MQSearch: Image Search by Multi-Class Query

Yiwen Luo[1] **Wei Liu**[1] **Jianzhuang Liu**[1]
Department of Information Engineering[1]
The Chinese University of Hong Kong
Hong Kong
{ywluo6, wliu5, jzliu}@ie.cuhk.edu.hk

Xiaoou Tang[1,2]
Visual Computing Group[2]
Microsoft Research Asia
Beijing, P.R. China
xitang@microsoft.com

ABSTRACT
Image search is becoming prevalent in web search as the number of digital photos grows exponentially on the internet. For a successful image search system, removing outliers in the top ranked results is a challenging task. Typical content based image search engines take an input image from one class as a query and compute relevance between the query and images in a database. The results often contain a large number of outliers, since these outliers may be similar to the query image in some way. In this paper we present a novel search scheme using query images from multiple classes. Instead of conducting query search for one image class at a time, we conduct multi-class query search jointly. By using several query classes that are similar to each other for multi-class query, we can utilize information across similar classes to fine tune the similarity measure to remove outliers. This strategy can be used for any information search application. In this work, we use content based image search to illustrate the concept.

Author Keywords
Image Retrieval, Ranking, Multi-Class

ACM Classification Keywords
H.5.2 User Interfaces: [Graphical user interfaces (GUI), Prototyping]; H.3.3 Information Search and Retrieval: [Clustering, Relevance feedback, Search Process]

INTRODUCTION
With the popularization of digital camera and mobile phone camera and the rapid development of the internet, people now have easy access to unlimited collections of digital images. It becomes increasingly important to design effective image search tools to help users to find images on the web more efficiently. One key challenge with existing image search engines, either surrounding text based or content based, is how to remove the large number of outliers in the top ranking search results effectively. Currently, all methods focus on using queries from one class of images to conduct the search. In this paper we present a novel idea using query images from multiple classes to conduct multi-class query search jointly. Since a key reason for outliers to exist is that the outliers are similar to the query image in some way that may cause the confusion. By using several query classes that are similar to each other, we can utilize information across similar classes to fine tune the similarity measure to remove outliers. This strategy can be used for any information search application. In this work, we use content based image search to illustrate the concept.

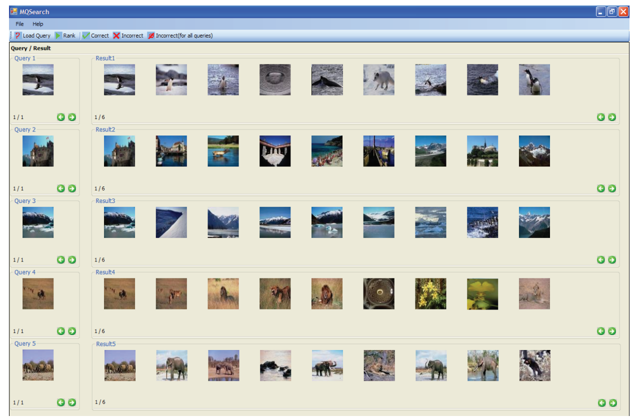


Figure 1: A screen shot of our system MQSearch.

Related Work
Content-based image search works this way: a query image is submitted to the system to show the user's "class of interest" and the system ranks the images in its database with a suitable metric and returns a list of images which are closest to the query [1], [2]. For general image search, image similarity is obtained based on image properties such as colors, textures, and shapes. The low-level features work in many cases. However, the user's idea of similarity is often in a high level where the similarity is hard to be measured with a metric. Thus, a large number of semantic outliers appear in the top ranking [3]. So far, it is still a difficult task to extract high level features from images. Therefore, relevance feedback is introduced to pursuit a result closer to the user's interest. During relevance feedback, the user labels a subset of images as positive and/or negative according to an abstract metric defined in the mind and the system refines its metric and displays another set of images hopefully closing the gap between the user's interest and the responses of the system. It is an iterative process until the found images con-

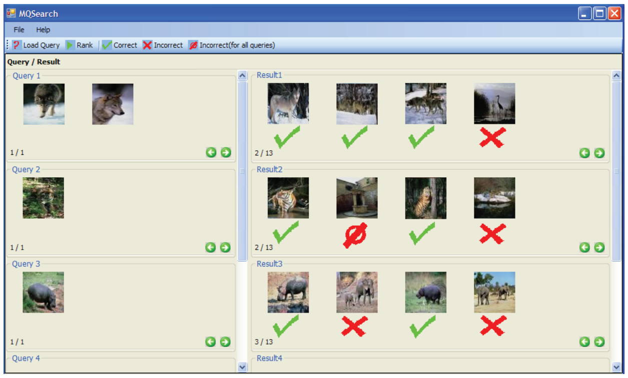


Figure 2: Queries grouping and image annotation.

verges to the user's class of interest [4], [5], [6], [7]. The performance of a relevance feedback system is usually evaluated as the number of the user's responses (or iterations) necessary to focus on the targeted class. It depends on how well the metric generalizes on the unlabeled images and how much information the user responses to the system. To release the user from the tiresomely long iterative process, we propose a re-ranking scheme which can remove outliers effectively in this paper.

Many web search techniques are designed for searching user interested contents on the internet, in which ranking web pages is a core procedure. The most popular search engine "Google" [9] accomplishes web page ranking using the PageRank algorithm that exploits the global rather than local hyperlink structure of the web [8]. Recently, ranking data objects (e.g. text and images) represented as vectors in the Euclidean space attracts much attention in image search research. The manifold ranking algorithm [11] is used to rank the data with respect to the intrinsic global manifold structure collectively revealed by a huge amount of the data. For many real world data types, the algorithm demonstrates superiority to traditional methods, which locally rank data simply by pairwise Euclidean distances or inner products. It motivates other practical applications, such as clustering through ranking [12] and ranking-based image retrieval [10]. However, one of the main weaknesses of the manifold ranking algorithms is its difficulties in preventing unrelated objects from appearing in the top ranking. This is because the ranking algorithm only considers the information of the query's class, but ignores the information of other classes. To address this problem, we propose a multi-class query ranking algorithm to rank data objects with concurrent queries from different classes.

Our Method

The contributions of our multi-class query image search system are described as follows.

1. Multi-query search

   Our system (Fig. 1), called MQSearch, introduces a novel search user interface(UI) for processing different querying tasks concurrently instead of doing them one by one. This search scheme is not only fast but also removes outliers in the top ranking more effectively. Our algorithm simultaneously ranks multiple classes of interest in queries by learning a ranking function corresponding to each query concurrently. Through enforcing graph regularization, the algorithm can also remove outliers effectively. In contrast, single query ranking only considers the information of the query's class, but ignores the information of the other classes.

2. Image annotation

   To cooperate with our multi-query search process, our system supports an novel UI for annotating the displayed images. In addition to just labeling the images as positive or negative as in previous feedback systems, we support more choices to the user: annotating the image as positive to its assumed class, positive to another class, negative to its assumed class, or negative to all the query classes. Such a feedback scheme fits to our multi-query search algorithm and improves the re-ranking efficiency.

3. Re-ranking

   After the user annotates the displayed images, images with positive annotation to some class will be considered as positive queries to this class and otherwise negative. The image annotation and re-ranking are combined to an iterative process and our experiments show that the number of iterations to converge is reduced than ranking each query one by one.

SYSTEM DESCRIPTION

Our system is divided into three parts: initial multi-query search, annotation of searched images, and image re-ranking.

Multi-Query Search

Our system allows the user to cluster multiple queries in different classes conveniently. As presented in Fig. 2, the user groups the query images into different categories in the left hand side boxes. The images can be moved from one group to another by drag-and-drop operation. Instead of searching for each class one by one, MQSearch supports multi-query multi-class search concurrently, returning the results corresponding to each class. Technical details of this multi-query ranking algorithm is described in Section Algorithm Description. Notice that all current image search systems allow only one class of query images as search query to search one class images.

Image Annotation

After the system gives a response to the user's queries, the UI supports the user to specify whether some images appearing in the displaying lists are outliers. For the queries in our system belonging to different classes, the feedback scheme is more complicated than that in a single query search system. Instead of specifying an image only as positive or negative to a query's class, we support three kinds of annotation: (i) an image is positive to its current class; (ii) an image is negative to its current class; (iii) an image is negative to all the query classes. Some image annotations are shown in the right hand side boxes of Fig. 2. The UI also supports the drag-and-drop operation to move an image from one class to another. The user provides the information for the next search process.

Image Re-ranking

Corresponding to the three kinds of annotation, the system gathers the information and searches again, when an image is annotated as positive to its current class, the system adds this image to the positive queries in its current class; when an image is annotated as positive to another query class, the system adds this image to not only the positive queries of the class specified by the user, but also the negative queries in its current class; when an image is labeled as negative to all the query classes, the system adds this image to the negative queries in all the classes.

ALGORITHM DESCRIPTION

Inspired by the manifold ranking algorithm in [11], we construct a weighted graph $G(V,E,W)$ that represents the intrinsic manifold structure of the data where each vertex(point) denotes an image represented by its feature vector. For convenience, we assume the graph is connected, which is satisfied if a sufficient number of edges are covered in E. The weight (affinity) matrix is formed by $W_{ij} = \exp(-d_{ij}^2/2\sigma^2)$ if $i \neq j$ and $W_{ii} = 0$, where d_{ij} is the distance between two points i and j. For each class C_k $(k = 1, \cdots, q)$, we define an initial score vector $\mathbf{y} = [y_1, \cdots, y_n]^T$, in which $y_i = 1$ if image i is a positive query in C_k, $y_i = -1$ if image i is a negative query, $y_i = 0$ if image i is not a query, and n is the number of images in the database. A negative query to C_k means that it is either positive to some other class or negative to all the classes the current queries are in.

The regularization framework proposed in semi-supervised learning can be modified to learn the current ranking function f by minimizing the following cost functional

$$\mathcal{Q}(f,f') = \frac{\alpha}{2}\sum_{i,j=1}^{n}\left(\frac{f'(\mathbf{x}_i)}{\sqrt{D_{ii}}} - \frac{f'(\mathbf{x}_j)}{\sqrt{D_{jj}}}\right)^2 W_{ij} + (1-\alpha)\left(\sum_{i=1}^{n}(f(\mathbf{x}_i)-y_i)^2 + \sum_{i=1}^{n}(f(\mathbf{x}_i)-f'(\mathbf{x}_i))^2\right), \quad (1)$$

where f' is the ranking function before the current annotation, α $(0 < \alpha < 1)$ is the regularization parameter that controls the trade-off between the smooth constraint and the fitting constraint, and $D_{ii} = \sum_{j=1}^{n} W_{ij}$ are the diagonal elements of the diagonal degree matrix D. The first term in eq. (1) enforces smoothness on f when it changes among nearby points, while the second term ensures f not to deviate largely from the initial scores $\{y_i\}_{i=1}^{n}$.

A more succinct form of eq. (1) is

$$\mathcal{Q}(\mathbf{f},\mathbf{f}') = \alpha\mathbf{f}'^T\mathcal{L}\mathbf{f}' + (1-\alpha)(\|\mathbf{f}-\mathbf{y}\|^2 + \|\mathbf{f}-\mathbf{f}'\|^2), \quad (2)$$

where $\mathbf{f} = [f(\mathbf{x}_1), \cdots, f(\mathbf{x}_n)]^T$ and $\mathbf{f}' = [f'(\mathbf{x}_1), \cdots, f'(\mathbf{x}_n)]^T$.

A closed form solution for $\mathbf{f}$ to $\min \mathcal{Q}(\mathbf{f},\mathbf{f}')$ can be obtained as

$$\mathbf{f}^* = (2I - (1-\alpha)(I-\alpha S)^{-1})^{-1}\mathbf{y}, \quad (3)$$

where $S = D^{-1/2}WD^{-1/2}$ symmetrically normalizes W and $\mathbf{f}^* = [f^*(\mathbf{x}_1), \cdots, f^*(\mathbf{x}_n)]^T$ stacks the final ranking scores.

Initially the user controls the values of $\mathbf{y}$ through specifying queries, and the system starts the ranking process. After a ranking comes out, the user changes $\mathbf{y}$ by annotating the displayed images. The system returns a new ranking corresponding to the user specified $\mathbf{y}$ again. In this way, the user and the computer interacts iteratively until the user is satisfied with the results.

Note that the normalized Laplacian $\mathcal{L} = I - S$ is highly sparse due to the k-NN graph $G(V,E,W)$ [13]. We can run a sparse eigenvalue procedure on $\mathcal{L}$ to obtain $\mathcal{L} \cong U\Lambda U^T$ in which $U \in \mathbb{R}^{n\times m}$, and $m \ll n$ is the number of eigenvalues kept for the computation. Eq. (3) is thus expressed as follows

$$\begin{aligned}\mathbf{f}^* &= (2I - (I + \frac{\alpha}{1-\alpha}\mathcal{L})^{-1})^{-1}\mathbf{y} \\ &\cong U(I + \frac{\alpha}{1-\alpha}\Lambda)(I + \frac{2\alpha}{1-\alpha}\Lambda)^{-1}U^T\mathbf{y} \\ &= T\mathbf{y}, \end{aligned} \quad (4)$$

where $T \in \mathbb{R}^{n\times n}$ can be pre-computed to speed up the computation, and $\mathbf{y}$ stays tuned to the user's interests.

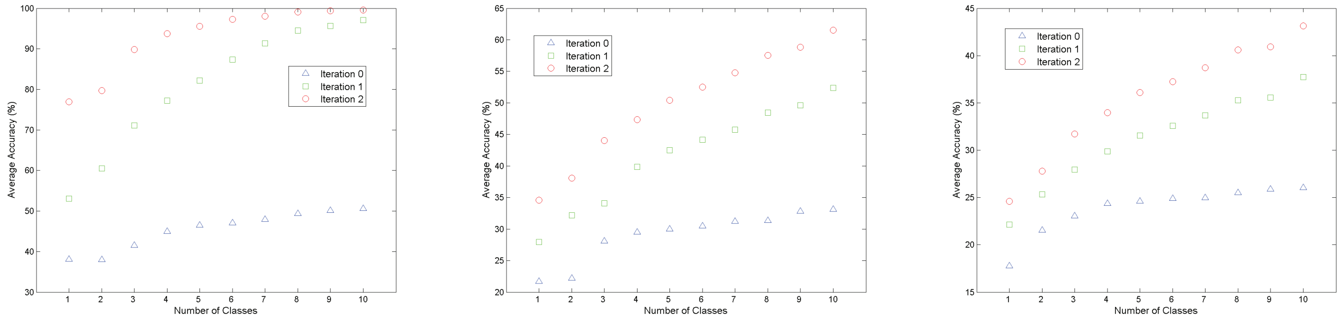
EVALUATION AND RESULTS

In this section, we demonstrate the effectiveness of our system using a large general image database Corel, which contains 200 categories each with 100 images. This database is generic and the images range from simple objects to natural scenes with complex background. We extract the low-level features, colors, and edges to represent each image. We use three types of color moments: mean, variance, and skewness in three different color channels. Thus the color information consists of 9 feature. We obtain the edges of an image by Canny edge detector, and quantize the histogram of the edge directions into 18 bins of every 20 degrees, resulting in 18 features to represent the edge information. Totally, we have 27 features to represent an image.

To evaluate the usability of our system, 50 volunteers aged between 18 and 30 from our university took part in the user study. Before the experiments started, each subject was given a short training on how to use the system.

Table 1: The average improvement on accuracy of Top-50 ranked images of multi-query search over single-query search.

Number of Queries	2	3	4	5	6
Improvement (%)	16.2	28.3	36.7	42.5	46.2
Number of Queries	7	8	9	10	
Improvement (%)	50.0	55.9	57.5	64.1	



(a) (b) (c)

Figure 3: (a) Average accuracy of top-10 ranked images from 1 class to 10 classes. (b) Average accuracy of top-30 ranked images from 1 class to 10 classes. (c) Average accuracy of top-50 ranked images from 1 class to 10 classes. Each class contains one query image.

Figure 4: Top ranked images of one class from the initial search (second row), iteration 1 (third row) and iteration 2 (fourth row). No outliers remain after iteration 2.

The experiments were carried out in a subset of Corel database, which contains 50 categories each with 100 images. The users did multi-query search concurrently in the system, and the number of queries varied from 1 to 10. In every iteration, the user annotated at most 10 images. 20 groups of queries were assigned to one user, which were randomly collected. We computed the average accuracy of top-10, 30, and 50 ranked images to evaluate the performance of our system. Fig. 3a, Fig. 3b and Fig. 3c clearly show that the average accuracy of top ranked images increases as the number of queries increases. The improvement of our multi-query search method over the single-query search is summarized in Table 1. Fig. 4 shows an example of outlier removal in the iterations.

CONCLUSION

This paper has proposed a novel interactive image search system to address the problem of multi-query searching simultaneously with efficient outlier removal. To accomplish this task, we develop a multi-query ranking algorithm and design an innovative interactive UI. The main idea of our multi-query ranking scheme is to derive ranking scores with a graph regularization technique based on the scores obtained by our manifold ranking algorithm. The UI supports the user to annotate current search results for the next round search. The experiments and user study show the efficiency of our multi-query image search system, the usability of the UI, and the power of the multi-query ranking algorithm.

ACKNOWLEDGMENTS

This work was supported by the Research Grants Council of the Hong Kong SAR (Project No. 414306).

REFERENCES

1. N. Boujemaa and F. Fleuret and V. Gouet and H. Sahbi. Visual content extraction for automatic semantic annotation of video news. In *the Proceedings of the SPIE Conference*, San Jose, CA, 2004.
2. A. Smeulders and M. Worring and S. Santini and A. Gupta and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Trans. on PAMI*, 22(12):1349-1380, 2000.
3. T. Gevers and A.W.M. Smeulders. Content-based image retrieval: An overview. In *G. Medioni and S. B. Kang, editors, Emerging Topics in Computer Vision*, Prentice Hall, 2004.
4. I. Cox and M. Miller and T. Minka and P. Yianilos. An optimized interaction strategy for bayesian relevance feedback. In *IEEE Conference on Computer Vision and Pattern Recognition*, Santa Barbara, CA, 1998.
5. Y. Rui and T. Huang and S. Mehrotra. Relevance feedback techniques in interactive content-based image retrieval. In *Storage and Retrieval for Image and Video Databases (SPIE)*, pages 25-36, 1998.
6. Y. Zhao and Y. Zhao and Z. Zhu. Relevance feedback based on query refining and feature database updating in cbir system. In *Signal Processing, Pattern Recognition, and Applications*, 2006.
7. X. Zhou and T. Huang. Relevance feedback in image retrieval: A comprehensive review. In *IEEE CVPR Workshop on Content-based Access of Image and Video Libraries (CBAIVL)*, 2006.
8. R. Albert and H. Jeong and A. Barabsi. Diameter of the world wide web. *Nature*, 401:130-131, 1999.
9. S. Brin and L. Page. The anatomy of a large scale hypertextual web search engine. In *WWW*, 1998.
10. J. He and M. Li and H. Zhang and H. Tong and C. Zhang. Manifold-ranking based image retrieval. In *ACM MM*, 2004.
11. D. Zhou and J. Weston and A. Gretton and O. Bousquet and B. Schölkopf. Ranking on Data Manifolds. In *NIPS*, 2003.
12. M. Breitenbach and G. Grudic. Clustering through ranking on manifolds. In *ICML*, 2005.
13. M. Hein and M. Maier. Manifold Denoising. In *NIPS*, 2006.