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, qortnw0326@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.

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:

\[ W_{td} = \alpha \cdot W_{tag} + \beta \cdot W_{tf-idf} + \gamma \cdot W_{sr} \]  

(Eq. 1)

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

TAG weight \( \alpha \), TF-IDF weight \( \beta \), SR weight \( \gamma \) 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.
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.

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


