

Image Ranking and Retrieving Analysis Using Machine Learning Approach

¹Neha Sunil Jankar, ²Dr. D. S. Bhosale

¹ Ashokrao Mane Group of Institutes, Kolhapur, Maharashtra, India

² PG Coordinator Department of Computer science Engineering, Ashokrao Mane Group of Institutes, Kolhapur, Maharashtra, India

Abstract—

The image retrieval system is used for browsing, searching and retrieving images from a large database of digital images. In retrieval, complexity of selecting a query object in single image query is high. The process of image search includes searching of images based on user given keywords. Most of the image retrieval techniques are based on text based image retrievals. But it has certain problems like images are time duplicated, low precision, and irrelevant. The purpose of this survey however, is to provide an overview of the functionality of temporary image retrieval systems in terms of technical aspects: querying, relevance feedback, features, matching measures, indexing data structures, and result presentation. We have reviewed two different techniques text based retrieval and rank using user clicks.

Keywords— Text, click, ranking, Learning, optimization

I. INTRODUCTION

The development of the multimedia technique and multimedia database are use to in image databases. The availability of digital images motivates research into automatic image retrieval. Image retrieval could be based on metadata or content. Most common methods of image retrieval search the images using associated metadata such as keywords and text. Another method is called as User clicks are integrated to textual features to make refinement of textual query. There are many ranking methods used for ranking the result namely Pair wise approach, List wise approach. Most of Proposed systems use web image database and classifiers to satisfy users demand for proper relevant image retrieval. Some of the systems used popular Word net ontology for automatically retrieved images which are relevant and irrelevant using inverted file method. Image database is increasing day by day, because searching images from large and diversified collection using image features as information is difficult and imperative problem. Image search is an important feature widely used in majority search engines, but the search engine mostly employs the text based image search. [1] The second approach to learning retrieval functions by analysing which links the users click on in the presented ranking. [2]

II. RELATED WORK

An approach that has proved extremely successful for document retrieval is learning to rank [4], where a ranking function is learnt, given either the pair wise preference relations or relevance levels of the training examples. Similar methods have also been proposed for ranking images [5].

There are many ranking methods used for ranking the result namely Pair wise approach, List wise approach:

The pair wise approach [5][6] minimizes above problem and successfully used in document retrieval. This approach collects document pairs from the ranking lists, and assigns a label to each pair that describes the relative relevance of the two documents. It then trains a classification model with the labeled data and adopts it for ranking.

List wise approaches [7][8] have been proposed to learn a ranking function by adopting separate lists as samples is explored to optimize errors in the classification of document pairs. This approach learns a ranking function by taking individual lists as instances and minimizing a loss function defined on the predicted list and the ground-truth list. Here the loss function is formulated on the predicted list and the ground-truth list.

III. TEXT BASED IMAGE RETRIEVAL SYSTEM

An image retrieval system is defined as a computer system for browsing, searching and retrieving images from a large database of digital images. These systems are useful in vast number of applications like engineering, fashion, travels and tourism, architecture etc. Because of the relative ease in understanding and processing text, commercial image-search systems often rely on techniques that are largely indistinguishable from text search. Thus we need a powerful image search engine which will organize and index the images available on web or large database in proper format. The image retrieval that is based on artificial notes labels images by using text firstly, in fact it has already changed image retrieval into traditional keywords retrieval [3]. The following figure shows the conventional approach for text based image retrieval system. The certain problems like images are time duplicated, low precision, and irrelevant are faces Text based image retrieval system. To avoid this drawback and rank the flexible image retrieval system need to use ranking using user click.

Conventional Approaches



Figure 1. Conventional approach for text based image retrieval system.

Proposed Approach

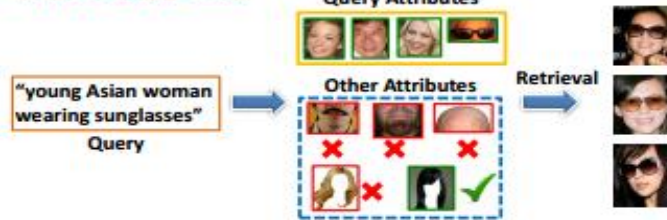


Figure 2. Sematic approach for text based image retrieval system.

IV. RANK USING USER CLICK MODEL

Click through data can be recorded with little overhead and without compromising the functionality and usefulness of the search engine. In particular, compared to explicit user feedback, it does not add any overhead for the user [2]. The rank user click model used to develop image search and ranking model for search engine by integrating user click features and visual features of the image. The system presented here can be used with present search engines that include the retrieval of images when user specifies some textual query. The images are searched from web and top-k results are displayed to user. At the first the accuracy of result obtained from search engine can be refined by integrating user clicks with previously given query text. Second approach is to generate image features like image hyper graph and then results are re-ranked by combining user clicks and image features.

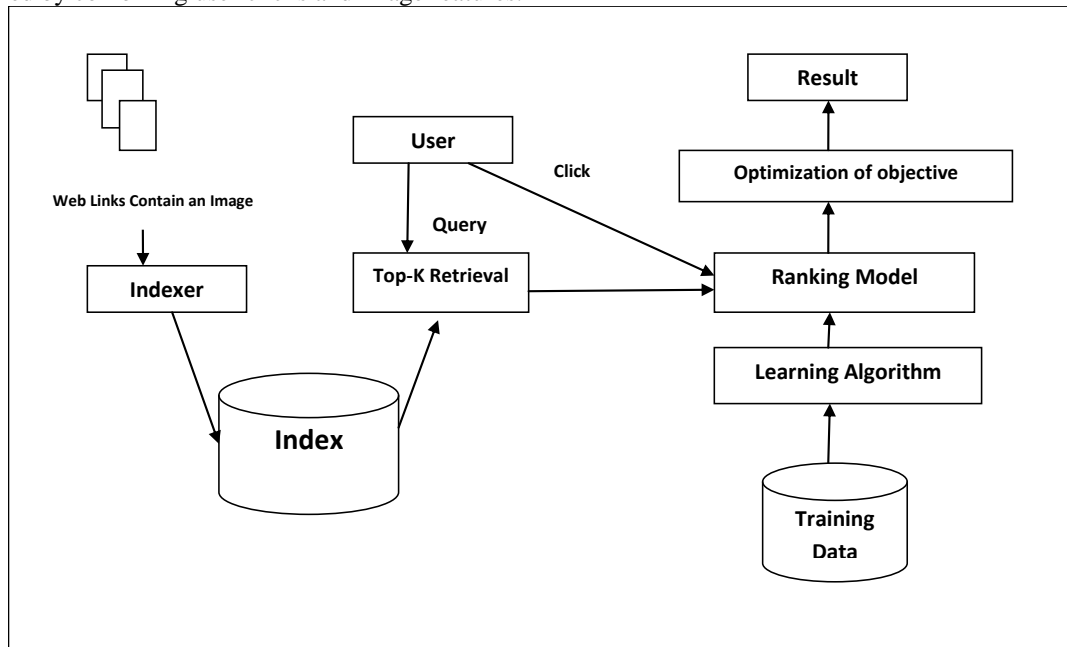


Figure 3: System Architecture for Rank using user click model

The figure shows architecture for image search and re-ranking includes same architecture as information retrieval system. For Image search user gives the textual query describing image, the search engine process the query and retrieves relevant web links that contains required result. The Indexer plays important role in the retrieval system as it provides indexing for retrieved links based on their popularity and arranged in descending order. After that only top-k results are shown to user.

This traditional image search will not perform well because textual information cannot describe accurate semantics of query. To achieve good image retrieval results machine learning approach is used. This learning model first performs training on dataset and forms a pair of result that gives query vector and relevance vector. The learning algorithm learns image features and click feature provided by user clicks. By integrating the visual features and click features, objective function is modified and hyper graph regularize and linear model are, respectively considered for two features. The combination of user clicks and image features produce more relevant results. The optimization of obtained result will be carried out by fast alternating linearization method (FALM). This method can alternately minimize two different approximations of the original objective function by keeping one function unchanged and linearizing the other.

V. CONCLUSIONS

In this paper we have reviewed and analyzed two different methods with data sets used to retrieve images. As considering results of different methods, we conclude that for better retrieval performance we must used the techniques to increase values of parameters like time duplicated, low precision, and irrelevant and Accuracy, which may lead to better results of retrieval performance

ACKNOWLEDGMENT

Support from work received from Department of Computer Science & Engineering of Ashokrao Mane Group of Institute, Vathar, Kolhapur, and Maharashtra, India.

REFERENCES

- [1] Y. Jing and S. Baluja, "VisualRank: Applying PageRank to Large-Scale Image Search", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 30, No. 11, 2008.
- [2] Joachims, 2002] Thorsten Joachims. Optimizing search engines using clickthrough data. In KDD, pages 133–142, 2002.
- [3] R. Brunelli and O. Mich, "Histograms Analysis for Image Retrieval," Pattern Recognition, Vol.34, No.8, pp1625–1637,2001.
- [4] H. Li, "Learning to rank for information retrieval and natural language processing," Synth. Lect. Human Lang. Technol., vol. 4, no. 1, pp. 1– 113, 2011.
- [5] S. Robertson and S. Walker, "Some simple effective approximations to the 2-poisson model for probabilistic weighted retrieval," in Proc. Annu. Int. ACM SIGIR Conf. Res. Dev. Inf. Retrieval, 1994, pp. 232–241.
- [6] K. Jarvelin and J. Kekalainen, "Cumulated gain-based evaluation of IR techniques," ACM Trans. Inf. Syst., vol. 20, no. 4, pp. 422–446, 2002.