Increase the Image Search Results by Using Flickr Tags

ShanBin Chan† Shin'ichi SATOH‡ and Hayato YAMANA†

†School of Fundamental Science and Engineering, Waseda University 3–4–1 Okubo, Shinjuku-ku, Tokyo, 169–8555, Japan

‡National Institute of Informatics 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, 101-8430, Japan

E-mail: †{chrisjan, yamana}@yama.info.waseda.ac.jp, ‡satoh@nii.ac.jp

Abstract Flickr is the largest online photo album in the world. Google crawls the contents on the Flickr, and Yahoo! also indexes the data on Flickr. Most people upload images to the Flickr with writing titles, descriptions, and tags, but some people upload images without adding any descriptions onto the images. These kinds of non-description images are not searchable by the users, because the search engine can only index the text contents. Although, human beings can easily indicate the objects or the meaning of an image, the search engine cannot identify the objects or concepts of a non-description image. The amounts of non-description images on Flickr are very huge. For example, the images titled with camera default filename (Ex: DSC_XXXX.jpg or IMG_XXXX.jpg) are over 300 million. If these non-description images can be searched, the image search database will be increased. This research uses the image recognition technology to predict the contents on the non-description images. We collect the sample images on ImageNet and use the concept hierarchical structure provided by WordNet to train the concept database. After training the concept database, we use the concept database and SVM methods to annotate the concepts for non-description images and add the annotated results onto the non-description images by using tagging system on Flickr. After the image search index updated, these non-description images will become searchable on the Yahoo! image search system.

Keyword Flickr, image recognition, SVM, image search, object extraction

1. Introduction

Accompanying by the cost down of digital camera, almost every people have a digital camera. Moreover, accompanying with the increase of network bandwidth, photo-sharing platforms become more and more popular. People can share their favorite photos with friends by using blog or online album services. Flickr is the largest online photo album in the world. Google crawls the contents on the Flickr, and Yahoo! also indexes the data on Flickr. Many image search results are contributed by Flickr. There are many studies about tag ranking[6], image hierarchical structure[1], and user behavior by using Flickr image data. But there are no researches talking about increasing the image search results by using Flickr image data. Our research will use the image recognition technology and use the Flickr APIs to increase the image search results on image search engines.

We found that there are many images are without title, descriptions, and tags on Flickr. And the title of many images is not meaningful, such as IMG_XXXX.jpg or DSC_XXXX.jpg. It is caused by that some people upload images without editing the title, descriptions, and tags. So we define these kinds of images as "non-description images". We use the "IMG" and "DSC" as the search keyword to search on the Flickr. There are over 300 million results, and these images are not meaningful for search engines. Search engine can only index the text contents, such as title, descriptions, and tags. Although, human beings can easily indicate the objects or the meaning of an image, but search engines cannot identify the objects or concepts of a non-description image. These kinds of non-description image are not searchable on search engines. If these non-description images can be searched, the image search database will be increased.

Our research uses the image recognition technology to predict the contents on the non-description images. We collect the sample images on ImageNet[9] and use the concept hierarchical structure provided by WordNet[18] to choose the positive and negative images for concept database training. After training the concept database, we use the concept database and SVM methods to annotate the concepts for non-description images and add the annotated results onto the non-description images by using Flickr tagging APIs. After the image search index updated, these non-description images will become searchable on the image search engines. Because Flickr is a part of Yahoo! services, Flickr provides contents to Yahoo! image search engine directly. Yahoo! image search engine indexes the contents from Flickr and updates the image search index in a specific period time. So our research chooses the Yahoo! image search engine as the candidate. After adding tags on non-description images, we monitor the image search

results on Yahoo! image search engine.

2. Methods

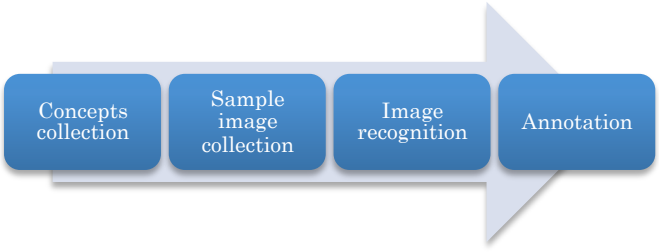

Fig 1. Research Procedure

Our research procedure is as Fig 1. Our methods are using the image recognition technology to extract the objects from non-description images, and recognize the objects, then annotate some keywords onto the non-description images. After annotating keywords, we use the API supported by Flickr to add these keywords as tags onto non-description images. When search engines index these tags on non-description images, these non-description images will become searchable.

Before image recognition, we must train a concept database for recognizing. First, we collect concepts from the photos on Flickr by using Flickr API. We do not crawl all the images on the Flickr, because the amounts of images are too huge. We use the "flickr.photos.getRecent" API to retrieve the recently upload photos randomly. Then we retrieve the tags of recently upload photos and we use the POS (part-of-speech) application[21] to tag the retrieved tags. And we store the image tags, counts, and POS tags to the database. We have already collected 176,303 concepts from July 2009. We will only use the noun as the concepts for sample images retrieval. The amounts of noun concepts are 21,309 from July 2009.

Second, we use these noun concepts as the queries and search the WordNet ID by using WordNet API. And then we use the WordNet ID to gather the sample image list from ImageNet. Then we retrieve the sample images from the sample image list. For each concept, we retrieve 200 images as sample images. And we delete the sample image whose image size is under 10Kbytes.

Third, we use the SVM method to train the concept database by using collected sample images. And then we use the concept database to annotate the concepts on the non-description images. Fig 2 is the procedure of the image recognition.

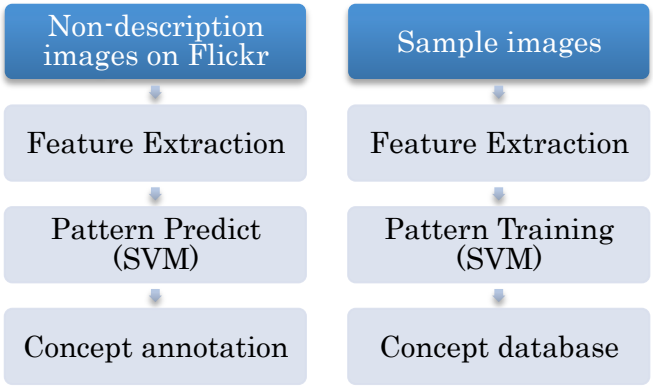

Fig 2. Image Recognition Procedure

We use the collected images as the sample images and extract the features on each sample image.[3][4] Then we implement the pattern training by using SVM. Before pattern training, we also build a hierarchical structure of collected concepts by using WordNet ID API. Fig 3 is a part of concept hierarchical structure. We randomly pick a concept from another sub tree as the negative sample images for SVM pattern training. For example, when we train the concept "house", we pick the "car" as negative sample images. After building the concept database, we extract the features on non-description images. Then we use the SVM to implement the pattern predicting and annotate concepts onto non-description images.

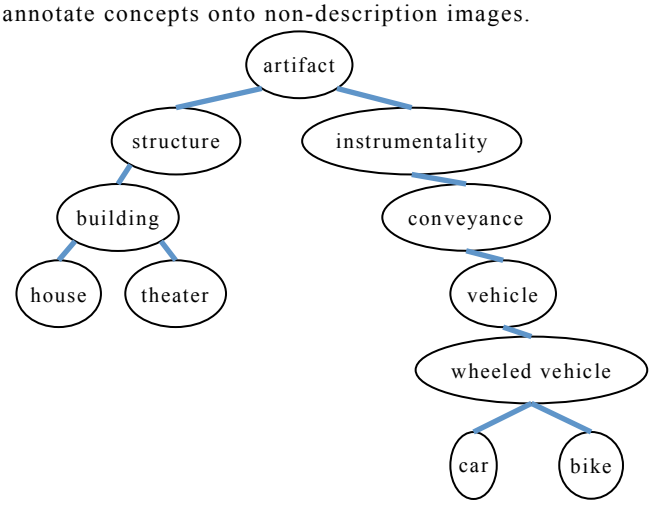

Fig 3. Apart of concept hierarchical structure

Finally, we use the Flickr API to add the annotate results onto the non-description images on Flickr. Fig 4 is the procedure of adding tags on non-description images.

Fig 4. Concept annotation Procedure

Because Flickr is a part service of Yahoo!, we need to get the authentication from Yahoo!. Then we can get the write access privilege on the non-description images. We also build a database to store the accessed logs. We store the date, time, photo urls, user id, photo id, and annotated tags into mysql database. When some users complain with this kind of concepts annotation on the their images, users can apply there user id or photo urls, then we can delete the annotated tags by executing a deleting script according to the access logs.

3. Experiments

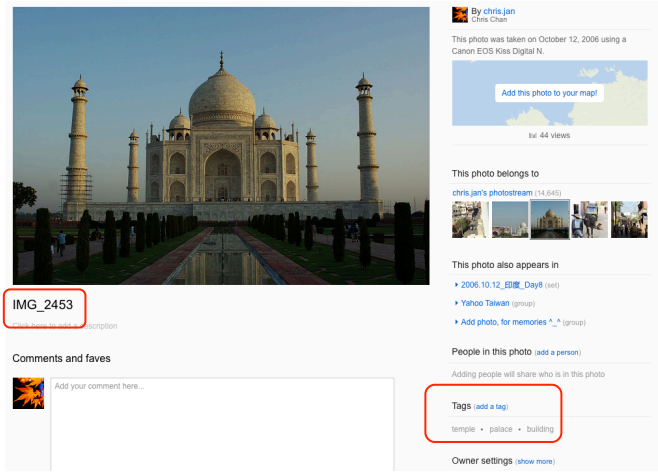

Fig 5. Sample of a non-description image

We take the "Taj Mahal" images for experimentation. Fig 5 is a non-description image of Taj Mahal. Human beings can easily identify that is an image of Taj Mahal. But search engines only can identify it is an image named "IMG_2453". If we input "Taj Mahal" as the query in the image search engine, this titled "IMG_2453" Taj Mahal image will not appear in the image search results.

According to our methods, this non-description image will be annotated as building, palace, and temple. Then we use the Flickr tagging API, we add three tags "palace, building, and temple" onto this non-description image. After search engines update the image results index, we try to use the "temple" as the query and set the search filter to the "flickr.com" and only search the images in the specific user album. This non-description image becomes searchable on Yahoo! image search shown as Fig 6.

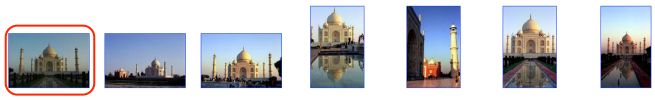

Fig 6. Search results of "temple" on Yahoo! image search

4. Conclusions

Our methods actually can increase the image search results. But the amounts of non-description images are really huge, and the trained concept database is also huge. We already collected 170 thousand concepts from Flickr and there are still over 300 million non-description images on Flickr. And these kinds of non-description images are still increasing. In this research we use only 1 pc to do the image recognition, and the image recognition process really takes a lot of time. If we try to annotate concepts on all 300 million non-description images, it needs over thousand years to complete the task.

Our future work is to enhance the performance of this annotate system by using decision tree algorithms[11] and distribution system (Ex: Hadoop). We are trying to implement the mapreduce algorithm[10] onto our works. Mapreduce procedure is shown as Fig 7.[14] We will divide the training job into mapper work and the reducer job will handle the image recognition process.

Another future work is to rank the annotate results. For one non-description image, there are many concepts in the image. When the collected concepts become larger, the annotate results will also be larger. But what is the main concept of a non-description image, we need a good tag ranking algorithm to identify the main idea of an image.

The sample image lists on ImageNet do not contain the landscape images. But many images on Flickr are about land scape images. We also try to build a sample image collection procedure that can gather landscape sample images.

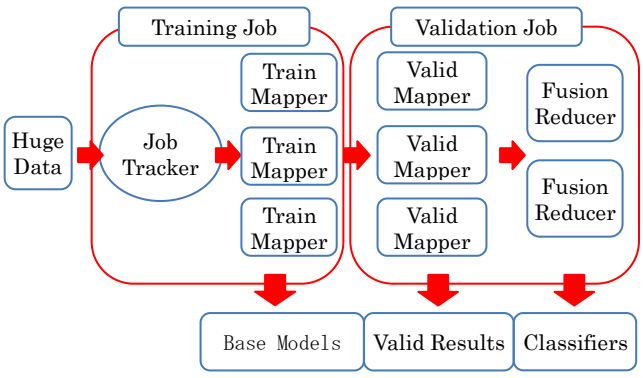

Fig 7. MapReduce framework

Acknowledgements

This research was supported by grant from JST “Multimedia Web Analysis Framework towards Development of Social Analysis Software,” Grant-in-Aid for Scientific Research(No. 18049068), and Ambient SoC Global COE Program of Waseda University” of the Ministry of Education, Culture, Sports, Science andTechnology, Japan.

References

[1] 秋間雄太，川久保秀敏，柳井啟司，“Folksonomy による階層構造画像データベースの構築”, MIRU 2010.

[2] Christopher Moretti, Karsten Steinhaeuser, Douglas Thain, and Nitesh V. Chawla, “Scaling Up Classifiers to Cloud Computers”, ICDM 2008

[3] David G. Lowe, “Object Recognition from Local Scale-Invariant Features”, Computer Vision 1999

[4] David G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints”, Computer Vision 2004

[5] Edward Y. Chang, Kaihua Zhu, Hao Wang, and Hongjie Bai, “PSVM: Parallelizing Support Vector Machines on Distributed Computers”, UCSB 2005

[6] Grigory Begelman, Philipp Keller, and Frank Smadja, “Automated Tag Clustering: Improving search and exploration in the tag space”, WWW 2006

[7] Hangjun Ye and Guangyou Xu, “Fast Search in Large-Scale Image Database Using Vector Quatization”, CIVR 2003, pp. 477-487

[8] Hao Zhang, Alexander C. Berg, Michael Maire, and Jitendra Malik, “SVM-KNN: Discriminative Nearest Neighbor Classification for Visual Category Recognition”, CVPR 2006

[9] ImageNet, http://www.image-net.org/

[10] James Z. Wang, Jia Li, and Gio Wiederhold, “SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries”, IEEE Trans. on Pattern Analysis and Machine Intelligence 2001, Vol.23 No.9

[11] Jeffrey Dean and Sanjay Ghemawat, “MapReduce: Simplified Data Processing on Large Clusters”, ACM 2008, Vol.51, pp. 107-113

[12] Jerry Ye, Jyh-herng Chow, Jiang Chen, and Zhaohui Zheng, “Stochastic Gradient Boosted Distributed Decision Trees”, CIKM 2009

[13] Jia Li and James Z. Wang, “Automatic Linguistic Indexing of Pictures by a Statistical Modeling Approach”, IEEE Transactions on Pattern Analysis and Machine Intelligence 2003, Vol.25 No.9

[14] Matei Zaharia, Andy Konwinski, Anthony D. Joseph, Randy Katz, and Ion Stoica, “Improving MapReduce Performance in Heterogeneous Environments”, OSDI 2008

[15] Rong Yan, Marc-Olivier Fleury, Michele Merler, Apostol Natsev, and John R. Smith, “Large-Scale Multimedia Semantic Concept Modeling using Robust Subspace Bagging and MapReduce”, LS-MMRM 2009

[16] Sergey Brin, “Near Neighbor Search in Large Metric Spaces”, VLDB 1995

[17] Tin Kam Ho, “The Random Subspace Method for Construction Decision Forests”, IEEE Trans. on Pattern Analysis and Machine Intelligence 1998, Vol.20 No.8

[18] Trevor Hastie and Robert Tibshirani, “Discriminant Adaptive Nearest Neighbor Classification”, IEEE Trans. on Pattern Analysis and Machine Intelligence 1996, Vol.18 No.6

[19] WordNet, http://wordnet.princeton.edu/

[20] Xiance Si and Maosong Sun, “Tag-LDA for Scalable Real-time Tag Recommendation”, J. of Information & Computational Science 2009, pp. 23-31

[21] Yoshimasa Tsuruoka and Jun'ichi Tsujii, “Bidirectional Inference with the Easiest-First Strategy for Tagging Sequence Data”, Proceedings of HLT/EMNLP 2005, pp. 467-474