

植基於 Adaboost 演算法之影像文字定位

翁偉鐘,曾斐君,*張榮吉

亞洲大學數位媒體設計學系研究生

*亞洲大學數位媒體設計學系助理教授

Email: fay0426@hotmail.com

摘要

從影像或視訊序列中將文字區塊擷取出來之用途顯見於相當多的應用領域。本論文擬提出由自然場景相片中擷取文字區塊的方法。在自然場景相片中欲得到一些文字資訊，必須先將影像中的文字區塊切割出來，以便後續文字的辨識，因此文字擷取的正確率對於文字的辨識有很大的影響。本篇研究論文，使用一個快速的連通元件演算法產生各個連通元件，並利用簡單的幾何特徵過濾大量的連通元件，接著對剩下的連通元件進行小波與紋理特徵的擷取，將其擷取的特徵送到 Adaboost 作為其輸入，然後利用 Adaboost 演算法所訓練出來的強分類器進行篩選判斷。本研究實驗結合離散小波與紋理特徵，使得在文字定位的擷取正確率達 94.65 % 以上；此外透過 Adaboost 演算法的快速收斂，更可以減少計算成本的浪費。

關鍵詞：文字偵測、離散小波、Adaboost

Abstract

Extracting text region from images or video streams can be widely used in various application fields. A new method for text extraction in natural scene background is proposed in this study. This proposed method employs the connected-component (CC) based text extraction skills to produce automatic detection and segmentation of text from the digital images. The fast algorithm for labeling connected component is developed to generate CCs. The wavelet and texture are extracted from the remaining CCs, and these features are provided to the Adaboost classifier as an input for classification. With the strong classifiers, the CCs can be easily categorized either as texts or non-text character. Due to this circumstance, CC can be categorized as the text and then produce the final extraction result. The result can successfully facilitate the precision rates of text extraction up to 94.65 % and the computational cost can be efficiently reduced through the speed of convergence of Adaboost algorithm.

Keywords: Text detection、Discrete Wavelet、Adaboost

1. 前言

隨著資訊的成長，數位影像往往具備了許多重要的文字資訊。一般街道場景的影像通常都包含許多標誌和廣告看板，這些文字能夠傳達給我們名稱、交通、商業及注意等重要訊息，透過文字偵測技術便可將數位影像中的資訊傳遞給使用者做進一步的運用，通常這種處理技術會先對該影像進行文字偵測與擷取等前置處理動作，再將所結取的結果輸入至文字辨識系統裡進行文字辨識，因此文字的偵測與擷取可是非常重要的一部分。

本研究主要是利用連通元件演算法去找出影像文字區塊的所在，針對文字的「排列」、「字型大小」及「文字與背景所呈現出的強烈對比」等性質去進行分析判斷。對於這些性質的應用技術包含以紋理為基礎及以連通元件為基礎的方法。

以紋理為基礎的方法是藉由文字在影像中為較特殊紋理的特性，通常在對比上文字與背景會有一定得差異，水平密度排列成周期性的變化，文字間的距離會呈等距，藉由這些紋理特性便可將文字擷取出來。最常見的紋理分析技術如下：

1. 嘉伯濾波器 (Gabor Filter)
2. 離散餘弦轉換 (Discrete Cosine Transform)
3. 小波轉換 (Wavelet Transform)
4. 空間域變異 (Spatial Variance)

這些方法是藉由影像中每個子區塊的頻率係數來判定是否為文字區塊，再利用數位影像處理中的形態學技術去對所得的區塊做運算。例如[15]對於背景較為複雜的影像進行文字擷取，假設對一張影像沿著水平方向去計算空間域的變異(Spatial Variance)，通常在有文字的地方變異程度較大，因此可以找出文字所形

成的邊緣，再透過一些限制做為成對組合的條件，譬如兩個邊投影到座標軸上，必須有大部分是重疊；上升邊緣需在下降邊緣上面；兩個邊緣中間不應有其他邊出現等等，就能將初步偵測所找到的邊緣組合成可能的文字區塊。

其他如 Wu 等人[14]提出一種多重尺度紋理分割的方法去偵測影像中文字，利用九個濾波器來求得九種紋理特徵，每一個像素在九種紋理特徵中的值會形成一特徵向量，採用 K-mean 分群法對這些向量作分群(Clustering)，便可切出類文字的區塊。這些區塊中能量係數較高的區塊將被視為可能的文字區塊。再利用文字與背景會成強烈對比的特性，使用二階高斯導數來找出明顯的邊，最後用一些相似的高度及排列過濾，並進行連接區塊計算。

另一種以連通元件為基礎的方法是藉由文字中邊緣的來排情況，或者相同顏色所組成的排列情況來尋找文字區塊，這個方法不僅可將文字區塊位置偵測出來，同時還能對其定位。例如運用在辨識貨物中具有文字資訊區塊的擷取[12]，利用量化過後影像中像素之間差異去找出潛在字元的邊界，依據邊界上的像素利用區域門檻值進行候選文字的選擇，再將具有相同灰階的潛在文字連結成連通元，最後用一些啟發式條件進行過濾誤判的連通元件區塊。

還有利用堆疊的方式偵測出所有連通元件[2]，經過三個階段的合併方式來進行連通元件合併，而這三個階段分別是利用元件之間彼此的重疊性、各元件之間的間隔大小與元件兩側鄰居間隔差異來決定兩個元件是否需要被合併。最後再利用字元區塊投映與鄰近元件位置差異檢測方式，來找出文字的所在區域。

綜觀上述本篇論文主要目的有下面三點：

- 自動偵測與擷取數位影像中文字的所在位置
- 結合連通元件與紋理技術
- 提高文字正確擷取率與系統效能。

2. 文字擷取技術

本研究包括兩大部分：影像前處理及分類器訓練，其中影像前處理部分包含二值化、邊緣偵測、連通元件等前處理技術。前處理之後

產生的候選區塊必需透過一些前置過濾，再輸入至分類器進行訓練，其訓練樣本影像包含了文字樣本和非文字樣本，接著對訓練樣本影像與候選的文字區塊進行特徵擷取，將特徵輸入到 Adaboost 裡進行分類器的訓練與判斷，隨後利用所得的強分類器去過濾掉非文字的區塊，並將文字區塊輸出，最後透過一些簡單的區塊連接技術將分散的文字區塊連接，並將其連接的結果當作為最終的擷取結果，而本研究系統架構如圖 1。

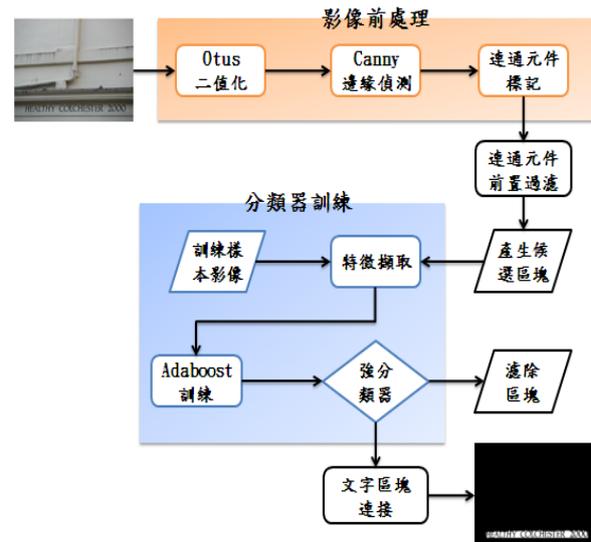


圖 1 系統架構

2.1 二值化

若欲將前景物件從背景裡抽取出來，使後續動作能夠變得較為簡單時，所需的方法就是影像二值化，觀察影像中灰階值分佈的情況，藉由這些分佈找出適當的門檻值。在本實驗中將採用自動計算門檻值的方法，稱之為 Otsu 二值化方法[5]。主要的方法為找出一個門檻值，能使各群集變異數的加權總合為最小，方法如下：

假設 T^* 為最佳門檻值，故利用這個 T^* 將影像中之像素分成兩個群，如圖 2 所示的 C1 和 C2。其中，C1 內的中任何一個灰階值 $f(x,y)$ ，必定滿足 $0 \leq f(x,y) \leq T^*$ ，而 C2 內中的任何一個灰階值則是滿足 $T^*+1 \leq f(x,y) \leq 255$ 。

在 Otsu 中提出了兩個條件，只要任何一者成立時即可決定 T^* 的值。

條件1. T^* 的決定會滿足 C1 與 C2 之間變異數

為最大(可看成 C1 與 C2 之間分的最開)。

條件2. T^* 的決定會滿足 C1 內變異數加上 C2 內的變異數之和為最小(C1 與 C2 間灰階值差異最小)。

以上二個條件只要有一個成立即可決定其 T^* 的值。由於目前並沒有很好的方法可以快速的找到 T^* ，因此在這裡需將 0~255 之間所有的灰階值帶入計算，再從最小的變異數中得到對應的灰階值 T^* ，即可決定出 C1 和 C2 之間的最佳門檻值。

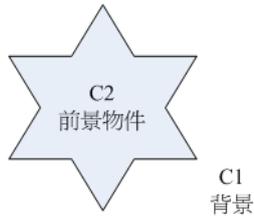


圖 2 影像分區示意圖

2.2 邊緣偵測

邊緣偵測的主要目的再找出物件與背景的分界，也就是找出兩個重疊物件之間的時間隔。本論文採用 Canny 邊緣偵測法步驟如下：

步驟1：高斯平滑濾波器，用來過濾雜訊

步驟2：計算梯度值與角度，利用 Sobel 算子去計算影像梯度值

步驟3：非最大值抑制法，上一步驟可以找到邊緣的位置，但是可能找到的邊緣不只一個像素，故此步驟將大於一個像素的邊緣，細化成一個像素之大小。

步驟4：雙門檻值演算法，由於步驟3可能會出現邊緣斷裂和刪除不正確邊界點的問題。因此在步驟4中提供了兩個門檻值來增加其偵測上的彈性，分別為 T_H 和 T_L 。當在步驟3所得到的各邊緣點的候選點中，若梯度值大於 T_H 的話，則判定為邊緣點。若邊緣點介於 T_H 和 T_L 之間，且八鄰接中包括一個以上的邊緣點，則會將此點補救回來，但若其值小於 T_L 時則視為非邊緣點。



圖 3 Canny 邊緣偵測(a)二值化影像 (b)邊緣偵測後

2.3 連通元件標記法

連通元件標記是透過像素與相連的像素連接，將整個影像中相連像素所連成的物件標示出來，通常相連的像素可分為四連通和八連通。本研究將使用 Suzuki[13]所提出的一個連通元件演算法，該方法能夠大幅減少標記所需花費的時間。

步驟1：先以第一次掃描得到初始標記，之後反覆進行逆向掃描和順向掃描更新標記，直到標記無改變即停止。

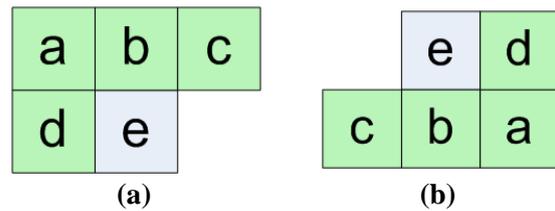


圖 4 連通元件掃描方式(a)順向(b)逆向

步驟2：第一次掃描的目的，用來求得初始標記，掃描的標記如公式 1、公式 2。其中 g 為標記影像， F_B 為非邊點， m 為標記， M_s 為鄰居的座標集合(除了 e 以外)， T 為標記相連表。當每次標記完 g 後都要更新標記相連表 T ，如公式 3

步驟3：第一次掃描後的掃描，此步驟分為逆向掃描和順向掃描及更新標記，分別如公式 4、公式 5 及公式 6。

完成連通元件標記法後，影像中的連通元件數量會非常多且大多是多餘的，如圖 5(a)所示，所以利用一些幾何特徵當作簡單的前置過濾條件，其前置條件如表 1。

$$g(x,y) = \begin{cases} F_B & \text{if } b(x,y) = F_B, \\ m, (m = m + 1) & \text{if } \forall \{i,j \in M_s\} g(x-i,y-j) = F_B, \\ T_{\min}(x,y) & \text{otherwise,} \end{cases} \quad (\text{公式 1})$$

$$T_{\min}(x,y) = \min_{\{i,j \in M_s\}} [T[g(x-i,y-j)]] \quad (\text{公式 2})$$

$$\begin{cases} \text{non-operation} & \text{if } b(x,y) = F_B \\ T[m] = m & \text{if } \forall \{i,j \in M_s\} g(x-i,y-j) = F_B, \\ T[g(x-i,y-j)] = T_{\min}(x,y) & \text{if } g(x-i,y-j) \neq F_B \end{cases} \quad (\text{公式 3})$$

$$g(x,y) \begin{cases} F_B & \text{if } g(x,y) = F_B \\ T_{\min}(x,y) & \text{otherwise} \end{cases} \quad (\text{公式 4})$$

$$T_{\min}(x,y) = \min_{\{i,j \in M\}} [T[g(x-i,y-j)]] \quad (\text{公式 5})$$

$$\begin{cases} \text{non-operation} & \text{if } g(x,y) = F_B \\ T[g(x-i,y-j)] = T_{\min}(x,y) & \text{if } g(x-i,y-j) \neq F_B \end{cases} \quad (\text{公式 6})$$

參數說明

W ：原始影像的寬。

H ：原始影像的高。

cW ：連通元件區塊的寬。

cH ：連通元件區塊的高。

cE ：連通元件區塊的邊緣點。



圖 5(a)連通元件標記結果(b)前置過濾結果

表 1 前置過濾規則表

過濾條件	過濾目的
$cW/cH > 2$ 和 $cE < 2(cW + cH)$ 或 $cH/cW > 1$ 和 $cE < 1.5(cW + cH)$	邊緣像素個數太少
$cW > W/2$ 或 $cH > H/2$	連通元件區塊太大
$cH \times cW < 50$	連通元件區塊面積太小
$8 \times cH < cW$ 或 $cH > 8 \times cW$	長寬比太大
$cW > 2cH$	連通元件區塊寬度異常
$cH < 16$	連通元件區塊高度太小

3. 分類器訓練

3.1 離散小波轉換

離散小波轉換的技術有許多種，其中以 Harr 小波轉換為最簡單的小波轉換(Wavelet Transform)，因此在本研究方法中採用了此離散小波轉換方式來獲得能量係數，將其所求得的能量係數做為影像特徵。

Haar 小波轉換會針對每一個相鄰的像素做處理，每一個相鄰像素的加總會存放在低頻頻道，也就是重要的頻道；而相鄰像素的相減，一張圖在平滑的區域的像素值之間差異並不大，所以相減之後的值並無太大的變化，但如果是在一張圖的邊緣區域，則相減之後差異很大，所以會造成強化邊緣的效果，而這些相鄰像素相減的值會存放在高頻頻道。

每一階段的 Haar 小波的轉換有兩個步驟，第一個步驟為水平切割，顧名思義是處理水平方向且由左至右的相鄰像素相加與相減，分別存放在低頻頻道與高頻頻道，水平切割完之後

緊接著第二步驟為垂直切割，它的處理順序為垂直方向且從上到下，運算的模式與水平切割相同，圖 6 為一階 haar 小波轉換的結果。



圖 6 Haar 轉換後

經過轉換後的影像有四個頻道，分別利用這四個頻道中的小波能量係數去做影像特徵值計算。其特徵共有 4 種如下：

特徵 1. 平均值

$$\mu = \frac{1}{M * N} \sum_{i=1}^M \sum_{j=1}^N s(i, j) \quad (\text{公式 7})$$

特徵 2. 標準差

$$\sigma^2 = \frac{1}{M * N} \sum_{i=1}^M \sum_{j=1}^N \sqrt{(s(i, j) - \mu)^2} \quad (\text{公式 8})$$

特徵 3. 能量

$$\frac{1}{M \times N} \sum_i \sum_j s(i, j)^2 \quad (\text{公式 9})$$

特徵 4. 二維動量

$$\sum_{i=0}^{255} (Z_i - \mu)^2 P(Z_i) \quad (\text{公式 10})$$

3.2 紋理

1973 年由 Haralick 等人[8]提出灰階共生矩陣 (Gray Level Co-occurrence Matrix)，利用統計分析的方式去計算影像像素的分佈情況，藉由檢視在相同的幾何位置上兩相對像素的相關性，以達到描述影像紋理的代表性特徵。

共生矩陣常見的表示為 $P(i, j, d, \theta)$ ，在本研究中， i 與 j 為小波係數值，而 d 則是指 i 與

j 之間的距離， θ 則是指在 i 與 j 之間的方向，其方向共分為 0° 、 45° 、 90° 和 135° 等四個角度如圖 7 所示。

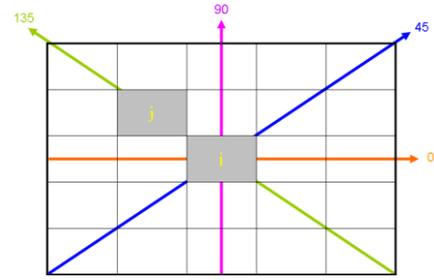


圖 7 共生矩陣的四個方向示意圖



圖 8 共生矩陣的例子

本研究採用了 Haralick[8]所提出的 14 種特徵的其中 4 種特徵去做為區塊的紋理特徵，並使用共生特徵矩陣去進行特徵分析，其特徵如下：

特徵 5. 同質性

$$\sum_i \sum_j \frac{P(i, j, d, \theta)}{1 + (i + j)^2} \quad (\text{公式 11})$$

特徵 6. 相關性

$$\sum_i \sum_j \frac{(i - \mu_x)(j - \mu_y)P(i, j, d, \theta)}{\sigma_x \sigma_y} \quad (\text{公式 12})$$

特徵 7. 對比度

$$\sum_i \sum_j (i - j)^2 P(i, j, d, \theta) \quad (\text{公式 13})$$

特徵 8. 熵

$$\frac{-1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N s(i, j)^2 \log s(i, j)^2 \quad (\text{公式 14})$$

因為共生矩陣有 4 個方向，故小波中的 4 個頻道都會有 16 種共生矩陣特徵，再加上前

面的小波係數特徵，每個區塊與訓練樣本都會有 80 種特徵，將這 80 種區塊特徵做為 Adaboost 的輸入。

3.3 Adaboost 演算法

本研究所採用的技術為 boosting 中的 Adaboost[7]，主要的想法是整合多個分類器學習效果，讓分類的精確度得以提高，其中某個機制是針對分類器所學習不好的訓練資料，再加以學習，如此便可能使得較不好學習的資料得以正確的被分類。

給予一樣本集： $(x_1, y_1), \dots, (x_m, y_m)$ ，其中 $x_i \in X, y_i \in Y = \{-1, +1\}$

初始化權重： $D_i(i) = 1/m$.

疊代次數 $t = 1, \dots, T$ ：

- 使用分佈樣本 D_t 去訓練弱分類器
- 得到一個分類器的假說 $h_t: X \rightarrow \{-1, +1\}$ ，

其錯誤率： $\varepsilon_t = \Pr_{i \sim D_t}[h_t(x_i) \neq y_i]$.

- 計算權重： $\alpha_t = \frac{1}{2} \ln\left(\frac{1 - \varepsilon_t}{\varepsilon_t}\right)$
- 更新：

$$D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} e^{-\alpha_t} > 1, & \text{if } h_t(x_i) = y_i \\ e^{\alpha_t} < 1, & \text{if } h_t(x_i) \neq y_i \end{cases}$$

$$= \frac{D_t(i) \exp(-\alpha_t y_i h_t(x_i))}{Z_t}$$

其中， Z_t 是一個正規化因子（使得 D_{t+1} 成爲一個分佈）。

輸出最終預測：

$$H(x) = \text{sign}\left(\sum_{t=1}^T \alpha_t h_t(x)\right).$$

圖 9 Adaboost 演算法

Adaboost 的演算法如圖 9 所示，一開始先給定一組訓練樣本集 $(x_1, y_1), \dots, (x_m, y_m)$ ，其中 x_i

代表的是訓練資料，而 y_i 則是屬於某些標籤的集合 Y 。在本研究中由於是兩類別的問題，因此假設 $Y = \{-1, +1\}$ 。Adaboost 會重覆執行 $t = 1, \dots, T$ 次疊代訓練，其中每次它所需學習的樣本會不斷的改變，重新被選取的樣本是依循著 $D_t(i)$ 這個權重關係，從此演算法中可以看出這個權重值的大小是隨著預測的對錯改變，也就是會將先前預測錯誤的樣本加重其權重值，使其錯誤的樣本再被選取的機率得以提升，並將先前預測正確的樣本減少其權重值，以降低重覆訓練的機率。

3.4 分類回歸樹

分類與迴歸樹 (Classification and Regression Trees, CART) 是在 1984 年由 Breiman 等學者所提出，利用二元遞迴分割 (Binary Recursive Partitioning) 的過程來分析龐大的資料集，CART 會依照預測變數與其相對的各項指標，將現有的訓練樣本劃分成數個已知的類別，並將其劃分的程序整理成一連串的規則 (Rule)。

本研究利用了下面的方法建立 CART 的節點，假設有一樣本集 $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$ ，在這裡各自的 x_j 屬於樣本空間 ($x \in R^n$)，即維度為 n 的實數向量，表示為 $\{(x_1^j, y_1^j), (x_2^j, y_2^j), \dots, (x_n^j, y_n^j)\}$ ，而且各自的標籤 (label) y^i 屬於有限的標籤空間 Y ，在本研究中，為二元分類之問題，故在這裡 $Y = \{1, -1\}$ 。

- 建立 CART 節點的方法：

步驟 1：對於每一特徵，找出一個適當的門檻值 θ ，使得分割樣本空間的錯誤率最低。

步驟 2：選擇具有最小錯誤的第 i 個特徵，並建立一個 CART 節點。

- (1) 設定分支條件。
- (2) 將分支的結果 (True/False) 連結到葉節點，並擁有各自的分類。

假設葉節點的錯誤指的是樣本被分類錯誤的機率，則當在樹的建立期間遭遇到錯誤分類時，我們將在葉節點上停止。建立整個 CART 樹的步驟如下：

● 建立 CART 樹的步驟：

- 步驟1：建立根節點(Root)。
- 步驟2：選擇錯誤率最大的葉節點。
- 步驟3：建立新的節點，使用與該節點相關的樣本。
- 步驟4：替換該葉節點。
- 步驟5：重複步驟 2-4 直到所有葉節點都沒有錯誤或達到預設停止條件。



圖 10 使用 Adaboost 過濾後的結果

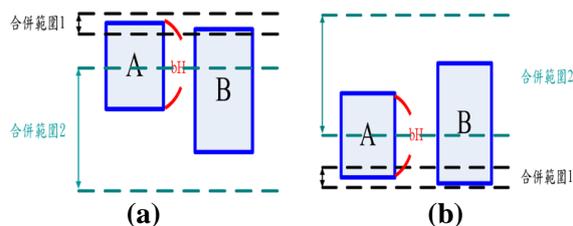


圖 11 基準線示意圖 (a) 基準線在上 (b) 基準線在下



圖 12 文字區塊連接的結果

3.5 文字區塊合併

最後本研究加入了文字區塊連接藉以來確認文字區塊的正確性，以彌補 Adaboost 可能產生的文字區塊遺失的問題，並將其結果做為最後的擷取結果。

先從某一文字區塊 A 開始進行掃描，其掃描方式是將區塊 A 的高度放大兩倍，由左到右進行水平方向的掃描。假使有一個連通元件 B 滿足了下面的任一條件時，將其視為同一個群集，並以遞迴的方式去完成所有的群集。

範圍1. 挑選具有較小高度(bH)的兩區塊，以 A 區塊做為基準線，上下延伸各 $bH/2$ 長。

範圍2. A 區塊在上：A 區塊的下界為基準，往上延伸 $bH/2$ 長，往下延伸 bH 長。A 區塊在下：A 區塊的上界為基準，往上延伸 bH 長，往下延伸 $bH/2$ 長。

濾除：

1. 不在同一文字線上的區塊。
2. 高度不滿足合併條件的區塊。
3. 獨立的文字區塊。

4. 實驗結果與分析

4.1 系統效能

本實驗中總共測試了 50 張影像，共 16621 個連通元件，其中文字連通元件的個數為 1224，而非文字連通元件之個數為 14997。

$$\text{召回率(Recall)} = \frac{Dc}{Dc + Du} \quad (\text{公式 15})$$

$$\text{準確率(Precision)} = \frac{Dc}{Dc + Dq} \quad (\text{公式 16})$$

其中 Dc 為文字區塊被分類正確的個數， Du 為文字區塊中被分類成非文字區塊的個數， Dq 為非文字區塊被分類成文字區塊的個數，其系統效能如表 2。

表 2 系統效能

資料 效能	召回率	準確率
訓練樣本	86.15%	90.16
測試樣本	92.76	94.65

本實驗主要測試的影像為數位影像中具有文字訊息之影像，本實驗中偵測與擷取的結果如圖 13。

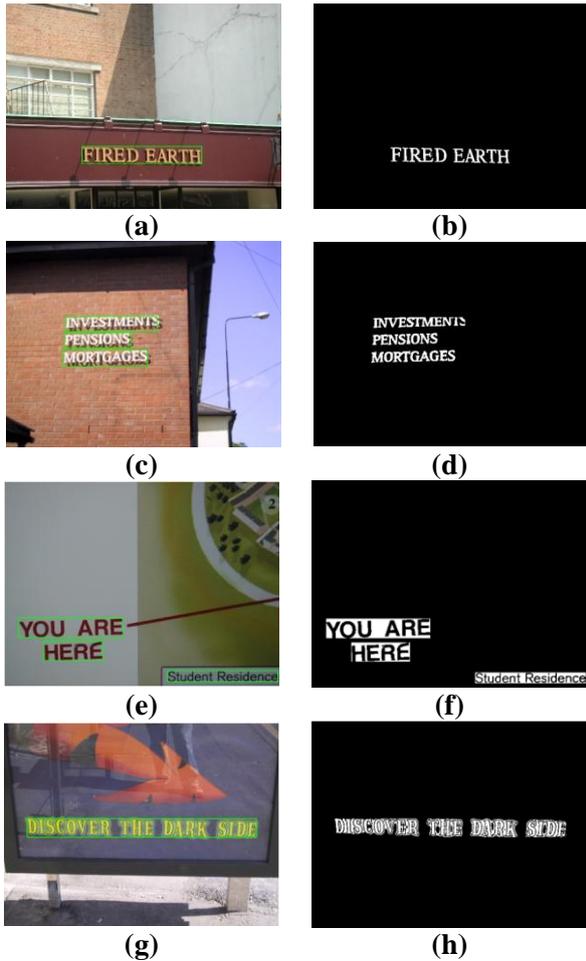


圖 13 實驗結果

4.2 文字偵測與擷取效能分析

在本實驗中，在某些特定的環境下，會導致前處理效果不佳以及最終的擷取結果令人不滿意。

1. 在二值化階段：對於環境光線、陰影、髒污等影像具有較多的限制，因此對圖中有陰影的部分明顯處理的較差，也造成最後二值化結果不佳，如圖 14。



圖 14 二值化不佳的例子(a)原始影像(b)不佳結果

2. 合併階段：由於進行文字區塊合併是依據在水平方向上周圍的區塊條件來判斷是否需要合併若文字的排列不是依照水平方向時，便會導致其最後結果的失敗，如圖 15。



圖 15 合併失敗(a)原始影像(b)失敗結果

5. 結論與未來展望

在本研究中，提出了一個以 Adaboost 演算法為基礎去建立一個快速的文字擷取系統，並結合連通元件標記法與紋理分析的優點，因此本系統同時具有文字定位與較佳特徵描述的能力；首先連通元件標記法去對數位影像中文字區塊做定位處理，接著再利用離散小波與紋理特徵去對所擷取的文字區塊進行分類判斷；最後本實驗的擷取正確率達 94.65 % 以上，在同一相關領域中這樣的擷取率已達不錯的水準；將這樣的擷取結果進一步輸入到 OCR 中，可以使 OCR 的辨識正確率更加的提升。

實驗中透過一些失敗與成功案例，了解未來的工作仍有許多可以處理的空間，例如影像中的陰影、反光等外在干擾因素的處理，都是值得去深入研究，因此本研究未來的工作將把其焦點放在如何去設計一個適應性的影像修補方式，使得後續的工作效能可以獲得提升。

參考文獻

- [1] 陳昭穎，“資料探勘技術於超音波旋轉肌肌群影像之診斷應用”，*國立屏東商業技術學院資訊管理學系*，2006。
- [2] 曾逸鴻，“文件影像之手寫中文字擷取技術”，*科學與工程技術期刊*，pp. 13-22，2005。
- [3] 吳上立、林宏墩，*C 語言數位影像處理(修訂版冊)*，全華書局，2006。
- [4] 連國珍，*數位影像處理*，儒林圖書公司，2004。
- [5] 鐘國亮，*影像處理與電腦視覺*，東華書局，

2004.

- [6] Canny, J. "A computational Approach to Edge Detection", *Pattern Analysis and Machine Intelligence*, Vol. 8, pp. 679-698, 1986.
- [7] Freund, Y., & Schapire, R. "A Decision-theoretic generalization of On-line Learning and an Application to Boosting". *Computer and System Science*, 55, 119-139, 1997.
- [8] Haralick, R. M., Shanmugam, K., & Dinstein, I. "Textural Features for Image Classification", *Systems, Man, and Cybernetics*, Vol. 3, pp. 610-662, 1973.
- [9] Kim, E. Y., K. Jung, Jeong, K. Y., & Kim, H. J. "Automatic Text Region Extraction using Cluster-based Templates", *Pattern Recognition and Digital Techniques*, pp. 418-421, 2000.
- [10] Kim, K. I., K. Jung, S. H. P., & Kim, H. J. "Support Vector Machine-based Text Detection in Digital Video", *Pattern Recognition*, Vol. 34, pp. 527-529, 2001.
- [11] Lee, E. R., Kim, P. K., & Kim, H. J. "Automatic Recognition of a Car License Plate using Color Image Processing", *IEEE International Conference on Image Processing*, Vol. 2, pp. 302-305, 1994.
- [12] Lee, C., M. "Automatic Extraction of Characters in Complex Scene Images", *Pattern Recognition*, Vol. 19, pp. 67-82, 1995.
- [13] Suzuki, K., Horiba, I., & Sugie, N. "Linear-time Connected-component Labeling Based on Sequential Local Operations", *Computer Vision and Image Understanding*, pp. 1-23, 2003.
- [14] Wu, V., Manmatha, R., & Riseman, E. M. "An Automatic System to Detect and Recognize Text in Images", *Pattern Analysis and Machine Intelligence*, Vol. 21, pp. 1224-1229, 1999.
- [15] Zhong, Y., Karu, K., & Jain, A. K. "Locating Text in Complex Color Image", *Pattern Recognition*, Vol. 28, pp. 1523-1535, 1995.