AN ENHANCED RELEVANCE FEEDBACK METHOD FOR IMAGE RETRIEVAL

LIM PEI GEOK

A project report submitted in partial fulfillment of the
requirements for the award of the degree of
Master of Science (Computer Science)

Faculty of Computer Science and Information System

Universiti Teknologi Malaysia

OCTOBER 2008

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. Images are the most widely used media type other than text to retrieve hidden information within data and it is used as a base for representing and retrieving videos, flash and other multimedia information. An efficient image retrieval tool needs to be developed to select the appropriate images from a digital images database in response to user queries. A content based image retrieval (CBIR) system has been proposed as an efficient image retrieval tool which the user can provide their query to the system to allow it to retrieve the user's desired image from the image database. However, there are several problems have been identified by previous researches such as semantic gap between high level query to low level features and human subjectivity. Therefore, relevance feedback mechanism has been introduced to integrate with CBIR system which intends to solve the problem of CBIR and indirectly increase the CBIR performance. Unfortunately, the traditional relevance feedbacks have some limitations that will decrease the performance of CBIR. In this study, the imbalance training set issue has been highlighted. Imbalance training set is an issue that the negative samples are overwhelming the positive samples during the relevance feedback process. As a result, insufficient training occurs and further degrades the performance of CBIR. To solve the problem, a representative image selection and user weight ranking methods have been introduced. Besides that, Support Vector Machine (SVM) has been proposed as a technique to aid the CBIR learning process. Through the learning process, the system will be able to adapt to different circumstances and situations. Finally, the experiment results reveal that the proposed method is better than traditional relevance feedback method which success improves the performance of CBIR.

ABSTRAK

Perkembangan teknologi komputer dan rangkaian web yang semakin giant telah meningkatkan penampungan informasi multi media. Selain daripada teks, gambar merupakan salah satu media yang sentiasa digunakan untuk memberi maklumat yang bermakna. Dengan itu, sistem *image retrieval* yang cekap perlu dikembangkan untuk memperolehi gambar-gambar yang diingini oleh pengguna daripada pangkalan data. Maka, *content-based image retrieval system* (CBIR) telah dicadangkan sebagai satu peralatan *image retrieval* iaitu pengguna diminta memberi gambar kepada sistem supaya sistem berupaya untuk memberi gambar-gambar yang diingini oleh pengguna. Akan tetapi, banyak masalah telah muncul dalam sistem ini dan ditentusahkan oleh para penyelidik. Oleh sebab itu, *relevance feedback* telah dicadangkan untuk digunakan dalam CBIR. Cadangan ini dianggap dapat menyelesaikan masalah-masalah yang telah ditentukan dan meningkatkan prestasi CBIR. Bagaimanupun, tradisional *relevance feedback* teknik menghadapi beberapa kesukaran dan menurun prestasi CBIR secara tidak langsung. Masalah ketidakseimbangan data latihan telah dititik beratkan dalam kajian ini. Masalah ketidakseimbangan data latihan muncul jika sampel negatif melebihi sampel positif dalam process *relevance feedback*. Oleh demikian, prestasi CBIR akan diturunkan sebab ketidakcukupan data latihan. Dalam projek ini, kaedah pemilihan gambar bermakna dan sistem penyusunan pemberat pengguna telah dicadangkan. Pada masa yang sama, *Support Vector Machine* (SVM) digunakan untuk membantu process pembelajaran sistem di mana proses ini membolehkan sistem menyesuaikan diri dalam situasi berlainan. Akhirnya, kajian ini menunjukkan kaedah yang dicadangkan telah mencapai keputusan yang lebih baik jika dibanding dengan kaedah tradisional *relevance feedback*.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Introduction

The rapid growth of the computer technologies and the advent of the World-Wide Web have increased the amount and complexity of multimedia information. This multimedia information includes digital images, video, audio, graphics and text data. Among the various types of media information, images are the most widely used other than text and it is used as a base for representing and retrieving videos, flash and other multimedia information (Rui et. al, 1998). Thus, images are considered as one of the prime media type to retrieve hidden information within data. At the same time, image database has been widely used in a wide range of application areas, such as advertising, medicine, security, and entertainment (Amato and Lecce, 2008). Therefore, an efficient image retrieval tool needs to be developed to select the appropriate images from a digital images database in response to user queries.

In general, an image retrieval system is a computer system for browsing, searching and retrieving images from a large database of digital images. Most traditional and common methods of image retrieval are utilized by adding metadata such as captioning, keywords, or descriptions to the images so that retrieval can be performed over the annotation words (Long et. al, 2003). Obviously, annotating images manually is a time-consuming, laborious and expensive task for large image databases, and is often subjective, context-sensitive and incomplete (Long et. al, 2003). As a result, there has been a large amount of research done on automatic image annotation to address this issue. Additionally, the increase in social web applications and the semantic web has inspired the development of several web-based image annotation tools.

Another method of image retrieval is content-based image retrieval (CBIR), which aims at avoiding the use of textual descriptions and instead retrieves images based on their visual similarity to a user-supplied query image or user-specified image features. CBIR uses visual contents to search images from large scale image databases according to users' interests. The visual contents of an image such as colour, shape, texture, and spatial layout have been used to represent and index the image. Recent retrieval systems have incorporated users' relevance feedback to modify the retrieval process in order to generate perceptually and semantically more meaningful retrieval results. Figure 1.1 shows the fundamentals of Content-Based Image Retrieval.

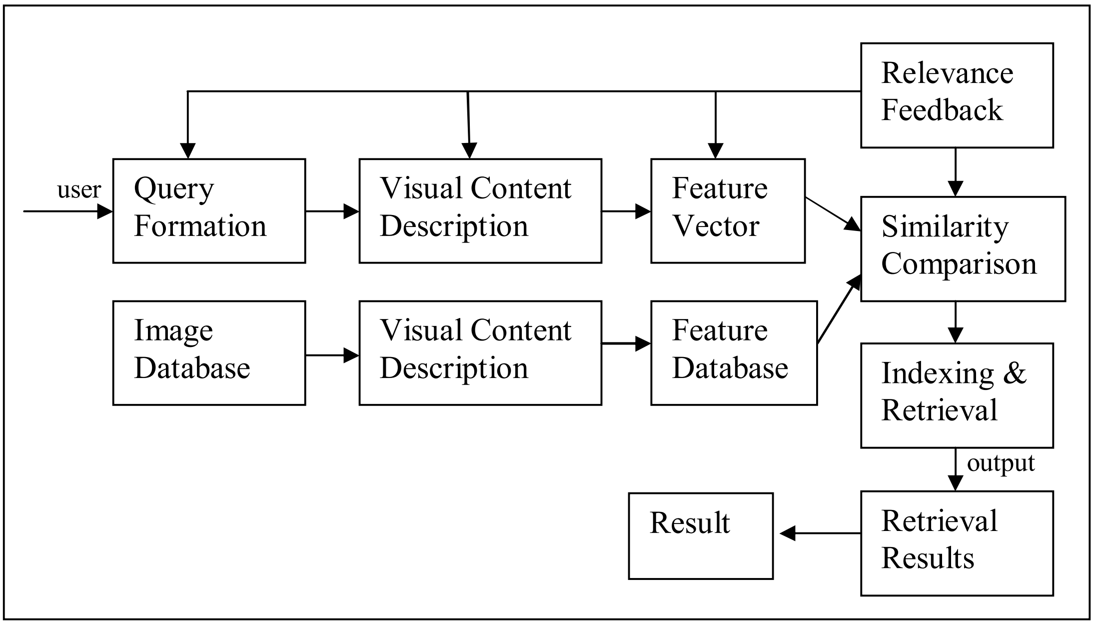

Figure 1.1: Diagram for content-based image retrieval system. (Long et. al, 2003)

As we mentioned above, image is the main media type used in the CBIR system. An image can be thought as a sequence of objects disposed on the background of the image. In general, image analysis is a process of extracting information from image to determine which part is relevant. Hence, the visual content of images will be analyzed to understand the meaning of its content and the relationship among these objects. In CBIR, every image carries a huge amount of information but only a small part of it is relevant for user target. The proverb says "One picture is worth more than thousand words". However, we do not know which part of the image told the user concerned information. Therefore, the goal of image analysis is to identify the user interest and relevant information within the information gathered and abandoned the irrelevant information to perform the suitable action.

Human perception of image similarity is subjective, semantic, and task-dependent. Although content-based methods provide promising directions for image retrieval, generally, the retrieval results based on the similarities of pure visual features are not necessarily perceptually and semantically meaningful. In addition, each type of visual feature tends to capture only one aspect of image property and it is usually hard for a user to specify clearly how different aspects are combined. To address these problems, interactive relevance feedback, a technique in traditional text-based information retrieval systems, was introduced. With relevance feedback, it is possible to establish the link between high-level concepts and low-level features (Rui et. al, 1998). Relevance feedback is a supervised active learning technique used to improve the effectiveness of information systems (Rui et. al, 1998). The main idea is to use positive and negative examples from the user 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 retrieval results based on the feedback and present a new list of images to the user. Hence, the key issue in relevance feedback is how to incorporate positive and negative examples to refine the query and/or to adjust the similarity measure (Long et. al, 2003).

1.2 Problem Background

Relevance feedback which is an interactive mechanism that involves a human as part of the retrieval process (Rui et. al, 1998). This technique involves the human and computer interaction to refine high level queries to low level features representations (Rui et. al, 1998). Human perception in relevance feedback process is considered as very subjective aspect which means that different persons or same person under different circumstances may perceive the same visual content differently (Rui et. al, 1998). As example, one person may be more interested in an image's color feature while another person may be more interested in texture feature on the same image. Even if both people are interested in texture, the way how they perceive the similarity of texture may be quite different (Rui et. al, 1998).

Hence, relevance feedback was introduced to reduce the semantic gap between low-level features and high level semantics and the subjectivity of human perception problems in CBIR (Tao et. al, 2006). Relevance feedback can iteratively refine the retrieval result by learning from user-labeled images until user obtain the target images. This mechanism has been shown as a powerful tool to improve the retrieval performance of CBIR systems (Tao et. al, 2006). The similarity between the images perceive by the humans does not necessarily correlated with the similarity between them in the features space. However, the objects existed in the images and how they correlated is more important to be used as a key point to calculate the similarity between the images (Kim et. al, 2007). The retrieved target image does not need to be exactly the same as the users provided query image but it should contain the similarity contents. As a result, relevance feedback capable to involve human to identify the user interest objects within the image and return a set of user target images to user.

Besides, how to incorporate positive (relevant) and negative (irrelevant) examples to refine the query and how to adjust the similarity measure according to the feedback are the main issues in the relevance feedback. Consequently, there are many early relevance feedback methods have been developed in recent years which includes query refinement (Rui and Huang, 1999) and re-weighting (Rui et. al, 1998).

These two approaches are combined to minimize the total distance between the positive examples and the refined query point with a refined similarity metric (Qin et. al, 2008). A real-time classification problem has appeared due to the relevance feedback in CBIR becomes an online learning problem regarding the positive samples (relevant images) and the negative samples (irrelevant images) as two difference groups (Tao et. al, 2006). Besides, it is difficult to retrieve the positive images which may distribute dispersively in the feature space, directly based on low-level feature similarity either refined or not (Qin et. al, 2008). Hence, a classifier or statistical learning technologies is needed to identify these groups which are positive and negative examples from each other in the feature space (Hong et. al, 2000).

Recently, there are many classification techniques are introduced to attack relevance feedback tasks such as Bayesian inference, Boosting, Support Vector Machines (SVM) and many other statistical learning technologies (Qin et. al, 2008). Among these classifiers, SVM-based techniques are considered as the most promising techniques. A SVM classifier can be learned from training data which consists of relevant and irrelevant images marked by users (Zhang et. al, 2001). Using the SVM-based classifier, the system can retrieve more images relevant to the query in the database efficiently. The SVM approach is considered a good candidate because of its high generalization performance without the need to add a priori knowledge, even when the dimension of the input space is very high (Zhang et. al, 2001). Besides, SVM has a good performance for pattern classification problems by minimizing the Vapnik-Chervonenkis dimensions and achieving a minimal structural risk (Tao et. al, 2006) (Hong et. al, 2000).

According to previous studies that have been done on SVM based relevance feedback, SVM normally treats the problem as a strict binary classification problem without notices an important issue of relevance feedback. For example, the imbalanced training set problem in which the negative instances significantly overwhelming the positive ones (Hoi et. al, 2004). This problem will cause the performance of SVM become poor and degrade the performance of CBIR system.

Even though relevance feedback intensively helps to improve the performance of the CBIR, most of the users do not like to label too many images during the feedback process (Qin et. al, 2008). As a result, the information that gained from the user is very limited. In addition, the training samples (user's labelled images) are insufficient due to the users would only label a few images and cannot label each feedback sample accurately all the time (Qin et. al, 2008) (Tao et. al, 2006). Consequently, SVM classifier will be unstable when the size of training samples is small (Tao et. al, 2006) (Kim et. al, 2007). Hence, the performance of SVM-based relevance feedback becomes poor when the number of labelled positive feedback samples is small (Tao et. al, 2006) (Kim et. al, 2007) (Qin et. al, 2008). As a result, the accuracy of SVM relevance feedback based CBIR system to retrieves target images will be decreased. Therefore, a new modified relevance feedback mechanisms need to be developing in order to increase the performance of CBIR system.

1.3 Problem Statement

Content-based image retrieval is a technique to retrieve images semantically relevant to the user's query image from an image database. As mentioned above, CBIR have two main problems which are semantic gap between high level concepts and low level features and the subjectivity of human perception. Regarding to the problems, relevance feedback has been introduced to solve the problems and improve the performance of CBIR system. Moreover, relevance feedback (RF) is a way for bridging this gap and scaling the performance in CBIR systems. There are appearing some of the research questions which are stated as below:

What kind of relevance feedback methods can be used to solve the semantic gap between low level feature and high level query and subjectivity of human perception in CBIR? How the performance of the relevance feedback based CBIR?

However, relevance feedbacks that highly rely on the human perception subjectivity are an issue in CBIR. As we know, different users under different situations will have different interest in the image even though the query image provided by the users is same. Therefore, the relevance feedback for CBIR should intend to capture the user's perception of an image's content and retrieve the images close to user preference. There are many studies and researches have showed that relevance feedback help the system to refine the feature's weight according to the user preference. Thus, system can reformulate the user query according to the relevant and irrelevant images marked by user (Cheng et. al, 2008). Although the relevance feedback can significantly improve the CBIR performance, user do not like to label too many images as feedback to system (Qin et. al, 2008). As conclusion, system will gain very limited information during the feedback process and the retrieval accuracy will degrade due to the limited feedback from user and the subjectivity of the human perception. There are few questions have been stated as below:

How the problems which appear in the relevance feedback based CBIR will degrade the performance of CBIR system? How to overcome these problems to improve the accuracy of image retrieval system?

Additionally, incorporate positive and negative examples to refine the query and similarity measure in relevance feedback can be key issues of CBIR. Recently, support vector machines (SVM) based relevance feedback is widely employed to solve the problems of the relevance feedback in CBIR. The SVM-based techniques are considered as the most promising techniques due to SVM classifier capable to learn from training data which is consists relevant and irrelevant images marked by users. Therefore, the used of SVM-based model is to find more relevant images from images database.

However, the imbalanced training set problem which the negative samples are overwhelming positive samples have degraded the performance of CBIR. Moreover, the performance of the SVM based relevance feedback become poor when the number of positive feedback is small (Kim et. al, 2007). This is primarily due to two main reasons. First, the SVM classifier is unstable when the size of training set is small. Second, there are usually many more negative feedback samples than the positive ones in relevance feedback process. As a result, there are some other questions have been explored which are stated as below:

How to solve the imbalance training set issue that exists in Relevance Feedback based CBIR which degrade the performance of CBIR system? What is the improvement of CBIR if the imbalance training set issue has been solved?

As conclusion, a modified design of relevance feedback based on CBIR is desired to solve the above mentioned problems.

1.4 Project Aim

This project aims to investigate the performance of different relevance feedback method in CBIR. The key issues of incorporate positive and negative examples to adjust the similarity measure according to the feedback will explore in this study. Besides that, the problems which appear in relevance feedback based CBIR which limit the performance of CBIR system will be determined. These problems including imbalance training set problem, classification problem, limited information from user, insufficient training set problem and weight adjustment issues. Among these problem, the imbalance training set problem which occurs when the negative samples is overwhelming the positive ones can significantly degrade the performance of CBIR system. The degrading of CBIR performance is mainly due to the over-representation of negative class that cause the performance of classifier becomes poor. This project tends to propose a modified CBIR model by using support vector machine-based relevance feedback in image retrieval. The performance of proposed method will be measure by using the standard information retrieval which are precision, recall and F1.

1.5 Objective

In order to accomplish the hypothesis of the study, few objectives have been identified as stated below.

1. To investigate the usage of CBIR by studying different kind of relevance feedback methods.
2. To improve the performance of relevance feedback based CBIR by proposing an enhanced relevance feedback method to solve the imbalance training set issue.
3. To compare the performance of proposed method with traditional relevance feedback methods by using standard information retrieval measurement which are precision, recall and F1.

1.6 Project Scope

The main focus of this study is on the problem and performance for the support vector machine based relevance feedback in image retrieval. The scopes for this project as follows:

1. The feature extraction part will focus on color and texture features extraction.
2. Support Vector Machine (SVM) will be used as classifier in this project.
3. This study will only focus on images as media type for the CBIR system.
4. The experiment will consist of five categories of data which are animal, building, flower, fruits and natural scene images.
5. Two traditional relevance feedback methods which are Relevance Feedback based CBIR by Cheng et al. (Cheng et. al, 2008) and Relevance Feedback based CBIR by Qin et al. (Qin et. al, 2008) will be used as performance comparison in this project.
6. The performance of this study will measure by standard information retrieval approach which are precision, recall and F1.

REFERENCES

A. Amato and V. D. Lecce (2008). A knowledge based approach for a fast image retrieval system. *Image and Vision Computing*, *Elsevier*.

B. Manjunath, P. Wu, S. Newsam, and H. Shin (2001). A texture descriptor for browsing and similarity retrieval. *Signal Processing Image Communication*.

C. H. Hoi, C. H. Chan, K. Z. Huang, M. R. Lyu, I. King (2004). Biased Support Vector Machine for Relevance Feedback In Image Retrieval. *In Proceedings of Intl. Joint Conf. on Neural Networks (IJCNN'04), Budapest, Hungary*.

D.C. Tao, X.O. Tang, X.L. Li, X.D. Wu (July 2006). 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.

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

G. Das and S. Ray (2007). A comparison of relevance feedback strategies in CBIR. *IEEE.*

F. Long, H.J Zhang and D. D. Feng (2003). Multimedia Information Retrieval and Management: Technological Fundamentals and Applications, chapter Fundamentals of Content-Based Image Retrieval. *Springer-Verlag, Berlin*.

H. Greenspan, G. Dvir, Y. Rubner (2000). Region Correspondence for Image Matching via EMD Flow. *IEEE*.

J. Ohm, F. Bunjamin, W. Liebsch, B. Makai, K. Müller, A. Smolic, D. Zier (2000). A Set of Visual Feature Descriptors and their Combination in a Low-Level Description Scheme. *Signal Processing: Image Communication 16,* 157-179.

J.R. Smith and S.F. Chang (7-10 May 1996). Automated binary texture feature sets for image retrieval. *Proc. ICASSP-96.*

K. Blekas, A. Likas, N. Galatsanos and I. Lagaris (2005). A Spatially-Constrained Mixture Model for Image Segmentation. *IEEE Transactions on Neural Networks,* vol. 16(2), 494-498.

L. Zhang, F. Lin, and B. Zhang (Oct 2001). Support Vector Machine Learning for Image Retrieval. *Proc. IEEE Int'l Conf. Image Processing,* vol. 2, 721-724.

M. Crucianu, M. Ferecatu, N. Boujemaa (October 10, 2004). Relevance feedback for image retrieval: a short survey. *Report of the DELOS2 European Network of Excellence (6th Framework Programme).*

M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafine, D. Lee, D. Petkovic, D. Steele, and P. Yanker (1995). Query by image and video content: The QBIC system. *IEEE Computer*.

P.C. Cheng, B.C. Chien, H.R. Ke, W.P. Yang (April 2008). A two-level relevance feedback mechanism for image retrieval. *Expert Systems with Applications*, Vol. 34, Issue 3, 2193-2200.

P.G. Foschi (October 2002). Feature extraction for image mining. *In Workshop on Multimedia Information Systems (MIS 2002), Tempe, Arizona*.

P.Y. Hong, Q. Tian, T.S. Huang 10-12 Sept. 2000. Incorporate support vector machines to content-based image retrieval with relevance feedback. *Image processing, 2000. Proceedings. 2000 International Conference,* Vol. 3, 750-753.

S.D. Cheng, X. Lan and H.L. Yan (October 2007). Image retrieval using both color and texture features. *The jounal of China Universities of posts and telecommunication, Supplement,* Vol. 14.

S.H Zhou, Y.V.Venkatesh, C. C. Ko (2000). Texture retrieval using tree-structured wavelet transform.

S. Liang and Z.X. Sun (2008). Sketch retrieval and relevance feedback with biased SVM classification. *Pattern Recognition Letters,* Vol. 29, 1733–1741.

S. Tong and E. Chang (October 2001). Support Vector Machine Active Learning for Image Retrieval. *Proc. of ACM Int. Conf. on Multimedia,* 107–118.

T. Qin, X.D. Zhang, T.Y. Liu, D.S. Wang, W.Y. Ma, H.J. Zhang (April 2008). An active feedback framework for image retrieval. *Pattern Recognition Letters*, Vol. 29, Issue 5, 637-6461.

W.Y. Ma and H. J. Zhang (1998). Benchmarking of image features for content-based retrieval. *IEEE.*

X.J. Qi and R. Chang (2007). Image retrieval using transaction-based and SVM-based learning in relevance feedback session. *M.Kamel and A.Campilho (Eds): ICIAR 2007, LNCS 4633,* 638-649 @ *Springer-Verge Berlin Heidelberg.*

Y. Chen, X. Zhou, and T. S. Huang (2001), One-Class SVM for Learning In Image Retrieval. *Proc. IEEE Int'l Conf. Image Processing,* 815-818.

Y. Liu, D. S. Zhang and G. Lu (August 2008). Region-Based Image Retrieval with High-Level Semantics using Decision Tree Learning. *Pattern Recognition,* 41(8), 2554-2570.

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

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