

# 通用水深資料編輯軟體 FREE 之研發

張逸中

私立致遠管理學院電腦與通訊學系 助理教授

## 摘要

水深資料處理過程中，尋找與剔除錯誤的原始實測資料是不能省略，也極為耗時費力的一項工作。目前較新穎的資料蒐集系統已開始配備有圖形化介面的編輯程式，可以檢視資料的空間分佈與連續性，並可在圖形介面上直接執行編輯工作，非常方便。但這類軟體多為特定系統而設計，格式的障礙使得類似的水深資料無法共享資源，十分不便。

本研究自行研發了一個可以通用於多種格式水深資料的編輯軟體 Format Recognizable Echo-sounding Editor(簡稱 FREE)。其主要特色是可以由使用者經過簡易的圖形介面操作定義自用檔案的格式，令程式辨識並將資訊儲存，爾後程式可以根據副檔名決定讀取方式，自動解讀已定義之所有格式，對於解決資料格式的障礙應有很大的幫助。此外，本軟體亦針對實務工作的需求，提供背景影像檢視、資料序列剖面檢視以及多種資料篩選工具，期能達到甚至超越一般商業軟體的功能與效能。此軟體目前定位為免費軟體(Freeware)，使用者可以免費合法自作者網站下載使用。

## 一、前言

水深資料處理過程中不能省略，也極為耗時費力的一項工作是：剔除藏身於數萬乃至數十萬正常測點之間的極少數錯誤資料。傳統上常用的方式是：繪製資料之連續剖面，由資料序列的連續性判斷是否異常，再切換到文字編輯器中進行實質編輯；這種方式必須不斷切換於文字與圖形繪製軟體之間，非常耗時也無法檢視資料在空間上的分布與連續性。目前少數資料蒐集系統已開始配備了圖形介面的編輯程式，可以檢視資料的空間分佈與連續性，並直接在圖形介面執行編輯工作，非常方便，但這類軟體多為特定資料系統而設計，格式上的障礙使得其用途受到限制。

本研究以筆者之前設計多項相關軟體的經驗，整合設計了一個可以通用於多種格式水深資料的編輯軟體 Format Recognizable Echo-sounding Editor(簡稱 FREE)。其主要特色是可以由使用者經過簡易的圖形介面操作定義格式，令程式『學習』辨識並將資訊儲存為檔案，爾後程式可以副檔名為依據，自動解讀已定義之所有格式，對於解決資料格式的障礙應有實質的幫助。

此外，本程式累積 Depth II [1,2]等類似軟體之使用者介面設計與使用經驗，其方便性與實用性方面已不遜於專業廠商提供之軟體；此次研發也特別著重於執行速度效能的

提升，在採用新的磁碟資料讀取模式以及繪圖方式之後，執行速度已獲得倍數的提升，對於處理數十萬筆以上的鉅量資料方便很多。本軟體以微軟公司最新版本之 Visual Studio 2005 [3,5]應用程式開發環境開發，在目前最常見的微軟視窗系統下可以獲得最佳的執行效能。

本軟體目前以免費軟體方式推廣使用，各單位可合法免費的自作者網站[4]取得軟體安裝檔；如有自行修改設計之需求，亦可函索原始程式碼。

## 二、資料格式分析器

所有水深資料結構都至少包含(X, Y, Z)三個部份，其中(X, Y)為來自定位系統的水平座標，如經緯度等等；Z為測深值。其餘附屬資料，如時間等等，與資料品質的檢測關係不大，因此我們的格式分析器(圖一)主要功能是建立(X, Y, Z)等三個欄位在資料錄(行)中的相對位置。首先選定需要分析的實質檔案，分析器會將其第一行的資料分別顯示於三個文字方塊中，使用者只需要分別用滑鼠分別選取所認知的 XYZ 三個欄位的資料即可。程式預設欄位是以字元的絕對位置為劃分基準，但是某些資料會以逗點或空白字元為劃分依據，此時可以用選項告知分析器，之後分析器會以第幾個欄位，而非字元位置來辨識資料。

這些辨識完成並儲存後，以後程式會依據所分析檔案的副檔名自動採取正確的辨識策略，使用者甚至不需再選擇檔案格式。格式資訊的數量可以增加到 50 個或者更多；副檔名的字數也可以使用 1~10 個字元，對於格式分類彈性更大。

## 三、資料處理流程

本軟體之內部資料處理流程如圖二所示。首先是從磁碟讀入檔案，在此有一個技術關鍵是程式以位元組陣列的方式，一次將整個檔案載入到記憶體內，而非使用一般的文字檔讀取方式，一次一行文字或一個變數。其差異是讀取資料時不必中斷磁碟的機械動作，停下來做資料解碼的工作。雖然目前的磁碟理論上可以一秒讀取數十 M 的資料，但事實上它們仍是具有慣性的機械動作，磁頭無法立即達到極速，斷續的磁碟讀取是非常緩慢的。以一般的數萬行水深資料檔來說，連續或逐行讀取，速度相差在十倍以上！掌握此項技巧是程式效能專業化與否的關鍵之一。

整檔讀取付出的設計代價是必須自行撰寫資料編碼的程式，本程式是先將資料在記憶體內由位元組陣列還原為字串資料；再使用前一節建立的資料格式資訊辨識 XYZ 資料，分別置入數值陣列，開始進行繪圖編輯等資料處理。處理過程中每筆 XYZ 資料有唯一的編碼，被刪除者以一個布林陣列加以標示，處理結束儲存資料時仍然回存最初讀取的，完整的字串資料，只是略過標示為刪除的資料行。

在資料繪圖技術方面，首先是地理座標與電腦繪圖座標(像素)之間的計算轉換，程式必須有相當的自動判斷能力，減少使用者的操作動作，譬如根據資料分佈、使用者的滑鼠操作決定適當的繪圖範圍等等；其次是資料水深值深淺的顏色表示，在本程式中每一次縮放顯示動作中，顏色都進行重新的最佳對比計算，也就是顯示中圖形的最深點顏色為純紅色，最淺為純藍色，如此可以讓資料異常點的顏色較為突出，易於鎖定並刪除。

另一個可以提升執行速度的重要繪圖技巧是充分使用記憶體背景繪圖，避免一般的前景繪圖，原因是一般的繪圖指令每完成一個動作會自動進行銀幕記憶體的複製更新，如果你要繪製數萬個點，就必須將記憶體區塊複製同樣的次數，相當耗時。相對的，記憶體繪圖在程式指令未指示貼圖複製之前不會自動做記憶體轉移，也就是可以全部劃完再一次轉移，根據實測這可以使繪圖加速約 3~5 倍！

#### **四、主要操作介面**

本程式主要操作介面如圖三，滑鼠操作的簡要說明直接寫在右側面板，包括：拖曳圖形、選取放大區域、資料點或資料剖面的點選等等；目前顯示的資料畫面在全區的相對位置繪製於右下角的導覽圖；顯示資料群的資訊在右側面板上方。值得特別注意的是右側中央的色階分佈柱狀圖，在資料群中沒有特別異常的太大或太小資料時，這個分布圖通常呈現常態分佈；相反的，如有異常值的時候分佈會偏向兩極，一端為多數正常的資料，另一端是異常的少數資料。此時在主畫面上多數的點顏色也會呈現全體偏藍或全體偏紅的景象，資料處理者因此不太可能忽略錯誤資料的存在，經由顏色的異常突顯，也很容易鎖定錯誤的位置，如圖四。

#### **五、時間序列剖面檢視**

在此類圖形化介面軟體出現之前，多數的資料除錯是在連續資料的剖面上進行，本程式也提供了類似的介面，如圖五。這個介面本身也是一個獨立的編輯物件，可以捲動檢視整個檔案的資料，並進行實質的編輯動作，編輯的結果也會同步顯示於主畫面。對於習慣傳統資料處理的熟練技術人員來說，這是個在處理經驗銜接上很好的設計。當然更好的功能是可以滑鼠在主畫面上點選任意點展示以那點為中心的剖面，使得時間序列資料與其空間分布可以有完美的切換檢視。

#### **六、資料篩選與稀釋**

除了異常值的編輯，常常我們也需要將資料依某些條件篩選，譬如只選某個區域或水深範圍的資料。在本軟體中的功能表內也附屬了許多此類介面，如圖六為 XYZ 資料篩選與空間稀釋資料的操作介面，後者可以選擇每個網格內只須擷取一定數量的實測點，以減少不必要的資料量。

## 七、文字編輯器

除了多樣化的圖形介面之外，相信隨時可以看到原始的文字資料也是許多資料處理者的心聲，本軟體也提供了完整的文字檢視與編輯介面，如圖七所示。與剖面檢視一樣，這也是可以實際進行編輯的工具，但是目前的版本必須回存磁碟之後再度開啟檔案方可與主畫面資料同步。

## 八、網格背景顯示

在實務上，測深工作的資料涵蓋密度常常是專業品質檢視的依據。眾所周知，一般的網格化軟體即使只用極少的實測點一樣可以產生網格資料、等深線圖甚至立體影像，但是其真實性必須從實測資料的分佈來檢驗，如果某些網格實際參考的實測值不在本身網格內，其可信度就不高了。因此本軟體設計了一個背景網格的檢視功能，使用者可以定義需要的網格長寬，通常也就是測量案必須達到的精度，然後程式會依據定義網格內的實測點平均深度顯示顏色，如果該網格範圍內並無實測點則保持背景色。

如圖三所示就是一個實例，可以清楚看到密集測量的測區內有一個網格方塊呈現背景的黑色，表示這裡沒有實測資料，這種空洞的有無多寡，可以讓資料處理者一眼就看出資料的概略品質；如果有連續的許多網格都缺少資料，如航線偏移造成的空窗地帶，也許就要安排補測了。

## 九、討論與結論

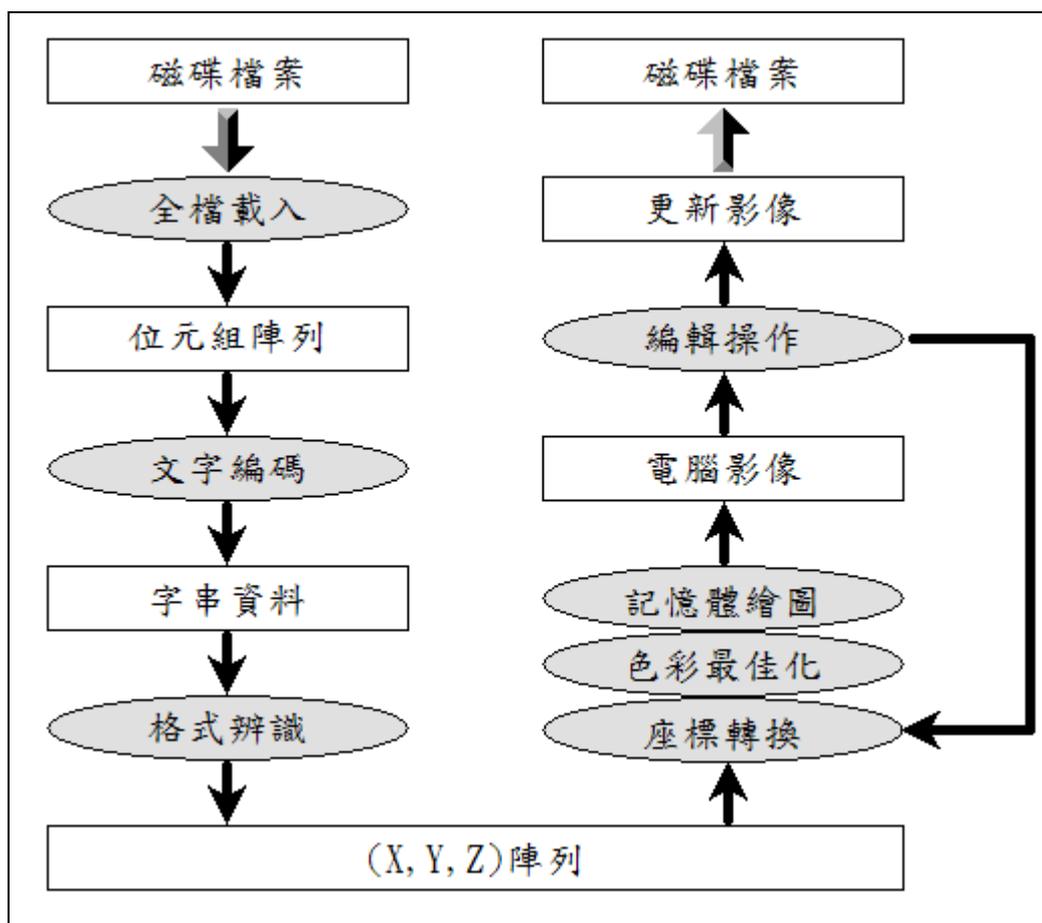
本研究是水下技術軟體實務研發的一個例子，隨著電腦科技的進步，測深工作依賴電腦軟硬體日深，專業軟體不僅是提昇工作品質效能所必須，也是提升整體水下科技研究所必須。目前國內測深相關軟硬體仍以外購為主，自行研發成果極少，因此可以自主改進工作品質與提昇研究水準的空間也受到限制，以長遠的角度分析，自行研發仍是必須嘗試的方向。筆者企盼本研究所得的技術經驗可供學界與業界參考，推動更多的本土技術研發。因此本軟體除了提供免費使用，如各單位有研發的需求，並可來函索取原始程式碼。

## 參考文獻

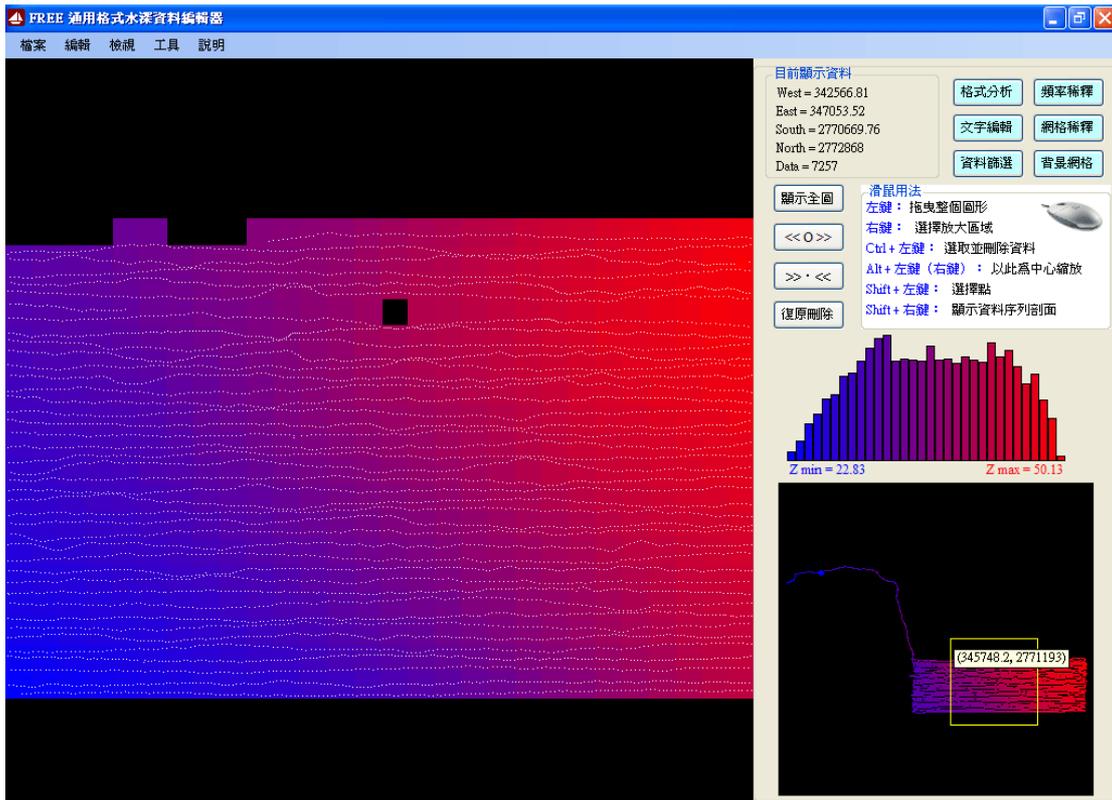
- 1.張逸中,2004,多音束水深資料編修軟體 Depth II，第六屆水下技術研討會暨國科會成果發表論文集，第三期 315-325 頁
- 2.張逸中,2003,即時測深與導航系統之研發，海洋技術季刊，第十三卷第三期 18-24 頁
- 3.Skibo, C., Young, M., Johnson, B., 2006, Working with Microsoft Visual Studio 2005. Microsoft Press.
4. <http://ycc.dwu.edu.tw/Research/FREE.htm>
5. <http://www.microsoft.com/taiwan/vstudio/express/>



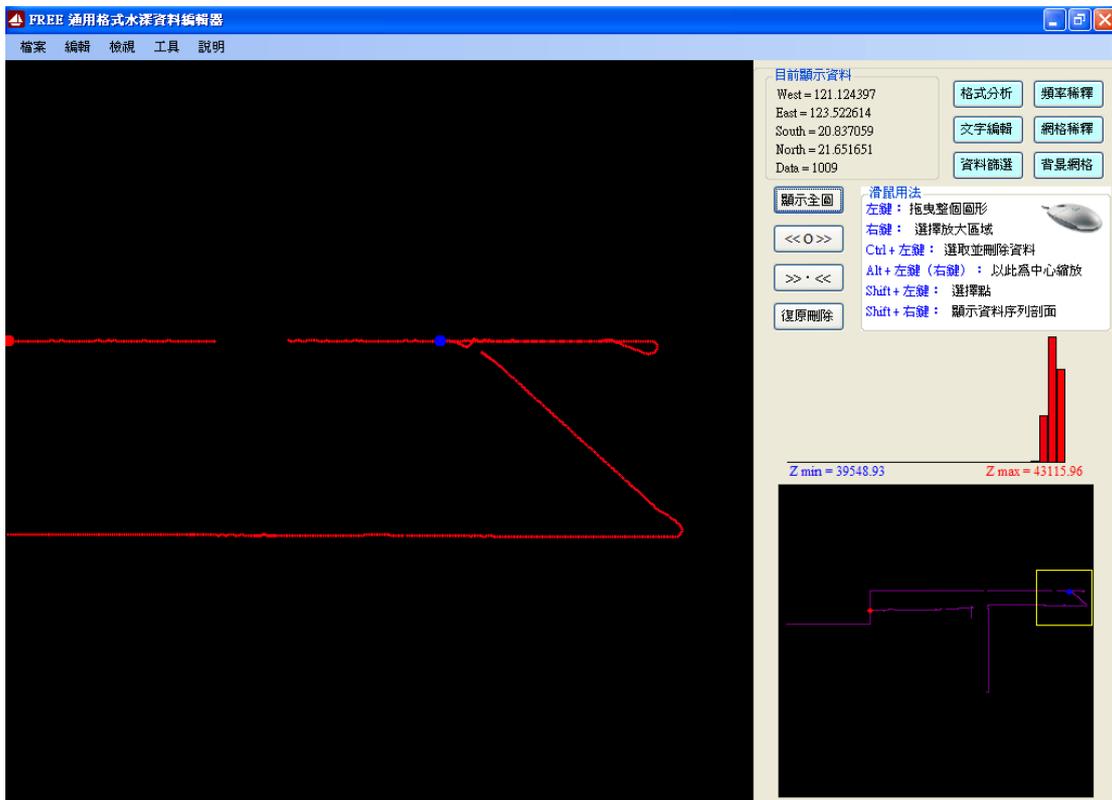
圖一、資料格式分析器



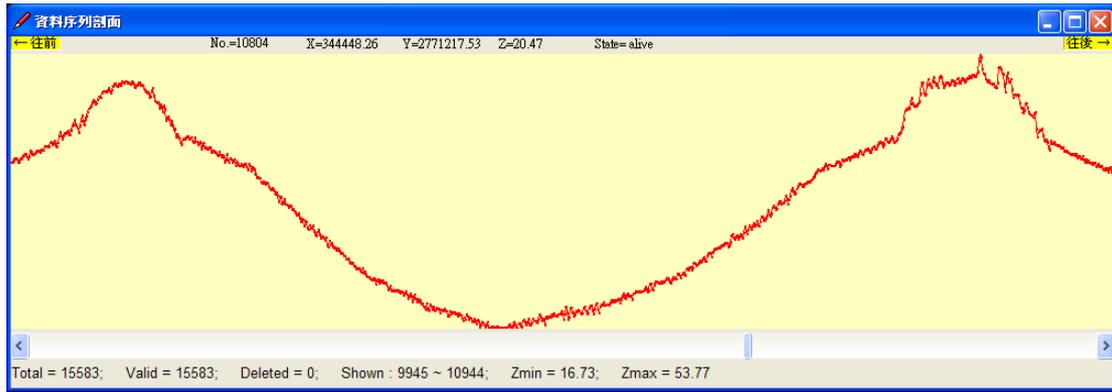
圖二、資料處理流程



圖三、程式操作主畫面



圖四、有過小的異常值存在時的景象



圖五、時間序列資料剖面

**資料篩選**

欄	條件
<input type="radio"/> X <input type="radio"/> Y <input checked="" type="radio"/> Z	<input type="radio"/> > <input type="radio"/> = <input type="radio"/> < <input checked="" type="radio"/> >= <input type="radio"/> <=

門檻值

**資料空間稀釋**

X網格式距 =

Y網格式距 =

每網格式資料最多

圖六、資料篩選介面

**文字編輯器**

檔案 編輯

346726.650	2770678.060	-46.95
346719.970	2770678.980	-46.97
346712.860	2770678.780	-46.75
346705.840	2770679.230	-46.70
346699.290	2770679.250	-46.73
346692.290	2770678.480	-46.90
346684.740	2770678.470	-46.73
346678.060	2770679.060	-46.27
346671.790	2770679.290	-46.45
346664.870	2770679.410	-46.57
346657.420	2770679.960	-46.30
346650.430	2770680.730	-46.13
346643.840	2770681.340	-46.17