

# 數位化筆記辨識管理系統

徐豐明、張家源、林銘哲、王守鈿、朱孝祥

朝陽科技大學資訊工程系

413 台中縣霧峰鄉吉峰東路 168 號

Tel: (04) 23323000 ext. 7633

Fax: (04) 23742375

E-mail: csie@mail.cyut.edu.tw

## 摘要

利用電腦強大的運算能力協助處理文件影像資料是當前最重要的電腦應用研究領域之一。許多針對字元的辨識方面的技術已被提出，但是這些技術要能有效的發揮必須有完善的處理，才能確保辨識的正確性，再加上有效的管理系統，使用者日後查詢便能有效提升閱讀以及工作效率，以此為目標，於是我們設計出可辨識又可以管理文件的數位化筆記辨識管理系統之研究。研究方面上，以文件的處理可以分成幾個部分：影像雜訊的處理，字元的分割，再透過各種字型的樣板去做比對，就能有效的辨識出我們所要的字元。

## Abstract

Use of the powerful computational capability to assist in handling computer data is the most important document imaging applications in the areas of computer PIH. Many view the identification of technical character has been proposed, However, these technologies must be able to function effectively improve the handling. To ensure the correctness of literacy. In the effective management system. Users future inquiries can effectively promote reading and Efficient, This objective. So we can design recognizable notebook digital identification document management systems management research. On the Research, a document processing can be divided into several sections: Image processing noise and Character segmentation. The model than to do it through a variety of fonts. We are effectively differentiate the characters.

## Keywords:

literacy, recognizable notebook digital identification document management systems, effective, Image processing noise.

## 1. 簡介

隨著時代的進步，影像處理的技術廣泛應用於日常生活當中，例如運用在文字辨識的部分，可藉由影像處理的技巧與辨識演算法，來處理含有文字影像的圖片並加以紀錄分析，可方便應用於文書處理中，利用辨識出的字元加以編輯，就可省去自行輸入文字的時間，可加速使用者完成的速度。文字

辨識系統可方便含有文字內容的圖檔的需求者，一個成功的文字辨識系統需要具備兩個條件，一是快速的辨識速度，另一個就是良好的辨識率，想要辨識出圖檔中的文字，必須先定位出文字的所在位置，在含有文字的圖片中找出文字的位置，然後再進行圖片中的文字字元切割，將文字的每個字元分割成個別獨立的字元出來，再來是字元辨識的部分，我們所採用的是樣版比對的辨識方法，將字元一一的辨識出來，並將辨識結果輸出。

## 2. 系統架構

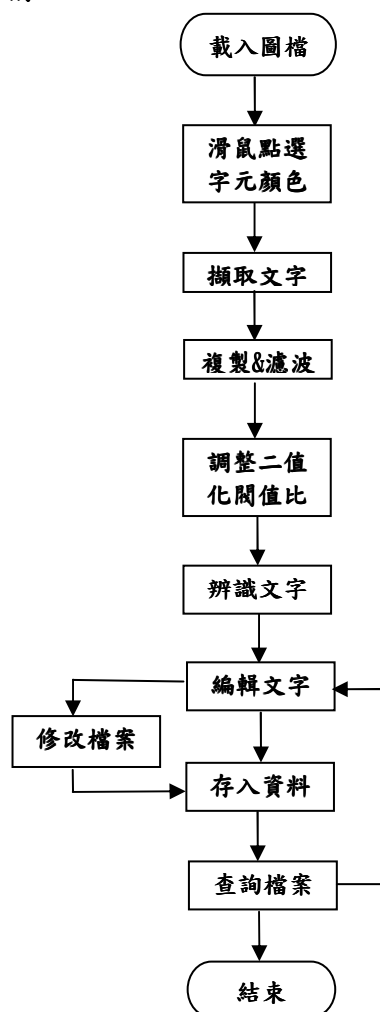


圖 1. 系統架構圖

### 3.系統功能

就整個文字辨識系統流程而言，

- (1) 影像前處理
- (2) 文字切割
- (3) 文字辨識
- (4) 資料庫整合

### 4.功能分析

#### 4.1.影像前處理

在正式辨識文字前，我們必須先針對取得的文字先做前處理的動作，這是由於原始圖檔大多包涵許多雜訊及干擾辨識的因素，因此，必須在正式辨識前，先行排除這些因素才行。

##### 4.1.1 灰階

當取得預辨識文字區塊有色彩時，可能會因為文字顏色和背景顏色相近而產生辨識誤差，諸如文字辨識錯誤、文字定位錯誤等誤差，所以當進行文字辨識前，必須先將預辨識影像進行灰階化前處理，使得辨識區域變成同一色系，減少辨識誤差率，而使用的灰階公式如下所示，將彩色影像轉化為強度為 0~255 灰階影像。

$$Y=0.299R+0.587G+0.114B$$

##### 4.1.2 二值化

在處理完彩色影像辨識問題後，預辨識文字區域會轉變成 0~255 色階的灰階影像，此時為使文字與背景分離出來必須採用二值化的動作，將字元與背景分離出來，也就是將之前得到預辨識的灰階文字區塊，經過二值化處理後，產生一個含文字的二值化影像，而採用的二值化演算法則是把個別的像素歸類為文字或背景。

二值化的動作也必須考慮介於區域化和辨識率之間的差距縮小的問題。在做文字擷取時，最後的階段就是將文字截取出來的結果，在進行文字識別的動作。而有關於文字識別 (OCR)，通常是直接使用現有的 OCR 演算法來解決影像識別的問題，就可以得到很好的效果，但是大多數的 OCR 演算法在處理文字識別時，多會是期望所輸入的資料影像，文字的筆劃是黑色的，而背景是白色的。這樣 OCR 影像辨識後，得到的結果會是最好的。所以在這裡我們就需要對於圖像作二值化前處理的動作，將彩色的影像轉為黑白的影像，而其中文字的筆劃會被強制轉換成黑色的，且背景轉換成白色的。所以演算法的目標就是朝這個方向前進。

二值化處理主要是將 256 灰階色彩轉換成二值 (不是黑即為白) 影像的工作，所有低於臨界值 (本系統將臨界值設為中間值 128) 的像素設定為黑色，而大於等於臨界值的像素都設定為白色，二值化影像最主要的功能是區分出影像中的物件與背景，以便萃取中物件影像資訊。

原始文件影像的灰階分布情形，大致以雙峰狀態呈現，只要適當的選取可以分離兩個區間的臨界值，便可以使圖文與背景分離。在大多數的圖文分類的研究報告中，我們發現圖形影像的二值化臨界值的選取相當重要，如果選取的臨界值不理想，那麼結果和預期目標就會有相當大的差異。例如：所選取的臨界值過大，影像經過二值化後，所呈現出來的影像會將無關聯的影像合併在一起；反之，選取的臨界值過小，便會造成影像的支離破碎，均會影像分類的準確率，因此在二值化技術的選擇上，應謹慎挑選。

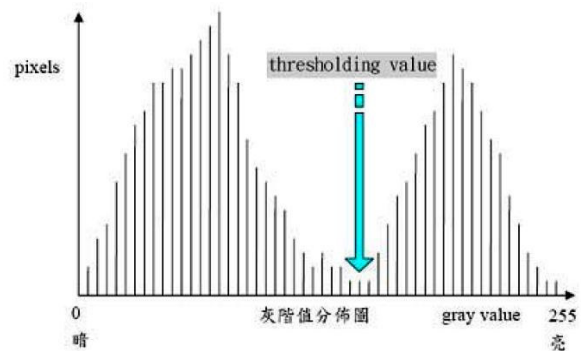


圖 2.灰階化之色階分布圖

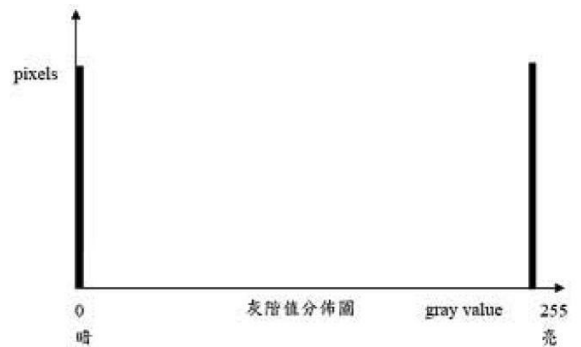


圖 3.二值化之色階分布圖

一張灰階影像利用臨界值來轉變成二值影像的方法，常見的臨界值取得方法是利用灰階總數直方圖 (Histogram) 找出適合的臨界值，如果像數強度大於此臨界值設為 255，小於等於此臨界值的像數則設為 0。其取得臨界值的方法有很多種類，臨界值最主要的目的是將正確區域與背景影像轉變成二值影像，可增加文字定位之正確率。而本專題採用較簡單和快速的平均值二值化法來進行。此方法中的臨界值取得是將各像素點的強度累加除以像素點個數可得強度介於 0~255 的臨界值，可利用此臨界值來進行二值化處理，進而初步去除非辨識區域。

## 4.2.文字切割

再進行完前處理後，此時待辨識文字以濾除大部分的雜訊，接下來則是要利用演算法，將文字中二值化後的白色及黑色畫素分別存成 0、1，方便系統將排列組合存成陣列以便做為辨識之用。

### 4.2.1 相連元件演算法

自從 1988 年 Lloyd Alan Fletcher 等人在針對圖文影像分類方面的研究中運用相連元件的方法後，在圖文影像分類低階處理方面大多使用這種概念。所謂相連元件 (connected component) 就是圖形經過二值化 (Binarization) 的轉換後，在二維矩陣內的像素質均轉換為只有 '0' 和 '255'，「0」代表黑色像素值，「255」代表白色像素值，接著運用 Nearest Neighbor Rule 的方法；當掃描到物件的第一個黑色像素值時，便會拜訪與它相鄰的八個鄰居 [8-neighbors]，凡是鄰居中有 '0' 的像素值時，均會予像素值連接，依此類推，即可將所有相連且為 '0' 的像素均連接在一起，如此相連完成後，每一個群組就是一個獨立物件。

所以我們在連接各黑色像素點方面就採用八相連 [8-connectivity] 的方式，以下圖 4 為例。

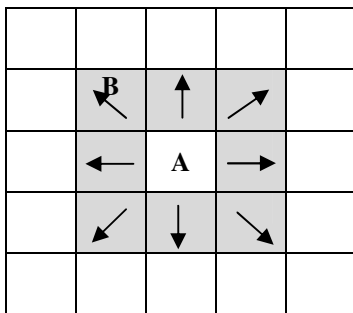


圖 4.八相連示意圖

其中 A 為我們目前掃描到的像素點，且 B 點可以在 A 點的上下左右或對角相鄰的位置上 (即圖 4 灰色部分)，假如 A 與 B 顏色相同，則判定 A 與 B 相連，如此重複步驟，即可將辨識字元轉換成 "0" "1" 矩陣圖。

### 4.2.2 文字定位

文字定位則是和相連元件演算法相輔相承，而文字定位的功能部份，則是採用掃描逐一文字中從上到下，從左到右的像素，直至找到第一個黑點為止，並把黑點位置存置陣列之中，再利用相連原件演算法的原理尋找八方位的鄰居，逐一搜尋，直至找不到黑點時，此時會再從下至上的再次掃描一次，然後將所取得的文字區域，以淺綠色的框線做為文字定位，再將框線內所得到的 "0" "1" 矩陣排列方式，送入系統中，做為樣板辨識。

### 4.2.3 文字正規化

經過影像前處理及文字定位後，我們再把定位

的文字區塊做正規化的處理，正規化的目的是為了使得定位的文字區塊能和文字樣版做點對點的一一比對，所以必須先將個別單一文字分別做成我們所需求的文字固定大小 (本系統的正規大小即為 32\*32，參考圖 5、圖 6)，以便於和文字樣版做比對。

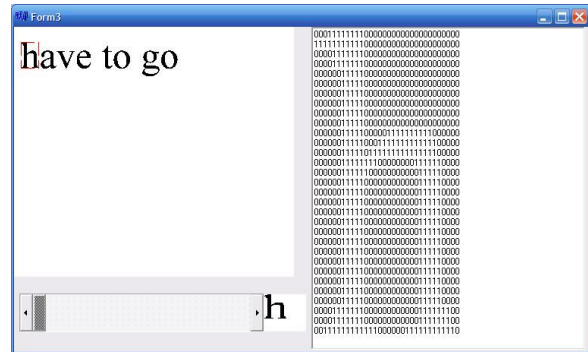


圖 5.秀出辨識後的首字元

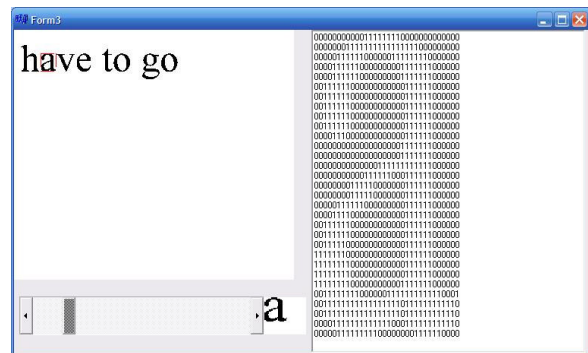


圖 6.秀出辨識後的次字元

## 4.3.文字辨識

就文字辨識而言，複雜的辨識方法雖然可以提升辨識率，但相對的也犧牲了辨識速度，對一個應用於文字辨識系統而言，除了辨識率外，辨識速度也是系統考量重要的因素之一。

本專題所研究的字元辨識法，希望具有低運算量與高辨識率的特性。在辨識率與辨識速度的雙重考量下，所以採用樣版比對法與相連元件演算法，以及文字定位來進行對字元的辨識，如此兼具速度與辨識率的系統，在實際的應用上必能達到以最低的硬體成本達到高效率的目的。

### 4.3.1 樣板比對法

樣板比對顧名思義即待辨圖樣與樣板間以某種相似性量測來決定像的程度。就幾何的觀點而言，相似性可視為兩個圖樣在特徵空間中的距離，常用者如：歐幾里得距離 (Euclidean metric)、餘弦距離量測 (cosine metric)、漢明距離量測 (Hamming metric)、絕對誤差量測。樣板比對法最大的優點在於辨識速度快，而其缺點為缺乏適應性只能辨識固定的字型，因此，本專題在系統規劃上，將樣板的擴充性也納入了考量，使得系統開發者可隨時隨地

加入新增的樣板以提高辨識率，所以可利用改良式的樣板比對法來進行文字的辨識進而達到有效的文字辨識。

樣板比對法的架構是將預先準備好的每個文字模板，把輸入的未知文字區塊與先前建立好的文字模板進行畫素重疊，比較兩張圖案在相同位置上畫素之異同。進而將相似性最高的文字模板輸出，即為我們要的答案。

而本專題採用的方式則是將之前利用相連文件辨識法，所輸出的代表黑色畫素的值轉成 1、代表白色畫素的值轉成 0，將文字轉成 32\*32 的文字樣版，存入系統指定的文字檔中，而當預辨識文字區塊經由先前影像前處理及文字定位及切割後，再至先前建立的文字樣版中，進行點對點的比對，把比對結果最接近的字元，在進行輸出。

### 4.3.2 文字排序

文字排序的功用在於，當使用者需要辨識兩行以上的文句時，文字的輸出會由於文字定位法的關係，會使得文字先由上至下、後由左至右的掃描，造成直排文字會先行輸出的錯誤，所以當遇到此情況時，就必須加入一道文字排序的功能程序，以防止文字在輸出時，發生排列錯誤，而文字排序的原理，請參考下圖所示。

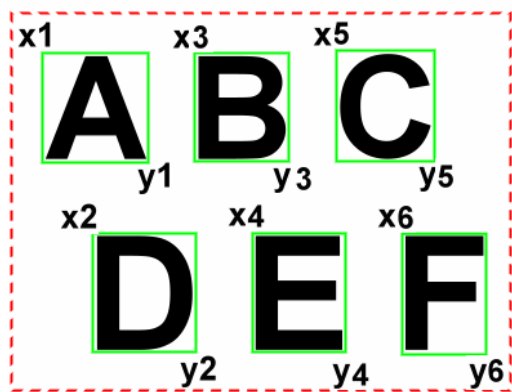


圖 7.文字排序示意圖

利用先前使用的相連元件演算法，由陣列中取出各像素點的座標值，並從中找出每字元的第一點黑色像素點座標位置，及最末點黑色像素點座標位置，然後再進行比較，以上圖 A 及 D 為例，因為 x1 與 x2、y1 與 y2 的 X 座標相近，且 x2 及 y2 的 Y 座標遠大於 x1 及 y1，所以我們可以知道 A 會在 D 的上面，以相同觀點來看 A 及 B，因為 x1 及 x3、y1 及 y3 的 Y 座標相同，且同時 x3 及 y3 的 X 座標會遠大於 x1 及 y1，所以我們可以確定 A 會在 B 的右邊，如此反覆動作可將所有文字輸出排序完成。

### 4.4.資料庫整合

在成功辨識完我們所需辨識之文字後，可以進入編輯區修改，編輯區中可從事基本的 Wordpad 功

能，例如字體轉換、字體縮放，字體粗細...等，也可蒐尋文章中的字句，並取代之。再做最後的修改及補充後，接著便可以存入資料庫中。

而資料庫最大的功用，就是能夠便於使用者整理辨識後的大量資料，達到筆記管理系統的功用，當日後辨識資料越來越多後，使用者就可以利用簡易的日期、檔名，或者關鍵字檢索，就可迅速找到需要的資料，當然也可以將資料再度修改、刪除以及列印等，如此結合資料庫系統，更能讓文字辨識系統發揮出更高效用。

## 5.系統展示

以下以圖示的方式，將程式做功能展示

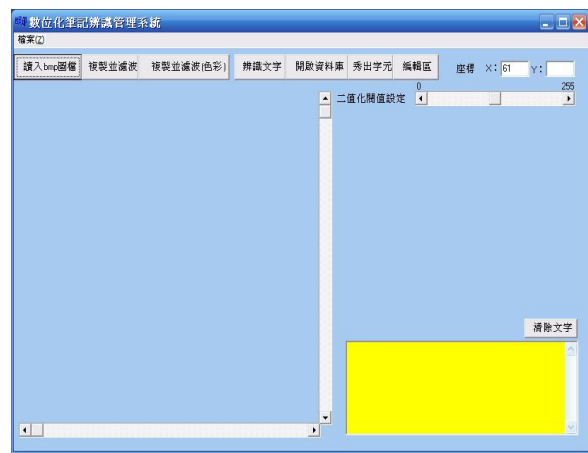


圖 8.程式主畫面

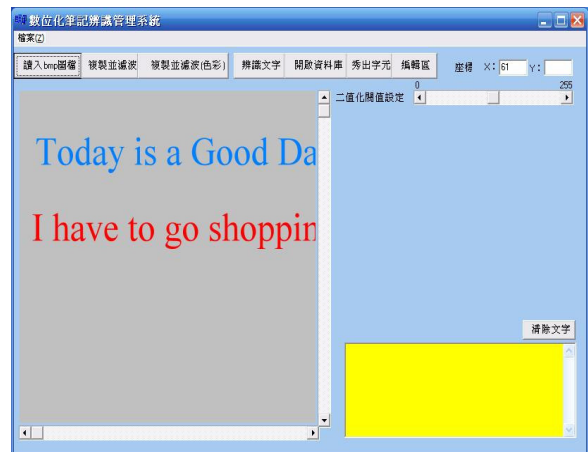


圖 9.載入圖檔

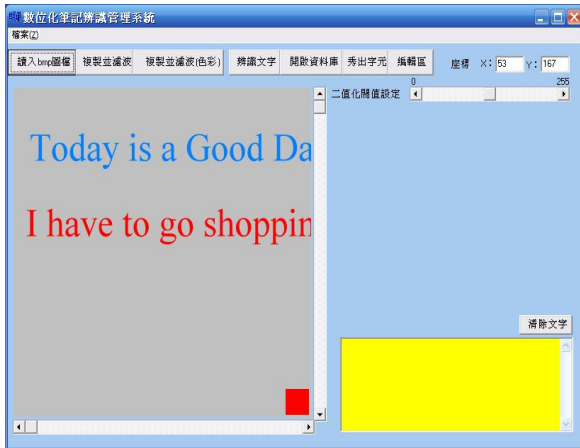


圖 10.點選預辨識的文字色彩

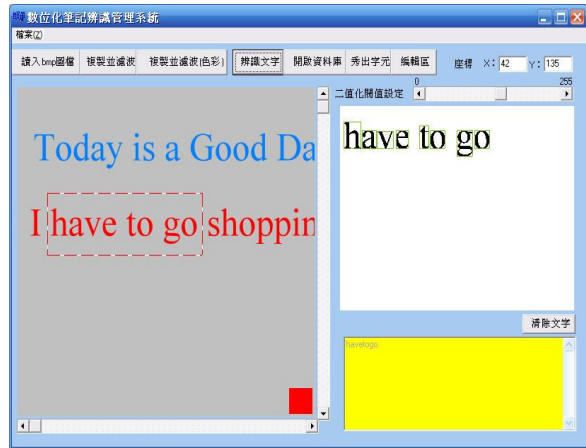


圖 13.經過文字定位及切割後，辨識出的可編輯文字

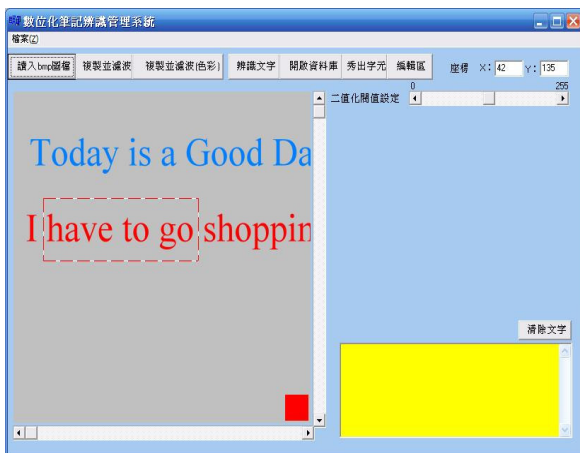


圖 11.圈選預辨識文字區塊，調整二值化閾值

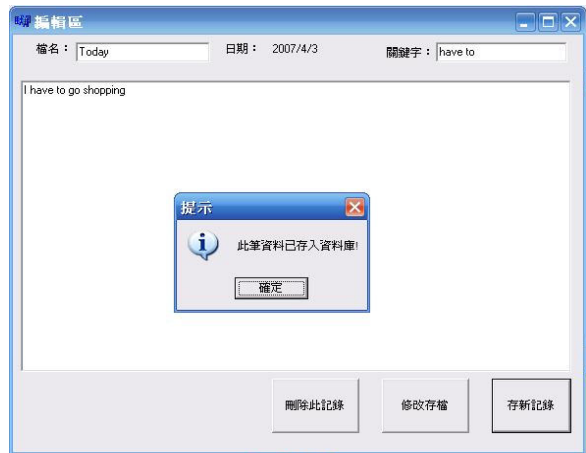


圖 14.打開編輯區，做修改刪減並存入資料庫

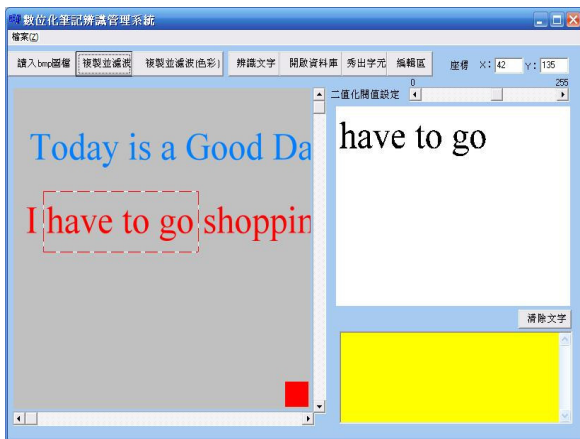


圖 12.顯示影像處理後的文字區塊



圖 15.檢視資料庫中的資料，以便使用者存取舊資料做修改

## 6.成果檢討

此專題已完成預定的功能要求，但還有一些地方需要再加強克服的地方，才能使系統更加的完備，在此分為二部分，分別討論本專題達到了那些功能，以及尚有那些問題需要再解決的。

## 6.1.已達到的功能

- (1) 可分別針對彩色圖像及黑白圖像中的文字，進行截取並辨識。
- (2) 能有效將辨識結果，做修改及刪減的文書工作。
- (3) 能將處理完成的文件存入資料庫，有效於日後的查閱及管理。

## 6.2.需要再克服的地方

- (1) 文字辨識率能提升部分仍有限，主要受限於樣板的寡，但過多的樣板，除了容易使辨識一些近似字容易混淆外，也會造成系統的負擔。
- (2) 對於輸入較不清晰的圖形，無法有效去除雜訊，以及在過小文字上的判定，容易產生錯誤，而使得辨識率無法有效提升。
- (3) 對於輸入大量的文字，無法進行文字轉換，使得傳換出的文字會有些許誤差，仍是需改進的地方。

## 7.結論與期許

在製作專題途中，曾遭遇到許多困難，在蒐集資料方面的困難，以及在編寫 BCB 方面，重新的學習，使得在一開始接觸時，備感棘手及不適應，繼而從中摸索學習，而漸漸達成預定的基本成果，雖然程式辨識結果尚未完美，但在製作專題初期時就認定文字辨識方面仍有所發展進步的空間，所以便想要去研究影像文字辨識的領域。

首先，在辨識度上相信應該仍可有所提升，原開發目的就是為了解決使用者在字字輸入的麻煩上而開發的，假使辨識度提升才能真正的方便到使用者的實用性，對於研究這次的影像文字辨識，使得我們對這領域有相當程度的認知，過程中學習到許多原理及 BCB 程式開發技術，也相當感謝豐明老師在專題中的指導，使我們在遭遇困難時有所方向，受益良多。

## 8.參考文獻

- (1) .李勁 著，精通 C++ Builder 6，文魁出版，2003 年 4 月。
- (2) .黃文吉 著，C++ Builder 與影像處理，儒林出版，2005 年 10 月。
- (3) .羅文圳 著，文字辨識系統之研究，專題研究報告。
- (4) .陳周造 著，Borland C++ Builder 入門與應用，博碩出版，1997 年。
- (5) .井上誠喜 著，C 語言數位影像處理，全華出版，2005 年 4 月。
- (6) .連國珍 著，數位影像處理，儒林出版，2002 年。
- (7) .洪國勝，張建原，洪月裡 著，Borland C++ 程式設計快樂上手，松崗出版，2001 年 2 月。
- (8) .陳周造，陳燦煌編著，C++ Builder 4 徹底研究，博碩文化出版，2001 年 12 月。
- (9) .蔡明志 著，Borland C++ 繪圖設計，松崗出版，1993 年 10 月。
- (10) .楊宗誌 著，C++ Builder 資料庫程式設計文魁資訊出版，2001 年。
- (11) .鐘國亮 著，影像處理與電腦視覺第二版，東華書局出版，2004 年 2 月。
- (12) .呂炎州 著，不需字元切割的車牌辨識法，靜宜大學碩士論文，  
<http://ethesis.lib.pu.edu.tw/ETD-db/ETD-search/getfile?URN=etd-0722104-225725&filename=etd-0722104-225725-2.pdf>，2004 年。
- (13) .李仁軍 著，中英文件混合影像的全自動分類系統，國防大學國防管理學院資訊研究所碩士論文，  
[http://etds.ncl.edu.tw/theabs/extQuery/detail\\_result.jsp?id=091NDMC1654004](http://etds.ncl.edu.tw/theabs/extQuery/detail_result.jsp?id=091NDMC1654004)，2003 年 6 月。
- (14) .蘇博太 著，從視訊影像抽取文字，輔仁大學資訊工程所碩士論文，2005 年 7 月。