Neurocomputing 204 (2016) 26–32

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

A framework of query expansion for image retrieval based on knowledge base and concept similarity

Yuanfeng He [a,*], Yuanxi Li [b], Jiajia Lei [a], C.H.C Leung [b]

[a] School of Electronic Information and Communications, Huazhong University of Science and Technology, 430074, China
[b] Department of Computer Science, Hong Kong Baptist University, Hong Kong

ARTICLE INFO

Article history:
Received 17 April 2015
Received in revised form
24 November 2015
Accepted 25 November 2015
Available online 9 April 2016

Keywords:
Image retrieval
Knowledge base
Query expansion

ABSTRACT

We study several semantic concept-based query expansion and re-ranking scheme and compare different ontology-based expansion methods in image search and retrieval. To improve the query expansion efficiency and accuracy, we employ the CYC knowledge base to generate the expansion candidate concepts, while filter and rank the expansion results by calculating concept similarities using the Semantic Relatedness Metrics. Using our knowledge-based query expansion in image retrieval, the efficiency and accuracy has been improved.

1. Introduction

The number of Web images is increasing at a rapid rate, and searching them semantically presents a significant challenge. Many raw images are constantly uploaded with little meaningful direct annotation of semantic content, limiting their capacity to be searched and discovered. Unlike in a traditional database, information in an image database is in visual form, which requires more space for storage, is highly unstructured and needs state-of-the-art algorithms to determine its semantic content.

The relative ineffectiveness of information retrieval systems [1–3] is largely caused by the inaccuracy with which a query formed by a few keywords models the actual user information need. Query expansion has proved to improve the effectiveness of information retrieval by automatically adding additional terms to the original query [4].

The typical Automatic Query Expansion [5] can be broken down into the four steps: preprocessing of data source, generation and ranking of candidate expansion features, selection of expansion features and query reformulation.

In [6] an original query is run using conventional information retrieval techniques [7]. Then, related terms are extracted from the top documents that are returned in response to the original query using statistical heuristics. This approach has been shown to be effective on some collections, but results on large collections of web data have been mixed. The work in [8] chooses expansion terms from past user queries directly, rather than using them to construct sets of full text documents from which terms are then selected.

Vast numbers of Web images are continuously added with few meaningful direct annotations of semantic content, limiting their search and discovery. While some Websites encourage tags or keywords to be included manually, such is far from universal and applies to only a small proportion of images on the Web. Research in image search has reflected the dichotomy inherent in the semantic gap [9], and is divided between two main categories: concept-based image retrieval and content-based image retrieval. The former focuses on retrieval by image objects and high-level concepts, while the latter focuses on the low-level visual features of the image. In order to determine image objects, the image often has to be segmented into parts. Common approaches to image segmentation include segmentation by region and segmentation by image objects. Segmentation by region aims to separate image parts into different regions sharing common properties. These methods compute a general similarity between images based on image properties, [10] and common examples of such properties are texture and color where these methods are found to be robust and efficient. Some systems use color, texture, and shape [11] as attributes and apply them for entire image characterization, and some studies include users in a search loop with a relevance feedback mechanism to adapt the search parameters based on user feedback, while various relevance feedback models and ranking methods for Web search have been developed [12].

Recent approaches for calculating image similarity using various image features, such as color similarity measure [13] can

* Corresponding author.
E-mail addresses: heyuanfeng@hust.edu.cn (Y. He), yxli@comp.hkbu.edu.hk (Y. Li), jjlei@hust.edu.cn (J. Lei), clement@comp.hkbu.edu.hk (C.H. Leung).

http://dx.doi.org/10.1016/j.neucom.2015.11.102
0925-2312/

ensure a certain degree of precision for image retrieval within a large image collection. The extracted features could be filled into local (like SIFT, SURF or DAISY) or global (like MPEG-7 color/texture descriptors, Edge Orientation Histogram or GIST) descriptors. Some studies [14] uses the Non-Subsampled Contourlet Transform (NSCT) detector combined with a DAISY descriptor to develop a robust interest point matching algorithm for infrared/visible images. As regards to the segmentation of color images, segmentation by object, on the other hand, is widely regarded as a hard problem, which if successful, will be able to replicate and perform the object recognition function of the human vision system; although progress on this front has been slow, some advances in this direction have nevertheless been made [15]. Makrogiannis et al. [16] proposed a multiresolution image segmentation scheme based on a graph-theoretic approach. Grady and Schwartz [17] treated image segmentation as a linear problem instead of the eigenvector approach to a graph-partitioning problem [18]. Tao et al. [19] developed a robust real-time approach for color image segmentation using the MS segmentation and the normalized cut (Ncut) [18] partitioning methods. Diplaros et al. [20] resorted to a spatially constrained EM algorithm to estimate the model parameters. In [21], semantic annotation of images combined with a region-based image decomposition is used, which aims to extract semantic properties of images based on the spatial distribution of color and texture properties. Such techniques have drawbacks, primarily due to their weak disambiguation and limited robustness in relation to object characterization. However, an advantage of using low-level features is that, unlike high-level concepts, they do not incur any indexing cost as they can be extracted by automatic algorithms. In contrast, direct extraction of high-level semantic content automatically is beyond the capability of current technology. Although there has been some effort in trying to relate low-level features and regions to higher level perception, these tend to be for isolated words, and they also require substantial training samples and statistical considerations [22]. These methods, however, have limited success in determining semantic contents in broad image domains. There are some approaches which exploit surrounding and associated texts in order to correlate and mine these with the content of accompanying images [23]. Text-based retrieval is often limited to the processing of tags, and no attempt is made to extract a thematic description of the picture. Some research focuses on implicit image annotations which involves an implicit, rather than an explicit, indexing scheme and, in consequence, augments the original indexes with additional concepts that are related to the query [24], necessitating the use of some probabilistic weighting schemes.

The presence of particular objects in an image often implies the presence of other objects [25]. In terms U and V, if only U is indexed, then searching for V will not return the image in the result, even though V is present in the image. The application of such inferences will allow the index elements Ti of an image to be automatically expanded according to some probability which will be related to the underlying ontology of the application.

In this proposed query expansion framework, we mainly focus on CYC Knowledge Base [26,27] to generate the candidates of query expansion. The CYC knowledge base (KB) is a formalized representation of a vast quantity of fundamental human knowledge: facts, rules of thumb, and heuristics for reasoning about the objects and events of everyday life.

As for redefining image indexing [25], the most popular way is to simplify the semantic knowledge into the semantic similarity between concepts, WordNet [28], is a large lexical database of English. Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept. It is one of these applications of semantic lexicon for the English language. Besides, its also a general knowledge base and commonsense reasoning engine. The purpose of the work is both to produce a combination dictionary-and-thesaurus that is more intuitively usable, and to support automatic text analysis and artificial intelligence applications.

In this paper, we propose a query expansion for image retrieval system using CYC knowledge base as expansion candidate generator, then employ WordNet as the refining tool to filter and rerank and expanded queries. With the proposed approach, the indexing accuracy and efficiency has been improved. In Section 2, related work has been introduced; Section 3 mainly describe how the system work; the experimental results are given in Section 4 with the conclusion is drawn in the last section.

2. Related works

To find the proper expansion candidates, it is required to measure the relatedness between the original query and the candidate queries. There are two types of relatedness measurement:

2.1. Topological similarity measurement

Four types of approaches are used to calculate topological similarity between ontological concepts or instances:

Edge-based similarity: Edge-based similarity measures are based mainly on counting the number of edges in the graph to get the path between two terms [29–31].

For example, Pekar et al. [31] proposed an approach to measure the taxonomic similarity between a and b by

$$T(a,b)=\frac{\delta(\text{root},c)}{\delta(a,c)+\delta(b,c)+\delta(\text{root},c)} \tag{1}$$

where $\delta(a,b)$ describes the number of edges on the shortest path between a and b.

Node-based similarity: in which the main data sources are the nodes and their properties.

For example, Resnik [32] proposed an approach to measure the concept similarity based on the notion of information content. The information content of a concept (term or word) is the logarithm of the probability of finding the concept in a given corpus.

$$\text{sim}(w_1,w_2)=\max_{c_1\in\text{sen}(w_1),c_2\in\text{sen}(w_2)}[\text{sim}(c_1,c_2)] \tag{2}$$

Pairwise similarity: Combining the semantic similarities of the concepts they represent, measure functional similarity between two instances. For example, the Pairwise Document Similarity considering symmetric similarity measures is defined as follows:

$$\text{sim}(d_i,d_j)=\sum_{t\in V}\omega_{t,d_i}\cdot\omega_{t,d_j} \tag{3}$$

where $\text{sim}(d_i,d_j)$ is the similarity between documents d_i and d_j, and V is the vocabulary set.

Groupwise similarity: Compare with Pairwise Similarity, Groupwise Similarity does not combining the semantic similarities of the concepts they represent, it calculates the similarity directly. For example, Jaccard coefficient [33] measures similarity between finite sample sets, and is defined as the size of the intersection divided by the size of the union of the sample sets:

$$J(A,B)=\frac{|A\bigcap B|}{A\bigcup B} \tag{4}$$

For Topological Similarity Measurement, the similarity calculations are mainly based on the taxonomy or topology, which are reflecting the knowledge of a large number of experts or users who built these taxonomy or topology. These types of relatedness measurement are very comprehensive but relatively static. In this