

# 國立交通大學

資訊科學系

碩士論文



果蠅腦細胞計數之自動方法設計

Toward designing an automatic method to count  
neuron cells in the fly brain

研究生：陳孟誌

指導教授：荊宇泰 教授

中華民國九十四年七月

果蠅腦細胞計數之自動方法設計

Toward designing an automatic method to count neuron cells in the fly brain

研究生：陳孟誌

Student : Meng-Chih Chen

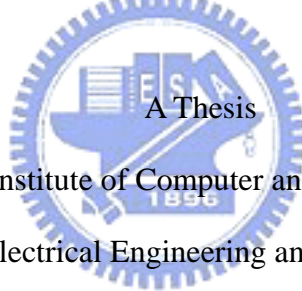
指導教授：荊宇泰

Advisor : Yu-Tai Ching

國立交通大學

資訊科學系

碩士論文



Submitted to Institute 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

July 2005

Hsinchu, Taiwan, Republic of China

中華民國九十四年七月

# 果蠅腦細胞計數之自動方法設計

學生：陳孟誌

指導教授：荊宇泰 博士

國立交通大學資訊科學系



在醫學和生物研究上，常常需要從特定影像上得到一些重要的資訊，例如血液中紅血球和瘧原蟲的比例或者是一隻果蠅腦中的腦細胞個數，要完成這個工作可以利用手動將影像上的資訊圈選出來，但隨著影像解析度的增加這個工作變的複雜且耗時，本論文提出利用 Matched Filter 和 k-means clustering 為基礎所發展的顆粒擷取方法。並將該方法應用在果蠅腦細胞的計數實驗上。

關鍵字：Particle picking，Matched Filter，k-means clustering

# Toward designing an automatic method to count neuron cells in the fly brain

Student : Meng-Chih Chen

Advisor : Dr. Yu-Tai Ching

Institute of Computer and Information Science

National Chiao Tung University



Abstract

We usually have to obtain some information important from specific image such as erythrocyte and the proportion of plasmodium in blood , or the number of neuron cells in the fly brain when we research in medical science and biology . This task can be carried out manually . However , as the image resolution approaches the atomic level , it become complex and time-consuming . In this thesis , we developed a particle picking algorithm based on Matched Filter and K-means clustering . The method was applied to count neuron cells in the fly brain .

Keyword : Particle picking , Matched Filter , k-means clustering

## 誌謝

本篇論文能得以完成，首先要感謝我的指導教授荆宇泰老師的指導，在這兩年內由於他的專業知識傳授和細心的指導才使我能順利完成碩士論文。同時也要感謝清華大學生物科技所的江安世老師和學弟永睿提供生物方面的專業知識和資料。另一方也要感謝口試委員江安世老師和謝昌煥老師對於論文謬誤處的指正而使這篇論文可以更加正確完整。感謝實驗室的學長、同學和學弟們，他們除了在學業知識上給我完全的支持外也在生活心靈上幫助我成長，更陪我度過這兩年的精華歲月。

最後我要感謝我的家人，謝謝他們在這兩年內給我的包容和支持，讓我得以安心的完成我的碩士論文，謝謝大家。



# 目錄

中文摘要	i
英文摘要	ii
致謝	iii
目錄	iv
圖表目錄	v
<b>第一章 序論</b>	<b>1</b>
1.1 顆粒擷取技術	1
1.2 顯微鏡取像技術	1
1.3 研究動機與目的	3
1.4 論文架構	3
<b>第二章 相關研究</b>	<b>5</b>
2.1 重疊和相連	5
2.2 低訊號雜訊比	7
2.3 幾何形狀特性	9
2.4 影像灰階分佈特性	10
2.5 結論	12
<b>第三章 3D 細胞影像分析計數方法</b>	<b>14</b>
3.1 Matched Filter	14
3.2 K-means Clustering	17
3.3 自動計數方法	18
3.3.1 前處理	18
3.3.2 尋找 Local Max	19
3.3.3 Group 分析	19
3.3.4 增加 Candidate	20
3.4 討論	21
<b>第四章 實驗結果及討論</b>	<b>22</b>
實驗資料組一	22
實驗資料組二	26
實驗資料組三	27
實驗結果數據	30
<b>第五章 結論及未來工作展望</b>	<b>31</b>
參考文獻	32
附錄	34

# 圖表目錄

圖 1-1 顯微鏡取像示意圖	3
圖 2-1 尋找初始圓和曲線線段	6
圖 2-2 從圖 2-1 分離右半邊的圓	6
圖 2-3 Leticia V. Guimarães 提出的 algorithm 結果	7
圖 2-4 Zeyun Yu 提出的 algorithm 結果	8
圖 2-5 Zeyun Yu 的實驗過程產生的相對結果	10
圖 2-6 Zeyun Yu 提出以幾何形狀為基礎的 algorithm 的結果	10
圖 2-7 目標影像	12
圖 2-8 目標影像直方圖	12
圖 2-9 pixel 的八個鄰居	12
圖 3-1 1D matched filter 範例	16
圖 3-2 2D 中偵測血管的 Matched Filter 例子	16
圖 3-3 matched filter 實際影像效果	17
圖 3-4 分群法的例子	18
圖 3-5 程式流程圖	18
圖 4-1 實驗資料組一部分影像切片	23
圖 4-2 matched filter mask 對影像的影響結果	23
圖 4-3 圖 4-1 經過 matched filter 之後的結果	24
圖 4-4 圖 4-1 經過程式計算後的結果 (選取機制)	25
圖 4-5 圖 4-1 經過程式計算後的結果 (著色機制)	26
圖 4-6 實驗資料組二部份切片影像	27
圖 4-7 圖 4-6 經過程式計算後的結果 (著色機制)	27
圖 4-8 實驗資料組三部份切片影像	28
圖 4-9 圖 4-8 經過程式計算後的結果 (著色機制)	29
圖 A-1 系統介面圖	34
圖 A-2 輸入視窗	35
圖 A-3 Connected File 儲存格式	36
圖 A-4 系統主要的兩個子視窗	36
表 4-1	30
表 4-2	30

# 第一章

## 序論

### 1.1 顆粒擷取技術

在影像處理和圖形識別的領域中，有一個領域的研究在探討如何正確有效的從一張影像中擷取我們想要的顆粒，也就是在一張影像中找出和我們想要找的顆粒大小形狀和紋理類似的區域，它有點類似於傳統影像處理的 Segmentation，但在應用和處理的技術上卻又有點不同，那就是顆粒擷取技術 (particle picking) 或稱為顆粒偵測技術 (particle detection)，這幾年來由於取像技術的進步，影像的解析度越來越高，以醫學上的應用來說，一張血液細胞的影像醫生會希望知道裡面有多少個白血球又或者在瘧疾患者的血液中紅血球和瘧原蟲的比例，這些資訊都可以利用顆粒擷取技術來得到，在[1-2]這幾篇論文就是針對這些需求所做的研究，另一方面，生物的研究也需要用到類似的技術，當生物專家對某些生物的特定部分組織或細胞作染色取像後他會希望知道染色的細胞數目跟形狀等等的資訊，這也就是本篇論文主要解決的問題，目前也已經有許多同類型的論文發表在一些知名的期刊上，我們將在第二章作較詳盡和完整的介紹。

### 1.2 顯微鏡取像技術

由於我們提出的方法針對的目標影像都是由顯微鏡取像得到的，所以在這一小節中我們大致介紹一些較基礎的顯微鏡取像方法和它的限制，其中包括光學顯微鏡 (Light microscopy)、螢光顯微鏡 (Fluorescence microscopy) 和共軛焦顯微鏡 (Confocal microsc-



opy)。

### I. 光學顯微鏡

一般來說，光學顯微鏡利用可見光中不被目標物體所吸收的光來成像，圖 1-1 左側為光學顯微鏡的示意圖，光從光源利用一個透鏡聚焦在目標物體上，能通過目標物體的光再經過兩個透鏡的放大和聚焦作用成像在偵測器上，利用光學顯微鏡來取像的影像其解晰度受限於光源的波長為  $0.2\mu\text{m}$ ，也就是藍光波長的一半。

### II. 螢光顯微鏡

螢光顯微鏡不同於光學顯微鏡，它利用螢光分子吸收特定波長的光並放射出較長波長的光來成像，圖 1-1 中間的圖為螢光顯微鏡的示意圖，光從光源經過一個濾波鏡後經由雙色向面鏡聚焦到目標物體上，目標物體上的螢光分子吸收光線後放出特定波長的光再穿過雙色向面鏡，之後會再經過另一個濾波鏡然後才聚焦在偵測器上，由於螢光顯微鏡的光源一樣為可見光，所以所取的影像解晰度一樣受限在  $0.2\mu\text{m}$ 。

### III. 共軛焦顯微鏡

共軛焦顯微鏡成像的方式約略跟螢光顯微鏡類似，是利用螢光分子放射出來的光線來成像，不同點在於光源的選擇，共軛焦顯微鏡利用的是擁有好的聚焦性的雷射來當作光源，圖 1-1 右側是共軛焦顯微鏡的示意圖，雷射光線經由雙色向面鏡後聚焦在目標物體上，物體裡的螢光分子吸收光線後再放射出其他波長的光線，放射出的光線經過雙色向面鏡後再經過一個針孔之後成像在偵測器上，由於使用光源不同，共軛焦顯微鏡的解晰度比起前述兩種顯微鏡會好很多。

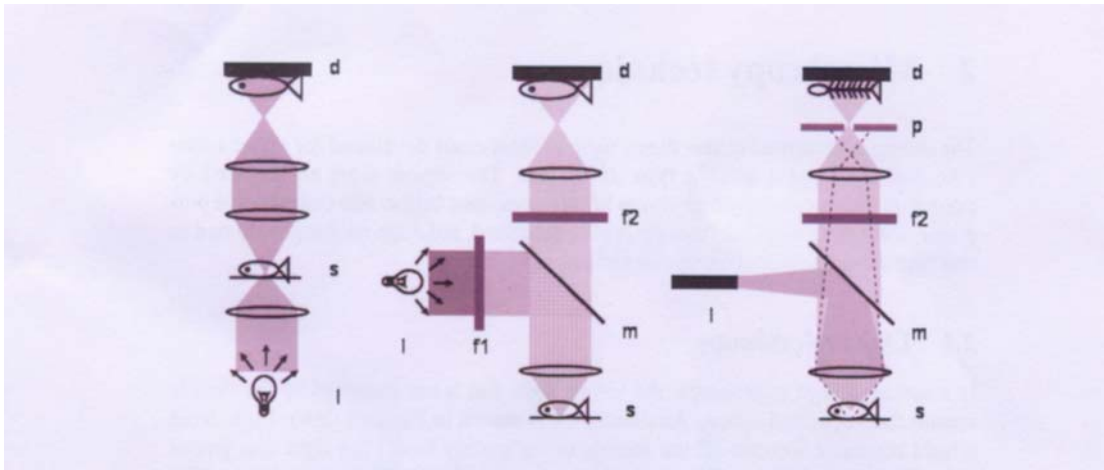


圖 1-1：光學顯微鏡(左)，螢光顯微鏡(中)，共軛焦顯微鏡(右)。所有顯微鏡都擁有一個光源(I)，一個樣本(s)，和一個偵測器(d)。除此之外螢光顯微鏡還有兩個濾波鏡(f1，f2)和一個雙色向面鏡(m)，而共軛焦顯微鏡則有一個濾波鏡(f2)，一個雙色向面鏡(m)和一個針孔(p)。



### 1.3 研究動機與目的

在生物和醫學的研究上，生物體某些特殊部位的細胞數目和位置在研究生物疾病上扮演著重要的角色，如今，由清華大學江安世老師所指導的團隊認為，果蠅腦袋中的細胞數目可以用來幫助判斷果蠅的疾病，例如老年癡呆等等，所以我們希望可以發展出一種方法，利用這種方法來計算一組 3D 的腦細胞影像中包含多少個細胞，更進一步希望可以正確的標定出細胞的中心位置，以期能幫助生物方面的研究發展。

### 1.4 論文架構

本論文共分五章，本章為論文之導論和基本技術之介紹。第二章將會整理出過去和顆粒擷取相關的一些論文研究。第三章則針對我們所設計的自動計數方法的技術和方法

來討論。第四章會顯示實驗結果和正確性。第五章則對未來的改進工作做一個整理和總結。



## 第二章

# 相關研究

在本章中，我們將探討過去一些顆粒擷取技術的研究，我們大致上將這些研究依照它們的特別處加以分類，並提出已經發表的論文來探討它們的做法。首先我們針對要解決的問題來分類，主要的問題分為：解決重疊和相連性的問題和解決低訊號雜訊比的問題，另一方面，我們依照解決問題的方法來分類，分為：依幾何形狀特性和依影像灰階分佈來解決問題的兩種分法，這四種分類只是為了突顯各篇論文的特點，但卻不是獨立的，像[2]雖然我們將它放在解決低訊號雜訊比的分類來討論，但它也算是依照影像灰階分佈的解決分法。

接下來我們在 2.1 節中會探討重疊和相連的問題，在 2.2 節中則會探討解決低訊號雜訊比的問題，然後在第 2.3 和 2.4 節中我們會討論依照幾何形狀和影像灰階分佈為基礎的方法，最後在 2.5 節中我們會做個總結。

### 2.1 重疊和相連

在 particle picking 的應用上,許多情況下會發生顆粒部分互相重疊或相連的情況，Leticia V. Guimarães 在[3]中提出一個以 polygonal approximation 為基礎的演算法來解決顆粒部份重疊相鄰的問題，首先利用 Ramer's Algorithm 找到位於顆粒邊上相距最遠的兩個點，a、b 這兩點的連線則成為初始圓的直徑，所以可以利用直徑將圓心和半徑定義出來，緊接著利用直徑將初始圓劃分為 A 和 B 兩個半圓，在每個半圓中找出落在顆粒邊

上且離圓心最近的點，分別得到  $\min\_A$  和  $\min\_B$ ，利用下面的準則來判斷是否將初始圓做更精確的劃分

If  $\min\_A \leq FS * r$  AND  $\min\_B \leq FS * r$

Then segment the curve

Else the object is a circle

這裡的  $FS$  所代表的是一個相似度的參數，其值介於 0 到 1 之間，在圖 2-1 的例子中， $\min\_A$  和  $\min\_B$  都小於  $FS * r$  所以這個初始圓會被作較精確的劃分，如圖 2-2 所示，首先先看右半邊的圓，利用相同的方法找出直徑和圓心，並分析新的  $\min\_A$  和  $\min\_B$ ，而左半邊的圓也做同樣的分析

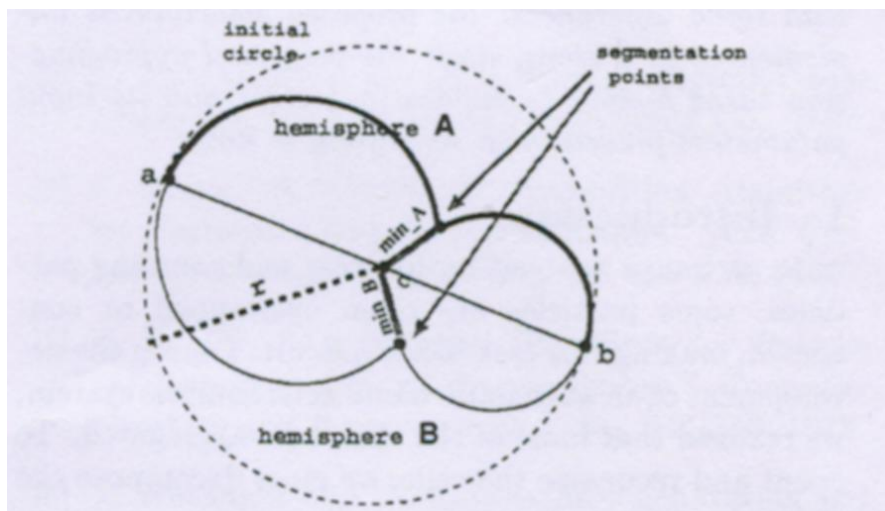


圖 2-1 : 找尋初始圓和曲線線段的點

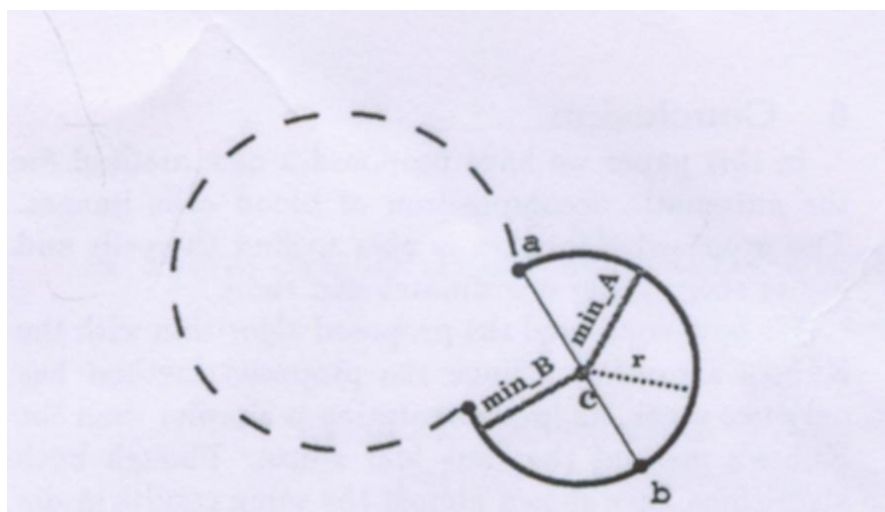


圖 2-2 : 從圖 2-1 分離右半邊的圓

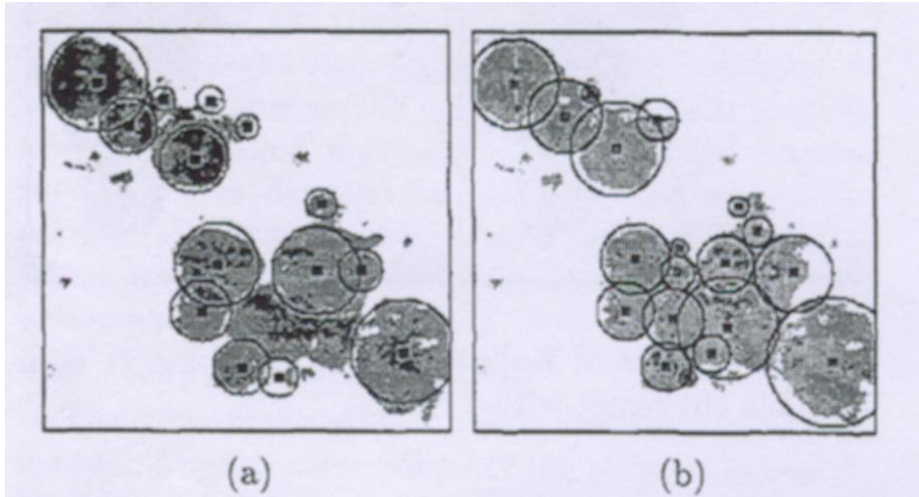


圖 2-3 : (a)Kubo's algorithm 的結果 (b) Leticia V. Guimarães 提出的 algorithm 的結果

## 2.2 低訊號雜訊比(signal-to-noise ratio)

在這部分的討論中是以一個低的訊號雜訊比(signal-to-noise ratio)為主要解決問題的對象，Zeyun Yu 在[4]中提出一個方法來處理類似的問題，首先設一個 threshold 值對影像作篩選,式子如下

$$h_0 = I_{\min} + a * (I_{\max} - I_{\min}) \quad (1)$$

$$g(\bar{x}) = h_0 - f(\bar{x}) \quad \text{if } f(\bar{x}) \leq h_0 \quad (2)$$

$$0 \quad \text{otherwise} \quad (3)$$

這裡的 $I_{\min}$ 和 $I_{\max}$ 是相對於整張影像灰階值的最大值和最小值,而在目標影像中,顆粒的灰階值比背景灰階值小,所以利用上述的threshold後的結果會得到每個點是顆粒一部分的機率值,接著模擬萬有引力的關係,將 2D影像上的訊號當成顆粒賦予一個正的質量而相對的背景則賦予零的質量,然後開始模擬萬有引力,

$$f(O_1; O_2) = \frac{c * m_1 * m_2}{d^2} \quad (4)$$

所以每個顆粒在影像上的任意點 A 所產生的力為



$$\left\| f(O; A) = \frac{c * m}{d^2} \right\| \quad (5)$$

由於一個點 A 所受的力不一定為唯一所以在點 A 的物體移動的方向決定在 A 點所受力的合力,經過推導可以將移動方式定義如下

$$A\vec{M} = \frac{1}{2} \min(1.0, \frac{\bar{m}}{m_0} * \|\bar{v}\|) * A\vec{B} \quad (6)$$

慢慢的較小的顆粒會因引力的關係聚在一起成為一個具有較大質量的新顆粒,直到整張影像所形成的引力空間達到平衡為止,在完成萬有引力模擬後找出來的數個顆粒中,可能有些是錯誤的,所以利用一些分析將錯誤的顆粒刪除,

1. 在被標定出來的一個區域方塊中,灰階值的變化量(variance)是第一個標準,如果這個灰階值的變化量小於一個事先給定的值,值將這個顆粒視為錯誤顆粒
2. 這個標準建構在假設顆粒會落在標定出來的區域方塊的中央部份,所以利用計算一個圓形的平均灰階值和計算方塊剩餘部份的平均灰階值來判斷,假設這兩個平均灰階值太接近則視為錯誤的顆粒

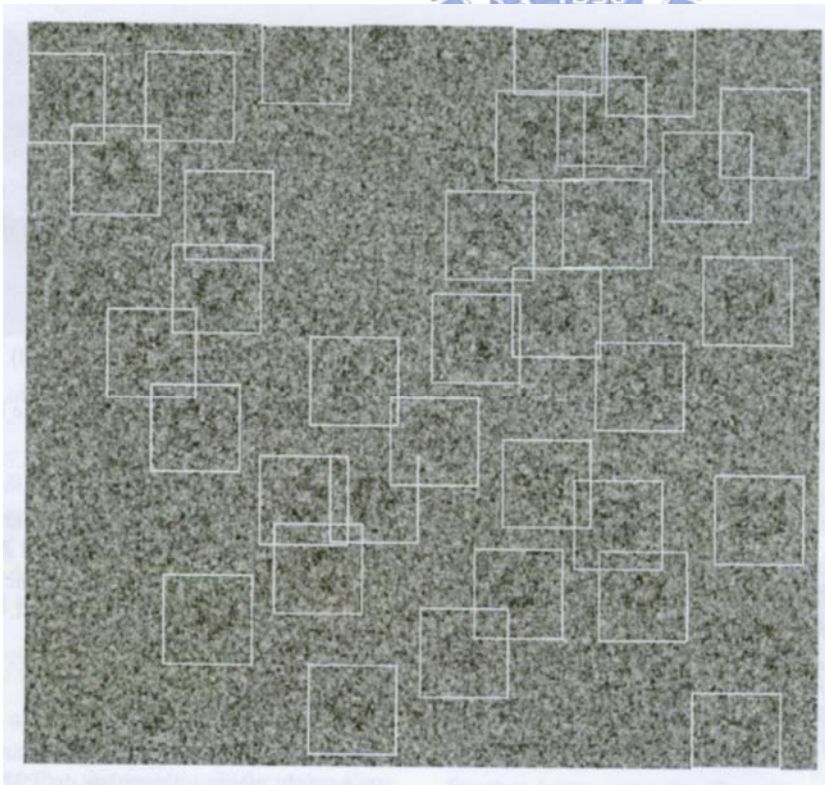


圖 2-4 : Zeyun Yu 所提出的 algorithm 的結果

## 2.3 幾何形狀特性

有些研究是利用幾何形狀的特性來實做 particle picking,例如 Zeyun Yu 在[5]中提出的方法，首先利用 Canny edge detector 將可能的邊找出來，然後利用找出來的邊做一次 distance transform，所得出來的值代表著這個 pixel 距離邊的距離，所以可以發現在 distance map 上的 local maxima 可能是圓形顆粒圓心發生的地方，所以將這些 pixel 當成初始圓的圓心，接下來，因為假設顆粒都是單獨存在而不會重疊或相連，所以判斷初始的圓是否有重疊，如果有就利用下面的準則去選擇留下來的顆粒

$$Adv(C) = \frac{1}{n} \sum_{|d(P,C)-Rd|\leq 1} DT(P) \quad (7)$$

其中 n 為符合  $|d(P,C)-Rd|\leq 1$  的 pixel 個數， $d(*,*)$  是兩個 pixel 的 Euclidean 距離， $DT(P)$  是 P 點在 Distance Map 上的值。



在做完刪減的動作後，剩下的顆粒要做一個 refinement 的動作，來減少誤差，[3]所用的方式如下

$$F(C) = \frac{1}{n} \sum_{|d(P,C)-Rd|\leq 1} P, NN(P) \quad (8)$$

其中 n 為符合  $|d(P,C)-Rd|\leq 1$  的 pixel 個數， $d(*,*)$  是兩個 pixel 的 Euclidean 距離，而  $NN(P)$  則為距離 P 最近的點， $P, NN(P)$  則代表從 P 到  $NN(P)$  的向量，顆粒的中心會依上述公式慢慢做調整直到  $F(C')$  的值小於一個給定的值為止。

另一方面，[3]中也提出找出矩形顆粒的方法，方法類似找出圓形顆粒的作法，但是因為矩形不像圓形,所以會有旋轉角度的問題,但因為矩形有四個邊角可以利用，所以可以利用 Voronoi diagram 來找出可能的邊角，利用邊角可以唯二初始出兩個可能的矩形，接下來一樣判斷顆粒是否有重疊，有則利用類似圓形的準則判別去留，之後再做 refi-



nement。

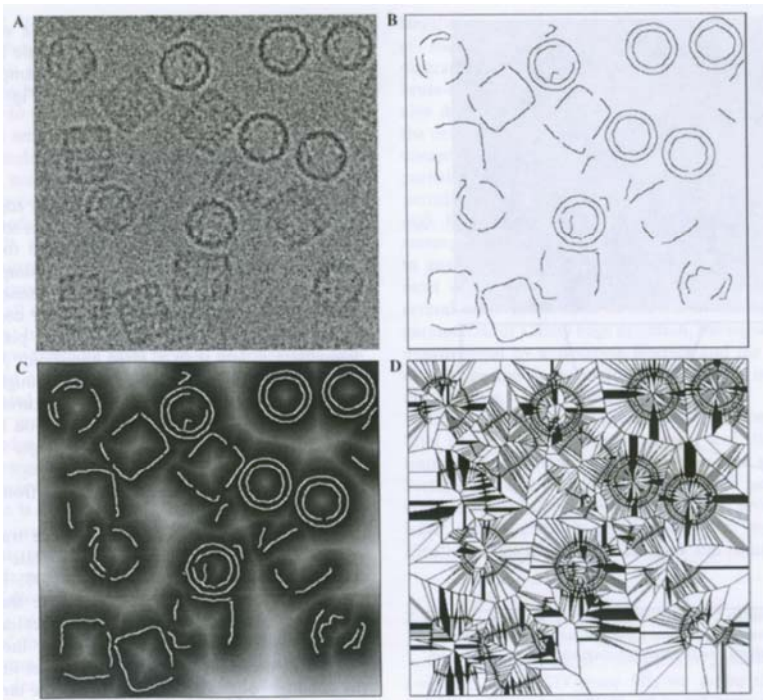


圖 2-5 : A.原始影像 B. Edge Map C. Distance transform and D. Voronoi diagram

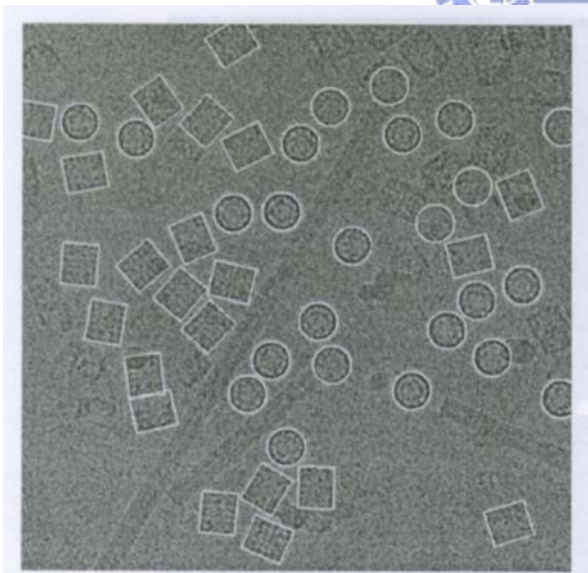


圖 2-6 : Zeyun Yu 提出以幾何形狀為基礎的 algorithm 的實驗結果

## 2.4 影像灰階分佈特性

這部份的研究是以灰階分佈特性為基礎來探討顆粒擷取技術，在[6]中首先會對整組影像做直方圖（histogram），從目標圖檔和直方圖來看可以發現背景和細胞之間的對比並不高且在灰階分佈統計圖中也只有一個峰（peak），另一方面細胞的形狀並不固定，所以[6]提出利用一個“區域調整式的臨界值方法”（local adaptive thresholding method）來達到分割和顆粒擷取的目的。

首先對目標影像做一個直方圖等化（histogram equalization），這個處理可以達到讓對比度提高的作用，緊接著再利用一個低通濾波器（low pass filter），這可以減低人為雜訊的影響，也降低每個細胞影像的灰階值分佈範圍。接下來開始實做區域調整式臨界值方法，在這個步驟中要初始一個臨界值的值，這個臨界值用兩個參數來表示，也就是平均值（ $M_l$ ）和標準差（ $S_l$ ），然後再從直方圖中峰（peak）的灰階值（ $P$ ）來得到初始值，公式如下：



$$M_l = P - \alpha * S_0 \quad (9)$$

$$S_l = 0 \quad (10)$$

$$T_l = M_l + S_l \quad (11)$$

其中 $S_0$ 是整張影像的標準差。經過初始動作後，再拿這個臨界值來做判斷，當在分析目標影像上的 pixel 時，灰階值小於臨界值則判斷為一個細胞的一部份，然後以這個 pixel 為基準，觀察它的八個鄰居，如圖 2-9，如果這些 pixel 有小於臨界值的就一樣會被標定為同一個細胞，而臨界值在此時則會被修正，修正公式如下：

$$M_u = \frac{M_l * N + P_{i,j}}{N + 1} \quad (12)$$

$$S_u = \sqrt{\left| \frac{(S_l^2 + M_l^2) * N + P_{i,j}^2}{N + 1} - M_u^2 \right|} \quad (13)$$

$$T_u = M_u + S_u \quad (14)$$

在臨界值調整完成後，繼續依照前述方法觀察週遭的鄰居和臨界值的關係，直到沒有

pixel 可以加進這個群體為止，當終止後，臨界值會被重設回初始狀態，再繼續尋找下一個細胞群體。

當掃描完整張目標影像後會有許多細胞群體被標定出來，接下來利用數學型態學來做一些修正後，觀察單一細胞群體，如果此細胞群體的大小在可以接受的大小範圍內則細胞計數數會加一，若大於可以接受範圍則利用以下公式來計算細胞數目的增加數

$$N = \frac{A_c}{A_s} \quad (15)$$

其中  $A_c$  是目前細胞群體的大小， $A_s$  則是標準的細胞大小，而  $N$  是細胞數目增加數。圖 2-7 為目標影像，圖 2-8 為目標影像的直方圖

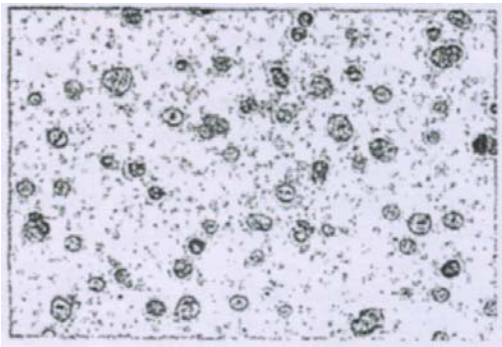


圖 2-7 目標影像

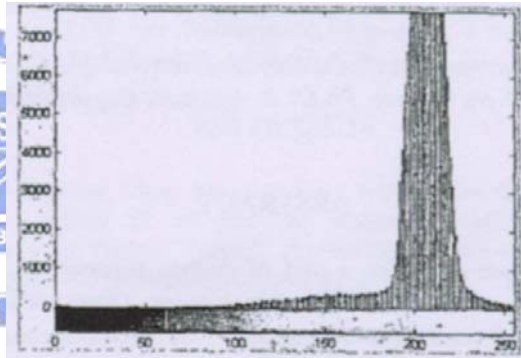


圖 2-8 目標影像直方圖

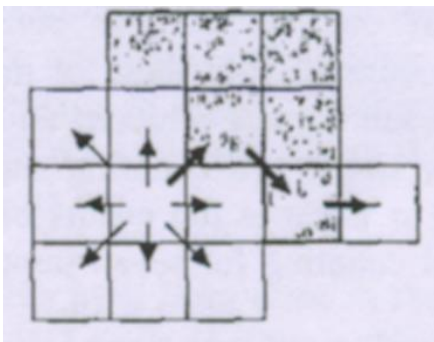


圖 2-9 pixel 的八的鄰居

## 2.5 結論

顆粒擷取技術所要處理的影像有許多種不同的情況，而且所要擷取的顆粒也以不同的形式展現，所以並沒有一個適用的方法（general solution），大部分的方法都是以特定影像和特定顆粒的問題來發展一個較好的方法（optimal solution），除了前幾節所提到的方法在[7]這篇中也提出以 Hough Transforms 的方法來解決這個問題，我們所提出來的方

法也是針對生物方面具已知大小型狀的三維顆粒為基礎所發展出來的方法。



## 第三章

# 3D 細胞影像分析記數方法

我們設計的細胞計數方法會以兩個已經被應用廣泛的技術為基礎：Matched Filter 和 k-means cluster，再發展出我們提出的方法，所以在第三章的前兩節我們會先將這兩個技術作一個概略性的介紹，接下來我們再提出我們設計的方法的概念和流程。

### 3.1 Matched Filter

影像前處理的目的，不外乎是減弱影像中不希望見到的雜訊，以及增強我們想做分析處理的影像特徵。在雜訊的消除部分，對於不同來源的影像往往有不同的做法，低通(low pass)、高通(high pass)、中值(median)、Sobel...等濾波器都是我們常用來消除雜訊的工具。

在加強特徵方面，Matched Filter 是一個常常被用到的濾波器，在[8]中有提到如何設計一個 Matched Filter 的 mask 和為什麼要這樣設計，接下來我們稍微介紹一下 Matched Filter，首先如果我們要在一張影像中找出它和另一張影像相似的地方，可以利用以下的計算方式來判斷相似度

$$\max_A |f - g|, \iint_A |f - g| \text{ 和 } \iint_A (f - g)^2$$

假如我們使用  $\iint_A (f - g)^2$  當做判斷的依據，則此式子可以分解為

$$\iint_A (f - g)^2 = \iint_A f^2 + \iint_A g^2 - 2\iint_A fg \quad (16)$$

現在假設  $\iint_A f^2$  和  $\iint_A g^2$  都已經是固定的值時，那我們的判斷依據會簡化為只剩下  $\iint_A fg$



這部份，所以當  $\iint_A fg$  越小我們的測量值會越大，也就是說越不可能是相似的地方。

接下來我們再利用 Cauchy-Schwarz 不等式來處理  $\iint_A fg$ ，由不等式我們可以將式子轉成  $\iint_A fg \leq \sqrt{\iint_A f^2 \iint_A g^2}$ ，且等號成立的條件為  $g = cf$ ， $c$  為一個常數。相同的，在離散的狀態下為  $\sum_i \sum_j f(i, j)g(i, j) \leq \sqrt{\sum_i \sum_j f(i, j)^2 \sum_i \sum_j g(i, j)^2}$ 。

現在我們將上述概念用到實際的應用上，要在一張影像  $g$  上找到一個樣本  $f$  發生的地方，可以將這個樣本在影像上的所有可能位置上移動，然後計算  $\iint_A fg$  來判斷，那麼上述的式子會變成

$$\iint_A f(x, y)g(x+u, y+v)dxdy \leq \sqrt{\iint_A f^2(x, y)dxdy \iint_A g^2(x+u, y+v)dxdy} \quad (17)$$

而且因為  $f$  在範圍  $A$  外的部份為零，所以左半邊的式子會等於

$$\iint_{-\infty}^{\infty} f(x, y)g(x+u, y+v)dxdy \quad (18)$$

這個式子也就是將  $g$  和  $f$  做 cross correlation，這是一個 linear shift-invariant 的操作，跟將  $f$  旋轉  $180^\circ$  後跟  $g$  做 convolve 是一樣的，接下來證明在某些假設下最好的 mask 會是  $f$  本身，這也就是被熟知的 matched filter theorem。

令影像為  $g = f + n$ ，其中  $n$  為雜訊，將一個 mask  $h$  作用在影像上，則經過 mask 處理過的影像會變成  $g' = h * g$  而雜訊也會變成  $n' = h * n$ ，那麼我們可以由[8]知道

$S_{n'n'} = S_m |H|^2$ ，所以現在的預期雜訊能量 (expected noise power) 為

$$E\{|h * n|^2\} = \iint S_{n'n'} = \iint S_m |H|^2 \quad (19)$$

現在假設雜訊為“white”，也就是說  $S_m$  是常數，所以預期雜訊能量為  $S_m \iint |H|^2$

另一方面，在點  $(x, y)$  的訊號能量為

$$|h * f|^2 = |IF(HF)|^2 = \left| \iint HF \exp[2\pi j(ux + vy)]dudv \right|^2 \quad (20)$$

由前面的內容我們知道右邊的式子會小於等於  $\iint |H|^2 \iint |F|^2$ ，因此在“white”雜訊的假設

下訊號能量和預期雜訊能量比為

$$\frac{|h * f|^2}{E\{h * n\}^2} \leq \frac{\iint |H|^2 \iint |F|^2}{S_m \iint |H|^2} = \frac{1}{S_m} \iint |F|^2 \quad (21)$$

而且等號成立在  $H = F * \exp[-2\pi j(ux + vy)]$  時，也就是  $h(\alpha, \beta) = f(x - \alpha, y - \beta)$ ，所以可以證明要使訊號和雜訊的比最高要使用要找的樣本本身。

在圖 3-1 到圖 3-3 中我們可以看到 matched filter 的範例和結果

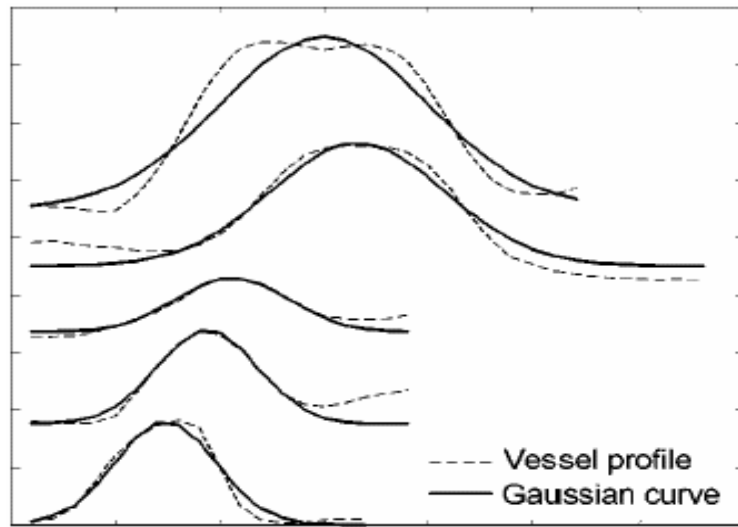


圖 3-1 1D matched filter 範例

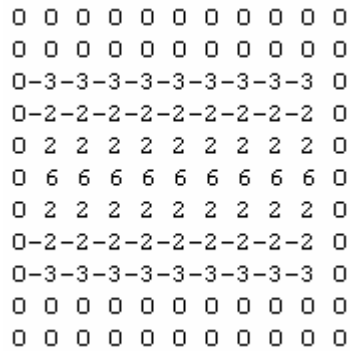


圖 3-2 2D 中偵測血管的 Matched Filter 例子

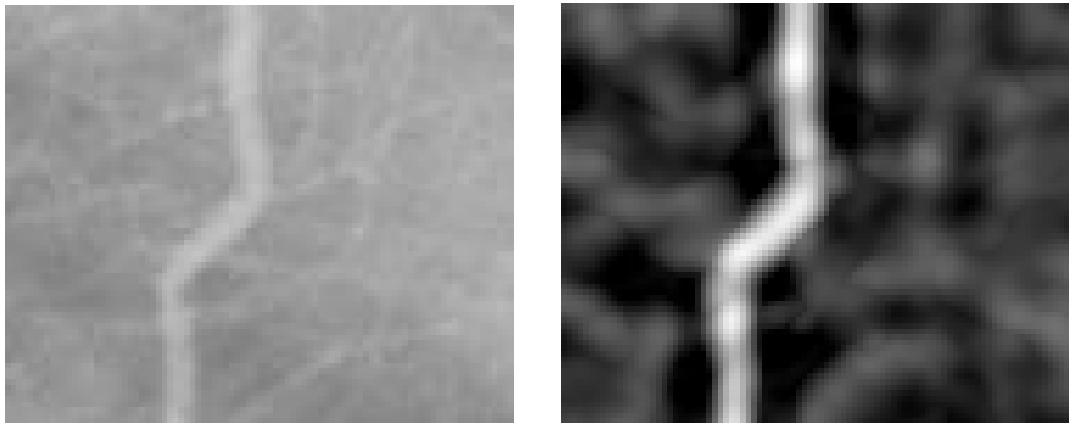


圖 3-3 matched filter 實際影像效果

(左)原始影像

(右)經過 Matched Filter 處理過的結果

## 3.2 K-mean Clustering

分群法 (Clustering) 在本質上和分類法 (Classifier) 有相同的行為，不同的地方在於分群法不需要人工給予的訓練資料，也就是說它是屬於非監督式 (unsupervised) 的分割法。為了補足人工訓練資料的資訊，分群法會在分割影像的過程中自動對現有類別的資訊做更新的動作。

常用的分群演算法有平常我們熟知的 k-means、fuzzy c-means 和 expectation-maximization，以 k-means 為例，假設預期的分群類別數目為  $k$ ，則 K-means algorithm 其演算過程可概述如下：

1. 隨機選取  $k$  個類別的類別中心。
2. 對每一個資料點  $x$ ，尋找與之最接近的類別中心，並將  $x$  加入該類別，隨即計算新的類別中心（該類別中原有的資料點加上  $x$  後的平均向量）。
3. 檢查每一個資料點目前與之最接近的類別中心是否和他的類別分配一致，如果不是，則回到步驟 2。

圖 3-4 也就是分群法的一個簡單的例子



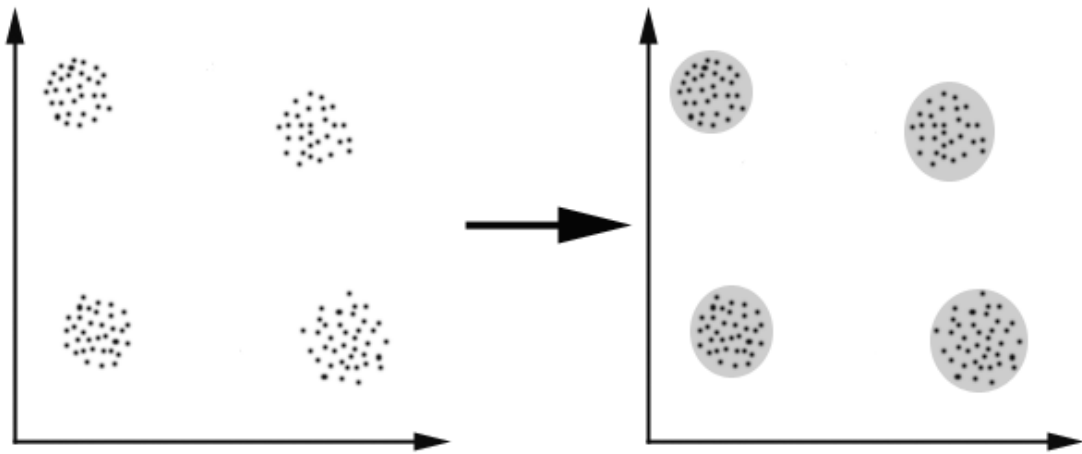


圖 3-4 分群法的例子

### 3.3 自動計數分法

我們實做的自動計數分法分為下面的幾個步驟：

- 前處理 (Matched Filter)
- 尋找 Local Max
- Group 分析
- 增加 Candidate

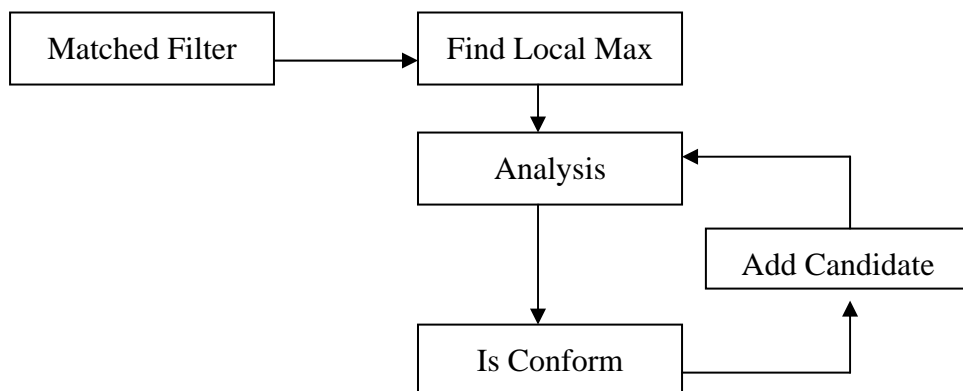


圖 3-5 程式流程圖

#### 3.3.1 前處理 (Matched Filter)

在我們設計的方法中，前處理是採用 3.1 節所提到的 Matched Filter，所以我們要設計在 Matched Filter 中扮演重要角色的 mask，目前我們希望在 3D 的影像資料中找到大小相近的細胞顆粒，所以我們在影像中找出五個較完整且單獨的細胞的中心位置，利用這個五個座標位置當成參數，再加上細胞的大小、xy 和 z 的解析度的比，程式會自動利用這五個細胞為範本產生一個平均的細胞灰階分佈的 mask，再利用這個 mask 對原始影像做處理，結束後即進入下一步驟。

### 3.3.2 尋找 Local Max

在 Matched Filter 的理論中，我們知道當影像中有和 mask 越類似的顆粒時經處理後的 response 會越強，所以在這一步驟中我們在影像中搜尋一個 3\*3\*3 的區塊中的 Local Max，再將這些 Local Max 經過連結性的測試 (Connected Component)，之後剩下的 Local Max 我們則預測為可能的細胞中心。



### 3.3.3 Group 分析

當我們找出一些可能的 Candidate 時，我們回到原始影像上設一個 Threshold 將可能的雜訊和部分低 Gray Value 的資訊去除，之後我們利用 3.2 節所提到的 K-means 分群法的概念將所有經 Threshold 後留下來的 pixel 歸類到離它最近的 Candidate 的 Group 中，經過分群後我們可以計算每個 Group 的空間平均值 (Mean) 和 Gray Value 的平均值 (Mean) 和變異數 (Variance)，並利用 CGAL 所提供的 Function 來計算每個 Group 的最小包含球 (Smallest Enclosing Sphere)，Smallest Enclosing Sphere 的概念是在空間中有一群散佈的點，我們要找出一個體積最小的球體將這些散佈的點全部包含在球體內，在我們的方法中利用 Smallest Enclosing Sphere 所計算出來的球半徑來判斷這個 Group 是否為符合我們期待的 cell 顆粒，

當所有 Group 的資訊都已計算完成後我們進入 IsConform 的 Function 中去判斷是否所有 Group 都已經符合我們所設定的條件，目前我們所設定的條件依照要找尋的 cell 顆粒的大小來給予，假設現在細胞大小為半徑 6 pixels（針對 xy 平面的解析度而言）的球體，則 Group 可能會有以下三種情況：

1. Group 的 Smallest Enclosing Sphere 半徑大於 8，且周遭沒有在適合距離內的鄰居，則我們判定這個 Group 為獨立存在。
2. Group 的 Smallest Enclosing Sphere 半徑大於 8，且有適合距離內的鄰居，這個情況又分為兩種情形，一種為鄰居中有半徑小於 4 的 Group，另一種則是鄰居半徑都大於 4。
3. Group 的 Smallest Enclosing Sphere 半徑介於 4 到 8 之間。

這三種情況在 Add Candidate 這個步驟中將會有相對應的處理方式。

#### 3.3.4 增加 Candidate

在上述的分析中，當程式的分析偵測到 Group 有 case 1 和 case 2 的狀況發生時程式就必須進入這個處理程序，這個處理程序針對不同的情況利用不同的分法來處理，以下為相對的處理方式：

1. 當這個情況發生時，我們將 Group 中的 note 屬性設定為 1，代表此 Group 為完整且單獨存在，在程式往後的處理中不會考慮到這個 Group。
2. 當這個情況發生時，我們必須判斷此 Group (A) 和鄰居的關係，首先我們利用 Delaunay triangulation 的概念再加上限制距離的條件來找尋週遭的鄰居，之後分析鄰居的 smallest enclosing sphere 看是否有半徑小於 4 的 Group (B)，若有則我們將 Group A 的中心在兩個 Group 的連心線上往遠離 Group B 的方向移動，原因是我們假設這種狀況會發生在 Group A 的初始中心點太偏向 Group B 因此造成執行分群法時部份原屬於 Group B 的 pixel 被歸類到 Group A，所以 Group A 的

半徑會大於正常值而 Group B 的半徑則小於正常值。另一種狀況則是週遭的鄰居半徑都大於 4，也就是說這些 Group 都是正常或是太大，那麼我們則是找出週遭半徑最大的 Group(B)，在 Group A 和 Group B 的連心線中心增加一個 Candidate，這樣處理的原因在於我們假設這種情況下 Group A 和其鄰居共同瓜分了他細胞的 pixel，所以我們希望可以在當中加入一個新的點使被瓜分掉的細胞可以重新被計算。

### 3.4 討論

在下一章我們將利用我們提出的方法來分析果蠅腦細胞的 3D 影像，在實驗過程中我們發現在 IsConform 和 AddCandidate 這兩個步驟會發生所需時間過長和增加太多細胞中心的問題，所以在下一章的實驗中我們先將這兩個步驟從實驗中移除，並討論實驗的結果。



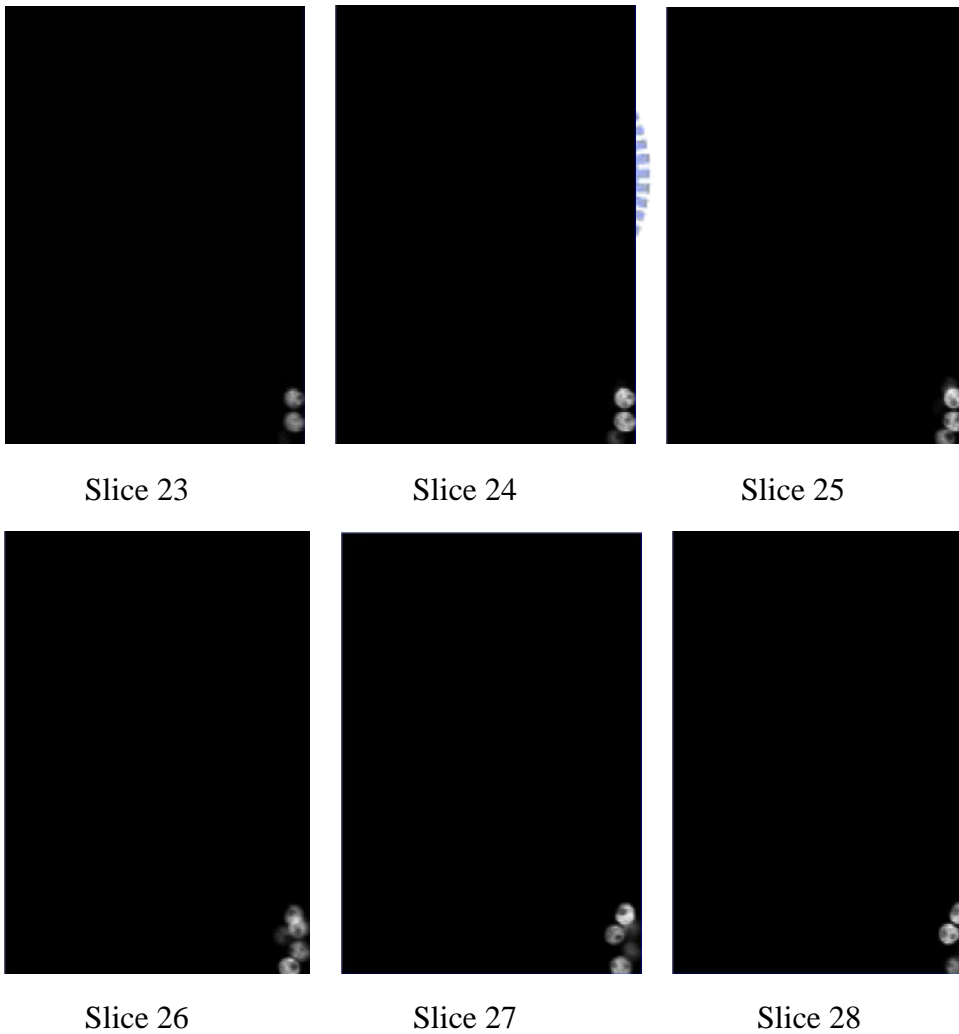
# 第四章

## 實驗結果及討論

在本章中我們會利用我們所提出來的的方法來針對幾組果蠅的腦細胞影像做實驗，並將實驗過程中的中間結果和問題加以討論。

實驗資料組一：

圖 4-1 為原始影像部分切片的影像，在這組測試影像中其取像的規格為 8 bits 的灰階範圍 (Gray value 0~255)，x-y 方向取像的距離為 0.3，z 方向取像距離為 0.5，影像大小為 140x205x62。



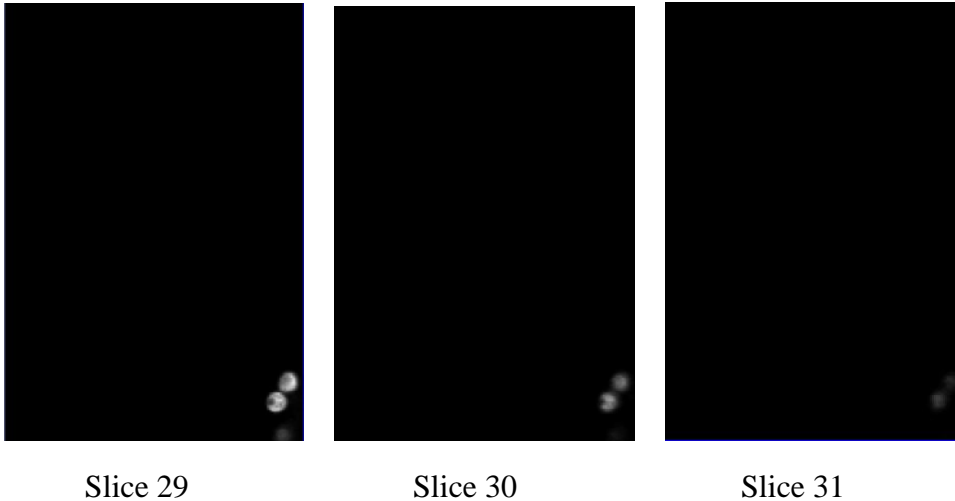


圖 4-1 : 實驗資料組一部分切片影像

依據影像特性和我們欲擷取的細胞大小(直徑為 13 x-y pixels 寬的球形顆粒),我們設計一個大小為 13x13x9 的樣版當成 Matched Filter 所使用的 mask, 實驗過程中第一次我們將 mask 的灰階分佈視為高斯分佈 (Gaussian distribution), 實驗結果如圖 4-2 所示, 我們會發現影像雖然有向細胞中心聚集的現象, 但沒有明顯的特徵訊號, 原因在於 mask 的灰階分佈和細胞的灰階分佈並不一致, 所以經過 Matched Filter 處理過後不會有明顯的特徵訊號。

經過第一次的實驗, 我們在第二次設計 mask 時直接從影像上選取五顆細胞的灰階分佈當成 mask 的灰階分佈, 實驗結果如圖 4-3 所示, 我們發現此次的結果較第一次為佳, 可以很明顯的發現有明顯的特徵訊號。



圖 4-2 :(左)原始影像 (中)利用 Gaussian distribution 的 mask 處理後的結果 (右)利用五顆細胞所設計出來的 mask 處理後的結果

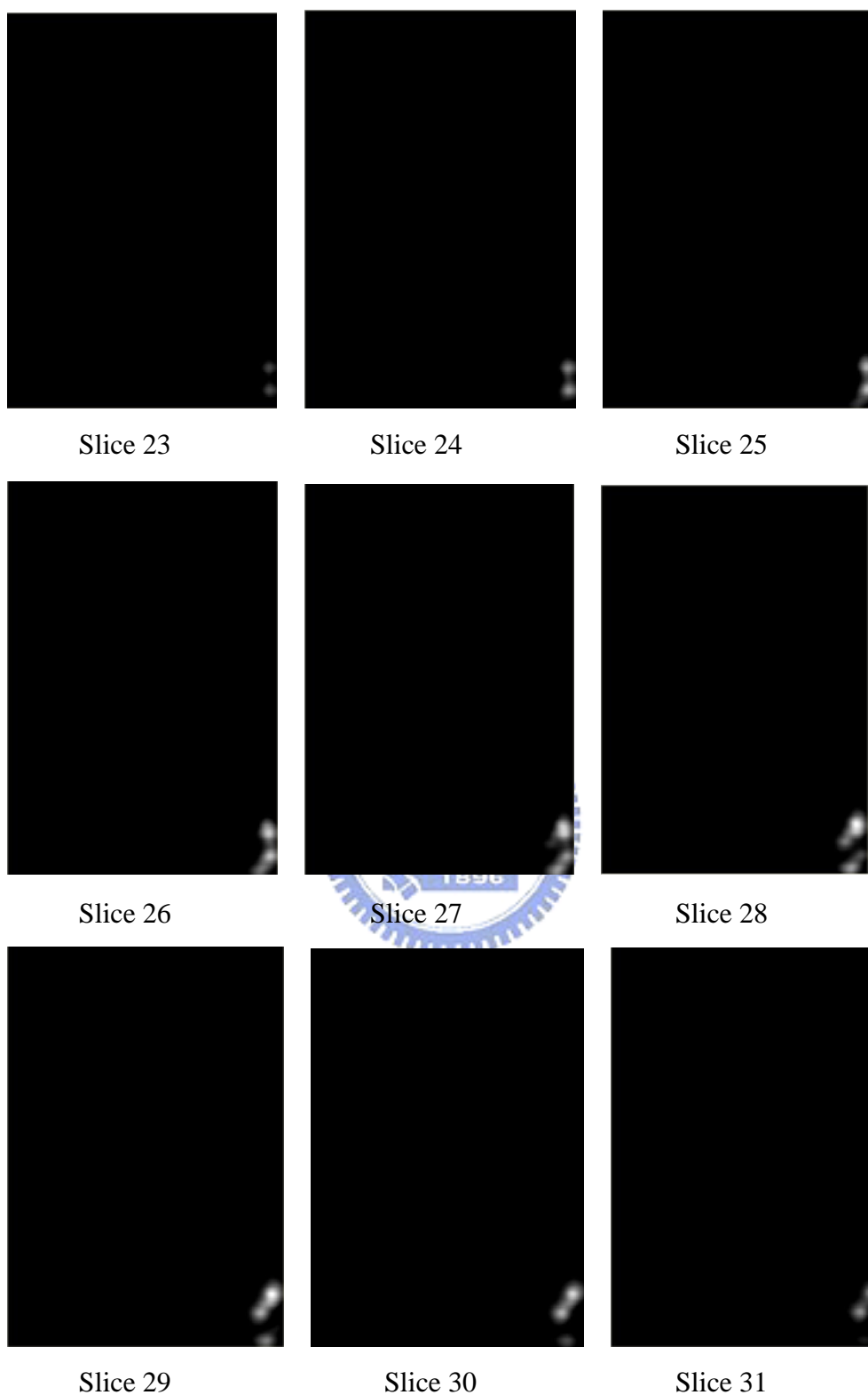


圖 4-3 : 圖 4-1 經過 matched filter 之後的結果

緊接著我們利用這些特徵訊號做了 clustering 的分析，結果如圖 4-4 所示，系統此時提供一個觀察的機制，使用者可以在影像上隨意點選特定的位置，系統會計算離此位置最近的 Group，並利用下面的方式呈現出來，圖中綠色圓圈為以找到的最近細胞 Group

中心為球心，直徑大小為 13 pixels 的球，用來表示粗略的細胞，藍色部分則為 pixel 灰階值大於一個臨界值且在分群過程中分屬於此 Group 的 pixel，紅色的點為我們找出來的細胞中心，另一方面，系統亦提供了著色的機制，系統可以自動將分屬不同 Group 的 pixel 加上特定的顏色來做全域的觀察，但由於著色色彩目前只使用了 13 種顏色，所以有時會發生臨接的兩個不同 Group 卻有相同的顏色，此時可以利用上述選取特定細胞的機制來做進一步的觀察，其結果如圖 4-5 所示

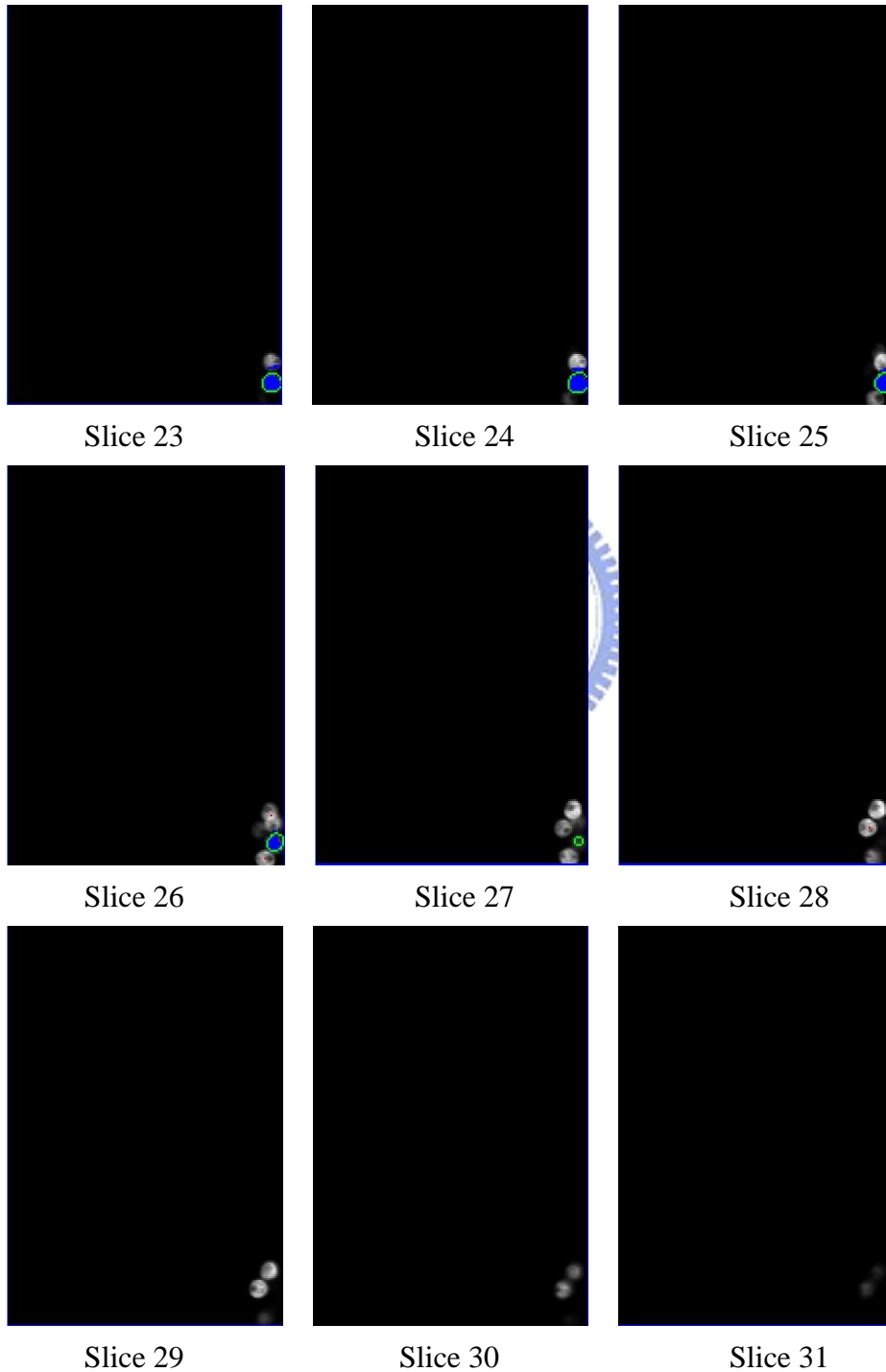


圖 4-4：圖 4-1 經過程式計算後的結果，並利用選取機制的結果



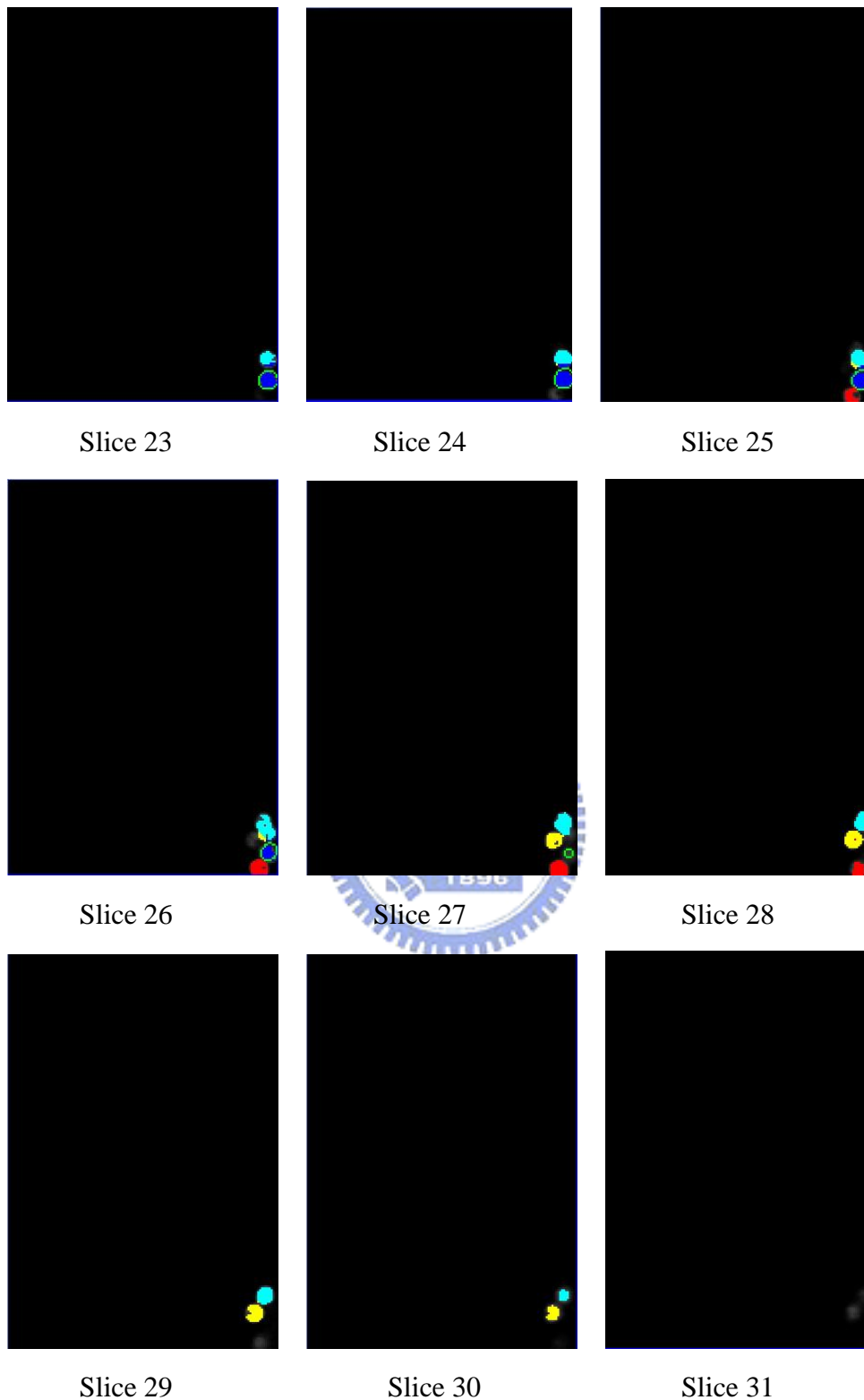


圖 4-5 : 圖 4-1 經過程式計算後的結果，並利用選取機制和著色機制的結果

實驗資料組二:

圖 4-6 為原始影像部分切片的影像，在這組測試影像中其取像的規格為 8 bits 的灰階範圍 (Gray value 0~255)，x-y 方向取像的距離為 0.3，z 方向取像距離為 0.5，影像大

小為 62x53x62。

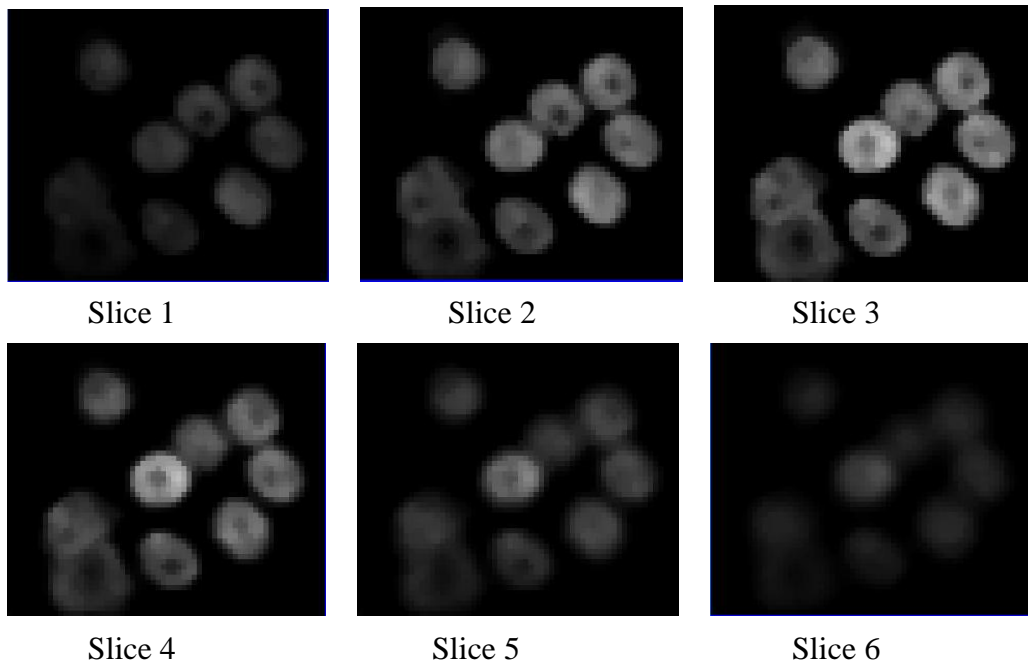


圖 4-6：實驗資料組二部分切片影像

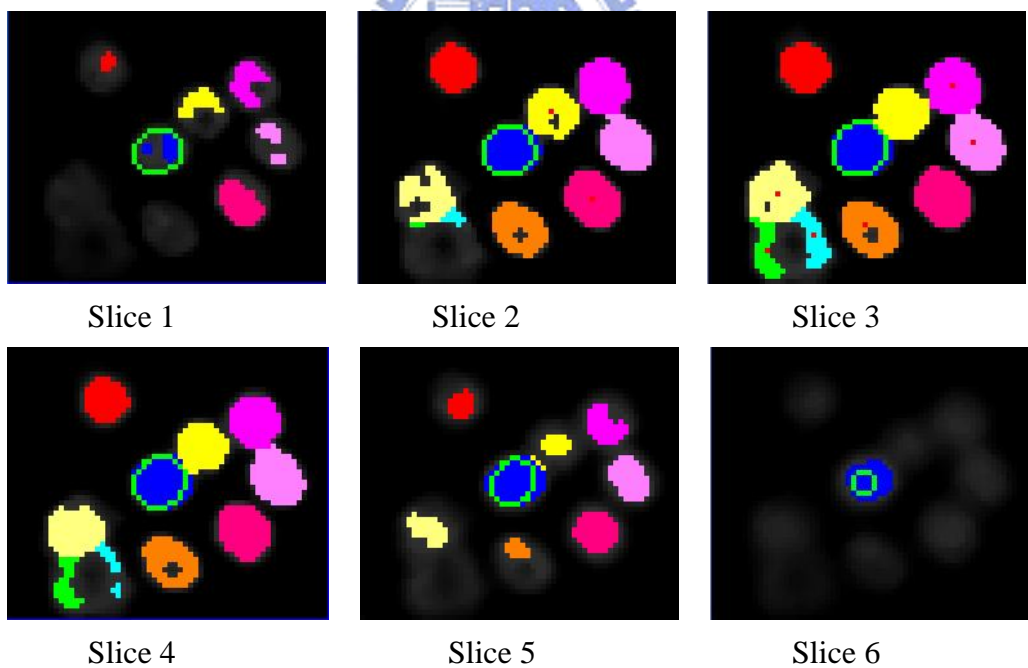


圖 4-7：圖 4-6 經過程式計算後的結果，並利用選取機制和著色機制的結果

實驗資料組三：

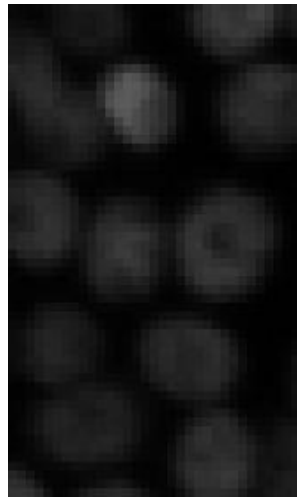
圖 4-8 為原始影像部分切片的影像，在這組測試影像中其取像的規格為 12 bits 的灰階範圍 (Gray value 0~4095)，x-y 方向取像的距離為 0.3，z 方向取像距離為 0.5，影像大小為 33x56x120。



Slice 4



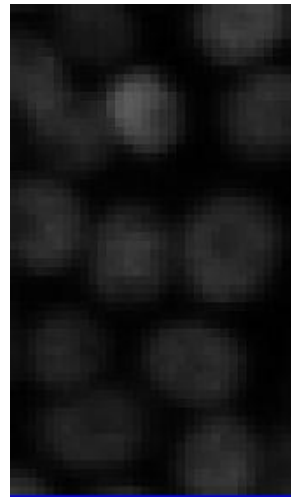
Slice 5



Slice 6



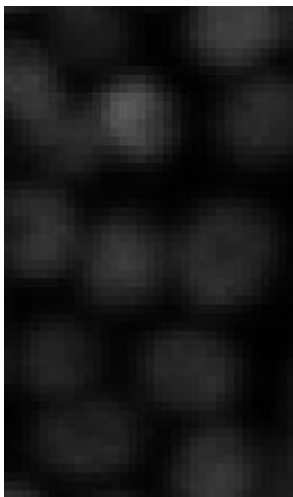
Slice 7



Slice 8



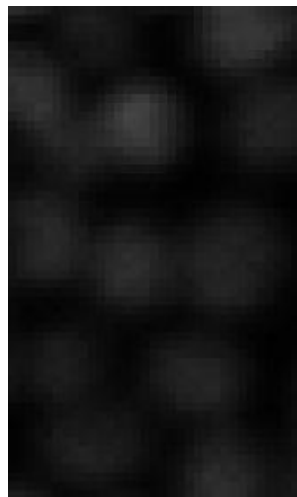
Slice 9



Slice 10

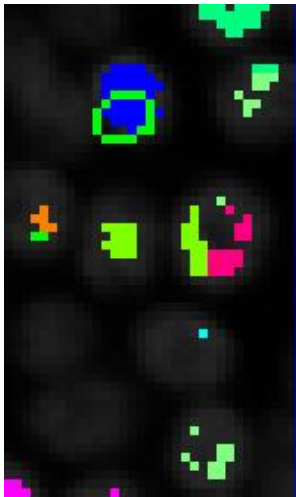


Slice 11

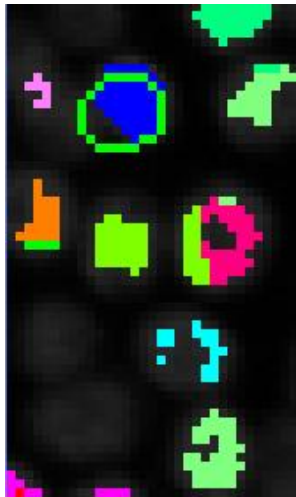


Slice 12

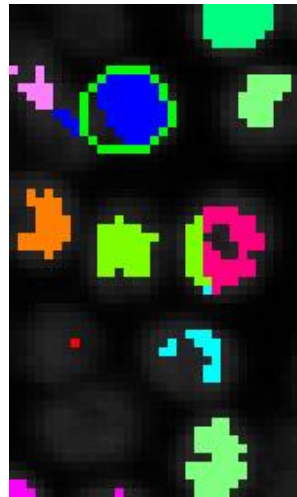
圖 4-8 : 實驗資料組三部分切片影像



Slice 4



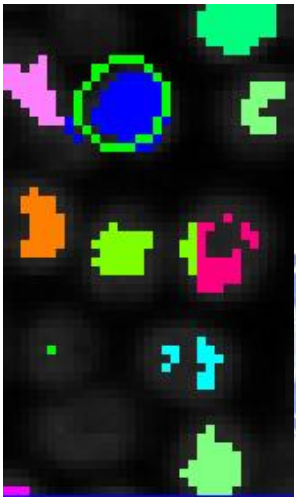
Slice 5



Slice 6



Slice 7



Slice 8



Slice 9



Slice 10



Slice 11



Slice 12

圖 4-9 : 圖 4-8 經過程式計算後的結果，並利用選取機制和著色機制的結果

## 實驗數據

### 測試環境

CPU : Pentium 4 3.0G

Ram : 3.0G

表 4-1

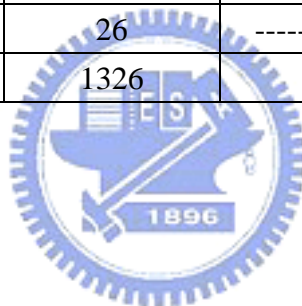
測試檔案	人工計算	程式計算(1)	程式計算(2)	生物其他方法	正確性
O1_1	5	4	5	-----	80%
O1_2	9	10	10	-----	90%
O1_3	88	91	110	-----	96.7%
O1	-----	1392	1782	1500	92.8%

時間：2 小時 47 分 49 秒

表 4-2

測試檔案	人工計算	程式計算(1)	生物其他方法	正確性
O2_3	15	14	-----	93.33%
O2_4	19	26	-----	73.1%
O2	-----	1326	1500	88.4%

時間：4 小時 54 分



## 第五章

# 結論及未來工作展望

本論文在介紹目前所發表過的一些顆粒擷取技術的方法後，我們提出以 Matched Filter 為基礎的方法，並應用在 3D 的生物影像上，雖然實驗結果尚稱良好但我們依舊在發展方法時和後續實驗中發現一些有待改進的地方。

以下我們列出有待改進和未來期待能達成的目標：

1. 在 matched filter 部分，以 matched filter 的理論基礎來看，所使用的樣本模型會對結果造成很大的影響，所以如何決定一個適當的樣本模型是我們第一個大考驗，目前我們使用的方式是讓生物專家從影像中挑選五個作為基準的細胞，利用這五個細胞做一個平均值來當成我們的樣本模型，我們希望這部分在未來的發展中能達到自動化，而不需經過人工訓練的步驟，另一方面，我們在實驗時發現當影像的灰階值分佈範圍變大時，理論上會有較多細節被表現出來，所以實驗結果應該較佳，可是我們發現當越多細節被表現出來時，我們在設計樣本模型時就得更加嚴謹，否則會如同實驗三的情況造成許多偏差。
2. 同樣的在 matched filter 這部份我們發現所費時間過長，效能問題也成為我們有待改進的地方。
3. 最後，在分析部份，我們發現目前所用的規則太過寬鬆，用於處理灰階值分佈楚於平坦的區域會造成計算過多的現象，未來應該要引進幾何形狀方面的規則或其他限制來改善計算過多的現象。

## 參考文獻

- 【1】 KAN JIANG , QING-MIN LIAO , SHENG-YANG DAI , “A Novel White Blood Cell Segmentation Scheme using Scale-Space Filtering and Watershed Clustering”, Processings of the Second International Conference on Machine Learning and Cybernetics , Xi'an , 2-5 November 2003.
- 【2】 Cecilia Di Ruberto , Andrew Dempster , Shahid Khan , Bill Jarra , “Segmentation of Blood Images Using Morphological Operators” , Pattern Recognition, 2000. Proceedings. 15th International Conference on Volume 3, 3-7 Sept. 2000 Page(s):397 - 400 vol.3
- 【3】 Leticia V. Guimarães , Altamiro A. Suzim and Junji Maeda , “A New Automatic Circular Decomposition Algorithm Applied to Blood Cells Image” , Bio-Informatics and Biomedical Engineering, 2000. Proceedings. IEEE International Symposium on 8-10 Nov. 2000 Page(s):277 - 280
- 【4】 Zeyun Yu and Chandrajit Bajaj , “A Clustering-Based Method for Particle Detection in Electron Micrographs.”
- 【5】 Zeyun Yu and Chandrajit Bajaj , “Detecting circular and rectangular particles based on geometric feature detection in electron micrographs.” , Journal of Structural Biology 145(2004) 168-180
- 【6】 Hazem Refai , Lun Li , T. Kent Teague , R. Naukam , “Automatic Count of Hepatocytes in Microscopic Images.” , Image Processing, 2003. Proceedings. 2003 International Conference on Volume 2, 14-17 Sept. 2003 Page(s):II - 1101-4 vol.3

【7】 Yuanxin Zhu , Bridget Carragher , Fabrice Mouche , and Clinton S. Potter , “Automatic Particle Detection Through Efficient Hough Transforms.” , Medical Imaging, IEEE Transactions on Volume 22, Issue 9, Sept. 2003 Page(s):1053 – 1062

【8】 Azriel Rosenfeld and Avinash C. Kak , Digital Picture Processing , Second Edition , 1982 , By Academic Press , Inc.





# 附錄 (Appendix)

以下附錄是為了方便使用者操作此系統所增列的說明：

圖 A-1 為系統介面圖，系統主要輸入檔案為 Original File、Matched File 和 Connected File 三種，輸入位置分別由 A、B、C 所標示

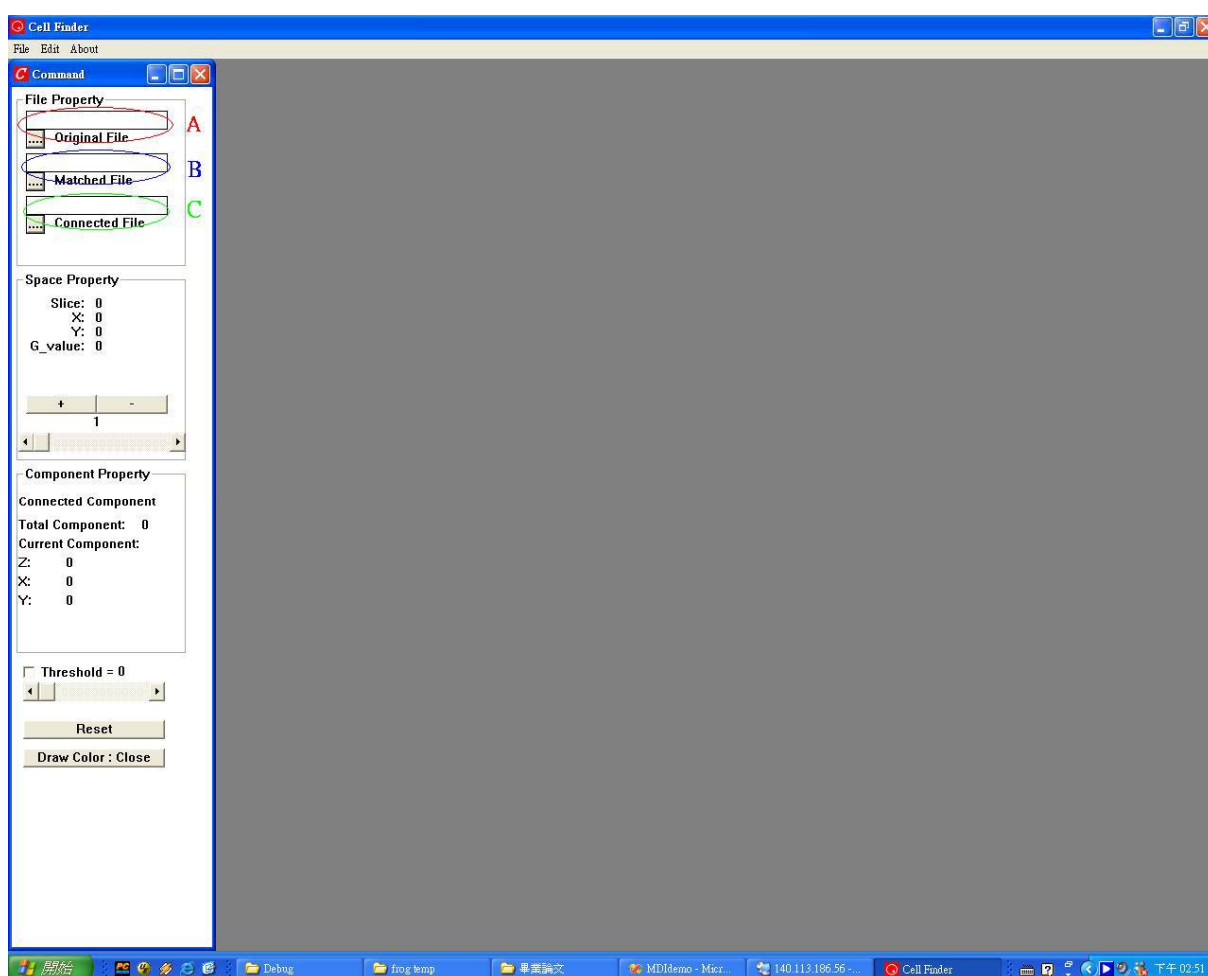


圖 A-1：系統介面圖

A：Original File

此處所讀入的檔案為我們想觀察的細胞影像的原使檔案，輸入時先擊點 A 處的按

鈕，之後便會出現如圖 A-2 的視窗，請選擇以 IPT 為副檔名的原始影像檔。

#### B : Matched File

此處讀入的檔案為 A 處檔案經過 Matched Filter 處理後的檔案，請務必先選擇 A 處所欲載入的檔案並且在 B 處選擇相對應的檔案，檔案載入方式如 A，副檔名亦為 IPT。

#### C : Connected File

此處讀入的檔案為 A 處檔案經過系統處理後的檔案，原始影像經過系統處理後可選擇以圖 A-3 的格式以副檔名為 TXT 的型式儲存起來，在下次可以直接讀入 Original File 和 Connected File 而不需再經過 Matched Filter 的處理。

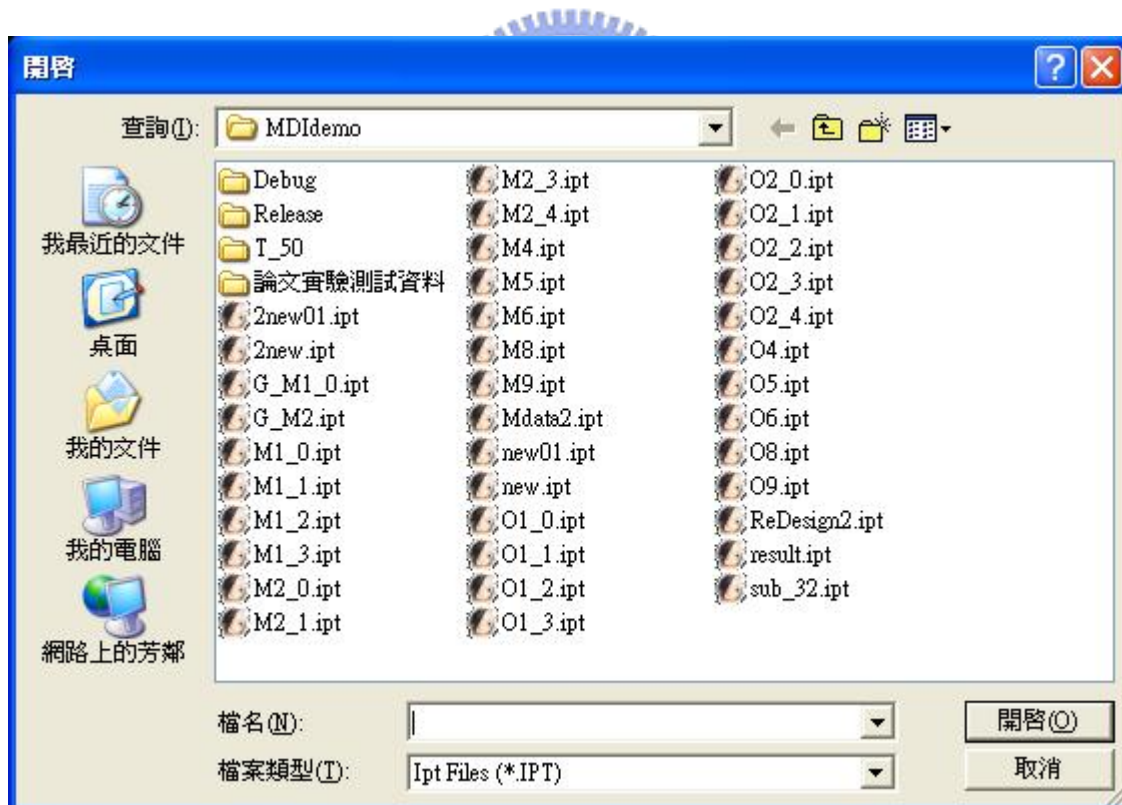


圖 A-2：輸入視窗（Original File 和 Matched File 請讀入 IPT 檔，Connected File 請讀入 TXT 檔）

複製 - C1.txt - 記事本							
檔案(F) 編輯(E) 格式(O) 檢視(V) 說明(H)							
<b>Local_Max</b>							
3	130	201	26				
5	128	186	28				
2	135	193	24				
4	133	179	26				
<b>Group_Map</b>							
-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1

圖 A-3：Connected File 儲存格式，Local\_Max 以下到 Group\_Map 為止為細胞中心位置和相對應所屬的 Group 編號，Group\_Map 以下為影像的 pixel 位置和所屬的 Group 編號

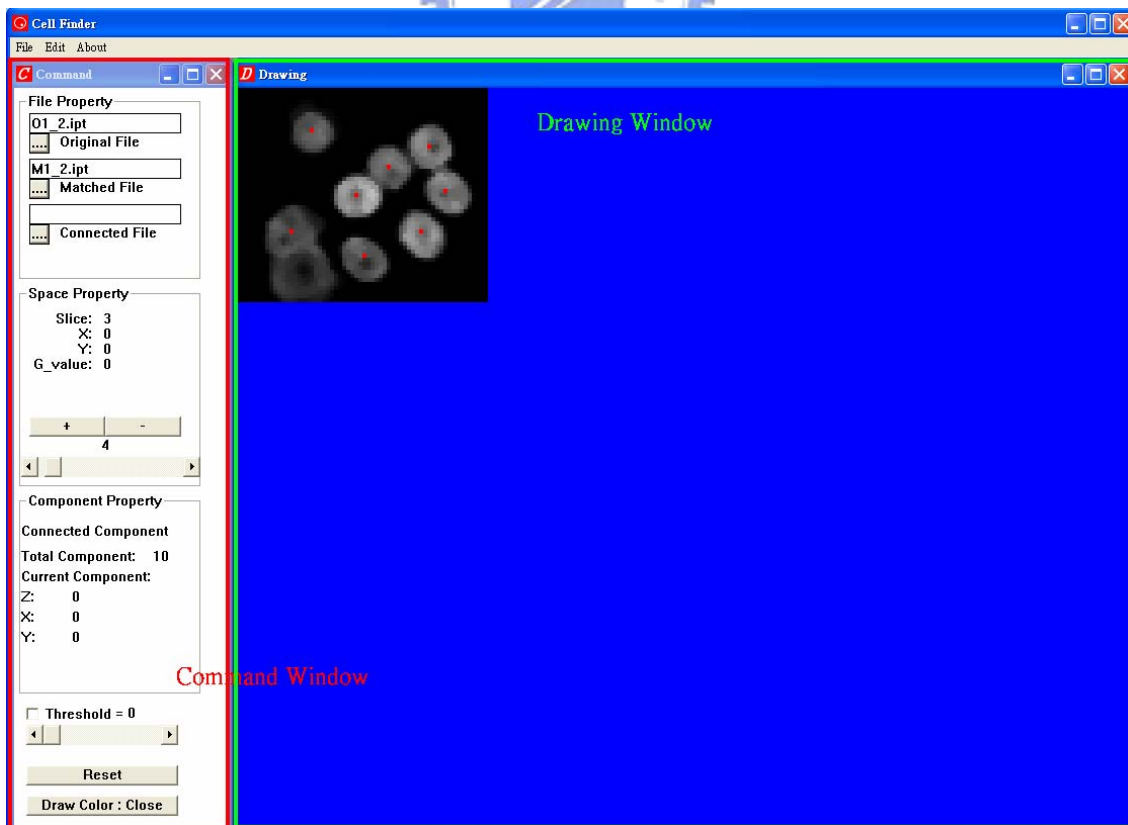
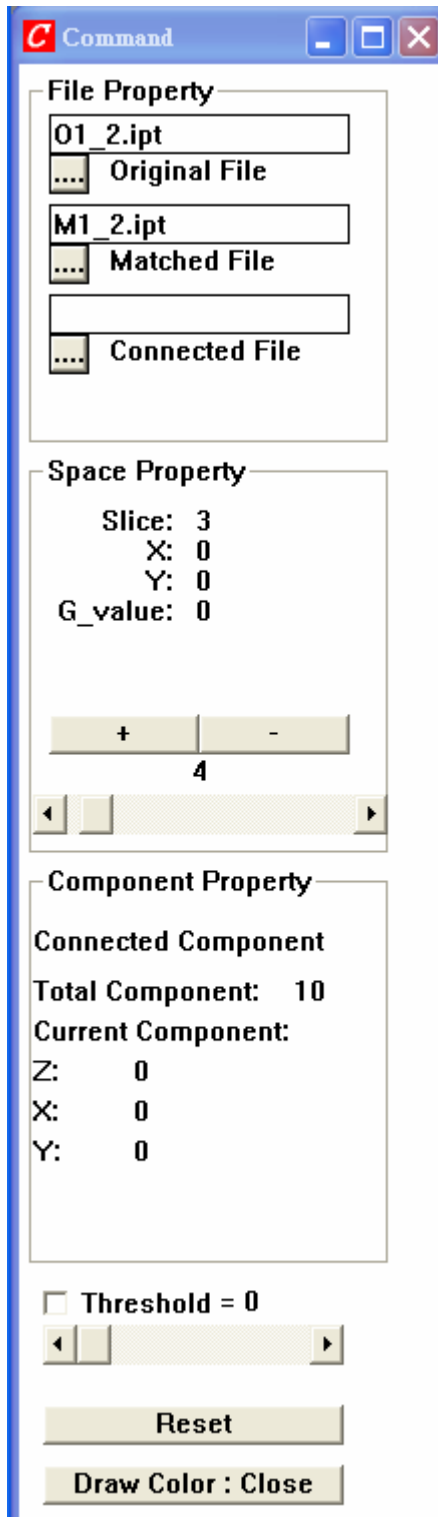


圖 A-4：系統主要的兩個子視窗 Command Window 和 Drawing Window

## Command Window



### File Property

第一個欄位為原始影像檔輸入位置

第二個欄位為經過 Matched Filter 處理後的檔案輸入位置

第三個欄位為結果檔案輸入位置

### Space Property

這個區塊為滑鼠點選位置的資料，Slice、X 和 Y 是目前的空間座標，G\_value 為這一點的灰階值，【+】和【-】兩個按鈕可以縮放顯示於 Drawing Window 的影像的大小，而在按鈕下方的數字則為縮放比例，捲軸為控制顯示的 Slice

### Component Property

這個區塊為計算後的資訊，Total Component 後的數字為系統計算出來的細胞數目，Current Component 下方的 X、Y 和 Z 顯示使用者用滑鼠點選後最近的細胞群中心位置

Draw Color 按鈕可以開啟或關閉著色機制