International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970)
Volume-3 Number-2 Issue-10 June-2013

Efficient Retrieval of Images for Search Engine by Visual Similarity and Re Ranking

Viswa S S

Abstract

Nowadays, web scale image search engines (e.g. Google Image Search, Microsoft Live Image Search) rely almost purely on surrounding text features. Users type keywords in hope of finding a certain type of images. The search engine returns thousands of images ranked by the text keywords extracted from the surrounding text. However, many of returned images are noisy, disorganized, or irrelevant. Even Google and Microsoft have no Visual Information for searching of images. Using visual information to re rank and improve text based image search results is the idea. This improves the precision of the text based image search ranking by incorporating the information conveyed by the visual modality. The typical assumption that the top- images in the text-based search result are equally relevant is relaxed by linking the relevance of the images to their initial rank positions. Then, a number of images from the initial search result are employed as the prototypes that serve to visually represent the query and that are subsequently used to construct meta re rankers .i.e. The most relevant images are found by visual similarity and the average scores are calculated. By applying different meta re rankers to an image from the initial result, re ranking scores are generated, which are then used to find the new rank position for an image in the re ranked search result. Human supervision is introduced to learn the model weights offline, prior to the online re ranking process. While model learning requires manual labelling of the results for a few queries, the resulting model is query independent and therefore applicable to any other query. The experimental results on a representative web image search dataset comprising 353 queries demonstrate that the proposed method outperforms the existing supervised and unsupervised Re ranking approaches. Moreover, it improves the performance over the text-based image search engine by more than 25.48%.

This work was supported in part by the Department of Computer Science and Engineering, PSG College of Technology, India (Autonomous, Government Aided Private Engineering College).

Viswa S S, Department of Computer Science and Engineering, PSG College of Technology, Coimbatore, India.

Keywords

Image search, Image Retrieval, Efficient Image Search, Image Re ranking.

1. Introduction

The existing web image search engines, including Bing, Google, and Yahoo!, retrieve and rank images mostly based on the textual information associated with the image in the hosting web pages, such as the title and the surrounding text. While text-based image ranking is often effective to search for relevant images, the precision of the search result is largely limited by the mismatch between the true relevance of an image and its relevance inferred from the associated textual descriptions.

The existing web image search engines, including Bing, Google, and Yahoo!, retrieve and rank images mostly based on the textual information associated with the image in the hosting web pages, such as the title and the surrounding text. While text-based image ranking is often effective to search for relevant images, the precision of the search result is largely limited by the mismatch between the true relevance of an image and its relevance inferred from the associated textual descriptions. To improve the precision of the text-based image search ranking, visual re ranking has been proposed to refine the search result from the text-based image search engine by incorporating the information conveyed by the visual modality.

Survey

Visual re ranking has become a popular research topic in both multimedia retrieval and computer vision communities since it provides possibilities for considering the visual modality in the existing image search engines in a lightweight fashion and without incurring scalability issues. Moreover, apart from the image search scenario, visual re ranking can also be used to improve the quality of the collected data in the process of automatically constructing training data from the web for object recognition. While various techniques including clustering , topic modelling , support vector machine (SVM), graph learning , etc. have been investigated for the purpose of creating visual search re rankers, all of the existing re ranking algorithms require a prior assumption regarding the relevance of the images in the initial, text-based search result. In the most widely used pseudo relevance feedback (PRF) assumption, the top- images of the initial result are regarded as pseudo relevant and used to learn a visual classifier for re ranking. Even though the PRF-based re ranking methods have been able to improve the precision over the initial text-based result in the past, the assumption that the top- images are equally relevant can still be seen as too rigorous to be satisfied well by any arbitrary text-based image search engine. Since the text-based image search is far from perfect (which is the reason to perform the re ranking in the first place), the top result will inevitably contain irrelevant images, which will introduce noise into the learning of re ranking models and which may lead to sub-optimal search results being returned after re ranking. In this sense, appropriately relaxing this assumption and redefining the re ranking approach accordingly has the potential to further improve the precision of the visual re ranking. In this paper we address this challenge by recalling the fact that image search engines usually optimize the system performance based on the relevance measures, such as normalized discounted cumulative gain (NDCG), which tend to emphasize differently on the results at different ranks. Hence, it can naturally be assumed that the images in the top result of each query at different ranks have different probabilities to be relevant to the query. This should be incorporated into the re ranking model for a more comprehensive utilization of the text-based search result.

Existing System

In the most widely used pseudo relevance feedback (PRF) assumption, the top- images of the initial result are regarded as pseudo relevant and used to learn a visual classifier for re ranking. Even though the PRF-based re ranking methods have been able to improve the precision over the initial text-based result in the past, the assumption that the top- images are equally relevant can still be seen as too rigorous to be satisfied well by any arbitrary text-based image search engine. Since the text-based image search is far from perfect (which is the reason to perform the re ranking in the first place), the top result will inevitably contain irrelevant images, which will introduce noise into the learning of re ranking models and which may lead to sub-optimal search results being returned after re ranking.

The other widely-adopted image search re ranking assumption is the cluster assumption, which says that the visually similar images should be ranked nearby. Based on this assumption, various graph-based methods have been proposed to formulate the image search re ranking problem. The main deficiency of this assumption is that it makes the visual similarity of images equal to the similarity of their relevance to the query. In addition, it omits to identify two images as equally relevant to the query if they are insufficiently visually similar to each other to each other.

2. System Overview

Problem Description

- Assumptions and necessity of redefining the
- re ranking approach
- Poor precision of the visual re ranking
- Ad hoc and suboptimal nature
- Poor mapping of visual similarity into
- relevance

Proposed System

In this project, the above problem is addressed challenge by recalling the fact that image search engines usually optimize the system performance based on the relevance measures, such as normalized discounted cumulative gain (NDCG), which tend to emphasize differently on the results at different ranks. Hence, it can naturally be assumed that the images in the top result of each query at different ranks have different probabilities to be relevant to the query. This should be incorporated into the re ranking model for a more comprehensive utilization of the text-based search result. Here, a prototype-based method is proposed to learn re ranking function from human labelled samples, based on the assumption that the relevance probability of each image should be correlated to its rank position in the initial search result. Based on the images in the initial result, visual prototypes are generated that visually represent the query. Each of the prototypes is used to construct a meta Re ranker to produce are ranking score for any other image from the initial list. Finally, the scores from all meta re rankers are aggregated together using a linear re ranking model to produce the final relevance score for an image and to define its position in the re ranked results list.

The linear re ranking model is learned in a supervised fashion to assign appropriate weights to different meta re rankers. Since the learned model weights are related to the initial text-based rank position of the corresponding image and not to the image itself, the re ranking model is *query-independent* and can be generalized across queries. Consequently, the proposed re ranking method can scale up to handle any arbitrary query and image collection, just like the existing visual re ranking approaches, even though supervision is introduced.

Enhancement

- Speeding up the Prototype-Set method variant while decreasing the precision degradation using online learning algorithms
- Making the system query-adaptive by automatically estimate the query-relative reliability and accuracy of each meta-re ranker and then incorporate it into the re ranking model.

Advantage

- Generic re ranking approach
- Higher precision of visual re ranking
- A graph within a graph is an "inset", not an "insert". The word alternatively is preferred to the word "alternately" (unless you really mean something that alternates).

3. Methodology

Modules

- Database modeling and system setup
- Text based search
- Constructing meta re rankers
  - Single-image prototype
  - Multiple-average prototype
- Performance analysis

Database Modelling and System Setup

Database modelling starts with an initial data collection and proceeds with activities in order to get familiar with the data. This includes operations to identify data quality problems, to discover first insights into the data or to detect interesting subsets to form hypotheses for hidden information.

- Collect initial data - acquire images from website with textual description
- Explore data - examine the "images" and "description" data
- Data Preparation - update the data i.e. image and related information to database

Besides this, the front end GUI for project is also designed.

Text Based Search

In this module an interface is designed that enables users to specify search criteria about an image of interest. The criteria are referred to as a search query. In the case of text search engines, the search query is typically expressed as a set of words that identify the desired concept that one or more documents may contain. Here the same concept is used but for image documents. There are several styles of search query syntax that vary in strictness. It can also search names within the search content as well as from its description too. Whereas this text based search require users to enter two or three words separated by white space and may result in unwanted results. Some existing search engines apply improvements to search queries to increase the likelihood of providing a quality set of items through a process known as query expansion. But this doesn't produce exact results for non-textual search. This existing functionality is implemented here.

Multiple-Average Prototype

In this module, we construct a prototype by first selecting the top images in the initial search result list and then by cumulatively averaging the features of all images ranked starting from the topmost position to the position, as illustrated in Figure 1.

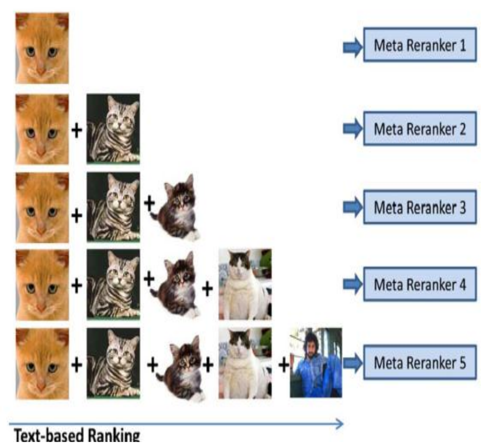

Figure 1: Multiple Image Prototype

In other words, the prototype can be defined as

$$P_i^{MA} = \frac{1}{i}\sum_{j=1}^{i} I_j.$$

Then, the prototypes can be employed to compute the scores of individual meta re rankers by again computing the visual similarity between a prototype and the image to be re ranked:

$$M^{MA}\left(I_j | P_i^{MA}\right) = S\left(I_j, P_i^{MA}\right).$$

4. Implementation and Experimental Results

Pre-processing of Database

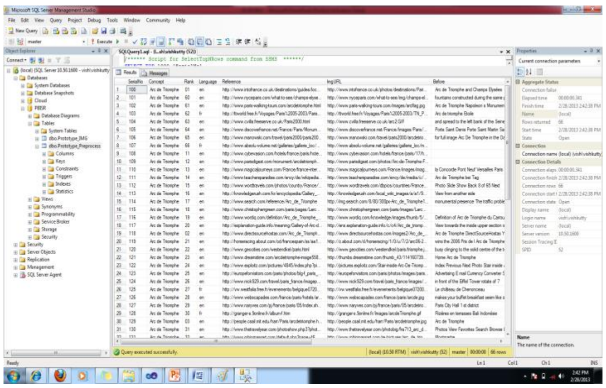

Figure 2: Pre-processing of the Database

This phase includes adding all the XML files with the respective image files. Microsoft SQL Server Management 28 is used.

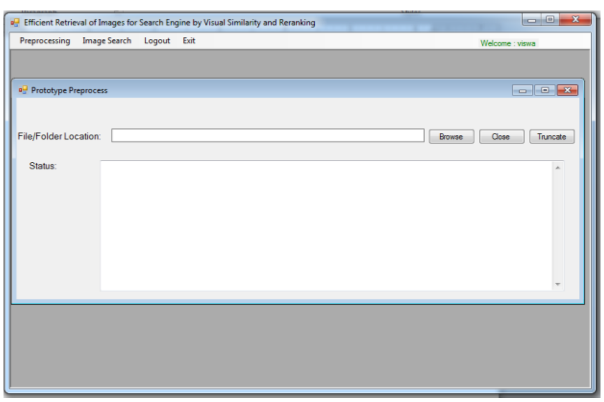

Figure 3: Updating the Database to the Search Engine

Search Engine and Re Ranking

User has the capability to retrieve the images, display the images with respective rank order, clear results, search multiple queries instantly.

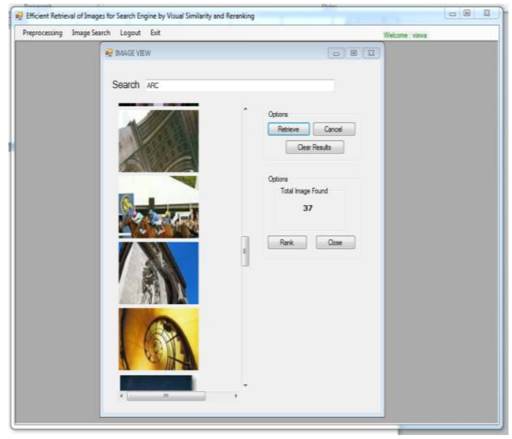

Figure 4: Query "arc"

First the user searches the query, here text based result queries are used. Example "arc", all images having surrounding text "arc" is displayed.

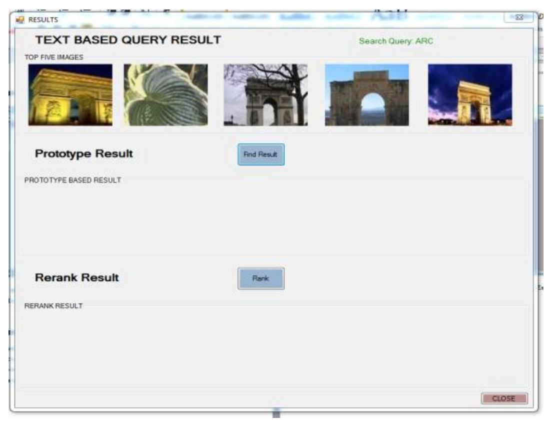

Figure 5: Displaying the top ranked images

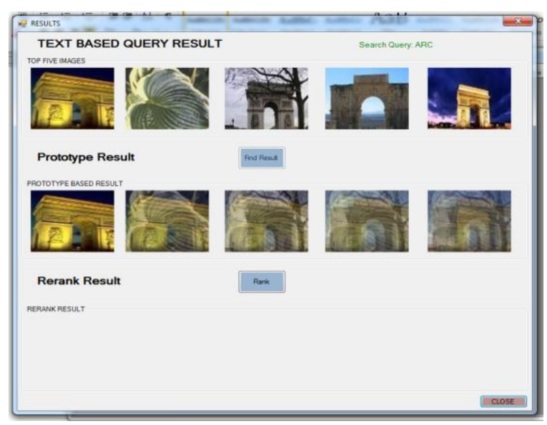

Figure 6: Construction of Meta re ranker

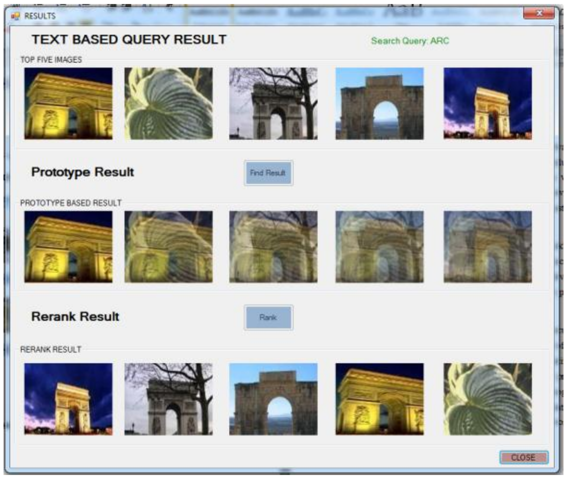

Figure 7: Final Re ranking of Images

Query is "arc" first the top ranked images are displayed, then a prototype re ranker is being constructed. From the meta re ranker each image from top 5 is being taken and fed into visual similarity function.
Visual Similarity is found by comparing the RGB match % and then the images are re ranked according to the percentage. Thus the negatively sampled images are pushed back, giving way to the relevant images.

Performance Analysis
We can observe from the results that Prototype-Average perform slightly worse than Prototype-Single. This seems to be in contradiction with the intuition since Prototype-Average is based on a more robust assumption. However, the analysis indicates that this result is reasonable since Prototype-Average is more moderate in changing the text-based search result than Prototype-Single. The performance comparison between Prototype-Single and Prototype-Average with the text-based search result serving as baseline.

This comparison shows that while Prototype-Average is Less effective in boosting the re ranking performance, it is also more robust in the cases when re ranking may lead to performance degradation, like in the last examples discussed in the previous section.

Performance Comparison
Performance comparison of various re ranking methods. The numbers in the brackets are the relative improvements of various methods over the text-baseline.

Table 1: Various Search Engine Performance Comparison

Methods	MAP	NDCG@10	NDCG@40
Text-baseline	0.569	0.682	0.633
Text-ranking	0.594 (+4.39%)	0.684 (+0.29%)	0.649 (+2.53%)
PRF [9]	0.658 (+15.64%)	0.772 (+13.20%)	0.718 (+13.43%)
Random walk [10][11]	0.641 (+12.65%)	0.766 (+12.32%)	0.704 (+11.22%)
Bayesian [12]	0.643 (+13.01%)	0.766 (+12.32%)	0.709 (+12.01%)
Supervised-reranking [4]	0.665 (+16.87%)	0.769 (+12.76%)	0.733 (+15.80%)
Query-relative [15]	0.666 (+17.05%)	0.768 (+12.61%)	0.729 (+15.17%)
Prototype-Single	0.678 (+19.16%)	0.804 (+17.89%)	0.750 (+18.48%)
Prototype-Average	0.669 (+17.57%)	0.794 (+16.42%)	0.742 (+17.22%)
Prototype-Set	0.703 (+23.60%)	0.826 (+21.11%)	0.777 (+22.75%)
Prototype-All+Query-relative	0.706 (+24.08%)	0.823 (+20.67%)	0.779 (+23.06%)
Prototype-Set+Text	0.714 (+25.48%)	0.835 (+22.43%)	0.787 (+24.33%)

Performance comparison of Prototype-Set and Text-baseline from the search engine. The query is arranged in the ascending order of the performance of Text-baseline.

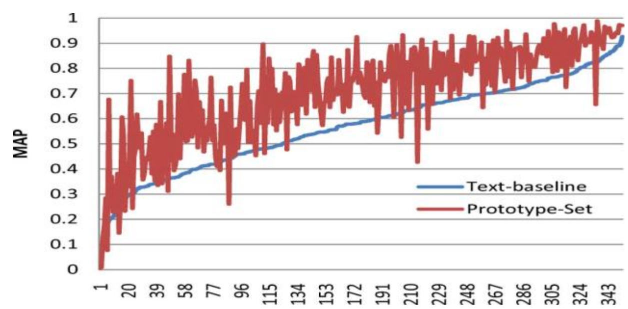

Figure 8: Performance Graph

5. Conclusion

In this project, we proposed a prototype-based re ranking framework, which constructs meta re rankers corresponding to visual prototypes representing the textual query and learns the weights of a linear re ranking model to combine the results of individual meta re rankers and produce the re ranking score of a given image taken from the initial text-based search result. The induced re ranking model is learned in a query-independent way requiring only a limited labelling effort and being able to scale up to a broad range of queries. The experimental results on the Web Queries dataset demonstrate that the proposed method outperforms all the existing supervised and unsupervised re ranking methods. It improves the performance by 25.48% over the text-based search result by combining prototypes and textual ranking features. A natural extension of the approach described in this paper would be to apply the proposed methods to learn concept models from image search engines in a semi-automatic fashion. Compared to the fully automatic methods, the semi-automatic approach could learn the concept models for any arbitrary concept much better and with only little human supervision.

While our proposed methods have proved effective for re ranking image search results, we envision two directions for future work to further improve the re ranking performance. First, we could further speed up the Prototype-Set method variant while decreasing the precision degradation. Since top images are incrementally added into the multiple-set prototypes to train the meta re rankers, one of the possible approaches in this direction is to utilize the online learning algorithms.

Second, although we assume that the rank position is generally correlated with the relevance value of the image found there, and while our results show that this assumption can be regarded valid in a general case, still deviations from this expectation can occur for individual queries. Hence, we could work on improving the proposed re ranking model to make it more query-adaptive. One possible approach here would be to automatically estimate the query-relative reliability and accuracy of each meta-re ranker and then incorporate it into the re ranking model. Another approach may be to learn the re ranking models for different query classes.

Acknowledgment

I am greatly indebted to Dr. R Venkatesan, Head of Department, Department of Computer Science and Engineering, PSG College of Technology, Coimbatore for granting me permission to undertake this project. I extend my profound gratitude to Mr B Vinoth Kumar, Assistant Professor, Department of Computer Science and Engineering, my Faculty Guide, for not only showing the path but also guiding me through the path. I thank him for his patience and encouragement he has given me.

References

[1] Yang, Linjun, and Alan Hanjalic. "Supervised reranking for web image search." In Proceedings of the international conference on Multimedia, pp. 183-192. ACM, 2010.

[2] Krapac, Josip, Moray Allan, Jakob Verbeek, and Frédéric Jurie. "Improving web image search results using query-relative classifiers." Computer Vision and Pattern Recognition (CVPR), pp. 1094-1101. IEEE, 2010.

[3] Krapac, Josip, Moray Allan, Jakob Verbeek, and Frédéric Jurie. "Improving web image search results using query-relative classifiers." Computer Vision and Pattern Recognition (CVPR), pp. 1094-1101. IEEE, 2010.

[4] Hsu, Winston H., Lyndon S. Kennedy, and Shih-Fu Chang. "Video search reranking through random walk over document-level context graph."15th international conference on Multimedia, pp. 971-980. ACM, 2007.

[5] Jones, Karen Sparck. "A statistical interpretation of term specificity and its application in retrieval." Journal of documentation 28, no. 1 (1972): 11-21.

[6] Tian, Xinmei, Linjun Yang, Jingdong Wang, Yichen Yang, Xiuqing Wu, and Xian-Sheng Hua. "Bayesian video search reranking." 16th ACM international conference on Multimedia, pp. 131-140. ACM, 2008.

[7] Wang, Xin-Jing, Wei-Ying Ma, Gui-Rong Xue, and Xing Li. "Multi-model similarity propagation and its application for web image retrieval." 12th annual ACM international conference on Multimedia, pp. 944-951. ACM, 2004.

[8] Yang, Jun, Yu-Gang Jiang, Alexander G. Hauptmann, and Chong-Wah Ngo. "Evaluating bag-of-visual-words representations in scene classification."International workshop on Workshop on multimedia information retrieval, pp. 197-206. ACM, 2007.

Viswa S S was born in Erode on 23.11.1992, is currently pursuing her Final Year in B.E Computer Science and Engineering in one of the prestigious Engineering Institutions in India, PSG College of Technology ranked 4th among top 10 Private Engineering Institutions (2012). Her interest and passion towards Computer Science grew right from her childhood and her perspective towards computer was different rather than playing games and watching movies. She utilized it for the real insight, how stuff works, understanding the know-how of various software and applications. Her Research Interests are Image Processing, Data Mining, Computer Graphics and Animation. One of her projects under the domain Image Processing has been published in the world`s renowned publisher, ELSEVIER in the year 2013. She has attended 2 International Conferences and 3 National Conferences and won Best paper award in one of the conference. She is a student member of Indian Society of Technical Education (ISTE), New Delhi. She has also served as a Joint Secretary, ISTE Students` Chapter PSG College of Technology, Coimbatore.