

國立交通大學

資訊科學與工程研究所

碩士論文

使用型態篩選法及動態部分函數作影像擷取

Image Retrieval Using Morphological Granulometry and Dynamic
Partial Function

研究生：顏佩君

指導教授：薛元澤 教授

中華民國九十五年六月

使用數學型態篩選法及動態部分函數作影像擷取
Image Retrieval Using Morphological Granulometry and Dynamic Partial
Function

研究生：顏佩君

Student : Pei-Jun Yan

指導教授：薛元澤

Advisor : Yuang-Cheh Hsueh

國立交通大學
資訊科學與工程研究所
碩士論文



Submitted to Institute of Computer Science and Engineering
College of Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Computer Science

June 2006

Hsinchu, Taiwan, Republic of China

中華民國九十五年六月

使用數學型態篩選法及動態部分函數作影像擷取

學生：顏佩君

指導教授：薛元澤

國立交通大學資訊學院資訊科學與工程系(研究所)碩士班

摘要

近年來，多媒體資料的大量成長。使得如何發展對這些資料作儲存、瀏覽、檢索以及擷取的技術成為一項重要的議題。在影像擷取的研究已有許多，但鮮少有使用到因為能即時應用而著稱的數學型態學。因此，本篇論文的目的是改善由 J.S.Wu 提出的 morphological primitives 的準確度，並且提出另一個型態運算子—型態篩選法作為影像的特徵。實驗結果顯示，morphological primitives 使用動態部分函數為相似度量測最佳有 30% 的準確度改善。而另一方面在大部分的影像中，granulometric histogram 比起 color histogram，color moment primitives，以及 morphological primitives 有較佳的準確度。

Image Retrieval Using Morphological Granulometry and Dynamic Partial Function

Student : Pei-Jun Yan

Advisors : Dr. Yuang-Cheh Hsueh

Institute of Computer Science and Engineering
College of Computer Science
National Chiao Tung University

ABSTRACT

In recent years, there is an explosive growth of multimedia data. How to develop techniques for storage, browsing, indexing, and retrieval them efficiently is a significant issue. There are many researches on image retrievals, but very few usage of mathematical morphology which is impressive because of making real-time applications possible. Our goal is to improve precisions of morphological primitives which was introduced by J.S.Wu and propose another morphology operator: morphological granulometry as an image feature. From experimental results, image retrieval using morphological primitives with dynamic partical function have improvement up to 30%. On the other hand, granulometric histogram has better precisions than those of color histogram, color moment primitives, and morphological primitives in most cases.

誌 謝

我在這裡要感謝我的指導教授 薛元澤教授，兩年來對我孜孜不倦的教誨，教導我研究學問的方法及待人處世道理，讓我畢生受益無窮，以及我的口試委員 張隆紋教授與 陳玲慧教授，二位老師不吝指教，讓這篇論文更加完善。

我還要感謝吳昭賢學長、莊逢軒學長、何昌憲學長、王蕙綾學姐、高薇婷學姐，給予我論文研究及寫作方面等的各種建議，感謝林明志同學、呂盈賢同學、劉裕泉同學、江仲廷同學、王慧縈同學在這兩年內與我共同努力，互相砥礪，陪我度過這段快樂的實驗室生活。

僅將此論文獻給我親愛的家人與朋友，我的父母、弟弟及妹妹，感謝他們在這段期間給我的關心、支持與鼓勵，祝福他們永遠健康快樂。



CONTENTS

ABSTRACT (CHINESE).....	i
ABSTRACT (ENGLISH)	ii
ACKNOWLEDGEMENT	iii
CONTENTS.....	iv
LIST OF FIGURES.....	vi
LIST OF TABLES	vii
Chapter 1 Introduction	1
1.1 MOTIVATION	1
1.2 BACKGROUND.....	2
1.2.1 Content-based image retrieval.....	2
1.2.2 Region-based image retrieval	2
1.3 ORGANIZATION OF THIS THESIS.....	3
Chapter 2 Previous Research	4
2.1 COLOR MODELS	4
2.1.1 RGB Color Model.....	4
2.1.2 HSI Color Model.....	5
2.1.3 YUV and YIQ Color Model	7
2.1.4 XYZ Color Model.....	7
2.1.5 L*a*b* Color Model.....	8
2.1.6 Comparison of above Color Models.....	9
2.2 COLOR FEATURES.....	10
2.2.1 Color Histogram	10
2.2.2 Color Moment Primitives.....	11
2.3 MORPHOLOGICAL OPERATOR	12
2.3.1 Basic definition.....	13
2.3.2 Dilation, Erosion, Opening ,and Closing of Binary Images	14
2.3.3 Extension to Grayscale Images.....	16
2.3.4 Morphological Gradient.....	18
2.3.5 Morphological Granulometry.....	19
2.3.6 Morphological Primitives.....	20
2.4 DISTANCE FUNCTIONS FOR SIMILARITY MEASURE	22
2.4.1 Intersection.....	22

2.4.2	Primitive Similarity	24
2.4.3	Minkowski-Like Metrics	25
2.4.4	Dynamic Partial Function	25
2.4.5	The DPF Enhancement	26
Chapter 3 The Proposed Method		29
3.1	OVERVIEW	29
3.2	MORPHOLOGICAL GRANULOMETRY	31
3.2.1	Distribution of Morphological Granulometry	31
3.2.2	Primitives of Granulometric Distribution	32
3.2.3	Histogram of Morphological Granulometry	33
3.3	SIMILARITY	38
3.3.1	Similarity for Morphological Granulometry	38
3.3.2	DPF for Primitives	38
Chapter 4 Experimental Results		40
4.1	EXPERIMENTAL ENVIRONMENT	40
4.2	EXPERIMENTAL RESULTS	42
4.2.1	Morphological Primitives	42
4.2.2	Granulometric Distribution and its Primitives	44
4.2.3	Granulometric Histogram	45
4.2.4	Comparisons	47
4.2.5	Example of Retrieval Results	50
4.2.6	Processing Time	54
Chapter 5 Conclusions and Future Works		56
5.1	CONCLUSIONS	56
5.2	FUTURE WORKS	56
REFERENCE		58

LIST OF FIGURES

Fig. 2-1 RGB color cube.	5
Fig. 2-2 HSI color model.....	6
Fig. 2-3 L*a*b* color model.....	8
Fig. 2-4 Examples of structuring elements.....	13
Fig. 2-5 Examples of (a) Translation and (b) Reflection.	14
Fig. 2-6 (a). Set A. (b) Structuring element B. (c). The Dilation of A by B. (d). The Erosion of A by B. (e). The Closing of A by B. (f). The Opening of A by B.....	16
Fig. 2-7 (a) original image (b)(c) closing and opening with 3*3 structuring element, respectively.	16
Fig. 2-8 (a) The original of Lena image (b) Dilation of Lena image (c) Erosion of Lena image.	18
Fig. 2-9 (a) The original of Lena image (b) Closing of Lena image (c) Opening of Lena image.....	18
Fig. 2-10 (a) The original of Lena image (b) Morphological gradient of Lena image	19
Fig. 2-11 Two different images with similar color histograms.	24
Fig. 3-1 Architecture of our retrieval system.....	30
Fig. 3-2 (a)-(e) Different processing methods.....	30
Fig. 3-3 Example of different image that has similar opening.....	34
Fig. 3-4 Openings of .Fig.3-3(a).....	35
Fig. 3-5 Openings of Fig.3-3(b).....	36
Fig. 3-6 Openings of .Fig.3-3(a) under YIQ.....	37
Fig. 3-7 Openings of .Fig.3-3(b) under YIQ.....	38
Fig. 4-1 Examples of the image database.	41
Fig. 4-2 The Precision-Recall curves when using morphological primitives. (a) The results of Type1 (b) The results of Type7 (c) The results of Type5.	44
Fig. 4-3 The precision comparison for each type of image database among the proposed methods, granulometric histogram, granulometric distribution, and primitives of granulometric distribution (R=20).....	45
Fig. 4-4 Comparison of GH under RGB and YIQ.....	47
Fig. 4-5 The Precision-Recall curves of different processing methods.	49
Fig. 4-6 The precision comparison for each type of image database among the proposed methods and other methods (R=20).	50
Fig. 4-7 Image retrieval using (a) Color Histogram (b) Color Moment Primitives (c) GH under YIQ with 1 structuring element (d) GH under RGB with 1 structuring element (e) Morphological Primitives with smallest-7 DPF (f) Morphological Primitives with smallest-8 DPF (g) Morphological Primitives for the top relevant 20 images.	54
Fig. 4-8 Histograms of processing time.....	55

LIST OF TABLES

Table 2-1 Characteristics of each color models.....	9
Table 4-1 Differences of precisions between proposed method GH and other methods (MP9, CMP, and CH)....	50



Chapter 1

Introduction

1.1 Motivation

In recent years, there is an explosive growth of multimedia data, such as digital images and video sequences. How to develop techniques for storage, browsing, indexing, and retrieval them efficiently is a significant issue. Thus, image retrieval has become an important research topic.

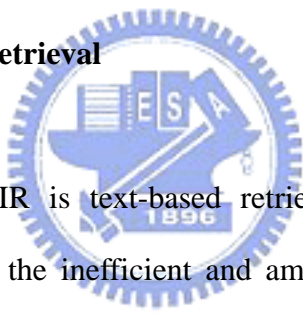
There are many diverse application areas of image retrieval like medicine, remote sensing, education, on-line information services, trademarks, color photographs, stamps, paintings, and traffic signs etc. Each application has its own property. Image retrieval is defined as automatically retrieving relevant images from an image database according to imagery features. Color, texture, and shape are the most frequently cited visual contents for image retrieval. In the MPEG-7 standard [11, 12], there are color, texture, shape, motion, localization, and face recognition features. The color descriptors in the standard include a histogram descriptor, a color structure histogram, a dominant color descriptor, and a color layout descriptor. The three texture descriptors include one that characterizes homogeneous texture regions, and another that represents the local edge distribution, and a compact descriptor that facilitates texture browsing. There are also many other features for image retrieval. Before extracting the features of an image, some methods of image processing such as image segmentation, color clustering, edge extraction, and vector quantization must be applied. If we retrieval a color image, we should emphasize color features more than others in order to get higher precision rate. Thus, to design the retrieval method, we should select the

proper features of image database.

In this thesis, we propose a color, shape and object size feature by using morphological granulometric histogram. Mathematical morphology is impressive because of its simple but effective operators and can be implemented in parallel, making real-time applications possible. It is also a well-known tool in image processing. The granulometry has been proven to be very useful in image segmentation. Therefore, we utilize the advantage of morphological operators to extract the granulometric histogram to represent the most important components in images.

1.2 Background

1.2.1 Content-based image retrieval



The first version of CBIR is text-based retrieval scheme. A query-by-image was proposed in 1990s because of the inefficient and ambiguous of textural annotation. CBIR relies on global properties of an image. When global features cannot represent local variations in the image, simple query-by-image may fail. Content-based image retrieval from digital libraries (C-BIRD), query by image content (QBIC), multimedia analysis and retrieval system (MARS) are examples of CBIR systems.

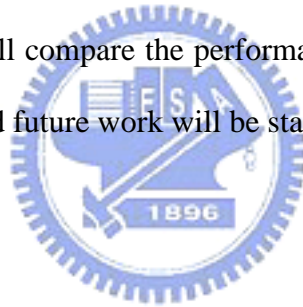
1.2.2 Region-based image retrieval

The emergence of region-based image retrieval (RBIR) was focus on solving the problem of inadequate representation of global features from CBIR. Because CBIR discards information of object location, shape, and texture. Therefore, RBIR overcomes the deficiencies of CBIR by representing images at the object-level and being more semantically.

There are some RBIR systems, such as the VisualSEEk, Blobworld, Integrate region matching (IRM), region of interest (ROI), and finding region in the pictures (FRIP) [13].

1.3 Organization of this Thesis

The remainder of this thesis is organized as follows. In chapter 2, we will survey the research of image retrieval and discuss some issues need to concern. Then, we will survey the concept of morphological operations and morphological granulometry. In chapter 3, we will present our methods include using granuloemtric distribution, primitives of granuometric distribution, and granulometric histogram to describe objects sizes. Then, we will define how to evaluate similarity value between two images. In chapter 4, we will experiment with different features. Then, we will compare the performance of our method with other method. In chapter 5, the conclusion and future work will be stated.



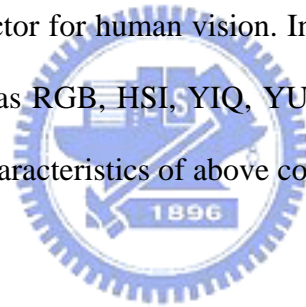
Chapter 2

Previous Research

We describe several color models in section 2.1, and color features which are used for image retrieval in section 2.2. In section 2.3, the basic morphological operators will be introduced. Finally, the distance functions for measuring image similarity are described in section 2.4.

2.1 Color Models

Color is one important factor for human vision. In this section, we introduce the formats of several color models, such as RGB, HSI, YIQ, YUV, XYZ, and $L^*a^*b^*$. We also list the results of [3] which give the characteristics of above color models.



2.1.1 RGB Color Model

In the RGB model, each color appears in its primary spectral components of Red, Green, and Blue. This model is based on a Cartesian coordinate system, its color subspace of interest is the cube shown in Fig. 2-1. It is a hardware-oriented model, which is most commonly used for color monitors and a broad class of color video cameras. Therefore, processing images in the RGB color space directly is the fast method. However it is not natural for human's perception. We can summarize by saying that RGB is ideal for image color generation, but its use for color description is much more limited.

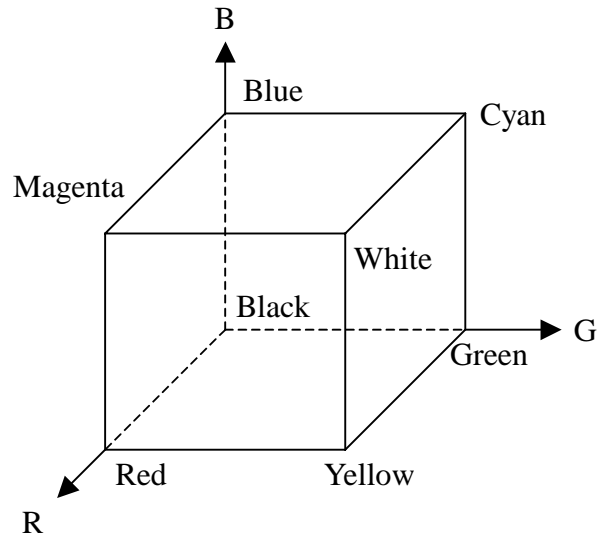


Fig. 2-1 RGB color cube.

2.1.2 HSI Color Model

The HSI color model is composed of Hue, Saturation, and Intensity. It is referred to HSV color model using the term Value instead of Intensity. Intensity is the brightness value, Hue and Saturation encode the chromaticity values. The HSI is very important and attractive for image processing since it represents colors similar to which human eye senses.

Given an image in RGB color format, the H component of each RGB pixel is obtained using the equation [1]

$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360 - \theta & \text{if } B > G \end{cases} \quad (1)$$

where

$$\theta = \cos^{-1} \left\{ \frac{\frac{1}{2}[(R-G) + (R-B)]}{\left[(R-G)^2 + (R-B)(G-B) \right]^{1/2}} \right\} \quad (2)$$

The saturation component is given by

$$S = 1 - \frac{3}{(R+G+B)} [\min(R, G, B)] \quad (3)$$

Finally, the intensity component is given by

$$I = \frac{1}{3}(R+G+B) \quad (4)$$

Fig. 2-2 shows the HSI model based on color circles. The Hue component describes the color itself in the form of an angle between [0,360] degrees. 0 degree mean red, 120 means green 240 means blue. 60 degrees is yellow, 300 degrees is magenta. The Saturation component signals how much the color is polluted with white color. The range of the S component is [0,1]. The Intensity range is between [0,1] and 0 means black, 1 means white.

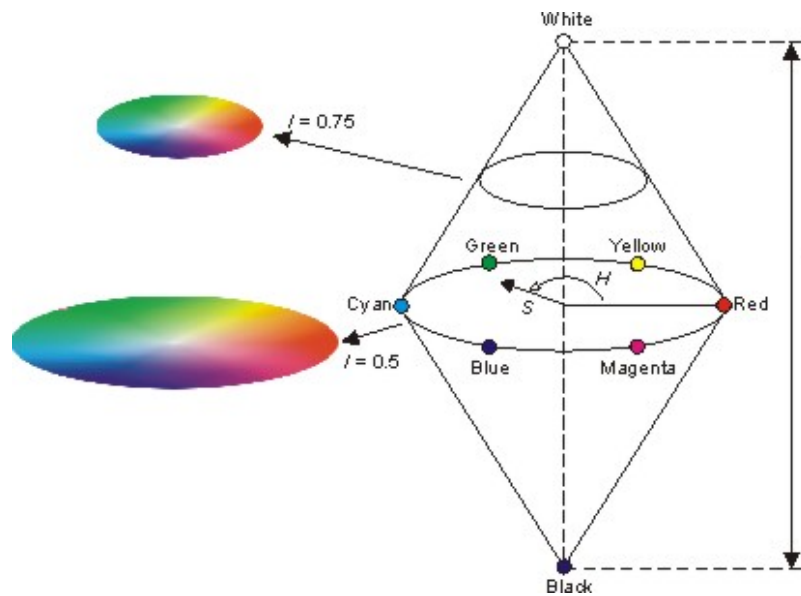


Fig. 2-2 HSI color model.

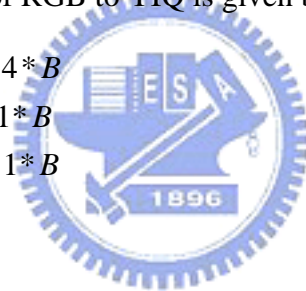
2.1.3 YUV and YIQ Color Model

In the YUV color model, Y represents the luminance of a color while I and Q represent the chromaticity of a color. The conversion of RGB to YUV is given by [2].

$$\begin{aligned} Y &= 0.30R + 0.59G + 0.11B \\ U &= 0.493*(B - Y) \\ V &= 0.877*(R - Y) \end{aligned} \quad (5)$$

The YIQ is used in NTSC color TV broadcasting. The original meanings of these names came from combinations of analog signals – I for in-phase chrominance, and Q for quadrature chrominance. The conversion of RGB to YIQ is given by [2].

$$\begin{aligned} Y &= 0.299*R + 0.587*G + 0.114*B \\ I &= 0.596*R - 0.275*G - 0.321*B \\ Q &= 0.212*R - 0.523*G + 0.311*B \end{aligned} \quad (6)$$



The luminance-chrominance color models (YIQ, YUV) are proven effective. Hence, they are also adopted in image compression standards such as JPEG and JPEG2000. The YUV color space is very similar to the YIQ color space and both were proposed to be used with the NTSC standard, but because the YUV do not best correspond to actual human perceptual color sensitivities. NTSC uses I and Q instead.

2.1.4 XYZ Color Model

The conversion of RGB to XYZ is given by [2].

$$\begin{aligned}
 X &= 0.607 * R + 0.174 * G + 0.200 * B \\
 Y &= 0.299 * R + 0.587 * G + 0.114 * B \\
 Z &= 0.000 * R + 0.066 * G + 1.116 * B
 \end{aligned}
 \tag{7}$$

2.1.5 L*a*b* Color Model

L*a*b* color model (also known as CIELAB) is a kind of visual uniform color system proposed by C.I.E. Its color model is shown in Fig. 2-3. The L* value represents luminance and ranges from 0 for black to 100. The maximum and minimum of value a* correspond to red and green, while b* ranges from yellow to blue.

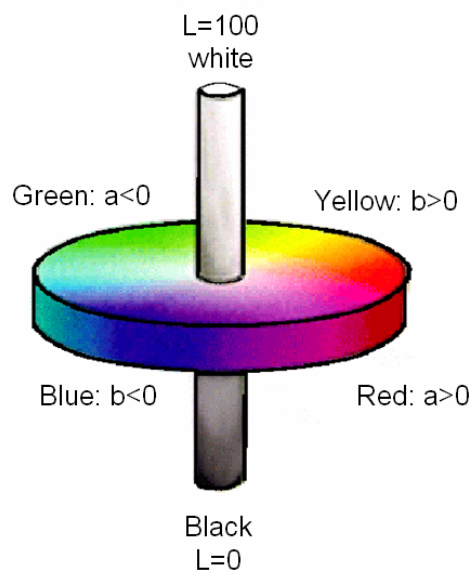


Fig. 2-3 L*a*b* color model

The conversion of XYZ to L*a*b* is given by [2].

$$\begin{aligned}
 L^* &= 116 \left(\frac{Y}{Y_n} \right)^{1/3} - 16 \\
 a^* &= 500 \left[\left(\frac{X}{X_n} \right)^{1/3} - \left(\frac{Y}{Y_n} \right)^{1/3} \right] \\
 b^* &= 200 \left[\left(\frac{Y}{Y_n} \right)^{1/3} - \left(\frac{Z}{Z_n} \right)^{1/3} \right]
 \end{aligned}
 \tag{8}$$

with X_n, Y_n, Z_n the XYZ values of the white point.

2.1.6 Comparison of above Color Models

Above color models has been claimed by many research that they are the most suitable color model for some reasons. In Table2-1, we list the comparison [3] of these color models in the following five terms.



- A. Is the color model device independent?
- B. Is the color model perceptual uniform?
- C. Is the color model linear?
- D. Is the color model intuitive?
- E. Is the color model robust against varying imaging conditions?

Table 2-1 Characteristics of each color models

	A	B	C	D	E
RGB	N	N	N	N	Dependent on viewing direction, object geometry, direction of the illumination, intensity and color of the illumination.
HSI	N	N	H:N	Y	H: Dependent on the color of the illumination.

			S: N		S: Dependent on highlights and a change in the color of the illumination
			I: Y		I: Dependent on viewing direction, object geometry, direction of the illumination, intensity and color of the illumination.
YUV and YIQ	Y	N	Y	N	Dependent on viewing direction, object geometry, highlights, direction of the illumination, intensity and color of the illumination.
XYZ	Y	N	Y	N	Dependent on viewing direction, object geometry, highlights, direction of the illumination, intensity and color of the illumination.
L*a*b*	Y	Y	N	N	Dependent on viewing direction, object geometry, highlights, direction of the illumination, intensity and color of the illumination.

2.2 Color Features



The most frequently cited visual features for image retrieval are color, texture, and shape. Among them, the color feature is most commonly used. It is robust to complex background and independent of image size and orientation. In this section, we will introduce two color features, color histogram, and color moment.

2.2.1 Color Histogram

Color Histogram is the most popular feature that is used in image retrieval. It is a global image feature. The advantages of color histogram are simple operation, invariant to translation, rotation, and scales of images.

A color histogram of an image counts the number of pixels with a given pixel value in red, green, and blue (RGB). To save storage and smooth out differences in similar but unequal images, the histogram is defined coarsely, with bins quantized to 8 bits, with 3 bits for each of red and green and 2 for blue.

2.2.2 Color Moment Primitives

The color distribution of an image can be characterized by color moments. Primitives of color moments [6] of an image can be viewed as a partial image feature. This feature contains local and global information of an image such that partly similar images will be retrieved.

The first color moment of the i -th color component ($i = 1,2,3$) is defined by

$$M_i^1 = \frac{1}{N} \sum_{j=1}^N p_{i,j} \quad (9)$$

where $p_{i,j}$ is the color value of the i -th color component of the j -th image pixel and N is the total number of pixels in the image. The h -th image moment, $h=2,3,\dots$, of the i -th color component is then defined as

$$M_i^h = \left(\frac{1}{N} \sum_{j=1}^N (p_{i,j} - M_i^1)^h \right)^{1/h} \quad (10)$$

Take the first H moments of each color component in an image s to form a feature vector, CT , which is defined as

$$\begin{aligned} CT &= [ct_1, ct_2, \dots, ct_z] \\ &= [\alpha_1 M_1^1, \alpha_1 M_1^2, \dots, \alpha_1 M_1^H, \alpha_2 M_2^1, \alpha_2 M_2^2, \dots, \alpha_2 M_2^H, \\ &\quad \alpha_3 M_3^1, \alpha_3 M_3^2, \dots, \alpha_3 M_3^H] \end{aligned} \quad (11)$$

where $Z=H*3$ and $\alpha_1, \alpha_2, \alpha_3$ are the weights for Y, I, Q.

Based on the above definition, an image is first divided into X non-overlapping blocks. For each block a , its h -th color moment of the i -th color component is defined by $M_{a,i}^h$. Then, the feature vector, CB_a , of block a is represented as

$$\begin{aligned} CB_a &= [cb_{a,1}, cb_{a,2}, \dots, cb_{a,Z}] \\ &= [\alpha_1 M_{a,1}^1, \alpha_1 M_{a,1}^2, \dots, \alpha_1 M_{a,1}^H, \alpha_2 M_{a,2}^1, \alpha_2 M_{a,2}^2, \dots, \alpha_2 M_{a,2}^H, \\ &\quad \alpha_3 M_{a,3}^1, \alpha_3 M_{a,3}^2, \dots, \alpha_3 M_{a,3}^H] \end{aligned} \quad (12)$$

From the above definition we can obtain X feature vectors. To speed up the image retrieval, we will find some representative feature vectors to stand for these feature vectors. To reach this aim, a progressive constructive clustering algorithm [7] is used to classify all CB_a s into several clusters and the central vector of each cluster is regarded as a representative vector and called the primitive of the image. The central vector, PC_k , of the k -th cluster is defined by

$$\begin{aligned} PC_k &= [pc_{k,1}, pc_{k,2}, \dots, pc_{k,Z}] \\ &= \frac{\sum_{j=1}^{n_k} CB_j^k}{n_k} = \left[\frac{\sum_{j=1}^{n_k} cb_{j,1}^k}{n_k}, \frac{\sum_{j=1}^{n_k} cb_{j,2}^k}{n_k}, \dots, \frac{\sum_{j=1}^{n_k} cb_{j,Z}^k}{n_k} \right] \end{aligned} \quad (13)$$

where $CB_j^k, j = 1, 2, \dots, n_k$, belongs to the k -th cluster and n_k is the size of the k -th cluster.

2.3 Morphological Operator

Mathematical morphology offers a powerful tool for extracting image components that

are useful in the representation and description of region shape, such as boundaries, skeletons, texture, and the convex hull. Mathematical morphology is a set-theoretic method. Sets in mathematical morphology represent the shapes of objects in an image. The operations of mathematical morphology were originally defined as set operations and shown to be useful for image processing.

In general, morphological approach is based upon binary images. In binary images, each pixel can be viewed as an element in Z^2 . Gray-scale digital images can be represented as sets whose components are in Z^3 , two components are the coordinates of a pixel, and the third corresponds to its discrete intensity value. The morphological operations input a source image and a structuring element which is another image usually smaller than source image. The structuring element is a predetermined geometric shape, and there are some common structuring elements as shown in Fig. 2-4.

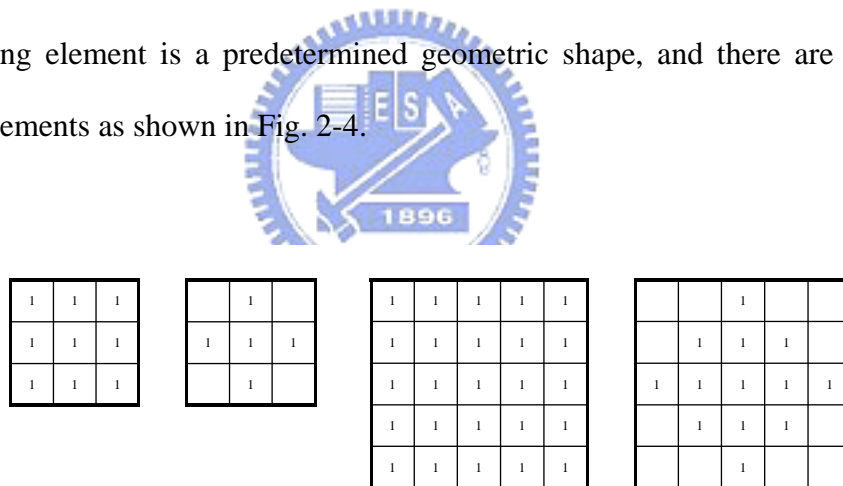


Fig. 2-4 Examples of structuring elements.

Here, we will discuss morphological operators in binary images [1, 7]. Given a source image A and a structuring element B in Z^2 , with components a and b respectively. And ϕ denotes the empty set.

2.3.1 Basic definition

The Translation of A by the point x in Z^2 , denoted $A_{\vec{x}}$, is defined by

$$A_{\vec{x}} = \{\vec{a} + \vec{x} \mid \forall \vec{a} \in A\} = A + \vec{x} \quad (14)$$

where the plus sign refers to vector addition.

And the Reflection of B , denoted \hat{B} , is defined as

$$B = \{-\vec{b} \mid \vec{b} \in B\} \quad (15)$$

The examples of Translation and Reflection are shown in Fig. 2-5.

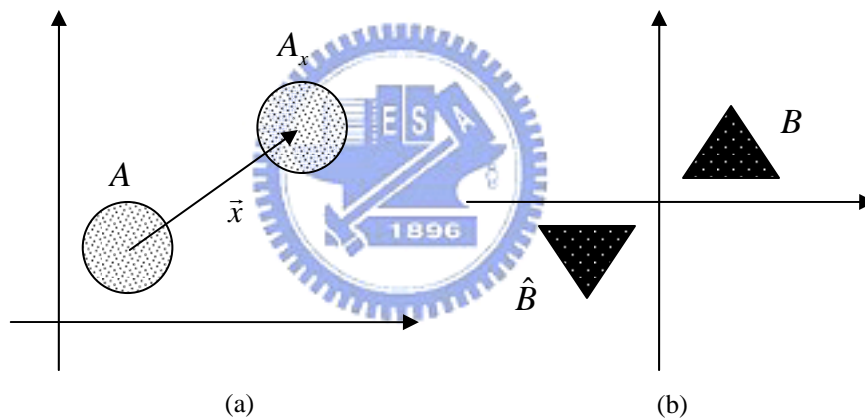


Fig. 2-5 Examples of (a) Translation and (b) Reflection.

2.3.2 Dilation, Erosion, Opening, and Closing of Binary Images

We introduce two of the fundamental morphology operations : Dilation and Erosion used in binary images, and two operators : Closing and opening that extended from Dilation and Erosion.

The Dilation of A by B , denoted $D_B(A)$, is defined as

$$D_B(A) = A \oplus \hat{B} = \{\bar{x} | \hat{B} + \bar{x} \cap A \neq \phi\} \quad (16)$$

where B is the structuring element.

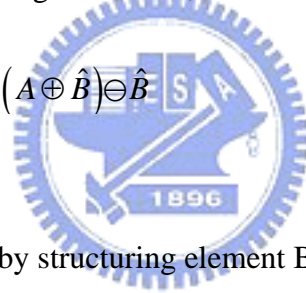
And the Erosion of A by B, denoted $E_B(A)$, is defined as

$$E_B(A) = A \ominus B = \{\bar{x} | B + \bar{x} \subset A\} \quad (17)$$

Fig. 2-6(c) and (d) are examples of Dilation and Erosion. The dilation of A by B is the set of all \bar{x} displacements such that \hat{B} and A overlap by at least one nonzero element. The erosion of A by B is the set of all points \bar{x} such that B translated by \bar{x} is contained in A.

The Closing of set A by structuring element B, denoted $C_B(A)$, is defined as

$$C_B(A) = A \bullet B = E_{\hat{B}}(D_B(A)) = (A \oplus \hat{B}) \ominus \hat{B} \quad (18)$$



And the Opening of set A by structuring element B, denoted $O_B(A)$, is defined as

$$O_B(A) = A \circ B = D_{\hat{B}}(E_B(A)) = (A \ominus B) \oplus B \quad (19)$$

The examples of closing and opening are shown in Fig. 2-6(e) and (f). The closing of A by B is simply the dilation of A by B, followed by the erosion of the result by \hat{B} . The opening of A by B is simply the erosion of A by B, followed by the dilation of the result by \hat{B} .

Another examples of closing and opening are shown in Fig. 2-7(b) and (c). From the figure, closing will connect thin lines and fill small holes and opening will leave objects larger than structuring element and subtract small thin lines.

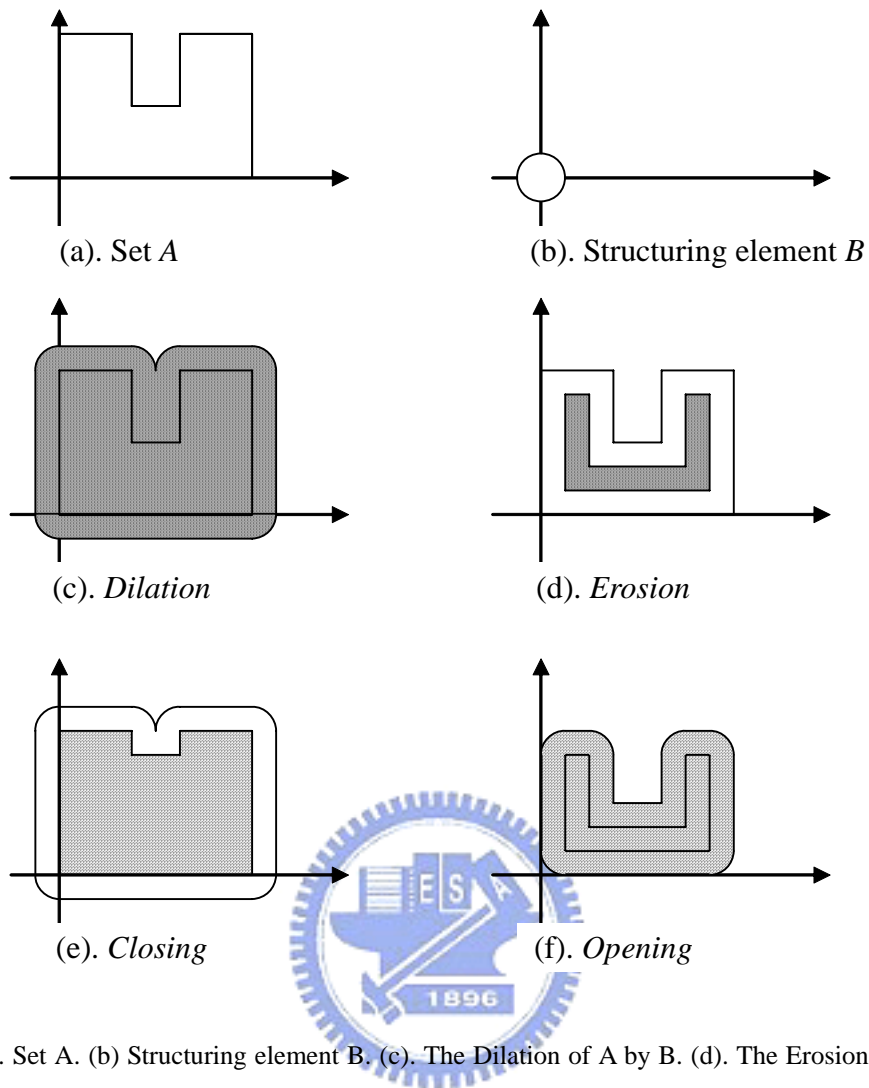


Fig. 2-6 (a). Set A. (b) Structuring element B. (c). The Dilation of A by B. (d). The Erosion of A by B. (e). The Closing of A by B. (f). The Opening of A by B.

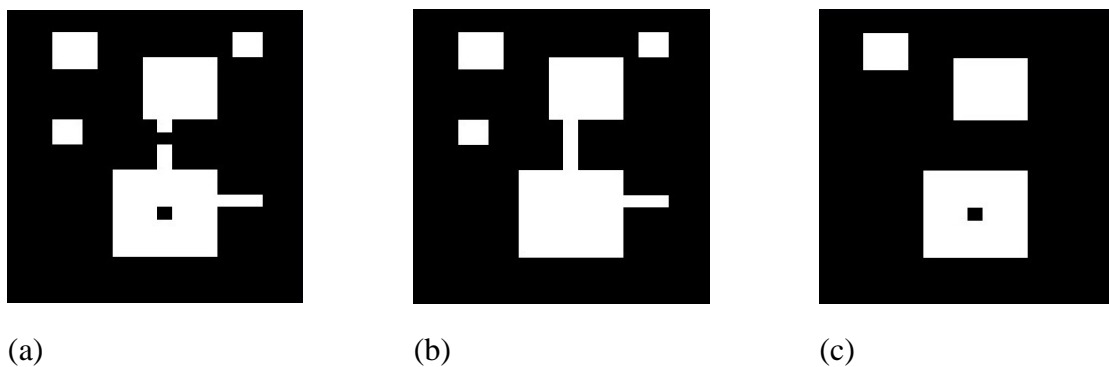


Fig. 2-7 (a) original image (b)(c) closing and opening with 3×3 structuring element, respectively.

2.3.3 Extension to Grayscale Images

In this section, we extend to gray-level images the basic operations of dilation, and erosion. Throughout the discussions that follow, we deal with digital image functions of the forms $f(x, y)$ and $b(x, y)$, where $f(x, y)$ is the gray-scale image and $b(x, y)$ is a structuring element.

Gray-Scale dilation of f by b , denoted by $f \oplus b$, is define as :

$$(f \oplus b)(s, t) = \max\{f(s-x, t-y) + b(x, y) \mid (s-x), (t-y) \in D_f; (x, y) \in D_b\} \quad (20)$$

where D_f and D_b are the domain of f and b , respectively. As before, b is the structuring element of the morphological process but note that b is now a function rather than a set. Because dilation is based on choosing the maximum value of $f + b$ in a neighborhood defined by the shape of the structuring element, the general effect of performing dilation on a gray-scale image is two-fold: (1) if all the values of the structuring element are positive, the output image tends to be brighter than the input; and (2) dark details either are reduced or eliminated, depending on how their values and shapes relate to the structuring element used for dilation. As illustrated in Fig 2-8 (b).

Gray-scale erosion of f by b , denoted by $f \ominus b$, is define as :

$$(f \ominus b)(s, t) = \min\{f(s+x, t+y) - b(x, y) \mid (s+x), (t+y) \in D_f; (x, y) \in D_b\} \quad (21)$$

where D_f and D_b are the domain of f and b . Because erosion is based on choosing the minimum values of $f - b$ in a neighborhood defined by the shape of the structuring element, the general effect of performing erosion on a gray-scale image is two-fold: (1) if all the values of the structuring element are positive, the output image tends to be darker than the input; and

(2) bright details are reduced or eliminated, depending on the used structuring element. As illustrated in Fig 2-8 (c).



Fig. 2-8 (a) The original of Lena image (b) Dilation of Lena image (c) Erosion of Lena image.

The usage of closing and opening is to smooth contours of objects. In gray-level images, opening is used for darker object objects with brighter background, and closing is used for brighter objects with darker background. As illustrated in Fig 2-9 (b) and (c).



Fig. 2-9 (a) The original of Lena image (b) Closing of Lena image (c) Opening of Lena image.

2.3.4 Morphological Gradient

The Morphological Gradient of an image, denoted G :

$$G = D_B(A) - E_B(A) = (A \oplus B) - (A \ominus B) \quad (22)$$

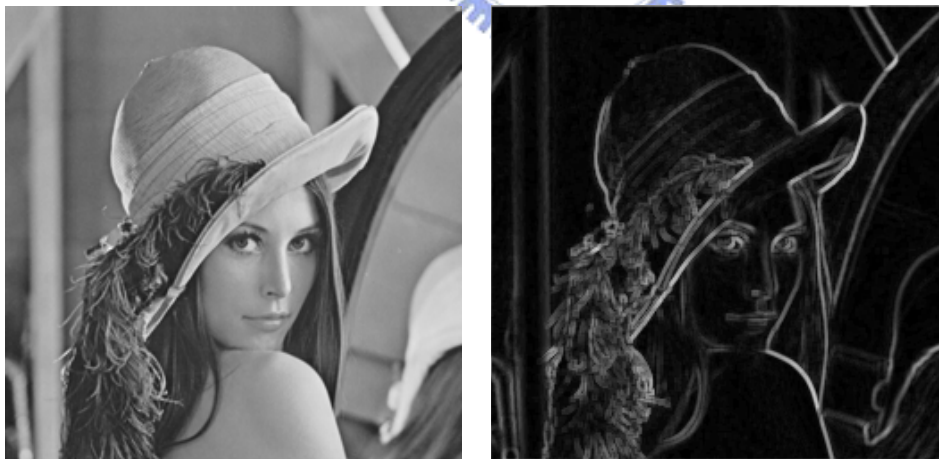
or

$$G = D_B(A) - A = (A \oplus B) - A \quad (23)$$

or

$$G = A - E_B(A) = A - (A \ominus B) \quad (24)$$

The morphological gradient highlights sharp gray-scale transitions in the source image. In other words, morphological can extract the boundary of an object. However, the morphological gradient is sensitive to the shape of the chosen structuring element. Here are examples of dilation, erosion, closing, opening, and morphological gradient of the gray-scale image, Lena, with the 3x3 structuring element as shown in Fig 2-10.



(a)

(b)

Fig. 2-10 (a) The original of Lena image (b) Morphological gradient of Lena image

2.3.5 Morphological Granulometry

Granulometries were introduced first by G. Matheron [8], and have been proven to be very useful in image analysis, finding application in image tasks such as texture classification, pattern analysis, and image segmentation. A granulometry can simply be defined as series of morphological openings with structuring elements of increasing sizes. The granulometric function maps each structuring element size to the number of image pixels removed during the opening operation with the corresponding structuring element. It is easy to extend the binary granulometry into grayscale form by replacing the binary opening with a grayscale opening.

There are three kinds of granulometric function [9]:

- 1 Number of particles of $\gamma_{S_k}(f)$.
- 2 Surface area of $\gamma_{S_k}(f)$, often called size distribution. This measure is typically chosen to be the area (number of ON pixels) in the binary case, and the volume (sum of all pixel values) in the grayscale case.
- 3 Loss of surface area between $\gamma_{S_k}(f)$ and $\gamma_{S_{k-1}}(f)$, called the granulometric curve, or pattern spectrum.

where γ is the binary morphological opening, f represents the original image, and $S_k, k = 1, 2, \dots$, is a sequence of opening structuring elements of increasing size.

2.3.6 Morphological Primitives

This feature is proposed by J-S Wu [10]. Before the morphological operation, first we transform RGB color model into YIQ model for the sack of human perception. Then we down-sample each image to remove the noise. Let $f(x, y) = (Y_{xy}, I_{xy}, Q_{xy})$ denote the color

value of the down-sampled image located at (x, y) . We define the following 9-tuple vector as the morphological context of a pixel located at (x, y) :

$$C_{x,y} = [c_{x,y}^1, c_{x,y}^2, \dots, c_{x,y}^9] \quad (25)$$

$$c_{x,y}^1 = D_b(Y_{x,y}) \quad c_{x,y}^2 = D_b(I_{x,y}) \quad c_{x,y}^3 = D_b(Q_{x,y})$$

$$c_{x,y}^4 = E_b(Y_{x,y}) \quad c_{x,y}^5 = E_b(I_{x,y}) \quad c_{x,y}^6 = E_b(Q_{x,y})$$

$$c_{x,y}^7 = G_b(Y_{x,y}) \quad c_{x,y}^8 = G_b(I_{x,y}) \quad c_{x,y}^9 = G_b(Q_{x,y})$$

where b is a structuring element, $D_b(Y_{x,y}), D_b(I_{x,y}), D_b(Q_{x,y})$ are the dilation of Y, I, Q color channels respectively; $E_b(Y_{x,y}), E_b(I_{x,y}), E_b(Q_{x,y})$ are the erosion of Y, I, Q color channels respectively; $G_b(Y_{x,y}), G_b(I_{x,y}), G_b(Q_{x,y})$ are the morphological gradient of Y, I, Q color channels respectively.



From the definition above, we get both the color and shape information of a pixel within the filter. We use large morphological gradient value to identify an edge pixel, and a small one to a pixel on a flat region. Then we get the color information in the neighborhood of the edge pixels. Therefore, two pixels with the same color value might not have the same context.

After morphological context extraction, there will be many similar contexts. So we use the following algorithm to cluster these contexts to small number of contexts which are more representative.

1. Initially, we extract the first pixel's context, and then this context is viewed as the center of the first cluster.
2. Extract next pixel's context, $C_{x,y}$, find the nearest cluster center $MP_{nearest}$ from

currently existing clusters. If the nearest distance is smaller than a threshold, T_d , then assign this context to the nearest cluster and update $MP_{nearest}$. Otherwise, construct a new cluster that the cluster center is $C_{x,y}$

3. Repeat .step2 until all pixels are processed.

In the above algorithm, $MP_{nearest}$ is defined by

$$\begin{aligned}
 MP_k &= [mp_k^1, mp_k^2, \dots, mp_k^9] \\
 &= \frac{\sum_{j=1}^{n_k} C_{x,y}^j}{n_k} \\
 &= \left[\frac{\sum_{j=1}^{n_k} c_{x,y}^1}{n_k}, \frac{\sum_{j=1}^{n_k} c_{x,y}^2}{n_k}, \dots, \frac{\sum_{j=1}^{n_k} c_{x,y}^9}{n_k} \right]
 \end{aligned} \tag{26}$$

where $C_{x,y}^k, j = 1, \dots, n_k$, is the context located at (x, y) and belongs the k -th context cluster.

Moreover, the Euclidean distance between $C_{x,y}$ and MP_k is defined as follows:

$$dis(C_{x,y}, MP_k) = \sqrt{\sum_{i=1}^9 (c_{x,y}^i - mp_k^i)^2} \tag{27}$$

2.4 Distance Functions for Similarity Measure

For measuring perceptual similarity between images, many distance functions have been used. Here we will introduce Minkowski-like metrics [4] and Dynamic Partial Function [5].

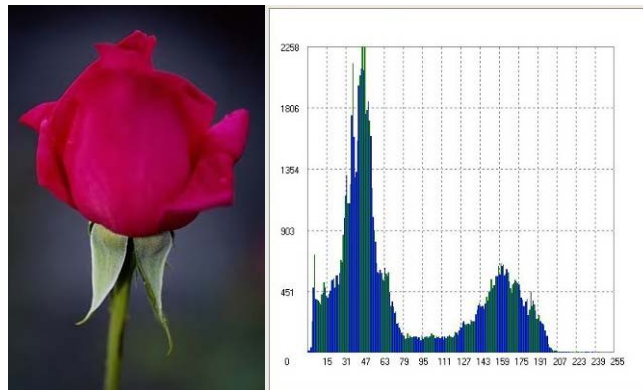
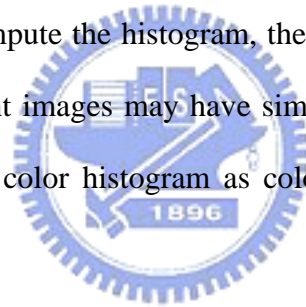
2.4.1 Intersection

Histogram intersection is the standard measure used for color histograms. First, a color histogram H_i is generated for each image i in the database. Next, each histogram is normalized, so that its sum equals unity. After this step, it effectively removes the size of image. The histogram is then stored in the database. For any selected model image, its histogram H_m is intersected with all database image histograms H_i , according to the equation

$$\text{intersection} = \sum_{j=1}^n \min(H_i^j, H_m^j) \quad (28)$$

where j denotes histogram bin j , with each histogram having n bins. The closer intersection value is to 1, the better the image match.

Although, it is fast to compute the histogram, the intersection value is sensitive to color quantization. And two different images may have similar color histograms as shown in Fig. 2-11. Therefore, only use the color histogram as color feature of image is not enough for image retrieval.



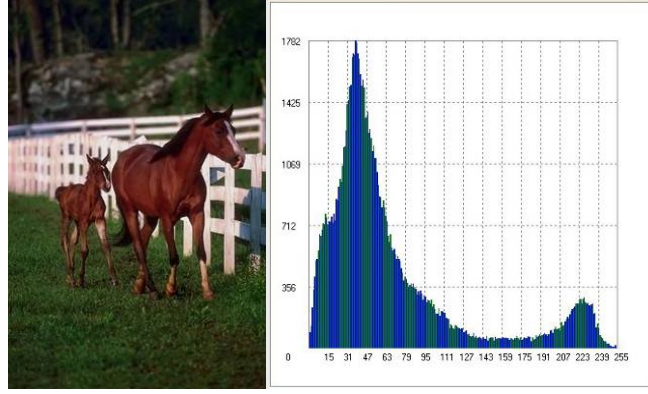


Fig. 2-11 Two different images with similar color histograms.

2.4.2 Primitive Similarity

First, we provide several definitions. The k -th primitive of a query image q is represented as:

$$PC_k^q = [pc_{k,1}^q, pc_{k,2}^q, \dots, pc_{k,Z}^q], \text{ where } k = 1, 2, \dots, m, \quad (29)$$

and m is the number of primitives in the query image. The λ -th primitive of a matching image s is denoted as

$$PC_\lambda^s = [pc_{\lambda,1}^s, pc_{\lambda,2}^s, \dots, pc_{\lambda,Z}^s] \quad (30)$$

The distance between PC_k^q and PC_λ^s is defined as follows:

$$D_{-}PC_{k,\lambda}^{q,s} = \sqrt{\sum_{i=1}^Z (pc_{k,i}^q - pc_{\lambda,i}^s)^2} \quad (31)$$

The minimum distance between $PC_k^{q,s}$ and all primitives of s is defined by

$$D_PC_k^{q,s} = \min_{\lambda} (D_PC_{k,\lambda}^{q,s}) \quad (32)$$

The distance between the query image q and the matching image s is defined by

$$D_PC^{q,s} = \sum_{k=1}^m n_k^q \times D_PC_k^{q,s} \quad (33)$$

where n_k^q is the size of the k-th cluster. The similarity measure between q and s is defined as

$$Sim^{q,s} = \frac{1}{D_PC^{q,s}} \quad (34)$$

Note that the larger $Sim^{q,s}$ a matching image has, the more similar it is to the query image.

2.4.3 Minkowski-Like Metrics

Each image can be represented by a p-dimensional feature vector (x_1, x_2, \dots, x_p) ,

where each dimension represents an extracted feature. $\Delta d_i = |x_i - y_i|$, $i=1,2, \dots, p$ is defined

as the feature distance in feature channel i . Minkowski-like distance functions are commonly used to measure similarity. The distance between any two images, X and Y, is defined as

$$D(X, Y) = \left(\sum_{i=1}^p \Delta d_i^r \right)^{1/r} \quad (35)$$

When $r=1$, above distance is the City-block or L1 distance. When we set $r=2$, it is the Euclidean or L2 distance. And, when r is a fraction, it defines a fractional function.

2.4.4 Dynamic Partial Function

To disagree with Minkowski-like metric, DPF does not consider all features for similarity measurement. If we define the similar data set

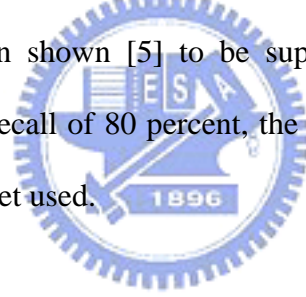
$$\Delta_m = \text{The smallest } m \text{ } \Delta d_i \text{'s of } (\Delta d_1, \Delta d_2, \dots, \Delta d_p) \quad (36)$$

then the dynamic partial distance between two image X and Y, is defined as

$$D(X, Y) = \left(\sum_{\Delta d_i \in \Delta_m} \Delta d_i^r \right)^{1/r} \quad (37)$$

where m and r are tunable parameters. When $m = p$, DPF degenerates to a Minkowski-like metric. When $m < p$, it counts the smallest m feature distances between two images.

Studies have shown [5] that the Minkowski approach does not model human perception. The performance of DPF has been shown [5] to be superior to that of widely-used distance functions, For instance, for a recall of 80 percent, the retrieval precision of DPF is shown to be 84 percent for the test data set used.



2.4.5 The DPF Enhancement

2.4.5.1 Thresholding Method

Here, we select the features that have a difference smaller than θ belong to the similar feature set A_θ . The Thresholding distance of image pair (X,Y) is then defined as

$$D(X, Y) = \left(\frac{1}{|A_\theta|} \sum_{\Delta d_i \in A_\theta} \Delta d_i^r \right)^{1/r} \quad (38)$$

2.4.5.2 Sampling Method

The main idea of Sampling-DPF is to try different m settings simultaneously and to guess that some of the sampling points can be near the optimal m . Suppose we sample N settings of m , denoted as m_n , $n = 1, \dots, N$. At each sampling m_n , a ranking list, denoted as $R_n(\Phi, X)$, is generated. Finally, the sampling method obtains N ranking lists and produces the final ranking, denoted as $R_r(\Phi, X)$.

Two methods are suggested, one is using the smallest rank as the final rank (SR). The other is using the average distance (or rank) as the final distance (AR). The AR distance is defined as

$$D_{AR}(X, Y) = \frac{1}{N} \sum_{n=1}^N \left(\sum_{\Delta d_i \in \Delta_m} \Delta d_i^r \right)^{1/r} \quad (39)$$

For the Sampling-Thresholding method, its AR distance is defined as

$$D(X, Y) = \frac{1}{N} \sum_{n=1}^N \left(\frac{1}{|A_{\theta_n}|} \sum_{\Delta d_i \in A_{\theta_n}} \Delta d_i^r \right)^{1/r} \quad (40)$$

2.4.5.3 Weighting Method

The Weighting method can be used with Sampling-DPF, Sampling-Thresholding, or other methods. It considers that different types of features can have different degrees of importance for measuring similarity. Let ω_i denote the weight of the i th similar feature. In the Weighted-Sampling-DPF method, the distance of image pair (X, Y) is defined as

$$D(X, Y) = \frac{1}{N} \sum_{n=1}^N \left(\sum_{\Delta d_i \in \Delta_m} \omega_i \Delta d_i^r \right)^{1/r} \quad (41)$$

and in the Weighted-Sampling-Thresholding method, it is defined as

$$D(X, Y) = \frac{1}{N} \sum_{n=1}^N \left(\frac{1}{|A_{\theta_n}|} \sum_{\Delta d_i \in A_{\theta_n}} \omega_i \Delta d_i^r \right)^{1/r} \quad (42)$$



Chapter 3

The Proposed Method

Overview of our retrieval system is described in section 3.1. In section 3.2, we will introduce a new feature, granulometric histogram, which represents the size distribution of an image. Its similarity is described in section 3.3.1. We also introduce the smallest DPF to be the similarity of Morphological Primitive and describe it in section 3.3.2.

3.1 Overview

Fig-3.1 shows the architecture of our retrieval system. Within it, processing and similarity measure are the key components. There are many processing methods for the purpose of image retrieval. In this thesis, we employ morphological primitive extraction, color moment primitive extraction, morphological granulometric distribution, and morphological granulometric histogram. We plot them in Fig-3.2.

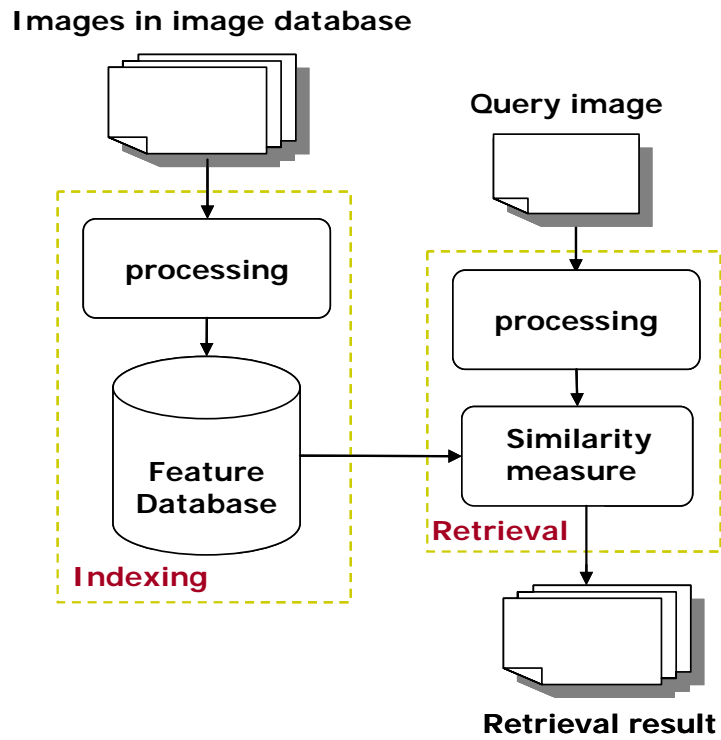


Fig. 3-1 Architecture of our retrieval system.

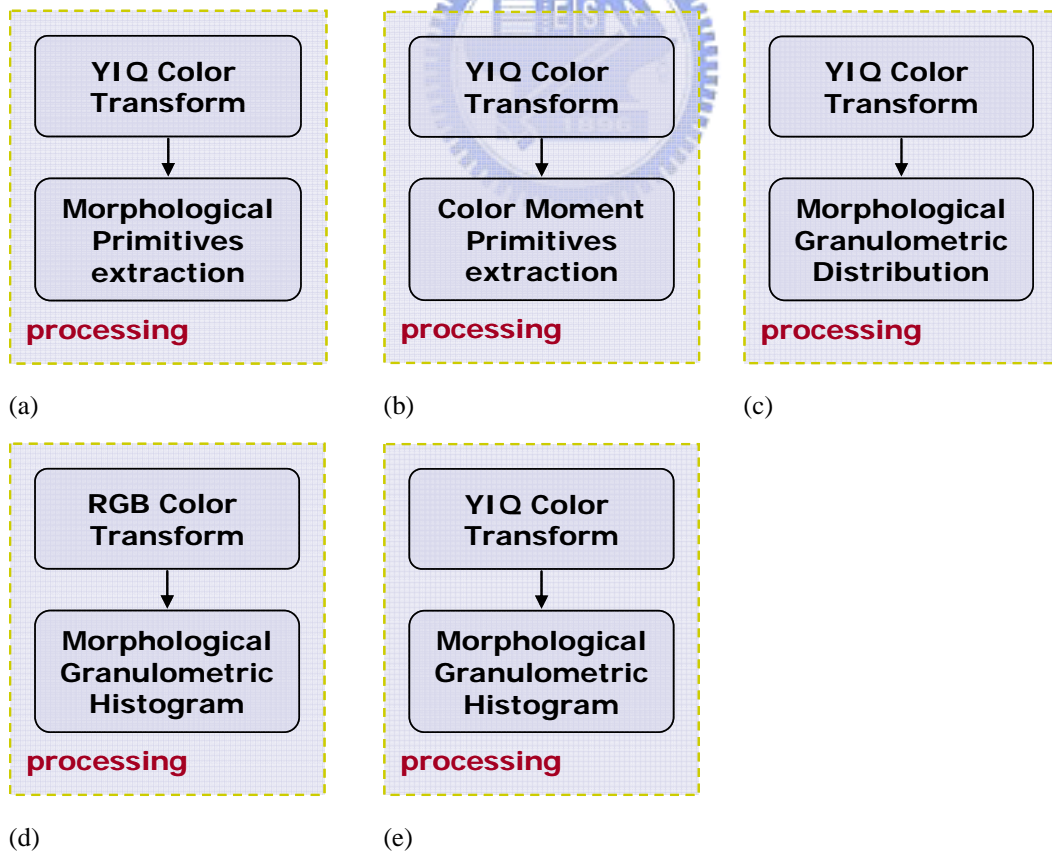


Fig. 3-2 (a)-(e) Different processing methods.

3.2 Morphological Granulometry

3.2.1 Distribution of Morphological Granulometry

Before applying morphological granulometry, we transform the down-sampled image from RGB to YIQ color model. Assume $C_{x,y} = (Y_{x,y}, I_{x,y}, Q_{x,y})$ be the pixel value at location (x, y) of the down-sampled image, and $B = \{b_1, b_2, \dots, b_L\}$ be a set of structuring elements with increasing sizes. According to section 2.3.5, we have chosen the pattern spectrum as the granulometry function.

First, we do a series of morphological openings, denoted $O_{b_i}(Y), O_{b_i}(I), O_{b_i}(Q)$, to each component of YIQ color space with structuring element $b_i, i = 1, \dots, L$. Therefore, we obtain $M = 3 * L$ openings for a single image.

Second, we compute the distribution of morphological granulometry as follows:

```
for(int i = 1; i <= L; i++)
{
     $DY_i = O_{b_{i-1}}(Y) - O_{b_i}(Y)$ 
     $DI_i = O_{b_{i-1}}(I) - O_{b_i}(I)$ 
     $DQ_i = O_{b_{i-1}}(Q) - O_{b_i}(Q)$ 
}
```

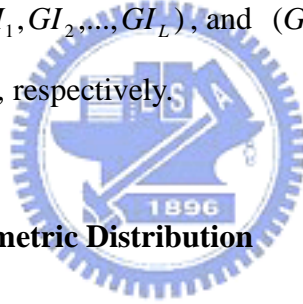
Here, $O_{b_0}(Y)$, $O_{b_0}(I)$, and $O_{b_0}(Q)$ represent Y , I , and Q , respectively.


```

for(int i = 1; i <= L; i++)
{
    GYi = 0;
    GIi = 0;
    GQi = 0;
    for all (x, y)
    {
        if (DYi(x, y) > 0) GYi++;
        if (DIi(x, y) > 0) GIi++;
        if (DQi(x, y) > 0) GQi++;
    }
    GYi / = number of pixles in Y
    GIi / = number of pixles in I
    GQi / = number of pixles in Q
}

```

Then, $(GY_1, GY_2, \dots, GY_L)$, $(GI_1, GI_2, \dots, GI_L)$, and $(GQ_1, GQ_2, \dots, GQ_L)$ are size distributions of Y , I , and Q , respectively.



3.2.2 Primitives of Granulometric Distribution

As morphological primitive mentioned in 2.3.6, we first transform RGB color model into YIQ model, then we down-sample each image. Let $f(x, y) = (Y_{xy}, I_{xy}, Q_{xy})$ denote the color value of the down-sampled image located at (x, y) . We define the following Z-tuple vector as the morphological granulometry context of a pixel located at (x, y) :

$$C_{x,y} = [c_{x,y}^1, c_{x,y}^2, \dots, c_{x,y}^Z] \quad (43)$$

$$c_{x,y}^1 = Y_{x,y} \quad c_{x,y}^2 = I_{x,y} \quad c_{x,y}^3 = Q_{x,y}$$

$$c_{x,y}^4 = O_{b_1}(Y_{x,y}) - Y_{x,y} \quad c_{x,y}^5 = O_{b_1}(I_{x,y}) - I_{x,y} \quad c_{x,y}^6 = O_{b_1}(Q_{x,y}) - Q_{x,y}$$

where $B = \{b_i \mid i = 1, \dots, L\}$ is a set of structuring elements with increasing size and

$Z = 3 * (1 + L)$. If $Z = 6$. $Y_{x,y}, I_{x,y}, Q$ are the pixel value of Y, I, Q color channels respectively; $O_b(Y_{x,y}), O_b(I_{x,y}), O_b(Q_{x,y})$ are the openings of Y, I, Q color channels respectively. Consequently, $c_{x,y}^4, c_{x,y}^5, c_{x,y}^6$ represent the pattern spectrum.

3.2.3 Histogram of Morphological Granulometry

The granulometric histograms are obtained by the following four steps.

First, we down-sampled the target image to remove the noise and reduce the computation time. Suppose $C_{x,y} = (R_{x,y}, G_{x,y}, B_{x,y})$ represents the pixel value at location (x, y) of the down-sampled image.



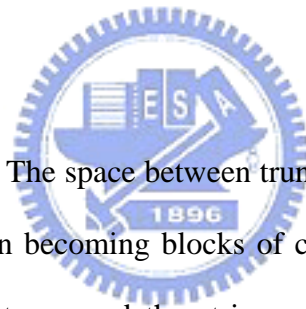
Second, let $B = \{b_1, b_2, \dots, b_L\}$ be a set of structuring elements with increasing sizes. We do a series of morphological openings, denoted $O_{b_i}(R), O_{b_i}(G), O_{b_i}(B)$, to each component of color space with structuring element $b_i, i = 1, \dots, L$. Therefore, we obtain $M = 3 * L$ openings for a single image.

Third, we count the histogram $H_j, j = 1, \dots, M$ of each opening, says the granulometric histograms. H_1, H_2, \dots, H_L are histograms of $O_{b_j}(R), j = 1, \dots, L$, respectively, $H_{L+1}, H_{L+2}, \dots, H_{2*L}$ are histograms of $O_{b_j}(G), j = 1, \dots, L$, respectively, $H_{2*L+1}, H_{2*L+2}, \dots, H_{3*L}$ are histograms of $O_{b_j}(B), j = 1, \dots, L$, respectively.

Fourth, we normalize the histogram. So $H_j = H_j / num$, $j = 1, \dots, M$ where $num = \sum_{i=0}^{255} H_j^i$ represents the size of down-sampled image. After this step, a histogram count turns into a probability of that color bin. Obviously, it removes the size of an image.

Morphological opening or closing could enhance, suppress or smooth some areas. Opening operation of a particular size of structuring element has the most effect on regions of the input image that contain particles of that size.

In RGB color space, it is easy to treat similar colors as different colors and then similar objects become different objects. Opening of different image may be similar under RGB color space.



Take Fig.3-3 for example. The space between trunks of trees and space separated by legs of horses in Fig.3-3(a) are then becoming blocks of colors after opening. The chinks in the leaves, color changing in the trees, and the stripes of the bus are causing blockings after opening. We may see those blocks in Fig.3-4 and Fig.3-5. These phenomenon leads to false retrieval results.

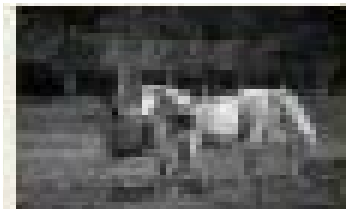


(a)

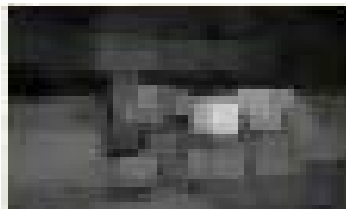


(b)

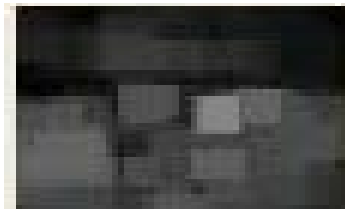
Fig. 3-3 Example of different image that has similar opening.



(a) R component



(b) opening of (a) with b_1



(c) opening of (a) with b_2



(d) G component



(e) opening of (d) with b_1



(f) opening of (d) with b_2



(g) B component



(h) opening of (g) with b_1



(i) opening of (h) with b_2

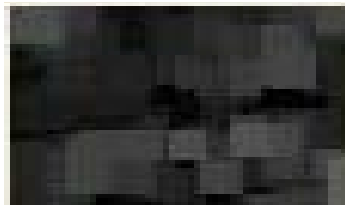
Fig. 3-4 Openings of Fig.3-3(a).



(a) R component



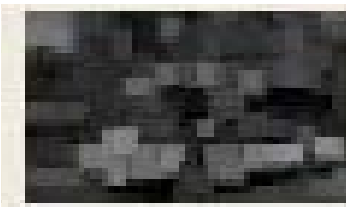
(b) opening of (a) with b_1



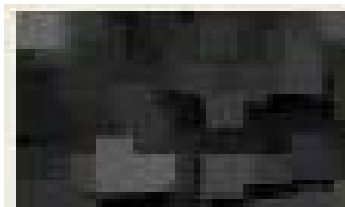
(c) opening of (a) with b_2



(d) G component



(e) opening of (d) with b_1



(f) opening of (d) with b_2

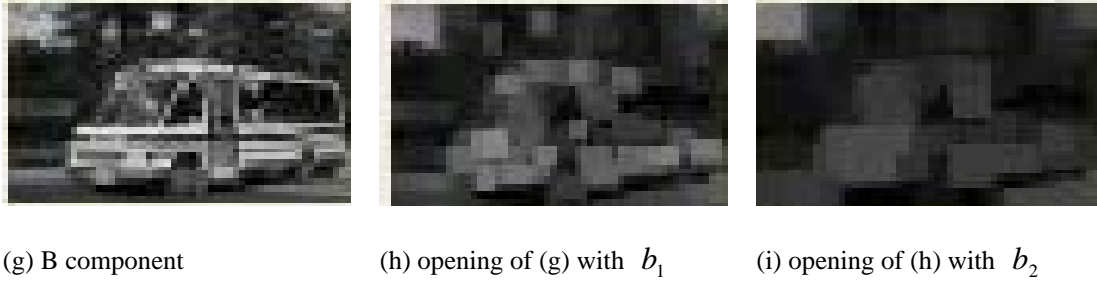
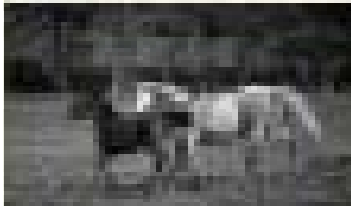


Fig. 3-5 Openings of Fig.3-3(b).

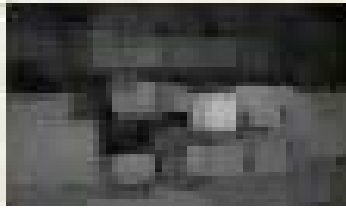
We improve this problem by changing RGB into YIQ color model for the sake of human perception and linear transform. In this color space, the false segmentation of one object with some similar colors can be avoided.

Use the same images mentioned above for example. The openings of Fig.3-3 under YIQ are listed in Fig-3.6 and Fig-3.7. Only Y component, the luminance part has similar blocking effect as which under RGB color model. We can see the blur shape of horse and bus with light gray in I and Q components.

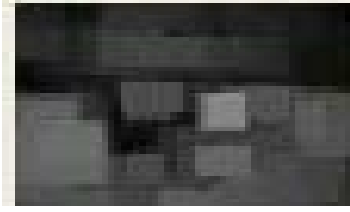
Therefore, the granulometry histogram gives more information about size of objects than luminance information of objects. If the size of objects in query image and that of matching images are closed, they will be retrieved as similar images. Since the histogram has been normalized, it is independent of image size. It depends only if the ratio of object size and image size are very different. That is, this histogram says dissimilar when the same object in different images has large variations in ratio of object size and image size, or says similar when different objects in different images have the same ratio of object size and image size.



(a) Y component



(b) opening of (a) with b_1



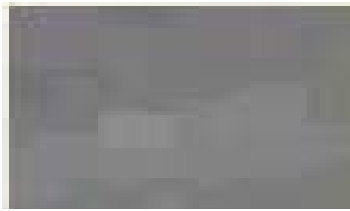
(c) opening of (a) with b_2



(d) I component



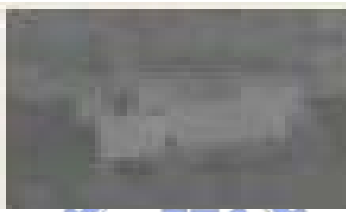
(e) opening of (d) with b_1



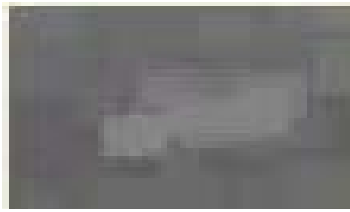
(f) opening of (d) with b_2



(g) Q component

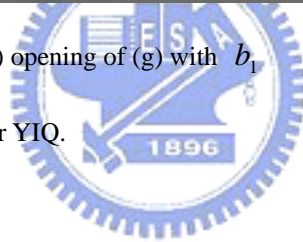


(h) opening of (g) with b_1



(i) opening of (h) with b_2

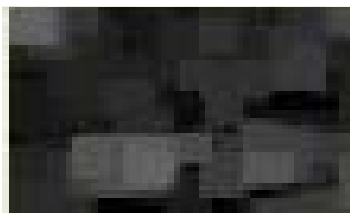
Fig. 3-6 Openings of .Fig.3-3(a) under YIQ.



(a) Y component



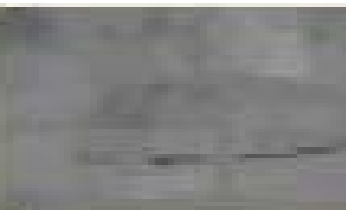
(b) opening of (a) with b_1



(c) opening of (a) with b_2



(d) I component



(e) opening of (d) with b_1



(f) opening of (d) with b_2



(g) Q component

(h) opening of (g) with b_1

(i) opening of (h) with b_2

Fig. 3-7 Openings of Fig.3-3(b) under YIQ.

3.3 Similarity

3.3.1 Similarity for Morphological Granulometry

We use absolute difference as the similarity measure for granulometric distribution. Therefore, similarity of granulometric distribution equals the sum of M absolute difference between target and matching images.

Besides, we choose intersection as the similarity measure for granulometric histogram. Hence, similarity of granulometric histogram equals M minus the sum of M histogram intersection.

3.3.2 DPF for Primitives

Here, we propose a new similarity measure for primitives, the DPF. The distance between two N -tuple primitive vectors is the sum of the smallest β bins where $\beta < N$. This idea comes from some problem of thresholding method. It is difficult to set a fixed threshold that is suitable for every image. And once a threshold T has been set, it is possible that none of the distance within the bins of the feature vectors is less than T . To make sure at least one feature bin has been considered, we choose some smallest bin distance as the important features.

Follow the definitions of primitive similarity in section 2.4.2. We only update the distance between PC_k^q and PC_λ^s into the following equation:

$$D_{-}PC_{k,\lambda}^{q,s} = \sqrt{\sum_{j=1}^{\beta} d_j^2} \quad (44)$$

where d_j is the smallest j th element of $\{pc_{k,i}^q - pc_{\lambda,i}^s \mid i = 1, \dots, N\}$. The rest are the same as the primitive similarity.



Chapter 4

Experimental Results

In this chapter, we will present some experimental results obtained by applying the proposed methods. We also compare the results with color histogram, morphological primitives [10] and color moment primitives [6].

4.1 Experimental environment

The 1000 test images are selected from Corel's database, and classified to 10 types of models, beach, buildings, buses, dinosaurs, elephants, flowers, horses, underwater worlds, and foods. Each type contains 100 images. Fig. 4-1 show some examples of each type from the database. The image size is either 384*256 or 256*384 pixels.

Images have similar object size in each of type 1, 4, 5, 6, 7, 8, and 9, respectively. Only type 1, 5, 9 have clean background. We define that an image has a clean background for there isn't too many variant colors or objects inside and with a large area of some similar colors. Objects in images of type 2, 3, 10 are very complicated. But every image has blue sky and ocean in type 2, has buildings in type 3, and has plates in type 10. Some images in type 3 also have blue sky.

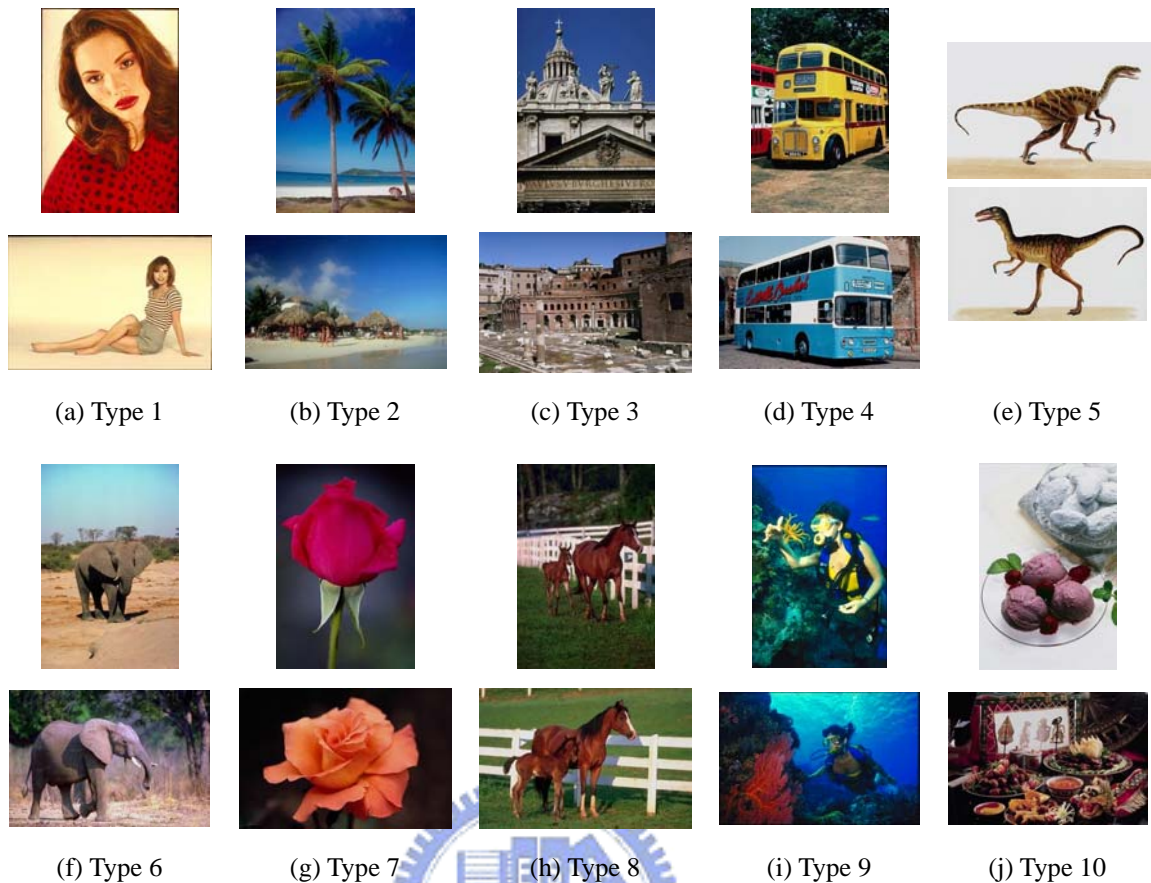


Fig. 4-1 Examples of the image database.

The proposed methods are implemented on PC with Pentium4 2.8GHz, RAM 1G. The operating system is Microsoft Windows XP SP2. The program was developed in the C++ language and compiled under Borland C++ Builder version 6.0. Our feature database is built in MySQL for the sake of its combination with PHP in the network.

The first experiment compares the results of morphological primitives (MP) using smallest DPF and original primitive similarity measure. The second experiment shows the results of morphological granulometric distribution (GD) and its primitives (GDP). The third experiment demonstrates the results of the granulometry histogram under RGB (GHrgb) and YIQ (GHyiq) color model. The fourth experiment compares granulometry histogram, granulometric distribution, primitives of granulometric distribution, morphological primitives

(MP), primitives of color moments (CMP), and color histogram (CH). And finally, we introduce the fifth experiment to verify the robustness of our image retrieval system.

4.2 Experimental results

We use precision and recall to measure our image retrieval system. The definitions of them are described in the following text. Assume that R is the number of retrieved relevant images, T is the total number of relevant images and K is the number of retrieved images. Then,

$$\text{Precision} = \frac{R}{K} \quad (45)$$

$$\text{Recall} = \frac{R}{T} \quad (46)$$

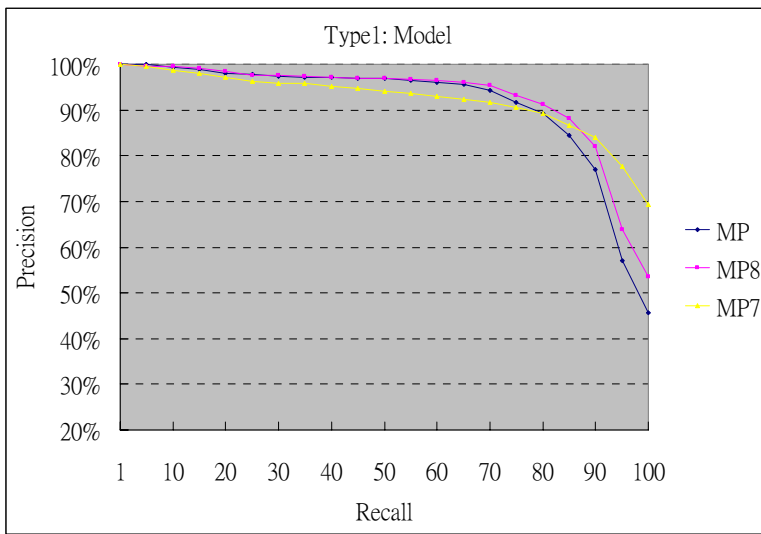


4.2.1 Morphological Primitives

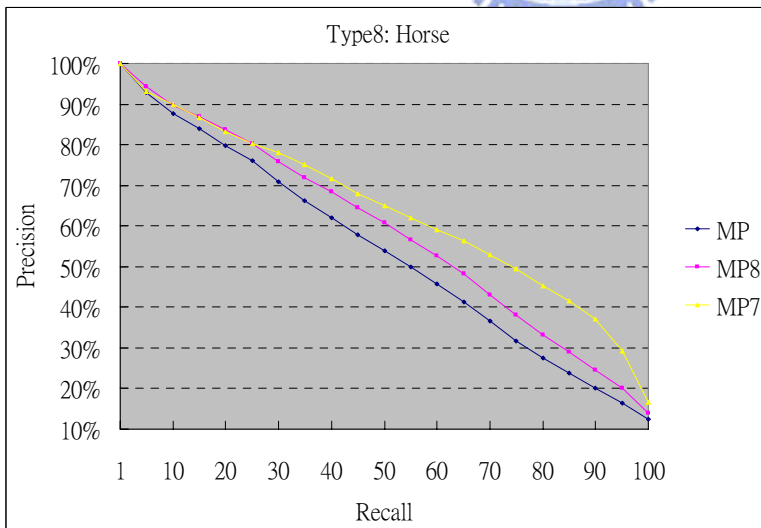
Morphological Primitive using smallest-8 DPF has improved the precision of all types in image database except type5 dinosaur. In Fig.4-2, MP means MP using similarity in section 2.4.2, MP8 and MP7 means MP using the smallest-8 and smallest-7 DPF, respectively. We only show the results of three types (model, dinosaur, and horse) in Fig.4-2.

In type1 model, the curve of MP7 is slightly lower than the other two curves when recall is less than 75, but larger than the other two when recall is larger than 75. When retrieve for the top 97, precisions of MP, MP8, and MP7 are 38%, 50%, 70%, respectively. Precision of MP7 for top 97 increases 32% to compare with MP. In type5 dinosaur, the curve of MP8 and MP7 are both lower than that of MP.

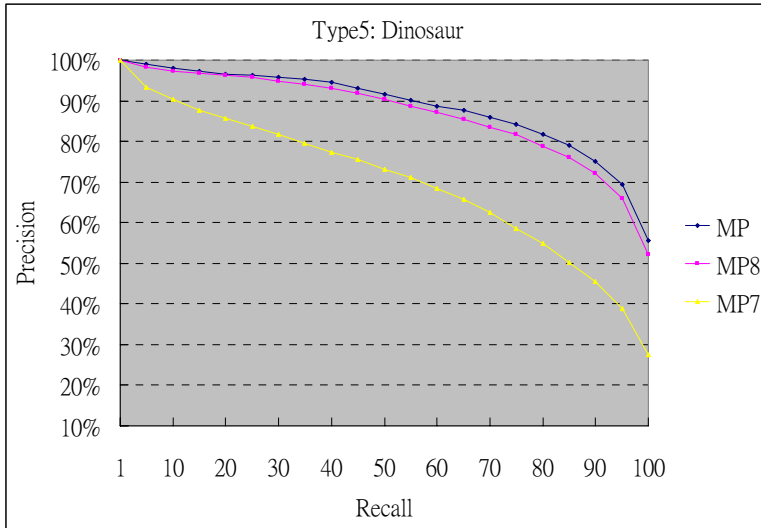
The consistency of regular object size and clean background upgrades the precision-recall curve. Although type1 and type5 both have clean background, models have two object sizes while dinosaurs have only one. There is space for improvement in type1 not in type5. That's why two types with similar options are not both upgraded when using DPF.



(a)



(b)



(c)

Fig. 4-2 The Precision-Recall curves when using morphological primitives. (a) The results of Type1 (b) The results of Type7 (c) The results of Type5.

4.2.2 Granulometric Distribution and its Primitives

We use five different numbers of sets of structuring elements to experiment on granulometric distribution. The precision-recall curves are very close with each other. Consequently, we choose the set of seven structuring elements to attain the average precision and take less computation time. Then, we try on its primitives with $Z = 6,9$. On average, precisions of GDPs are better than those of GDs, but all lower than those of GHs (see next section). Precision curves in Fig. 4-3 show that GH has the best performance.

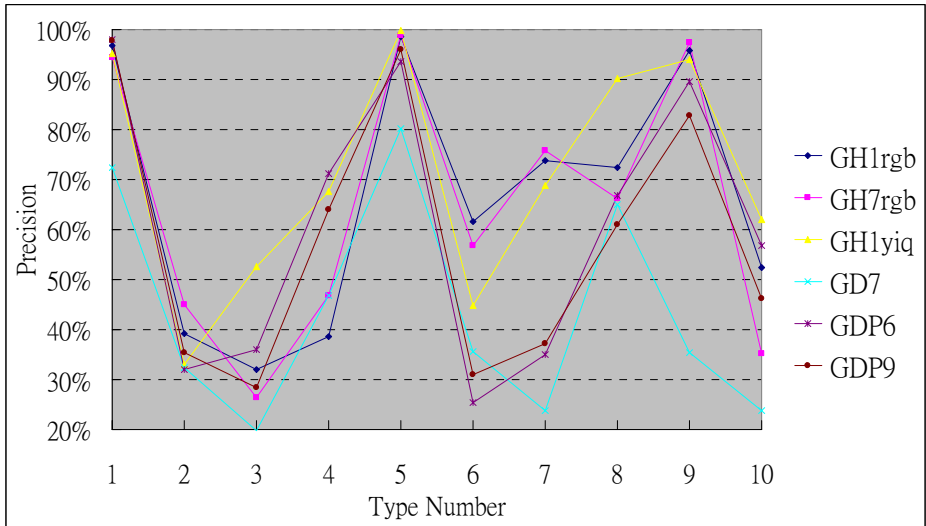
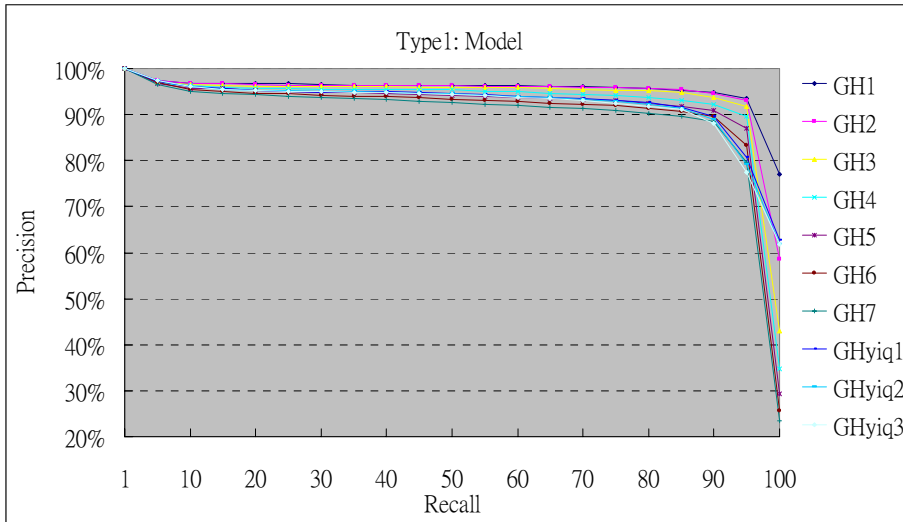


Fig. 4-3 The precision comparison for each type of image database among the proposed methods, granulometric histogram, granulometric distribution, and primitives of granulometric distribution (R=20).

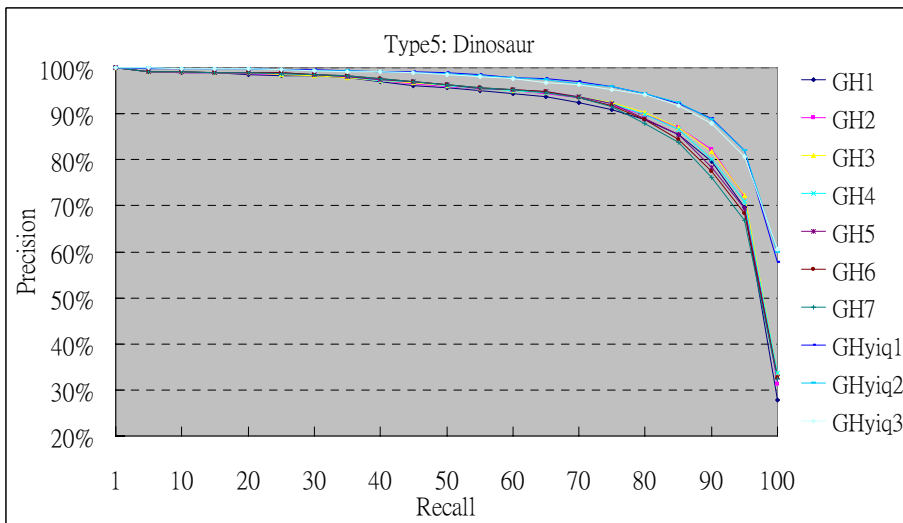
4.2.3 Granulometric Histogram

We suppose $M = 1, 2, \dots, 7$ for GH under RGB, and $M = 1, 2, 3$ for GH under YIQ. In Fig. 4-4, GH_k , $k = 1, 2, \dots, 7$ means GH under RGB with $M = k$, and GH_{yiqk} , $k = 1, 2, 3$ means GH under YIQ with $M = k$.

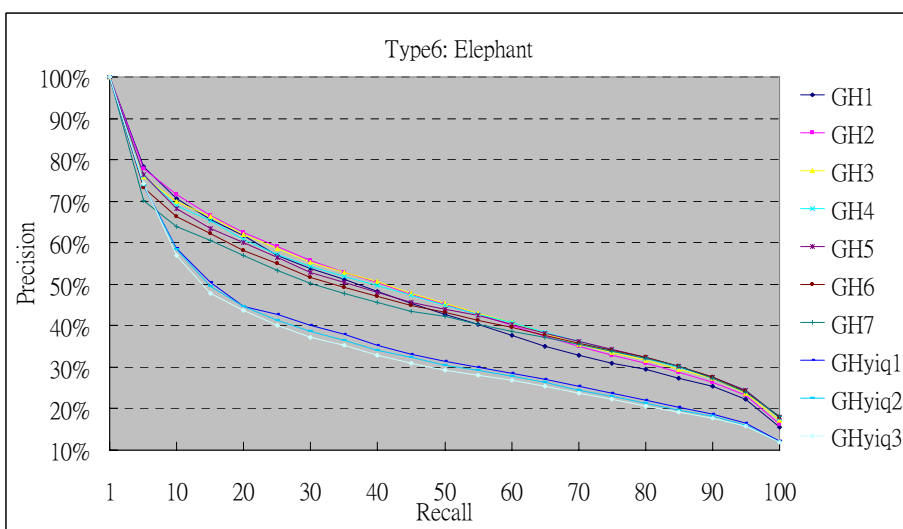
From the results in Fig. 4-4, we may conclude that it is more suitable to take GH under YIQ as an object size feature than GH under RGB in most cases. Precision-recall curves of images with variant object size like type2 beach and type 9 under water will be degenerated. The sizes of two types, flowers and elephants, are very close. Therefore, the mismatching of these two type images is caused. The first type models is a little bit degenerated for the reason that images of models have not unique but two major sizes. The rest of types are all superior to those under RGB.



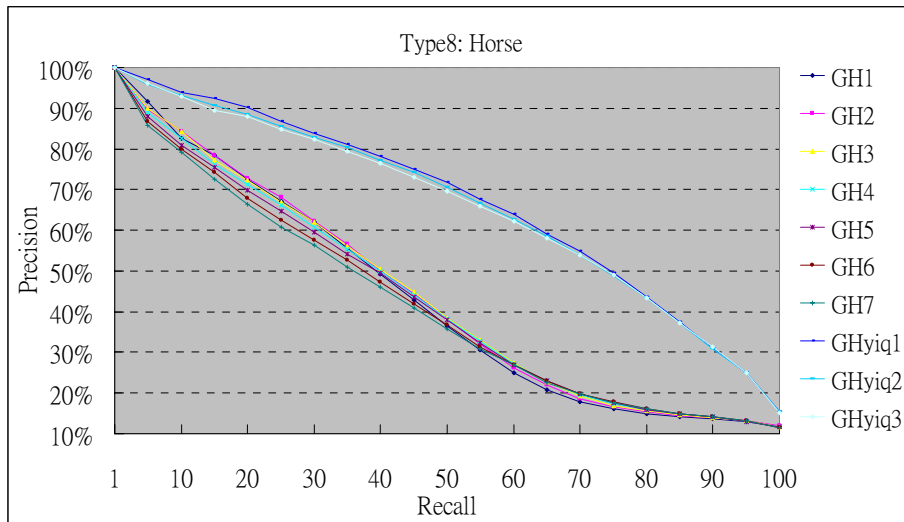
(a)



(b)



(c)



(d)

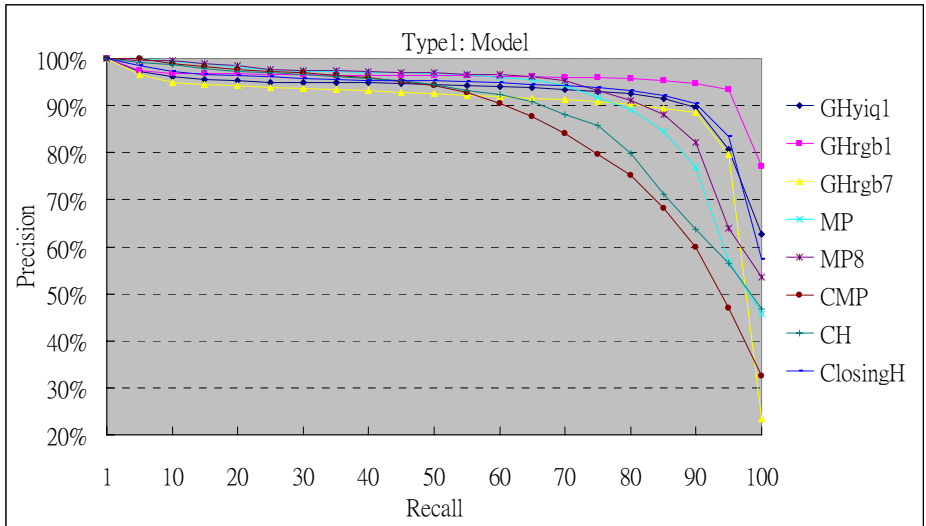
Fig. 4-4 Comparison of GH under RGB and YIQ.

4.2.4 Comparisons

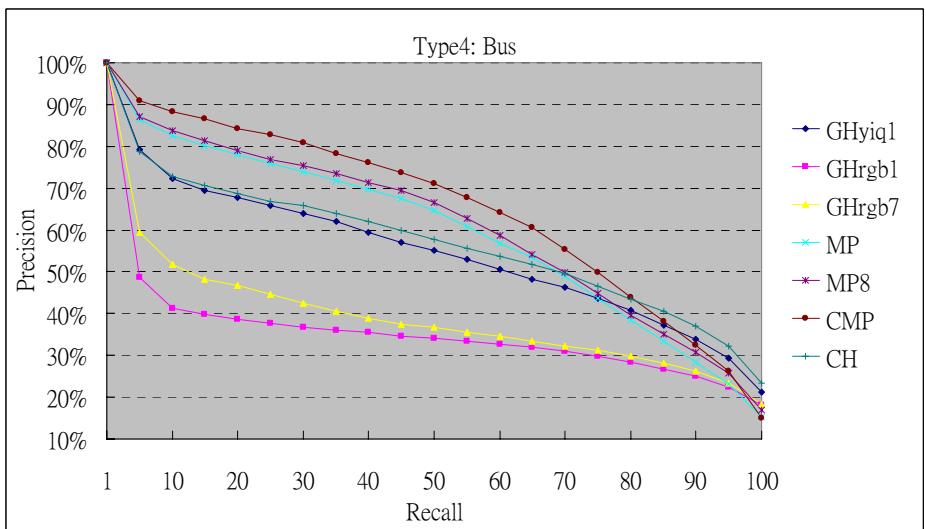


To show the performance of the proposed methods, the retrieval results are compared with those using morphological primitives, color moment primitives, and color histogram. Fig. 4-5 shows the precision-recall curves of each type in image database using different method. And Fig. 4-6 list the precisions of each method when $R = 20$. Due to the factor of object size, GH is a good feature for type1, 5, 6, 7, and 9. It is only below CH but above all other methods for type 3 and 8. And those are lower than other methods for type 2, 4 and 10.

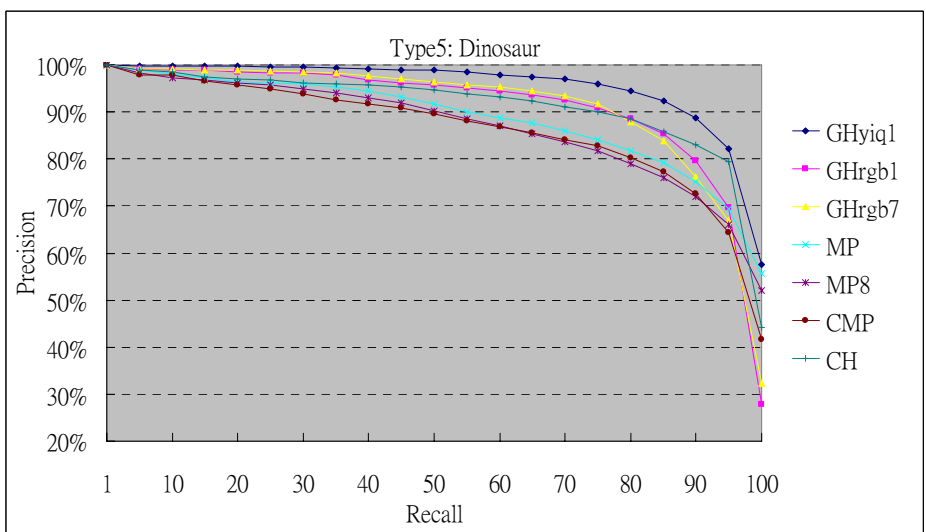
We may see more clearly by compute the difference of precisions between the proposed method GH and other methods which are listed in Table 4-1. The best improvement of GH when compared with MP, CMP, and CH is 47%, 38%, and 37%, respectively, and the worst degradation is 5%, 14%, and 17%, respectively.



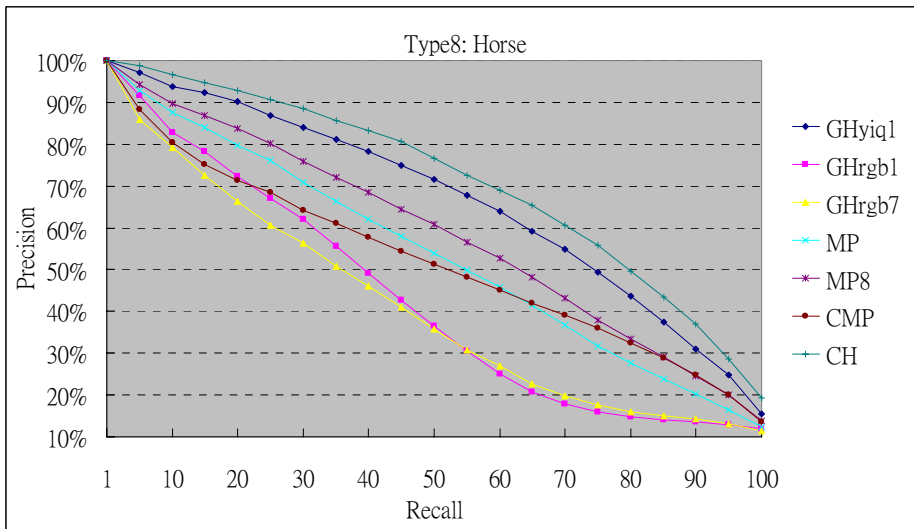
(a)



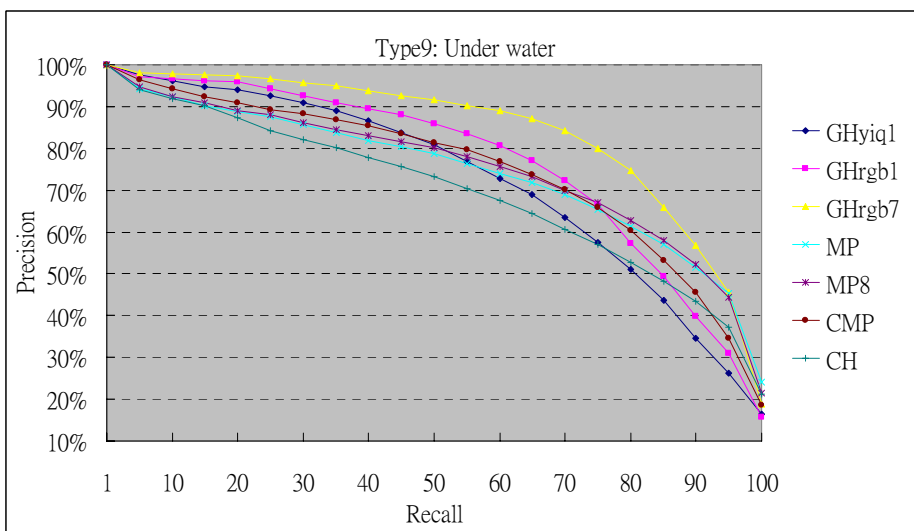
(b)



(c)



(d)



(e)

Fig. 4-5 The Precision-Recall curves of different processing methods.

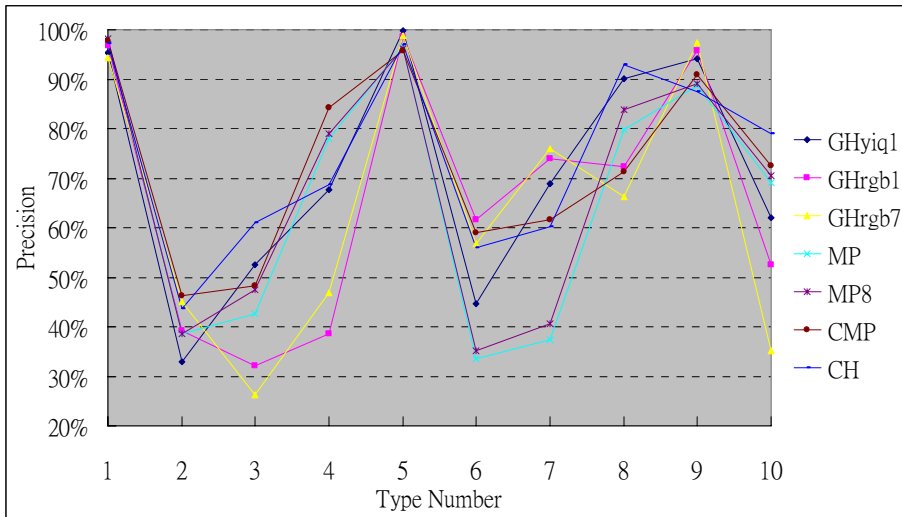


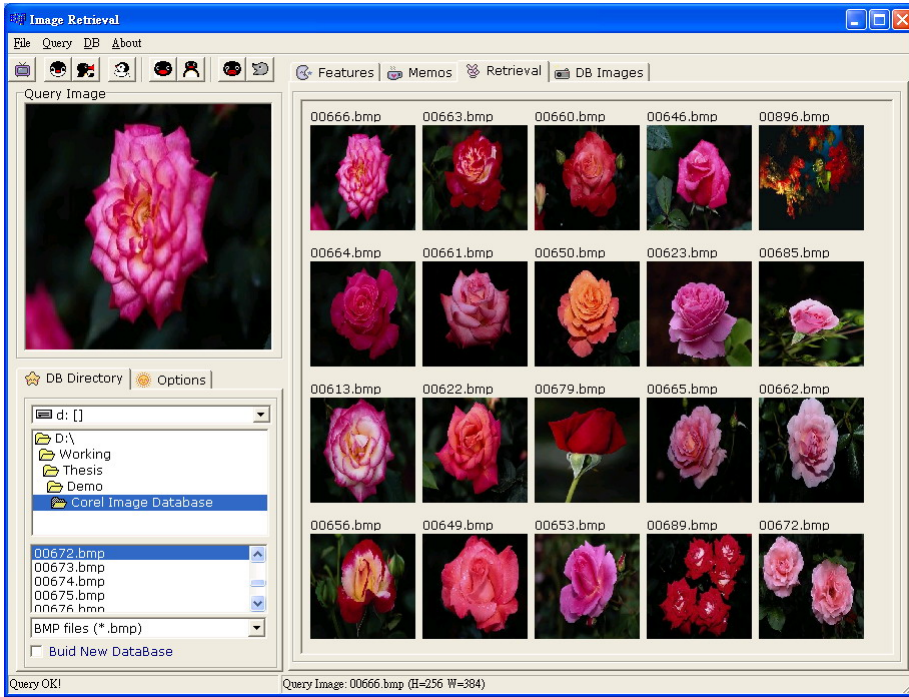
Fig. 4-6 The precision comparison for each type of image database among the proposed methods and other methods (R=20).

Table 4-1 Differences of precisions between proposed method GH and other methods (MP9, CMP, and CH).

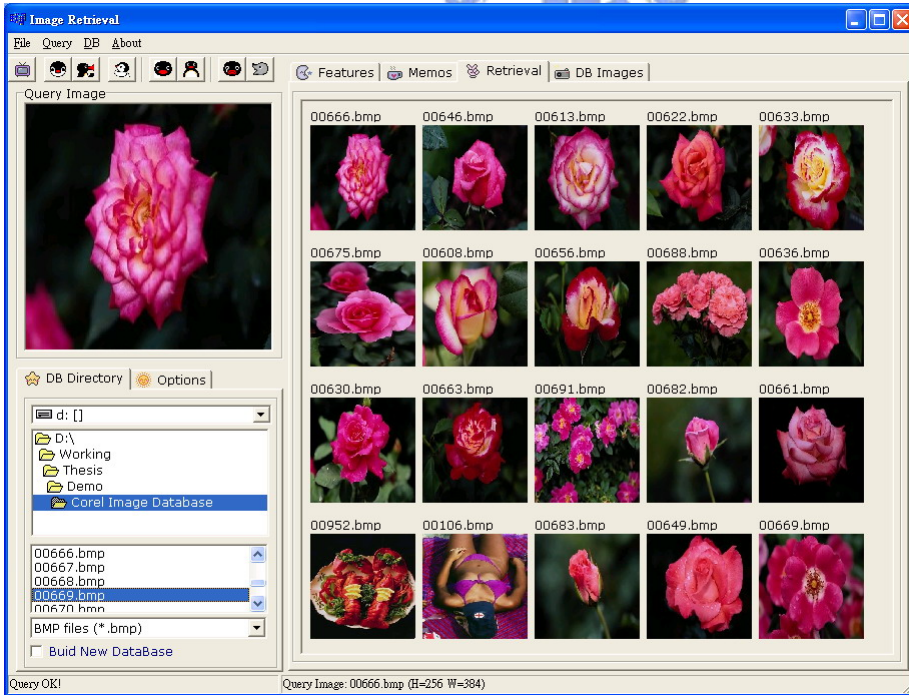
Type	MP9		CMP		CH	
	max	min	max	min	max	min
1	47%	-2%	25%	-3%	37%	-2%
2	11%	0%	22%	-1%	3%	-2%
3	15%	0%	16%	0%	0%	-11%
4	11%	-5%	7%	-11%	1%	-3%
5	4%	-3%	37%	0%	14%	0%
6	28%	0%	29%	0%	6%	-6%
7	36%	0%	38%	0%	18%	0%
8	27%	0%	21%	0%	0%	-6%
9	18%	0%	11%	-13%	24%	-2%
10	9%	0%	2%	-14%	0%	-17%

4.2.5 Example of Retrieval Results

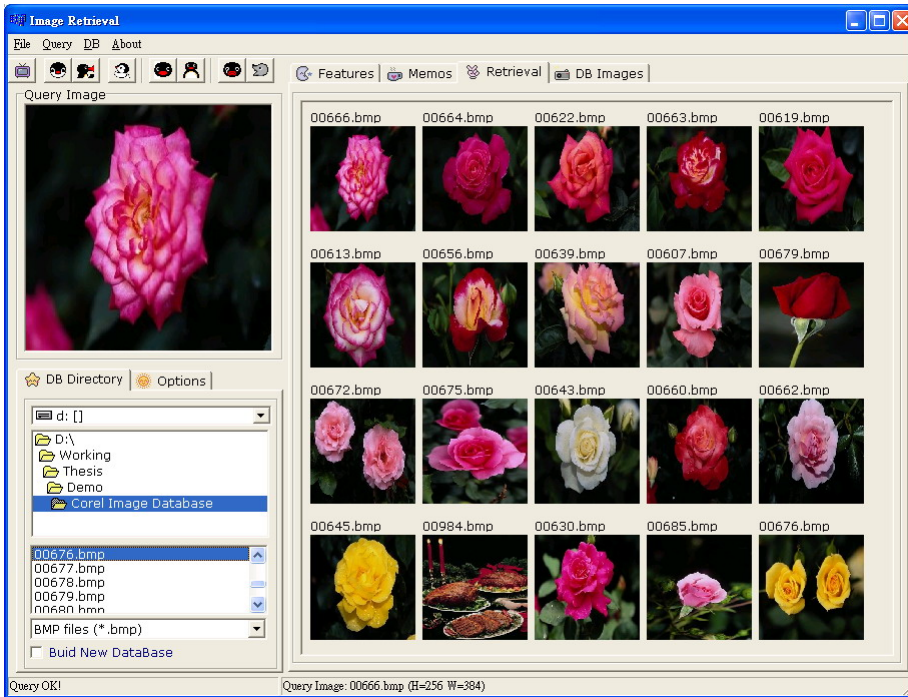
We give an example of a flower image by retrieve the top relevant 20 images with methods mentioned above and show the results in Fig. 4-7. Precisions of CH, CMP, GH with YIQ, GH with RGB, MP7, MP8, and MP9 are 95%, 90%, 95%, 100%, 95%, 100%, and 90%, respectively.



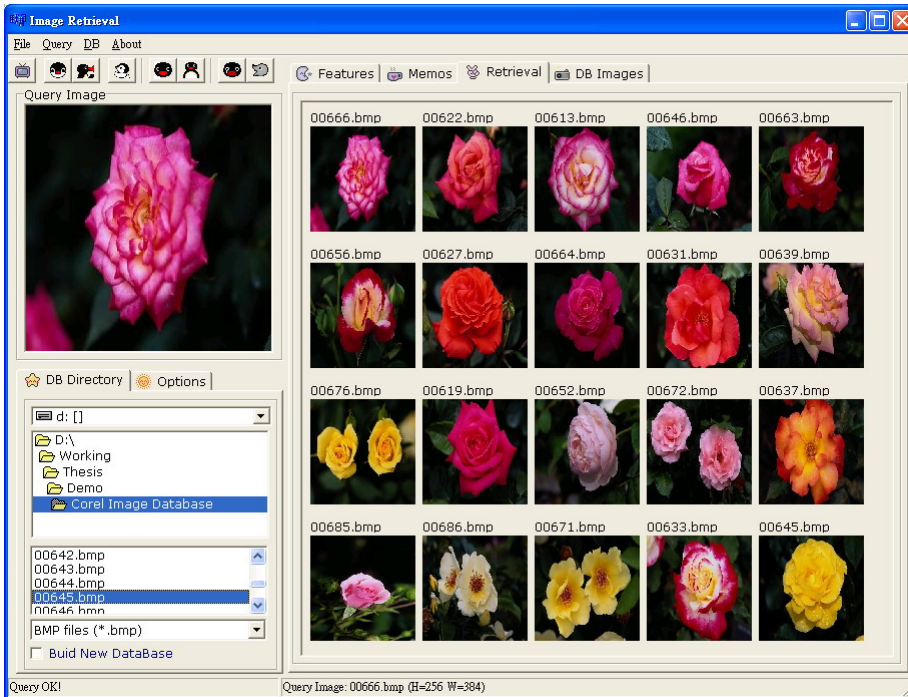
(a) Color Histogram



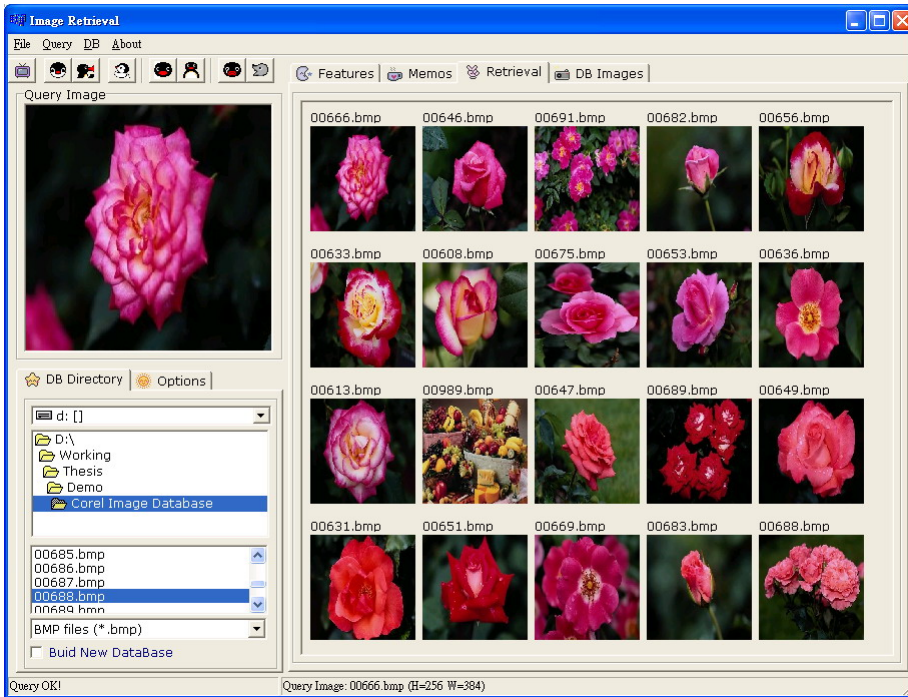
(b) Color Moment Primitives



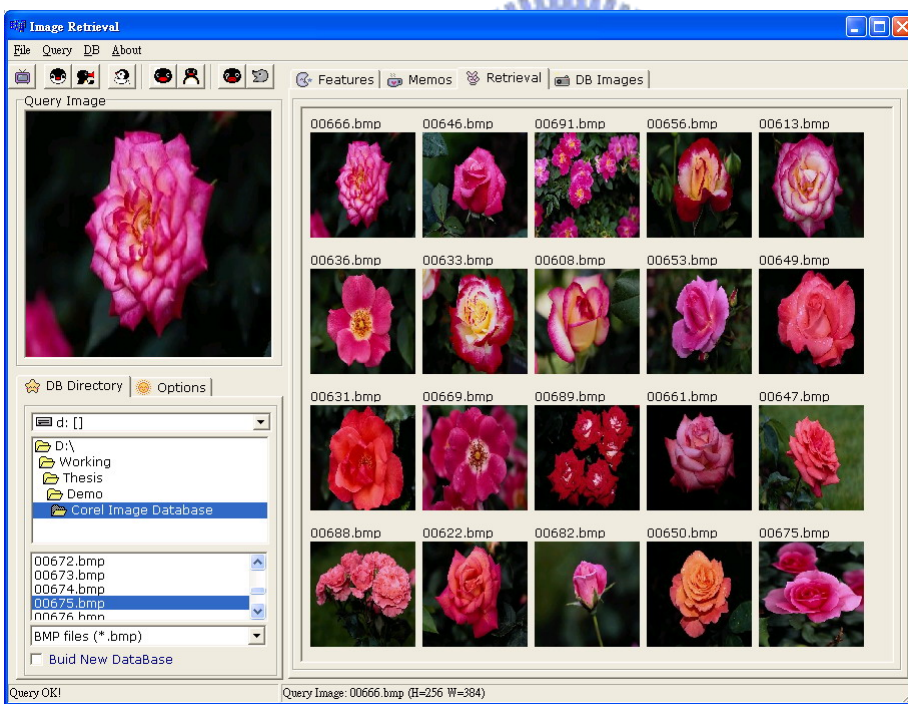
(c) GH under YIQ with 1 structuring element



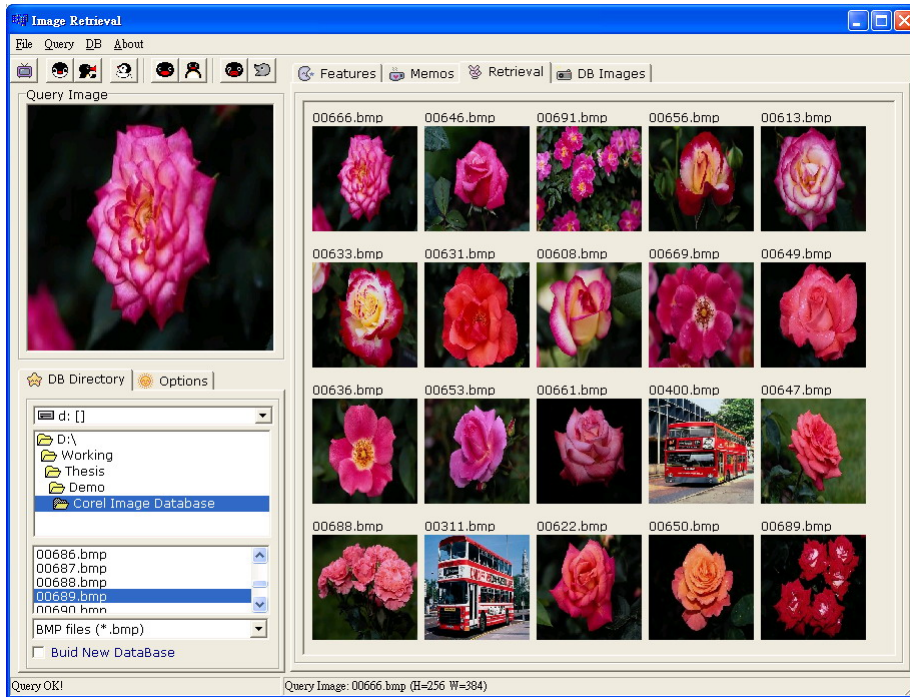
(d) GH under RGB with 1 structuring element



(e) Morphological Primitives with smallest-7 DPF

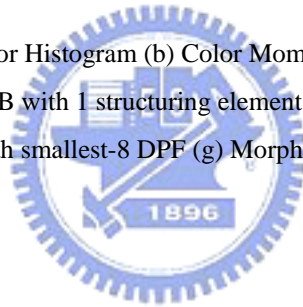


(f) Morphological Primitives with smallest-8 DPF



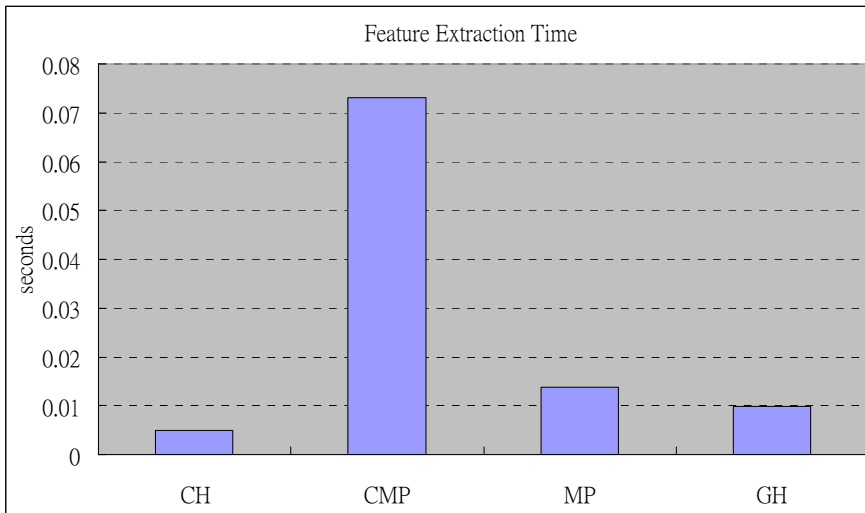
(g) Morphological Primitives

Fig. 4-7 Image retrieval using (a) Color Histogram (b) Color Moment Primitives (c) GH under YIQ with 1 structuring element (d) GH under RGB with 1 structuring element (e) Morphological Primitives with smallest-7 DPF (f) Morphological Primitives with smallest-8 DPF (g) Morphological Primitives for the top relevant 20 images.

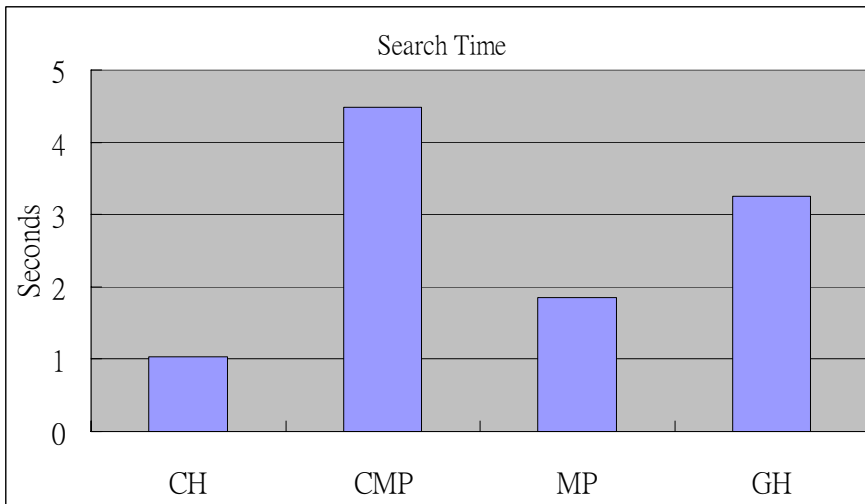


4.2.6 Processing Time

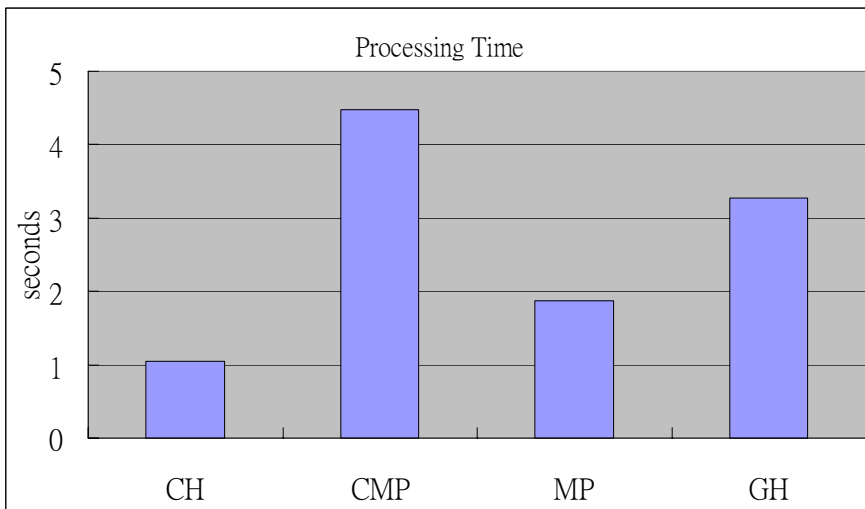
In Fig. 4-8, we show the average feature extraction time and average similarity measure time of methods discussed above. Our method takes less feature extraction time than MP and CMP, but almost three times of search time of CH since GH does intersection of three granulometric histogram.



(a)



(b)



(c)

Fig. 4-8 Histograms of processing time.

Chapter 5

Conclusions and Future Works

5.1 Conclusions

In this study, we propose a new method of image retrieval using morphological granulometric histogram. We also propose a new method of morphological primitives with dynamic partial function. By computing the granulometry histogram, we get the color, shape, and object size feature of an image. And it can represent a feature of an image effectively. On the other hand, the similarity, dynamic partial function, improves the precision of morphological primitives.



From the results shown in Chapter 4, we can conclude that our method has some advantages as described below.

1. Morphological operators involve simple logical operations and can be implemented in parallel, making real-time applications possible. [10]
2. Morphological granulometric histograms can represent features of images effectively.
3. Our proposed method reduces time to feature extraction of morphological primitives and color moment primitives.
4. Our proposed method performs good precisions.

5.2 Future works

We introduce the future works in the following topics. First, construct a hybrid image retrieval system that using our proposed method after a region segmentation. Second, seek for

faster similarity for histogram. Third, the relevant feedback shall be designed to make the interaction with users. Finally, to construct an on-line search engine one can use the technique of network and image retrieval.



REFERENCE

- [1] Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing", Prentice Hall, 2001.
- [2] Ze-Nian Li, Mark S. Drew, "Fundamentals of Multimedia", Prentice Hall, 2004.
- [3] Theo Gevers, "Color in Image Search Engines", Principles of Visual Information Retrieval, editor Michael S. Lew, Springer-Verlag, London, ISBN 1-852333-381-2, February 2001.
- [4] M.W. Richardson, "Multidimensional Psychophysics," Psychological Bull., vol. 35, pp. 639-660, 1938.
- [5] Arun Qamra, Yan Meng, and Edward Y. Chang, Senior Member, IEEE, "Enhanced Perceptual Distance Functions and Indexing for Image Replica Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27, No. 3, pp.379-391, March 2005
- [6] J.-L. Shih and L.-H. Chen, "Color image retrieval based on primitives of color moments", IEE Proceedings Vision, Image and Signal Processing, pp370-376, 2002.
- [7] C. R. Giardina and E. R. Dougherty, "Morphological Methods in Image and Signal Processing," Prentice Hall, New Jersey, 1988.
- [8] Matheron, G., "Randoms sets and integral equation", New York, Wiley, 1978.
- [9] Cecilia Di Ruberto, Andrew Dempster, Shahid Khan, bill Jarra, "Segmentation of Blood Images Using Morphological Operators", IEEE, pp397-400. 2000.
- [10] Jao-Shian Wu, "Color Image Retrieval Based on Morphological Primitives", master thesis in NCTU, 2003.
- [11] B.S. Manjunathe, Jens-Rainer Ohm, Vinod V. Vasudevan, and Akio Yamada, "Color and Texture Descriptors", IEEE Transactions on circuits and systems for video technology, Vol. 11, No. 6, pp703-715, 2001.

[12] Jose M. Martinez, “MPEG-7 Overview”, ISO/IEC JTC1/SC29/WG11 N4980, 2002.

[13] Byoung Chul Ko and Hyeran Byun, “FRIP: A Region-Based Image Retrieval Tool Using Automatic Image Segmentation and Stepwise Boolean and Maching”, IEEE transactions on multimedia, Vol. 7, No. 1, pp.105-113, 2005.

