Chinese Journal of Electronics
Vol.17, No.3, July 2008

Pseudo Relevance Feedback for Searech-Based Image Annotation*

ZHAO Yufeng, ZHAO Yao and ZHU Zhenfeng

(Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China)

Abstract — Search-based image annotation (SBIA) is a promising way of Automatic image annotation (AIA). In this paper, Pseudo relevance feedback (PRF) is employed to SBIA (referred as PRF-SBIA) for effectively promoting the annotation performance. Given an un-annotated image as the query, PRF is triggered out for collecting more relevant images from the annotated images. With these relevant images, PRF-based probability model is built up to characterize the hidden relation between the visual content of images and the textual keywords. Furthermore, based on the text-based retrieval technique, a regularized factor with respect to the reliability of keywords is proposed to re-rank the annotation list in each round of PRF. The experimental results reveal that high annotation accuracy can be achieved via PRF-SBIA.

Key words — Search-based image annotation, PRF (Pseudo relevance feedback), PRF-based probability model, Regularized factor.

I. Introduction

Automatic image annotation (AIA), which is to explore some keywords to describe the content of image, has been the key issue for the further development of Content-based image retrieval (CBIR). Many algorithms have been proposed and testified that AIA can greatly improve the performance of CBIR. The Translation model (TM) was the first milestone of AIA, which was treated as a translation process from a set of region tokens to a set of concepts[1]. Then, the Cross-media relevance model (CMRM)[2] imported the relevance language models into AIA. Based on CMRM, Continuous space relevance model (CRM)[3] and Multiple Bernoulli relevance model (MBRM)[4] were proposed to improve the annotation performance. Recently, as a data-driven approach, Search-based image annotation (SBIA) turned up to be effective to enhance the performance of AIA. Given an annotated keyword and an un-annotated image as the query, Wang et al apply the Search result cluster (SRC) algorithm to build a three-layer annotation model[5]. Later, it was improved with Ref.[6] to annotate the web personal images.

One merit of SBIA is at semantic level since the final annotation list is mined from the textual keywords of the relevant images. Another one is scalable for a larger image database. However, since too much irrelevant images are brought into the relevant images, the annotation list is improper or incomplete. As a result, the annotation precision is unsatisfactory in Refs.[5, 6]. That is, the searched relevant images are a common and crucial part in SBIA. To address on this aspect, a novel SBIA scheme based on pseudo relevance feedback (PRF-SBIA) is proposed to improve the annotation performance. On the assumption that most top-ranked searched images are relevant, PRF can serve as an effective means to mine more relevant images without user's intervention. Firstly, given an un-annotated image as the query, PRF can iteratively search and thus accumulate as more relevant images as possible. Secondly, with these relevant images, PRF-based probability model is constructed to characterize the hidden relation between the content of images and the textual keywords. Thirdly, according to the text-based retrieval technology, a regularized factor with respect to the reliability of keywords is proposed to re-rank the annotation list. Compared with the previous SBIA methods, the first merit of PRF-SBIA is to mine more relevant images for annotating with less computational cost involved. Another one is that the textual property of keywords is taken into consideration to improve the reliability of annotation.

The rest of this paper is organized as follows. Section II introduces PRF-based probability model and its theory basis. The correspondences of the model will be dealt with in Section III. In Section IV, the performance evaluation of the proposed method is exhibited. At last, we conclude the paper.

II. PRF-Based Probability Model

Given an un-annotated image I_q, the aim of AIA is to find several keywords w^* that maximize the conditional probability distribution $p(w|I_q)$. We have:

$$w^* = \arg\max_{w} p(w\|I_q) \tag{1}$$

Then, the key problem is how to estimate $p(w|I_q)$. On the assumption that an un-annotated image can be annotated by the keywords of its relevant images, SBIA is proposed from

*Manuscript Received Aug. 2007; Accepted Feb. 2008. This work is supported in part by the National Natural Science Foundation of China (No.60602030, No.90604032), the National 863 Program of China (No.2007AA01Z175), PCSIRT (No.IRT0707) and Research Foundation of BJTU (No.2005SM013, No.2005SZ005).

the viewpoint of search and mining process. Firstly, given an un-annotated image I_q as the query, some relevant images are searched from the annotated images. Then, a few keywords w^* are mined from the annotated keywords of these relevant images for best representing the concept of the query. The procedure is reformulated as:

$$\begin{aligned} w^* =& \arg\max_W p(w|I_q) \\ \approx& \arg\max_W p(w|\Psi_q)p(\Psi_q|I_q) \\ \approx& \arg\max_W \textstyle\sum_{I_r\in\Psi_q} p(w|I_r)\cdot p(I_r|I_q) \end{aligned} \tag{2}$$

where Ψ_q is a set of relevant images to the query I_q. $p(w|I_r)$ is the probability of the keyword w annotating the relevant image I_r. $p(I_r|I_q)$ simulates the search process and denotes the similarity between two images. Obviously, if the relevant image set Ψ_q is more accurate, the annotation performance will be promoted greatly. Thus, PRF-SBIA is proposed to make sure that the searched results for annotation are as precise as possible with less computational cost.

As a self-supervised learning manner, the basic idea of PRF is to hunt automatically more representative query from the top-ranked images for next round of search. As a result, some images missed in the previous round can be searched so as to improve the overall search performance. Hence, by iteratively searching and progressively accumulating more relevant images, PRF-SBIA can boost the search quality and further mine accurately semantic annotation. Moreover, with the searched relevant images, PRF-based probability model is developed to represent the hidden association between the content of images and the textual keywords. In theround of PRF, let $p(w|I_r^{(i)})$ be the probability of the keywordannotating the relevant image $I_r^{(i)}$, and $p(I_r^{(i)}|I_q^{(i)})$ be the similar probability between the relevant image $I_r^{(i)}$ and the query $I_q^{(i)}$. Then, PRF-based probability model is defined by:

$$\begin{aligned} p(w|I_q) =& \textstyle\sum_i p_i(w|I_q) \\ \approx& \textstyle\sum_i p(w|\Psi_q^{(i)})p(\Psi_q^{(i)}|I_q) \\ \approx& \textstyle\sum_i\sum_{I_r^{(i)}\in\Psi_q^{(i)}} p(w|I_r^{(i)})\cdot p(I_r^{(i)}|I_q^{(i)}) \end{aligned} \tag{3}$$

where $\Psi_q^{(i)}$ is the set of the searched relevant images in theround of PRF. The query $I_q^{(i)}$ is initiated as the un-annotated image $I_q^{(0)} = I_q$. In particular, the relevant images, which appear repeatedly in some rounds of PRF, reflect the fact that they are more significant and scatter around the query in feature space. For this, the conditional probabilities of repeated images are summed over the rounds of PRF to highlight them. In addition, each round of PRF is treated equally in PRF-based probability model.

Generally, the annotations can be obtained according to the conditional probability $p(w|I_q)$, which is adopted by many previous annotation models. However, based on the research of text-based retrieval, the keywords of the relevant images do not take on the consistent annotation reliability. Thus, in the i-th round of PRF, the regularized factor denoted by $\lambda_i(w)$ is recommended to deliver the reliability of the keywordto be annotated. Then, a refined PRF-based probability model is defined to re-rank the annotation list:

$$\begin{aligned} p(w|I_q) =& \textstyle\sum_i \lambda_i(w)\cdot p_i(w|I_q) \\ \approx& \textstyle\sum_i \lambda_i(w)\cdot p(w|\Psi_q^{(i)})\cdot p(\Psi_q^{(i)}|I_q) \\ \approx& \textstyle\sum_i \lambda_i(w)\cdot \sum_{I_r^{(i)}\in\Psi_q^{(i)}} p(w|I_r^{(i)})\cdot p(I_r^{(i)}|I_q^{(i)}) \end{aligned} \tag{4}$$

The key issue is to estimate the correspondences of the refined PRF-based probability model effectively after k-NN relevant images are selected to be searched at each round of PRF.

III. Correspondences of PRF-Based Probability Model

For the refined PRF-based probability model, it is to learn accurately the conditional probability distribution $p_i(w|I_q)$ and the regularized factor $\lambda_i(w)$.

1. The conditional probability distribution

In the i-th round of PRF, the conditional probability distribution $p_i(w|I_q)$ is determined by $p(w|I_r^{(i)})$ and $p(I_r^{(i)}|I_q^{(i)})$, which are evaluated via k-NN searched relevant images.

Firstly, the term $p(w|I_r^{(i)})$ denotes the probability of the keywordannotating the relevant image $I_r^{(i)}$. To simplify the computation, $p(w|I_r^{(i)})$ is defined as:

$$p(w|I_r^{(i)}) = \begin{cases} 1, & \text{if the keyword } w \text{ appearing in image } I_r^{(i)} \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

Secondly, the term $p(I_r^{(i)}|I_q^{(i)})$ denotes the similarity between the relevant image $I_r^{(i)}$ and the query $I_q^{(i)}$. Based on the similarity measurement of images, it is given by:

$$p(I_r^{(i)}|I_q^{(i)}) = \exp(-D(I_r^{(i)}, I_q^{(i)})) \tag{6}$$

where $D(I_r^{(i)}, I_q^{(i)})$ is the Euclidian distance between the relevant image $I_r^{(i)}$ and the query $I_q^{(i)}$.

2. The regularized factor

In this paper, the regularized factor $\lambda_i(w)$ is proposed to deliver the reliability of the keywordto be annotated. To measure the regularized factor $\lambda_i(w)$, $tf-idf$ weighted scheme in text-based image retrieval is concerned on. Especially, a keyword can be taken as a term, and all keywords appeared in each round of PRF can be deemed as a document. Let tf_i be term frequency, which denotes the times of the keywordappearing in one round of PRF, idf be inverse document frequency of the keyword w in all rounds of PRF. Then, we have:

$$\lambda_i(w) = tf_i \times idf = tf_i/\ln(n+1) \tag{7}$$

where n is the total rounds of PRF containing w. The regularized factor $\lambda_i(w)$ ensures that an available keyword is assigned an impartial weight even if the appearing times are less. The regularized factor is used to re-rank the initial annotation list in each round of PRF so as to boost the reliability of annotation.

3. Query refining

Another important issue is to find out automatically more representative query from some top-ranked images, which is named by query refining. For the i-th round of PRF, we define the new query as mean of relevant images via k-NN search.

$$I_q^{(i+1)} = \frac{1}{k}\sum_{r=1}^{k} I_r^{(i)} \tag{8}$$

where $I_q^{(i+1)}$ is the refined query for the i+1-th round of search, which is initiated as the un-annotated image I_q. Finally, according to the Eq.(4), the most five top-ranked keywords are selected for annotating the un-annotated image I_q.

IV. Experimental Results and Analysis

We test the proposed algorithm on the Corel dataset from Barnard *et al.*[1], which is extensively used as basic comparative data for recent research works in AIA. The experimental data set comprises 5,000 images, in which 4,500 images are used as training set and the remaining 500 images as testing set. In HSV color space, 36-D color histogram is extracted for representing the global visual feature of each image. Each image is annotated with 1 to 5 keywords, and totally 374 keywords are used for the annotated images. In addition, we also measure the image annotation performance by using the annotation recall and precision as $recall = B/C$ and $precision = B/A$, where A is the number of images automatically annotated with the given keyword in the top 5 of the returned word list; B is the number of images correctly annotated with that keyword in the top 5 returned word list; and C is the number of images having that keyword in ground truth annotation[4]. In addition, the parameters in the experiments are set to be $k = 20$ and $i \leq 3$.

1. Comparison with other annotation models

To show the effectiveness of PRF-SBIA, we compare it with other related annotation models: MBRM[4], and SBIA[5]. Table 1 illustrates average precision and recall of each keyword on the experimental dataset. The results of MBRM are obtained from Ref.[4]. Since the experimental dataset of SBIA is different from ours, it is implemented on our dataset.

Table 1. Performance comparison of the related models

Models	MBRM	SBIA	PRF-SBIA
#words with $recall > 0$	122	153	188
Results on all 374 keywords			
Average per-word recall	0.25	0.33	0.39
Average per-word precision	0.24	0.27	0.31

Table 1 represents that the performance of search-based annotation methods, i.e. SBIA and PRF-SBIA, is better than MBRM. Compared with SBIA, PRF-SBIA exhibits a gain of 5% in recall and 3% in precision. The number of words with positive recalls increases by 35%. It is believed that two reasons lead to the improvement of annotation. One is that more relevant images are mined in PRF-SBIA than that of SBIA. Another one is to exploit the regularized factor for re-ranking the annotation list.

2. Effectiveness of k-NN

The numberof searched relevant images is a crucial parameter in the refined PRF-based probability model. To show its effectiveness, Fig.1 reports the average precision and recall for 374 keywords with various.

As demonstrated in Fig.1, we can find that the best result is achieved when k is set to be 20. The benefit will keep being improved until k reaches 20; on the contrary, the benefit will gradually reduce when it exceeds 20. For this fact, the main reason is that the assumption of PRF, which corresponds to k most top-ranked images taken as relevant images, is not well satisfied. However, it can be better satisfied when k is set to be 20 in each round of PRF. Another reason is that the total number of relevant images is no more than 90 in the annotated images. Hence, when k value enlarges, more irrelevant images will be involved in the relevant images. Based on the optimal annotation performance, k is set to be 20 in our experiment.

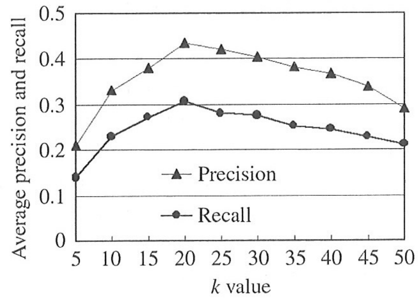


Fig. 1. Effectiveness of various k

3. Influence of the rounds of PRF

The iterative number i of PRF plays also an important role on the searching efficiency. As shown in Fig.2, the annotation performance can nearly reach a stable state with $i = 3$, and no obvious improvement is achieved with $i > 3$. However, with increasing the round i of PRF, the computational cost will go up greatly. Thus, the round i of PRF is set to be 3 with lower computation complexity.

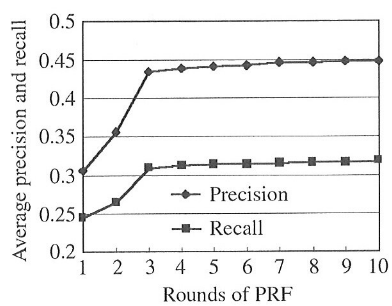


Fig. 2. Influence of the iterative number of PRF

Table 2. Performance for 10 most frequent keywords

Keywords	Without λ		With λ	
	Precision	Recall	Precision	Recall
nest	0.3581	0.2810	0.5714	0.8714
tiger	0.4308	0.6000	0.5385	0.9000
stone	0.6000	0.8000	0.9231	0.8714
water	0.2901	0.7328	0.3769	0.8448
plane	0.6553	0.4800	0.9474	0.7200
window	0.4182	0.5385	0.6667	0.4500
garden	0.4818	0.6000	0.6667	0.6000
Average	**0.3234**	**0.4032**	**0.4681**	**0.5258**

4. Evaluation of regularized factor

To illustrate the effectiveness of regularized factor $\lambda_i(w)$ for annotation, the performance evaluation for 10 most frequent keywords is given in Table 2.

For each keyword in Table 2, we can find that the annotation performance is improved greatly when the regularized factor $\lambda_i(w)$ is taken into the PRF-based probability model. The reason is that the regularized factor can leverage the weight of keywords contained in the searched relevant images. In particular, it can enhance the weight of the useful keyword with little appearing times.

V. Conclusions

In this paper, we have developed a novel annotation scheme referred as PRF-SBIA. The experimental results show that PRF-SBIA explores more relevant images so as to improve the annotation performance. Furthermore, the regularized factor is proposed to build the refined PRF-based probability model and to further promote the annotation reliability. PRF-SBIA is scalable since the annotated images can be enlarged arbitrarily. Thus, PRF-SBIA is a promising way to fulfill the task of AIA with less tedious human activity involved.

References

[1] P. Duygulu, K. Barnard, N. de Freitas, D.A. Forsyth, "Object recognition as machine translation: learning a lexicon for a fixed image vocabulary", *Proc. of IEEE Conference on Computer Vision*, Copenhagon, Denmark, pp.97–112, 2002.

[2] J. Jeon, V. Lavrenko, R. Manmantha, "Automatic image annotation and retrieval using cross-media relevance models", *Proc. of ACM Special Interest Group on Information Retrieval*, Toronto, Canada, pp.119–126, 2003.

[3] R. Manmatha, V. Lavrenko, J. Jeon, "A model for learning the semantics of pictures", *Proc. of IEEE Conference on Neural Information Processing Systems*, Whistler, British Columbia, Canada, 2003.

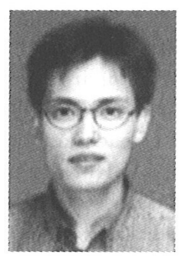
[4] S.L. Feng, R. Manmatha, V. Lavrenko, "Multiple Bernoulli relevance models for image and video annotation", *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, Washington DC, USA, pp.1002–1009, 2004.

[5] X. Wang, L. Zhang, F. Jing, W. Ma, "AnnoSearch: image auto-annotation by search", *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, New York, USA, pp.1483–1490, 2006.

[6] C.H. Wang, F. Jing, L. Zhang, H.J. Zhang, "Scalable search-based image annotation of personal images", *ACM Conference on Multimedia Information Retrieval*, pp.269–277, 2006.

ZHAO Yufeng is Ph.D. in the Institute of Information Science at Beijing Jiaotong University. She has worked in content-based image retrieval, machine learning and pattern recognition. (Email: snowmanzhao@163.com)

ZHAO Yao is a professor in the Institute of Information Science at Beijing Jiaotong University. He has worked in watermarking, image and video coding, content-based image retrieval. (E-mail: yzhao@bjtu.edu.cn)



ZHU Zhenfeng is an instructor in the Institute of Information Science at Beijing Jiaotong University. He has worked in content-based image retrieval, machine learning and pattern recognition. (Email: zhfzhu@bjtu.edu.cn)