

國立交通大學

多媒體工程研究所

碩士論文

為投射神經在果蠅葷狀體萼端建立座標系統之研究

Research on constructing coordinate system for projection neuron in

Drosophila calyx

研究生：廖栢毅

指導教授：荊宇泰 教授

中華民國九十七年一月

為 投 射 神 經 在 果 蠅 葷 狀 體 萼 端 建 立 座 標 系 統 之 研 究
Research on constructing coordinate system for projection neuron in
Drosophila calyx

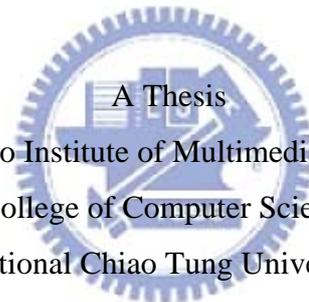
研 究 生：廖栢毅

Student：Bo-Yi Liao

指 導 教 授：荆宇泰

Advisor：Yu-Tai Ching

國 立 交 通 大 學
多 媒 體 工 程 研 究 所
碩 士 論 文



A Thesis
Submitted to Institute of Multimedia Engineering
College of Computer Science
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master
in
Computer Science

January, 2008

Hsinchu, Taiwan, Republic of China

中 華 民 國 九 十 七 年 一 月

為投射神經在果蠅葷狀體萼端建立座標系統之研究

學生：廖栢毅

指導教授：荊宇泰 博士

國立交通大學多媒體工程研究所

摘 要

人類大腦的奧秘一直是各領域學者深感興趣的主題，因為大腦掌管了生命中許多重要的部份。果蠅腦研究更是現在研究大腦的熱門主題，因為果蠅有和人類類似的學習、記憶行為。目前果蠅的嗅覺神經網路前端已經有許多發現，但是更深入果蠅腦內部的區域卻仍然充滿了神秘色彩。我們的實驗資料來自清華大學江安世老師實驗室。利用他們清晰的果蠅腦部資料，我們使用主成分分析做迴歸去分析嗅覺投射神經。然後在果蠅腦的葷狀體萼端定義一個座標系統，用以定位其軸突末梢的空間座標。經由我們的實驗，發現到大多數相同根源的投射神經的軸突末梢分佈有相似之處，希望我們的實驗可以帶來更多研究果蠅腦的想法。

Research on constructing coordinate system for projection neuron in
Drosophila calyx

Student: Bo-Yi Liao

Advisor: Dr. Yu-Tai Ching

Institute of Multimedia Engineering
National Chiao Tung University

ABSTRACT



The secret of the human brain has been a topic that many scholars deeply interested in all the time, because the brain has been an important role in our whole life. Research on Drosophila brain is very popular now, because they have the same behavior, like learning, memory as human beings. Now there are many discoveries about the front olfactory neural network of Drosophila, but the deeper area is still mysterious. Our experimental data supplied by Dr. Chiang AS's laboratory of Tsing Hua University. With the clear Drosophila brain data, we utilize Principal Component Analysis (PCA) to do regression for olfactory projection neuron (PN). Then, we construct a coordinate system for terminal bouton (TB) of PN in calyx of mushroom body (MB). We find out similar distributions for most homologous PNs through our experiments, and hope that it would bring more ideas for research on Drosophila brain.

誌 謝

首先要感謝我的指導教授荊宇泰老師，兩年多來因為老師耐心的傳授和指導，我才能順利完成這份論文。同時也要感謝口試委員江安世教授、謝昌煥博士和林志陽學長指正論文的謬誤處，並且提供許多專業上的建議及鼓勵，使此論文更加正確完整。

感謝 MIP-Lab 的大家給予我許多知識上和精神上的支持。謝謝我週遭的好朋友們 Jo, Verio, Wanchen, et al. 最後特別感謝我的家人在這段時間對我的包容、支持與鼓勵，讓我可以安心的完成碩士學業。



目 錄

中文摘要	i
英文摘要	ii
誌 謝	iii
目 錄	iv
圖目錄	vi
表目錄	viii
一、	緒論	1
1.1	研究動機	1
1.2	研究目標	2
1.3	論文架構	3
二、	實驗想法及文獻回顧	4
2.1	前言	4
2.2	Least squares regression	7
2.3	More regression methods	8
三、	實驗流程	9
3.1	實驗簡介	9
3.2	取出 PN 的初始資料	10
3.3	正交迴歸分析	12
3.3.1	概論	12
3.3.2	Total least squares	12
3.3.3	Principal components analysis	13
3.4	取出 TBs 及座標轉換	16
3.5	改進初始資料	20
3.6	實驗流程圖	21

四、	實驗結果	22
4.1	實驗資料	22
4.2	實驗成果	25
五、	未來展望	43
5.1	自動化	43
5.2	延伸研究	43
參考文獻	45



圖目錄

圖【1-1】	果蠅嗅覺神經網路	3
圖【2-1】	管道間圖一	4
圖【2-2】	管道間圖二	5
圖【2-3】	管道間圖三	5
圖【2-4】	PN 分佈圖	6
圖【2-5】	least squares method residual	7
圖【3-1】	取 PN 初始位置圖	10
圖【3-2】	GVF Snake on PN	11
圖【3-3】	residual 取法比較圖	13
圖【3-4】	分析出 PN 的直線與平面	14
圖【3-5】	新原點取法與位置	15
圖【3-6】	球形 TB	17
圖【3-7】	長條不規則形 TB	17
圖【3-8】	座標轉換圖	18
圖【3-9】	選取 TB 後並顯示	20
圖【3-10】	實驗流程圖	21
圖【4-1】	D-L-188&221	26
圖【4-2】	DA1-L-85&114	26
圖【4-3】	DC1-L-24&82	27
圖【4-4】	DC2-L-122&176	27
圖【4-5】	DL1-L-44&49	28
圖【4-6】	DL3-L-115&139	28
圖【4-7】	DM1-L-283&284	29
圖【4-8】	DM2-L-40&46	29
圖【4-9】	VA1d-L-132&134	30
圖【4-10】	VA2-L-20	30

圖【4-11】	VA3-L-172&203	31
圖【4-12】	VA4-L-58	31
圖【4-13】	VA5-L-255&268	32
圖【4-14】	VA7I-L-275&278	32
圖【4-15】	VC1-L-69	33
圖【4-16】	VC2-L-76&106	33
圖【4-17】	VM4-L-12	34
圖【4-18】	D 群組圖一	35
圖【4-19】	D 群組圖二	36
圖【4-20】	D 群組圖三	36
圖【4-21】	V 群組圖一	37
圖【4-22】	V 群組圖二	38
圖【4-23】	V 群組圖三	38
圖【4-24】	All PN 圖一	40
圖【4-25】	All PN 圖二	40
圖【4-26】	All PN 圖三	41
圖【4-27】	All PN 圖四	41
圖【4-28】	All PN 圖五	42



表目錄

表【4-1】	實驗資料詳表	22-25
--------	--------------	-------



第一章

緒論

1.1 研究動機

人類的大腦一直是各領域的學者深感興趣的主題，解開大腦的奧秘除了對於一些腦神經疾病，如阿茲海默症(Alzheimer's Disease, AD)、帕金森氏症(Parkinson's Disease, PD)等，可以有更深的認識以求找出治療的辦法外；探索大腦的能力並且開發其無窮的潛力一般相信會有足以改變人類發展的力量。但是人腦中約有一千億個神經細胞，結構極為複雜，且人類的生命週期長，加上礙於道德法規的限制，也難以直接取得人腦的神經資料進行研究，因此科學家們便拿其他小型動物的大腦代替，如小白鼠，兔子及果蠅(Drosophila)。其中果蠅因為繁殖容易，生命週期短，且和人類一樣有學習、記憶的能力，此外有許多的人類疾病基因在果蠅中也可以發現，因此果蠅腦成為研究人腦的一個先驅目標，目前世界上有許多大型的研究團隊都正致力於解開果蠅腦的秘密。

2004 年的諾貝爾生理暨醫學獎的得主 - Richard Axel 及 Linda Buck - 他們的研究[1]發現果蠅嗅覺神經細胞表面有許多的嗅覺受體，這些嗅覺受體接受各種氣味刺激後，和氣味分子結合產生對應的排列組合，進而發送出不同的神經訊號傳遞到腦中。這些訊號之後經由許多嗅小球(Glomeruli)的重新組合後再次發送，透過嗅覺投射神經(Projection Neuron, PN)傳到果蠅腦中的蕈狀體(Mushroom Body, MB)，並接著往更深一層的側角區(Lateral Horn, LH)傳送。果蠅腦中的 MB 約由兩千五百個神經細胞組合而成，國內外已經有不少研究顯示 MB 可能是腦中訊號的聚集處理部位，也是研究的重要區域。今年三月份清華大學江安世老師的團隊

發表在 Cell 期刊上的論文[2]，就是關於果蠅腦的嗅覺神經地圖，透過進一步發現蕈狀體的花萼端(MB calyx)中嗅覺投射神經分布的影像地圖，可以看清楚這些神經的分布結構，揭開以往不清楚的後段神經細胞生長分布圖。

1.2 研究目標

目前果蠅嗅覺的機制大致研究到了嗅覺神經通過嗅小球到達 MB 的事實，但是這些嗅覺訊號和接受的神經細胞之間對應的關係還不是很清楚。為了將這些經由嗅覺受體接收來的訊號進一步分析，除了要先看清楚神經的分布之外，我們想利用科學計算的方法將這些神經的位置加以分析，也許會是一個有用的分類方式，並且期望對以後的果蠅腦神經圖譜的定位及分析有所幫助。

我們的研究目標是建構這些 PN 的一個空間座標在 MB calyx 中，並且利用這個座標找出這些嗅覺投射神經伸展到 MB calyx 裡面的軸突末梢(Terminal Boutons, TBs)分布。下圖是果蠅的嗅覺神經分布簡圖[3]，紅色圓圈裡就是 MB 還有其通過的 PN(紅色線)，也就是這篇論文關注的區域。

圖【1-1】符號解釋：

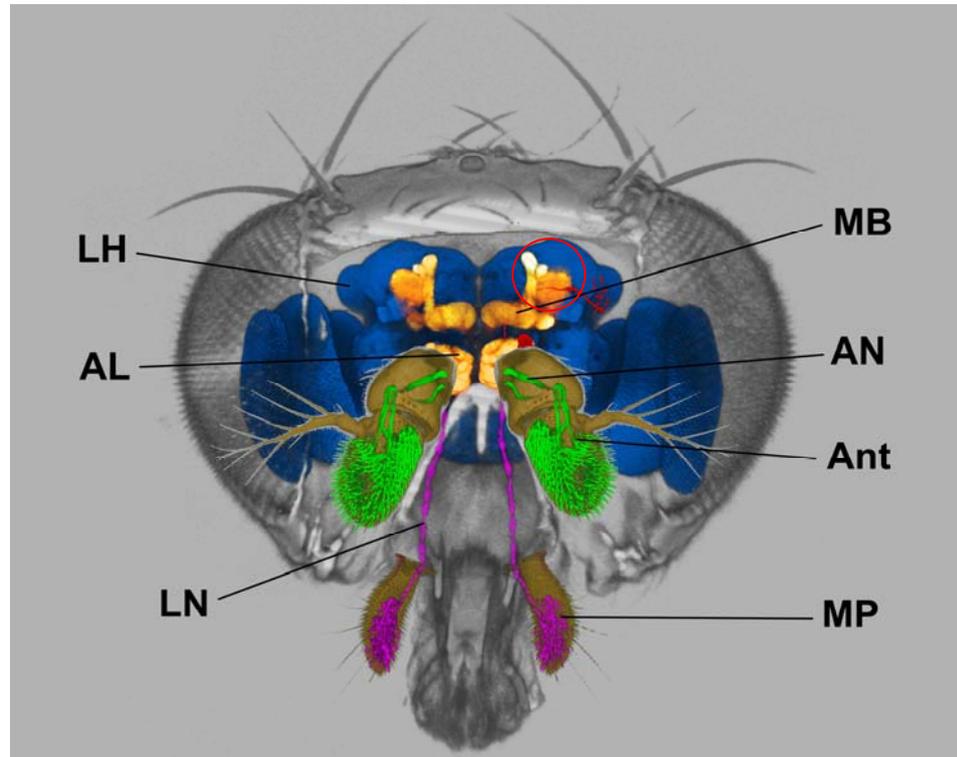
Ant: antenna, AN: antennal nerve

MP: maxillary palp, LN: labial nerve

AL: antennal lobe, MB: mushroom body, LH: lateral horn

嗅覺訊號經由綠色和紫色的神經元傳送至 AL，再由紅色的 PN 經過 MB 再傳

送到到 LH。



圖【1-1】果蠅嗅覺神經網路

1.3 論文架構

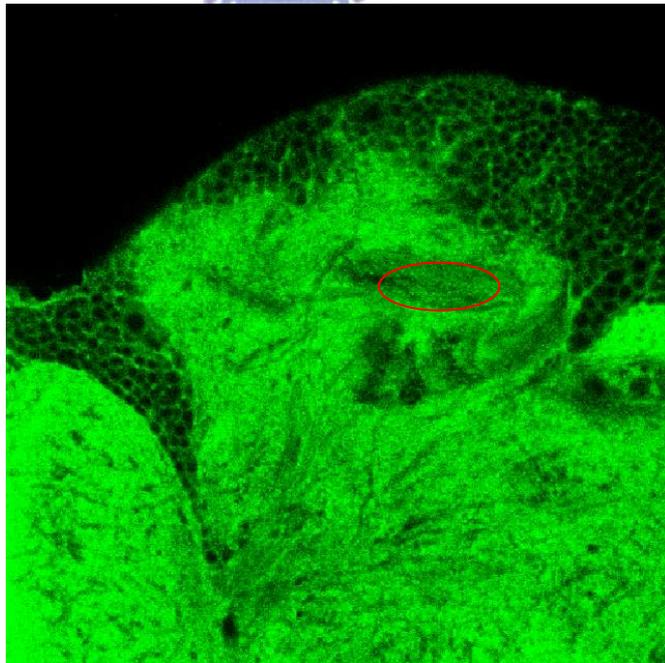
透過第一章的引導，可以知道相關的研究進展及我們為什麼要進行這項研究。接下來在第二章介紹我們的實驗想法、文獻回顧和一些延伸的方法。第三章會介紹我們的實驗的方法和實驗流程。第四章為實驗的結果，和幾個特殊的例子。最後一章則是結論和這個研究的未來發展。

第二章

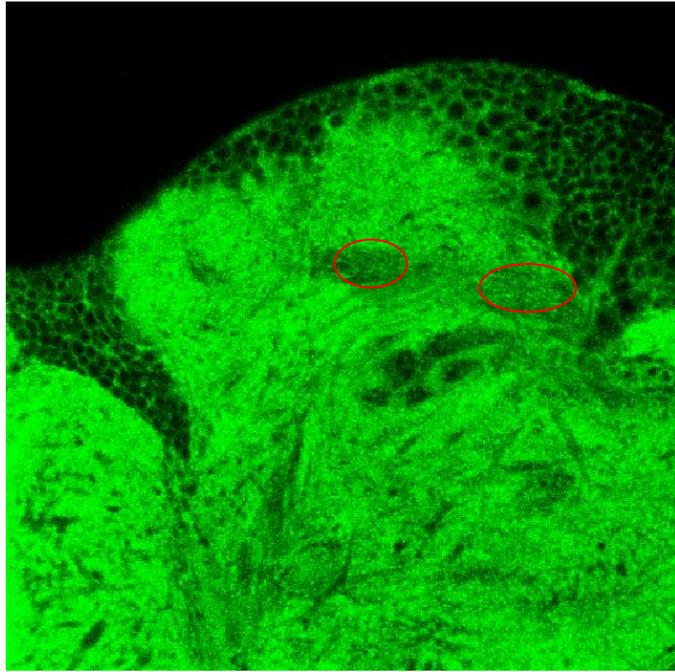
實驗想法及文獻回顧

2.1 前言

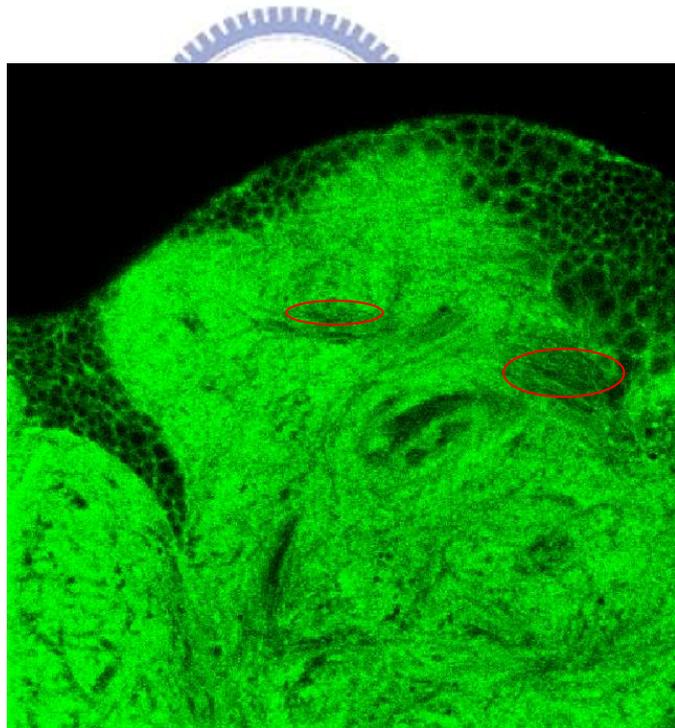
在第一章中我們提到想要在果蠅腦中建置一個座標系統，方便我們用來定義神經細胞或軸突末梢的位置，但是在我們看過的果蠅腦資料中，光用肉眼實在難以抓出其規則，不僅神經生長的樣子、長度都有差別，大腦中的細胞分布也相當複雜，甚至在光學取樣的時候放置的角度也有明顯差別，使得單純處理影像更加困難。因此我們想到必須從大腦中找出一樣的特徵來處理。我們幾經觀察後發現果蠅的嗅覺 PN 在經過 MB calyx 時會有一段類似的生長趨勢，我們用連續幾張圖來表示。紅色橢圓形裡的深色陰影就是 PN 會通過的地方，這三張圖的陰影部分連結起來像一個彎曲的通道，形狀很像一條管道間，讓 PN 順著它生長：



圖【2-1】管道間圖一



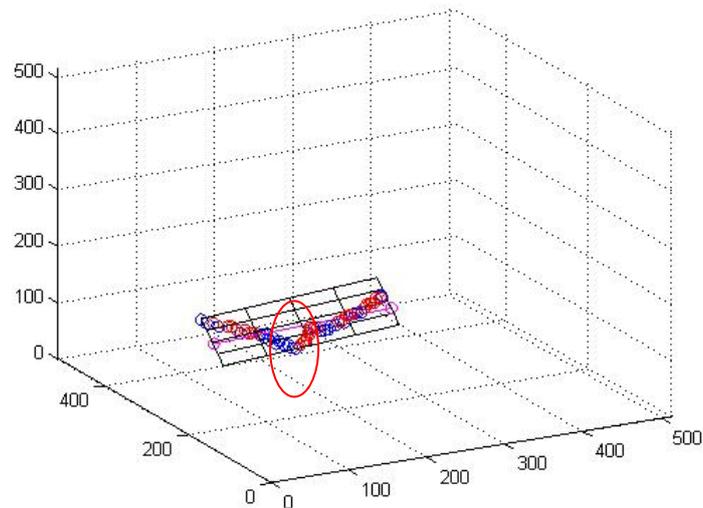
圖【2-2】管道間圖二



圖【2-3】管道間圖三

上面三張圖陰影部分在每組大腦資料都有相似的存在。

雖然 PN 生長的形狀是相似的，不過在管道間裡，PN 還是可以任意的曲折生長，並不一定是單純的曲線，我們取出的 PN 資料就是一群亂數的點集合，不過大致上會有像下圖一般的分布，紅色橢圓圈的那段彎曲的地方，就是我們說的管道間彎曲處，這裡彎曲的程度不盡相同，但是幾乎所有的 PN 都有：



圖【2-4】PN 分佈圖

因為每個果蠅腦資料都可以找到這些相似的 PN 生長分布圖，所以我們想到利用迴歸分析的方法處理這些 PN 的座標資料，找出這條 PN 的最佳 fitting line，和最佳 fitting plane 再來加以利用，所以我們想到統計學常用的迴歸分析 (Regression analysis)。

Regression 這個用語的出現最早可以追溯到十九世紀，當時是用來描述一種生物上的遺傳現象：子代異常特徵往往和其父代的異常特徵不甚相似，反而和其更前代的祖先的異常特徵比較相近。之後這個詞才慢慢被引伸到跟統計相關的研究，近代更常利用電腦幫忙分析大量的資料。

2.2 Least squares regression

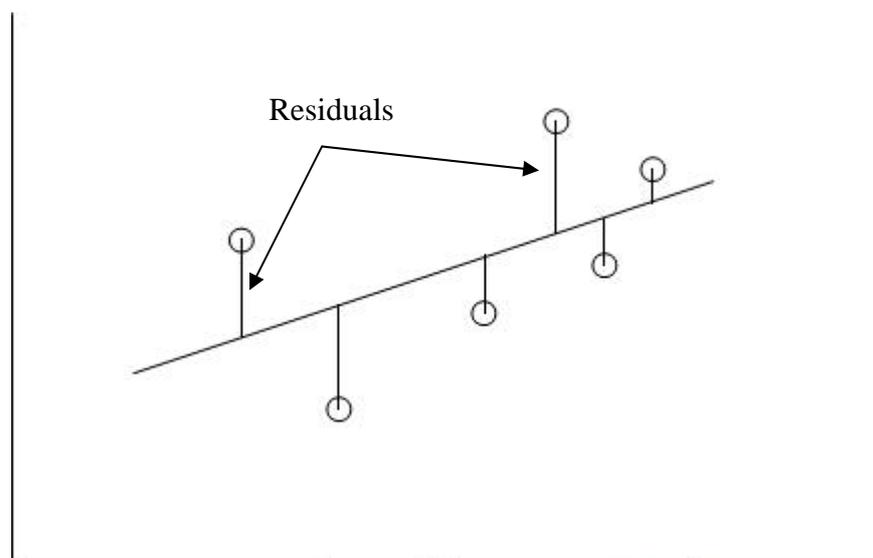
最基本的 regression 我們首先會想到的是 simple linear regression。假設我們有一組 independent (or, explanatory) variables: x_i ；根據 x_i 我們預測出一組對應的 dependent (or, response) variables: y_i 。

下式為 simple linear regression 的 general form:

$$y_i = \alpha + \beta x_i + \varepsilon_i \quad (1)$$

其中 α 為 intercept； β 為 slope； ε 為 error term。我們可以利用 least square method 來解這個問題。Least square method 歷史久遠，最早由 Legendre 在 1805 年提出，之後 Gauss 在 1809 年出版的書中也提到這個方法，並且宣稱 least squares method 他早在 1795 年就用過了。

least squares method 主要功用是使所有的 residuals 總和為最小，如下圖有六個 residual:



圖【2-5】least squares method residual

並解得 α , β 的最佳預測值：

$$\hat{\beta} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} \quad (2)$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \quad (3)$$

其中 \bar{x} 和 \bar{y} 為 x_i 和 y_i 的平均數。

2.3 More regression method

然而 least squares 方法有個缺點就是受到 outliers 的影響太大，也就是明顯距離整體 data 值較遠的孤立點。有 outliers 出現，會使 least squares method 結果偏頗、效率降低。所以後來由 Hampel et al. 在 1986 年提出 robust regression[4]，概念是給 outliers 較低的權重(weight)，甚至忽略，以解決 least square 的缺點，但是要付出比較多的計算時間，當然關於 weight function 是有很多種設計方法。

某些特別的資料特性中沒有清楚的相依關係，例如工程學或自然科學常見的資料，這時便不能單單只是將 residual 定義為延著某個軸的資料到 fitting function 的距離；更直覺的想法應該是資料到 fitting function 的垂直最短(perpendicular)距離才是我們要計算的。這便是 total least squares regression 的想法，Y. Nievergelt 在 1994 年對這個方法有一個很好的闡述[5]。後來的 errors-in-variables model[6] 也是這個想法的延伸。

由於我們的實驗資料正符合 total least squares method 的使用特性，所以我們下一章會詳細說明實驗過程。

第三章

實驗流程

3.1 實驗簡介

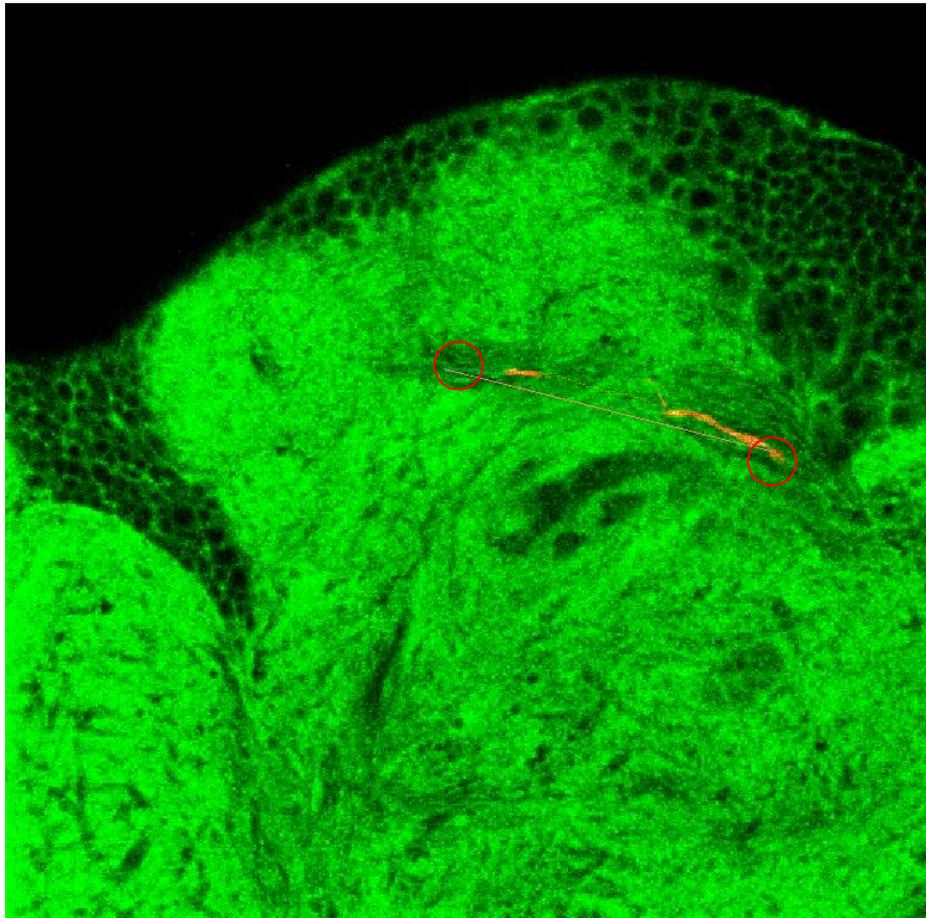
我們的實驗目的是要在果蠅腦中建置一個座標系統，利用這個座標系統來描述果蠅的嗅覺 PN 延伸到大腦中 MB calyx 裡面的 TBs 分佈情形，期望實驗結果可以讓腦科學家做進一步的研究。根據第二章的想法，我們必須先對每組通過 MB calyx 上端的 PN 建構出一個新的座標系統，接著抓出此 PN 的 TBs 的生長位置，經由座標轉換後得到這些 TBs 的新座標，再將各組 TBs 座標放在新的座標系統一起觀察，希望從中發現不同組的 TBs 有沒有特定的群聚現象，甚至會不會有特定的 PN 會生長到 MB calyx 的特定區域。

我們的實驗主要分成三大部分：

1. 選取初始點配合 GVF Snake 方法取出初始資料中的 PN 座標資料。
2. 利用 Orthogonal Regression using PCA 方法計算出該 PN 最佳直線及平面，計算出所需要的三維向量參數。
3. 點取出該 PN 的 TBs 座標，利用座標轉換計算出新的座標位置。再利用 OPGL 繪出在螢幕上，可以做任意角度的旋轉，方便我們整合不同組的 TBs 資料作觀察。

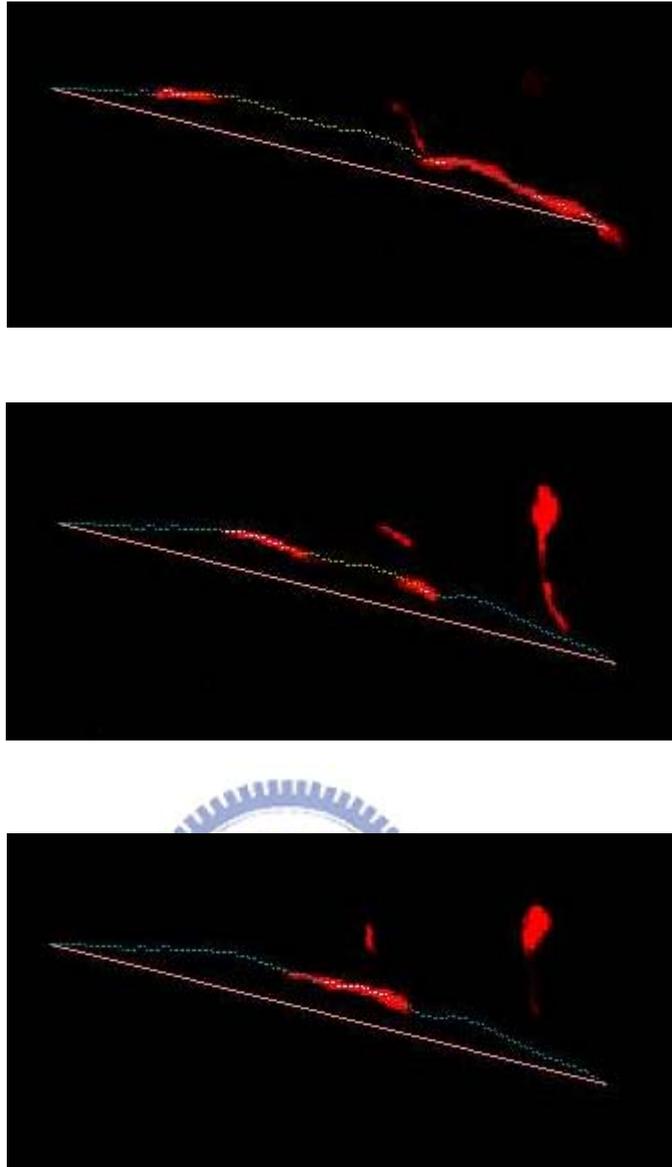
3.2 取出 PN 的初始資料

首先我們必須處理原始資料，取出我們需要的 PN 座標位置，這裡我們先用人工點選初始點，接著利用 GVF Snake 的方法將他們收斂到 PN 的中心線上。因為每組資料的差異性，我們必須儘量取出一致的初始資料，所以我們盡量選擇從 MB 要進入 calyx 的周圍開始抓取我們的 PN，如下圖紅色圓圈位置所示：



圖【3-1】取 PN 初始位置圖

下面是將 PN channel 單獨抽離出來看，紅色就是 PN 本身，圖中彎曲的虛線就是我們取的 PN 位置軌跡線，可以看的出來，都有抓到 PN 的走向和形狀。



圖【3-2】GVF Snake on PN

我們簡單介紹一下 GVF Snake 這個方法，Gradient Vector Flow (GVF) Snake 是在 1998 年由 C. Xu 和 J.L. Prince 提出[7]。主要是利用 GVF 產生的 vector field 來當作 Snake 中的 external force 計算出 snake 的最小 energy 之後，會使 snake 曲線逼近到圖形中央。

3.3 正交迴歸分析

3.3.1 概論

我們所使用的迴歸分析法是 MathWorks, Inc[8]中提到的一種方法，稱作 orthogonal linear regression using principal components analysis。Orthogonal regression 的意思其實和我們在第二章有提到的 total least squares regression 類似，也就是要使 data 到 fitting function 之間為垂直最短距離，也是一種適合用在工程資料上的迴歸分析法。而 principal components analysis(PCA)也是一種資料分析方法，特別適合用在取出 low dimension subspace 的資料。我們這裡是利用 PCA 來做 orthogonal linear regression，接下來我們會介紹這兩種方法。

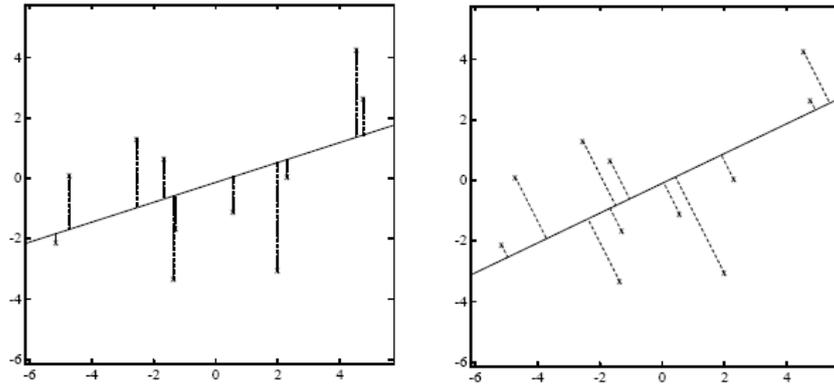


3.3.2 Total least squares

我們取出的PN位置座標因為是人工選取這裡會有一定程度的errors產生，再經過GVF Snake決定後，便無法說有哪一個維度是精確的independent or dependent data，如此一來用least squares regression是沒有說服力的。因為我們的三維座標資料都有error成分，也都是相同單位型態的資料，所以我們這裡選擇total least squares regression (TLS)來做逼近。

TLS相較於least squares regression最大的不同在於，TLS所要的是每個維度到預測函數的直線最短距離，而不是只取預測資料的誤差距離。

下面的圖可以明顯的看出不同，左邊是least squares，右邊是TLS[9]。



圖【3-3】residual取法比較圖

我們的實驗就是要做出右圖的效果，所以我們接下來利用PCA來達成。

3.3.3 Principal components analysis

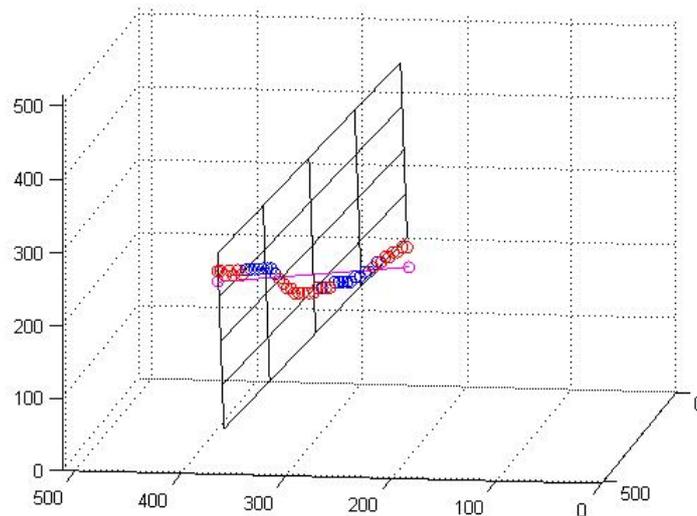
主成分分析(Principal Components Analysis, PCA)，主要的功用是將一群多維的資料簡化成幾組線性獨立的組合變數，期待用比較少的維度就可以表現出原本資料最多的變異度。其實PCA和TLS都是想把 N 維度的資料降低維度，而且期待找出的新維度 P 上的資料能描述一定程度 N 的原始資料分布，這兩種方法都有 minimizing orthogonal distance 投射到原始的 N 維度上資料的意涵。

觀察我們取出的PN資料，特性是有相當的共線性特徵和我們實驗目的很明顯的是要找到兩個組成單位；一個最佳平面並取出它的法向量和一個最佳直線的方向向量。所以我們利用PCA來取出我們想要的principal component，我們需要的最佳平面要取出兩個主成分feature vectors，而最佳直線則是取出一個主成分feature vector。

PCA主要有下列幾個步驟：

1. 計算每個維度資料的平均值。
2. 計算出共變異數矩陣(covariance matrix)。
3. 求出共變異數矩陣的eigenvalues及eigenvectors。
4. 依照eigenvalues的大小取出principal component的feature vector。

經過主成分分析計算，我們可以得到一組feature vectors是形成我們要找的最佳平面的兩個basis vector，加上另一組basis也就是此平面的normal vector，所以可以得到平面的方程式。最佳直線的計算也是類似的方法，只是只取一組feature vector為direct vector，便可以寫出我們要的直線方程式。下圖可以觀察到我們的直線其實就已經是在平面上，因為其實在算平面的PCA時，也早就算出了我們的直線的主成分了。下圖可以看到，粉紅色直線就是我們的最佳直線，本身已經位在平面上，其實也就是最佳平面的第一個主成分。



圖【3-4】分析出PN的直線與平面

起初我們是取PN投射到這直線上的頭尾兩個端點，算出中心為之後我們要的新座標原點，由於我們的PN座標資料，有些生長的很扭曲，會造成這個直線在某幾組資料上傾斜角度上特別不同，這會使的實驗結果不好，也就是TBs在空間上的位置差異甚大。

我們嘗試了許多種的原點取法，我用下圖來說明各種取法的差異性：



圖【3-5】新原點取法與位置

圖【3-5】中，紅色直線為我們計算的最佳直線，若是取這紅線中心為新原點，大部份不會取到PN上面，並且根據PN彎曲的程度不同，也會影響紅線中心距離PN的遠近都不同，這當然也就使的依據紅線中心為新原點計算出的TBs位置倍受影響，這樣的新原點顯然並不可靠。

黃色橢圓區域是我們計算PN資料後，找到的最彎曲的位置來當作新的原點，但是PN資料的複雜性實在太大，很快的我們發覺這樣的取法也不適當，雖然找到的新原點的確都是在PN上面，但是由於最彎曲處可能出現在整個PN資料的任何地方，不能做為依據。

最後藍色橢圓的區域，我們取切到PN的兩端點後再取其中點作為新原點，這個區域是PN最接近MB calyx時的一小段區段，通常都是比較平滑的曲線甚至接近直線，這個區段有兩個特性，讓我們最後決定使用這個新原點。一是這個區段在每個PN上都找得到，而且將這區段逼近PN後，取出的新原點幾乎都是在PN的身上，這符合我們的期待；二來我們觀察後發現在這個區段附近，比起其他部份會有較多的神經分支，也就是延伸出去形成TBs的branch。由於我們正是要計算這些TBs的分佈，所以利用這個位置的中心為新原點是我們最後的決定。最後我們有了一個最佳直線、最佳平面的向量參數，和一個新原點。

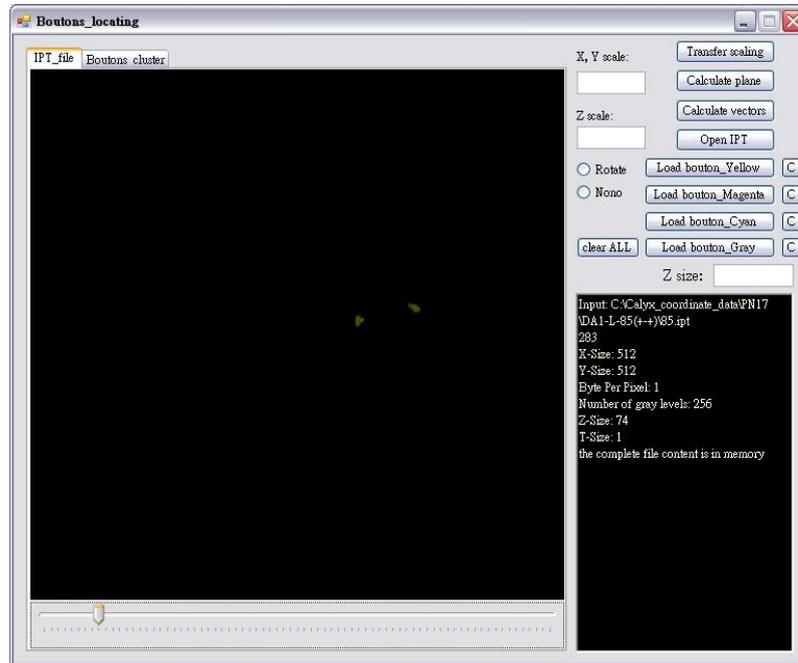
3.4 取出 TBs 及座標轉換



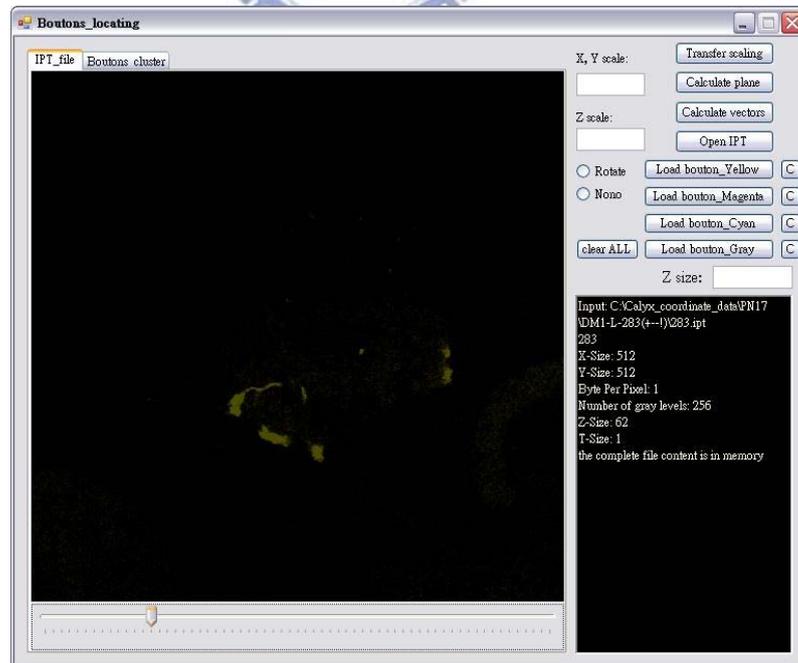
接下來我們要開始著手取出 PN 上的所有 TBs，我們觀察發現基本上一條 PN 上面的 TBs 數量並不會太多，加上我們還沒有發現夠好的演算法可以自動取出所有 TBs，所以我們在這裡做一個軟體介面如圖【3-6】、【3-7】，讓使用者可以自行在 IPT_file 的頁面上點選出 TBs。在點選出 TBs 後，我們會同時進行座標轉換並顯示在 Boutons_cluster 的頁面上。

果蠅嗅覺 PN 上的 TBs 大致有兩種型態，一種是近似球形、另一種是長條不規則形，後者常常分佈範圍較大。如下圖【3-6】可以看到大多是球形的 TB，這樣的 TB，我們就是選取中間的張數或是球形範圍最大的那張圖作為選取的依據，我們用這一點來表示這個 TB。而圖【3-7】就是長條型的 TB 分布，這裡我們就選取此 TB 最大分布範圍的那張圖來選取 TB，也就是盡量在中間將 TBs 以多個點

表現出來。用這樣的方式來選取 TBs，當然會有人為的誤差，我們用這兩個分類法希望盡量減少這個部份的誤差。

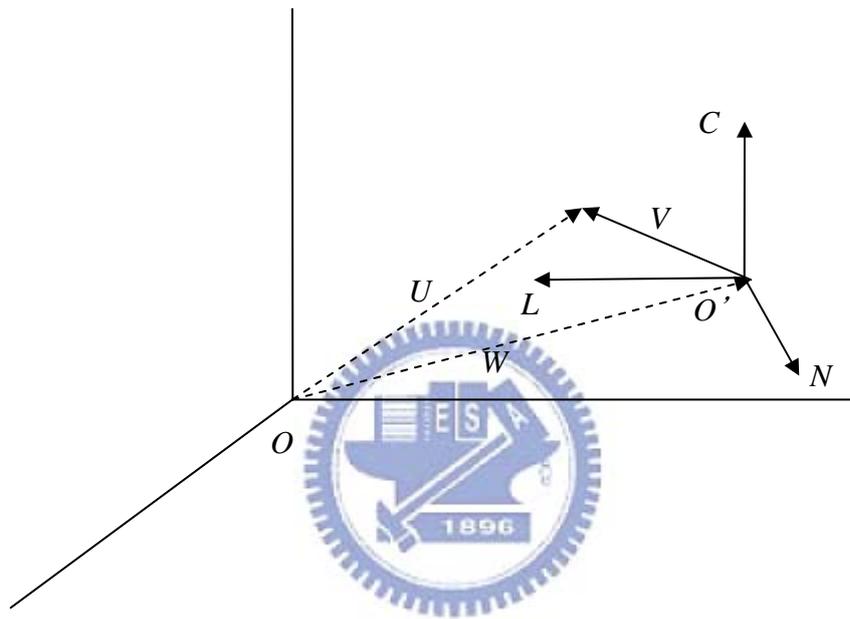


圖【3-6】球形 TB



圖【3-7】長條不規則形 TB

在 3.3 中，我們已經計算出建構新座標系統所需要的三個向量，分別是最佳 fitting line 方向向量假設為 L 、最佳 fitting plane 的法向量假設為 N 和兩者的外積向量 $L \times N$ 假設為 C ，以及我們在新座標的原點 O' 。接著我們就可以把剛才點選出來的 TBs 座標進行座標轉換。



圖【3-8】座標轉換圖

假設點選出來的 TB 向量在原座標系為 U ，而新座標的原點向量為 W ，所以我們知道在新座標上面的 TB 向量應為 V

$$V = U - W \quad (4)$$

接著我們先將 L, N, C 先化為單位向量，因為這是新座標上的三維座標單位，也方便之後描述新座標時的轉換運算。假設轉換後的三個單位向量為 $L_u = (L_{ux}, L_{uy}, L_{uz})$, $N_u = (N_{ux}, N_{uy}, N_{uz})$, $C_u = (C_{ux}, C_{uy}, C_{uz})$ ，令 $V = (V_x, V_y, V_z)$ ，接著我們就可以列

出一個聯立方程式， a, b, c 為我們想求的新座標常數：

$$V_x = aL_{ux} + bN_{ux} + cC_{ux}$$

$$V_y = aL_{uy} + bN_{uy} + cC_{uy}$$

$$V_z = aL_{uz} + bN_{uz} + cC_{uz} \quad (4)$$

令矩陣 $M = [(L_{ux}, N_{ux}, C_{ux}), (L_{uy}, N_{uy}, C_{uy}), (L_{uz}, N_{uz}, C_{uz})]$ ，我們可以將上一個式子改寫為：

$$V = (a, b, c) * M \quad (5)$$

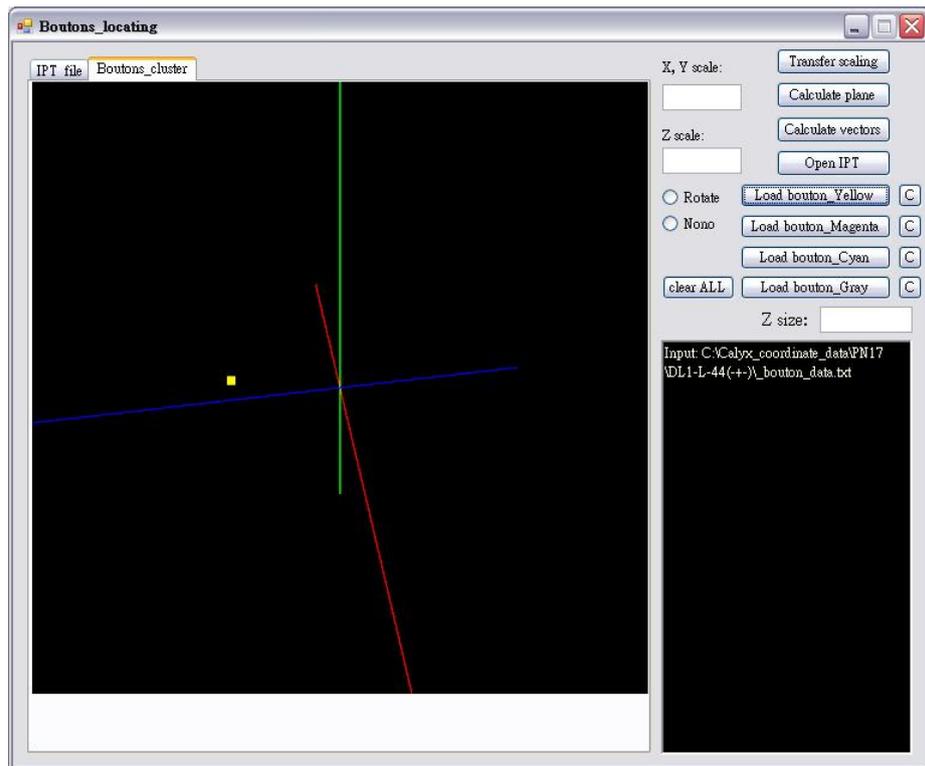
利用反矩陣性質可得 $(a, b, c) = V * M^{-1}$ ，令矩陣 $M^{-1} = [(L^{-1}_{ux}, N^{-1}_{ux}, C^{-1}_{ux}), (L^{-1}_{uy}, N^{-1}_{uy}, C^{-1}_{uy}), (L^{-1}_{uz}, N^{-1}_{uz}, C^{-1}_{uz})]$ 所以可解出 a, b, c 為：

$$a = V_x * L^{-1}_{ux} + V_y * L^{-1}_{uy} + V_z * L^{-1}_{uz}$$

$$b = V_x * N^{-1}_{ux} + V_y * N^{-1}_{uy} + V_z * N^{-1}_{uz}$$

$$c = V_x * C^{-1}_{ux} + V_y * C^{-1}_{uy} + V_z * C^{-1}_{uz} \quad (6)$$

我們對於某個 TB 向量 U ，轉換到新座標的位置便是 (a, b, c) ，使用者每輸入一個 TB 點位，我們就會即時進行上列的轉換，並且同步更新顯示在我們程式的第二頁畫面上(紅、綠、藍線為新座標軸)，如下圖黃色點代表一個 bouton 新座標，顯示在 Boutons_cluster 的頁面上：



圖【3-9】選取 TB 後並顯示

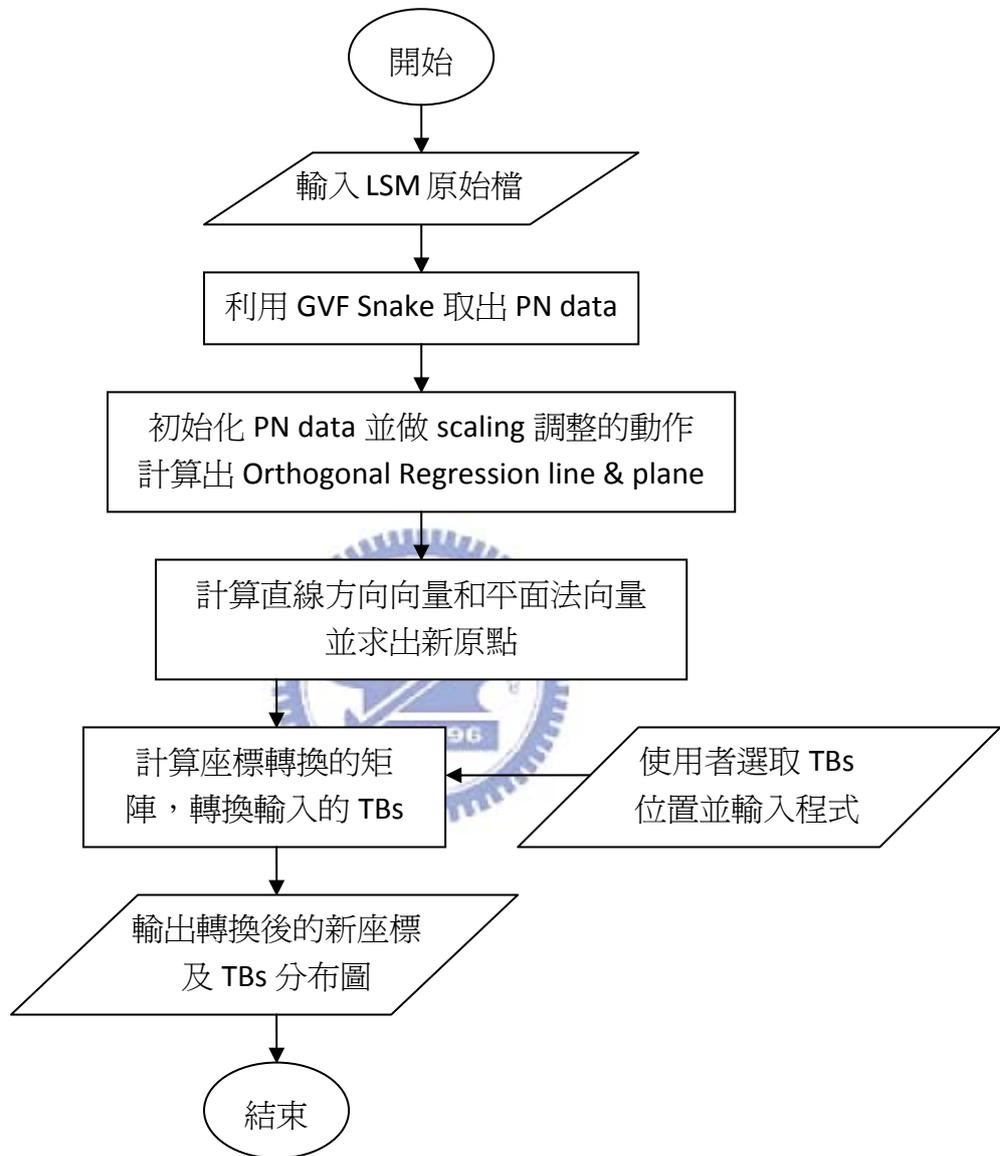
3.5 改進初始資料



在我們處理初始資料時發現每個 LSM 檔案的 Z 方向 resolution 比 X, Y 方向要小的多，大約只有 60 張到 75 張之間，而 X, Y 卻是 1024 x 1024。此外每個組資料的 scaling 大小也不相同，這對實驗結果影響很大，因為在原始資料裡，TBs 的位置在 Z 軸上可能差距僅有四、五張 Z 維度；但是我們如果 scaling 不對，轉換後的新座標差距就不僅僅是四、五倍了，而要再乘上放大的倍數。

最後我們的實驗是以 X-Y scaling 為 0.17 為標準，在做 orthogonal regression 前就先把所有資料的三個維度都調成一致。

3.6 實驗流程圖



圖【3-10】實驗流程圖

第四章

實驗結果

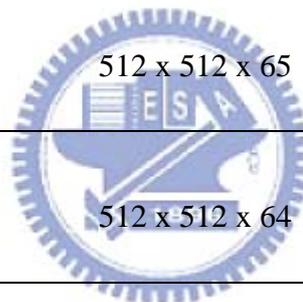
4.1 實驗資料

我們的實驗資料是由清華大學江安世老師實驗室提供的果蠅腦圖像，原始的 data 是每組有兩個 channel 的 LSM 檔案，主要內容是經過染色的一組果蠅嗅覺 PN 和果蠅大腦部份組織這兩個 channel，總共有 17 組不同編號的資料，其中每個編號裡面有好幾個 PN，各有數組左右腦的資料，因為我們的 data 都是由右下往左上生長的 PN，所以我們給予增加代號為-L。基本上每個編號取兩組 PN 來進行我們的實驗，其中 VA2, VA4, VC1, VM4 裡面只有一組為 L，所以只選取一個 PN 來進行實驗(表【4-1】標記*的部份)。

下面是我們所取的 30 個資料列表，X, Y 解晰度皆由 1024 降為 512，另外還有 X, Y 的 scaling 數值：

Data_ID	Resolution	X, Y scaling
D-L-188	512 x 512 x 57	0.17 μ m
D-L-221	512 x 512 x 57	0.18 μ m
DA1-L-85	512 x 512 x 74	0.18 μ m

DA1-L-114	512 x 512 x 65	0.18 μ m
DC1-L-24	512 x 512 x 53	0.18 μ m
DC1-L-82	512 x 512 x 57	0.19 μ m
DC2-L-122	512 x 512 x 54	0.19 μ m
DC2-L-176	512 x 512 x 68	0.17 μ m
DL1-L-44	512 x 512 x 65	0.20 μ m
DL1-L-49	512 x 512 x 64	0.19 μ m
DL3-L-115	512 x 512 x 70	0.18 μ m
DL3-L-139	512 x 512 x 74	0.17 μ m
DM1-L-283	512 x 512 x 62	0.16 μ m
DM1-L-284	512 x 512 x 62	0.17 μ m
DM2-L-40	512 x 512 x 74	0.21 μ m



DM2-L-46	512 x 512 x 60	0.18 μ m
VA1d-L-132	512 x 512 x 62	0.17 μ m
VA1d-L-134	512 x 512 x 66	0.17 μ m
*VA2-L-20	512 x 512 x 71	0.20 μ m
VA3-L-172	512 x 512 x 61	0.18 μ m
VA3-L-203	512 x 512 x 71	0.15 μ m
*VA4-L-58	512 x 512 x 50	0.19 μ m
VA5-L255	512 x 512 x 78	0.19 μ m
VA5-L-268	512 x 512 x 64	0.18 μ m
VA7I-L-275	512 x 512 x 39	0.19 μ m
VA7I-L-278	512 x 512 x 44	0.18 μ m
*VC1-L-69	512 x 512 x 68	0.19 μ m

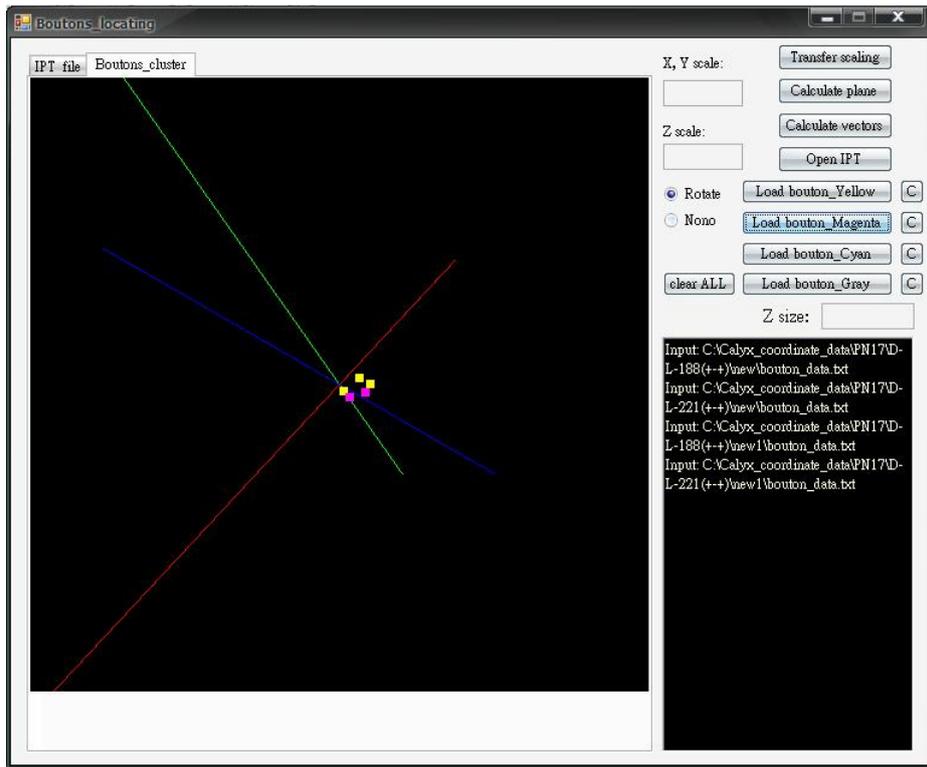
VC2-L-76	512 x 512 x 52	0.16 μm
VC2-L-106	512 x 512 x 55	0.18 μm
*VM4-L-12	512 x 512 x 60	0.17 μm

表【4-1】實驗資料詳表

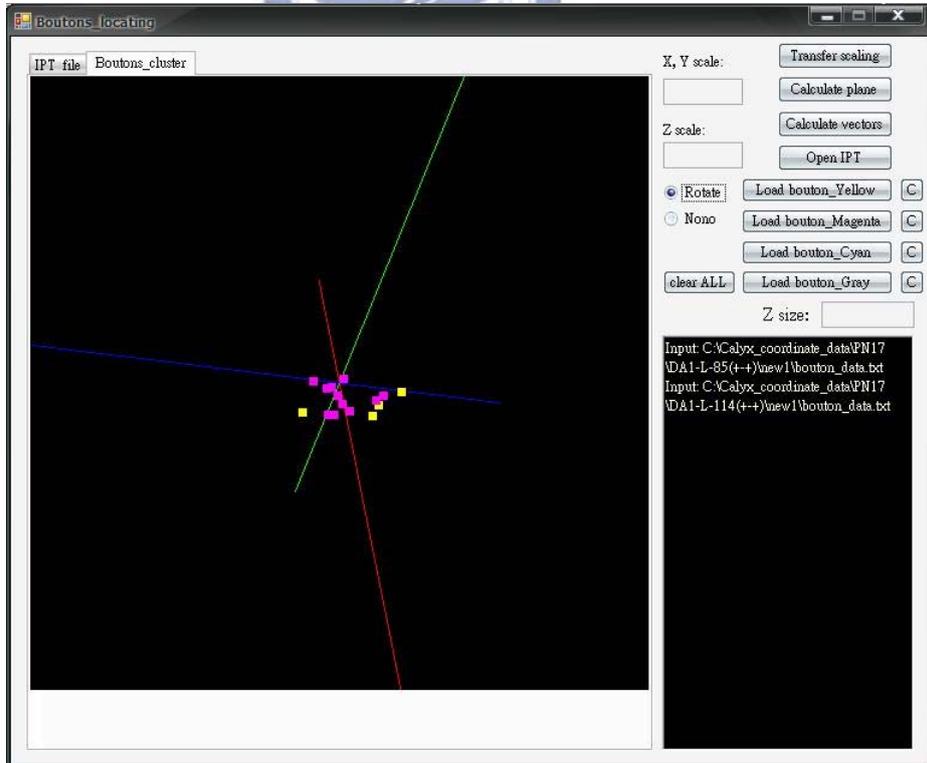
4.2 實驗成果

接著是上面各組 Data 經過選取 TBs 後的三維座標分布圖，因為圖的數量太多，所以我們以同一組編號(L 前面的序號相同者為同組)只選取一張角度的顯示圖來觀察，基本上都是用初始的畫面角度來呈現，但是有些圖因為 TBs 的位置會超出初始畫面，所以我們會旋轉角度來得以顯示全部的 TBs。其中圖中黃色(Yellow)的點為每組中第一個 PN 的 TBs 位置；紅紫色(Magenta)為第二個 PN 的 TBs 位置(例如，DA1-L-85 黃色，則 DA1-L-114 為紅紫色)，如果是單一 PN 就單用青綠色(Cyan)表示。在每張圖中有一個三維的座標軸，其中紅色線為轉換後新座標的 X 軸；綠色線為 Y 軸；藍色線為 Z 軸。

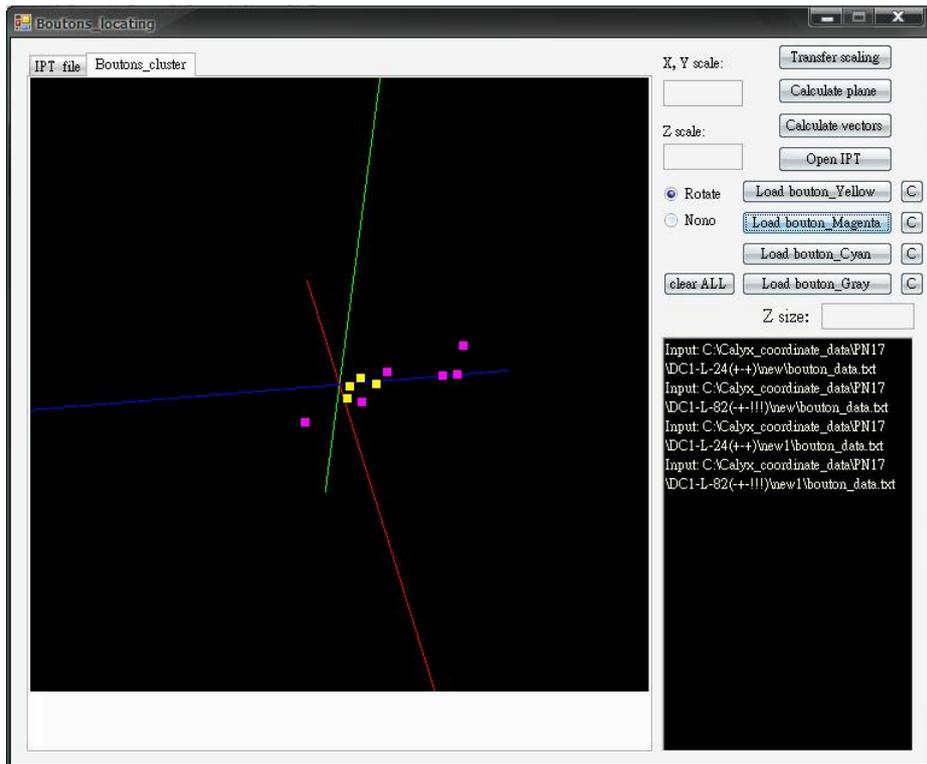
圖共有 17 張分別對應 17 組編號，根據表【4-1】的順序分組列出如下：



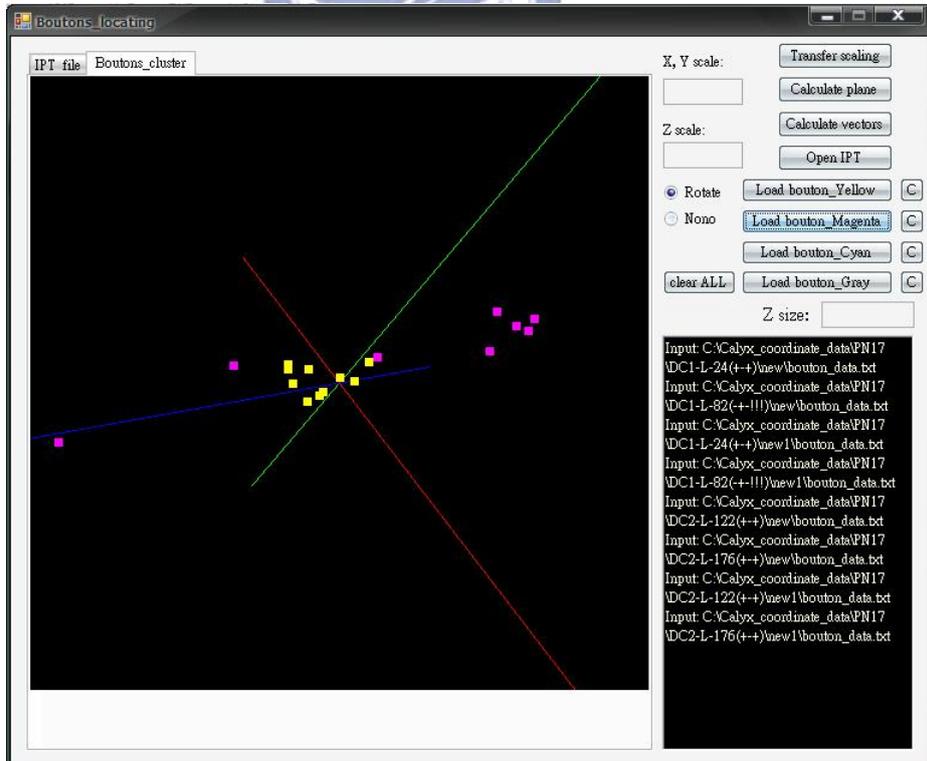
圖【4-1】D-L-188&221



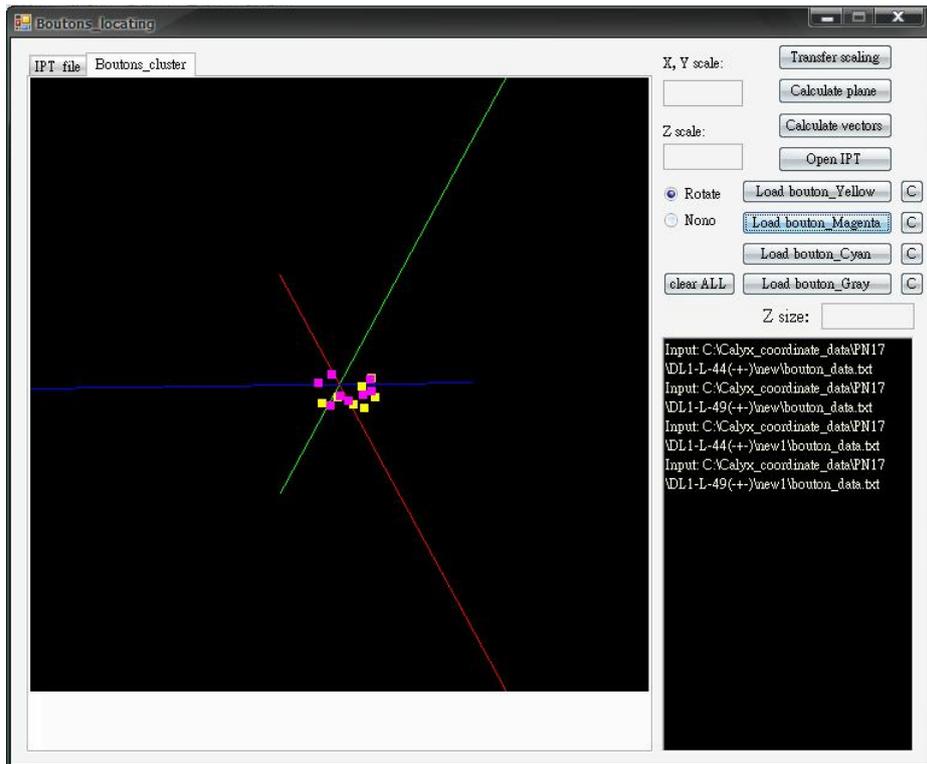
圖【4-2】DA1-L-85&114



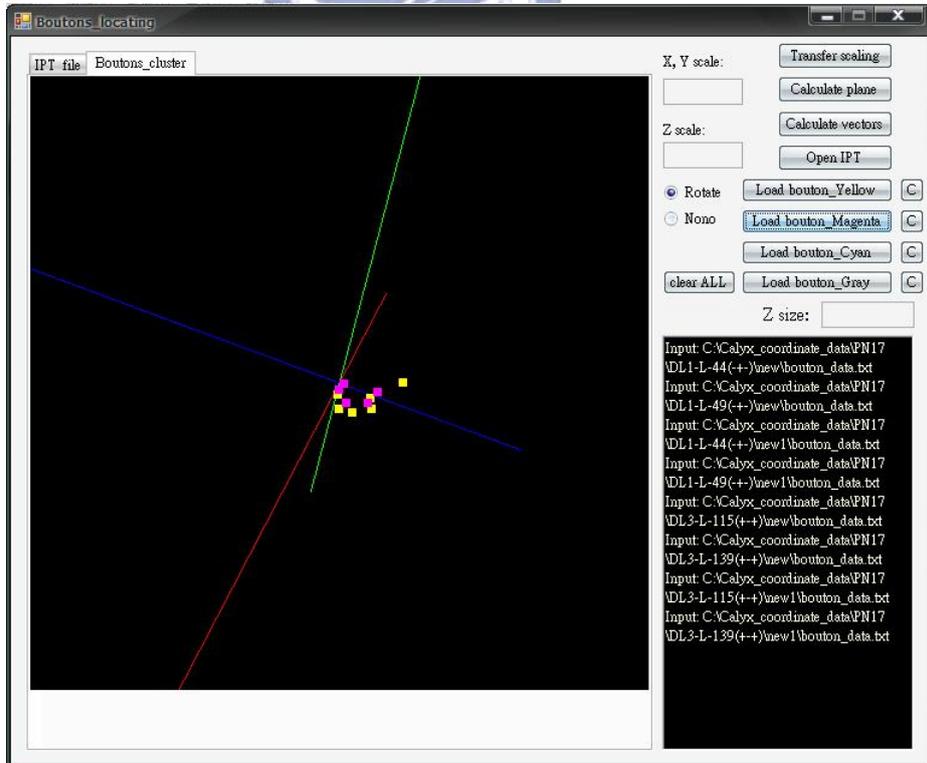
圖【4-3】DC1-L-24&82



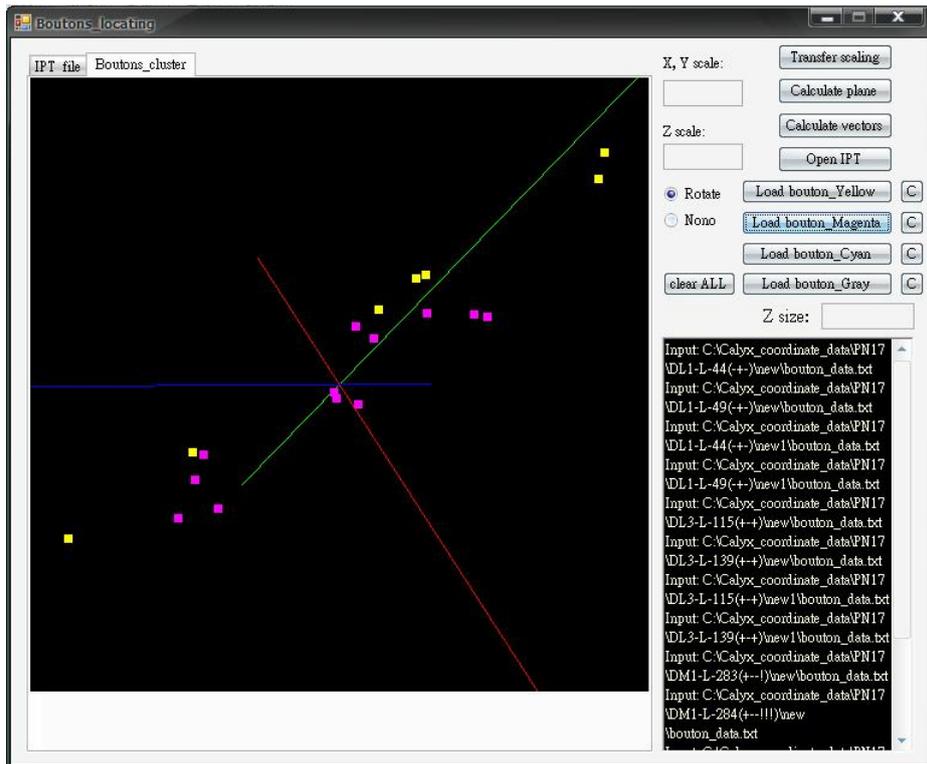
圖【4-4】DC2-L-122&176



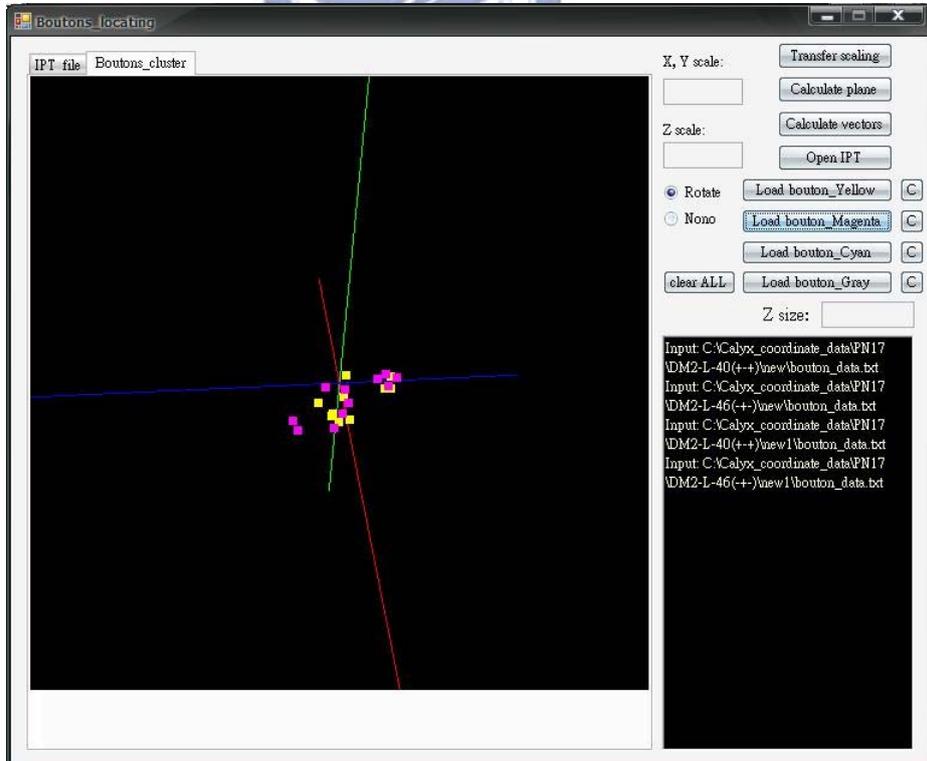
圖【4-5】DL1-L-44&49



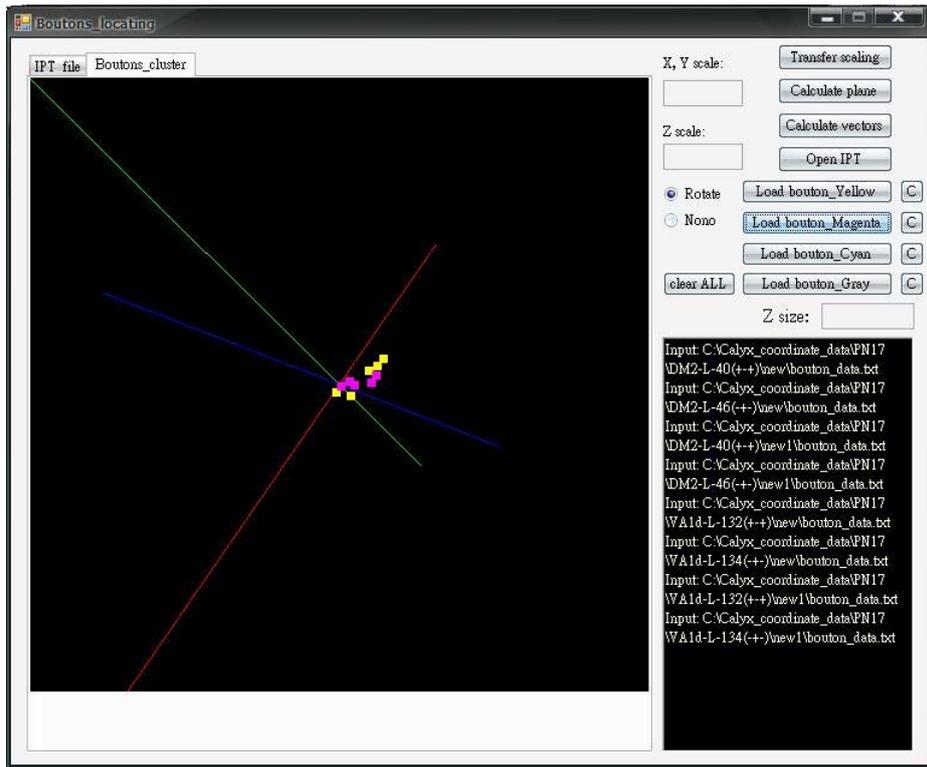
圖【4-6】DL3-L-115&139



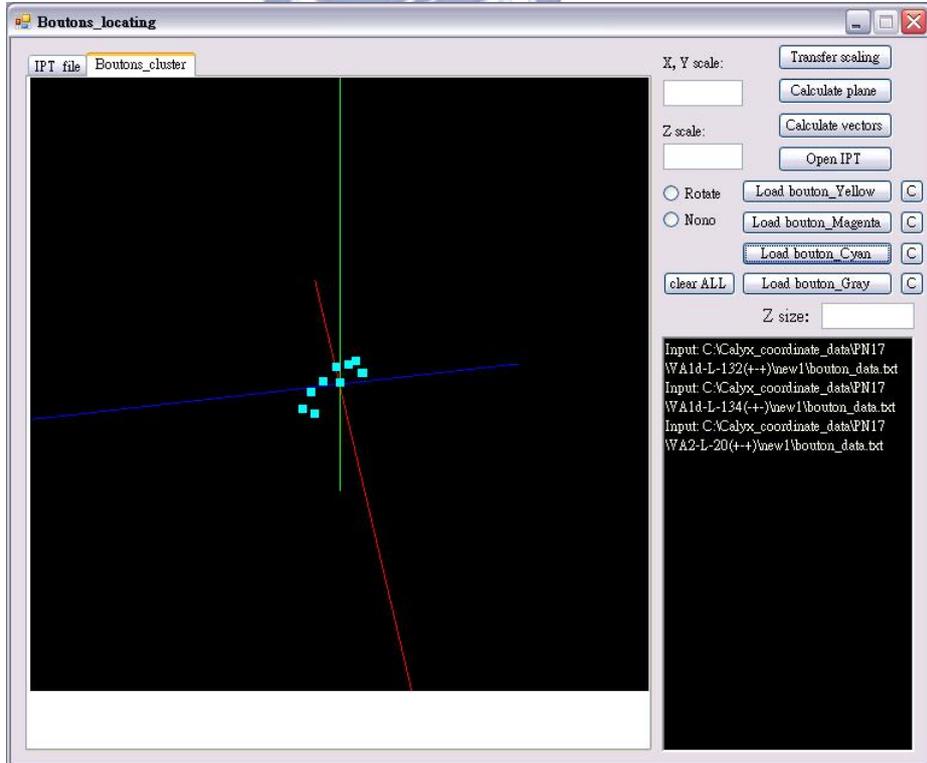
圖【4-7】DM1-L-283&284



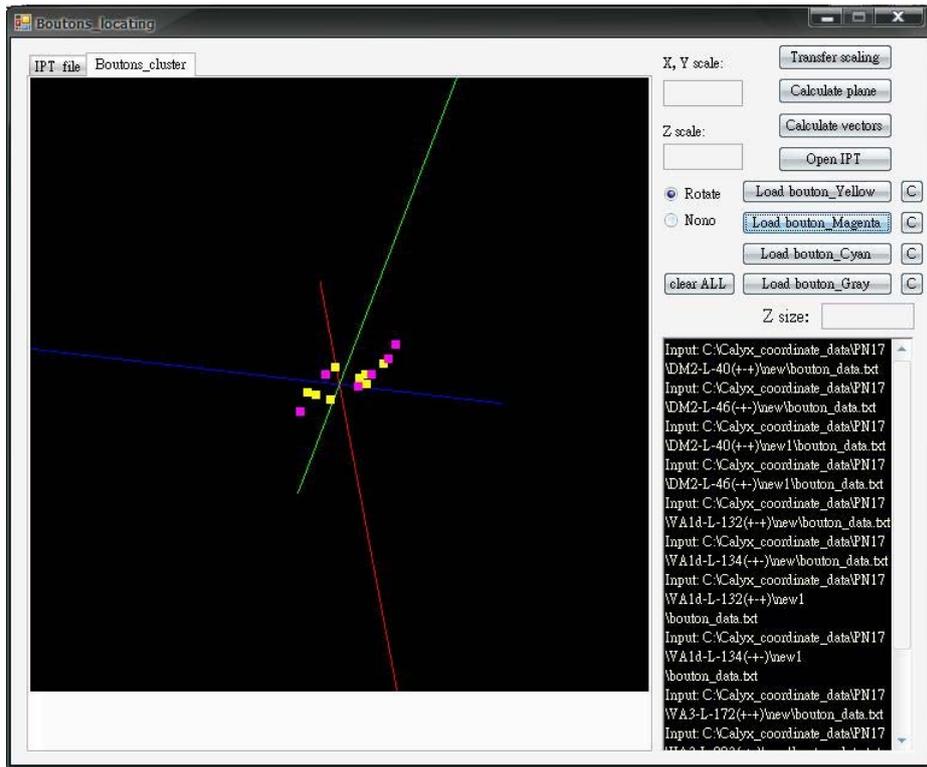
圖【4-8】DM2-L-40&46



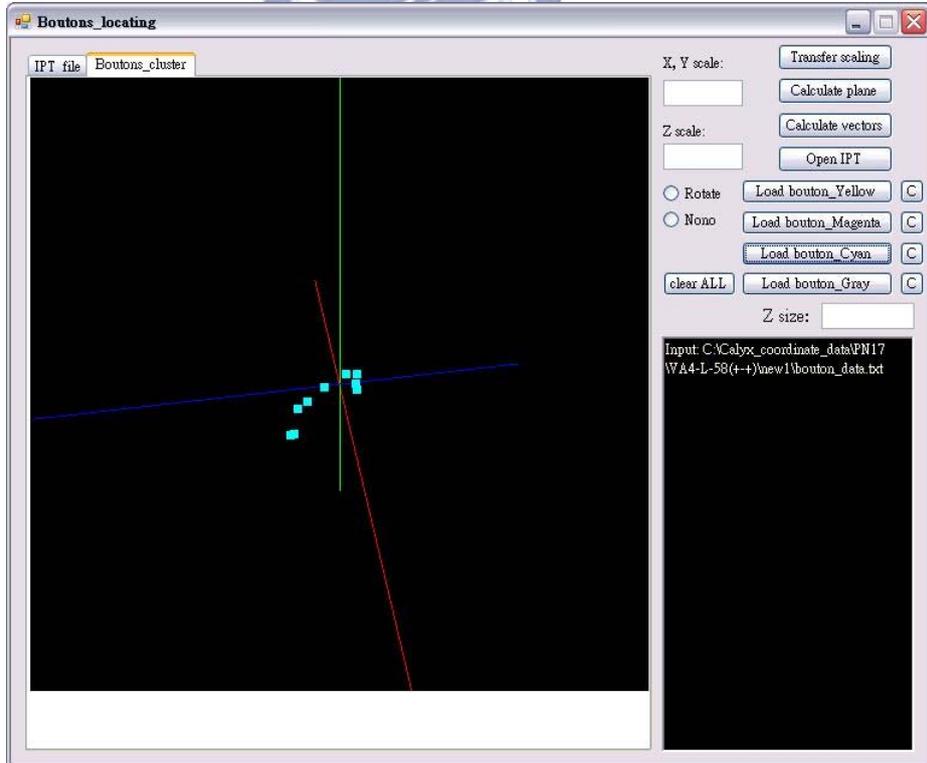
圖【4-9】VA1d-L-132&134



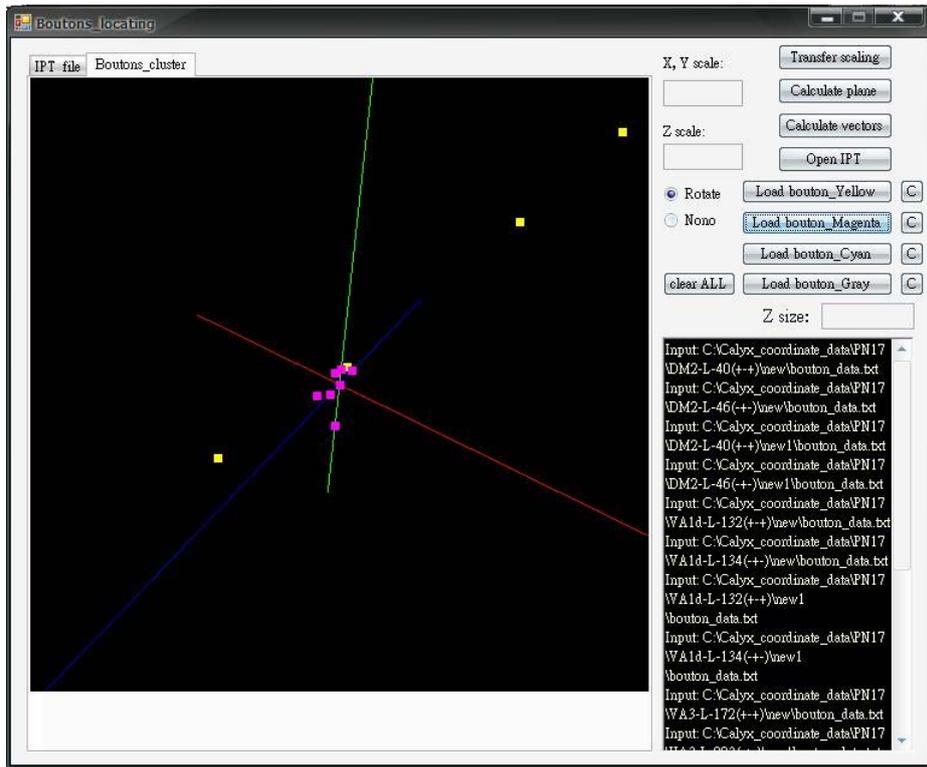
圖【4-10】VA2-L-20



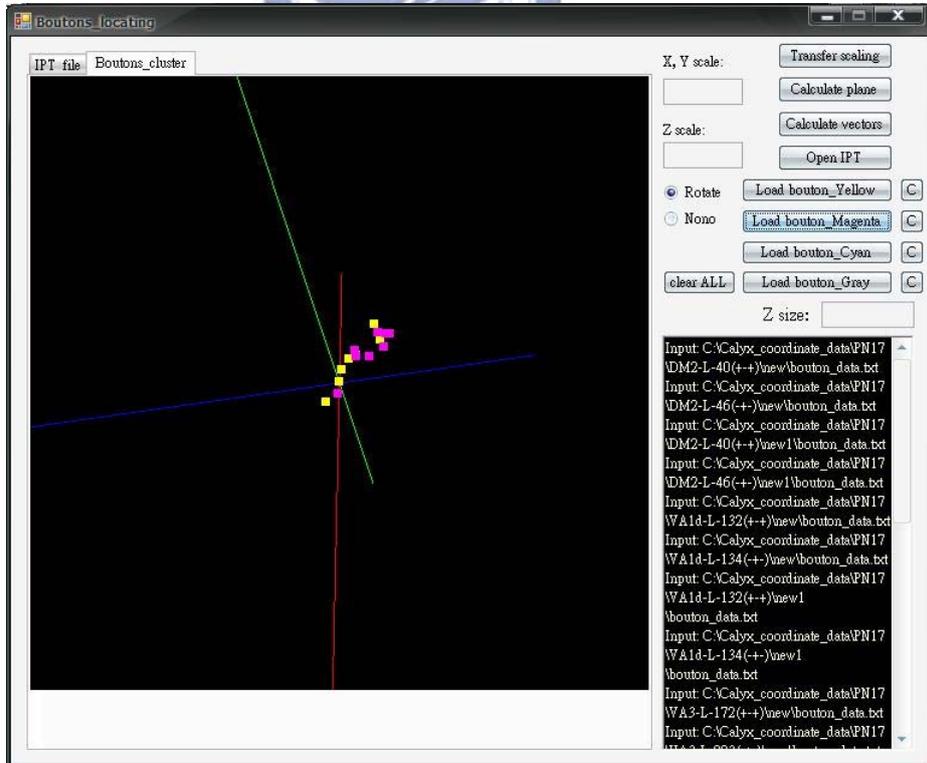
圖【4-11】VA3-L-172&203



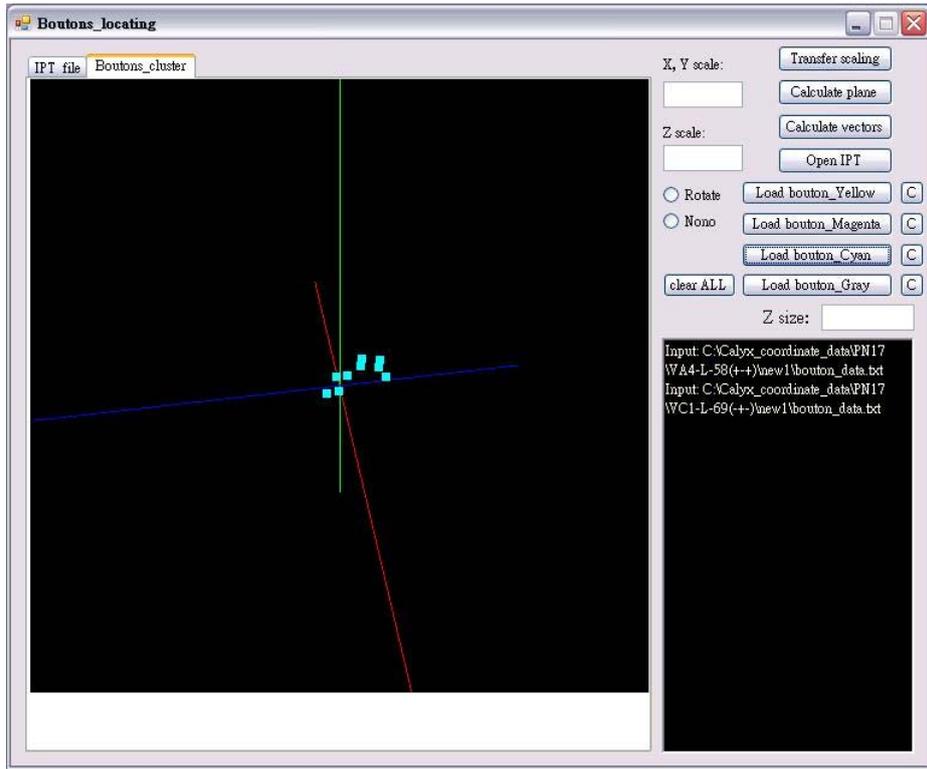
圖【4-12】VA4-L-58



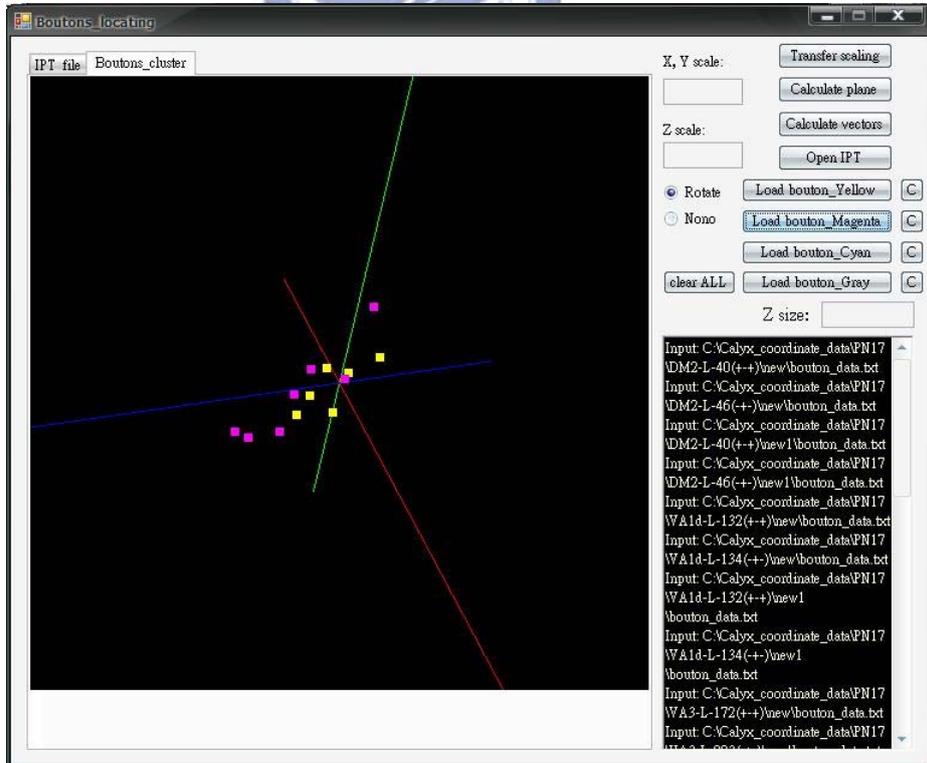
圖【4-13】VA5-L-255&268



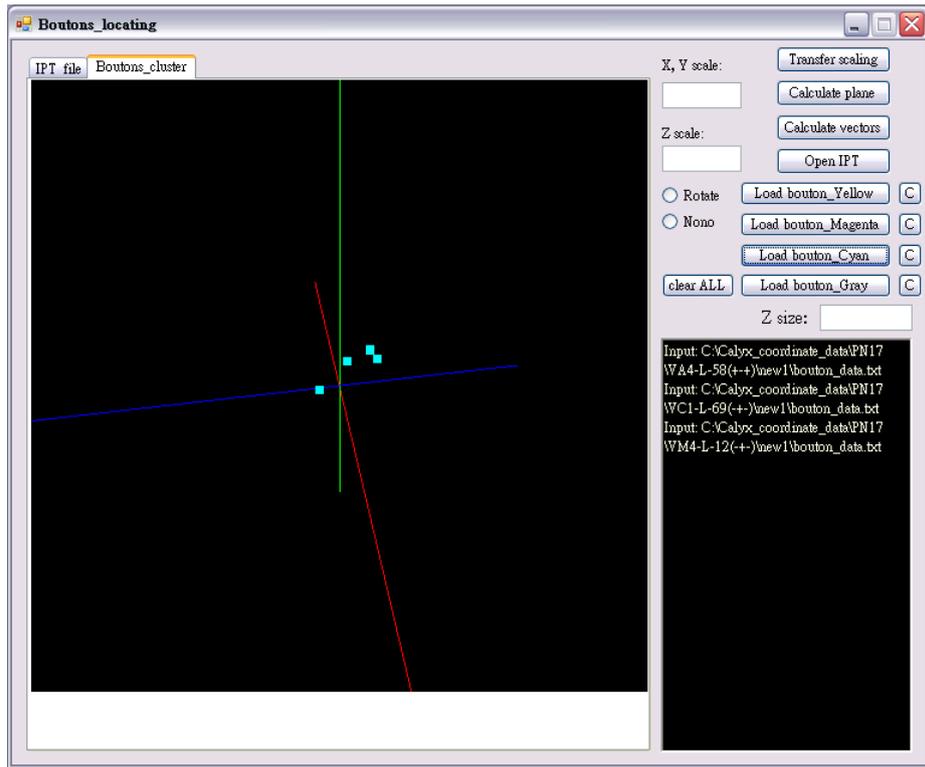
圖【4-14】VA7I-L-275&278



圖【4-15】VC1-L-69



圖【4-16】VC2-L-76&106



圖【4-17】VM4-L-12

上面一共有十七組，三十個 PN，我們可以發現大部分同組 PN 的 TBs 會分佈在相近的位置(圖【4-1】、圖【4-5】、圖【4-6】、圖【4-8】、圖【4-9】、圖【4-11】、圖【4-14】)，或是有類似的分佈結構(圖【4-6】、圖【4-7】、圖【4-8】、圖【4-9】、圖【4-11】、圖【4-16】)。但是其中有兩組 PN(圖【4-4】、圖【4-13】)做出來的分佈結果差異特別大，會造成這樣的差異有兩個可能原因：可能是這 PN 本身的生長曲線特別不同，造成實驗結果差異；或者可能是在一開始分組上就不該是同一個編號，這需要生物學者的進一步研究分析。

接下來我們將編號相近的 PN 聚集在一起觀察，我們利用原始資料的編號來分組。我們將 D 開頭和 V 開頭分開，再細分成下面兩組。

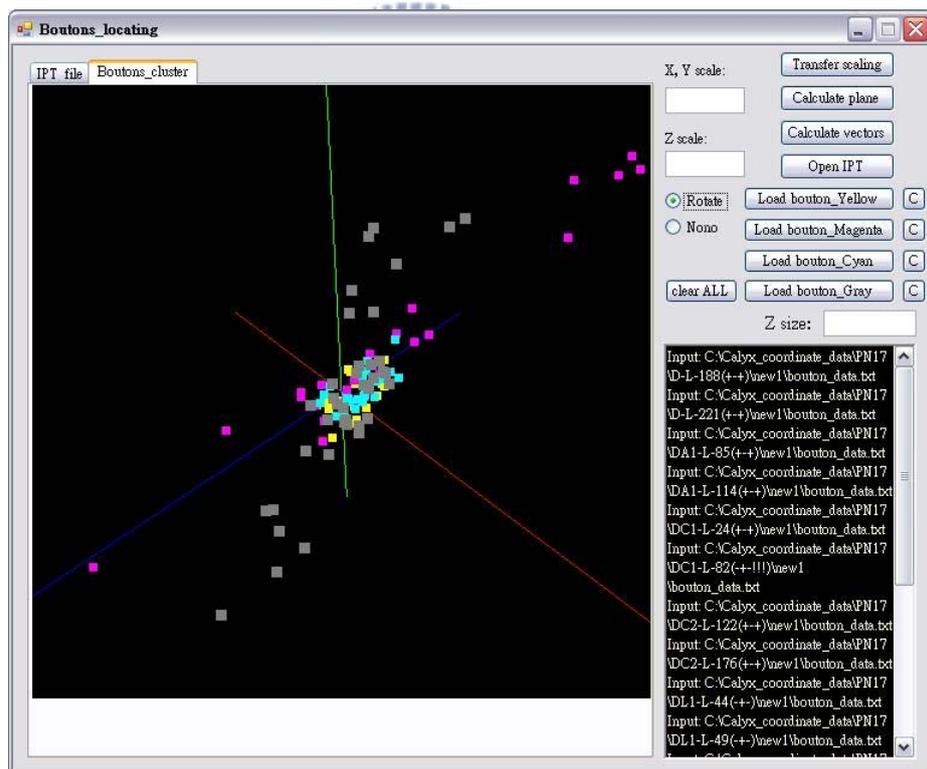
D 組：D-L-188, D-L-221, DA1-L-85, DA1-L-114(黃色).

DC1-L-24, DC1-L-82, DC2-L-122, DC2-L176(紅紫色).

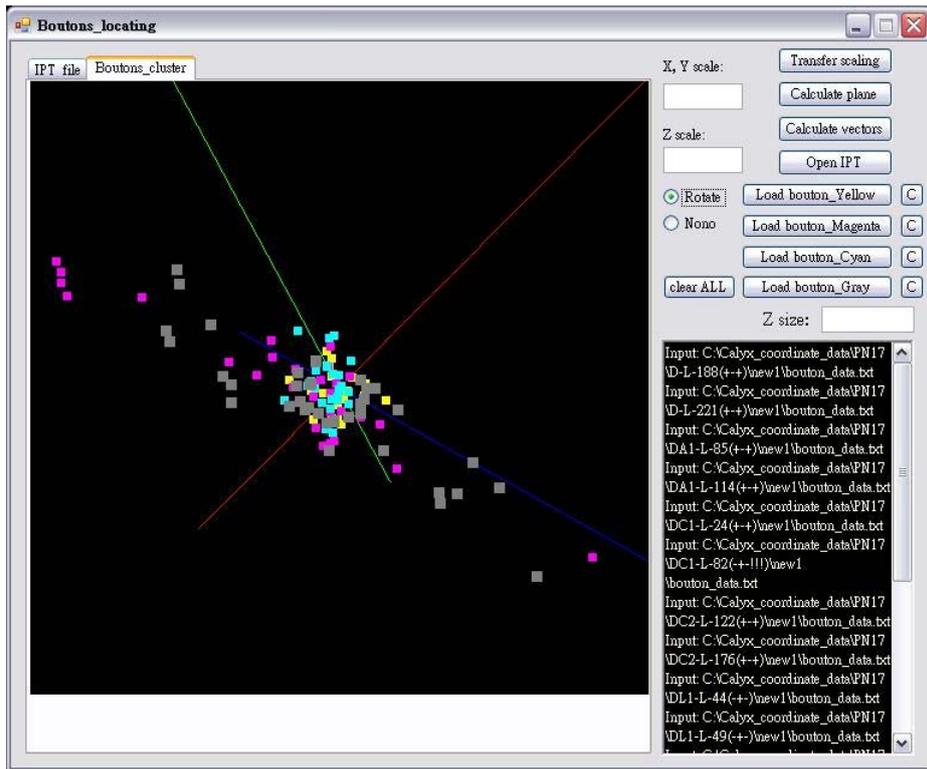
DL1-L-44, DL1-L-49, DL3-L-115, DL3-L-139(青綠色).

DM1-L-283, DM1-L-284, DM2-L-40, DM2-L-46(灰色)

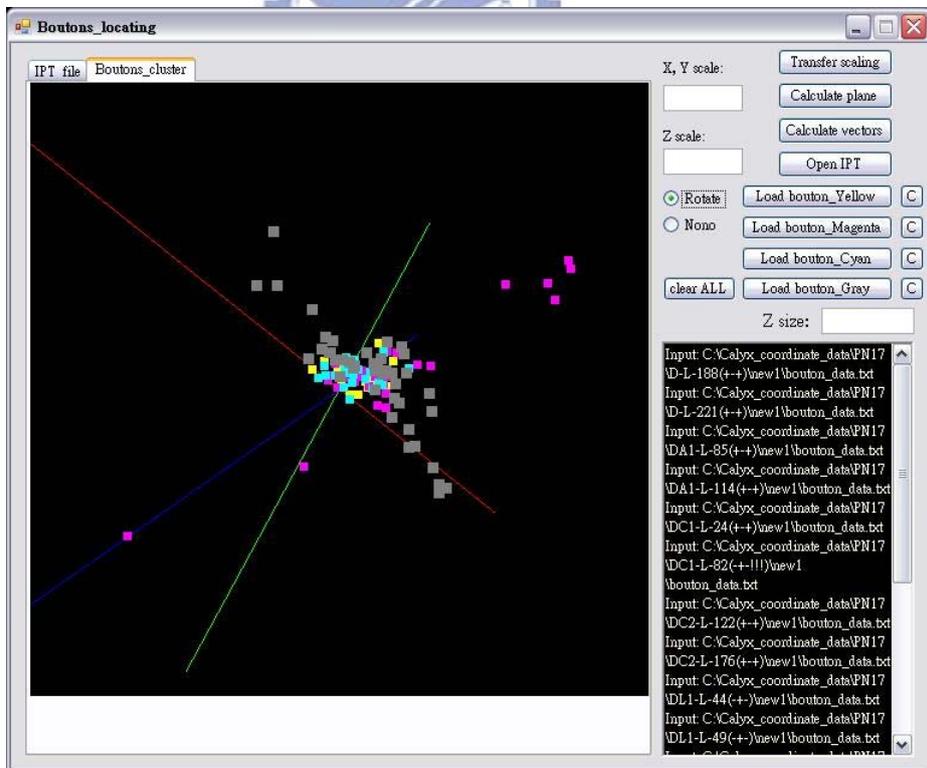
我們取三個角度來看看這些分佈：



圖【4-18】D 群組圖一



圖【4-19】D 群組圖二



圖【4-20】D 群組圖三

我們可以發現 D 組中的四個顏色分佈差異明顯，灰色有向外對稱發散的分佈；紅紫色則是其中有一個 PN 分佈特殊，比較像灰色的分佈法；而黃色和青綠色分佈範圍相對比較集中在原點附近。

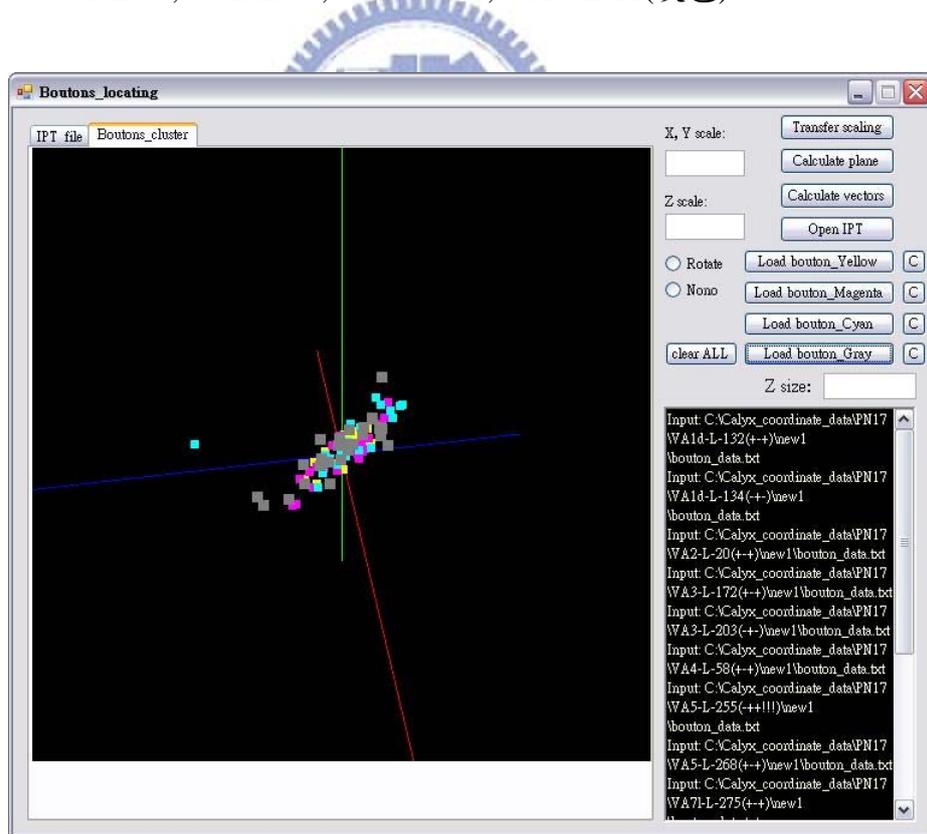
再來是編號 V 開頭的 14 個 PN 組合圖：

V 組：VA1d-L-132, VA1d-L-134, VA2-L-20(黃色).

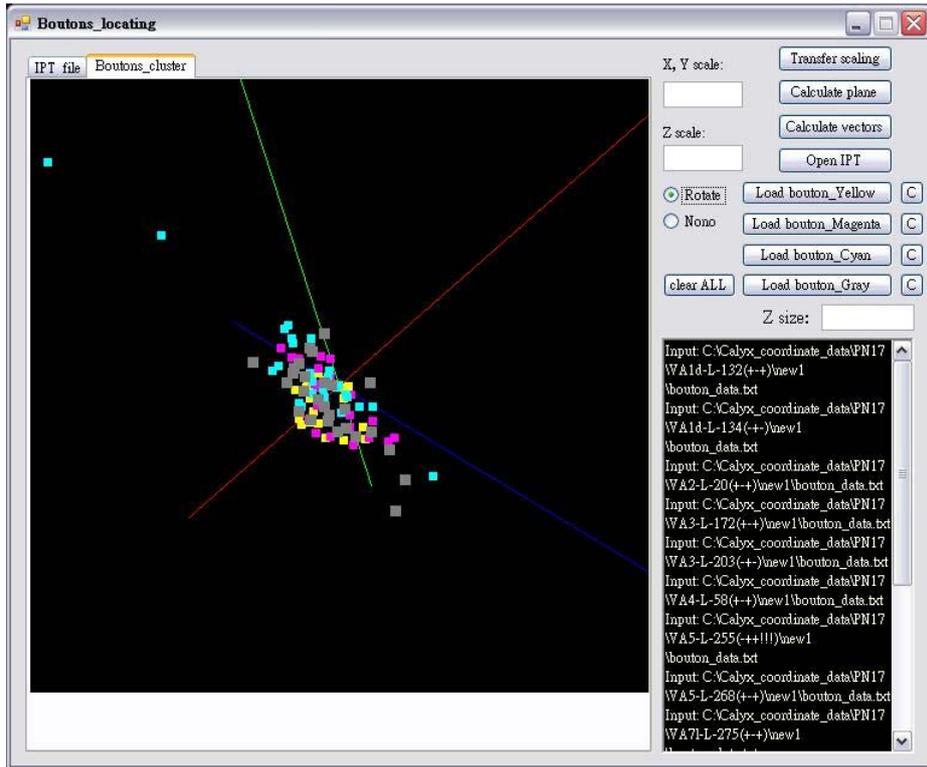
VA3-L-172, VA3-L-203, VA4-L-58(紅紫色).

VA5-L-255, VA5-L-268, VA7l-L-275, VA7l-L-278(青綠色).

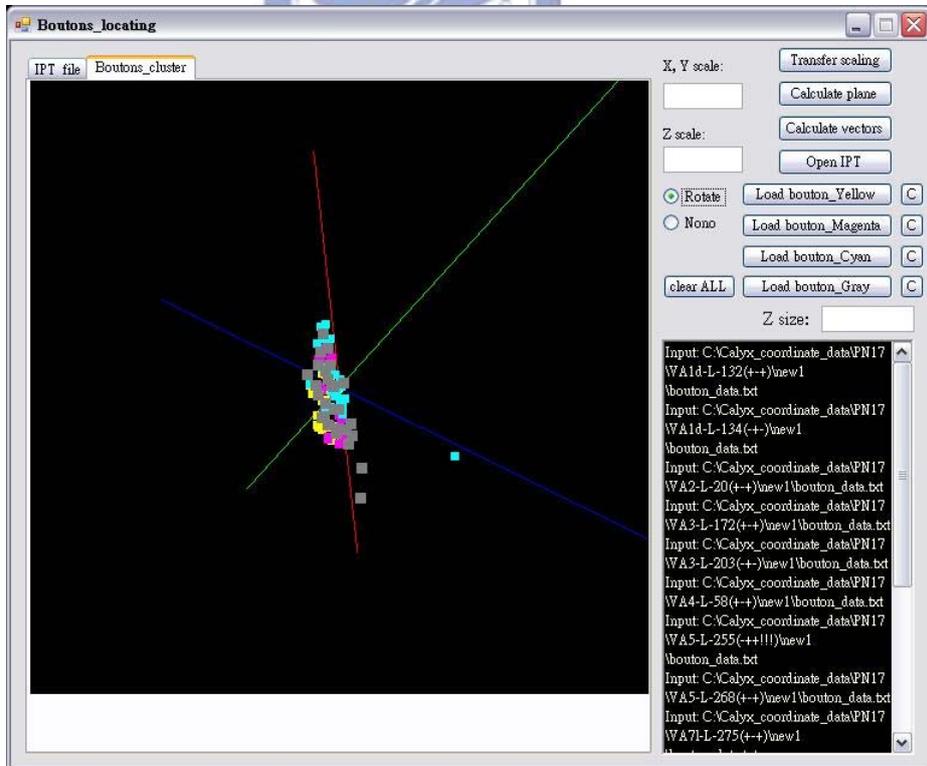
VC1-L-69, VC2-L-76, VC2-L-106, VM4-L-12(灰色).



圖【4-21】V 群組圖一



圖【4-22】V 群組圖二



圖【4-23】V 群組圖三

V 組的 TBs 可以看出分佈的範圍和密集度都比 D 組別更為明顯集中，其中黃色和紅紫色的分佈較集中於座標原點附近，結構也相似。而青綠色和灰色點分佈較為零散。

最後我們將這全部 17 組的 TBs 放在同一個座標系上，看看是不是有特殊的分布情形。因為我們的程式只有利用四種顏色點來表示 TBs，能區分的種類有限，我們還是按原始資料的編號來分成四組顯示，分組如下：

D-L-188, D-L-221, DA1-L-85, DA1-L-114, DC1-L-24, DC1-L-82, DC2-L-122, DC2-L-176(黃色).

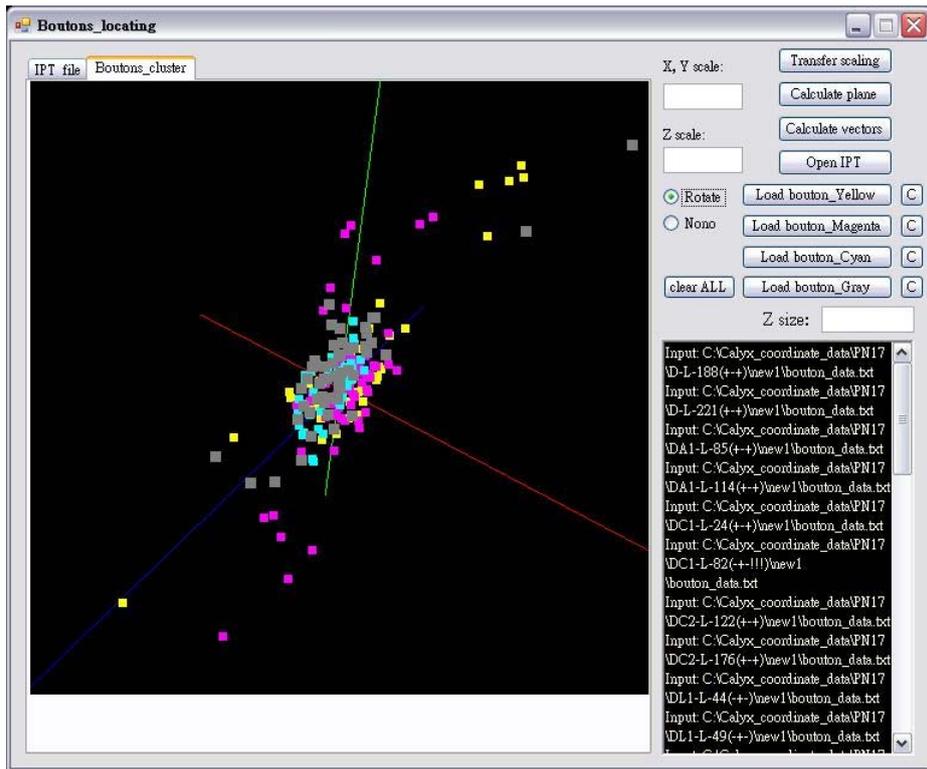
DL1-L-44, DL1-L-49, DL3-L-115, DL3-L-139, DM1-L-283, DM1-L-284, DM2-L-40, DM2-L-46(紅紫色).

VA1d-L-132, VA1d-L-134, VA2-L-20, VA3-L-172, VA3-L-203, VA4-L-58(青綠色).

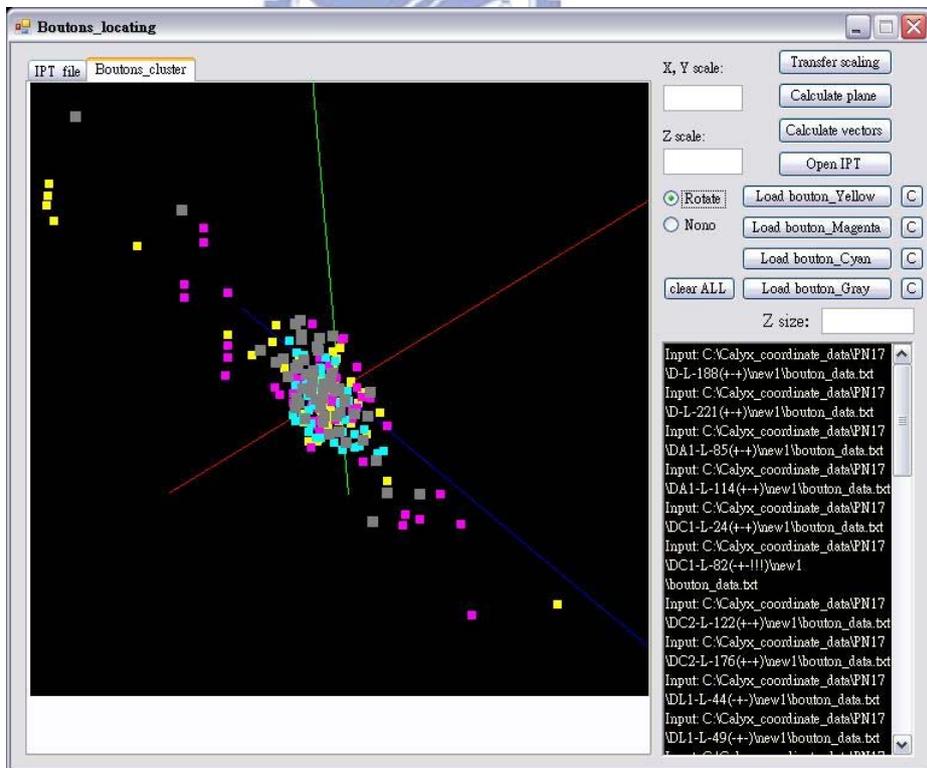
VA5-L-255, VA5-L-268, VA7I-L-275, VA7I-L-278, VC1-L-69, VC2-L-76, VC2-L-106, VM4-L-12(灰色).



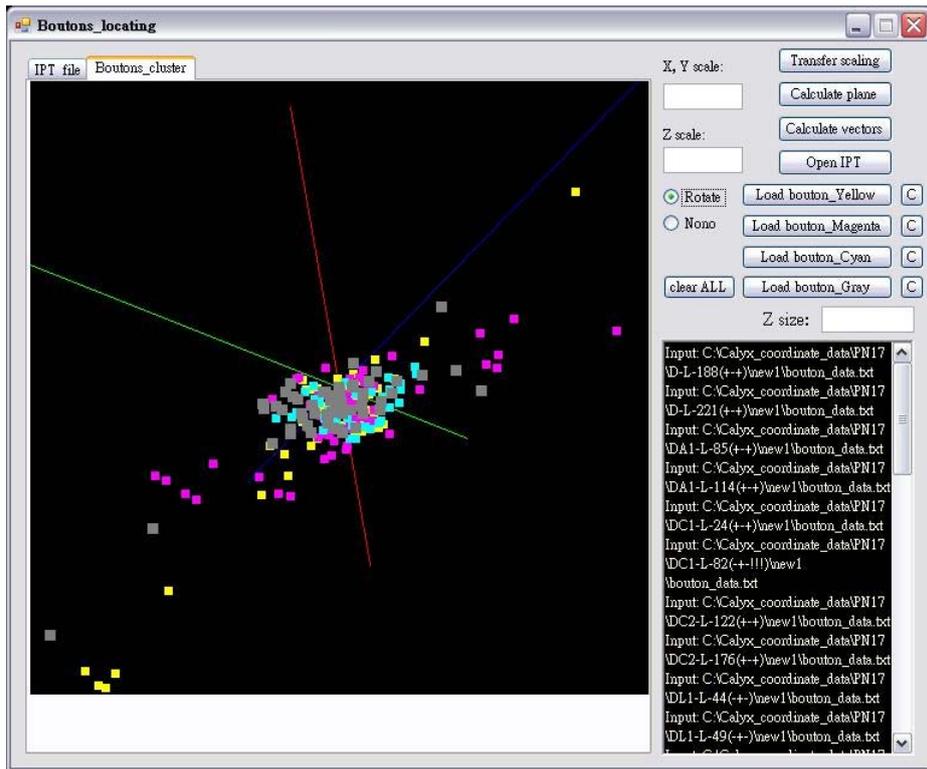
總共取五張比較全面的角度分佈圖：



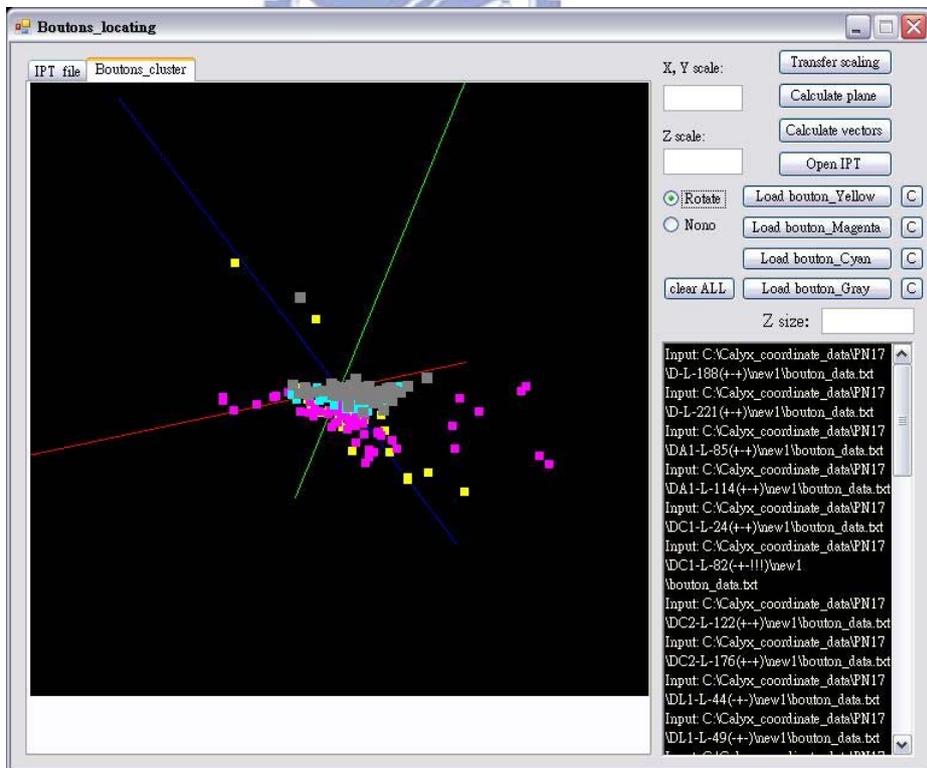
圖【4-24】All PN 圖一



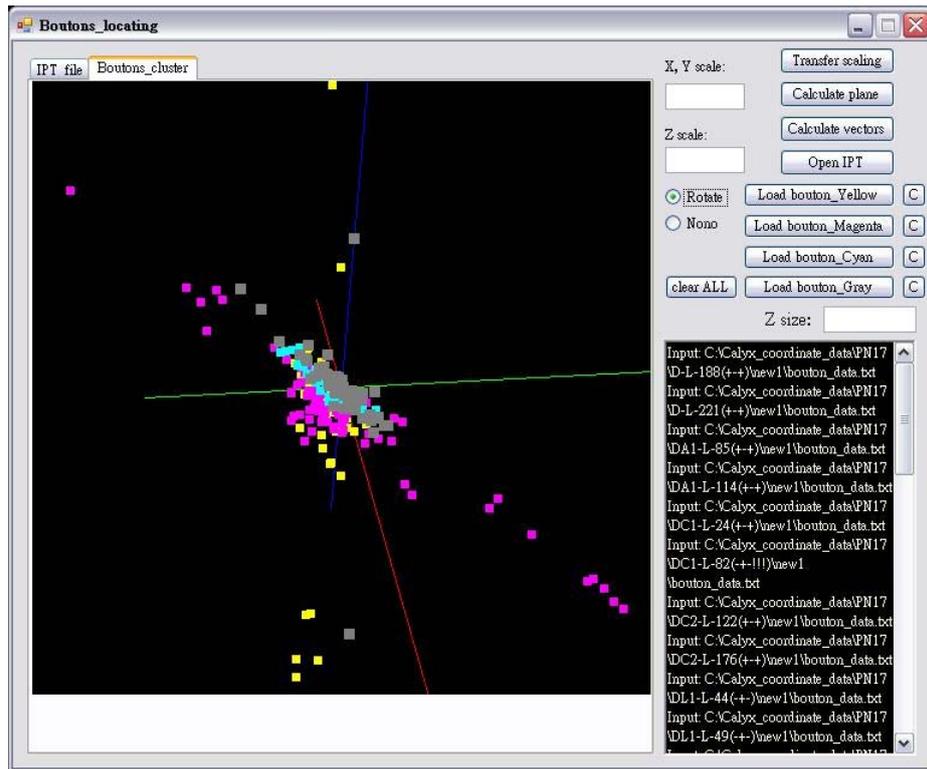
圖【4-25】All PN 圖二



圖【4-26】All PN 圖三



圖【4-27】All PN 圖四



圖【4-28】All PN 圖五

經由最後這組圖可以看出，有絕大部分的 TBs 在原點附近聚集成一團狀，這可能就是果蠅嗅覺 PN 在 MB calyx 中主要分佈的中心區域，青綠色和灰色的分佈大多密集在這個區域之中，甚至有分層的感覺。而紅紫色和部分黃色的分布主要在這區域的較邊緣的地帶，甚至伸出兩個明顯較遠的 TBs 區域。圖中有比較遠的獨立端點，照理說也是在 MB calyx 裡面，這可能是我們找尋的座標還不是這麼準確，但是不管座標的位置如何移動，這些相對孤立的點，還是會相當明顯，這也是比較特殊的 TBs 位置。

我們想出來的這個方法目前是可以將 TBs 經過一定的轉換，放到一樣的座標系中，是一種可能的分析 PN 的方法，這個方法也可以運用在轉換整條 PN 的定位上。在我們實驗的過程中，除了第一次啟動計算 fitting plane 的小程式要花十幾秒之外，其餘部分 CPU 計算都很快捷，不需要考慮。

第五章

未來展望

5.1 自動化

我們的實驗中有兩個部份是必須用人工介入的：一、初始 PN 資料的選取及取出目前是用半手動半自動的方式；二、嗅覺 PN 上的 TBs 目前是用手動選出。人工的介入會有幾個缺點，一來不同人的實驗結果會有誤差，而且實驗效率上也比較慢，未來我們需要的是自動化的萃取出 PN 座標和自動定位 TBs 的演算法，由於腦部資料的複雜度很高，要自動分析出想要的部位並不容易。目前我們實驗室也正在發展這兩個程式，完成後就可以將整個定位實驗自動化，並且大大減少實驗人為的誤差和時間。



5.2 延伸研究

想要知道嗅覺神經在果蠅腦中 MB calyx 的分布，其實整個 calyx 的區域和外型也十分重要，假如我們可以取出一個標準的 calyx 再將我們找到的 TBs 甚至整條 PN 放進去，可以有機會找出這些 PN 在 MB calyx 發生訊息傳遞的可能區域，進一步發現更後端的神經分類。所以如何找出標準的 MB calyx 是一個可以發展的題目。

再來因為大腦中神經的類型繁多，可以再分類的細目更多，我們的實驗中在呈現實驗結果時單用顏色區別，重疊部分會不太清楚，如何利用電腦圖學將這些資料清楚的呈現在一個三維的立體座標，讓生物學家可以進一步利用分析，這也是一個需要研究的地方。

我們利用電腦技術，分析出了許多腦部資料，這些資料如果可以整理後做成資料庫，以後有需要相關的資料可以方便大家從資料庫中搜尋、取出比對，也可以讓大家上傳資料，因為資料量很大而且格式也很多種，要做成資料庫也是個挑戰。這樣隨著資料庫的擴大，可以做神經的自動分類和比較，這樣才不會白白浪費之前所做過的努力。



參考文獻

- [1] Buck, L. and Axel, R., “Odorant Receptors and the Organization of the Olfactory System.” *Cell*. 1991; 65: 175-187.
- [2] Lin H. H. et al. “A Map of Olfactory Representation in the Drosophila Mushroom Body.” *Cell*. 2007; 128: 1205-1217.
- [3] <http://brc.life.nthu.edu.tw/index.html> FlyCircuit Database.
- [4] Hampel et al. “*Robust Statistics: The Approach Based on Influence Functions.*” Wiley, 1986, 2005.
- [5] Y. Nievergelt, “Total least squares: state of the art regression in numerical analysis,” *SIAM Review*, 36, 1994, pp. 258-264.
- [6] S. V. Huffel, P. Lemmerling, “*Total Least Squares and Errors-in-Variables Modeling: Analysis, Algorithms and Applications.*” Dordrecht, The Netherlands: Kluwer Academic Publishers, 2002.
- [7] C. Xu, J.L. Prince, “Snakes, shapes, and Gradient Vector Flow,” *IEEE Trans, on Image Process.* (March 1998) 359-369.
- [8] <http://www.mathworks.com/> The MathWorks, Inc.
- [9] de Groen, P. “An Introduction to Total Least Squares,” *Nieuw Archief voor Wiskunde, Vierde serie, deel 14*, 1996, pp. 237-253.

- [10] Golub, G. H., and van Loan, Ch. F. “An analysis of the total least squares problem,” *SIAM J. Number. Anal.*, vol. 17, 1980, pp. 883-893.
- [11] D.P. Berrar, W. Dubitzky, M. Granzow, eds. “Singular value decomposition and principal component analysis,” *A Practical Approach to Microarray Data Analysis*, Kluwer: Norwell, MA, 2003, pp. 91-109.
- [12] <http://web.neurobio.arizona.edu/Flybrain/html/index.html> FLYBRAIN.

