Visual Re-ranking Through Greedy Selection and Rank Fusion

Bin Lin, Ai Wei, and Xinmei Tian(✉)

Department of Electronic Engineering and Information Science,
University of Science and Technology of China, Hefei, Anhui, China
{lb1991,clytieai}@mail.ustc.edu.cn, xinmei@ustc.edu.cn

Abstract. Image search re-ranking has proven its effectiveness in the text-based image search system. However, traditional re-ranking algorithm heavily relies on the relevance of the top-ranked images. Due to the huge semantic gap between query and the image, the text-based retrieval result is unsatisfactory. Besides, single re-ranking model has large variance and is easy to over-fit. Instead, multiple re-ranking models can better balance the biased and the variance. In this paper, we first conduct label de-noising to filter false-positive images. Then a simple greedy graph-based re-ranking algorithm is proposed to derive the resulting list. Afterwards, different images are chosen as the seed images to perform re-ranking multiple times. Using the rank fusion, the results from different graphs are combined to form a better result. Extensive experiments are conducted on the INRIA web353 dataset and demonstrate that our method achieves significant improvement over state-of-the-art methods.

Keywords: Image search re-ranking · Rank fusion · Greedy graph-based re-ranking

1 Introduction

The importance of image searching has gained more and more attention to the society due to the explosive growth of the social media and multimedia information. Most web image search engines rely on textual information to determine the relevance between images and search keywords. Due to the lack of visual information and the lack of context of images, the searching results are unsatisfying at most of the time. Therefore, image re-ranking, which incorporates visual features of images to improve text-based image-searching, is introduced as the post-process of core search. It is defined as re-ordering the visual documents based on the initial text-based search results and their visual patterns [1]. With the help of the image re-ranking techniques, the quality of image search engine can be improved to a certain extent.

Image search re-ranking is mainly based on two assumptions. (1) The top ranked images are expected to possess the same semantic meaning with the query. (2) Images relevant to the query are expected to share similar visual

Q. Tian et al. (Eds.): MMM 2016, Part I, LNCS 9516, pp. 289–300, 2016.
DOI: 10.1007/978-3-319-27671-7_24

patterns more often than the irrelevant ones [2]. The key aspect of image re-ranking is the descriptive ability of the image feature and the robustness and strength of the image re-ranking model.

Researchers have proposed many informative local and holistic features to dig out the image information. Krupac et al. extracted SIFT features [4] on the dense grid and applied the BOW (bag of visual words) model [5] to present the image visual information [3]. Wang and Hua proposed an intuitive way to analyze the spatial distribution of color for desired images, and color map was applied for their interactive system to enhance the text-based image search [6]. Besides, other non-visual features can also help us determine the relevance between different images. Click information, as an example, can describe the relevance between images and queries accurately. Yu et al. employed the click data combined with the multi-model sparse coding method to build the image retrieval system [7]. Apart from that, multi-view features are also powerful features which can be combined with click features in the re-ranking model. Yu et al. proposed a re-ranking method using both click constraints and multi-view features to improve the retrieval performance [8]. Besides, different models are applied to conduct the re-ranking procedure. Yan et al. adopted pseudo relevance feedback which assumed that the top-ranked images were the few relevant ones. Those pseudo-relevant samples were further used in SVM to classify the remaining images into different classes [9]. Motivated by the well-known PageRank technique, Jing and Baluja proposed the Visual-Rank algorithm to treat images as the visual pages and analyze the visual link structures among images [10]. Tian et al. treated the re-ranking as a global optimization problem and proposed a Bayesian framework to derive the re-ranking model [11,12]. Yang and Hanjalic were inspired by the learning-to-rank paradigm and derived a re-ranking function in a supervised way from the human-labeled training data [13]. Luo and Tao proposed a manifold regularized multi-task method to learn a discriminative subspace to deal with multiple labels, thus images with different labels are divided [14]. Data mining techniques are also applied in image re-ranking model. Deng et al. proposed a weakly supervised multi-graph learning based on the mining of the intrinsic attributes among the instances [15]. Liu et al. proposed a noise-resistant graph and performed a graph ranking scheme to improve web image search results [16].

There are deficiencies existing in these re-ranking systems. First, among most of the re-ranking models, the credibility of images ranked on the top of the list contributes a lot to the system. However, the image set is usually filled with noisy samples, thus the performance of the retrieval system is often degraded. Therefore, specific de-noising method should be applied to avoid this circumstance. Second, we adopt graph-based learning, which can capture the intrinsic manifold structure underlying the images in the query, to replace traditional feature learning method as our re-ranking model. However, using just one graph for re-ranking is easy to over-fit for its low bias and high variance. Therefore, rank fusion method should be introduced to solve this problem.

We propose a framework where simple label de-noising is performed before re-ranking. Then a simple yet effective greedy graph-based re-ranking method is proposed for each query. Finally, multiple graphs are combined to reduce the high variance at the expense of a small increase in the bias and some loss of interpretability. It can significantly boost the performance. Extensive experiments are conducted on a web image data set to prove the necessity of our re-ranking architecture.

The remainder of this paper is organized as follows: Sect. 2 introduces our architecture of algorithm, including feature extraction, label de-noising, greedy selection, and rank fusion. Section 3 describes the experiments on the benchmark dataset which prove the effectiveness of our methods. Section 4 concludes the paper and raises a suggestion for future work.

2 Effective Visual Re-ranking

The key of image re-ranking is to find images relevant to the query and re-rank those images to the top of the list. With the large semantic gap between image and textual query, it is extremely hard to rank images only basing on their textual data. A large number of false positive samples are ranked on the top of the list as in Fig. 1, which makes it harder to re-rank images.

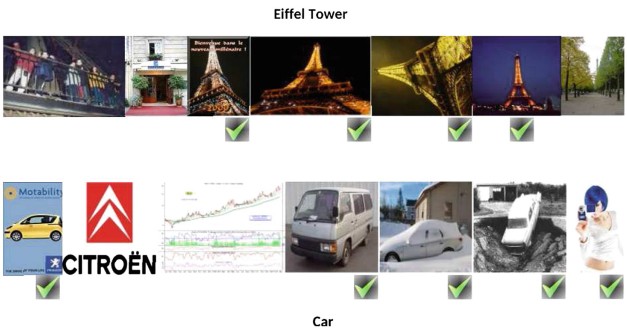

Fig. 1. Top-7 ranked images returned by a text-based image search engine for two queries: "Eiffel tower" and "car", ordered left to right. Query-relevant images are marked by the ticking sign. It illustrates that there are many irrelevant images lying on the top of the list.

Our architecture mainly consists of four parts. (1) Descriptive feature extraction and similarity calculation. We extract the informative holistic features from images and calculate the similarity matrix for each query. (2) Label de-noising. We define a simple confident criterion to evaluate the relevance between image and query. Then images with high confidence scores will replace those top-ranked images. We re-order the initial list by their confidence scores and create the "re-arranged list". (3) Graph-based re-ranking. After the pre-filtering, the "re-arranged list" is more relevant to the query. We select the image ranked 1st from the "re-arranged list" as the seed image and a simple graph-based greedy algorithm is proposed to incorporate the image which shares the highest similarity with the "seed image" as the relevant image. (4) Multiple graph fusion. Different image is selected as the "seed image" and we perform the graph-based re-ranking several times to generate multiple graphs. Rank fusion method is applied to combine different results into a more reasonable one.

2.1 Informative Feature Extraction

In order to capture the similarities between images, we first extract a powerful feature, which is the DCNN (deep convolutional neural network) feature, to present the holistic information of the image. Convolutional neural network was proposed by Le-Cun et al. [17] to solve the handwritten digit recognition problem. The convolutional neural network shows its superior ability in imitating the human biological vision system. The image can be well described using this network. Krizhevsky et al. [18] proposed a deep convolutional neural network with millions of parameters and applied it on the ImageNet dataset.

In this paper, DCNN feature is extracted on the network trained on the ILSVRC-2012 (ImageNet Large-Scale Visual Recognition Challenge) dataset. We select the 4096-dimensional vector which is induced from the fc7 level. The feature is normalized as in Eq. 1

$$\mathbf{x} = [x_1, x_2, .., x_N], x_k = \frac{\sqrt{f_k}}{\sum\limits_{m=1}^{N} f_m} \tag{1}$$

where N is the dimensionality of the vector and f is the input DCNN vector.

The measurement of similarity between images is crucial to image re-ranking system as well. We calculate the chi-square distance to show the visual distance between image i and image j as in Eq. 2

$$d_{ij} = \frac{\sum\limits_{k=1}^{N} \frac{(x_k^i - x_k^j)^2}{x_k^i + x_k^j}}{2}. \tag{2}$$

where x_k^i is the k-th feature of image i.

An inverse proportional function is applied as in Eq. 3 to transfer the distance into similarity score. s_{ij} denotes the similarity between image i and j and d_{ij} denotes the chi-square distance between two feature vector of images i and j. $\lambda = 0.5$ is chosen to avoid the situation where $d_{ij} = 0$.

$$s_{ij} = \frac{1}{d_{ij} + \lambda} \tag{3}$$

2.2 Label De-noising

In this part, our main goal is to filter out those false positive samples which are ranked on the front of the list. The procedure of filtering those images is based on a simple confidence score counting. After label de-noising, those outliers will be ranked lower, which significantly increase the reliability of the list.

Our label de-noising method is based on an intuition that relevant images share similar visual patterns with each other more often than irrelevant images [19]. We assume that the relevant images take up the majority of the query set. For each image, we sum up their similarities with every other image in the list. Thus we calculate the confidence score $C(I_j)$ for image i as in Eq. 4

$$C(I_j) = \sum_{i=1, i \neq j}^{N} \mu_i s_{ij}, \tag{4}$$

where s_{ij} denotes the similarity score between image i and image j and μ_i is a damping value which demonstrates the importance of the similarity with image i. When i is small, μ_i should be larger to increase the weight that come from the top-ranked images of the initial list because these images are more likely to share the same visual information with the query.

Images with high confidence scores are similar to most of the images in the list, which show their high probability to be relevant to the query. We simply move images with high confidence scores to the front of the list and thus the "re-arranged list" is created. The false positive samples will be removed from the front of the list due to their low confidence scores.

2.3 Graph-Based Re-ranking

The similarity matrix can be treated as the graph where each node represents a single image and the connectivity of a node reflects its visual similarity to others. Thus graph-based re-ranking is applied to derive the result image list.

After label de-noising, the high-ranked images are more credible. We treat the top-ranked image as the "seed image". Instead of finding images relevant to the query, we can simplify the task into finding images relevant to the "seed image" [20]. We apply a simple yet effective greedy algorithm to deal with the graph-based re-ranking problem.

Our goal is to find images which share more similarity with the "seed image set". At first, the "seed image set" only includes one image which is ranked 1st on the list. Then the image sharing the largest similarity value with images in the "seed image set" is included to the set. This process is iteratively conducted until all images in the query are included to the set. The algorithm is listed as in Algorithm 1.

The advantage of this method is that we only focus on images which are similar to the certificated relevant images. As long as those "certificated images" are truly relevant to the query, this method is fully reliable. Our experiment shows that the more credible the candidate images are, the more accurate the average precision will be.

Algorithm 1. Image Greedy Selection.

Input:

The initial image list, $L = [I_i]_{i=1}^{N}$;

The seed image set, $V = \emptyset$;

A given query, Q;

Output:

The image sequence set, Seq;

1: We start by choosing a seed image I_{seed} from L and add this image to the seed image set V.

2: Greedy Selection. We find the candidate image by calculating $I_{candidate} = \max_{i \in L} \sum_{j \in V} s_{ij}$ where s_{ij} denotes the similarity value between two images i and j.

3: $V = V \bigcup I_{candidate}$, $L = L - I_{candidate}$.

4: We iteratively conduct Step 2 and Step 3 until all images in the query set are included into set V.

5: We order images by their sequence of being added to the image set V and get the sequence list Seq.

6: **return** Seq;

2.4 Multiple Graph Fusion

Instead of using only one graph to derive the result lists, multiple graphs are generated to raise the credibility of the re-ranking results. In our greedy algorithm aforementioned, the result is closely related to the relevance of the seed image. In order to reduce the high variance, other high-ranked images on the lists are selected to be the "seed image". Therefore multiple graphs are generated using these different seed images. Different ranks are derived from these graphs and simple rank fusion method is applied to derive the final result. Three classic rank fusion methods are introduced.

Borda Fusion. The Borda fusion is a simple fusion method which turns the ranking information into the score [21]. For those top-ranked visual documents, higher scores are allocated to them. The Borda count function is defined as in Eq. 5.

$$s = \frac{num - R}{num - 1} \tag{5}$$

where num is the number of images in the list and R is the rank of this image. We sum up the ranking score for each image and order the list by their scores.

Condorcet Fusion. The Condorcet voting algorithm is a majoritarian method which specifies that the winner of the election is the candidate(s) that beats or ties with every other candidate in a pair-wise comparison [22]. For each iteration, we find a single image which beat every other image in a pair-wise comparison, then we exclude this image on the list and repeat the iteration until there is no image on the list. We order these images by their sequence of being excluded from the list.

RRF Fusion. Reciprocal Rank Fusion simply sorts the documents according to a naive scoring formula as in Eq. 6 [23].

$$RRFscore(d \in D) = \sum_{r \in R} \frac{1}{k + r(d)} \tag{6}$$

where k is a fixed number of 60. This formula derived from facts that while high-ranked documents are more important, the importance of lower-ranked documents still exists. This method is straightforward and effective. We order these images by their RRF scores.

3 Experiments

In this section, several experiments are conducted on a web image dataset called Web353 to prove the effectiveness of our re-ranking strategy.

3.1 Dataset

Our experiment is conducted on a diversified dataset - INRIA web353 dataset, which was collected by Krapac et al. [3]. This dataset includes 353 queries, where the original textual query is also included. For 80 % of queries, there are more than 200 images. Each image is resized to 150×150 pixels square. The ground-truth relevance label for every image is divided to two levels, which are "relevant" and "irrelevant". The 353 queries are diverse in topics, covering "object" items (e.g., "flag" and "car"), "celebrity" (e.g., "Justin Timberlake" and "will smith"), and abstract terms (e.g., "tennis court"), as shown in Fig. 2. Queries are diverse in ratio of relevance as well. In all, there are about 43.86 % images in this dataset labeled as relevant samples.

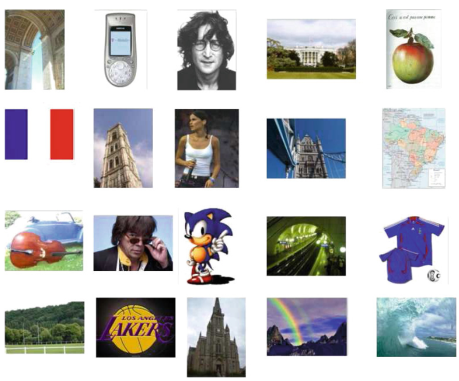

Fig. 2. Example pictures in INRIA web353 dataset.

3.2 Performance Metrics

Average precision (AP), which can reflect the occurrence of the relevant images, is adopted as our criterion to measure the effectiveness of the algorithm [18]. Average precisions are calculated at several truncation levels of T, i.e. $AP@T$; $T = \{5, 10, 20, 40, 60, 80, 100\}$, which reflect the precision for the top-T-ranked images, and they are defined in Eq. 7.

$$AP@T = \frac{1}{Z_T} \sum_{i=1}^{T} [precision(i) \times rel(i)] \tag{7}$$

$rel(i)$ is a binary function which reflects the relevance of the ith-ranked image. The precision value is the precision of top i-ranked images.

3.3 Experiments for Label De-noising

The “Search Engine” result is conducted based on the textual information on the meta-data file. It is listed as the baseline result for re-ranking.

To evaluate the power of label de-nosing, we compare this method with the Visual-Rank method [10]. From the Table 1, it is clear that after label de-noising, the MAP result is higher than Visual-Rank when T is larger than 20. Furthermore, we conduct Visual-Rank algorithm on the “re-arranged list”, the result is even better.

Table 1. MAP for evaluating label de-noising.

T	Search engine	Visual-Rank	“Re-arranged list”	“Re-arranged list” +Visual-Rank
5	0.611	0.799	0.743	**0.771**
10	0.553	0.743	0.715	**0.742**
20	0.503	0.656	0.676	**0.699**
40	0.452	0.552	0.633	**0.650**
60	0.431	0.557	0.612	**0.628**
80	0.426	0.567	0.605	**0.620**
100	0.431	0.581	0.610	**0.624**
ALL	0.569	0.679	0.699	**0.710**

3.4 Experiments for Greedy Selection

To evaluate the power of greedy selection, we propose two different schemes to test its performance.

Scheme 1: We treat the image on the top of the initial list as the “seed image” and then perform the greedy selection.

Table 2. The MAP result for greedy selection using different schemes.

T	Search engine	Scheme 1	Scheme 2	Scheme 3
5	0.611	0.670	0.773	**0.794**
10	0.553	0.664	0.754	**0.767**
20	0.503	0.655	0.728	**0.739**
40	0.452	0.639	0.689	**0.694**
60	0.431	0.629	0.667	**0.671**
80	0.426	0.627	0.656	**0.660**
100	0.431	0.634	0.657	**0.662**
ALL	0.570	0.715	0.736	**0.740**

Table 3. The number of graphs for graph fusion.

T	1	2	3	4	5	6	7	8	9	10
5	0.794	**0.818**	0.811	0.814	0.817	0.810	0.809	0.804	0.809	0.806
10	0.767	0.786	0.784	0.786	**0.789**	0.781	0.783	0.783	0.780	0.780
20	0.739	0.751	0.752	0.753	**0.757**	0.755	0.754	0.752	0.751	0.751
40	0.694	0.700	0.703	0.704	**0.708**	0.707	0.704	0.703	0.702	0.702
60	0.671	0.676	0.677	0.679	**0.683**	0.681	0.679	0.679	0.678	0.677
80	0.660	0.663	0.665	0.666	**0.670**	0.668	0.667	0.667	0.666	0.666
100	0.662	0.664	0.667	0.667	**0.670**	0.669	0.667	0.667	0.667	0.666
ALL	0.740	0.742	0.744	0.745	**0.748**	0.746	0.745	0.745	0.744	0.743

Scheme 2: We treat the image on the top of list after label de-noising as the "seed image" and then perform the greedy selection.

After simply conducting the greedy selection, the MAP for $AP@T$ gains a lot for every T. Adding label de-nosing can significantly improve the performance. $MAP@ALL$ achieves as high as 73.60 %. In comparison to Visual-Rank method, greedy selection beats it when T is larger than 20. But when T is small, the performance of greedy selection is quite poor. The reason for that is this method hugely depends on the relevance of images on the top. Visual-Rank performs poorly when T is large. But when we only focus on the small Ts, Visual-Rank achieves the best result. In order to get a better performance, we can sacrifice the time to perform Visual-Rank first and conduct the greedy selection based on the "Visual-Rank list". The result is listed as Scheme 3 in Table 2. And the performance is elevated as expected.

3.5 Experiments for Rank Fusion

Using one single graph to perform the re-ranking is of huge variance. Thus multiple graphs are introduced to boost the performance.

The Number of Graphs for Rank Fusion. First the number of graphs for each query is needed to confirm. Since employing Visual-Rank first and then conduct greedy selection can achieve a better result, we select the seed image from the "Visual-Rank" list. And Borda fusion is conducted first to evaluate the parameter. We evaluate the results while choosing the graph number ranging from 1 to 10. Table 3 shows that five is an appropriate number to choose since it beats other numbers when T is larger than 10.

Different Rank Fusion Methods. We perform Borda count method, Condorcet Fusion and RRF scoring method to test their performance. Five is selected as the number of graphs.

Table 4 shows that the RRF method is the best among these three methods. And the MAP@ALL achieves 74.94 %, which is one percent point higher than the situation when we don't conduct rank fusion.

3.6 Experiments to Compare with the State-of-Art Methods

In this part, we compare our results with state-of-the-art re-ranking methods, including pseudo relevance feedback [9], Bayesian re-ranking [11], query relative

Table 4. The experiment for testing different rank fusion methods.

T	Borda	Condorcet	RRF
5	0.817	0.817	**0.824**
10	0.789	0.792	**0.796**
20	0.757	0.759	**0.762**
40	0.708	0.709	**0.710**
60	0.683	0.683	**0.685**
80	0.670	0.670	**0.671**
100	0.670	0.671	**0.672**
ALL	0.748	0.749	**0.750**

Table 5. The experiment for comparing the state-of-the-art methods.

Methods	MAP
Search engine	0.569
PRF [9]	0.658
Bayesian [11]	0.665
Query relative [3]	0.666
Two-stage learning [24]	0.705
Noise-resistant [16]	0.736
Our method	**0.750**

re-ranking [3], two stage learning [24], and noise-resistant graph ranking [16]. Since most of them only reported the results on MAP@ALL, we compare the results under this metric. The results are listed in Table 5. Our method demonstrates better performance than the state-of-the-art re-ranking methods.

4 Conclusion

In this paper, we propose a simple graph-based re-ranking model. The contributions of our proposed method lie in three aspects. (1) A simple yet effective label de-noising method is conducted to filter out those false-positive images for each query. (2) A simple greedy selecting scheme is performed to fastly and accurately find the images relevant to the query. The experiment results show its superior performance over the Visual-Rank algorithm. (3) Different "seed images" are selected to conduct the greedy selection to avoid the huge biased. Then rank fusion is conducted to further improve the re-ranking performance. Extensive experiments are conducted on a web image dataset, which show the effectiveness of our method. There are many avenues for future explorations. To begin with, our greedy selection method is quite raw and straightforward. More delicate and carefully designed methods can be applied to further improve the selection performance. In addition, only the holistic information of the image is employed. Other useful feature can be applied to enrich the re-ranking model.

Acknowledgment. This work was supported by the 973 project under Contract 2015CB351803, by the NSFC under Contracts 61390514 and 61201413, by the Youth Innovation Promotion Association CAS No. CX2100060016, by the Fundamental Research Funds for the Central Universities No. WK2100060011 and No. WK2100100021, and by the Specialized Research Fund for the Doctoral Program of Higher Education No. WJ2100060003.

References

1. Mei, T., Rui, Y., Li, S., Tian, Q.: Multimedia search re-ranking: a literature survey. ACM Comput. Surv. **46**(3), 1–38 (2014)
2. Wang, X., Qiu, S., Liu, K., Tang, X.: Web image re-ranking using query-specific semantic signatures. IEEE Trans. Pattern Anal. Mach. Intell. **36**(4), 810–823 (2014)
3. Krapac, J., Allan, M., Verbeek, J., Jurie, F.: Improving web image search results using query-relative classifiers. In: Computer Vision and Pattern Recognition, pp. 1094–1101 (2010)
4. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision **60**(2), 91–110 (2004)
5. Jun, Y., Jiang, Y.G., Hauptmann, A.G., Ngo, C.-W.: Evaluating bag of-visual-words representations in scene classification. In: Proceedings of the International Workshop on Multimedia Information Retrieval, pp. 197–206 (2007)
6. Wang, J., Hua, X.-S.: Interactive image search by color map. ACM Trans. Intell. Syst. Technol. **3**(1), 12–37 (2011)

7. Yu, J., Rui, Y., Tao, D.: Click prediction for web image re-ranking using multimodal sparse coding. IEEE Trans. Image Process. **23**(5), 2019–2032 (2014)
8. Yu, J., Rui, Y., Chen, B.: Exploiting click constraints and multiview features for image re-ranking. IEEE Trans. Multimedia **16**(1), 159–168 (2014)
9. Yan, R., Hauptmann, A., Jin, R.: Multimedia search with pseudo relevance feedback. In: Image and Video Retrieval, pp. 238–247 (2003)
10. Jing, Y., Baluja, S.: Visualrank: applying pagerank to large-scale image search. IEEE Trans. Pattern Anal. Mach. Intell. **30**(1), 1877–1890 (2008)
11. Tian, X., Yang, L., Wang, J., Yang, Y., Wu, X., Hua, X.-S.: Bayesian video search re-ranking. In: Proceedings of the 16th ACM International Conference on Multimedia, pp. 131–140 (2008)
12. Tian, X., Yang, L., Wang, J., Wu, X., Hua, X.-S.: Bayesian visual re-ranking. IEEE Trans. Multimedia **13**(4), 639–652 (2011)
13. Yang, L., Hanjalic, A.: Supervised re-ranking for web image search. In: Proceedings of the International Conference on Multimedia, pp. 183–192 (2010)
14. Luo, Y., Tao, D., Geng, B., Xu, C., Maybank, S.J.: Manifold regularized multitask learning for semi-supervised multilabel image classification. IEEE Trans. Image Process. **22**, 523–536 (2013)
15. Deng, C., Ji, R., Liu, W., Tao, D., Gao, X.-B.: Visual re-ranking through weakly supervised multi-graph learning. In: International Conference on Computer Vision, pp. 2600–2607 (2013)
16. Liu, W., Jiang, Y.-G., Luo, J., Chang, S.-F.: Noise resistant graph ranking for improved web image search. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 849–856 (2011)
17. Le Cun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D.: Handwritten digit recognition with a backpropagation network. In: Advances in Neural Information Processing Systems, pp. 396–404 (1990)
18. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)
19. Morioka, N., Wang, J.: Robust visual re-ranking via sparsity and ranking constraints. In: Proceedings of the 16th ACM International Conference on Multimedia, pp. 533–542 (2011)
20. Zhang, S., Yang, M., Cour, T., Yu, K., Metaxas, D.N.: Query specific rank fusion for image retrieval. IEEE Trans. Pattern Anal. Mach. Intell. **37**(4), 803–815 (2015)
21. Aslam, J.A., Montague, M.: Models for metasearch. In: Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 276–284 (2001)
22. Montague, M., Aslam, J.A.: Condorcet fusion for improved retrieval. In: Proceedings of Acmcikm, pp. 538–548 (2002)
23. Cormack, G.V., Clarke, C., Buettcher, S.: Reciprocal rank fusion outperforms condorcet and individual rank learning methods. In: Proceedings of SIGIR, pp. 758–759 (2009)
24. Yang, L., Hanjalic, A.: Learning to re-rank web images. IEEE Multimedia **20**, 13–21 (2013)