Enhanced Visual Search Using Image Re-Ranking

Raviraj S. Patkar
PG Student IOK-COE
Computer Engineering Department
Shikrapur, Pune
Maharashtra, India

S. Pratap Singh
Asst. Prof. & Dean IOK-COE
Computer Engineering Department
Shikrapur, Pune
Maharashtra, India

Swati V. Birje
Asst. Prof. RMCET
ExTC Department
Ambav (Devrukh)
Maharashtra, India

ABSTRACT
Image re-ranking is being used by commercial search engine giants like Google and Bing as an effectual way to improve the precision of image search results. Current methods mostly focus on text based information like xml tags and meta tags. Image redundancy has still been a major area of concern. It is very difficult to understand users expected search results only by providing text based keywords. It also leads to ambivalent and obstreperous results which are far from satisfactory. It is therefore important to use visual information in order to solve the ambiguity in text-based image search. This paper proceeds through various image re-ranking techniques, approaches and methodologies for a quick review and then proposes a enhanced visual search technique using SVM and KNN algorithms.

Keywords
Image re-ranking, precision, xml tags, meta tags, SVM, KNN.

INTRODUCTION
The exponential growth of the Web suggest that users are becoming more and more dependent on the search engines as well as ranking methods to discover information relevant to their needs. Typically, users expect to find such information in the top-ranked results, and more often than not they only look at the document snippets in the first few result pages and then they give up or reformulate the query. Information retrieval is the activity of obtaining information resources relevant to an information need from a collection of information resources. Searches can be based on metadata. An information retrieval process begins when user enters a query into the system. Queries are formal statements of information needs, for example search strings in web search engines. In information retrieval a query does not uniquely identify a single object in the collection. Instead, several objects may match the query, perhaps with different degrees of relevancy. Although image search has become a popular feature in many search engines the majority of image searches use little, if any, image information to rank the images. Instead, commonly only the text on the pages in which the image is embedded (text in the page body, anchor-text, image name, etc) is used. There are three reasons for this: first, text-based search of web pages is a well studied problem that has achieved a great amount of real-world success. Second, a fundamental task of image analysis is yet largely an unsolved problem: human recognizable objects are usually not automatically detectable in images. Although certain tasks, such as finding faces [15] [16] and highly textured objects like CD covers [17], have been successfully addressed, the problem of general object detection and recognition remains open. Few objects other than those mentioned above can be reliably
detected in the majority of images. Third, even for the tasks that are successfully addressed, the processing required can be quite expensive in comparison to analyzing the text of a webpage. Not only do the signal-processing algorithms add an additional level of complexity, but the rapidly increasing average size of images makes the simple task of transferring and analyzing large volumes of data difficult and computationally expensive.

The problem of answering a query without image processing is that it can often yield results that are inconsistent in terms of quality. For example, the query “Eiffel Tower” submitted to image search on Google.com returns good results as shown in Figure 1. However, the query for “McDonalds” returns mixed results as shown in Figure 2; the typical expected yellow “M” logo is not seen as the main component of an image until results 6 and 13.

**LITERATURE REVIEW**

In this Section, we survey related work on image re-ranking.

**Circular Re-ranking**

In [1], Ting Yao proposed Multi-modal graph-based and circular re-ranking techniques. It captures more than one feature of image for more accurate re-ranking results. The basic idea of circular re-ranking is to facilitate interaction among different modalities through mutual reinforcement. In this way, the performance of strong modality is enhanced through communication with weaker ones, while the weak modality is also benefited by learning from strong modalities. These methods do not always compete but can complement each other.

**(pLSA) for mining visual categories through clustering of images**

Fergus R. et al., [2], employed probabilistic Latent Semantic Analysis (pLSA) for mining visual categories through clustering of images in the initial ranked list and which extends pLSA (as applied to visual words) to include spatial information in a translation and scale invariant manner. Candidate images are then re-ranked based on the distance to the mined categories. Self-re-ranking seeks consensus from the initial ranked list as visual patterns for re-ranking.

**Information Bottleneck (IB) re-ranking**

Hsu et al., [4], Smeulders et al., [9] employed information bottleneck (IB) re-ranking to find the clustering of images that preserves the maximal mutual information between the search relevance and visual features. The IB re-ranking method, based on a rigorous Information Bottleneck (IB) principle which finds the optimal image clustering that preserves the maximal mutual information between the search relevance and the high-dimensional low-level visual features of the images in the text search results. Among
all the possible clustering’s of the objects into a fixed number of clusters, the optimal clustering is the one that minimizes the loss of mutual information (MI) between the features and the auxiliary labels.

**Crowd Re-ranking**

Richter et al.[10], employed an crowd re-ranking, it is similar to self-re-ranking except that consensus is sought simultaneously from multiple ranked lists obtained from internet resources and further formulate the problem as random walk over a context graph built through linearly fusing multi-modalities for visual search.

Liu et al. [5], suggested a re-ranking paradigm by issuing query to multiple online search engines. Based on visual word representation, both concurrent and salient patterns are respectively mined to initialize a graph model for randomized walks based on re-ranking. Different from self- and crowd-re-ranking, example-based re-ranking relies on a few query examples provided by users for model learning.

**Multimedia search with pseudo relevance feedback**

Yan et al, [7], Tao Mai et al, [11] employed classifiers are learnt by treating query examples as positive training samples while randomly picking pseudo-negative samples from the bottom of initial ranked list. The classifiers which capture the visual distribution of positive and negative samples are then exploited for re-ranking.

**Optimizing video search re-ranking via minimum incremental information loss.**

Liu et al. [6], proposed query examples are utilized to identify relevant and irrelevant visual concepts, which are in turn be employed to discover the rank relationship between any two documents using mutual information for correcting ranking of document pairs.

**PROPOSED WORK**

**Motivation**

Current approaches focus on the mining of recurrent patterns from different means, such as by random walk [3], external knowledge [6], and classifier learning [7], web image search engines retrieve and rank images mostly based on surrounding text features.

Image redundancy is still a problem area concerned as it is difficult to interpret users search intention only by query keywords and this leads to ambiguous and noisy search results which are far from satisfactory. It is important to use visual information in order to solve the ambiguity in text-based image retrieval.

**Objectives**

**Capture users intention:** Capture users intention to search relevant images with minimum efforts.

**To enlarge image pool:** Enlarge image pool to retrieve more relevant images.

**Keyword expansion:** Expand the query keywords in order to remove ambiguity or doubtfulness of user.

**Integrate various visual features:** Integrate various visual features in order to compute the similarities between the query image and other images.

**To improve retrieval precision:** Improve retrieval precision by retrieving images those are similar to query from a large database.

**To remove noisy search results:** Use visual features to remove noisy search results.

**Experimental analysis:** Analyze proposed work with various parameters like precision rate, image retrieval time etc.

**Methodology**

Our approach will consist of following phases in order to re-rank the images and provide improved search results.
1. Database Modeling and System Setup
2. Feature Extraction
3. Visual Similarity
4. Relevant image search
5. Combining Visual and Textual Similarities

**Proposed Architecture**

The proposed system in Fig 3 will make use of Support Vector Machine (SVM) and K-nearest neighbors (KNN) algorithms. The image retrieved from query keyword will be subjected to feature extraction. Three features are considered over here; Color, Texture and Shape. Factors under consideration of these features is as shown in Table 1. The SVM classifier is used to group the images into different categories based on extracted features. The KNN classifier helps to remove noise and display relevant images from the categories. A More enhanced result is then formulated using the keyword expansion.

**Table 1. Visual Features**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Factors to be considered</th>
</tr>
</thead>
</table>
| Color   | 1. Color Space  
          | 2. Color Correlation 
          | 3. Coherence Vector 
          | 4. Histogram |
| Texture | 1. Tamura Feature 
          | 2. Gabor Filter |
| Shape   | 1. Moment Invariant 
          | 2. Turning Angles 
          | 3. Polynomial approximation 
          | 4. Fourier Descriptors |

**Detailed information of algorithms**

**Support vector machine (SVM)**

SVM algorithm is used for the classification of both linear and non-linear data. It uses a non-linear mapping to transform the original training data into a higher dimension. Within this new dimension, it searches for the linear optimal separating hyper-plane. Hyperplane is a decision boundary separating the tuples of one class from another. In case of support vector machine an object is viewed as a n-dimensional vector and such objects are separated with the help of hyperplane. The SVM finds the maximum separating hyperplane. The hyperplane with the larger margin to be more accurate at classifying than the hyperplane with the smaller margin[18].

The goal of SVM is try to address the nearest distance between a point in one class and a point in the other class being maximized and draw a hyperplane to classify two categories as clearly as possible.
K-Nearest neighbors (KNN)

Nearest neighbor classifiers are based on learning by analogy[18]. In KNN an object is classified by the “distance” from its neighbors, with the object being assigned to the class most common among its k distance-nearest neighbors. If k = 1, the algorithm simply becomes nearest neighbor algorithm and the object is classified to the class of its nearest neighbor. "Closeness" is defined in terms of a distance metric, called Euclidean distance. Distance is a key word in this algorithm, each object in the space is represented by position vectors in a multidimensional feature space. It is usual to use the Euclidean distance to calculate distance between two vector positions in the multidimensional space. The training process for KNN consists only of storing the feature vectors and class labels of the training samples[19].

CONCLUSION

Various surveys conducted on image re-ranking conclude that text based methods used for image retrieval are not sufficient enough to provide precise results and thus an additional visual similarity and keyword expansion domain would help to understand the users expected results.

The domain of image retrieval and image re-ranking has offered a vast research scope for exploration and also for innovation. Methods surveyed and proposed in this paper will be beneficial for researchers as an overview of the amount of work that has been done in this area.

REFERENCES


