



자동 이미지 태깅에 관한 연구

A Research on Automatic Image Tagging

저자 (Authors)	전우경, 이일병 Woogyoung Jun, Yillbyung Lee
출처 (Source)	한국정보과학회 학술발표논문집 39(1D) , 2012.6, 85-87 (3 pages)
발행처 (Publisher)	한국정보과학회 KOREA INFORMATION SCIENCE SOCIETY
URL	http://www.dbpia.co.kr/Article/NODE01907181
APA Style	전우경, 이일병 (2012). 자동 이미지 태깅에 관한 연구. 한국정보과학회 학술발표논문집, 39(1D), 85-87.
이용정보 (Accessed)	대구가톨릭대학교 203.250.33.*** 2018/01/12 22:22 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독 계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

자동 이미지 태깅에 관한 연구

전우경[○] 이일병

연세대학교 컴퓨터과학과 인공지능연구실

woogyoun@yonsei.ac.kr, yillbyunglee@yonsei.ac.kr

A Research on Automatic Image Tagging

Woogyoun Jun[○] Yillbyung Lee

Artificial Intelligence Lab., Dept. of Computer Science, Yonsei University

요 약

최근 모바일 기기는 물론 디지털 카메라, SNS의 발전으로 인하여 매일 방대한 양의 디지털 이미지가 생성된다. 따라서 효과적이고 신뢰도 있는 인덱싱 기법과 탐색 기법이 요구되고 있다. 이미지 태깅은 효과적이고 신뢰도 있는 이미지 탐색에 큰 연관관계가 있다. 본 연구에서는 여러가지 이미지 태깅 기법들을 서베이하고 자동 및 반 자동 이미지 태깅 기법들에 대하여 알아본다.

1. 서론

최근 모바일 기기는 물론 디지털 카메라, SNS 등의 급속한 발전에 따라 노트북, 스마트폰 등 많은 종류의 디바이스로 언제 어디서든 인터넷에 접속이 가능케 되었다. 유비쿼터스 컴퓨팅은 정보 공간과 물리적인 공간으로 구성되는데, 이 때문에 언제 어디서든 인터넷에 접속이 가능한 것이다. 이러한 것들이 디바이스 사용의 어려움을 줄여주고, 간결하게 해주고, 잘 정리될 수 있게 해준다[1].

모바일 디바이스들의 발전으로 인하여 많은 양의 이미지들이 생성이 되고 있는데 따라서 성공적인 인덱싱과 이미지 탐색 기법이 중요하다. 내용기반 검색 기법은 가장 대표적인 방법이지만 이미지 콘텐츠를 검색하는 것은 아직도 발전해야 할 분야이다. 이미지 탐색을 위한 가장 기본적인 접근 방법은 키워드 기반 검색인데, 이미지 콘텐츠 자체에 의미 있는 메타데이터를 생성하는 것은 아직도 부족하다. 태깅은 사진의 효용 가치를 충분히 높일 수 있을 것이고, 사진의 검색을 위해서도 효과적인 것이다[2].

이미지 태깅은 내용기반 이미지 검색을 위한 전문적인 기법이라고 볼 수 있다. 수동적인 이미지 태깅은 특히나 모바일 디바이스에서는 많은 시간과 비용을 필요로 한다. 사진문맥, SNS문맥 등과 같이 모바일 장치를 위한 최적의 문맥들이 있는데 이러한 문맥들은 사진의 실용적인 의미를 얻기 위해 결합 되곤 한다. 이렇게 생성된 결과는 태그로서 제안되기도 한다. 태그는 이벤트, 시간, 사람, 장소 등 많은 정보를 포함할 수 있다[3].

2장에서 관련 연구를 소개하고 3장에서 이미지 태깅 기법에 대하여 소개를 한다. 4장에서는 문맥기반 반 자동 이미지 태깅 기법을 소개하고 5장에서 결론을

맺는다.

2. 관련연구

이미지 태깅과 이미지 탐색은 최근 아주 중요한 연구 분야 중에 하나이다. 이미지 태깅과 탐색에서 텍스트 태깅 역시 중요한 역할을 차지한다. 이미지에 대한 효율적인 라벨링은 최근 연구가 활발히 진행 중에 있다. [4]에서 저자는 개인적인 사진 콜렉션에 중점을 두고 GPS와 시간 메타데이터 정보를 문맥정보로 활용한다. 사진은 콜렉션이나 그룹기반으로 태깅이 된다. [5]에서는 모바일 디바이스에서의 사진 태깅 기법을 다루고 있다. 공간적이고 임시적인, 그리고 대중적인 사진들이 자동 태깅을 위해 쓰였다. [6]에서는 사진이 태깅 될 때 사진의 종류에 따라 다른 방법들이 쓰여야 한다고 설명하고 있다. [7]에서는 태깅 제안에 있어서 두 가지 유형을 가지고 태깅 하는 모델을 제안하고 있다. 그것은 특정 사용자의 문맥 또는 커뮤니티 네트워크 중 하나가 될 수 있다. [8]에서는 모바일 디바이스에서 사진 공유를 위한 모델이 소개된다. 이는 웹 서버 컴퓨팅 및 사용자 상호작용을 기반으로 태깅 제안을 한다.

3. 이미지 태깅 기법

3.1 텍스트 정보를 활용한 태깅 기법

인터넷에는 수많은 이미지들이 존재한다. [9]에서, 그 이미지들을 분류하고 적절히 검색할 수 있도록 이미지의 백그라운드 정보가 이미지 태깅을 위해 사용된다. 자동적으로 웹의 이미지 태깅을 할 수 있다. 또한 비슷한 문맥이 페이지 레이아웃 분석 기법을 통하여 이미지에 부여될 수 있다. 하지만 이미지 검색에

있어서 정확도는 떨어질 것인데, 그 이유는 첫째, 이미지는 누구나 웹에서 사용될 수 있고 그 웹 페이지와 이미지는 표준이 없기 때문이며 둘째, 웹 이미지는 의미론적인 측면에서 상당히 범위가 넓다. 셋째, 웹 이미지의 품질의 차이가 상당히 두드러지게 나타나기 때문이다. 사용자는 원하는 결과를 얻기 위해서 검색 결과 전체를 살펴보아야 한다.

3.2 수동 태깅 기법

수동 태깅에서는 사용자가 이미지에 맞는 특정 키워드를 직접 입력 해야 한다. 수동 태깅은 사실상 가장 정확도가 높고 신뢰성이 높은 방법이다. 하지만 상당한 시간과 노력이 소요되며 해당 태그가 사용자에게 기억되지 않을 수도 있다.

3.3 온톨로지 기반 이미지 태깅 기법

온톨로지와 같은 시멘틱 웹 기반 태깅 방법 역시 태깅을 위해 사용될 수 있다. 온톨로지는 개요와 요약등을 위한 설계라고 볼 수 있다. 온톨로지는 컬렉션등을 정의한다. 온톨로지 기반 이미지 태깅의 초점은 이미지들의 상관관계를 따지는 것에 있다. 이미지 태깅을 위한 3단계 구조[10] 기법이 제안되고 있다. 아랫단에 의해 이미지의 저차원 특징이 선택된다. 이 특징들은 중간 단에서 시멘틱하게 키워드와 맵핑된다. 이 키워드들은 상단에서 스키마와 온톨로지에 연결된다.

3.4 반 자동 이미지 태깅 기법

저 차원 특징과 고차원적인 인간 시멘틱에는 많은 차이가 존재한다. 그래서 반 자동 이미지 태깅에서는 사용자의 참여를 필요로 한다.

[10]에서, 반 자동 이미지 태깅을 위해 사용자가 참여하는 머신 러닝 기법을 설명하고 있다. 사용자의 피드백이 존재하는 머신 러닝 기법은 태깅의 정확도를 상승시킨다. 머신 러닝 기법은 일관성이 있고, 비용적인 측면에서도 효과적이며 지능화된 태깅 방법이다. Intelligent Image Indexing Web Service(I3WS)가 이 방식에 사용된다.

[11]에서는 이미지의 시멘틱 의미를 습득하기 위해서 이미지는 콘텐츠에 의해 분해된다. 이미지의 결과 분류는 구조적으로 정의 구조의 가장 루트에 위치하게 된다. 키워드의 시퀀스는 이미지를 태깅하기 위해 사용된다. 저 차원 특징추출 단계에서 만들어진 학습 예제들이 입력으로 사용된다. 저 차원과 고차원 사이의 차이를 줄여주기 위해서 사용자 참여가 필요한 것이다.

[12]의 반 자동 이미지 태깅 기법은 자동 태깅 기법의 효율과 수동 태깅 방법의 정확도의 조합이다. 사용자가 피드백을 제공 해야 하며 이 방법은 세 가지의 구성요소로 구성된다. 쿼리를 날려주는 부분과, 이미지 브라우저, 그리고 피드백을 할 수 있는 인터페이스다. 사용자가 쿼리를 입력하며 랭크 된

이미지가 브라우저에 표현되고, 사용자는 각각의 이미지에 피드백을 줄 수 있고 그 피드백은 이미지 랭크에 영향을 미친다.

3.5 자동 이미지 태깅 기법

[13]에서의 자동 이미지 태깅 기법은 이미지 분할 알고리즘이 사용된다. 이미지 태깅을 위해 전역적 특징을 사용한다. 이 모델링 프레임워크는 'kernel smoothing'이라는 비모수 밀도 판단방법에 기반한다.

[14]의 방법에서는 이미지의 학습 셋이 자동 태깅을 위해 사용된다. 이미지를 묘사하기 위해서 어휘가 사용된다. 트레이닝 셋에 속한 이미지의 도움으로 인해서 태그생성에 있어서 구체적인 이미지의 짐작이 가능한 것이다.

[15]에서는 단어의 상관관계를 위해 단어가 사용되는데, 이미지의 특징이 불충분한 경우가 있기 때문이다. 단어의 상관관계를 따지기 위해 단어가 사용되는 것을 통합시키기 위해 이미지 태깅의 가능성을 예측할 필요가 있다. 이 방법은 언어 모델을 사용하며 단어 개연성을 포함하게 되는데, 개연성이란 태그를 위한 단어가 얼마나 가능성이 있느냐는 것이다.

[16]에서는 기존의 태깅 기법을 개선했다. 조건부 개연성을 개선하여 더욱 정확한 태그를 제공한다. 그 영향으로 인해서 태그는 최종 태그로 가장 높은 가능성을 가지고 있게 된다. 이미지 쿼리를 위해 기존의 이미지 태그 방법이 사용된다. 그러면 지원자의 태그가 다시 랭크 되고 가장 높은 랭크의 태그만이 가장 높은 가능성을 가지게 된다.

표 1. 이미지 태깅 기법의 비교

기법	장점	단점
수동 태깅	높은 신뢰도	노력, 시간비용
반자동 태깅	수동보다 능률적	피드백 인터페이스 필요
자동 태깅	빠른 태깅 속도	신뢰도 낮음

4. 문맥 기반 반자동 이미지 태깅 기법

유비쿼터스 컴퓨팅 환경에서 문맥의 세 가지 중요한 측면이 지역, 정체, 그리고 시간이다. 한가지 더 생각해볼 수 있는 것은 이벤트라고 볼 수 있다. 그래서 사진의 네 가지 중요한 특징은 시간, 지역, 사람, 그리고 이벤트라고 할 수 있다. 세 종류의 문맥은 사진 문맥, 개인적인 문맥 그리고 소셜네트워크 문맥 이다.

사진 문맥은 사진을 찍을 때, 카메라의 시간이 사진에 기록되고 시간은 EXIF라고 불리는 사진 메타데이터에 저장되는데, 그 EXIF정보를 의미한다. 비슷하게 모바일 장치가 GPS모듈을 탑재하고 있다면 위치 정보 역시 저장된다. 이러한 방법으로 임시적이고 공간적인

콘텐츠가 얻어진다. 이 데이터들과 함께 EXIF정보에는 작가정보도 저장할 수 있다.

개인의 장비에 따라서, 추가적인 사용자 문맥 정보가 사용될 수 있다. 모바일 디바이스에서의 일반적인 기능인 캘린더와 연동하여 캘린더에 저장된 일정에 촬영된 사진은 그 이벤트와 연관성이 있을 수 있다. 사진에서의 최고 문맥은 바로 이벤트라고도 할 수 있다. 캘린더와의 연동으로 자동 이미지 태깅이 수월하게 된다. 이것을 개인적인 문맥이라고 한다.

소셜네트워크 문맥은 SNS의 가장 큰 목적 중에 하나는 바로 사진 공유인데 바로 이 소셜 문맥이 태깅에 사용될 수 있다. 소셜 컨텍스트 기반 사진 공유 시스템에서, 사용자간의 친분은 이미지의 암시적인 태그라고 볼 수 있다. 이러한 태그는 서버 쪽에서 추가되어야 한다. 소셜 관계는 사진 공유 시스템에서 상당히 중요하다. 시스템에서 친구를 추가하는 방법에는 첫째로 이름 검색에 의한 추가, 둘째 사진에 의한 추가, 이메일 주소록에 의한 불러오기 방법 등이 있다.

모바일 디바이스를 이용한 입력 방법은 상당히 제한적이기 때문에 태깅 방법은 상당히 간단하다. 태깅 알고리즘은 최대한 간결해야 한다. 모바일 장치에서 태깅을 하기 위해서는 메타데이터 분석-개인적인 문맥수집-태깅-업로드 의 과정을 따른다. 사진이 업로드 되기 위해 선택 되었을 때, 시스템은 사진의 EXIF정보를 읽어 들인다. 캘린더를 통해 사진이 찍힌 시간에 대한 이벤트를 읽어오고 결과는 태그 제안 목록에 추가된다. 사진이 업로드 되면, 서버는 GPS와 기타 메타데이터를 수집하고 소유자 정보 등을 태그한다. 서버는 사진을 모으고 썸네일을 생성한다. 그동안 시간인덱스, 사용자 인덱스, 이모티콘 인덱스, 지역 인덱스, 관련성 인덱스, 이벤트 인덱스와 같은 사진의 다차원 인덱스를 생성한다. 이 인덱스는 이미지를 검색하는데 사용된다.

5. 결론

수동 태깅은 특히 모바일 디바이스에서 비용적, 시간적 측면에서 비 효율적이다. 본 연구에서는 반 자동, 자동 이미지 태깅 기법에 대해서 논하였다. 반 자동 태깅 기법이 정확도 측면에서 사용자가 태깅에 관여를 하고 있다면 다른 어떠한 기법 보다 좋은 태깅 기법이라는 것을 보여주고 있다. 모바일 디바이스의 문맥 정보를 이용하여 이미지 태깅을 하는 반 자동 기법을 본 연구에서 논하였다. 머신 러닝 기법이 이미지 태깅 기법과 접목된다면 이미지 태깅은 보다 지능적이고 정확한 자동 태깅 기법이 될 것이다.

6. 참고문헌

[1] Wei Liu, Xue Li, and Daoli Huang "A Survey on Context Awareness", Computer Science and Service System (CSSS),

International Conference on, 29 June 2011 IEEE.

[2] B. Shevade, H. Sundaram, and L. Xie. "Modeling Personal and Social Network Context for Event Annotation in Images". In JCDL 2007, ACM Press (2007).

[3] Shuangrong Xia, Xiangyang Gong, Wendong Wang, and Ye Tia "Context-Aware Image Annotation and Retrieval on Mobile Device", 2010 IEEE.

[4] L. Cao, J. Luo, H. Kautz, and T. S. Huang. "Image Annotation within the Context of Personal Photo Collections Using Hierarchical Event and Scene Models", In IEEE Multimedia 2009 11(2), 208–219.

[5] W. Viana, J. B. Filho, J. Gensel, M. Villanova-Oliver, and H. Martin. "PhotoMap: From location and time to context-aware photo Annotations", In Journal of Location Based Services 2008 2(3), 211–235.

[6] M. Ames, and M. Naaman. "Why We Tag: Motivations for Annotation". In proc. CHI 2007, ACM Press (2007), 971–980.

[7] U. WESTERMANN and R. JAIN. "Toward a Common Event Model for Multimedia Applications", In IEEE Multimedia 2007 14(1), 19–29.

[8] M. Davis, N. V. House, J. Towle, S. King, S. Ahern, C. Burgener, Perkel, M. Finn, V. Viswanathan, and M. Rothenberg. "MMM2: Mobile Media Metadata for Media Sharing", Ext. Abstracts CHI 2005, ACM Press (2005), 1335–1338.

[9] Asullah Khalid Alham, Maozhen Li, Suhel Hammoud and Hao Qi "Evaluating Machine Learning Techniques for Automatic Image Annotations", 2009 IEEE.

[10] O. Marques, and N. Barman, "Semi-Automatic Semantic Annotation of Images Using Machine Learning Techniques", Proc. of ISWC, pp. 550–565, 2003.

[11] J. Vompras, and S. Conrad, "A Semi-Automated Framework for Supporting Semantic Image Annotation". Proc. of ISWC, pp. 105–109, 2005.

[12] L. Wenyin, S. Dumais, Y. Sun, H. Zhang, M. Czerwinski and B. Field, "Semi-Automatic Image Annotation", Proc. Of INTERACT, pp. 326–333, 2001.

[13] A. Yavlinsky, E. Schofield, and S. M. Rüger, "Automated Image Annotation using Global Features and Robust Nonparametric Density Estimation", Proc. of CIVR, pp. 507–517, 2005.

[14] Jeon, V. Lavrenko, and R. Manmatha, "Automatic Image Annotation and Retrieval using Cross-Med Relevance Models", Proc. ACM SIGIR, pp. 119–126, 2003.

[15] R. Jin, J. Chai, and L. Si, "Effective Automatic Image Annotation via a Coherent Language Model and Active Learning", Proc. of ACM Conference on Multimedia, pp. 892–899, 2004.

[16] C. Wang, F. Jing, L. Zhang, and H. Zhang, "Content-Base Image Annotation Refinement", Proc. of CVPR, 2007.