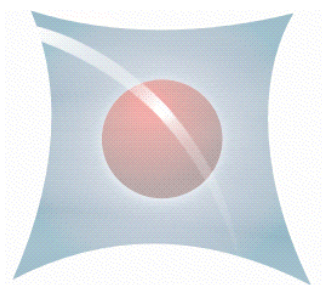


國立中興大學

生物產業機電工程學系
專題研究報告

影像處理與手勢辨識



指導教授：黃國益 博士

學生：董璟和

標若安

指導教授：_____博士

中華民國 105 年 1 月

摘要

本專題利用影像處理技術建立一套能夠辨識簡易手勢的系統。於預設為深色背景的控制環境下，利用隨手可得的 Webcam 擷取影像，將背景預先拍攝後統計背景的像素灰階分布，作為該次取像的二值化門檻參考值。擷取之手勢影像經由影像前處理、手勢輪廓外型萃取、指蹼判斷進行手勢辨識，最後於辨識軟體的視窗上顯示最終判斷結果。

指蹼係指手指間相連的蹼狀皮，本專題定義手指間凸包線切線長與對應垂直深度之比值為指蹼係數，依此特徵值作為分類基準，可辨識出影像中不同的指蹼形狀，分別為剪刀、石頭、布等三種不同手勢。樣本數包括剪刀手勢 24 張、石頭手勢 21 張、布手勢 21 張，由試驗結果得知，其總體辨識率為 78.8%。

目錄

摘要.....	2
目錄.....	3
圖目錄.....	5
表目錄.....	6
第一章 前言.....	7
1.1 研究動機.....	7
1.2 研究目的.....	7
第二章 文獻探討.....	8
2.1 背景分離.....	8
2.2 手勢之辨識.....	8
第三章 基本原理.....	9
3.1 基本影像處理原理.....	9
3.1.1 二值化運算.....	9
3.1.2 膨脹運算.....	10
3.1.3 侵蝕運算.....	10
3.1.4 補洞運算.....	11
3.1.5 直方圖等化.....	12
3.2 鏈碼搜尋.....	12
第四章 研究方法.....	15
4.1 手勢影像處理流程.....	15
4.2 RGB 圖層分離.....	16
4.3 直方圖等化與二值化.....	17
4.4 去除雜訊.....	18
4.5 鏈碼追蹤 (Chain Code Following).....	18
4.6 凸包 (Convex Hull).....	19
4.7 計算指蹠係數與分類手勢.....	19
第五章 設備與樣本.....	21
5.1 研究設備與樣本.....	21
5.1.1 硬體設備.....	21
5.1.2 軟體設備.....	21
5.1.3 實驗樣本.....	21
5.2 UI 操作.....	22
5.2.1 手勢影像擷取.....	22
5.2.2 使用方法.....	22

第六章 結果與討論.....	25
6.1 影像處理結果.....	25
6.2 結果分析.....	27
第七章 結論與未來展望.....	29
7.1 結論.....	29
7.2 未來展望.....	29
參考文獻.....	30

圖目錄

▲圖 1	二值化運算.....	9
▲圖 2	膨脹運算.....	10
▲圖 3	侵蝕運算.....	10
▲圖 4	補洞示意圖.....	11
▲圖 5	直方圖等化示意圖.....	12
▲圖 6	八方鏈碼示意圖：起始點 S	13
▲圖 7	凸包示意圖.....	13
▲圖 8	凸包輪廓搜尋示意圖.....	14
▲圖 9	辨識流程圖.....	15
▲圖 10	原始影像.....	16
▲圖 11	R 圖層分離.....	16
▲圖 12	針對 R 圖層進行直方圖等化.....	17
▲圖 13	二值化處理後之二值影像.....	17
▲圖 14	去除雜訊後之影像.....	18
▲圖 15	手勢影像之輪廓鏈碼，小圓為搜尋起點.....	18
▲圖 16	對輪廓進行凸包之結果示意圖.....	19
▲圖 17	指蹼係數 L_i 與 D_i 位置	20
▲圖 18	由程式介面查看指蹼係數.....	20
▲圖 19	程式初始 UI	22
▲圖 20	點選 Format	23
▲圖 21	裝置設定.....	23
▲圖 22	執行 Check 並辨識後的結果	24
▲圖 23	輪廓萃取完整而辨識正確的影像.....	26
▲圖 24	影像陰影部分造成輪廓侵蝕.....	28
▲圖 25	二值化處理後形態不完整.....	28
▲圖 26	手指間隙合併、背景未正確分離之示意圖.....	28

表目錄

表 1.1.a	辨識結果統計.....	25
表 1.1.b	各手勢影像之辨識率.....	25

第一章 前言

1.1 研究動機

現代教學及簡報方式逐漸拋棄傳統板書而使用投影、大型觸控螢幕等電子設備，但是仍不乏需要直接碰觸到實體控制器才能夠進行翻頁的控制。在科技發展趨勢下，遙控手段已不滿足於紅外線、藍芽等方式，坊間 XBOX 等體感偵測遊戲主機導入機器視覺。相較於按鈕等硬體操作介面，肢體語言在使用上更加直覺，許多大型科技龍頭如 Intel、Apple、Sony、Microsoft 等等公司，相繼投入研發手勢操控專利，從已普及的觸控螢幕延伸到 3D 力度觸控甚至 4D 空間偵測。

手部是人體部位中最靈活的形態表現器官，不同的手勢便能有效地成為獨立指令，藉由機器視覺可達成非接觸式的控制手段。本專題利用影像處理技術及程式撰寫工具 Visual Basic 6.0 達到簡單手勢的辨識方法。

1.2 研究目的

針對靜態的擷取影像，以不使用現成辨識軟體或函式庫為前提，進行辨識軟體的撰寫。撰寫軟體應能辨別剪刀、石頭、布等三種手勢。進一步發展即時、動態式的辨識軟體。

第二章 文獻探討

2.1 背景分離

與背景分離之文獻如下所述。Cheung 等人[1] 將動態物件偵測與靜態背景分離運用在交通監視系統之研究。將監視器連續畫面經過相減運算，可得到變動中的影像範圍。並對此類變動像素進行形狀特徵萃取，分離出車輛之形狀。吳等人[2] 研究手臂大範圍動作手勢用於遙控音樂，動作如機場旗號人員的動作，混和雙臂平舉與上舉。背景分離部分採用預存背景，將即時影像與背景相減之方式進行軀幹的尋找。黃等人[3]將影像轉換為 YCbCr 色彩空間，進行適當訓練資料建立在直方圖上，再經由貝氏分類器(Bayes classifier)將膚色與背景進行分離。

2.2 手勢之辨識

吳等人[2] 將變動像素個數積分至 X 軸與 Y 軸，根據累積像素之寬窄範圍可判斷為何種旗號動作。陳等人[4]運用特製資料手套，同時運算手指彎曲動作與影像處理，藉由資料手套內的加速度感測器讀取手勢，再配合影像處理進行雙重驗證，但多一層手套會導致負重增加、靈活度下降。

第三章 基本原理

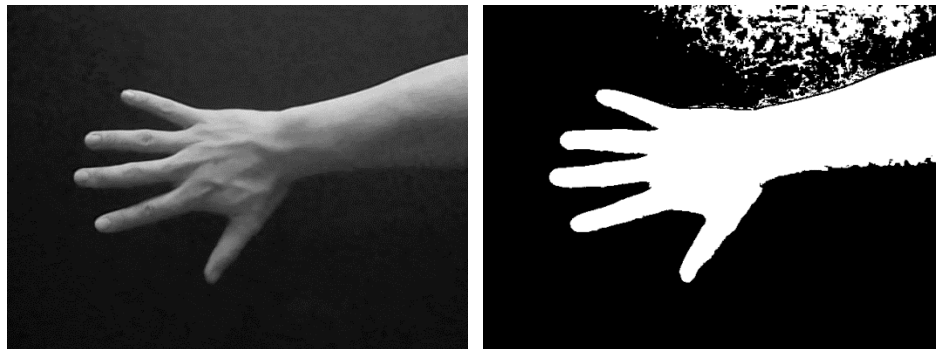
3.1 基本影像處理原理

本專題使用的基本影像處理運算[7] 包括二值化(binary) 、膨脹(dilation) 、侵蝕(erosion)、補洞運算及直方圖等化(Histogram Equalization)，其基本原理如下：

3.1.1 二值化運算

二值化運算是將灰階影像以一指定灰階值為門檻，將其他灰階值轉換 0 與 255 所組成的二值影像。令原始影像 $f(x,y)$ 及門檻值(Threshold Value)為 T ，且介於 $[0,255]$ ， $f_t(x,y)$ 為 $f(x,y)$ 經二值化運算後的影像(如圖 2.1 所示)，其定義如下所示：

$$f_t(x,y) = \begin{cases} 255, & \text{if } f(x,y) \geq T \\ 0, & \text{if } f(x,y) < T \end{cases}$$



(a) 灰階影像

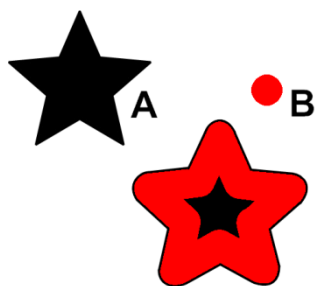
(b) 二值影像

▲圖 1. 二值化運算

3.1.2 膨脹運算

膨脹運算主要用於二值化影像，將定義為物體之灰階值像素點往鄰近的像素點擴大，如圖 2 所示，A 為物體，B 為遮罩，以 B 中心走訪 A 之邊界，圖形向外擴大，其定義如下：

$$A \oplus B = \{x | (\hat{B}_x) \cap A \neq \Phi\}$$

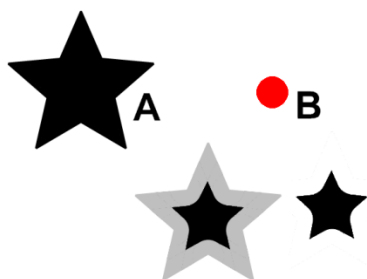


▲圖 2 膨脹運算

3.1.3 侵蝕運算

侵蝕運算主要用於二值化影像，將定義為物體灰階值之像素點由邊緣向內進行收縮，如圖 3 所示，A 為前景物體，B 為遮罩，以 B 中心走訪 A 之邊界，圖形向內縮減，其定義如下：

$$A \ominus B = \{x | (B_x) \subseteq A\}$$

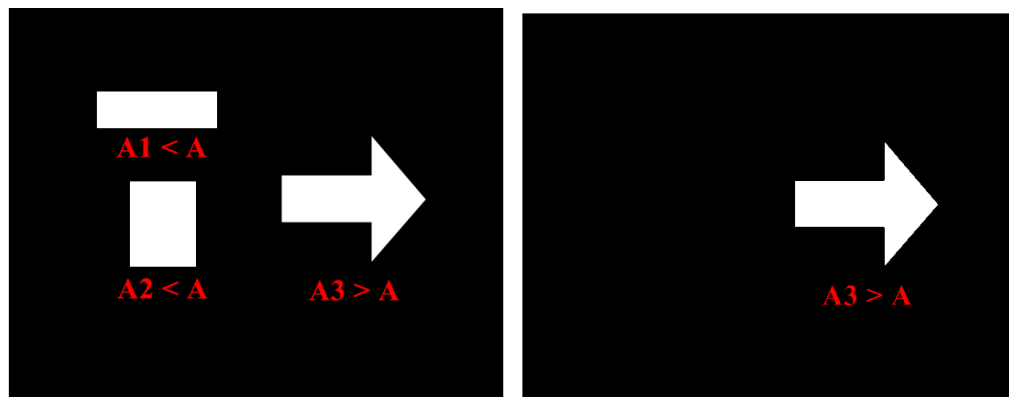


▲圖 3 侵蝕運算

3.1.4 補洞運算

補洞運算為物體與背景的灰階值替代原理，應用於二值化影像，將物體影像的雜訊或孔洞進行填補。將影像中的灰階值 0 視為主體，灰階值 255 視為孔洞，以面積像素點個數作為補洞之門檻依據，進行補洞運算，如圖 4 所示，步驟如下：

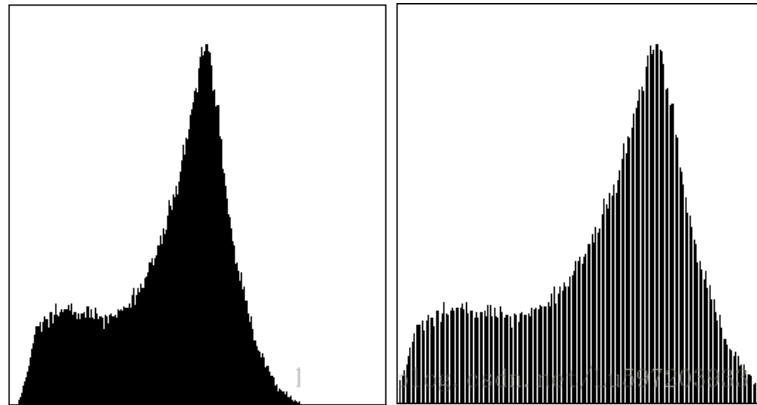
- (1) 將灰階值 0 視為主體，灰階值 255 之物體視為背景孔洞。
- (2) 計算孔洞面積 A_i ， $i=1\sim N$ 。
- (3) 設定補洞之面積門檻值 A 。
- (4) 若孔洞面積 A_i 小於門檻值 A ，則將孔洞之灰階值以主體灰階值 0 替代；反之，若孔洞面積大於門檻值則不改變其灰階值，如此即可完成補洞運算。
- (5) 灰階值 255 之主體之孔洞同理以灰階值 255 填補。



▲圖 4 補洞示意圖

3.1.5 直方圖等化

直方圖等化之概念為將影像灰階值 $[f_{\min}, f_{\max}]$ 擴展為 $[0, 255]$ 。本專題針對影像進行灰階值統計，影像中最亮之灰階值為 f_{\max} ，最暗灰階值為 f_{\min} ，對此 $[f_{\min}, f_{\max}]$ 區間擴展為 $[0, 255]$ ，如圖 5 所示。



▲圖 5 直方圖等化示意圖

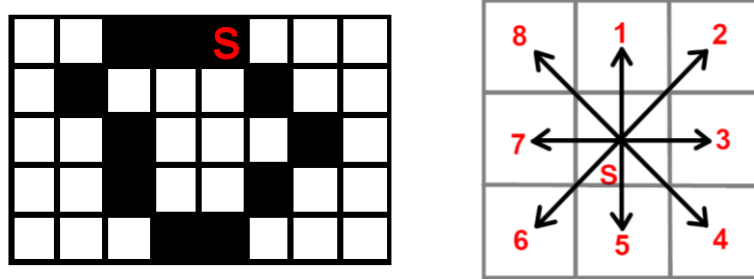
3.2 鏈碼搜尋

對於數位影像而言，物體邊界的描述是很重要的，若能紀錄物體邊界上之座標點即可對各座標點進行運算，由此得到形狀特徵。本專題使用八近鄰的邊界追蹤法紀錄手勢輪廓座標。

鏈碼搜尋輪廓方法如下：

- (1) 取得輪廓外緣上一像素點為起點後，鄰八個像素以順時針方向排成序列給予 1、2、3.....7、8 的編號，如圖 6。八近鄰兩兩比較灰階值，若灰階值有變化表示為輪廓點，若灰階值無變化則依序列比較下一點，直到判斷為輪廓為止。
- (2) 取得下一點輪廓後，記錄方向序列與其二維座標留用。
- (3) 因為邊緣順序特性，在走訪下一個點之鄰八格時，將起始序列+2，避開重覆部分，繼續以順時鐘方向尋找下一個邊緣像素。

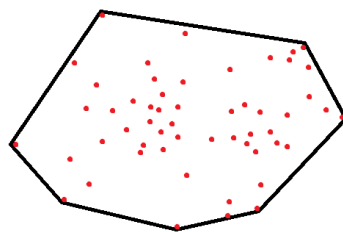
- (4) 假使判斷下一輪廓點座標與起點相同，則表示已完成走訪。
- (5) 一個圖案完成八方鏈碼走訪後，可用連續的方向序列碼來描述其外型。在不改變圖案比例大小的狀態下，不論存在於圖面的何處，其鏈碼均相同，可用做特徵或辨識的方法。



▲圖 6 八方鏈碼示意圖：起始點 S

3.3 凸包運算

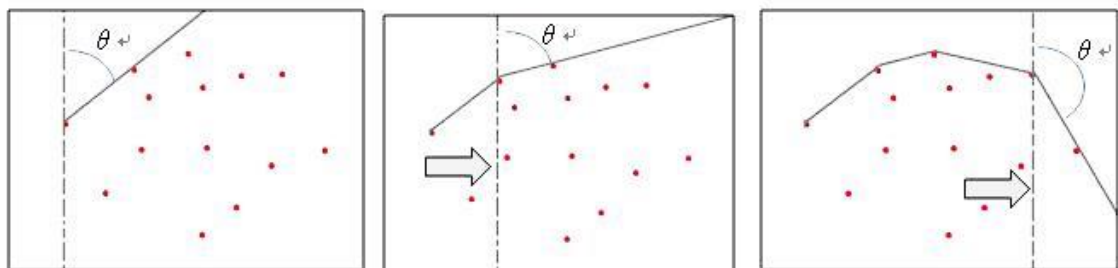
凸包之定義為：在二維空間中有一群散佈的點，凸包為可涵蓋這些點的最小封閉表面。其定義使凸包只會呈現凸出而不會產生凹陷形狀，如圖 7 所示。凸包運算可以擴展至三維空間應用，定義不變。



▲圖 7 凸包示意圖

本專題所使用之凸包演算法改寫自 Andrew 的凸包演算法(Monotone ChainConvex Hull Algorithm)[6]，係由餘弦定理中的三角型三個已知邊長可得任一指定內角角度之特性，如圖 8 所示。演算法步驟如下：

- (1) 將所有輪廓點由左而右，由上而下進行排序。
- (2) 選取影像最左邊的像素點為起始點，以垂直線為基準，計算往後每個像素點方向與垂直線的夾角 θ 大小，夾角最小者為凸包點，記錄下個像素點座標值 $H(x,y)$ ($i=0\sim N$)，與此點的鏈碼值 H_i ，並將起始點移至此點，計算往後每個像素點方向與垂直線的夾角。
- (3) 直到最後一個點，凸包上半部即記錄完成。
- (4) 依據步驟(1)之排列順序，由右至左反序運算即可得到下半部凸包，記錄像素點座標時是接續上半部凸包序列，序列整體是按照凸包點順時鐘排列而成。
- (5) 直到最左側掃描結束，即得到完整的凸包影像。

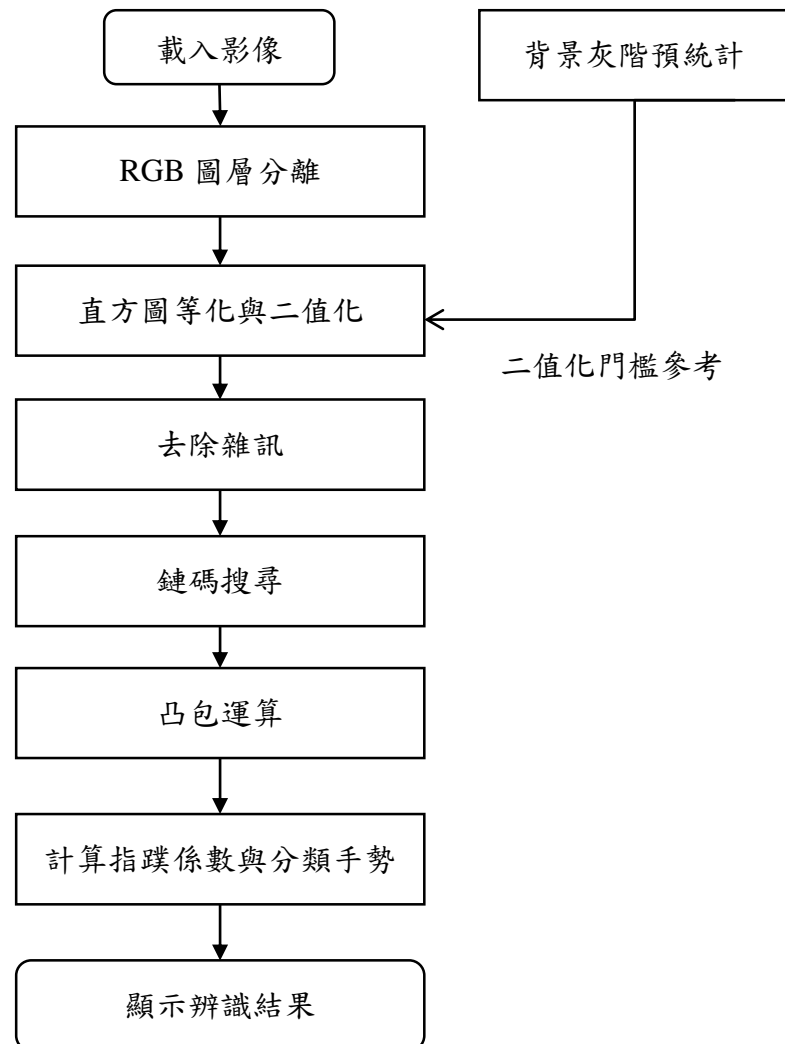


▲圖 8 凸包輪廓搜尋示意圖

第四章 研究方法

4.1 手勢影像處理流程

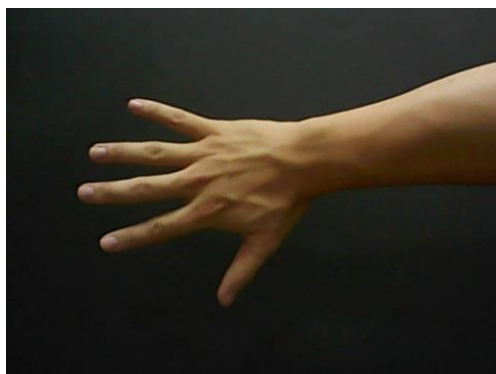
本專題的影像擷取後，經由一系列處理方法包括 RGB 圖層分離、直方圖等化、二值化、去除雜訊、鏈碼搜尋、凸包運算、計算指蹼係數與分類手勢。完整流程圖如圖 9 所示，各方法詳細內容如下：



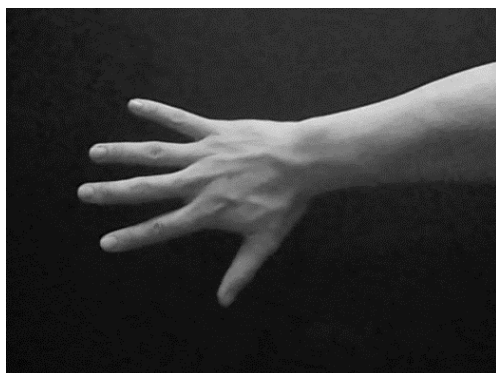
▲圖 9 辨識流程圖

4.2 RGB 圖層分離

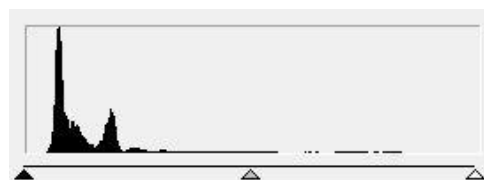
將由 webcam 擷取獲得之彩色影像(如圖 10 所示) 轉換為紅(Red)、綠(Green)、藍(Blue)三圖層。本專題背景灰階預統計係採用 G 圖層，另將 R 圖層分離(如圖 11 所示) 並針對 R 圖層進行後續處理。程式起始時，



▲圖 10 原始影像



(a) R 圖層

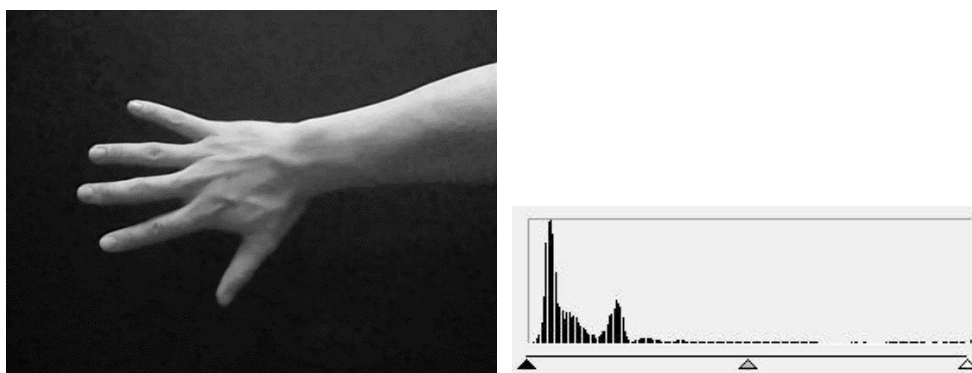


(b)R 圖層之灰階分布直方圖

▲圖 11 R 圖層分離

4.3 直方圖等化與二值化

針對 R 圖層進行直方圖等化，如圖 12 所示。將程式一開始時統計背景灰階結果作為二值化的參考依據，對前述影像作二值化處理，如圖 13 所示。.



(a) 直方圖等化後之影像

(b) 等化後之直方圖分佈

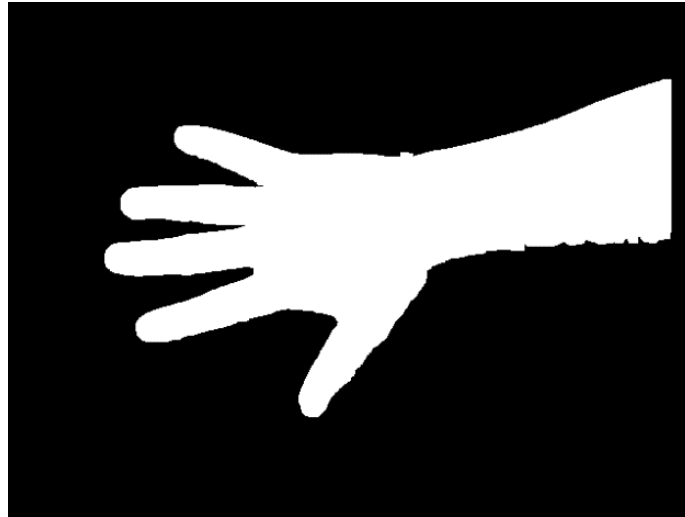
▲圖 12 針對 R 圖層進行直方圖等化



▲圖 13 二值化處理後之二值影像

4.4 去除雜訊

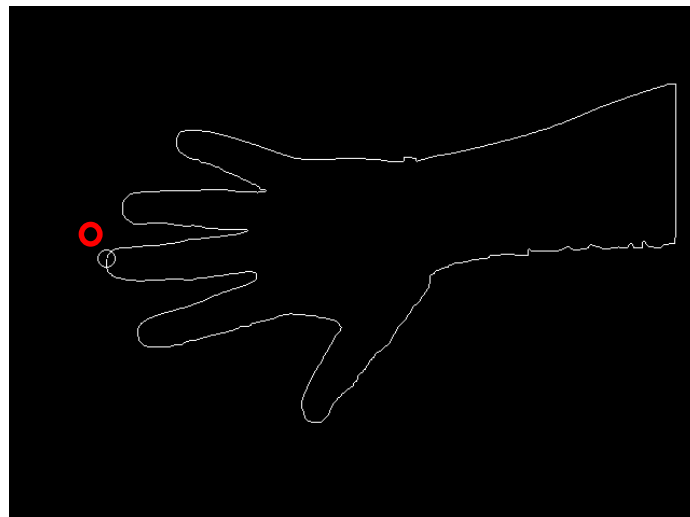
本專題利用膨脹(dilation) 與侵蝕(erosion) 運算針對手勢影像邊緣鋸齒狀進行填補，並以補洞運算消除背景與手勢影像內之雜訊，處理結果如圖 14 所示。



▲ 圖 14 去除雜訊後之影像

4.5 鏈碼追蹤 (Chain Code Following)

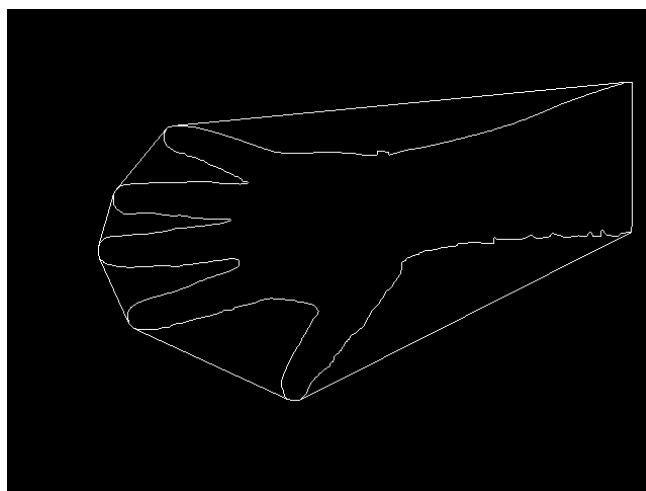
影像處理技術中，鏈碼為可以表示物體外型的方法，本專題使用八近鄰鏈碼追蹤記錄手勢輪廓的每個像素點之座標，以供後續運算，如圖 15 所示。



▲ 圖 15 手勢影像之輪廓鏈碼，小圓為搜尋起點

4.6 凸包 (Convex Hull)

對前述鏈碼追蹤所紀錄的輪廓進行凸包運算，如圖 16 所示。依據鏈碼追蹤取得的外輪廓座標陣列及其個別的序列編號、凸包點的座標陣列進行整合，可得到連接手指間端的凸包切線段與該凸包點的鏈碼之間的對應關係。



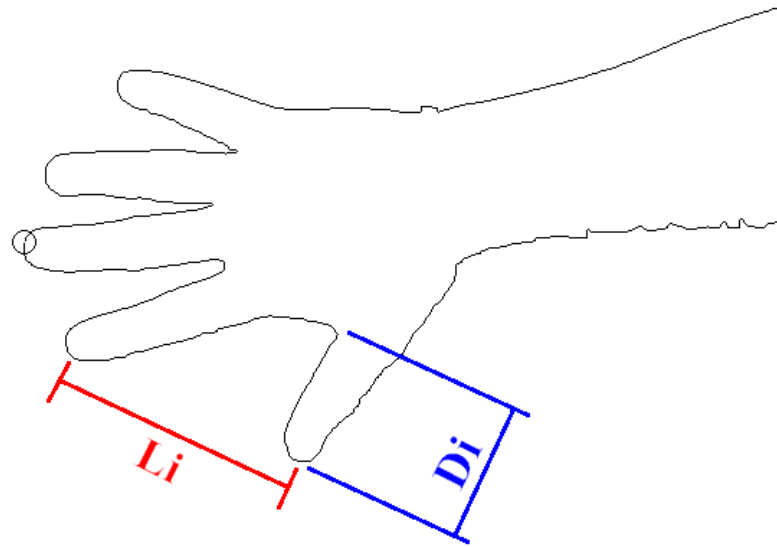
▲圖 16 對輪廓進行凸包之結果示意圖

4.7 計算指蹠係數與分類手勢

計算每一段凸包切線長度 L_i , $i=1\sim n$ 與間隙的最大深度 D_i , $i=1\sim n$ ，兩者之比，稱為指蹠係數，見圖 17。此係數在手指伸出的間隙凹下距離相較於手腕及收合手指的邊緣凹陷為大，以係數大小判斷為手指之間隙或為其他凹下邊緣，其他之定義包含：輪廓侵蝕不完整、手腕內縮處、收合手指的殘留輪廓、手指與拳的切線段。

$$\text{指蹠係數} = D_i / L_i$$

將每張手勢影像中指蹠係數的數值統計，如圖 18 所示，能得到指蹠與其他凹下邊緣的分界，進而計算此手勢影像有幾個「指蹠」。由於基本手勢的指蹠數目皆異，因此得出影像的指蹠數目就能判斷為何種手勢。



▲圖 17 指蹠係數 L_i 與 D_i 位置

凸包點編號	指蹠係數
0	.15
1	.00
2	.00
3	.03
4	.44
5	.00
6	.00
7	1.14
8	1.83
9	.00
10	.07
11	1.35
12	.32
13	.62
14	.22
15	.00

▲圖 18 由程式介面查看指蹠係數

第五章 設備與樣本

5.1 研究設備與樣本

5.1.1 硬體設備

本專題擷取手勢影像所使用的系統硬體設備為:個人電腦與網路攝影機。選用網路攝影機而非 CCD 攝影機主要因為 CCD 攝影機主要為擷取高解析度的影像，在手勢辨識上並不需要如此規格，且 CCD 攝影機價格昂貴。因此本專題選用物美價廉的網路攝影機作為擷取工具。硬體設備詳細規格如下方:

1. 個人電腦 : CPU Intel Core i5-3210M , ACER 筆記型電腦 E1-571G
2. 網路攝影機 : Logitech C170 羅技網路攝影機

5.1.2 軟體設備

本專題所設計之系統建置於 Windows7 專業版作業系統上，使用 Microsoft Visual Basic 6.0 與影像函式庫 MIL 5.12 (Matrox Imaging Library 5.12) 進行程式之撰寫與設計。軟體詳細規格如下方:

1. 作業系統: Microsoft Windows 7，經測試可相容於 Windows 8.1、Windows 10
2. 程式語言: Microsoft Visual Basic 6.0
3. 影像處理函式庫: MIL 2.12，Matrox 出品

5.1.3 實驗樣本

本專題的樣本為控制環境下擷取之 66 張影像，樣本手勢包含剪刀、石頭、布等三種及多種旋轉角度。本研究仍在陸續蒐集影像中，包含數字手勢 一、三、四、六等，可提供於後續的研究進行測試。

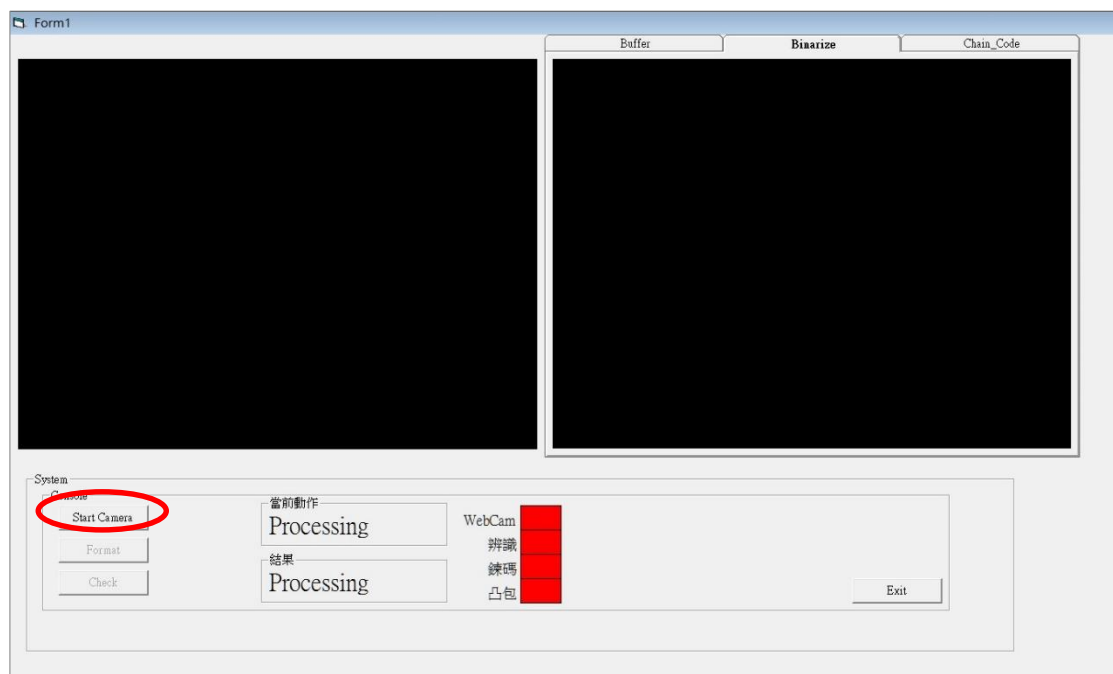
5.2 UI 操作

5.2.1 手勢影像擷取

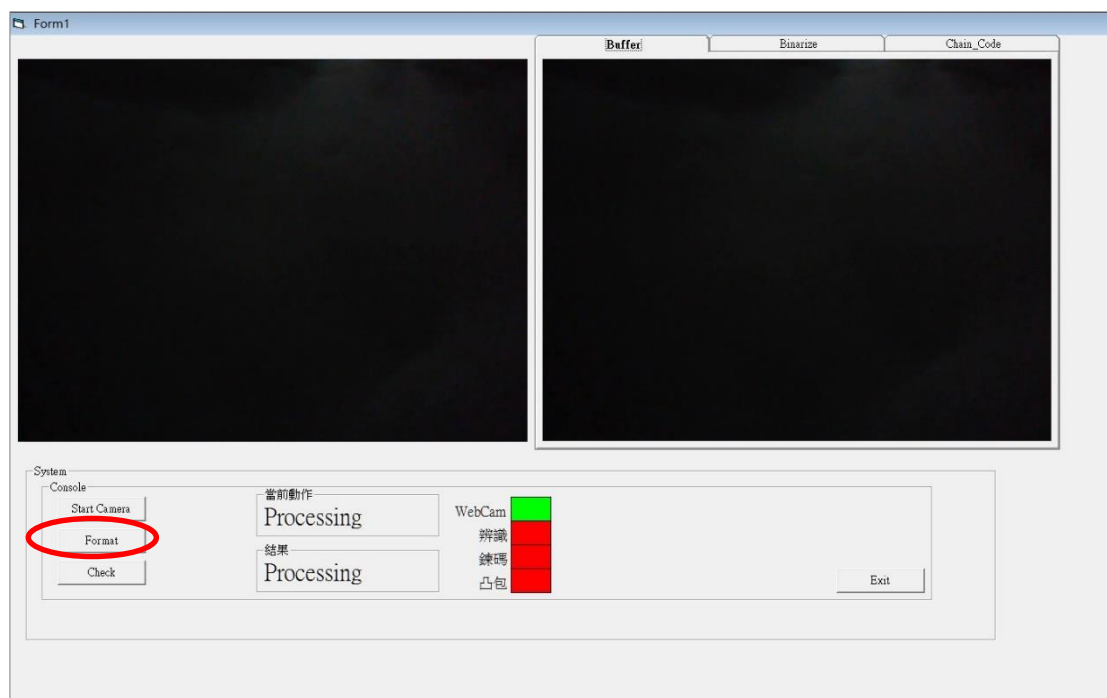
本專題之目的為手勢輪廓之萃取與處理而非手部表面細節，因此以 Webcam(型號 Logitech C170)擷取之影像即可作為本專題使用，影像大小為 640×480 pixels。透過程式啟動 Webcam 後，程式自動把影像存取成 bmp 檔，之後再進行辨識。

5.2.2 使用方法

點選 Start Camera (圖 19)，啟動 Webcam 後,等待 Format 按鈕 ON 並點選 (圖 20)，點選裝置設定內的亮度設定為最低，白平衡設定為最高(圖 21)。之後點選 Start Camera，等待 Check 按鈕 ON，再放入手勢，按下 Check(圖 22)。程式便會進行辨識。



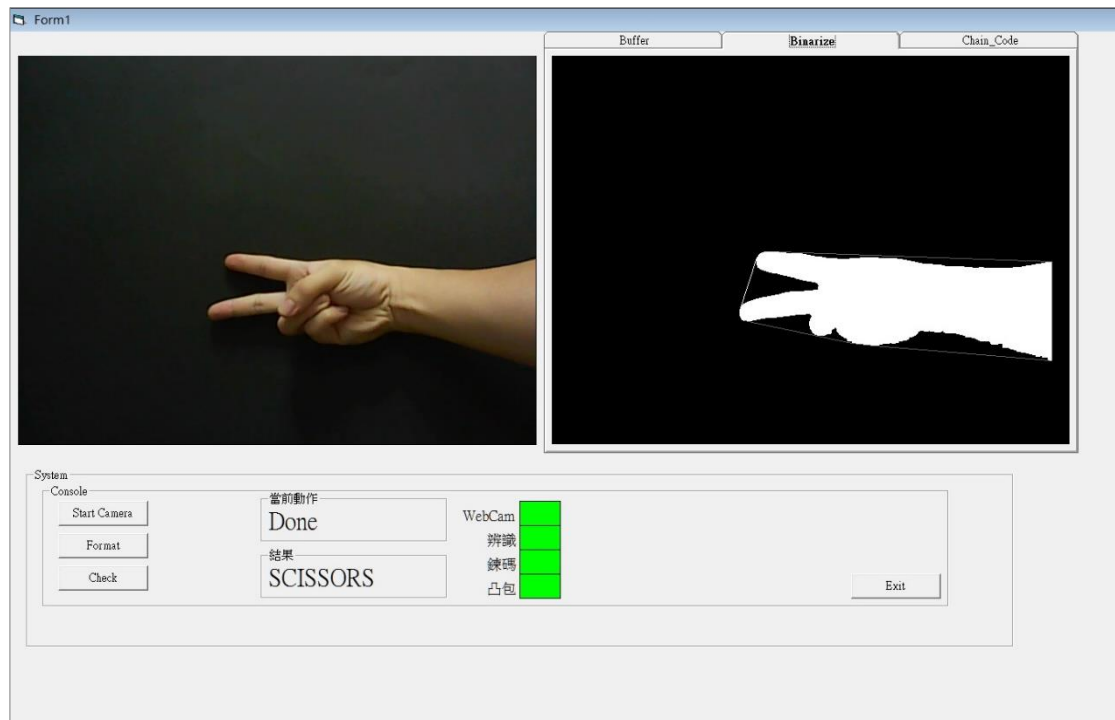
▲圖 19 程式初始 UI



▲圖 20 點選 Format



▲圖 21 裝置設定



▲圖 22 執行 Check 並辨識後的結果

第六章 結果與討論

6.1 影像處理結果

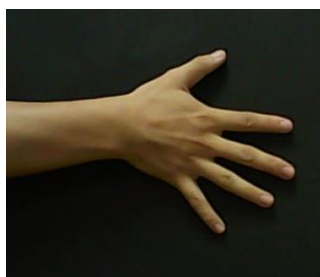
本專題系統對旋轉 20 度內且手指不重疊的手勢辨識率為 77.8% (66 張基本手勢影像中有 52 張正確，見表 1.1)。正確辨識定義為指蹠位置與個數同時標示正確，如圖 23 所示。若是指蹠係數錯誤或位置標示錯誤則歸類為辨識錯誤，常發生於手指間隙合併或重疊之影像，如圖 26 所示。

表 1.1.a 辨識結果統計

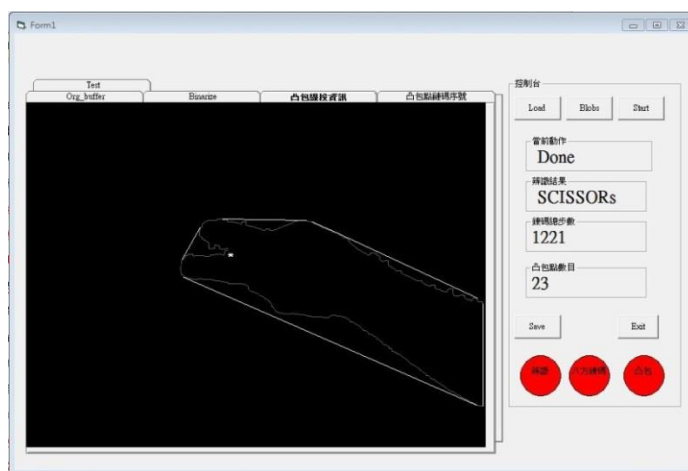
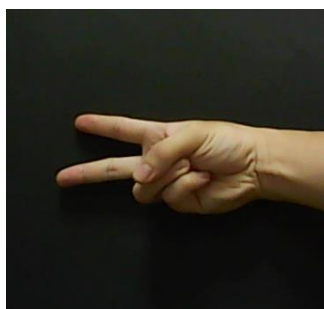
	剪刀	石頭	布	總數 67
辨識正確	16	17	19	52
辨識錯誤	8	4	2	14
合計	24	21	21	辨識率 77.8%

表 1.1.b 各手勢影像之辨識率

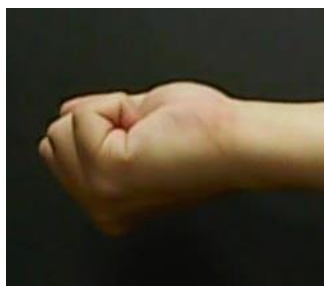
	剪刀	石頭	布
辨識率	66.7%	81.0%	90.5%



(a) 布手勢正確辨識結果



(b) 剪刀手勢正確辨識結果



(c) 石頭手勢正確辨識結果

▲圖 23 輪廓萃取完整而辨識正確の影像

6.2 結果分析

辨識錯誤之原因為於二值化處理與去除雜訊過程造成手勢輪廓缺陷與形態上之破損，因此無法正確計算出指蹠係數，對上述所提問題以下分別進行說明。

(1) 陰影造成二值化手勢形態不完整

因為影像拍攝環境受日光燈影響，日光燈與拍攝平面之夾角約為 40 度角，缺少正面光，手勢近地面處易形成陰影，如圖 24 所示。即便有經過直方圖等化仍會於二值化步驟造成輪廓破損，雖然膨脹運算能稍微彌補破損之輪廓，但此方法不適用所有影像。

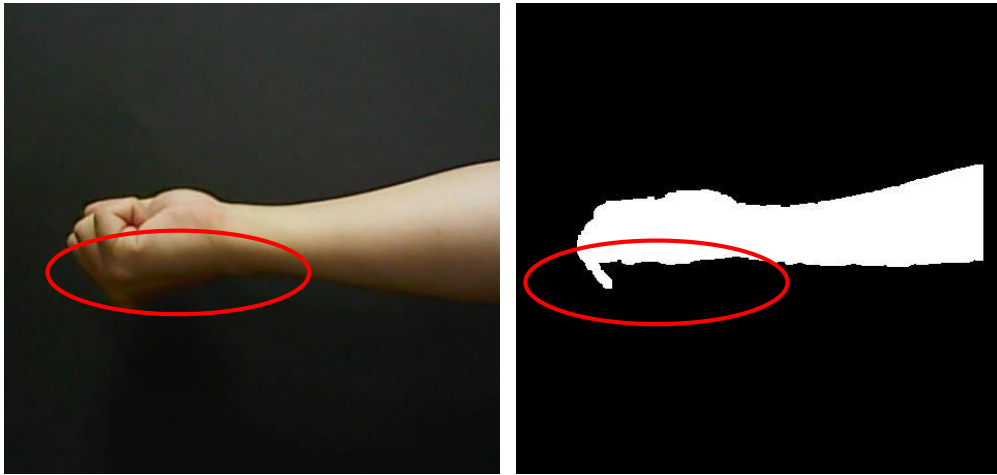
由表 1.1 可知剪刀及石頭手勢辨識率偏低，按照影像處理流程步驟查看影像，歸納出一個主要原因，即剪刀與石頭手勢影像包含掌心握拳處時，拳心有大片陰影導致二值化處理過程無法保留完整手勢形態，如圖 25 所示。

本專題之針對整張影像進行二值化處理，除造成手勢破碎外，少數影像之背景無法以二值化處理去除，如圖所示。後續將加入區域二值化與 OTSU 方法，使系統自動取得影像各區域的最佳二值化門檻值，並於拍攝正面適當補光。除此之外也可以採取每次擷取兩張，一張預先儲存背景，一張手勢影像對背景進行相減運算即可得出手部影像。

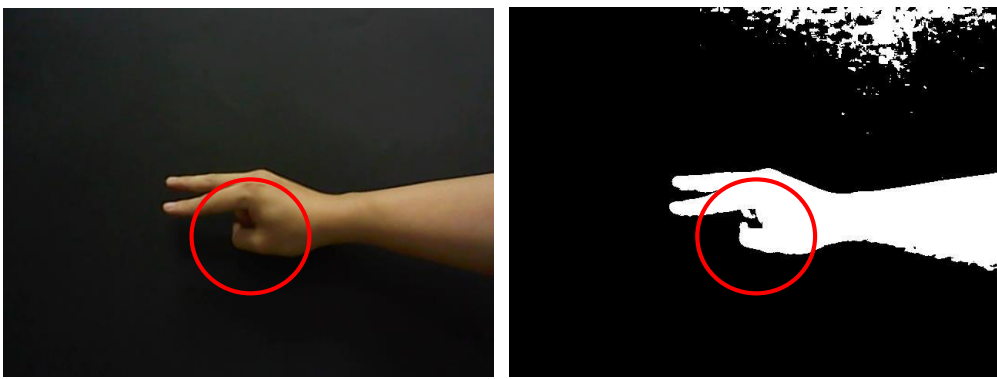
(2) 手指縫隙合併

載入之影像皆以相同流程處理，當原始影像之手指間隙較小時，經過膨脹運算後所得的手指間隙呈現相連、合併之狀態，如圖 26 所示，注意原始影像中食指與中指中間確實有間隙，但進行二值化及膨脹運算後食指及中指合併。本專題利用手指間距與深度計算指蹠係數，缺少正確手指輪廓會影響指蹠位置與個數的判斷。

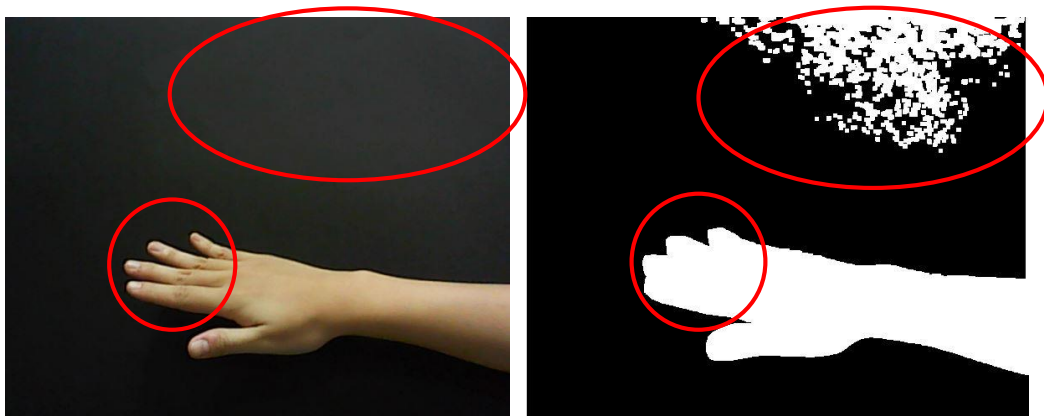
對於圖形內部的黑色點狀雜訊而言，膨脹是很簡單的解決方法，背景的白色點狀雜訊用斷開(Opening)亦能很輕易地去除，唯處理強度、次數與順序需要多方考慮與嘗試。



▲圖 24 影像陰影部分造成輪廓侵蝕



▲圖 25 二值化處理後形態不完整



(a)原始影像

(b)經前處理後之影像

▲圖 26 手指間隙合併、背景未正確分離之示意圖

第七章 結論與未來展望

7.1 結論

本專題後續研究將針對上述分析作出應對，首要是將 OTSU 方法自動選取最佳門檻值加進系統流程中，以提升二值化處理後輪廓的完整度，另外幾點需改進如下：

1. 系統介面新增控制項與監看表格以監控影像處理細節
2. 增加拍攝至少 15 人的基礎手勢影像，總樣本數需有 300 張以上
3. 建立系統數據庫，以實際資料庫統計最佳指蹠辨識分界

7.2 未來展望

在未來，遙控之趨勢定會走向無硬體無接觸式感測器，因此運用機器視覺取代傳統感測器就更顯得重要。另外，手勢、人臉辨識等在坊間已有許多開源函式庫可供參考，可說是日新月異，發展相當快速。

目前，本專題仍在靜態的影像處理分析階段，為了能夠實際應用於控制，期望改為動態即時辨識，並且增加可辨識之手勢數量，使程式本身可用性更加靈活。

參考文獻

- [1] Sen-Ching S. Cheung, Chandrika Kamath. Robust techniques for background subtraction in urban traffic video. EURASIP Journal on Applied Signal Processing. 2005. 1: 2330-2340.
- [2] 吳怡明。手勢辨識應用於遙控音樂播放系統。碩士論文。2009。國立台灣科技大學電機工程系。
- [3] 黃雅軒、王勻駿。基於階層式時序記憶的多角度手勢辨識方法。碩士論文。2011 中華大學資訊工程學系。
- [4] 陳治宇。虛擬滑鼠：以視覺為基礎之手勢辨識。碩士論文。2003。國立中山大學資訊工程學系。
- [5] 陳旻廷。以資料手套輸入裝置之手勢操控虛擬人物系統之建構。碩士論文。2007。中原大學工業工程學系。
- [6] A.M.Andrew, Another efficient algorithm for convex hulls in two dimensions, 1979. 9(5): 216-219.
- [7] Gonzalez, Rafael C., Richard E. Woods, Digital image processing. 3rded. Taipei:Pearson Education Taiwan Ltd.