Modified Relevance Feedback for Content Based Image Retrieval Using Support Machine

Pei-Geok Lim
(anggie@gmail.com)
Intelligent Software Engineering Laboratory,
Faculty of Computer Science & Information Systems,
Universiti Teknologi Malaysia,
81310 UTM Skudai, Johor, Malaysia

Ali Selamat
(aselamat@utm.my)
Intelligent Software Engineering Laboratory,
Faculty of Computer Science & Information Systems,
Universiti Teknologi Malaysia,
81310 UTM Skudai, Johor, Malaysia

Abstract: The rapid growth of the computer technologies and the advent of World Wide Web have increased the amount and the complexity of multimedia information. Content-based image retrieval (CBIR) system has been developed as an efficient image retrieval tool whereby user can provide their query to the system to allow it to retrieve the user's desired image from image database. However, the traditional relevance feedback of CBIR has some limitations that decrease the performance of the CBIR system that are the imbalance of training set problem, classification problem, limited information from user problem, and insufficient training set problem. Therefore, in this paper we propose a modified relevant feedback method to support the user query and user profiles based on the weight ranking of the image retrieved. The Support Vector Machine (SVM) has been used to support the learning process in order to reduce the semantic gap between user and the CBIR system. From the experiments, the proposed learning method has enabled the users to improve their search results based on the precision and accuracy.

Keywords: Image Retrieval, Content-based Image Retrieval (CBIR), Relevance Feedback, and Support Vector Machines (SVM).

I. INTRODUCTION

Images are considered as the prime media type to be used to retrieve hidden information within data. In general, an image retrieval system is a computer system for browsing, searching, and retrieving images from a large digital image database. Early trends witnessed image retrieval utilizing some method of adding metadata such as captioning, keywords, or descriptions to the images [1] so that retrieval can be performed over the annotation words. The issue of subjectivity of human perception here means that the perception of different persons or the same person for the same image may vary under different circumstances [2]. Beside the human subjectivity issue, similarity is another issue that is highly focused on in CBIR. This issue has affected the retrieved results based on the similarities of pure visual features that are not necessarily perceptual and semantically meaningful. In addition, each type of visual feature tends to capture only one aspect of the image property and it is usually hard for a user to specify clearly how different aspects are combined [3]. Hence, there are semantic gaps between the low-level features and the high level query in the CBIR system. The semantic gap expresses the discrepancy between the low-level features that can be readily extracted from the images and the descriptions that are meaningful for the users. To solve these problems, an interactive relevance feedback which involves the interaction between human and system was introduced. Relevance feedback is a supervised active learning technique which uses the positive and negative examples feedback from the users to improve system performance. For a given query, the system first retrieves a list of ranked images according to a predefined similarity metrics. Then, the user marks the retrieved images as relevant (positive examples) or not relevant (negative examples) to the query. The system will refine the retrieved results based on the feedback and present a new list of images to the user. This process will go through several iterations until the user is satisfied with the retrieved result.

II. RELATED WORKS

The traditional CBIR relevance feedback techniques include query refinement [2] and re-weighting [4]. However, both techniques did not deliver satisfying performances for CBIR due to several issues. The most important issue in relevance feedback is how to incorporate positive and negative examples to refine the query and/or to adjust the similarity measure [5]. On the other hand, classification is another issue that needs to be considered by the CBIR domain. However, classification problem occurred in relevance feedback for CBIR when the learning problem regarding the positive samples (relevant images) and the negative samples (irrelevant images) are presented as two difference groups [6]. In this situation, it becomes difficult to retrieve the positive images which may disperse in the feature space; as a result, it is difficult to retrieve them directly based on low-level feature similarity whether they are refined or not [7]. Thus, a classifier or statistical learning technology is needed in order to identify these groups into positive and negative examples in the feature space [5]. During the feature similarity measure part, it will find the similar texture region based on the query image from the set of images in the database. A similarity comparison within the texture feature of query image and the database images will be conducted in this part. At the end of this part, the images that are over certain predefine similarity index threshold will be retrieved into the relevance feedback part for classification purposes. After the feature similarity part, a set of possible images will be retrieved by the system and sent to the user. The user will determine and mark the images as relevant or irrelevant in the relevance feedback part. The selected images that are marked by the users will be considered as relevant images and treated as input for support vector machine for adaptive learning purpose. This process will be repeated iteratively until the user is satisfied with the feedback images that are retrieved by the system. Finally the preferences of the user will be captured and acknowledged by the system. The system will automatically retrieve the relevant images based on the user's query.

III. TEXTURE EXTRACTION METHOD

After obtaining the user's query, the query image will go through a pre-processing process. In the pre-processing process, the image will be resized to 24 x 24 dimensions. The purpose of resizing is to help reduce the computations and complexity. This will also cause some of the information being lost after the image has been resized. After that, the image will be transformed into a gray scale image. After the pre-processing process is completed, the texture feature will be extracted from the gray scale image. Each vector in the texture feature represents the index of each pixel in the image. Each obtained decimal magnitude of pixel will be thresholded with an identified fixed threshold value. In this case, the threshold value is set at 100. A binary texture vector will be obtained after the thresholding process.

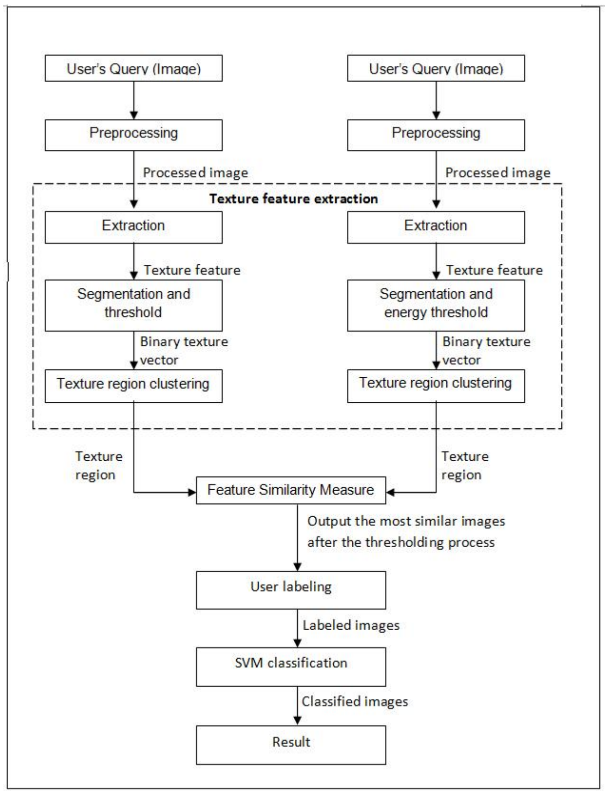

Fig. 1: Methodology of the CBIR by using SVM.

Equation (1) shows how the binary texture vector will be obtained from the gray scale image.

$$b_k[m,n] = \begin{cases} 1 & if \quad y_i[m,n] \geq \mathrm{T}_k \\ 0 & otherwise \end{cases} \qquad (1)$$

where $b_k[m,n]$ = binary value of each pixel with m and n coordinate

$y_k[m,n]$ = texture feature vector of each pixel with m and n coordinate

T_k = threshold

Texture segmentation is the process by which the image is split into different regions of homogeneous texture. By using the binary feature set that is obtained after thresholding process, the texture segmentation will take place to segment the binary feature set to some significant texture set with lower dimension. This segmentation mechanism can help to reduce the computational burden and at the same time produce some significant texture feature set. Therefore, some of the insignificant texture vector or regions will be eliminated. In this project, the segmentation mechanism uses the threshold method which the segmented area will be output as binary one if that area contains 50% of binary one. Otherwise, the segmented area will be output as zero if it contains less than the threshold value.

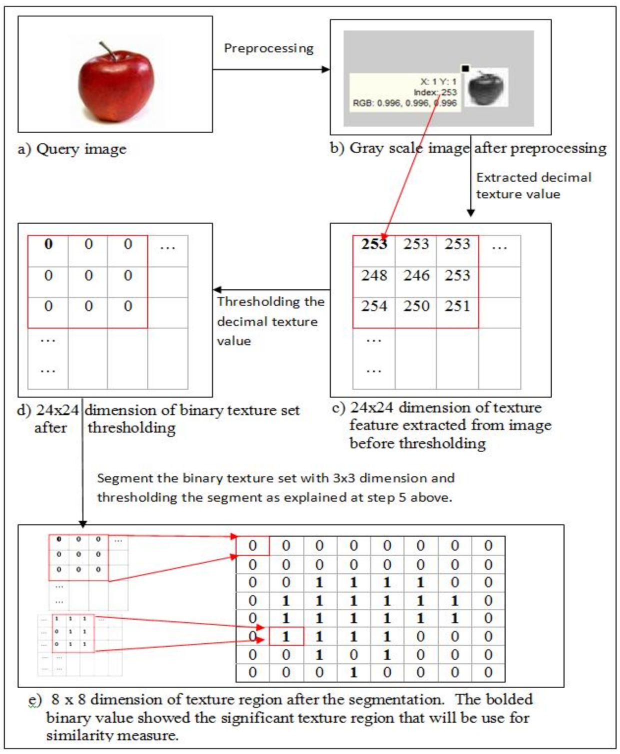

Fig. 2: The texture feature extraction mechanism.

1. FEATURE SIMILARITY MEASURE

The feature similarity measure will compare the similarity percentage between the texture regions of two images. Therefore, the similarity percentage shows how similar an image is from the image database to the query image. As a result, an output list that shows the similarity percentage to an input query will be obtained as shown in Figure 1. From the output list, the image which has the similarity percentage more than 40% will be retrieved for the user to mark, as shown in Figure 2.

2. SVM BASED RELEVANCE FEEDBACK

As we know, the relevance feedback process involves the interaction between human and the system, and this process will go through several iterations until the user is satisfied with the output images. In our experiment, relevance feedback process is aimed to evaluate the accuracy of CBIR system with the usage of the SVM classifier and without using of SVM classifier. A list of retrieved images from the feature similarity measurement will be given to the user. User will mark the images that are relevant to the query images. The unmarked images among these retrieved images will be considered as irrelevant images. Both relevant and irrelevant retrieved images will be used as training samples in classification process. Support Vector Machine (SVM) is used to learn the pattern of the relevant and irrelevant images within the training samples and classify the unlabeled images within the image database. In the end, the images from the image database were being classified as relevant or irrelevant images to the query image by the classifier. The new ranked list of similar images according to the result of the SVM will be retrieved in the next iteration of relevance feedback process.

Table 1: Type of dataset that use for experiment

Dataset	Category	Size
Dataset 1 (30 Samples)	Fruits	10
	Natural Scene	10
	Building	10
	Total	**30**
Dataset 2 (60 Samples)	Fruits	20
	Natural Scene	20
	Building	20
	Total	**60**
Dataset 3 (90 Samples)	Fruits	30
	Natural Scene	30
	Building	30
	Total	**90**

IV. EXPERIMENTAL RESULTS

In this study, the experiment uses three types of image dataset which are dataset 1, dataset 2, and dataset 3. Each dataset consists of three data categories which are fruits, natural scene, and building, as shown in Table 1. The main purpose of these dataset is to evaluate the performance of texture feature extraction. The results of this experiment will be evaluated using the accuracy rate measurement as shown in Eq. (2). In this experiment, the accuracy rate is calculated as the percentage of the images being identified as relevant or irrelevant image correctly by the system and human over the total of images in the database.

Accuracy Rate =

$$\frac{\textit{Total retrieved images that correctly identified as relevant or irrelevant images}}{\textit{Total image samples}} \qquad (2)$$

Table 2: Similarity measurement for image database of dataset 1.

Category	Fruits	Natural Scene	Building
Comparison Result	0.47826	0.34783	0
	0.3913	0.65217	0.52174
	0.47826	0.26087	0.086957
	0.56522	0.17391	0
	0.21739	0.34783	0.34783
	0	0.30435	0.17391
	1	0.13043	0.043478
	0	0.73913	0.043478
	0.21739	0.17391	0.86957
	0.78261	0.26087	0.3913
Possibly retrieval by system	5	2	2

V. RESULT AND DISCUSSION

The experiments were conducted using the dataset that are shown in Table 1. The experiments were conducted under two different approaches which are relevance feedback without SVM and relevance feedback with SVM. We conducted the first experiment using relevance feedback without SVM. After the texture feature extraction process, the query images and database images would go through the feature similarity measurement. The percentage of texture region similarity within query images and database images would be measured. Images that were over a certain predefine similarity index threshold would be retrieved by the system for the user to do the user labelling process (relevance feedback process). In this experiment, we predefined the similarity threshold as 0.5. The images database with over 0.5 similarity index would be retrieved by the system and sent for the user to label them. Table 2 shows the similarity measurement result using dataset 1 (30 image samples). We found that there were 5 fruits images, 2 natural scene images, and 2 building images that were considered similar and retrieved by the systems respectively, as shown in Figure 3. This retrieval result is evaluated using the accuracy formula. Similar experiments had been conducted using two other dataset which are dataset 2 (60 image samples) and dataset 3 (90 image samples). Figure 5 shows the comparison of the accuracy measurement for the approach of relevance feedback with SVM and without SVM, using different size of dataset which are 30, 60 and 90 image samples.

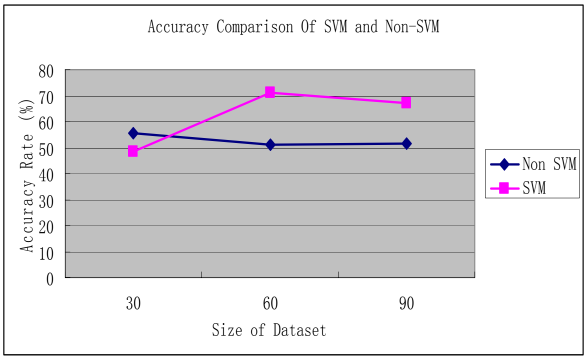

Fig. 3: Comparison of accuracy measurement for non-SVM and SVM according to different size of dataset.

We notice that for the relevance feedback without the SVM approach, the accuracy rate decreases as the size of dataset increases. This experiment result illustrates that the system could not perform adaptive learning and even insufficient (pls insert insufficient for what) especially when the number of images expanded. Meanwhile, on the average, the system also does not achieve a satisfactory performance with results lower than 60 percent accuracy rate. The experiment results illustrate that a better relevance feedback technique is necessary in order to improve the CBIR performance.

VI. CONCLUSIONS

In this paper, we propose a relevance feedback using the SVM learning method to retrieve images according to user's preferences. This proposed method has been used to support the learning process in order to reduce the semantic gap between the user and the CBIR system. The results of experiments have shown the improvement of the search result based on the precision and accuracy.

ACKNOWLEDGEMENTS

This work is supported by the Ministry of Science & Technology and Innovation (MOSTI), Malaysia and Research Management Center, University of Technology Malaysia (UTM) under the Vot 79200.

REFERENCES

[1] F. Long, H. Zhang and D. F. David. "Multimedia Information Retrieval and Management- Technological Fundamentals and Applications, chapter Fundamentals of Content-Based Image Retrieval," Springer, 2003.

[2] Y. Rui, T. S. Huang, M. Ortega, and S. Methrotra. "Relevance feedback: A power tool in interactive content-based image retrieval," *IEEE Trans. On Circuits and Systems for Video Techonology*, 8(5), 644-655, 1998.

[3] D.H. Kim, J.W. Song, J.H. Lee and B.G. Choi, "Support Vector Machine Learning for region-based image retrieval with Relevance Feedback," *ETRI Journal. Volume 29, Number 5*, October 2007.

[4] Y.Rui and T.S.Huang. "A novel relevance feedback techniques in Image retrieval," *In: Proc. 7th ACM Conf. on Multimedia, pp. 67–70*, 1999.

[5] D.C. Tao, X.O. Tang, X.L. Li, X.D. Wu, "Asymmetric Bagging and Random Space for Support Vector Machines-Based Relevance Feedback in Image Retrieval," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.28, n07, July 2006.

[6] T. Qin, X.D. Zhang, T.Y. Liu, D.S. Wang, W.Y. Ma, H.J. Zhang. "An active feedback framework for image retrieval," *Pattern Recognition Letters*, *Volume 29, Issue 5, Pages 637-6461*, April 2008

[7] P. Hong, Q. Tian, T.S. Huang. "Incorporate support vector machines to content-based image retrieval with relevance feedback," *Image processing, 2000. Proceedings. 2000 International Conference on Volume 3, Page: 750-753 vol.3*, 10-12 Sept. 2000.