

# 含多個 ROI 之乳房攝影影像存取系統

楊富華

中山醫學大學應用資訊科學學系  
jordanyang529@yahoo.com.tw

徐麗蘋

中山醫學大學應用資訊科學學系  
apple@csmu.edu.tw

## 摘要

本論文目的即在於提出一可自動存取與查詢影像相似之含多個 ROI 乳房 X 光片之系統，以幫助醫師搜尋具有相似鈣化點病徵的乳房 X 光片影像來輔助醫師作出更佳的診斷結果。因乳房鈣化點之分佈情況及形狀為乳房鈣化診斷之最重要依據，故本論文提出以擷取 ROI 鈣化點分佈之空間關係及其內所含鈣化點之形狀兩方面之特徵來建立可表示含多個 ROI 之乳房 X 光影像的表示法，接著為了符合醫生不同之查詢需求，亦提出依據此影像表示法的一套具有彈性之相似度衡量標準及方法。本論文所提出的乳房 X 光影像資訊系統能自動搜尋並提供醫師與新病患具有類似病徵之乳房 X 光影像，以輔助醫生利用該乳房 X 光影像來對新病患作出更佳的病情診斷及治療規劃。實驗結果顯示所提出之方法與人類視覺感知一致且具有良好之精確度並具滿足不同查詢需求之彈性。

關鍵詞：乳房 X 光攝影、微鈣化、Extended 9D-SPA、ROI、形狀特徵

## 壹、緒論(Introduction)

由於乳房 X 光攝影在偵測乳癌初期的症狀也就是偵測微鈣化點方面有非常好的辨識率，尤其是在無法以觸診或一些乳癌症狀較不明顯的狀況下更能夠顯現出乳房 X 光攝影用來偵測乳癌的優點，但也因為如此醫生便開始大量的使用乳房 X 光攝影來作乳癌的偵測，但是當檢測的人數越來越多時，相對的由乳房 X 光片也快速的增加，無形之中也增加了醫師對於檢驗影像上的負擔，如此一來有可能會造成醫師作出錯誤的決定，當醫生在診斷病情時可能需要類似相關的乳房 X 光片作為參考之用，但因為數量龐大造成找尋類似乳房 X 光的時間拉長，非常費時，並鑑於醫師在乳房 X 光中所圈出的 ROI 鈣化點是乳癌初期的重要特徵之一，且一般病例皆含有多個鈣化群(ROI)，但含多個 ROI 之影像之辨識及存取，在方法之設計及複雜度皆截然與含單一 ROI 之影像不同，所以本論文所研究的重點就是在設計一自動搜尋含多個 ROI 之類似的乳房 X 光片之方法，使能實際應用於輔助醫生作出更正確的診斷。

## 貳、相關文獻探討

茲將與本論文相關之研究情形與分析分為三個方面來探討，首先是以內容為基礎之影像檢索系統Content-Based Image Retrieval System(CBIR)，此為影像處理中的一個重要主題，另外，本論文主要將Extended 9D-SPA 表示法加以修改使能應用在含多個ROI之乳房X 光片的擷取上，再將鈣化點之形狀特徵整合在內而建立，因此在此先說明Extended 9D-SPA 表示法；最後則是微鈣化點類別以及偵測的相關介紹；微鈣化點的出現是判斷乳癌初期的一個重要的特徵。

### 一、以內容為基礎的影像擷取

一般傳統的影像資料庫建置方式往往是利用文字來描述影像(Text-Based)，其既無效率且又相當主觀，往往會因為描述者不同，而對相同影像做出不同的關鍵性註解，當然，也會影響影像查詢結果的正確性。現今之趨勢是以影像內容為基礎來進行影像分析、特徵擷取之後再和影像資料庫中的影像作比對，最後則輸出相似的影像，此即為CBIR(Content-Based Image Retrieval)[5, 6]，其便是減少上述視覺感官與文字描述之間的差距，針對影像資料的特徵，如顏色(color)[7, 8]、形狀(shape)[9, 10]、紋理(texture)[11]等關鍵資訊，進行資料庫的建置、檢索或是查詢，而使得影像的處理更具客觀性。

### 二、Extended 9D-SPA 表示法

一個優良的影像表示法除了編碼的內容必須簡單，還不能受到影像平移或是比例變化的影響，用唯一的方式表示所欲擷取的影像，此外，尚兼具有相似度查詢以及空間推論等功能，而Extended 9D-SPA[19]表示法就是一個綜合以上優點的空間關係表示法，且此表示法之特性為以其方向區域之定義方法能簡單地一併將鈣化點ROI的大小以及與其它物件間的距離等因素考慮進去，並且能以簡單的編碼記錄了乳房X光影像中各個鈣化點ROI的重心方位、起點資訊、終點資訊、方位關係以及拓撲關係，因此本論文之影

像表示法為將Extended 9D-SPA 表示法加以修改以應用在乳房X 光片的空間關係特徵擷取上，再將鈣化點之形狀特徵整合在內而建立，以下說明Extended 9D-SPA 表示法的編碼、處理方式以及如何適當地應用在含多個ROI之乳房X 光影像。首先，要以圖片中的物件分佈空間位置來辨識圖片，第一個步驟是去定義兩個物件間的方位關係。每個方位關係包含一個參考物件和一個目標物件。關於每個方位關係，對於參考物件的空間被劃分成八個非重複的區域。這八個區域分別為東、東北、北、西北、西、西南、南、東南；各個區域被分配一個從1到8的整數。有兩種方式可以用來分割空間：

第一種分割方式：

在參考物件與目標物件的MBR沒有重疊的情形下。兩條分割線的角度為45度，分別通過了參考物件的MBR的東北角與東南角，在分割線所包圍的東區設為整數1。

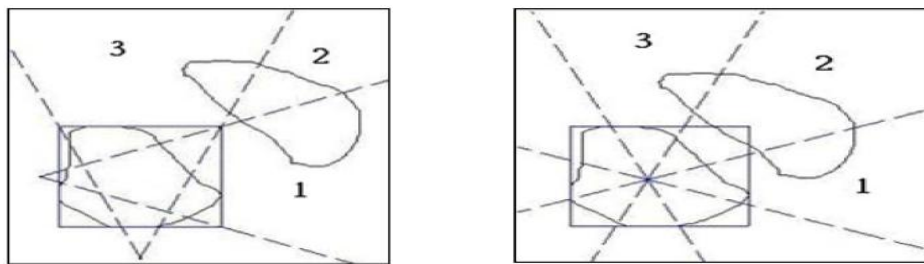


圖 1 (a)第一種空間分割方式

(b) 第二種空間分割方式

北區的分割線也是用相同的方法，設為整數3。而東北區域在東區與北區之間，以整數2來表示。

第二種分割方式：

在參考物件與目標物件的MBR有部分重疊的情形下。空間由參考物件的重心向外分散的分割線切割成8個均等個區域。

不論用哪種方式來分割空間，方位區域可沿著各個區域的中線進一步被畫分成兩個區域，按照逆時鐘的順序，稱為第一部分與第二部分。假設所有的圖片均已完成了影像分割的程序。在說明Extended 9D-SPA 前，茲先將相關符號之定義說明如下：一個圖片  $P$  包含  $n$  個物件  $(O_1, O_2, \dots, O_n)$ ，將圖片  $P$  編碼為四個元素的集合形式： $R = \{(O_{ij}, T_{ij}, D_{ij}, D_{ji}) \mid O_i, O_j \in P, \text{ and } 1 \leq i < j \leq n\}$ ，其中

- $O_{ij}$ ：表示一對物件對  $(O_i, O_j)$ 。
- $T_{ij}$ ：表示  $O_i$  與  $O_j$  兩物件之間的拓撲關係。
- $D_{ij}$ ：表示  $O_i$  與  $O_j$  兩物件之間的方位關係，以物件  $O_j$  為參考物件。
- $D_{ji}$ ：表示  $O_i$  與  $O_j$  兩物件之間的方位關係，以物件  $O_i$  為參考物件。

以下針對集合  $R$  的各個特徵的編碼方式及原則說明如下：

第一個元素  $O_{ij}$ ：稱為物件對  $(O_i, O_j)$  的編碼，其中

$$O_{ij} = \frac{(j-1)(j-2)}{2} + i$$

第二個元素  $T_{ij}$ ：分別以數字0~7表示以下八種不同的拓撲關係碼。0-代表「分離」、1-代表「相鄰」、2-代表「部分重疊」、3-代表「覆蓋」、4-代表「被覆蓋」、5-代表「包含」、6-代表「被包含」以及7-代表「相等」。

第三及第四個元素  $D_{ij}$  與  $D_{ji}$  :  $D_{ij}$  為物件  $O_i$  與物件  $O_j$  之間的方位關係以  $O_j$  為參考物件,  $D_{ij}$  包含 5 個元素 ( $C_{ij}, f_{ij}, b_{ij}, s_{ij}, e_{ij}$ ), 其中

- $C_{ij} \in \{k | 1 \leq k \leq 8\}$ , 顯示目標物件  $O_i$  的重心位於參考物件  $O_j$  的哪個方位區域。
- $f_{ij}$  (或  $b_{ij}$ ) 顯示額外的區域個數由逆時鐘(順時鐘)的順序橫跨幾個區域, 從以  $O_j$  為參考物件的矩心區域開始,  $0 \leq f_{ij}, b_{ij} \leq 8$ 。
- $s_{ij}$  (或  $e_{ij}$ ) 顯示關於目標物件  $O_i$  的起點(終點)的資訊(見圖 2)
  - 在  $f_{ij}=b_{ij}=0$  的情況下:
    - 當整個物件  $O_i$  位於物件  $O_j$  方位區域的第一部分時,  $s_{ij}=e_{ij}=0$ 。
    - 當整個物件  $O_i$  位於物件  $O_j$  方位區域的第二部分時,  $s_{ij}=e_{ij}=1$ 。
    - 當物件  $O_i$  橫跨物件  $O_j$  方位區域的中線時,  $s_{ij}=0, e_{ij}=1$ 。
  - 在  $f_{ij} \neq 0, b_{ij} \neq 0$  的情況下:
    - 當物件  $O_i$  橫跨  $O_i$  的起點方位區域之中線,  $s_{ij}=1$ ; 若物件  $O_i$  沒有橫跨  $O_i$  的起點方位區域之中線,  $s_{ij}=0$ 。
    - 當物件  $O_i$  橫跨  $O_i$  的終點方位區域之中線,  $e_{ij}=1$ ; 若物件  $O_i$  沒有橫跨  $O_i$  的終點方位區域之中線,  $e_{ij}=0$ 。

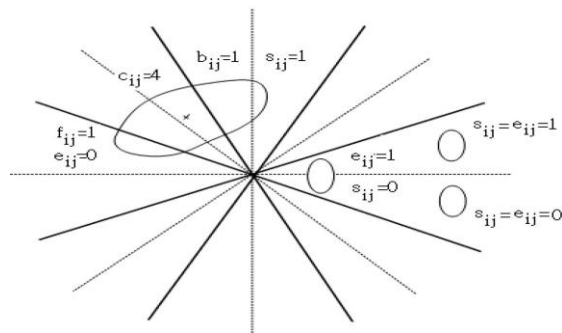


圖 2  $f_{ij}$ 、 $b_{ij}$ 、 $s_{ij}$  以及  $e_{ij}$  的圖例

### 三、微鈣化

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。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, 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[20]。



圖3 微鈣化之形狀種類

### 參、影像存取方法

本論文主要將 Extended 9D-SPA 表示法加以修改使能應用在含多個 ROI 之乳房 X 光片的擷取上，再將鈣化點之形狀特徵整合在內而建立，因此在此說明本論文方法主要包含的三個部份：Extended 9D-SPA 表示法在系統中的應用、Extended 9D-SPA 表示法之修改、相似度計算修正等。

### 一、 Extended 9D-SPA 表示法在系統中的應用

醫師想擷取乳房 X 光片的相關影像時，依照病理與發生病灶的相對位置來搜尋擷取所需的乳房 X 光片之相關影像，所以必須建構 QBE(Query By Example)或是使用簡圖來當作擷取的範本來擷取影像，所以不適合以紋理來做為特徵，乳房攝影 X 光片為灰階度影像且醫師所畫之 ROI 皆以圓形為主，故也不適合使用形狀及顏色來當作特徵；另外，由於每個人的人體構造均不相同，所以每張乳房 X 光片的影像中，鈣化群的位置也不會完全相同，因此不適合以上述的第二種類型作來作為特徵，以物件間的空間關係之特徵來建立資料庫，才能不受到人體構造因素的影響，順利擷取出所需的乳房 X 光片，故本論文將以空間關係來作為資料庫索引所需之特徵。

### 二、 Extended 9D-SPA 表示法之修改

在 Extended 9D-SPA 表示法中，定義了影像資料庫中包含  $n$  個物件( $O_1, O_2, \dots, O_n$ )，而此  $n$  個物件在同一張圖片中不能重複出現，表是每張圖片中所包含的每個物件都必須存在於資料庫，否則就須將其新增到資料庫，但資料庫中的物件會因此而不斷的增加，造成使用上很大的限制，也就是同一類型的物件將無法利用 Extended 9D-SPA 表示法來加以搜尋，大大地減低了 Extended 9D-SPA 表示法的實用性，所以若不修正 Extended 9D-SPA 表示法的編碼方式，就無法將 Extended 9D-SPA 表示法應用在乳房 X 光片的影像擷取資料庫上。本論文的乳房 X 光片影像資料庫設計將同一 ROI 類型的物件當作是同一物件，以克服 Extended 9D-SPA 表示法中同一物件不能同時存在於同一張影像中的限制。

在本系統中，物件  $O_i$  的編碼  $i$  不再用以表示物件的編碼，而是用以表示 ROI 的類型編碼，其中

$$i = 1 + \sum_{k=1}^M b_k T_k \quad (1)$$

$M$  表示所有鈣化點形狀的類型總數， $T_k$  則表示鈣化點形狀的二元編碼，表 1 為鈣化點形狀的二元編碼表，當  $b_k = 1$  時，表示此 ROI 存在該形狀的鈣化點，反之，若該形狀的鈣化點不存在，則  $b_k = 0$ ，由上述公式可看出， $O_i$  的編碼不會重複，例如：當一個 ROI 中存有長橢圓、三角形、小圓點三種形狀的鈣化點時，此 ROI 的類型編碼即為  $(1+2^0+2^2+2^4)=22$ 。 $O_1$  為本系統最主要的參考物件—乳頭，乳房 X 光片攝影所得的 X 光片總會受到不同的攝影儀器和病患自身的乳房大小、胸型等因素的影響，因此，不同的 X 光片裡，所呈現的鈣化群或腫塊的相對位置將會有所差異，以乳頭做為參考物較不易受上述因素的影響。本論文即採用 L. Shen [23]所提及的 12 種鈣化點類型作為標準，這 12 種鈣化點的類型可依需要進行調整，再重新編碼，並不會因此而影響到系統的運作方式。



表 1 鈣化點形狀類型的二元編碼表

|          |          |       |       |       |       |
|----------|----------|-------|-------|-------|-------|
| 五角形      | 長橢圓      | 矩形    | 三角形   | 圓形    | 小圓點   |
| $2^5$    | $2^4$    | $2^3$ | $2^2$ | $2^1$ | $2^0$ |
| 其他 2     | 其他 1     | 星行    | 海鷗形   | V 字形  | 長條形   |
| $2^{11}$ | $2^{10}$ | $2^9$ | $2^8$ | $2^7$ | $2^6$ |

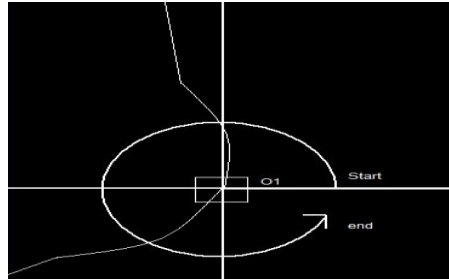


圖 4 同類型鈣化群排序策略

Extended 9D-SPA 表示無法呈現出 ROI 中同時出現兩個以上的物件擁有相同種類編碼的情形，因此，本研究設計先把影像中擁有相同種類編碼的物件  $O_i$  全部找出，依據物件與乳頭間的方位角和距離關係，如圖 4 所示，以乳頭為中心，正東方為  $0^\circ$ ，依逆時鐘的方向計算出個物件的方位角，此時，若正巧有物件出現在同一角度，則以距離乳頭近的誤驗為優先，對這些物件進行排序，再依序編出其所對應的物件對編碼  $O_{ij}$ ，其中本論文將 Extended 9D-SPA 表示法中物件對的  $O_{ij}$  編碼修改如下：

$$\begin{cases} O_{ij} = \frac{(j-2) \times (j-1)}{2} + 1 + \alpha \times C_2^{O_{max}} & \text{if } i < j \\ O_{ji} = -(i + \alpha \times O_{max}) & \text{if } i = j \end{cases} \quad (2)$$

$\alpha$  代表排序後的  $O_i$ ，在同一類型的物件間之順位關係，由 0 開始。 $O_{max}$  為系統中所有物件編碼  $O_i$  的最大值，也就是資料庫中 ROI 類型的種數。若影像中有出現相同的物件編碼時，使用 Extended 9D-SPA 表示法的物件對編碼方式會出現重複的編碼，但編碼不能重複，故本系統將其另加上其物件對的最大值  $O_{max}$ ，以消除物件對編碼重複的現象，在未來的相似度計算中，只要將其除以物件對的最大值  $O_{max}$ ，即可還原得到原始的物件編碼  $O_i$  及順位  $\alpha$  合物件編碼  $O_i$  或  $O_j$ ，在相似度計算上提供了很大的幫助。如圖 5 所示，同時出現三個物件類型編碼為 2 的物件，排序後，較上方的新物件對編碼將恰同於未經排序而編得的物件對編碼，較下方的新物件對編碼則會依其對應的編碼方式，以原物件對編碼加上物件對編碼的最大值  $O_{max}$  作為其物件對編碼，以此類推，其結果如表 2 所示。

表 2 修改後的 Extended 9D-SPA 表示法之編碼結果

| $\alpha$ | $O_i$ | $O_j$ | $O_{ij}$                   | $D_{ij}$    | $D_{ji}$    | $T_{ij}$ |
|----------|-------|-------|----------------------------|-------------|-------------|----------|
| 0        | 1     | 2     | $1+0 \times C_2^{O_{max}}$ | (8,0,0,0,1) | (4,0,0,0,1) | 0        |
| 1        | 1     | 2     | $1+1 \times C_2^{O_{max}}$ | (1,0,0,0,1) | (5,0,0,0,1) | 0        |
| 2        | 1     | 2     | $1+2 \times C_2^{O_{max}}$ | (1,1,0,0,0) | (5,0,0,1,1) | 0        |
| 0        | 2     | 2     | $-(2+0 \times O_{max})$    | (2,0,0,0,1) | (6,0,0,0,1) | 0        |
| 1        | 2     | 2     | $-(2+1 \times O_{max})$    | (3,0,0,0,0) | (7,0,0,0,0) | 0        |
| 2        | 2     | 2     | $-(2+2 \times O_{max})$    | (4,1,0,0,0) | (8,0,0,0,1) | 0        |

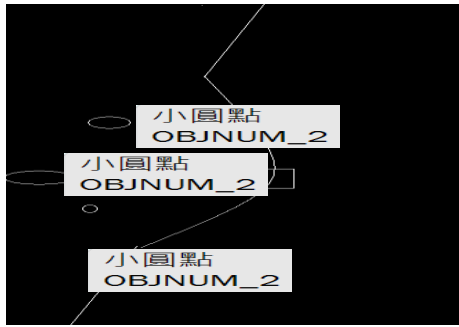


圖 5 修改後的 Extended 9D-SPA 表示法編碼

### 三、 相似度計算修正

由於本論文所有的物件編碼和物件對編碼都已改變，故在相似度的計算上也有些改變，以使 Extended 9D-SPA 表示法中相似度的彈性能夠發揮其應有的作用。

在此說明本論文所定義的符號，如下：

$R_p$ (或  $R_q$ )：圖片 p 或是圖片 q 的修改後之 Extended 9D-SPA 表示法。

$t^p$ (或  $t^q$ )： $R_p$ (或  $R_q$ )內的子項目。

$t.O$ ：項目 t 內的物件對編碼。

$t.O_\alpha$ (或  $t.O_\beta$ )：還原後的物件對編碼。

$t.D_1$ ：在項目 t 內，以  $O_j$  為參考物件的位置關係編碼。

$t.D_2$ ：在項目 t 內，以  $O_i$  為參考物件的位置關係編碼。

$t.T$ ：項目 t 內的拓撲關係。

$t.D_1(j)$ (或  $t.D_2(j)$ )： $t.D_1$  (或  $t.D_2$ )編碼中的第 j 個位元。

$x(k)$ ：x 二元碼中的第 k 個位元。

$O_{\max}$ ：物件編碼的總數。

$C_2^{O_{\max}}$ ：物件對編碼的最大值。

$R_q = \{t_1^q, t_2^q, t_3^q, \dots, t_m^q\}$ 、 $R_p = \{t_1^p, t_2^p, t_3^p, \dots, t_m^p\}$  分別為資料庫中的圖片 p 與欲查詢的圖片 q 之修改後的 Extended 9D-SPA 表示法，若物件對編碼  $t.O$  大於零，則此物件對編碼除以物件對編碼的最大值之商即為  $\alpha$  (圖片 q 排序後的  $O_i$  在同一類型的物件間之順位關係)與  $\beta$  (圖片 q 排序後的  $O_i$  在同一類型物件間之順位關係)，餘數則為物件對  $O_{ij}$  原來的編碼，若物件對編碼  $t.O$  小於零，此物件對編碼的絕對值除以物件種類的數目之商即為  $\alpha$  與  $\beta$ ，餘數則為相同物件對的編碼，以此公式來還原物件編碼相同前，每組的順序與原始編碼。

$$\begin{cases} \alpha = \frac{t.O}{C_2^{O_{\max}}} & , \quad t.O_\alpha = t.O \bmod C_2^{O_{\max}} \quad \text{if } t.O > 0 \\ \alpha = \frac{t.O}{O_{\max}} & , \quad t.O_\alpha = t.O \bmod O_{\max} \quad \text{if } t.O < 0 \end{cases} \quad (3)$$

“m”代表 Extended 9D-SPA 表示法中圖片  $q$  的物件對個數，“n”代表 Extended 9D-SPA 表示法中圖片  $p$  的物件對個數，“match”代表  $R_q$  與  $R_p$  之間的最佳對應關係，使  $S_D$  的計算結果最大，並假設  $\text{match}(t_i^q \cdot O_\alpha) = t_i^p \cdot O_\beta$ ， $1 \leq i \leq k$ ，而  $k = \min\{m, n\}$ 。

$R_q$  與  $R_p$  的位置相似度將改以下列公式計算：

$$S_D(R_p, R_q) = \frac{\sum_{i=1}^k S_w(\text{match}(t_i^q \cdot O_\alpha), t_i^p \cdot O_\beta)}{m + n} \quad (4)$$

$O_{ij}$  大於零時：

$$\begin{aligned} S_w(\text{match}(t_i^q \cdot O_\alpha), t_i^p \cdot O_\beta) \\ = S_D(\text{match}(t_i^q \cdot O_\alpha), t_i^p \cdot O_\beta, D_1) + S_D(\text{match}(t_i^q \cdot O_\alpha), t_i^p \cdot O_\beta, D_2) \end{aligned}$$

另外，當  $O_{ij}$  小於零時，計算如下：

$$\begin{aligned} S_w(\text{match}(t_i^q \cdot O_\alpha), t_i^p \cdot O_\beta) \\ = \text{MAX} \left\{ \begin{array}{l} S_D(\text{match}(t_i^q \cdot O_\alpha), t_i^p \cdot O_\beta, D_1) + S_D(\text{match}(t_i^q \cdot O_\alpha), t_i^p \cdot O_\beta, D_2) \\ S_D(\text{match}(t_i^q \cdot O_\alpha), t_i^p \cdot O_\beta, D_2) + S_D(\text{match}(t_i^q \cdot O_\alpha), t_i^p \cdot O_\beta, D_1) \end{array} \right\} \end{aligned}$$

其中

$$S_D(x, y) = \begin{cases} 1 & \text{if } \sum_{k=1}^{16} x(k) \cup y(k) = 0 \\ \frac{\sum_{k=1}^{16} x(k) \cap y(k)}{\sum_{k=1}^{16} x(k) \cup y(k)} & \text{otherwise.} \end{cases}$$

符號“ $\cap$ ”代表邏輯運算原“AND”，符號“ $\cup$ ”代表邏輯運算原“OR”。相似值越大者表示圖片  $p$  與圖片  $q$  越相近，如圖 6 所示，圖片  $q$  有三個類型相同的 ROI，其編碼為 8，物件對編碼則由上節提及的新物件對編碼方式所編成，故需重建  $\alpha$  值與物件對編碼，圖片  $p$  也是依此步驟進行，再根據符合的相似度來篩選，依照相似度的計算公式計算，其中

$$\begin{aligned} \text{match}(t^q \cdot 22_0) = t^p \cdot 22_0 \cdot \text{match}(t^q \cdot 22_2) = t^p \cdot 22_1 \cdot \text{match}(t^q \cdot -8_1) \\ = t^q \cdot -8_0 \end{aligned}$$

$$S_D(p, q) = \frac{(1+1) + (1+1) + (1+1)}{6+3} = \frac{6}{9} \cong 0.67$$

$$S_T(p, q) = \frac{2 \times (1+1+1)}{6+3} = \frac{6}{9} \cong 0.67$$

以上結果即為圖片  $q$  與圖片  $p$  的相似度，由於本系統考慮 ROI 的拓樸關係通常為分離，有較少的相鄰、部分重疊、覆蓋、包含與被包含，因此在 match 時並不使用拓樸關係作為主要依據，而是使用物件的空間關係來作為主要依據，當兩張圖完全相同時，相似度為 1。Match 兩張圖的關係集合時，系統將考慮以計算  $S_D$  恰為最大時，相似度將與原本 Extended 9D-SPA 表示法的相似度相同，因此也擁有 Extended 9D-SPA 相似度的彈性。



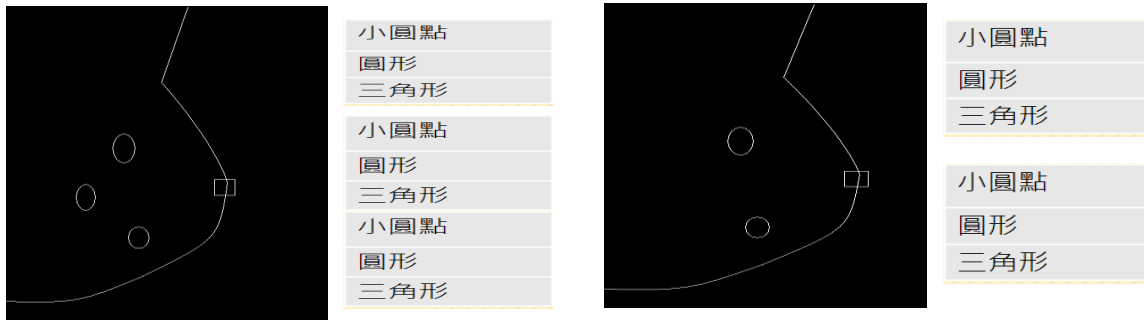


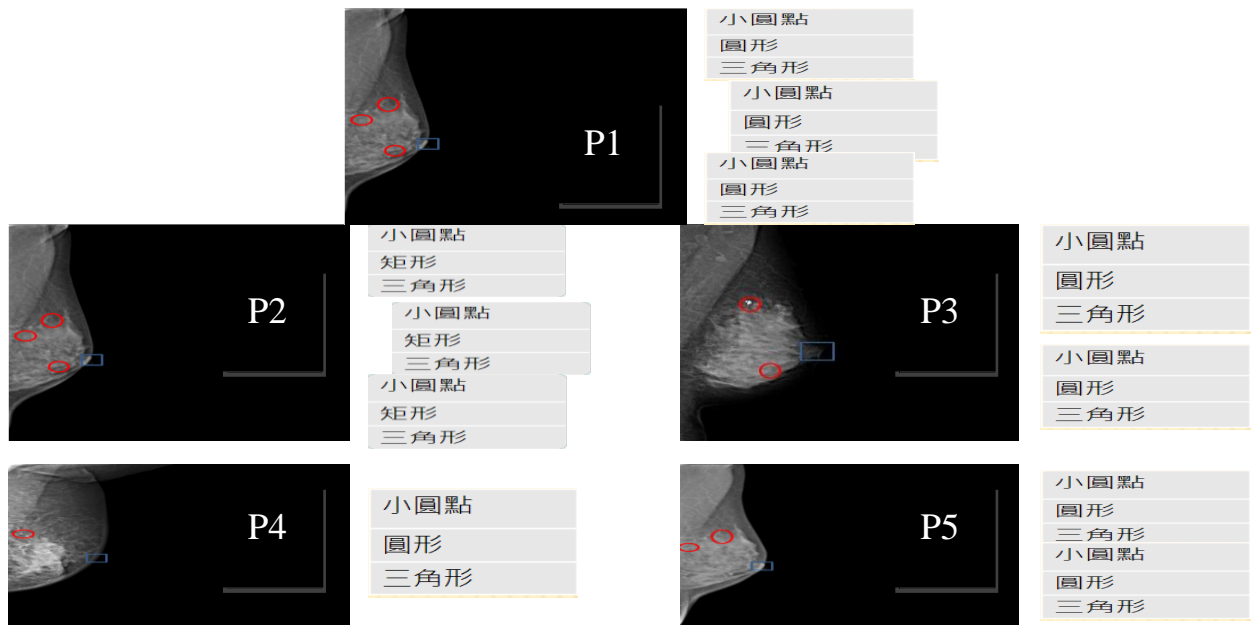
圖 6 相似度計算圖例

表 3 編碼結果

| $\alpha$ | $O_{ij}$ | $D_{ij}$    | $D_{ji}$    | $T_{ij}$ |
|----------|----------|-------------|-------------|----------|
| 0        | 22       | (8,0,0,1,1) | (4,0,0,1,1) | 0        |
| 1        | 22       | (1,0,0,0,1) | (5,0,0,0,1) | 0        |
| 2        | 22       | (2,0,0,0,1) | (6,0,0,0,1) | 0        |
| $\alpha$ | $O_{ij}$ | $D_{ij}$    | $D_{ji}$    | $T_{ij}$ |
| 0        | -8       | (2,0,1,0,0) | (6,0,0,0,1) | 0        |
| 1        | -8       | (3,0,0,0,1) | (7,0,0,0,1) | 0        |
| 2        | -8       | (4,0,0,0,1) | (8,0,0,0,1) | 0        |

| $\beta$ | $O_{ij}$ | $D_{ij}$    | $D_{ji}$    | $T_{ij}$ |
|---------|----------|-------------|-------------|----------|
| 0       | 22       | (8,0,0,1,1) | (4,0,0,1,1) | 0        |
| 1       | 22       | (2,0,0,0,1) | (6,0,0,0,1) | 0        |
| $\beta$ | $O_{ij}$ | $D_{ij}$    | $D_{ji}$    | $T_{ij}$ |
| 0       | -8       | (3,0,0,0,1) | (7,0,0,0,1) | 0        |

#### 肆、實驗結果及討論



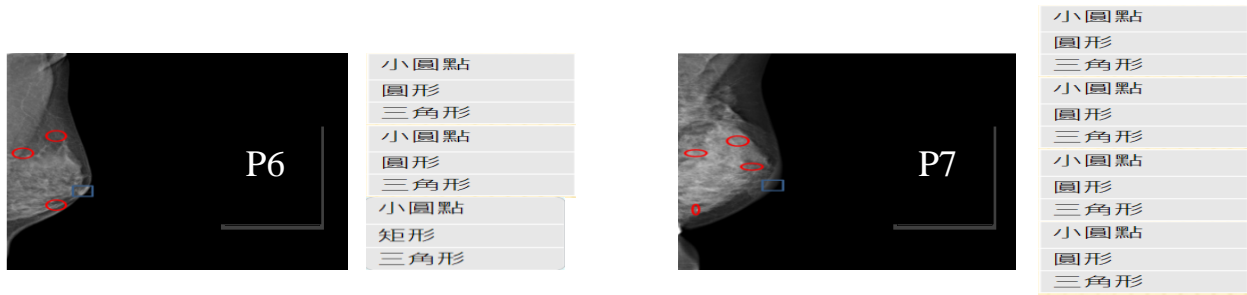


圖 7 ROI 已編碼之乳房 X 光片

其中， $P_1$  是我們要查詢的圖片， $P_2 \sim P_7$  是我們查詢的範圍， $P_1$  之 ROI 編碼都為 8(小圓點+圓形+三角形)， $P_2$  之 ROI 編碼都為 14(小圓點+矩形+三角形)， $P_3$  之 ROI 編碼都為 8(小圓點+圓形+三角形)， $P_4$  之 ROI 編碼為 8(小圓點+圓形+三角形)， $P_5$  之 ROI 編碼都為 8(小圓點+圓形+三角形)， $P_6$  之 ROI 編碼分別為 8(小圓點+圓形+三角形)、8(小圓點+圓形+三角形)、14(小圓點+矩形+三角形)， $P_7$  之 ROI 編碼都為 8(小圓點+圓形+三角形)，這七張圖片的 Extended 9D-SPA 表示法之編碼結果各為表 4 至表 10：

表 4  $P_1$  的 Extended 9D-SPA 表示法之編碼結果

| $\alpha$ | $O_{ij}$ | $D_{ij}$    | $D_{ji}$    | $T_{ij}$ |
|----------|----------|-------------|-------------|----------|
| 0        | 22       | (8,0,0,0,1) | (4,0,0,0,1) | 0        |
| 1        | 22       | (1,0,0,0,0) | (5,0,0,0,0) | 0        |
| 2        | 22       | (2,0,0,0,1) | (6,0,0,0,1) | 0        |
| $\alpha$ | $O_{ij}$ | $D_{ij}$    | $D_{ji}$    | $T_{ij}$ |
| 0        | -8       | (2,0,1,0,0) | (6,0,0,0,1) | 0        |
| 1        | -8       | (3,0,0,0,1) | (7,0,0,0,1) | 0        |
| 2        | -8       | (4,0,0,0,1) | (8,0,0,0,1) | 0        |

表 5  $P_2$  的 Extended 9D-SPA 表示法之編碼結果

| $\beta$  | $O_{ij}$ | $D_{ij}$    | $D_{ji}$    | $T_{ij}$ |
|----------|----------|-------------|-------------|----------|
| 0        | 79       | (8,0,0,0,1) | (4,0,0,0,1) | 0        |
| 1        | 79       | (1,0,0,0,0) | (5,0,0,0,0) | 0        |
| 2        | 79       | (2,0,0,0,1) | (6,0,0,0,1) | 0        |
| $\alpha$ | $O_{ij}$ | $D_{ij}$    | $D_{ji}$    | $T_{ij}$ |
| 0        | -14      | (2,0,1,0,0) | (6,0,0,0,1) | 0        |
| 1        | -14      | (3,0,0,0,1) | (7,0,0,0,1) | 0        |
| 2        | -14      | (4,0,0,0,1) | (8,0,0,0,1) | 0        |

表 6  $P_3$  的 Extended 9D-SPA 表示法之編碼結果

| $\beta$ | $O_{ij}$ | $D_{ij}$    | $D_{ji}$    | $T_{ij}$ |
|---------|----------|-------------|-------------|----------|
| 0       | 22       | (8,0,0,0,1) | (4,0,0,0,1) | 0        |
| 1       | 22       | (2,0,0,0,1) | (6,0,0,0,1) | 0        |
| $\beta$ | $O_{ij}$ | $D_{ij}$    | $D_{ji}$    | $T_{ij}$ |
| 0       | -8       | (3,0,0,0,1) | (7,0,0,0,1) | 0        |

表 7  $P_4$  的 Extended 9D-SPA 表示法之編碼結果

| $\beta$ | $O_{ij}$ | $D_{ij}$    | $D_{ji}$    | $T_{ij}$ |
|---------|----------|-------------|-------------|----------|
| 0       | 22       | (1,0,0,0,0) | (5,0,0,0,0) | 0        |

表 8  $P_5$  的 Extended 9D-SPA 表示法之編碼結果

| $\beta$ | $O_{ij}$ | $D_{ij}$    | $D_{ji}$    | $T_{ij}$ |
|---------|----------|-------------|-------------|----------|
| 0       | 22       | (8,0,0,0,1) | (4,0,0,0,1) | 0        |
| 1       | 22       | (1,0,0,0,0) | (5,0,0,0,0) | 0        |
| $\beta$ | $O_{ij}$ | $D_{ij}$    | $D_{ji}$    | $T_{ij}$ |
| 0       | -8       | (2,0,1,0,0) | (6,0,0,0,1) | 0        |

表 9  $P_6$  的 Extended 9D-SPA 表示法之編碼結果

| $\beta$ | $O_{ij}$ | $D_{ij}$    | $D_{ji}$    | $T_{ij}$ |
|---------|----------|-------------|-------------|----------|
| 0       | 22       | (8,0,0,0,1) | (4,0,0,0,1) | 0        |
| 1       | 22       | (1,0,0,0,0) | (5,0,0,0,0) | 0        |
| 2       | 79       | (2,0,0,0,1) | (6,0,0,0,1) | 0        |
| $\beta$ | $O_{ij}$ | $D_{ij}$    | $D_{ji}$    | $T_{ij}$ |
| 0       | -8       | (2,0,1,0,0) | (6,0,0,0,1) | 0        |
| 1       | -14      | (3,0,0,0,1) | (7,0,0,0,1) | 0        |
| 2       | -14      | (4,0,0,0,1) | (8,0,0,0,1) | 0        |

表 10  $P_7$  的 Extended 9D-SPA 表示法之編碼結果

| $\beta$ | $O_{ij}$ | $D_{ij}$    | $D_{ji}$    | $T_{ij}$ |
|---------|----------|-------------|-------------|----------|
| 0       | 22       | (8,0,0,0,1) | (4,0,0,0,1) | 0        |
| 1       | 22       | (8,0,0,0,1) | (4,0,0,0,0) | 0        |
| 2       | 22       | (1,0,0,0,0) | (5,0,0,0,0) | 0        |
| 3       | 22       | (5,1,0,0,0) | (1,1,0,0,0) | 0        |
| $\beta$ | $O_{ij}$ | $D_{ij}$    | $D_{ji}$    | $T_{ij}$ |
| 0       | -8       | (7,1,0,0,0) | (3,1,0,0,0) | 0        |
| 1       | -8       | (1,0,0,0,0) | (5,0,0,0,0) | 0        |
| 2       | -8       | (2,0,0,0,1) | (6,0,0,1,1) | 0        |
| 3       | -8       | (1,1,0,0,0) | (5,1,0,0,0) | 0        |
| 4       | -8       | (3,0,0,0,0) | (7,0,0,0,0) | 0        |
| 5       | -8       | (3,0,0,0,0) | (7,0,0,0,0) | 0        |

我們以  $P_1$  查詢圖片，而  $P_2, P_3, \dots, P_7$  為資料庫中圖片，因此分別針對配對為  $(P_1, P_2)$ 、 $(P_1, P_3)$ 、 $(P_1, P_4)$ 、 $(P_1, P_5)$ 、 $(P_1, P_6)$ 、 $(P_1, P_7)$  等做相似度之計算，其結果分別如下：

$(P_1, P_2)$ ：

因兩張圖之 OBJECT\_NUM 互不相符

$$S_D(P_2, P_1) = \frac{0}{6+6} = \frac{0}{12}$$

相似度為 0。

$(P_1, P_3)$ ：

$$\text{match}(t^{P_1}.22_0) = t^{P_3}.22_0, \text{match}(t^{P_1}.22_2) = t^{P_3}.22_1, \text{match}(t^{P_1}.-8_1) = t^{P_3}.-8_0$$

$$S_D(P_3, P_1) = \frac{(1+1)+(1+1)+(1+1)}{6+3} = \frac{6}{9}$$

相似度為  $\frac{2}{3}$ 。

$(P_1, P_4)$ ：

$$\text{match}(t^{P_1}.22_0) = t^{P_4}.22_0$$

$$S_D(P_4, P_1) = \frac{(1+1)}{6+1} = \frac{2}{7}$$

相似度為  $\frac{2}{7}$ 。

同理分別可得  $(P_1, P_5)$ 、 $(P_1, P_6)$  及  $(P_1, P_7)$  等相似度，我們得到如表 11 之結果：

表 11 資料庫圖片  $P_2 \sim P_7$  與查詢圖片  $P_1$  之相似度結果

|       | $P_2$  | $P_3$  | $P_4$  | $P_5$  | $P_6$  | $P_7$  |
|-------|--------|--------|--------|--------|--------|--------|
| $P_1$ | 0.0000 | 0.6667 | 0.2857 | 0.6667 | 0.5000 | 0.3750 |

我們可得知  $P_3$  和  $P_5$  與  $P_1$  的相似度最高，由以上結果可得知本論文所提出的方法與人類視覺感知一致。乳房 X 光片的相似度會因為其 ROI 數目或其鈣化點不同而有明顯的差異，若兩張 X 光片有相同數量的 ROI 且鈣化點在同樣的位置但其鈣化點的形狀不同，相對的相似度也會減低。

## 伍、結論

本論文是將 Extended 9D-SPA 表示法多方面加以修改使其能應用在乳房 X 光片之影像擷取上。本系統修改了 Extended 9D-SPA (Extended 9 Direction Spanning Areas) 表示法及其相似度計算方法，並將鈣化點重新編碼，讓原本只能應用在單一個 ROI 上的 Extended 9D-SPA 表示法能夠應用在多個 ROI 上，如此便能更實際應用於臨床上，讓醫師在乳房 X 光片的影像擷取上能更正確的擷取出所需的影像，以供診療之參考。實驗結

果顯示所提出之方法與人類視覺感知一致且具有良好之精確度並具滿足不同查詢需求之彈性。

未來希望能將影像資料庫建置完成，如此能夠更落實影像擷取資料庫的功能。此外，期望未來能完成使用者介面部分，讓系統更加人性化，使用起來能更方便。

### 參考文獻

1. [http://www.doh.gov.tw/CHT2006/DM/SEARCH\\_RESULT.aspx](http://www.doh.gov.tw/CHT2006/DM/SEARCH_RESULT.aspx)
2. 黃俊忠、任玄，“利用 X 光乳房攝影產生之紋理特徵影像在腫瘤偵測上之研究，” 中央大學資工系碩士在職專班畢業論文，2006。
3. [http://www.mmh.org.tw/taitam/gen\\_su/teach/breast.asp](http://www.mmh.org.tw/taitam/gen_su/teach/breast.asp)
4. Riloff, E., and Hollaar, L., “Text Databases and Information Retrieval,” ACM Computing Surveys, Vol. 28, No. 1, pp. 133-135, 1996.
5. 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, Vol. 22, No. 12, pp. 1349-1380, 2000.
6. Yoshitaka, A., and Ichikawa, T., “A Survey on Content-based Retrieval for Multimedia Databases,” IEEE Trans. Knowledge and Data Engineering, Vol. 11, No. 1, pp. 81-93, 1999.
7. Qiu, G., “Indexing Chromatic and Achromatic Patterns for Content-based Colour Image Retrieval,” Pattern Recognition, Vol. 35, No. 8, pp. 1675-1686, 2002.
8. Berretti, S., Bimbo, A.Del., and Vicario, E., “Spatial Arrangement of Color in Retrieval by Visual Similarity,” Pattern Recognition, Vol. 35, No. 8, pp. 1661-1674, 2002.
9. Chi, Y., and Leung, K.H., “ALSBIR : A Local-structure-based Image Retrieval,” Pattern Recognition, Vol. 40, pp. 244-261, 2007.
10. 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., Vol. 43, No. 2, pp. 488-502, 2007.
11. Huang, P.W., and Dai, S.K., “Image Retrieval by Texture Similarity,” Pattern Recognition, Vol. 36, pp. 665-679, 2003.
12. Huang, P.W., and Jean, Y.R., “Using 2D C+-string as Spatial Knowledge Representation for Image Database Systems,” Pattern Recognition, Vol. 27, No. 9, pp. 1249-1257, 1994.
13. Lee, Anthony J.T., and Chiu, H.P. “2D Z-string: A New Spatial Knowledge Representation for Image Databases,” Pattern Recognition Letters, Vol. 24, pp.3015–3026, 2003.
14. Lee, S. Y. and Hsu, F. J., “Spatial Reasoning and Similarity Retrieval of Images Using 2D C-String Knowledge Representation,” Pattern Recognition, Vol. 25, No. 3, pp. 305-318, 1992.
15. Nabil, M., Shepherd, J., and Ngu, A.H.H., “Picture Similarity Retrieval Using 2D Projection Interval Representation,” IEEE Trans. Knowl. Data Eng, Vol. 8, No. 4, pp.533-539, 1996.
16. Chang, S.K., Jungert, E., and Li, Y., “Representation and Retrieval of Symbolic Pictures Using Generalized 2D Strings,” Technical Report, University of Pittsburg, 1988.



17. Chang, C.C., "Spatial Match Retrieval of Symbolic Pictures," *Journal of Information Science and Engineering*, Vol. 7, pp. 405-422, Dec. 1991.
18. Huang, P.W., and Lee, C.H., "Image Database Design based on 9D-SPA Representation for Spatial Relations," *IEEE Transactions on Knowledge and Data Engineering*, Vol.16, No. 12, pp. 1486-1496, Dec. 2004.
19. [http://www.doh.gov.tw/cht2006/index\\_populace.aspx](http://www.doh.gov.tw/cht2006/index_populace.aspx).
20. Lin, T.P., and Chen, J.C.," Feature Extraction and Differentiation of Microcalcifications and Tumors of Mammogram," 國立雲林科技大學資訊管理系碩士畢業論文, 2007.
21. Kim, J. K., Park, H.W. "Statistical Textural Features for Detection of Microcalcifications in Digitized Mammograms", *IEEE Transaction on Medical Imaging*, Vol.18, No.3, March 1999.
22. 羅見順, 國立成功大學電機工程所博士班論文"乳房 X 光攝影之微鈣化群電腦輔助診斷系統", 2000.
23. Chin-Ling Ho, and Chang-Ju Yang , "The Improved Image Indexing Architecture Based on 9D-SPA," *International Journal of Advanced Information Technologies (IJAIT)*, Vol. 5, No. 1, 2011.
24. Verma, B., "Novel Network Architecture and Learning Algorithm for the Classification of Mass Abnormalities in Digitized Mammograms," *Artificial Intelligence in Medicine*, January, 2008.
25. Tourassi, Georgia, D., Ike, Robert, Singh, Swatee, and Harrawood, Brian, "Evaluating the Effect of Image Preprocessing on an Information-Theoretic CAD System in Mammography," *Academic Radiology*, May 2008.
26. Lin, T.P., and Chen, J.C.," Feature Extraction and Differentiation of Microcalcifications and Tumors of Mammogram," 國立雲林科技大學資訊管理系碩士畢業論文, 2007.
27. [http://www.doh.gov.tw/CHT2006/DM/SEARCH\\_RESULT.aspx](http://www.doh.gov.tw/CHT2006/DM/SEARCH_RESULT.aspx)
28. [http://www.mmh.org.tw/taitam/gen\\_su/teach/breast.asp](http://www.mmh.org.tw/taitam/gen_su/teach/breast.asp)
29. Mehmet Fatih Akay," Support vector machines combined with feature selection for breast cancer diagnosis," *Expert Systems with Applications*, Vol. 36, pp. 3240–3247, 2009.
30. Kunio Doi, "Computer-aided diagnosis in medical imaging: Historical review, current status and future potential," *Computerized Medical Imaging and Graphics*, Vol. 31, pp. 198–211, 2007.
31. Seiji Kumazawa et al., "An Investigation of Radiologists' Perception of Lesion Similarity: Observations with Paired Breast Masses on Mammograms and Paired Lung Nodules on CT Images," *Academic Radiology*, Vol. 15, pp. 887–894, 2008.

# Image Information Retrieval for Mammograms Containing multiple ROIs

Fu-Hua Yang

Department of Applied Information Sciences, Chung Shan Medical University, Taiwan  
jordanyang529@yahoo.com.tw

Li-Pin Hsu

Department of Applied Information Sciences, Chung Shan Medical University, Taiwan  
apple@csmu.edu.tw

abstract

In this paper, we have developed an efficient and practical database system for recognizing and retrieving patterns containing multiple ROIs in mammographic images that represent pathological processes. The purpose of this paper is to present the physician with a selection of similar cases with known diagnoses. There are several procedures carried out, first, due to the distribution and the shape of calcification are important bases to diagnose that the calcification is benign or malignant, we proposed a representation of mammographic image containing multiple ROIs, which can capture the spatial relationships and the shape of calcification. Secondly, we established the indexing method and an image database system with automatic indexing and retrieval capabilities in mammographic images. Experimental results showed that this image retrieval system is powerful in terms of accuracy and flexibility. Finally, the method proposed here can assist the physician in diagnosis of breast cancer by presenting the similar cases with known diagnoses.

keywords: mammographic image, micro-calcification, Extended 9D-SPA, ROI, shape feature.