International Journal of Computer Science and Engineering Communications,
Volume.5, Issue.4 (2017): Page.1648-1655
www.ijcsec.com

Query Specific Fusion For Image Retrieval System

K**. Divya**[1]**, R.Rajarathinam**[2]**, R.Latha**[3]
[1]*Research Scholar, St.Peter's University, Chennai.*
[2]*Asst.Prof. & Head., Dept. of Computer Applications, St.Peter's University, Chennai*
[3]*Prof. & Head., Dept. of Computer Applications, St.Peter's University, Chennai*
divyasan29@gmail.com

Abstract: *This paper presents a new majority voting technique combines the two basic modalities of Web images textual and visual features of image in a re-annotation and search based framework. The proposed framework considers each web page as a voter to vote the relatedness of keyword to the web image, the proposed approach is not only pure combination between image low level feature and textual feature but it take into consideration the semantic meaning of each keyword that expected to enhance the retrieval accuracy. The proposed approach is not used only to enhance the retrieval accuracy of web images; but also able to annotate the unlabeled images.*

Keywords: *Web Image Retrieval, Ontology, Data Mining, Image Clustering, Extraction, Ranking.*

1. Introduction

With the advent of the Internet, billions of images are now freely available online [A.Torralba]. The rapid growth in the volume of such images can make the task of finding and accessing image of interest, overwhelming for users. Therefore, some additional processing is needed in order to make such collections searchable in a useful manner. The current Web image retrieval search engines, including Google image search, Lycos and AltaVista photo finder, use text (i.e., surrounding words) to look for images, without considering image content [M.L.Kherfi,Z.Gong]. However, when the surrounding words are ambiguous or even irrelevant to the image, the search based on text only will result in many unwanted result images. In contrast, Content-Based Image Retrieval (CBIR) systems are proposed to utilize only low-level features, such as color, texture and shape, to retrieve similar images. In general, the bottleneck to the efficiency of CBIR is the semantic gap between the high level image interpretations of the users and the low level image features stored in the database for indexing and querying. In other words, there is a difference between what image features can distinguish and perceives from the image [Z.Alemu]. This paper try to narrow the semantic gap problem and enhance the retrieval precision by fusing the two basic modalities of Web images, i.e., Textual context (usually represented by keywords) and visual features for retrieval. This paper presents a new majority voting technique that considers each web page as a voter to vote the relatedness of keyword to the web image, the proposed approach is not only pure combination between image low level feature and textual feature but it take into consideration the semantic meaning of each keyword that expected to enhance the retrieval accuracy. This paper is organized as follows; the related work presented in section 2, while, section 3 presents in details the architecture of the proposed approach, while section 4 concludes this research.

2. Related work

Current research in web image retrieval suggested a joint use existing Textual context and visual features can provide a better retrieval results [N.Xiony,J.Hou]. The simplest approach for this method is based on counting the frequency-of-occurrence of words for automatic indexing. This simple approach can be extended by giving more weights to the words which occur in the alt or src tag of the image or which can occur inside the head tag or any other important tags of the HTML document. However, purely combination of traditional text-based retrieval and content-based retrieval is not adequate to deal with the problem of image retrieval on the WWW. The first reason is that there is already too much clutter and irrelevant information on the web pages. These semantic features are less accurate than annotating text. The second reason is due to the mismatch between the page author's expression and the user's understanding and expectation.This problem is similar to the subjectivity of image annotation. The third reason is due to the difficulty to find out the relationship between low-level features and high-level features. The second approach takes a different stand and treats images and texts as equivalent data. It attempts to discover the correlation between visual features and textual words on an unsupervised basis, by estimating the joint distribution of features and words and posing annotation as statistical inference in a graphical model. For example image retrieval system based on decision trees and rule induction was presented in [R.C.F.Wong] to annotate web image using combination of image feature and metadata, while in [N.Xiong], a system that automatically integrate the keyword and visual features for web image retrieval by using association rule mining technology. These approaches usually learn the keywords correlations according to the appearance of keywords in the web page, and the correlation may not reflect the real correlation for annotating Web images or semantic meaning of keywords such as synonym [H.Xu]. Ontology-based image retrieval is an effective approach to bridge the semantic gap because it is more focused on capturing semantic content which has the potential to satisfy user requirements better [V.Mezaris,O.Murdoch]. While semantically rich ontology addresses the need for complete descriptions of image retrieval and improves the precision of retrieval. However, the lack of text information which affects the performance of keyword approach is still a problem in text ontology approach. Ontology works better with the combination of image features [H.Wary].This paper presents a new framework for web image retrieval search engine which relay not only on ontology to discover the semantic relationship between different keywords inside the web page but also propose a new voting annotation technique extract the shared semantically related keywords from different web pages to eliminate and solve the problem of subjectivity of image annotation of traditional approaches and enhance the performance of the retrieval results by taking the semantic of the correlated data into consideration.

3. Architecture of the Proposed Web Image Retrieval Search Engine Prototype

3.1. General Overview

Due to there are many unrelated keywords are associated to the web image, in order to enhance the retrieval process of web image, it should be decrease or remove these keywords. The proposed approach trying to solve this shortcoming of most of current systems by proposing a voting technique which based on frequent item set mining to get the relation between low level feature of image content and high level feature. The more frequent the keyword occurs is the key of measuring keyword correlation. If two or more keywords appearing together frequently with an image can be considered as being highly relevant to each other [Y.Yang]. By considering a near-duplicate (visually similar) images as a transaction and its associated keywords as the items in the transaction, it is very natural to discover correlations between image low level feature and keywords by applying association rules mining model [R.Agrawal].

3.2. System Components

The proposed search engine is composed mainly on two phase, preprocessing phase and semantic search phase. The preprocessing phase is responsible for data and image collection and semantic annotation, while semantic search phase is responsible for image retrieval. The next sections explain in details each phase and its components.

3.2.1 Preprocessing Phase

The preprocessing phase has the following main modules: These modules are: (1) Web Crawler module, (2) Parser module, (3) Image Processing module, (4) NLP module, (5) Voting Module. Figure 1 shows these modules. Each module from these modules will be composed to a set of functions in terms of system functionality. The following section of the research contains the description of each module and its functions in details.

(1) Web Crawler Module

There are lots of images available on the Web pages. In order to collect these images, a crawler (or a spider, which is a program that can automatically analyze the web pages and download the related pages). Instead of creating a new ontology from scratch, extend WordNet, the well-known word ontology, to word-image ontology, WordNet is one of the most widely used, commonly supported and best developed ontologies. In WordNet, different senses and relations are defined for each word. Will use Wordnet to provide a comprehensive list of all classes likely to have any kind of visual consistency. Do this by extracting all non-abstract nouns from the database, 75,062 of them in total, by collecting images for all nouns, have a dense coverage of all visual forms. Figure 2 shows a repeat region comprised of four blocks in repeated style. Each block contains a restaurant record that includes four attributes: picture , restaurant name, location description, and rating. Extract three lists from this region: a list of restaurant names, a list of location description , and a list of ratings, and ignore images.

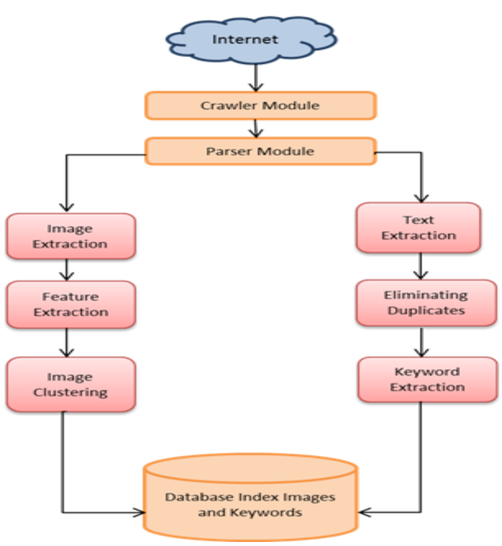


Fig 1. The Proceeing Phase

Fig 2. An example repeat region in a webpage

To extract these lists, first detect repeat regions in webpages based on vision-based DOM trees . Here a repeat region is the region that includes at least two adjacent or nonadjacent blocks, e.g., *M* blocks, with similar DOM and visual structures. Then extract all leaf HTML nodes within each block, and group them by their tag names and display styles. In the above example, all restaurant names have the same tag name (<a>) and displaying style (in blue color), and they can be grouped together. Each group usually contains *M* nodes. Each two of them are from different blocks. At last, for each group, we extract all text from its nodes as a list. Named this kind of pattern as *REGION*.
For a list extracted from a repeat region, choose the owest common ancestor element of all blocks of the repeat region as a container node.

(2) HTML Parsing Module

In our system, we use HTML Parser to transform html documents into DOM tree. the DOM Tree is based webpage segmentation algorithm that automatically segments web pages into sections, with each section consisting of a web image and its contextual information (i.e. image segment), and then extract the text and images by traversing through the DOM tree. . First, we generate the DOM tree for each web page containing the web images. From the bottom, visual objects like image, text paragraph are identifier as basic elements. The tags such as <TABLE>, <TD>, <TR> and <HR> are used to separate the different content passages.

3) Natural Language Processing (NLP) Module

The raw text of the document is treated separately as well. In order to extract terms from text, classic Natural Language Processing (NLP) techniques are applied. This module responsible for these functions:

Stop Words Removing:

Removal of non-informative categorization. In this scenario, a pre defined "stop list", which consist of hundreds of less meaningful high-frequency words(e.g., prepositions and conjuctions), is employed to eliminate irrelevance information in text documents. Such a method in quite advantageous for improving the accuracy of the search and reducing the redundancy of the computation.
Weighted lists as query facets because:
(1) An individual list may inevitably include noise. For example, the first item of the first list in table 2, i.e, "watch brands", is noise. It is difficult to identify it without other information provided;
(2) an individual list usually contains a small number of items of a facet and thus it is far from complete.
(3) many lists contain duplicated information. They are not exactly same, but share overlapped items. To conquer the above issues.

Word Stemming: By stemming of word, it will be changed into the word's basic form. The documents are first parsed into words. Second the words are represented by their stems, for example 'walk', 'walking' and 'walks' would be represented by the stem 'walk'.

Keyword Extraction: After stemming each word are then weighted using a normalized (tf-idf). At the end, the text part of the document is represented simply by a set of keywords and weights.

(4) Image Processing Module

This module is responsible for performing the function that are related to the image, the next section explain these functions in details.

Image Clustering

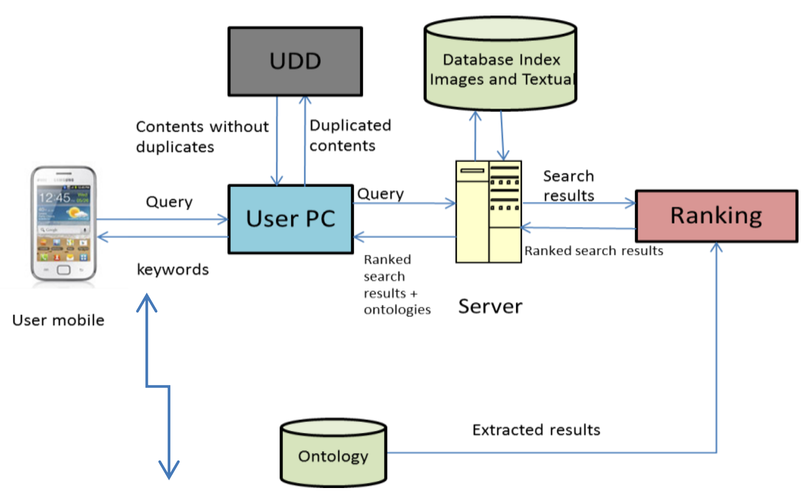
The k-means algorithm is used to perform the clustering process. This choice was mainly motivated by the comparably fast processing of the k-means algorithm compared to other unsupervised clustering. K-means can be described as a partitioning method. It is an unsupervised clustering method that provides k clusters, where k is fixed a priori. K-means treats each observation in data as an object having a location in space. It finds a partition in which objects within each cluster are as close to each other as possible, and as far from objects in other clusters as possible. First k points are chosen as centroids (one for each cluster). The next step is to assign every point from data set to the nearest centroid.

Feature Extraction Process

The color feature is extracted using color histogram method. Color histogram is popular because they are trivial to compute, and tend to be robust against small changes in object rotation and camera viewpoint. The color histogram represents an image by breaking down the various color components of an image and extracts the three histograms of RGB colors; Red (HR),Green (HG),and Blue(HB), one for each color channel by computing the occurrences of each color (histogram). After computing the histogram, the histogram of each color is normalized because the images are downloaded from different sites which maintain images with different size.

(5) Ranking Module

In the previous modules all images have been initially annotated and visually clustered. Due to the primarily annotation error, the target image may be primarily annotated with error keyword. The underlying problem that we attempt to correct is that annotations generated by probabilistic models present poor performance as a result of too many "noisy" keywords. By "noisy" keywords, we mean those which are not consistent with the rest of the image annotations and in addition to that, are incorrect. Our assumption in this module is that if certain images in the database are visually similar (located in one cluster) and semantically related to the candidate annotations, the textual descriptions of these images should also be related together. If there is in the cluster image label does not have similarity in semantic to annotations to other images, this means that this image was initially annotated with error keyword. In this module, we attempt to find the proper label for each cluster based on semantic analysis of the different candidate image labels inside the cluster and then using data mining technique to find association rule between each different keyword and the cluster to select most appropriate keyword to this cluster. This module consist these functions;



Concept Extraction

In this function the concept (lemma) of the keyword will be extracted from wordnet, then the similarity between the different keywords are measured, this process performed because the image may be described by different keywords ind different web pages but the meaning is related. Two main relationships between keyword are analyzed; taxonomy and Partonomy. Taxonomy divides a concept into species of kinds (e.g. Car and bus are types of vehicle), while Partonomy divides the concept as a whole into different parts (e.g. Car and wheel). For example, such analysis might show that car is more like a bus than it is a tree, due to the fact that car and bus share vehicle as an ancestor in the WordNet noun hierarchy.

Image Label Ranking

Our assumption is that if certain images in the database are visually similar to the target image and semantically related to the candidate annotations, the textual descriptions of these images should also be correlated to the target image. If the target image label does not have similarity in semantic to the candidate image label, this means that the target image was annotated with error keyword.

One of the typical data mining functions is to find association rules among data items in a database. To discover the Association between the high-level concept and low-level visual features of images, we need to quantify the visual features by clustering, because the concept space is discrete while the visual feature space is continuous in general. Therefore, we aim to associate the concepts and the visual feature.

Dentitions

1) imageset: A set of one or more images
2) k-images X = {x1, …, xk}
3) count of X: Frequency or occurrence of instance of image X
4) support , s, is the fraction of transactions that contains X (i.e., the probability that a transaction contains X)
5) An imageset X is frequent if X's support is no less than a minsup threshold
6) support, X ^ Y , probability that a transaction contains XuY
7) confidence, c, conditional probability that a transaction having X also contains Y
8) An association rule is a pattern that states when X occurs,Y occurs with certain probability.

The steps of annotation process based association rule:

(1) Scan the transaction DB to get the support S of each concept Ki and visual cluster Ci, and select those concepts and clusters with support greater than user specified minimum support.

(2) Construct the transaction database D and the basic candidate 2-itemsets based on the existing inverted file. We do not start from 1-itemset because the visual features are very high dimensional and the associations between concepts which are single modality association rules are much stronger than the associations between concepts and low-level features or low-level visual clusters. If starting from 1-itemset, the concepts and visual feature cluster are equally treated, and then most of the created 2-itemsets based on 1-itemset are concept and concept, but few of concept and visual feature cluster. Our goal is not the association between concept and concept. We are interested in the association between concepts and visual feature clusters. Therefore, only the imageset containing one concept and one visual feature cluster are considered. The existing inverted file relates the concepts to their associated images.

(3) For each concept ki in the cluster cj, calculate the support between concept ki and cluster cj.
Supp (Ki, Cj)= Count(Ki,Cj) / Size(Cj). Where count(ki) is the frequency of occurrence of concept ki in the cluster cj; while size(cj) is the total number of visual images in the cluster.
(4) All imageset that have support above the user specified minimum support are selected. These imageset are added to the frequent imageset.
(5) For each imageset in the frequent imageset, calculate the confidence between concept ki and cluster cj.conf (Ki, Cj)= Count(Ki,Cj) / count(ki)
Where count(ki) is the frequency of occurrence of concept ki in the database;
(6) The rules that have confidence >= minimum Confidence are selected to strong rule. Order all frequent imagset in the strong rule according to their confidence, and then select the concept with highest confidence as a label to the associated cluster.

Ontology Reasoning

Ontology reasoning is the cornerstone of the semantic web, a vision of a future where machines are able to reason about various aspects of available information to produce more comprehensive and semantically relevant results to search queries. Rather than simply matching keywords, the web of the future will make use of ontology to understand the relationship between disparate pieces of information in order to more accurately analyze and retrieve images. Most image retrieval method always assumes that users have exact the mind searching goal in mind. However, in the real world application, the case is that users do not clearly know what they want. Most of the times, they only hold a general interest to explore some related images. The ontology reasoning is based on the semantic associations between keywords. This is achieved by finding which concepts in the ontology relate to a keyword and retrieving information about each of these concepts. By this module the ontology is used for quickly locating the relevant semantic concept and a set of images that are semantically related to the user query are returned.

4. Conclusion and Future Work

After a review of existing techniques related to web image retrieval, that point out that these methods are not powerful enough to retrieve efficiently relevant images including semantic concepts.The propose an architecture that combines semantic annotation, data mining and visual features which are collected from different web pages that share visually similar images not from single web page as traditional approaches.. This system use visual ontology, which is a concept hierarchy, is built according to the set of annotations. In the retrieval process to suggest more results that are related to the user's query.Currently,continue to develop the proposed framework and to look for the best appropriate algorithms and methods to compute interesting relevant descriptive metadata and suited visual ontology.

REFERENCES

[1]Y. Alemu, J. Koh, and M. Ikram, "Image Retrieval in Multimedia Databases: A Survey" Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2009.

[2]R. Agrawal, T.S Imielin, A.T. Swami, "Mining Association Rules Between Sets of Items in Large Databases", SIGMOD Rec, 1993, Vol. 22, No.2, pp. 207–216.

[3]Z. Chen, L. Wenyin, F. Zhang, H. Zhang, "Web Mining for Web Image Retrieval", Journal of the American Society for Information Science and Technology - Visual based retrieval systems and web mining archive, Vol. 52, No. 10, 2001, pp. 832-839.

[4]Z. Gong, Q. Liu, "Improving Keyword Based Web Image Search with Visual Feature Distribution and Term Expansion", Journal Knowledge and Information Systems Vol. 21, No.1, 2009.
[5]R. He , N. Xiong, L.T. Yang, and J. H. Park, "Using Multi-Modal Semantic Association Rules to Fuse Keywords and Visual Features Automatically for Web Image Retrieval", Information. Fusion , 2010.
[6]J. Hou, D. Zhang, Z. Chen1, L. Jiang, H. Zhang, and X. Qin, "Web Image Search by Automatic Image Annotation and Translation", IWSSIP 17th International Conference on Systems, Signals and Image Processing, 2010.
[7]M.L. Kherfi, D. Ziou, A. Bernardi, "Image Retrieval from the World Wide Web: Issues, Techniques, and Systems", ACM Computing Surveys Vol.36, No. 1, 2004, pp. 35–67.
[8]V. Mezaris, I. Kompatsiaris, and M. G. Strintzis, "An Ontology Approach to Object-Based Image Retrieval", in Proc. IEEE ICIP, 2003.
[9]O. Murdoch, L. Coyle and S. Dobson, "Ontology-Based query Recommendation as A Support to Image Retrieval". In Proceedings of the 19th Irish Conference in Artificial Intelligence and Cognitive Science. Cork, IE. 2008.
[10]A. Torralba, R. Fergus, W.T. Freeman, "80 Million Tiny Images: A Large Data Set for Nonparametric Object and Scene Recognition" Pattern Analysis and Machine Intelligence, IEEE, Vol. 30, No. 11, 2008, pp. 1958 – 1970
[11]R.C.F. Wong and C.H.C. Leung,"Automatic Semantic Annotation of Real-World Web Images", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 30, No. 11, 2008, pp. 1933-1945.
[12]H. Wang, S. Liu, L.T. Chia, "Does Ontology Help in Image Retrieval ?: A Comparison Between Keyword, Text Ontology and Multi-Modality Ontology Approaches", Proceedings of the 14th annual ACM International Conference on Multimedia, 2006, pp. 23-27.
[13]H. Xu, X. Zhou, L. Lin, Y. Xiang, and B. Shi, "Automatic Web Image Annotation via Web-Scale Image Semantic Space Learning", APWeb/WAIM 2009, LNCS 5446, pp. 211–222, 2009
[14]Y. Yang, Z. Huang, H. T. Shen, X. Zhou," Mining Multi-Tag Association for Image Tagging", Journal World Wide Web Archive, Vol. 14, No.2, 2011.