

國立交通大學

資訊科學系

碩士論文

一個基於 MPEG-7 紋理描述器

之紋理檢索的研究



A Study on Texture Retrieval based on

MPEG-7 Texture Descriptors

研究生：林冠伶

指導教授：陳玲慧 教授

中華民國九十三年六月

一個基於 MPEG-7 紋理描述器之紋理檢索的研究

A Study on Texture Retrieval based on MPEG-7 Texture
Descriptors

研究生：林冠伶

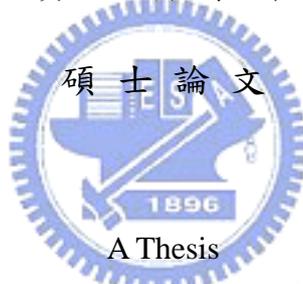
Student : Kuan-Ling Lin

指導教授：陳玲慧

Advisor : Dr. Ling-Hwei Chen

國立交通大學電機資訊學院

資訊科學系



Submitted to Department of Computer and Information Science
College of Electrical Engineering and Computer Science
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master

In
Computer and Information Science

June 2004
Hsinchu, Taiwan, Republic of China

中華民國九十三年六月

一個基於 MPEG-7 紋理描述器之紋理檢索的研究

研究生：林冠伶

指導教授：陳玲慧博士

國立交通大學電機資訊學院 資訊科學系碩士班

摘 要

由於多媒體資料的大量增加，如何有效的管理及檢索多媒體資料變成為一個重要的研究領域。其中在影像的檢索方面，可以利用顏色、形狀和紋理等特徵來作為檢索的依據。在本論文中就以紋理作為研究的方向，來作為影像檢索的依據。

本論文的主要目的在於利用 MPEG-7 中所制訂的統一格式，來萃取紋理影像的特徵，其中使用了同質性紋理影像描述器、紋理影像瀏覽描述器和邊界直方圖描述器三種特徵描述來檢索影像。

同質性紋理影像描述器是使用賈伯過濾器等來加強特定紋理方向和紋理大小的訊號，並計算在各個頻道的能量強度。紋理影像瀏覽描述器也是利用賈伯過濾器等來萃取紋理的方向性，並利用找出的方向性經過投影和自相關函數來找出紋理大小，並判斷規則度。而邊界直方圖描述器則是找於影像區塊中的邊界型態，統計成直方圖來作為特徵。

本論文實際使用了上述三種描述器，可以檢索出相似的紋理影像，但是每種描述器仍有它使用上的限制和缺點。同質性紋理影像描

述器只適合使用在有同質性較高的紋理影像上；紋理影像瀏覽描述器對於非 30 倍數的角度較不敏感，容易出現誤差；邊界直方圖描述器只適用於有明顯邊界分佈出現的影像。



A Study on Texture Retrieval based on MPEG-7 Texture Descriptors

Student : Kuan-Ling Lin

Advisor : Dr. Ling-Hwei Chen

Institute of Computer and Information Science
National Chiao Tung University

Abstract

Because the multimedia data is largely increased, how to management and retrieval multimedia information becomes an important study topic. For images, we can use color, shape and texture to do retrieval. In this thesis, we use texture to do image retrieval.

The main goal of this thesis is to extract the features of images based on MPEG-7 format. This standard provides three descriptors, homogeneous texture descriptor, texture browsing descriptor and edge histogram descriptor, to do image retrieval.

Homogeneous texture descriptor applies the Gabor filters for some direction or scale and computes the energy in each channel. Texture browsing descriptor also applies the Gabor filters to extract the directions of texture. Than the projections along these directions and autocorrelation function are used to find scales and to compute the regularity of texture.

Edge histogram descriptor is used to find the edge type in each image block and then to establish the edge histogram of an image.

This thesis uses the three descriptors described above to design an image retrieval/browsing system. Experimental results show that it can retrieval most similar texture image. But these descriptors have some limits and disadvantages. Homogeneous texture descriptor can only apply to homogeneous images. Texture browsing descriptor is non-sensitive to the angles that are not the multiple of 30 degrees, and this will make errors. Edge histogram descriptor can only apply to the images with obvious edges.



誌 謝

首先要感謝我的指導老師 陳玲慧教授在這二年間對我的教導，使我能得到許多的專業知識與技能，還有待人處事方面的經驗。還要感謝實驗室的瑞祥學長、民全學長和萱聖學長在我遇到困難時給我的建議；感謝一起畢業的同學雅茹和業承在我挫折時給我的鼓勵；還要謝謝學弟們給我精神上的支持，沒有這些人的幫忙我就無法順利完成這篇論文。此外，還要感謝口試委員李坤龍博士和黃仁俊教授給予的建議和指導，讓我的論文能更加完善。

最後要感謝我的父母和妹妹，他們在我背後所給予的支持，是我最大的動力，也是我繼續努力的原因，將這篇論文獻給我最親愛的家人。



目 錄

中文摘要	i
英文摘要	iii
誌謝	v
目錄	vi
圖目錄	vii
表目錄	viii
一、	緒論	1
1.1	研究動機	1
1.2	文獻探討	2
1.3	論文架構	4
二、	同質性紋理影像描述器	5
2.1	定義	5
2.2	特徵萃取	6
2.3	比較方法	9
2.4	實驗結果	9
三、	紋理影像瀏覽描述器	12
3.1	定義	12
3.2	特徵萃取	13
3.2.1	主要方向估計	13
3.2.2	大小估計	14
3.2.3	規則度估計	15
3.3	瀏覽方法	17
3.4	實驗結果	18
四、	邊界直方圖描述器	20
4.1	定義	20
4.2	特徵萃取	21
4.3	產生邊界直方圖	23
4.4	比較方法	24
4.5	實驗結果	25
五、	結論	27
參考文獻	29

圖 目 錄

圖 2.1	計算 HTD 所使用之頻道規劃	6
圖 2.2	表示在極座標上的 5×6 的賈伯過濾器	8
圖 2.3	以 HTD 作檢索的結果	10
圖 3.1	紋理影像的規則度	18
圖 3.2	紋理影像的方向性	18
圖 3.3	紋理影像的大小	19
圖 3.4	令 $v_1=2$ 且 $v_2=3$ 或 $v_4=3$ ，利用 TBD 瀏覽的結果	19
圖 4.1	子影像和影像方塊的定義	21
圖 4.2	子方塊的切割方式	22
圖 4.3	邊界偵測的過濾係數	22
圖 4.4	將子影像分類為 13 種半全域影像	24
圖 4.5	所有的邊界直方圖	24
圖 4.6	以 EHD 檢索結果 (一)	25
圖 4.7	以 EHD 檢索結果 (二)	26
圖 4.8	以 EHD 檢索結果 (三)	26



表 目 錄

表 2.1	半徑分量所用之參數.....	7
表 2.2	角度分量所用之參數.....	8
表 2.2	HTD 的統計結果.....	11
表 3.1	規則度所代表的意義.....	17
表 3.2	方向性所代表的意義.....	17
表 3.3	紋理大小所代表的意義.....	17
表 5.1	三種紋理影像描述器的比較.....	28



第一章 緒論

1.1 研究動機

目前對於影像資料的檢索，大部分都是在影像資料上加入文字描述來記錄影像中所包含的內容，然後在檢索的時候輸入文字來描述所要求的影像內容。但是並不是所有網路上的影像資料都有文字的描述，因此輸入範例圖片來做為檢索的依據（query by example），更能有效的檢索所有的影像資料。

紋理影像（texture）是一種由大量的相似圖樣（pattern）構成的特殊影像，例如：水波、木紋和編織品的花樣等；也可以將紋理影像形容為許多基本的圖樣元素（primitive）所組成的集合。圖樣元素決定了紋理影像的外觀，所以我們可以利用圖樣元素的排列方式定義紋理影像的方向（direction）、大小（scale）和規則度（regularity），並以此來檢索相似的紋理影像或瀏覽具有相同方向、大小和規則度的紋理影像。在以往的研究當中，大家各自使用了自己的方法來萃取紋理影像的特徵[1-4]，雖然各有優點，但是在網際網路上便無法統一作檢索的動作，而需要為網路上的每一張影像重新依照自己的方法萃取特徵，需要作大量即時運算，因此花費了許多額外的時間和電腦資源。

由 ISO/IEC 所制訂的標準 MPEG-7[6,7]，就是為了因應網際網路上的檢索需要而產生，以固定的格式儲存影像或聲音的特徵資料，影

像的製作者可以預先為影像儲存 MPEG-7 所定義的特徵資料，提供檢索系統使用，以達到快速檢索的效果。

我們依照 MPEG-7 中所規定的格式，萃取出紋理影像的特徵，以此來檢索資料庫中的紋理影像，並比較影像間的相似度來找出與所輸入的範例圖片最為相似的結果。

1.2 文獻探討

在過去對於紋理影像的研究中，也有多種萃取紋理特徵的方法，以作為檢索或瀏覽的標準。例如[5]中所提出的方法，是以紋理影像的規則度作為檢索的標準，在此篇論文中使用圖樣元素的相似度和圖樣元素排列的規則度來判斷整張紋理影像的規則度。

首先，算出紋理影像上在 (i, j) 位置的亮度變化梯度 (intensity gradient)，在 X 軸方向表示為 $g_x(i, j)$ ，在 Y 軸方向表示為 $g_y(i, j)$ ，然後再計算梯度域 (gradient field) 上的自相關函數 (autocorrelation function)，令 $c(k, l)$ 為在 (k, l) 位置上的相關值 (correlation value)：

$$c(k, l) = \frac{\sum_i \sum_j g_x(i, j)g_x(i+k, j+l) + g_y(i, j)g_y(i+k, j+l)}{\sqrt{(\sum_i \sum_j g_x(i, j)^2 + g_y(i, j)^2)(\sum_i \sum_j g_x(i+k, j+l)^2 + g_y(i+k, j+l)^2)}}$$

如果圖像元素的排列擁有高度的重複性，在經過梯度和自相關函數的計算之後，也會有重複性的結果出現；若是一張毫無規則的隨機影像，則只有在原點的位置會有較高數值出現。

只要不是完全隨機的影像，我們可以假設圖樣元素是一個平行四邊形的區塊，而圖樣元素也會依照此平行四邊形的兩個方向來排列。為了要找出這兩個方向，我們利用之前的梯度自相關函數的結果，找出除了原點之外的最大峰值 (peak)，由原點到此峰值位置的向量就是我們找到的第一個方向。因為是平行四邊形的兩個方向，所以第二個方向和第一個方向會有一個適當的夾角，我們假設這個夾角的範圍是在 $\pi/4$ 到 $3\pi/4$ 之間，所以我們在和第一個方向夾角為 $\pi/4$ 到 $3\pi/4$ 的範圍內尋找最大的峰值，第二個方向就是由原點到此峰值之向量。



找到方向之後，沿著方向對梯度自相關函數作剖面 (profile)，若是圖樣元素重複性出現，則在剖面圖上峰值出現的位置會有相近的間距，所以我們從得到的剖面圖上，找到最高的峰值 (通常也是第一個峰值)，令其與原點的距離為 p ，數值為 C_1 ，並在與原點距離為 $2p$ 的位置附近找尋第二個峰值，令其數值為 C_2 ，然後給予影像各項數值：

$$\text{排列規則度} = \begin{cases} \frac{C_2}{C_1} & \text{when } C_2 \geq 0 \\ 0 & \text{when } C_2 < 0 \end{cases}$$

$$\text{圖樣元素相似度} = \frac{C_1}{\text{排列規則度}} = \frac{C_1}{\frac{C_2}{C_1}} = \frac{C_1^2}{C_2} \quad \text{if } C_2 \geq 0$$

$$\text{整張影像重複度} = \text{圖樣元素相似度} \times \text{排列規則度} = \frac{C_1^2}{C_2} \times \frac{C_2}{C_1} = C_1$$

我們便可以利用這三個數值來判斷一張影像的規則度，並與其他影像作比較。

在這篇論文中，雖然可以正確地判斷出紋理影像的規則度，針對規則度將紋理影像加以分類，但是對於檢索系統來說，並無法分辨擁有相同規則度的紋理影像間的差異度，所以我們還需要利用其他的特徵來檢索紋理影像。

1.3 論文架構

在第一章中，我們敘述了本論文的研究背景和研究動機，在接下來的第二到四章中，我們將介紹 MPEG-7 所定義的三種描述紋理影像特徵的方法，第二章是同質性紋理影像描述器，第三章是紋理影像瀏覽描述器，而第四章是邊界直方圖描述器，在章節中將會詳細介紹這三種方法的實作細節和我們所得到的實驗結果。

第二章 同質性紋理影像描述器 (Homogeneous Texture Descriptor)

2.1 定義 (Definition)

同質性紋理影像描述器 (HTD, Homogeneous Texture Descriptor) [8] 是利用紋理影像的方向 (directionality) 和大小 (scale) 來萃取紋理影像的特徵，透過賈伯過濾器 (Gabor filter) 加強在頻率域 (frequency domain) 上不同的頻道 (frequency channel) 上的能量，並計算頻道所含有的平均能量 (mean energy) 和能量標準差 (energy deviation) 來記述紋理影像的特徵；另外還以空間域 (spatial domain) 上影像的平均值和標準差作為特徵。

我們將頻率平面 (frequency plane) 依方向和半徑切割成 30 個頻道，如圖 2.1 所示，頻道在角度上是以 30 度等量分割為 6 個方向，而在半徑上則是以二分之一為底等比分割為 5 個半徑，然後計算在每個頻道中的平均能量和其偏差值，將其紀錄在描述器中：

$$\text{HTD} = [f_{DC}, f_{SD}, e_1, e_2, \dots, e_{30}, d_1, d_2, \dots, d_{30}]$$

其中 f_{DC} 和 f_{SD} 是影像在空間域上的平均值和標準差，而 e_i 和 d_i 是在第 i 個頻道中的平均能量和能量標準差。

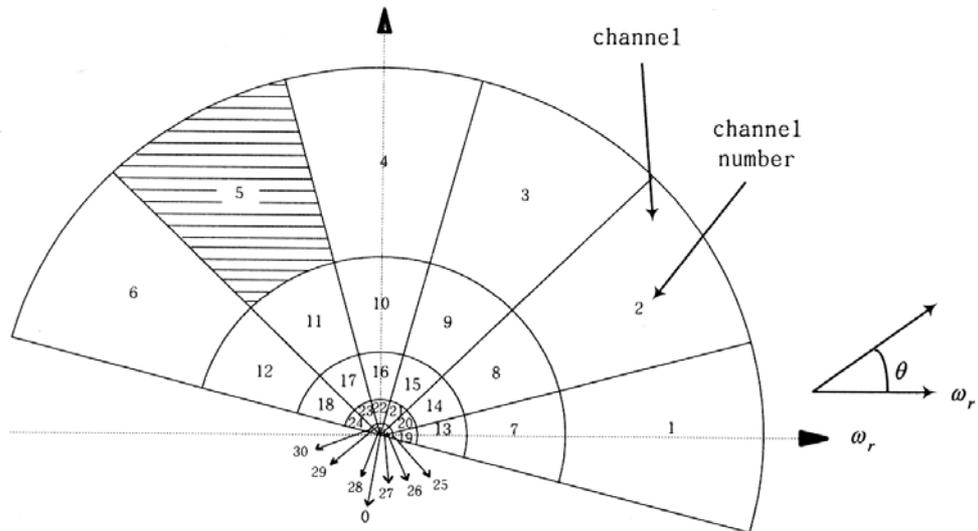


圖 2.1 計算 HTD 所使用之頻道規劃

2.2 特徵萃取 (Feature Extraction)

描述器中的 f_{DC} 和 f_{SD} 可以用影像的像素值直接計算平均值和標準差，而剩下的元件則可在頻率域上使用賈伯過濾器[3]，以圖 2.1 中的 6 個方向分割（以 r 表示）和 5 個半徑分割（以 s 表示）共 30 個頻道來計算，方法如下：

- 第一步：將輸入的影像經過快速傅立葉轉換（FFT，Fast Fourier Transform）轉換到頻率域中，並以極座標的方式表示為 $F(\omega, \theta)$ 。
- 第二步：以賈伯過濾器處理 $F(\omega, \theta)$ ，找出紋理影像的方向特性。

賈伯過濾器可以加強在某個特定方向和大小上的亮度，所以如果我們所使用的方向和大小，剛好和影像所具有的方向和大小相同，則可以得到最大的亮度，可以用於尋找影像所含有的方向性和大小。

在頻率域上，我們所使用的公式如下：

$$H_i(\omega, \theta) = G_{s,r}(\omega, \theta)F(\omega, \theta), i = 6 \times s + r + 1$$

$H_i(\omega, \theta)$ 是賈伯過濾器處理之後的結果， i 是頻道的索引

(index)， $G_{s,r}(\omega, \theta)$ 則是在第 s 個半徑分割和第 r 個方向分割上的賈

伯函數 (Gabor Function)：

$$G_{s,r}(\omega, \theta) = \exp\left[-\frac{(\omega - \omega_s)^2}{2\sigma_s^2}\right] \exp\left[-\frac{(\theta - \theta_r)^2}{2\sigma_r^2}\right]$$

在賈伯函數中的標準差 σ_s 和 σ_r 可表示為：

$$\sigma_s = \frac{B_s}{2\sqrt{2\ln 2}}, \quad \sigma_r = \frac{B_r}{2\sqrt{2\ln 2}}$$

B_s 和 B_r 是在半徑分量和角度分量的間距，表 2.1、2.2 列出所有在賈伯函數中所用到之參數。

半徑索引 (s)	0	1	2	3	4
中心頻率 (ω_s)	$\frac{3}{4}$	$\frac{3}{8}$	$\frac{3}{16}$	$\frac{3}{32}$	$\frac{3}{64}$
間距 (B_s)	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$
σ_s	$\frac{1}{4\sqrt{2\ln 2}}$	$\frac{1}{8\sqrt{2\ln 2}}$	$\frac{1}{16\sqrt{2\ln 2}}$	$\frac{1}{32\sqrt{2\ln 2}}$	$\frac{1}{64\sqrt{2\ln 2}}$

表 2.1 半徑分量所用之參數

角度索引 (r)	0	1	2	3	4	5
中心頻率 (ω_r)	0°	30°	60°	90°	120°	150°
間距 (B_r)	30°	30°	30°	30°	30°	30°
σ_r	$\frac{30^\circ}{2\sqrt{2\ln 2}}$					

表 2.2 角度分量所用之參數

依據上面的數據，我們可以將賈伯過濾器使用在極座標表示的頻率空間上，如圖 2.2。

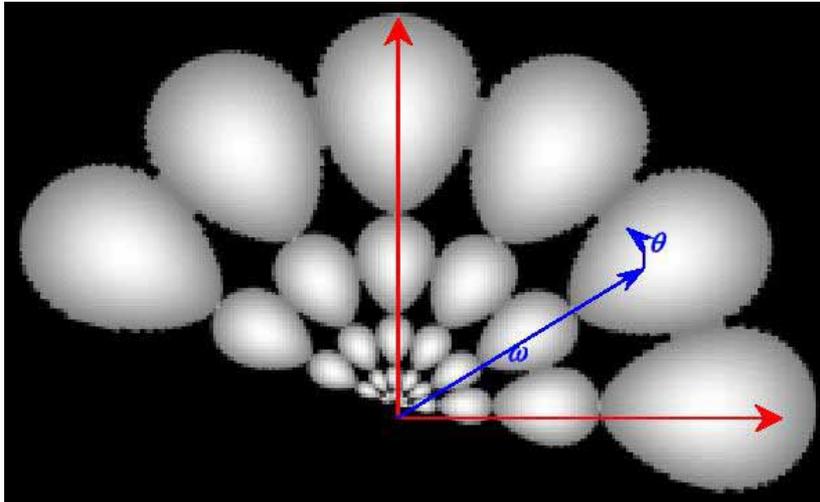


圖 2.2 表示在極座標上的 5x6 的賈伯過濾器

- 第三步：計算平均能量 e_i 和能量標準差 d_i

$$e_i = \log[1 + p_i] \text{ , 其中 } p_i = \sum_{\omega} \sum_{\theta} H_i^2(\omega, \theta)$$

$$d_i = \log[1 + q_i] \text{ , 其中 } q_i = \sqrt{\sum_{\omega} \sum_{\theta} [H_i^2(\omega, \theta) - p_i]^2}$$

依照上述三個步驟，我們可以得出 30 個能量平均值和 30 個能量標準差，再加上像素平均值和像素標準差，這 62 個數值可以構成同質性紋理影像描述器，使 $HTD = [f_{DC}, f_{SD}, e_1, e_2, \dots, e_{30}, d_1, d_2, \dots, d_{30}]$ ，我們

以HTD來描述一張紋理影像的特徵，並比較兩張紋理影像之間的相似度。

2.3 比較方法 (Matching Method)

在上一小節中，我們從紋理影像中擷取出了 HTD 值，以 62 個數值來代表一張紋理影像。為了描述二張紋理影像間的差異，我們定義兩張紋理影像的差異值 (distance)，並將之用來找尋相似的紋理影像。

差異值計算方式定義為計算兩張紋理影像 HTD 數值的差，如下：

$$distance(HTD_{query}, HTD_{Database}) = \sum_k \left| \frac{HTD_{query}(k) - HTD_{Database}(k)}{\alpha(k)} \right|$$

$HTD_{query}(k)$ 表示所要查詢的紋理影像的第 k 個 HTD 值， $HTD_{Database}(k)$ 表示資料庫中紋理影像的第 k 個 HTD 值，而 $\alpha(k)$ 是正規化參數，可以使用資料庫中所有紋理影像的第 k 個 HTD 值所計算出的標準差，或是自訂的任一數值，如 $\alpha(k) = 1$ 。

影像之間的差異值越小，代表二張影像特徵間的相似度越高，二張影像的性質也越相近。反之，則影像差異性越大。

2.4 實驗結果 (Experimental results)

我們以 Brodatz[9]所列出的 112 張紋理影像作為基礎資料庫，對於這 112 張基礎影像分別作亮度調整、放大、縮小、平移等調整，使每一張影像另外產生 8 張類似的紋理影像，所以整個資料庫一共有

112×9=1008 張紋理影像。

而在這 112 張基礎影像中，我們選擇 43 張屬於同質性的紋理影像，和其延伸影像共 387 張，來作為 HTD 的檢索資料庫。

圖 2.3 便是一個例子，排名第一的影像也是我們輸入作檢索的範例影像，在 HTD 差異值最小，即相似度最高的前 10 名影像中，有 8 張影像是和範例影像同一來源的相似影像。

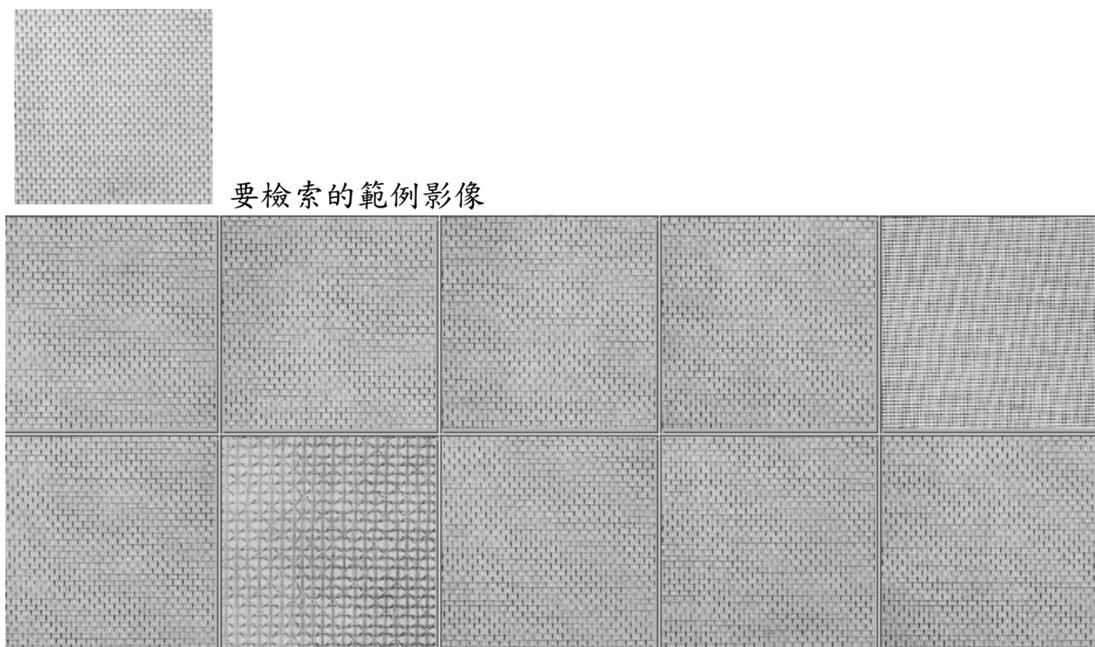


圖 2.3 以 HTD 作檢索的結果

對於整個資料庫作檢索統計的結果，我們以精確度 (precision) 和檢索率 (recall) 來做為測量的標準：

$$\text{精確度} = \frac{\text{搜尋到的正確影像數目}}{\text{所有搜尋到的影像數目}}$$

$$\text{檢索率} = \frac{\text{搜尋到的正確影像數目}}{\text{應該搜尋到的正確影像數目}}$$

我們搜尋 10 張與輸入影像最相似的紋理影像當作結果，以人工將資料庫中的 43 種紋理影像分類成相似的集合並將同一集合中的影像當作所應搜尋到的正確影像數目。對於我們所選擇的 HTD 資料庫，我們所得到的統計結果如表 2.3。

精確度(p)	檢索率(r)
0.72	0.32

表 2.2 HTD 的統計結果



第三章 紋理影像瀏覽描述器 (Texture Browsing Descriptor)

3.1 定義 (Definition)

紋理影像瀏覽描述器 (TBD, Texture Browsing Descriptor) 是以人類視覺所感受到的特徵表示紋理影像，共有 5 個參數表示方向性 (directionality)、大小 (scale) 和規則度 (regularity)：

$$\text{TBD} = [v_1 \ v_2 \ v_3 \ v_4 \ v_5]$$

$v_1 \in \{1,2,3,4\}$ ： v_1 表示紋理影像的規則程度，越大的值代表越是規則的紋理影像。

$v_2, v_4 \in \{0,1,2,3,4,5,6\}$ ： v_2, v_4 表示紋理影像主要的兩個方向性，所計算出的方向性準確度取決於紋理影像規則的程度，在此以 30° 為一間距分割 180° 為 6 個區間，分別以 1-6 表示，0 則表示該紋理影像不具有方向性。

$v_3, v_5 \in \{1,2,3,4\}$ ： v_3, v_5 表示兩個紋理影像主要方向上的紋理大小，而就像方向性一樣，所計算出的紋理大小準確度也是取決於紋理影像規則的程度，在此將半徑維度以二分之一等比切割成 4 個區間，數值越大所表示的大小也越大。

3.2 特徵萃取 (Feature Extraction)

我們使用如第二章所述之賈伯過濾器對紋理影像做前置處理，不同的是，在此章中以 4 個半徑分量來取代之前所使用的 5 個半徑分量，方向分量則維持不變為 6 個，並且在空間域中做後續的處理。

3.2.1 主要方向估計 (Dominant Direction Estimation) [10,11]

以賈伯過濾器對於 6 個方向分量和 4 個半徑分量作處理，我們可以得到 24 張濾後影像 (filtered image)，在第 s 個半徑分量和第 k 個方向分量上的濾後影像表示為 $W_{sk}(x,y)$ 。對 24 張濾後影像，我們可以計算出一個具有 24 個數值的直方圖 (histogram)：

$$H(s,k) = \frac{N(s,k)}{\sum_{k=1}^6 N(s,k)}, s = 1, \dots, 4, k = 1, \dots, 6$$

$N(s,k)$ 是在 $W_{sk}(x,y)$ 中像素值大於臨界值 t_s 的像素數目， $t_s = \mu_s + \sigma_s$ 。 μ_s 和 σ_s 分別是在半徑分量 s 上，六個方向分量的濾後影像 ($W_{sk}(x,y)$, $k = 1, \dots, 6$) 加起來的像素平均值和標準差。

如此我們可以得到一張具有 24 個數值的直方圖，根據賈伯過濾器的特性，在直方圖上最大值 (peak) 的位置 $H(s_0, k_0)$ ，其所代表的方向分量 k_0 就可能是紋理影像所具有的方向性。我們可以進一步利用計算對比值來加強對於主要方向的估計準確度，對比值 $C(s,k)$ 可由下式計算：

$$C(s,k) = 0.5 \times [2 H(s,k) - H(s,k-1) - H(s,k+1)]$$

假設我們所找到的最大值為 $C(s_0, k_0)$ ，若是 $C(s_0+1, k_0)$ 或 $C(s_0-1, k_0)$ 也是找到的區域最大值，則將此 k_0 所對應的方向分量訂為所要找的第一個方向性 DO_1 。再依此方法找第二大的對比值，就可以找到第二個代表方向性 DO_2 。

我們依照上述步驟計算出的代表方向性，依照 $v_2 = DO_1, v_4 = DO_2$ 分別設為TBD中的方向參數 v_2 、 v_4 。

3.2.2 大小估計 (Scale Estimation)

算出紋理影像的主要方向之後，我們利用這兩個方向來計算在濾後影像上的投影 $P_H^{(sk)}$ 和 $P_V^{(sk)}$ ：

$$P_H^{(sk)}(l) = \sum_x \sum_y W_{sk}(x, y) \delta(x \cos \theta_{v_2} + y \sin \theta_{v_2} - l)$$

$$P_V^{(sk)}(l) = \sum_x \sum_y W_{sk}(x, y) \delta(x \cos \theta_{v_4} + y \sin \theta_{v_4} - l)$$

$$\delta(x) = \begin{cases} 1 & , \text{if } x = 0 \\ 0 & , \text{otherwise} \end{cases}$$

$W_{sk}(x, y)$ 代表在第 s 個半徑分量和第 k 個方向分量上的濾後影像， θ_{v_2} 和 θ_{v_4} 則表示在 v_2 、 v_4 方向上的角度值， l 是在投影上的位置。

然後，對於每個投影函數我們計算其正規自相關函數 (NAC, Normalized Autocorrelation Function)：

$$NAC(k) = \frac{\sum_{m=k}^{L-1} P(m-k)P(m)}{\sqrt{\sum_{m=k}^{L-1} P^2(m-k) \sum_{m=k}^{L-1} P^2(m)}}$$

$NAC(k)$ 代表投影函數上距離 k 的兩點之間的相關性， $P(l)$ 為投影函數 $P_H^{(sk)}(l)$ 或 $P_V^{(sk)}(l)$ ，而 L 是投影函數的長度， $NAC(k)$ 可以找出函數的週期性，假設函數的週期為 k_0 ，則 $NAC(k_0)$ 的值會最大，而在此投影函數的週期所代表的意義就是紋理影像的大小。

對於每個計算出的 $NAC(k)$ ，我們找出所有的區域峰值(local peak)和區域谷值 (local valley)，並記錄其位置及大小。假設共找到 M 個峰值和 N 個谷值，並以 $p_posi(i)$ 和 $p_magn(i)$ ($i=1,2,\dots,M$) 表示峰值的位置和大小， $v_posi(i)$ 和 $v_magn(i)$ ($i=1,2,\dots,N$) 表示谷值的位置和大小。然後計算在這個投影函數上的對比值：

$$contrast = \frac{1}{M} \sum_{i=1}^M p_magn(i) - \frac{1}{N} \sum_{i=1}^N v_magn(i)$$

則我們可以找出TBD中的紋理大小參數[11]，使得 v_3 等於擁有最大對比值的 $P_H^{(sk)}$ 所代表的半徑分量， v_5 等於擁有最大對比值的 $P_V^{(sk)}$ 所代表的半徑分量。

3.2.3 規則度估計 (Regularity Estimation)

利用在上一小節所算出的峰值數列 $p_posi(i)$ ，我們先對於濾後影像的投影做一個分析篩選。分別計算相鄰的兩個峰值間距離的平均值 dis ，和間距標準差的平方根 std ，然後計算一個檢驗值 $g = std / dis$ ，如果計算出來的 g 值小於我們事先選定的臨界值 T_p (T_p 實驗值為 0.33)，

則此投影便可認為是候選者。

對於這些通過檢驗的候選者，我們依照下列規則加以分類，並給予一個信任值：

- C_1 ：當目標候選者的鄰近方向分量或是鄰近半徑分量，至少有一個候選者存在時，便歸於此類，並給予信任值 $V_1 = 1.0$ 。

- C_2 ：當目標候選者所在的方向分量或是半徑分量，有其他候選者存在，但不是在鄰近的位置時，便歸於此類，並給予信任值 $V_2 = 0.5$ 。

- C_3 ：當目標候選者所在的方向分量和半徑分量，都沒有任何其他候選者存在時，便歸於此類，並給予信任值 $V_3 = 0.2$ 。

如此一來，每一個候選的投影都有一個相對應的信任值，規則度的估計方法便是將所有候選者的信任值總和起來：

$$M = \sum_{i=1}^3 N_i \times V_i \quad , \quad N_i \text{ 表示屬於 } C_i \text{ 這個分類的後選者數量}$$

令 $M_{img} = M_H + M_V$ ， M_H 是在 v_2 方向上，由上式所算出的信任度總和， M_V 是在 v_4 方向上，由上式所算出的信任度總和。然後由 M_{img} 的數值來估計規則度[11]，我們是用一個非等距的區間值[3 6 15]來給予紋理影像規則度，也就是說：

$$\text{令 } v_l = 0, \text{ 當 } M_{img} < 3$$

$$\text{令 } v_l = 1, \text{ 當 } 3 \leq M_{img} < 6$$

$$\text{令 } v_l = 2, \text{ 當 } 6 \leq M_{img} < 15$$

$$\text{令 } v_l = 3, \text{ 當 } 15 \leq M_{img}$$

3.3 瀏覽方法 (Browsing Method)

輸入我們所想要瀏覽的影像方向性、規模度或是規則度，即可從每張影像的 TBD 中找到符合條件的影像。

$$\text{TBD} = [v_1 v_2 v_3 v_4 v_5]$$

v_1 表示規則度， v_2 、 v_4 表示方向性， v_3 、 v_4 表示紋理大小，每個參數所代表的含意如表 3.1-3.3

規則度	表示意義
0	不規則
1	有點規則
2	規則
3	十分規則

表 3.1 規則度所代表的意義

方向性	表示意義
0	沒有方向性
1	0°
2	30°
3	60°
4	90°
5	120°
6	150°

表 3.2 方向性所代表的意義

紋理大小	表示意義
0	細小的
1	中等的
2	粗大的
3	十分粗大的

表 3.3 紋理大小所代表的意義

3.4 實驗結果 (Experimental results)

和 2.4 節使用相同的資料庫 Brodatz[9]，由 112 張基礎影像延伸出 1008 張的紋理影像資料庫。

圖 3.1 顯示了影像不同的規則度，數值越大則規則度越高，反之，則影像越呈現亂數分佈。

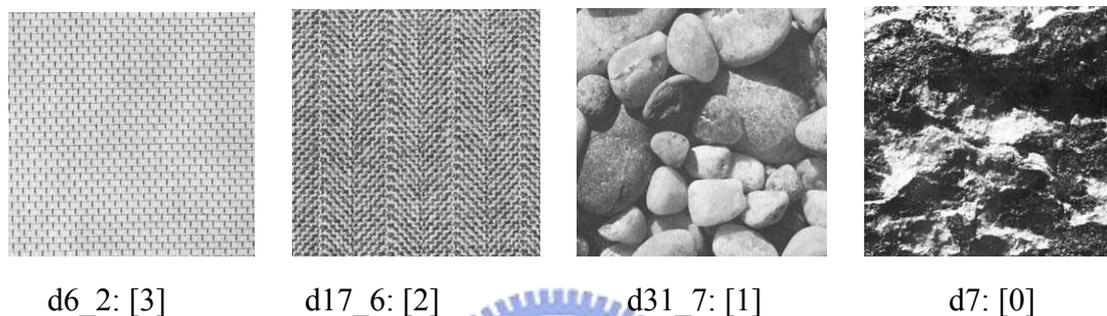


圖 3.1 紋理影像的規則度

圖 3.2 顯示的每張影像所具有的兩個主要方向性，1 是代表 0° ，2 是代表 30° ，依此類推，若是兩個數值皆相同，則表示該影像只有一個主要方向性。

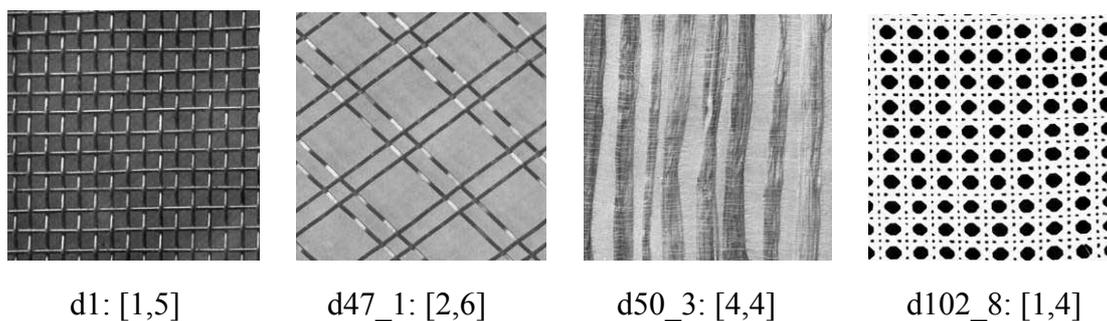


圖 3.2 紋理影像的方向性

圖 3.3 表示影像的紋理大小，也是在兩個主要方向上圖樣重覆出現的週期大小，數值越大紋理影像越粗糙，反之，則紋理影像越細小。

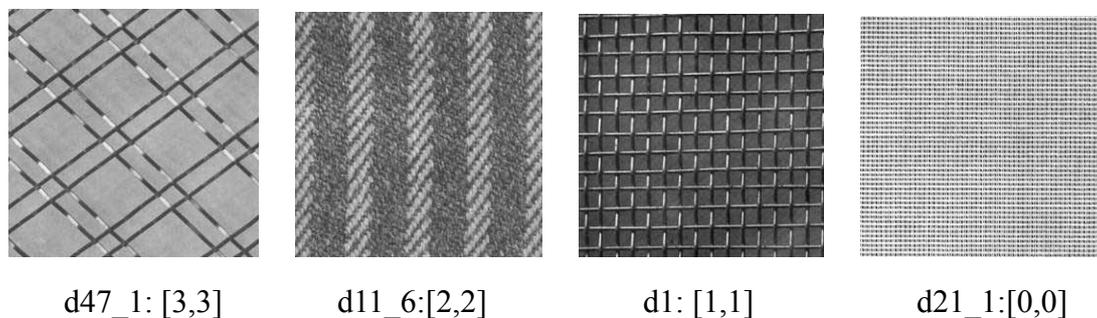


圖 3.3 紋理影像的大小

我們以規則度為 2，方向性為 4 (90°) 的條件，瀏覽資料庫中的影像之後，從中選擇 10 張顯示為圖 3.4。

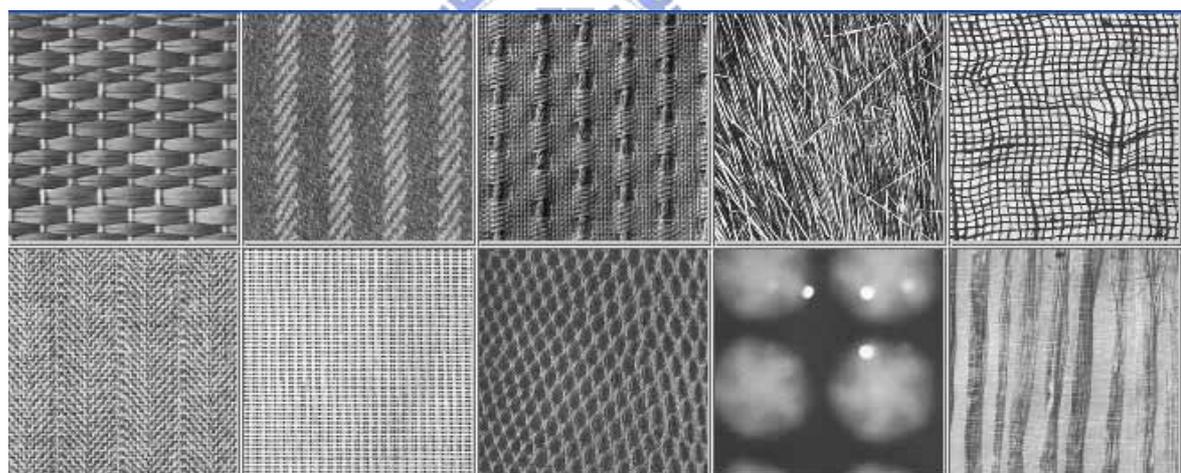


圖 3.4 令 $v_1=2$ 且 $v_2=4$ 或 $v_4=4$ ，利用 TBD 瀏覽的結果

第四章 邊界直方圖描述器 (Edge Histogram Descriptor)

4.1 定義 (Definition)

邊界直方圖描述器 (EHD, Edge Histogram Descriptor) 是以空間域中的邊界分佈情況來描述紋理影像[12,13]。描述器將紋理影像如圖 4.1 分割成 16 個子影像 (sub-image)，並將每個子影像中的邊界分佈情形以直方圖 (histogram) 表示。我們將子影像中的邊界歸類為五種，分別是四種方向性邊界 (垂直、平行、45° 對角線、135° 對角線) 以及一種無方向性的邊界，計算在子影像中各種邊界的出現次數，繪製成直方圖，所以 16 個子影像各有一個具有 5 種邊界型態的直方圖，這 16 個直方圖可以合併成一個具有 80 個代表區域邊界 (local edge) 的直方圖。

除了區域邊界的直方圖之外，EHD 還含有全域 (global) 和半全域 (semi-global) 邊界直方圖，全域直方圖是計算整張影像的各種邊界數目。半全域是將 16 個子影像以 13 種方法分群，再對每一種分群計算每種邊界的數目。所以 EHD 可表示為：

$$\text{EHD} = [h_l(0), \dots, h_l(79), h_g(0), \dots, h_g(4), h_s(0), \dots, h_s(64)]$$

$h_l(i)$ 為區域邊界直方圖， $h_g(i)$ 為全域邊界直方圖， $h_s(i)$ 為半全域邊界直方圖。

4.2 特徵萃取 (Feature Extraction)

為了萃取方向性邊界的特徵，我們必須將子影像再切割為更小的影像方塊 (image-block)，如圖 4.1 所示，將影像切割成不重疊的正方形區塊，並取得每個小區塊中的邊界資訊。

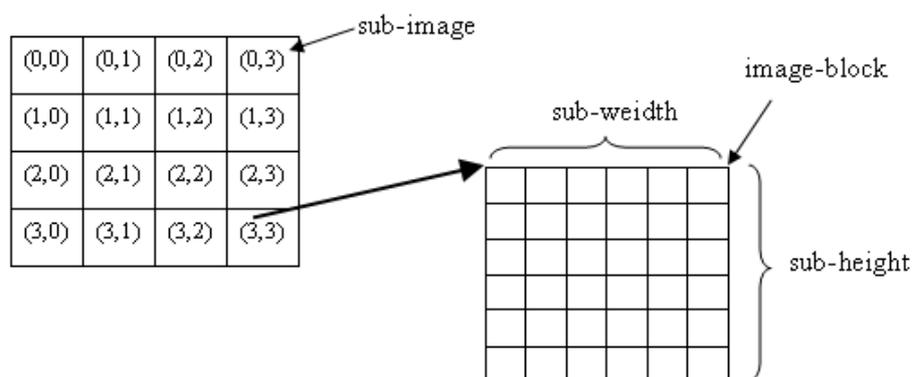


圖 4.1 子影像和影像方塊的定義

每個影像方塊的大小由我們所給定的方塊數量及影像大小決定，所以在不同大小的影像上，影像方塊的大小也有所不同，計算方法如下：

$$x = \frac{\text{sub_width}}{m}, y = \frac{\text{sub_height}}{m} \quad (4.1)$$

$$\text{block_x} = \left\lfloor \frac{x}{2} \right\rfloor \times 2, \text{block_y} = \left\lfloor \frac{y}{2} \right\rfloor \times 2 \quad (4.2)$$

sub_width 是子影像的寬度， sub_height 是子影像的高度，我們設定將子影像切割成 $m \times m$ 個影像方塊，(4.1) 式中算出來的 x 和 y 就是影像方塊的寬和高，但是為了後面邊界資訊的計算需要，影像方塊

的大小必須是 2 的倍數，所以在 (4.2) 式中將 x 、 y 調整為方塊長、寬 ($block_x$ 、 $block_y$)，使之成為 2 的倍數。

由前面所算出的方塊長寬將影像切割之後，對於每個影像方塊我們還需要將它切割成四個子方塊 (sub_block)，如圖 4.2，然後對於四個子方塊分別計算亮度平均值。

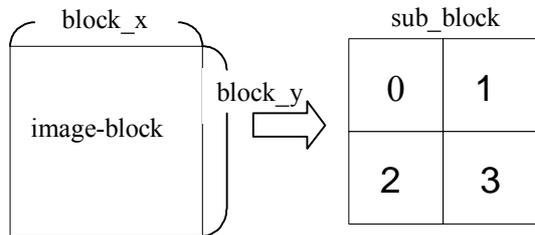


圖 4.2 子方塊的切割方式

在第 (i, j) 個影像區塊的第 k 個 ($k=0,1,2,3$) 子區塊中，計算其亮度平均值 $A_k(i, j)$ 。

對於每個影像區塊，我們使用圖 4.3 中的過濾係數來判斷影像區塊所屬的邊界型態，將過濾係數和影像區塊的子區塊亮度平均值對應相乘，而得到每種邊界型態的強度值。

1	-1	1	1	$\sqrt{2}$	0	0	$\sqrt{2}$	2	-2
1	-1	-1	-1	0	$-\sqrt{2}$	$-\sqrt{2}$	0	-2	2

a) ver_edge_filter b) hor_edge_filter c) dia45_edge_filter d) dia135_edge_filter e) nond_edge_filter

圖 4.3 邊界偵測的過濾係數

在第 (i, j) 個影像區塊的五種邊界強度值計算方式如下：

$$ver_edge_stg(i, j) = \left| \sum_{k=0}^3 A_k(i, j) \times ver_edge_filter(k) \right|$$

$$hor_edge_stg(i, j) = \left| \sum_{k=0}^3 A_k(i, j) \times hor_edge_filter(k) \right|$$

$$dia45_edge_stg(i, j) = \left| \sum_{k=0}^3 A_k(i, j) \times dia45_edge_filter(k) \right|$$

$$dia135_edge_stg(i, j) = \left| \sum_{k=0}^3 A_k(i, j) \times dia135_edge_filter(k) \right|$$

$$nond_edge_stg(i, j) = \left| \sum_{k=0}^3 A_k(i, j) \times nond_edge_filter(k) \right|$$

找出五個邊界強度值中的最大值，如果這個最大值大於或等於我們所選定的臨界值 (Th_{edge})，則這個影像區塊屬於最大值所對應的邊界型態，如果該最大值小於臨界值，則此影像區塊屬於無邊界的型態。

4.3 產生邊界直方圖 (Edge Histogram Generation)

在 4.2 節中，我們找出了每個影像區塊所屬的邊界型態。在此節中，我們將接著統計每個子影像中，屬於各個邊界型態的影像區塊個數成為直方圖。由 16 個子影像和 5 種邊界型態，我們可以得到 80 個區域邊界直方圖 (h_l)。

鑑於只使用區域邊界直方圖無法有效的達到檢索的目的，所以我們增加了全域和半全域的邊界直方圖來達到更精確的檢索結果。全域

邊界直方圖即為計算整張影像中屬於各種邊界型態的區塊個數，所以會有 5 個全域邊界直方圖 (h_g)。

另外我們還將 4x4 的子影像，依照垂直、水平和鄰近三種方法歸類成 13 種半全域影像，如圖 4.4 所示。然後計算每個半全域影像中各種邊界型態的直方圖，即可得到 65 個半全域邊界直方圖 (h_s)。

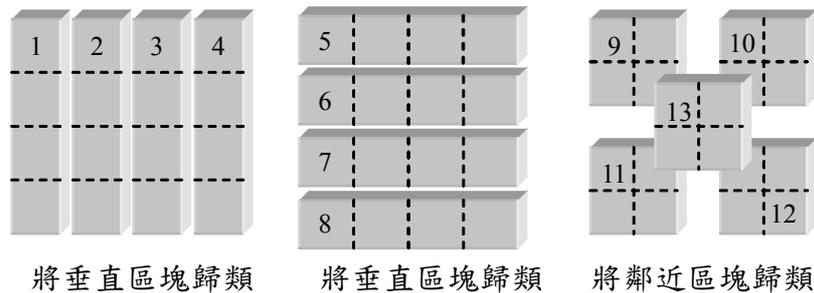


圖 4.4 將子影像分類為 13 種半全域影像

4.4 比較方法 (Matching Method)

在 4.3 節中，我們得到了一組 EHD，共 150 個邊界直方圖 (80 個區域邊界直方圖 + 5 個全域邊界直方圖 + 65 個半全域邊界直方圖) 來表示影像的邊界特徵，如圖 4.5 所示。

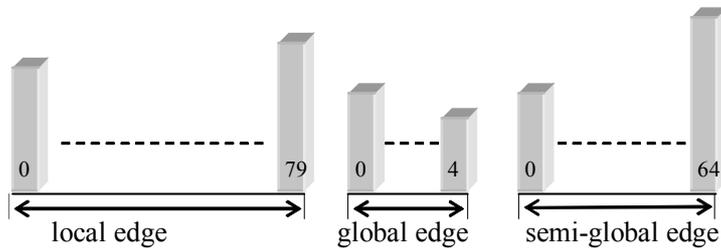


圖 4.5 所有的邊界直方圖

而我們比較兩張影像的方法就是將兩組 EHD 對應相減，以求出兩張影像間的差異度，但是考慮到只有 5 個全域邊界直方圖，在全部

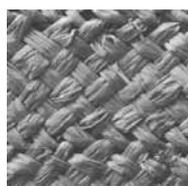
150 個邊界直方圖中的影響度太低，所以在全域邊界直方圖前加上 5 倍的權重，增加它的影響度。給定兩張影像 A、B，其差異度 $D(A,B)$ 計算為：

$$D(A,B) = \sum_{i=0}^{79} |h_l^A(i) - h_l^B(i)| + 5 \times \sum_{i=0}^4 |h_g^A(i) - h_g^B(i)| + \sum_{i=0}^{64} |h_s^A(i) - h_s^B(i)|$$

4.5 實驗結果 (Experimental results)

也是以同樣的資料庫 Brodatz[9]來作為檢索的依據，由 112 張基礎影像延伸出 1008 張的紋理影像資料庫。下面分別是幾個檢索的結果，只要有明顯邊界出現的影像，大致上都可以找到相似的紋理影像。

圖 4.6 是以 EHD 檢索的結果，與範例影像差異度最低的 10 張影像中，有 9 張是來自同一來源（包含範例影像自己），精確度為 0.9，檢索率為 0.89。



要檢索的範例影像

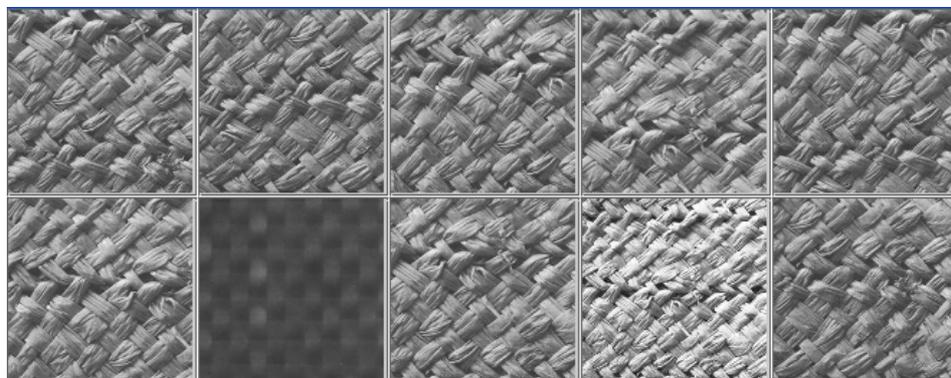


圖 4.6 以 EHD 檢索結果 (一)

圖 4.7 是第二個以 EHD 檢索的結果，與範例影像差異度最低的 10 張影像中，有 10 張是來自同一來源或相似的紋理影像（包含範例影像自己），所以精確度為 1，檢索率為 0.56。



圖 4.7 以 EHD 檢索結果 (二)

圖 4.8 是第三個以 EHD 檢索的結果，與範例影像差異度最低的 10 張影像中，有 9 張是來自同一來源或相似的影像（包含範例影像自己），所以精確度為 0.9，檢索率為 0.5。



圖 4.8 以 EHD 檢索結果 (三)

第五章 結論

MPEG-7 提供了統一的格式來定義紋理影像之特徵，只要所有的影像檔案中都含有 MPEG-7 的特徵資訊在裡面，就能在網際網路上做快速的影像資料檢索。但是為了能全自動且快速的產生這些特徵資訊，因此 MPEG-7 所提供的方法並不是檢索效果最好的一種。MPEG-7 所提供的三種方法皆有它的限制和缺點。

同質性紋理影像描述器因為使用賈伯過濾器，主要在萃取方向上的特徵，所以只適合使用在有明確方向性、規則度較高的紋理影像上，若是紋理排列混亂、規則度低的影像，檢索的效果就會較為不理想。

紋理影像瀏覽描述器因為將方向角度分為 6 個，以 30 度為一個區間，所以若是非 30 倍數的角度，像是 45 度，就會產生混亂的情形，有可能找到 30 度或是 60 度的結果，也因此會造成瀏覽結果的錯誤。

邊界直方圖描述器只適用於有明顯邊界分佈出現的影像，若是影像的紋理是較為平滑的分佈，就無法正確的判斷邊界型態，產生理想的邊界直方圖，而造成檢索的困難。

表 5.1 是三種描述器在各方面的比較。

	HTD	TBD	EHD
參數個數	62	5	80
所使用的特徵	在頻率域中，經過賈伯過濾器所得到的反應	在空間域中，經過賈伯過濾器所得到的反應之投影	在子影像中，邊界的分佈情形
應用領域	影像／視訊之檢索	影像／視訊之瀏覽	影像／視訊之檢索
適用對象	同質性的影像 織品的圖樣	同質性的影像 織品的圖樣	同質性的影像 自然影像 素描圖 剪貼圖

表 5.1 三種紋理影像描述器的比較



參考文獻

- [1] D. C. He and L. Wang, "Textural Filters Based on The Texture Spectrum," Pattern Recognition, Vol. 24, No. 12, pp. 1187-1495, 1991.
- [2] F. Liu and R. Picard, "Periodicity, Directionality and Randomness: World Features for Image Modeling and Retrieval," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 18, No.7, pp. 722-733, 1996.
- [3] B. S. Manjunath and M. Y. Ma, "Texture Features for Browsing and Retrieval of Image Data," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 18, No.8, pp.837-842, 1996.
- [4] H. C. Lin, L. L. Wang, and S. N. Yang, "Extracting periodicity of a regular texture based on autocorrelation functions," Pattern Recognition Letters, Vol. 18, pp. 433-443, 1997.
- [5] J. G. Leu, "On indexing the periodicity of image texture," Image and Vision Computing, Vol. 19, pp. 987-1000, 2001.
- [6] Text of ISO/IEC 15938-3 Multimedia Content Description Interface -Part 3: Visual. Final Committee Draft, ISO/IEC/JTC1/SC29/WG11, Doc.N4062, Mar. 2001.
- [7] MPEG-7 Visual Experimentation Model (XM), Version 10.0, ISO/IEC/JTC1/SC29/WG11, Doc. N4063, Mar. 2001.
- [8] Y. M. Ro, M. C. Kim, H. K. Kang, B. S. Manjunath and J. W. Kim, "MPEG-7 Homogeneous Texture descriptor," ETRI Journal, **23**(2), pp. 41-51, 2001.
- [9] P. Brodatz, Textures: A photographic Album for Artists & Designers, Dover Publications, New York, 1966.

- [10] A. Branca, M. Tafuri, G. Attolica and A. Distanto, "Directionality Detection in Compositional Textures," Proceedings of International Conference on Pattern Recognition, Vol.II, pp. 830-834, 1996.
- [11] P. Wu, B. S. Manjunath, S. Newsan and H. D. Shin, "A Texture Descriptor for Browsing and Similarity Retrieval," J. Signal Processing: Image Communication, 16, pp. 33-43, 2000.
- [12] C. S. Won and D. K. Park, "Image block classification and variable block size segmentation using a model-fitting criterion," Optical Engineering, 36(8), pp. 2204-2209, 1997.
- [13] D. K. Park, Y. S. Jeon, C. S. Won and S. J. Park, "Efficient use of local edge histogram descriptor," Proceedings of ACM International Workshop on Standards, Interoperability and Practices, Marina del Rey, pp52-54, California, USA, 2000.

