

工程圖影像字元萃取實作

曾伊秀 陳文儉*

大葉大學資訊工程學系

51591 彰化縣大村鄉學府路 168 號

*cwj@mail.dyu.edu.tw

摘要

早期工程圖影像以紙本方式來保存，需利用掃描儀器數位化建檔儲存電腦後，可以進一步在工程圖影像開發相關的應用。本文將實作工程圖影像中的字元萃取。工程圖影像經過二值化處理可把背景與物件分離，依此影像中所含特徵，利用串接應用進行斷裂線段之填補，並找出影像中線段與文字的所有可能群聚區域，再以形態學影像處理的運算來重建影像。最後，利用連通物件標記法以及一些判斷條件來分離出字元與線段來提高萃取的正確率。實驗結果顯示本文所用的方法可以成功萃取出工程圖影像字元的位置，獲得相當高的萃取正確率。

關鍵詞：字元萃取，工程圖影像，二值化，形態學，連通物件標記法

Implementation of Character Extraction for Engineering Drawing Images

I-HSIU TSENG and WEN-JAN CHEN*

Department of Computer Science and Information Engineering, Da-Yeh University

No. 168, University Rd., Dacun, Changhua 51591, Taiwan, R.O.C.

*cwj@mail.dyu.edu.tw

ABSTRACT

Engineering drawing images are saved in paper form initially. We can then scan those images and save them on a computer if we wish to subsequently develop some of their applications. This paper explores the implementation of character extraction for engineering drawings. An engineering drawing image is binarized to separate the background from the object. For engineering drawings, we use cascading applications to fill lines and use connections of points to group lines and writing regions. Then, two forms of morphological image processing, namely dilation and erosion, are applied to rebuild it. Extraction errors result when characters and lines are close together. To solve this problem, connected component labeling and judgment conditions are used to separate characters and lines, thus improving extraction accuracy. Experimental results demonstrate that this method can be used to identify character locations for engineering drawing images, which produces higher quality simply structured engineering drawing images.

Keywords: Character extraction, Engineering drawing images, Thresholding, Morphology, Connected



component labeling

一、前言

早期的工程圖影像皆仍以紙本方式來保存，經由儀器將紙本影像掃描並結合電腦建檔儲存，目前工程圖影像已經逐步數位化。

不同的工程圖影像籍，例如：地籍圖、建築圖、手繪圖、印刷文件等，其中線段的長度、粗細及文字尺寸的形式皆不一致，如何從這些工程圖影像中萃取出重要的特徵值，以提供形狀辨識研究上所需的依據，這是工程圖影像自動檢測系統的重要研究項目。

由於工程圖內含了數量龐大的線段，其中線段的寬度會有不均勻的分佈，且因資料保存的不善或掃描儀器在影像品質會選擇不同的解析度之過程，產生許多線段斷裂的現象，綜合上述情形，若想在人類視覺系統上取得線段是容易的，但在數位影像處理中，則為複雜且繁多的演算法。

本研究動機及主要目的，希望能夠將工程圖影像中線段的長度、粗細及文字上尺寸的形式不一致，加以萃取出來，並進一步以高效率的方式來存取，另外如果可以正確將工程圖的字元萃取出來，在某些工程圖的中，字元可能是個人資訊如姓名，就可以去做個人資訊的隱藏或處理。

目前在而在工程圖影像相關研究文獻較少，車牌定位辨識 [1, 8-9] 的相關研究中，作法有許多相似的地方，所以本文引用車牌辨識的方法來實作工程圖影像中字元的部分萃取出來。

本文共分為五節：第一節闡述研究動機與背景；第二節將介紹本文研究中所應用到的相關理論及方法；第三節詳述工程圖字元萃取的系統架構設計與實現；第四節為實驗結果的呈現與比較；最後為本文的結論。

二、相關研究

本節將介紹本文中所應用的相關理論及方法，包括影像二值化、形態學、連通物件標記法及串接等影像處理技術。

(一) 二值化

影像二值化的主要功能為分離出背景 (background) 與前景 (foreground) [2, 5]，舉凡圖形辨識、字元識別，皆可利用此方式將一個複雜影像圖面做簡單化。

處理的過程中，通常使用門檻值 (threshold) 設定來決定背景和物件，而門檻值主要是利用直方圖分析，在影像中找出單一門檻值，將所有低於門檻值之像素視為背景，反之若高於門檻值之像素，則視為物件。

(二) 形態學

形態學影像處理基本上是一個集合觀念上的理論和應用，可用來解決影像切割、雜訊的抑制、特徵萃取等影像處理問題，提供影像處理問題上強有力之方法，斷開 (opening) 與閉合 (closing) 運算為影像處理的重要工具 [7, 10]。

連通物件標記法 (connected component labeling, CCL) 是一種區塊標記的方法，從二值化影像作群聚分類，將影像中相聯在一起的區塊之像素給予同一個標記，視為同一物件 [4, 6]。在介紹連接物件標記演算法之前，先說明四鄰點 (4-neighbor points) 與八鄰點 (8-neighbor points) 之連通物件法。

取出一為 3×3 遮罩範圍的 9 個像素，以 5 為中心點，其上、下、左、右，如圖 1(b) 黑色位置，稱為四鄰點；以 5 為中心點，其周圍的像素點，如圖 1(c) 黑色位置，則稱為八鄰點。

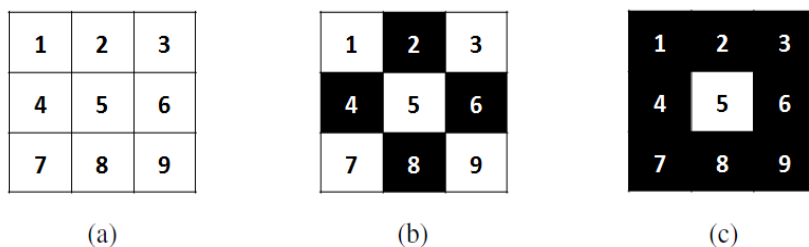


圖 1. 鄰點示意圖



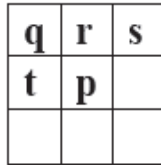


圖 2. 八鄰點示意圖

本文使用八鄰點來做連通物件標記法如圖 2 所示，其標記步驟如下：

先由左至右，再由上而下掃描影像中每一個像素點， p 為影像中之像素點， $\{q、r、s、t\}$ 則為 p 的鄰近像素點，分別會有以下三種情形：

1. 當掃描到 p 為白點，則不給予任何編號，繼續往下掃描。若掃描到 p 為黑點，且 $\{q、r、s、t\}$ 皆為白點，則給 p 一個新的編號後繼續往下掃描。
2. 若掃描到 p 為黑點，且 $\{q、r、s、t\}$ 其中有任一個以上為黑點，則找出為黑點像素位置之最小編號給 p ，並將像素位置為黑點的每一個編號紀錄為同一集合，接著繼續往下掃描。

當掃描完成後，再將紀錄為同一集合的編號做合併。以圖 3 為例，圖 3(b) 為連通物件標記中三種情形的結果，它形成了 3 個集合分別為 $\{1、2、3\}$ 、 $\{4、5\}$ 、 $\{6、7、8\}$ ，其中編號 1、編號 2、編號 3 被紀錄為同一個物件，編號 4、編號 5 紀錄為同一個物件，以下依此類推，經由此關係，在執行步驟 3 合併每個物件來完成連通物件標記法，如圖 3(c) 所示。

像素點的串接主要目的是將兩個分開之像素點，利用設

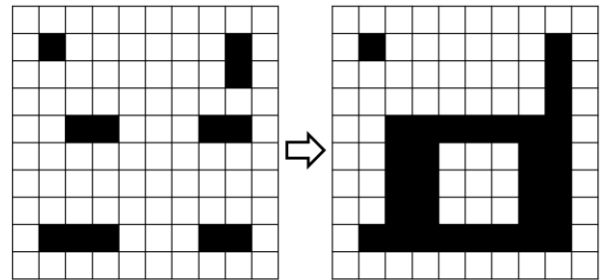


圖 4. 像素點的串接示意圖

定的門檻值 T ，使兩個分開之像素點距離小於門檻值 T 的部位強制串接為線段 [3]，以下將介紹串接的步驟：

1. 設定一距離門檻值 T 後，整張影像由左至右，由上而下掃描影像中水平方向，計算水平方向同一列中左右相鄰點之間的距離，只要兩點之間的距離是小於門檻值 T ，則強制將兩點串接為線段。
2. 整張影像由上而下，由左至右掃描影像中垂直方向，計算垂直方向同一列中上下相鄰點之間的距離，只要兩點之間的距離是小於門檻值 T ，則強制將兩點串接為線段。

整張影像之像素點的串接示意圖，如圖 4 所示，門檻值 T 設為 5。

三、系統設計與實現

(一) 系統設計架構

本文所使用測試之工程圖是經數位化後影像，再由電腦進行工程圖之字元萃取，主要步驟包含：二值化、像素點的串接、連通物件標記法、斷開運算、細部處理與字元萃取。

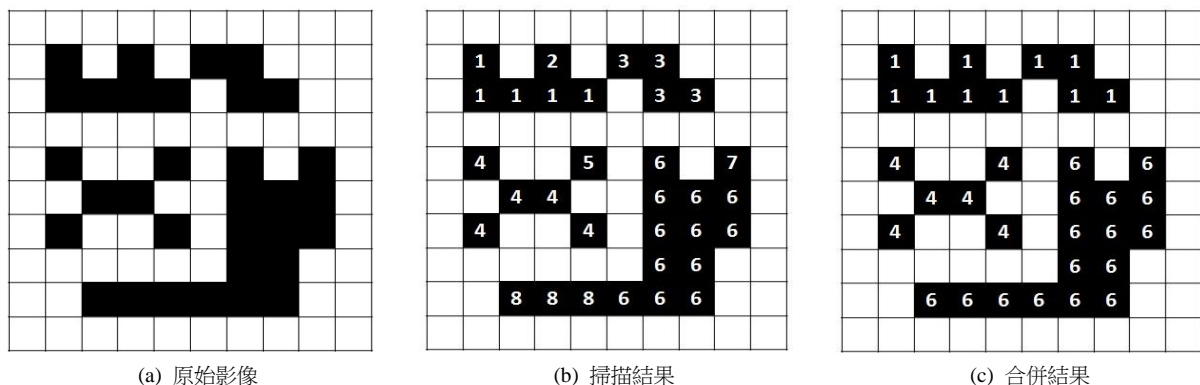


圖 3. 連通物件標記法示意圖



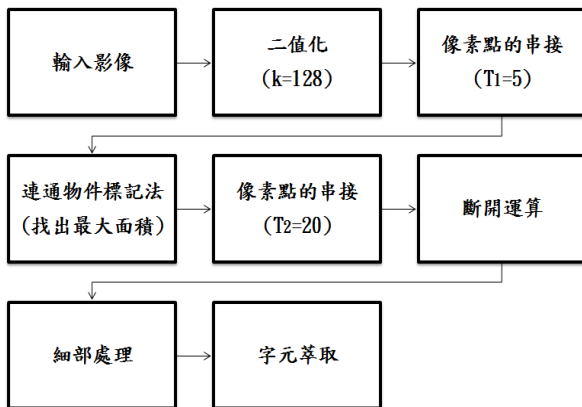


圖 5. 系統流程圖

首先，原圖經二值化的程序分離出影像中背景和物件，由於本研究之工程圖影像較為單純，所以門檻值設定為 128，接著利用像素點的串接應用找出字元與線段的區塊位置，根據這些區塊進行細部處理，分離出所需要之部位完成字元萃取，系統主要流程如圖 5 所示。

(二) 像素點的串接應用

原圖經二值化的程序分離出影像中背景和物件，根據物件群聚區域分布，發現線段群聚的區域較為狹窄或細長，利用此特性，本文提出像素點的串接與形態學中之斷開運算結合應用來進行字元萃取，稱為方法一。

方法一主要目的是將影像中狹窄或細長部分截斷，即工程圖上線段的部分，所以在進行斷開運算之前，先找出影像中線段與文字所有可能群聚的區域，距離門檻值 T 設為 20，利用像素點的串接，只要兩點之間的距離小於距離門檻值 T ，則強制將兩點串接為線段，使群聚區域填補成為一整個區塊。由於文字部分所形成的區塊，不會因為斷開運算而被截斷，所以留下來的部分約略就是字元所在的區塊，接著利用這些區塊，即可將字元萃取出來，其流程如圖 6 所示。

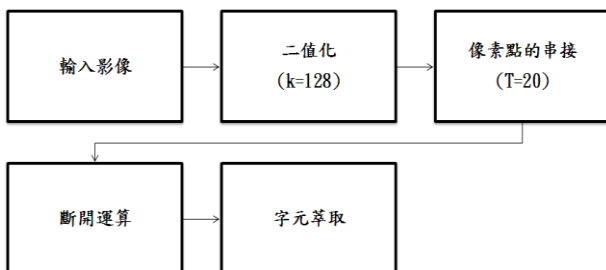


圖 6. 方法一流程圖

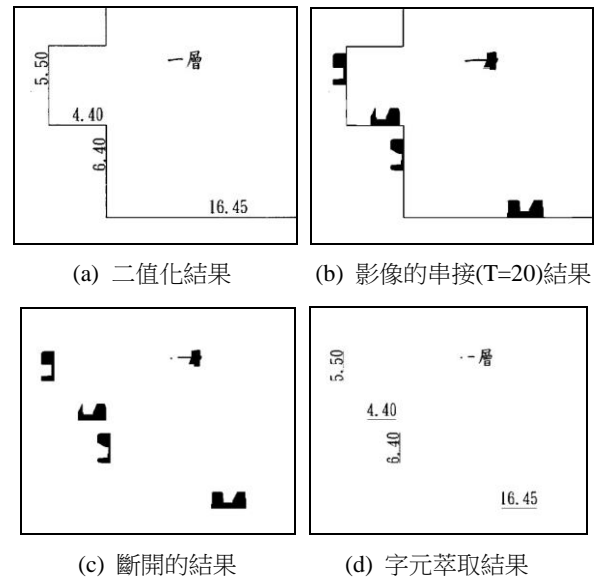


圖 7. 方法一實驗結果

經由方法一萃取結果發現，字元群聚的區域如果遇到狹窄或細長部分，仍會因為斷開運算而被截斷，以圖 7(b)為例，影像中的中文字元「一」，因為群聚區域為一個細長部分，而周圍群聚區域也較為不密集，所以像素點串接後的區塊，經過斷開運算被視為線段而截斷。

由於工程圖影像中線段群聚分佈所圍起來部分，會包含整張影像，根據此特性，本文提出利用連通物件標記法找出影像中最大面積的物件，即為線段群聚主要部分，來排除那些離線段較遠且群聚區域較為不密集的部分，進而改善方法一當中誤判之現象，稱為方法二。

方法二主要目的為利用連通物件標記法找出影像中最大面積的物件，但是影像在進行二值化後留下線段斷裂是不可避免的情形，所以在進行連通物件標記法之前，為了要正確找出連通物件標記法後影像中最大面積的物件，利用像素點的串接來進行填補，距離門檻值 $T1$ 設為 5，只要兩點之間的距離小於距離門檻值 T ，則強制將兩點串接為線段，使之前因為二值化產生線段斷裂的部分填補起來，CCL 後之結果先與上述 $T=5$ 的串接結果將影像中狹窄或細長部份區塊截斷分離標示出來後，再做第二次像素點的串接，距離門檻值 $T2$ 設為 20，接著段開運算後留下來的部分約略就是字元所在的區塊，最後利用這些區塊及上述狹窄或細長部份區塊，將字元萃取出來，流程如圖 8 所示，其實驗結果如圖 9 所示。



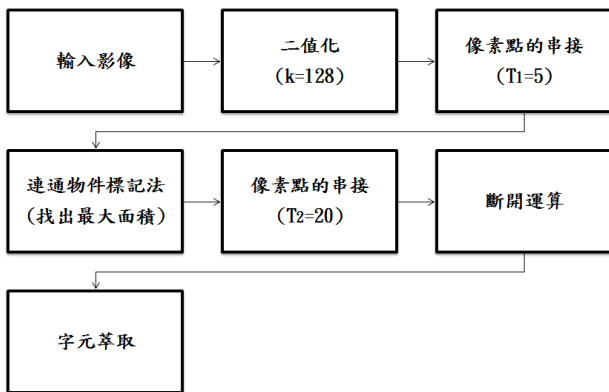


圖 8. 方法二流程圖

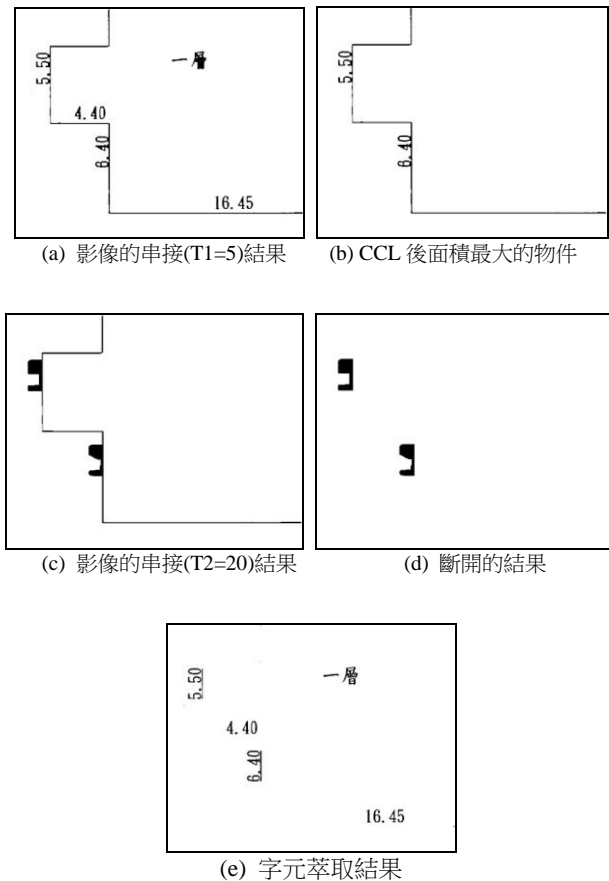


圖 9. 方法二實驗結果

(三) 細部處理

由圖 9(e) 可以發現，留下來的部分約略就是字元所在的區塊，雖然萃取結果在字元部份成功的被萃取出來，但是因字元與線段相當接近，而將線段部分一併視為字元，造成字元萃取後工程圖上的失誤。

為了解決上述問題，細部處理的部分利用連通物件標記法以及一些判斷條件來分離出字元與線段，使線段部分不會

被誤判為字元，稱為方法三其細部處理流程如圖 10 所示。

首先利用斷開運算結果影像以連通物件標記法進行標記，以圖 9(d) 為例，此影像可分成兩個物件，接著找出包含這兩個物件所形成之最小矩形的影像，如圖 11(a) 和圖 11 (b) 兩個影像分別作處理。

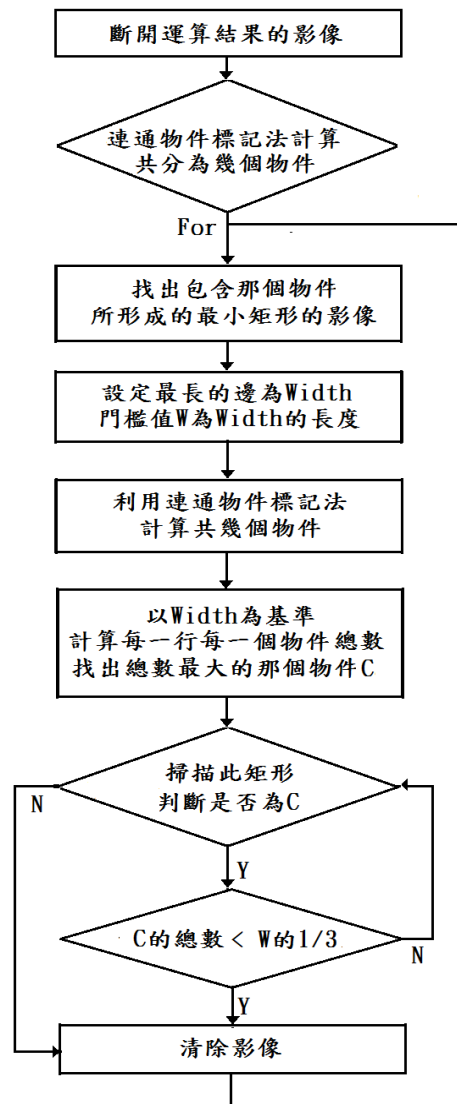


圖 10. 細部處理的流程圖

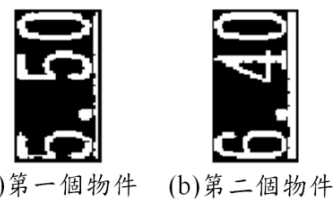


圖 11. 經 CCL 後產生的最小矩形影像



以圖 11(b)第二個物件為例，找出兩邊長分別為 34 像素與 67 像素，因此設定最長的邊 67 像素為影像 Width，令 Width 長度為門檻值 $W = 67$ ，接著以圖 11(b)進行連通物件標記法來標記，共可分為四個子物件，如圖 12 所示。

以四個子物件垂直方向計算每一行每一個子物件的投影總數如圖 13 所示，並找出所有行投影總數的，並將其物件編號訂為 C，圖 12 範例為子物件 2。接下掃描整張影像，將已經標記的區塊用兩個條件來篩選出線段，篩選條件如下：

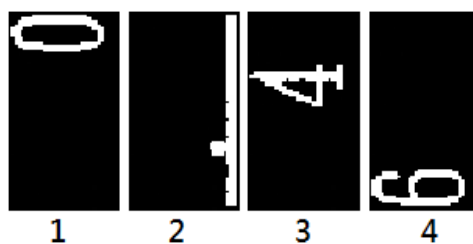


圖 12. 經 CCL 後產生的 4 個子物件

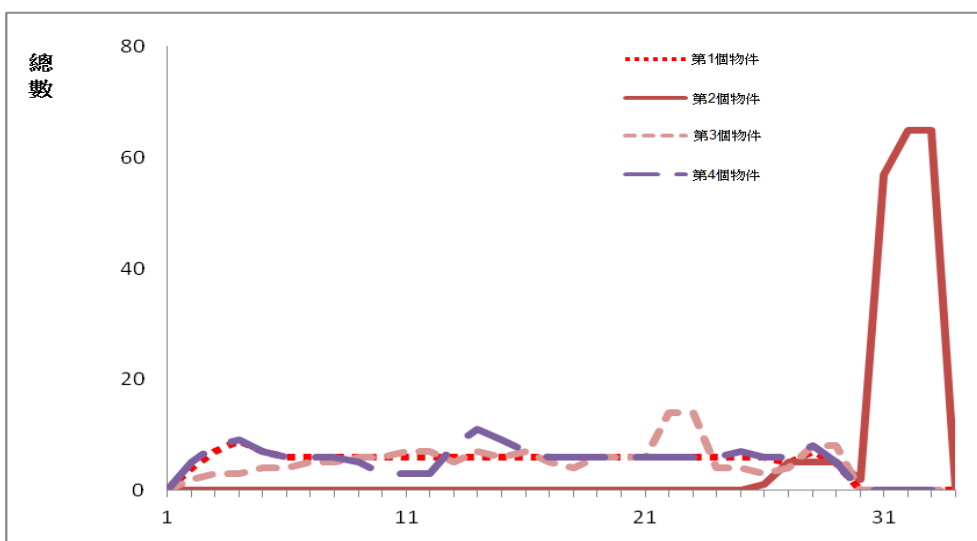


圖 13. 子物件的投影總數

1. 只要不為最大的子物件 C，則將影像清除。
2. 為子物件 C，且行投影總數 < 門檻值 W 的 $1/3$ ，則將行影像清除。

經由上述兩個條件將篩選出影像被誤判的線段部分，這樣就可以得到影像中字元與線段正確的位置如圖 14，只要知道字元或線段的位置，即可萃取所需要之部位。

(四) 字元萃取

根據方法三中細部處理判斷條件來分離出誤判的線段部分以及方法二中離線段較遠而被排除的字元部分，這樣就可得到整張影像中字元與線段正確的位置，只要知道字元或線段的位置，即可利用簡單的邏輯運算符號 AND 和 XOR 運算，進而萃取所需要之部位，將整張影像字元的位置與二值化影像做 AND 運算即可得出字元萃取結果如圖 15 所示，而做 XOR 運算即可得出字元萃取後工程圖如圖 16 所示。

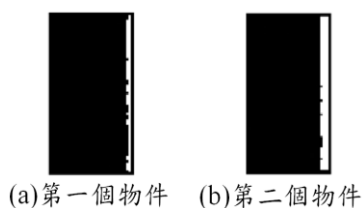


圖 14. 篩選結果

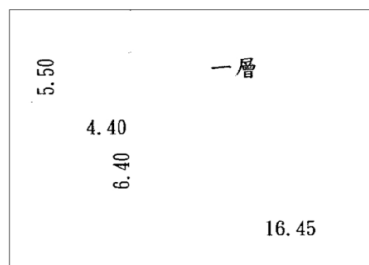


圖 15. 字元萃取結果



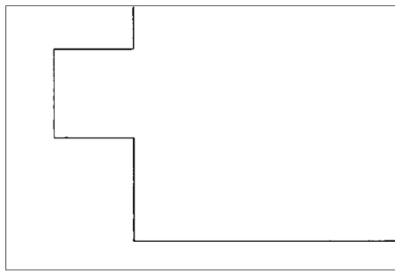


圖 16. 字元萃取後的工程圖

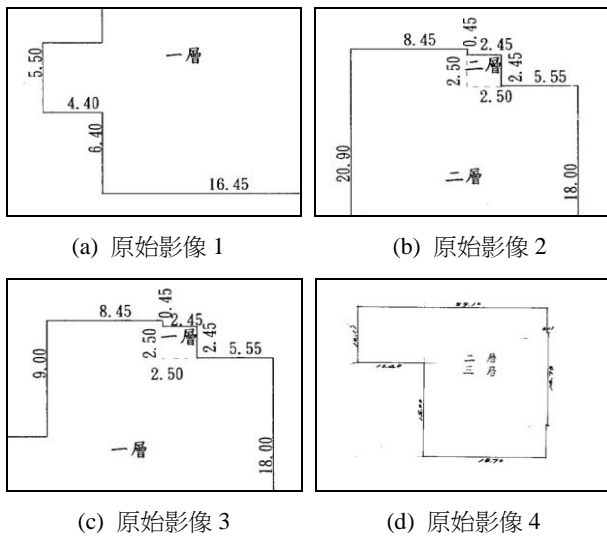


圖 17. 測試原始影像

四、系統實作結果

本文所測試之工程圖為紙本經由數位化後影像，測試樣本皆採用解析度 640×480 像素大小的影像，所有使用之工程圖影像由光特資訊科技股份有限公司所提供。

本節我們將使用 4 張工程圖測試影像如圖 17，透過本文所述的系統架構：方法一、方法二以及方法三來進行萃取結果比較，其表現差異如表 1 所示。可以看出方法一表現結果較差，而方法二在細部處理功能較差，最後的方法三在細部處理功能有良好的效能。

並以人工手動方式進行 640×480 像素大小的影像字元萃取做為標準，利用簡單的邏輯運算符號 XOR 和 NOT，統計出字元部分的總數，計算出萃取結果之正確率，如表 2 為字元萃取正確率。另外我們也利用簡單的邏輯運算符號 XOR 和 NOT，統計字元萃取後工程圖部分的總數並計算正確率，如表 3 為字元萃取後工程圖正確率，圖 18~21 分別為

測試影像 1~4 的字元萃取比較，圖 22~25 為字元萃取後工程圖比較。

整體實驗可以看出由於方法二沒有做細部處理，對於工程圖影像中，如果字元較接近線段的區塊有很多時，就會與方法三的效果差異很大，就較不適合使用。方法三對於一般的工程圖影像都有不錯效果，但是如果字元集中且黏著線段則仍無法完全清除線段，未來可以進一步研究解決這問題。

表 1. 三種方法表現差異表

	方法一	方法二	方法三
二值化	O	O	O
像素點的串接(T1=5)	X	O	O
連通物件標記法	X	O	O
像素點的串接(T2=20)	O	O	O
斷開運算	O	O	O
細部處理	X	X	O

表 2. 字元萃取的正確率

	影像 1	影像 2	影像 3	影像 4	平均
方法一	90.04%	91.95%	89.79%	81.11%	88.22%
方法二	95.25%	94.18%	93.89%	84.12%	91.86%
方法三	100%	99.73%	99.71%	96.29%	98.93%

表 3. 字元萃取後工程圖的正確率

	影像 1	影像 2	影像 3	影像 4	平均
方法一	88.42%	76.61%	70.77%	86.13%	80.48%
方法二	92.09%	81.26%	74.33%	80.28%	81.99%
方法三	99.93%	97.25%	94.06%	95.57%	96.70%

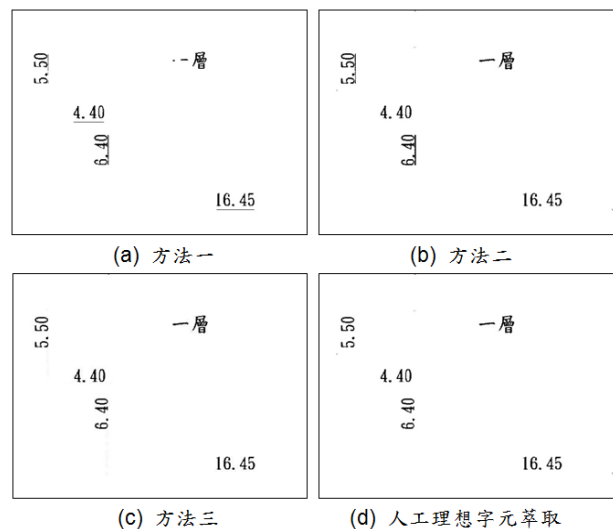


圖 18. 影像 1 字元萃取比較



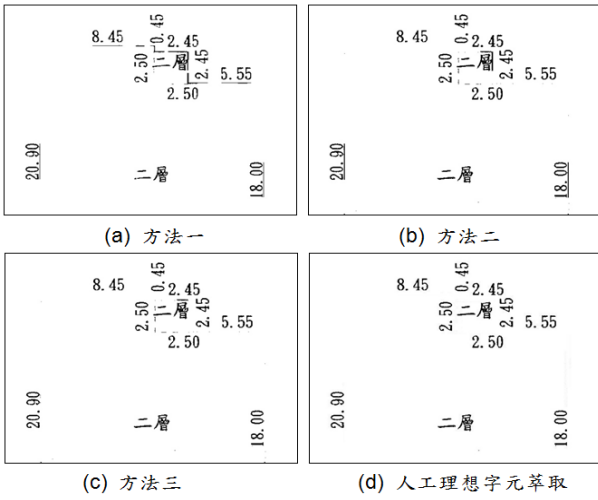


圖 19. 影像 2 字元萃取比較

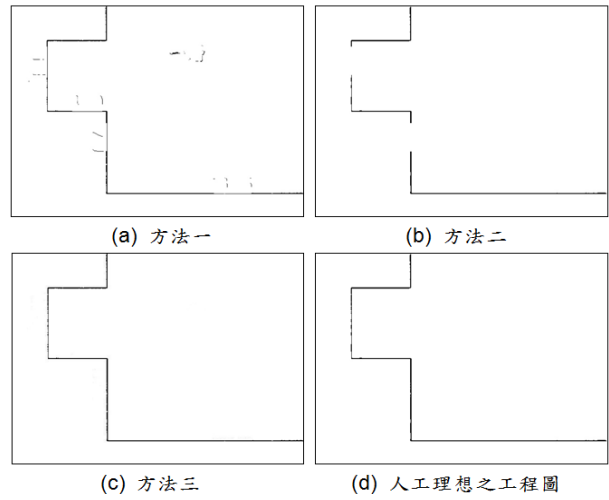


圖 22. 影像 1 字元萃取後工程圖比較

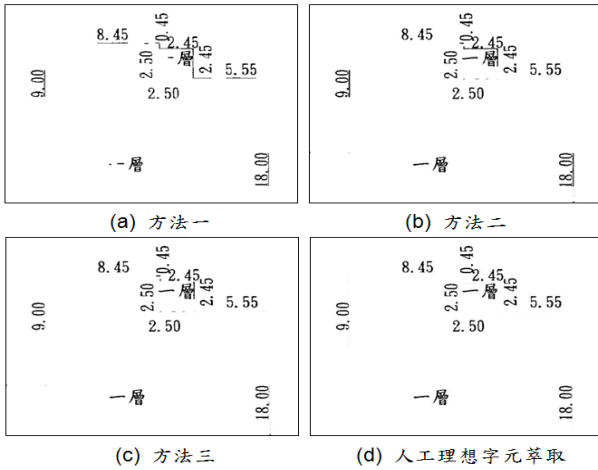


圖 20. 影像 3 字元萃取比較

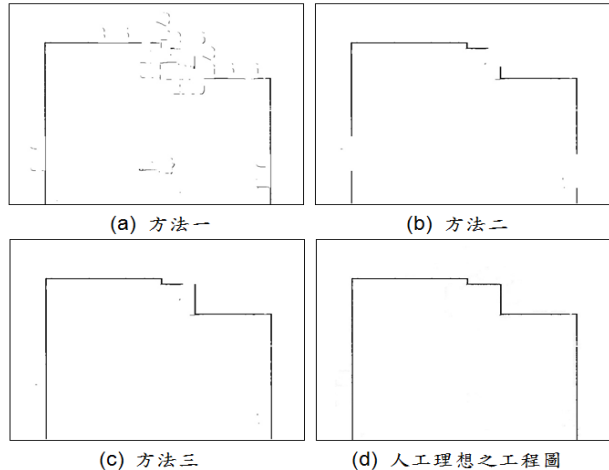


圖 23. 影像 2 字元萃取後工程圖比較

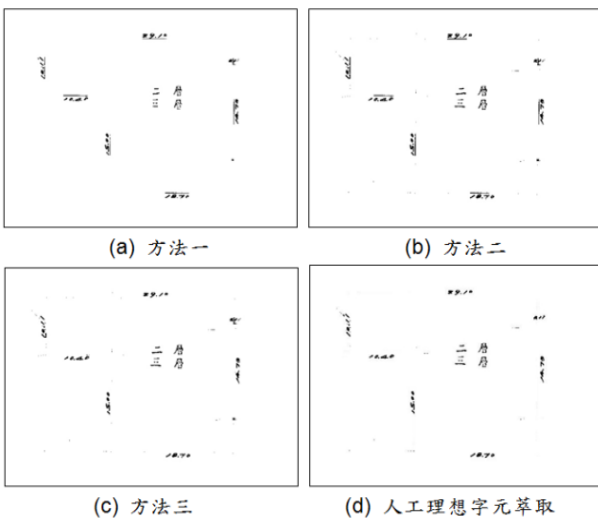


圖 21. 影像 4 字元萃取比較

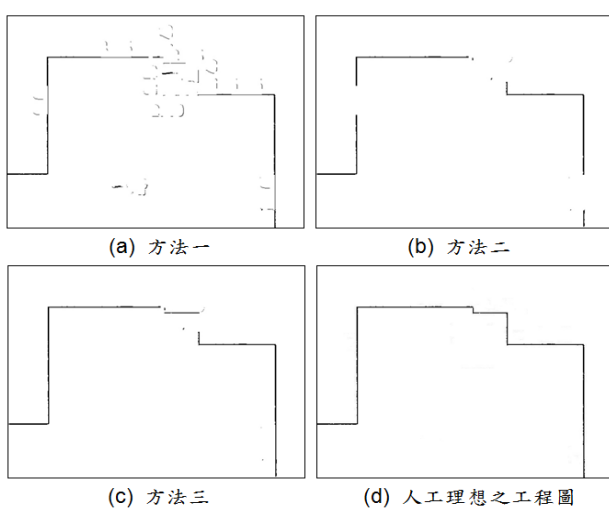


圖 24. 影像 3 字元萃取後工程圖比較



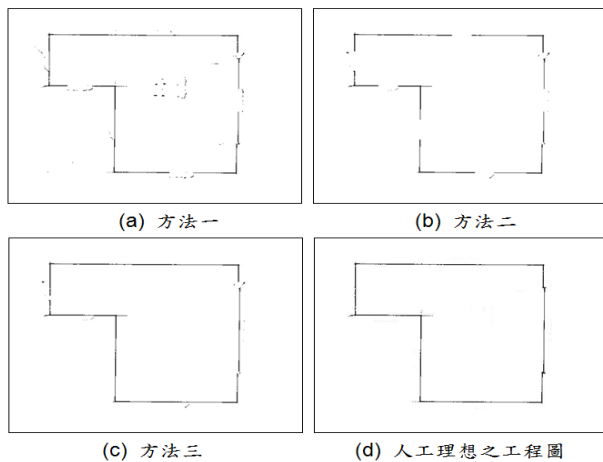


圖 25. 影像 4 字元萃取後工程圖比較

五、結論

本文針對工程圖包含了大量線段與文字資訊之特性，實現一個工程圖影像字元萃取的雛型。首先，我們建立出影像的串接應用，解決影像在進行二值化後留下線段斷裂的情形，以及根據影像中線段與文字可能群聚的區域，將群聚區域填補成為一整個區塊，減少運算。另外為了解決萃取部分會因字元與線段相當接近，而將線段部分一併視為字元，造成萃取上的失敗，提出細部處理來降低萃取的失誤。根據實驗結果，本文可以在結構簡單之工程圖上，獲得良好的萃取品質，提供工程圖影像相當多的實作經驗。

參考文獻

1. 邱智國 (民 96)，車牌辨識系統的研製，國立台灣科技大學電機工程系碩士論文。
2. 梁世昌 (民 99)，工程圖掃描影像之線向量化，大葉大

學資訊工程系碩士論文。

3. 陳翔傑 (民 94)，自動化車牌辨識系統設計，國立中央大學電機工程研究所碩士論文。
4. 楊昌祐 (民 100)，中文扭曲文件影像之還原與文字切割，台灣科技大學電子工程系碩士論文。
5. 鍾國亮 (民 91)，影像處理與電腦視覺，東華書局，台北。
6. Gonzalez, R. C., R. E. Woods and S. L. Eddins (2004) *Digital Image Processing Using MATLAB*, Prentice-Hall, Upper Saddle River, New Jersey.
7. Hsieh, J. W., S. H. Yu and Y. S. Chen (2002) Morphology-based license plate detection from complex scenes. *IEEE Journal of Electronic Imaging*, 11(4), 507-516.
8. Lee C. H., K. L. You and Y. P. Lin (2010) Dynamic real-time license plate recognition. *Journal of Technology*, 25(2), 151-165.
9. Sedighi A. and M. Vafadust (2011) A new and robust method for character segmentation and recognition in license plate images. *Expert Systems with Applications*, 38(11), 13497-13504.
10. Suryanarayana, P. V., S. K. Mitra, A. Banerjee and A. K. Roy (2005) A morphology based approach for car license plate extraction. *Proceedings of India Conference*.

收件：107.07.17 修正：107.10.03 接受：107.12.05

