Chapter 1

Text-Based Image Retrieval

ABSTRACT

The role of textual keywords for capturing the high-level semantics of an image in HTML document is studied. It is observed that the keywords present in HTML documents can be effectively used for describing the high-level semantics of the images appear in the same document. Techniques for processing HTML documents and Tag Ranking for Image Retrieval (TRIR) is explained for capturing semantic information about the images for retrieval applications. A retrieval system returns a large number of images for a query and hence it is difficult to display the most relevant images in top results. This chapter presents newly developed method for ranking the images in Web documents based on the properties of HTML TAGS in web documents for image retrieval from WWW.

INTRODUCTION

The technological advancements of Internet have increased the population of images, which demands effective retrieval mechanisms. Search engines have become indispensible tools for retrieving relevant images from WWW. In general, there are two flavors of retrieval namely, Text Based Image Retrieval (TBIR) and Content Based Image Retrieval (CBIR). The CBIR systems are suitable only for domain specific applications. It is also difficult to capture the semantics of images using low-level features alone and the semantic gap posed between the query image and database images remains as challenge. The extracted semantic features of the query image may not effectively match with
the database images. Additionally, the cost of time associated for extracting image features among large database is also very high. In a typical CBIR system, the query image is matched with database images and the retrieval set is ranked using a suitable similarity measure. The image retrieval can be broadly classified as Text-Based Image Retrieval and Content-Based Image Retrieval. Below, the functionality of each of these methods are discussed in detail with suitable example.

**Text-Based Image Retrieval**

The Text-Based Image Retrieval is relatively old method, where a text or sequence strings are provided as input query. Say for example, “an elephant with human riding on it”. Thus, the query keyword may provide information about image name, date of adding, deleting, modifying, etc. However, the text based query ignores unexpressed feelings, emotions and conveys same meaning in an unusual way. In addition, the synonyms, hyponyms, homonyms and misspellings of the query text tend to increase the semantic gap. A typical TBIR system is depicted with various blocks for better understanding.

In the above Figure 1, the user submits text based query and is being processed by query processing unit. The stop words are removed and textual keywords are stemmed for making the query text suitable for precise search process. A suitable similarity measure is used for calculating the distance between the query and indexed text for ranking the database text and corresponding image. The indexed text-DB houses all the textual keywords based on an indexing structure (say inverted index) so that only a part of the database is searched. The indexed image-DB has all the corresponding images.

*Figure 1. A typical TBIR system*
Dynamics and Simulation of General Human and Humanoid Motion in Sports
Veljko Potkonjak, Miomir Vukobratovic, Kalman Babkovic and Branislav Borovac
(pp. 998-1022).
www.igi-global.com/chapter/dynamics-simulation-general-human-
humanoid/49432?camid=4v1a

Enhancing Rating Prediction by Discovering and Incorporating Hidden User
Associations and Behaviors
Ligaj Pradhan (2019). International Journal of Multimedia Data Engineering and
Management (pp. 40-59).
www.igi-global.com/article/enhancing-rating-prediction-by-discovering-and-
incorporating-hidden-user-associations-and-behaviors/232181?camid=4v1a