

A Novel Retrieval Refinement and Interaction Pattern by Exploring Result Correlations for Image Retrieval

Rongrong Ji, Hongxun Yao, Shaohui Liu, Jicheng Wang, and Pengfei Xu

VILAB, School of Computer Science, Harbin Institute of Technology, No.92,
West Dazhi Street, Harbin, 150001, China
{rrji,yhx,shaohl}@vilab.hit.edu.cn

Abstract. Efficient retrieval of image database that contains multiple predefined categories (e.g. medical imaging databases, museum painting collections) poses significant challenges and commercial prospects. By exploring category correlations of retrieval results in such scenario, this paper presents a novel retrieval refinement and feedback framework. It provides users a novel perceptual-similar interaction pattern for topic-based image retrieval. Firstly, we adopts Pairwise-Coupling SVM (PWC-SVM) to classify retrieval results into predefined image categories, and reorganizes them into category-based browsing topics. Secondly, in feedback interaction, category operation is supported to capture users' retrieval purpose fast and efficiently, which differs from traditional relevance feedback patterns that need elaborate image labeling. Especially, an Asymmetry Bagging SVM (ABSVM) network is adopted to precisely capture users' retrieval purpose. And user interactions are accumulated to reinforce our inspections of image database. As demonstrated in experiments, remarkable feedback simplifications are achieved comparing to traditional interaction patterns based on image labeling. And excellent feedback efficiency enhancements are gained comparing to traditional SVM-based feedback learning methods.

Keywords: image retrieval, image classification, relevance feedback, support vector machine, pairwise coupling, bagging.

1 Introduction

Content-based image retrieval (CBIR) is a technique that effectively retrieves images based on their visual contents [1]. Over the past decade, many efforts have been carried out to enhance the retrieval precision and recall of CBIR systems [2]. However, few attentions are concerned on investigating result correlations and representing retrieval result in a more perceptual similar manner. Better representation leads to more efficient and effective user browsing, which can inspire user passions in interaction engagement. Traditional CBIR methods simply return to user the most similar retrieval results without any post-processing. The correlations between query results is out of consideration.

In many cases, the image databases are consisted of image categories that can be predefined to certain degree, in which image category is known beforehand but only a

very limited number of category images can be obtained for training. In such case, fixed image pre-classification is not a feasible solution for user browsing since the classification precision and recall are very limited. However, online classification of retrieval results, which can assist user navigation by topic-based image representation, is feasible. For instance, images in many medical databases are of finite kinds of imaging modalities and anatomic regions. Similar examples can be found in, e.g., remote-sense satellite image database (landforms can be predefined), art museum painting collections (consists of portraits, landscape, tachisme paintings et. al.), sports news image database (sports image such as soccer, baseball, tennis can be predefined beforehand). In such scenario, the design of CBIR browsing and interaction pattern can be improved by integrating domain-specific knowledge. Especially, in such cases, users may request category search and wish the search result to be represented in a topic-specific manner. Similar but not identical researches can be found in textual retrieval [16] [17]. In unsupervised scenario, clustering-driven retrieval refinement is an ad hoc research area [18]. For instance, Vivisimo (<http://vivisimo.com>) is a search engine that automatically clusters search results into representative categories.

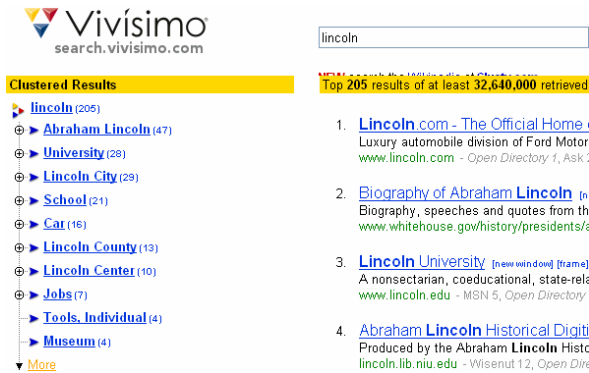


Fig. 1. Clustering-Based retrieval result representation in Vivisimo

To the best of our knowledge, clustering & categorizing-based refinement in image retrieval is also becoming a new research hot-spot. For instance, some interactive search engines participating in the TRECVID benchmark offer such functionality. And some of them (e.g., Columbia, IBM) also feature cluster-based re-ranking of results. Chen [12] developed a cluster-based image retrieval system: CLUE to automatically cluster content-based search results by unsupervised learning. And Lee [13] indexed objects based on shape and grouped retrieval results into a set of clusters, in which cluster is represented by a prototype based on the learning of the users' specific preference.

Furthermore, the relevance feedback mechanism can be further simplified and improved using predefined category knowledge in above-mentioned scenario. To integrate human supervision into retrieval, traditional relevance feedback schemes [3-5, 14] need users to manually label certain amount of positive/negative example

images. However, in many cases, users are reluctant to provide sufficient feedback images, since this labeling procedure is tedious and burdensome. Thus it affects the efficiency of RF learning algorithms. As a result, the poor RF learning results would cause users to end their retrieval tasks with a failure.

To address these issues in above-mentioned scenario, this paper proposes a novel retrieval refinement and feedback learning framework, which leads to a new user browsing and interaction pattern. Our target is to facilitate user browsing and enhance interaction efficiency. Firstly, pre-categorized images are utilized for Pairwise-Coupling SVM training, which produces a set of one-to-one category classifiers. In retrieval, the initial similarity ranking results (Top k most similar images) are classified into their corresponding image categories. Both top k ranking results and top m ranking categories are returned to the users. In RF procedure, category selection and labeling is available (No selection indicates the appearing image categories are all negative examples). The image similarities are re-ranked using the user-provided positive and negative image categories. Especially, an asymmetry bagging SVM (ABSVM) network is adopted in category-based RF learning. Finally, in a successful retrieval, the query image is added into its positive labeled category in training collection. Fig.2 shows our refinement system framework.

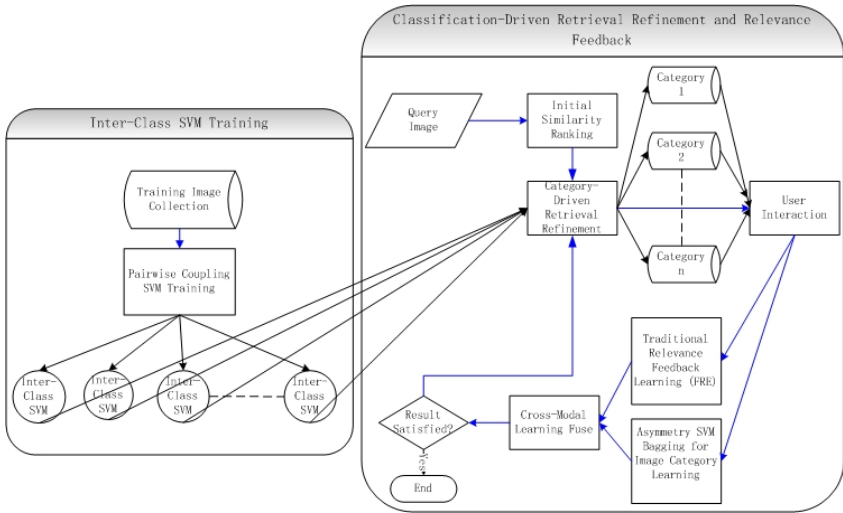


Fig. 2. Category-Based Retrieval Refinement and Relevance Feedback Framework

Three innovations distinguish our category-based retrieval framework from the other CBIR schemes:

1. Classification-based retrieval refinement (PWC-SVM) to group and represent results in a perceptual similar topic-based interface.
2. Support category-level feedback interaction to efficiently and effectively capture users' retrieval purpose.

3. Adopt asymmetry bagging SVM network to address the issues of sample insufficiency and sample asymmetry in feedback learning.

The rest of this paper is organized as follows: Section.2 proposes our retrieval result refinement algorithm based on Pairwise-Coupling SVM. Section.3 presents our classification-based RF learning strategy using asymmetric SVM bagging network. The experimental results are shown in Section.4. And finally this paper concludes in Section 5.

2 Category-Based Retrieval Result Refinement Using Pairwise-Coupling SVM

Our refinement algorithm concerns on category search in the database which consists of image categories that can be predefined to some degree. Such scenario is prevalent in many specific image databases (Medical image database, gene image database, museum painting collections). Since the image category can be defined beforehand, category information can be adopted to design a refinement scheme to group and represent results in a perceptual similar topic-based interface, which will largely simplify and accelerate user retrieval.

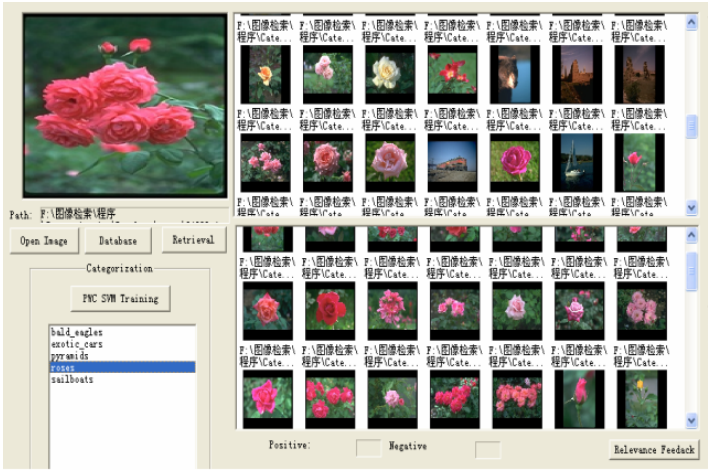


Fig. 3. Category Refinement System Interface

Assume that m image categories can be defined beforehand in a category-based image database. We can utilize this category information to train decision rules that classify retrieval result images into their corresponding categories. Consequently, the retrieval results is consisted of not only top k relevant images but also their top n frequently appeared categories. In RF operation, users can either mark positive/negative images or just simply select their interesting image category. As shown in Fig.3, the initial retrieval result is shown in the top right image window.

Once clicking a category in the left textual list (which presents the name of the top n category), an image browsing window would be presented to the users in the bottom right image window, which contains the retrieval results with the selected category. Users can further select misclassified image inside each category. In this scenario, an active user can participate as much as possible to achieve his goal (Select both feedback categories and the misclassified images); while a “lazy” user without specified knowledge only needs to select positive category (one operation) to conduct the relevance feedback learning.

Our scheme adopts a Pairwise-Coupling SVM (PWC-SVM) framework to classify images to their corresponding categories. After initial retrieval, the top k most similar images are further classified into their corresponding categories using a PWC-SVM framework (Tab.1). And the top n most frequently appeared categories are returned to the users.

In such framework, SVM is selected as the inter-class classifier. SVM is an effective binary classifier and widely used in the RF learning of content-based image retrieval [5] [7]. It aims at separating two classes in feature space by an optimal hyperplane, where the maximum geometric margin gains [6]. But as a binary classifier, it is unsuitable to be applied directly to image categorization scenario, which is a multiple class classification problem. Generally, there are two strategies adopted to extend binary classifiers to multiple-class situation:

1. One-To-All: For each class, train an intra-class classifier whose classification result indicates the confidence (scope $[0,1]$) that whether the test sample belongs to this class or not. For each test sample, these classifiers of all categories are used to gain the classification results, in which the class with the highest output is selected as the classification result.
2. One-To-One: For each category, train $(m-1)$ pairwise classifiers (inter-class) between current class and the rest classes (Totally m classes). For each test sample, the confidence level of each class is the majority voting result of its corresponding $(m-1)$ classifiers (each between $(0, 1)$). The class with highest accumulated outputs is selected as the classification result. This is called PWC (Pairwise-Coupling), which combines the outputs of all classifiers to form prediction. And it has gained more concerns due to its good generalization ability [9, 10].

In our category-based retrieval framework, the Pairwise-Coupling (one-to-one) strategy is adopted to extend binary SVM into multiple-class scenario (Tab.1). There are totally C_m^2 SVM classifiers to be trained.

Table 1. PWC SVM Retrieval Result Classification

<p>Input: retrieval result images $I_1 \dots I_k$, pairwise SVM classifiers $C_{01}, C_{02}, \dots, C_{m(m-1)}$</p> <p>For each retrieval result image I_i</p> <p>Begin</p> <p>Utilize its $(m-1)$ corresponding SVM classifiers $C_{i1}, C_{i2}, \dots, C_{i(m-1)}$ to classify I_i to its corresponding category</p> <p>End</p> <p>Output: The top n categories with the n highest accumulated classification values among the result images.</p>

3 Category-Level Feedback Learning Using Asymmetric Bagging SVM Network

The distinction between our RF learning algorithm and traditional schemes lies in: users' category selection may be the only interaction information available for retrieval system. Furthermore, in such scenario, the training examples (images from one positive category and other $(n-1)$ negative categories) are strongly asymmetric.

As pointed out by Tao [5], the SVM classifier is unstable on a small training set. And its optimal classification hyperplane may be biased when the numbers of the

Table 2. Asymmetry SVM Bagging for Image Category Confidence Calculation

<p>Input: Positive Image Category C_p and its corresponding training images $I_{p1}...I_{pe}$, Negative Image Category C_{n_i} ($i=1$ to n, and $C_{n_i} \neq$ positive category), and their corresponding training images $I_{n1_1}...I_{n1_f}...I_{nm_1}...I_{nm_f}$. classifier C (SVM)</p> <p>For bootstrapping interval k from 1 to n</p> <p>Begin</p> <p> Train SVM C_k between positive image category C_p and negative image category C_{n_k}</p> <p>End</p> <p>Construct the final classifier C_{final} by aggregation of all C_k. Its output (rank 0 to $(m-1)$) indicates the confidence level that an image belongs to the positive category</p> <p>End</p> <p>Output: C_{final} to determine the category belonging of each image</p>
--

Consequently, the category-level correlations are integrated into traditional image-level RF learning schemes to re-rank resulting images as follows:

Table 3. Classification-Based Similarity Re-ranking

<p>Input: images in the database: $I_1...I_h$, category classifier: C_{final}</p> <p>For each image I_i ($i = 1$ to h)</p> <p>Begin</p> <p> Using a traditional RF learning method to re-rank its similarity S_i of the query image (In our experiments both FRE [11] and SVM are investigated)</p> <p> Calculate its confidence level Ca_i which indicates to what degree this image belongs to the positive image category.</p> <p> Re-calculate the similarity of S_i: $S_i = S_i / (Ca_i + 1)$</p> <p>End</p> <p>Re-Rank all the images in the database and return to the user the most similar k images</p> <p>Output: the most similar k images and their most frequently appeared image category calculated using Tab.2</p>
--

positive and negative samples are asymmetric. To solve this problem, an asymmetry SVM bagging network is adopted for category-level relevance feedback learning. The asymmetry bagging idea [5] is adopted to overcome the problem of sample asymmetry in SVM training. Different from Tao, our algorithm utilizes this bagging procedure to image category classification. And such classification is further integrated into traditional RF learning algorithms to enhance system performance.

4 Experimental Results and Discussion

Our experiments are conducted on a subset of COREL image database, with over 2000 general purpose images belonging to 20 classes (100 each). 16 images of each topic (Totally 320 images) are randomly selected, half for training and half for query test.

Our experimental evaluation aims at verifying the efficiency of our categorization-based retrieval refinement and interaction scheme. Consequently , for simplicity, a 264-dimensional feature vector is extracted from each image, in which the former 256 dimension features are the auto-correlogram [15] in RGB color space (We quantize RGB color space into 64 bins and based on which the pixel correlation within distance 1,3,5,7 are extracted and calculated to form a 256-dimentional feature.) and the later 8 dimension features are the features of texture co-occurrence matrix [8] (weighted mixed by 8:256).

Table 4. Average RF Operations in Top 100 Results

Category	Traditional RF			Category	Our Method		
	1st	2nd	3rd		1st	2nd	3rd
eagle	49	67	83	train	12	15	14
car	25	43	66	pyramid	8	7	10
tiger	28	41	57	sunrise	11	13	13
fish	35	52	56	sailboat	16	23	23
rose	48	64	75	bear	21	17	19

Our first experiment is the demonstration of our PWC-SVM based category-level retrieval refinement scheme. Tab.4 presents the average RF operation in top 100 retrieval results between traditional image-level RF pattern and our category-based method. As shown in Tab.4, users' average RF operations can be greatly reduced comparing to traditional RF strategies. Our method only needs users to label image categories and the misclassified images in each category.

Our second experiment is the demonstration of ABSVM network in category-based feedback learning. To measure its performance, two classical and efficient baseline feedback learning algorithms are implemented: 1. Rui's feature reweighting (FRE) RF learning method [11] and 2. SVM Based RF learning. Both of these two methods are conducted over the same data scope as our ABSVM category RF. They need users to manually label all positive/negative images in top 100 returning images (The tedious RF operations can be demonstrated by their average operation times in left of Tab.4).

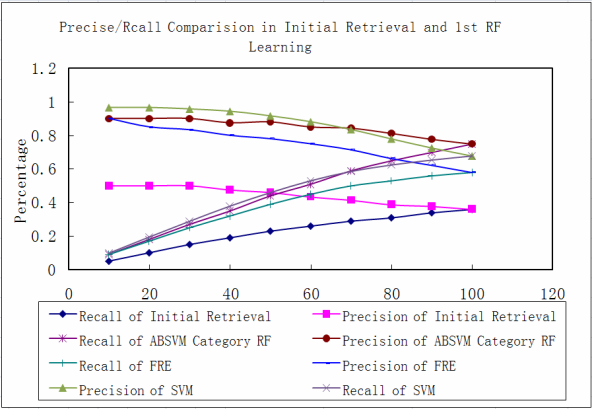


Fig. 4. P-R Comparison in Initial Retrieval and 1st RF Learning

As shown in Fig.4-6, in despite of much fewer RF operations and mis-labeling images, our ABSVM category-RF learning scheme is still excellent comparing to SVM-based image-level RF methods, which is commonly used in state-of-art systems (In image-level SVM, elaborated manually labeling is demanded in top 100 images, which is extremely tedious and burdensome). Promising retrieval Precision & Recall enhancements are gained comparing to traditional RF schemes (both FRE and SVM). In addition, from users' point of view, this pattern is more semantically meaningful and straightforward.

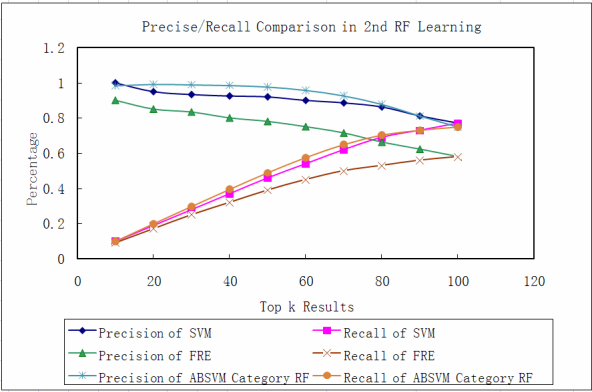


Fig. 5. P-R Comparison in 2nd RF Learning

It should be noted that other well-explored learning and ranking schemes can be easily integrated into our refinement and category-RF frameworks to further enhance system performance., due to its good extensibility and generalization ability.

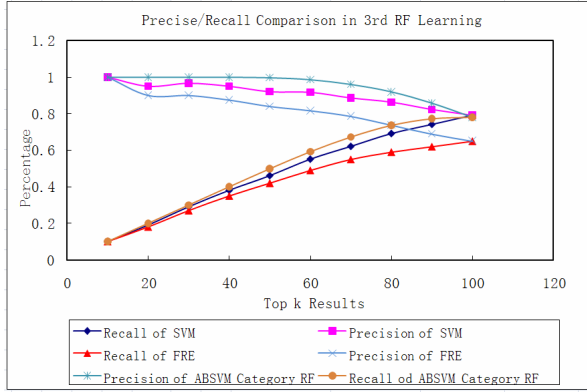


Fig. 6. P-R Comparison in 3rd RF Learning

5 Conclusion

This paper proposes a novel classification-driven retrieval refinement and interaction pattern by exploring result correlations for category image retrieval. Our framework supports topic-based perceptual similar result representation and category-level feedback operation. Especially, a PWC-SVM strategy is proposed for category-based result post-processing and an asymmetry SVM bagging network is integrated into traditional RF learning methods to enhance learning efficiency in category search. Our scheme provides a general idea for category search in domain-specific image database, which can be easily integrated by other sophisticated or state-of-art feature extraction or feedback learning schemes to further enhance system performance.

Acknowledgments. This research is supported by State 863 High Technology R&D Project of China, No. 2006AA01Z197; the Program for New Century Excellent Talents in University (NCET-05-03 34); Natural Science Foundation of China (No. 60472043) and Natural Science Foundation of Heilongjiang Province (No.E2005-29).

References

1. Smeulders, A., Worring, M., Santini, S., Gupta, A., Jain, R.: Content-Based Image Retrieval at the End of the Early Years. *IEEE Trans. on Pattern Analysis and Machine Intelligence* 22(12), 1349–1380 (2000)
2. Veltkamp, R.C., Tanase, M.: Content-Based Image Retrieval Systems: A Survey, Technical report UU-CS-2000-34, Department of Computing Science, Utrecht University 34 (October 2000)
3. Gong, Y., Zhang, H.J., Chua, T.C.: An image database system with content capturing and fast image indexing abilities. In: *Proc. IEEE Int. Conference on Multimedia Computing and Systems*, Boston, 14-19 May, pp. 121–130 (1994)

4. Su, Z., Zhang, H., Li, S.: Relevance Feedback in Content-Based Image Retrieval: Bayesian Framework, Feature Subspaces and Progressive Learning. *IEEE Tran. on Image Processing* 12(3), 8 (2003)
5. Tao, D., Tang, X., Li, X., Wu, X.: Asymmetric Bagging and Random Subspace for Support Vector Machines-Based Relevance Feedback in Image Retrieval. *IEEE Trans. on Pattern Analysis and Machine Intelligence* 28(7) (July 2006)
6. Burges, J.C.: A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery* 2(2), 121–167 (1998)
7. Tong, S., Chang, E.: Support Vector Machine Active Learning for Image Retrieval. In: *Proceeding of ACM International Con. on Multimedia*, pp. 107–118 (2001)
8. Haralick, R.M., Shanmugam, K., Dinstein, I.: Texture features for image classification. *IEEE Tran. on System Man Cybern* 3, 610–621 (1973)
9. Rahman, M.M., Bhattacharya, P., Desai, B.C.: A Framework for Medical Image Retrieval using Machine Learning & Statistical Similarity Matching Techniques with Relevance Feedback, *IEEE Trans. on Information Tech. in Biomedicine* (accepted for future publication)
10. Wu, T., Lin, C.J., Weng, R.C.: Probability Estimates for Multi-Class Classification by Pairwise Coupling. *Int. Journal on Machine Learning Research* 10(5), 975–1005 (2004)
11. Rui, Y., Huang, T.S., Mehrotra, S., Ortega, M.: Relevance Feedback: A Power Tool for Interactive Content-based Image Retrieval. *IEEE Trans. Circuits and Systems for Video Technology* 8(5), 644–655 (1998)
12. Chen, Y., Wang, J.Z., Krovertz, R.: CLUE: Cluster-Based Retrieval of Image by Unsupervised Learning. *IEEE Trans. on Image Processing* 14(8), 1187–1201 (2005)
13. Lee, K.-M., Nike Street, W.: Cluster-Driven Refinement for Content-Based Digital Image Retrieval. *IEEE Trans. on Multimedia* 6(6), 817–927 (2004)
14. Tao, D., Tang, X., Li, X., Rui, Y.: Direct Kernel Biased Discriminant Analysis: A New Content-based Image Retrieval Relevance Feedback Algorithm. *IEEE Trans. on Multimedia* 8(4), 716–727 (2006)
15. Huang, J., Kumar, S.R., Mitra, M., Zhu, W.J., Zabin, R.: Image Indexing using Color Correlogram. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, San Juan, PR, January, pp. 762–768 (1997)
16. Cui, H., Heidorn, P.B., Zhang, H.: An Approach to Automatic Classification of Text for Information Retrieval. In: *The 2nd ACM/IEEE-CS Joint Conference on Digital Libraries*, Portland, Oregon
17. Karanikolas, N., Skourlas, C., Christopoulou, A., Alevizos, T.: Medical Text Classification based on Text Retrieval techniques. In: *MEDINF 2003*, Craiova, Romania, October 9 - 11 (2003)
18. Liu, X., Gong, Y., Xu, W., Zhu, S.: Document clustering with cluster refinement and model selection capabilities. In: *Proceedings of ACM SIGIR, SESSION: Clustering*, Tampere, Finland, pp. 191–198 (2002)