 높이기 위해 웹 문서에 포함되어 있는 텍스트 정보뿐만 아니라 이미지의 시각 정보를 함께 이용하고 있다. 하지만 이러한 방법들은 모두 웹 문서가 가지고 있는 다양성을 고려하지 않아 이미지 주석이나 분류의 정확성의 향상 폭이 크지 않다는 문제점을 갖고 있다.

<13> 이에 본 출원인은 이미지를 포함하고 있는 웹 문서의 텍스트 정보와 시각 정보를 함께 이용하여 WWW상에서 수집한 이미지에 대한 주석 선택의 정확성을 높일 수 있는 방법을 제안하고자 한다.

발명이 이루고자 하는 기술적 과제

<14> 전술한 문제점을 해결하기 위한 본 발명의 목적은 텍스트 질의 이미지 검색을 위한 웹 이미지 데이터베이스 구축 시, 웹 문서 분석 결과를 이용하여 수집된 이미지에 자동으로 정확한 주석을 다는 방법을 제공하는 것이다.

발명의 구성 및 작용

<15> 전술한 기술적 과제를 달성하기 위한 본 발명의 특징은 텍스트 질의 이미지 검색을 위한 웹 이미지 데이터베이스 구축 시, 웹에서 수집된 이미지에 대하여 자동으로 주석을 부여하는 주석 추출 방법에 관한 것으로서, 상기 자동 주석 추출 방법은,

<16> (a) 다수의 훈련용 데이터를 이용해 자동 주석 규칙을 생성하는 단계,

<17> (b) 상기 (a) 단계에서 생성된 자동 주석 규칙을 이용하여 웹에서 수집된 이미지에 자동으로 주석을 다는 단계를 구비하고,

<18> 상기 (a) 단계는 (a1) 자동 주석 규칙을 생성하기 위해 다수의 웹 문서에서 임의의 이미지와 임의의 단어 간의 관계 속성을 추출하여 훈련용 데이터들을 준비하는 단계, 및 (a2) 상기 훈련용 데이터들을 이용하여 결정 트리 기반의 자동 주석 규칙들을 생성하는 단계로 이루어지며,

<19> 상기 (b) 단계는, (b1) 수집하고자 하는 이미지가 포함되어 있는 웹 문서에서 복수개의 후보 주석들을 선택하여 수집하고자 하는 이미지와 후보 주석들 간에 관계 속성들을 추출하는 단계, 및 (b2) 상기 자동 주석 규칙을 이용하여 (b1) 단계에서 추출한 이미지와 후보 주석 간의 관계 속성을 평가하여 최종 주석을 결정하는 단계로 이루어진다.

<20> 전술한 특징을 갖는 자동 주석 추출 방법에 있어서, 상기 훈련용 데이터는 웹 문서의 <IMG> 태그 및 강조 태그로부터 얻는 정보들, 이미지와 단어간의 구조적인 정보 분석으로부터 얻는 정보, 웹 문서의 전체적인 특징으로부터 얻는 정보 들 중 적어도 하나 이상을 선택적으로 포함하는 것이 바람직하다.

<21> 이하, 첨부된 도면을 참조하여 본 발명의 바람직한 실시예에 따른 결정 규칙 기반의 웹 이미지 자동 주석 추출 방법을 구체적으로 설명한다. 본 발명의 바람직한 실시예에 따른 결정 규칙 기반의 웹 이미지 자동 주석 추출 방법은 주석의 정확성을 높이기 위하여 웹 문서 구조 분석에 의한 주석 단계와 시각 정보에 의한 필터링 단계의 두 단계 과정으로 주석을 단다. 웹 문서 구조 분석에 의한 주석 단계에서는 웹 문서에 포함되어 있는 이미지와 텍스트 간의 관계를 표현할 수 있는 속성을 정의하고, 이를 바탕으로 주석 선택을 위한 결정 트리(Decision Tree) 기반의 규칙을 생성한다. 생성된 규칙에 의해 선택된 주석의 정확율은 평균 82% 이상의 높은 수치를 보여준다. 정확율을 더욱 높이기 위해 수행되는 필터링 단계는 개념별로 이미지들을 모아 복수 개의 시각 정보를 기반으로 클러스터링하고 개념과 연관 정도가 낮은 클러스터들을 필터링시킨다. 이러한 필터링 과정을 거친 후, 주석의 정확율은 평균 92%가 된다.

<22> 도 1은 본 발명의 바람직한 실시예에 따른 결정 규칙 기반의 웹 이미지 자동 주석 추출 방법을 순차적으로 나타내는 흐름도이다. 이하, 도 1을 참조하여 본 발명의 바람직한 실시예에 따른 웹 이미지 자동 주석 추출 방법을 구체적으로 설명한다. 도 1을 참조하면, 본 발명의 실시예에 따른 결정 규칙에 기반한 웹 이미지 자동 주석 방법은 크게 결정 규칙 생성 단계(단계 10) 및 이미지에 대한 주석 결정 단계(단계 20)로 이루어진다.

<23> 상기 결정 규칙 생성 단계(단계 10)는 웹 문서에 포함되어 있는 임의의 단어가 임의의 이미지의 주석이 될 수 있는지의 여부를 판단하기 위한 규칙이 존재하지 않으므로 다수의 훈련용 데이터를 이용해 자동 주석 규칙을 생성하여 데이터베이스화하는 단계이다.

<24> 상기 결정 규칙 생성 단계를 설명하기 위해서, 웹 문서에 포함되어 있는 임의의 이미지 o 와 임의의 단어 c 를 가정한다. 우선, 단어 c 가 이미지 o 의 주석이 될 수 있는지의 여부를 판단하기 위해 이미지와 단어 간의 관계 속성을 정의한다. 도 2는 결정 규칙을 생성하기 위한 훈련 과정과 자동 주석을 다는 과정에서 이용되는 같은 웹 문서에 포함되어 있는 임의의 이미지와 임의의 단어 간의 관계 속성들을 정의한 내용의 일예를 나타낸 도표이다. 실제로 이미지를 포함하고 있는 웹문서는 이미지와 함께 포함되어 있는 임의의 단어들과 연관되어 있는지의 여부를 알려줄 수 있는 많은 단서들을 포함하고 있다. 그 단서들은 웹 문서를 분석하여 얻을 수 있으며, 크게 ① <IMG> 태그로부터 얻을 수 있는 단서, ② 강조 태그로부터 얻을 수 있는 단서, ③ 이미지와 단어간의 구조적인 정보 분석으로부터 얻는 단서, 그리고 ④ 웹 문서의 전체적인 특징으로부터 얻을 수 있는 단서 등의 네가지 카테고리로 나누어 생각할 수 있다. 이하, 전술한 네가지의 카테고리에 대하여 설명한다.

<25> 첫번째로 이미지를 직접적으로 표현하는 <IMG> 태그로부터 직접 얻을 수 있는 정보들을 이용하는 것이다. HTML 문서에서 <IMG> 태그의 src, alt 속성들과 함께 나오는 단어들은 매우 중요한 의미를 갖는다. 특별히 alt 속성은 웹 브라우저가 이미지를 표현할 수 없을 경우에 대체 수단으로서 만들어졌기 때문에 alt 속성과 함께 쓰인 단어는 이미지와의 연관 정도가 매우 높다고 할 수 있다. 두번째로 텍스트를 강조하기 위하여 사용되는 태그와 함께 쓰인 단어들을 고려해야 한다. HTML 문서에서의 <Hx>, <B>, <BIG>, <EM>과 같은 태그는 함께 쓰인 텍스트를 크게 하거나 굵은 글씨체로 표현하기 때문에 보통 강조의 의미로 사용되고 중요한 의미를 갖는 것이 일반적이다.

<26> 도 2에 도시된 속성 A1부터 속성 A6까지는 c가 특별한 태그 또는 태그의 속성과 함께 쓰였는지의 여부를 표현한다. c가 이미지 o를 표현하고 있는 <IMG> 태그의 alt 속성과 함께 나왔는지, 문서의 타이틀이나 URL에 포함되어 있는지 등을 표현한다.

<27> 세번째로 HTML 문서안에서 이미지와 단어간의 구조적인 정보를 고려해야 한다. 일반적으로 HTML 코드의 구조적 특징은 HTML이나 XML 문서의 논리적인 구조를 표현하는 DOM 트리에 기반하여 분석한다. DOM 트리는 각 태그들을 트리의 노드로 하여 구성되기 때문에 각 태그들간의 관계를 파악하는데 도움이 된다. 도 2에 도시된 속성 A7부터 속성 A9까지는 HTML 코드의 DOM 분석 결과로부터 얻을 수 있는 o와 c간의 거리 정보를 다양한 형태로 표현한다. 도 2에 도시된 속성 A10은 B(c) 의 크기를 의미하는데, 그 크기가 크면 다른 많은 단어들을 포함하고 있을

확률이 높아져 이미지와의 연관 정도가 작아져야 할 것이다. 또한, 도 2에 도시된 속성 A11부터 속성 A14까지는 *o* 자체에 대한 정보들로서, 사이트 탐색을 위한 화살표, 버튼 등의 아이콘 이미지, 각종 광고 이미지 등 검색 대상이 되기에 부적합한 이미지들을 수집 대상에서 제외하는데 이용할 수 있다.

<28> 마지막으로 웹 문서 전체의 속성을 고려해야 한다. 만일 웹 문서에 단 한개의 이미지만이 포함되어 있다면, 그 문서의 URL이나 타이틀이 이미지를 설명하는 좋은 단서가 될 수 있을 것이다. 웹 문서 전체에 대한 속성으로서 단어의 출현 빈도수(A15)와 아이콘 이미지를 제외한 이미지의 개수(A16)를 고려한다. 단어 빈도수가 높다는 것은 그 단어가 웹 문서에서 중요함을 의미하고 정렬되어 있는 이미지의 개수가 많다는 것은 이미지 수집에 적절한 사이트임을 의미할 수 있다.

<29> 전술한 내용을 토대로 하여, 본 발명의 바람직한 실시예에 따른 자동 주석 부여 방법에 사용되는 훈련용 데이터는 , 임의의 이미지 *o*와 임의의 주석 *c* 쌍인 < *이미지(o), 주석(c)* >에 대하여 전술한 16가지 관계 속성에 근거하여 <A1, A2, A3, …, A15, A16, SelectImage=yes or no >으로 구성된다. 여기서, A1~A16은 도 2의 16개의 속성에 대한 값들의 나열이고, SelectImage는 대상이 되는 이미지와 주석의 선택 여부를 표현하는 목표 개념(Target Concept)을 의미한다. 예를 들면, 임의의 웹 문서의 중앙에 큰 'tiger' 이미지가 있고, <IMG> 태그의 alt 속성에 'tiger'라는 단어가 포함되어 있는 경우의 훈련용 데이터는 <yes,no,no,no,no,no,yes, 1, 203, 120019, 12120, center, jpeg, 1, 11, 51, SelectImage=yes> 로 표현될 수 있다.

<30> 결정 규칙 생성 단계(단계 10)에서 자동 주석 규칙은 위의 과정을 거쳐 <A1, A2, A3, …, A15, A16, SelectImage=yes or no >의 형태로 준비된 다수의 훈련용 데이터에 의해 구축된다. C4.5 알고리즘 기반의 결정트리(Decision Tree) 는 단어와 이미지의 내용이 일치하는 긍정적 샘플들과 포함된 단어와 이미지의 내용이 일치하지 않는 부정적 샘플들을 이용하여 구축된다. 결정트리가 지나치게 훈련용 데이터에 의존적이 되는 것을 막기 위해 가지치기 과정을 거쳐 규칙을 생성한다.

<31> 따라서, 전술한 결정 규칙 생성 단계는, (a-1) 자동 주석 규칙을 생성하기 위해 다수의 웹 문서에서 임의의 이미지와 임의의 단어 간의 관계 속성을 추출하여 훈련용 데이터들을 준비하는 단계, 및 (a-2) 상기 (a-1) 단계에서 준비된 훈련용 데이터들을 이용하여 결정 트리 기반의 자동 주석 규칙들을 생성하는 단계로 요약할 수 있다. 위와 같은 결정 규칙 생성 단계(단계 10)는 규칙을 생성하기 위해 초기 한 번만 수행한다.

<32> 다음, 상기 이미지에 대한 주석 결정 단계(단계 20)는 상기 결정 규칙 생성 단계에서 생성된 자동 주석 규칙을 이용하여 웹에서 수집된 이미지에 자동으로 주석을 다는 단계이다.

<33> 상기 주석 결정 단계(단계 20)를 설명하기 위해, *n*개의 이미지와 *m*개의 단어를 포함하고 있는 웹 문서 *D*를 가정하자. 우선 *D* 에서는 *n* x *m*개의 테스트 해야 할 이미지와 후보 주석 쌍 $<o_i, c_j>$ ($1 \le i \le n$, $1 \le j \le m$)을 만들 수 있다. 여기에 이미지와 가장 가까운 visual keyword의 대표 키워드도 추가되어 총 *n* x (*m*+1)개의 테스트 쌍이 만들어진다. 각 $<o_i, c_j>$ 쌍들에 대해 c_j 가 o_i 의 주석이 될 수 있는 지의 여부를 판단하기 위해 (a-1) 단계에서처럼 관계 속성들을 추출한다. 이미지 o_i의 주석은 (a)단계에서 생성된 규칙에 의해 목표 개념이 "yes"로 평가된 $<o_i, c_j>$의 c_j로 결정된다.

<34> 따라서, 전술한 주석 결정 단계(단계 20)를 요약하면, (b-1) 수집하고자 하는 이미지가 포함되어 있는 웹 문서에서 복수개의 후보 주석들을 선택하여 수집하고자 하는 이미지와 후보 주석들 간에 관계 속성들을 추출하는 단계, 및 (b-2) 결정 규칙 생성 단계에서 생성된 자동 주석 규칙을 이용하여 (b-1) 단계에서 추출한 이미지와 후보 주석 간의 관계 속성을 평가함으로써 최종 주석을 결정하는 단계로 이루어진다.

<35> 상기와 같은 본 발명에 따른 텍스트 질의 이미지 검색을 위한 웹 이미지 데이터베이스 구축 시, 이미지의 시각 정보와 웹 문서 분석 결과를 이용하여 수집된 이미지에 자동으로 정확한 주석을 다는 방법에 따르면 평균 85%의 정확률을 유지하며 자동으로 이미지에 주석을 달 수 있다.

<36> 이상에서 본 발명에 대하여 그 바람직한 실시예를 중심으로 설명하였으나, 이는 단지 예시일 뿐 본 발명을 한정하는 것이 아니며, 본 발명이 속하는 분야의 통상의 지식을 가진 자라면 본 발명의 본질적인 특성을 벗어나지 않는 범위에서 이상에 예시되지 않은 여러 가지의 변형과 응용이 가능함을 알 수 있을 것이다. 그리고, 이러한 변형과 응용에 관계된 차이점들은 첨부된 청구 범위에서 규정하는 본 발명의 범위에 포함되는 것으로 해석되어야 할 것이다.

발명의 효과

<37> 본 발명에 따른 방법의 성능을 평가하기 위하여 표 1의 6개 사이트를 시작으로 BFS(Breadth First Search)를 사용하여 15,185장의 이미지를 수집하여 실험하였고, 성능 평가 방법으로는 정확률(Precision)과 재현률(Recall)을 사용하였다.

표 1

<38>

Seed Sites
http://www.junglewalk.com
http://nationalzoo.si.edu
http://www.freefoto.com
http://www.hickerphoto.com
http://www.amusetoi.com
http://www.indianwildlifeportal.com

<39> seed 사이트 특성에 의해 동물들이나 자연 풍경을 표현하는 이미지들이 주로 수집되었고, 각 개념별 최소 200장 이상 수집된 'tiger', 'bird', 'dog','snake','cat','flower','mountain'의 8가지 개념에 대하여 평가하였다. 성능 평가를 위한 Ground Truth를 만들기 위해 15,835장의 이미지들 중에서 8개의 개념에 해당하는 이미지들을 수동으로 골랐으며 성능 평가를 위해 트레이닌 및 테스팅에 사용한 <이미지,주석> 샘플 갯수는 4,641개이다. 결정 트리를 만들기 위해 사용한 훈련용 이미지 집합은 8개의 개념에 해당하는 <이미지,주석>쌍 중 무작위로 1,000개를 선택하여 사용하였다. 마지막으로 프로그램 제작에 사용한 라이브러리는 DOM 트리 분석을 위하여 MSHTML을 사용하였고 저급 수준 정보 추출을 위하여 MPEG-7 XM 코드를 사용하였다.

<40> 본 실험에서는 훈련을 통해 만들어진 규칙 중 가지 치기후 남은 27개의 규칙이 테스팅하는데 사용되었다. 도 3은 임의의 선택된 1,000개의 훈련용 <이미지,주석> 쌍의 관계 속성 샘플들로 훈련하여 만들어진 규칙 중 실제 분류하는데 가장 많이 사용된 규칙의 일부이다. Rule 15는 주어진 단어가 alt 속성과 Title 에 나타나지 않고 웹문서에 포함된 이미지의 개수가 너무 많을 경우, 그 단어와 이미지는 연관되어 있지 않을 가능성이 매우 높음을 의미한다. 반면에, Rule 67에서처럼 주어진 단어가 alt 속성과 이미지 이름과 함께 나타남과 동시에 이미지와의 물리적 거리가 63이하인 곳에서도 나타날 경우, 주어진 단어가 이미지의 주석일 확률이 매우 높음을 알 수 있다.

<41> 도 4는 전술한 8개의 개념에 대하여 본 발명의 자동 주석 부여 방법에 따라 결정 트리에 의해 주석된 결과와 종래의 ImageRover의 결과를 비교한 그래프이다. ImageRover는 널리 알려진 웹 이미지 검색 시스템 중 하나로서 HTML 태그, 이미지와 텍스트간의 거리, 그리고 단어 출현 빈도수들을 고려하여 주석에 가중치를 주어 주석을 선택한다. 이와 같이 몇가지 특징에 대해 임의로 정해진 가중치를 부여하는 방법은 본 발명에 따른 방법과 같이 훈련용 데이터에 의해 결정 트리를 만드는 방법과 좋은 비교 대상이 될 수 있다. 본 발명에 따른 방법과의 비교 실험을 위해, ImageRover에서 본 발명에 따른 방법을 사용하여 이미지와 함께 나온 각 단어에 가중치를 계산하여 [0,1] 사이로 정규화한 후 0.5 이상의 가중치를 갖는 단어들만 이미지에 대한 주석으로 선택하였다. 도 4의 (a)는 정확률을 비교한 그래프로서, 도 4의 (a)를 통해 본 발명에 따라 결정 규칙을 이용한 주석 방법이 종래의 ImageRover에 비해 매우 높은 정확률을 보여줌을 알 수 있다. 그리고, 도 4의 (b)는 재현률을 비교한 그래프로서, 도 4의 (b)를 통해 본 발명에 따라 결정 규칙을 이용한 주석 방법이 종래의 ImageRover보다 높고 안정적인 재현률을 보여줌을 알 수 있다.

<42> 전술한 바와 같이, 본 발명에 의하여, 텍스트 질의 이미지 검색을 위한 웹 이미지 데이터베이스 구축 시, 이미지의 시각 정보와 웹 문서 분석 결과를 이용하여 만들어진 관계 속성을 기반으로 생성된 규칙을 이용하여 기존의 방법에 비하여 향상된 정확성을 가진 자동 주석이 가능하다.

도면의 간단한 설명

<1> 도 1은 본 발명의 바람직한 실시예에 따른 결정 규칙 기반의 웹 이미지 자동 주석 추출 방법을 순차적으로 나타내는 흐름도이다.

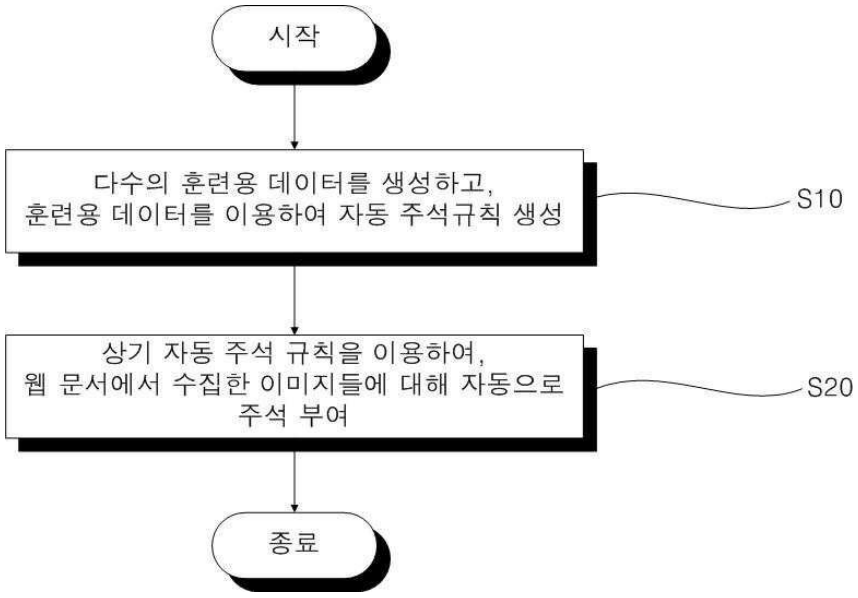
<2> 도 2는 본 발명의 바람직한 실시예에 따른 결정 규칙 기반의 웹 이미지 자동 주석 추출 방법에 있어서, 결정 규칙을 생성하기 위한 훈련 과정과 자동 주석을 다는 과정에서 이용되는 임의의 이미지와 임의의 단어 간의 관계 속성들을 정의한 내용을 예시적으로 도시한 도표이다.

<3> 도 3은 본 발명의 바람직한 실시예에 따른 결정 규칙 기반의 웹 이미지 자동 주석 추출 방법에 있어서, 임의의 선택된 1,000개의 훈련용 <이미지,주석> 쌍의 관계 속성 샘플들로 훈련하여 만들어진 규칙 중 실제 분류하는데 가장 많이 사용된 규칙의 일부를 예시적으로 도시한 것이다.

<4> 도 4는 본 발명의 바람직한 실시예에 따른 결정 규칙 기반의 웹 이미지 자동 주석 추출 방법의 성능을 입증하기 위하여, 8개의 개념에 대하여 본 발명의 자동 주석 추출 방법에 따른 결과와 종래의 ImageRover의 결과를 비교한 그래프이다.

도면

도면1



도면2

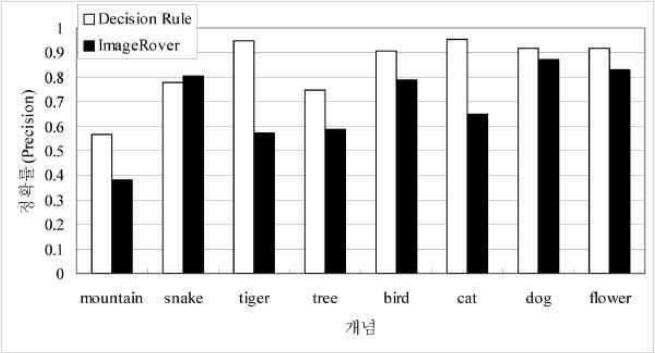
ID	Attribute	Possible Values
A1	$c \in \Psi_{ALT}$	{yes, no}
A2	$c \in \Psi_{Hx}$	{yes, no}
A3	$c \in \Psi_{IMG}$	{yes, no}
A4	$c \in \Psi_{TITLE}$	{yes, no}
A5	$c \in \Psi_{URL}$	{yes, no}
A6	$c \in \Psi_B \bigcup ..., \Psi_{STRONG}$	{yes, no}
A7	$B(o) \in P(B(c))$	{yes, no}
A8	$f_l(B(c), B(o))$	continuous
A9	$f_p(B(c), B(o))$	continuous
A10	$f_s(B(c))$	continuous
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

Notations
o : HTML 문서에 나온 임의의 이미지
c : o 와 함께 나온 임의의 단어 (명사)
Ψ_{Tag} : 주어진 tag와 함께 나온 명사들의 집합
$B(\cdot)$: 주어진 이미지 또는 단어를 포함하고 있는 DOM 노드(블록)
$P(\cdot)$: 주어진 DOM 노드의 Parent 노드
$C(\cdot)$: 주어진 DOM 노드의 Child 노드
$f_l(\cdot,\cdot)$: 주어진 두 DOM 노드들간의 edge 개수 (논리적인 거리)
$f_p(\cdot,\cdot)$: 주어진 두 DOM 노드들간의 브라우저에서 보여지는 거리 (물리적 거리)
$f_s(\cdot)$: 주어진 DOM 노드의 크기

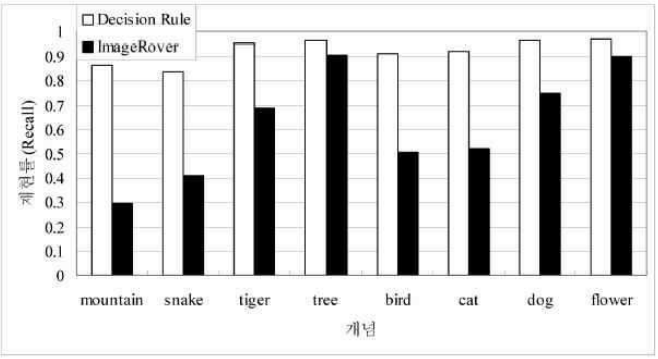
도면3

Rule 15
if A1=no & A4=no & 72<A16 **then**
  SelectImage = *no*
end if
Rule 67
if A1=y &A3=y & A9<63 **then**
  SelectImage = *yes*
end if

도면4



(a) 정확률(Precision)



(b) 재현률(Recall)