Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, 18-21 August 2005

MULTIPLE LAYAR KERNEL-BASED APPROACH IN RELEVANCE FEEDBACK CONTENT-BASED IMAGE RETRIEVAL SYSTEM

KIEN-PING CHUNG, CHUN-CHE FUNG

School of Information Technology, Murdoch University, Perth, Australia
E-MAIL: k.chung, l.fung@murdoch.edu.au

Abstract:

Relevance feedback has drawn intense interest from many researchers in the field of content-based image retrieval (CBIR). In recent years, kernel-based approach has been a popular choice for the implementation of the relevance feedback based CBIR system. This is largely due to its ability to classify patterns with limited sample data. Since most of the kernel approaches reported have been treating the input as a long flat vector, such arrangement may increase the chances of "polluting" the feature element that uniquely identifies the selected image group. This paper proposes a two layer kernel configuration with an objective to improve the retrieval accuracy. While the performance of the two configurations is similar in certain conditions, the proposed configuration has shown to superior when dominant feature element exists that is capable to uniquely identify the selected image group.

Keywords:

Content-Based Image Retrieval System (CBIR); Relevance Feedback; Kernel Method; Kernel Bias Discriminant Analysis (KBDA)

1. Introduction

Content-based image retrieval (CBIR) system has been one of the most active areas of research in recent years. One of the main reasons is due to its potential in many commercial applications such as facial recognition for security surveillance system, or, in computer aided diagnostic systems for medical applications. The other reason for the high level of activities may be due to the fact that such system is relatively new, and most of the systems reported in the literatures are still in the prototype stage. Issues such as semantic gap and indexing structure are still largely unanswered.

Recently, relevance feedback in CBIR system has gained much attention from the research community. It is a strategy that invites interactive inputs from the user to refine the query for subsequent retrieval. This approach generally starts from prompting users to search the system via keywords, image examples or a combination of both. The system then prompts the user to select the relevant images from the search results. After the user selected the images, the system will refine the original query by analyzing the common features among the selected images. This process is continued iteratively until the target is found. The selection of the common features will be the most appropriate for applications incorporating intelligent technologies such as neural network and fuzzy logic. This is due to the need for fine-tune and modification of the process with human input. Evolutionary computation techniques could also be used in optimizing the process.

Over the past ten years, relevance feedback has been evolved from a simple machine learning problem such as frequency of occurrence [1], Bayesian classification [2], and now, to the more popular kernel based approach [3-6]. Kernel technique has long been used in the statistic pattern recognition applications. Its recent popularity gain in CBIR applications is mostly due to its ability to analyze small sample data and the classification of non-linear data. The integration of kernel based approach with discriminant analysis reported by authors such as Tao and Tang [3], and, Zhou and Huang [6] have shown improvement from the earlier proposals.

In a general CBIR system, it is impossible to know what feature model/s can be used to capture the unique identity of certain groups of images. Hence, one idea is to employ as many image features as possible in hope that one has the ability to capture the unique feature of the targeted images. Such idea introduces problems if the image features are treated as a cascade of one flat vector. Such arrangement may increase the chances of "polluting" the feature element that uniquely identifies the selected image group.

Inspired by the kernel based discriminant analysis as introduces by Zhou and Huang [6] and the hierarchical framework as proposed in MARS [1] system, this paper proposes a kernel-based discriminant analysis hierarchical relevance feedback framework for the content-based image retrieval system. The idea is to project each image feature separately to a feature space and calculate the Euclidean distance of these features in the new space with reference to the positive centroid. Similarly, the calculated distance of each feature vector is again treated as inputs to another feature space, and thereby calculating the Euclidean distance of each image and finally ranking the images. Such arrangement has the advantage of treating each image feature independently and thus, increases the weight alias with the more important feature elements.

This paper will first provide a brief description on the background of the theory, and follow by the description of the proposed framework. The paper will then look at the experiment conducted on the proposed method and lastly, conclusion will be drawn based on the experiment finding.

2. Background

2.1. Kernel Bias Discriminant Analysis (KBDA)

The idea of KBDA [6] is to transfer data from the original space to a new feature space that can best discriminate the positive from negative given samples. This is done be applying a set of weight vectors W that maximize the ratio between the positive covariance matrix S_x^ϕ and the biased matrix S_y^ϕ. The problem can be expressed as:

$$W_{opt} = \arg\max_{w} \frac{\left\| W^T S_y^\phi W \right\|}{\left\| W^T S^\phi W \right\|} \qquad (1)$$

The positive covariance matrix S_x^ϕ and the biased matrix S_y^ϕ are defined as:

$$S_y^\phi = \sum_{i=1}^{Ny} \left(\phi(y_i) - m_x^\phi\right)\left(\phi(y_i) - m_x^\phi\right)^T \qquad (2)$$

$$S_x^\phi = \sum_{i=1}^{Ny} \left(\phi(x_i) - m_x^\phi\right)\left(\phi(x_i) - m_x^\phi\right)^T \qquad (3)$$

where $\{x_i = 1,\ldots,N_x\}$ denote the positive examples, $\{y_i = 1,\ldots,N_y\}$ are the negative examples given, and, Φ is the kernel mapping function. One can view Equation (1) as a problem of generalized eigenanalysis where the optimal eigenvectors associated with the largest eigenvalues are the weight factor for the new feature space. By knowing the weight, one can now project the new input pattern z onto the new space:

$$new\_space = w^T \phi(z) \qquad (4)$$

2.2. Proposed Method

Figure 1 illustrates the abstract computation model of the proposed method. It is essentially what Zhou and Huang [6] proposed except the authors of this paper have restructured the analysis of the input vector into two layer. In this configuration, each image feature vector is processed separately by individual KBDA module. The outcome of the low-level KBDA module is the Euclidean distances of projected point with reference to the positive centroid. These outcomes serve as inputs to another KBDA module which will yield a new point in the final projected space. This new point will be used to compute the final Euclidean distance for ranking purpose. The authors called this configuration as the hierarchical kernel biased discriminant analysis (HKBDA).

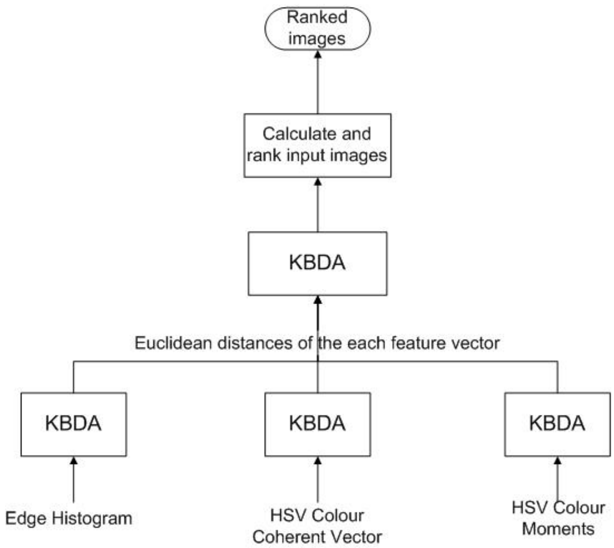

Figure 1: The proposed HKBDA framework.

The steps of HKBDA can be summarized as follows:

1. Project each different feature vectors of the given positive and negative examples into a new feature space by using KBDA.
2. In the new space, calculate the Euclidean distances of each example from the positive centroid.
3. Similar to Step 1, project the calculated distances to another new feature space by using KBDA.
4. In this new space, return the points corresponding to the Euclidean nearest neighbors from the positive centroid. Wait for the user feedback then go to Step 1.

3. Experiment and Evaluation

3.1. Can The Data Be "Polluted"?

Figure 2 and **Figure 3** show the results produced from KBDA and HKBDA. In this experiment, 24 images from Corel digital library were selected with three labeled as positive images with a common theme and the non-selected images as negative samples. These images are selected with prior knowledge that the two image groups can be separated by one of the feature elements in HSV color coherent vector (CCV) [7]. The first plot on

Figure 2 shows the Euclidean distances produced by KBDA when only CCV is applied in analyzing the images. In this plot, the three selected positive images clearly have shorter Euclidean distances with reference to the positive centroid than all the other negative images. On the second plot, HSV moments [8] has also been added to the same system and the plot showing the Euclidean distances between the positive selected images and two other negative images have diminished. On the last plot, another feature, global edge histogram [9], has again been added to the system. By now, the Euclidean distance of one of the negative images is shorter than one of the selected positive images. Together, these plots clearly show the unique feature identifying the two groups of images have been polluted by the extra image features added to the system.

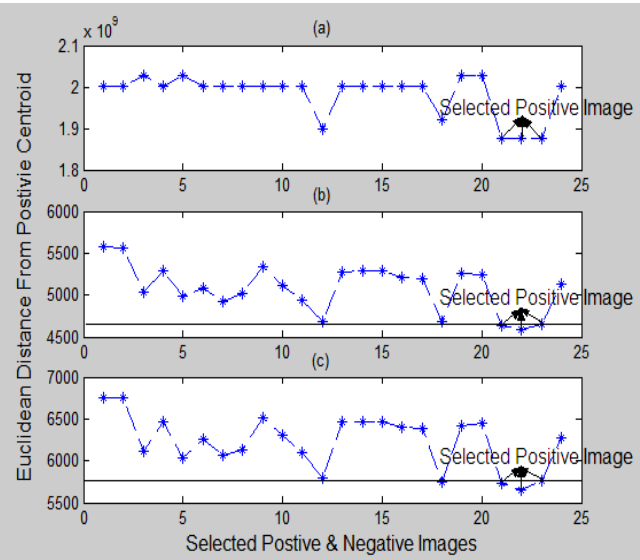

Figure 2: Euclidean distances computed using KBDA. (a), distance calculated by using HSV CCV as image feature. (b). distances calculated by using HSV CCV and moments as image features. (c), distances calculated by using HSV CCV, HSV moments and global edge histogram as features.

The same procedures have been applied HKBDA and the results are shown in **Figure 3**. The plots have shown that although the scales of the plots vary, the shape of the curves has not changed. More importantly, regardless of the additional features, the three selected positive images have the shortest Euclidean distances on all three plots.

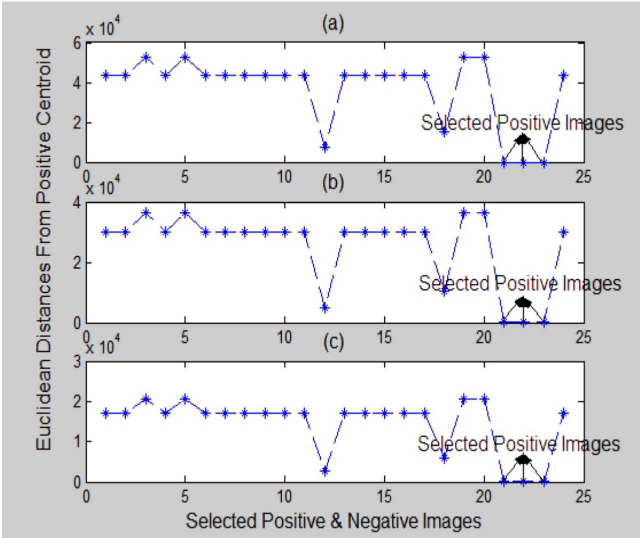

Figure 3: Euclidean distances computed using HKBDA.

3.2. The Prototype System

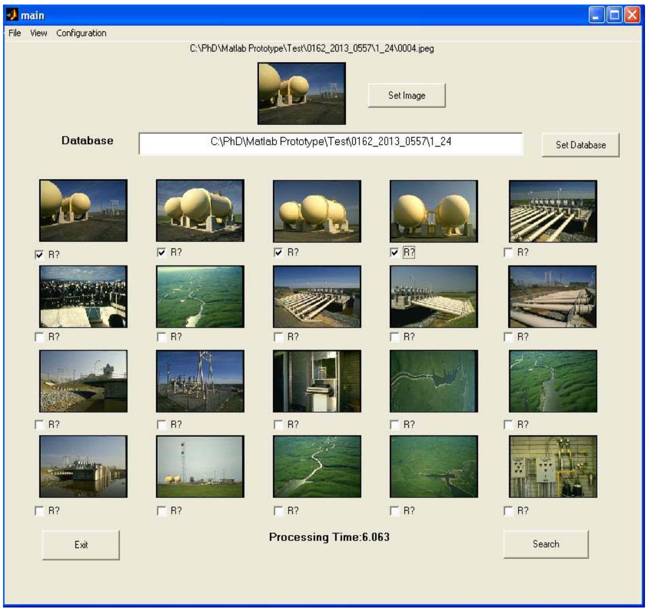

Figure 4: User interface for the prototype system.

To evaluate the performance of the proposed approach, the authors have designed and implemented a prototype image retrieval system using Matlab. The user interface is shown in **Figure 4**. This prototype system provides the user with the ability to query the image database via an image sample. After the first retrieval iteration, users can select the relevant image while ignoring the non-relevant images. The system will label the selected images as positive images while treating the ignored images as negative images. The retrieval procedure of the prototype system is as follows:

1. User inputs a query image;
2. The 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. User selects the positive images and the rest will be automatically labelled as negative images.
5. Query and all images in the database are projected in the kernel space based on the HKBDA approach.

For this experiment, Radial Basis Function (RBF) is used as the kernel for both configurations. The purpose is to see how the two configurations perform under similar settings. As for the image features, the authors have selected the water-filling edge histogram [10], HSV colour coherent vector [7], HSV histogram, global edge direction histogram [9], HSV colour moments [8] and colour intensity histogram. Together, 66 feature elements have been used for this testing. It is the authors' intention to use as many feature elements as possible. This is specifically done by polluting the important features with noise.

It is extremely difficult to test and verify the retrieval result if one is to let the system performing search on a database. In addition, relevancy of an image can be subjective. Thus, the authors have developed a set of testing strategies to ensure the test result can be easily compared and measured. The idea is to let a user to select the relevant images from a group of images. After the user's selection, the system is to re-rank this group of images, and ideally, the user selected images will have a higher rank than the other images. The exact testing steps are as follow:

1. User inputs a query image.
2. User selects an image database.
3. On the first iteration, the system will retrieve and rank the image based on the Euclidean distance measure on the included image features. In this iteration, the weight is set equally to all the features.
4. User selects the relevant images from the retrieved images.
5. Base on the current retrieved images, the system will re-calculate Euclidean distances of each image using the two configurations and re-rank the already retrieved images accordingly.

3.3. Results of the Prototype System

The performance of the proposed approach is evaluated according to the retrieval accuracy and distance ratio between the closest negative image and the furthest positive image. The bigger the ratio implies the further the positive images are from the negative images, and hence, better separation. The accuracy of the retrieval is calculated by dividing the number of correctly identify positive images by the number of positive images selected by users.

Table 1, Retrieval result from KBDA

Concept	Accuracy	Distance ratio
Tank	50%	0.0259
Landscape with river	78%	0.9973
Bird	100%	1.0193
Mouse	50%	0.9617
Yellow flower	50%	0.9988

Table 2, Retrieval result from HKBDA

Concept	Accuracy	Distance ratio
Tank	75%	0.8571
Landscape with river	100%	3.02
Bird	100%	8.322
Mouse	50%	0.0698
Yellow flower	50%	0.0737

Table 1 and 2 are the retrieval results gathered from 130 images from the Corel image database. The images were retrieved and classified under five different themes. These results show both configurations performed poorly on the image groups 'mouse' and 'yellow flower'. The relatively low accuracy in both tables indicate both configurations are unable to distinguish the selected images from the negative labeled images. As for the other three groups of images, HKBDA out-performed KBDA in both retrieval accuracy and distance ratio. After careful analysis on the image groups, the authors have found the reason for the poor performance of the two image groups is because the image analysis algorithms included are unable to extract the feature/s that clearly distinguish the two image groups. This observation implies that in terms of retrieval accuracy and classification ability, HKBDA out-performs KBDA when the images can be identified by a unique feature.

4. Conclusion

A new multi-layer kernel based framework for the relevance feedback content-based image retrieval system has been introduced in this paper. The proposed framework combines kernel based disciminant approach with multi-layer analysis framework. Using this framework, an improvement has been shown in retrieval and classification ability as compare to the original framework in which the input was treated as a flat vector.

In this paper, the authors have only applied the proposed framework to a kernel-based algorithm. More detailed study is going to be conducted by the authors to analyze the characteristic of such framework further by applying it to several other kernel-based algorithms. One of the future directions is to incorporate relevance feedback in the implementation of computational intelligent techniques for the tasks of image retrieval. The other area of research will be profiling and personalization of the user in order to provide first level of “implicit” relevancy feedback.

Acknowledgement

This project is supported by a grant awarded by the 2004 Murdoch University Research Excellence Grant Schemes (REGS). Kien Ping is also a recipient of the Murdoch Research Scholarship (MURS) from 2005.

References

[1] Y. Rui, T. S. Huang, and S. Mehrotra, "Content-Based Image Retrieval with Relevance Feedback in MARS," Proceedings of the International Conference on Image Processing, Washington, DC, October 26-29 1997.

[2] I. J. Cox, M. L. Miller, S. M. Omohundro, and P. N. Yianilos, "PicHunter: Bayesian Relevance Feedback for Image Retrieval," Proceedings of the 13th International Conference on Pattern Recognition, Vienna, Austria, August 25-29 1996.

[3] D. Tao and X. Tang, "A Direct Method to Solve the Biased Discriminant Analysis in Kernel Feature Space for Content Based Image Retrieval," Proceedings of IEEE International Conference Acoustics, Speech, and Signal Processing, Montreal Canada, May 19 2004.

[4] S. Tong and E. Chang, "Support Vector Machine Active Learning for Image Retrieval," Proceedings of 9th ACM International Multimedia Conference, Ottawa, Cananda, September 30 - October 05 2001.

[5] G.-D. Guo, A. K. Jain, W.-Y. Ma, and H. J. Zhang, "Learning Similarity Measure for Natural Image Retrieval With Relevance Feedback," *IEEE Transactions on Neural Networks*, vol. 13, pp. 811-820, 2002.

[6] X. S. Zhou and T. S. Huang, "Small Sample Learning During Multimedia Retrieval using BiasMap," IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, United States, December 2001.

[7] G. Pass, R. Zabih, and J. Miller, "Comparing Images Using Color Coherence Vectors," The 4th ACM international conference on Multimedia, Boston, Massachusetts, United States, November 18-22 1997.

[8] M. Stricker and M. Orengo, "Similarity of Color Images," Storage and Retrieval for Image and Video Databases III, San Diego/La Jolla, CA, USA, February 5-10 1995.

[9] D. K. Park, Y. S. Jeon, and C. S. Won, "Efficient Use of Local Edge Histogram Descriptor," Proceedings of the 2000 ACM workshops on Multimedia, Los Angles, Calfornia, United States 2000.

[10] 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.