Paper:

A Feature Vector Approach for Inter-Query Learning for Content-Based Image Retrieval

Kien-Ping Chung and Chun Che Fung

School of Information Technology, Murdoch University
South Street, Murdoch, Perth, Australia, 6150
E-mail: {k.chung, l.fung}@murdoch.edu.au

[Received April 20, 2006; accepted July 29, 2006]

Use of relevance feedback (RF) in the feature vector model has been one of the most widely used approaches to fine tuning queries for content-based image retrieval (CBIR). We propose a framework that extends RF to capturing the inter-query relationship between current and previous queries. Using the feature vector model, this avoids the need to "memorize" actual retrieval relationships between actual image indexes and the previous queries. This approach is suited to image database applications in which images are frequently added and removed. In the previous work [1], we developed a feature vector framework for inter-query learning using statistical discriminant analysis. One weakness of the previous framework is that the criteria for exploring and merging with an existing visual group are based on two constant thresholds, which are selected through trial and error. Another weakness is that it is not suited to mutually interrelated data clusters. Instead of using constant values, we have further extended the framework using positive feedback sample size as a factor for determining thresholds. Experiments demonstrated that our proposed framework outperforms the previous framework.

Keywords: content-based image retrieval system, inter-query learning, relevance feedback, statistical discriminant analysis

1. Introduction

In the last decade, query tuning using relevance feedback (RF) has gained much attention in content-based image retrieval (CBIR) R&D, largely due to RF's ability to refine user queries through a sequence of interactive sessions. Different approaches [2] to RF in CBIR have yielded certain success, but most researches have focused on query tuning in *a single* retrieval session, i.e., *intra-query* learning. In contrast, *inter-query learning*, i.e., *long-term learning*, analyzes the relationship between current and past retrieval sessions. By accumulating knowledge learned from previous sessions, inter-query learning improves the retrieval performance of the current and future sessions. One may view that inter-query is an extension of intra-query. Although intra-query in CBIR has been studied for the last decade, inter-query in CBIR has only begun to attract interests and has yet to be fully explored.

We developed an inter-query learning framework based on statistical discriminant analysis to represent characteristics of a visual group during a retrieval session [1]. This approach is suited to database applications in which images are regularly added and removed because it avoids the need to establishing relationships between individual images in a database, a common approach in most inter-query learning frameworks. A weakness of our previous framework is that the criterion for exploring and merging with an existing visual group is based solely on two constant thresholds determined through trial and error, and this approach is not suited to mutually interrelated data clusters.

Here, we extend the previous framework using available feedback samples as the factor for determining thresholds for cluster merging and exploration. We start by briefly reviewing current approaches to inter-query learning for content-based image retrievals. We then review our previous work and present our proposed approach. We then report results from experiments on the implemented framework and state our conclusions.

2. Background

According to the literature, inter-query learning is based on two approaches. The first originated in conventional text search and retrievals whose goal is to analyze the relationship between images and related queries through user retrieval patterns. This assumes that if two query retrieval patterns are identical, images retrieved by the queries must be semantically similar. Examples of this approach are latent semantic indexing (LSI) [3] and statistic correlation [4]. This approach has been extended [5–7] by attempting to analyze the hidden semantic relationship among indirectly related images groups. The weakness in this is that the image database must be relatively static making maintenance complicated if images are frequently added or removed from the database.

The second inter-query approach is based on the feature vector model involving changing the feature vector coordinate scale to bring similar images closer. The feature vector coordinate is transformed through weighting or kernel matrix transformation. The feature vector model approach is widely used in short-term learning. The support vector machine (SVM) [8, 9] and self-organizing map (SOM) [10] are the most widely used machine learning tools for image clustering for inter-query learning. In the feature vector model approach, images retrieved from a query session are captured as a clustered group and information on retrieved images is stored in the "user log". This approach avoids the need for "memorizing" actual retrieval relationships between actual images and the previous queries. However, applying SVM in data clustering can be tricky and requires non-linear data approximation such as pre-image calculation [11, 12] to transform data back to the original feature space before clusters are merged. Instead of using SVM [1], uses statistical discriminant analysis to analyze the inter-query relationship in the visual feature space. In statistical discriminant analysis, two data clusters can be merged by relatively simple algebraic mathematics. Statistical discriminant analysis is detailed in the next section.

3. Previous Framework

In previous work [1], we proposed a framework for inter-query learning in content-based image retrieval system. The framework was based on kernel-biased discriminant analysis as proposed by Zhou and Huang [13] for short-term learning in an image retrieval. Unlike the short-term learning framework, our proposed framework captures retrieval information using a data cluster consisting of a center point, a boundary value and the transformation matrix used to transform the extracted image feature from the original feature space to a new feature space. If an image falls within the boundary, the image becomes part of this group of images. The boundary distance is the distance of the furthest positive sample from the center. The proposed inter-query learning analysis locates the data cluster able to identify and contain most the current positive samples. The searching criterion is written as:

$$criteria = \frac{N_{n_i}}{N_p} \quad \dots\dots\dots\dots\dots \quad (1)$$

where N_p denotes the total number of positive samples gathered during the feedback cycle, and N_{n_i} is the number of positive samples that fall within the boundary of cluster i.

Once clusters within the user log are identified, the system decides how the group is to be used. The system either merges feedback samples with the identified cluster to form a new query point. Alternatively, the system creates multiple query points created in short-term learning using feedback samples and the query point created by the identified group from the long-term learning. The decision to create a merged group or to separate query points is based on search criteria (Eq. (1)). If the value of criteria exceeds a certain threshold, then a new merge group or two separate queries are created. This is designed to enable the user to further explore related visual groups having similar semantic content.

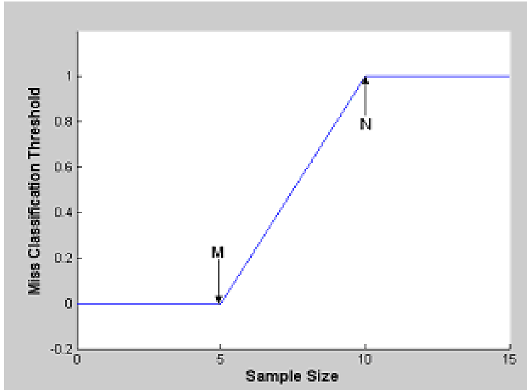

Fig. 1. Threshold for missing classification when $M = 5$ and $N = 10$.

At the end of the retrieval session, the user log is updated. The user log updating policy is similar to the query expansion policy discussed above, but this search criteria differs from criteria as shown in Eq. (1). Negative feedback samples are used to determine the data cluster most suitable for updating. Negative feedback samples provide additional information on the discrimination of selected data clusters for feedback samples. This criterion is written as:

$$criteria = \frac{N_{si}}{N_t} \quad \dots\dots\dots\dots\dots \quad (2)$$

where N_t denotes the total number of samples gathered during the retrieval session, and N_{si} is the number of positive samples that fall within the boundary of cluster i. Using feedback samples, the system searches through groups to find the most appropriate group. A new group is created if the search criterion is less than the threshold, or the current group is updated with samples. Similarly, the rule for merging two clusters depends on the threshold in Eq. (2) and the Euclidean distances between the two positive centre points.

4. Proposed Framework

Our previously proposed framework had room for improvement. Thresholds for exploring or merging clusters were determined through trial and error, and exploring and merging criteria were based on the number of samples falling within the cluster. Since the boundary of the cluster is based on the furthest positive sample from the center, the discriminant approach cannot clearly separate the positive sample furthest away from the negative samples.

Most importantly, thresholds are constant and based only on the ratio of samples falling within the cluster without considering actual sample size which is important to consider when dealing with any statistical analysis. If the threshold as listed in Eq. (1) is set to 0.5, this implies that the cluster is selected if more than 50% of positive samples fall within the cluster. A huge difference in implication lies between scenarios in which "*2 samples have been labeled positive, and 1 of them falls within the cluster*" and "*20 samples have been labeled positive, and 10 of them fall within the cluster.*" In these two scenarios, if the cluster is explored, the second scenario will receive preference. This issue is resolved by only enabling the system to search clusters after the number of positive samples gathered exceeds a minimum preset value. This, however, prevents system from exploring the cluster in earlier search cycle.

To resolve these three issues, we propose Eqs. (3) and (4) for replacing for criteria in Eqs. (1) and (2).

$$\alpha = \frac{N_{ir}}{N_r} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (3)$$

$$threshold = \begin{cases} 0, & 0 \subseteq R \subseteq M \\ \dfrac{R-N}{N-M}, & M \subset R \subseteq N \\ 1, & N \subset R \end{cases} \quad \ldots \quad (4)$$

where α is the misclassification factor and parameter N_{ir} is the number of negative labeled samples in whose Euclidean distances to the positive center are smaller than average Euclidean distances of positive samples, N_r is the number of positive labeled samples whose Euclidean distances are smaller than average Euclidean distances of positive samples. This minimizes the effect of poor discrimination between furthest positive samples and negative samples.

For the threshold in Eq. (4), R is the number of positive samples gathered from each relevance feedback cycle and parameters M and N are boundary sample sizes for the threshold for setting 0 or 1. As an alternative to the previous framework, the newly proposed threshold is no longer a constant, it is now depends on sample size. When the threshold is set to 0, it implies that zero classification is allowed. When the threshold is set to 1, it allows that half of the samples are falsely classified so that the threshold for exploring the cluster is harsher when the sample size is small. As the sample size increases, the rule for exploring the cluster becomes more lenient. **Fig. 1** depicts the value of the threshold when the minimum and maximum sample sizes are set to 5 and 10.

The decision for exploring is, in short, "*exploring the cluster only when the misclassification factor is smaller than or equal to the threshold.*" Similarly, the same threshold is used along with similarity Euclidean distance calculation to determine if two clusters are to be merged.

5. Experiments

To evaluate the performance of our proposal, we have designed and implemented a prototype CBIR using the Matlab simulation software toolbox. This prototype enables the user to query the image database with an image sample. After the first retrieval iteration, users select relevant images while ignoring non relevant images. The system labels selected images as positive while treating ignored images as negative. Retrieval is as follows:

1. The user inputs a query image.
2. Visual features of the query are extracted by the system.
3. All images in the database are sorted in ascending order based on the distance of dissimilarity.
4. The top 20 images with the highest rank are displayed.
5. The user selects positive images and the rest are automatically labeled as negatives.
6. A new query is created from feedback samples based on the proposed decision algorithm. The equations for the proposed algorithm are given in Eqs. (3) and (4).
7. The top 20 images not labeled by the user are ranked and displayed.
8. Return to step 5 for the next retrieval cycle.
9. Update the user log at the end of the retrieval session.

5.1. Experiment Environment

To test the performance of our proposal, we implemented three systems. They are (i) our proposal, (ii) our previous framework, and (iii) a short-term learning framework based on [13]. To evaluate the validity of the experiment, the environment and parameters used by all three systems are identical. Image features and generalized eigenvector calculation are the same and the same parameters are used in the kernel transformation algorithm for all three systems based on analysis of shape [14], texture [15], and RGB and HSV color histograms of images. Each feature is composed of a number of elements. A total of 48 feature elements are used. The Gaussian Radial Basis Function (RBF) is the kernel transformation matrix for the KBDA approach based on literatures on CBIR [13, 16] reporting that RBF yields the best accuracy of all kernel transformation approaches.

5.2. Experiment Procedure and Data

This experiment uses 500 images from the Corel image database of which 300 are classified under seven themes and each consists of several different visual groups. The inter-relationship of each theme is shown in **Fig. 2**. The themes "bird" and "cat" are subsets of "animal," "fish" is the subset for both "animal" and "water," and "water" consists of "water scene." "Yellow flower" is independent of all other themes. The interrelationship between each theme emulates the complexity of the semantic relationship between each object in the real world.

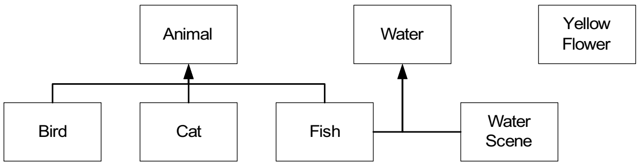

Fig. 2. Relationship between themes in test data set.

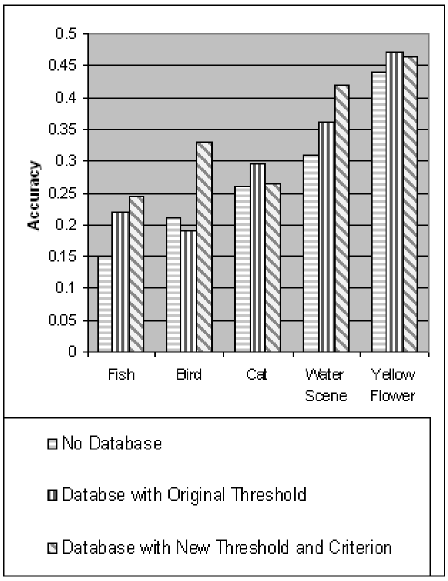

Fig. 3. Retrieval performance of three frameworks with five independent themes.

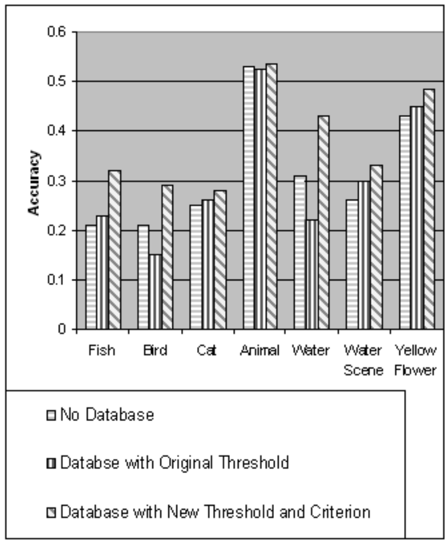

Fig. 4. Retrieval performance of three frameworks with more complex relationship between themes.

The retrieval performance of the frameworks was measured via four different tests. The first two tests involved selecting all themes independent of each other, i.e., "bird," "fish," "cat," "water scene," and "yellow flower" were selected in the first two tests. Tests were generated by randomly selecting 300 positive labeled images from each theme as an input entry point to the system. The same data set was then applied in the second test with a different random sequence to ensure consistency of test results. The last two tests were done using all themes (**Fig. 2**). These four tests measured the performance of the two frameworks under a simple and a more complex cluster relationship. Retrieval accuracy is used as the main factor for comparing system performance.

5.3. Test Results and Discussion

Figure 3 shows the average accuracies of the five independent themes after two random sequences of 300 feedback sessions. While short-term learning has the worst retrieval performance, test results with the new criterion and threshold value are only marginally better than results with the original threshold value. In fact, the framework with the original threshold has better retrieval performance in the themes "cat" and "yellow flower." Overall, test results generated from the new threshold are consistently better than results from the short-term learning framework. The retrieval performance of the framework with the original threshold is compatible in some areas but significantly worse in others. This is clearer (**Fig. 4**) when a more complex relationship between each theme is introduced. In **Fig. 4**, the framework with a new threshold maintains advantages over the other two frameworks, the performance of the framework with the original threshold in themes "bird," "water," and "animal" is worse than the short-term learning framework.

6. Conclusion

We have introduced a new criterion and threshold in a statistical discriminant analysis framework for inter-query learning in CBIR. The proposed criterion and threshold are more flexible than the previously used constant threshold because the new threshold is no longer constant but dependent on the size of positive samples, which may differ for different cases. The proposed criterion is based on the number of samples whose Euclidean distances to the positive centre are smaller than average Euclidean distances of positive samples. This minimizes the effect of poor discrimination between furthest positive samples and negative samples. Test results demonstrated that the accuracy retrieval performance for the proposed criterion and threshold is better than the original threshold from a previous proposal. Currently, we use only a *flat* structure in grouping images. Using a complex hierarchical structure to organize these image groups, which are semantically similar, yet visually different, would more accurate retrieval results.

References:

[1] K.-P. Chung, K. W. Wong, and C. C. Fung, "Reducing User Log Size in an Inter-Query Learning Content-Based Image Retrieval (CBIR) System with a Cluster Merging Approach," Interna-

tional Joint Conference on Neural Networks, Vancouver, Canada, pp. 2163-2170, 2006.
[2] X. S. Zhou and T. S. Huang, "Relevance Feedback in Image Retrieval: A Comprehensive Review," ACM Multimedia Systems Journal, Vol.8, pp. 536-544, 2003.
[3] D. R. Heisterkamp, "Building a Latent Semantic Index of an Image Database from Patterns of Relevance Feedback," Proceedings of International Conference on Pattern Recognition, 4, pp. 134-137, 2002.
[4] M. Li, Z. Chen, and H. J. Zhang, "Statistical Correlation Model for Image Retrieval," Pattern Recognition, Vol.35, pp. 2687-2693, 2002.
[5] J. Han, K. N. Ngan, M. Li, and H. J. Zhang, "A Memory Learning Framework for Effective Image Retrieval," IEEE Transactions on Image Processing, Vol.14, pp. 511-524, 2005.
[6] P.-Y. Yin, B. Bhanu, K.-C. Chang, and A. Dong, "Improving Retrieval Performance by Long-term Relevance Information," Proceedings of 16th International Conference on Pattern Recognition, Vol.3, pp. 533-536, 2002.
[7] W. Jiang, G. Er, and Q. Dai, "Multi-Layer Semantic Representation Learning for Image Retrieval," The IEEE International Conference on Image Processing, 2004.
[8] I. Gondra and D. R. Heisterkamp, "Summarizing Inter-Query in Content-Based Image Retrieval via Incremental Semantic Clustering," Proceedings of the International Conference on Information Technology: Coding and Computing, Vol.2, pp. 18-22, 2004.
[9] I. Gondra, D. R. Heisterkamp, and J. Peng, "Improving the Initial Image Retrieval Set by Inter-Query Learning with One-Class SVMs," Proceedings of the 3rd International Conference on Intelligent Systems Design and Applications, pp. 393-402, 2004.
[10] C. C. Hang, In Computer Science and Engineering, The Chinese University of Hong Kong, Hong Kong, p. 104, 2004.
[11] C. Burges, "Simplified Support Vector Decision Rules," Proceedings of the 13th International Conference on Machine Learning, 1996.
[12] J. T. Kwok and I. W. Tsang, "The Pre-Image Problem in Kernel Methods," Proceedings of the 12th International Conference on Machine Learning, 2003.
[13] X. S. Zhou and T. S. Huang, "Small Sample Learning During Multimedia Retrieval using BiasMap," IEEE Conference on Computer Vision and Pattern Recognition, Vol.1, pp. 11-17, 2001.
[14] X. S. Zhou and T. S. Huang, "Edge-based Structural Features for Content-based Image Retrieval," Pattern Recognition Letters, Vol.22, pp. 457-468, 2001.
[15] M. Stricker and M. Orengo, "Similarity of Color Images," Storage and Retrieval for Image and Video Databases III, SPIE Proceedings Vol.2420, pp. 381-392, 1995.
[16] L. Wang, K. L. Chan, and P. Xue, "A Criterion for Optimizing Kernel Parameters in KBDA for Image Retrieval," IEEE Transactions on Systems, Man, and Cybernetics, Vol.35, pp. 556-562, 2005.

Name:
Kien-Ping Chung

Affiliation:
Ph.D. Candidate, School of Information Technology, Murdoch University

Address:
South Street, Murdoch, Perth, Australia, 6150

Brief Biographical History:
1995 Received a Bachelor degree in Computer Science from Murdoch University
1999 Graduated with Second Upper Class Honour degree in Computer System Engineering from Curtin University of Technology
1999- Member of the Golden Key Society

Main Works:
• "Reducing User Log Size in an Inter-Query Learning Content-Based Image Retrieval (CBIR) System with a Cluster Merging Approach," International Joint Conference on Neural Networks, Vancouver, Canada, pp. 2163-2170, July, 2006.

Membership in Academic Societies:
• Institute of Electrical and Electronics Engineers (IEEE)

Name:
Chun Che Fung

Affiliation:
Associate Professor, School of Information Technology, Murdoch University

Address:
South Street, Murdoch, Western Australia, Australia 6150

Brief Biographical History:
1982-1988 Department of Electronic and Communication Engineering, Singapore Polytechnic
1989-2002 School of Electrical and Computer Engineering, Curtin University
2003 Joined Murdoch University

Main Works:
• "Simulated-annealing-based economic dispatch algorithm," I.E.E. Proceedings, Part C, Generation, transmission and distribution, Vol.140, No.6, pp. 509-515, 1993.
• "Modular Artificial Neural Network for Prediction of Petrophysical Properties from Well log data," I.E.E.E. Transactions on Instrumentation and Measurement, Vol.46, No.4, pp. 1295-1300, 1997.
• "The STAG Oilfied formation evaluation: a Neural Network Approach," Australian Petroleum Production & Exploration Journal, Vol.39, part 1, pp. 451-460, 1999.

Membership in Academic Societies:
• Senior Member, Institute of Electrical and Electronics Engineers (IEEE)
• Member, Institute of Engineer Australia (IEAust)
• Member, Australian Computer Society (ACS)