International Journal of Scientific & Engineering Research, Volume 7, Issue 2, February-2016
ISSN 2229-5518

Click Prediction For Web Image Reranking Using Multimodal Sparse Coding

Ankush R. Deshmukh, Asst. Prof. Pushpanjali M. Chouragade

Abstract—For the improvement of the performance of a text-based image search, Image reranking is a effective method. There are two reasons for which the reranking algorithms are limited and they are: One is that the data that is associated with images is not matched with the actual visual content and the second reason is that the reextracted visual features do not accurately describe the meaningful similarities between images. The relation of retrieved images to search queries has been more precisely described by user clicks, in recent years. However, a the lack of click is the data critical problem for click-based methods, since users have clicked a small number of web images. Therefore, the solution to this problem is by predicting image clicks. A multimodal hypergraph learning-based sparse coding method is proposed for image click prediction, and apply click data that has been obtained to the reranking of images. To build a group of manifolds, a hypergraph is adopted. A hyperedge present in a hypergraph is the edge that connects a set of vertices, and preserves the constructed sparse codes. The weights of different modalities and the sparse codes are obtained by an alternating optimization procedure. Finally, to describe the predicted click as a click or no click, a voting strategy is used from the images that was corresponding to the sparse codes. Image reranking algorithms are used to improve the performance of graph-based the use of click prediction is shown by an additional image reranking experiments on real world data that is beneficial.

Index Terms— sparse code, Image reranking, manifolds, click.

1 Introduction

In image reranking, both textual and visual information are combined in order to return improved results to the user. Pseudo-relevance-feedback (PRF) is the most existing re-ranking method tool where a proportion of the toped ranked images are assumed to be relevant, and subsequently used to build a model for reranking. Graph based re-ranking and Baysian based re-ranking promotes low-rank images by receiving reinforcement from related high-rank images, but are limited because irrelevant high-rank images are not demoted causing both implicit and explicit re-ranking methods to suffer from the unreliability of the original re- ranking list, since the textual information cannot accurately describe the semantics of the queries.

Sparse coding is a widely used signal processing method and performs well in applications for e.g. signal decomposition, signal reconstruction and signal denoising. Although statistically unrelated bases like Fourier or Wavelets have been widely adopted, an overcomplete basis has been adopted the latest trends, in which the number of basis vectors is greater than the dimensionality of the input vector. A signal can be described by a set of bases using a very small number of nonzero elements. Many applications need this compact representation of signals. Sparse coding and image features, signals are is adopted as an efficient technique for feature reconstruction in computer vision. It has been mostly used in many different applications, such as, face recognition, image classification , image restoration and image annotation.

In this paper, the problem of click prediction through sparse coding is solved. Based on a group of web images with clicks known as a codebook, and a new image without any clicks, sparse coding is used to choose as few basic images as possible from the codebook in order to linearly reconstruct a new input image while the reconstruction errors are minimized. A voting strategy is utilized to predict the click as a binary event. The over complete characteristics of the codebook gives the guarantee that the sparsity of the reconstruction coefficients. However, in addition to sparsity, the over completeness results in similar web images that has being described by fully distinct sparse codes, and unstable performance in image reconstruction; clicks are thus not successfully predicted. To solve this issue, to add an additional locality preserving term is one of the possible solution to the formulation of sparse coding for e.g. Laplacian sparse coding (LSC), in which a locality-preserving Laplacian term is added to the sparse code.

2 Previous Work Done

Detailed submission Existing reranking algorithms were limited for two main reasons: 1) the textual meta-data associated with images often mismatched with their actual visual content and 2) the extracted visual features do not accurately describe the meaningful similarities between images.

- **Multimodal learning for web images**:

[C. Snoek 2005] proposed the methods of multimodal feature fusion classified into two categories, namely early fusion and late fusion. It has been shown that if an SVM classifier is used, late fusion tends to result in better performance. [M. Wang 2009] provided a method to integrate graph representations generated from multiple modalities for the purpose of video annotation [B. Geng 2009] have integrated graph representations using a kernelized learning approach. This approach integrates multiple features into a graph-based learning algorithm for click prediction.

- **Graph-based learning methods**

Graph-based learning methods have been widely used in the fields of image classification, ranking and clustering. In these methods, a graph is built according to the given data, where vertices represent data samples and edges describe their similarities. The Laplacian matrix [M. Belkin 2003] is constructed from the graph and used in a regularization scheme. The local geometry of the graph is preserved during the optimization, and the function is forcefully smoothed on the graph. However, a simple graph-based method cannot capture higherorder information. Unlike a simple graph, a hyperedge in a hypergraph links several (two or more) vertices, and thereby captures this higher-order information. Hypergraph learning has achieved excellent performance in many applications.[R. Zass 2008] utilized the hypergraph for image matching using convex optimization. Hypergraphs have been applied to solve problems with multilabel learning [L. Sun 2008] and video segmentation. [Z. Tian 2009] provided a semi-supervised learning method named HyperPrior to classify gene expression data, by using biological knowledge as a constraint. [Y. Huang 2010] proposed a hypergraph-based image retrieval approach.

3 Literature Survey

[M. Wang 2008] proposed Optimizing multigraph learning: Towards a unified video annotation scheme. The description of learning with hypergraphs: clustering, classification and embedding proposed by [D. Zhou 2006] presented the overview of hypergraph utility on reranking method. [R. Zass 2008] utilized the hypergraph for image matching using convex optimization. Hypergraphs have been applied to solve problems with multilabel learning [L. Sun 2008] and video segmentation. [M. Wang 2009] provided a method to integrate graph representations generated from multiple modalities for the purpose of video annotation [B. Geng 2009] have integrated graph representations using a kernelized learning approach.

4 Proposed Methodology

IJSER style is to In this paper, a novel method named multimodal hypergraph learning-based sparse coding is proposed for click prediction, the predicted clicks to re-rank web images have been applied. Both strategies of early and late fusion of multiple features are used in this method through three main steps.

- A web image base with associated click annotation is constructed, collected from a commercial search engine which records clicks for each image such that the images with high clicks are strongly relevant to the queries. These two components form the image bases.
- The proposed objective function considers both early and late fusion. The early fusion directly concatenates the multiple visual features, and is applied in the sparse coding term. Manifold learning term accomplishes in the late fusion. For web images without clicks, hypergraph learning is implemented to construct a group of manifolds, thus preserving the smoothness locally using hyperedges. Unlike a graph that has an edge between two vertices, a set of vertices are connected by the hyperedge in a hypergraph. Common graph-based learning methods only consider the pairwise relationship between two vertices, and ignores the higher-order relationship among three or more vertices. Using this term can help the proposed method preserve the local smoothness of the constructed sparse codes.

An alternating optimization procedure is conducted to expansion of proving complement nature of various modalities. To predict if an input image will be clicked or not, based on its sparse code a voting strategy is then adopted. The obtained click is then integrated within a graph-based learning framework to achieve image re-ranking.

5 Implementation of Proposed Methodology

If you are using Word The important contributions:

- The search engine derived images are effectively utilized annotated with clicks, and successfully predict the clicks for new input images which are not with clicks. Depending upon the obtained clicks, the images are reranked, a strategy which could be beneficial for improving commercial image searching.
- A novel method named multimodal hypergraph learning-based sparse coding is proposed. Both early and late fusion are used by this method in multimodal learning. By simultaneously learning the performance of sparse coding performs significantly.
- Comprehensive experiments are conducted to empirically analyze the proposed method on real-world web image datasets, collected from a commercial search engine. Internet users clicks their corresponding click. The effectiveness of the proposed method has been demonstrated.

Please note that math equations might need to be reformatted from the original submission for page layout reasons. This includes the possibility that some in-line equations will be made display equations to create better flow in a paragraph. If display equations do not fit in the two-column format, they will also be reformatted. Authors are strongly encouraged to ensure that equations fit in the given column width.

5.1 Tools Used

Search engines like Google, Live and Yahoo

6.2 Advantages

- The re-ranking approach conducted on a real world data set not only states the usefulness of click-through data, which is not only viewed as the footprints of user behavior, in understanding user intention, but it also verifies the importance of query dependent fusion weights for multiple modalities.
- It performs better than other modality methods.

6.3 Disadvantages

In this paper only image search relevance is taken into consideration while image diversity is to be discovered further.

7 Results

Experiments conducted on a real world data set not only describes the usefulness of click-through data, which can be viewed as theimage of an user behavior, in understanding user intention, but also verify the importance of query dependent fusion weights for multiple modalities.Based on a gradient method, a proper combination of modality weights is learnt adaptively and query-dependently.

Conclusion

Image diversity is a factor in search performance by enhancing the diversity of re-ranked images by dupliAn image search rerankingalgorith is presented called click-based relevance feedback, by exploring the use of click through data and the function of multiple modalities. Based on the gradient method, a proper combination of modality weights are described adaptively and query-dependently.

Future Work

Image diversity is a factor in search performance by enhancing the diversity of re-ranked images by duplication detection or other applicable method.

References

[1] J.S. Bridle, "Probabilistic Interpretation of Feedforward Classification Network Outputs, with Relationships to Statistical Pattern Recognition," *Neurocomputing – Algorithms, Architectures and Applications,* F. Fogelman-Soulie and J. Herault, eds., NATO ASI Series F68, Berlin: Springer-Verlag, pp. 227-236, 1989. (Book style with paper title and editor)

[2] Y. Gao, M. Wang, Z. J. Zha, Q. Tian, Q. Dai, and N. Zhang, "Less is more: Efficient 3D object retrieval with query view selection," IEEE-Trans. Multimedia, vol. 13, no. 5, pp. 1007–1018, Oct. 2011.

[3] S. Clinchant, J. M. Renders, and G. Csurka, "Trans-media pseudorelevance feedback methods in multimedia retrieval," in Proc. CLEF,2007, pp. 1–12.

[4] L. Duan, W. Li, I. W. Tsang, and D. Xu, "Improving web image searchby bag-based reranking," IEEE Trans. Image Process., vol. 20, no. 11,pp. 3280–3290, Nov. 2011.

[5] B. Geng, L. Yang, C. Xu, and X. Hua, "Content-aware Ranking forvisual search," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit.,Jun. 2010, pp. 3400–3407.

[6] B. Carterette and R. Jones, "Evaluating search engines by modeling therelationship between relevance and clicks," in Proc. Adv. Neural Inf.Process. Syst., 2007, pp. 1–9.

[7] G. Dupret and C. Liao, "A model to estimate intrinsic document

[8] relevance from the clickthrough logs of a web search engine," in Proc.ACM Int. Conf. Web Search Data Mining, 2010, pp. 181–190.

[9] G. Smith and H. Ashman, "Evaluating implicit judgments from imagesearch interactions," in Proc. WebSci. Soc., 2009, pp. 1–3.

[10] M. Shokouhi, F. Scholer, and A. Turpin, "Investigating the effectivenessof clickthrough data for document reordering," in Proc. Eur. Conf. Adv.Inf. Retr., 2008, pp. 591–595.

[11] V. Jain and M. Varma, "Random walks on the click graph," in Proc.ACM SIGIR Conf. Res. Develop. Inf. Retr., 2007, pp. 239–246.

[12] V. Jain and M. Varma, "Learning to re-rank: Query-dependent imagere-ranking using click data," in Proc. Int. Conf. World Wide Web, 2011,pp. 277–286.

[13] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery fromincomplete and inaccurate measurements," Commun. Pure Appl. Math.,vol. 59, no. 8, pp. 1207–1223, 2006.

[14] A. Eriksson and A. van den Hengel, "Efficient computation of robustlow-rank matrix approximations in the presence of missing data usingthe l1 norm," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit.,Jun. 2010, pp. 771–778.