International Journal of Scientific Engineering and Research (IJSER)
www.ijser.in
ISSN (Online): 2347-3878

Review for Feature Based Image Re-Ranking

Nishakumari Ashokkumar Lodha[1], Minakshi Somanath Bagad[2]

[1]Assistant Professor, Department of Computer Science, G.H.R.I.E.M, Jalgaon, Maharashtra, India

[2]M.E (C.S.E) Student, Department of Computer Science, G.H.R.I.E.M, Jalgaon, Maharashtra, India

Abstract: ***Search reranking is considered as a best and common way to improves retrieval precision. The images are retrieved using the associated textual information, such as surrounding text from the web page. The performance of such systems mainly relies on the relevance between the text and the images. However, they may not always match well enough, which causes noisy ranking results. For instance, visually similar images may have very different ranks. So reranking has been proposed to solve the problem. Image re-ranking, as an effective way to improve the results of web-based image search however the problem is not trivial especially when we are considering multiple features or modalities for search in image and video retrieval. This paper suggests a new kind of reranking algorithm, the circular reranking, that supports the mutual exchange of information across multiple modalities for improving search performance and follows the philosophy of strong performing modality could learn from weaker ones.***

Keywords: Image reranking, Image retrieval, Modality, Visual search

1. Introduction

Searching for relevant images from large scale community databases given a query term is an important task. The image ranking approach represents an image collection as a graph that is built using multimodal similarity measures based on visual features and user tags. To improve the performance of this image search image re-ranking technology is used. Search re-ranking is regarded as a common way to boost retrieval precision. The problem nevertheless is not trivial especially when there are multiple features or modalities to be considered for search, which often happens in image and video retrieval. Different re-ranking algorithms are available in computer world which gives different precisions. Formally; the definition of the re-ranking problem with a query image is as follows. The re-ranking process is used to improve the search accuracy by reordering the images based on the multimodal information extracted from the initial text-based search results, the auxiliary knowledge and the example image. The auxiliary knowledge can be the extracted visual features from each image or the multimodal similarities between them.

1.1 Example image based reranking

Lui et al. proposed a new re-ranking scheme [9]: after query by keyword, user can click on one image, which is the image desired by the user. Then the image search engine re-ranks the images according this query image: those that are visually similar to query images are top ranked. However, this method is based on direct comparison of the example image and each image in the ranking list. Therefore, noisy results usually appear.

1.2 Graph-based Semi-supervised Learning [7]

In this re-ranking first, the images returned by a text-based search engine are re-ranked according to their distances to the query image, and the distances are used as the initial ranking scores. Second, a graph-based semi-supervised Learning algorithm is applied to propagate the scores between images. However it may produce noisy result. The final scores have the following properties: (1) they are consistent across visually similar images (2) they are close to the initial scores (3) the example query images have high scores. This problem can be formulated as graph-based semi-supervised learning.

Three different dimensions are considered for image re-ranking: self-reranking, crowd-reranking by exploiting online crowd sourcing knowledge, and example-based reranking by leveraging user-provided queries.

1.3 Co-Reranking

Co-reranking for image search [4] jointly explores the visual and textual information. Co-reranking couples two random walks, while reinforcing the mutual exchange and propagation of information relevancy across deferent modalities. The mutual reinforcement is iteratively updated to constrain information exchange during random walk. As a result, the visual and textual reranking can take advantage of more reliable information from each other after every iteration.

1.4 Self-reranking

It aims to improve the initial performance by only mining the initial ranked list without any external knowledge [1], [16], [18]. For example, Hsu et al. formulate the reranking process as a random walk over a context graph, where video stories are nodes and the edges between them are weighted by multimodal similarities [16]. Fergus et al. first perform the visual clustering on initial returned images by probabilistic Latent Semantic Analysis (pLSA), learn the visual object category, and then rerank the images according to the distance to the learned categories [1].

1.5 Example-reranking

This dimension of reranking leverages a few query examples (e.g., images or video shots) to train the reranking models [13]. The search performance can be improved due to the external knowledge derived from these examples. For example, Yan et al. and Schroff et al. view the query examples as pseudo-positives and the bottom-ranked initial results as pseudo-negatives [13]. A reranking model is then built based on these samples by Support Vector Machine (SVM). Liu et al. use the query examples to discover the relevant and irrelevant concepts for a given query, and then identify an optimal set of document pairs via an information theory [13]. The final reranking list is directly recovered from this optimal pair set.

1.6 Crowd-reranking

It is characterized by mining relevant visual patterns from the crowd sourcing knowledge available on the Internet. For example, a recent work first constructs a set of visual words based on the local image patches collected from multiple image search engines, explicitly detects the so-called salient and concurrent patterns among the visual words, and then theoretically formalizes the reranking as an optimization problem on the basis of the mined visual patterns [9]. However, it is observed that most of existing reranking methods mainly exploit the visual cues from the initial search results.
Various visual search re-ranking methods are as follow

1.7 Traditional Image Re-ranking

Major web image search engines have adopted the strategy which works as given a query keyword input by a user, a pool of images relevant to the query keyword are retrieved by the search engine according to a stored word-image index file. By asking the user to select a query image, which reflects the user's search intention, from the pool, the remaining images in the pool are re-ranked based on their visual similarities with the query image. The word-image index file and visual features of images are pre-computed offline and stored [15], [2]. Visual features must be saved. The web image collection is dynamically updated. If the visual features are discarded and only the similarity scores of images are stored, whenever a new image is added into the collection and we have to compute its similarities with existing images, then the visual features need to be computed again[4].
The main online computational cost is on comparing visual features. To achieve high efficiency, the visual feature vectors need to be short and their matching needs to be fast.

1.8 Click Boosting

The Click Boosting technique which is a straightforward way of re-ranking search results based on click data [1]. This technique promotes all of the clicked images, sorted in descending order according to the number of clicks, to the top. The original ranking is used to break ties as well as to rank all images that have not been clicked.

1.9 Gaussian Process Re-ranking using Click Data

This algorithm works as follows. Once a query has been issued, the top thousand results from the baseline search engine are retrieved and features are extracted. We then identify the set of clicked images and perform dimensionality reduction on all the feature vectors. A Gaussian Process regressor [1] is trained on the set of clicked images and is then used to predict the normalized click counts (pseudo-clicks) for all images. Re-ranking is then carried out on the basis of the predicted pseudo-clicks and the original ranking score.

1.10 Circular reranking

The basic idea of circular reranking 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. Circular reranking takes advantages of both pattern mining and multi-modality fusion for visual search. More importantly, modality interaction is taken into account, on one hand to implicitly mine recurrent patterns, and on the other, to leverage the modalities of different strength for maximizing search performance.

2. Literature survey

Wei et al, [3], proposed a concept-driven multi-modality fusion (CDMF), explores a large set of predefined semantic concepts for computing multi-modality fusion weights in a novel way. In CDMF, the query-modality relationship is decomposed into two components that are much easier to compute: query-concept relatedness and concept-modality relevancy.

Fergues et al,[5], 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 reranked based on the distance to the mined categories. Self-reranking seeks consensus from the initial ranked list as visual patterns for reranking.

Richter et al,[7], employed an crowd-reranking is similar to self-reranking except that consensus is sought simultaneously from multiple ranked lists obtained from Internet resources and further formulated the problem as random walk over a context graph built through linearly fusing multi-modalities for visual search.

Tan et al, [8], proposed an agreement-fusion optimization model for fusing multiple heterogeneous data. The leveraged rank agreement mined from multiple lists iteratively to update the weights of modalities until reaching an equilibrium stage.

Liu et al. [9] suggested a reranking 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 reranking.

Kennedy et al, [12] proposed a query class dependent search models in multimodal retrieval for the automatic discovery of query classes. This scheme starts by predefining query classes, then learning of weights in offline conducted on the query class level. During search, a given query is routed into one of the predefined classes, and the learnt weights are directly applied for fusion.

Hsu et al, [18], employed information bottleneck (IB) reranking to find the clustering of images that preserves the maximal mutual information between the search relevance and visual features. Multi-modality fusion based on weighted linear fusion is widely adopted. Broadly, we can categorize the existing research into adaptive [15], and query-class-dependent fusion [9].

Wilkins et al, [20], proposed a multi-modal data for video Information Retrieval, models the change of scores in a list to predict the importance of a modality. Specifically, the gradual (drastic) change of scores indicates the difficulty (capability) of a modality in distinguishing relevant from irrelevant items, and fusion weights are thus determined accordingly.

3. Conclusion and Future Work

This paper presents a survey on various Reranking algorithms that were proposed by earlier researches for the better development in the field of Image Processing. Various algorithms and methods discussed above will help in developing efficient and effective re-ranking for image processing. In the future scope, a comparative study of various algorithms will be presented for circular re-ranking. Circular re-ranking provides information exchange and reinforcement for visual search re-ranking for images. Particularly, the placement of modalities in the circular framework which could lead to the highest possible retrieval gain in theory for search re-ranking.

References

[1] Vidit Jain and Manik Varma, “Larning to rerank query Dependent Image Re-rnking using Click Data” 2011
[2] C.A. Bhatt and M.S. Kankanhalli, “Multimedia Data Mining: State of the art challenges,” Multimedia tools Appl. Vol.51, no.1, pp.35-76, 2011.
[3] X-Y. Wei, Y-G. Jiang and C.W. Ngo, “Concept Driven Multi-modality fusion for video search,” IEEE Trans. Circuits syst. Vide Technol., vol.21, no.1, pp. 62-73, Jan. 2011
[4] T. Yao, T. Mei and C.W. Ngo, “Co-reranking by mutual Reinforcement for image search,” ACM Int. Conf. Image Video Retr.2010, pp. 34-41
[5] R. Fergus, L. Fei-Fei, P. Perona and A. Zisserman, “Learning object categories from Google’s image Search,” in Proc. IEEE Int. Conf. Comput, Vis Oct.2005, pp.1816-1823
[6] P.K. Atrey, A. Hossain, A.E. Saddik and M.S. Kankahalli, “Multimodal Fusion for Multimedia Analysis: A Survey,” Springer Multimedia Syst., vol. 16 No.6, pp. 345-379,200.
[7] F. Richter, S.Romberg, E. Horster and R. Lienhart, “Muti-modal Ranking for Image Search on community Database,” ACM SIGMM Int. Workshop Multimedia Inf. Retr., 2010 pp.63-72
[8] H.K. Tan and C.W. Ngo, “Fusing Heterogenous Modalities for Video and image Re-ranking,” ACM Int.Conf. Multimedia Ret., 2011 pp, 15, TRECVID
[9] Y. Liu, T. Mei and X.-S. Hua, “Crowd Reranking Exploring Multiple Search Engines for Visual Search Reranking,” in Proc. ACM Special Interested Group Inf. Retr. 2009, pp.500-507
[10] M. Wang, L. Yang and X.S.Hua, “MSRA-MM: Bridging Rengtonsearch and Tech. industrial societies for Multimedia information retrieval, “Microsoft Washington, DC, Rep. MSR-TR-2009-30, 2009
[11] Y.S. Jing and S. Baluja, “Visual rank: Applying page rank to large-scale image search,” Trans. Pattern Anal. Mach. Intel., vol.30 no. 11. pp. 1877-1890 Nov.2008
[12] L. Kennedy, S.F. Chang and A. Natsev, “Query adaptive Fusion for Multimodal Search,” Proc. IEEE vol.96, no.4 pp. 567-588, Apr. 2008
[13] Y.Liu, T. Mei, X.Wu and X.-S. Hua’ “Optimizing video search reranking via minimum incremental information loss,” in Proc ACM Int. Workshop Multimedia Inf. Retr., 2008, pp. 253-259
[14] D. Zhou, S.A. Orshanskiy, H.Zha and C.l. Giles, “Co-ranking authors and documents in a heterogeneous network,” IEEE Int.Conf. Data Mining, Dec.2007, pp.739-744
[15] W. Hsu, L. Kennedy and S.F. chang, “Reranking methods for visual search” IEEE Multimedia, vol.14, no. 3, pp.14-22, Jul.-Sep.2007
[16] W. Hsu, L. Kennedy and S.F. Chang, “Video Search re-ranking through random Walk over document–level Context graph,” ACM int. Conf. Multimedia, 2007, pp. 971-980
[17] P. Wilkins, T. Adamek, G.J.F. Jones, N.E.O’ Connor and A.F. Smeaton “Trecvid2007 Experiments at Dublin City University,” NIST TRECVID Workshop, 2007, pp.1-8
[18] W. Hsu, L. Kennedy and S.F. Chang, “Video Search Reranking via Information Bottleneck Principle,” ACM Int. Conf. Multimedia, 2006, pp. 35-44
[19] A.F. Smeaton, P. Over and W. Karaaji, “Evaluation campaigns and trecvid,” ACM SIGMM Int. Workshop Multimedia Inf. Retr., 2006, pp.321-330
[20] P. Wilkins, P. Ferguson and A.F. Smeaton, “Using Score distribution for query–time fusion in Multimedia retrieval,” ACM SIGMM Int. Workshop Multimedia Inf. Retrieval 2006, pp.51-60