A Hierarchical Nonparametric Discriminant Analysis Approach for a Content-Based Image Retrieval System

Kien-Ping Chung and Chun Che Fung

School of Information Technology, Murdoch University, Perth, Australia
Centre for Enterprise Collaboration in Innovative Systems
E-MAIL: k.chung, l.fung@murdoch.edu.au

Abstract

This paper proposes a Hierarchical Nonparmetric Discriminant Analysis (HNDA) content-based image retrieval (CBIR) system for E-Business applications. It has the potential to become an important and integral component for future e-Business applications. Developments in CBIR have drawn interest from many researchers and practitioners in recent years. The challenge is how to retrieve the most appropriate or relevant images at the fastest speed. To increase the retrieval speed, most of the systems pre-process the stored images and extract out the essential features. Such scheme only works well for the server type database system. Such approach is not feasible for systems that analyze images in real-time. In this paper, a hierarchical multi-layer statistical discriminant framework is proposed. The system is able to select the most appropriate features by analyzing the newly received images, and then apply a Relevance Feedback (RF) approach to improve the retrieval accuracy. As the number of features being analyzed is less, an improvement in performance is achieved.

1 Introduction

Image provides an important dimension in e-Business. It provides much richer information and content to the clients and users than text alone. While video and animation may extend the dimension further, such files will have greater demands in terms of storage and bandwidth requirements.

In general, an image retrieval system is a computer system for browsing, searching and retrieving images from a large database or multiple databases of digital images. While such activities can be easily based on textual descriptions of the images in the form of keywords or "tags", manual naming and labeling of the images is laborious and time-consuming. It can also be ambiguous as different people may interpret the key words or features differently. Content-based image retrieval (CBIR) aims at reducing the need for textual description and to provide the most appropriate images automatically. In addition, Relevance Feedback (RF) can be used to refine the search and to define the scope of the search or retrieval.

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 the explosion in the number of images available in digital format. Combining this factor along with the easy to access World-Wide Web (WWW), digital images become very easy to obtain by any person. This can be used in conjunction within the context of e-Business for B2C, B2B, P2P and C2C applications.

CBIR can be defined as a system that retrieves an image based on the semantic or visual content of the query which can either in the format of text or image. The images in these systems are often indexed in the form of text, certain image feature vectors or a combination of both. To enhance the retrieval speed, most of the system often pre-processed the images stored in the database. However, if one is to retrieve images from WWW, unless one is accessing common databases such as Yahoo or Google, it is often required to process all the download images in real time. In this case, the feature extraction speed becomes crucial. And ideally, the system should only use feature extraction algorithms that best discriminate between the positive and the negative label images.

However, in a general CBIR system, it is impossible to know what feature model/s should be used to capture the unique identity of certain groups of images. Hence, one idea is to employ as many image features as possible and only select the most appropriate features to classify the images. In general, most of the systems treat the extracted features as a cascade of one *flat* vector. In this arrangement, it is rather difficult and messy to determine the discriminant ability of each extracted feature.

Inspired by the discriminant analysis as introduces by Tao and Tang [1] and the hierarchical framework as proposed in MARS [2] system, this paper proposes a nonparametric discriminant analysis hierarchical relevance feedback framework for the content-based image retrieval system. In particular, the system is aimed to perform in real-time with access to images from the World Wide Web for e-Business applications. As explain later, such

Proceedings of the 2005 IEEE International Conference on e-Business Engineering (ICEBE'05)
0-7695-2430-3/05 $20.00 © 2005 IEEE

arrangement provides the system with the ability to perform automatic feature selection during the retrieval process.

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

2 Background

2.1 Nonparametric Discriminant Analysis (NDA)

Nonparametric discriminant analysis [1] is an extension of biased-discriminant analysis (BDA) [3] in that both techniques are biased toward the positive examples. However, unlike BDA, NDA does not require all positive samples to come from a single Gaussian distribution. The data transformation can be non-linear. The objective of the NDA is to apply a set of transform vector W that maximizes the ratio between the positive covariance matrix S_x and the biased matrix S_y. The problem can be expressed as:

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

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

$$S_y = \sum_{i=1}^{Ny} \left(y_i - m_{yi}^{kx}\right)\left(y_i - m_{yi}^{kx}\right)^T + \sum_{i=1}^{Nx} \left(x_i - m_{yi}^{ky}\right)\left(x_i - m_{yi}^{ky}\right)^T \qquad (2)$$

$$S_x^{\phi} = \sum_{i=1}^{Nx} \left(x_i - m_{xi}^{kx}\right)\left(x_i - m_{xi}^{kx}\right)^T \qquad (3)$$

where $\{x_i = 1,\ldots,N_x\}$ denote the positive examples, $\{y_i = 1,\ldots,N_y\}$ are the negative examples given. 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:

$$new\_space = wz \qquad (4)$$

2.2 Proposed Method

Figure 1 illustrates the abstract computation model of the proposed method. It is based on the proposal by Tao and Tang [1] except the authors of this paper have restructured the analysis of the input vector into two layers. In this configuration, each image feature vector is processed separately by individual NDA module. The outcome of the low-level NDA module is the Euclidean distances of projected point with reference to the positive centroid. These outcomes serve as inputs to another NDA 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 nonparametric discriminant analysis (HNDA).

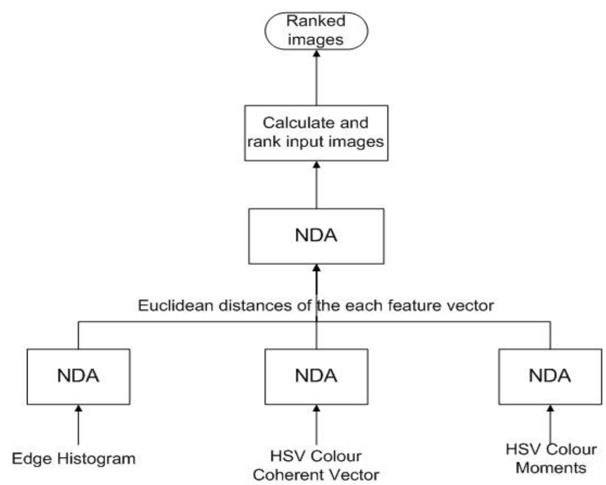

Figure 1. The proposed HNDA framework.

The steps of HNDA 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 NDA.
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 NDA.
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.

2.3 Feature Selection Criteria

As discussed in the previous section, the outcome from the low-level module shown in **Figure 1** is the Euclidean distances of the projected point with reference to the positive centroid. Hence, the smaller the value the closer it

Proceedings of the 2005 IEEE International Conference on e-Business Engineering (ICEBE'05)
0-7695-2430-3/05 $20.00 © 2005 **IEEE**

is to the positive centroid. By applying the discriminant analysis as discussed in Section 2.1, one will expect the bigger the ratio the better it is to discriminate the positive from the negative samples. The discriminant ratio can be shown as follow:

$$ratio = \frac{S_y}{S_x} \tag{5}$$

2.4 Overview of the Proposed System

In this paper, it is proposed that a system will act as an independent agent from the major search engines. Instead of storing the images in a database, the system only has to analyze the collected images from other search engines and output only those are relevant. Like any other search engines, the image retrieval process is triggered by search query performed by the user. This system does not require a large database, but at the cost of slower feedback response time. To accelerate the retrieval process, feature selection becomes a crucial factor in this system.

Figure 2 is the use-case diagram for the proposed Web content-based image retrieval system. It shows that the retrieval process begins by user entering a search phrase or keyword into the system's user interface. The system then alias with the internet search engine to download all the images related to the search phrase. Using the rank provided by the search engines, the system will then retrieve the first batch of images. User will then select the relevant images. From the label samples, the system will select the appropriate features for classifying the positive and negative images. The selected features will then be used to process and rank all the downloaded images, and the user can again select the appropriate images to trigger the next retrieval process. The process will refine subsequent searches and provide the most appropriate images.

3 Experiment and Evaluation

To evaluate the performance of the proposed HNDA framework, the authors have selected results or data from the water-filling edge histogram algorithm [4], HSV color coherent vector [5], HSV histogram , global edge detection algorithm [6], HSV color moments [7] and color intensity histogram as features for the system. It is the author's intention to use as many feature as possible with an expectation that at least one feature will possess the character that can uniquely describing the selected positive label images. In addition, this is also specifically designed to demonstrate the proposed framework's feature selection capability.

Understandably, it is 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 that the test results 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 test procedure is 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.

3.1 Results of the Prototype System

The performance of the proposed approach is evaluated according to the retrieval accuracy and distance ratio as defined in (5). A big ratio implies that the positive images are further away 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 results from HNDA

Concept	Accuracy	Ratio	Feature Selected
Landscape	89%	5.8828	3/7
Tank	100%	11058	2/7
Bird	100%	1.7548e+007	2/7
Mouse	100%	3975	3/7
Orange Flower	100%	2023	2/7
Yellow Flower	100%	1.7153e+003	1/7

Table 2. Retrieval result from NDA

Concept	Accuracy	Distance ratio
Landscape	89%	19.12
Tank	100%	4815
Bird	100%	3.8594e+006
Mouse	100%	40
Orange Flower	80%	16
Yellow Flower	100%	1.7132e+004

Table 1 and 2 are the retrieval results gathered from 410 images from the Corel image database. The images were retrieved and classified under six different themes. The result shows that while the retrieval performance of both algorithms is similar, HNDA has the advantage in using fewer features for obtaining the similar retrieval accuracy.

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 the statistical discriminant approach with multi-layer analysis framework. Using this framework, an improvement has been shown in retrieval speed as compared to the original framework in which the input was treated as a flat vector. The testing shows the potential of this framework to be used in e-Business applications and systems where only the raw images are stored in the database. World-Wide-Web is one of the typical applications.

This paper has shown that the test results for image groups that are visually similar. One of the future directions is to incorporate this framework with images that are similar semantically but differ visually. In addition, this paper only uses retrieval accuracy as the feature selection criteria. Research effort will be spent on incorporating the complexity of the feature extraction algorithm into the feature selection decision making process.

5 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.

6 References

[1] D. Tao and X. Tang, "Nonparametric Discriminant Analysis in Relevance for Content-based Image Retrieval," Proceedings of the 17th International Conference on Pattern Recognition, Cambridge, United Kingdom, August 23 - 26 2004.

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

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

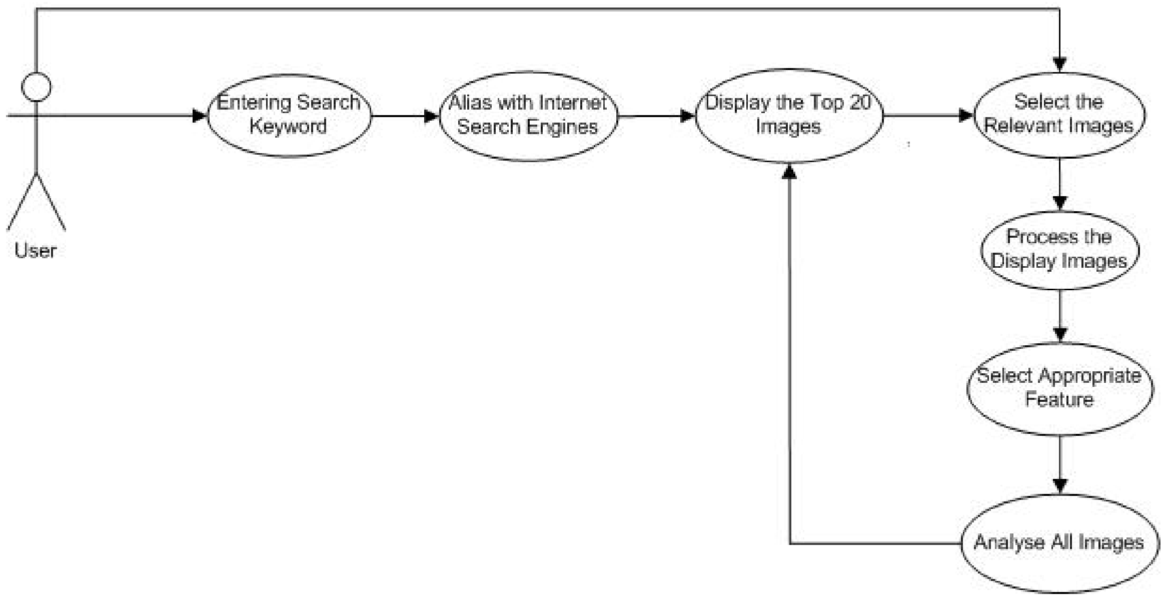

Figure 2. Use case diagram for the proposed content-based image retrieval system.

0-7695-2430-3/05 $20.00 © 2005 IEEE