

乳房鈣化點形狀辨識系統

謝孟哲 徐麗蘋 魏聖原
中山醫學大學 中山醫學大學 中山醫學大學
醫學資訊學系 醫學資訊學系 醫學資訊學系
morearts@hotmail.com.tw apple@csmu.edu.tw s510668@hotmail.com

黃麒安 陳重光
中山醫學大學 中山醫學大學
醫學資訊學系 醫學資訊學系
alan7618778@hotmail.com mindplayer@livemail.tw

摘要

本研究針對乳房 X 光片影像中的鈣化點作形狀分類，進而開發出一套乳房鈣化點形狀辨識系統，由於鈣化點在乳房 X 光片影像中的面積通常很小，使一般現有的形狀判定分類技術都無法得到有效的結果，因此本研究搜尋出一些適用於鈣化點且有具體成效之形狀特徵方法，藉此建立出對應的決策樹模式，用以區分出不同形狀的鈣化點以輔助醫師對乳房鈣化點良惡性之判斷。本研究結果可直接加入含多個 ROI 乳房攝影影像搜尋系統，使其得以自動化運作且完整建置，讓醫師在診斷乳房 X 光片時，可以透過此系統找到相關的影像與資訊，以輔助醫師做正確診斷與治療之規劃。

關鍵詞：微鈣化、形狀特徵、不變矩、似圓性、乳房 X 光攝影影像

Abstract

When most mammograms have been digitized, how to effectively assist the doctors to diagnose a large number of materials rapidly has become an important task. Since the micro-calcifications are tiny, most of recognition methods are not suitable for them. This research focuses on developing techniques that extract and represent the shape features of calcification in an abstract and efficient way in mammographic images.

Because the shapes of calcification are important bases to diagnose that the calcification is benign or malignant, there are two main purposes of this paper. First, this paper is to recognize the shapes of

calcification automatically. Moreover, the paper is also to incorporate the system proposed here with the system of image information retrieval in mammograms, based on spatial relations established in last project we carried out. With those two systems, the whole system of image information retrieval in mammograms will be completely implemented and its performance will be automatic to facilitate the physician in diagnosing on breast cancer. There are several procedures carried out; first, we extract and represent the shape features of calcification in mammographic image containing multiple ROIs. Secondly, we establish a new classification method based on the decision tree model. Some experimental results we conducted indicate that the new recognition method of calcification in mammographic images is well-performed in terms of accuracy and time complexity. Finally, we believe that this paper will shed new light on the medical society in terms of mammogram image database and diagnosis on breast cancer.

Keywords: micro-calcification shape features, first moment, circularity, mammographic image.

1. 緒論

乳房 X 光攝影在偵測乳癌初期的微鈣化點症狀有非常好的辨識率，尤其在無法以觸診診斷或一些乳癌症狀較不明顯的狀況下，更能凸顯出乳房 X 光攝影偵測乳癌的優點。因此，醫生開始大量的使用乳房 X 光攝影進行乳癌的偵測，當檢測數量越來越多時，同時乳房 X 光片的數量亦快速的增加，在無形之中增加了

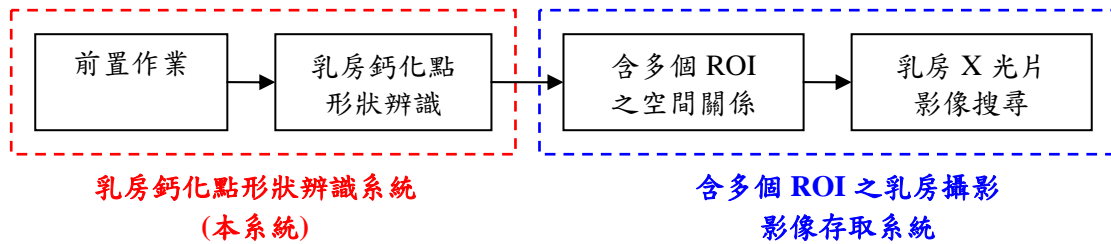


圖 1 含多個 ROI 乳房攝影影像之自動搜尋系統架構圖

醫師在檢驗影像上的負擔，如此一來會造成醫師作出錯誤的決定的可能。當醫生在診斷病情，需要相關或類似的乳房 X 光片作為參考之用時，會因為影像數量的龐大，造成尋找類似乳房 X 光片的時間拉長，效率不佳且浪費時間。因此，如何有效的讓醫師在診斷時，取得符合要求的乳房 X 光片影像作為診斷之輔助以減少誤判，是非常重要且迫切的。由於乳房鈣化點面積過小，使以一般內容為基礎的影像擷取技術 (CBIR, Content-Based Image Retrieval)[2,3] 中的低階影像特徵如色彩 (Color)[4,5]、紋理(Texture)[25]、形狀(Shape)[6,7] 等方法皆不適用，有鑑於此，學長們研製了一套使用空間關係作為影像特徵的含多個 ROI 之乳房攝影影像存取系統，使用時需以人工方式標記各個乳房鈣化點 ROI 的位置與形狀，才能進行影像查詢。由於使用人力標記鈣化點的方法相當不便且費時，為了改善這點，本論文提出一個乳房鈣化點的形狀辨識系統，使系統可以完全自動化對乳房 X 光片影像進行鈣化點的形狀辨識。在形狀辨識上，主要可以分為以區域為基礎的描述 (Region-Based Description) 及以邊緣為基礎的描述 (Boundary-Based Description) 兩大類，但因鈣化點面積過小，造成鈣化點在形狀辨識時，使用這兩大類中的方法大多都無法有效取得準確的特徵值進行形狀辨識，所以本論文的主要目的，為研究出適用於鈣化點的形狀分類方法，使系統可以準確分類出鈣化點的形狀，以輔助醫師針對乳房鈣化點良惡性之判斷，另一方面，本研究結果可直接引入含多個 ROI 乳房攝影影像搜尋系統使其得以自動化運作且完整建置(如圖 1)，以作為醫生進行正確的診斷及治療規劃的輔助。

2. 文獻探討

2.1 以內容為基礎的影像擷取 (Content-Based Image Retrieval)

近年來，電腦與網路快速發展普及，使數

位影像的產生與使用高速增加，以往利用文字建置資料庫、查詢檢索之方式 (Text-Based Image Retrieval, TBIR)，已無法滿足使用者在視覺感官上之使用需求。傳統的檢索方式是以文字為主，影像在儲存時會以文字型式同時儲存檔名、標題或關鍵字等與影像相關的描述，並將這些文字當作索引，使用者輸入文字在資料庫中搜索影像時，便能找到與文字相匹配的影像。如此，所有圖片都需要以人工方式建立索引，對日益龐大的影像資料庫來說，無疑是相當費時費力且相當主觀的事情，常會因為影像描述者的不同，對相同影像做出不同的關鍵性註解，影響使用者輸入文字搜索影像的結果。為瞭解決文字查詢產生的搜索盲點，也有專家提出加入關連性回饋 (Relevance feedback techniques) 的方法 [1]，但使用單純的文字進行描述，確實無法給予使用者較直觀的視覺感受。因此，許多的研究工作者開始研究以影像內容 (image content) 為查詢基礎的影像資料庫搜尋系統 (Content-Based Image Retrieval, CBIR) [2,3]，用以取代原先以文字為基礎的查詢方法。

CBIR 以影像取代文字當作輸入，藉由擷取輸入影像的一些特徵值或向量，與資料庫中儲存的影像比對相對特徵並產生相似度值，然後將相似值高的影像及相關資料回傳給使用者。如此，CBIR 在適當的特徵擷取中，便能減少視覺感官與文字描述之間的差距，而影像資料的特徵主要可分為高階與低階特徵，低階部分使用顏色 (color)[4,5]、形狀 (shape)[6,7]、紋理 (texture)[25] 等當作影像特徵，高階部分採用如：空間關係，即兩兩物件在平面上的分佈狀況，可用 Spatial Reasoning (空間推理)、Pictorial Query (圖像搜尋)、Similarity Retrieval (圖形相似擷取) 三種查詢方式來判斷圖形在空間的狀況。針對這些特徵進行資料庫的建置、檢索或是查詢，可使得影像的處理更具客觀性，藉以取代文字為基礎的查詢方法。

2.2 鈣化點之偵測

目前在偵測乳房微鈣化組織的議題上已有許多研究進行探討，Lin 採用灰階角度分析對整張乳房 X 光影像進行 ROI 區塊切割[7]，並且使用多碎型分析進行 ROI 的初步篩選。邊緣偵測部份以 SOBEL 邊緣檢測配合區域動態門檻值使鈣化點的邊緣更容易被偵測到。最終使用梯度角度計算來檢測鈣化點是否為一個封閉曲線，用以作為辨識鈣化點的依據。實驗結果在辨識方面有不錯的成效，但容易出現因為有特殊亮點而忽略鄰近鈣化點的情況。

林子平提出偵測影像鈣化點的方法[13]，首先使用 SOBEL 來檢視影像明暗強弱，便能針對微鈣化與附近組織灰階值差異較大之特性。再來利用二值化處理，將邊緣特徵較明顯的部分篩選出來。另外雜訊及亮點部分，可利用數學型態學的侵蝕運算消除掉；細微接縫處能利用擴張運算來填補，使得鈣化點偵測正確率提高。實驗結果在鈣化點偵測敏感度上較佳，準確度上效果不彰。

T.Balakumaran 和 ILA.Vennila[24]藉由第三階矩(偏斜度)和第四階矩(峰度)找出 ROI，接著利用小波轉換將訊號依細節與相近度分為不同層，再藉由小波工具分為多個可辨識的組件，最後經 Foveal Algorithm 使得各別組件容易被提取。由於鈣化點經由 X 光照射會呈現微小顆粒亮點，這些亮點作為奇點，它的訊號頻率較高，故可藉由小波轉換提取出鈣化點。

2.3 鈣化點形狀辨識

目前辨識物件形狀之特徵主要分為以區域為基礎的描述(Region-Based Description)及以邊緣為基礎的描述(Boundary-Based Description)兩大類。而目前絕大部分的形狀辨識方法皆不適用於微小面積之物件。由於鈣化點皆非常的微小，因此研發出適用於鈣化點形狀的特徵表示法是非常需要的。

2.3.1 以區域為基礎的描述(Region-Based Description)

以區域為基礎的描述，可敘述二維物件的圖元(pixels)分配狀況，不僅能處理單一封閉的影像內容物件外形，也能處理一個由多個不相關物件所組成的複雜影像資料。以區域為主的

描述子，以處理物件區域為主，所以對非影像內容物件區域的雜訊及獨立點圖元，也有處理與隔離的能力。

以區域為基礎的描述方法有 MPEG-7 ART[9]、多階層邊界特徵向量(multi-layer eigenvector 簡稱 MLEV)[9,10]、Zernike moment[8-10]。

<1>MPEG-7 ART[9]是 MPEG-7 所提出以影像內容物件外形區域為主的徑角轉換(Angular Radial Transformation 簡稱 ART)描述子。ART 描述子是一個定義在極座標系統的單位圓盤(Unit disk in polar coordinates)的 2-D 複數形態的轉換。

ART 特徵適用於檢索對稱性較好的物件，常被用在圖形中有空缺或是由多個形狀所組合成的圖形如商標。但在現實生活中並非所有物體都有對稱性，在非對稱形狀的檢索上，效果較差。

<2>多階層邊界特徵向量(Multi-layer eigenvector)表示法簡稱 MLEV[9,10]，是屬於以影像內容物件外形區域(Region-based)為主的表示法。將影像物件外形區域所有圖元(pixel)位置，均視為二維向量；利用影像物件區域所有圖元二維向量資料集，計算其物件質量中心(center of mass)，再利用共變矩陣(Covariance matrix)求得整個影像物件外形區域的二個特徵向量(eigenvector)。

二個特徵向量表示整個物件區域圖元位置相對於質心變異量最大及最小的二個方向。接下來利用二個特徵向量切割整個影像物件區域，得到四個次區域(sub-regions)，再利用上述方法切割四個次區域，每個次區域得到四個次次區域。在切割過程中獲得四個參數(parameter)，四個參數滿足旋轉、平移、映射及等比例放大或縮小屬性。

<3>Zernike 多項式是無窮多項式的完全集(complete set)[8-10]，而這些多項式在單位圓內為彼此正交，因此透過利用 Zernike moments，能夠取得對於圖像旋轉不變性的特徵，主因就在於 Zernike moments 之特徵為相互正交，使得擷取出之特徵為相互獨立，並且能取出對於圖像較有代表性之全域特徵。

在計算 Zernike moments 分為三個步驟，第一是計算影像的徑向函數(radial polynomials)，接著計算 Zernike moments 的基底函數如公式，最後一步驟則是將影像投影到 Zernike 的基底函數。

Zernike moments 如需區別出極相似的影

像特徵時，必須提高計算之階數以提高特徵擷取效能，反而造成所花費的時間大幅增加，因此在很多影像辨識的研究上 Zernike moments 往往只當作是輔助的角色。

2.3.2 以邊緣為基礎的描述 (Boundary-Based Description)

傅立葉描述子 (Fourier Descriptor) [16,18-20]的方法首先由 Zahn T.和 Roskies R.Z.提出[15]，而之後也有研究者推出一種基於面積函數和一種基於弦長函數的傅立葉描述子 [17]傅立葉描述子使用在描繪閉合圖形的輪廓，我們把圖形中的輪廓視為一個封閉的曲線，並在二維 xy 坐標平面上以曲線上任意點為起始點，沿著曲線逆時針方向或順時針方向追蹤一圈，得到 $s(k)=[x(k),y(k)]=0,1,2,\dots, N-1$ 的座標序列，而我們把這些二維座標序列轉化為複數座標，即式(1)

$$s(k) = x(k) + jy(k) \dots\dots\dots (1)$$

其中 $k = 0,1,2,\dots,N-1$ ，這種將目標形狀的輪廓線表示成輪廓線函數成為形狀特徵圖 (shape signature)。計算這 N 個點的離散傅立葉變換係數 作為傅立葉描述子 (Fourier Descriptor)，即式(2)

$$a(u) = \frac{1}{N} \sum_{k=0}^{N-1} s(k) e^{-j2\pi uk / N} \dots\dots\dots (2)$$

由於 $a(u)$ 是由一個連續的座標序列所產生的傅立葉描述子，因此也可以利用質心距離特徵來做為轉換傅立葉描述子之依據，方法為將原序列之 $s(k)=[x(k),y(k)]$ 視為 $s(k)=[CCDK,0]$ 。當計算出傅立葉描述子 $a(u)$ 後，利用式(3)

$$a(u) = [a(u)] e^{-j\theta(u)} \dots\dots\dots (3)$$

將 $a(u)$ 轉換為極座標，如式(4)

$$|a(u)| = \sqrt{F^2(u) + G^2(u)} \dots\dots\dots (4)$$

$|a(u)|$ 稱為傅立葉轉換之頻譜 (Spectrum)

就是他的大小 (Magnitude)，且轉換後的傅立葉描述子為前後對稱，因此只需使用 $u/2$ 個描述子即可重現整個傅立葉序列。

鏈碼 (Chain Code) [21-23] 是由 Freeman 首先提出，常用於曲線編碼或是物體邊界的判知，Freeman 鏈碼使用了八個方向值作為八鄰接點，一個經過二值化的影像輪廓曲線，由曲線起始點 a_1 順著輪廓曲線到終點 a_n ，表示各個點座標及其相對位置關係如圖 2，

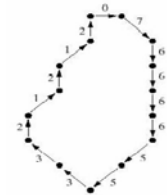
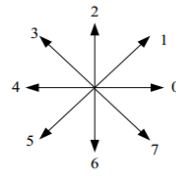


圖 2 各個點座標及其相對位置關係

曲線長度也可由下列方程式得到式(5)

$$L = K(n_e + n_o \sqrt{2}) \dots\dots\dots (5)$$

其中 K 為單位長度之常數； n_e 為鏈碼中偶數鄰接點個數； n_o 為鏈碼中奇數鄰接點個數；採用鏈碼表示一個影像輪廓比一般使用矩陣表示影像節省更多位元元，影像越大或曲線在影像所佔的比例越小越適合使用鏈碼來表示可使數據量大為減少。

以鈣化點的形狀分類來說，因為乳房鈣化點在影像上面積過小，使得不論以區域為基礎 (Region-Based) 或是以邊緣為基礎 (Boundary-Based) 的絕大多數方法，並不適用於鈣化點上，因為無法取得小面積形狀較為準確的特徵值，使得形狀的分類不易，因此需要另外研究搜尋出其他能適用於鈣化點上的方法。

2.3.3 適用小面積的特徵表示法

余信賢在 2005 年提出一個使用決策樹的分類方法 [26]，主要採用似圓性、不變矩、距離角度標記法、閉合運算，對鈣化點影像進行運算當做特徵值，並以貝氏分類器進行鈣化點影像分類，因為不同特徵值對圖型的分類效果不同，所以使用分類演算法加以輔助，分類出圓形 (C)、五角形 (P)、三角形 (T)、矩形 (R)、長橢圓形 (E)、V 字形 (V)、海鷗形 (G)、星形 (S)、長條形 (B)。實驗中在分類上都有相當高的準確率，但在一些邊緣角度不明顯的鈣化點上，有可能出現誤判的情形。

2.4 乳房鈣化點之醫學診斷相關資訊

乳房細胞產生鈣化是一種現象或結果，並非一種特定疾病。且高達八成五的乳房鈣化為良性，只有一成五的鈣化可能會合併癌症。即是鈣化不一定為乳癌，乳癌也不一定會出現鈣化。一般來說，將會根據鈣化的形狀特徵來判斷是否會產生惡性腫瘤。

在乳房內的組織細胞產生了不正常的分

裂、繁殖，形成的惡性腫瘤，稱之為乳癌。乳癌產生之後，不只會影響局部身體器官，也會有機會轉移至其他部位，影響身體狀況，導致破壞身體器官的正常運作，更嚴重將至危害到生命。細胞在乳房組織內的變化，與乳癌的形成密切相關。當在乳房組織內有出現微鈣化點，即是乳癌早期發現的指標，微鈣化點偵測便是為此檢測的一項技術。

2.4.1 乳房病理學

乳房組織主要由乳小葉、輸乳管、脂肪組織、血管及淋巴管所構成的，大部分乳癌是乳管(80%)所產生的，小部分來自乳小葉(10~15%)，極少部分是脂肪組織、血管及淋巴管所發生。

乳癌的類型主要可分為：乳房原位癌、浸潤性腺管癌、浸潤性小葉狀乳癌

乳房原位癌可分成乳管內癌 (DCIS) 和原位小葉癌 (LCIS)，其中以乳管內癌(DCIS)較為常見，被認為屬前癌性或乳管不正常細胞早期生長階段，約占乳腺癌的 5~10 %。原位性乳癌的診斷意義，是指癌細胞為非侵略性的，亦不會轉移。

浸潤性腺管癌是最常見的侵略性乳癌，佔所有乳癌百分之 70~80 %。浸潤性乳癌的臨床特徵，在乳房周邊會有不整齊的硬塊，從乳房放射線攝影等影像檢查所見，在腫瘤部位常發現微形鈣化。

浸潤性小葉狀乳癌約佔所有惡性、侵略性乳癌的 5 %。此類腫瘤在乳房放射線攝影中不容易顯現，且易有多發性的傾向。

乳房的性質可分為緻密型乳房(dense-glandular)、脂肪型乳房(fatty)及線狀脂肪型乳房(fatty-glandular)三類。在 X 光片中，緻密型乳房以白色顯現。脂肪型乳房看起來是灰色偏透明。腺體脂肪型乳房的乳房組織逐漸減少，脂肪組織逐漸增多，在 X 光影像中為兩者各佔一半。

2.4.2 醫學上 12 種鈣化點分類

醫師在乳房 X 光中所圈出的 ROI 鈣化點是乳癌初期的重要特徵之一，在 1997 年美國全國乳癌與子宮頸癌早期偵測計劃(The National Breast and Cervical Cancer Early Detection Program)中指出乳房 X 光攝影是早期偵測乳癌最有效的方法，乳房微鈣化點

的由來是因為乳汁的主要成分為鈣，而乳汁在沉澱後產生硬化或組織壞死而產生微鈣化點，微鈣化在乳房影像中的特徵是以細微的亮點顯現出來，通常是以單一或叢集方式呈現，簡單的辨認方式為灰階值較高之部分(較亮)，目前醫學上將微鈣化分為 12 種(如圖 3)，前 10 種為良性，後 2 種為惡性。其鈣化的各種說明如下[14]：



圖 3 微鈣化圖例

- 1、管腺內碎片之鈣化
- 2、脂肪壞死之微鈣化(中心透光)
- 3、小囊腫中的沉澱物鈣化
- 4、小葉囊中的結石
- 5、囊腫之邊緣鈣化
- 6、纖維性瘤之早期沉澱物
- 7、纖維性瘤的大沉澱物
- 8、血管鈣化
- 9、皮膚鈣化
- 10、分泌器官疾病的鈣化
- 11、腺內癌之多形沉澱
- 12、粉刺癌中細線形的鈣化

3. 研究方法

3.1 似圓性(Circularity)

透過圖形的邊界與面積的關係來計算出圖形接近圓形的程度，來求出二值化圖形中的似圓性。如圖 4 所示，我們使用各個不同的(2x2)的二值化單元遮罩與影像中的鈣化點進行比對計算，並統計出不同遮罩在影像中出現的次數，作為計算面積與邊長之用。其中，在計算時是以遮罩做為單位，一個遮罩邊長為 4，面積為 1，如圖 5。

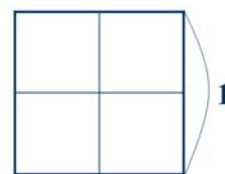


圖 5 遮罩計算

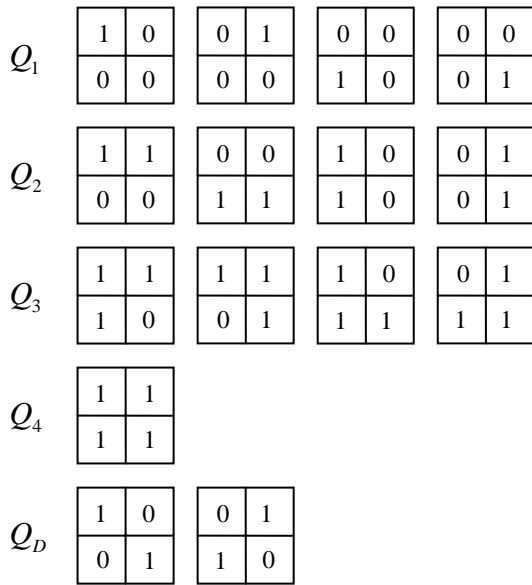


圖 4 單元遮罩

周長公式為：

$$P = [n\{Q_1\} + n\{Q_2\} + n\{Q_3\} + 2n\{Q_d\}] \dots\dots\dots (6)$$

面積公式為：

$$A = \frac{1}{4} [n\{Q_1\} + 2n\{Q_2\} + 3n\{Q_3\} + 4n\{Q_4\} + 2n\{Q_d\}] \dots\dots\dots (7)$$

根據上述所求出的參數，我們可以定義出物件的似圓性(Circularity)為式(8)

$$Cir = \frac{4\pi A}{P^2} \dots\dots\dots (8)$$

實驗中發現，鈣化點影像在進行平移、大小比例調整時，似圓性值變化不大，但在旋轉時卻會有差異，此問題在矩形尤其明顯，如表 1，原因是圖型進行旋轉後，如圖 6 所示，在計算邊長與面積時，像素的排列造成傾斜的邊長長度遠大於垂直的邊長長度，使似圓性值不穩定。

表 1 旋轉測試

	原圖	旋轉 45 度
測試影像		
似圓性值	0.7854 (P=118, A=870.25)	0.3654 (P=173.5, A=875.25)

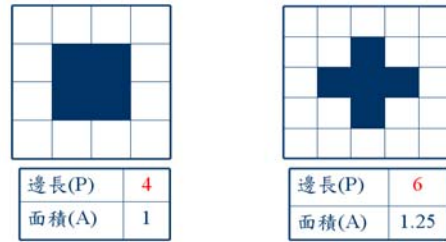


圖 6 遮罩計算

針對旋轉所造成的似圓性變化大的問題，我們在圖形進行似圓性計算前進行正規化處理，作法是取出圖形長軸，並統一將長軸旋轉成與 X 軸平行，藉以解決因旋轉造成的似圓性值變化。

在所有檢測圖形裡，其似圓性分佈情形，如圖 7。

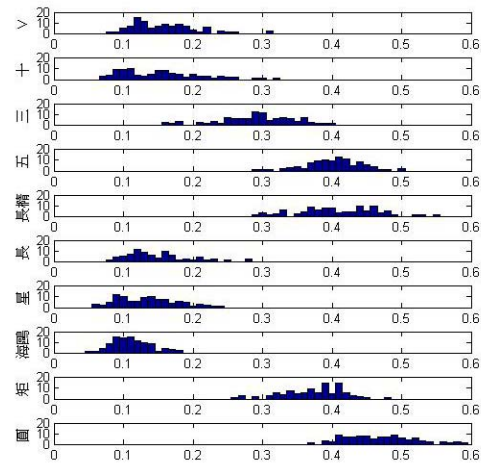


圖 7 鈣化點形狀其似圓性分佈圖形

似圓性數值界於 1~0 之間，其中越接近圓形的圖形，似圓性的值越接近 1，似圓性對於具有明顯內凹角的形狀有很強的區分能力，本研究依統計將十種鈣化點形狀透過似圓性的值分成四類：

- 1、圓形
- 2、圓形、橢圓形、五角形、矩形
- 3、三角形、橢圓形、五角形、矩形、十字形、V 字形
- 4、三角形、星形、海鷗形、十字形、長條形、V 字形

3.2 不變矩(Moment Invariants)

二值化影像 B(m,n)區域的(p+q)階矩可以定義為式(9)

$$M_{pq} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} m^p n^q B(m,n) \dots\dots\dots (9)$$

定義完 B(m,n)的矩之後，同理可定義該二值化影像的中心矩為式(10)

$$\mu_{pq} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (m-\bar{x})^p (n-\bar{y})^q B(m,n) \dots\dots\dots (10)$$

其中

$$\bar{x} = \frac{M_{1,0}}{M_{0,0}}, \bar{y} = \frac{M_{0,1}}{M_{0,0}} \dots\dots\dots (11)$$

而正規化的中心矩，其定義為式(12)

$$\eta_{p,q} = \frac{\mu_{p,q}}{\mu_{0,0}^r} \dots\dots\dots (12)$$

其中

$$r = \frac{p+q}{2} + 1 \dots\dots\dots (13)$$

而正規化的中心矩，可推導出式(14)(15)

$$\phi_1 = \eta_{2,0} + \eta_{0,2} \dots\dots\dots (14)$$

$$\phi_2 = (\eta_{2,0} - \eta_{0,2})^2 + 4\eta_{1,1}^2 \dots\dots\dots (15)$$

不變矩的特性是不會受到平移、旋轉及大小比例改變時影響，但是有時候不變矩的數值動態範圍很大，因此計算不變矩的時候，通常使用。不變矩可以判斷圖形的分佈情形，如果圖形的分佈較趨於質心的兩端時，不變矩的數值會越大，例：V 字形、長條形、海鷗形；當圖形的分佈較接近質心時，則數值越小，例：圓形、矩形、五角形、橢圓形。

例如：帶入式(14)，各鈣化點形狀之不變矩值如表 2。

表 2 各鈣化點形狀之一階不變矩值

				
V 字形	十字形	三角形	五角形	長條形
-0.9904	-1.2658	-1.4914	-1.7644	-0.4925
				
橢圓形	星形	海鷗形	矩形	圓形
-1.6503	-1.3586	-0.8699	-1.711	-1.8309

在所有檢測圖形裡，其實際分佈情形(如圖 8)。

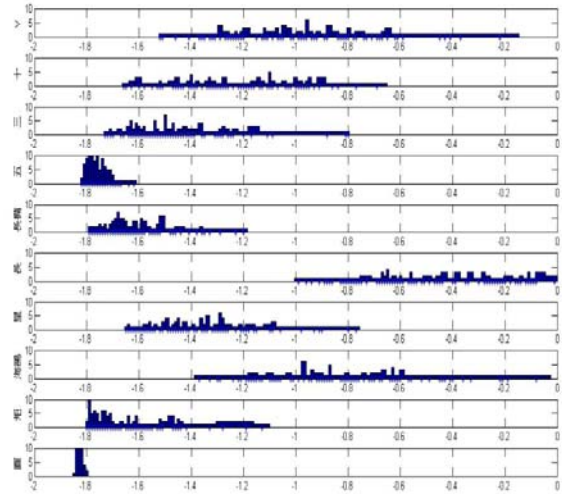


圖 8 各鈣化點形狀之一階不變矩分佈圖

由上述分佈圖我們將鈣化點形狀分為 5 類

- 1、圓形、五角形
- 2、矩形、橢圓形、五角形、三角形、十字形
- 3、矩形、星形、橢圓形、五角形、三角形、十字形
- 4、矩形、海鷗形、星形、橢圓形、三角形、十字形、V 字形
- 5、海鷗形、星形、長條形、三角形、十字形、V 字形

在這 10 種鈣化點形狀中，我們可以清楚知道圓形、矩形和五角形的質心處於中心，其不變矩值小，易於和其他圖形區別開來。但對於 V 字形、十字形、三角形、長條形、星形還有海鷗形來說，相互重疊部分相當多，導致區分結果並不順利。

另外我們亦使用第二不變矩(帶入式(15))進行測試得到圖 9。

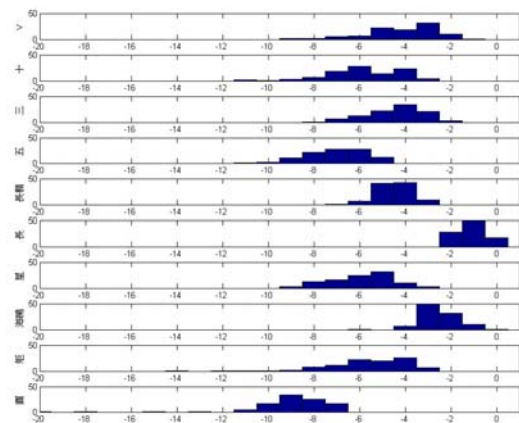


圖 9 各鈣化點形狀之二階不變矩分佈圖

大多數鈣化點形狀的第二不變矩值重複不分相當密集，導致分辨率不佳，所以在本論文中只採用第一不變矩的部份進行分析。

3.3 距離角度標記法(Distance-Angle Signature)

距離角度標記(Distance-Angle Signature, 簡稱 DAS), 在作標記法之前會將原圖做正規化, 其目的是為了調整圖片的大小, 並讓所有圖片大小作統一, 正規化的處理方式為, 先將圖片中的鈣化點質心與邊的最大距離 d 取出, 再將圖片進行縮放。其縮放倍率為: $25/d$ 。選擇縮放倍率為 $25/d$ 的原因乃依據鈣化點的一般大小使縮放倍率不會太大, 目的在使此鈣化點之大小正規化為其質心與邊之最大距離為 25, 如圖 10。

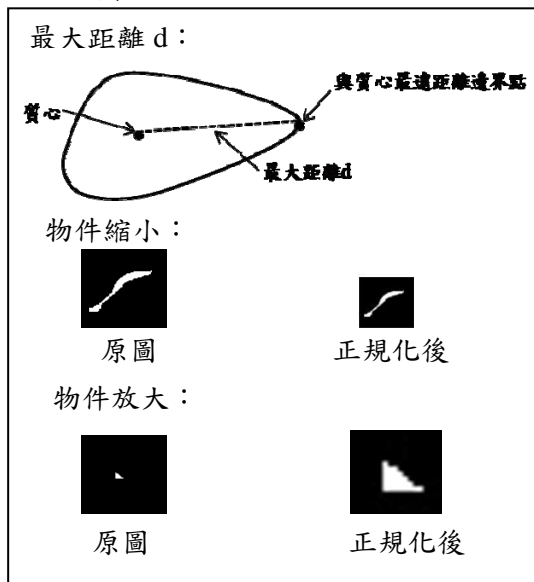


圖 10 正規化前後之圖形

將作完正規化的圖片再作後續處理, 如圖 11 所示, 首先將物件邊緣上各點與物件質心之距離作蒐集, 蒐集下來的距離資料從物件最右方開始順時鐘排列, 並在圖表上呈現, 再利用圖表中波峰的個數區分物件的種類。

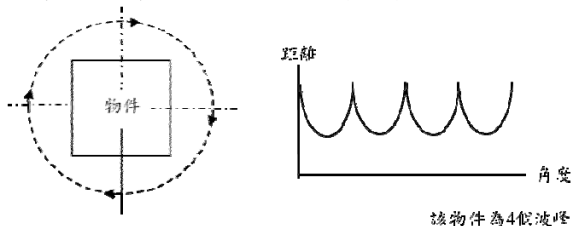


圖 11 利用斜率的標記法示意圖

3.3.1 利用斜率變化排序 :

首先我們利用斜率的方法來作處理, 將所有邊緣點與質心點的距離和斜率記錄下來, 在利用斜率變化排序, 在物件質心位置在物件區域之外部的圖片效果不佳, 如: V 字形、海鷗形、十字形、長條形。以上圖形都會產生同時偵測到兩個邊的距離, 會有許多重複的斜率, 在這邊將同斜率距離較小的值消去, 只保留最大值, 如圖 12。

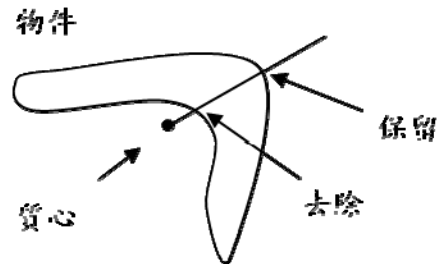


圖 12 利用斜率的標記法重複擷取值處理示意圖

質心位在物件區域之外部的圖形上會發生兩種狀況 :

- 1、一是上述提到的處理方式, 會將同一斜率非同一點的最大值留下其餘較小值被去除, 即是只取一邊的值。
- 2、另一是當斜率是鄰近但在圖片上並非連續的點將會殘存在結果上, 導致圖表會產生連續波動的值如圖 13, 進而引響判斷結果。

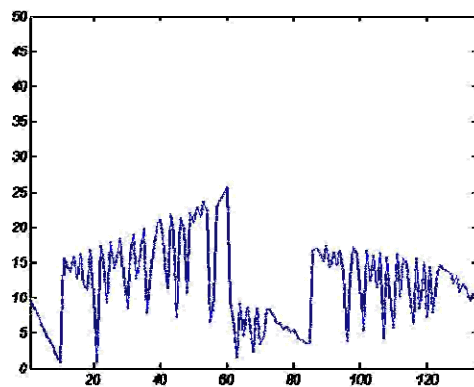


圖 13 質心在物件外的圖形利用斜率標記法時圖表會產生連續波動的值

3.3.2 利用角度依序擷取 :

因為以斜率進行擷取時會遇到以上問題我們嘗試改用以角度擷取鈣化點的方法做處理, 另外以角度為擷取鈣化點的方法做處理,

能夠將上面斜率所產生的問題解決，角度的方法在重複值的處理方式會比斜率容易很多，並不會造成許多重複值或鄰近點不同斜率的問題，其所作的處理程式第一步與斜率的做法相同，也是先作相同的正規化，再後續處理。首先找出物件的質心位置，再以質心位置為軸心，整張圖進行順時針旋轉。旋轉的度數上若以每1度作旋轉會造成太多雜訊，一次作大變動的旋轉則會造成失真的後果，在測試階段時以一次角度為5度作旋轉較適宜。實作過程是先作一次擷取後，再5度旋轉，再作一次擷取來進行，直到360度結束為止。擷取的部分，利用一條通過質心位置的水平線段，擷取出最右邊外圍邊界位置的點作下紀錄，如圖14。

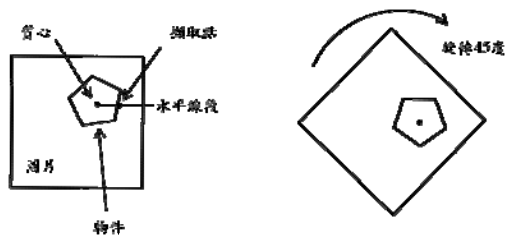


圖 14 擷取點與旋轉示意圖

隨後再將收集的擷取點，同樣也是從最右方起始點作順時針排序，在以圖表上作呈現。

經過評估後，我們選用角度的方式來做主要方法，其原因是利用內插值來為圖表上作平滑的處理，缺點是怕會有重點遺漏值，相反地其優點是在鈣化點質心屬於外部的圖形分辨較佳，其圖表的呈現，會比斜率的方式更加平滑，即是角度的方式比斜率的方式效果更好。

十種鈣化點形狀經距離角度標記法處理後其結果如表3：

表 3 10種鈣化點形狀分類

鈣化點形狀	波峰個數	鈣化點形狀	波峰個數
圓形	3~4	矩形	4
三角形	3	長條形	2
橢圓形	2	V字形	3
星形	5	十字形	4
五角形	5	海鷗形	2~3

因此依波峰個數可將鈣化點形狀分類如表4：

表 4 10種鈣化點形狀依其波峰個數分類

波峰個數	鈣化點形狀
2	海鷗形、長條形、橢圓形
3	海鷗形、三角形、圓形、V字形
4	矩形、圓形、十字形
5	星形、五角形

3.4 鈣化點質心線段與邊緣相交數

上述有談及到的，兩種標記法在相同角度或斜率時，會在直線上的不同位置出現物件與背景的值，利用計算物件邊緣的相交處出現次數作為鑑別度的分析處理。在斜率的部分上，由於是用遮罩進行擷取的，在計算邊緣相交點的部分時容易會重複擷取，而難以實行。因此決定使用以角度的方式來進行這項特徵處理。

使用此特徵主要是針對質心位置不確定在內或外的鈣化點物件，這些圖形是使用距離角度標記法較難辨識的圖形，例如：V字形、海鷗形、長條形。在進行處理時會有質心位置不同以及毛邊抖動影響計算的問題，我們將會使用一些影像處理的方法作前處理，再進行分析。

影像前處理部份，我們嘗試使用開啟、關閉運算還有骨架化作前處理。在使用開啟、關閉運算進行前處理時發現，使用較小遮罩時，並無法有效的去除毛邊，而使用較大遮罩時，容易將較細的物件消除造成形狀改變，且無法解決質心位置(物件內、外)產生的計算量差異，故不用該影像處理方法進行前處理步驟。使用骨架化後，可以使質心位置恆定在物件外部，解決的質心位置不穩定的狀況，在判斷處理上也較容易。

如圖15所示，對一張圖片計算其經過水準質心線段所偵測到的相交點個數，以改變旋轉角度的方式作整張圖的旋轉，其中旋轉時以1度作為單位，因此可以得到180組線段所擷取到的相交點數目，最後以最大數目為該張圖片作代表。然而，實驗中得到的相交點最大值沒辦法確切的將種類分類出來，因為在計算時容易遇到骨架化後的零碎雜訊，產生零碎的較大相交數，造成相交數不準確，因此經過實驗後，我們使用加權值計算後再作辨識。

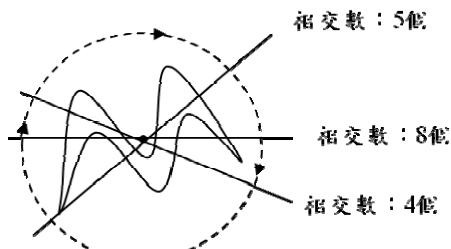


圖 15 鈣化點質心線段與邊緣相交示意圖

以下面例子來說明加權的計算方式：

例：圖片結果如下

相交點個數	次數
2	50
4	30
6	80
8	20

$$2 \times 50 = 100$$

$$4 \times 30 = 120$$

$$6 \times 80 = 480$$

$$8 \times 20 = 160$$

次數總共 180

故該圖片之相交數的代表值為 6

3.5 關閉運算(Closing)

數學形態學(Mathematical Morphology)在影像的處理中，常用來取得圖形的形狀特徵及消除雜訊，其中包含了擴張(Dilation)、侵蝕(Erosion)、開啟(Opening)、關閉(Closing)四個基本運算元，以遮罩對圖元進行運算，現以二值化之影像為例，將其演算法介紹如圖 16：

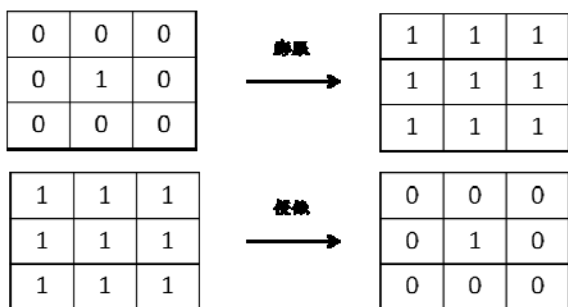


圖 16 膨脹侵蝕示意圖

- 膨脹運算：進行膨脹運算時，若該圖元點為白色時，則運算後會使其圖元點及其周圍圖元點皆變為白色。
- 侵蝕運算：進行侵蝕運算時，若該圖元及其周圍像速點皆為白色時，則運算後該圖元保

持白色，否則該圖元會變為黑色。

- 開啟運算：開啟運算是由侵蝕與膨脹兩運算元組合而成，處理過程中，會先進行侵蝕運算，再將結果做膨脹運算。
- 關閉運算：關閉運算是由侵蝕及膨脹兩運算元組合而來，但與開啟運算相反，處理過程為先膨脹後侵蝕。

在二值化的影像處理中，膨脹與關閉運算常被用來進行填補凹缺與強化邊緣，而侵蝕與開啟運算則常被用來消除雜訊與切斷較細的連接邊。





















我們的系統主要採用數學形態學的關閉運算，運用關閉運算可以將影像中的凹陷填滿的效果，對 V 字形、海鷗形、十字形等具有明顯內凹的圖形進行填補，再計算經過關閉運算後的面積與原面積的差異百分比，以此做為分類依據。由表 5 為不同形狀經過關閉運算後所得到的面積差異百分比，結果顯示會產生比較明顯差異的圖形主要為 V 字形、海鷗形、星形與十字形。

在實驗中，發現關閉運算在某些開口較開的內凹處會因角度關係無法有效填補，因此在進行運算時，我們讓影像每旋轉 15 度進行一次關閉運算，再以最大的面積差異百分比做為分類依據。如表 6 中，我們以每 30 度做一次關閉運算做為範例，發現在 30 度與 120 度時，面積差異百分比最大，故採用 13.07% 做為該圖的面積差異百分比。

表 6 不同角度時的面積差異百分比

旋轉角度	0	30	60
圖形			
與原面積的差異百分比	0%	13.07%	12.62%
旋轉角度	90	120	150
圖形			
與原面積的差異百分比	0%	13.07%	12.62%
最大面積差異百分比=13.07%			

表 5 不同形狀經過關閉運算後的面積差異百分比

形狀	圓形	三角型	矩形	長橢圓形	五角形
原圖形					
關閉運算後之圖形					
與原面積的差異百分比	0%	1.49%	3.87%	0%	0.55%
形狀	長條形	V字形	海鷗形	星形	十字形
原圖形					
關閉運算後之圖形					
與原面積的差異百分比	0%	10.71%	6.12%	20.34%	15.52%

3.6 以分類演算法(Algorithm of Classification)判斷鈣化點形狀

以下為各種方法區分開來的結果:

似圓性
 > 0.54 圓形
 0.40~0.54 圓形、橢圓形、五角形、矩形
 0.28~0.40 V字、十字形、三角形、橢圓形、五角形、矩形
 0~0.28 V字、十字形、三角形、長條形、星形、海鷗形
 不變矩
 < -1.8016 圓形、五角
 -1.8016 ~ -1.7324 矩形、橢圓形、五角形、三角形、十字形
 -1.7324 ~ -1.6236 矩形、星形、橢圓形、五角形、三角形、十字形
 -1.6236 ~ -1.1302 矩形、海鷗形、星形、橢圓形、三角形、十字形、V字

> -1.1302

海鷗形、星形、長條形、三角形、十字形、V字

距離角度標記法(DAS)

波峰個數=2 長條形、橢圓形、海鷗形
 波峰個數=3 三角形、V字、海鷗形、矩形、圓形
 波峰個數=4 矩形、圓形、十字形
 波峰個數=5 星形、五角形

鈣化點質心線段與邊緣相交(相交法)

相交數=2 長條形
 相交數=4 V字
 相交數=6以上 海鷗形

關閉運算(CLOSE)

0~0.15 三角形； >0.15 V字、海鷗形
 0~0.13 三角形； >0.13 V字
 0~0.12 矩形； >0.12 十字形

經由上述 5 種方法之特性分析可規劃出分類演算法圖，如圖 17

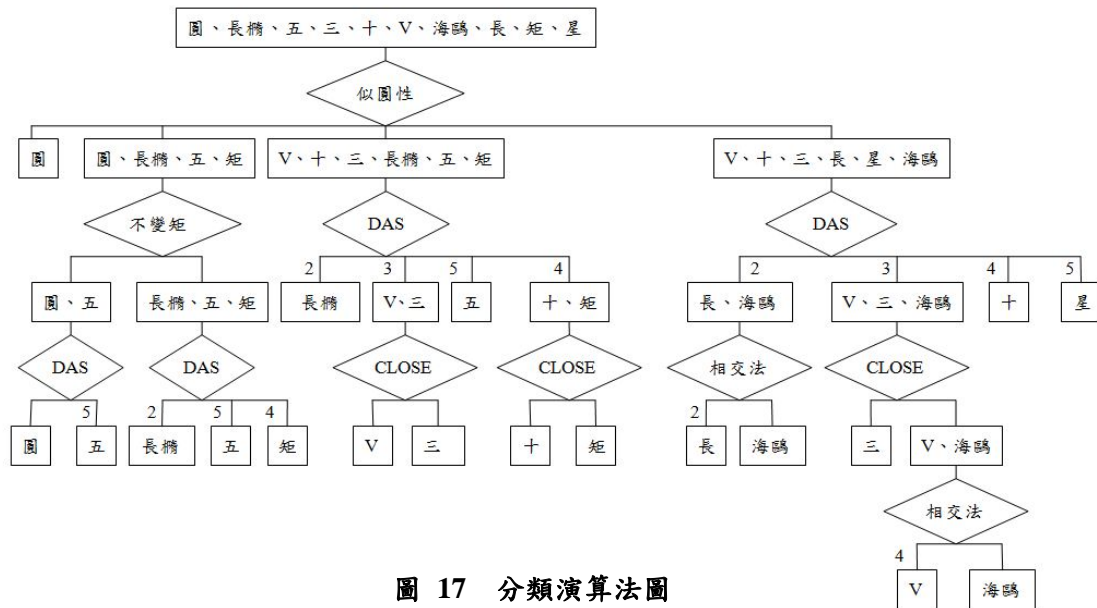


圖 17 分類演算法圖

4. 實驗結果及討論

本論文針對 3000 張鈣化點影像進行圖形辨識，每一種形狀的影像各佔 300 張，本系統利用多種方法來分辨出鈣化點的形狀，方法有：似圓性、不變矩、距離角度標記法、穿透以及 Close，來構成分類演算法，並使用 3000 張鈣化點影像進行實驗與數據分析。

針對分類演算法的部份，我們使用一張經過前處理後的 X 光影像進行實際操作，做為一個簡單範例。

4.1 分類過程

首先我們輸入了一張經過前處理後的 X 光影像進入系統如圖 18 所示，依照分類演算法設定進行分析，以下針對未知物件在分類演

算法中的判斷流程進行簡單的範例與圖解。根據設定，在分類演算法第一步遇到似圓性分析。



圖 18 未知物件之影像

在計算前，系統會先找出通過物件質心的最長軸，並將最長軸旋轉成與水準軸平行。似圓性分析可以求得圖形與圓的相似程度，經過似圓性運算後，如圖 19 所見，得到未知物件的似圓性值為 0.1552，依圖 20 的分類演算法判斷未知物件屬於似圓性值 0~0.26 區間的 V、十、三、長、星、海鷗類別。

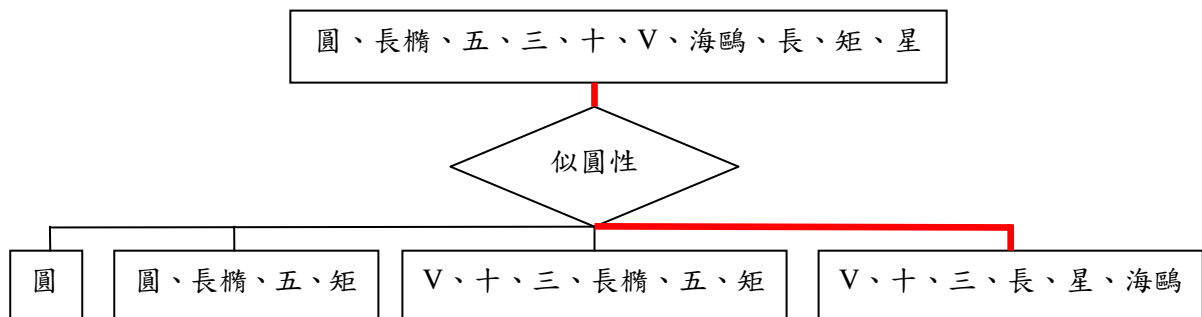


圖 20 似圓性分類演算法圖

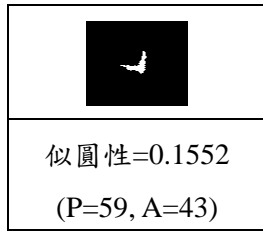


圖 19 未知物件經旋轉後的似圓性值

在判斷為 V、十、三、長、星、海鷗類別後，系統使用 DAS 進行分析，在進行 DAS 前，會先將物件正規化成固定大小(如圖 22)再進行運算，並產生圖 23 的波形圖，X 軸為角度，Y 軸為邊緣與質心的距離，之後計算未知物件的波峰個數，得到波峰個數值 3，並依此判斷未知物件為 V、三、海鷗類別，如圖 21 所示。

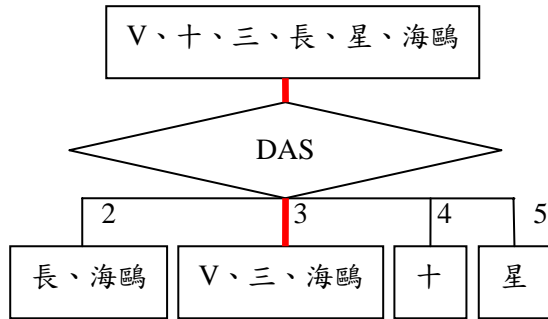


圖 21 距離角度法分類演算法圖



圖 22 經正規化的未知物件

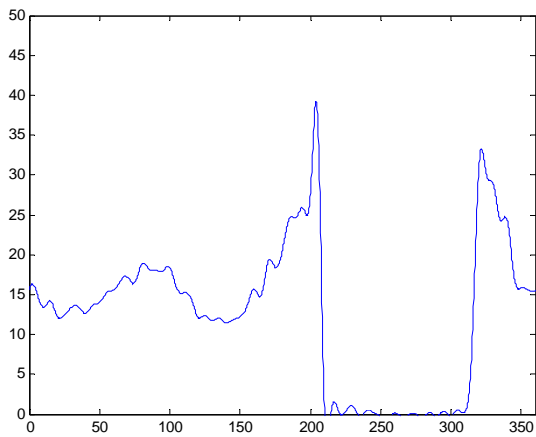


圖 23 未知物件的距離角度標記法之波形圖

V、三、海鷗類別中，因 V 字形與海鷗形都是具有明顯內凹的圖形，因此使用形態學中

的關閉運算將內凹的部份填滿，並計算填補過後的面積差異百分比，藉此與三角形進行區分。在此步驟中，如表 7 所示未知物件經計算後得到 26.95%，屬於明顯具有內凹的圖形，即圖 24 分類演算法圖中的 V、海鷗類別。

表 7 未知物件經關閉運算後之面積差異百分比

原圖	經關閉運算後
167	212
與原面積的差異百分比	
$(212-167)/167=26.95\%$	

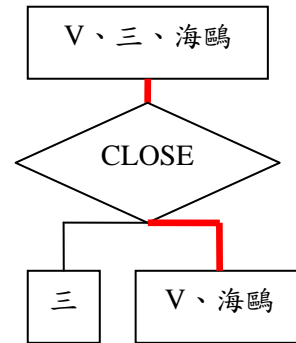


圖 24 關閉運算分類演算法圖

最後步驟，分辨未知物件為 V 字形或海鷗形，在此處使用鈣化點質心線段與邊緣相交進行判斷，V 字形在鈣化點質心線段與邊緣相交分析中值固定為 4，海鷗則為 8 或更多，經過計算後得到未知物件的鈣化點質心線段與邊緣相交值為 4，故判斷未知物件為 V 字形，如圖 25。

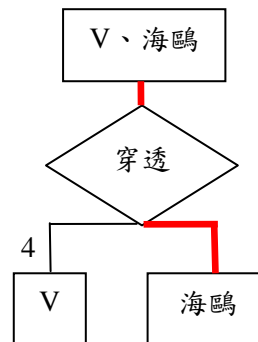


圖 25 穿透分類演算法圖

表 9 TP/FP/FN/TN 示意表

圖形		實際	
		陽	陰
系統 回傳	真	TP (true positive, 真陽性)	FP (false positive, 假陽性)
	假	FN (false negative, 假陰性)	TN (true negative, 真陰性)

TP:代表系統在全部的圖形中判斷是該形狀裡面，實際上也是該形狀。

FP:代表系統在全部的圖形中判斷是該形狀裡面，實際上不是該形狀。

FN:代表系統在全部的圖形中不被判斷是該形狀裡面，實際上是該形狀。

TN:代表系統在全部的圖形中不被判斷是該形狀裡面，實際上也不是該形狀。

敏感度(Sensitivity):又稱真陽性率,即實際為陽性被判定為陽性的機率。計算公式: $TP/(TP+FN) \times 100\%$,本論文代表實際是該形狀,且系統也判斷為該形狀的機率,此值越大,說明系統判斷越靈敏。

特異度(Specificity):又稱真陰性率,即實際為陰性被判定為陰性的機率,計算公式: $TN/(FP+TN) \times 100\%$,本論文代表實際不為該形狀,且系統也判斷不是該形狀的機率,此值越大,說明系統判斷越精確。

表 10 特異度敏感度

	Sensitivity(敏感)	Specificity(特異)
長條	90.33%	99.07%
橢圓	94.66%	98.52%
三角	93.00%	98.59%
矩形	86.66%	99.56%
五角	89.67%	99.59%
V字	95.33%	99.37%
十字	99.00%	99.33%
星形	98.00%	100%

海鷗	92.33%	99.07%
圓形	96.33%	99.70%
平均	93.53%	99.28%

經由數據可以求得表 10 中的各形狀特異度與敏感度,其中本系統對十字形、星形具有極佳的敏感度(即表 8 中的準確度),因為它們都擁有明顯的凸出與內凹特徵,使距離角度標記法可以準確的抓出波峰個數,而其他形狀也都能達到 90% 以上的敏感度,至於五角與矩形容易因為某些角度不明顯,使距離角度標記法不易測出其波峰,進而而有誤判的情形發生,此部份在下一章會再深入探討。特異度部份,指的是實際並非該形狀,且系統也判斷不是該形狀的機率,數據中所有形狀都有極高的特異度,也是本系統擁有辨識能力極佳的表現。

4.3 實驗結果之討論

實驗中,我們遇到了一些瓶頸,在判斷三角形、矩形、五角形、圓形時,會有誤判結果產生,其原因如下:

- 1、三角形:如圖 26 中三角形為鈍角三角形,偵測時因為鈍角不是明顯的角度,使波峰較難呈現,導致無法有效的偵測出來。此外鈍角三角形在形狀上較扁長,在肉眼上和程式上可能會被判斷為長條形或橢圓形。

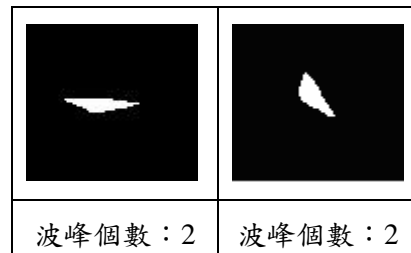


圖 26 分類錯誤圖形一

- 2、矩形:如圖 27,某些矩形在邊角的地方會偏向圓弧狀,使波峰個數出現 2 或 3 的情況,導致被誤判為橢圓形或三角形。



圖 27 分類錯誤圖形二

3、圓形及五角形：如圖 28，論文的分類演算法中，圓與五角被分到同一個集合時，可能因為波峰個數恰好為 5，而將圓形誤判為五角形；同時如圖 29 所示，也有可能因為五角形的角度不明顯，使波峰數不等於 5，而被誤判為圓型。

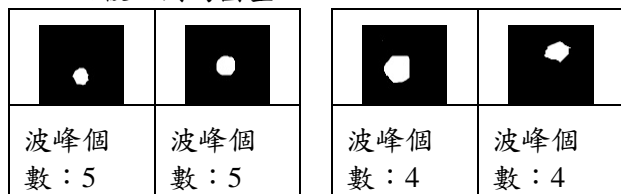


圖 28 分類錯誤圖形三

圖 29 分類錯誤圖形四

我們發現只要在邊緣上的角度不明顯，在程式上都相當容易被誤判為其他形狀，此情形不單會發生在上述鈣化點形狀，亦有可能發生在其他的形狀。

5. 結論及未來研究方向

由於乳房 X 光影像的普及，使得醫生在檢驗影像上負擔加重，為了輔助醫生做乳癌的偵測，所以我們決定研究一套適用於鈣化點的形狀分類方法，利用似圓性、不變矩、關閉運算、距離角度標記法以及鈣化點質心線段與邊緣相交數進行檢測，再依據我們所研究建立出的分類演算法進行分類，就可以準確分類各 X 光影像中的鈣化點形狀，並將此提供給醫生做輔助診斷的數據。在未來研究的方向，我們想嘗試針對 4.3 節實驗討論所述，辨識率較低之情形進一步深入研究，使其辨識率提高，並且找出其鈣化點之其他特徵，並加以整合建置為一乳房攝影影像資訊查詢系統，使能全自動化地輔助醫師做正確診斷與治療規畫。

感謝

本研究感謝 101 年度國科會計畫(含多個 ROI 乳房攝影影像之自動搜尋系統 NSC 101-2221-E-040 -011)經費及資源上之支持。

參考文獻

- [1] Riloff, E., and Hollaar, L. "Text Databases and Information Retrieval," ACM Computing Surveys, 1996, Vol. 28, No. 1, pp. 133-135.
 [2] Smeulders, A.W., Worring, M., Santini, S.,

Gupta, A., and Ramesh, J. "Content-based Image Retrieval at the End of the Early Years," IEEE Trans. Pattern Analysis and Machine Intelligence, 2000, Vol. 22, No. 12, pp. 1349-1380.

- [3] Yoshitaka, A., and Ichikawa, T. "A Survey on Content-based Retrieval for Multimedia Databases," IEEE Trans. Knowledge and Data Engineering, 1999, Vol. 11, No. 1, pp. 81-93.
 [4] Qiu, G. "Indexing Chromatic and Achromatic Patterns for Content-based Colour Image Retrieval," Pattern Recognition, 2002, Vol. 35, No. 8, pp. 1675-1686.
 [5] Berretti, S., Bimbo, A. Del., and Vicario, E. "Spatial Arrangement of Color in Retrieval by Visual Similarity," Pattern Recognition, 2002, Vol. 35, No. 8, pp. 1661-1674.
 [6] Chi, Y., and Leung, K.H. "ALSBIR: A Local-structure-based Image Retrieval," Pattern Recognition, 2007, Vol. 40, pp. 244-261.
 [7] Lin, W.C., Chang, Y.C., and Chen, H.H. "Integrating Textual and Visual Information for Cross-language Image Retrieval: A Trans-media Dictionary Approach," Inf. Process. Manage., 2007, Vol. 43, No. 2, pp. 488-502.
 [8] 李育強、楊智雁、呂學承、吳偉誠，「運用內部切割提昇 Zernike Moments 之特徵辨識效能」，南台科技大學資訊工程系，資訊技術與應用研討會摘要論文集，民國 99 年。
 [9] Hong-Jun Wei "Region-Based Descriptor For Trademark Retrieving," Kaohsiung I-Shou University in Partial Fulfillment of the Requirements for the Master degree in Information Engineering, July, 2006.
 [10] 段立娟，「形狀特徵的編碼描述研究綜述」，北京工業大學計算機學院，計算機科學，第 34 卷第 8 期，民國 96 年。
 [11] Jun-Yue Lin "Detection of Microcalcification on Mammograms Based on Gradient Angle Feature," Department of Information Management Chaoyang University of Technology, 2005.
 [12] 王濤、胡事民、孫家廣，「基於顏色-空間特徵的圖像檢索」，清華大學計算機科學與技術系，Journal of Software，第 13 卷第 10 期，民國 91 年。
 [13] 林子平，「乳房 X 光片影像之微鈣化及腫瘤特徵擷取及辨識」，國立雲林科技大學資

- 訊管理系碩士畢業論文，民國 96 年。
- [14] 柯穎昌，” Breast Mass Detection in Digital Mammograms using Textural Features”，國立中正大學電機工程研究所，民國 97 年。
- [15] Zahn T. and Roskies R.Z. “Fourier descriptors for plane closed curves,” IEEE Transactions on Computers, 1972, 21(3), pp.268-281.
- [16] 王聖凱，” Segmentation of the Gallbladder Wall Based on Fourier Descriptor”，東海大學資訊工程學系碩士論文，民國 97 年。
- [17] EI-ghazal A, Basir and Belkasim S. “Farthest point distance: A new shape signature for Fourier descriptors,” Signal Processing: Image Communication, 2009, 24(7), pp.572-586.
- [18] 張明旭，” A Novel System for Lepidoptera Recognition on Natural Image”，交通大學多媒體工程研究所學位元論文，民國 98 年。
- [19] 王斌，「一種基於多級弦長函數的傅立葉形狀描述子」，Chinese Journal of Computers，第 33 卷第 11 期，民國 99 年，頁 2387-2396。
- [20] 賴忠延，「運用傅立葉描述子完成唇形分類」，義守大學電子工程學系碩士班論文，民國 100 年。
- [21] 邊修弘，”New Intra and Inter Shape Coding Algorithms for MPEG-4”，國立中山大學資訊科學與工程學系碩士班論文，民國 91 年。
- [22] 張志鵬，”Segmented Vertex Chain Coding with Huffman Coding”，朝陽科技大學資訊與通訊工程學系碩士班論文，民國 97 年。
- [23] 何海濤,宋健，「基於鏈碼趨勢表的交通燈形狀識別方法」，Computer Engineering，第 37 卷第 15 期，民國 100 年，頁 158-160。
- [24] T.Balakumaran and ILA.Vennila “Detection of Microcalcification Clusters in Digital Mammograms using Multiresolution based Foveal Algorithm,” World Congress on Information and Communication Technologies, 2011.
- [25] Huang, P.W., and Dai, S.K. “Image Retrieval by Texture Similarity,” Pattern Recognition, 2003, Vol. 36, pp. 665-679.
- [26] 餘信賢，”A New Classification Method based on Calcification Shapes in Mammographic Images”，國立中興大學資訊科學研究所碩士學位論文，民國 94 年。