LOGO IMAGE
DATABASE RETRIEVAL

Oleh

Tan Piet Gek

Disertasi ini dikemukakan kepada
UNIVERSITI SAINS MALAYSIA

Sebagai memenuhi sebahagian daripada syarat keperluan
untuk ijazah dengan kepujian

SARJANA MUDA KEJURUTERAAN (KEJURUTERAAN ELEKTRONIK)

Pusat Pengajian Kejuruteraan
Elektrik dan Elektronik
Universiti Sains Malaysia

May 2006

ABSTRACT

Due to the increasing number of registered trademarks, it is significant to have a retrieval system which is able to detect and retrieve the similar or existing trademarks. In this method, images are retrieved by their contents: color, texture, shape, or objects. Thus, the degree of similarity between query images and images in databases can be measured by color distribution, texture distribution, shape similarity, or object presence between the two images. In this project, I proposed three types of retrieval methods: color, shape boundary, and area measurement. A set of 20 color logo images are used as database images and a set of 10 scanned images and example images copied from website as query images. Color-based retrieval is done by computing the average correlation value between query and database images on their respective R, G and B components. Shape boundary-based method is based on Fourier Descriptors representation. The similarity measure is working based on string similarity function. For area measurement method, object areas are calculated as estimation of total on pixels to represent the whole image. Images with least area difference are considered to be more similar to each other. The overall performance of the retrieval system gives acceptable results. A Graphical User Interface (GUI) is also developed in this project for image retrieval purpose.

ABSTRAK

Disebabkan bilangan tandaniaga berdaftar yang semakin bertambah, adalah penting untuk mempunyai suatu system yang berupaya untuk mengesan dan memperoleh semula tandaniaga berdaftar yang telah wujud atau yang serupa. Untuk sistem ini, imej adalah diperoleh semula melalui kandungan: warna, textur, bentuk atau objek. Oleh itu, tahap kesamaan antara imej input dan imej pangkalan data boleh diukur oleh taburan warna, taburan textur, kesamaan dalam bentuk, atau objek yang wujud antara dua imej tersebut. Dalam projek ini, saya mencadangkan tiga jenis cara untuk perolehan semula imej, iaitu warna, garis luar bentuk, dan kiraan luas. Satu set yang terdiri daripada 20 imej logo berwarna digunakan sebagai imej pangkalan data dan 10 imej lain digunakan sebagai imej pertanyaan. Perolehan semula imej melalui cara warna dilakukan melalui pengiraan nilai purata korelasi antara imej pertanyaan dengan imej pangkalan dalam komponen warna R (merah), G (hijau), dan B(biru) masing-masing. Cara berdasarkan garis luar bentuk pula adalah dilakukan melalui fungsi *string similarity*. Bagi cara kiraan luas pula, luas bagi objek dikira dengan menganggar bilangan pixel ‘on’. Imej yang mempunyai perbezaan luas yang paling kecil berbanding dengan imej pangkalan adalah dianggap paling serupa dengan imej pertanyaan. Di samping itu, antaramuka grafik pengguna (GUI) juga dibangunkan bagi tujuan proses untuk perolehan semula imej.

ACKNOWLEDGEMENTS

For this final year project, I have gained so much, both in terms of wisdom and knowledge in the past few months. For that, I would like to dedicate this section to everyone that involved and helping me during doing this project and make it an unforgettable experience.

Firstly, I would like to extend my highest gratitude and heartiest thanks to my project supervisor, Dr. Khoo Bee Ee, for her important guidance and invaluable advices to me towards achieving the objectives of this project.

I would also like to thanks some of my course mates and school mates, Ms. Fong Wai Mei, Ms. Yap Shu Shin for their impressive ideas and the effort to guide and help me on how to use MATLAB and Image Processing Toolbox. I experienced much and gained a giant step forward to better skills and abilities.

A token of appreciation is also dedicated to my family members, friends and all the staffs in PPKEE (Pusat Pengajian Kejuruteraan Elektrik & Elektronik) that have been given me lot's of guidance, advices, support and make my days in doing this research bright and colorful.

Thanks for the partnership, friendliness and assistants given by you all. Thank you very much.

TABLE OF CONTENTS

LIST OF FIGURES

PAGE

LIST OF TABLES

PAGE

CHAPTER I

INTRODUCTION

Trademarks/Logos are considered valuable intellectual properties and also an important factor to successful business. Since more and more trademarks are registered and protected, the task of registering a newly designed trademark has become more and more difficult. Therefore, certain retrieval techniques must be applied to avoid inadvertent infringement of copyright.

Figure 1.1 below shows some examples of registered trademarks found in Malaysia.

Toyota Car	**Hyundai Car**	**Malaysia Airline System**	**NTV7**

Figure 1.1: Examples of Registered Trademarks

In traditional trademark retrieval systems, the trademarks are first annotated with keywords that are then used for retrieval. Anyway, there are some limitations for text-based image annotation. As a result, it is more desirable to allow the user to input a query by providing sketches or example images and then the system is smart enough to automatically extract features from the query image and use them for comparison in order to search for similar trademarks in the database. The retrieval system is working based on few basic features that designed by the designer, such as color, shape, texture and so on. These are the features that can be easily sensed and differentiated by human vision.

1.1 Basic Concept on Image Retrieval

Image retrieval has been a very active research area within two major research communities – database management and computer vision. These research communities study image retrieval from two different angles. The first is primarily text-based, whereas the second relies on visual properties of the data [V. Castelli and L. D. Bergman, 2002].

1.1.1 Text-Based Retrieval

Text-based image retrieval can be traced back to the late 1970s. Images were annotated by key words and stored as retrieval keys in traditional databases. Although keywords are the most useful in retrieving images that a user wants, sometimes the keyword approach is not sufficient. This process requires a lot of labor in order to assign keywords. Also, the annotation for the same trademark may not be consistent with different users. The other, probably more essential, results from the difficulty of capturing the rich content of images using a small number of keywords, a difficulty which is compounded by the subjectivity of human perception.

1.1.2 Content-Based Retrieval

By nature, images contain unstructured information which makes it a difficult task to be processed and retrieved. Therefore, techniques that seek to index this unstructured visual information are grouped under content-based retrieval.

In the early 1990s, the emergence of large-scale image collection results in the propose method of content-based image retrieval (CBIR) as a way to overcome the difficulties of text-based annotation. Content-based retrieval is actually a type of retrieval by similarity. Images are automatically indexed by summarizing their visual contents through automatically extracted quantities or features such as color, shape, or texture. CBIR system processes the features of image content to establish similarities between images [V. Castelli and L. D. Bergman, 2002].

For low semantic level images, users are concentrating on the basic perceptual features of visual content (dominant colors, texture patterns, relevant edges and 2D shape, and so on). Typical application domains for this level include retrieval of trademark images.

Three major visual perspectives (color, shape, and feature) are generally used to extract features of images, including logo/trademarks. Simultaneous use of different representations often improves retrieval effectiveness, but it also increases the dimensionality of the search space, which reduces retrieval efficiency, and has the potential for introducing redundancy, which can negatively affect effectiveness.

1.2 Application of Content-based Image Retrieval System

A typical content-based image retrieval system is depicted as in Figure 1.2 below. The image collection database contains raw images for the purpose of visual display. The visual feature repository stores visual features extracted from images needed to support content-based image retrieval. The text annotation repository contains key words and free-text descriptions of images. Multidimensional indexing is used to achieve fast retrieval and to make the system scalable to large image collections.

The retrieval engine includes a query interface and a query-processing unit. The query interface, typically employing graphically displays and direct manipulation techniques, collects information from users and displays retrieved results. The query-processing unit is used to translate user queries into an internal form, which is then submitted to the Database Management System (DBMS). Moreover, in order to gap the bridge between low-level features and high-level semantic meanings, users are usually allowed to communicate with the search engine in an interactive way [V. Castelli and L. D. Bergman, 2002].

Figure 1.2 below shows the architecture of a content-based image retrieval system [V. Castelli and L. D. Bergman, 2002].

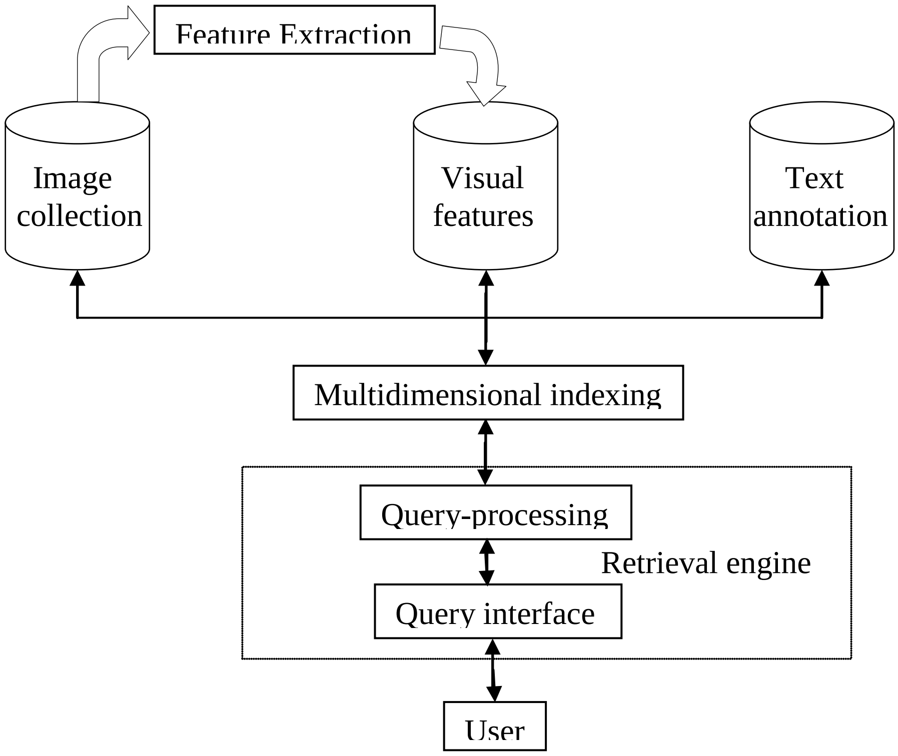


Figure 1.2 : An image retrieval system architecture.

1.3 Types of Content-Based Queries

There are different types of queries typically used for content-based search. The search methods used for image databases differ from those of traditional databases. Exact queries are only of moderate interest and, when they apply, are usually based on metadata managed by traditional database management retrieval-by-similarity. The user search, expressed through one of a number of possible user interfaces [V. Castelli and L. D. Bergman, 2002].

Similarity queries are grouped into three main classes:

a. Range search. Find all images in which feature 1 is within range r_1, feature 2 is within range r_2, and …, and feature n is within range r_n. This type of query requires a complex interface or a complex query language, such as SQL. An example is to find all images showing a circle of $size_{min}$ and $size_{max}$ within a given region.

b. k-Nearest-Neighbor Search. Find the *k* most similar images to the template. It relies on the definition of a similarity function based on pre-defined criteria, which is distance function. Scoring functions are commonly normalized between zero and one; with the best match query has the highest score. An example is to find a set of 10 tumors that are most similar to a specified example, in which the similarity is defined in terms of location, shape, and size, and return the corresponding images.

c. Within-Distance (or α-cut). Find all images with a similarity score better than α with respect to a template, or find all images at distance less than *d* from a template. It relies on a distance or scoring function. An example is to find all images containing tumors with similarity scores larger than α_o with respect to an example provided.

A query can be processed in either classification or abstraction mode. In the classification mode, the database is searched for the specified query primitives that have a spatial extent such as certain shapes, colors or other features. For certain objects that cannot be decomposed into components that are comprised of these primitives, abstraction of that particular feature of the image is needed.

1.4 Interactive Retrieval Interface

A query interface to the application is the processes of sending requests and receiving results. An interactive retrieval interface allows the user to formulate queries. Query-by-example interfaces are designed to help the user of a CBIR system in formulating queries. A Graphical User Interface (GUI) is developed for user to interact with this retrieval system. In most cases, users may not have an exact idea of what they want. To help users quickly form a query, there are three types of tools generally used: browsing, sketching, and feature editing. For this project, I proposed a method of trademark retrieval using query by set of example images.

1.5 Query Formulation

Users tend to use three kinds of research strategies. One of them is search by association when they have no specific aim other than find interesting things. Another class of users aims to search for a specific image. Users may have a specific image in mind and the target is interactively specified as similar to a group of given examples, which are useful for art, industrial components or catalogues. The third class of applications may be the category search. Users aim to retrieve an image from a specific class (catalogues of varieties, for example). The goal is to have a database of images, and then given a sample, find the images in the database that are most similar to it. These are all examples of image retrieval from database, with either no specific aim or targeted at certain specified images for certain purposes.

1.6 Objective and Scope of Project

The main objective of this project is to study and determine the importance in logo images analysis and representation that are useful for digital image processing in retrieval system. The second objective is to design the image retrieval system based on the pre-defined representation. For this project, the images involved are a set of color logo images which are available in Malaysia. Trademark image retrieval is useful to users for the purpose of checking if a similar trademark already exists.

The scope of this project is to study the logo image retrieval scheme, which is for the purpose of compare query images and retrieve the most similar database logo images from database. This proposed retrieval scheme is also compared with another algorithm. Twenty logo images were selected to be used as database images in this project. Their sizes are of dimension 60x60.

1.7 Report Organization

This report contains six chapters. Chapter I is an introduction to image retrieval system including the basic knowledge on image retrieval and applications, types of retrieval, query interface and formulation on image retrieval.

Chapter II contains literature review which summarized the researches done on image retrieval in logo imaging based on some journals and literature studies. Chapter III describes the proposed image retrieval methods including the descriptions and the algorithm of the retrieval scheme. Chapter IV generally describes the use of MATLAB and the GUI design for the logo image retrieval system.

Chapter V shows the experimental results and analysis for the project. Chapter VI is the conclusion and suggestion for future works.

CHAPTER II

LITERATURE REVIEW

In this project, preparation by searching for references or literature reviews are necessary in order to proceed further to start the project. Journals and literatures are search through website, reference books and any kind of material providing useful information.

According to A. Folkers and H. Samet, they proposed method of image retrieval using Fourier descriptors on a logo database. Their retrieval system enables pictorial specification of queries in an image database. The queries are processed in either classification or abstraction method. The database contains both the classification and the abstraction for each image component, where in some cases the classification may be unknown. Each image component is approximated with the pre-defined shape and the quality of the approximation is evaluated to be classified, using matching similarity level (msl), extent similarity level (esl), contextual similarity level (csl) and spatial similarity level (sml). In their method, the experimental results show both classification and abstraction queries produced acceptable results. Abstraction approach also shows that it has a greater flexibility in choosing an arbitrary shape of the components of the query image. [A. Folkers and H. Samet, 2002]

Shinfeng D.Lin, B.Y.Hsu and X. L. Yang proposed trademark retrieval system using a new region based shape description method: the distance-angle pair-wise histogram. The combination of the angle and distance histogram is rotation, scale and translation invariant. The 2D histogram generated from angle and distance histogram is matrix-like, the distance between two arbitrary pair-wise distance-angle histograms is computed by Euclidean distance. Some queried images are modified or added noise. From observation, the proposed method returns better results than moment-based methods. Zernike Moment Magnitudes also not perform well while querying non-symmetric symbols or trademarks. These show that the proposed method achieves good results and outperforms moment based retrieval methods in general cases [Shinfeng D.Lin et al., 2002]

Multi-modal retrieval of trademark images using global similarity is proposed by S. Ravela and R. Manmatha. Images are characterized and retrieved using associated text and visual appearance. Geometric features used are image shape index (a ratio of curvatures of the three dimensional intensity surface) and local orientation of gradient are computed from images. Two images are said to be similar if they have similar distributions of such features. Global representation uses histograms of local curvature and phase capture the distribution of local features. Query histograms are compared with the one stored, and the images are then ranked and displayed. This technique gives good results, and is reasonably tolerant to view variations. In conclusion, they claim that global representations are better constructed by representing the distribution of robustly computed local features [S. Ravela and R. Manmatha, 1998].

Jusub Kim proposed method of shape-based image retrieval in logo databases. He claims that shape similarity is challenging since good segmentation and recognition algorithms are necessary in retrieving images and the states of the algorithms are still primitive. This retrieval system is dealing with black-white images. A bounding border is added around a logo and the logo features included the negative feature between the bounding border and the original logo. Local and global shape features are computed for each component using several shape descriptors. Inner shape information in each connected component of the image is necessary to be extracted, based on two interior model, transparent patch and opaque patch. The sample queries test shows reasonable results [Jusub Kim]

Another approach of trademark retrieval is proposed by Wing Ho Leung and Tsuhan Chen, which is using contour-skeleton stroke classification. For this method, query by sketches is used. User can search for similar trademarks by providing a rough sketch instead of keywords. Trademark images are first filtered to then are segmented into regions based on pixel connectivity. For each region, a decision is made about whether thinning or edge extraction should be applied. Stroke tracing is then performed to extract the sketch of the trademark. The query sketch that will be compared with those extracted sketches from the database images to retrieve the similar trademarks. Shape estimation is then used for determine the primitive shape type the query falls in. The similarity score between the strokes is calculated after finding the correspondence based on both spatial order of the strokes in the x and y directions and the feature distance of the strokes. By classifying the sketch into skeleton strokes and contour strokes, the retrieval performance is better than simply skeleton strokes or simply contour strokes. [W. H. Leung and T. Chen, 2002].

According to A. Soffer, and H. Samet, they proposed logo similarity matching using negative shape features. Border is added around the logos that are not enclosed in a simple shape. Logos are segmented. Local and global shape features are computed for each component. Four global shapes descriptors and three local shape descriptors are used to compute the similarity between two logos. The similarity measure is the weighted Euclidean distance between the feature vectors of these two representative components. In conclusion, adding border to logos and characterizing them with both positive and negative shape features for logo similarity matching using the similarity measure is effective both for finding logos that belong to the same class and for finding a logo given an edited version of it [A. Soffer and H. Samet, 2001]

CHAPTER III

PROPOSED RETRIEVAL METHOD

There are various approaches to develop an Image Retrieval System (IRS). Most shape databases have been applied to trademark databases where global shape features and possibly color and texture, are used.

In my project, I proposed a method of trademark retrieval using query by example of images, including scanned images and images copied from website. The retrieval method is based on color, boundary-based descriptions (Fourier descriptors) and shape measurement, area. The query image is extracted and processed to compute the similarity between the query and the database trademarks. The similarity scores are ranked and the database trademarks with the highest similarity score and the second highest rank are retrieved.

3.1 Image and Digital Image Processing

An image may be defined as a two-dimensional function $f(x, y)$, where x and y are spatial coordinates, and the amplitude of f at any pair of coordinates (x, y) is called the intensity or gray level of the image at that point. When x, y, and the amplitude values of f are all finite, discrete quantities, the image is known as digital image. A digital image has been discretized in spatial coordinates and in brightness, which is represented by a 2-dimensional array or a series of 2-dimensional arrays, one for each color band. A single pixel represents a value of either light intensity or color.

The field of digital image processing refers to processing digital images by means of a digital computer. A digital image is composed of a finite number of elements, each of which has a particular location and value. These elements are referred to as picture elements, image elements, pels, and pixels. Pixel is the term most widely used to denote the elements of a digital image [R. C. Gonzalez et al., 2004].

The following is a representation for a digital image:

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

Images are processed to obtain information beyond what is apparent given the image's initial pixel values.

3.2 Pixel

Generally, all digital images can be expressed as an N x M metrics pixel value. The color value in each pixel is unique and is combined to be the image. Origin point or starting point at coordinate (*x*, *y*) for digital image starts from the upper left metrics, where

$\boldsymbol{i = x} \qquad \mathbf{0 \leq \boldsymbol{i} \leq N} \quad @ \quad \mathbf{1 \leq \boldsymbol{i} \leq N}$

$\boldsymbol{j = (M-y)} \qquad \mathbf{0 \leq \boldsymbol{j} \leq M} \quad @ \quad \mathbf{1 \leq \boldsymbol{j} \leq M}$

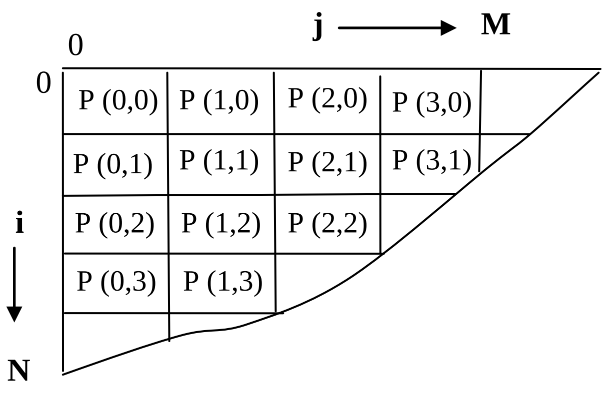


Figure 3.1: Pixel Indexing In An Image Metrics.

3.3 Image processing

Two steps are important and sufficient to describe the content of an image. The first one is the image-content operations that transpose the data into another spatial data array, in three units: color, texture and geometry. They may be characterized by: $f(x) = g \circ i(x)$, where $i(x)$ is the image, g is an operator on images and the resulting image filed is $f(x)$. In that way, image retrieval with image processing must enhance aspects in the image data relevant to the query and to reduce the remaining aspects.

3.4 Feature Extraction

Feature extraction is the basis of CBIR. Features can be categorized as general or domain-specific. General features typically include color, texture, shape, texture, shape, sketch, spatial relationships, and deformation, whereas domain-specific features are applicable in specialized domains such as human face recognition or fingerprint recognition. For logo images retrieval, low-level features such as colors, textures, and shape of objects are widely used [V.Castelli and L. D. Bergman, 2002].

3.4.1 Color

Color is one of the most recognizable elements of image content and is widely used in image retrieval of its invariance with respect to image scaling, translation, and rotation. Generally, color descriptors are relatively easily extracted and matched. Therefore they are well-suited for content-based query. The key issues in color feature extraction include the color space, color quantization, and the choice of similarity function. [1]

Color makes the image $i(x)$ take values in a color vector space. The commonly used color spaces include RGB (red, green, and blue), YCbCr, HSV (hue, saturation, and value), NTSC (luminance, hue, and saturation), CMY (cyan, magenta, and yellow), CMYK (cyan, magenta, yellow, and black) and HSI (hue, saturation, and intensity). There are formulas for conversion of color spaces from RGB to other color spaces, and from other color spaces back to RGB [R. C. Gonzalez, 2004].

3.4.1.1 RGB color representation

RGB color space is represented as a three-dimensional coordinate system, with the axes corresponding to the red, green, and blue contributions, respectively. RGB color representation is a good choice when there is no variation in the perception of an object. It describes the image in its literal color properties and it is sufficient in the case of art-paintings and color composition of photographs, for example. For retrieval applications, this space is helpful thanks to its opponent color representation, in using the opponent color axes (R-G, 2B-R-G, R+G+B). The advantage is that axes are invariant to changes in illumination intensity and shadows.

An individual color is encoded as a coordinate within this RGB space. Thus, a color consists of three elements: a red value, a green value, and a blue value, at different ratio. The following figure shows a RGB color cube with each displayable color corresponds to a location within a three-dimensional color cube, based on Cartesian coordinate. The origin, (0, 0, 0), where each color coordinate is 0, is black. The point at (255, 255, 255) is white, representing an additive mixture of the full intensity of each of the three colors. Points along the main diagonal - where intensities of each of the three primary colors are equal - are shades of gray. The color yellow is represented by the coordinate (255, 255, 0), or a mixture of 100% red, plus 100% green, and no blue.

Figure 3.2 shows the RGB color cube with the coordinates of each color at each location. The vertices of the cube are the primary (red, green, and blue) and secondary (cyan, magenta, and yellow) colors of light.

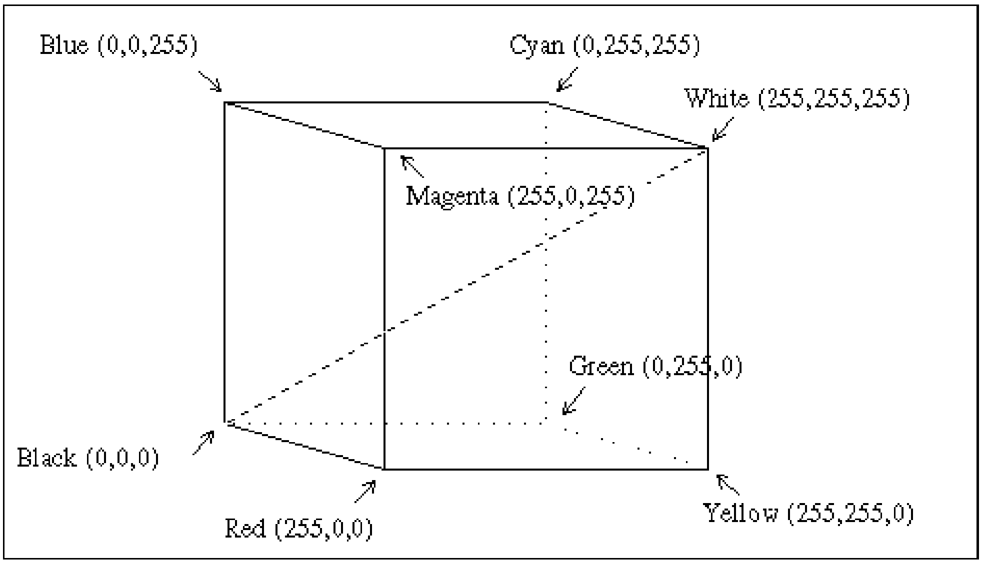


Figure 3.2: RGB Color Cube

RGB can represent more than 16 million colors (which is often called *True Color*). The calculation is as below:

256 red values x 256 green values x 256 blue values = 16,777,216 colors

Figure below shows an example of Malaysia Airline System (MAS) logo with the corresponding Red, Blue, and Green components.

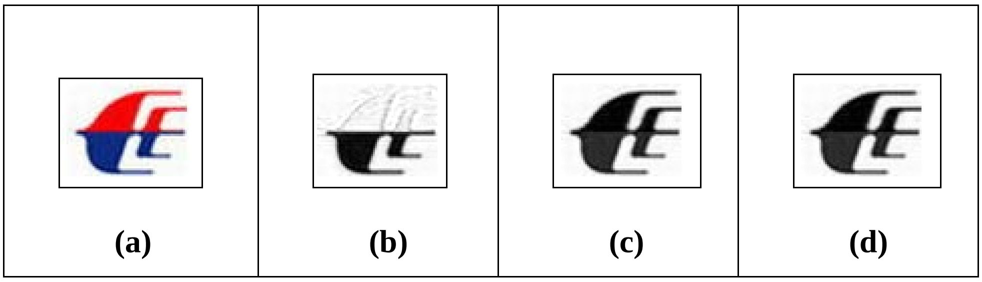


Figure 3.3: An example of logo image and the extracted three color component images.

a) original logo b) extracted Red component

c) extracted Green component d) extracted Blue component

3.4.1.2 Color for Image Retrieval

For retrieval method by color, correlation is used to compute the two-dimensional correlation coefficient between two matrices or vectors of the same size. Therefore, we are using query images and database images of same sizes.

Correlation therefore can be measured between the query image and database images. The relative measurement of the correlation is defined as:

$$Correlation = \frac{\sum_m \sum_n (A_{mn} - \tilde{A})(B_{mn} - \tilde{B})}{\sqrt{\left[\sum_m \sum_n (A_{mn} - \tilde{A})^2\right]\left[\sum_m \sum_n (B_{mn} - \tilde{B})^2\right]}} \quad \text{----- (1)}$$

A_{mn} = original R/ G/ B component value

$\tilde{A}$ = original R/ G/ B component average value

B_{mn} = extracted R/ G/ B component value

$\tilde{B}$ = extracted R/ G/ B component average value

where $|\,Correlation\,| \leq 1$.

The average correlation value of the three components is defined as follow:

$$Average\,Correlation = \frac{Correlation\,R + Correlation\,G + Correlation\,B}{3} \quad \text{----- (2)}$$

Query images and database images are of sizes 60x60. Therefore, the correlation is actually the relative measurement of the corresponding location of pixel value between query image and database images, based on Cartesian coordinate in image. The correlation value is ranging from 0 to 1. If the value of correlation is 0, the two variables compared are totally different. On the other hand, if the value of correlation is 1, the two variables compared are exactly the same or identical. This means that the higher the correlation value (between 0 and 1), the more similar the two comparable variables. In this case, the images retrieved with higher correlation shows that they are more similar to the query image.

In this project, I used clustering schemes where I classify the database images into few groups with different color. If the highest correlation value is in certain class, then the second rank image to be retrieved is also from the same class. Correlation measures the degree of similarity in RGB color components. The database color images and the input query images are extracted to Red (R), Green (G), and Blue (B) components, respectively. The correlation values are then sum up and divide by 3 to get the average correlation value of the RGB components. After extracting the three components, the images are then compared based on the extracted pixel values.

This concept is roughly described by the following figure.

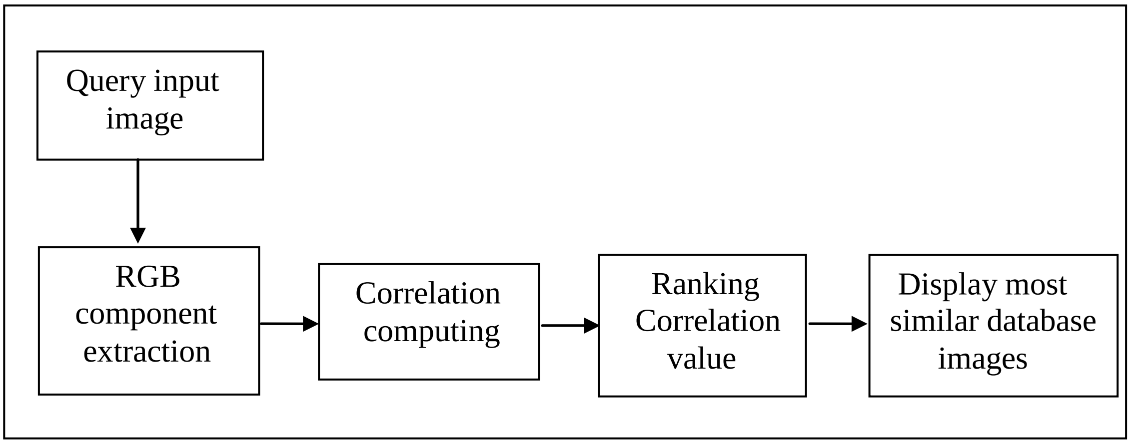


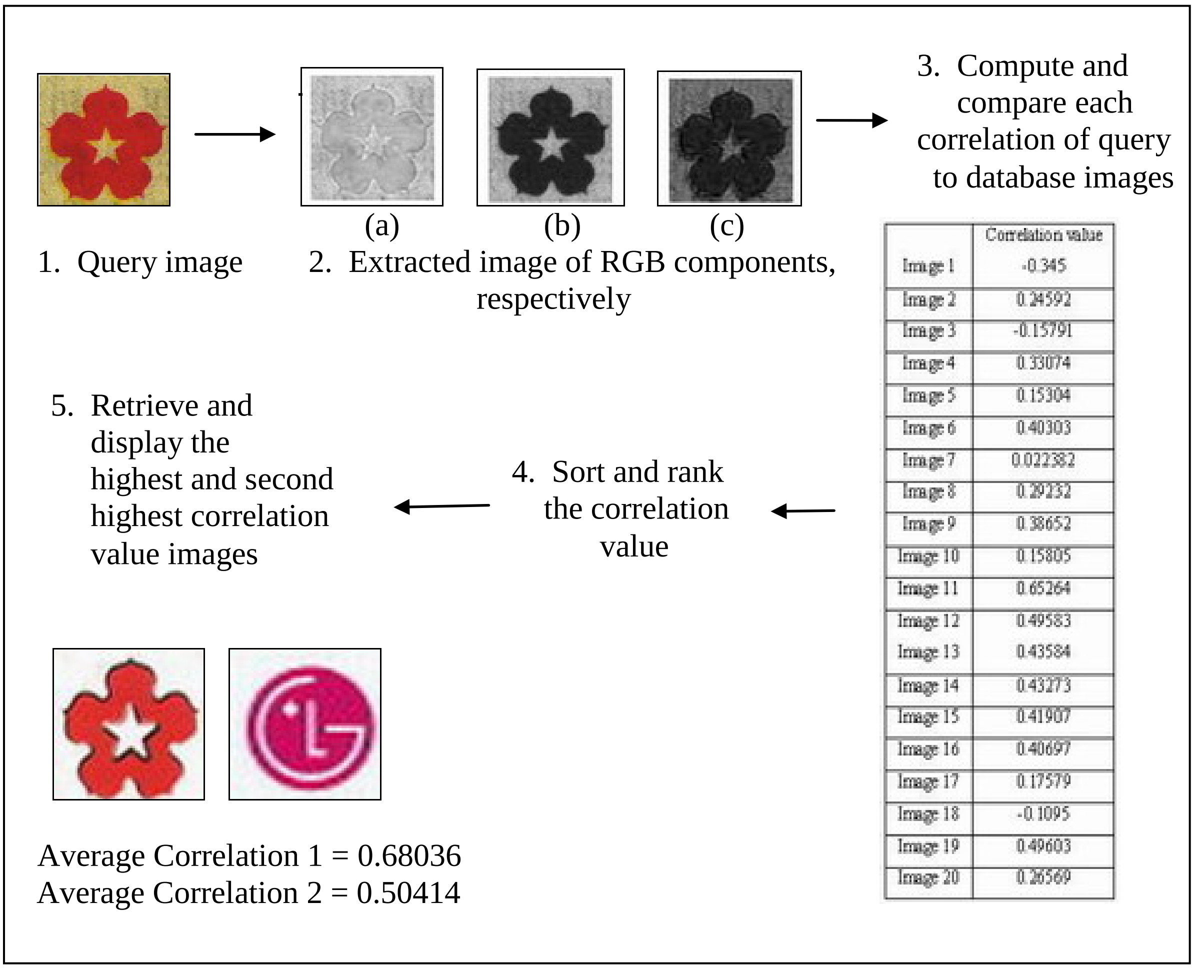
Figure 3.4: Rough concept of color-based retrieval

The color-based retrieval procedures are as follows as describe in detail in Figure 3.5:

Step 1 - User insert query images to the system through Graphical User Interface (GUI).

Step 2 - Correlation values are computed for the three components of query image with respect to each of database images.

Step 3 - Two images with the highest value and the second highest will be retrieved and displayed.



	Correlation value
Image 1	-0.345
Image 2	0.24592
Image 3	-0.15791
Image 4	0.33074
Image 5	0.15304
Image 6	0.40303
Image 7	0.022382
Image 8	0.29232
Image 9	0.38652
Image 10	0.15805
Image 11	0.65264
Image 12	0.49583
Image 13	0.43584
Image 14	0.43273
Image 15	0.41907
Image 16	0.40697
Image 17	0.17579
Image 18	-0.1095
Image 19	0.49603
Image 20	0.26569

Figure 3.5: Detail flow-chart of color-based retrieval

3.5 Shape

Two major steps are involved in shape feature extraction. They are object segmentation and shape representation.

3.5.1 Object Segmentation

Image retrieval based on object shape is considered to be one of the most difficult aspects of content-based image retrieval because of difficulties in low-level image segmentation and the variety of ways to given three-dimensional (3D) object can be projected into 2D shapes. Several segmentation techniques have been proposed so far and include the global threshold-based technique, the region-growing technique, the texture-based technique, the color-based technique, and the model-based technique. Generally speaking, it is difficult to do a precise segmentation owing to the complexity of the individual object shape, the existence of shadows, noise and so on. Anyway, this is the essential process for the image to be useful for further processing.

3.5.2 Boundary-Based Representation

For boundary-based representation, boundary data is extracted as (x, y) coordinates in the image space and needs to be represented in a form suitable for archiving, indexing, and similarity matching. The representation is based on the outer boundary of the shape. A shape representation method converts a dense 2D representation of a boundary, i.e., the (x, y) coordinates of boundary points, into a form that has the following properties:

- *Invariance*: two boundaries that have the same shape should have the same representation.
- *Uniqueness*: two boundaries with different boundaries should have different representations.
- *Stability*: a small change in the boundary should cause a small change in its representation.

- *Efficiency*: the representation should be computationally efficient to compute and store.
- *Ease of implementation*: the representation should be the least complex of competing methods to implement.
- *Computation of shape properties*: the representation should support computation of shape properties such as symmetry, area, perimeter, etc.
- *Meaningful representation*: The representation should retain properties of the shape that are meaningful to the application.

These requirements may be extended to include matching of partial boundaries or specific local regions in the boundary. In addition to the shape representation properties identified above, other properties more relevant to retrieval are described as below:

- *Geometric Invariance* (also known as Similarity Invariance): the representation should be invariant to rotation, translation, and scaling.
- *Compact representation*: essential shape information should be maintained in a significantly reduced manner.
- *Fast matching speed*: efficient shape feature computation and fast retrieval in a large image database.
- *High quality image retrieval*: retrieved images should be relevant to the inquirer.

The commonly used descriptors of this class include the chain code, the Fourier Descriptors, UNL Fourier, and so on. For my project, I proposed the method of shape representations by its boundary (outline) by Fourier descriptors.

3.5.3 Region-Based Representation

This type of representation describes the object in image based on the entire shape region. Descriptors of this class include Moment Invariants (MI), Zernike Moments (ZMM), the morphological descriptor, and Pseudo Zernike Moments.

3.5.4 Combined Representation.

We may consider the integration of several basic representations such as moment invariants with the Fourier descriptor or moment invariants with the UNL descriptor.

3.6 Fourier descriptors

One of the well established methods for describing shape of closed curve is the Fourier descriptors, which serve powerful boundary-shape representation tools. However, there is a limit to expressing an object by its boundary because the boundary itself does not represent inside shape feature of the object. Each image of the logo database is decomposed into its image components, i.e., its connected components. As an abstract representation of image components, we use Fourier descriptors, which are invariant against translation, scale, rotation, and their starting point. We retain the phase which contains essential information about the contour of the image component.

3.6.1 Overview of the Fourier descriptors methods

- Transformation to the tangent space

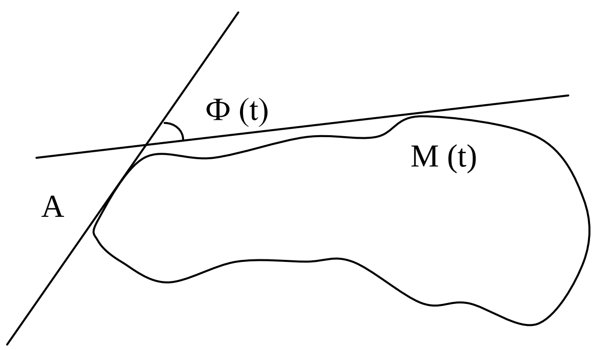


Figure 3.6: Closed-curve Fourier description

The outline of the shape is seen as a closed curve, described by its arc length **s** from an origin **A**. We normalize this parameter so that its sum over the whole curve is equal to **2.Pi**, it is the parameter **t** as shown below. We define a function of **t**, called **Phi**, that gives the angular variation between the tangent at the origin **A** and the tangent at position **s**.

$$\Phi(t) = \phi(t) - t \text{ with } t = \frac{2\pi s}{L}$$

This function is real, continuous, and periodical with a period **2.Pi**. Therefore, it can be described by a Fourier series:

$$\Phi(t) = \sum_{k=0}^{\infty} a_k \exp(-jkt)$$

The set of modules of the coefficients **a(k)** is called Fourier descriptors. They stay constant no matter one of the following transformations: translation or rotation of the shape, change of scale or of origin.

One can compare two shapes by comparing subsets of the Fourier descriptors, beginning with the lower order coefficients, and then using higher order coefficients. To make the description simpler, one can get rid of the Fourier descriptors of higher order. But that can cause much trouble because when a modification is introduced in the set of Fourier descriptors, the new shape described is usually no longer a curve. Some conditions allow ensuring a correct closure of the curve, but they seem very restrictive.

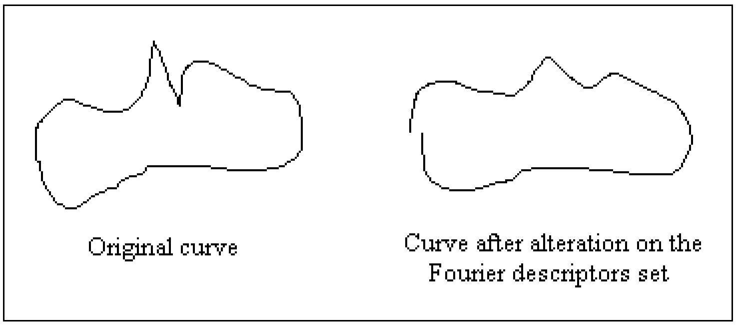


Figure 3.7: Original closed-curve and curve after alteration on the Fourier descriptors set

Therefore, a second method using a complex function will usually be preferred.

3.6.2 Complex Fourier descriptors

Granlund introduced Fourier descriptors using complex representation in 1972. This method ensures that a closed curve will correspond to any set of descriptors.

The shape is now described by a set of **N** vertices {**z(i) : i = 1,…,N**} corresponding to **N** points of the outline. The Fourier descriptors {**c(k) : k = -N/2 + 1,…,N/2**} are the coefficients of the Fourier transform of **z**:

$$z_i = \sum_{k=-\frac{N}{2}+1}^{\frac{N}{2}} c_k \exp(2\pi j \frac{ki}{N})$$

The inverse relationship exists between **c(k)** and **z(i)**:

$$c_k = \frac{1}{N} \sum_{i=1}^{N} z_i \exp(-2\pi j \frac{ik}{N})$$

The Fourier descriptors have useful properties:

The range of **k** can be restricted to {**-N/2 +1, N/2**}: according to Shannon's theorem, the highest frequency is obtained for **k = N/2**, and any **c(k)** with **k** greater than **N/2** would be redundant since we use a discrete representation of the outline.

The descriptors **c(k)** describe the frequency contents of the curve: a value of **k** close to zero will describe low frequency information, an approximative shape, and the higher frequencies will describe details.

- For **k = 0**, **c(k)** represents the position of the center of gravity of the shape. This term is not interesting for the shape description. Without this term, the description will not be affected by a translation of the shape.
- The first frequency component, **c(k)** for **k = 1**, describes the size of the shape. If all the other components are set to zero, the shape becomes a circle

(in fact, a **N**-sided polygon). One can use this component to normalize the set of Fourier descriptors, so they remain constant after a homothety on the shape.

- The other frequency components will make alterations on the circle described by **c(1)**. The descriptors **c(k)** have opposite effects for positive and negative values of **k**:
  - For positive values of **k**, the effect of **c(k)** is to “push” the circle at **k–1** points periodically placed on the circle. This means the curve will be bent towards the outside of the circle at **k–1** points.
  - For negative values of **k**, the effect of **c(k)** is to “pull” the circle at **1-k** points periodically placed on the circle. The curve will be bent towards the inside of the circle this time.
  - The phase of **c(k)** is a rotation of this perturbation, which means it describes the place on the circle where the action is performed.

Complex Fourier descriptors allow scalability in describing a curve. If we keep only a low frequencies subset of descriptors, we get a curve that just approximates the outline of a shape. By increasing the number of components in the description, high frequencies are also rendered, and sharp curves or details can be generated.

This scalability was already available with the tangent method, but the curve generated was not closed any more after the alteration of the Fourier descriptors set. The complex method does not have this drawback. This method also has its weaknesses: for instance, one cannot guarantee that a simple closed curve will be rendered the good way after truncating the set of Fourier descriptors. Sometimes, the curve generated has one or more crossovers.