CONTENT BASED RETRIEVAL USING COLOUR AND TEXTURE OF WAVELET BASED COMPRESSED IMAGES

by

IRFAN AFIF ABDUL FATAH

Thesis submitted in fulfillment of the requirements
for the Degree of
Master of Science

MARCH 2008

CONTENT BASED RETRIEVAL USING COLOUR AND TEXTURE OF WAVELET BASED COMPRESSED IMAGES

IRFAN AFIF ABDUL FATAH

UNIVERSITI SAINS MALAYSIA

2008

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisor, Dr. Putra Sumari, whose comments and numerous suggestions have helped me in developing and formulating the ideas in this study. His encouragement and generous attention have provided me with the confidence I needed to undertake this task.

I would also like to thank my co-supervisor, Dr. Abdullah Embong for his help and invaluable advice.

A very special thank you to all my family members especially my beloved parents, Dr. Abdul Fatah Che Hamat and Pn. Nor Eynizan Hassan for their constant support and love during the period of my postgraduate study.

Last but not least, I would also like to thank my colleagues from the Multimedia Research Group. Their help and useful comments has helped made this thesis possible. Not forgetting all my friends whom have provided me help and support during the time of this thesis writing, my gratitude goes out to all of you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PUBLICATIONS & SEMINARS

DAPATAN SEMULA BERDASARKAN KANDUNGAN MENGGUNAKAN WARNA DAN TEKSTUR BAGI IMEJ-IMEJ YANG DIPADATKAN BERDASARKAN WAVELET

ABSTRAK

Permintaan yang tinggi terhadap penggunaan dapatan semula imej telah menggalakkan pembangun aplikasi multimedia untuk mencari cara untuk mengurus dan mencari imej dengan lebih efisien. Dapatan semula imej berdasarkan kandungan seperti warna dan tekstur merupakan suatu cabaran. Warna dan tekstur merupakan maklumat abstrak yang terkandung didalam sesebuah imej. Penyampaian maklumat tersebut secara baik adalah amat penting untuk mencapai hasil dapatan yang baik. Teknologi pemadatan berasaskan Wavelet merupakan suatu kaedah baru untuk memadat dan menyimpan data bergambar yang banyak didalam ruang storan yang terhad. Wavelet mempunyai kuasa pemadatan yang lebih baik berbanding teknik-teknik pemadatan terdahulu.

Didalam tesis ini, kami mencadangkan suatu enjin untuk mendapatkan imej-imej yang dipadatkan berasaskan Wavelet berdasarkan kandungan warna dan tekstur imej berkenaan. Maklumat warna disampaikan melalui penggunaan histogram dengan bilangan bin yang rendah. Sebuah kaedah pemetakan saiz-tetap (Fixed-size partitioning) berserta pemberat telah dicadangkan. Kaedah ini membahagikan imej kepada beberapa kawasan untuk digunakan semasa carian imej. Maklumat tekstur disampaikan melalui tenaga tekstur. Dapatan semula warna dilakukan dengan menggunakan jarak kuadratik (Quadratic Distance) manakala bagi tekstur, jarak Euclidean (Euclidean Distance) digunakan. Sebuah

prototaip sistem yang dinamakan CERWACI telah dibangunkan untuk menguji kaedah yang dicadangkan. Pengujian dilakukan menggunakan kaedah ujian kejituan dan ingat kembali (Precision and Recall) dan ujian masa dapatan semula keatas set imej ujikaji yang pelbagai.

Kami mendapati 16 bin merupakan bilangan bin terbaik untuk mempersembahkan maklumat warna. Dapatan semula warna menggunakan pemetakan saiz-tetap berserta pemberat memberi nilai kejituan dan ingat kembali yang lebih tinggi berbanding tanpa pemberat. Faktor pemberat telah berjaya mengatasi masalah nilai jarak yang tidak konsisten. Bagi dapatan semula tekstur, kami telah mengurai (decompose) imej dari 1 hingga 7 peringkat. Kami mendapati 5 peringkat penguraian menghasilkan nilai kejituan dan ingat kembali yang lebih tinggi berbanding peringkat-peringkat lain. Turutan dapatan semula tekstur turut dikaji. Dapatan semula tekstur yang dilakukan keatas hasil dari perbandingan warna menghasilkan nilai kejituan dan ingat semula yang lebih tinggi berbanding dapatan semula keatas semua imej didalam pangkalan data.

CONTENT BASED RETRIEVAL USING COLOUR AND TEXTURE OF WAVELET BASED COMPRESSED IMAGES

ABSTRACT

The growing demands for image retrieval in multimedia field such as crime prevention, health informatics and biometrics has pushed application developers to search ways to manage and retrieve images more efficiently. Retrieving images based on content such as colour and texture is still a challenging issue. Colour and texture are the abstract information embedded in an image. Representing these information properly is crucial in order to achieve better retrieval results. Furthermore the recent wavelet based image compression technology has been seen as a new way to store millions of pictorial data within the limited space of the hardware capabilities. Wavelets have been proven to be superior in terms of compression compared to previous compression methods.

In this thesis, we propose an engine to retrieve wavelet based compressed images based on its colour and texture features. Colour information is represented using colour histogram with low bin count. A fixed-size image partitioning method with weight is proposed. This method divides the image into several blocks to be used during image retrieval. We represent texture by using texture energies. Colour and texture retrieval was done using Quadratic Distance and Euclidean Distance respectively. A prototype system called CERWACI which uses the proposed retrieval engine has been developed to test this approach. We evaluate the effectiveness by applying Precision and Recall tests and retrieval time on various sets of test images.

We found out that 16 bins was the best bin count to represent colour information. Colour matching using fixed-size partition with weight provided better precision and recall scores compared to fixed-size partition without weight. The weight factor managed to correct the inconsistent distance value faced by the fixed-size partition without weight. For the texture feature, we tested decomposing the image using 1 to 7 levels of decomposition. We found out that 5 levels of decomposition provided better precision and recall compared to other levels. The order of texture retrieval was also investigated. Texture retrieval performed on results from colour matching provided better precision and recall scores compared to texture retrieval performed on the whole compressed image in the database.

CHAPTER 1
INTRODUCTION

1.0 Introduction

Since the last few years, systems working with retrieving large amounts of multimedia data have been growing rapidly. Systems such as search engines, e-business systems, online tutoring system, GIS, and image archive are among few to name [38]. These systems involve retrieving multimedia data based on pictorial content. In the image archive for example, a simple query such as searching for bird with yellow feathers requires the system to be able to find all images in the database which contains a bird with yellow feathers. This is a challenging task since it requires the system to browse every single image in the database and compare it to the query image. Manual browsing the database to search for identical images would be impractical since it takes a lot of time and requires human intervention. A more practical way is to use Content based image retrieval (CBIR) technology. CBIR has provided an automated way to retrieve images based on the content or features of the images itself. The CBIR system simply extracts the content of the query image matches them to contents of the search image.

Finding similar images is indeed a challenging task since thousand of images are involved. System such as Medical imaging, Remote sensing & GIS, Cultural heritage, Painting and Arts, Image archive database and Surveillance systems rely on massive images stored in their database. These images are usually stored in compressed form to preserve storage space. Automated searching of similar images based on its content is a challenging task since it deals with pictorial content and the result are not always accurate. During retrieval, normally images that are not similar to the query are also retrieved and presented to the users. This is because pictorial content such as colour is abstract type and contributes to false retrieval. Efficient retrieval algorithm is crucial to retrieve images in such systems.

1.1 Content Based Image Retrieval

Content Based Image Retrieval or CBIR is defined as a process to find similar picture or pictures in the image database when a query image is given. Given a picture of an apple, the system should be able to present all similar images of an apple in the database to the users. This is done by extracting the features of the images such as colour, texture and shape. These image features is used to compare between the query image and images in the database. A similarity algorithm is used to calculate the degree of similarity between those two images. Images in the database which has similar images features to the query image (acquiring the highest similarity measure) is then ranked and presented to the user.

CBIR is basically a two step process which are Feature Extraction and Image Matching (also known as feature matching) [30]. Feature Extraction is the process to extract image features to a distinguishable extent. Information extracted from images such as colour, texture and shape are known as feature vectors. The extraction process is done on both query images and images in the database. Image matching involves using the features of both images and comparing them to search for similar features of the images in the database.

Using multiple feature vectors to describe an image during retrieval process increases the accuracy when compared to the retrieval using single feature vector. For example, searching of image based on its colour and texture provides a better result than using a single colour feature since two features are now used as indicator during matching process. In this research we focus on the method and strategies to retrieve images by using both colour and texture feature vectors to produce more accurate results.

1.1.1 Spatial Information Distribution Factor

Spatial information distribution of an image refers to the distribution of colour within the image. Let us take the scenery picture of a forest with blue sky for example. In the picture, basically, the bottom part of the picture contains green trees and the top part of the image contains blue colour representing the sky. If we were to retrieve the images based only on the whole colour distribution of the image without any spatial knowledge, many unrelated images which might share the same colour distribution will be retrieved. The spatial distribution where blue is mostly located at the top and green is mostly located at the bottom of the picture is an important factor when matching the images during retrieval. This is done by separating the image into homogenous regions and using each region as a matching criterion during retrieval. In this thesis, we study the spatial distribution with the aim to increase the result accuracy of image retrieval.

1.1.2 Texture Representation

Texture of an image refers to properties that represent the surface of an object. Texture contains important information about structural arrangement and visual pattern of an objects surface, such as bricks, grasses, and fabric. It also describes the relationship of the surface to the surrounding environment. The challenge when dealing with texture feature in CBIR is the method of representation which is to find a way to describe texture information in a form suitable and easy for computer to interpret. There are many approaches in representing texture. The approaches could be categorized into statistical method, geometric methods, model based methods and signal processing methods. Among these we focus on the signal processing method to represent texture, specifically we choose to use Wavelet Transform. In our work, we use Wavelet Transform to decompose the image. The subband energy of the decomposed image is used as the texture index.

1.2 Wavelet Based Compression

In the recent years, many algorithm and techniques has been exploited to compress images efficiently. Among these algorithms are the Discrete Cosine

Transform, (DCT) Discrete Fourier Transform (DFT), Karhunen-Love Transform (KFT) and Discrete Wavelet Transform (DWT). Wavelet Transform has been proven to provide better compression rates compared to other transforms. For example, JPEG2000 compression which uses wavelet transform provides superior compression performance compared to other compression standards. At lower bit rates, compared to JPEG, images compressed using JPEG2000 has less visible artefacts and almost no blocking effects. Compression artefacts is the result of an aggressive data compression scheme applied to an image that discards some data which is determined by an algorithm to be of lesser importance to the overall content but which nonetheless discernible.

Wavelet Transform (WT) is capable of decomposing the image spectrum into several frequency bands with little correlation between the bands as possible. The decomposition of the image into subbands (low frequency subbands and high frequency subbands) is done by passing the image data through a series of analysis filter banks. Any of the resulting subbands can further be re-inputted to the analysis bank for further down sampling as many stages as desired. The different subbands can be down sampled due to their lower bandwidth as compared to the original image.

The superior compression performance of Wavelets and its advantages over other compression methods has motivated us to solely work on content based retrieval of wavelet based compressed images.

1.3 Problem Statement

The problem involves entering a wavelet-based compressed image as a query into a software application that is designed to employ CBIR techniques in extracting visual properties, and matching them. This is done to retrieve images in the database that are visually similar to the query image. An effective way to retrieve image contents is needed. The new retrieval process must be able to retrieve image based on the image's low-level content without the aid of keywords or textual descriptions.

1.4 Objectives of the Research

This thesis aims to propose a Wavelet-based compressed image retrieval system using low-level features, colour and texture. Among other objectives of this thesis are:

- To define a method to represent spatial colour information of wavelet based compressed images. We aim to capture spatial colour information by investigating partitioning techniques in the image.
- To define an algorithm to represent texture for wavelet based compressed image. We investigate the use of wavelet transform, image decomposition and sub-band technology for the texture representation.
- To define a retrieval engine to retrieve wavelet based compressed images based on colour and texture features
- To develop a prototype system to evaluate the retrieval effectiveness of the proposed method by using various test images.

The scope of this study is limited to the content-based image retrieval domain regarding image feature extraction, image indexing and feature matching between query images and search images using only colour and texture features. This study uses images with dimension of 384 x 256 pixels in landscape perspective.

1.5 Research Contribution

The contribution of the thesis is as follows:

- We propose a fixed-size partitioning method in colour extraction stage. The method partitions the image into 4 sub-regions (2 by 2 blocks) with equal size. This method is designed to represent colour spatial information within an image.
- We also propose the weight factor strategy used in feature matching stage. Weight factor is designed to support the fixed-size partitioning method. Weight factor is applied to solve the inaccurate 4 blocks distance measure of an image in the feature matching stage.
- An engine of content based retrieval of wavelet based compressed image. The engine is designed to use colour and texture features. The engine contains two main components which are the feature extraction module and feature matching module. A Content Based Retrieval of Wavelet based Compressed Image system prototype called CERWACI was built based on the engine. This prototype is built to implement and test the new proposed method. The prototype system built can be implemented into

many areas of multimedia intensive computing such as multimedia systems which utilizes massive wavelet-based compressed images.

1.6 Thesis Overview

In this thesis, we shall discuss in detail about Content Based Image Retrieval, Wavelet Transform and its significance and benefits towards the development of future multimedia applications. Chapter 1 gives the introduction of the research. It outlines the objectives and contributions of the research. It also gives the overview of other chapters in this thesis.

Chapter 2 will review current trends and methods in existing CBIR systems. This will cover about the CBIR in depth, colour & texture definition Wavelet transform, and a brief description of the JPEG2000 architecture.

In Chapter 3, we shall explain on the retrieval engine design and methodology we choose to use in this research. This chapter will explain the methods of representation and matching for both colour and texture feature. In this chapter, the proposed system architecture will be presented. Each module and key functions in the prototype system architecture will be explained in detail. Chapter 3 also presents the retrieval methods used for both colour and texture features. The indexing method and distance measure used will be described in this chapter.

Chapter 4 will discuss the implementation of the CERWACI prototype system. This chapter will focus on detailed implementation of each key component of the prototype system. This will cover the choice of development software and tools used to create the system. The user interface design and test data to be used will also be discussed in this chapter.

Chapter 5 will discuss the evaluation of the prototype system. In this chapter, we will explain the test method used and how to evaluate the prototype system. A series of experiments are carried out to evaluate the effectiveness of this system for use with different images. We will also observe and calculate the performance of the system.

Finally Chapter 6 will discuss about the achievements and limitation of the proposed method. We shall also lookup suggestions on future development and any possible extensions of the prototype system.

CHAPTER 2
RESEARCH BACKGROUND

2.0 Introduction

Before we move on to the design of the Content Based Retrieval of Wavelet Based Compressed Images (CERWACI) system, we would like to study some subject which will provide basic information for the consideration and justification for the construction of the CERWACI system. That information will facilitate the understanding of following chapters for readers.

Chapter 2 will focus mostly on the background of the Content Based Retrieval of Wavelet based Compressed Image research. The most important part of this research is the Content Based Image Retrieval (CBIR) technology and methods used to retrieve colour and texture information. Therefore, in this chapter, we will focus on studying some basic concepts of CBIR. We will discuss recent techniques and methods being applied in existing multimedia application which uses CBIR. We shall also study some examples of existing CBIR system and understand how they work.

First, we will study about content based image retrieval. We shall then study some basic digital image fundamentals. The following sub-chapters will discuss about the methods used to retrieve colour and texture information. We shall also discuss the usage of wavelets as a tool to compress images. We shall first cover some basic concepts of the Wavelet Transform, and then move on to study its various application in image processing, especially in still image compression and texture retrieval. The JPEG2000 standard which is the latest state-of-the-art compression standard using Wavelet Transform, will also be discussed briefly in this chapter.

2.1 Content-Based Image Retrieval (CBIR)

CBIR is the retrieval of images based on visual features such as colour, texture and shape [32]. Before CBIR was widely used to retrieve images, researchers relied heavily on text-based retrieval [2, 34].

Text-based retrieval method was the most widely used method to retrieve images because of its simplicity. It was also easily implemented. Through this method, the user is required to key in the search keyword describing the image desired. The retrieval process is carried out by matching the query keyword with the index kept in the database. However, this method has it's drawbacks as it is not feasible to be implemented in large-scale databases [12, 35]. This is because the system administrator is required to key in keywords describing each image into the database to create indexes. As this process must be done manually, it becomes time consuming and labour intensive. Furthermore, text-based description tends to be incomplete, imprecise and inconsistent in specifying visual information [7]. An example of search engine which uses text-based retrieval is Lycos [9], created by Dr. Michael Mauldin from Carnegie Mellon University in 1994.

2.1.1 Definition of CBIR

CBIR is the application of computer vision to aid the image retrieval process of searching for digital images in large database based on the comparison of low level features of images. The search is carried out by using contents of the image themselves rather than relying on human-inputted metadata such as caption or keyword describing the image. Compared to text-based retrieval systems, CBIR is more feasible in large-scale databases and is usually used in environments which require fast retrieval and real-time operations. Softwares which implements CBIR are known as *content-based image retrieval systems* (CBIRS).

CBIR came to the interest of researchers as it offers the ability to index images based on content of the image itself [34]. CBIR retrieves images based on visual features such as colour, texture and shape [32]. In this method, colour, shape and texture of an image are classified automatically or semi-automatically with the aid of human classifier. Retrieval results are obtained by calculating the similarity between the query and images stored in the database using predefined distance measure. The results are than ranked according to the highest similarity score.

The ideal CBIRS from the viewpoint of the user would be a 'semantic' type of retrieval, where users are able to make query sentences such as "find pictures of birds" or "find all pictures of Bob". However, this type of open-ended task is quite difficult for a computer to perform. A hummingbird and an eagle would totally be different to a computer. Even all images of ‘Bob’ would not always be in the same pose. Therefore, current CBIRS generally make use of lower level features such as colour, texture and shape. Even though some systems take advantage of high-level features such as the case in face recognition systems, however not all CBIRS are generic, some are designed to perform basic retrieval operations whereas some handle more specific tasks.

2.1.2 CBIR framework

In 2000, Remco and Mirela conducted a survey on recent CBIRS [41]. They surveyed how user queried the system, whether relevance feedback was available, what features were used and how features from query images and database were matched. From their work, they have concluded that most CBIRS would follow a similar design to the framework depicted in Figure 2.1 [41]. The figure shows that a graphical user interface is used to handle users query. The Query formulation could be done in three ways, which is by direct query, Query by example, or by simply browsing the database. An integral part of any CBIRS is the feature extraction module. This module uses specific algorithms to extract visual features from images. Another important module is the Index construction module which creates indexes in the database. Careful implementation of this module could save much time when retrieving images from large scale databases.

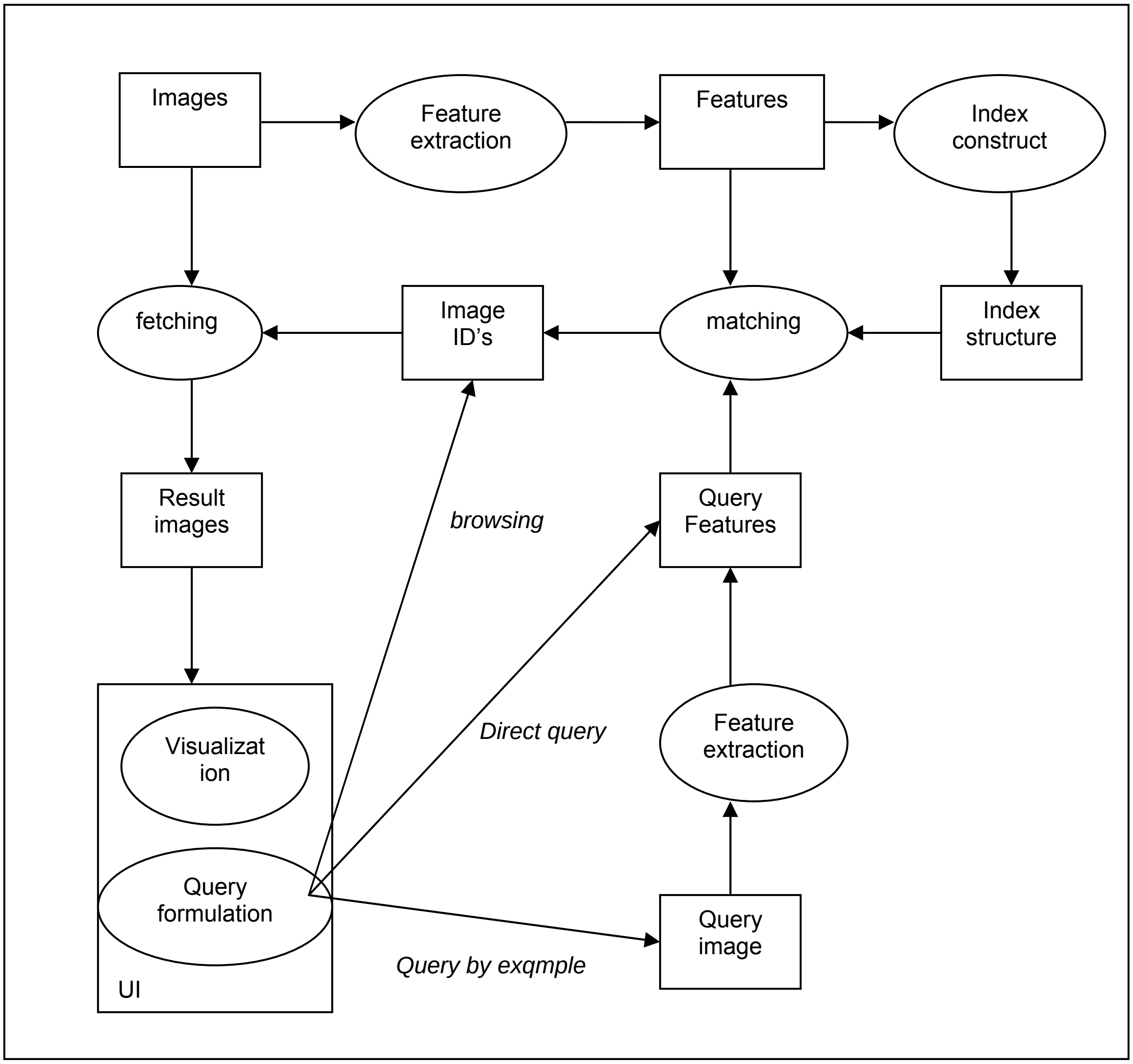

Figure 2.1: Content-based image retrieval framework.

Different implementations of CBIR make use of different types of user queries. Query by example (QBE) for example, requires the user to provide a query image to the system. The system will then extract low level features from the query image, and find for similar images from the database.

In Query by sketch (QBS), users draw a rough approximation of the image they are looking for. This task can be accomplished by creating blobs of flatcolour or drawing a rough sketch of an object in the image and colouring it. The system will then locate images whose layout matches the sketch drawn. Other methods include specifying the proportions of colours desired. For example, users can enter values for each red, green and blue intensity values or by using percentage values (red 20%, green 10%, blue 70%).

2.1.3 CBIR system examples

CBIR plays a major role in many multimedia-based applications. In the areas of crime prevention, CBIR is used for automatic face recognition systems. CBIR is also widely used in security purposes especially those related to biometrics detection systems such as finger print or retina matching [35]. Another area where CBIR is widely used is in medical diagnosis. CBIR systems are used to aid diagnosis by identifying similar past cases from databases of medical images. CBIR is also used to protect intellectual properties. In trademark image registration, new candidate mark is compared with existing marks to ensure no risk of confusing property ownership [6, 13, 28]. In some cases, CBIR have even been used to detect nudity or pornographic images by law enforcement agencies [43].

Many CBIR systems currently exist and many more are being developed. Some of the existing CBIR systems are listed below. A brief description of the methods use for retrieval is also included.

- **Query By Image Content, QBIC**

  QBIC [29] was developed by IBM Almaden Research Centre. It allows users to graphically pose and refine queries based on multiple visual properties such as colour, texture and shape [42]. It supports queries based on input images, user-constructed sketches, and selected colour and texture patterns [32]. For colour and texture queries, QBIC provides users with a colour and texture sampler. The percentage of a desired colour in an image is adjusted by moving a slider [41]. QBIC uses two distance measure to match between query and search images which are weighted Euclidean Distance and Quadratic Distance.

- **VIR Image Engine**

  VIR Image Engine is an extensible framework for building content based image retrieval systems created by Virage Inc. It enables image retrieval based on primitive attributes such as colour, texture and structure. When comparing two images, it examines the pixels in each image and performs an analysis process, deriving image characterization features [42]. A similarity score is computed using distance function defined within the primitive [41].

- **VisualSEEK**

  VisualSEEK were both developed by the Image and Advanced Television Lab, Columbia University. It supports colour and spatial location matching as well as texture matching [42]. To query images, the user sketches a number of regions, positions and dimensions them on the grid and selects a colour for each region [41]. The user could also indicate boundaries for location and size and/or spatial relationships between regions.

- **NeTra**

  NeTra was developed by the Department of Electrical and Computer Engineering, University of California. It supports colour, shape spatial layout and texture matching, as well as image segmentation [42]. Images in the database were segmented into regions of homogenous colour. Of those regions, colour, texture shape and spatial location features were extracted.

- **MARS or Multimedia Analysis and Retrieval System**

  MARS [25] was developed by the Department of Computer Science, University of Illinois and further developed at Department of Information and Computer Science, University of California. It supports colour, spatial layout, texture and shape matching [42]. Users could formulate complex queries using Boolean operators. The desired features could be specified either by example or direct query [41].

Table 2.1 lists down the comparison between the feature types used by CBIR systems described earlier.

Table 2.1: Comparison between image features used in QBIC, VIR, VisualSEEK, NeTra and MARS CBIR systems.

	Key word	Colour					Texture				
		Dominant Colour	Fixed Subimage Histogram	Average Colour Vector	Global Histogram	Other	Atomic Texture Feature	Wavelet/ Fourier Transform	Edge Statistics	Random Fields	Other
QBIC	√	-	-	√	√	-	√	-	-	-	-
VIR	√	-	-	-	-	√	-	-	-	-	√
VisualSEEK	-	√	-	-	-	-	-	-	-	-	-
NeTra	-	-	-	-	-	√	-	√	-	-	-
MARS	√	-	√	-	√	-	√	√	-	-	-

The comparison in Table 2.1 focuses on keyword, colour and texture features only. Some CBIR systems also uses shape features but this is not shown in table 2.1. None of the CBIR systems shown above currently retrieve wavelet-based compressed images.

2.2 Digital image fundamentals

A digital image is represented by a two-dimensional image with a finite set of digitized values, called pixel elements or pixels. Pixels are the smallest individual element in an image, which contains the intensity values of a specific colour for any specific point in the image. A digital image is made up of a finite number of columns and rows of pixels. Digital images could be modelled by using an image function. An image function is a mathematical representation of a two-dimensional image as a function of two spatial variables as defined in Equation 2.1,

$$I_{xy} = f(x, y) \tag{2.1}$$

where I_{xy} is the intensity value for pixel at coordinates (x, y) in the image given by the image function $f(x, y)$. Assume a digital image, *A* with *M* x *N* size. Figure 2.2 shows the representation of image *A*.

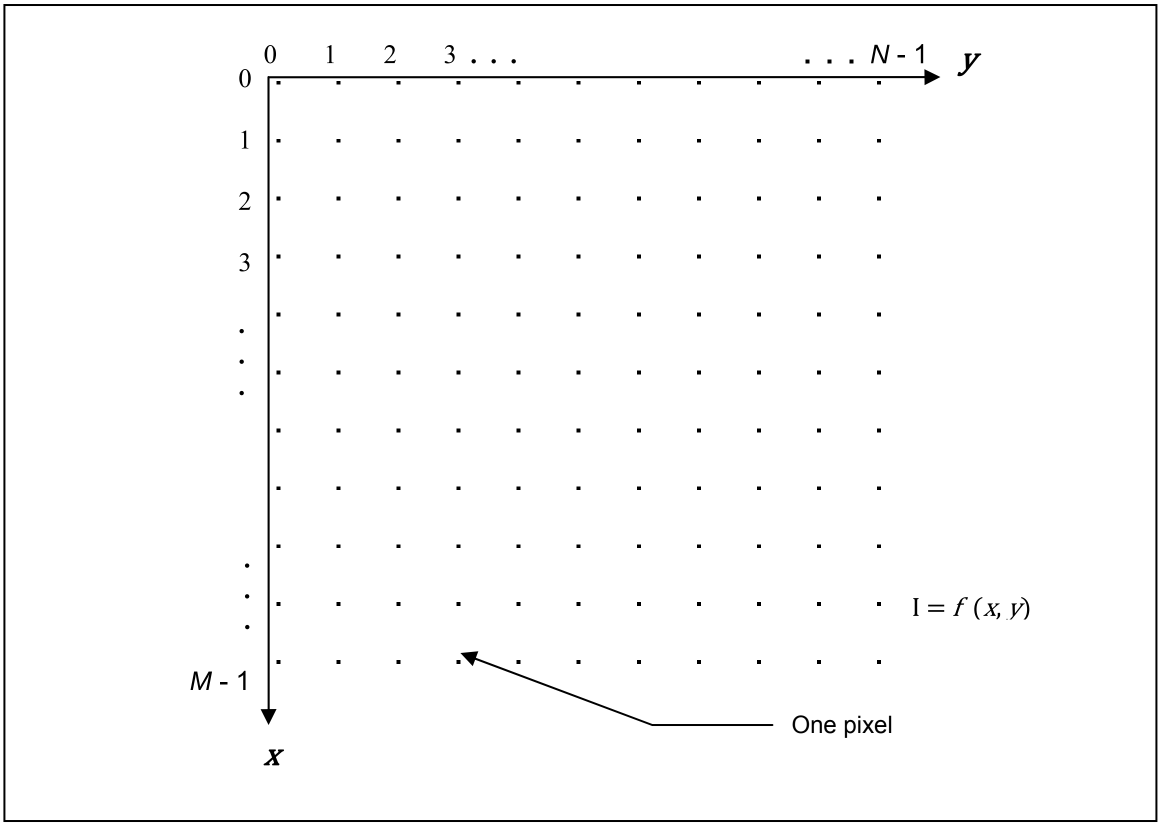

Figure 2.2: Representation of image *A*. The intensity, I for pixel at coordinate (x, y) is given by the image function $f(x, y)$.

The digital image could also be represented by a matrix form as shown in Equation 2.2

$$A = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0, N-1) \\ f(1,0) & f(1,1,) & \dots & f(1, N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \dots & f(M-1, N-1) \end{bmatrix} \quad (2.2)$$

In this representation, A is viewed as a $M \times N$ matrice of pixels. Each pixel is defined by the image function $f(x, y)$ where x and y are the row and columns of the matrice.

2.2.1 Image Domain Processing

Image domain processing falls into 2 main categories which are spatial domain methods and frequency domain methods.

Spatial domain method refers to the process of applying image processing techniques on the image plane itself. This method uses direct manipulation of pixels and its neighbouring pixels in an image.

Frequency domain methods uses signal processing techniques to enhance images. The processing techniques are based on modification of the Fourier transform to represent an image.

2.3 Colour

One of the most important features that make possible the recognition of images by human is colour. Colour is a property that depends on the reflection of light to the eye and the processing of that information in the brain. Light is defined as electromagnetic radiation with wavelength which is visible to the eye (visible light). Electromagnetic waves are made from different wavelengths; the chromatic spectrum spans electromagnetic spectrum from approximately 400nm to 700nm [35]. We detect colours as combinations of the three primary colours which are red, green and blue. We use colour everyday to tell the difference between objects, places, and the time of the day [14]. Colours are defined in three dimensional colour spaces. These could either be RGB (Red, Green and Blue) or HSV (Hue, Saturation and Value).

Most image formats such as JPEG, BMP, and GIF use the RGB colour space to store information [45]. The RGB colour space could be defined as a unit cube with red, green and blue axes. Figure 2.3 shows the RGB colour cube.

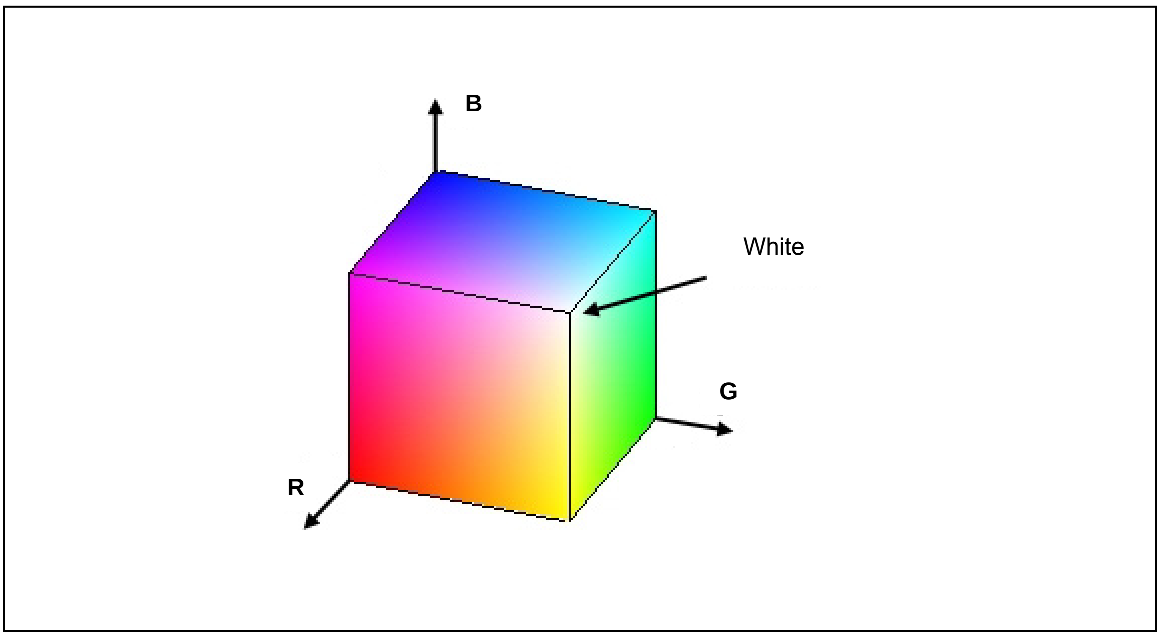

Figure 2.3: The RGB colour cube.

When all three coordinates are set to zero the colour is perceived as black. When all three coordinates are set to 1 the colour perceived is white [45]. The other colour spaces operate in similar fashion but with different perception.

The HSV colour space for example uses a conical form to represent its colour space. In this representation, the hue, H component is depicted by the angle around the axis. The saturation, S is represented by the distance from the centre of a circular cross-section of the cone. The value, V, is the distance from the pointed end of the cone. In some cases, the HSV colour space uses a cylindrical representation, but still follows the same definitions for H, S and V components. Figure 2.4 shows the HSV colour cone.

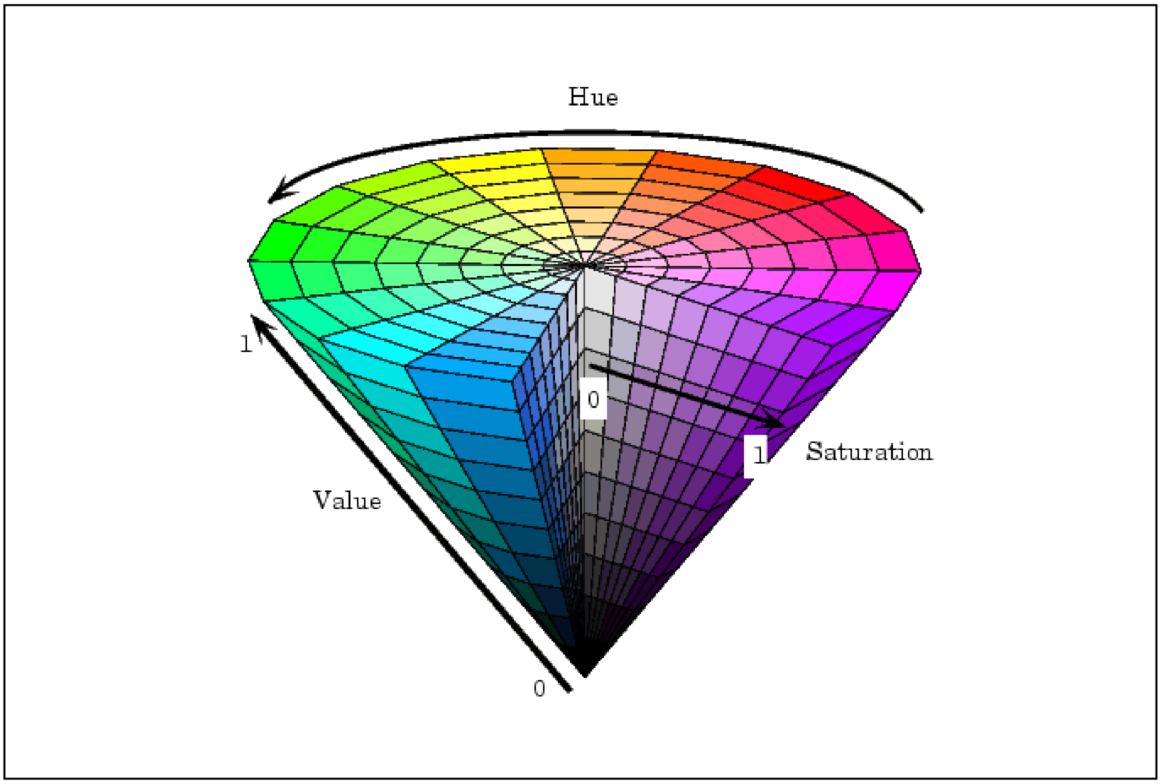

Figure 2.4: The HSV colour cone.

The RGB colour space might be easier to work with compared to HSV colour space as it is easily computed. However in most CBIR systems, the RGB space is often not used as it has the major deficiency of not being perceptually uniform [33]. HSV colour space is dependent on the human perception of hue, saturation and intensity value. Therefore colour perception is better expressed in the HSV compared to the RGB colour space Other colour spaces such as HSV, CIE-LAB, CIE-LUV and Munsell offer better perceptual uniformity. They represent the three colour variants which characterize colour, which are hue, lightness and saturation with equal emphasis [33].

The transformation from RGB colour space to HSV is accomplished through Equations 2.3, 2.4 and 2.5 as given below.

$$h(r,g,b) = \cos^{-1}\left(\frac{\frac{1}{2}\big((r-g)+(r-b)\big)}{\sqrt{\big((r-g)^2+(r-b)(g-b)\big)}}\right) \tag{2.3}$$

$$s(r,g,b) = 1 - \frac{1}{r+g+b}\min(r,g,b) \tag{2.4}$$

$$v(r,g,b) = \frac{r+g+b}{3} \tag{2.5}$$

Where h, s, v, r, g and b represent hue, saturation, value, red, green and blue intensity values respectively.

2.3.1 Methods of colour representation

The main method of representing colour information of images in CBIR systems is through a colour histogram. A colour histogram is a type of bar graph, where each bar represents a particular colour of the colour space being used.

The bars in a colour histogram are referred to as bins and they represent the x-axis. The number of bins depends on the number of colours there are in an image. The y-axis denotes the number of pixels there are in each bin. In other words, it shows how many pixels in an image are of a particular colour. An example of an image together with its corresponding colour histogram is shown in Figure 2.5

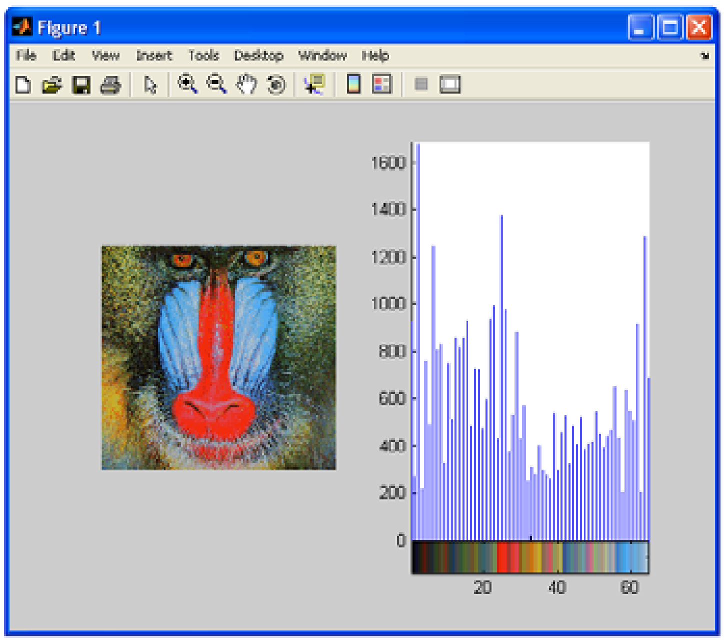

Figure 2.5: An example of an image and its corresponding colour histogram.

A histogram could also be viewed in numerical form. Usually a digital image would have a colour map to store all colours used to represent that image. The colour histogram would be constructed using information from this colour map. A colour map in numerical form is stored in a table with three columns. Each row represents a colour of a bin. The row is composed of three coordinates of the colour space. If we take a RGB for an example, the first coordinate