

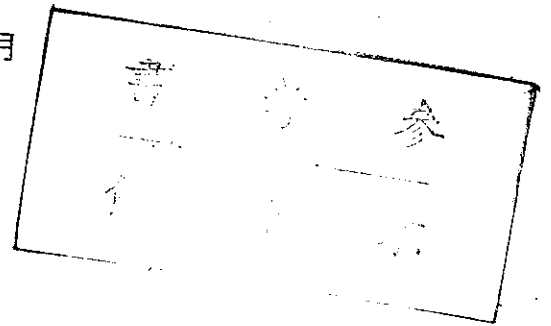
TR-87-002

重慶物體的辨認

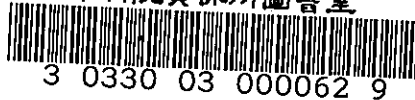
中央研究院資訊研究所

陳弘明 · 黃俊雄

中華民國七十六年一月



中研院資訊所圖書室



0062

摘要：

自然界中的事物通常是相互重疊的，然而人的眼睛能將它們正確的分辨出來，因此重疊物體的識別就成爲電腦視覺研究及工廠自動化中一項重要課題。

本篇研究報告主要是先用 Back Lighting (背後打光) 使 3D 物體經由影像感應器形成二元化影像，再利用物體影像之輪廓經由 Hough Transform 將重疊物體一一識別出來。

本研究在 Machine Intelligence Inc. 出品之 DS-100 上做實驗，取三種綠絲釘混在一起照像，根據實驗結果，其識別率與物體重疊的程度成反比，主要原因是重疊部份愈多，其可得到的邊點資料就相對減少。如果不重疊則由 DS-100 所提供 SRI 方法經由學習過程可將物體一一識別出來。

## 第一節：前言

美國機械智慧 (Machine Intelligence Inc.) 公司曾發展出一套 Vision Development System，簡稱 DS-100 [1] 其主要是先用物體在影像感應器上形成二元化影像，再利用影像之面積，連接性及橢圓等的一些特性來辨別物體。但其對於重疊物體的辨認卻不能達成。因此我們希望用另一種方法使得其不但有個識別的功能，且對重疊物體的識別也做得到。我們估且叫「邊點識別法」，其核心是 Hough Transform。邊點識別法主要有下列幾個部份：

### 一、找出物體的邊點：

我們可以掃描整個影像找到第一個邊點，或利用 DS-100 其所提供物體的質心位置，及最長徑的長度，角度來找出第一個邊點，然後利

用這個邊點找出整個物體的邊點，詳細在第三節中說明。

二、找出物體相對參考點的向量：

我們可以在影像平面座標上任意設定一點作為參考點（通常選質心或邊點  $x$  之方向及  $y$  之方向座標的平均值），求出邊點相對參考點的向量，（詳細於本文第三節中介紹）將其作為爾後判斷物體的依據，估且稱其為參考向量，而擁有這些參考向量之物體我們稱之為參考體。

三、利用 Generalized Hough Registration 之方法 [2]，找出參考體在重疊物體中最可能配對的部份，詳細於本文第四節中介紹。Hough Transform 之用途非常廣，交大蔡文祥博士曾用在印章鑑別 [3] 及清大許文星博士曾用在中文字識別 [4]。有關 Hough Transform 之誤差分析及花樣請看參考資料 [5-9]。

四、找出配對最成功的參考體。（詳見第五節）

五、取出參考體所屬之邊點。（詳見第五節）

六、有關的改進方法介紹於第六節，對於識別率幾乎可達百分之百的成果。

在正式介紹邊點識別法之前我們先於第二節中介紹一些 DS-100 的功能，可供有興趣的人作一個比較參考。

第二節：DS-100 的功能

DS-100 Development System 是用來發展 VS-100 Vision System（一種商業應用的二分化視覺系統）的一種有 Basic 語言能力的系統 [1]。DS-100 可直接利用光筆 (light pen) 或利用 Basic 程式來呼叫 VS-100 的一些程序或函數以便處理二元化的影像。

VS-100 可以接受一個灰度 (gray level) 影像，且將其二分化。同

時 VS-100 也可作連接性分析 (connectivity analysis) 而將物體的輪廓畫出來。VS-100 本身有一套訓練程序 (training routine) 可用來訓練其識別物體之能力。其主要是依據物體本身的投影 (經由背面打光) 以及其二元化後相對應橢圓等特性所產生的一些據數 (parameter) 作為比對的標準。每次對影像之一物體作比對，系統即算出該物體所屬之據數的值，然後將存在 file 中的所有原始資料一一比對，即可找出配對最成功的原型 (prototype)。這些方法是由 Stanford Research Institute 所發展出來的。如果這時訓練程序事先被呼叫則 VS-100 會將該物的數據值和原型的數據值的平均值作為原型的比對據數。其若干據數的關係如圖 1，其他的據數如圖 2，對應每個據數值，VS-100 都有一個程序供 Basic 程式呼叫，以便於程式進行中隨時取得據數的值，如圖 3。

很明顯的，這些功能用來識別重疊的物體是不夠的，因此在下面幾節將介紹的方法是利用物體投影的邊點資料來作重疊物體的識別。

### 第三節：如何建立原型特徵 (feature of prototype)

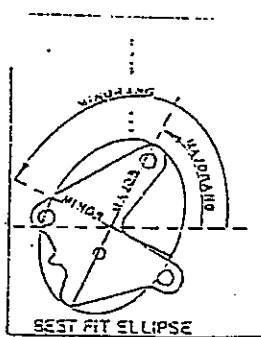
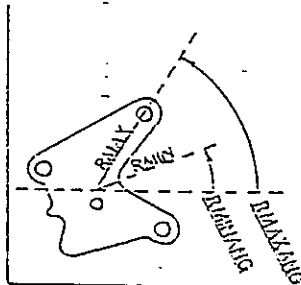
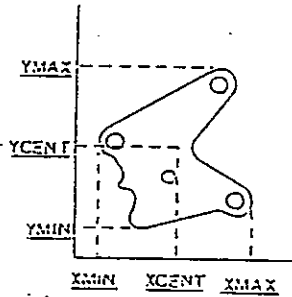
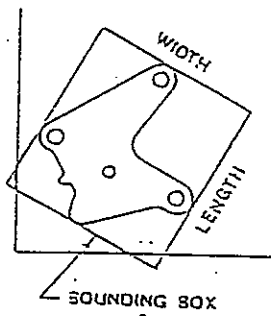
首先我們必須將原型的邊點 (boundary) 找出來。第一個邊點的找法可利用 DS-100 所提供物體質心位置及最長徑，最短徑的長度、角度，再利用簡單的公式：

$$(x,y)=(x_1,y_1)+(\text{length of radians}) * (\cos \theta, \sin \theta)$$

然後在座標  $(x,y)$  為中心點找一個  $3*3$  的陣列，如中心點其上下左右四點至少有一點為白點，則  $(x,y)$  為邊點，否則需作適度的調整，調整的方法詳見第五節。

其他邊點的找法是以目前邊點為中心點找一個  $3*3$  的陣列

$$\{w(i,j), 0 \leq i \leq 2, 0 \leq j \leq 2\},$$



ANGMOO - The difference between RMAXANG and RMINANG.

AREA - Area, in calibrated units, of a blob.

AVRAD - The average of the radius of a blob from its centroid to its perimeter.

AXRATIO - The ratio of the minor to the major axes of the equivalent ellipse.

EQUIVALENT ELLIPSE - That ellipse whose second moments equal the second moments of a given blob.

BLOB - A connected region in a digitized image, which may be black or white. A blob may fully surround other blobs (holes).

BOUNDING BOX - The smallest rectangle aligned with the axes of the equivalent ellipse that completely encompasses a blob.

BOXAREA - The area of the bounding box of a blob.

BOXRATIO - The ratio of boxarea to totalarea.

CENTER OF AREA - The center of a blob. Any straight line drawn through this point will bisect the blob into regions of equal area. The center of area (xcen and ycen) is determined by a blob's first moments,  $I_x$  and  $I_y$ .

CENTROID - Center of area.

CGDIST (-2) - The distance, squared, of a blob's center of area to the origin.

COLOR - Whether a blob is (0) black or (-1) white.

FIRST MOMENTS -  $I_x$  and  $I_y$ , where the sums are over all pixels in a blob. These are used to calculate the center of area of a blob.

HOLESAREA - Total area (in calibrated units) of the holes in a blob.

HOLERATIO - Ratio of hole area to total area.

LENGTH - The length of the bounding box that is aligned with the major axis of the equivalent ellipse.

LENRATIO - The ratio of the width to the length of the blob's bounding box.

MAJOR - Length of the major axis of the equivalent ellipse of a blob.

MAJORANG (THETA) - The angle of the major axis of a blob's equivalent ellipse relative to the x-axis. Theta is unique within a range of 180 degrees.

MINOR - Length of the minor axis of the equivalent ellipse of a blob.

MINORANG - The angle of the minor axis of a blob's equivalent ellipse relative to the x-axis. Minorang is unique within a range of 180 degrees.

NCELLS - Number of pixels in a blob.

NHOLES - Number of holes in a blob.

ORIENTATION - Direction of a blob with respect to the positive x-axis.

ORIGIN - The point (xzero, yzero).

PEROUNG -  $(4 * pi * area) / (perimeter ** 2)$ .

PERIMETER - Length of the perimeter of a blob.

PIXEL - A picture element; one entry in the matrix representing the digitized image.

PPDA - The ratio of the square of a blob's perimeter to its area.

RADRATIO - The ratio of rmin to rmax.

RMAX - The length of the maximum radius from the centroid to the perimeter.

RMAXANG - The angle of a blob's maximum radius.

RMIN - The length of the minimum radius from the centroid to the perimeter.

RMINANG - The angle of a blob's minimum radius.

SECOND MOMENTS -  $I_x^2$ ,  $I_y^2$ , and  $I_{xy}$ , where the sums are over the pixels in a blob. These are used to calculate the equivalent ellipse of a blob.

THETA - Majorang

TOTALAREA - The total area of a blob, including the area of its holes.

圖 1 DS-100 之物體特徵表。

OPERATING PARAMETER	DESCRIPTION:	DEFAULT:
Rejection Interval	A measure of how close a blob must be to some prototype to be called an instance of that prototype.	4
Right Window	Column number defining the right edge of the picture to be processed.	largest column
Threshold	The threshold value of the currently selected camera.	50
Top Window	Row number defining the top edge of the picture to be processed.	0
XSIZE	Length in the x-direction of a single pixel for the currently selected camera.	1.00
XZERO	The x-coordinate of the origin of the coordinate system.	0
YSIZE	Length in the y-direction of a single pixel for the currently selected camera.	1.00
YZERO	The y-coordinate of the origin of the coordinate system.	0

OPERATING PARAMETER	DESCRIPTION:	DEFAULT:
Bottom Window	Row number defining the bottom edge of the picture.	largest row number supported
Calibration Size	Diameter of the object used during calibration.	1
Camera Number	Current camera number (1 or 2)	1
Image Number	Number of the current image buffer.	1
Processing Time	Time in milliseconds required to complete the most recent Analyze New Image or Reanalyze Current Image command.	N/A
Left Window	Column number defining the left edge of the picture to be processed.	N/A
Lines in Linear Image	Number of lines defined to complete an image taken with the linear camera.	256 L
Minimum Blob Pixels	Area (in pixels) of the smallest acceptable blob.	10
Recognition Typeout	Number of items displayed in the lower left corner of the screen upon an attempt at recognition of an object.	3

圖 2. DS-100 操作參數表。

VISION/COMMAND CROSS REFERENCE

VISION FEATURE:

ANALYZE NEW IMAGE  
 ANGMOD  
 AREA  
 AVRAD  
 AXRATIO  
 BOTTOM WINDOW  
 BXRATIO  
 CALIBRATE CENTER  
 CALIBRATION SIZE  
 CAMERA NUMBER  
 CGDIST  
 COLOR  
 COMBINED ANALYSIS  
 CONCURRENT PROCESSING  
 CONNECTIVITY ANALYSIS  
 DEFERRED LINEAR IMAGE START  
 DISPLAY BANNER  
 ERROR RESTART  
 EXPANDED FEATURE DISPLAY  
 FIRST MOMENTS  
 FIXED LINEAR IMAGE SIZE  
 HOLEAREA  
 HOLE RATIO  
 IMAGE EXPANSION  
 IMAGE NUMBER  
 INCREMENTAL LINEAR IMAGE ANALYSIS  
 INSUFFICIENT STORAGE RESTART  
 KEEP ALL BLOBS  
 LEFT WINDOW  
 LENGTH  
 LENRATIO  
 LINES IN LINEAR IMAGE  
 LOWER PIXEL LIMIT  
 MAJOR  
 MAJORANG  
 MINIMUM BLOB PIXELS  
 MINOR  
 NBLOBS  
 NCELLS  
 NEGATIVE  
 NHOLES  
 NNOIST  
 NOISE SUPPRESSION

VISION CALL REF:

VPIC, VPIC  
 VFEAT  
 VFEAT, VSIZE, YAREA  
 VFEAT  
 VFEAT  
 VGET, VSET  
 VFEAT  
 VGETOP, VTURN  
 VGET, VSET  
 VGET, VSET  
 VFEAT  
 VFEAT, VCOLOR  
 VGETOP, VTURN  
 VGETOP, VTURN  
 VGETOP, VTURN  
 VGETOP, VTURN  
 VGETOP, VTURN  
 VGETOP, VTURN  
 VGETOP, VTURN  
 VGETOP, VTURN  
 VFEAT  
 VFEAT  
 VGETOP, VTURN  
 VGET, VSET  
 VGETOP, VTURN  
 VGETOP, VTURN  
 VGETOP, VTURN  
 VGET, VSET  
 VFEAT  
 VFEAT  
 VGET, VSET  
 VGET, VSET  
 VFEAT, VLIPE  
 VFEAT  
 VGET, VSET  
 VFEAT, VLIPE  
 VGET, VCOUNT  
 VFEAT  
 VGETOP, VTURN  
 VFEAT, VHOLES  
 VFEAT  
 VGETOP, VTURN

OBJECTS ONLY PROCESSED  
 ORIENTATION  
 OUTLINE BLOBS  
 PERIMETER  
 PERIMETERS  
 PEROUND  
 PIXEL COUNT  
 PPDA  
 PROCESS ALL BLOBS  
 PROCESSING TIME  
 RADRATIO  
 REANALYZE CURRENT IMAGE  
 RECOGNITION  
 RECOGNITION TYPEOUT  
 REJECTION INTERVAL  
 RIGHT WINDOW  
 RMIN  
 RMAX  
 RMINANG  
 RMAXANG  
 SECOND MOMENTS  
 SELECT BLOB OUTLINE  
 STROBE  
 THETA  
 THRESHOLD  
 TOTALAREA  
 TOTALCELLS  
 TRAINING  
 UPPER PIXEL LIMIT  
 WIDTH  
 XCENT  
 XDIFF  
 XMAX  
 XMIN  
 XSIZE  
 XZERO  
 YCENT  
 YDIFF  
 YMAX  
 YMIN  
 YSIZE  
 YZERO

VGETOP, VTURN  
 VFEAT  
 VGETOP, VTURN,  
 VFEAT  
 VGETOP, VTURN,  
 VFEAT, VPERNO  
 VGET  
 VFEAT  
 VGETOP, VTURN  
 VGET  
 VFEAT  
 VYANAL, VQANAL  
 VGETOP, VTURN  
 VGET, VSET  
 VGET, VSET  
 VGET, VSET  
 VFEAT, VRAD  
 VFEAT, VRAD  
 VFEAT, VRAD  
 VFEAT, VRAD  
 VGETOP, VTURN  
 VGETOP, VTURN  
 VGETOP, VTURN  
 VFEAT, VTHETA, VLIPE  
 VGET, VSET  
 VFEAT, VSIZE  
 VFEAT  
 VGETOP, VTURN  
 VGET, VSET  
 VFEAT  
 VFEAT, VPOS  
 VFEAT  
 VFEAT, VXTENT  
 VFEAT, VXTENT  
 VGET, VSET  
 VGET, VSET  
 VFEAT, VPOS  
 VFEAT  
 VFEAT, VXTENT  
 VFEAT, VXTENT  
 VGET, VSET  
 VGET, VSET

，自上一個邊點的位置開始以反（正）時鐘的方向掃描，找到下個邊點。如圖 4，若現在邊點為  $w(1,1)$ ，上一個邊點為  $w(0,1)$  方向是反時鐘掃描，則下一個找到的邊點為  $w(2,1)$ 。要完整找出整個物體的邊點，掃描方向須一致沿著白點（或黑點）部份，所以當其違反這原則時我們必須改變其掃描方向。圖 5 中，我們由高亮點（high light）來表示邊點。

有了邊點我們可算出其相對於質心的參考向量，如圖 6，如此就完整的建立了參考體的原型特徵了。

#### 第四節：識別重疊物體

識別重疊物體主要分為三個步驟：

- 1 找出重疊物體的邊點：這和上一節的方法一樣。
- 2 找出最好的參考體在最適當度下之質心位置（參考點的位置）：  
角

我們用的方法是 Generalized Hough Registration.

將上節所求得的參考向量全部反向，沿著重疊物的邊點，以其為中心，各向量自四面八方射出，其向量所指到的位置作加 1 的記錄，反覆對不同的參考體及每旋轉  $5^\circ$  作一次，如此我們可得到一個 4 維的陣列，其四個據數分別是  $x$  座標 (0-255)， $y$  座標 (0-255)，角度 ( $0^\circ$ - $360^\circ$ ) 及參考體的編號 (1-參考體的數目)。在這個 4 維陣列中以一個  $5*5$  的陣列，在不同的參考體，不同的角度  $\theta$  及不同的  $x, y$  座標上掃描，找出其 25 個位址所記錄值的總和，最後可得到在某個參考體，某個角度及某個位置時，其記錄值總和最大，則此點為某個參考體，旋轉  $\theta$  角度後之參考點的位置。

- 3 最後的工作則是將屬於參考體的邊點自重疊物中去掉，用殘餘



	0	1	2
0	0	1	0
1	0	1	1
2	0	1	1

圖 4.  $w(1,1)$  為目前邊點位置，  
 $w(0,1)$  為上一個邊點位置，  
 $w(2,1)$  為下一個邊點位置，  
 方向是反時鐘方向。

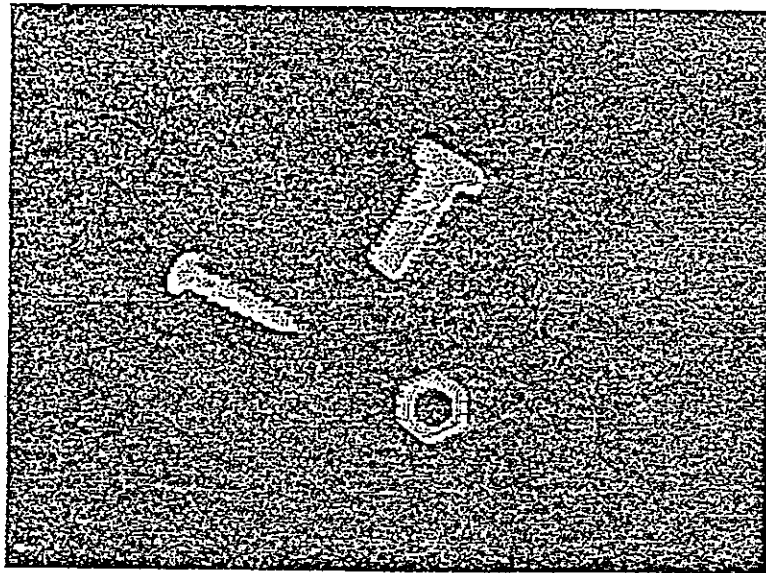


圖 5. Boundary 用 High light 表示。



的邊點繼續 1 2 的動作，直到邊點全部去掉為止。

去掉邊點的方法是將參考體之參考向量全部旋轉  $\theta$  度，以 2 中找出的參考點為中心，其向量所指的位置的邊點即為欲去掉之邊點。當然，由於誤差的關係，我們不一定能找到欲去掉之點。調整的方法詳如第五節。

## 第五節：改進方法

### 1 Registration 的改進：

在向量旋轉的運算中，不可避免一些誤差，以至使得參考向量無法落於正確位址，影響到參考點的正確性，因此我們在作 registration 時不僅對參考向量所指之位址作加 1 的記錄，而是對其 3\*3 鄰點亦作加 1 的記錄。

### 2 調整邊點的位置：

在抽取參考體邊點和尋找第一個邊點時，都可能無法正確落在邊點的位置，因此我們用一個 3\*3 的陣列，保持向量的方向不變，找其鄰近邊點中距離最近的點，作為其將被抽取的點。如圖 7。

### 3 減少 registration 的次數：

重疊物不一定佔了整個影像畫面，故在作 registration 及尋找最大記錄值時對整張影像處理，是非常浪費時間的。我們在尋找重疊物的邊點時間時找出 HX ( X 方向最大之座標值 )，HY ( Y 方向最大之座標值 )，LX ( X 方向最小之座標值 )，LY ( Y 方向最小之座標值 )，故只須對

$$LX \leq X \leq HX \quad \text{and} \quad LY \leq Y \leq HY$$

之中的點作 registration，尋找最大的記錄值，可減少很多的 computing time。

	0	1	2
0	1	1	0
1	1	0	0
2	1	0	0

圖 7.  $w(1,1)$  是向量所指的位置，並非邊點，參考向量之角度  $22.5^\circ \leq \theta \leq 67.5^\circ$  則  $w(2,0)$  為其矯正後之點。

## 第六節：實驗結果

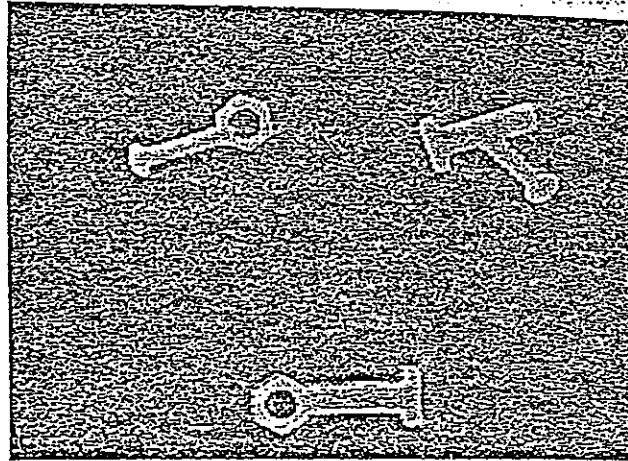
圖 8 和圖 9 是我們對同一張影像的兩個重疊物作識別的結果，影像座標系是由左上角為原點漸增加的。

我們可看出圖 8 的一號重疊物（用高亮度表示其邊點者）是由兩種不同螺絲釘組成，在我們的 algorithm 中的第一個 loop 找出 1 號螺絲釘其位置是 (141,206) ，且旋轉角度  $\theta = 4.10122$  （徑度），將 1 號螺絲釘之邊點自重疊物中去掉，則第二個 loop 找出 2 號螺絲釘其位置是 (107,208) ，且其旋轉角度為 1.83246 （徑度）。

注意在第一個 loop 中 2 號螺絲釘之位置為 (108,209) 與第二個 loop 找出的位置很接近，這表示我們可以一次就將重疊物的各個參考體全找出來，因其牽涉到複雜的數學模式的建立與說明，這將是我們研究的下一個目標。

也許你覺得奇怪，在圖 8 第一個 loop 和第二個 loop 找出的 2 號螺絲釘，其  $\theta$  為何不同？這是因為 2 號螺絲釘是一對稱之六邊形，因其不只一個角度能得到最大的 register score，圖 9 是我們對第 2 號重疊物（在右上方）的識別。其結果亦很好。

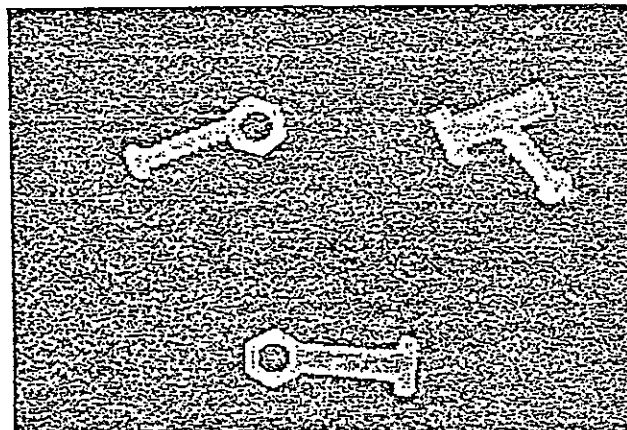
圖 8.



BLOB 1 CX= 129 CY= 208  
BOUNDRY ELEMENT:

BLOB 1 HAS 179 BOUNDRY ELEMENT  
PROTOTYPE 1 ,THETA= 4.10122 REGIST SCORE= 632 X= 141 Y= 206  
PROTOTYPE 2 ,THETA= .8726 REGIST SCORE= 332 X= 108 Y= 209  
PROTOTYPE 3 ,THETA= 1.04712 REGIST SCORE= 322 X= 145 Y= 204  
THE BEST PROTOTYPE IS 1 X= 141 Y= 206 THETA= 4.10122 REGIST SCORE= 632  
PROTOTYPE 1 LOCATES AT 141 206  
PROTOTYPE 1 ,THETA= 1.83246 REGIST SCORE= 203 X= 111 Y= 197  
PROTOTYPE 2 ,THETA= 1.83246 REGIST SCORE= 313 X= 107 Y= 203  
PROTOTYPE 3 ,THETA= 5.41012 REGIST SCORE= 168 X= 103 Y= 199  
THE BEST PROTOTYPE IS 2 X= 107 Y= 208 THETA= 1.83246 REGIST SCORE= 313  
PROTOTYPE 2 LOCATES AT 107 208

圖 9.



BLOB 2 CX= 195 CY= 69  
BOUNDRY ELEMENT:

BLOB 2 HAS 214 BOUNDRY ELEMENT  
PROTOTYPE 1 ,THETA= 1.3089 REGIST SCORE= 664 X= 189 Y= 61  
PROTOTYPE 2 ,THETA= .61082 REGIST SCORE= 231 X= 201 Y= 61  
PROTOTYPE 3 ,THETA= 6.10819 REGIST SCORE= 433 X= 205 Y= 79  
THE BEST PROTOTYPE IS 1 X= 189 Y= 61 THETA= 1.3089 REGIST SCORE= 664  
PROTOTYPE 1 LOCATES AT 189 61  
PROTOTYPE 1 ,THETA= 0 REGIST SCORE= 277 X= 202 Y= 75  
PROTOTYPE 2 ,THETA= 1.7452 REGIST SCORE= 184 X= 212 Y= 86  
PROTOTYPE 3 ,THETA= 6.19545 REGIST SCORE= 375 X= 205 Y= 80  
THE BEST PROTOTYPE IS 3 X= 205 Y= 80 THETA= 6.19545 REGIST SCORE= 375  
PROTOTYPE 3 LOCATES AT 205 80

參考資料：

- 1) DS-100 technique manuals, Machine Intelligence Inc., California, USA.
- 2) Ballard, D. H. "Generalizing the Hough Transform to detect arbitrary shapes" Pattern Recognition 13, 2, 1981, 111-122.
- 3) 蔡文祥、陳嘉玫：變形印文自動鑑別。交大計算機工程研究所碩士論文（民國75年）。
- 4) M. Y. Chen, W. H. Hsu and F. H. Cheng, "An application of Hough Transform to recognition of handwritten Chinese characters", Proceedings of International Computer Symposium 1986, R.O.C., pp.719-727.
- 5) S. D. Shapiro, "Transformations for the computer detection of curves in noisy pictures", Computer Graphics and Image Processing, 4, 1975, 328-338.
- 6) L. S. Davis, "Hierarchical generalized Hough transforms and line-segment based generalized Hough Transforms", Pattern Recognition, vol. 15, 1982, 277-285.
- 7) T. M. van Veen and F. C. A. Groen, "Discretization errors in the Hough Transform", Pattern Recognition, vol. 14, 1981, 137-145.
- 8) C. M. Brown, "Inherent bias and noise in the Hough Transform", IEEE trans. on PAMI, vol. PAMI-5, 1983, 493-505.
- 9) H. Wechsler and J. Sklansky, "Finding the rib cage in chest radiographs", Pattern Recognition, vol. 9, 1977, 21-30.