Re-Ranking Technique using HCC based Similarity and Typicality Process

Dr. A. Nagamalleswar Rao[1], L. Mahesh[1], L. Prasanna Lakshmi[1]

[1]Dept. of CSE, Sree Dattha Institute of Engineering and Science, Hyderabad, Telangana, India

ABSTRACT

In image search re-ranking, a major problem restricting the image retrieval development is an intent gap, which is a gap between user's real intent and query/demand representation, besides well-known semantic gap. In the past, for achieving effective web image retrieval, classifier space or feature space is explored at a time by researchers. Visual information and images initial ranks with single feature are only considered in conventional re-ranking techniques for measuring typicality and similarity in web image retrieval, while overlooking click-through data influence. For image retrieval, various image features aggregation shows its effectiveness in recent days. But, uplifting the best features impact for a specific query image presents a major challenge in computer vision problem. In this paper, based on web query, features are assigned with weights, where different weights are received by different queries in ranked list. IABC algorithm used to compute weights is a data-driven algorithm and it does not require any learning. At last, in a web, color and texture features are fused using fusion and these features are extracted with respective modalities. A HypergraphConstruction Clustering (HCC) re-ranking with click-based similarity and typicality procedure termed as HCCCST is used in re-ranking technique. Its operation is depends on selection of click-based triplet's and a classifier is used for integrating multiple features into a unified similarity space. The web image search re-ranking performance is greatly enhanced using proposed technique.

Keywords: Click through data, re-ranking, HypergraphConstruction Clustering (HCC), Weight based Multi-Feature Fusion (WMFF), Texture Feature and ColorFeature.

1. INTRODUCTION

The phenomenal growth in the internet sector, have made every device across globe connected. The development in electronic gadgets [1] has pulled the devices in each and every sector of industry from education, mechanical inspections, medical sector and automobile industries. The low cost devices and on demand pricing of cloud have integrated the digital devices and cloud computing platform [2]. With the limited capacities of the devices enforces the storage of multimedia into secondary and cloud storages. The digital images generated are used by various users across the globe [3]. The doctors might the image of MRI on network with other doctors across globe for opinion [4]. In automobile industries, remote inspection of quality of parts requires transferring several images in short time frame. The images might contain various sensitive information such as patient disease status, financial details in image etc [5]. The challenge here arises to share the images across the globe with maintain the integrity and confidentiality of image. Some traditional approaches annotated the images and stored in a database. The annotation method was cumbersome [6] as it would require lot of human efforts in annotating each image in a huge dataset. Later it was developed by adding multiple tag names for images to perform a query operation more flexible. But all the traditional approaches required human intervention. All the retrieval operation [7]was based on the tags and annotations attached to the images in dataset. The human error in Meta tagging and annotation would fail the entire retrieval operations. Tag based image Retrieval was major development in this field, where the features of images were considered in retrieval operations which made the more accurate operations. To address the challenges of space and manipulation of images, all the images were compressed to some standard formats such as Jpeg and gif. The feature extractions were performed on these compressed standards itself [8]. The challenge was to use semi trusted cloud servers for storing the images and sharing across the various users without compromising the original images and by maintaining the privacy of users and images.The various TBIRalgorithms [9] majorly focused on feature extraction and indexing them for fast retrieval of images similar to query images,our work we majorly focus on generating the features by color and texture differentiation using Color Co-occurrence Matrix for colors and Gabor wavelet transformation for texture. The model was tested on Corel 10K dataset with five major categories.

The major contributions of this work as follows:

- Weight based Multi-Feature Fusion technique is used for fusion of color features and Texture features respectively.
- Hypergraph Construction Clustering re-ranking with click-based similarity and typicality procedure is used in re-ranking technique.

The remaining part of paper is organized as; section 2 gives the comprehensive literature survey carried in the domain. Section 3 demonstrates the workflow of proposed model. Section 4 discusses the experimental design and results. Section 5 concludes the paper.

2. LITERATURE REVIEW

Related works are categorized in two dimensions: search with click-through data and visual search re-ranking.Tian et al [10-11] implemented a local-global discriminative dimension reduction algorithm, where discriminative information and local geometry of labeled images are transferred into entire image database for learning sub manifold. Proposed active re-ranking scheme’s effectiveness is demonstrated via experimentation on real web image search dataset and synthetic datasets. The proposed scheme includes local-global discriminative dimension reduction algorithm and active sample selection technique based on structural information.Zhang et al [12-13] leveraged click-through data, which is “implicit” user feedback for understanding the query in an effective manner. In this work proposed a novel re-ranking algorithm, termed as click-based relevance feedback. User search intentions are identified in this algorithm by emphasizing click-through data successfully. It also learns query-dependent fusion weights adaptively for multiple modalities by leveraging multiple kernel learning algorithms. Around 11.62% initial search results are enhanced by proposed re-ranking technique as shown in results of experimentation and for most kind of queries like top, middle and tail queries, various existing techniques are outperformed by this technique. Liu and Mei [14] stated that globally optimum ranked list will not be produced if classification performance is not good enough. They formulated an optimization problem from re-ranking, where, if any two arbitrary documents in the list are ranked properly based on relevance, an optimum ranked list will be produced. For computing relevant relation of every pairwise re-ranking, introduced two pairwise re-ranking techniques called, Exclusion Pairwise re-ranking (EP-re-ranking) and Difference Pairwise re-ranking (DP-re-ranking). At last, final ranked list is recovered by exploring Round Robin criterion. Over text search baseline and other re-ranking techniques, consistent enhancement is shown using this technique.Duan et al [15] implemented a Generalized Multi-Instance (GMI) setting for this application. The retrieval performance is enhanced using developed GMI-SVM classifier, where, labels from bag level are propagated to instance level. All the bags are ranked using proposed bag ranking technique based on defined bag ranking score. Top ranked bags are utilized as pseudo positive training bags, where few irrelevant images which are not having any association with texture query are sampled randomly for computing pseudo negative training bags. With manually labelled training bags which are derived from relevance feedback, better performance can be achieved using proposed GMI-SVM as demonstrated in experimentation.

3. Research Method

The proposed system outline is presented in figure 1. The system operates in an offline mode where all the collected image data, their descriptive tags are aligned to a database, with the user updation and defined tags for each image. This process of database creation is carried out in the training phase. In the test phase, a new image is given for updation to the database, for which tags are defined by the user. he comprehensive literature survey reviewed the need of secured TBIR operations over cloud by encrypting images before outsourcing and need to address the latency challenge in searching in encrypted images. So we designed a model that can address a challenge of latency by performing the search operations through indices and preventing the sensitive information stored in images. The figure 1 demonstrates the workflow of our proposed model. Our model has three sub components. The comprehensive literature survey reviewed the need of secured TBIR operations over cloud by encrypting images before outsourcing and need to address the latency challenge in searching in encrypted images. So we designed a model that can address a challenge of latency by performing the search operations through indices and preventing the sensitive information stored in images. The figure 1 demonstrates the workflow of our proposed model. Our model has three sub components. The functional operation of suggested architecture and each of the operational unit function is outlined below.

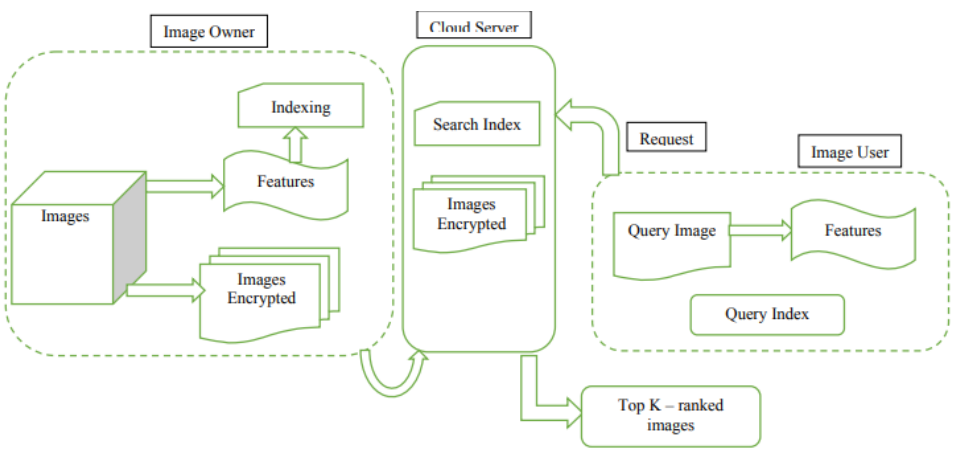

Figure 1: Workflow of Proposed Model

The image owner has set of images and wants to distribute securely to its authorized users across the similar or different networks. The Image owner extracts the features from each image in the database and assigns a indexing for its computed feature set. The image owner encrypts all the images and outsources the encrypted images and search indices to the third party cloud server. The cloud server is responsible for storing the encrypted images along with the search indexes. The cloud server is also responsible for providing immediate responses for user queries by replying them with the encrypted images by taking the search indices from the image users. The image users authorized to retrieve the encrypt images from cloud server by generating a trapdoor to query image and decrypt the images with a valid key provided by the image owner.

PROPOSED METHODOLOGY

An effective as well as less time consumed web image retrieval process is focused in this work. In the first stage, colors are extracted from given keyword text using RGBtoLAB techniques. Gray level Co-occurrence Matrix (GLCM) is used for extracting features. Then, WMFF is used for fusing the extracted color and texture features. Based on web query, weights are assigned for features, where different weights are assigned to ranked list of various queries. For re-ranking, proposed a novel image similarity learning algorithm called HCCCST. For multiple features, after learning similarity metric for grouping semantically and visually similar images as same clusters proposed aHCC. Cluster typicality and within-clusters image typicality in descending order are computed for computing final re-rank list.

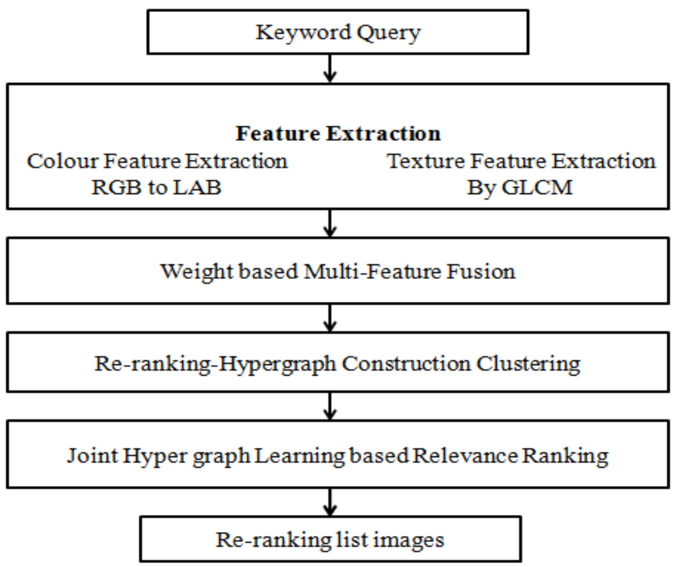

Figure 2: overview of the proposed architecture

A new direction has emerged forimproving the search performance and termed as image search Re-ranking. In this technique, text based searching function outcomes are reordered using visual information. From initial search function outcomes, this work considers image feature's visual characteristics, which follows the ranking function construction and for ranking, functions are used for reordering the images conclusively. Image's visual similarity are included for performing image re-ranking and for getting better results from image and also for obtaining multiplicity of image result. In image search process, re-ranking objective is to upgrade text based search's impact which requires image Tag extraction, which is a visual. Re-ranking also considers visual Tags of images. For confining user's goal with least human intervention, significant image search technique is proposed for searching related images via necessitation of single click by user over images, which are initially searched.

3.1Feature extraction

Extracting an effective features and represent them in an effective way is very important to improve the retrieval performance of tag based image retrieval system.Color is the most effective feature as it is simple and robust. Also, the frequency domain features are observed to be more effective in image representation due to its continuous monitoring property of image variation. These frequency features are called textural features. The color features are extracted using color moment technique in L*a*b* color space, texture features are extracted using wavelet packet transform and tag features are extracted using TF-IIF method

In feature extraction, shape, texture and color are the three major feature descriptors applied. Here, texture and color extraction are mainly focused. Following section presents, brief discussion about it. After the specification query, this entire process takes place. Keyword type query is specified in this work. In any Tag-based image retrieval approach, visual feature extraction plays a major role. In general, features include both visual features like texture, shape, color and text based features like keywords and annotations. Further, in visual feature range, features are classified as lowlevel and high-level features.

3.1.1 Color feature extraction using color moments:

Examining images based on the colors they contain is one of the most widely used techniques because it can be completed without regard to image size or orientation. The color features are extracted using color moment technique as it more effective and simple method. Color moments differentiate images based on probability distribution of color which are characterized by a number of moments. The first, second and third order moments are the three central moments which represents an image's color distribution. There are different color models to represents colors of the images such as RGB, HSV, L*a*b*. Therefore, colors can be represented by three different channels based on color space models. The CIE L*a*b* color space model is adopted due to its perceptually uniform color space which is defined by International Commission of Illumination. It is perceptually uniform because Euclidean distance between colors represents the color difference perceived by the human eye. Also, it is very efficient in representing even small color differences.

3.1.2 Texture feature extraction:

Textures are analyzed using a statistical approach called GLCM. Spatial relationships between pixels are reflected using these textures. The GLCM functions are used for describing the texture of an image by computing how certain values are possessed by frequent pixel pairs and in a definite spatial relationship, appear in image, producing a GLCM, followed by arithmetical measures extracted from matrix. A GLCM is generated using a gray co matrix function. It computes how often a pixel with intensity I appears in specific spatial relationship with a pixel with intensity j. Spatial relationship is described as interest pixel, pixels lying in its direct right. In consequential GLCM, every available element (i, j) is merely an aggregation of pixels count with value i, which is indicated in spatial relationship with a pixel with value j present in input image. Shape, color, Texture, Entropy, Energy, Intensity, correlation, Homogeneity and contrast features are extracted in this work. Similarity measurement is done using these features.

3.1.3 Weight based Multi-Feature Fusion (WMFF)

In image search re-ranking, when compared with single feature, multi-feature fusion shows its effectiveness in mining relevant recurrent visual patterns of multiple features. Various aspects of image Tag are represented using multiple features, where same image's semantics are shared potentially. Direct computations of weights are difficult due to different scaling of similarity scores from various features.

3.2. Hypergraph construction

The 'm' different Wp (p = 1, 2, m), are obtained after learning click-based multi-feature similarity procedure and image similarity is measured.For detecting relevant recurrent patterns according to learnt similarity metrics, HCC for grouping semantically as well as visually similar images into same clusters. From data, pairwise-affinity matrix is constructed easily according to learnt image similarity metrics and relevant operation is performed directly. In spectral clustering, prior knowledge is used in HCC. From domain knowledge, pairwise constraints are used in this. Between two semantically as well visually similar images, pairwise constraints are represented as cannot-links (in different classes) and as must-links (in the same class). For every must link (i, j) pair, $sim_{ij} = sim_{ji} = 1$ is assigned and for every cannot-link (i, j) pair, $sim_{ij} = sim_{ji} = 0$ is assigned. In the graph, connection between two semantically as well as visually similar images with high similar features is made, if SSC is used to cluster images in feature vector via t-nearest neighbor graph representation. Between image pair, similarity is changed to 0 for cannot-link constraints and in graph, edge between images pair is broken. This study applies only must-link constraints. Algorithm 1 describes the Hypergraph spectral clustering algorithm details. For a specified images $I1,In$, created l pairwise must-link constraints,similarity function expressed in (1) is used for obtaining similarity matrix 'S'.

$$sim_{ij} = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \qquad (1)$$

Where, scaling parameter is represented as σ, which is used to measure similarity between two images. Sparse matrix is formed by modifying S, where, for every image in S, only t nearest neighbors is maintained. Then, in S, applied the l Pairwise constraints, Normalized spectral clustering algorithm is followed in steps 5 to 10 [1]. As in [1], for a specified $n \times m$ web image with n samples and m features, generated a $n \times q$ image subspace (q < d) as,

$$q = q_{min} + [\alpha\,(q_{max} - q_{min})] \qquad (2)$$

Where, a uniform random variable is represented as α, which lies between 0 to 1, subspace's lower bound is represented as $qmin$ and upper bound is represented as $qmax$. 0.75d is said as $qmin$ and 0.85d is said as $qmax$. Assume a cluster ensemble as $\pi = \pi_1, \pi_u$ with u clustering solutions. On every subspace dataset, applied this clustering for computing clustering results (u). Fixed clusters count u is used and one clustering solution is represented by every $\pi_i = u_{i1}, u_{ik}$. Between clusters and images, according to crisp associations, generated the Cluster Association Matrix 'AM', it has m*u clusters and n images. If Iibelongs to a cluster uj , AM(I_i, u_j) = 1, $i = 1, \dots, n$; $j = 1, \dots, m$, otherwise AM(I_i , u_j) = 0. In RAM (I_i,u_j), new association values are estimated for generating a Refined Cluster Association matrix 'RAM', if AM (I_i,u_j) = 1. Between $uj$and other clusters to which Iiprobably belongs to them each other, similarity function is represented as RAM (I_i,u_j). From a clusters weighted graph, similarity between any clusters can be obtained in cluster ensemble. At last, for computing final clustering solution, on RAM, applied the spectral clustering.

3.3 Joint Hyper graph Learning based Relevance Ranking

Cluster Typicality: The relative cluster similarity sim(u) and initial cluster confidence confd (u) are used for deciding cluster typicality. Where, clusters are represented using u. With respect to specified query, in this cluster, overall sample relevance is measured using initial cluster confidence. Combination of click score and initial score are used for defining initial cluster confidence as,

$$confd(u) = \sum q - \frac{r(x_1)}{n} + \frac{C_1}{sum\ (c)} \qquad (3)$$

Where, images count to be re-ranked is represented as n, x's initial ranked order is represented as r (x), in initial ranked list, from all clicked images, click counts summation is represented as sum(c). On the other side, image representativeness to a specified query topic is reflected using typicality. There will be similarity between images in the same cluster, if they have good typicality. So, ratio between intra-cluster and inter-cluster similarity defines relative cluster similarity.Within cluster u, between images, average similarity is represented using the first term in the above mentioned expression, in the same cluster, between any image, average similarity is represented using the second term. According to above two measures, defined the cluster typicality CT (u) as a linear weighted combination.

$$CT(u) = Y + \ conf(u) + (1 - Y) * sim(u) \qquad (4)$$

Local Typicality: Local density ld(x), image initial confidence con(x) are used for computing local typicality. In cluster u, image x's local density $ld(x|u)$ is computed using Kernel Density Estimation (KDE).

4. RESULTS AND DISCUSSION

Experimental settings are introduced in this section with results. In order to validate the proposed re-ranking techniques effectiveness, overall performance is specified through multiple features integration, usage of various distance metrics with comparison. With different metrics, available Click-Based Multi-feature Similarity Learning (CMSL), Support Vector Machine (SVM) based web image retrieval, Multiple Support Vector Machine with Kernel Learning (MSVM-KL) and proposed HCCCST algorithm are compared.For comparison of performances, metrics like accuracy, NDCG, mean Average Precision (mAP), Average Precision (AP) are used in this research work.

4.1 Results comparison

Results of different metrics discussed above are compared based on various image retrieval techniques.

Table 1: various image retrieval techniques with respect to average precision (AP)

Method	Retrieval methods -Average Precision (AP)			
	CMSL[12]	SVM[14]	MSVM-KL[16]	HCCCST
AP	0.8712	0.8892	0.9128	0.9318
mAP	0.8625	0.8812	0.9076	0.9363
Accuracy (%)	85.74	87.45	90.78	94.75

From Table 1 and Figure 3 shows the overall performance of different re-ranking approaches on dataset. On the whole, the proposed HCCCST outperforms other methods, and the improvements are consistent. Using proposed HCCCST re-ranking approach the NDCG values are improved significantly.

The AP performance comparison between different image retrieval techniques are shown in Table 1. It indicates that, existing methods are outperformed by proposed HCCCST and for various images counts; consistent and stable performance is exhibited using this technique. Around 0.9318 AP value is attained in proposed HCCCST, whereas, 0.8712 is attained in CMSL, 0.8892 is attained in SVM and 0.9128 is attained in MSVM-KL for 100 image set as shown in that figure. When compared with existing techniques, high AP value can be attained using proposed HCCCST technique as depicted in implemented and validated results.

The mAP performance comparison between different image retrieval techniques are shown in figure 3. It indicates that, existing methods are outperformed by proposed HCCCST and for various images counts, consistent and stable performance is exhibited using this technique. Around 0.9363 mAP value is attained in proposed HCCCST, whereas, 0.8625 is attained in CMSL, 0.8812 is attained in SVM and 0.9076(as shown in table 3) is attained in MSVM-KL for 100 image set as shown in that figure. When compared with existing techniques, high mAP value can be attained using proposed HCCCST technique as depicted in implemented and validated results.

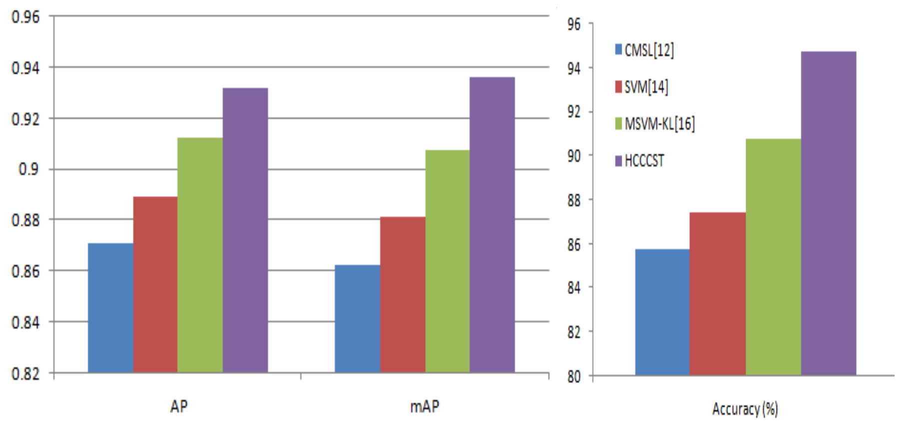

Figure 3: Performance comparison

Accuracy comparison between different image retrieval techniques are shown in figure 5. Around 94.75% of accuracy can be attained in proposed HCCCST, whereas, 85.74% is attained in CMSL, 87.45% is attained in SVM and 90.78% is attained in MSVM-KL for 100 image set as shown in that figure. When compared with existing techniques, high accuracy can be attained using proposed HCCCST technique as depicted in implemented and validated results.

5. CONCLUSION AND FUTURE WORK

Issues related to leveraging click-through data for reducing image search intent gap are studied in this paper. Separate extraction of colour and texture from given keyword text is performed in the initial stage. WMFF is used for fusing the extracted color and texture features. Based on web query, weights are assigned for features, where different weights are assigned to ranked list of various queries. From image set, feature vector ‘p’s weight value is used. A HCCCST is entirely adopted in this proposed re-ranking technique for guiding image typicality and similarity learning. For multiple features, after learning similarity metric for grouping semantically and visually similar images as same clusters proposed aHCC. Cluster typicality and within-clusters image typicality in descending order are computed for computing final re-rank list. With respect to metrics like accuracy, NDCG, mAP, AP is used for comparing proposed HCCCST with various available re-ranking techniques. Superiority and availability of proposed HCCCST is demonstrated using extensive results. Moreover, on feature-weighting, its effectiveness is also exploited in computer vision problems like image segmentation, object detection, person re identification.

REFERENCES

1. Mei T., Y. Rui, S. Li, and Q. Tian, "Multimedia search reranking: A literature survey," ACM Comput. Surv., vol. 46, no. 3, p. 38, 2014.
2. Hsu W. H., L. S. Kennedy, and S.-F. Chang, "Reranking methods for visual search," IEEE Multimedia, vol. 14, no. 3, pp. 14–22, 2007.
3. Hsu W. H., L. S. Kennedy, and S.-F. Chang, "Video search reranking through random walk over document-level context graph," in Proc. 15th ACM Int. Conf. Multimedia, 2007, pp. 971–980.
4. Yao, T., Ngo, C.W. and Mei, T., 2013. Circular reranking for visual search. IEEE Transactions on Image Processing, 22(4), pp.1644-1655.
5. Cui J., F. Wen, and X. Tang, "Real time google and live image search re-ranking," in Proc. 16th ACM Int. Conf. Multimedia, 2008, pp. 729–732.
6. Tian, X., Tao, D., Hua, X.S. and Wu, X., 2010. Active reranking for web image search. IEEE Transactions on Image Processing, 19(3), pp.805-820.
7. Zhang C., J. Y. Chai, and R. Jin, "User term feedback in interactive text-based image retrieval," in Proc. 28th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, 2005, pp. 51–58.
8. DupretG. and C. Liao, "A model to estimate intrinsic document relevance from the clickthrough logs of a Web search engine," in Proc.3rd ACM Int. Conf. Web Search Data Mining, 2010, pp. 181–190.
9. DupretG. E. and B. Piwowarski, "A user browsing model to predict search engine click data from past observations," in Proc. 31st Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr., 2008, pp. 331–338.
10. Zhang Y., X. Yang, and T. Mei, "Image search reranking with querydependent click-based relevance feedback," IEEE Trans. Image Process., vol. 23, no. 10, pp. 4448–4459, 2014.
11. Tang J., R. Hong, S. Yan, T.-S. Chua, G.-J. Qi, and R. Jain, "Image annotation by kNNsparse graph-based label propagation over noisily tagged Web images," ACM Trans. Intell. Syst. Technol., vol. 2, no. 2, p. 14, 2011.
12. Tang J., Z. Li, M. Wang, and R. Zhao, "Neighborhood discriminant hashing for large-scale image retrieval," IEEE Trans. Image Process., vol. 24, no. 9, pp. 2827–2840, 2015.
13. Liu Y., T. Mei, M. Wang, X.Wu, and X.-S. Hua, "Typicality-based visual search reranking," IEEE Trans. Circuits Syst. Video Technol., vol. 20, no. 5, pp. 749–755, 2010.
14. Liu Y. and T. Mei, "Optimizing visual search reranking via pairwise learning," IEEE Trans. Multimedia, vol. 13, no. 2, pp. 280–291, 2011.
15. Duan L., W. Li, I. W. H. Tsang, and D. Xu, "Improving Web image search by bag-based reranking," IEEE Trans. Image Process., vol. 20, no. 11, pp. 3280–3290, 2011.