

致理技術學院

致理技術學院資訊網路技術系實務專題報告

資訊網路技術系 實務專題報告

室內環境人臉偵測

室內環境人臉偵測

指導老師：高楊達 老師

學生：施博銘(29534505)

呂柏緯(29534507)

林鼎軒(29534527)

張寄園(29534528)

九十六年十二月

中華民國 96 年 12 月

致理技術學院

資訊網路技術系 實務專題報告

室內環境人臉偵測

指導老師：高楊達 老師

學生：施博銘(29534505)

呂柏緯(29534507)

林鼎軒(29534527)

張寄園(29534528)

本成果報告書經審查及口試合格特此證明。

指導老師：_____

中華民國 96 年 12 月

專題研究授權書

本授權書所授權之專題研究為_____

共_____人，在致理技術學院資訊網路技術系 _____學年度第_____學期完成資網實務專題。

專題名稱：_____

同意 不同意

本組同學共_____人，皆同意著作財產權之論文全文資料，授予教育部指定送繳之圖書館及本人畢業學校圖書館，為學術研究之目的以各種方法重製，或為上述目的再授權他人以各種方法重製，

不限地域與時間，惟每人以一份為限。

上述授權內容均無須訂立讓與及授權契約書。依本授權之發行權為非專屬性發行權利。依本授權所為之收錄、重製、發行及學術研發利用均為無償。上述同意與不同意之欄位若未勾選，該組同學皆同意視同授權。

指導教師姓名：

專題學生簽名：

(親筆正楷)

學號：

(務必填寫)

中華民國 年 月 日

誌 謝

首先誠摯的感謝指導老師高楊達老師，老師教學嚴謹，為人務實認真，在老師悉心的教導使我們得以一窺影像處理領域的深奧，在我們遇到瓶頸時，也適時地提供建議及思考方向，讓我們如亦師亦友般的情誼，在求學的路上可以循著正確的方向邁進。

此外特別提出對本系系主任及系上所有師長的感謝，感謝他們這些日子以來的陪伴，並給予每位專題組員適時的指導，對學生於專題研究室，甚至是各項實驗的儀器也能給予很大的協助，讓學生能無後顧之憂地隨時使用專題研究室。

感謝本組組員們的配合與協助，能運用自身之知識與經驗，在遭遇困難時能以冷靜的態度與實事求是的精神面對，在困難時彼此激勵與幫助，由於大家互相配合，努力耕耘及孜孜不倦的學習精神，才得以完成這份專題。

最後還要感謝家人、同學、學長與朋友們，在這段期間內所給予的支持與鼓勵。在此，僅以此論文獻給所有幫助過，指導過我們的人。

摘 要

人臉偵測方面技術在各影像領域佔有一席之地，在視聽娛樂、照相留念、戶口資料管理進而保全系統皆需要以此技術為基礎建設出完整的系統。本論文以 C++ Builder 6.0 開發人臉偵測操作介面來實現影像人臉偵測的技術，以室內影像人臉為主要實驗方向，雖然不受外界光線的影響，但在室內中也會受到燈光強弱與陰暗遮蔽來左右偵測準確度。影像人臉偵測中的困難大多是如何判別移動物體為人體，又如何過濾膚色、過濾背景，之後就是從偵測過濾後的膚色上搜尋人臉、辨識特徵等等。實驗中為了克服種種困難，我們將以 YCbCr 色彩空間來偵測膚色，利用 Extra 軟體把影像切割成一張張的圖片，之後再利用 C++ Builder 6.0 開發的介面系統來過濾背景及異物，再讓系統依據特徵辨識與預設條件來判別是否符合人臉條件。

關鍵字：人臉偵測、過濾膚色、過濾背景、辨識特徵、YCbCr 色彩空間

目 錄

專題研究授權書.....	i
誌 謝.....	ii
摘 要.....	iii
目 錄.....	iv
圖 目 錄.....	vi
表 目 錄.....	ix
第壹章 緒論	1
第一節 研究動機與目的	1
第二節 研究重點[1]	3
第三節 研究範圍與報告書架構.....	6
第貳章 室內人臉偵測之理論與技術探討	8
第一節 YCbCr 色彩空間偵測膚色	8
第二節 HSV 色彩人臉區域偵測.....	10
第三節 YCbCr 與 HSV 之差異.....	12
第四節 技術 1 探討 人臉偵測技術研究	13
第五節 技術 2 探討 彩色視覺技術於多人人臉辨識上的應用	20
第六節 技術 3 探討 利用膚色及區域極小值作人臉特徵是否遮蔽之偵測判斷.....	28
第參章 室內環境人臉偵測之方法	35
第一節 偵測移動物體	35
(一) 乾淨背景的取得.....	35
(二) 影像差值計算.....	37
第二節 膚色偵測.....	40
第三節 影像二值化 (IMAGE BINARY)	43
第四節 膨脹與侵蝕運算	46

第五節	連通元件(CONNECTED COMPONENT)	49
第六節	條件判斷搜尋候選臉區域	56
第七節	灰階化(GRAY LEVELS).....	60
第八節	搜尋眼嘴之位置	62
	(一) 二值化	62
	(二) 中間值濾波(MEDIAN FILTER)[14]	63
	(三) CONNECTED-COMPONENT 並做條件判斷.....	66
第九節	正規化(NORMALIZE)	68
第十節	支撐向量機器 SVM (SUPPORT VECTOR MACHINE).....	70
第肆章	實驗結果與分析	74
第一節	實驗結果.....	74
第二節	效能分析.....	78
第三節	系統實現的困難與待改處	85
第伍章	結論	87
參考文獻	90

圖 目 錄

圖 1 專題研究架構圖。	7
圖 2 HSV 的錐形色彩空間示意圖。	10
圖 3 膚色分布圖：(a)HSV 色彩空間；(b)YCbCr 色彩空間。	12
圖 4 彩色邊緣偵測流程架構圖。	14
圖 5 Entropic Thresholding 技術結果圖。	15
圖 6 邊緣偵測後結果比較。	15
圖 7 人臉長寬比例校正後之結果。	16
圖 8：(a)膚色比例校正後之結果；(b)邊緣比例校正後輸出結果。	16
圖 9 過濾背景雜訊結果圖。	17
圖 10 偵測結果不受人臉角度影響。	18
圖 11 影像尺寸不一導致邊緣偵測效果不彰。	19
圖 12 受光線影響膚色偵測失敗。	19
圖 13 輸入原始影像。	20
圖 14 膚色在 HSV 之分布圖。	21
圖 15 二值化處理後的影像。	22
圖 16 開閉運算後的影像。	23
圖 17 連通成分後的影像。	24
圖 18 雙線性內插法在原圖上的格子點。	25
圖 19：(a)原始影像；(b)人臉候選區域；(c)求嘴巴轉正角度；(d)轉正後結果。	26
圖 20 人臉比例表示圖。	27
圖 21 臉部偵測後結果。	27
圖 22 人臉追蹤簡易架構圖。	28
圖 23 移動物偵測：(a)前一張影像 I_{t-1} ；(b)現在的影像 I_t ；(c)連續兩張影像灰階值相減；(d)將(c)透過門檻值取二元化；(e)將(b)邊緣化；(f)取(d)跟(e)的交集。	29
圖 24 長短軸比率 1.2 的橢圓與實際人臉。	30
圖 25 搜尋視窗：(a)原影像；(b) Sobel 邊緣化後；(c)偵測後與搜尋視窗；(d)把(c)之搜尋視窗套用到(b)。	31
圖 26 橢圓模組比對：左為橢圓模組；右為欲比對的影像。	31
圖 27 橢圓模組比對過程與結果：(a)~(g)為橢圓模組比對過程；(h)比對後找到分數最高的位置。	32
圖 28 4 個點所形成的橢圓模組與樣板。	33
圖 29 橢圓模組與樣板比對追蹤人臉過程與結果。	34
圖 30 單人圖乾淨背景。	36

圖 31 雙人圖乾淨背景。	36
圖 32 單人影像原圖。	38
圖 33 雙人影像原圖。	38
圖 34 單人影像相減。	39
圖 35 雙人影像相減。	39
圖 36 切割後的純膚色圖片範例。	41
圖 37 單人膚色過濾。	42
圖 38 雙人膚色過濾。	42
圖 39 圖影像二值化函數轉換。	44
圖 40 單人二值化。	45
圖 41 雙人二值化。	45
圖 42 8 鄰居範例圖。	47
圖 43 單人影像膨脹。	48
圖 44 雙人影像膨脹。	48
圖 45 以連通技術搜尋像素點。	50
圖 46 以堆疊來實現連通技術搜尋像素點之範例圖 1。	51
圖 47 以堆疊來實現連通技術搜尋像素點之範例圖 2。	51
圖 48 以堆疊來實現連通技術搜尋像素點之範例圖 3。	52
圖 49 連通技術搜尋像素點之完成圖。	53
圖 50 以連通技術搜尋像素點 A 區域完成。	53
圖 51 以連通技術搜尋像素點 B 區域完成。	54
圖 52 單人之 Connected Component 並計算出其寬、高之值。	54
圖 53 雙人之 Connected Component 並計算出其寬、高之值。	55
圖 54 單人影像中刪除不符合高、寬比之比值在 0.8~2.5 之間的區域。	57
圖 55 雙人影像中刪除不符合高、寬比之比值在 0.8~2.5 之間的區域。	57
圖 56 單人影像中刪除不符合高、寬之值各在 20 以上之區域。	58
圖 57 雙人影像中刪除不符合高、寬之值各在 20 以上之區域。	58
圖 58 單人影像中刪除不符合像素點佔區域的 50% 以上之區域。	59
圖 59 雙人影像中刪除不符合像素點佔區域的 50% 以上之區域。	59
圖 60 單人候選臉區域灰階化。	61
圖 61 雙人候選臉區域灰階化。	61
圖 62 以門檻值區分灰階亮度之分布圖。	63
圖 63 中間值濾波範例圖。	64
圖 64 單人候選臉區域二值化後經中間值濾波之結果。	65
圖 65 雙人候選臉區域二值化後經中間值濾波之結果。	65
圖 66 單人候選臉 Connected-Component 後並做條件判斷。	66
圖 67 雙人候選臉 Connected-Component 後並做條件判斷。	67
圖 68 單人候選臉區域正規化之放大示意圖。	69

圖 69 雙人候選臉區域其中之一正規化之放大示意圖。	69
圖 70 單人影像完成人臉偵測。	72
圖 71 雙人影像完成人臉偵測。	72
圖 72 室內環境人臉偵測方法之流程圖。	73
圖 73 實驗後人臉偵測結果。	77
圖 74 Video1 實驗後人臉未被偵測出來。	79
圖 75 Video3 實驗後左邊的人臉未被偵測出來。	80
圖 76 Video1 與 Video2 實驗後錯誤接收率 FAR 出現情況。	82
圖 77 Video3 實驗後實驗後錯誤接收率 FAR 出現情況。	83
圖 78 人臉偵測與網路結合成居家監護系統。	89

表 目 錄

表 1 實驗數據表。.....	75
表 2 沒有加 dilation 的實驗數據表。.....	84

第壹章 緒論

第一節 研究動機與目的

在經濟繁榮的現代，科技與知識都在突飛猛進，然而在道德觀的低落與貧富差距大的情況下，造成治安每況愈下，不管是擄人勒索、強盜搶劫等重大的刑事案件，還是各種手段的詐欺、詐騙案件，都突顯了縱使有再多的警察人員也有顧及不到的治安死角，所以人臉辨識方面的保全系統重要性也跟著提高許多。

保全系統不外乎是身份識別系統，身分識別系統也由相當多技術所組成。這些技術在近年來也大為風行，其技術與原理的領域也有相當多的專業人才在規劃研究，例如：影像偵測系統、指紋辨識系統、人臉辨識系統、晶片辨識系統等等。其中人臉辨識系統的應用可以相當的廣泛，例如：提款機的錄影系統、出入門禁的監視器系統、停車場、百貨公司及便利商店的監控系統、甚至是無需接觸式(hands-off)的身份確認機制等等[12]，若人臉辨識結合視訊監控系統，更可進行身分比對與確認[25]。精確的偵測出人臉位置，並辨識是否為許可出入的人員，運用人臉辨識技術，對於犯罪偵防上都會有大大的提高效率。

近年來，人臉偵測技術因有相當多人投入研究已有相當程度的進展，而如果只是以單純的膚色來做辨別，當背景顏色與膚色相近時會造成系統誤判，或者有其他造成系統誤判為人臉的影像，如動物的臉、膚色的身體、顏色與膚色相近的物體或背景、形狀像人臉的物品等，又或者人臉因為角度不正或被物品遮蔽等等，上述種種原因皆會造成系統難以判斷而出現錯誤的結果，所以其精確度實在有限。如何配合其他更多的技術來進行影像人臉的偵測以提高其精確度。諸如在多數的人臉中，如何正確的偵測到此區域內是否有人臉存在就是本文的重點所在。

本專題的目的就是在影像中偵測人臉，先針對影像中的移動物體，經由膚色偵測過濾後搜尋出人臉為基礎，建立起完善的影像人臉偵測技術，以便讓其他專業技術人員以此技術建立起更完善的人臉辨識系統，用以核對身份，確定是否符合資料中的人，這樣不僅可以節省管理方面的人力、物力，避免人為疏失，提高精確度與工作效率，還對大樓、停車場、機密空間的安全性又多了一層的保障，所以人臉偵測與辨識研發的越完整，越妥當，越能使其提供更優質的服務與貢獻，故此人臉偵測有其研發的必要性，可看出在社會上是與人們生活息息相關的，在未來的延展性必定不可限量。

第二節 研究重點[1]

我們研究的重點是室內影像人臉偵測，判定畫面中是否存在著人臉。在室內影像人臉偵測中，起先都是偵測移動物體，但之後會區分為以過濾膚色來偵測人臉或以偵測到的人頭區域來辨識人臉，兩種方法各有好壞。辨識膚色固然會被許多因素所左右結果，而偵測人頭區域來辨識人臉就不須經過辨識膚色這項難度[10]，但本專題主要著重在膚色偵測，且利用過濾膚色來偵測人臉也可免去製作橢圓模組偵測人頭區域的繁雜步驟。而膚色偵測之成功關係到許多因素，舉凡移動中的物體、光的亮度、臉的角度、背景顏色、形狀與人臉相似的物件等等都可能影響程式判斷的正確性，所以我們主要依循五個要素來作為判別是否為人臉的依據：

要素一：移動物體

畫面中移動的物體雖大部分為人體，但也有可能是貓狗之類的動物，或是丟擲中的物品，如何判別為人體就是這技術初步困難所在，但人體的形狀大小皆有其固定比例，可以此為條件來加以判別，刪除不必要的範圍。

要素二：膚色偵測

人體的膚色有分很多種，但是紅、黑、白、黃主要的四種人種膚色差異過大，並在光的強弱下膚色些許改變，種種因素將會造成偵測人臉上的困難度，所以我們只針對黃種人來研究。

要素三：膚色偵測的區域是否有人臉

倘若畫面中的物體判斷為人體時，如何經由膚色偵測過濾後從上面偵測出人臉就是一項考驗，人臉雖有其形狀比例，可以此為參考條件來加以判別，但是人臉有可能因角度、光線、物體遮蔽等等影響到判別的正確率，故解決此困難將是本技術中的一環。

要素四：人臉五官

人臉皆有其五官，這就是人臉辨識上的可靠依據，偵測時，可留下鼻子與耳朵，而眼睛與嘴巴將會被過濾，這時就可依據其五官位置來判斷是否為人臉。

要素五：特徵

人臉形狀雖有差異，但都有一定的輪廓，如雙眼間の間距、鼻子與眼睛的位置比例、嘴巴的大小、嘴巴與鼻子的位置比例等皆有一定的高寬比例，而且還要考慮到照片距離遠近所造成人臉形狀的大小，所以可以依據膚色及人臉形狀判斷出人臉的雛型來。

有了上面五項要素，就可以作為偵測人臉的依據，在我們人眼來分

辨這五項要素是輕而易舉的事，但是如果由系統來自行分辨，就要去注意會造成誤判的種種問題，像人臉大小、特徵、角度、亮度、背景等，這些都是我們需要去研究、探討及改良的重點。

往後系統將會運用多樣技術來加強人臉偵測的正確性，如 YCbCr 色彩轉換、正規化、二值化、開閉運算(膨脹與侵蝕技術)、連通技術等等，本論文在後續也會一一探討這些相關技術。

第三節 研究範圍與報告書架構

本次專題研究的範圍，剛開始時先從室內影像去偵測是否有人體存在，從人體做膚色偵測，把屬於膚色的部分留下，其它的過濾掉，再來就依照人臉或五官的固定比例來判別圖像中的膚色是否為人臉部份，如人臉長寬比、眼睛與鼻子的比例，是的話就繼續重複步驟，看是否有第 2 個以上的人臉存在，直到全部都判斷完為止。

本篇論文細分許多細節，以五個重要章節貫穿全文。

第 1 章為緒論，主要在於簡述人臉偵測之用處、說明本專題之目的及解釋人臉偵測的困難與方法。

第 2 章為室內影像人臉偵測之理論與技術探討，說明目前各種人臉偵測之技術方向與理論。

第 3 章為本論文所使用之室內人臉偵測的方法，包含著型態影像學各種原理與技術。

第 4 章為本專題技術成果，以 C++ Builder 6.0 開發操作介面來呈現。

第 5 章為結論，對本專題做個最後的總結，並期許未來更進一步的研究重點與方向。

其它如致謝、摘要、參考文獻等將不在此敘述。

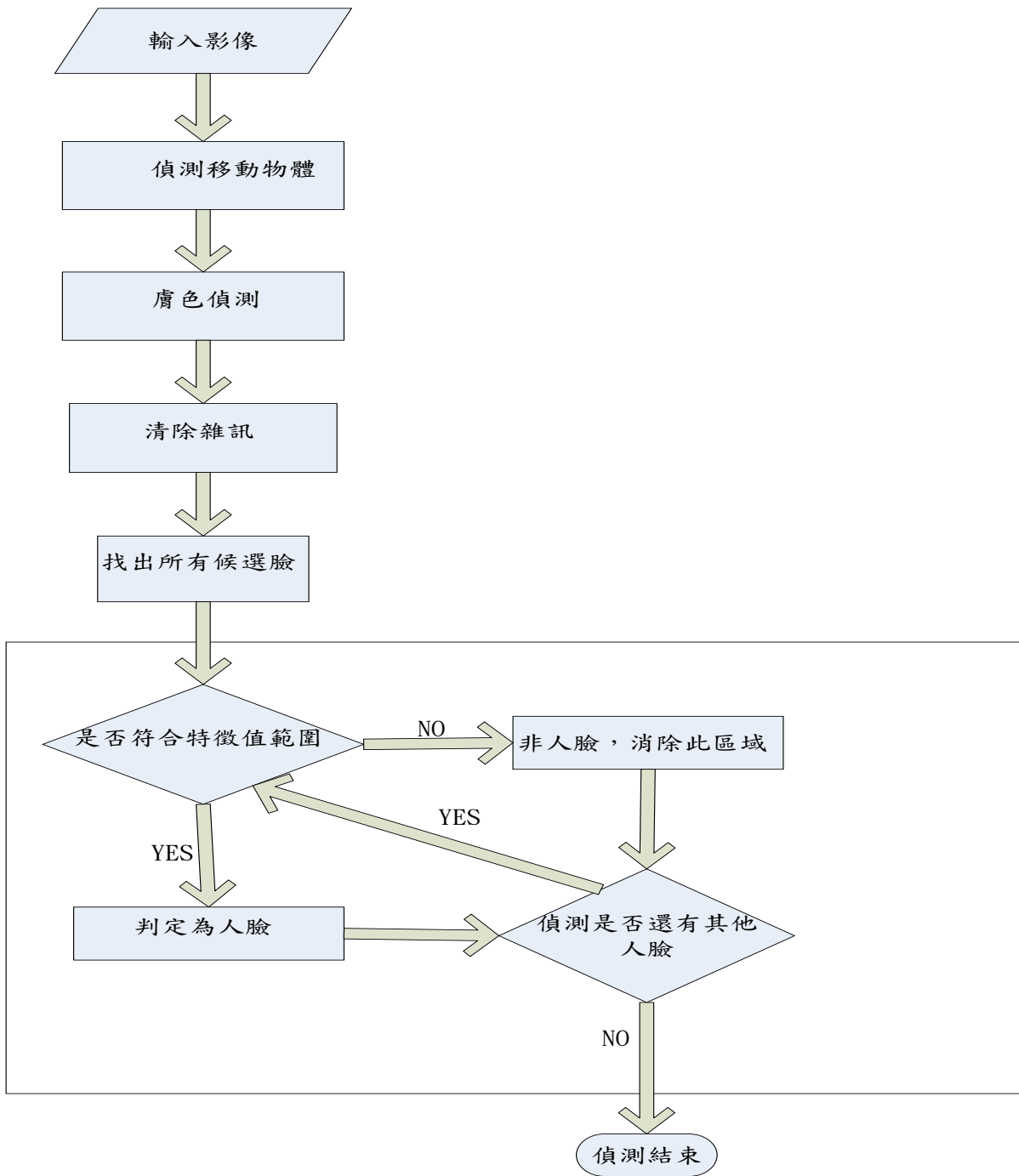


圖 1 專題研究架構圖。

第貳章 室內人臉偵測之理論與技術探討

第一節 YCbCr 色彩空間偵測膚色

RGB 是電腦螢幕顯示色彩的標準，但並不是用來描述色彩的好方法。因為人的膚色與背景顏色可能會很相似，所以如何把膚色跟背景顏色分離必須要加以研究，但若使用 RGB 色彩空間來表示的話，要將膚色和背景顏色區分開來判別的話準確性並不高[1]，且容易受光線強弱及陰影影響，所以影像中亮度若有微小變化，RGB 色彩空間變會出現明顯差異[11]，因此我們可以利用另一個色彩空間：YCbCr 色彩空間來加以區分膚色及背景顏色。

YCbCr 又稱 YPbPr，它是由 YUV 系統修改後所形成的色彩空間[7][9]。YCbCr 主要是因為人類對亮度變化的敏感度要比顏色改變還來的大，所以利用增減顏色範圍來增加亮度範圍的原理所創造的。YCbCr 是常用在視訊或影像傳輸的色彩模式，在電腦視覺技術上應用相當廣泛，MPEG(Moving Pictures Experts Group)、H.236 視訊壓縮、CCD Camera(Charge Coupled Device Camera)等都是使用到 YCbCr 色彩技術，圖片格式 JPEG(Joint Photographic Experts Group)檔即是以 YCbCr 為基礎

所創造的，他不僅在運算及處理上較為簡單，也較不易受到干擾[8]。膚色分布在 CbCr 上有不錯的集中效果，各種人種的膚色在 CbCr 上差異也不大，主要是利用亮度與色度分離的方法，因此只要將三維 RGB 轉成二維 CbCr 就可以降低膚色分類上的複雜程度。其轉換公式如下所示：

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.209 & 0.587 & 0.114 \\ -0.168 & -0.331 & 0.5 \\ 0.5 & -0.418 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix} \quad (1)$$

其中 YCbCr 各為 8bits(0~255)的編碼，YCbCr 的意思分別是亮度 Y(luminance)和彩色度 Cb(blueness)、Cr(redness)。Y 之值為 16 到 235，共 220 個等級，數值越大越亮，16 是黑色，235 是白色，並將其值以 ± 16 作為修改，使其值可至(0~255)[7]。彩度 Cb 代表偏藍色向量，Cr 代表偏紅色向量[34]。YCbCr 色彩空間在轉換上來得比其他色彩空間較為快速，且由於亮度 Y 部分被抽離彩度資料後，降低對於彩度資料的干擾，因此只要在彩色度 Cb 及 Cr 上不斷做調整及修改就可以偵測出膚色了[1]。

第二節 HSV 色彩人臉區域偵測

HSV 為色相 H(Hue)、飽和度 S(Saturation)、明暗度 V(Value)這三個單字的縮寫所組成，色相 H(Hue)是全彩的屬性，如紅、綠、藍、橘等各種顏色，飽和度 S(Saturation)是色彩被白色稀釋的程度，飽和度越高，色彩中包含的白色成分越低[2]。明暗度 V(Value)是指明暗的程度，顏色越亮明暗度越高。將 HSV 色彩畫起來的模型有如錐形[3]，如圖 2 所示：

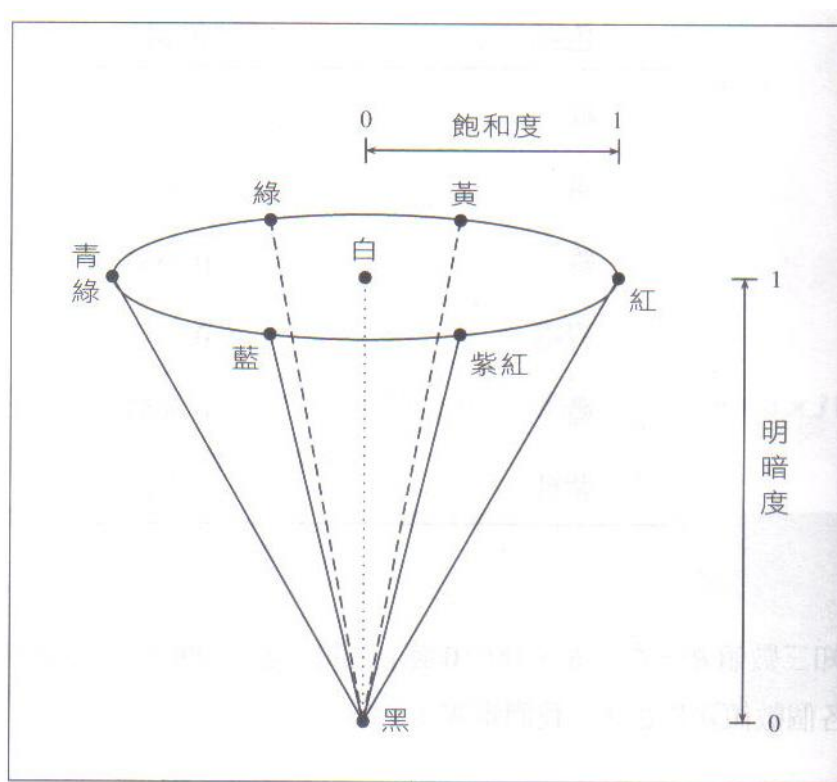


圖 2 HSV 的錐形色彩空間示意圖[3]。

因皮膚與背景的颜色有可能會很相近，且 RGB 色彩空間容易受光線強弱及陰影影響，所以影像中亮度若有微小變化，便會出現明顯差異，故使用 RGB 色彩空間將不容易把膚色和背景顏色分開，所以可採用 HSV 色彩空間，此色彩轉換可以將亮度(luminance)和色度(chrominance)做良好的分離[2]。兩者的轉換公式如下：

$$H1 = \cos^{-1} \left\{ \frac{(R-G) + (R-B)}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \right\} \quad (2)$$

$$H = \begin{cases} H1 & \text{if } B \leq G \\ 360^\circ - H1 & \text{if } B > G \end{cases} \quad (3)$$

$$S = \frac{\text{Max}(R, G, B) - \text{Min}(R, G, B)}{\text{Max}(R, G, B)} \quad (4)$$

$$V = \frac{\text{Max}(R, G, B)}{255} \quad (5)$$

其中色相 H(Hue)，是用來區分顏色的主要成分，也是偵測人臉的憑藉，S(Saturation)代表飽和度，它與顏色明暗度 V(Value)之值皆在 0 到 1 之間。

第三節 YCbCr 與 HSV 之差異

HSV 與 YCbCr 比較起來，會有更準確的偵測結果，因為他是用色相 H(Hue)、飽和度 S(Saturation)和明暗度 V(Value)來區分出膚色和非膚色區域，在分類上較為細膩，但 HSV 在轉換上並不像 YCbCr 是用線性方式轉換，所以 HSV 在轉換上也較耗費時間[4] [5]。但 YCbCr 在感光上的反應並不會像 HSV 這麼敏感，對不同的膚色在使用上有較好的效果，由圖 3 可看出，膚色在 YCbCr 色彩空間的分布情形較 HSV 色彩空間來得集中，可善用此特性將膚色與非膚色分離開來[9]。

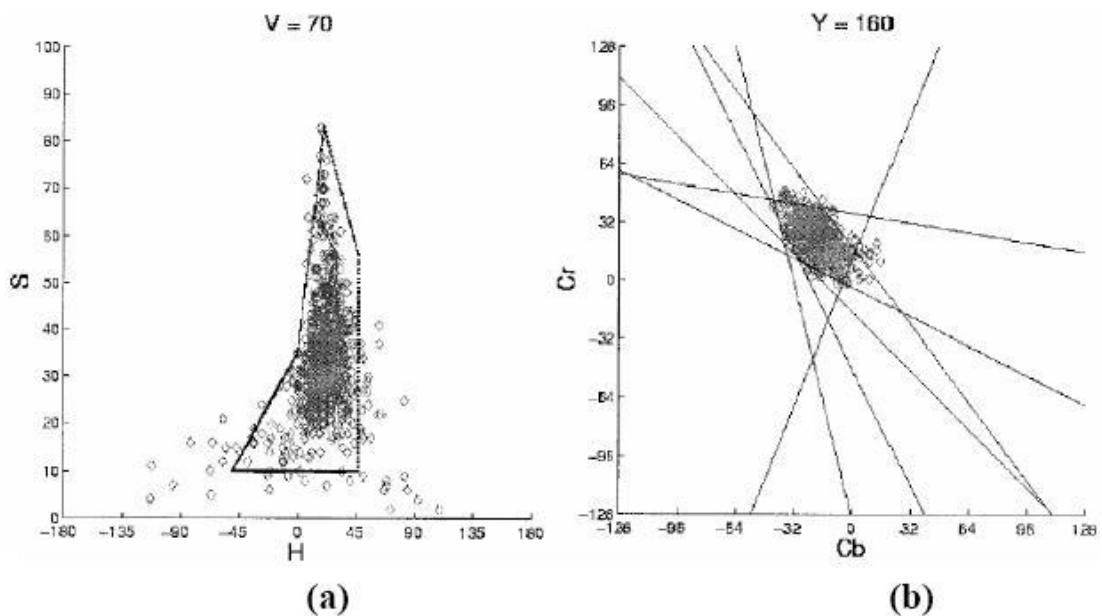


圖 3 膚色分布圖：(a)HSV 色彩空間；(b)YCbCr 色彩空間。

第四節 技術 1 探討 人臉偵測技術研究[1]

此技術主要是採取特徵偵測方式為主要方法，如膚色、邊緣與人臉形狀為主來偵測畫面中是否有人臉存在。技術中是以 YCbCr 的色彩模式，因為在運算上及處理上都較為簡單，且不易受到干擾。由於亮度(Y)的部分被抽離色彩度後，降低對於色彩度的干擾，因此僅採用色彩度成分(CbCr)。

輸入影像經過 YCbCr 轉換後，經過統計可發現膚色分部都集中在 $R_{Cb}[96, 131]$ 、 $R_{Cr}[131, 173]$ 這組範圍內，可利用此一範圍多次分析測試後找出一組可靠門檻值，往後將以此門檻值為基礎作為偵測人臉之依據。

本論文技術是利用多向性邊緣偵測的優點，可以偵測水平、垂直、斜邊。輸入影像後，經過 YCbCr 轉換後偵測斜邊，再經過邊緣偵測後就輸出影像。

邊緣偵測是利用 Sobel Filter 搭配 4 種不同的遮罩值，往各方向比較強度，比較後的值來代表該點向量。但因處理後雜訊過大，將以 Entropic Thresholding 之技術來處理，此技術目的是為了取一個適合影像內容的最佳門檻值，如圖 4 所示，輸入影像後，將 Y、Cb、Cr 各自分離，並透過計算來產生一門檻值，此門檻值可藉著執行邊緣與非邊緣像素計算

Entropy 總合，使邊緣與非邊緣像素資訊量達到最大。結果如圖 5 所示。

之後再以 4 元膨脹的技術來降低雜訊干擾。如圖 6 所示。

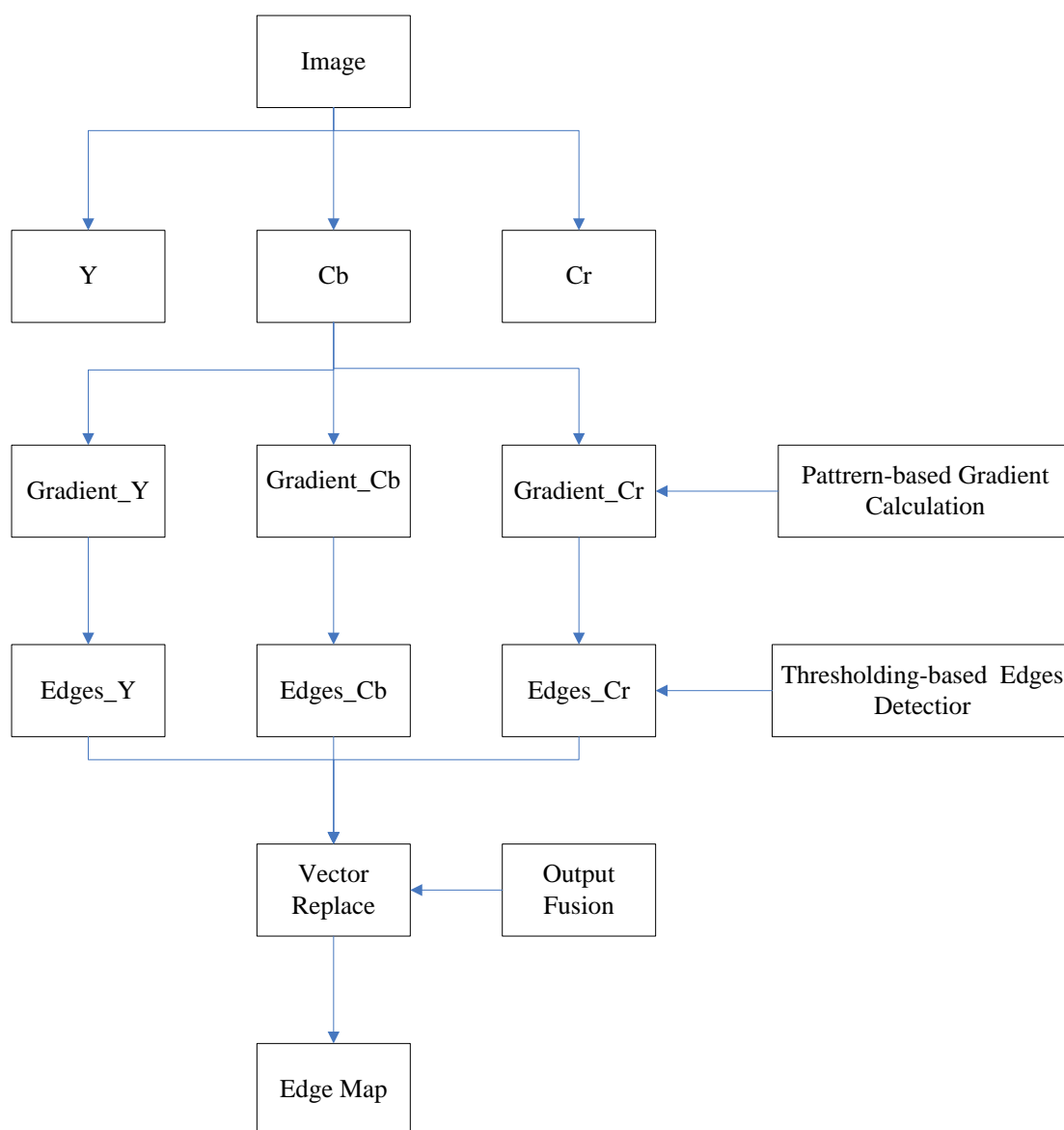


圖 4 彩色邊緣偵測流程架構圖。

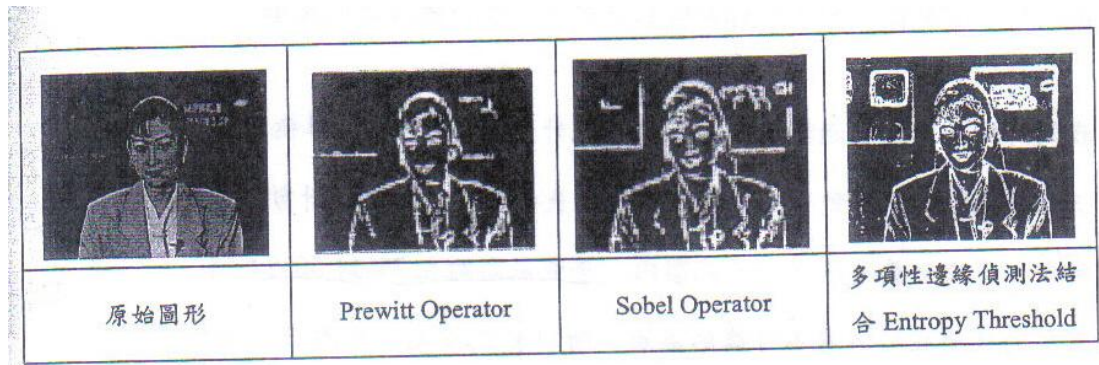


圖 5 Entropic Thresholding 技術結果圖。



圖 6 邊緣偵測後結果比較。

本論文技術提出，人臉的長度與寬度分布比例大約在 1.20 到 1.91 左右，而其他部分則是不規則狀的，利用此一特性就可先初步刪除非人臉部分的區域。如圖 7 表示結果。

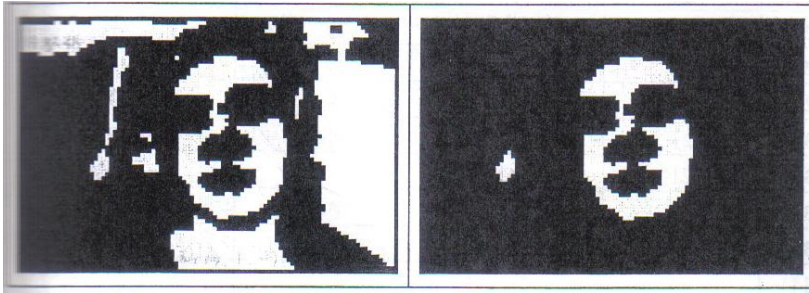


圖 7 人臉長寬比例校正後之結果。

另外，第二步刪除的方法，便是在一個區域畫面中主要有兩個成分，一是純膚色區域，二是含有邊緣的膚色區域，且這兩種有一定的比例存在，此技術根據統計表算出膚色區塊比例介於 0.35 到 0.8 之間。同理，人臉亦可分為純膚色跟邊緣區塊，根據統計人臉區域內邊緣比例介於 0.06 到 0.45 之間，如圖 8 表示結果。

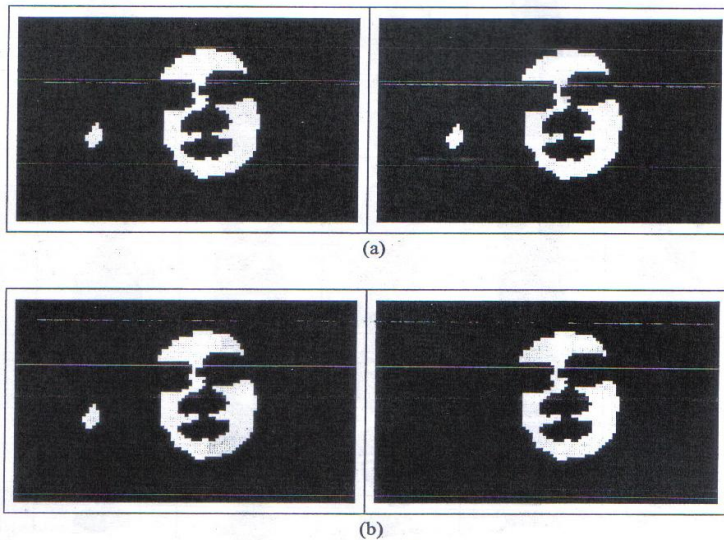


圖 8：(a)膚色比例校正後之結果；(b)邊緣比例校正後輸出結果。

此技術之優點有三：

1. 透過人臉認知的臉部特徵來過濾背景雜訊，正確性會較其他演算法來的高，如圖 9 所示。
2. 不受膚色的干擾，不會將非人臉部分的膚色也列入判斷。
3. 不受人臉角度的影響，如圖 10 所示。

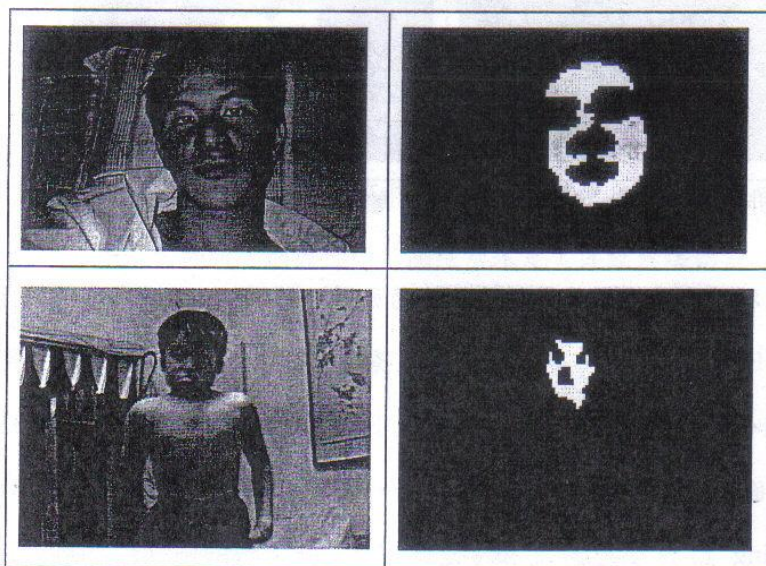


圖 9 過濾背景雜訊結果圖。



圖 10 偵測結果不受人臉角度影響。

但此技術將有兩個缺點存在，第一是如果輸入影像尺寸過小，而此技術的邊緣偵測大小範圍卻是固定的，將會造成固定區塊的比例大幅提升，導致區塊被刪除，使膚色區塊沒有連續性或是與其他膚色區塊連結在一起，使得整個圖像與原始圖像差異過大，如圖 11 所示。第二是因為光線照射影響下，使膚色頻譜產生位移，以至於膚色不能正確偵測出來，如圖 12 所示。雖此技術有些許缺點存在，但根據作者所提出的數據，偵測準確率依然高達 84%，所以有其參考的價值存在。

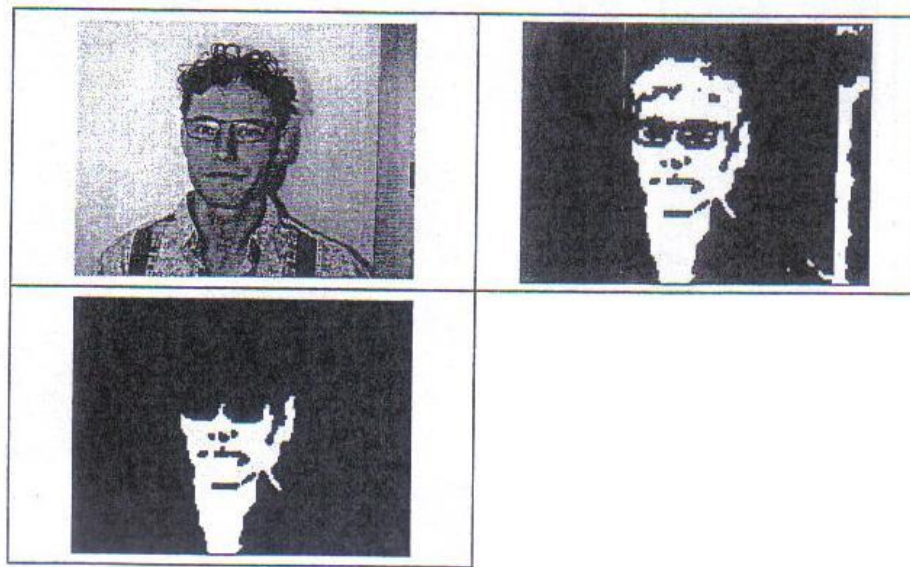


圖 11 影像尺寸不一導致邊緣偵測效果不彰。



圖 12 受光線影響膚色偵測失敗。

第五節 技術 2 探討 彩色視覺技術於多人人臉辨識上的應用

[2]

此技術結合了膚色偵測，影像處理，和幾何規則判定等方法。為了讓人臉特徵擷取的過程能成功，最後將會對人臉候選區域進行角度轉正的動作。

在人臉偵測時，因為影像中包含許多複雜的訊息，所以將影像大致區分為人臉區域與背景區域。

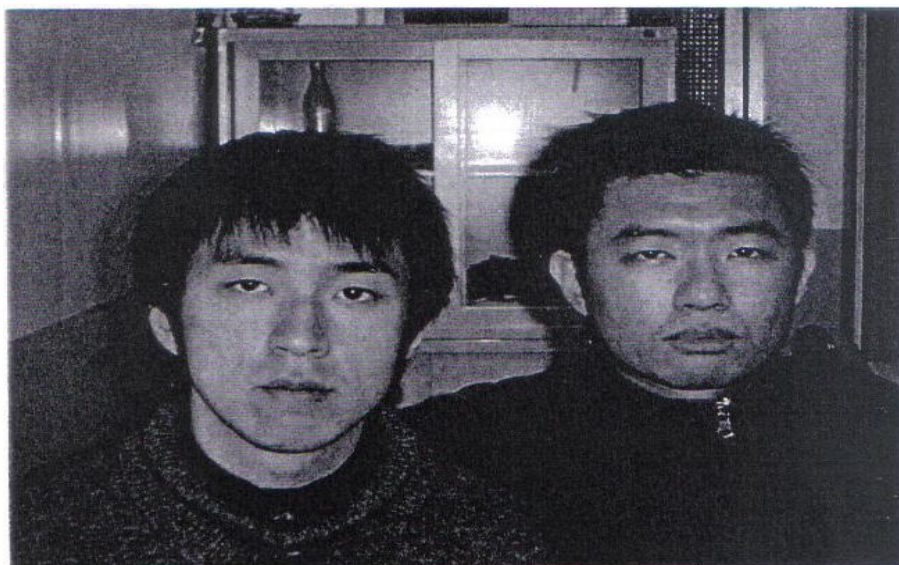


圖 13 輸入原始影像。

因皮膚與背景的颜色可能會很相近，且 RGB 色彩空間易受光線影響將不容易把膚色和背景颜色分開，所以採用 HSV 色彩空間。將 RGB 轉

換為 HSV 後，接下來就是將皮膚顏色與背景顏色分離。此技術將 S 和 V 座標值放大 100 倍，可發現膚色分部都集中在某範圍內，如圖 14 所示，此技術利用這一特性做二值化處理。

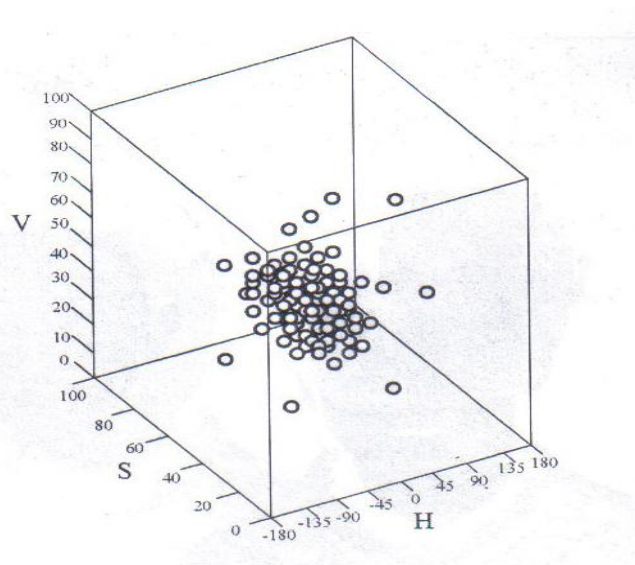


圖 14 膚色在 HSV 之分布圖。

要將這兩個顏色有效分離，可採用臨界值法將影像做二值化處理。

作法是以一些公式來定義出膚色在 HSV 的範圍，公式如下：

$$\begin{aligned}
 S &\geq 10 \\
 V &\geq 25 \\
 S &\leq -H - 0.1V + 110 \\
 H &\leq -0.4V + 75 \\
 S &\leq \begin{cases} 0.08(100 - V)H + 0.5V & \text{if } H \geq 0 \\ 0.5H + 35 & \text{if } H < 0 \end{cases}
 \end{aligned} \tag{6}$$

利用這些公式可以找到圖像中的像素點皆屬於皮膚顏色，經處理後的結果，黑色部分為 0，白色部分為 1，此技術將這些像素點分為這兩種，符合就是 0，不符合就是 1。



圖 15 二值化處理後的影像。

由圖 15 可看出二值化後的影像分為人臉與非人臉兩個區域，之後在用開閉運算對影像加以處理，消除雜訊並使嘴巴部分作為之後角度的計算。經由閉合運算，即先膨脹後侵蝕，一些人臉外的小雜訊都被消除了，此技術的膨脹與侵蝕運算都是針對二值化影像中的白色部分來處理，是

先做閉合運算再做斷開運算(先侵蝕後膨脹)，目的是為了使嘴巴部分更明顯的顯示出來。結果如圖 16 所示。

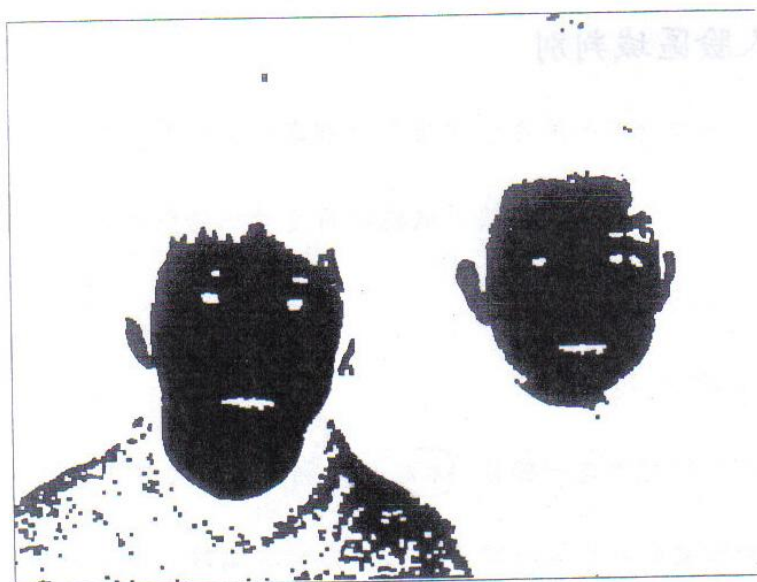


圖 16 開閉運算後的影像。

在特徵擷取這方面，此技術是利用相對位置來處理，為了要準確的尋找相對位置，必須先將人臉做轉正的處理，而為了轉正則需要用到連通成分的技术。

像素與像素間的連接是許多影像物體描述意義之簡化概念，此即為相連元件(Connected-Component)。連通成分技术是使用資料結構中的堆疊(stack)原理來處理，目的是為了讓二值影像中屬於同一物體區域的所有像素都能夠連接起來，如此便能知道像素是屬於哪個區域，並知道此

區域之面積大小，此方法是利用 8 個鄰接方式將鄰接點連接在一起，讓所有連接的點都被標示為一個物體，未來此技術在偵測膚色的過程中都會用到連通成分技術。

此技術設定了像素數目大小的範圍，範圍太小將立即刪除，只要符合條件便判斷屬於人臉區域，經連通成分後只留下三個部分，如圖 17 所示。

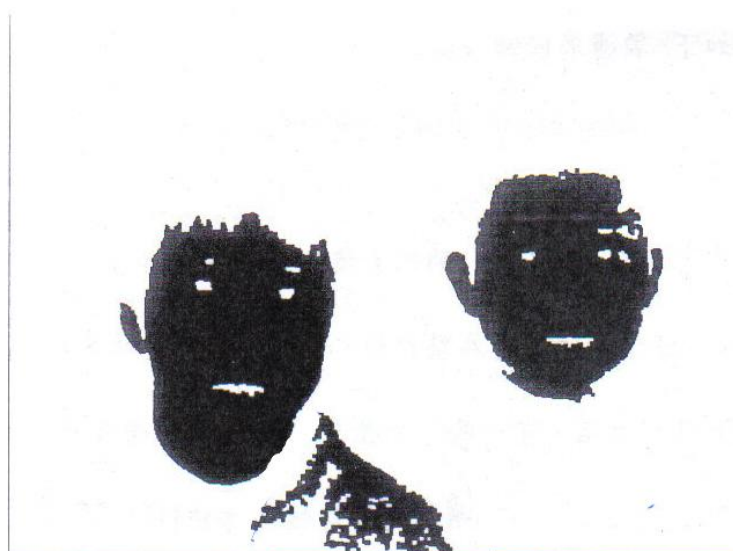


圖 17 連通成分後的影像。

完成連通成分後，接下來便是對人臉部分做轉正，此技術是利用嘴巴在人臉區域中的所有封閉缺洞(白色部分)中所佔的區域是最大，而嘴巴的兩端點也就是嘴角聯成一條線，可依據此特性來轉正整個人臉，並

可求取轉正人臉所需的 θ 角。而轉正的技巧使用雙線性內插法，示意圖如圖 18 所示。

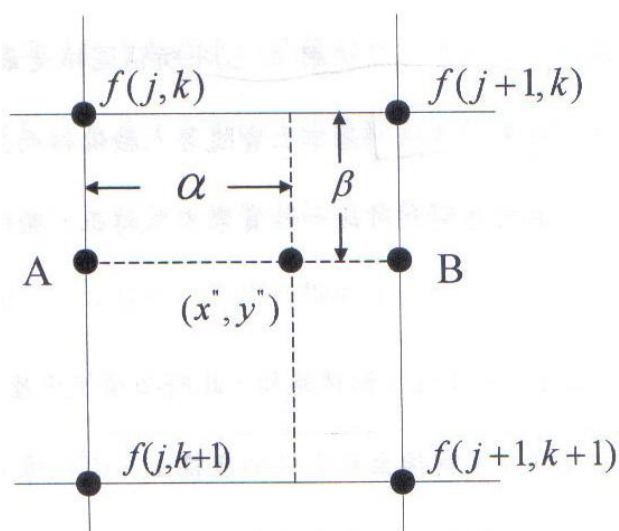


圖 18 雙線性內插法在原圖上的格子點。

求得角度並轉正，一連串的過程如下圖 19 所示。



(a)

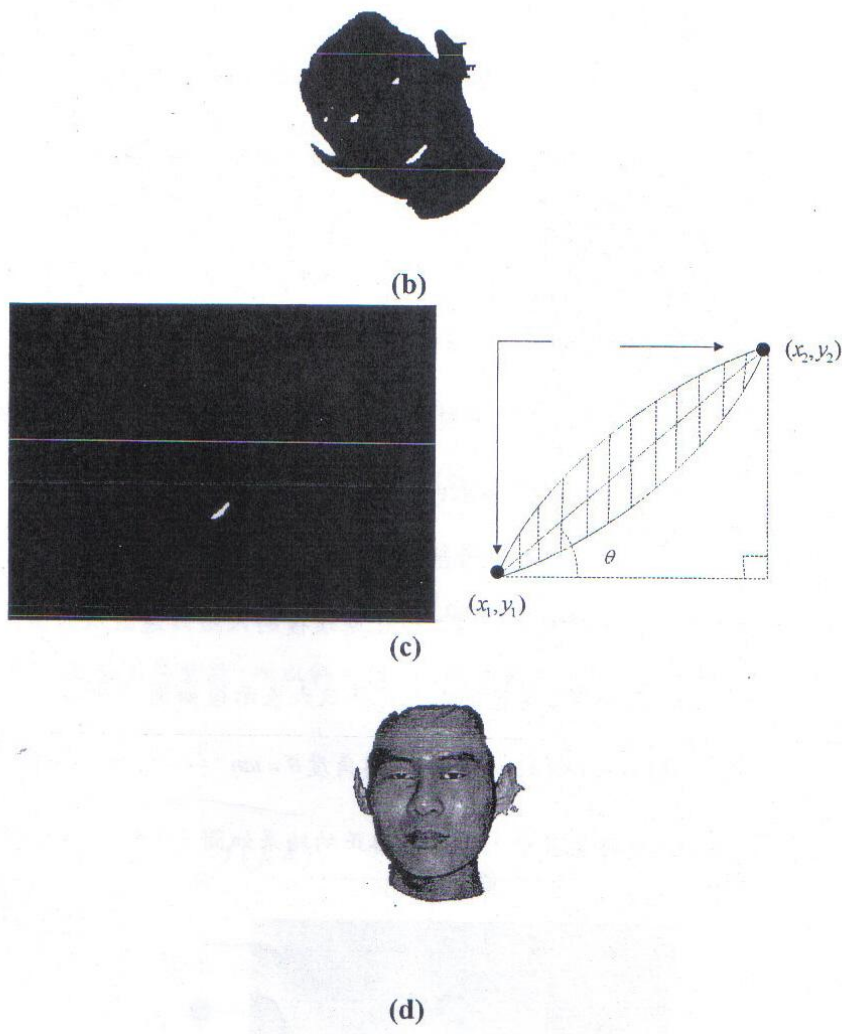


圖 19：(a)原始影像；(b)人臉候選區域；(c)求嘴巴轉正角度；(d)轉正後結果。

最後當每個留下的區域都做轉正後，為了判斷剩下的區域是否為臉部，因此設定了一個臉部的長寬比例，這個長寬比例的高度必定大於寬度，公式如下：

$$0.3 \leq \frac{Height}{Width} \leq 0.8 \quad (7)$$

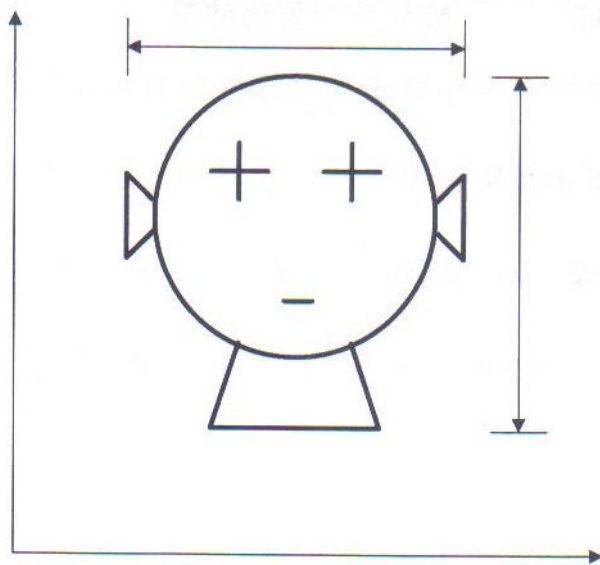


圖 20 人臉比例表示圖。

最後偵測出來的結果如下圖 21。



圖 21 臉部偵測後結果。

第六節 技術 3 探討 利用膚色及區域極小值作人臉特徵是否 遮蔽之偵測判斷[6]

這項技術是用來偵測並追蹤移動中的人臉，大多技術都會使用到膚色偵測，但他是在移動物體上依照臉部比例特徵來搜尋人臉。這項技術有兩個問題會產生：1. 人臉位置不知道；2. 人臉大小不知道。因此為了降低失誤的可能性，先對輸入影像做移動物偵測，可以縮小搜尋人臉的範圍，也可以依照特徵判斷是否為人臉。如圖 22 所示。

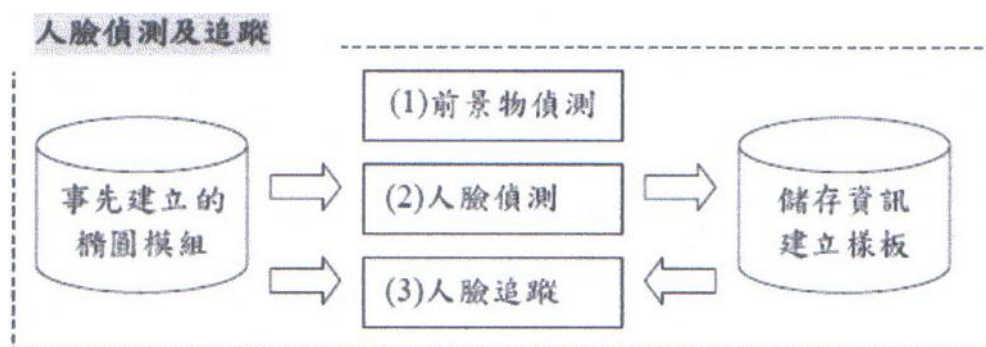


圖 22 人臉追蹤簡易架構圖。

而偵測移動物體的方法，此技術選擇使用連續影像相減。首先將輸入影像轉成灰階影像，再跟前一張影像的灰階化相減取絕對值，再透過一個門檻值把輸入影像分前景點與背景點，以求出二元化影像，再將原

影像邊緣化後與二元化影像取交集，即可偵測出移動物體。如圖 23 所示。

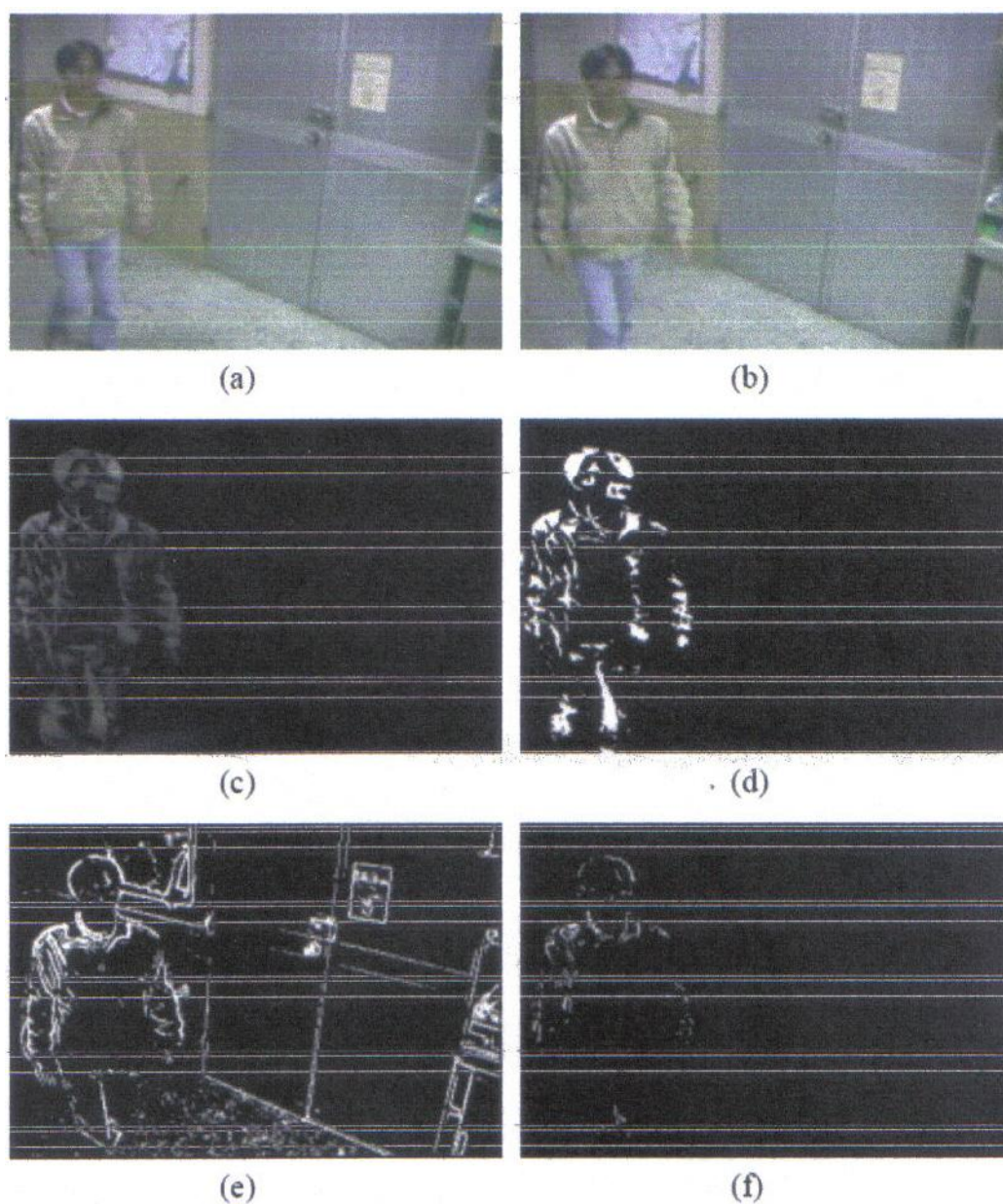


圖 23 移動物偵測：(a)前一張影像 I_{t-1} ；(b)現在的影像 I_t ；(c)連續兩張影像灰階值相減；(d)將(c)透過門檻值取二元化；(e)將(b)邊緣化；(f)取(d)跟(e)的交集。

在搜尋人臉方面是先建立橢圓模組，因為人臉輪廓與橢圓形相似，而橢圓模組中，的 a 表示短軸長度， $1.2a$ 表示長軸長度，可根據不同的 a 值創造出不同的橢圓。如圖 24 所示。



圖 24 長短軸比率 1.2 的橢圓與實際人臉。

而搜尋的方式是先將輸入影像邊緣化後，利用 Sobel 方法找出邊緣化影像，如圖 25 所示，在結合前景偵測的結果來縮小搜尋範圍，最後在此範圍內，以適當的橢圓模組在範圍內一一比對，最左上角的點為基準點，套入橢圓模組後將計算比對的分數紀錄下來，再往下一個點移動，不斷的重複直到移動到右下角為止，分數最高的即為偵測到的人臉位置。如圖 26 所示，白色點為 1，黑色點為 0，比對過程如圖 27 所示。

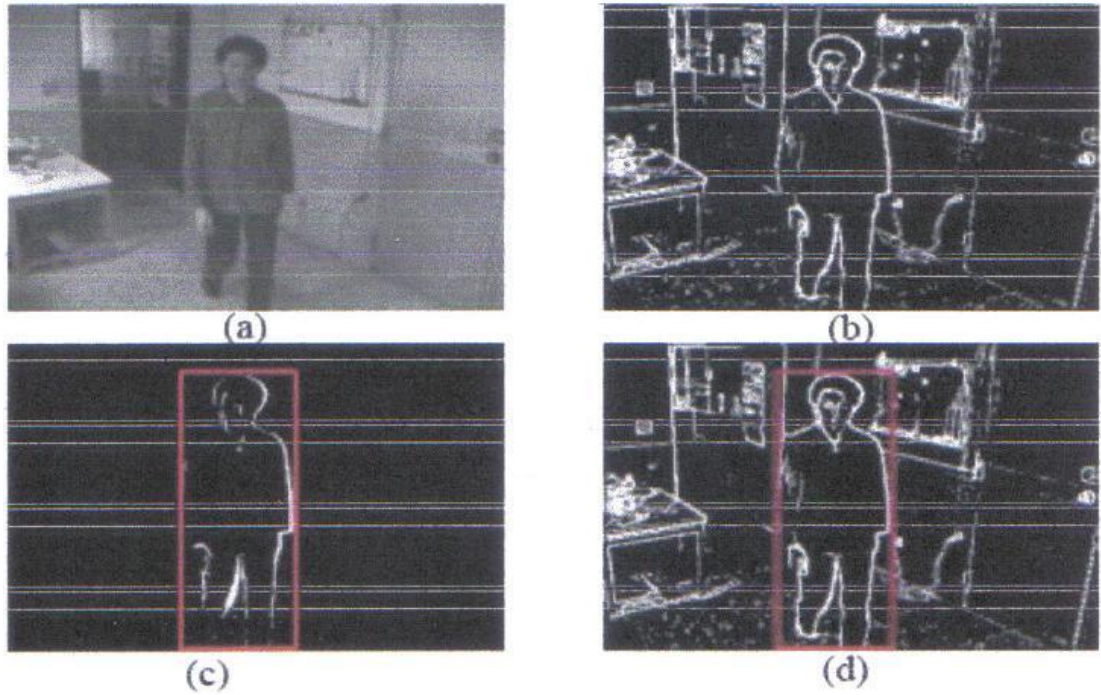


圖 25 搜尋視窗：(a)原影像；(b) Sobel 邊緣化後；(c)偵測後與搜尋視窗；(d)把(c)之搜尋視窗套用到(b)。

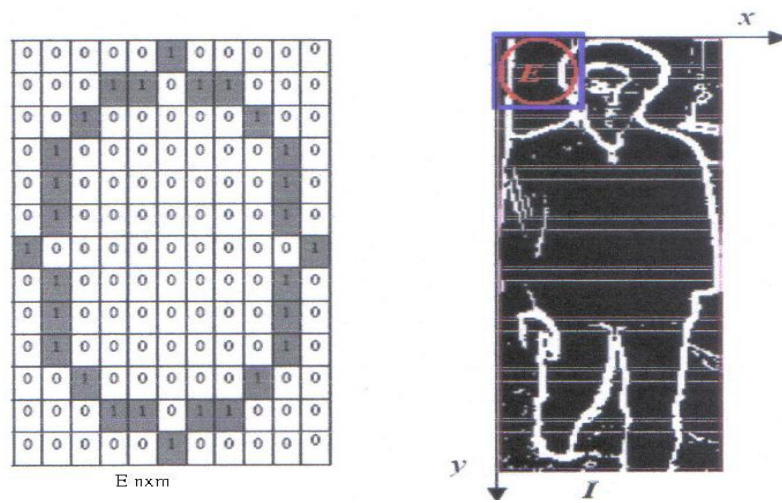


圖 26 橢圓模組比對：左為橢圓模組；右為欲比對的影像。

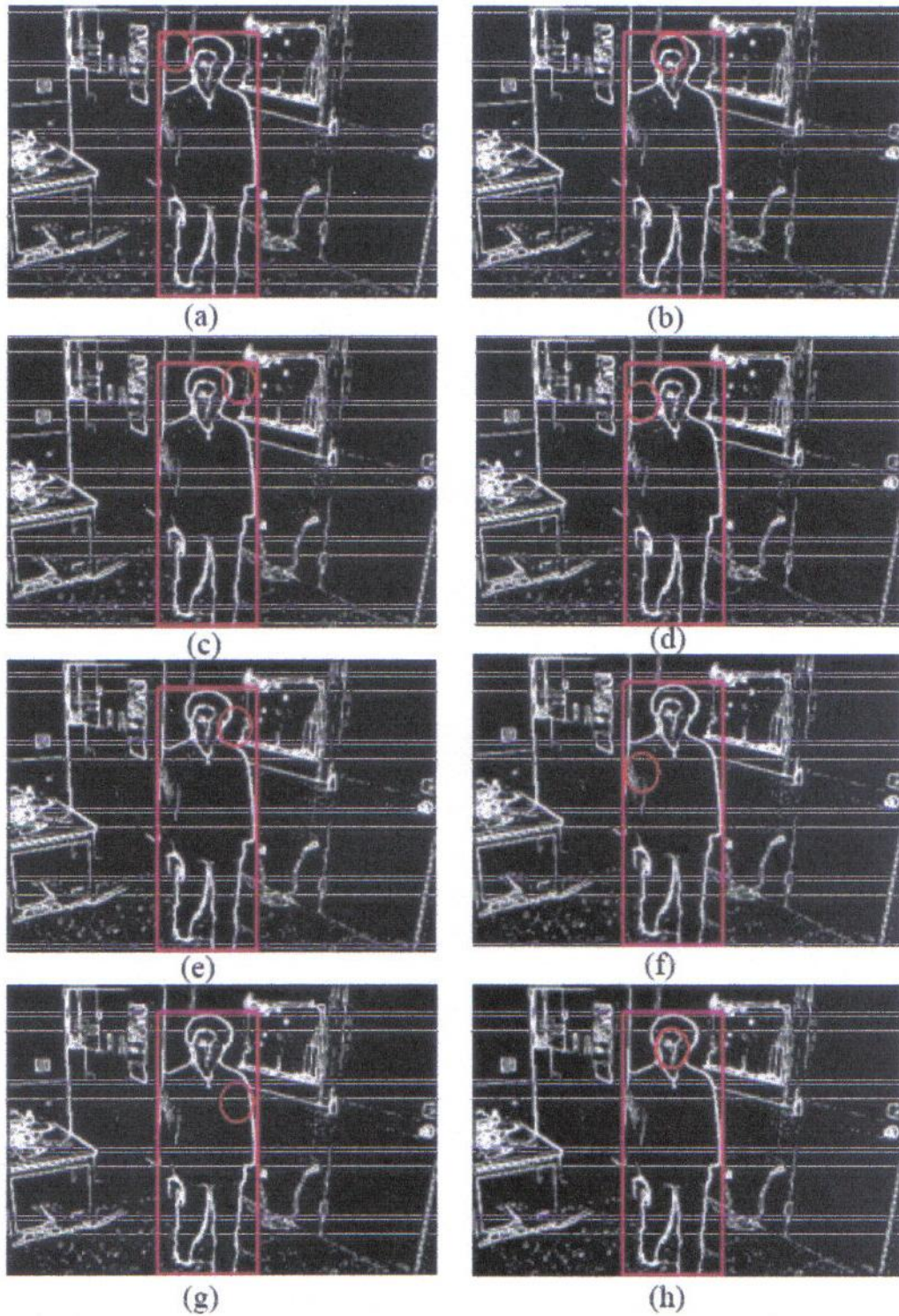


圖 27 橢圓模組比對過程與結果：(a)~(g)為橢圓模組比對過程；(h)

比對後找到分數最高的位置。

找到人臉位置後，接下來辨識追蹤的方法。由於物體在移動時可能會影響到拍攝造成人臉位置的遠近，以至於看起來時大時小，此時就必須對樣板做正規化的動作，如圖 28 所示，橢圓模組是由 4 個點所建立出來，而此 4 個點所形成的矩形範圍，就是此技術採用的樣板大小，所以只需改變樣板大小即可達到正規化。比對的方法選擇 SAD(Sum of Absolute Differences)方法，它主要是直接比較差異度，且速度最快，可節省時間，比對結果如圖 29 所示。

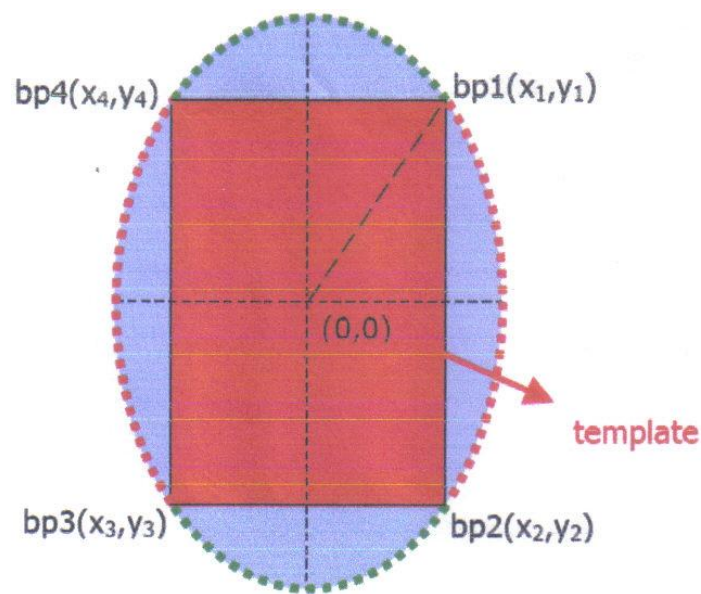


圖 28 4 個點所形成的橢圓模組與樣板。



圖 29 橢圓模組與樣板比對追蹤人臉過程與結果。

第參章 室內環境人臉偵測之方法

第一節 偵測移動物體

偵測移動物體常見的方法有兩種，一為連續影像相減，另一個是與乾淨背景相減，連續影像相減只要與前一張影像相減找出差異即可[26][29]，但前景影像若差異過大或過小，將導致相減出來誤差過大或無法明顯表現出來，而背景相減方法只要事先建立好乾淨之環境背景圖，再讓前景與背景相減，可以依照環境地點而更改背景，簡單且正確性高，故我們選擇以背景相減來偵測移動物體[9]。但不論方法為何，皆不適合攝影機本身也在移動時的狀況，因攝影機移動中背景也跟著改變，兩張影像差異會過大，無法將正確目標偵測出來[26][29]。

(一) 乾淨背景的取得

在室內影像人臉偵測中，我們先將拍攝的影片以每秒 30 張 images 切割出來，因為在室內雖然光線固定，但實際上光源通常是不均勻的，不均勻的光源產生強光與陰影使人臉偵測難度增高[27]，且系統對光的

敏感度比人的肉眼還來的強烈，所以將對全部的純背景圖之每個像素值做一個取平均值的動作，來製作出乾淨之背景圖，如圖 30、31 所示，之後就是偵測移動物體。



圖 30 單人圖乾淨背景。



圖 31 雙人圖乾淨背景。

(二) 影像差值計算

在拍攝機器固定及室內光源變化不大的條件下，外加移動物體的速度一般而言並不會太快，因此，我們將對每張影像進行影像相減來偵測移動物體。

首先，我們以乾淨的環境背景圖來當作第 0 張，再讓每張影像和背景做影像相減的動作，我們先在單人和雙人的影像中各取 1 張圖，分別如圖 32、33 所示，經由影像相減後藉以定位出此影像內之移動物體區域，系統會在影像相減後在畫面中留下移動物體區域，其它非移動物體的背景環境將以白色表示，如此一來，可以刪除掉大部分的背景環境，不必擔心背景顏色與人臉相近，還可以縮小搜尋範圍，使人臉偵測更加快速、便利，偵測移動物體將顯得很實用[31]，相減結果分別如圖 34、35 所示。



圖 32 單人影像原圖。



圖 33 雙人影像原圖。

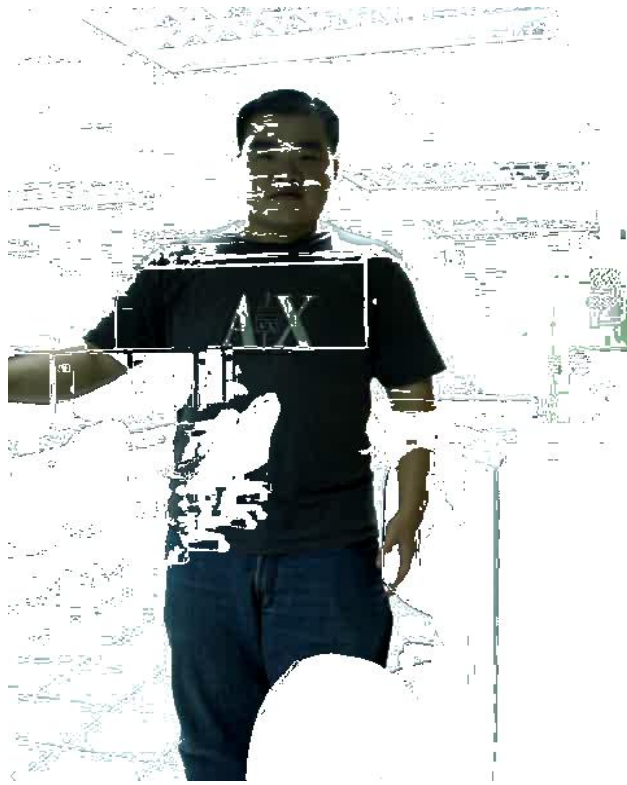


圖 34 單人影像相減。



圖 35 雙人影像相減。

第二節 膚色偵測

在畫面中所留下移動物體區域，我們可以利用膚色偵測的技術來過濾大部分非人臉的區域，雖然尚無法立即判斷出人臉所在位置，但卻已經為我們大幅度縮小了所需偵測之候選臉的範圍。

在膚色偵測的技術中，如第貳章第一節所述，因 RGB 並無法把膚色與非膚色區域有效的分離出來，所以可利用色彩空間轉換來達到膚色分離的效果。如第貳章所敘述，色彩空間的模式有很多種，其中 YCbCr 是以線性方式轉換，轉換速度較快，對光線的敏感度也較低，不易受光線影響，且對膚色的集中有不錯效果，方便把膚色分離出來，因此我們選擇以 YCbCr 色彩空間來進行色彩轉換[25]。

YCbCr 分別為 Blueness(Cb) 與 Redness(Cr) 兩個色彩度以及 Luminance(Y) 亮度所組成[12][33]，CbCr 的模式下對膚色有著不錯的集中效果，主要是以色彩度和亮度分離的方法，將三維 RGB 轉換成二維 CbCr 就可以降低膚色分類上的複雜程度，其轉換公式如公式 (1) 所示。

我們先針對室內環境下的圖片進行手動膚色切割，切割出數百張的純膚色圖片，如圖 36 所示。再利針對這些圖片計算出 CbCr 的平均值，以此平均值做為膚色偵測用的範圍值，將膚色保留下來，其餘非膚色的

部分將以白色表示，影像經膚色偵測後結果分別如圖 37、38 所示。

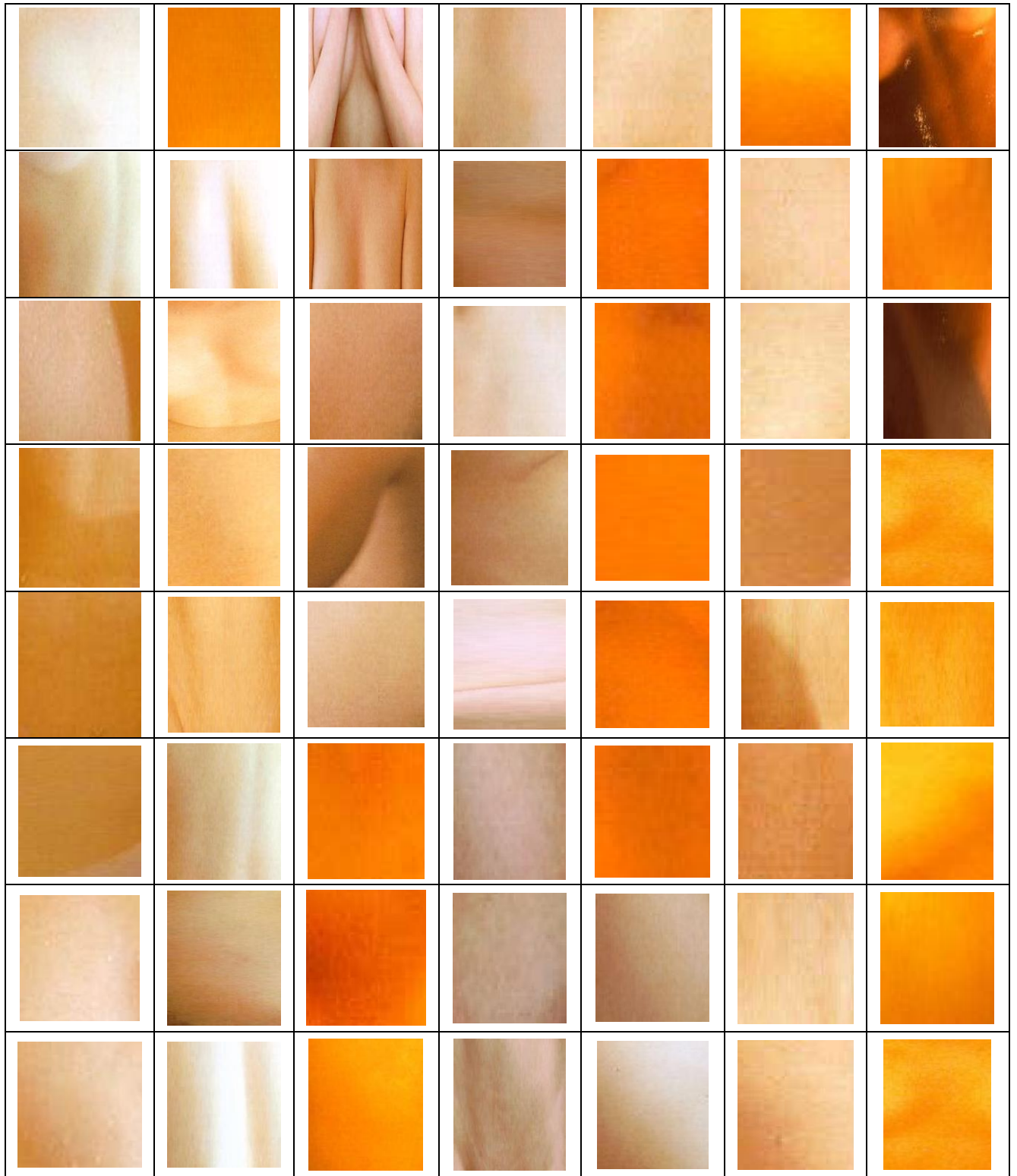


圖 36 切割後的純膚色圖片範例。



圖 37 單人膚色過濾。



圖 38 雙人膚色過濾。

第三節 影像二值化(Image Binary)[2]

影像經膚色偵測後，再來就是將膚色做二值化(Binary)。在影像處理的應用上，臨界值法在影像的分割上是重要的一環，目的是設定臨界值大小，凡是影像灰階度高於此值，便設為白色，低於此值便設為黑色，有效的將膚色與非膚色部分分離，有效的將影像分為黑與白，使影像資料簡化以利往後系統運作[30][31]。

單一臨界值法是目前最常用來處理影像二值化的轉換方式，假設一個二維灰階值影像的函數為 $f(x, y)$ ，是以 0(黑色)到 255(白色)為其值， x 與 y 則是影像座標，經過二值化後影像的函數為 $g(x, y)$ ，其函數轉換表示如圖 39 所示，則二值化的單一臨界值法的轉換公式如公式(6)所示。

$$g(x, y) = \begin{cases} 255 & \text{if } f(x, y) \geq T \\ 0 & \text{if } f(x, y) \leq T \end{cases} \quad (6)$$

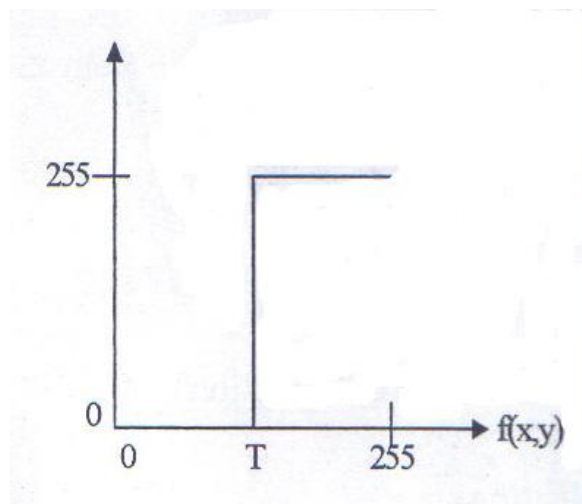


圖 39 圖影像二值化函數轉換。

其中 T 便是代表臨界值，尋找臨界值的方法便是計算出整個欲轉換區域的灰階值後再取平均值，此平均值便當作是臨界值，與 N 分別代表影像的寬度與長度，臨界值求法如公式(7)所示。灰階值為 $0\sim 255$ 共 256 組色階， $0\sim T$ 畫分為 0， $T\sim 255$ 劃分為 255。影像經二值化處理後，膚色區域值為 0(黑色)，非膚色區域值為 255(白色)[25]，影像二值化結果分別如圖 40、41 所示。

$$T = \frac{1}{M * N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \quad (7)$$



圖 40 單人二值化。



圖 41 雙人二值化。

第四節 膨脹與侵蝕運算[2]

在型態影像學中，所有的應用都是以下列四種運算混合使用，如此便可達到欲達成之理想效果。

在偵測人臉區域時，有時會遇到待測區域之像素點不完整，可能是區域中像素太多，或是區域中像素破碎凌亂不堪，此時為了讓我們可以更準確的偵測出人臉區域，我們可利用膨脹(dilation)與侵蝕(erosion)來增減區域之像素點。

膨脹運算會以一點為原點，往它的八鄰居向外擴張，讓原本的像素點往外擴張長出一層邊，將外型均勻的加粗，宛如其名「膨脹」一般；侵蝕運算會讓原本的像素點往內慢慢消失，將外型均勻的縮小變細，顧名思義即為侵蝕。如圖 42 所示，假設 X 為中心點，A~H 為其 8 鄰居，其膨脹與侵蝕公式分別如公式 8 和公式 9 所示， \cap 代表 AND 運算， \cup 代表 OR 運算。

A	B	C
D	X	E
F	G	H

圖 42 8 鄰居範例圖。

$$X = A \cup B \cup C \cup D \cup E \cup F \cup G \cup H \quad (8)$$

$$X = A \cap B \cap C \cap D \cap E \cap F \cap G \cap H \quad (9)$$

影像之開閉運算即是使用膨脹與侵蝕兩種方法的運算組合，開運算 (opening) 是將影像先侵蝕運算再經過膨脹運算來處理，可清除影像之雜訊小點點，使物體輪廓平滑，消除細小峽部以及截斷窄的細頸 [13]，如圖 52、53；閉運算 (closing) 則是先使用膨脹運算再經由侵蝕運算來處理，可用來填補影像內之輪廓上的破洞或將一些狹窄的中斷部分和長細缺口連接起來 [13]。本系統之型態影像使用方式，因之前步驟做過影像相減後再經由膚色偵測的處理，造成了影像較為殘缺不堪，為了配合後續之條件判斷，我們將先做一次膨脹運算來處理，如圖 43、44。



圖 43 單人影像膨脹。



圖 44 雙人影像膨脹。

第五節 連通元件(Connected Component)

像素與像素間的連接是許多影像物體描述意義之簡化概念[2]，此即為連通元件(Connected Component)。連通的目的是為了讓二值化影像中像素屬於相同區域的連接起來，如此一來，便可以清楚的區分出每個像素分別在哪個區域，並可獲得各區域的長、寬大小，以利於後續之條件判斷。連通元件的方法是利用 8 鄰居鄰接的方式，利用堆疊的原理將相鄰的像素點連接起來，所有連接起來的像素都會被判斷為位在同一塊區域上。

舉例說明，如圖 42 所示，以 X 為中心點，A~H 為其 8 鄰居，若找到像素點時便設為 1，當 X=1 時，則會開始往 A、B、C、D、E、F、G、H 搜尋，沒有先後順序，只要其中一個不為 0，便會與 X 合併，形成同一個區域。

我們以圖 45 的 A 區域為例，連通技術將以堆疊的原理，由左上到右下的順序來搜尋區域，如圖 46(a)所示，首先搜尋到(2, 2)為 1，便將(2, 2)丟入圖 46 (b)的堆疊中，再以(2, 2)為中心點去找其 8 鄰居，並將(2, 2)自 stack 中 pop 出。接下來搜尋到(2, 3)、(3, 2)、(3, 3)三個點為 1，依序放入 stack 中，再以(3, 3)為中心去找其 8 鄰居，並將(3, 3)自 stack

中 pop 出，stack 如圖 47 (c)所示。之後再依序將(3, 3)的 8 鄰居為 1 的(3, 4)、(4, 2)、(4, 3)、(4, 4)一個個 push 進入 stack 中，再依序一個個 pop 出來，並以其為中心點去搜尋 8 鄰居中是否有相連之像素點，步驟如之前所述，如圖 48 所示。

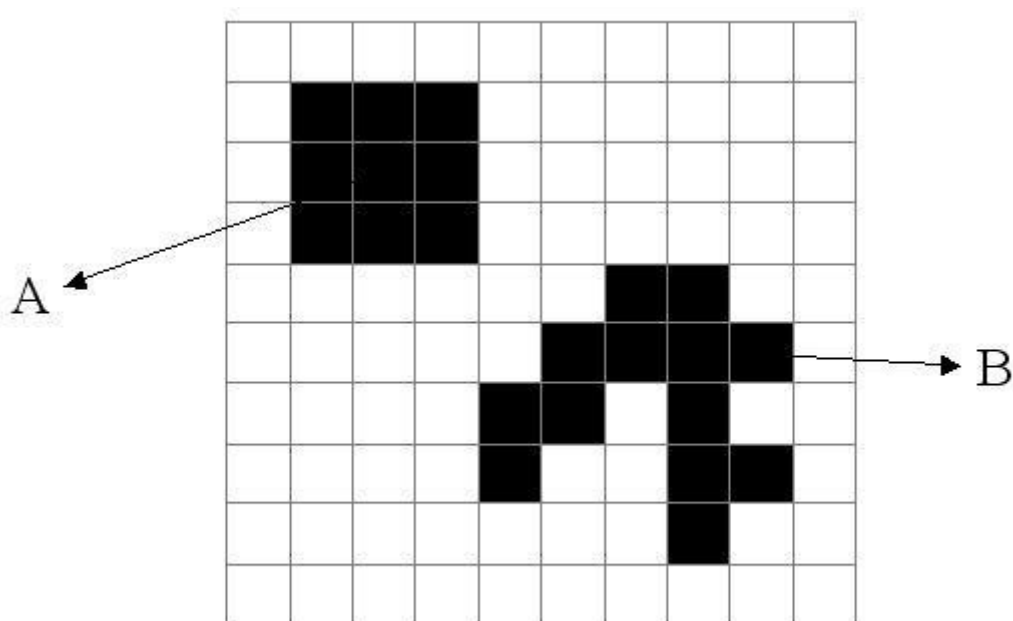


圖 45 以連通技術搜尋像素點。

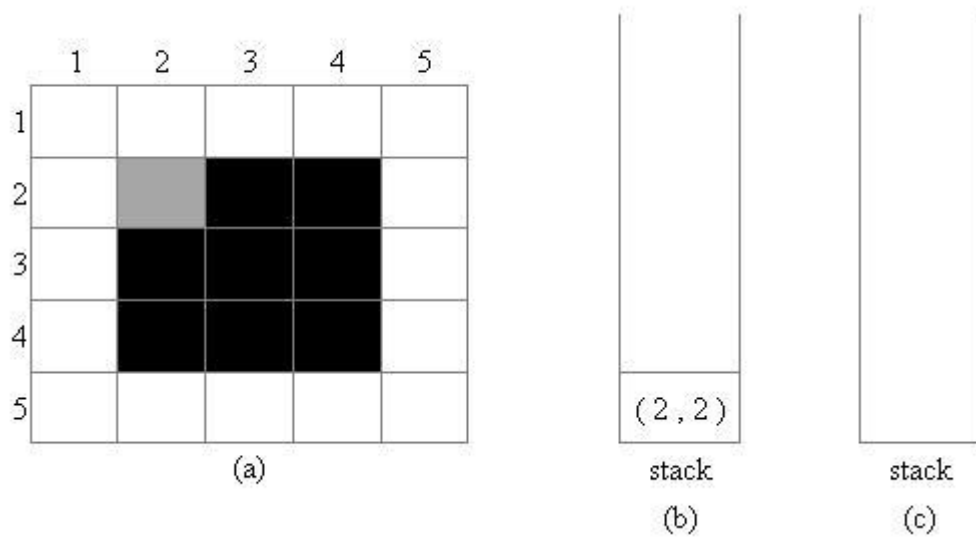


圖 46 以堆疊來實現連通技術搜尋像素點之範例圖 1。

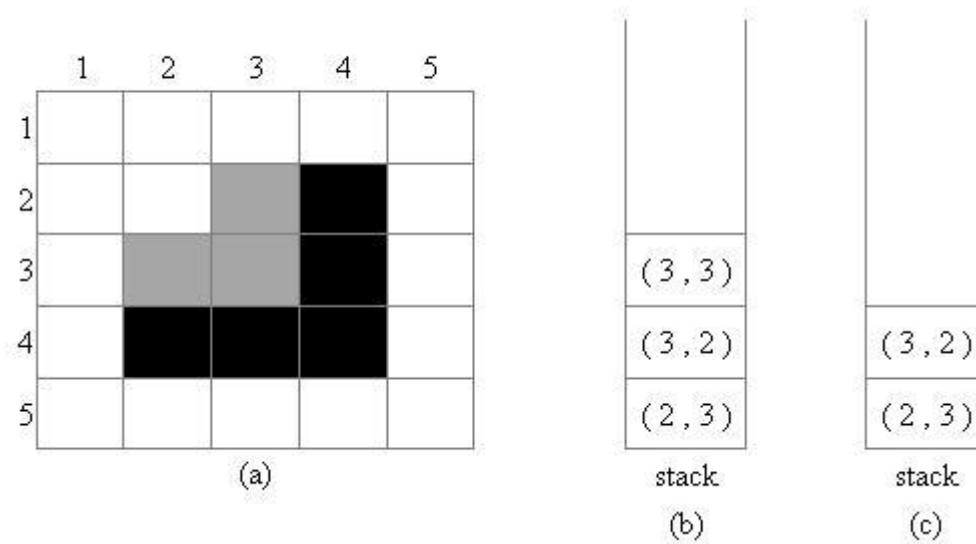


圖 47 以堆疊來實現連通技術搜尋像素點之範例圖 2。

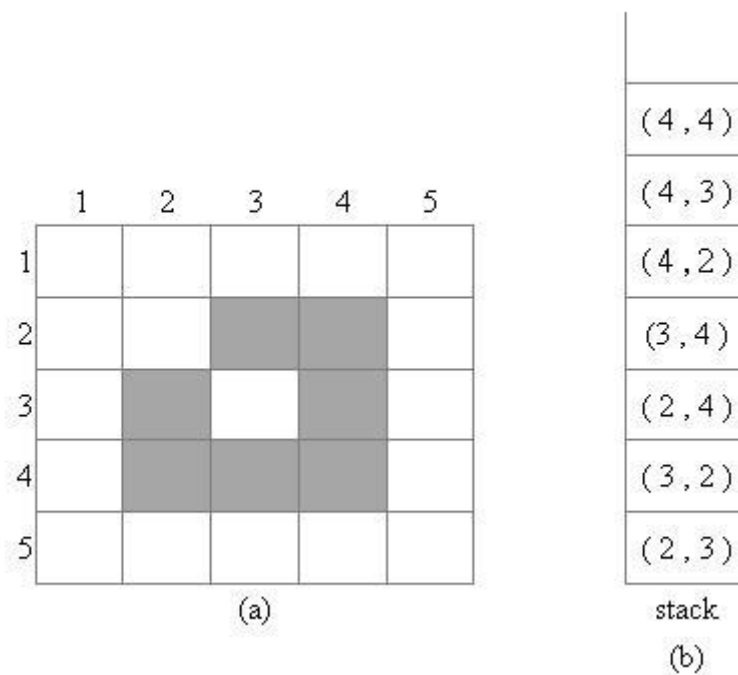


圖 48 以堆疊來實現連通技術搜尋像素點之範例圖 3。

完成後，將偵測到整塊 A 區域為相連的像素，並將其用標示起來，如圖 49 所示。當 A 區域搜尋完成後，如圖 50 所示，再由左上到右下重新搜尋一次，B 區域做法同上所述，完成後如圖 51 所示。我們將膨脹後的圖做連通技術，如圖 52、53 所示，可偵測出其長、寬之值以利後續條件判斷是否存在人臉區域。

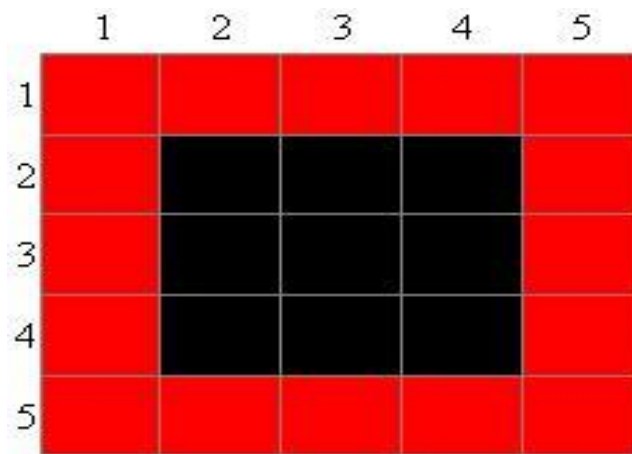


圖 49 連通技術搜尋像素點之完成圖。

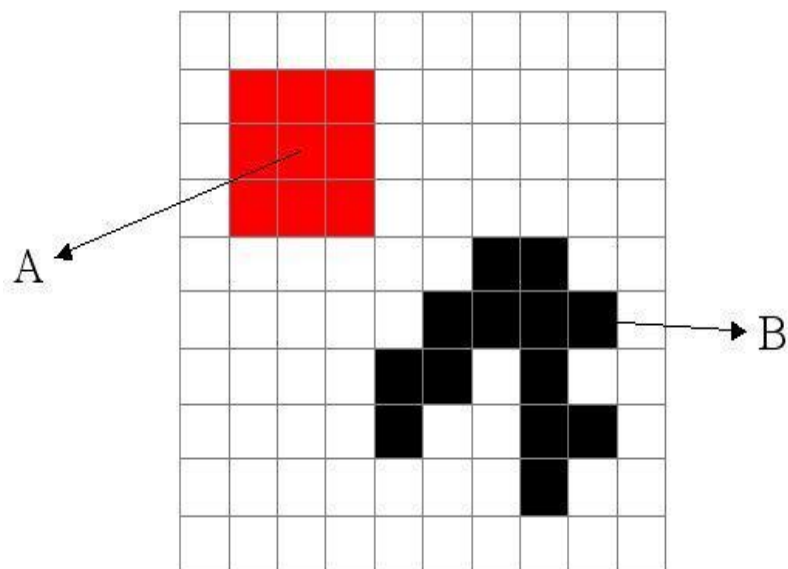


圖 50 以連通技術搜尋像素點 A 區域完成。

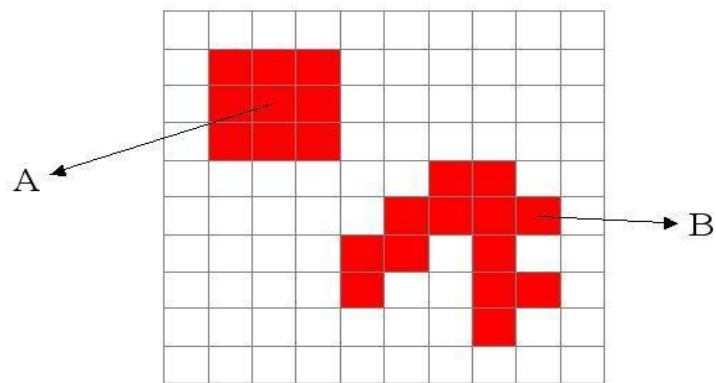


圖 51 以連通技術搜尋像素點 B 區域完成。

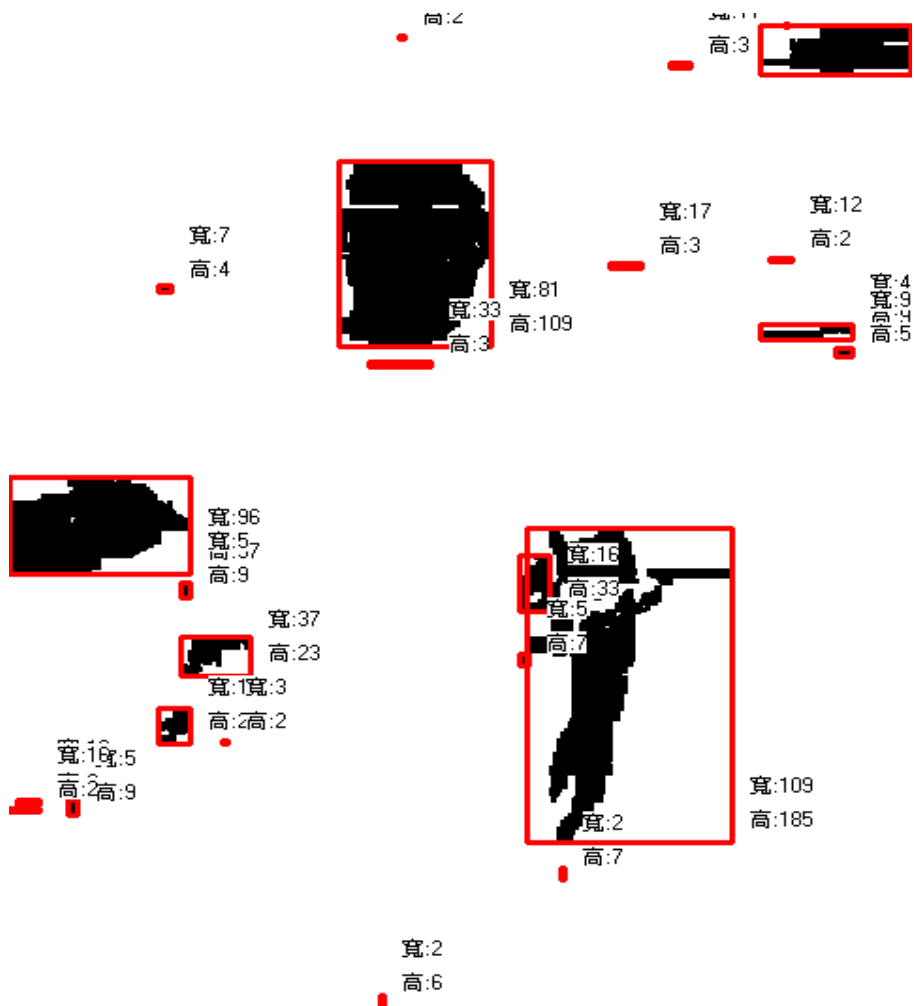


圖 52 單人之 Connected Component 並計算出其寬、高之值。

第六節 條件判斷搜尋候選臉區域

以 Connected Component 找出相連區域後，為了進一步刪除非人臉之區域，我們將設定條件來刪除與人臉差異過大之區域，條件判斷方式有三個：

1. 區域高、寬比之比值在 0.8~2.5 之間的區域，如圖 54、55，不符合之區域將被刪除。
2. 區域高、寬之值各在 20 以上之區域，如圖 56、57，不符合之區域將被刪除。
3. 區域像素點佔區域的 50% 以上，如圖 58、59，不符合之區域將被刪除。

只要不符合上述三個判斷條件之區域，我們就將其刪除，符合的將留下來列為候選臉區域。

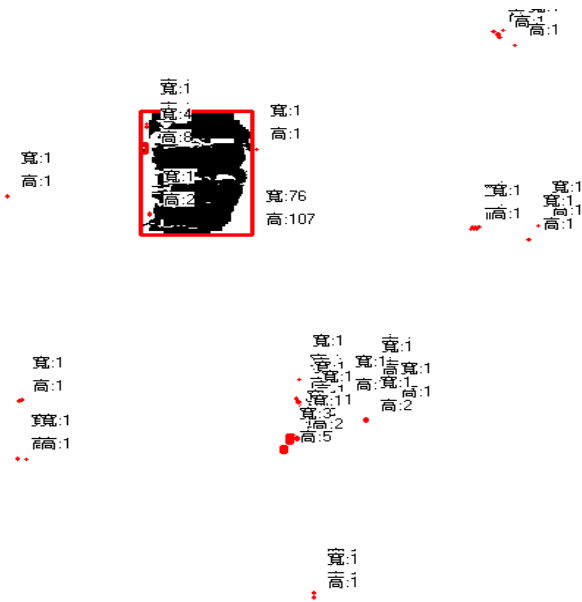


圖 54 單人影像中刪除不符合高、寬比之比值在 0.8~2.5 之間的區域。

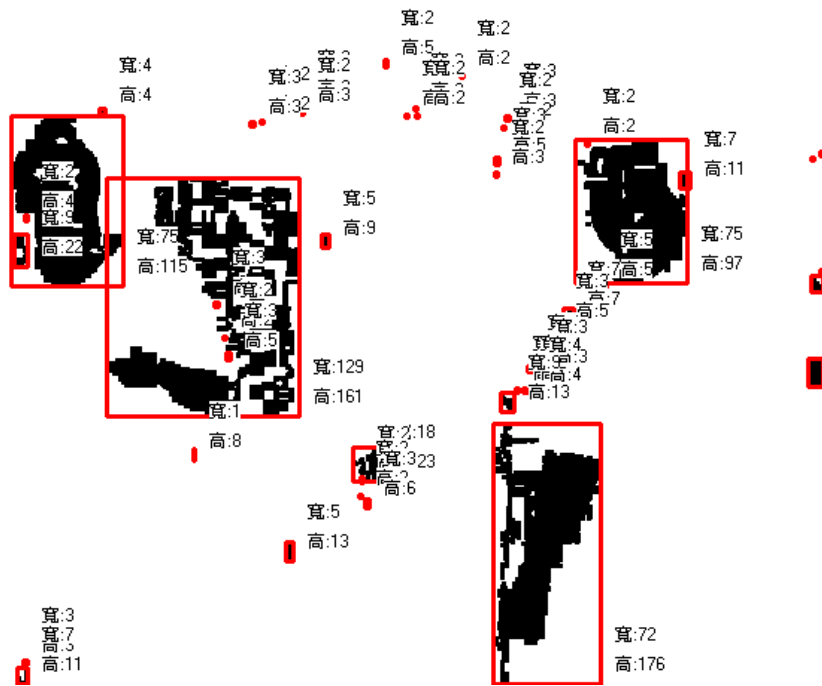


圖 55 雙人影像中刪除不符合高、寬比之比值在 0.8~2.5 之間的區域。

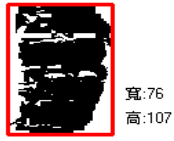


圖 56 單人影像中刪除不符合高、寬之值各在 20 以上之區域。

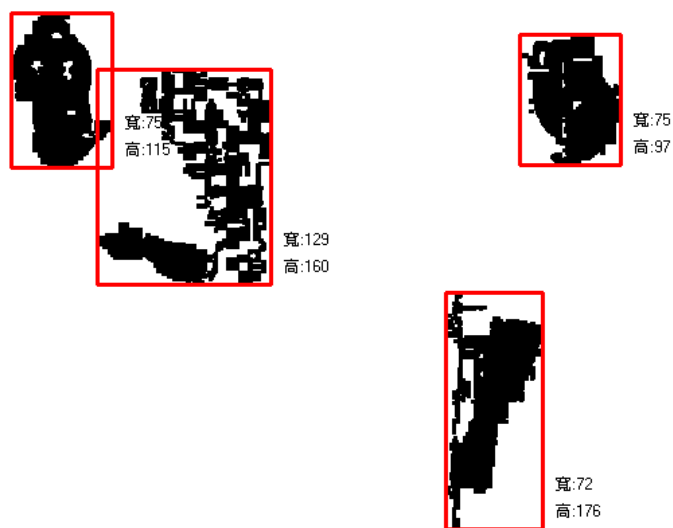


圖 57 雙人影像中刪除不符合高、寬之值各在 20 以上之區域。

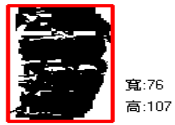


圖 58 單人影像中刪除不符合像素點佔區域的 50%以上之區域。



圖 59 雙人影像中刪除不符合像素點佔區域的 50%以上之區域。

第七節 灰階化(gray levels)

在經過一連串的條件判斷後，剩下的區域即為候選臉區域，表示可能為人臉所在之位置，但並非完全正確，所以我們將進一步的使用型態影像學中的灰階化(gray levels)來做判斷，增強影像中 useful 資訊，使品質提高，以便肉眼觀察與分析，讓系統做進一步的處理[22][26]。

在臉部的特徵上，眉毛、瞳孔、嘴唇在灰階化後的顯示上都比其他區域來的更黑更暗，可利用此特性來區分臉部內五官或非五官之區域。倘使將整張 image 做灰階化(gray levels)的動作，因為背景亮度不同，某些地方也會變暗，將影響到臉部的五官偵測，所以我們只針對可能為臉部區塊做灰階化的動作。

灰階化主要是用來區分色彩上的亮度，前面所敘述的 YCbCr 色彩空間中的 Y，便是在對亮度做處理的工作，其轉換公式如公式(10)所示，其轉換結果如圖 60、61 所示。

$$Y = 0.299 \times R + 0.587 \times G + 0.114 \times B \quad (10)$$

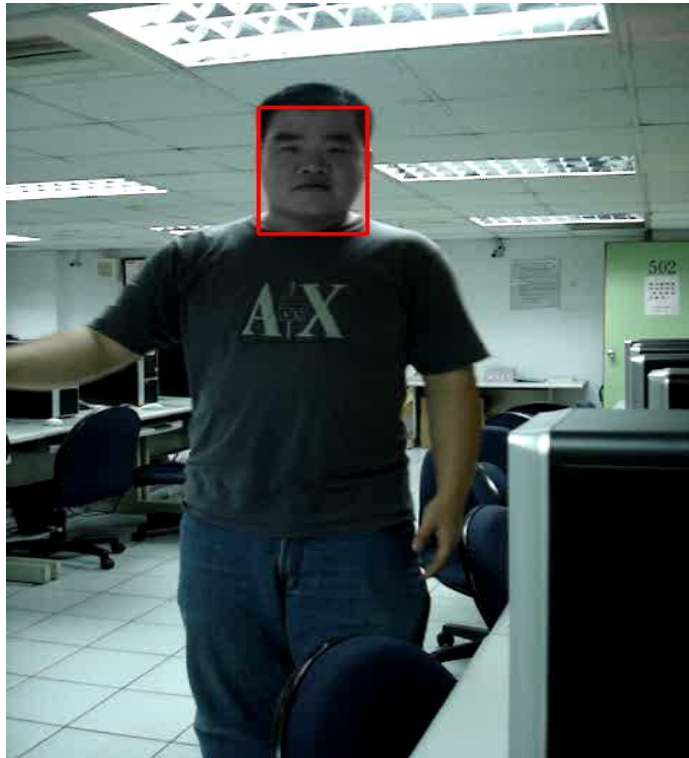


圖 60 單人候選臉區域灰階化。



圖 61 雙人候選臉區域灰階化。

第八節 搜尋眼嘴之位置

為了更進一步偵測出剩下的區域是否為人臉，我們將依據人臉之五官特徵來加以判別。人臉的特徵如眼睛、鼻孔、嘴巴、眉毛等都是偵測人臉所用的參考方向[25]，因為人臉的五官有其固定位置，形狀很穩定，不像臉部邊緣容易受到解析度與表情而產生改變[17]，而鼻孔因臉部的角度可能遭到遮蔽，造成搜尋上的困難，因此只選擇眼睛與嘴巴部分[34]，在前面經由候選臉區域灰階化後，我們將再進行以下步驟來加以搜尋眼、嘴之位置。

(一) 二值化

灰階化完成後，我們將利用灰階化特性之亮度值做二值化動作讓眼睛與嘴巴之區域更明顯地區分出來。一般以灰階值來解釋亮度大多以位元組(Byte)來表示像素中的色階，所以每個像素的亮度等級分為 256 組色階，0 為黑色，255 為白色[22][23]。我們設立一門檻值 T ，取 $0 \sim T$ 設為黑色， $T \sim 255$ 設為白色，將色彩的亮度作為一個區分，將可有效的區隔出亮與暗的差異[21]，如圖 62 所示。

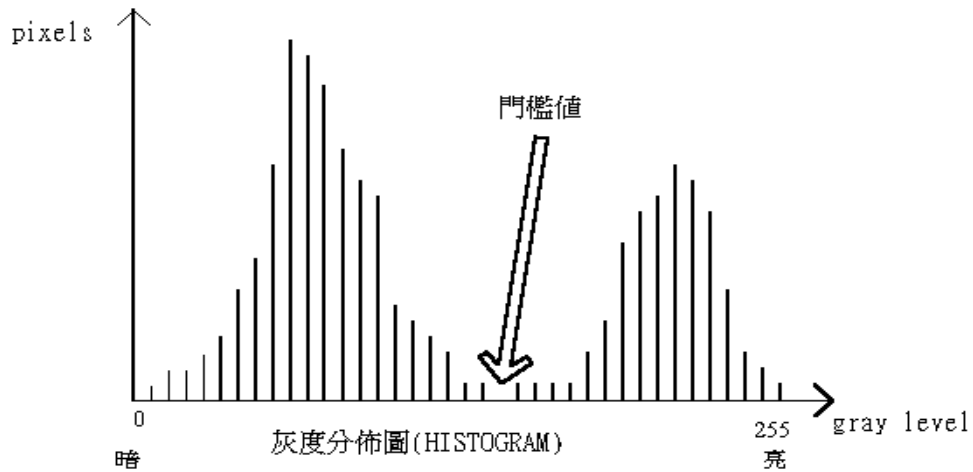


圖 62 以門檻值區分灰階亮度之分布圖。

(二) 中間值濾波(median filter)[14]

經由二值化後，此時臉部將會出現一些雜訊，可經由中間值濾波(median filter)技術來去除雜訊，增加影像識別的成功率。

在數位影像處理中，常藉由濾波器來去除雜訊，其中中間值濾波技術可以很平均的把影像重點部份表現出來[31]。中間值濾波是將一像素與鄰近區域的像素依序做灰階值排列，再取其中間(medium)灰階值(gray scale level)來取代目前像素的灰階值，使其更接近鄰域像素值[30]。我們舉例說明中間值濾波處理方式。

例如在一個 3×3 的區塊內，其排序後的灰階值中的第5個值為其中間值；在 5×5 的區塊內，其排序後的第13個灰階值為其中間值，以此類推。

假設有一個 3×3 區域，如圖63所示，其區域內像素值為(80, 10, 30,

35, 10, 50, 5, 100, 35), 經由中間值濾波排序後為(5, 10, 30, 35, 35, 35, 50, 80, 100), 中間值為35, 取代原本的10, 5×5 區塊中也是相同作法。消除雜訊後得到結果如圖64、65所示。

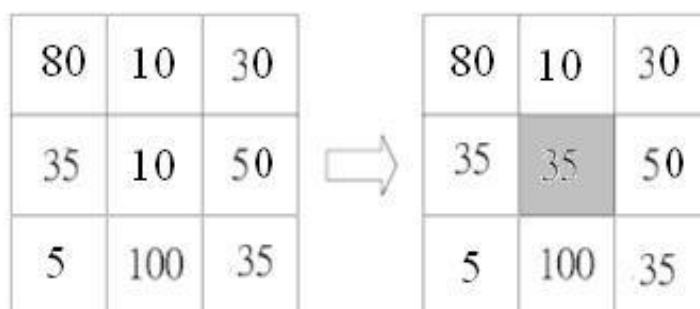


圖 63 中間值濾波範例圖。



圖 64 單人候選臉區域二值化後經中間值濾波之結果。



圖 65 雙人候選臉區域二值化後經中間值濾波之結果。

(三) Connected-Component 並做條件判斷

再對圖 64、65 做 Connected-Component，計算出區域的寬、高後，針對寬 > 高、寬 ≥ 5 、高 ≥ 3 且區域與邊界相鄰距離 ≥ 5 (單位皆為 pixel) 等 4 個條件來加以判斷不符合之區域，只要區域不符合上述之四個條件，表示其不符合眼、嘴在臉部之位置條件，便不選取。找出眼、嘴之位置後將其選取用紅色框框出，如圖 66、67 所示。



圖 66 單人候選臉 Connected-Component 後並做條件判斷。



圖 67 雙人候選臉 Connected-Component 後並做條件判斷。

第九節 正規化(Normalize)

人臉偵測必須要有其合理的適用範圍，因此要達到提高偵測率的目標，必需要一有個適當的人臉影像前置作業機制，用以減少輸入的影像資料間的差異性，去除不必要的雜點，以得到標準的影像形式[18][23]。

影像正規化分為影像大小正規化和色彩強度正規化[32]，因影像皆已做二值化動作，只有黑與白，故我們主要目的是做正規化大小的動作，將影像轉換成統一的大小尺寸規格。

正規化是指將原本長寬不一的所有影像，使其長寬變為相同，並將色彩影像轉換為灰階化影像[24][28]。當我們將每張影像搜尋出候選臉區域後，每張影像之候選臉區域大小並不一致，必須將其轉成大小一致的影像，也就是說對於所有長與寬為 $m \times n$ 的影像都將它正規化成 $k \times k$ 的影像。如此一來，在做影像的比較時，才能站在同一個基準點上[19]。

我們再針對圖 66、67 剩餘的區域找出其中心點之位置 O ，將所有的中心點位置取平均值，找出候選臉區域之中心點位置，以此中心點向外找寬、高比為 2 比 3 之比例的範圍，如公式(11)所示，將此範圍取 20×20 的正規化動作，如圖 68、69 所示。

$$H = W = (2 \times O) \div 3 \quad (11)$$



圖 68 單人候選臉區域正規化之放大示意圖。



圖 69 雙人候選臉區域其中之一正規化之放大示意圖。

第十節 支撐向量機器 SVM (Support Vector Machine)

支持向量機 SVM(Support Vector Machines)是貝爾實驗室研究人員 V. Vapnik 等人在對統計學習理論多年的研究基礎上發展起來的一種全新的機器學習演算法，SVM 並不像傳統的學習機器以經驗風險最小化 (Minimize Empirical Risk) 為目標，而是以結構風險最小化 (Minimize Structural Risk) 為目標[20]。

顯性特徵的人臉偵測方式通常以統計分析(statistical analysis)及機器學習(machine learning)這兩樣技術的相關特徵來區分是否為人臉，而 SVM 就是在統計學習理論的基礎上發展起來的新一代學習演算法[13]，在文本分類、手寫識別、圖像分類、生物資訊學或辨識等領域中獲得了較好的應用，為近年來非常熱門的一項技術[20]。

在前面步驟所得如圖 66、67 的結果，我們已經過了初步的過濾，但因為人臉中，眼睛、嘴巴區塊搜尋精確度較不穩定，我們將可能是人臉的候選臉區域交由後續處理，為了讓可能的區域有更高的判斷結果，我們引用了支撐向量法(Support Vector Machine)[15][16]做最後的驗證。支撐向量法在近幾年來是一個相當成功的分類器，此分類器將在低維度的資料無法簡單的分類問題，映射到高維度空間，在高維度空間內，以線性

可分割的問題加以分類。由於此研究的問題很單純的只有兩類(人臉/非人臉)，我們可使用兩類之線性可分割支撐向量法加以處理。利用此特性，將前面步驟中，我們已經找到可能是人臉中的眼睛、嘴巴的位置，往外擴張一定距離，找到一個能框住人臉的主要器官影像，並正規化成一定大小，再送入支撐向量機器內進行分類。在使用支撐向量機器之前，我們必須先對機器加以訓練，首先，我們先以人工的方式，切割出各 300 張人臉與非人臉的區塊，並將這 600 張影像正規化成 20×20 的大小，送入支撐向量機器內加以訓練參數並儲存。在此研究中，我們使用的核心函數為放射基礎函數(radial basis function)，由於訓練的人臉/非人臉樣本有限，無法代表所有可能的樣本空間，因此給予機器在學習過程中有適當的懲罰值是必要的。初步的實驗中，我們給予懲罰值為 1000 時，可得到不錯的結果，最後偵測結果如圖 70、71 所示，即使有其他非人臉區域被選為候選臉區域，經由 SVM 技術驗證後便能有效的刪除。

最後，我們再將本章所有的實驗步驟做個完整的流程圖，如圖 72 所示。



圖 70 單人影像完成人臉偵測。



圖 71 雙人影像完成人臉偵測。

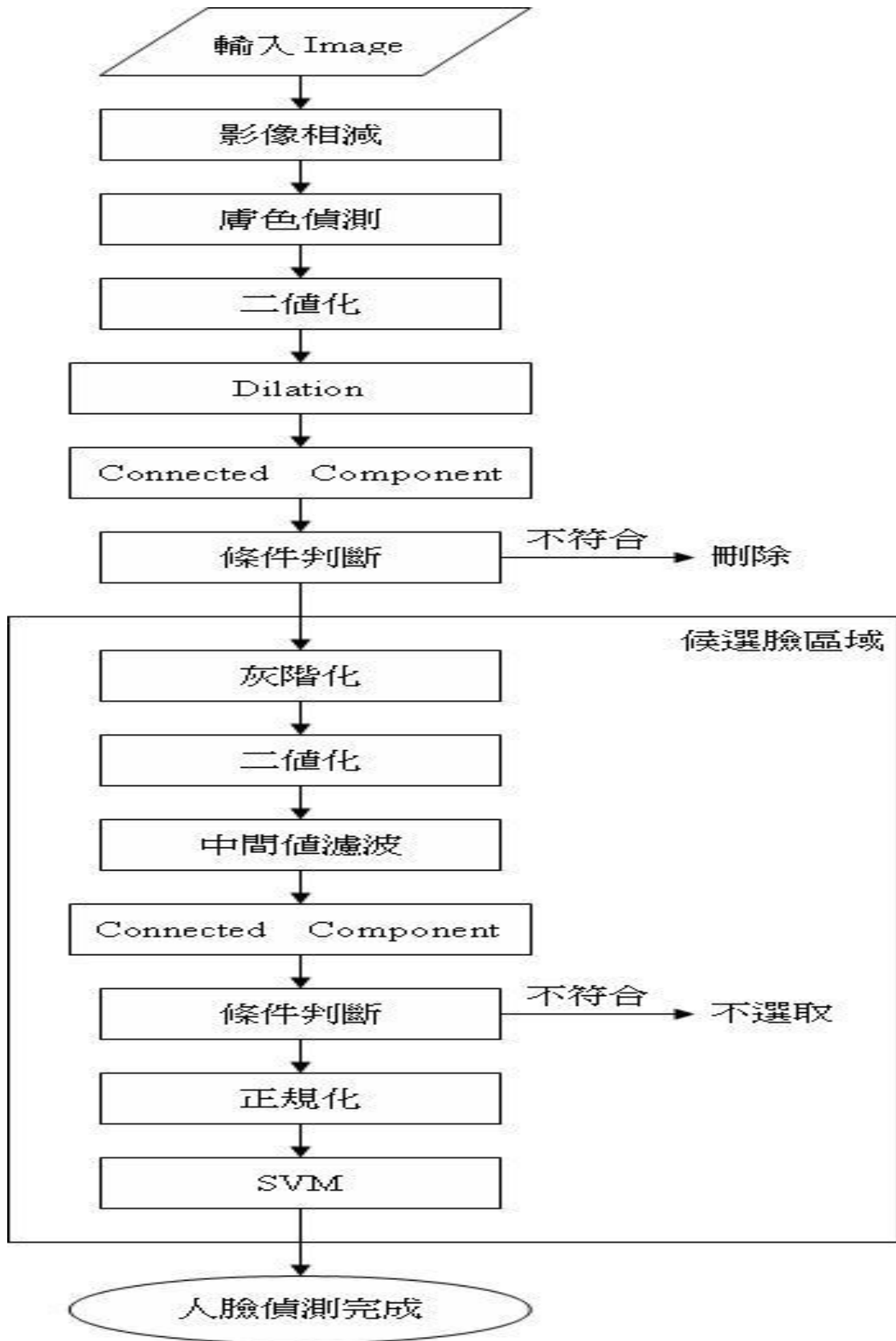


圖 72 室內環境人臉偵測方法之流程圖。

第肆章 實驗結果與分析

第一節 實驗結果

系統在順利完成後，對許多影片做實驗也得到了不錯的數據，我們取其中三段影片的實驗結果，分為 4 欄來分析數據，分別是：

1. Correction(正確率)：人臉被正確偵測到。
2. 錯誤接收率 FAR(False Acceptation Rate)：非人臉區域卻被偵測為人臉，這指標相當重要，因為有些演算法號稱有 100% 成功率，可是錯誤率卻相當高[13]。
3. 錯誤拒絕率 FRR(False Rejection Rate)：是人臉卻未被偵測到。
4. Total(總人頭數)：每張影像中人頭數加起來的總和。

如表 1 所示，實驗結果展示在圖 73。圖 73 左邊是展示一個最佳的情況下偵測出人臉，而右邊是展示非最佳情況（角度不正、人臉太小、太暗等）下依然有效地偵測出人臉來。我們擷取每張影像當作測試結果，最後並將其合成為 AVI(Audio Video Interleav)格式的影片，撥放時將可經楚看出人臉清楚的被偵測到。我們是用數位相機 SONY DSC-W100、8.1 Mega Pixels 來拍攝影片，電腦 CPU 是使用 ASUS P5GD1-VM Pentium D

2.8GHZ。

實驗結果與數據可看出，雖然並非每段影片的人臉偵測 Correction 都能達到零錯誤，多少會受到光線陰暗、人臉角度等影響，但還是得到一個不錯的實驗結果。

表 1 實驗數據表。

	Correction	FAR	FRR	Total
Video1	93%	10%	7%	41
Video2	92%	30%	8%	161
Video3	85%	13 %	15%	202


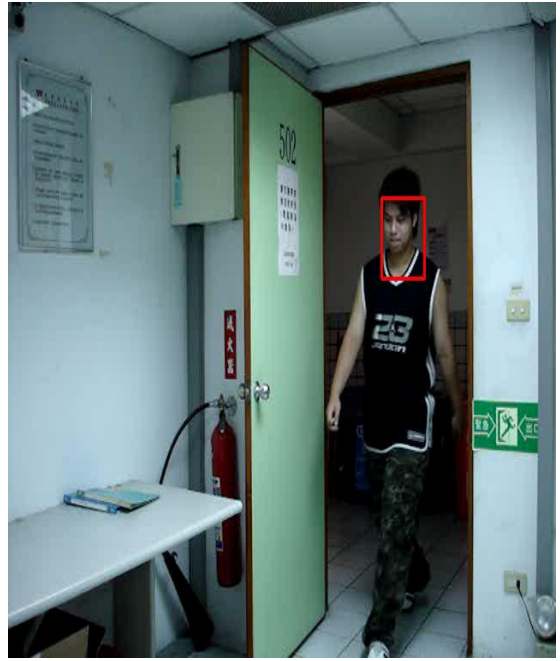
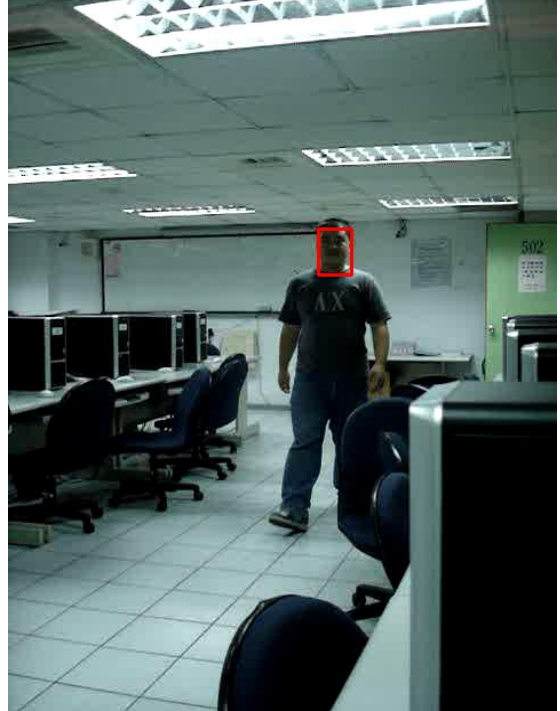
	偵測成功範例	人臉太小、太暗、角度不正等情況下偵測成功
Video1		
Video2		



圖 73 實驗後人臉偵測結果。

第二節 效能分析

雖然成功率頗高，但實驗中也有許多失敗的案例，我們取一些較具代表性的來一一分析其原因。

如圖 74 左上圖所示，人臉未被偵測出來，依據右上圖圈起來的部位顯示，在移動時，因光線角度的影響，讓背景某些區域受移動物陰影遮蔽，以致於影像相減後也被保留下來，且又與頭部相連起來，而左下圖中，此區域顏色又與膚色相近，故無法有效進行膚色偵測。右下圖已經做到條件判斷，但其區域不符合像素點佔區域的 50% 以上，所以區域被刪除，人臉未被偵測出來。

圖 75 中左上圖左邊的人臉未被偵測到，因影像相減再經由膚色偵測後，臉部已破碎不堪，如右上圖所示。即使經由膨脹，在左下圖 Connected Component 後，仍沒有達到高、寬比之比值在 0.8~2.5 之間，故右下圖中人臉被刪除。

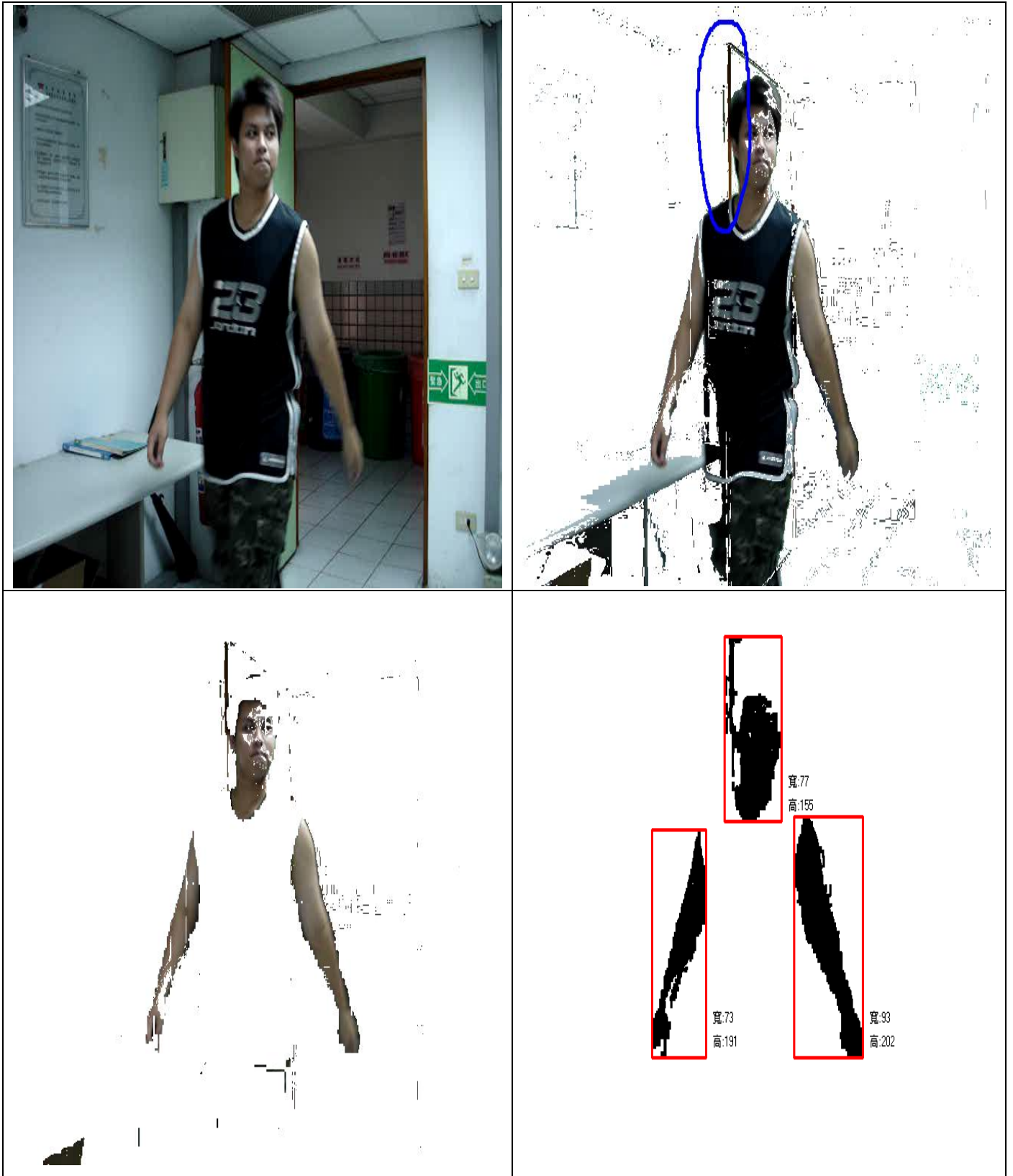


圖 74 Video1 實驗後人臉未被偵測出來。

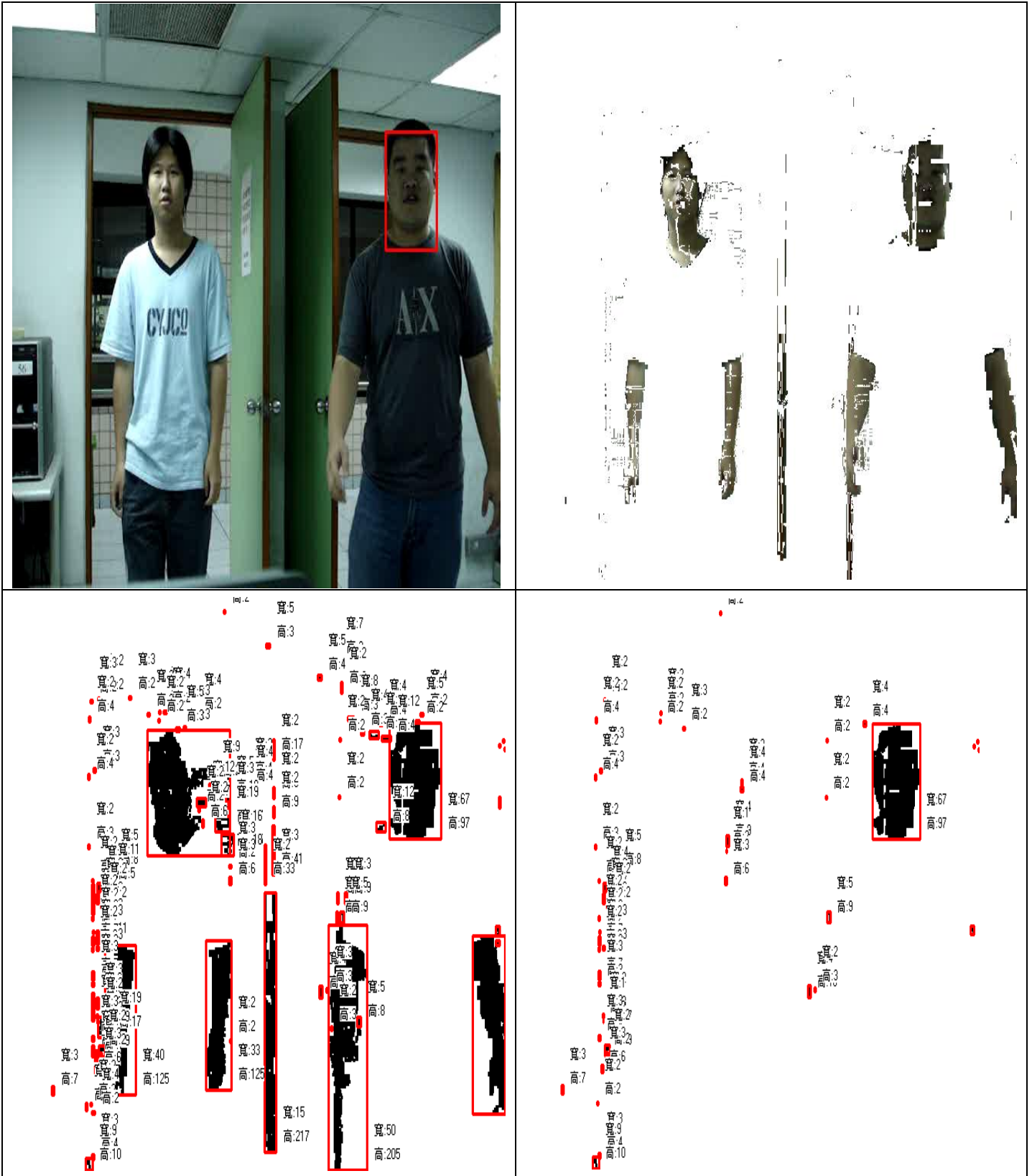


圖 75 Video3 實驗後左邊的人臉未被偵測出來。

圖 76 中 Video1 與 Video2 中也有背景與手臂被偵測為人臉的情況，圖 77 Video3 也是相同的情形，分析其原因，在前面三個條件判斷時非人臉區域大小恰好符合條件，以致於沒有被刪除，倘使為了讓這些非人臉區域能順利刪除，勢必要將條件判斷設的更為嚴苛，如此一來，將會影響到人臉區域之偵測，人臉區域將有可能被一併刪除。所以依據情況來手動設定條件值，是本系統較為不方便處，也是未來需要改善的方向。

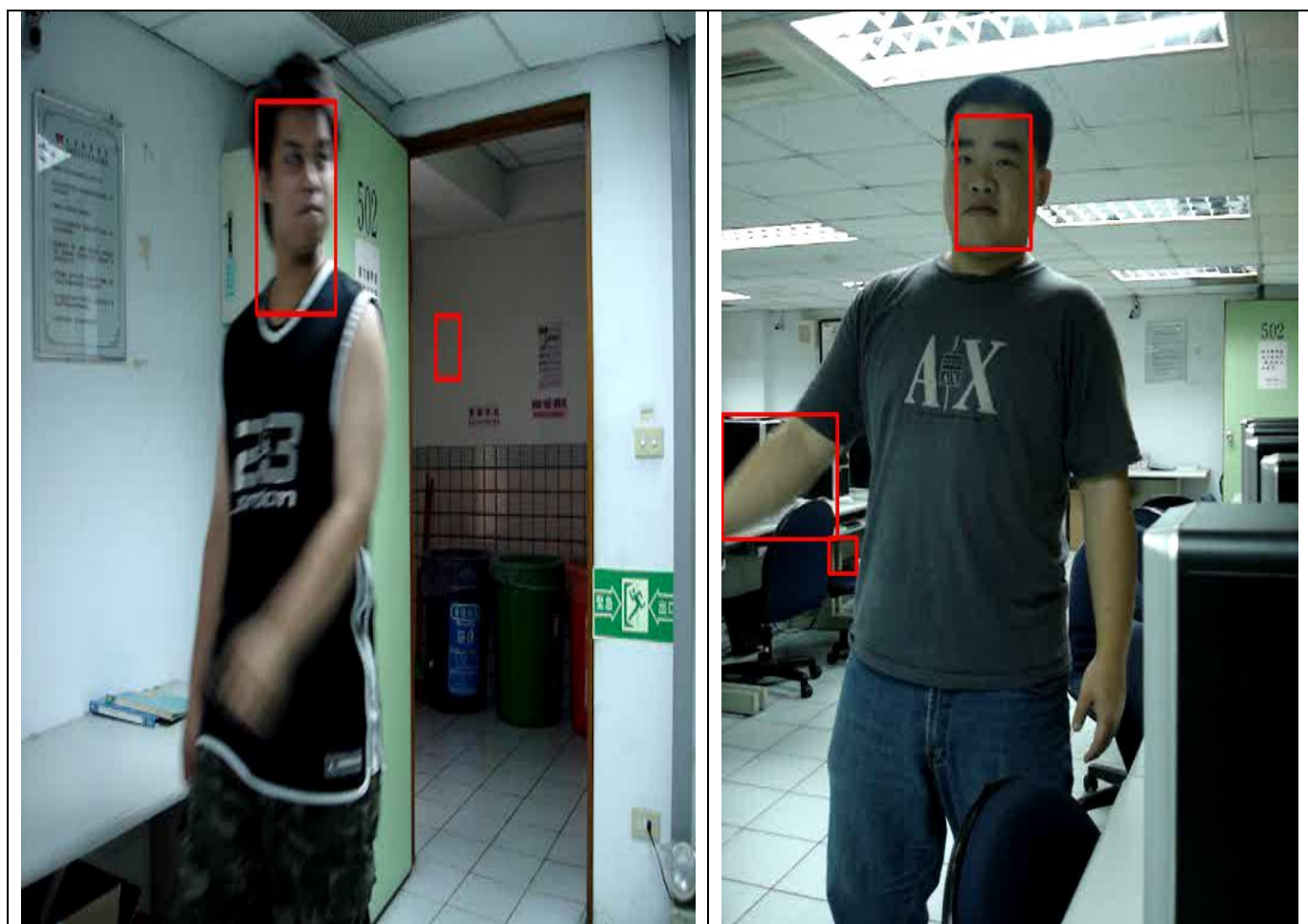
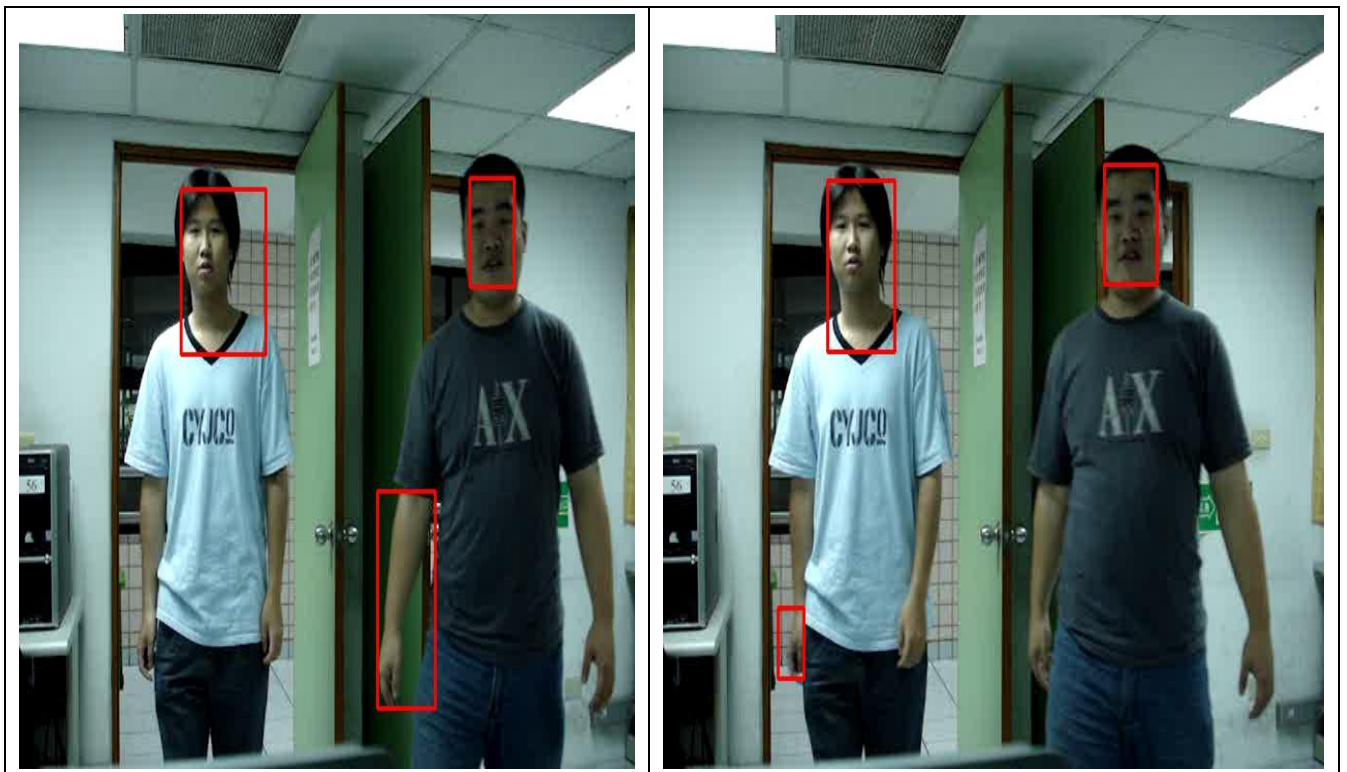




圖 76 Video1 與 Video2 實驗後錯誤接收率 FAR 出現情況。



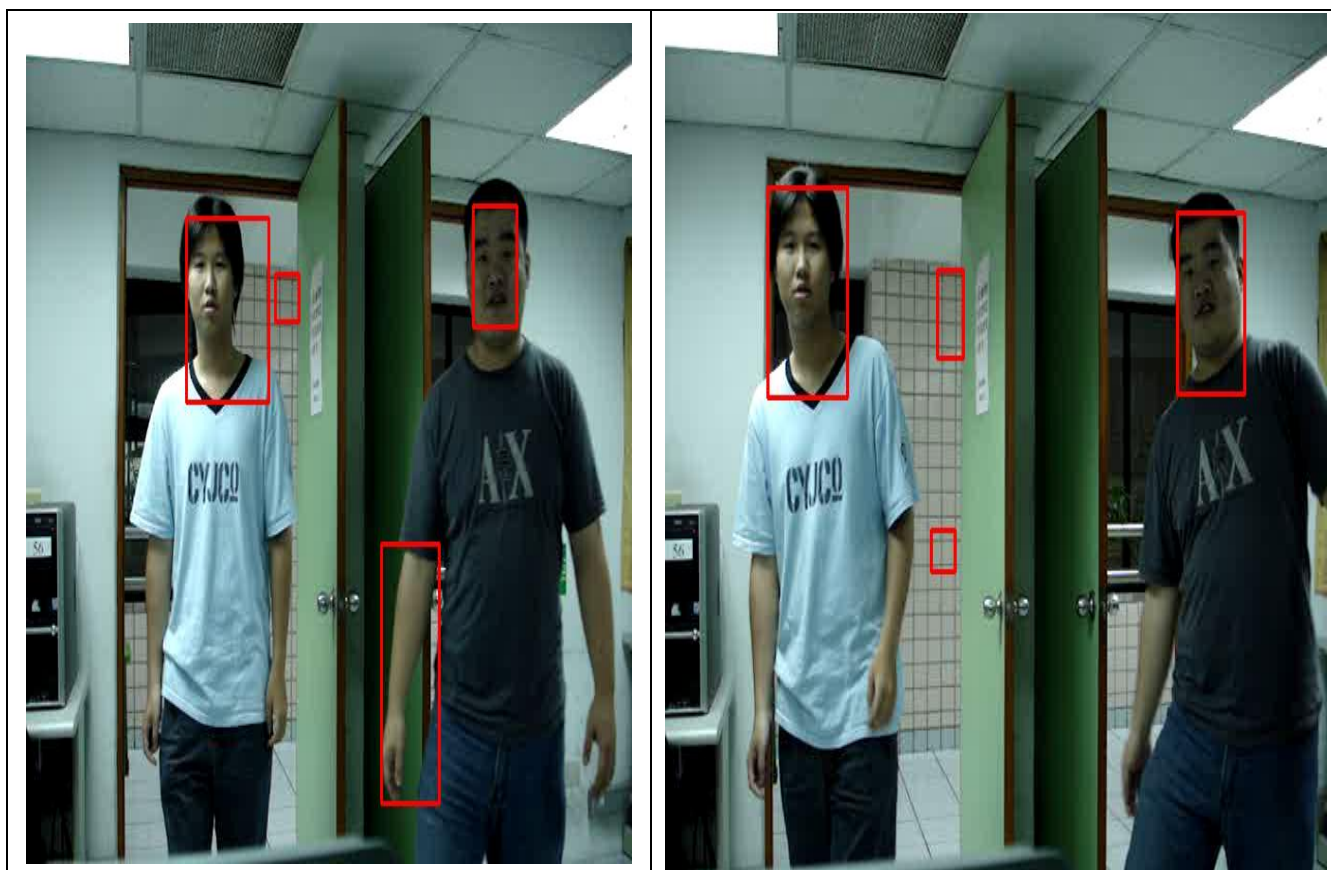


圖 77 Video3 實驗後實驗後錯誤接收率 FAR 出現情況。

實驗中，為了改善背景也被偵測出來的情況，我們針對沒有加 Dilation 的技術下做實驗，因為 Dilation 後讓某些背景的像素點跟著增加，因此可能讓這些背景區域也符合了條件判斷，以致於沒有順利刪除。實驗結果發現錯誤接收率 FAR 確實下降不少，可是正確率也跟著下降，為此，我們確定必須要做 Dilation 才可保證 Correction 有一定的水準。

表 2 沒有加 Dilation 的實驗數據表。

	Correction	FAR	FRR	Total
Video1	85%	6%	15%	41
Video2	90%	19%	10%	161
Video3	80%	12 %	20%	202

第三節 系統實現的困難與待改處

系統完成後，尚有許多設計上困難的地方與待改進的空間，諸如下列所述：

1. Dilation 有助於破碎的人臉區域進行修補，但也造成背景膨脹或人臉與背景相連，條件判斷的過程將受影響。
2. 系統參數值需視情況手動修正，無法自動修正是系統中較為不方便的地方。
3. 中間值濾波可消除臉部區域細小雜訊，但也影響到整個系統的速度較為延緩。
4. 在系統實現的過程中，雖然背景固定、光源穩定，但由於系統對光的敏感度特別高，連帶的也影響到了後續的判斷，造成某些人臉區域未被偵測出來或非人臉卻被偵測為人臉區域。
5. 多人的人臉偵測正確率比單人的來的較低一些，其原因可能為畫面大小固定的情況下，越多人入鏡人臉將越小，造成後續人臉特徵之判別較為不易。
6. 系統中也只能針對黃種人的膚色來做偵測，若要偵測出其他人種的膚色，必須要先建立起膚色的資料庫，並且導入演算法來加以比對判

別，但有可能會因為資料庫龐大，不斷的進行資料比對，影響到整體系統的執行速度。

7. 背景中若有與膚色相近的顏色，可利用系統中影像相減的技術來排除，倘若待測的人身上若穿著顏色近膚色系的衣服，如橘、紅、黃等顏色，或身上並未穿著衣物(打赤膊)，將會造成系統膚色偵測上的困難。
8. 五官若有其他裝飾(眼鏡、唇環、鼻環)，或影片中人體移動速度太快，也會影響到系統偵測的正確率。

在本專題研究中，遭遇了上述不少困難點，很難有效的達到百分之百的完美偵測率，許許多多的困難處短期內也難以解決。但相對來講，本專題研究已建構起一個人臉偵測的基礎，未來將有許多改善的方向，以期後續的研究能使人臉偵測或人臉辨識系統能達到一個更完善的空間。

第五章 結論

在本研究中，我們利用了影片中物體移動的特性，針對室內環境進行人臉偵測，並透過 YCbCr 的彩色空間，對偵測到移動區域進行膚色過濾，保留下可能為人臉的區塊。實驗過程中發現，除了人臉的區域被保留下之外，還有許許多多的雜訊亦殘留下來，因此，針對每個相連區域，使用連通成分(Connected -Component)技術，找出個別獨立區域的寬、高、及該區塊內屬於膚色像素點數等資訊加以初步過濾。雖然，經過初步過濾後，已大量排除了非人臉區域，但還是有誤判的情況，例如人的四肢，膚色相近之背景等區域，因此我們考慮到人臉上重要的特徵：五官。

因為人臉的五官有其固定位置，形狀很穩定，在經過灰階取亮度做二值化後，五官這些位置會被標示為非膚色的像素(本專題標示黑色)，經由在剩下的可能為人臉的候選臉區域中，找出眼睛、嘴巴的相對位置，可進一步提高該區域是否為人臉的區域。

最後，我們再利用近來相當成功的分類器—支撐向量機器(Support Vector Machine)加以分類，即使前面步驟將不是人臉的區域也判斷為候選臉區域，經由 SVM 技術後便能大幅改善，有效降低錯誤接收率 FAR，提高正確率及降低誤判率。最後的結果顯示，本專題之可行性及實用性質

相當高，測試數段影片後的成功率多半高達 80%~90% 以上，系統執行速度也頗為迅速。

在本專題完成後，雖然成功率已經頗高，但還有許多進步的空間，如人種膚色的問題、影像中人的位置太遠造成人頭過小的解決問題、人臉面對鏡頭的角度非正臉等，都是些可以再加強的方向。

微軟總裁比爾·蓋茲說過：「生物辨識技術將是二十一世紀重要技術之一。」，生物辨識包含人臉辨識、指紋辨識、虹膜辨識。人臉偵測更是人臉辨識中的首要課題。

本系統未來可應用在相當多的領域，如停車場安全、科技大樓保全、居家生活老人或小孩之安全、教室內學生人數統計，以及後續之人臉辨識等。

未來的趨勢將是不斷的以科技代替人力，在網路發展的普及下，台灣已有 67.21% 的人口，約 1 千 5 佰萬人(2006，財團法人網路資訊中心) 在使用網路，大陸更高達 12.3%，約 1.62 億(2007，CNNIC) 的網路使用人口，若將人臉偵測技術與網路結合，將可發展出遠端監控系統，如居家生活老人與小孩之照護，利用特定動作(跌倒、手勢辨識) 發生或陌生人進入場所時，發出訊號來警告遠端監控者，如圖 77 所示。人臉偵測的好壞影響著後續的系統研發，期望著後續研究多方嘗試，以達到最佳的效

果與實用性。

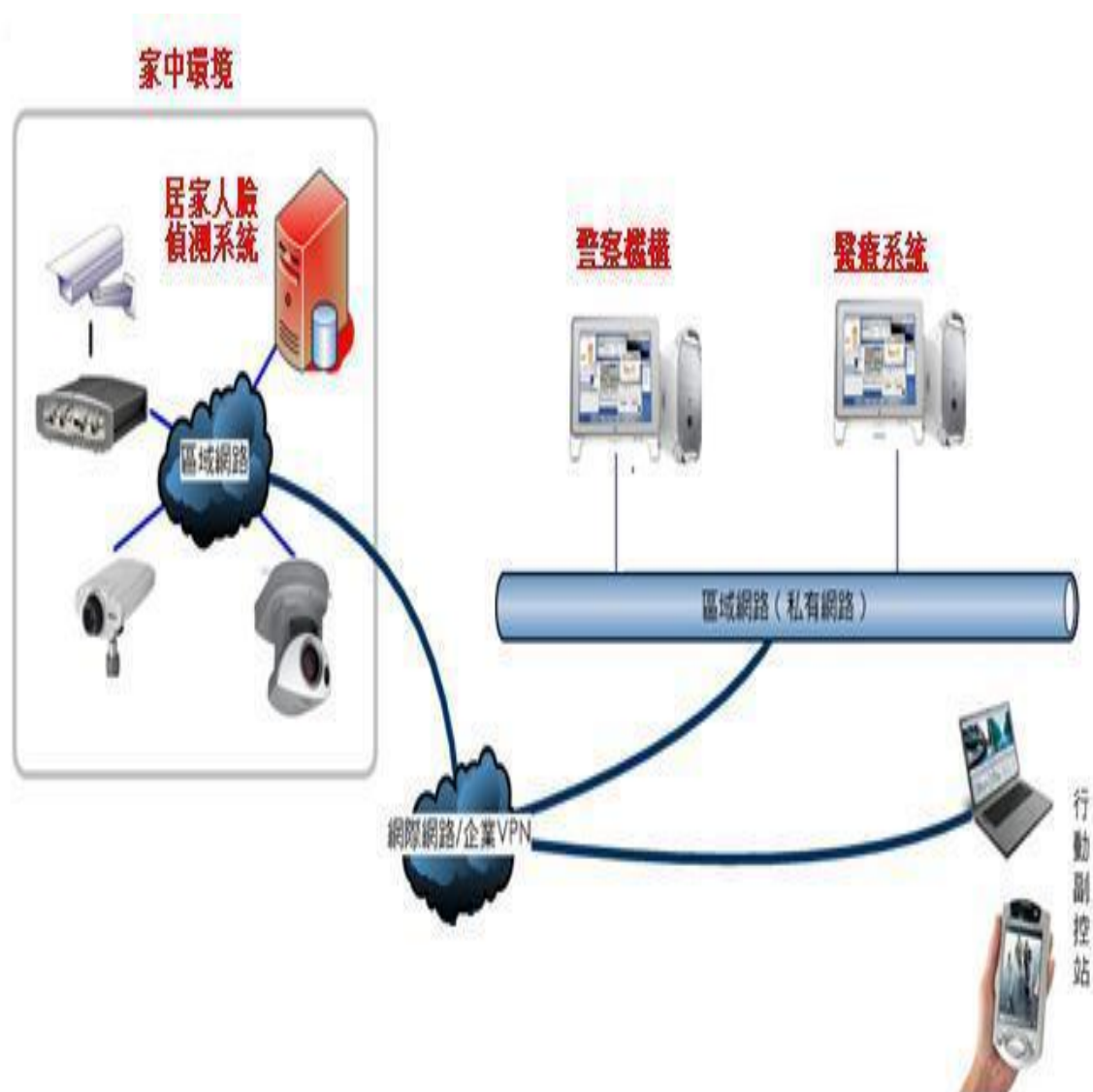


圖 78 人臉偵測與網路結合成居家監護系統。

參考文獻

- [1] 馬炳義，民國 95 年 6 月，人臉偵測技術研究，義守大學資訊工程研究所碩士論文。
- [2] 黃建邦，民國 94 年 7 月，彩色視覺技術於多人臉辨識上的應用，崑山科技大學電機工程系碩士論文。
- [3] 鍾國亮 教授，民國 95 年 3 月三版，影像處理與電腦視覺，東華書局。
- [4] 蔡奇男，民國 93 年 6 月，以小波類神經網路實現彩色影像之臉部偵測，朝陽科技大學資訊工程研究所碩士論文。
- [5] 何俊德，民國 93 年 6 月，基於影像與文字之網頁內容方法分類之研究，朝陽科技大學資訊管理研究所碩士論文。
- [6] 林煌山，民國 92 年 6 月，利用膚色及區域極小值做人臉特徵是否遮蔽之偵測判斷，國立中央大學資訊工程研究所碩士論文。
- [7] 謝明娟、連珍慧，94 學年度，即時人臉影像偵測系統，逢甲大學通訊工程系。
- [8] 胡冠宇，民國 93 年 7 月，基於膚色之裸體影像偵測之研究，國立成功大學工程科學所碩士論文。

- [9] 鄭凱方，民國 94 年 6 月，人臉可辨識度計算用於監控系統中人臉正面最佳影像判定，國立中央大學資訊工程研究所碩士論文。
- [10] 羅卿如、鄭惠文、謝朝宗，智慧型視訊監控系統之研究與應用，華梵大學資訊管理系。
- [11] 吳明芳¹、魏育誠¹、詹慧珊²，人臉特徵之身分確認的研究，¹崑山科技大學電機系、²台南女子技術學院美容系。
- [12] 吳鴻鈞，民國 95 年 6 月，一個以半徑對稱性來加強人臉偵測可靠度的方法，元智大學資訊工程研究所碩士論文。
- [13] 吳宣諭，民國 95 年 6 月，運用膚色過濾的人臉偵測法，國立台灣科技大學電機工程所碩士論文。
- [14] 陳文淵、古凱達、劉正忠，使用離散餘弦轉換與串長編碼技術之撲克牌影像辨識系統，國立勤益技術學院電子工程系。
- [15] C. Cortes and V. Vapnik, "Support Vector Networks", Machine Learning, (1995), Vol. 20, pp. 273-293.
- [16] C. J. C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition", Data Mining and Knowledge Discovery, (1998), Vol. 2, pp. 121-167.
- [17] 何信瑩、施東和，民國 84 年，複雜背景中人臉偵測快速演算法則，逢甲大學資訊工程系碩士論文。
- [18] 黃泰祥，民國 93 年 6 月，具備人臉追蹤與辨識功能的一個智慧型

- 數位監視系統，中原大學電子工程學系碩士論文。
- [19] 許福特，民國 89 年 6 月，結合語意分類與部份搜尋的影像搜尋系統，逢甲大學資訊工程所碩士論文。
- [20] 林群雄、吳建樺，使用多類支持向量機混合三角偵測做人臉辨識，國立台北大學資訊管理研究所。
- [21] 孫榮男，民國 92 年 6 月，應用主要分量分析與類神經網路於人臉辨識，國立高雄第一科技大學電腦與通訊工程系碩士論文。
- [22] 黃敏峰，民國 92 年 6 月，人臉追蹤法應用於監控系統之研究，國立成功大學電機工程學系碩士論文。
- [23] 陳同孝、陳明江，植基於中文文字區塊數目之中文文件比對影像搜尋引擎，國立台中技術學院資訊管理系。
- [24] 羅竣威¹、謝清祿²，三種主軸成分法於數位影像特徵擷取與判斷之研究，¹ 國立屏東科技大學機械系碩士生、² 生物系統工程系副教授。
- [25] 曾郁展，民國 94 年 6 月，DSP-Based 之即時人臉辨識系統，國立中山大學電機工程系碩士論文。
- [26] 李姿慧，人臉追蹤法應用於監控系統之研究，國立雲林科技大學電機工程系。

- [27] 施宏政¹、葉榮木¹、蔡俊明²，應用模糊理論之人臉膚色光線補償，¹ 國立台灣師範大學機電科技研究所、² 台北市立師範學院數學資訊教育系。
- [28] 陳嘉雄，民國 91 年 7 月，多角度人臉之二階段辨識，國立東華大學資訊工程研究所碩士論文。
- [29] 王健權，民國 91 年 7 月，以數位影像處理搭配動態攝影機做即時人臉追蹤之研究，國立成功大學電機工程學系碩士論文。
- [30] 楊正宏¹、莫雄偉¹、陳昭和¹、羅錦興²，眼控系統使用於摩斯碼之應用，¹ 國立高雄應用科技大學電子工程系、² 國立成功大學電機工程系。
- [31] 蔡博智，民國 91 年 7 月，影像追蹤方法應用在監控系統之研究，中原大學機械工程學系碩士論文。
- [32] 陳同孝、黎廣福、謝鑫棟，植基於邊緣影像處理之影像比對搜尋引擎，國立台中技術學院資訊管理系。
- [33] 黃奕翔、洪志豪、黃崇恩，PR final project face detection，台灣大學資訊工程學系。
- [34] 許家瑋、蕭志豪、王柏文，民國 95 年 12 月，公仔產生系統，崑山科技大學資訊管理系。

