Received August 23, 2019, accepted September 14, 2019, date of publication September 18, 2019, date of current version October 23, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2942142

Aggregation of Multiple Pseudo Relevance Feedbacks for Image Search Re-Ranking

WEI-CHAO LIN
Department of Information Management, Chang Gung University, Taoyuan 333, Taiwan
Healthy Aging Research Center, Chang Gung University, Taoyuan 333, Taiwan
Department of Thoracic Surgery, Chang Gung Memorial Hospital, Linkou 333, Taiwan

e-mail:viclin@gap.cgu.edu.tw

This work was supported in part by the Ministry of Science and Technology of Taiwan under Grant MOST 106-2410-H-182-024, in part by Chang Gung Memorial Hospital, Linkou, under Grant CMRPD3I0031, and in part by the Healthy Aging Research Center, Chang Gung University from the Featured Areas Research Center Program within the Framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan under Grant EMRPD1I0481 and Grant EMRPD1I0501.

ABSTRACT Image retrieval effectiveness can be improved by pseudo relevance feedback (PRF), which automatically uses top-k images of the initial retrieval result as the pseudo feedback. Since there are several different strategies for performing PRF leading to different search results, in this paper we focus on image search re-ranking by search result aggregation as a hybrid approach. In particular, different combinations of the original retrieval result with the result of PRF and the result of PRF by pseudo positive and negative feedbacks, using a strategy based on the Borda count, are compared. Our experiments, carried out on the NUS-WIDE-LIT and Caltech 256 datasets, demonstrate that search result aggregation can provide better retrieval performance than PRF. Specifically, the combination of the original result and the result of PRF by pseudo positive feedback performs the best.

INDEX TERMS Image retrieval, re-ranking, pseudo relevance feedback, Borda count.

I. INTRODUCTION

Content-based image retrieval (CBIR) focuses on automatically extracting low-level image features, such as color, texture, and shape, to index images. For the stage of retrieving images, the indexed image features are used for the similarity measurement. In this framework, each image is represented as a visual feature vector, which is a point in a high dimensional feature space. Therefore, the similarity between two images is based on the distance between their feature vectors. This indicates that two similar images are likely close to each other in the feature space. In other words, images close to the query can be simply regarded as they have similar contents to the query.

However, the semantic gap problem usually occurs in CBIR. It is the gap between the low-level features automatically extracted and measured for their similarities by computers and the high-level concepts or semantics in users' minds. In practice, CBIR systems are difficulty to effectively match users' needs [1]. Therefore, the retrieval performance of CBIR systems cannot achieve a high level of effectiveness.

The associate editor coordinating the review of this manuscript and approving it for publication was Lefei Zhang.

To cope with this problem, a post-retrieval or image re-ranking step can be employed. Image re-ranking can be defined as follows. For Web image search, given the initial text search, it returns a set of images ordered with descending similarity. The image re-ranking process is executed to reorder the images according to their visual similarities in order to improve the initial retrieval result [2], [3].

The pseudo relevance feedback (PRF) is a simple and widely used approach for improving initial search results. In traditional relevance feedback (RF), users are required to manually provide positive and/or negative feedbacks to the systems [4].

On the contrary, according to the initial retrieval result, PRF assumes that a significant fraction of top and bottom ranked images can be used as pseudo positive and negative feedbacks to the systems, respectively. According to these feedbacks, a model can be constructed to re-rank the search result set [2], [5], [6]. Particularly, this approach can iteratively re-modify the query vector based on the pseudo positive and/or negative feedbacks to move the query toward more relevant images and away from irrelevant ones in the multidimensional vector space where the images are represented by vectors [7].

Although PRF avoids the user-in-the-loop process required in RF, the top-retrieved images used as the pseudo relevance feedback set may contain some noise meaning that their semantic contents are not similar to the query one. In addition, for some query results many samples in the relevant set can be very similar or even (near) duplicate [8].

Another post-retrieval approach is based on combining multiple search results by search result aggregation. The idea behind this approach comes from the metasearch model in information retrieval where a query is introduced to a number of different search engines. Next, several lists of pages returned by different search engines are merged and the resulting ranked list is presented to the user [9]. Particularly, for a specific query, multiple search results are not necessarily obtained from different search engines but from different responses of a single search engine or retrieval model to different queries, such as textual and image features. One representative example is video retrieval by multiple modalities, where image features, audio signals, face detection, and caption information, can be used to improve text-based video search systems [2], [10].

Different from video retrieval, the aim of this paper is to present a simple but effective fusion approach for enhancing the retrieval performance of conventional CBIR systems. This approach is based on search result aggregation by combining multiple results including the original result itself and the results of PRF procedures for image re-ranking of these initial results. Specifically, the retrieval results of PRF include the result based on pseudo positive feedback alone and the result based on pseudo positive and negative feedbacks. In addition, the Borda count is applied as the combination method for search result aggregation.

This paper is organized as follows. Section II overviews related literatures including image re-ranking, pseudo relevance feedback, and search result aggregation. Section III presents the aggregation approach and experimental results. Finally, Section IV concludes the paper.

II. LITERATURE REVIEW

A. IMAGE RE-RANKING

In CBIR systems, images are first of all indexed based on their visual features as an off-line feature extraction process. Then, during retrieval, users can search images by providing some example(s) to the system to search similar images as the query by example approach. The images having similar visual features are retrieved and they are ranked based on the level of similarity to the query image. However, the initial search results often have a certain level of noise. Image re-ranking, which reorders images based on related textual and/or visual features, can be employed to improve initial image retrieval results [11].

In the literature, image search re-ranking approaches can be classified into off-line supervised learning and on-line unsupervised learning based methods. The off-line supervised learning (or classification) based approaches, which are similar to the task of 'learning-to-rank' [12], they are based on training some specific classification technique to assign a more relevance score to each image.

For the example of Duan *et al.* [13], relevant images are clustered by using both textual and visual features and each cluster is treated as a 'bag' and the images in the bag are denoted as 'instances'. In particular, positive and negative bags are used for classifier training, and a ranking score is used to rank all the bags.

Huang *et al*. [14], visual saliency and visual consistency are integrated for re-ranking where visual saliency assumes that salient images are often relevant to the users' query since they are easier to catch user's eyes and visual consistency is based on the concept that visually similar images are closely related to the search query. Particularly, the saliency model is trained to assign retrieved images some relevance scores in order to classify them into salient and cluttered classes, respectively.

On the other hand, Jain and Varma [15] believe that the clicked images corresponding to a query are mostly relevant to the query. Therefore, their re-ranking method promotes images, which are likely to be clicked, to the top of the ranked list. Particularly, the normalized click count for each image is predicted, and it is combined with the original ranking score for image re-ranking.

Wang *et al.* [3] focuses on automatic off-line learning scheme to associate different visual semantic spaces with different query keywords based on keyword expansions. Specifically, the visual image features are mapped into their corresponding visual semantic spaces to get semantic descriptors. For on-line re-ranking, the similarities between images are measured by the semantic signatures.

On the contrary, for on-line unsupervised learning based approaches, Wang *et al.* [16] propose the so called ContextRank procedure, which considers the difference of importance between target areas (i.e. the main objects) and background areas (i.e. the regions without the main objects) in images. That is, if visual words are close enough within an image; their links are constructed for intra-image context. On the other hand, two images having the same visual words are constructed by combining both feature similarity and spatial consistency for inter-image context. As a result, the score of images are the sum of scores of visual words in each image.

Yang *et al.* [17], an ordinal re-ranking approach is proposed, which adjusts the initial ranking list based on co-occurrence patterns, i.e. ordinal relationships between target semantics and low-level features extracted from the initial ranking list. Their experimental results show the outperformance of this approach over some representative baselines, such as RankSVM [18] and ListNet [19](Cao et al., 2007), in terms of effective and efficient video retrieval.

Besides dealing with image visual features alone, some works focus on multiple features or modalities simultaneously for search re-ranking. For example, Gao *et al.* [6] combine textual (i.e. social tags) and visual features to improve the keyword-based image search. Li *et al.* [20] use visual and geo-tags (i.e. GPS information) for example-based image retrieval. Yao *et al.* [21] consider visual, textual, and bag-of-words features for boosting retrieval precision.

Although the re-ranking search results by supervised learning based re-ranking approaches are promising, Liu and Mei [10] have shown that they cannot perform optimally and require very high computational complexity during off-line training and on-line computation for re-ranking. In addition, the re-training process is needed when more and more images are continually stored in the image database, which is usually the case in Web image search. Moreover, it is very difficult to collect enough training data and train classifiers for every possible concepts or classes in terms of Web scale image search re-ranking [22]. Therefore, we focus on image re-ranking through on-line unsupervised learning in this paper. Two related approaches are described hereafter.

B. PSEUDO RELEVANCE FEEDBACK

In addition to the above mentioned studies, image re-ranking can be simply approached by pseudo relevance feedback (PRF), which can be regarded as an on-line unsupervised learning approach. It is based on the concept of query point movement. More specifically, the query vector is iteratively modified by pseudo relevant and irrelevant feedbacks to move the query toward more relevant images and away from irrelevant ones in the feature space [4].

Relevance feedback can be defined as follows. Given an image database containing n images and an interface is provided for users to issue queries by image examples. Let Q be the query example and I an image in the database, which are denoted as $Q = (x_1, x_2, \ldots, x_m)$ and $I = (y_1, y_2, \ldots, y_m)$, in which there are m low-level visual features extracted.

However, from the user's viewpoint, not every retrieved image is semantically relevant to the query image Q. When the retrieval result cannot reach to user satisfaction, he/she can execute the iterative RF process. After some relevance feedback(s) are provided by users, a new retrieval result based on a new ranking list of top similar images can be produced, which can increase the level of user's satisfaction [4], [23].

The Rocchio algorithm is one representative approach for query vector modification [24]. The query is reformulated as the modified query $\vec{q}_m$ by

$$\vec{q_m} = \alpha\,\vec{q_0} + \beta \frac{1}{|D_r|} \sum_{\vec{d_j} \in D_r} \vec{d_j} - \gamma \frac{1}{|D_{nr}|} \sum_{\vec{d_j} \in D_{nr}} \vec{d_j} \qquad (1)$$

where q_n is the original query vector, D_r and D_{nr} are the relevant and non-relevant image sets, respectively, and α, β, and γ are the parameters to adjust the related weights for the three component, i.e. q_n, D_r, and D_{nr}.

In particular, α is the weight to make the original query vector to move toward a specific direction, and β and γ are the weights to reflect the level of importance of the relevant and irrelevant feedback sets, respectively. Then, the new reformulated query represented by the modified query vector is issued to the system to re-search similar images to it based on Euclidean metric.

Due to the limitations of RF that requires users to provide positive/negative feedbacks, pseudo relevance feedback (PRF) can be considered to automate the manual part of RF. In PRF, it begins when the relevant images are retrieved based on an initial query. According to the initial retrieval result, a fraction of the top-k ranked images is assumed as relevant to the query, which is called as pseudo-positive and some low ranked images are regarded as pseudo-negative. After the pseudo-positive and negative images are identified, the Rocchio algorithm is used to accomplish the RF process [7], [25].

One major advantage of performing PRF is its high computational efficiency for Web image search. However, although the re-ranking results by PRF may provide better retrieval effectiveness than initial search results, the major problem of PRF is that the pseudo feedback set may contain some images whose semantic contents are not similar to the query one. That is, some visually similar images in the feedback set are not semantically related to the query. Moreover, for Web image search, many images in the feedback set are mostly duplicate or near duplicate because of the high similarity of their visual features.

Note that, in reality, a good search result should not only contain relevant images to the query, but also cover a wide range of topics. In other words, the retrieval results should be relevant and diverse, which belong to the relevance-based re-ranking and diversified re-ranking problems, respectively [26]. In this paper, we mainly focus on the relevance-based re-ranking problem. This is because PRF is one of relevance-based re-ranking methods, which assigns higher rank to more relevant images without considering the coverage of diversified topics.

C. SEARCH RESULT AGGREGATION

Aggregation of multiple search results is another approach for image re-ranking. It is based on combining the ordered results from different retrieval methods or systems by rank aggregation [27]. This approach can be regarded as metasearch in information retrieval [9]. That is, the same query is issued to multiple search engines, where different retrieval results can be obtained. Next, different ranked lists produced by these engines are combined. As a result, the new re-ranked list is likely to increase the precision and the wide range of topics of the resulting list.

The Borda count procedure is a conventional method for combining multiple retrieval results. It optimally satisfies all the required symmetry properties for information retrieval [9], [28].

The Borda count procedure is a generalization of the majority vote. First of all, each voter (i.e. the retrieval system) ranks a fixed set of c candidates (i.e. the set of retrieved images). For each voter, c point is assigned to the top ranked candidate, $c - 1$ point for the second ranked candidate, and so on. Note that for some candidates that are unranked by the voter, remaining points are divided evenly among the unranked candidates. Finally, the fixed set of c candidates is ranked

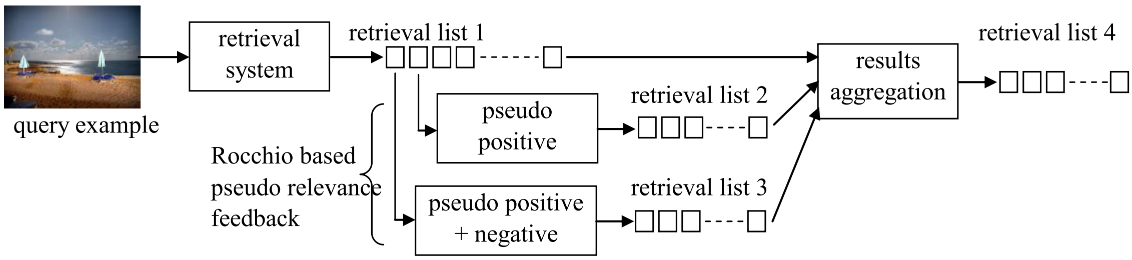

FIGURE 1. The proposed aggregation approach.

by their total points, and these points are used to re-rank the list [29].

Suppose that there are three retrieval results A, B, and C where their ranked lists of images for a given query are $A = (a, c, b, d)$, $B = (b, c, a, e)$, and $C = (c, a, b, e)$.

In total, five distinct images are retrieved, which are a, b, c, d, and e. The Borda count (BC) of each distinct image is computed by summing their Borda count values in individual results BC_A in the retrieval result (A) as follows:

$$\begin{aligned}
BC(a) &= BC_A(a) + BC_B(a) + BC_C(a) = 5 + 3 + 4 = 12 \\
BC(b) &= BC_A(b) + BC_B(b) + BC_C(b) = 3 + 5 + 3 = 11 \\
BC(c) &= BC_A(c) + BC_B(c) + BC_C(c) = 4 + 4 + 5 = 13 \\
BC(d) &= BC_A(d) + BC_B(d) + BC_C(d) = 2 + 0 + 0 = 2 \\
BC(e) &= BC_A(e) + BC_B(e) + BC_C(e) = 0 + 2 + 2 = 4
\end{aligned}$$

Finally, the five distinct images are re-ranked by their Borda counts. That is, the final ranked list of images is $c > a > b > e > d$.

III. THE AGGREGATION APPROACH

The proposed aggregation approach is based on the Borda count procedure to combine different retrieval results obtained by different methods to produce the final re-ranking result. Particularly, three different retrieval methods are considered, which include the original retrieval result without performing PRF, the re-ranked retrieval result by pseudo positive feedback, and the re-ranked retrieval result by pseudo positive and negative feedbacks.

Figure 1 shows the aggregation approach. When a user provides a query image to the retrieval system, the initial ranked retrieval result (i.e. retrieval list 1) is obtained based on the similarity between the query vector and the other feature vectors in the database. Next, top-k pseudo relevance feedback from retrieval list 1 is performed with the Rocchio algorithm. In particular, the top-k pseudo positive feedback set is used alone, resulting in a new ranked retrieval list (i.e. retrieval list 2). At the same time, the top-k pseudo positive and negative feedback sets are used together to produce another ranked retrieval list (i.e. retrieval list 3). Finally, the retrieval lists 1, 2, and 3 are combined by means of the Borda count procedure to produce the final re-ranking result (i.e. retrieval list 4).

According to the descriptions of Section 2.3, suppose that the original retrieval result (i.e. list 1) for a given query is $A = (a, c, b, d)$. Then, after performing PRF for list 1 based on pseudo positive feedbakc and pseudo positive + negative feedbacks, two different retrieval resutls (i.e. list 2 and list 3) can be obtained, represented by $D = (b, a, c, e)$, and $E = (c, a, e, b)$. In total, four different aggregation results by the Borda count procedure can be produced, which are list 1 + list 2, list 1 + list 3, list 2 + list 3, and list 1 + list 2 + list 3 (i.e. list 4), respectively.

The re-ranked results based on these four aggregation methods are as follows:

- list 1 + list 2 + list 3: (a, c, b, e, d)
- list 1 + list 2: (a, b, c, d, e)
- list 1 + list 3: (a, c, b, d, e)
- list 2 + list 3: (a, c, b, e, d)

For the following experiments, these single and aggregation retrieval methods are compared in terms of retrieval effectiveness and efficiency.

IV. EXPERIMENTS

A. EXPERIMENTAL SETUP

In this paper, two datasets are used for the experiments discussed below. The first dataset is based on NUS-WIDE-LITE[1] [30], which is a smaller version of the NUS-WIDE dataset. It contains a large number of real-world images downloaded from Flickr.[2] Particularly, 69 concepts (i.e. classes) are selected, which are composed of 22156 images, where each concept contains at least 50 images. The second dataset is based on Caltech 256.[3] It contains 257 object categories, which are composed of 30607 images. For each category, we randomly select 10 images as the query images. For image feature representation, the 500-D BoW feature based on the SIFT descriptor (BoW) [13] is extracted from each image.

In addition, we use the Euclidean distance similarity measure for the retrieval system since it is the most widely used distance function in image retrieval systems. For the pseudo positive and negative feedback sets, the top-20 highest ranked images and top-20 lowest ranked images of each query are used, respectively [31]–[33].

About the Rocchio parameters, since Moschitti [34] show that $\beta = \gamma$ is the best setting in text retrieval and Tsai *et al.* [35] compare different parameter settings for PRF in images,

[1] http://lms.comp.nus.edu.sg/research/NUS-WIDE.htm

[2] Flickr is an image hosting and video hosting website (https://www.flickr.com/).

[3] http://www.vision.caltech.edu/Image_Datasets/Caltech256/

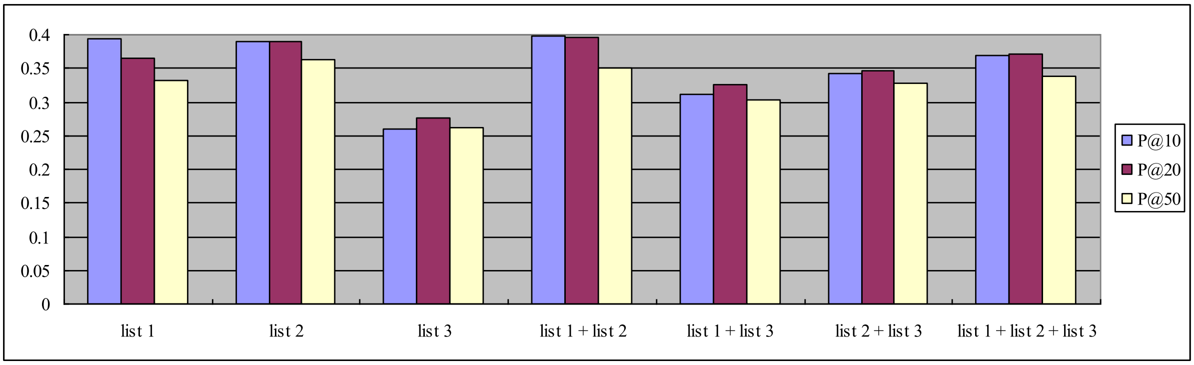

FIGURE 2. **The re-ranking performances over NUS-WIDE-LITE.**

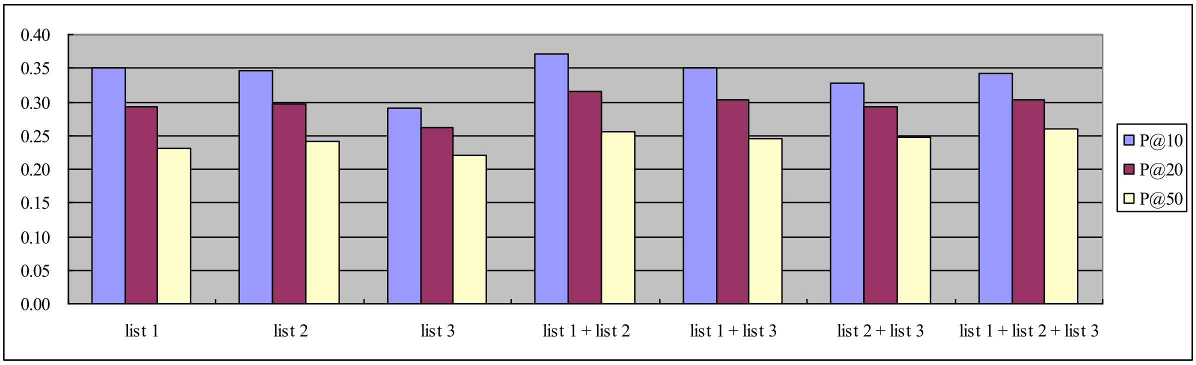

FIGURE 3. **The re-ranking performances over Caltech 256.**

they find that α, β, and γ set 1, 05, and 0.5, respectively can provide reasonably well performance over different image feature representations. Therefore, we follow the parameter setting of [35]. Moreover, for retrieval efficiency, feedback iteration is only performed once to collect the retrieval lists 2 and 3. We found that the performance gradually degrades when more feedback iterations by PRF are executed, which may be affected by some images in the pseudo feedback set whose semantic contents are different from the query one. Therefore, other results of PRF obtained with different feedback iterations are not compared.

Finally, the top 100 retrieved images from the retrieval lists 1, 2, and 3 are used for result aggregation by Borda count. This is because the re-ranking of all the retrieved images will lead to a very large computational cost, and will affect the query response time. Moreover, for each query result, users are usually not concerned with the lower ranked images. Therefore, re-ranking the top 100 retrieved images from different retrieval results is enough.

Consequently, in our aggregation approach, there are four possible combinations of three retrieval lists by Borda count, which are (1) list 1 + list 2, (2) list 1 + list 3, (3) list 1 + list 2 + list 3, and (4) list 2 + list 3.

B. EXPERIMENTAL RESULTS

Figures 2 and 3 show the re-ranking results of different aggregation strategies for the NUS-WIDE-LITE and Caltech 256 datasets, respectively. For the results of single retrieval methods (i.e. lists 1, 2, and 3), we can observe that the original result is slightly better than the one obtained by PRF in terms of P@10, which are 39.42% vs. 39% over NUS-WIDE-LITE and 35.02% vs. 34.67% over Caltech 256. However, performing PRF based on pseudo positive feedback can provide slightly better P@20 and P@50 rates for the two datasets. However, there is no significant difference between list 1 and list 2 (i.e. $p > 0.05$ by t-test).

On the other hand, for the aggregation result the combination of list 1 and list 2 (i.e. list 1 + list 2) performs best, which provides the highest rates of P@10 and P@20 over both datasets. In particular, PRF based on pseudo positive feedback (i.e. list 2) performs best in terms of P@50 and the second best in terms of P@10 and P@20 over NUS-WIDE-LITE. For Caltech 256, the combination of list 1 and list 3 performs the second best in terms of P@10, P@20, and P@50.

Table 1 lists the average retrieval performances of these seven methods. As we can see, the aggregation of lists 1 and 2 is the best choice for image search re-ranking, which performs significantly better than the others (i.e. $p < 0.05$).The exception is list 2 over the NUS-WIDE-LITE dataset where the performance difference between list 1 + list 2 and list 2 is 0.04% only.. It is interesting to note that, except for the combination of lists 1 and 2, other aggregation results do not perform better than the original retrieval result and PRF by pseudo positive feedback over NUS-WIDE-LITE. However, for the Caltech 256 dataset, the aggregation

TABLE 1. Average retrieval performances of P@10, P@20, and P@50.

	list 1	list 2	list 3	list 1 + list 2	list 1 + list 3	list 2 + list 3	list 1 + list 2 + list 3
NUS-WIDE-LITE	36.41%	38.13%	26.59%	38.17%	31.35%	33.84%	35.99%
Caltech 256	29.14%	29.5%	25.76%	31.46%	30%	28.98%	30.14%

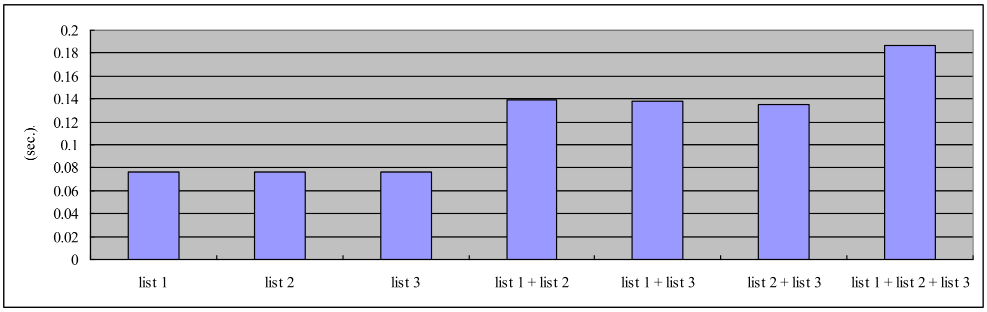

FIGURE 4. Average query response time over NUS-WIDE-LITE.

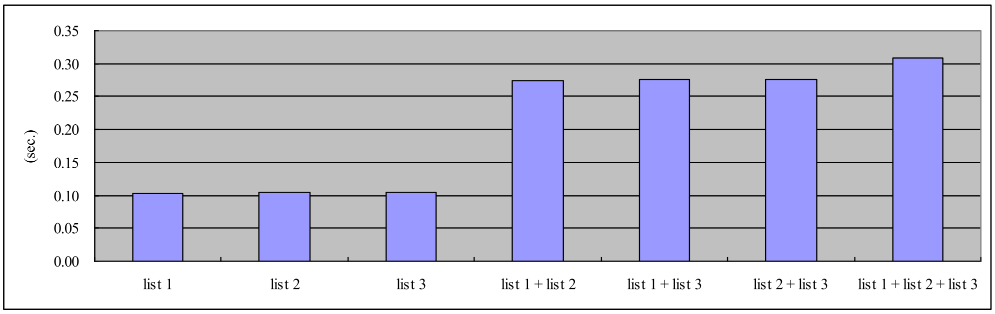

FIGURE 5. Average query response time over Caltech 256.

strategies, except for list 2 + list 3, outperform the three single retrieval methods. This performance difference could be affected by the image content. This may be because NUS-WIDE-LITE contains about half real (single object) and half abstract (non-single object) categories. For example, a harbor image is an abstract category that is composed of several single objects. On the other hand, categories in Caltech 256 are all single objects. This indicates that search result aggregation is much more suitable for the images containing single objects.

Furthermore, we examine the average query response times of different aggregation strategies applied on the NUS-WIDE-LITE and Caltech 256 datasets.[4] Note that each query response time is measured from issuing a query image example to the final retrieval results retrieved. Figures 4 and 5 show obviously that combining multiple retrieval results takes longer time than using single retrieval results. However, although the best performance with list 1 + list 2 requires 0.139 second over NUS-WIDE-LITE and 0.275 second over Caltech 256, which is higher than list 1, list 2, and list 3, for users this performance difference is very small because it is difficult to clearly differentiate between them during retrieval. In other words, the query response time by list 1 + list 2 is still efficient enough. In addition, if we consider fewer images for re-ranking, the query response time will become much shorter.

Therefore, according to the retrieval effectiveness and efficiency points of view, combining the original retrieval result and the one obtained by performing PRF based on pseudo positive feedback can provide the best re-ranking result.

V. CONCLUSION

In this paper, we present a hybrid approach for image search re-ranking in order to improve retrieval effectiveness of CBIR systems and PRF. This approach is based on search result aggregation. Particularly, three kinds of single retrieval results are combined, which are the original retrieval result,

[4]The software is based on Matlab 7 on an Intel Pentium 4 computer, with a 2.8GHz CPU, and 20GB RAM.

the result of PRF by pseudo positive feedback, and the result of PRF by pseudo positive and negative feedbacks.

Two datasets including a small version of NUS-WIDE-LITE and Caltech 256 (c.f. Section 3.1) are used in our experiments. The experimental results show that search result aggregation can provide better retrieval performance than the single retrieval results. Specifically, the combination of the original result and the result of PRF by pseudo positive feedback performs the best. On the other hand, although the average retrieval time per query by search result aggregation is longer than the single retrieval methods, the query response time is still very short to re-rank top 100 images of different retrieval results, i.e. 0.139 second for the NUS-WIDE-LITE dataset and 0.275 second for the Caltech 256 dataset.

For future work, several issues could be considered as the factors affecting the image re-ranking result. Since there are different distance functions, such as cosine measure, for image retrieval, they can be used for further comparisons. Similarly, different image features, such as color histogram and wavelet texture, can be extracted from images for image feature representations. In addition, some supervised learning techniques, such as co-training [36], can be employed for search result cooperation. Finally, larger image datasets can be used to examine the scalability of the aggregation approach.

REFERENCES

[1] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, ''Content-based image retrieval at the end of the early years,'' *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 12, pp. 1349–1380, Dec. 2000.

[2] W. H. Hsu, L. S. Kennedy, and S. F. Chang, ''Reranking methods for visual search,'' *IEEE MultimediaMag.*, vol. 14, no. 3, pp. 14–22, Jul. 2007.

[3] X. Wang, K. Liu, and X. Tang, ''Query-specific visual semantic spaces for Web image re-ranking,'' in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 857–864.

[4] X. S. Zhou and T. S. Huang, ''Relevance feedback in image retrieval: A comprehensive review,'' *Multimedia Syst.*, vol. 8, no. 6, pp. 536–544, Apr. 2003.

[5] K. S. Lee, W. B. Croft, and J. Allan, ''A cluster-based resampling method for pseudo-relevance feedback,'' in *Proc. ACM Int. ACM SIGIR Conf. Res. Develop. Inf.*, 2008, pp. 235–242.

[6] Y. Gao, M. Wang, Z. J. Zha, J. Shen, X. Li, and X. Wu, ''Visual-textual joint relevance learning for tag-based social image search,'' *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 363–376, Jan. 2013.

[7] Y. Rui and T. S. Huang, ''A novel relevance feedback technique in image retrieval,'' in *Proc. ACM Int. Conf. Multimedia*, 1999, pp. 67–70.

[8] Q. Xing, Y. Zhang, and L. Zhang, ''On bias problem in relevance feedback,'' in *Proc. ACM Int. Conf. Inf. Knowl. Manage.*, 2011, pp. 1965–1968.

[9] J. A. Aslam and M. Montague, ''Models for metasearch,'' in *Proc. ACM 24th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf.*, 2001, pp. 276–284.

[10] Y. Liu and T. Mei, ''Optimizing visual search reranking via pairwise learning,'' *IEEE Trans. Multimedia*, vol. 13, no. 2, pp. 280–291, Feb. 2011.

[11] T. Mei, Y. Rui, S. Li, and Q. Tian, ''Multimedia search reranking: A literature survey,'' *ACM Comput. Surv.*, vol. 46, no. 3, 2014, Art. no. 38.

[12] T. Joachims, ''Optimizing search engines using clickthrough data,'' in *Proc. ACM Int. Conf. Knowl. Discovery Data Mining*, 2002, pp. 133–142.

[13] L. Duan, 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, Nov. 2011.

[14] J. Huang, X. Yang, X. Fang, W. Lin, and R. Zhang, ''Integrating visual saliency and consistency for re-ranking image search results,'' *IEEE Trans. Multimedia*, vol. 13, no. 4, pp. 653–661, Aug. 2011.

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

[16] S. Wang, Q. Huang, S. Jiang, L. Qin, and Q. Tian, ''Visual ContextRank for Web image re-ranking,'' in *Proc. ACM Int. Workshop Large-Scale Multimedia Retr. Mining*, 2009, pp. 121–128.

[17] Y. H. Yang, W. H. Hsu, and H. H. Chen, ''Online reranking via ordinal informative concepts for context fusion in concept detection and video search,'' *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 12, pp. 1880–1890, Dec. 2009.

[18] T. Joachims, ''Training linear SVMs in linear time,'' in *Proc. ACM Int. Conf. Knowl. Discovery Data Mining*, 2006, pp. 217–226.

[19] Z. Cao, T. Qin, T.-Y. Liu, M.-F. Tsai, and H. Li, ''Learning to rank: From pairwise approach to listwise approach,'' in *Proc. Int. Conf. Mach. Learn.*, 2007, pp. 129–136.

[20] J. Li, X. Qian, Y. Tang, L. Yang, and T. Mei, ''GPS estimation for places of interest from social users' uploaded photos,'' *IEEE Trans. Multimedia*, vol. 15, no. 8, pp. 2058–2071, Dec. 2013.

[21] T. Yao, C. W. Ngo, and T. Mei, ''Circular reranking for visual search,'' *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1644–1655, Apr. 2013.

[22] L. Kennedy and M. Naaman, ''Generating diverse and representative image search results for landmarks,'' in *Proc. ACM Int. Conf. World Wide Web*, 2008, pp. 297–306.

[23] W.-C. Lin, Z.-Y. Chen, S.-W. Ke, C.-F. Tsai, and W.-Y. Lin, ''The effect of low-level image features on pseudo relevance feedback,'' *Neurocomputing*, vol. 166, pp. 26–37, Oct. 2015.

[24] J. Rocchio, ''Relevance feedback in information retrieval,'' in *The SMART Retrieval System—Experiments in Automatic Document Processing*, G. Salton, Ed. Upper Saddle River, NJ, USA: Prentice-Hall, 1971, pp. 313–323.

[25] R. Yan, A. Hauptmann, and R. Jin, ''Multimedia search with pseudo-relevance feedback,'' in *Proc. Int. Conf. Image Video Retr.*, 2003, pp. 238–247.

[26] X. Tian and D. Tao, ''Visual reranking: From objectives to strategies,'' *IEEE Multimedia*, vol. 18, no. 3, pp. 12–20, Mar. 2011.

[27] S. Zhang, M. Yang, T. Cour, K. Yu, and D. Metaxas, ''Query specific fusion for image retrieval,'' in *Proc. Eur. Conf. Comput. Vis.*, 2012, pp. 660–673.

[28] C. Dwork, R. Kumar, M. Naor, and D. Sivakumar, ''Rank aggregation methods for the Web,'' in *Proc. ACM Int. Conf. World Wide Web*, 2001, pp. 613–622.

[29] R. Nuray and F. Can, ''Automatic ranking of information retrieval systems using data fusion,'' *Inf. Process. Manage.*, vol. 42, pp. 595–614, May 2006.

[30] T.-S. Chua, J. Tang, R. Hong, H. Li, Z. Luo, and Y. Zheng, ''NUS-WIDE: A real-world Web image database from National University of Singapore,'' in *Proc. ACM Int. Conf. Image Video Retr.*, 2009, Art. no. 48.

[31] G. Giacinto and F. Roli, ''Bayesian relevance feedback for content-based image retrieval,'' *Pattern Recognit.*, vol. 37, no. 7, pp. 1499–1508, 2004.

[32] T. Deselaers, R. Paredes, E. Vidal, and H. Ney, ''Learning weighted distances for relevance feedback in image retrieval,'' in *Proc. Int. Conf. Pattern Recognit.*, Dec. 2008, pp. 1–4.

[33] J.-H. Su, W.-J. Huang, P. S. Yu, and V. S. Tseng, ''Efficient relevance feedback for content-based image retrieval by mining user navigation patterns,'' *IEEE Trans. Knowl. Data Eng.*, vol. 23, no. 3, pp. 360–372, Mar. 2011.

[34] A. Moschitti, ''A study on optimal parameter tuning for Rocchio text classifier,'' in *Proc. Eur. Conf. Inf. Retr.*, 2003, pp. 420–435.

[35] C.-F. Tsai, Y.-H. Hu, and Z.-Y. Chen, ''Factors affecting Rocchio-based pseudorelevance feedback in image retrieval,'' *J. Assoc. Inf. Sci. Technol.*, vol. 66, no. 1, pp. 40–57, 2014.

[36] F. Wei, W. Li, W. Wang, and Y. He, ''iRANK: An interactive ranking framework and its application in query-focused summarization,'' in *Proc. ACM Int. Conf. Inf. Knowl. Manage.*, 2009, pp. 1557–1560.

WEI-CHAO LIN received the Ph.D. degree from the University of Sunderland, U.K. He is currently an Associate Professor with the Department of Information Management, Chang Gung University, Taoyuan, Taiwan, the Healthy Aging Research Center, Chang Gung University, and the Department of Thoracic Surgery, Chang Gung Memorial Hospital, Linkou, Taiwan. He has published over 30 journal articles, such as *Information Sciences*, *Journal of Systems and Software*, *Applied Soft Computing*, *Neurocomputing*. His current research interests include data mining, machine learning, and information retrieval.