

소셜 이미지 분류를 위한 클러스터링 알고리즘 기반 트레이닝 집합 획득 기법의 비교

정진우, 이동호*

한양대학교 컴퓨터공학과

e-mail : selphyr@hanyang.ac.kr, dhlee72@hanyang.ac.kr

A Study on Comparison of Clustering Algorithm-based Methods for Acquiring Training Sets for Social Image Classification

Jin-Woo Jeong, Dong-Ho Lee

Dept. of Computer Science and Engineering, Hanyang University

요 약

최근, Flickr, YouTube 와 같은 사용자 참여형 미디어 공유 및 검색 사이트가 폭발적으로 증가하면서, 이를 멀티미디어 정보 검색 서비스에 효과적으로 활용하기 위한 다양한 연구들이 시도되고 있다. 특히, 이미지에 할당되어 있는 태그를 이용하여 이미지를 효과적으로 검색하기 위한 연구가 활발히 진행 중이다. 그러나 사용자들에 의해 제공되는 소셜 이미지들은 매우 다양한 범위와 주제를 가지고 있기 때문에, 소셜 이미지들의 분류 및 태그 할당을 위한 트레이닝 집합의 획득이 쉽지 않다는 한계점을 가지고 있다.

본 논문에서는 데이터 군집화를 위한 클러스터링 알고리즘들 중 K-Means, K-Medoids, Affinity Propagation 을 활용하여 소셜 이미지 집합으로부터 트레이닝 집합을 획득하기 위한 방법들을 살펴본다. 또한, 각 알고리즘으로부터 획득한 트레이닝 집합을 이용하여 소셜 이미지를 분류한 결과를 비교 분석한다.

1. 서론

최근, Flickr, YouTube 와 같은 사용자 참여형 미디어 공유 및 검색 사이트가 폭발적으로 증가하면서, 이를 멀티미디어 정보 검색 서비스에 효과적으로 활용하기 위한 다양한 연구들이 시도되고 있다.

Flickr[1]와 같은 소셜 이미지 공유 사이트는 사용자들이 직접 업로드 하는 이미지 및 태그를 바탕으로 이미지의 공유 및 검색 서비스를 제공하고 있다. 이러한 소셜 이미지 데이터베이스는 실시간으로 추가되는 다양한 이미지 정보를 기반으로 최근 이미지 검색의 중요한 원천으로서 새롭게 주목 받고 있다.

태그를 이용한 이미지 검색을 위한 연구는 크게 1) 태그 랭킹 2) 태그 추천 및 태그 수정 3) 태그 분류와 같은 분류로 구분될 수 있다.

[2,3,4,5]의 연구는 소셜 이미지에 할당되는 태그를 이미지에 적합한 순서에 따라 랭킹하기 위한 연구들이다. 태그는 사용자에게 의하여 어떠한 제약 없이 할당되는 키워드이기 때문에, 이미지와의 연관성 및 중요성을 보장할 수 없다. 이들 연구는 각 태그와 이미지 간의 연관성을 분석하기 위하여 kernel density

estimation 기반의 scoring 기법, neighbor voting 기법, WordNet 기반 의미 관계 분석 기법 등을 제안하였다.

[6,7]의 연구는 토픽 모델에 기반하여 이미지에 할당되어 있는 태그를 이미지에 보다 적합한 태그로 수정·추천하기 위한 기법들을 제안하였다.

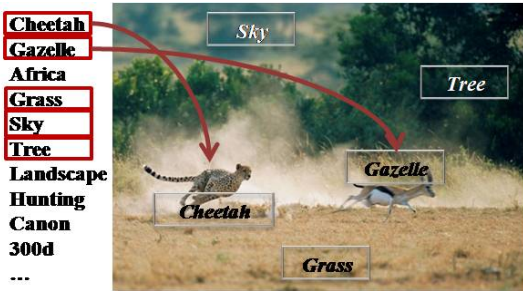
[8, 9, 10]의 연구는 태그를 기반으로 이미지를 분류하거나, 이미지의 특정 영역을 분류하기 위한 연구들이다. 특히, [9, 10]의 연구는 (그림 1)과 같이, 이미지에 할당되어 있는 태그들을 이미지의 영역에 매핑함으로써 이미지 내 개념정보들의 공간적 정보를 활용하여 보다 효과적인 이미지 검색 서비스를 제공하기 위한 연구를 진행하고 있다. 태그를 이미지 내 영역에 매핑하기 위해서는 우선적으로 해당 태그와 이미지 내 영역간의 연관성을 분석해야 한다. 즉, 특정 태그가 가질 수 있는 대표적인 시각적 특징정보를 획득한 후, 이미지 내 영역의 시각적 특징정보들과 비교하여 가장 적합한 영역을 추출할 수 있어야 한다. 기존의 이미지 어노테이션을 위한 연구에서는 키워드에 해당하는 대표적인 시각적 특징정보를 획득하기 위하여 일련의 트레이닝 집합을 준비하고, 이를 바탕으로 기계학습 분류자를 학습시키거나 규칙기반 분류자를 생성하였다[11, 12]. 그러나 소셜 태그는 매우 다양한 주제를 포함하고 있기 때문에 기존의 연구들과 같이

*교신 저자

고정적인 트레이닝 이미지 집합을 획득하여 분류자를 학습시키는 것은 불가능하다. 따라서, 이미지에 할당되어 있는 태그의 대표적인 특징들을 학습하기 위해선 실시간으로 트레이닝 집합을 획득하기 위한 방법이 필요하다.

본 논문에서는 데이터 군집화를 위한 클러스터링 알고리즘들 중 K-Means, K-Medoids, Affinity Propagation을 활용하여 소셜 이미지 집합으로부터 트레이닝 집합을 획득하기 위한 방법들을 살펴본다. 또한, 각 알고리즘으로부터 획득한 트레이닝 집합을 이용하여 소셜 이미지를 분류한 결과를 비교 분석하였다.

본 논문의 구성은 다음과 같다. 2 장에서는 각각의 데이터 클러스터링 알고리즘에 대하여 간략히 살펴본 후, 각 알고리즘을 활용하여 실시간으로 트레이닝 집합을 획득하는 방법에 대하여 서술한다. 3 장에서는 각 알고리즘으로부터 획득한 트레이닝 집합을 기반으로 이미지 분류 실험을 수행한 결과를 비교 분석한다. 마지막으로, 4 장에서 결론 및 향후 연구에 대하여 기술한다.



(그림 1) 태그기반 영역분류의 예

2. 트레이닝 집합 획득을 위한 클러스터링 알고리즘

본 논문에서는 데이터 군집화 알고리즘 중 K-Means, K-Medoids, Affinity propagation 알고리즘을 활용하여 실시간 트레이닝 이미지 집합 획득 과정을 수행하였다. 본 장에서는 각각의 클러스터링 알고리즘에 대하여 살펴본 후, 각 알고리즘을 활용하여 트레이닝 집합을 획득하기 위한 과정에 대하여 기술한다.

2.1 K-Means

K-Means 알고리즘은 클러스터 기반 비계층적 군집화 알고리즘의 하나로, 데이터 집합으로부터 k 개의 클러스터를 발견하기 위한 알고리즘이다. K-Means 알고리즘의 기본적인 작업 절차는 알고리즘 1과 같다.

Algorithm 1 K-Means Clustering

Input : initial K points
Output: K clusters
1. Choose K initial centroid points
2. Assign all data points to the closest centroid point
3. Re-compute the centroid point of each cluster
4. Repeat 3-4 until centroid of each cluster is not changed

K-Means 알고리즘은 수식 (1)과 같이 계산될 수 있는 전체 클러스터 결과의 오류를 최소화하기 위한 알

고리즘으로 생각될 수 있으며, 여기서 각 클러스터의 centroid p_i 는 식(2)와 같이 클러스터의 mean 포인트로 계산될 수 있다.

$$\text{Error} = \sum_i^K \sum_{j \in C_i} |p_i - j|^2 \quad (1)$$

$$p_i = \frac{1}{N_i} \sum_{x \in C_i} \vec{x} \quad (2)$$

2.2 K-Medoids

K-Means 알고리즘에서 각각의 centroid는 식 (2)와 같이 각 클러스터의 중앙점으로 결정된다. 즉, 결정된 centroid가 실제 데이터 포인트일 필요는 없다. K-Medoids 알고리즘은 K-Means 알고리즘과 달리 각각의 클러스터를 대표할 수 있는 representative point (medoid)를 선정하고 이를 기반으로 클러스터링을 수행한다.

K-Medoids 알고리즘은 클러스터의 대표자인 medoid가 실제 데이터 포인트이며, 이를 기반으로 클러스터 오류를 최소화하기 위한 과정을 처리하는 것 외에는 K-Means 알고리즘(알고리즘 1)과 유사하다.

2.3 Affinity Propagation

Affinity propagation 알고리즘은 데이터 포인트간의 유사도(similarity)를 반영하는 메시지들을 교환함으로써 클러스터를 구축해나가는 알고리즘으로서, 각 데이터 포인트로 하여금 자신의 대표자(exemplar)를 선택하게 하고, 동일 대표자를 가지는 데이터 포인트를 하나의 군집으로 인식하는 알고리즘이다. 식 (3)과 식 (4)는 affinity propagation 알고리즘에서 데이터들간 교환되는 메시지인 responsibility와 availability를 나타내고 있다. 각 데이터 포인트 i 의 대표자 $C(i)$ 는 식 5와 같이 $r(i, k) + a(i, k)$ 를 최대화 시켜주는 데이터 포인트 k 로 선정된다.

$$r(i, k) = s(i, k) - \max_{k' \neq k} [s(i, k') + a(i, k')] \quad (3)$$

$$a(i, k) = \begin{cases} \sum_{i' \neq i} \max[0, r(i', k)], & \text{for } k = i \\ \min[0, r(k, k) + \sum_{i' \neq \{i, k\}} \max[0, r(i', k)]] & \text{for } k \neq i \end{cases} \quad (4)$$

$$C(i) = \text{argmax}_k r(i, k) + a(i, k) \quad (5)$$

2.4 클러스터링 기반 실시간 트레이닝 집합 획득

본 논문에서는 특정 태그의 대표적 특징 벡터를 획득하기 위하여, 알고리즘 2와 같이, 소셜 이미지 공유사이트인 Flickr로부터 해당 태그와 연관된 이미지들을 획득한 후, 각각의 클러스터링 알고리즘을 적용하여 태그에 적합하다고 판단되는 이미지들을 트레이닝 집합으로 제공한다.

Algorithm 2 Online Training Example Acquisition

Input : tag
Output: a set of training samples for the tag
1. Collect images from Flickr with respect to the inputted tag
2. Generate N clusters through clustering algorithm from collected images
3. Return top k largest clusters

3. 이미지 분류 성능 평가

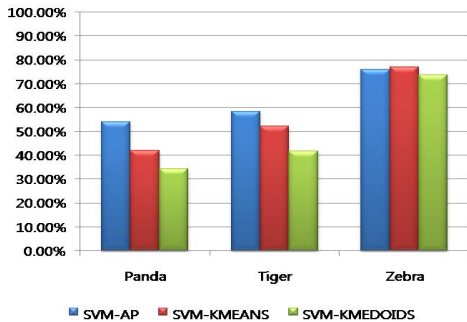
본 장에서는 각각의 클러스터링 알고리즘을 이용하여 구축된 트레이닝 집합을 이용하여 이미지 분류자를 학습한 후, 테스트 집합을 분류한 실험결과를 비교 분석한다.

3.1 실험 환경

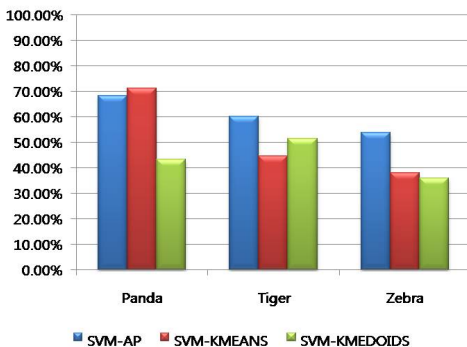
본 논문에서는 ‘Tiger’, ‘Zebra’, ‘Panda’ 세 개의 태그에 대하여 트레이닝 집합을 획득하고, 획득된 집합을 이용하여 이미지 분류자를 학습시킨 후, 이미지 분류 성능을 평가하였다. 각 태그 별로 300 개의 이미지를 Flickr 로부터 다운로드 후, 총 15 개의 클러스터를 생성하도록 각각의 클러스터링 알고리즘의 파라미터를 조정하였다. 그 후, 클러스터의 크기를 기반으로 클러스터를 정렬한 후, 약 100~120 개의 트레이닝 이미지를 제공 할 수 있도록 상위 클러스터들을 선택하였다. ‘Panda’, ‘Tiger’, ‘Zebra’ 카테고리의 이미지 분류에 사용된 테스트 집합은 각각 992, 1176, 1042 개의 이미지를 포함하고 있다.

3.2 이미지 분류 성능

본 연구에서는 2.4 절의 알고리즘을 통하여 획득한 트레이닝 집합을 바탕으로 SVM 과 Bayes classifier 를 학습 시킨 후, 각각의 이미지 분류 결과의 precision 과 recall 을 비교하였다.



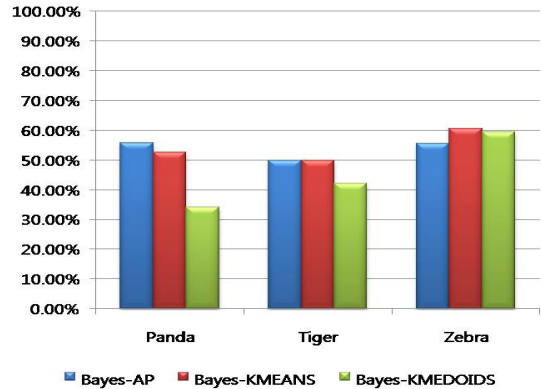
(그림 2) SVM 을 이용한 분류 결과의 precision 비교



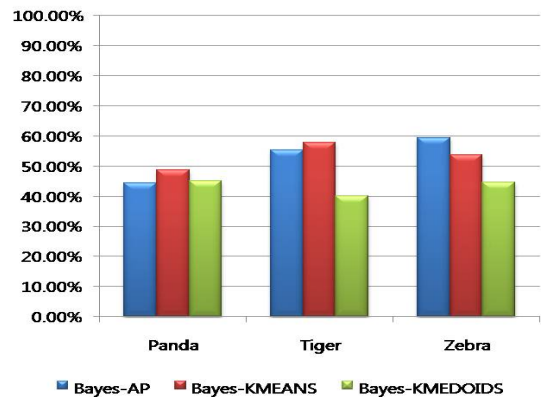
(그림 3) SVM 을 이용한 분류 결과의 recall 비교

(그림 2)와 (그림 3)은 각각의 클러스터링 알고리즘을 이용하여 획득한 트레이닝 집합을 이용하여 SVM 을 학습시킨 후, 이미지 분류 실험을 수행한 결과를 보이고 있다. 그림에서 볼 수 있듯이, Affinity propagation 을 통하여 획득한 트레이닝 집합으로부터

학습된 SVM 분류자(SVM-AP)가 K-Means 및 K-Medoids 를 통하여 학습된 SVM 보다 평균적으로 16.5%, 22.4%의 높은 precision 과 recall 을 보이고 있다.



(그림 4) Bayes 를 이용한 분류 결과의 precision 비교



(그림 5) Bayes 를 이용한 분류 결과의 recall 비교

(그림 4)와 (그림 5)는 각각의 클러스터링 알고리즘을 이용하여 획득한 트레이닝 집합을 이용하여 Bayes 분류자를 학습한 후, 이미지 분류 실험을 수행한 결과를 보이고 있다. 그림에서 볼 수 있듯이, Affinity propagation 을 통하여 획득한 트레이닝 집합으로부터 학습된 Bayes 분류자(Bayes-AP)가 K-Means 및 K-Medoids 를 통하여 학습된 분류자 보다 평균적으로 7.3%, 7.6%의 높은 precision 과 recall 을 보이고 있다.

3.3 대표 클러스터 이미지

본 절에서는 각 클러스터링 알고리즘을 이용하여 획득된 ‘Tiger’ 태그에 해당하는 트레이닝 이미지 집합의 예와 트레이닝 이미지 중 적합하지 않은 이미지의 예를 살펴본다.

- K-Means



(그림 6) K-Means 트레이닝 집합의 예

(그림 6)은 K-Means 알고리즘을 통하여 획득한 트레이닝 집합의 예이다. 좌측은 적합한 트레이닝 이미지, 우측은 적합하지 않은 트레이닝 이미지의 일부를 보여주고 있다. K-Means 트레이닝 집합에는 25.5%의 부적합 이미지가 포함되었다.

• K-Medoids



(그림 7) K-Medoids 트레이닝 집합의 예

(그림 7)은 K-Medoids 알고리즘을 통하여 획득한 트레이닝 집합의 예이다. K-Medoids 트레이닝 집합에는 24%의 부적합 이미지가 포함되었다.

• Affinity propagation



(그림 8) Affinity propagation 트레이닝 집합의 예

(그림 8)은 affinity propagation 알고리즘을 통하여 획득한 트레이닝 집합의 예이다. Affinity propagation 트레이닝 집합에는 8.9%의 부적합 이미지가 포함되었다.

4. 결론 및 향후 연구

본 논문에서는 데이터 군집화를 위한 클러스터링 알고리즘들 중 K-Means, K-Medoids, affinity propagation을 활용하여 소셜 이미지 집합으로부터 트레이닝 집합을 획득하기 위한 방법들을 살펴보았다. 또한, 각 알고리즘으로부터 획득한 트레이닝 집합을 이용하여 소셜 이미지를 분류한 결과를 비교하였다. 실험을 통하여 affinity propagation 기반의 클러스터링을 이용한 트레이닝 집합이 특정 태그에 대하여 보다 적합한 이미지 집합을 구축하고 있음을 확인할 수 있었다. 그러나 본 논문에서 사용한 트레이닝 집합 획득 기법은 구축된 클러스터 중 크기가 가장 큰 k 개의 클러스터를 우선적으로 선택하였기 때문에, 특정 클러스터에 태그에 매우 적합한 이미지들이 포함되어 있다 하더라도 개체수가 적을 경우 선택될 수 없다는 한계점을 가지고 있다.

향후 시각 특징정보뿐만 아니라 이미지에 할당된 연관 태그와의 관련성을 고려하여 보다 적합한 트레

이닝 집합을 추출하기 위한 연구를 수행할 예정이다. 또한, 이미지에 할당되어 있는 태그를 이미지 내의 영역으로 매핑하기 위한 ‘영역기반 태깅’을 위하여, 특정 이미지와 태그가 주어졌을 경우 해당 이미지를 고려하여 태그의 컨텍스트를 판별하고 이에 기반하여 트레이닝 집합을 획득하기 위한 연구도 수행할 예정이다.

5. Acknowledgement

본 연구는 정보통신산업진흥원의 IT/SW 창의연구 과정의 연구결과로 지식경제부와 마이크로소프트에 의해 지원된 과제로 수행되었음 (NIPA-2010-C1810-1002-0012)

참고문헌

[1] Flickr, <http://www.flickr.com>
 [2] 홍현기, 정진우, 이동호, "이미지 태그 간의 의미적 관계에 기반한 효율적인 태그 랭킹 기법", 데이터베이스 연구, 제 26 권 제 3 호, pp.97-117, 2010. 12
 [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] Xu, H., Wang, J., Hua, X.-S., Li, S., 2009. Tag refinement by regularized LDA, ACM Multimedia' 09, pp. 573-576
 [7] Liu, D., Hua, X.-S., Wang, M., Zhang, H.-J., 2010. Retagging social images based on visual and semantic consistency, Proceedings of the 19th international conference on World wide web. ACM, Raleigh, North Carolina, USA.
 [8] 조수선, "의미적 연관태그와 이미지 내용정보를 이용한 웹 이미지 분류", 한국 인터넷 정보학회 11 권 3 호, pp.15 - 24
 [9] Xiaobai Liu et al., "Label to region by bi-layer sparsity priors", Proceedings of the 17th ACM international conference on Multimedia, pp.115-124, 2009
 [10] Hao Xu et al., "Image Search by Concept Map", SIGIR 2010, pp. 275 -
 [11] Jin-Woo Jeong et al., "Automatic Extraction of Semantic Relationships from Images Using Ontologies and SVM Classifiers", International Conference on Multimedia Contents Analysis and Mining 2007, Jun, China
 [12] Kyung-Wook Park et al., "OLYBIA : Ontology-based Automatic Image Annotation System using Semantic Inference Rules", The 2007 Database System For Advanced Application (DASFAA), APRIL 5-9, 2007, Bangkok, Thailand