



## 자동 이미지 태깅 기술 동향

이시형\* 노용만\*

멀티미디어 기기의 보급과 웹 애플리케이션의 확대에 인해서 온라인을 통한 사용자 생성 콘텐츠(User Created Contents, UCCs) 공유가 활발해지고 있다. 특히, 웹을 통한 이미지 공유는 사용자들이 가장 활발하게 이용하고 있는 분야 중에 하나이다. 효율적인 이미지 검색을 위한 태깅 기술은 이미지의 색상, 질감 그리고 형태 등의 분석을 통한 내용 기반 특징(content-based features)과 이미지 획득 과정에서 저장된 정보(Exchangeable Image File Format: EXIF)를 중심으로 연구되었다. 하지만, 소셜 웹 사이트(social web site)의 등장으로 사용자에게 의한 태그(tag)를 중심으로 대용량 이미지를 위한 자동 이미지 태깅 기술이 연구되고 있다. 향후 멀티미디어 콘텐츠 활용에 핵심이 될 자동 이미지 태깅 기술의 연구개발 동향을 알아보고, 앞으로의 연구 전망을 살펴본다. ☐

### 목 차

- I. 서 론
- II. 연구개발 동향
- III. 자동 이미지 태깅 기술 발전 전망
- IV. 결 론

### I. 서 론

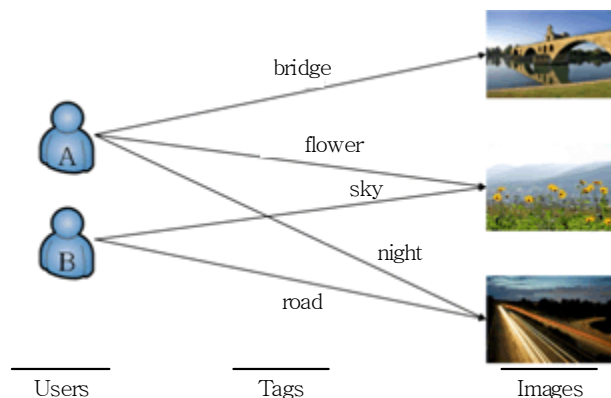
멀티미디어 기기의 확산, 스토리지 가격 하락은 사용자에게 의해서 생성된 이미지의 양을 급격하게 증가시켰다. 이런 이미지는 Flickr 나 Facebook 과 같은 소셜 웹 사이트를 통해서 공유되고 있다. 대표적인 이미지 공유 사이트인 Flickr 에는 2008 년 11 월에 30 억 장의 이미지가 공유되고, 매분 수천 장의 새로운 이미지가 등록되고 있다[1]. 이처럼 웹을 통한 이미지 공유가 활발해지면서, 대용량 이미지를 효율적으로 인덱싱하고 관리하는 방법은 중요한 문제로 인식되고 있다.

이미지 태깅 기술은 조직 체계가 갖추어지지 않은 대용량의 이미지 정보를 처리하여 사용자가 원하는 정보에 접근할 수 있도록 하는 기술을 의미한다 [2]. 키워드에 의한 이미지 태깅 방법은 가장 널리

\* KAIST 전기전자 영상및비디오시스템연구실

이용되는 방식이다. 이 방법은 사용자가 키워드를 쿼리로 입력하면 매칭되는 색인어를 가지고 있는 이미지를 검색 결과로 보여준다. 이 방법은 데이터베이스 안에 있는 이미지들이 적절한 색인어로 기술이 되어 있고 사용자가 색인어에 대한 사전 지식이 있을 때, 효율적인 검색을 보장한다. 사용자가 키워드를 정확하게 알고 있지 못할 경우 검색이 불가능하고, 쿼리 키워드와 정확하게 일치하는 이미지만 검색할 수 있기 때문에 쿼리 키워드와 유사한 내용을 가진 이미지는 검색하기 어렵다. 이미지의 내용 기반 인덱싱(content-based indexing)은 내용 기반 특징을 통해서 이미지 인덱싱을 보장하기 때문에, 키워드 기반 인덱싱 방법의 한계를 해결할 수 있다. 데이터베이스 내의 모든 이미지는 색상, 질감 그리고 형태 등으로 대표되는 내용 기반 특징을 바탕으로 기술된다. 사용자가 검색 엔진으로 쿼리 이미지를 전달하면, 쿼리 이미지와 내용 기반 특징이 유사한 이미지가 검색 결과로 보여진다. 그리고 이미지 획득 당시의 정황 정보(context information)를 기술하고 있는 EXIF 정보 또한 이미지 인덱싱에 활용된다.

최근 웹 2.0의 출현을 바탕으로 협력 태깅(collaborative tagging) 또는 소셜 태깅(social tagging)에 대한 관심이 높아지고 있다. 협력 태깅은 (그림 1)과 같이 온라인을 통해서 공유되고 있는 리소스에 사용자가 태그를 다는 행위를 의미한다. 대부분의 소셜 웹 사이트는 협력 태깅을 통해서 생성된 폭소노미(folksonomy)를 이미지 인덱싱에 활용하고 있다. (그림 1)과 같이 동일한 이미지라도 다양한 사용자에 의해서 많은 태그가 등록될 수 있기 때문에, 기존의 키워드 기반 이미지 인덱싱 방법에 비해서 효율적이다. 하지만 대부분의 소셜 웹 사이트가 사용자에 의한 수동적 태깅만 지원하기 때문에 충분한 양의 태그가 등록되어 있지 않다. 연구 결과에 의하면, Flickr에 등록된 60%의 이미지에는 3개 이하의 태그만 등록되어 있다[3]. 그래서 수동으로 등록된 태그로는 효율적인 이미지 검색을 보장하기 어렵다. 자동 태깅은 온라인을 통해서 공유되



(그림 1) 협력 태깅의 개념도

고 있는 이미지에 적합한 태그를 사용자에게 자동으로 추천하는 기술로써, 대용량 이미지 검색을 위한 효율적인 태깅 기술로 고려되고 있다.

본 고에서는 먼저 키워드 및 내용 기반 특징을 통해서 이루어지고 있는 이미지 태깅 기술과 관련 기술의 동향을 소개하고, 태그 추천 기술을 통한 자동 태깅 기술의 동향을 소개한다.

## II. 연구개발 동향

### 1. 이미지 검색 기술 동향

1990 년대부터 콘텐츠의 검색과 태깅을 위해서 콘텐츠의 내용 특징에 기반을 둔 연구가 활발하게 진행되었다. 하지만 콘텐츠의 내용 특징만으로는 개인 사용자의 검색 의도를 신뢰성 있게 보장할 수 없는 한계(semantic gap)를 가지고 있다[4]. 현재까지 이미지 검색을 위한 프로토타입 시스템으로는 SIMBA, Blobworld, FIDS, KIWI, MetaSEEK, Photobook, QBIC, PicToSeek, VisualSEEK 등이 있다. 미국이나 유럽이 이미지 검색 시장을 주도하고 있으며 상용 시스템뿐만

<표 1> 이미지 인덱싱 연구 및 기술 개발 내용

연구 수행 기관	연구 개발 내용 및 결과	국적
IBM	<ul style="list-style-type: none"> <li>- 텍스트 정보와 영상의 특징 정보를 연동하는 멀티미디어 검색 시스템 개발(IBM Multimedia Search and Retrieval System)</li> <li>- 영상 내에서 특정 화면을 찾는 프로토타입의 검색엔진 개발(Marvel)</li> <li>- 자동적으로 영상의 구조 정보를 요약하는 시스템 개발(CueVideo)</li> <li>- 영상의 내부정보를 바탕으로 시맨틱 컨셉을 모델링하는 기술 개발(Statistical and Semantic Multimedia Content Analysis: SLAM)</li> <li>- 영상내부의 특징정보를 활용한 검색엔진 개발(Query by image content: QBIC)</li> </ul>	미국
Microsoft	<ul style="list-style-type: none"> <li>- 얼굴 인식 기술을 사용한 이미지 검색엔진(MS Live Image Search-Microsoft live labs)</li> <li>- 사용자가 촬영한 이미지의 객체를 이용하여 검색할 수 있는 프로토타입 개발(World-Wide Media eXchange-Interactive Visual Media Group)</li> <li>- 위치 정보(GPS)를 이용하여 이미지를 태깅하고 검색할 수 있는 시스템 개발(lncdn)</li> </ul>	미국
MIT University	<ul style="list-style-type: none"> <li>- 영상의 통계적인 특징의 기반한 내용 기반의 이미지 검색엔진 개발(Photobook)</li> </ul>	미국
Amsterdam University	<ul style="list-style-type: none"> <li>- 특징 분석을 통한 컨셉 기반의 영상 검색 시스템 개발(MediaMill-Intelligent Systems lab)</li> <li>- 객체 기반의 비디오 검색 시스템 개발(VideoQ)</li> </ul>	네덜란드
Columbia University	<ul style="list-style-type: none"> <li>- Web 을 기반으로 한 영상 탐색 시스템 개발(VisualSEEK)</li> <li>- 이미지의 특정 영역의 색상, 질감, 위치 정보를 사용하는 이미지 검색 시스템 개발(NeTra-Image Processing &amp; Vision Research Labs)</li> </ul>	미국
Riya.com	<ul style="list-style-type: none"> <li>- 영상의 비주얼 정보를 통한 이미지 검색엔진 개발(Riya.com)</li> <li>- 영상의 비주얼 정보를 통하여 상품을 검색하는 검색엔진 개발(Like.com)</li> <li>- 개인영상에서 존재하는 인물 검색엔진 개발(Riya.com)</li> <li>- 이미지의 색상과 질감을 바탕으로 한 이미지 검색엔진 개발</li> </ul>	미국
Tiltomo.com	<ul style="list-style-type: none"> <li>- 이미지의 색상과 질감을 바탕으로 한 이미지 검색엔진 개발</li> </ul>	미국

아니라 다양한 연구 프로젝트를 통하여 콘텐츠 의미 분석 기술 연구를 시도하고 있다. 지금까지 연구되어 온 이미지 인덱싱 기술 내용은 <표 1>에 종합하였다.

상용화된 이미지 검색 시스템으로는 일반적인 이미지 검색 시스템인 Google Image Search, Like.com, Picsearch 등이 있으며, Like.com 과 Picsearch 는 사진의 색상이나 질감 등의 정보를 검색에 이용한다. 소셜 웹 사이트를 대표하는 Flickr, Facebook, Myspace, Webshots 은 온라인을 통해 공유되는 이미지를 관리하기 위해 키워드, 태그에 기반한 검색을 제공한다. <표 2>에서 상용 이미지 검색 시스템의 특징에 관해 종합하였다.

<표 2> 상용 이미지 검색 시스템의 특징

상용 이미지 검색 시스템	메타데이터 특징	이미지 검색 방법의 특징
Google Images	<ul style="list-style-type: none"> <li>- 사용자가 기술한 이미지의 제목, 설명, 그리고 태그를 사용</li> <li>- 메타데이터는 size, type, color, category, domain 정보 기술</li> </ul>	<ul style="list-style-type: none"> <li>- 텍스트 기반이고 내용기반 검색은 지원하지 않음</li> </ul>
Yahoo! Image Search	<ul style="list-style-type: none"> <li>- 사용자가 기술한 이미지의 제목, 설명, 그리고 태그를 사용</li> <li>- 메타데이터는 size, color, domain, adult 정보 기술</li> </ul>	<ul style="list-style-type: none"> <li>- 텍스트 기반이고 내용기반 검색은 지원하지 않음</li> <li>- 성인 영상물을 필터링하는 검색 기술 제공</li> </ul>
Picasearch	<ul style="list-style-type: none"> <li>- 사용자가 기술한 이미지의 제목, 설명, 그리고 태그를 사용</li> <li>- 메타데이터는 type, color, size 정보 기술</li> </ul>	<ul style="list-style-type: none"> <li>- 텍스트 기반이고 내용기반 검색은 지원하지 않음</li> </ul>
Like.com	<ul style="list-style-type: none"> <li>- 전문가가 기술한 이미지의 제목, 설명, 그리고 태그를 사용</li> <li>- 메타데이터는 shape, color 정보 기술</li> </ul>	<ul style="list-style-type: none"> <li>- 텍스트 기반과 내용기반 검색 지원</li> <li>- 검색 결과를 개선하기 위해서 영상의 색상, 형태를 이용하는 내용기반 검색 제공</li> </ul>
Flickr.com	<ul style="list-style-type: none"> <li>- 사용자가 기술한 이미지의 제목, 설명, 태그, 그리고 이미지의 EXIF 정보를 사용</li> <li>- 메타데이터는 location, upload date, EXIF 정보 기술</li> </ul>	<ul style="list-style-type: none"> <li>- 텍스트 기반과 내용기반 검색 제공</li> <li>- 이미지의 콘텍스트 정보(획득 위치, 카메라의 종류)를 통한 검색 제공</li> </ul>

## 2. 이미지 메타데이터 기술 동향

메타데이터는 콘텐츠의 속성들을 기술하는 정보로서 ‘데이터에 대한 데이터’로 설명할 수 있다. 메타데이터는 콘텐츠의 내용에 관한 정보, 다른 정보 자원과의 관계, 콘텐츠의 지적 특성에 관한 사항, 콘텐츠의 형식 및 생성 일자 등 콘텐츠를 구분할 수 있는 사항을 구조화하고 이 구조를 통해서 콘텐츠의 특징을 기술한다[5]. 메타데이터를 통해서 콘텐츠 정보를 기술하고 관리하면 효율적으로 콘텐츠를 검색할 수 있기 때문에 이미지 콘텐츠를 위한 메타데이터 연구도 활발하게 연구되었다.

이미지를 기술하기 위한 대표적인 메타데이터는 EXIF, DIG35 그리고 MPEG-7 등이 있다.

EXIF는 디지털 카메라로 이미지를 획득할 때 발생하는 정보를 저장하는 규격을 의미한다[6]. 이 정보는 실제 이미지 파일 내부에 저장되기 때문에 각각의 이미지는 고유의 EXIF 정보를 가지게 된다. EXIF는 이미지 획득 당시의 셔터 스피드, 노출 보정, 조리개 넘버, ISO 정보, 플래시 동작 정보, 시간과 위치 등을 저장할 수 있다.

DIG35는 DIG(Digital Imaging Group)에 의해서 체계화된 메타데이터로서 상호호환성, 확장성, 범용성, 독립적인 이미지 파일 포맷, 일관성, 인터넷 적용성을 목적으로 개발되었다[7]. DIG35는 크게 기본 이미지 파라미터 정보, 이미지 생성 정보, 콘텐츠 기술 정보, 히스토리 정보 그리고 지적재산권 정보로 구분할 수 있다. 기본 이미지 파라미터 정보는 이미지의 포맷, 크기, 색 공간 등의 이미지 자체에 대한 내용이 기술되고, 이미지 생성 정보는 이미지의 생성 시간, 생성자 그리고 생성에 사용된 장비에 대한 정보가 기술된다. 이미지에 기술된 인물이나 사물 그리고 사건에 관한 내용은 콘텐츠 기술 정보를 통해서 기술되고, 이미지가 수정된 히스토리 정보는 히스토리 정보를 통해서 기술한다. 이미지의 소유자, 소유권 등의 지적재산권에 관련된 내용은 지적재산권 정보를 통해서 기술된다.

MPEG-7은 멀티미디어 콘텐츠에 대한 구조 정보와 의미 정보를 기술할 수 있는 국제 표준 규격으로 메타데이터를 통해서 효율적인 콘텐츠 접근을 제공하고 있다. MPEG-7 기반 멀티미디어 기술 구조(Multimedia Description Schemes: MDS)는 멀티미디어 콘텐츠에 대한 정보 표현을 구조화 하고, MPEG-7 비주얼 기술자(Visual descriptor)는 이미지의 색상, 질감, 형태 등의 내용 기반 특징을 자동으로 추출하고 표현하는 방법을 제공함으로써, 내용 기반 이미지 검색을 가능하게 한다[8]. MPEG-7 MDS는 크게 콘텐츠 메타데이터, 콘텐츠 기술, 콘텐츠 구조, 사용자 상호작용으로 구성되어 있다. 멀티미디어 콘텐츠의 구조적, 의미적 특징 정보는 콘텐츠 기술을 통해서 기술되고, 콘텐츠의 제목, 장르 등의 제작과 포맷, 크기, 사용 정보 등의 정보와 유통, 활용에 대한 정보는 콘텐츠 메타데이터를 통해서 기술된다. 콘텐츠 구조는 콘텐츠의 집합을 구조화 및 모델링하고, 콘텐츠 소비에 관련한 기록과 선호도에 대한 정보는 사용자 상호작용을 통해서 기술된다. <표 3>은 DIG35와 MPEG-7 요소를 기준으로 위에서 설명한 EXIF, DIG35 그

<표 3> 주요 이미지 메타데이터 비교

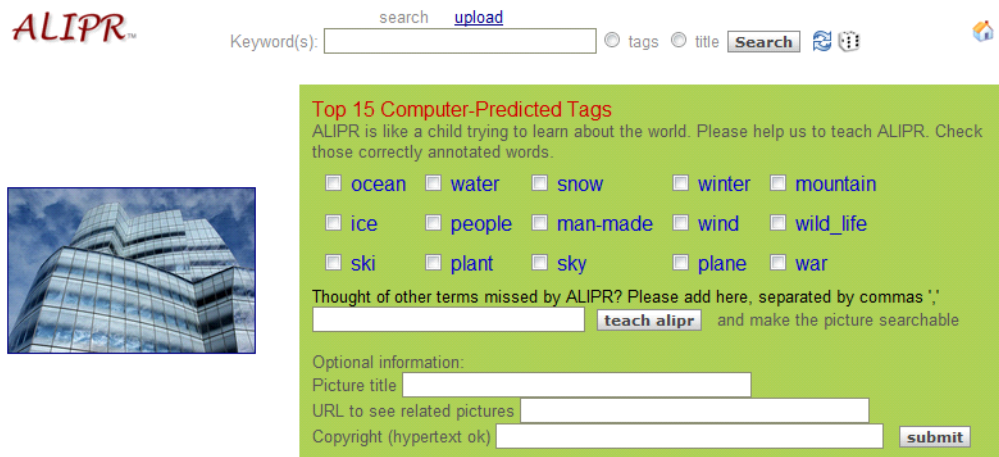
구분	EXIF	DIG35	MPEG-7
기본 이미지 정보(size, format, etc)	O	O	O
이미지 생성 정보(creation date, time, etc)	O	O	O
이미지 의미 정보(who, what, etc)	X	O	O
이미지 내용 기반 특징 정보(color, shape, etc)	X	X	O
이미지 히스토리 정보(editing history, etc)	X	O	O

리고 MPEG-7 기술 내용을 간략하게 정리한 것이다.

### 3. 자동 이미지 태깅 기술 소개

<표 2>와 같이 텍스트에 기반한 이미지 검색 방법은 가장 보편적인 이미지 검색 방법이다. 사용자는 이미지에 나타난 의미 정보를 인지하고 그 정보를 태그로 등록한다. 그리고 등록된 태그는 검색 엔진에 의해서 이미지 검색 과정에 사용된다. 이런 방법은 이미지에 표현된 의미가 사람에 의해서 표현되기 때문에 비교적 정확하게 이미지의 의미가 표현될 수 있다. 하지만, 대용량의 데이터에 대하여 사람이 일일이 태깅하는 것은 불가능한 일이며, 일관성을 잃기 쉽다는 단점이 있다.

자동 이미지 태깅은 입력 이미지와 태그 사이의 상관관계를 분석하여, 사용자에게 자동으로 입력 이미지에 적절한 태그를 추천하는 기술이다. 자동 이미지 태깅은 사용자에게 의한 수동 태깅에서 발생하는 문제를 줄일 수 있다. 학습 태그 추천(supervised tag recommendation)은 기계 학습 방법을 통해서 이미지와 태그 사이의 의미적 일관성을 유지시킨다. 또한, 사용자의 역할이 이미지 관찰, 적절한 태그 생각 그리고 태그 등록의 과정에서 추천된 태그를 선택하는 과정으로 축소되기 때문에 이미지 태깅에 걸리는 시간이 줄어든다. (그림 2)는 미국 펜실베이니아 주립대에서 개발한 자동 태그 추천 시스템 ALIPR(Automatic Photo Tagging and Visual Image Search)이다. 그림의 왼쪽 이미지는 사용자가 입력한 이미지를 나타내고, 그림의 오른쪽은 개발된 시스템에 의해서 자동으로 추천된 15 개의 태그를 나타낸다. 사용자는 추천된 태그 중에서



**ALIPR™** search upload

Keyword(s):  tags title Search

**Top 15 Computer-Predicted Tags**  
 ALIPR is like a child trying to learn about the world. Please help us to teach ALIPR. Check those correctly annotated words.

☐ ocean ☐ water ☐ snow ☐ winter ☐ mountain  
☐ ice ☐ people ☐ man-made ☐ wind ☐ wild\_life  
☐ ski ☐ plant ☐ sky ☐ plane ☐ war

Thought of other terms missed by ALIPR? Please add here, separated by commas ','  
 **teach alipr** and make the picture searchable

Optional information:  
 Picture title   
 URL to see related pictures   
 Copyright (hypertext ok)  **submit**

© alipr.com 2006-2007 Patent Pending. All rights reserved. Do NOT upload objectionable images. Pictures may be subject to copyright.

(그림 2) 상용화된 자동 태깅 서비스

입력 이미지와 적합한 태그를 선택함으로써 이미지 태깅에 소요되는 시간을 줄일 수 있다.

자동 이미지 태깅 기술은 크게 트레이닝 데이터베이스를 이용한 방법과 폭소노미를 이용한 방법으로 구분된다. 트레이닝 데이터베이스를 통한 자동 태깅 기술은 전문가를 통해서 태깅된 트레이닝 데이터베이스의 이미지를 활용하여 입력 이미지에 대한 태그를 사용자에게 추천한다. 트레이닝 데이터베이스를 활용한 자동 태깅 기술은 분류기를 통한 방법과 확률 모델에 기반한 방법으로 구분된다. 분류기를 통한 방법은 이미지와 태그 사이의 관계를 분류기를 통해서 획득한다. 분류기를 통한 방법은 먼저 분류기 훈련 과정을 통해서 이미지에 존재하는 의미의 특징을 사진에 나타난 색상이나 질감 등으로 패턴을 학습시킨다. 훈련 과정을 통해서 분류기가 ‘나무’, ‘바다’ 등의 의미가 학습되면, 입력 이미지로부터 이미지의 색상이나 질감을 추출하고 이를 분류기의 입력으로 사용하여 분류기를 통해서 이미지에 존재하는 의미의 유무를 판단하게 된다. 그리고 분류기를 통해서 입력 이미지에 존재한다고 판단이 된 태그 정보를 사용자에게 추천한다. SVMs(Support Vector Machines)[9]-[11]과 Bayes Point Machines[12]을 분류기로 사용한 연구가 진행되었다. 분류기는 하나의 의미에 대해서만 훈련되기 때문에, 분류기를 통한 자동 태깅 방법을 통해서 추천될 수 있는 태그의 숫자는 훈련된 분류기의 수와 동일하다. 그래서 실제 이미지에 존재하는 무한한 태그를 추천하기 위해서는 무한히 많은 수의 분류기를 훈련시켜야 한다는 한계가 있다.

확률 모델에 기반한 방법은 입력 이미지와 트레이닝 데이터베이스에 있는 이미지 사이의 조건부 확률과 태그가 트레이닝 데이터베이스에서 관찰된 확률을 기반으로 태그와 이미지 사이의 관계를 예측한다. 입력 이미지와 트레이닝 데이터베이스의 이미지 사이의 조건부 확률은 가우시안 분포(Gaussian distribution)로 가정된 이미지의 내용기반 특징의 분포에 의해서 예측된다. 그리고 트레이닝 데이터베이스의 이미지에서 태그를 관찰할 확률은 다항 분포(Multinomial distribution)[13] 또는 다중 베르누리 모델(Multiple Bernoulli Model)[14] 등을 통해서 예측된다. 이와 같은 트레이닝 데이터베이스를 통한 방법은 고정된 트레이닝 데이터베이스에 의존하여 태그를 추천하기 때문에, 입력 이미지에 추천될 수 있는 태그의 종류가 트레이닝 데이터베이스에 등록된 태그 종류에 한정된다.

최근 폭소노미를 통한 자동 태깅 방법이 활발하게 연구되고 있다. 이 방법은 온라인 이미지 공유 사이트를 이용하는 다양한 사용자들의 태깅 활동을 통해서 생성된 폭소노미를 통해서 무한한 종류의 태그를 추천할 수 있다. [15]는 내용 기반 특징 비교를 통해서 입력 이미지와 폭소노미에 존재하는 이미지 사이의 유사도를 예측한다. 그리고 시각적으로 유사한 이미지에 반복적으로 등록된 태그를 사용자에게 추천해주는 방법을 제시한다. 사용자가 이미지에 태그를 등록할



때 빈번하게 동시에 등록되는 태그가 존재한다. 예를 들어서 Flickr.com 에는 ‘바다’와 ‘태양’이 같이 등록된 이미지가 ‘바다’와 ‘스핑크스’가 같이 등록된 이미지보다 약 550 배가 더 많다. 이렇게 태그 사이의 관계를 분석하여 기존에 달려 있는 태그를 확장하기 위한 태그 추천 연구도 진행되고 있다[16]. 뿐만 아니라 태그 사이의 관계 분석을 통해서 사용자가 입력한 태그와 관계가 높은 태그를 실시간으로 추천하고, 사용자가 입력한 태그의 종류가 변화할 때마다 추천 태그가 변화하는 사용자 맞춤 동적 태그 추천 방법도 연구되고 있다[17]. 그리고 분류기를 통해서 입력 이미지에 대한 초기 태그를 추천하고, 추천된 초기 태그와 폭소노미에 등록된 태그와의 관계를 분석하여 추가적으로 다양한 태그를 추천하는 방법도 연구되고 있다[18].

### III. 자동 이미지 태깅 기술 발전 전망

최근 자동 이미지 태깅 기술은 폭소노미를 통한 태그 추천 기술을 중심으로 연구되고 있다. 폭소노미는 다양한 사람들에 의해서 생성된 집합적 지식(collective knowledge)으로 인식되고 있다. 매 분 수 천장의 이미지와 태그가 등록되는 폭소노미는 무한한 수의 이미지가 존재할 뿐만 아니라, 무한한 종류의 태그를 추천할 수 있는 데이터베이스이다. 그러므로 앞으로도 폭소노미를 중심으로 자동 태그 추천 기술이 활발하게 이루어질 것이다.

최근 연구는 폭소노미에 등록된 태그 중에서 상당 부분이 이미지와는 전혀 상관 없는 노이즈라는 관찰을 보고하고 있다[19]. 이는 크게 다음의 2 가지 요인에 기인된 것으로 생각할 수 있다. 첫째, 사용자가 태그를 등록할 때는 이미지 표면에 드러난 의미뿐만 아니라 사용자의 경험, 사전 정보로 이미지를 재생산해내서 이해한 의미도 태그로 등록하는 성향이 있다[18]. 둘째, 사람들은 이미지를 이벤트 단위로 저장하고 관리하는데, 같은 이벤트에서 획득한 사진에는 동일한 태그를 등록하는 특징이 있다. 예를 들어서, 바닷가 근처 휴양지에서 찍은 사진 모두에 ‘바다’라는 태그를 등록하는 성향이 있다. 이렇게 발생된 노이즈 태그는 전체 태그의 약 40% 정도를 차지한다[19]. 이런 폭소노미 내부의 노이즈를 처리하지 않고 태그 추천 시스템을 개발한다면, 태그 추천 성능을 보장하기가 힘들 것이다. 그래서 최근 노이즈를 고려한 자동 태깅에 대한 연구가 시작되고 있다[20].

사용자들은 다양한 멀티미디어 디바이스를 통해서 온라인 이미지 공유 사이트가 제공하는 서비스를 이용하고 있다. 최근에는 핸드폰이나 디지털 카메라에서 직접 온라인 이미지 공유 사이트로 이미지를 업로드하는 응용 기술도 개발되고 있다. 다양한 멀티미디어 환경을 고려한 실시간으로 태그를 추천 기술에 대한 연구가 진행될 것으로 보인다. 개인용 컴퓨터에서 실시간 태



그 추천을 보장하기 위해서 효율적인 이미지의 내용 기반 특징 추출 방법 및 태그 추천 알고리즘은 진행되고 있다[17]. 이런 연구는 이미지의 내용 기반 특징을 효율적으로 추출하는 것에서부터 폭소노미에 등록된 이미지와 태그를 분석하는 기술까지 향후 다양한 각도로 연구될 것으로 보인다.

#### IV. 결 론

본 고에서는 최근 멀티미디어 기기의 확산, 스토리지 가격 하락에 의해서 급격하게 증가한 사용자 생성 이미지를 효율적으로 관리하는 방법을 중요한 문제로 인식하고 이와 관련된 이미지 태깅 기술 동향을 살펴보았다.

온라인 웹 사이트를 통해서 공유되는 대용량의 이미지는 내용 기반 특징에 의한 방법과 키워드에 의한 방법으로 자동 태깅이 이루어지고 있다. 내용 기반 태깅은 색상, 질감 그리고 형태 등의 이미지의 내용 기반 특징을 통해서 쿼리 이미지와 유사한 내용 기반 특징을 가지는 이미지를 자동으로 인덱싱한다. 키워드에 의한 방법은 사용자가 키워드를 쿼리로 입력하면 매칭되는 키워드를 가지고 있는 이미지를 인덱싱한다. EXIF, DIG35 그리고 MPEG-7 메타데이터는 이미지 자동 인덱싱을 위해서 사용되는 다양한 이미지 정보(내용 기반 특징, 이미지 획득 정보 등)를 저장하는 구조를 제공한다.

키워드 기반 이미지 인덱싱 방법은 가장 널리 사용되고 있는 방법이다. 이 방법은 데이터베이스 안에 있는 모든 이미지들이 색인으로 기술이 되어 있을 때, 효율적인 검색을 보장한다. 하지만, 급격하게 증가하고 있는 이미지를 사용자가 수작업으로 관련된 키워드를 등록하는 것은 많은 시간이 요구되는 일이다. 자동 이미지 태깅 기술은 이런 한계를 극복할 수 있는 기술로써 많은 연구가 진행되고 있다. 현재는 다양한 태그를 추천할 수 있는 폭소노미를 기반으로 하는 태그 추천하는 기술이 활발하게 이루어지고 있으며, 주로 입력 이미지와 폭소노미에 존재하는 이미지 사이의 유사도 그리고 폭소노미에 존재하는 태그 사이의 관계 분석을 통해서 연구가 진행되고 있다. 향후에는 폭소노미를 활용한 실시간 태그 추천 기술뿐만 아니라, 폭소노미에 존재하는 노이즈를 고려한 태그 추천 기술이 필요할 것으로 분석된다.

#### <참 고 문 헌>

- [1] Flickr blog, "3 Billion!", November 2008. Available on <http://blog.flickr.net/en/2008/11/03/3-billion/>

- [2] 이지연, “이용자 관점에서 본 이미지 색인의 객관성에 대한 연구”, 정보관리학회지, 제 19 권 제 3 호, 2002. 9, pp.123-144.
- [3] B. Sigurbjörnsson and R. van Zwol, “Flickr Tag Recommendation based on Collective Knowledge”, Proc. of 17th International Conference on World Wide Web(WWW), April 2008, pp.327-336.
- [4] R.C. Veltkamp, M. Tanase, “Content-Based Image Retrieval Systems: A survey”, Technical Report UU-CS-2000-34, 2002.
- [5] 고영만, “메타데이터 표준화와 메타데이터 레지스트리”, 국회도서관보, 제 42 권 제 11 호, 2005. 11, pp.18-26.
- [6] Japan Electronics and Information Technology Industries Association, “Exchangeable image file format for digital still cameras: Exif Version 2.2”, April 2002.
- [7] Digital Imaging Group, “DIG35 Specification-Metadatas for Digital Images. Version 1.1”, April 2001.
- [8] Manjunath B.S. et al., Introduction to MPEG-7: John Wiley and Sons, 2002.
- [9] S. Yang, S. Kim and Y. M. Ro, “Semantic Home Photo Categorization”, IEEE Trans. Circuits and Systems for Video Technology, Vol.17, No.3, March 2007, pp.324-335.
- [10] S. Yang, S. K. Kim, K. S. Seo, Y. M. Ro, J. Kim and Y. S. Seo, “Semantic categorization of digital home photo using photographic region templates”, International Journal of Information Processing and Management, Vol.43, No.2, March 2007, pp.503-514.
- [11] S. Yang, R. M. Ro, “Photo Indexing Using Person-Based Multi-feature Fusion with Temporal Context”, Proc. International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, Nov. 2007, pp.257-262.
- [12] E. Chang, G. Kingshy, G. Sychay, G. Wu, “CBSA: Content-Based Soft Annotation for Multimodal Image Retrieval using Bayes Point Machines”, IEEE Trans. Circuits and Systems for Video Technology, Vol.13, No.1, January 2003, pp.26-38.
- [13] V. Lavrenko, R. Manmatha, and J. Jeon, “A Model for Learning the Semantics of Pictures”, Proc. of 7th Annual Conference on Neural Information Processing Systems, Vol.16, 2003, pp.553-560.
- [14] S. L. Feng, R. Manmatha, V. Lavrenko, “Multiple Bernoulli Relevance Models for Image and Video Annotation”, Proc. of International Conference on Computer Vision and Pattern Recognition, Vol.2, No.2, June 2004, pp.1002-1009.
- [15] O. Kucuktunc, S.G. Sevil, A.B. Tosun, H. Zitouni, P. Duygulu, and F. Can, “Tag Suggestr: Automatic Photo Tag Expansion using Visual Information for Photo Sharing Websites”, Proc. of International Conference on Semantic and Digital Media Technologies(SAMT), Dec. 2008, pp.61-73.
- [16] B. Sigurbjörnsson and R. van Zwol, “Flickr Tag Recommendation based on Collective Knowledge,” Proc. of 17th International Conference on World Wide Web(WWW), April 2008, pp.327-336.
- [17] J. Li and J. Z. Wang, “Real-Time Computerized Annotation of Pictures”, IEEE Trans. Pattern

- Analysis and Machine Intelligence, Vol.30, No.6, June 2008, pp.985-1002.
- [18] R. Mörzinger, R. Sorschag, G. Thallinger and S. Lindstaedt, "Automatic Image Annotation using Visual Content and Folksonomies", Multimedia Tools and Applications, Vol.42, No.1, March 2009, pp.97-113.
- [19] T. Chua, J. Tang, R. Hong, H. Li, Z. Luo, Y. Zheng, "NUS-WIDE: A Real-World Web Image Database from National University of Singapore", Proc. of ACM International Conference on Image and Video Retrieval. July 2009, pp.223-232.
- [20] S. Lee, S. H. Min, Y. B. Lee and Y. M. Ro, "Measurement of Tag Confidence in User Generated Contents Retrieval", Proc. International Society for Optical Engineering, Vol.7259, Jan. 2009, pp.1-6.

\* 본 내용은 필자의 주관적인 의견이며 NIPA의 공식적인 입장이 아님을 밝힙니다.