영상의 자동 주석: 멀티 큐 통합

신성윤[*] · 안은미[*] · 이양원[*]

[*]군산대학교

Images Automatic Annotation: Multi-cues Integration

Seong-Yoon Shin[*] · Eun-Mi Ahn[*] · Yang-Won Rhee[*]

[*]Kunsan National University

E-mail : syshin{ywrhee}@kunsan.ac.kr, qortnwl0326@nate.com

ABSTRACT

All these images consist a considerable database. What's more, the semantic meanings of images are well presented by the surrounding text and links. But only a small minority of these images have precise assigned keyphrases, and manually assigning keyphrases to existing images is very laborious. Therefore it is highly desirable to automate the keyphrases extraction process. In this paper, we first introduce WWW image annotation methods, based on low level features, page tags, overall word frequency and local word frequency. Then we put forward our method of multi-cues integration image annotation. Also, show multi-cue image annotation method is more superior than other method through an experiment.

키워드

Keyphrases Extraction, Multi-cue Image Annotation,

Ⅰ. Introduction

In the earlier image retrieval systems, images are annotated manually by text descriptors. There are two disadvantages with this approach. The first is a considerable level of human labor is required for manual annotation. The second is the annotation inaccuracy due to the subjectivity of human perception. To overcome these two disadvantages in text-based retrieval systems, content-based image retrieval was introduced[1]. In CBIR, images are indexed by their visual features, such as color, texture, shape. Though many sophisticated algorithms[2][3][4] have been designed to describe color, shape and texture features, these algorithms cannot adequately model image semantics and have many limitations when dealing with broad content images databases.

Ⅱ. Multi-cue Annotation

WWW images locate in structural, networking documents, so the importance of a word can be indicated by its location, frequency. There are two patterns for multi-cues integration annotation.

(1) Linear Integration

In this paper, we use the below multi-linear integration formula Eq. 1 to compute word weight in documents:

$$\underline{\hspace{4cm}} \quad \underline{\hspace{2cm}} \quad \text{(Eq. 1)}$$

The importance of token t in document d consists of three cues' weight :

TAG weight $\underline{\hspace{1.5cm}}$, TF-IDF weight $\underline{\hspace{1.5cm}}$, SR weight $\underline{\hspace{1cm}}$, a, β, γ are influence of each weight. After the weight is computed, some (for instance, 4) are chose to be the annotation keyphrases to the image.

(2) Tactic Integration

Tactic integration use each cue to select keyphrases orderly. We generally apply TAG to narrow the keyphrases range down, then we apply TF-IDF and SR to filter in candidate set. Fig 1 is the flow chart of multi-cues tactic integration.
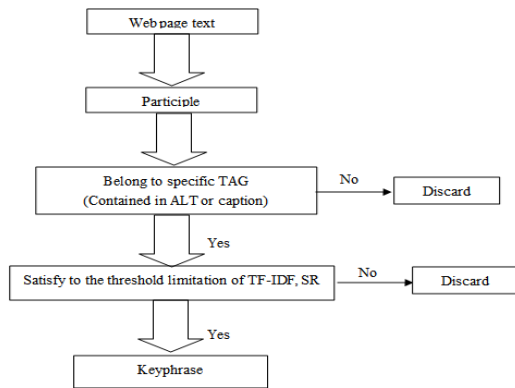
Fig 1. Multi-cues tactic integration flow chart

TF-IDF threshold limitation can select threshold according to experiments, or simply order the tokens and take the first few ones as keyphrases.

Ⅲ. Experiment

The indexer then extracted index entries from the abstract data of these analysed web pages. These index entries are actually keyphrases filtered by Meta tag, TF-IDF tag and SR tag. We selected the accessory Meta tags to the image and adapted participle program to get participles. Together with TF-IDF and SR, this multi-cues integration algorithm (Eq. 1) was then adapted to get keyphrases of the image. We usually take the first several ones with high W(t, d)(Eq. 1) as keyphrases, and the average number of correct keyphrases found in each corpus was recorded and is presented below in Fig. 2.
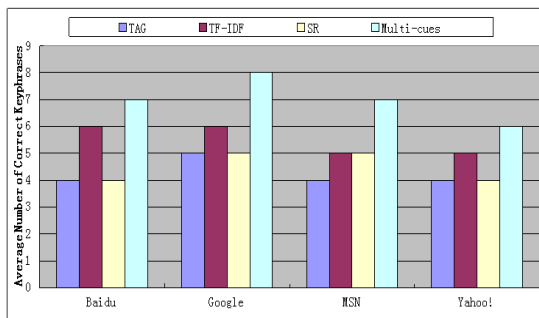

Fig 2. Comparison of TAG, TF-IDF, SR and Multi-cues based Annotation

Fig. 2 shows that the integration of TAG, TF-IDF and SR feature improves the success of the multi-cues integration annotation algorithm by between 25% and 60% in these four test corpora.

Ⅳ. Conclusion

In this paper, we first introduce WWW image annotation methods, based on low level features, page tags, overall word frequency and local word frequency. Then we put forward our method of multi-cues integration image annotation. Also, show multi-cue image annotation method is more superior than other method through an experiment.

The multi-cues integration algorithm shows initial promise as an indicator of semantic keyphrases of the web images. The latent semantic automatic keyphrase extraction that causes the improvement with the usage of multi-cues is expected to be preferable.

참고문헌

[1] Ritendra Datta, Dhiraj Joshi, Jia Li and James Z. Wang, ``Image Retrieval: Ideas, Influences, and Trends of the New Age,'' ACM Computing Surveys, vol. 40, No. 2, pp. 1-77, 2008.

[2] J.Cox, et al. PicHunter: Bayesian relevance feedback for image retrieval [C]// Proceedings of 13th International Conference On Pattern Recognition, Vienna, pp. 361-369, 1996.

[3] Y. Rui, T. S. Huang, S. F. Chang, Image Retrieval: Current Technologies, Promising Directions and Open Issues [J]. Journal of Visual Communication and Image Representation, 10(4), pp. 39-62, 1999.

[4] Theo Gevers and Arnold Smeulders. Pictoseek: Combining color and shape invariant features for image retrieval. IEEE Transactions on image Processing, 9(1):102119, January 2000.