Chapter 4
An Image Retrieval Model Combining Ontology and Probabilistic Ranking

Lisa Fan
University of Regina, Canada

Botang Li
University of Regina, Canada

ABSTRACT

The demand for image retrieval and browsing online is growing dramatically. There are hundreds of millions of images available on the current World Wide Web. For multimedia documents, the typical keyword-based retrieval methods assume that the user has an exact goal in mind in searching a set of images whereas users normally do not know what they want, or the user faces a repository of images whose domain is less known and content is semantically complicated. In these cases it is difficult to decide what keywords to use for the query. In this chapter, we propose a user-centered image retrieval method based on the current Web, keyword-based annotation structure, and combining ontology guided knowledge representation and probabilistic ranking. A Web application for image retrieval using the proposed approach has been implemented. The model provides a recommendation subsystem to support and assist the user modifying the queries and reducing the user's cognitive load with the searching space. Experimental results show that the image retrieval recall and precision rates are increased and therefore demonstrate the effectiveness of the model.

INTRODUCTION

With the development of the Internet and database techniques, Information Retrieval (IR) has become very popular (Ricardo, B.Y. & Berthier, R.N., 1999). As a powerful form of delivering information, multimedia data are frequently used in many applications. Techniques for effectively dealing with multimedia databases are useful and in demand. During the past decade, online image retrieval has become one of the most popular topics on the Internet. The number of images available in online repositories is growing dramatically. For example, Flickr.com is hosting more than 50 million member-submitted images on their website (Terdiman, D., 2009), and Google claims that they have indexed more than 880 millions images since 2004 (Google, 2009). It is expected that the number of images found in personal collections, publications and archives will continue to grow at an exponential rate. Given the recent explosion of interest in social networking, largely driven by Myspace, Facebook and YouTube, one can expect to see image searching playing a very important role in the future. As a result, the demand for efficient and effective image retrieval, searching, and browsing methods will also increase significantly. Demand will come from users in various domains, including medicine, publishing, architecture, crime prevention, and fashion.

Image retrieval is a human centered task. Images are created by people and are ultimately retrieved and used by people for human related activities. The typical method of image retrieval using mostly by the industry is to create a keyword-based query interface above the media indexing database (Agosti & Smeaton, 1996). There are two major problems in keyword-based image retrieval. The first one is the retrieval quality problem from the search result. The keyword annotation of image documents has low capability to analyze semantic relations among keywords, such as synonym, homonym and antonym. Taking the topics of images as an example, it is nearly impossible to

include all the synonyms of the topic keywords in the annotation for every image. The reality is that if the images are annotated with keywords having same meanings with users input but in different terms, those images are not able to be retrieved by the keyword-based retrieval system. The second problem is that keyword-based search methods always assume that users have the exact searching goal in their minds (Hyvonen, Saarela & Viljanen, 2003). However, in the real world application, the case is that users normally do not know what they want. Most of them only hold a general interest to explore the images, and have a vague knowledge about the domain topic. As a result, a recommendation or a support subsystem, helping users to modify their queries, is needed.

Semantic Web technologies have been expected to improve the quality of information retrieval on the Web (Berners-Lee, Hendler & Lassila, 2001) & (Berners-Lee). In this paper, we proposed a hybrid model which is using a Web Ontology-based reasoning component and combining Bayesian Network model to improve the quality of image retrieval. Our proposed method returns more query keywords as recommendations which are semantically related to the user input keywords so that it can assist the users to explore more relevant images.

RELATED WORK

Image Retrieval is a large and active research area of computer and information science. A summary review of the literature shows an exceptionally active community of researchers in this area. Smeulders et al. reviewed more than 200 research papers prior to 2000 (Smeulders, Worring, Santini, Gupta & Jain, 2000). Rui et al. have summarized more than 100 research papers (Rui, Huang & Chang, 1999). Recently, Datta et al. have surveyed about 300 papers, mostly published between 2000 and 2007 (Datta, Joshi, Li & Wang, 2008).

Related Content

Video Shot Boundary Detection for Video Indexing

Waleed E. Faragand Hussein Abdel-Wahab (2005). *Video Data Management and Information Retrieval
(pp. 193-216).*

www.irma-international.org/chapter/video-shot-boundary-detection-video/30766

Context Window Based Co-occurrence Approach for Improving Feedback Based Query
Expansion in Information Retrieval

Jagendra Singhand Aditi Sharan (2015). *International Journal of Information Retrieval Research (pp. 31-
45).*

www.irma-international.org/article/context-window-based-co-occurrence-approach-for-improving-feedback-based-query-expansion-in-information-retrieval/132490

A Noval Approach for Object Recognition Using Decision Tree Clustering by Incorporating Multi-
Level BPNN Classifiers and Hybrid Texture Features

Upendra Kumar (2024). *International Journal of Information Retrieval Research (pp. 1-31).*

www.irma-international.org/article/a-noval-approach-for-object-recognition-using-decision-tree-clustering-by-incorporating-multi-level-bpnn-classifiers-and-hybrid-texture-features/338394

Knowledge Discovery From Massive Data Streams

Sushil Kumar Narang, Sushil Kumarand Vishal Verma (2018). *Information Retrieval and Management:
Concepts, Methodologies, Tools, and Applications (pp. 1508-1534).*

www.irma-international.org/chapter/knowledge-discovery-from-massive-data-streams/198612

Energy-Efficient Virtualized Scheduling and Load Balancing Algorithm in Cloud Data Centers

J. K. Jeevithaand Athisha G. (2021). *International Journal of Information Retrieval Research (pp. 34-48).*

www.irma-international.org/article/energy-efficient-virtualized-scheduling-and-load-balancing-algorithm-in-cloud-data-centers/280525