ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 3, Issue 11, November 2016

A Hybrid Approachfor Web Image Reranking

[1] Shah Shraddha R [2] Nitin N. Patil
R. C. Patel Institute of technology

Abstract: **-- For the improvement of the performance of a text-based image search, Image reranking is a efficientmethod. There are two reasons for which the reranking algorithms are limited and they are: One is that the data that is connectedwith images is not coordinatedwith the actual visual content and the second reason is that the reextracted visual features do not exactlydescribe the consequentialsimilarities between images. The relativeof retrieved images to explorequeries has been more correctlydescribed by user clicks, in recent years. However, the lack of click is the data critical problem for click-based methods, since users have clicked a small number of web images.Consequently, the solution to this problem is byguess image clicks. A multimodal hypergraph learning based sparse coding method is proposed for image click prediction, and be relevantclick data that has been obtained to the reranking of images. To build a group of manifolds, a hypergraph is adopted. A hyperedge exist in a hypergraph is the edge that connects a set of vertices, and conservethe constructed sparse codes. The weights of dissimilarmodalities and the sparse codes are obtained by an irregularoptimization procedure. Finally, to describe the predicted click as a click or no click, a voting strategy is used from the images that was matchingto the sparse codes. Image reranking algorithms are used to progressthe performance of graph-based the use of click prediction is exposedby an supplementaryimage reranking experiments on real world data that is useful.**

***Keywords*:-- Image Reranking, Click based method, Manifold Learning, Multimodal hypergraph learning, sparse code, Pseudo relevance feedback**

I. INTRODUCTION

In image reranking, both textual and visual in sequenceare collectivein order to return enhancedresults to the user. Pseudo-relevance-feedback (PRF) is the most existing re-ranking method tool where a quantityof the toped ranked images are assumed to be relatedand subsequently used to build a model for reranking [1]. Graph based re-ranking and Baysian based re-ranking promotes low-rank images by gettingsupportfrom correlatedhigh-rank images, but are restrictedbecause irrelevant high-rank images are not demoted causing both inherentand unambiguousre-ranking methods to go throughfrom the unreliability of the innovative reranking list, since the textual information cannot accurately describe the semantics of the queries. Sparse coding is a extensivelyused signal processing method and performs well in applications for e.g. signalbreakdown, signal restorationand signal denoising. Although statistically unrelated bases like Fourier or Wavelets have been generallyadopted, an over entirebasis has been adopted the newesttrends, in which the numeralof basis vectors is greater than the dimensionality of the input vector. A signal can be described by a set of basisusing a very small number of nonzero elements. Many applications requirethis condensedillustrationof signals. Sparse coding and image features, signals are is adopted as an well-organizedtechnique for feature restorationin computer vision. It has been mostly used in many different applications, such as, facedetection, image classification, image restoration and image annotation [1].

In this paper, the difficultyof click prediction throughoutsparse coding is solved. Based on a group of web images with clicks recognizedas a codebook, and a new image without any clicks, sparse coding is used to preferas few basic images as possible from the codebook in regulateto 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 [2]. 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 unbalanced performance in image reconstruction; clicks are thus not successfully predicted. To resolvethisproblem, 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].

II. PREVIOUS WORK DONE

Completesubmission existing reranking algorithms were restrictedfor two main reasons: 1) the textual meta-data associated with images often unequalwith their actual visual content and 2) the extracted visual features do not exactlydescribe the meaningful similarities between images [3].

A. Multimodal learning for web images

Snoekhave proposed the methods of multimodal feature fusion classified into two categories, namely early fusion and late fusion. It has been exposedthat if an SVM classifier is used, late synthesistends to result in better performance [2]. Wang have provided a techniqueto incorporategraph representations generated from manymodalities for the reason of video annotation B. Geng have integrated graph representations by means ofa kernelized learningtechniqueThis method integrates many features into a graph-based learning algorithm for click prediction [3].

B.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 illustrate their similarities. The Laplacian matrix M. Belkin is constructed from the graph and used in a regularization scheme. The local geometry of the graph is conservedthroughthe optimization, and the function is forcefully smoothed on the graph. However, a simple graph-based method cannot confinehigherorder information. Unlike a simple graph, a hyperedge in a hyper-graph links several (two or more) vertices, and thereby captures this higher-order information. Hypergraph learning has achieved excellent performance in severalapplications [3].

III. LITERATURE SURVEY

M. Wang estimated the Optimizing multigraph learning: Towards a unified video annotation scheme. The description of learning with hypergraphs: clustering, classification and embedding proposed by D. Zhou presented the over-view of hypergraph utility on reranking method [4].

R. Zass determined the hypergraph for image matching using con-vex optimization. Hypergraphs have been applied to solve problems with multilabel learning and video segmentation [5] . M. Wang 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 [6].

Winston H. Hsu et al. have proposed classification based PRF tool which has been shown to expandsimple text search result in both text and image retrievals. Top result will referred as positive and other are negative. The methods are conducted on the TRECVID 2003-2005 data set. In which pseudo labeling strategies used such as binary, normalized rank and score stretching [6].

Vidit Jain suggested the technique referred as Gaussian regression model. These models guessthe normalized click count for every image and combineit with original ranking score. In which significantly expandthe performance of Bing image search engine on a extensiverange of tail queries. Due to the compositesemantics of pseudo-clicks estimates, it is likely that they are a non-linear function of the features used to signifythe data [6].

IV. PROPOSED METHODOLOGY

A novel method named multi-modal hypergraph learning-based sparse coding is re-implemented with hybrid approach 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 based 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 [7].

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 con-struct 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 [8]. 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 frame-work to achieve image reranking [9].

V. IMPLEMENTATION

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.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 [10]. 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. 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 [10].

A. Texture

The ability to retrieve images based on texture helps to make the difference between the similar colors for example consider a query image which consists of both sky and sea which are blue in color or image that consists of grass and leaves so retrieving images based on texture helps in making difference between those similarities. It describes how the picture is composed physically. The technique used to retrieve image based on texture is wavelet transform technique.

Wavelet means nothing but small wave. Wavelet change is a procedure of turning a signal to series of wavelets. It transforms the picture into multi scale representation with both spatial and frequency attributes. The main objective is to compute the pixel power of the pictures. In this technique the image is divided into 4-sub bands each of various frequency that is low-low, lowhigh, high-low, high-high. After getting the sub-band on which the frequencies are concentrated most and used for further processing. The energies of that sub-band is obtained. Within the range of energies, the distance between the exhibited query picture and every picture is ascertained. The pictures with lesser separation distance are placed at the top most place [11].

Guiying Li proposed to defined texture is a repeated pattern of information or arrangement of the structure with regular intervals. In a general sense, texture refers to surface characteristics and appearance of an object given by the size, shape, density, arrangement, proportion of its elementary parts. A basic stage to collect such features through texture analysis process is called as texture feature extraction. Due to the signification of texture information, texture feature extraction is a key function in various image processing applications like remote sensing, medical imaging and content- based image retrieval.

There are four major application domains related to texture analysis namely texture classification, segmentation, synthesis and shape from texture. Texture classification produces a classified output of the input image where each texture region is identified with the texture class it belongs. Texture segmentation makes a partition of an image into a set of disjoint regions based on texture properties, so that each region is homogeneous with respect to certain texture characteristics. Texture synthesis is a common technique to create large textures from usually small texture samples, for the use of texture mapping in surface or scene rendering applications [11].

The shape from texture reconstructs three dimensional surface geometry from texture information. For all these techniques, texture extraction is an inevitable stage. A typical process of texture analysis is shown in Figure.

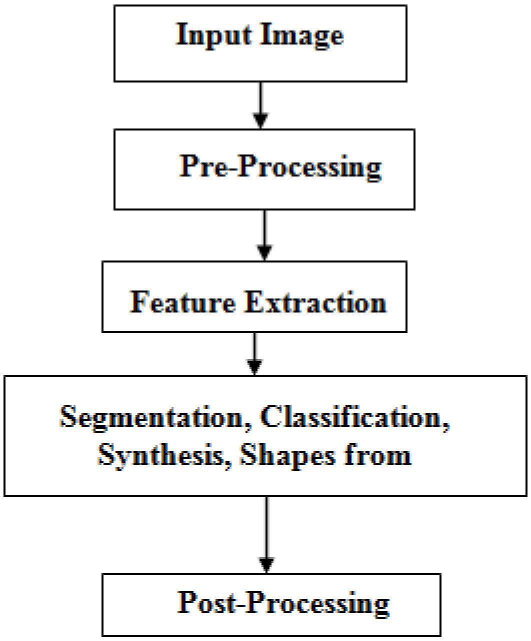

Fig.1. Texture Feature Process

Primitive Length Texture FeaturesCoarse textures are represented by a large number of neighboringpixels with the same gray level, whereas a small number represents finetexture [12]. A primitive is a continuous set of maximum number of pixels in thesame direction that have the same gray level. Each primitive is defined by itsgray level, length and direction. Letrepresents the number ofprimitives of all directions having lengthr and gray levela. AssumeNM,beimage dimensions,Lis the number of gray levels,rNis the maximumprimitive length in the images andKis the total number of runs. It is givenby the Equation

$$\sum_{a=1}^{L}\sum_{r=1}^{N} B(a,r)$$

Short primitive emphasis =

$$\frac{1}{K}\sum_{a=1}^{L}\sum_{r=1}^{N}\frac{B(a,r)}{r^2}$$

Long primitive emphasis =

$$\frac{1}{K}\sum_{a=1}^{L}\sum_{r=1}^{N} B(a,r)^2$$

Gray level uniformity =

$$\frac{1}{K}\sum_{a=1}^{L}\left[\sum_{r=1}^{N} B(a,r)r^2\right]^2$$

Primitive length uniformity =

$$\frac{1}{K}\sum_{a=1}^{L}\left[\sum_{r=1}^{N} B(a,r)\right]^2$$

Primitive percentage =

$$\frac{K}{\sum_{a=1}^{L}\sum_{r=1}^{N} rB(a,r)} = \frac{K}{MN}$$

B.Haar Wavelet

The Haar orthogonal system begins with Φ (t),the characteristic function of the unit intervalIt is clear that Φ (t) and Φ (t - n), n ≠ 0, n∈Z are orthogonal since their product is zero. It is also clear that {ϕ (t – n)} is not a complete orthogonal system in L^2 (R) since its closed linear span V_o consists of 2 piecewise constant functions with possible jumps only at the integers. The characteristic function of (0,1/2), for example, with a jump at 1/2, cannot have a convergent expansion.The ϕ (t) is usually called the scaling function in wavelet terminology while Ψ(t) is the mother wavelet. Haar Wavelet transform is used to calculate the feature vectors of textured images.Some Bold Advantages of Haar Wavelet as below [13] :

1) In terms of computation time, it produced Best performance
2) Speed of Computation is very high.

3) Haar Wavelet Transformations deals with Simplicity inworking.
4) It is an efficient method for Image compression.

C. *Hybrid Approach*

Texture Feature and Haar wavelet both image processing feature are combinly word with the multimodal hypergraph learning algorithm and produced more relevent and accurate results by using the sparce code. An image texture is a set of metrics calculated in image processing designed to quantify the perceived texture of an image. Image texture gives us information about the spatial arrangement of color or intensities in an image or selected region of an image.In mathematics, Haar wavelet is a sequence of rescaled square shaped functions which is together form a wavelet family or basis. Wavelet analysis is similar to Fourier analysis in that it allows a target function over an interval to be represented in terms of anorthonormal basis.

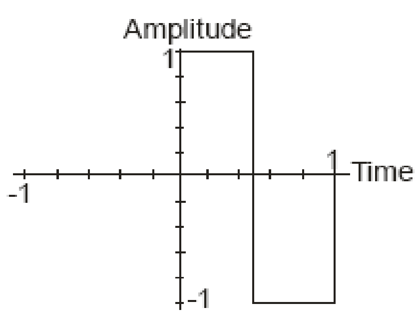

Fig. 2. The scaling function

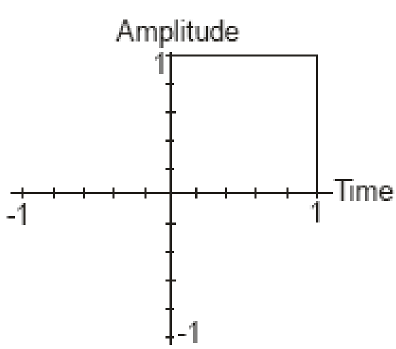

Fig.3. Mother wavelet for the Haar system.

VI. RESULTS

Experiments conducted on a real world data set not only describes the usefulness of click-through data, which can be viewed as the image 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. We randomly select images to form the image bases and test images. Since different queries contain a different number of images, it would be inappropriate to find a fixed number setting for different queries. Therefore, we choose different percentages of images to form the image bases. we conduct experiments to show the effects of different parameters. For all methods, we independently repeat the experiments five times with randomly selected image bases and report the averaged results. According to the experimental results, we observe that nearly all the used methods effectively improve on baseline comparative results. Our method, MHL, achievedthe best results for click prediction, with hypergraph based method performing better than other single graph based methods [14].

Table I

Performance comparisons of classification accuracy (%) for click predictionswith a fixed size of image bases, and variedSize of test image sets. The comparison includes a comparison of MHL, MGL, SHL, SGL, SC, KNN AND GP. The size of theTest image set is varied from among [10%, 30%, 50%], and the size of the image base is fixed at 75%

	MHL	MGL	SHL(A)	SGL(A)	SGL(L)	SC(A)	SC(L)	KNN	GP	MHL Approach	Hybrid Approach
10	64.5	63.5	61.1	59.3	61.3	59.6	61.3	60.2	57.3	69.1	71.2
30	65.9	64.6	61.6	58.8	62.5	60.8	61.6	61.5	58.8	66.7	69.3
50	66.4	65.4	63.4	61.4	63.1	60.5	61.7	61.7	59.4	67.3	70.0

IFERP
connecting engineers... developing research

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 3, Issue 11, November 2016

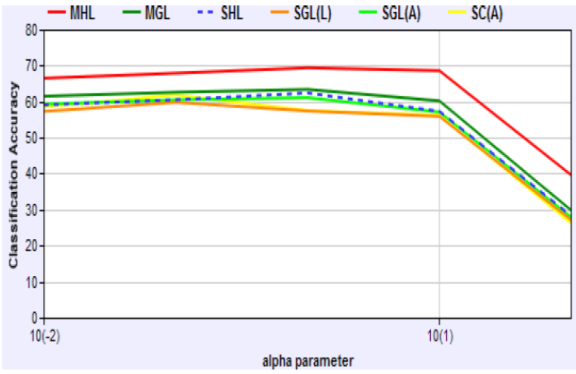

Fig. 4 Graphical representation of classification accuracy
Vs Alpha parameter

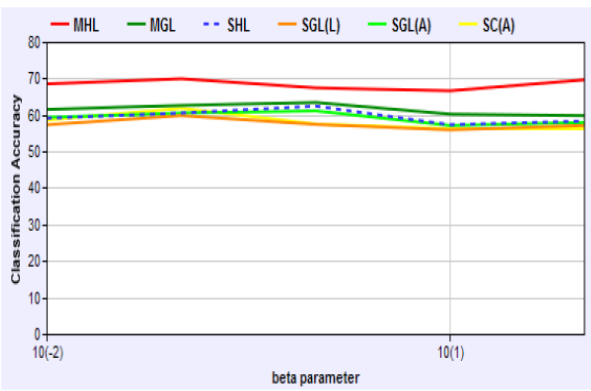

Fig. 5. Graphical representation of classification accuracy Vs beta parameter

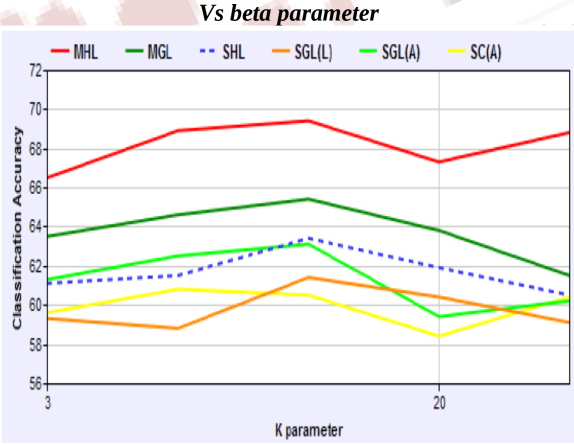

Fig. 6. Graphical representation of classification accuracy Vs K parameter

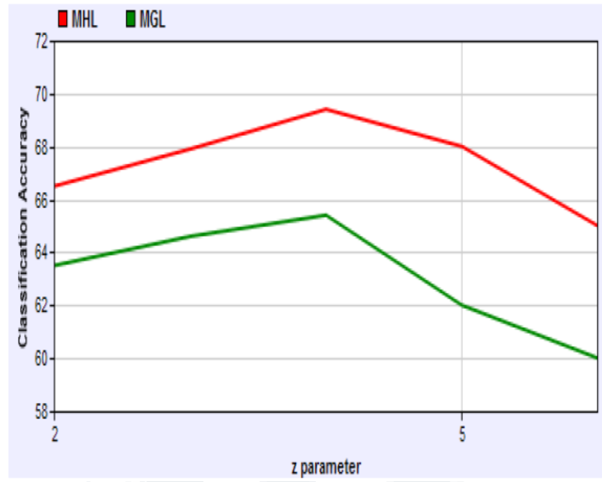

Fig. 7. Graphical representation of classification accuracy Vs Z parameter

The performance of all above methods which shows in table ISuch as MHL, MGL, SHL, SGL, SC, KNN, GP, MHL approach and Hybrid approach. In table I test image set is varied from among [10%, 30%, 50%] and size of image base is fixed. Hybrid approach which consist two bold features one is Haar wavelet second is texture feature both features are merged and combinely worked in algorithm, Produce more Accurate results which shows in table I [14]. When test image set is at 10% in that stage MHL result is 69.1 at same time result of hybrid approach is raised with 71.2. When test image set is at 30% in that stage MHL result is 66.7 at same time result of hybrid approach is raised with 69.3. When test image set is at 50% in that stage MHL result is 67.2 at same time result of hybrid approach is raised with 70.0. By the overall resulting values and graphical representation shows as test image set increased hybrid approach result is more accurate as compare to MHL.

VII. CONCLUSIONS

Image reranking is efficient for improving the performance of image search. Haar-waveletof computation time, it produced Best performance and speed of Computation is very high Texture feature is very important to identyfy object or region of interest in specific image.In which multimodal hypergraph learning based sparse coding method for click prediction of images use haar-wavelet as well as texture feature are combinely work through MHL and produce more accurate results. In future prospect 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]Bridle, John S. "Probabilistic interpretation of feed forward classification network outputs, with relationships to statistical pattern recognition." InNeurocomputing, pp. 227-236. Springer Berlin Heidelberg, 1990.

[2] Gao, Yue, Meng Wang, Zheng-Jun Zha, Qi Tian, Qionghai Dai, and Naiyao Zhang. "Less is more: efficient 3-D object retrieval with query view selection." IEEE Transactions on Multimedia 13, no. 5 (2011): 1007-1018.

[3]Clinchant, Stephane, Jean-Michel Renders, and Gabriela Csurka. "Trans-media pseudo-relevance feedback methods in multimedia retrieval." InWorkshop of the Cross-Language Evaluation Forum for European Languages, pp. 569-576. Springer Berlin Heidelberg, 2007.

[4] Duan, Lixin, Wen Li, Ivor Wai-Hung Tsang, and Dong Xu. "Improving web image search by bag-based reranking." IEEE Transactions on Image Processing 20, no. 11 (2011): 3280-3290.

[5] Gengtí, Bo, Linjun Yang, Chao Xu, and Xian-Sheng Hua. "Content-aware ranking for visual search." In Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, pp. 3400-3407. IEEE, 2010.

[6] Carterette, Ben, and Rosie Jones. "Evaluating search engines by modeling the relationship between relevance and clicks." In Advances in Neural Information Processing Systems, pp. 217-224. 2008.

[7] Candes, Emmanuel J., Justin K. Romberg, and Terence Tao. "Stable signal recovery from incomplete and inaccurate measurements." Communications on pure and applied mathematics 59, no. 8 (2006): 1207-1223.

[8] Dupret, Georges, and Ciya Liao. "A model to estimate intrinsic document relevance from the clickthrough logs of a web search engine." InProceedings of the third ACM international conference on Web search and data mining, pp. 181-190. ACM, 2010.

[9] Smith, Gavin, and Helen Ashman. "Evaluating implicit judgements from image search interactions." (2009).

[10] Shokouhi, Milad, Falk Scholer, and Andrew Turpin. "Investigating the effectiveness of clickthrough data for document reordering." In European Conference on Information Retrieval, pp. 591-595. Springer Berlin Heidelberg, 2008.

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

[12] Jain, Vidit, and Manik Varma. "Learning to re-rank: query-dependent image re-ranking using click data." In Proceedings of the 20th international conference on World wide web, pp. 277-286. ACM, 2011.

[13] Porwik, Piotr, and Agnieszka Lisowska. "The Haar-wavelet transform in digital image processing: its status and achievements." Machine graphics and vision 13, no. 1/2 (2004): 79-98.

[14] Eriksson, Anders P., and Anton Van Den Hengel. "Efficient computation of robust low-rank matrix approximations in the presence of missing data using the L1 norm." In CVPR, vol. 1, no. 2, p. 4. 2010.