

Intent based Image Ranking for Web Search Reranking

Kshitija H. Ghadge¹, Sunil M. Sangve²

^{1,2}Department of Computer Engineering, Zeal College of Engineering, Pune, India

Abstract: New approach is presented for reranking of the images. While searching images on the web to increase the accuracy of the image search result Image search reranking is used. Image search reranking is an effective approach to refine the initial text-based image search result. In the reranking process, image visual contents are used to refine the initial text based result. The image re-ranking restricts users' attempt to just one-click feedback is a useful way to improve search results and its interaction is simple enough. One click event of the user is considered as user intention. We propose a feature extraction algorithm named Hypersphere-based Relevance Preserving Projection (HRPP) and a ranking function called Hypersphere-based Rank (H-Rank). A main challenge in the research of image Reranking is that the similarities of the visual features do not perfectly associate with actual semantic meaning of images which intends users search goals. H-Rank is an effective ranking algorithm to sort the images by their distances to the hypersphere center. Moreover, to capture the user's intent with minimum human interaction, a reversed KNN algorithm is proposed as One-Click-based HRPP (OC-HRPP). Finally, OC-HRPP algorithm and H-Rank algorithm form a new ISR method.

Keywords: Image search reranking, feature extraction, hypersphere based relevance preserving projection (HRPP), H-rank, Support Vector Machine (SVM), K nearest neighbour (KNN) algorithm, Visual Content.

I. Introduction

The exponential growth of image content has led to flourishing research and commercial activities in image search. The search engines search images mostly by using the text associated with the images like title of the images. This is often good to search relevant images using text associated with the image but the problem is that text based image retrieval do not take in to account the visual contents of the images or visual similarity of the image. Also in the text based image retrieval there is mismatch between textual features and the actual visual contents. However this kind of image search approaches often return some noisy results on the top of the ranking list because the text can not entirely reflect the image visual contents. To address this problem and to boost the performance of web image search and overcome the semantic gap between text information and image content, image search re-ranking is done, which refines the initial ranking orders by mining visual content or leveraging some auxiliary knowledge. A new promising direction has emerged named as image search Reranking which applies visual information to reorder the text based search results to improve the search performance. This Kind of approach first extract image's visual features from the initial search results and then build the ranking function with labeled training data and finally reorders the images with the ranking function. For example yang *et al.* were the first one to employ learning to rank technique in the search Reranking task. Ranking SVM and ListNet are the two popular learning to rank algorithms and employed to build the ranking function then rerank the initial search results.

Recently many methods have been proposed for image search reranking, which can be divided into three categories: classification-based, clustering-based, and random walk-based.

The first category is classification-based. It simplifies reranking as a classification problem. There are normally three steps: 1) select training samples from initial text-based search results; 2) train a classifier with the selected samples; and 3) reorder all samples according to predictions given by the trained classifier. In the first step, pseudo relevance feedback (PRF) is often utilized. PRF is a concept introduced from text retrieval. It assumes that a fraction of the top-ranked documents in the initial search results are pseudo-positive. Alternatively, uses the query images or example video clips as positive samples. The pseudo-negative samples are selected from either the lowest ranked samples in initial result or the database with the assumption that few samples in the database are relevant. In step 2), different classifiers, such as SVM Boosting, and Ranking SVM, can be adopted. Although the classifiers are effective, sufficient training data are demanded to achieve satisfactory performance since a lot of parameters need to be estimated.

The second category is clustering-based. In, it each sample is given a soft pseudo label according to the initial text search result, and then the Information Bottleneck principle is adopted to find optimal clustering which maximizes the mutual information between the clusters and the labels. Reranked list is achieved by ordering the

clusters according to the cluster conditional probability firstly and then ordering the samples within a cluster based on their local feature density. This method achieves good performance on the named-person queries, while it is limited to those queries which have significant duplicate characteristic.

The third category is random walk-based. A graph is constructed with the samples as the nodes and the edges between them being weighted by visual similarity.

Then, reranking is formulated as random walk over the graph and the ranking scores are propagated through the edges. To leverage the text search result, a “dongle” node is attached to each sample with the value fixed to the initial text ranking score. The stationary probability of the random walk process is adopted as the reranked score directly.

Methods like Clustering[7], topic modeling[6]-[2], Support Vector machine (SVM)[8], graph learning[9]-[10] have been investigated for visual re-ranking. All these require prior assumption regarding the relevance of images in initial text based search result. Also Top N search results can also contain irrelevant images which introduce noise. To overcome these drawbacks here we propose a system to address a ranking problem in web image retrieval that is we propose a system to re-rank images returned by image search engine as an initial search result. Re-ranking of images is done by incorporating, Visual aspects, Visual similarity of the images, to maximize relevancy of image results & to achieve diversity of image results. The goals of proposed system are Retrieve image results that are relevant and more approximate to query images. In image search re-ranking our aim is to refine the text based search result which requires mining image visual content. That is taking into account the image visual contents while reranking.

Also the precision of search result is less. So text based image retrieval has achieved limited success, and it is less approximate so that ISR has been receiving great attention. The two important steps in ISR are extracting low level features of the images, low level features means number of pixels, colors, curves, shapes etc. to represent image visual contents. For that purpose we propose a Hypersphere-based Relevance Preserving Projection (HRPP) algorithm. It transforms original high dimensional feature space into low dimensional hypersphere space by preserving the relevance relationship with the images. The second step in ISR is ranking function called Hypersphere-based Rank (H-Rank). H-Rank algorithm ranks the initially searched images by calculating their distances to the hypersphere center. The design of ranking function is another key issue in ISR, it can be categorized into classification-based methods, graph-based methods, and learning-to-rank-based methods. Classification-based methods first train a classifier with the training data directly from the initial search results, and then reorder all the documents by the relevance scores predicted from the classifier. Support vector machine (SVM) is typically employed in it. However, this method takes ranking as a classification problem and neglects their differences, thus generally performs poorer than the other two kinds of methods. Graph-based methods are often implemented from a random walk or Bayesian perspective.

To capture the user’s intent with minimum human interaction, we propose a reversed KNN algorithm to search pseudo-relevant images by requiring that the user gives only one click on the initially searched images. We call the HRPP method with reversed KNN as One-Click-based HRPP (OC-HRPP).

We bring the idea of hypersphere in one-class classification to transform the original visual features and reorder the images by assuming that the relevant images are located close to the hypersphere center while the irrelevant images are away from the center.

II. Literature Survey

A. Multimodal graph-based reranking:

A recently proposed novel ISR method exploring multiple modalities. The approach simultaneously learns relevance degrees, weights of modalities, and the distance metric with its scaling for each modality.

B. Context Reranking:

A typical graph-based ISR method which formulates the process as a random walk over a context graph, where images are nodes and the edges between them are weighted by multimodal similarities.

C. Bayesian reranking:

This approach is based on bayesian framework. It maximizes the ranking score consistency among visually similar data examples while minimizes the ranking distance that represent the disagreement between the objective ranking list and the initial text based result.

D. RANGE based reranking:

This method is learning to rank based ISR method which Utilizes feature embedding algorithm to represent visual features and ranking SVM algorithm to modal the ranking function.

E. WEBSEIC Reranking:

This approach uses high density region in the image feature space to identify related images and reranking the images by combining text based and image based ranks.

F. Information bottleneck based reranking

This is an ISR approach, employs an information bottleneck Principle to find optimal clustering of images. It preserves the maximal mutual information between the search relevance and image visual features in the text search results. Then for cluster ranking cluster probability is obtained. Atlast to rerank the relevant images within each cluster KDE based on visual features are used.

G. Generalized multi-instance based Reranking

This is the multi instance learning based reranking framework, which first clusters the relevant images using both textual and visual features.then treat each cluster as an instance.then finally formulates ISR as a multi-instance learning problem to perform reranking.

H. Random Walk Re-ranking:

This method perform reranking as a random walk on the document level context graph. Context graph is a graph with nodes are representatives of documents and edges between them represents the multimodal contextual similarity between two credentials. Assume that we have N nodes which represent the video stories. The N nodes are the N credentials obtained in the initial search results. The traversing of graph is initialized from one node and based on the multimodal similarity connecting the credentials and the original text scores in the initial search result, it traverse to the next node.

I. Minimum Incremental Information Loss (MIIL) :

This video search reranking involves two parts- Learning and Reranking. In the knowledge process several query examples are provided for each textual query. This is the first pair wise approach that does visual search reranking. The example images are paired with sample images randomly selected from most important search result. Learning is used to find out which information is relevant and which information is irrelevant. Concept uncovering is processed on this example pairs to form the relevant and unrelated information.

J. Pairoff Wise Reranking:

Reranking and crowd reranking are the two methods used by pair wise reranking. Like MIIL reranking it also has two parts learning and Reranking. First we input the textual query and get the primary ranked list then this query is feed into the web search engine to get asituate of image search result.

K. Video Search Reranking:

Multimedia search [3] on the distributed sources are generally result in videos which are manifested beyond the textual modality. To exploit such contextual patterns and retain the simplicity of the keyword-based search, novel reranking methods to leverage the recurrent patterns to improve the initial text search results. Context reranking along with the context graph, where video stories are nodes and the edges between them are weighted by multimodal con- textual similarities.

L. Visual Search Reranking:

The aim of visual search reranking is to improve the text-based image search[5] with the help of analysis of visual content. The interestingness of the topic comes mainly from the fact that the search reranking is an unsupervised process and therefore has the potential to scale better than its main alternative, namely the search based on offline-learned semantic concepts.

The random walk method has few major disadvantages [3] first, it does not exactly retain the mean position of the vorticity in free space. Next, the computed solutions are noisy due to the statistical errors. The statistical errors could mask the effects of varying the control parameters. The statistical errors can also cause

symmetric flows to turn asymmetric erroneously. To reduce the statistical errors requires a very large number of vortices.

As in MILL [1] there are $N!$ Possible reranked lists for the given initial search results with the size of N . Therefore, it is impractical to compare with each other when N is extremely large. Pair off method either has limited applicability to the specific queries, the desire of the specific user interfaces, or it has the limited detection performance.

One major problem impacting performance [1] is the mismatches between the actual content of image and the textual data on the web page. The problem with these methods is the reliability of the obtained pseudo-positive and pseudo-negative images is not guaranteed.

The image reranking problem in the Bayesian framework[10] i.e. maximizing the ranking score consistency among visually similar video shots while minimizing the ranking distance, which represents the disagreement between the objective ranking list and the initial text-based. Different from existing point-wise ranking distance measures, which compute the distance in terms of the individual scores, two new methods are proposed by them to measure the ranking distance based on the disagreement in terms of pair-wise orders. Specifically, hinge distance penalizes the pairs with reversed order according to the degree of the reverse, while preference strength distance further considers the preference degree.

III. Implementation Detail

A. System Architecture:

We apply the idea of hypersphere in one-class classification to transform the original visual features. The proposed ISR method is named as H-Reranking, It has four stages:

- 1) Image feature extraction, HRPP algorithm is used for this purpose.
- 2) Images labeling, reserved K nearest neighbors i.e. KNN algorithm is used for image labeling.
- 3) To capture user's intention an interest algorithm is used.
- 4) Feature embedding with the proposed OC-HRPP algorithm, it actually transforms low level features of the image into hypersphere space.
- 5) Image search reranking with the proposed H-Rank algorithm, considering the distance to the hypersphere center for example, take the query "fish" when we submit it to a web image search engine, an initial text-based search result is returned to the user. The result is not satisfactory because of some noisy or unwanted images are also returned as top results. To improve the retrieval performance, the proposed H-Reranking method is used. Original low-level features are first extracted to represent their visual contents. And then, some relevant and irrelevant images are labeled manually or automatically for the implementation of the proposed OC-HRPP and H-Rank algorithm. Next, the embedding features are obtained with OC-HRPP algorithm, by which the images are actually transformed into a hypersphere space. Finally, the images are reordered with H-Rank algorithm by the distance to the hypersphere center. The hypersphere center is obtained with the method of support vector data description (SVDD).

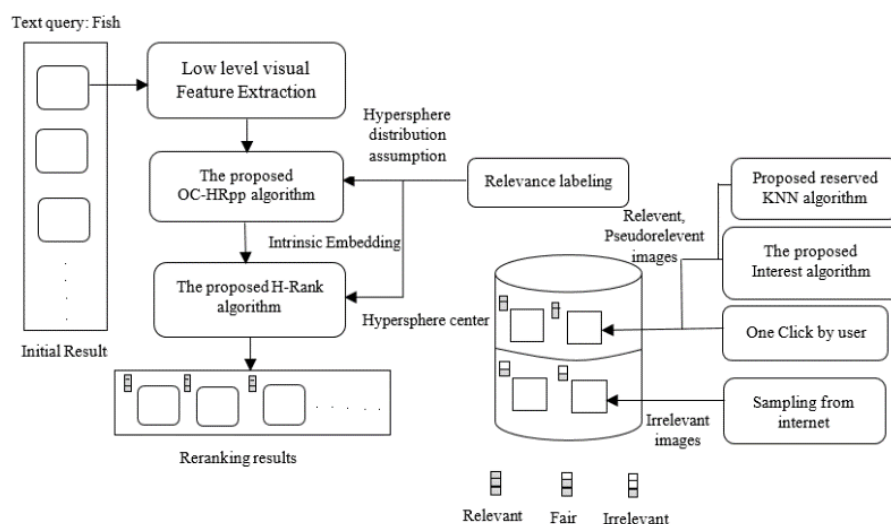


Fig.1. System architecture

B. Algorithm:

1. Take the text query as input.
2. From initial text based result extract the low level Features of the image using HRpp algorithm.
3. Using reversed KNN images are labeled as relevant, Irrelevant, fair.
4. Feature embedding is done using OC-HRpp Algorithm.
5. Interest algorithm stores interest of each and every User into their corresponding accounts.
6. Images are reordered with H-rank algorithm.

C. Mathematical Model :

Input: Text based query. for example “fish”.

Output: Reranked images.

- 1) Let I be the initial result image set.

$$I = \{i_1, i_2, i_3, \dots, i_n\}$$

Where i_1, i_2, i_3 are n number of different images from initial result image set.

2) Processing

- Feature Extraction- $F(I)$
F be the set of features of images $F = \{f_1, f_2, f_3, \dots, f_n\}$
Where f_1, f_2, f_3 are the number of extracted features of each image.
- Image labeling as R=relevant image set, S= irrelevant image set, p= pseudo relevant image set.
- Feature Embedding of R set.

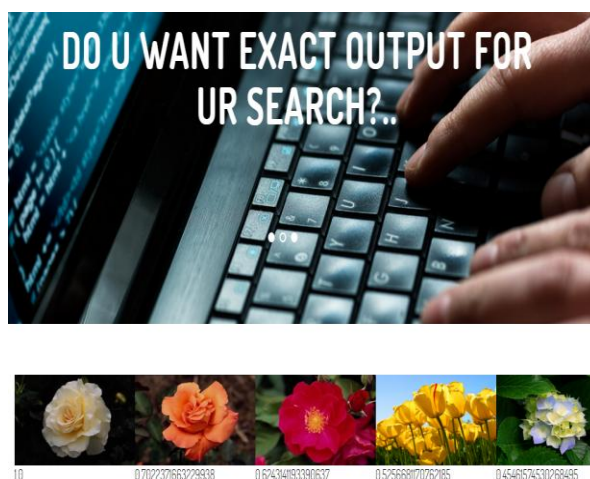
Reordering of images in set R by their distance to the hypersphere center.

3) Output

- H be the set of images obtained after feature extraction, image labeling,
- Feature embedding and reordering of images $H = \{h_1, h_2, h_3, h_4\}$.
- Where h_1, h_2, h_3, h_4 are top most relevant images obtained after reranking.

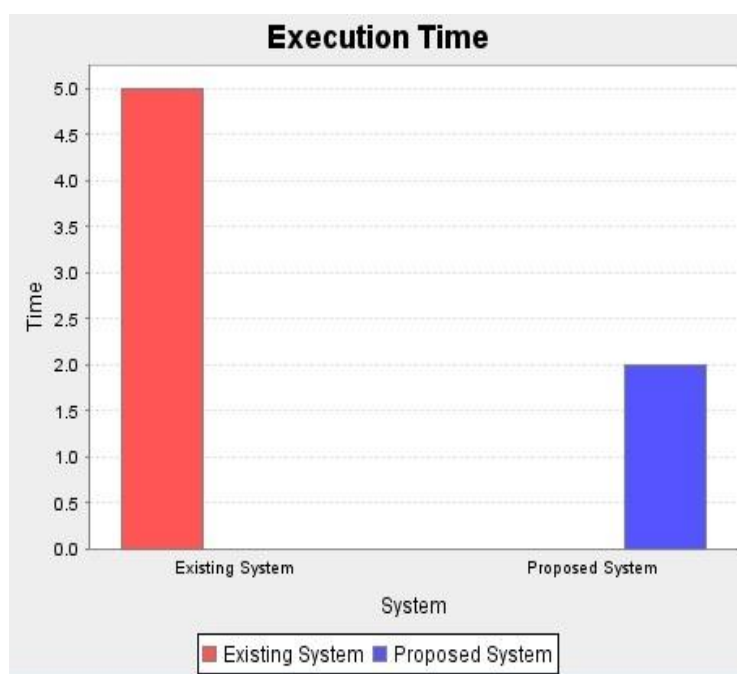
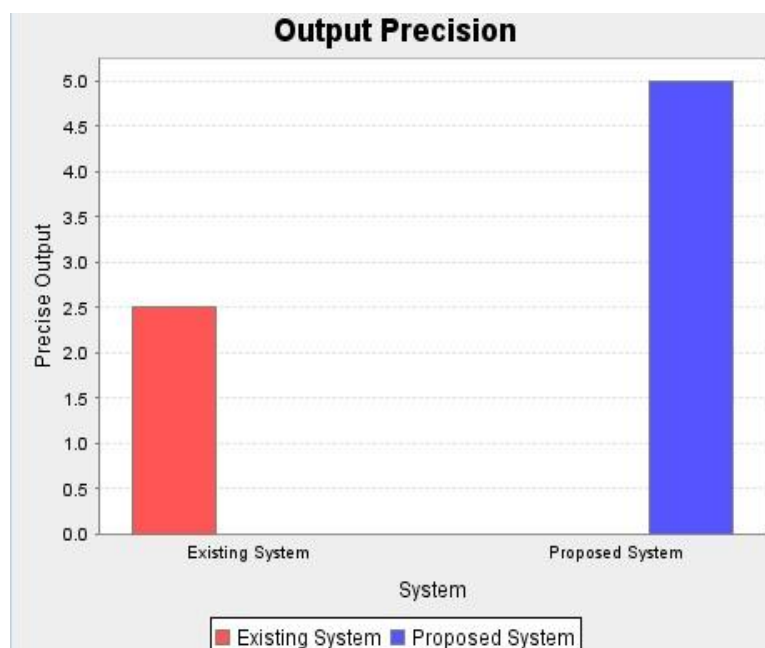
IV. Results and Discussion

In this section we demonstrate effectiveness of our system by comparing our reranking method with different reranking methods on the relatively small dataset and also analyze the effect of different parameters on the performance.



Above final result is for search category flower and user intention white flower.

Following graphs shows the Precision and Execution Time comparison of our system and existing system.



V. Conclusion

Here the reranking framework, which constructs image rerankers corresponding to visual prototypes representing the image features and learns the weights of a reranking model to combine the results of individual meta rerankers and produce the reranking score of a given image taken from the initial text-based search result. The induced reranking model is learned in a query-independent way requiring only a limited labeling effort and being able to scale up to a broad range of queries. The experimental results on the Queries dataset demonstrate that the proposed method outperforms all the existing supervised and unsupervised reranking methods. It improves the performance over the text-based search result by combining prototypes and textual ranking features. We can apply the proposed system efficiently, also the time complexity of the algorithms we are using is better than existing systems. So it will help us to get most relevant images in image search process.

Acknowledgment

Authors would like to thank the researchers as well as publishers for making their resources available and teachers for their guidance. We also thank the college authority for providing the required infrastructure and support.

References

- [1] Zhongji, Yanwei Pang and Xuelong Li, "Relevance Preserving Projection and Ranking for Web Image Search Reranking," *IEEE Trans on Image Processing* 2015.
- [2] M. Fritz and B. Schiele, "Decomposition, discovery and detection of visual categories using topic models," in Proc. CVPR, 2008.
- [3] W. H. Hsu, L. S. Kennedy, and S.-F. Chang, "Video search reranking through random walk over document-level context graph," in Proc. ACM Multimedia, 2007.
- [4] L.-J. Li and L. Fei-Fei, "OPTIMOL: Automatic online picture collection via incremental Model learning," *Int. J. Comput. Vision*, 2009.
- [5] L. Yang and A. Hanjalic, "Supervised reranking for web image search," in Proc. ACM Multimedia, 2010.
- [6] R. Fergus, L. Fei-Fei, P. Perona, and A. Zisserman, "Learning object categories from Google's image search," in Proc. ICCV, 2005, *IEEE Computer Society*.
- [7] W. H. Hsu, L. S. Kennedy, and S.-F. Chang, "Video search reranking via information bottleneck principle," in Proc. ACM Multimedia, 2006.
- [8] R. Yan, A. G. Hauptmann, and R. Jin, "Multimedia search with pseudorelevance feedback," in Proc. CIVR, 2003.
- [9] Y. Jing and S. Baluja, "Visualrank: Applying pagerank to large-scale image search," *IEEE Trans. Pattern Anal. Mach.Intell.*, Vol. 30, No.11, pp. 1877–1890, Nov. 2008.
- [10] X. Tian, L. Yang, J. Wang, Y. Yang, X. Wu, and X.-S. Hua, "Bayesian video search reranking," in Proc. ACM Multimedia, 2008.