Journal of Information & Knowledge Management
Vol. 19, No. 1 (2020) 2040003 (11 pages)

DOI: 10.1142/S0219649220400031

Efficient Implicit Content-based Image Re-ranking Approach

Saed Alqaraleh[*,‡] and Omar Ramadan[†,§]
[*]*Computer Engineering Department*
Hasan Kalyoncu Üniversity Havaalanı
Yolu Üzeri 8. Km, Şahinbey, Gaziantep, Turkey

[†]*Department of Computer Engineering*
Eastern Mediterranean University
Via Mersin 10, 99628 Famagusta
North Cyprus, Turkey
[‡]*saed.alqaraleh@hku.edu.tr*
[§]*omar.ramadan@emu.edu.tr*

Published 12 March 2020

Abstract. This paper presents a new image re-ranking approach that can implicitly improve the retrieved images based on the file's contents and some user-specific actions. In more detail, multiple descriptors are used to describe image files accurately and they do not require user intervention or tuned parameters. Furthermore, each of these descriptors has a weight, which affects the file rank. Unlike existing approaches, descriptor weight is assigned dynamically and changes from one file to another based on the percentage of differences found by the descriptor. Hence, the developed weight mechanism improves the chance of getting the required files significantly. The performance of the developed approach is investigated through several experiments and it has been observed that the approach has the ability of showing the most relevant files at the top of the query results and increases the percentage of the retrieved relevant files.

Keywords: Re-ranking algorithm; implicit ranking; image search engines; content-based retrieval.

1. Introduction

In recent years, search engines are mainly using images metadata and keywords for indexing purposes. Such a mechanism is very efficient for indexing text and webpages. However, when this mechanism is used for indexing multimedia files in general and images in particular, its efficiency can be decreased, as in many cases, as there may be no relation between the contents of the multimedia files and their metadata such as names or their surrounding text. This may lead the outcome of a query to contain a large number of irrelevant files. Ranking query results is basically ordering the results (websites in most cases) into a list that will be displayed to the user (Zhu *et al.*, 2008; Chen *et al.*, 2011; Lu *et al.*, 2013; Singhal and Srivastava, 2016). One of the methods that have recently been frequently used to overcome this problem is the re-ranking approach. In general, re-ranking works on re-ordering the query results in a way that tries to show the most relevant files at the top of the new list (Lafferty and Zhai, 2001; Yang *et al.*, 2009; Liu, 2011; Jain and Varma, 2011; Dali *et al.*, 2012). This study focussed on the content-based re-ranking approach. In other words, it uses the contents of the images, which we believe can contain useful information that should be used and can significantly improve the order of the image query results.

The main aim of this study is twofold. The first one includes: (1) Investigating the performance of some well-known descriptors for improving the overall performance of the image retrieval systems. (2) Investigating the performance of an ensemble system that uses the studied descriptors. Then, our next aim was to develop an approach that can help users who are willing to find certain images through any search engine to obtain the required files as easy and fast as possible.

The main contribution of the work presented in this paper is introducing a new image re-ranking approach which is based on the image contents and some implicit user actions, i.e. downloading, copying a file or a part of a file or spending more than the required duration (N) in checking the file. In addition, multiple descriptors that can precisely describe the content of image files have been used. Furthermore, the approach uses multiple threads to extract the representation of each image using the selected descriptors. Unlike existing re-ranking approaches, one of the main contributions of this paper is that a mechanism has been proposed to assign a weight to each of the used descriptors. This weight is dynamically calculated and changes from one file to another based on the descriptors' ability to distinguish between the files.

The remaining of this paper is organised as follows: Section 2 discusses the related literature. Section 3 presents the proposed re-ranking approach (PRA). The experimental evaluation and analysis are presented in Sec. 4. Finally, conclusions and future works are given in Sec. 5.

2. Related Works

In recent years, multiple content-based re-ranking approaches have been developed to improve the quality of the retrieved image files (Jain and Varma, 2011; Krizhevsky *et al.*, 2012; He *et al.*, 2016; Yang *et al.*, 2016; Jing *et al.*, 2018; Yuan *et al.*, 2019). In the work of Jain and Varma (2011), a click-based re-ranking approach was introduced. This approach re-ordered the original results based on their similarity to the clicked image. However, while users are viewing the obtained images, some of the clicked images may appear to be irrelevant. Hence, in this case, the top of the re-ranked list will contain absolutely irrelevant files.

Yuan *et al.* (2019) introduced a re-ranking framework with an updateable image pool. In general, the image pool is constructed by collecting images while following the rules of maximizing diversity. Then, the distance between the feature vectors for the image pool and the gallery is calculated. Next, the rank score of partial elements of the initial ranking lists is calculated to define the re-ranking list.

Yang *et al.* (2016) proposed an image re-ranking that used click-through data. Briefly, this algorithm integrates multiple features into a unified similarity space. Then, spectral clustering is used to group visually and semantically similar images into the same clusters. Finally, the re-ranked list is obtained by calculating click-based image typicality for the clusters and their contents as well in the descending order.

A hyper-graph-based semi-supervised ranking method for image re-ranking was presented by Jing *et al.* (2018). This approach is built based on the assumption that visually similar images should have similar ranking scores. In addition, a graph construction approach that incorporates relevance information from labelled samples and pseudo-relevance degree from unlabelled samples was developed. Then, it incorporates both the hyper-graph regulariser and the prior pairwise preferences information into a unified ranking learning framework.

Yousuf *et al.* (2018) introduced a content-based image system that used visual words fusion of SIFT and LIOP descriptors. As stated by Yousuf *et al.* (2018), the main reason for selecting these descriptors is due to the fact that SIFTs can handle scale changes and invariant rotations, while local intensity-order pattern outperforms SIFT in handling illumination changes.

Li *et al.* (2017) developed a multi-view image re-ranking based on user relevance feedback and multiple features. The main advantage of this approach is that it incorporates some heterogeneous property features to exploit their complementarities in representing image files.

Qian *et al.* (2017) proposed a ranking system for tag-based image retrieval. In more detail, this system constructs a tag graph, and then the topic is predicted for each tag. Finally, the query retrieved results are obtained using both inter-community and intra-community ranking methods.

In the last few years, many researchers have attempted to adopt convolutional neural network (CNN) to improve the performance of multimedia retrieval. Overall, CNN models such as GoogleNet (Szegedy *et al.*, 2015), ResNet (He *et al.*, 2016), AlexNet (Krizhevsky *et al.*, 2012) and VGGNet (Simonyan and Zisserman, 2014) were able to achieve state-of-the-art performance in many of the image retrieval tasks such as classification and object detection. For instance, AlexNet was one of the first deep networks that overcame the classification traditional methodologies. In addition, VGGNet can be considered as one of the most preferred choices for extracting features from images. For instance, Qi *et al.* (2018) proposed a multi-graph-based non-negative feature embedding framework for the image. In this framework, multiple image features are extracted by AlexNet and NIN model which will be embedded into a unified latent space. Then this system works on finding an optimized combination of multiple Laplacian matrices to approximate the intrinsic manifold automatically. Another CNN semantic re-ranking system was presented in by Wang *et al.* (2019). This system trains two CNN models, one for classification of sketches and the other for the natural images. The main idea is to capture the semantic information of both the sketches and natural images using deep learning. Then, category information is used for re-ranking the initial retrieval results.

3. The Proposed Image Re-ranking Approach

In the proposed re-ranking approach, the user actions will be detected and used to improve the order of the retrieved relevant files. It has been assumed that whenever a user performs one of the following actions on any of query results (files), it means that this file, which we will refer in this paper as Target, is related to the required ones: (a) A file is downloaded; (b) a file or a part of the file is copied; (c) the duration of checking a file has exceeded the required duration (N), where N will be specified by the system administrator.

Then, the selected file (Target) will be analysed and the query results will be re-ordered depending on their similarity with Target. In general, the proposed approach has some offline operations, such as pre-processing and features extraction, and some online operations, i.e. calculating files' similarity and re-ordering the query results. The details of these operations are summarized in the following.

3.1. *Adding and/or updating the database*

This process can consider as offline operation, and whenever a new file is added to the database, the file will be pre-processed and its features will be extracted and saved. The details of these operations are explained below.

3.1.1. *Pre-processing operations*

In general, pre-processing aims to improve the image data by eliminating the undesired distortions and at the same time enhancing some main components of the image. In this paper, the median filter which has the ability to efficiently keep the image detail at a reasonable cost is used. In addition, if the image contains text, the de-skewing approach of Chen and Ni (2011) is used to remove any skew.

3.1.2. *Extracting file features*

Selecting the features to represent a file is still one of the main challenges in content-based image retrieval, where selecting few representative features is not sufficient and may degrade the system while using a large number makes the system complex and time-consuming. All existing feature extraction techniques have some limitations. In other words, the main limitation of the developed approach and similar ones is that there is not a single technique that can obtain the best feature vector that can represent the image. In the following, the feature extraction process is described. Based on our preliminary experiments, the following descriptors are used to represent the images accurately:

(1) Scalable Colour Descriptor (SCD): This mainly describes the colour distribution in an image by obtaining the colour histogram, and can provide the global

colour features when it is measured over the entire image (Lux and Chatzichristofis, 2008).

(2) Edge Histogram Descriptor (EHD): This has the ability to describe the distribution of the edges in an image (Lux and Chatzichristofis, 2008), and helps in finding the semantic meaning of the image contents (Balan and Devi, 2012).

(3) Joint Composite Descriptor (JCD): This is a combination of the colour and edge directivity descriptor (CEDD) (Chatzichristofis and Boutalis, 2008a) and the fuzzy colour and texture histogram descriptor (FCTH) (Chatzichristofis and Boutalis, 2008b). In general, CEDD and FCTH descriptors describe the file's visual contents by combining colour and texture information. In addition, it has been found by Chatzichristofis *et al.* (2010) that the JCD can obtain quite a good efficiency for image retrieval.

As a result, each image file is represented by three vectors containing the descriptor outputs as follows:

$$F_{\mathrm{SCD}} = [f_1, f_2, \ldots, f_{255}]^T, \quad F_{\mathrm{EHD}} = [f_1, f_2, \ldots, f_{80}]^T,$$
$$F_{\mathrm{JCD}} = [f_1, f_2, \ldots, f_{168}]^T.$$

It is worth mentioning that the extracted features will be normalised and adjusted to be in the same range.

3.2. *Re-ordering the query results*

This is an online operation that works by monitoring the user behaviour and whenever the user performs one of the predefined actions, i.e. downloading, copying or spending more than the required duration (N) with a Target file. The query results will be re-ranked by applying the following steps:

(1) Calculate the distance between the Target file (T) and the other files (Y) by using the Euclidean formula

$$D_{\mathrm{dscr}_j}(T, Y) = \sqrt{\sum_{i=1}^{n} |f_{(i)}(T) - f_{(i)}(Y)|^2} \quad \text{for } j = 1, 2, 3, \tag{1}$$

where the descriptors for image files are SCD, EHD, and JCD and the value of n equals to 255, 80, 168, respectively. It is worth mentioning that the proposed re-ranking algorithm can work with any distance matrix.

(2) Assign a weight for each descriptor that specifies the descriptor influence on the file rank. This is based on the similarity of the descriptor vector of Y and T files. In this paper, the sum of weights of the used descriptors is assumed to be 1, i.e. $\alpha + \beta + \gamma = 1$, where α, β and γ are the weights of the used descriptors, which are computed as

$$\alpha = \frac{D_{\mathrm{dscr}_1}(T, Y)}{T_D(T, Y)}, \tag{2}$$

$$\beta = \frac{D_{\mathrm{dscr}_2}(T, Y)}{T_D(T, Y)}, \tag{3}$$

$$\gamma = \frac{D_{\mathrm{dscr}_3}(T, Y)}{T_D(T, Y)}, \tag{4}$$

where T_D is the total distance between T and Y files computed as

$$T_D(T, Y) = D_{\mathrm{dscr}_1}(T, Y) + D_{\mathrm{dscr}_2}(T, Y) + D_{\mathrm{dscr}_3}(T, Y). \tag{5}$$

(3) The rank of the file (Y) is finally computed as follows:

$$\mathrm{Rank}(Y) = \alpha N^{Y}_{D_{\mathrm{dscr}_{(1)}}} + \beta N^{Y}_{D_{\mathrm{dscr}_{(2)}}} + \gamma N^{Y}_{D_{\mathrm{dscr}_{(3)}}}. \tag{6}$$

(4) Finally, in order to ensure the rank value to be in the range of [0, 1], the rank of file (Y) need to be normalised. The normalised file rank is denoted as N^Y_R and computed as

$$N^Y_R = \frac{R(Y) - R_{\min}}{R_{\max} - R_{\min}}, \tag{7}$$

where $R_{\min}$ and $R_{\max}$ are, respectively, the smallest and the largest assigned rank values.

4. Experimental Study

In this section, the performance of the proposed re-ranking approach is investigated through multiple experiments. In addition, the R-precision (R_{Prec}) (Craswell, 2009), which is the precision value obtained for the top retrieved R documents, is used and computed as

$$R_{\mathrm{Prec}} = \frac{r}{R}, \tag{8}$$

where r is the number of relevant documents among the top-R retrieved ones. In other words, this matrix finds and represents the fraction of the retrieved files that are relevant to the user query.

Regarding the used datasets, the first 50 files of the Google results for a predefined set of image queries are collected and used as the dataset for the first experiment. While in the second experiment, the first 100 files of the Google results are used. Finally, for the third experiment, a dataset which included 25,000 images is used, and 25 image queries that belonged to multiple topics and subjects have been constructed and used.

4.1. *Experiment 1: Performances of SCD, EHD, JCD and the proposed approach*

In this experiment, the SCD, EHD and JCD descriptors and the proposed approach are used to re-order the first 50 files of the Google results for a set of image queries.

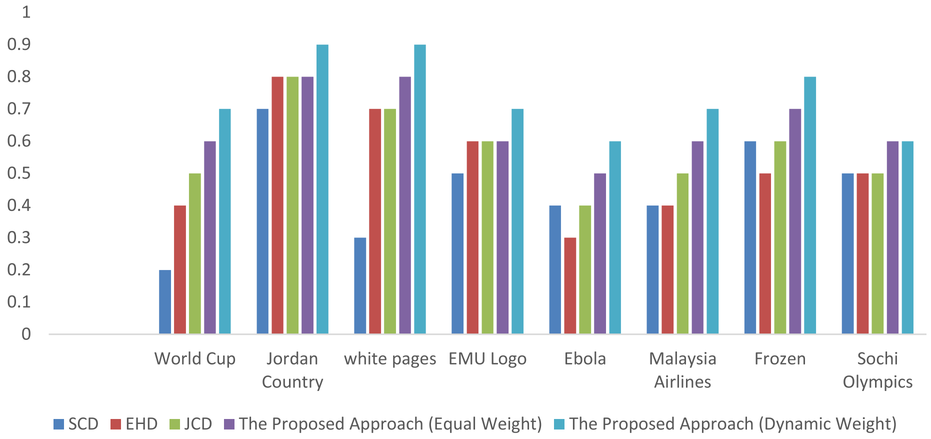

Fig. 1. The 10-precisions of SCD, EHD, JCD and the proposed approach for re-ranking a set of queries.

Figure 1 shows the 10-precisions of SCD, EHD, JCD and the proposed approach. As mentioned before, in the proposed approach, a weight is assigned for each descriptor; the effect of assigning equal or dynamical weights has also been investigated in this experiment. As shown in Fig. 1, the proposed approach outperforms the SCD, EHD and JCD descriptors. In addition, it has been found that dynamically assigning the descriptor weights is much better than the equal weights assignment.

4.2. *Experiment 2: Positions of the most relevant image files*

This experiment aims at finding the position of the relevant files for specific queries. Two independent evaluators are used to decide on the file relevancy. In more detail, the positions of the first five relevant files using Google and Yahoo search engines are found, and the results of this part are shown in Fig. 2. In addition, as shown in Fig. 3,

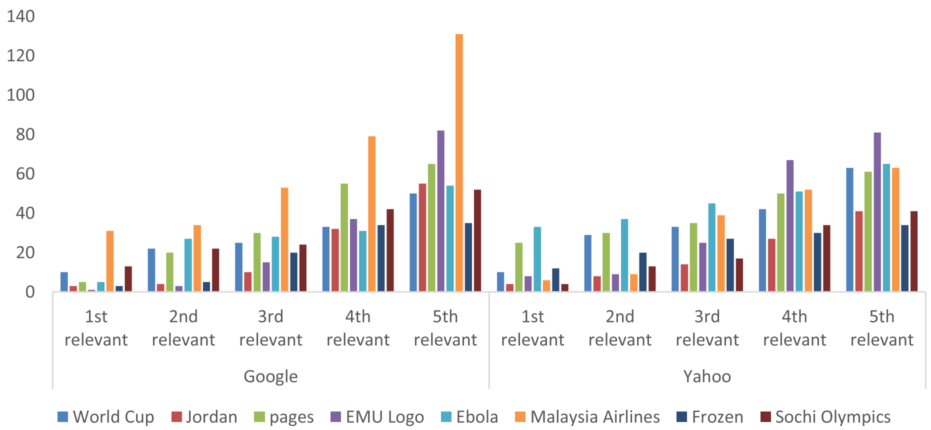

Fig. 2. The positions of first five relevant files for specific queries obtained using Google and Yahoo search engines.

the positions of first 10 relevant files for the same queries are obtained after re-ordering the first 100 files of the Google results using JCD, which are found to be the best compared to the other descriptors as shown in experiment 1 and the proposed re-ranking approach. Overall, it can be seen from Fig. 2 that the relevant files using Google and Yahoo can be distributed among a large number of results. Hence, the user has to check a large number of results to get the required files and such a process can be time-consuming. On the other hand, it is clear from Fig. 3 that the proposed re-ranking approach successfully re-orders the original results and shows the most relevant ones at the top of the list.

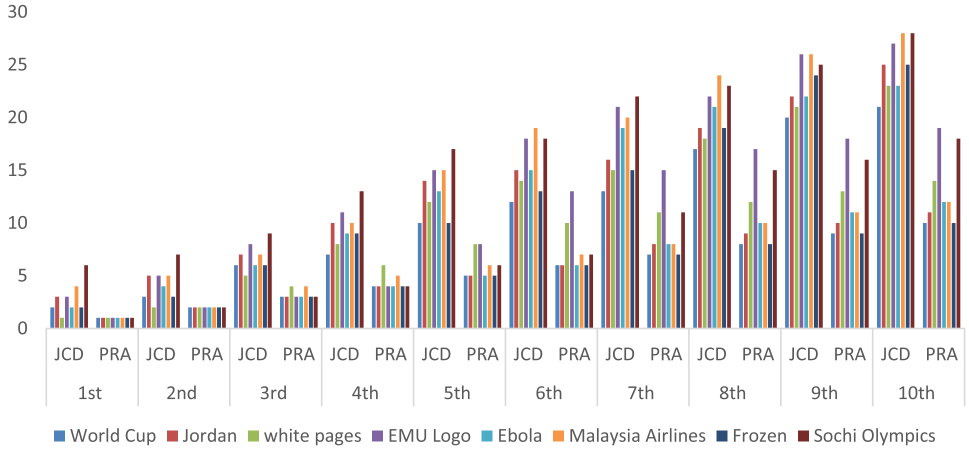

Fig. 3. The positions of first 10 relevant files for specific queries using JCD and the developed re-ranking approach in this study (PRA).

Table 1. The R-precision values for 25 image queries.

Query #	R-precision			Query #	R-precision		
	$R = 10$	$R = 20$	$R = 30$		$R = 10$	$R = 20$	$R = 30$
Query 1	0.8	0.75	0.8	Query 14	0.7	0.8	0.833
Query 2	1	0.95	0.85	Query 15	0.8	0.9	0.8
Query 3	1	1	1	Query 16	0.9	0.8	0.833
Query 4	1	0.9	0.866	Query 17	0.8	0.9	0.8
Query 5	0.9	0.85	0.833	Query 18	0.9	0.8	0.833
Query 6	1	0.95	0.9	Query 19	1	0.9	0.7
Query 7	0.9	0.8	0.8	Query 20	0.9	0.7	0.8
Query 8	1	1	0.9	Query 21	0.9	0.7	0.8
Query 9	0.8	0.75	0.8	Query 22	0.7	1	0.8
Query 10	0.9	0.75	0.8	Query 23	0.8	0.9	0.9
Query 11	0.8	0.833	0.8	Query 24	0.7	0.8	0.9
Query 12	0.9	0.8	0.9	Query 25	0.8	0.8	0.8
Query 13	0.9	0.9	0.9				

4.3. *Experiment* 3: *Scalability of the proposed re-ranking approach*

In this experiment, the performance of the developed re-ranking approach for a real image dataset is investigated. The used dataset includes 25,000 images, and 25 image queries have been constructed and used. The R-precision for the developed re-ranking approach is computed for $R = 10$, 20 and 30 by selecting a Target file for each query. Table 1 show the R-precision values for image queries. It is clear from this experiment that even when the database includes different subjects, the developed re-ranking approach is able to provide relevant retrieved files with increased accuracy and it shows the most relevant files at the top of the query results.

5. Conclusions and Future Works

Image retrieval is still a challenging field. One of the improved mechanisms is content-based re-ranking. A new image re-ranking approach that does not require any user intervention and implicitly can improve the retrieved images based on the file's contents and some user-specific actions is introduced. In addition, multiple descriptors where each has a dynamic weight that affects the file rank are used to describe image files accurately. Several experiments were conducted to evaluate the performance of the developed approach and we observed that the developed approach has the ability to show the most relevant files on top of the query results and can increase the percentage of the retrieved relevant files. On the other hand, similar to all existing re-ranking approaches, the main limitation of the approach is that it requires extra processing time as it will start processing after performing the regular ranking step.

This work can be extended by taking advantage of significant improvement achieved by CNN models in computer-based systems; for instance, the possibility of integrating a CNN model to assign the weight of the used descriptors can be investigated. In addition, adopting the developed approach for supporting other types of multimedia files can be another direction for future work.

References

Balan, S and T Devi (2012). Design and development of an algorithm for image clustering in textile image retrieval using color descriptors. *International Journal of Computer Science, Engineering and Applications*, 2(3), 199–211.

Chatzichristofis, SA and YS Boutalis (2008a). CEDD: Color and edge directivity descriptor: A compact descriptor for image indexing and retrieval. In *ICVS 2008: Computer Vision Systems*, Lecture Notes in Computer Science, Vol. 5008, pp. 312–322. Berlin: Springer.

Chatzichristofis, SA and YS Boutalis (2008b). FCTH: Fuzzy color and texture histogram-a low level feature for accurate image retrieval. In *Proceedings of the 2008 Ninth International Workshop on Image Analysis for Multimedia Interactive Services*, pp. 191–196. Piscataway: IEEE.

Chatzichristofis, SA, YS Boutalis and M Lux (2010). Combining color and spatial color distribution information in a fuzzy rule based compact composite descriptor. In *ICAART 2010: Agents and Artificial Intelligence*, Communications in Computer and Information Science, Vol. 129, pp. 49–60. Berlin: Springer.

Chen, C and J Ni (2011). Median filtering detection using edge based prediction matrix. In *IWDW 2011: Digital Forensics and Watermarking*, Lecture Notes in Computer Science, Vol. 7128, pp. 361–375. Berlin: Springer.

Chen, S, F Wang, Y Song and C Zhang (2011). Semi-supervised ranking aggregation. *Information Processing & Management*, 47(3), 415–425.

Craswell, N (2009). *R*-precision. In *Encyclopedia of Database Systems*, p. 2453. New York: Springer US.

Dali, L, B Fortuna, TT Duc and D Mladenić (2012). Query-independent learning to rank for RDF entity search. In *ESWC 2012: The Semantic Web: Research and Applications*, Lecture Notes in Computer Science, Vol. 7295, pp. 484–498. Berlin: Springer.

He, K, X Zhang, S Ren and J Sun (2016). Deep residual learning for image recognition. In *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778. Piscataway: IEEE.

Jain, V and M Varma (2011). 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. New York: ACM.

Jing, P, Y Su, C Xu and L Zhang (2018). HyperSSR: A hypergraph based semi-supervised ranking method for visual search reranking. *Neurocomputing*, 274, 50–57.

Krizhevsky, A, I Sutskever and GE Hinton (2012). ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems 25*, pp. 1097–1105. Red Hook, NY: Curran Associates.

Lafferty, J and C Zhai (2001). Document language models, query models, and risk minimization for information retrieval. In *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 111–119. New York: ACM.

Li, J, C Xu, W Yang, C Sun and D Tao (2017). Discriminative multi-view interactive image re-ranking. *IEEE Transactions on Image Processing*, 26(7), 3113–3127.

Liu, T-Y (2011). *Learning to Rank for Information Retrieval*. Berlin: Springer Science & Business Media.

Lu, M, Y Huang, M Xie and J Liu (2013). Rank hash similarity for fast similarity search. *Information Processing & Management*, 49(1), 158–168.

Lux, M and SA Chatzichristofis (2008). Lire: Lucene image retrieval: An extensible Java CBIR library. In *Proceedings of the 16th ACM International Conference on Multimedia*, pp. 1085–1088. New York: ACM.

Qi, S, X Wang, X Zhang, X Song and ZL Jiang (2018). Scalable graph based non-negative multi-view embedding for image ranking. *Neurocomputing*, 274, 29–36.

Qian, X, D Lu, Y Wang, L Zhu, YY Tang and M Wang (2017). Image re-ranking based on topic diversity. *IEEE Transactions on Image Processing*, 26(8), 3734–3747.

Simonyan, K and A Zisserman (2014). Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556 [cs.CV].

Singhal, R and SR Srivastava (2016). Enhancing the page ranking for search engine optimization based on weightage of in-linked web pages. In *Proceedings of the 2016 International Conference on Recent Advances and Innovations in Engineering*, pp. 1–5. Piscataway: IEEE.

Szegedy, C, W Liu, Y Jia, P Sermanet, S Reed, D Anguelov and A Rabinovich (2015). Going deeper with convolutions. In *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9. Piscataway: IEEE.

Wang, L, X Qian, Y Zhang, J Shen and X Cao (2019). Enhancing sketch-based image retrieval by CNN semantic re-ranking. *IEEE Transactions on Cybernetics*. doi:10.1109/TCYB.2019.2894498.

Yang, X, T Mei, Y Zhang, J Liu and SI Satoh (2016). Web image search re-ranking with click-based similarity and typicality. *IEEE Transactions on Image Processing*, 25(10), 4617–4630.

Yang, Y, D Xu, F Nie, J Luo and Y Zhuang (2009). Ranking with local regression and global alignment for cross media retrieval. In *Proceedings of the 17th ACM International Conference on Multimedia*, pp. 175–184. New York: ACM.

Yousuf, M, Z Mehmood, HA Habib, T Mahmood, T Saba, A Rehman and M Rashid (2018). A novel technique based on visual words fusion analysis of sparse features for effective content-based image retrieval. *Mathematical Problems in Engineering*, 2018, 2134395.

Yuan, M, D Yin, J Ding, Z Zhou, C Zhu, R Zhang and A Wang (2019). A multi-image joint re-ranking framework with updateable image pool for person re-identification. *Journal of Visual Communication and Image Representation*, 59, 527–536.

Zhu, Y, N Xiong, JH Park and R He (2008). A web image retrieval re-ranking scheme with cross-modal association rules. In *Proceedings of the 2008 International Symposium on Ubiquitous Multimedia Computing*, pp. 83–86. Piscataway: IEEE.