UNIVERSITY OF SCIENCE AND TECHNOLOGY

COLLEGE OF GRADUATE STUDIES AND ACADEMIC ADVANCEMENT

Faculty of Computer Science and Information Technology

Relevance Feedback Method for Content Based Image Retrieval

By

Mohamed Ahmed Mohamed Abdelkreem

A Thesis

Submitted to the College of Graduate Studies and Academic Advancements

in Partial Fulfillment of the Requirement for the Degree of Master of Science in

Information Technology

Supervisor

Dr. Ali Ahmed Alfaki abdalla

Abstract

Content-based image retrieval (CBIR) systems allow the user to retrieve images similar in visual content to a given query from an image database. Image features are color, texture and shape used to search relevant images. The relevance feedback retrieve the most relevant images and provide a weight of preference for each relevant image and it calculated by the features similarity. Relevance feedback is a powerful query modification technique in the field of content-based image retrieval, which considerably improves the quality of the term-weighting retrieval system. Feedback refers to an interactive process that helps to improve the retrieval efficiency. This study applies the concept of relevance feedback. The relevance feedback method achieves better performance and accuracy in retrieving images.
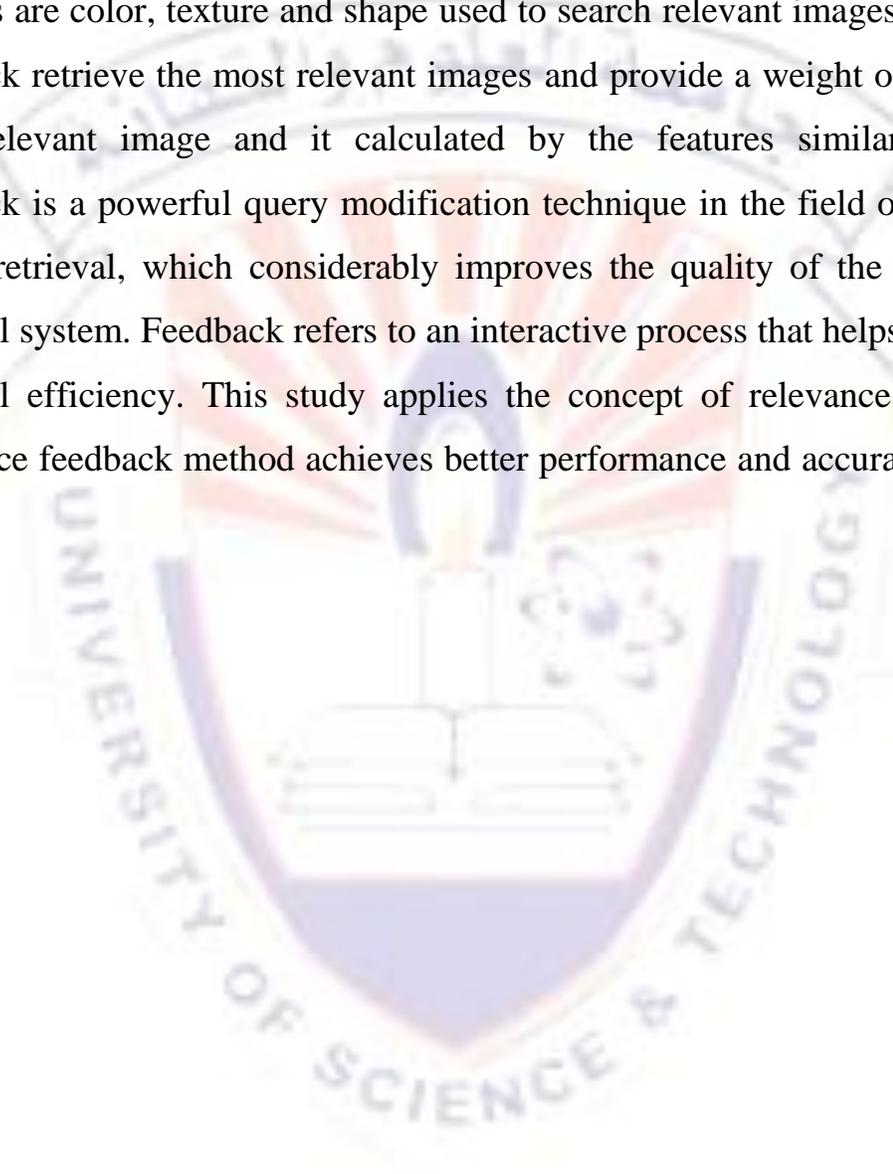
المستخلص

استرجاع الصور عبر محتوى الصورة (CBIR) هي أنظمة تسمح للمستخدم باسترجاع الصور الأكثر تشابه في المحتوى لاسترجاع الصورة المستعلم عنها من قاعدة بيانات الصور. الخواص الصورة تتكون من اللون والشكل وتستخدم هذه الخواص للبحث عن الصور ذات الصلة أو الصورة التي تحمل خواصا مشابهه. ال (Relevant feedback) هي استرجاع الصور الأكثر صلة بالصورة المستعلم عنها وتوفير الوزن لكل صورة ذات صلة، وكل ذلك يتم عبر تشابه الخواص داخل الصورة. ال (Relevant feedback) هي تقنية ذات أهمية قوية في مجال استرجاع الصور القائم على المحتوى ، حيث تساهم بشكل كبير على تحسين نظام الاسترجاع و تساعد على تحسين كفاءة الاسترجاع، ولتحقق أفضل أداء ودقة في استرجاع الصور.

Introduction

Data retrieval systems of the modern technology used in many areas including image retrieval system. The image retrieval system is an important and necessary because it increase the size of image databases [1], and it needs in all fields including commerce, government, academics, hospitals, crime prevention, surveillance, engineering, architecture, journalism, fashion and graphic design, and historical research [2] . Content-based image retrieval (CBIR) systems retrieve images from a database according to the visual contents of image such as color, shape, texture, and spatial layout to represent the image [3] .

1.1.1 Text-Based Image Retrieval

The TBIR is a technique used the text to store information about images. This technique has disadvantages that because it use a single feature to retrieve images and taking long time on the retrieval. Image may contain many features, including color, texture, shape, etc. [6] .These features are extracted from images and convert them into digital ratios and saved to making access to image, each image features describe what's inside the image.

1.1.2 Content-Based Image Retrieval

Content-based image retrieval (CBIR), also known as query by image content is the application of computer vision techniques to image retrieval problem, that is, problem of searching for digital images in large databases. It aims to finding images of interest from a large image database using the visual content of the images. "Content-based" means that the search will analyze the actual contents of the image rather than the metadata such as keywords, tags, and/or descriptions associated with the image. The term 'content' in this context might refer to colors, shapes, textures, or any other information that can be derived from the image itself.
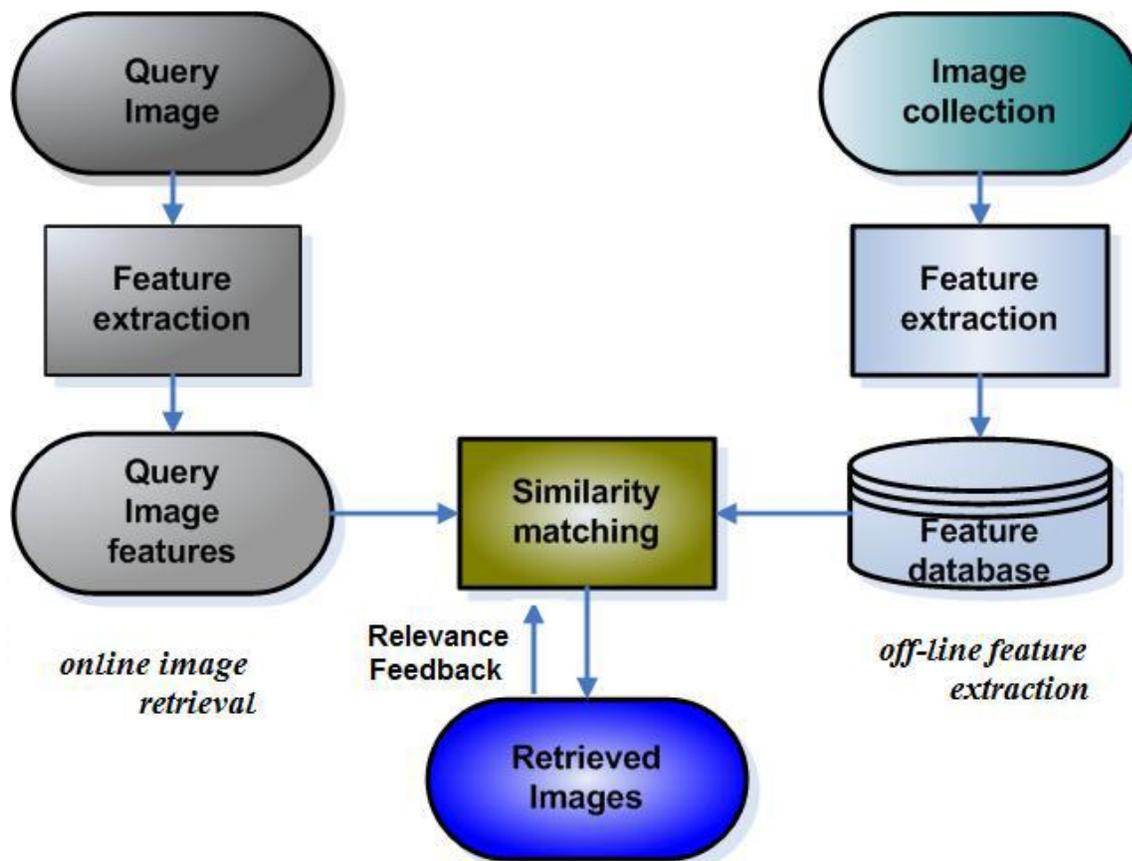
Figure (1.1): CBIR System [14] .

1.2 Applications of CBIR Systems

The public area of the image processing like (CBIR) it's a diverse applications such as Internet, multimedia, medical image archives, and crime prevention [4] .The CBIR its more advanced techniques used in image retrieval because its use more than one of image features, such as text, color and shape. Previously was technique use the text as information for describes the content of image [5] , this technique called TBIR 'Text Based Image Retrieval'.

1.2.1 Medical Applications

The use of CBIR can result in powerful services that can benefit biomedical information systems. Three large domains can instantly take advantage of CBIR techniques: teaching, research, and diagnostics . Clinicians usually use similar cases for case-based reasoning in their clinical decision-making process. In the medical field, some ailments require the medical

practitioner to search and review similar X-rays or scanned images of a patient before proffering a solution.

1.2.2 Digital Libraries

There are several digital libraries that support services based on image content. One example is the digital museum of butterflies, aimed at building a digital collection of Taiwanese butterflies. This digital library includes a module responsible for content-based image retrieval based on color, texture, and patterns.

1.2.3 Crime Prevention

One of the main jobs of police is to identify and arrest criminals in the country. However, to achieve that, the department of security investigation must identify the identity of criminals as fast as possible and with a high accuracy rate. Day after day, the crime rate is increasing so that the police must deal with a large number of criminals images that stored in a database. Once a new image is arrived, it must be compared with all of these images to classify it correctly. It is clear that, doing this job manually takes a long time so, the need for criminal recognition system is strongly highlighted here.

1.2.4 Web Searching

The most important application, however, is the Web, as a big fraction of it is devoted to images, and searching for a specific image is indeed a daunting task. Numerous commercial and experimental CBIR systems are now available, and many web search engines are now equipped with CBIR facilities, as for example Alta Vista, Yahoo and Google. Today it is estimated that there are 30 billion images in Image shack, Facebook holds 35 billion photos and Corp's Photo Bucket has 10 billion photos .

1.2.5 Other Applications

In the commerce department, before trademark is finally approved for use, there is a need to find out if such or similar ones ever existed. In architectural and engineering design, image database exists for design projects, finished projects, and machine parts. In publishing and advertising, journalists create image databases for various events and activities such as sports, buildings, personalities, national and international events, and product advertisements. In historical

research, image databases are created for archives in areas that include arts, sociology, and medicine .

1.3 Research Background

Relevance feedback is a supervised learning technique refines the distance between the query and the database images. Distances calculated by similarity of the high level queries represented by low level features. Relevance Feedback developed to speed up the search process, the user submits to the relevance feedback system a query image that is an example of the image of interest, begin from the query, the relevance feedback system assigns a score to the images in the database the score start related to a similarity measure between the images and the query, a number of best scored images are returned to the user that show them as relevant or not.

1.4 Research Problem

Content-based image retrieval many algorithms developed to retrieve the required image from the image database, development of an efficient and effective algorithm has been proposed for image retrieval. This research aim is to apply relevance feedback method as an effective solution to improve performance of content-based image retrieval.

1.5 Research Objectives

The main objective of this research is to improve the effectiveness of an image retrieval system, including the following objectives:

 - ➢ To extract image features

 - ➢ To retrieve top similar images from each class and use top results to modify the new queries
 - ➢ To apply relevance feedback based on query modification mentioned above.

1.6 Thesis Organization

This research consists of five chapters, the first chapter gives about the introduction to content-based image retrieval, second chapter reviews the literature review and previous work in the (CBIR) field, the third chapter there is an overview of the research methodology, the fourth chapter introduces the results and discussions, finally in the fifth chapter include the conclusion and future work.