

國立臺灣師範大學機電科技學系
碩士論文

指導教授：葉榮木 博士
蔡俊明 博士

自適性臉部特徵擷取的動態人臉偵測
**Dynamic Face Detection via Adaptive Face Features
Extraction**



研究生：李宗岳 撰

中華民國 95 年 7 月

國立臺灣師範大
學機電科技學系

碩士論文

自適性臉部特徵擷取的人臉偵測系統

李宗岳 撰

中華民國九十五年

書背

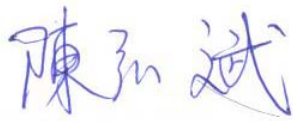
國立台灣師範大學機電科技學系

研究生：李宗岳

碩士論文：自適性臉部特徵擷取的動態人臉偵測

經審查合格，特予證明

論文口試委員



陳弘斌 博士

國立台南護理專科學校共同科數理組助理教授



蔡俊明 博士

國立台北教育大學數學暨資訊教育學系助理
教授 本論文指導教授



葉榮木 博士

國立臺灣師範大學機電科技學系教授
本論文指導教授

主任：



中華民國九十五年七月二十九日

國立臺灣師範大學

博碩士論文授權書

(提供授權人裝釘於紙本論文書名頁之次頁用)

本授權書所授權之論文為授權人在國立臺灣師範大學 機電科技

系所 機電光整合 組 94 學年度第 2 學期取得 碩 士學位之論文。

論文題目：自適性臉部特徵擷取的動態人臉偵測

指導教授：葉榮木 博士 蔡俊明 博士

茲同意將授權人擁有著作權之上列論文全文(含摘要)，授權本校圖書館及國家圖書館，提供讀者基於個人非營利性質之線上檢索、閱覽、下載或列印。

讀者基於非營利性質之線上檢索、閱覽或下載、列印上列論文，應依著作權法相關規定辦理。

指導教授：葉榮木 博士 蔡俊明 博士

授權人：李宗岳

研究生姓名：李宗岳 (請親筆正楷簽名)

中華民國 95 年 08 月 28 日

致謝

碩士班兩年中承蒙指導教授葉榮木博士和蔡俊明博士，在研究領域上或是生活學習上，皆能適時地給予督促與指導。以及論文內容細心校稿、逐字斧正，使得本論文內容更趨完善。另外更要感謝口試委員：陳弘斌博士給予本論文指導與建議，使本論文得以順利完成，在此獻上十二萬分的謝意與敬意。

研究期間蒙實驗室許溥鑫、黃津超、沈世評學長對於學業上的指導。此外，柏智、則謙、誠鈺同學，勝均、誌仁、偉力，智仰學弟，在學業上的討論及生活上的幫忙與陪伴，使我在這二年間留下美好及深刻的回憶。並在此特別感謝Delphi K.Top 程式討論區提供我在程式寫作的學習環境。

最後要感謝我的家人，李吉和先生和余青蓉女士和佳倩姐姐宗杰弟弟，一路上在背後默默的關心、支持與鼓勵我，讓我可以順利完成碩士班的學習。僅將此論文獻給曾經陪伴過我的每一個人，願他們都能順心平安。

摘要

人臉偵測最近幾年來迅速地受到大家的重視，並且廣泛地應用在各種領域，諸如『人臉辨識』、『視訊會議』、『數位監控系統』、『智慧家庭系統』、『智慧型人機介面』等等。此類的應用程式之首要工作即是找出是否有人臉存在，並找出其所在之區域。換句話說，一個可靠且快速的人臉偵測法是相當重要的。

本篇文章提出一個簡單且有效的方法來偵測人臉的位置，由於複雜的背景嚴重影響到人臉偵測的成果。我們首先使用膚色分割和連通成分的方法找出人臉候選區、再使用色彩分析的方法從人臉候選區中去尋找臉部特徵(眼睛、嘴唇)。使用色彩的資訊可以減少檢測時間，不過在不同的光線下會有誤判的情況產生，所以我們用自適性的閾值來找出臉部特徵，針對不同的影像用不同閾值去做判斷，最後利用眼睛和嘴唇的幾何關係找到正確的人臉位置。

我們的實驗證明本論文所提出的方法確實可行且可靠，在少許的條件限制下，對於環境採光變化、不同大小之人臉、及複雜背景等問題都可以有效的處理。除了在臉部特徵被遮蔽的情況下會有偵測不到的狀況，其他情況均有良好的偵測率。

關鍵詞：人臉偵測、色彩空間、膚色、臉部特徵。

Abstract

The face detection receives attention rapidly in recent years, and widely apply to various kinds of fields, such as 『face identity』, 『Video Conference』, 『surveillance systems』, 『intelligent family system』, 『intelligent human-machine interface』 ...etc. The primary work of the application program is to determine face existence, and to locate face position. In other words, it is quite important that a certain and fast face detection method.

This paper describes a simple and effective method to detect position of faces, because the face detection effect was affected seriously by the complex background. First, we use the skin color segmentation and connected component method to find out the face candidate and then use color-analysis to search the features of faces (eyes, lips) from face candidates. Using color information to detect faces can save time, but there will cause wrong detection while under different light situation, so we use an adaptive threshold value to find out the face features, and use different threshold values to determine face features of different images, the geometry relation between eyes and lips is applied to locate correct face at last.

Our experiment results prove that our method is feasible and effective. While a little conditions are limited, we can detect the face effectively in difference light conditions, size of faces and complex background. All situations have good detection rate except the face feature is covered.

Key words: face detection; color space; skin color; face feature

目錄

摘要.....	I
Abstract.....	II
目錄.....	III
圖目錄.....	V
表目錄.....	VII
第一章 緒論.....	1
1-1 前言.....	1
1-2 研究動機與目的.....	3
1-3 系統架構.....	5
1-5 論文架構.....	6
第二章 人臉偵測的相關研究.....	7
2-1 色彩空間.....	7
2-2 相關研究.....	11
第三章 自適性閾值的臉部特徵偵測.....	19
3-1 膚色分割.....	21
3-2 去除雜訊.....	25
3-3 取得連通成份與人臉候選區.....	29
3-4 眼睛和嘴唇的偵測.....	34

3-5 幾何判斷.....	43
第四章 實驗與分析.....	45
第五章 結論.....	55
參考文獻.....	57

圖目錄

圖 1.1 人臉偵測的流程.....	5
圖 2.1 光三原色.....	7
圖 2.2 HSV 的色彩空間.....	9
圖 2.3 YCbCr 的色彩空間.....	10
圖 2.4 人臉偵測的分類.....	11
圖 3.1 人臉偵測的流程圖.....	20
圖 3.2 膚色群聚的範圍在不同的亮度.....	22
圖 3.3 膚色分割.....	24
圖 3.4 中值濾波器.....	25
圖 3.5 膨脹侵蝕的結果.....	27
圖 3.6 經過中值濾波器和型態學運算的結果.....	28
圖 3.7 區域填充.....	30
圖 3.8 區域填充的結果.....	31
圖 3.9 連通成分.....	32
圖 3.10 人臉候選區.....	32
圖 3.11 連通成分的結果.....	33
圖 3.12 偵測臉部特徵的範圍.....	35
圖 3.13 黑色增強的結果.....	36

圖 3.14 不同亮度下的眼睛候選區.....	37
圖 3.15 自適性眼睛偵測流程圖.....	38
圖 3.16 眼睛偵測的結果.....	39
圖 3.17 不同亮度下的嘴唇候選區.....	40
圖 3.18 嘴唇偵測的流程圖.....	41
圖 3.19 嘴唇偵測的結果.....	42
圖 3.20 臉部特徵的幾何距離.....	44
圖 4.1 由近拍到遠的實驗結果.....	47
圖 4.2 由遠拍到近的實驗結果.....	47
圖 4.3 較亮情況的實驗結果.....	49
圖 4.4 較暗情況的實驗結果.....	50
圖 4.5 頭由左轉到右的實驗結果.....	52
圖 4.6 帶眼鏡的實驗結果.....	52
圖 4.7 多人的實驗結果.....	53

表目錄

表 2.1 相關文獻綜合比較表.....	16
表 3.1 不同亮度值下膚色群聚範圍表.....	23
表 3.2 眼睛和嘴唇的範圍座標表.....	34
表 3.3 不同亮度下唇色的 RGB 值.....	39
表 4.1 實驗中所使用的硬體規格.....	45
表 4.2 不同大小人臉的實驗結果.....	46
表 4.3 不同亮度情況的偵測結果.....	48
表 4.4 其他其況下的人臉偵測率.....	51
表 4.5 不同系統的比較.....	54

第一章 緒論

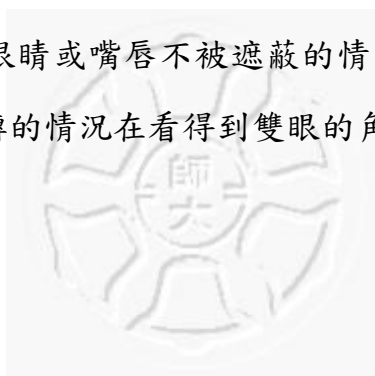
1-1 前言

人臉偵測一直以來都是電腦視覺的重要課題之一，人臉偵測的技術，可應用在很多的方面，如警察局檔案中的通緝犯之照片比對用、監視系統、人臉辨識及人臉影像資料庫的管理上。為了維護公眾安全都會在公共場所安裝監視攝影機，不過這些監視攝影機通常都只有攝影卻沒有錄影的功能，另一種是有錄影功能卻只能在之後拿來當作指認兇手的工具，沒有辦法達到即時警告的作用。對於入侵者的人臉偵測就是相當重要的。所以如何有效且正確的人臉偵測就是現在許多人研究的目標。

另外在影像處理技術上，為了獲得人臉在影像中的位置則需搭配人臉定位技術，關於臉部定位技術的研究中雖然已經有許多的方法被提出[1]。然而在影像中偵測人臉的位置是相當困難的，因為很難預知影像中人臉的大小以及數量，再加上背景複雜和光線的變化會影響到偵測的結果，因為膚色在彩度空間的分佈具有相當高的一致性，故而大多數早期的作法，都是以定義膚色分佈為基礎。但是不論是採用那一種顏色空間，如果一開始定義了較大的偵測膚色範圍，那就會遇到很多誤判的背景顏色與人臉結合在一起，這將導致人臉的無法判別。但如果一開始定義了較小的膚色偵測範圍，可能會導致有些人臉膚色的偵測失敗，未能測出人臉所在。所以在背景顏色與膚色相似，或是背景顏色較複雜的環境下，以這樣的技術來偵測人臉所在，會受到很多的限制。在人臉偵測的問題上，除了膚色特徵以外，還需要其它特徵的配合使用，以提昇人臉偵測的可靠性。而大多數的人臉偵測技術，都先定義人臉膚色的分佈範圍，再從影像中篩選出這些色點，然後用形態學(morphology)上的擴張原理，將這些不連續的色點連接起來，再判別其外形是否相似橢圓，然後再運用人臉特徵的辨識技術，來確定是否是人臉的位

置。類似這樣的技術，在使用上容易受到限制，不能在較複雜的背景環境中使用。故而這樣的方法常應用在比較限制的情況下。例如要限制在單一且固定位置的背景顏色，人臉要在影像的中央位置，人臉正面要朝向正前方，不能戴有類似圍巾或是頭巾等物品。因為這些條件都會破壞人臉偵測的形狀，致使無法偵測到人臉的所在位置。

為了達到快速偵測人臉，所以本文使用膚色在色彩空間中會群聚在一起的特性去切割出膚色的位置，為避免類似的膚色的背景以及人的手腳被當作人臉，所以再從這些位置中去尋找只屬於臉部的特徵例如：眼睛、嘴唇…，最後再利用兩眼和嘴唇幾何的關係判斷是否為真的人臉，這樣對於在複雜的背景之下也可以判斷是否為人臉，因為不是使用橢圓偵測所以對於戴圍巾或是頭巾的實驗者只要在眼睛或嘴唇不被遮蔽的情況下都可以達到良好的偵測，同樣的對於頭部旋轉的情況在看得到雙眼的角度之下都可以偵測。



1-2 研究動機與目的

隨著社會經濟的蓬勃發展與犯罪率的節節高昇，個人安全的保障越來越受到重視。過去常用的方式，是使用名牌、磁條卡、或密碼來辨識身分，不過這些方法，總難以避免卡片遺失、被盜用、或忘記密碼等情形，所以最好的方式，是依據每個人特有的生理特徵，如人臉、指紋、掌紋、血管、虹膜、語音等；或是行為模式，如簽名、步伐、敲打鍵盤(keystrokes)。這些以人類生理特徵或行為模式為辨識依據的方法，統稱為生物認證系統(Biometrics)，已成為具有廣大市場需求的工作。

在這麼多種特徵中，又以人臉最具有獨特與方便取得的特性。對人類來說，看臉認人本來就是相當直覺且容易的方式，對電腦來說，只要安裝簡易的CCD 或數位像機，就可以讓使用者在很自然的情況下擷取臉部影像來進行辨識。此技術可輕易整合在自動提款機、個人電腦、個人數位助理及IA產品，廣泛應用於網路電子商務、住戶門禁安全、信用卡市場、數位監控、犯罪偵查、及電腦遊戲的人機介面上。

人臉偵測是許多生物認證系統的前置作業，它是利用圖形識別的方法自動偵測出影像中每個人臉部的位置，以作為後續移動追蹤、行為分析、或身分確認的應用。其演算法會直接影響整個系統的執行效能，因此必須具備足夠的彈性與穩定性。已有很多關於人臉偵測的研究報告，主要都是以人臉膚色分割為主，以此來找出人臉的位置所在，然而使用膚色的方法常遇到問題是容易受到光線的影響以及膚色的背景的影響。所以本篇提出可以改善膚色分割缺點，其研究重點可分為兩大部分： 1.膚色偵測技術的改善 2.人臉特徵的偵測和擷取的方法。

1. 膚色偵測技術的改善：

我們的方法之中我們有統計不同亮度膚色的Cb、Cr範圍，我們可以根據這些pixels的亮度大小給予不同的範圍，即使在光源變化的情況之下我們也

可以有效的偵測膚色，接下來因為亮度的不同臉部特徵的色彩資訊也會改變，所以我們加入自適性的閾值來分割出臉部特徵(眼睛、嘴唇)，每張不同的影像都會對眼睛和嘴唇分別給予一個適當的閾值，使我們的方法在非極端的光源下也可以進行偵測。(這裡的極端是指亮度小於50和大於200，因為在極端的亮度中膚色呈現黑色或白色以不屬於皮膚的顏色。)

2.人臉特徵的偵測和擷取的方法：

就是在膚色的分佈的範圍內，研究如何找出正確合適的眼睛及嘴的候選區，有別於之前使用固定的閾值，根據亮度的不同給予不同的閾值，然後再根據眼睛和嘴的中心點，所構成的幾何關係，找出最合適的眼睛和嘴的配對。以這樣的方式來找出真正的人臉並取出。



1-3 系統架構

本論文系統分為兩大部份，整個系統概觀如圖1.1所示。第一部份是使用兩種不同的輸入管道對人臉膚色範圍的定位。首先離線統計人臉膚色在不同亮度範圍的色彩空間所存在的Cb、Cr範圍，然後當系統接收到輸入的影像，便會根據色彩資訊來轉換並且切割出可能有人臉存在的區域，並且運用一些形態學演算法和中值濾波器加以去除雜訊。本系統第二部分是使用色彩的資訊來分割出臉部特徵(眼睛、嘴巴)，我們提出自適性的閾值來分割出眼睛和嘴巴，因為使用固定的閾值來做判斷並不適用每張輸入的影像，要考慮每張影像的明暗程度的不同給予不同的閾值。我們是根據臉部特徵占人臉面積的比例來判定閾值，第三章會有詳細的說明。最後我們偵測出來的臉部特徵候選區，再依照臉部五官的幾何關係去定義出正確的眼睛、嘴巴，進而定義出正確的人臉位置。

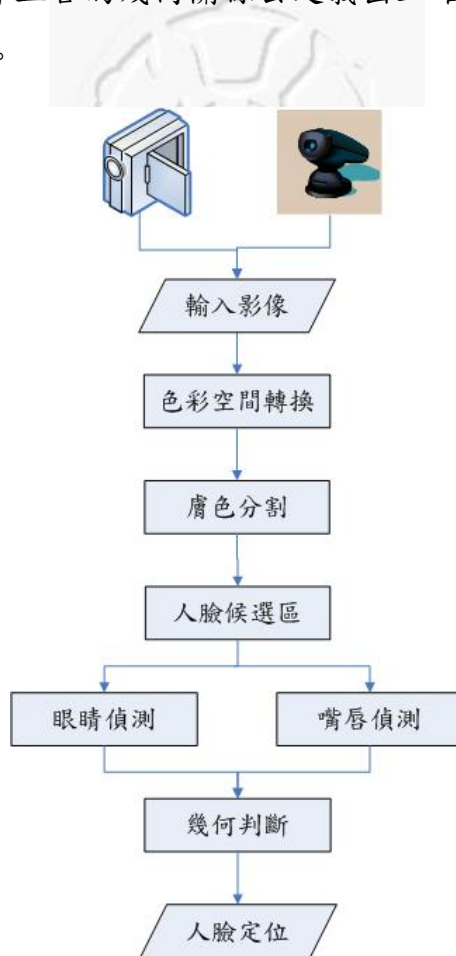


圖1.1 人臉偵測的流程

1-4 論文架構

本論文其他章節的分配如下：第二章是介紹人臉偵測的相關研究，闡述一些人臉偵測常遇到的問題，再簡單的介紹我們常見的色彩空間以及之前人臉偵測的文獻。第三章則詳細說明我們所使用的發方法介紹我們偵測人臉候選區所使用的型態學運算(morphology operations)、中值濾波器(median filter)和連通成分(connected components)，先找出人臉候選區，再使用自適性閾值的臉部特徵偵測，找出眼睛和嘴巴，最後使用幾何判斷的方式定義出正確的臉部特徵進而定義出正確的人臉。第四章呈現出我們的實驗結果，我們不同的情況做了幾個不同的實驗，然後對錯誤的實驗結果加以解釋。第五章會對整個論文內容歸納出一些結論與對未來研究之建議。



第二章 人臉偵測的相關研究

2-1 色彩空間

要找出膚色區域必須先定義什麼是皮膚顏色，首先要選擇一種合適的色彩空間，這樣才能用色彩參數將皮膚顏色具體化描述。在過去的許多研究中，曾經用過的色彩空間有RGB、HSV、YCbCr、XYZ、YIQ、與YES等，由於環境光線變化對膚色的影響很大，通常會選擇能夠從色彩中分離出亮度因素的色彩空間，其中最常見的是RGB、HSV與YCbCr色彩空間，分別介紹如下：

2-1-1 正規化RGB(normalized RGB)

人眼感測器只能檢測到波長介於400nm至700nm之間的可見光，這些感測器稱為圓錐細胞，具有三種不同的類型：紅光圓錐細胞、綠光圓錐細胞和藍光圓錐細胞。如果單波長的光可見，這三種感測器的相對反應能使我們鑒別出光的顏色。該現象極具實用價值，因為這意味著我們只需簡單地按不同比例將上述三種光混合，就能產生各種顏色的光。這就是著名的三色原理，它在彩電系統中獲得了廣泛的應用。

RGB為最普遍使用的色彩空間之一。其中R代表紅色，G及B則各為綠色與藍色。此系統和我們熟知的(色)光三原色相似，愈混合後會得到愈明亮的色光。如圖2.1。

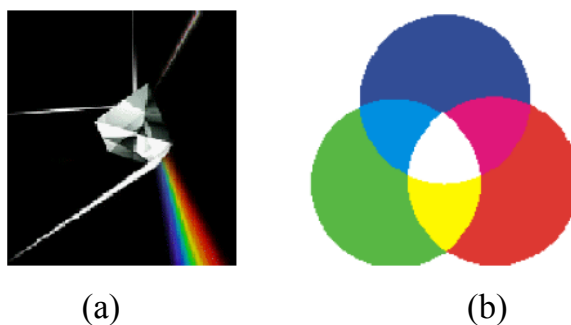


圖2.1 光三原色(a)由白光所分出之三原色光; (b)由三原色所混合出之顏色。

(資料來源：COULEUR.ORG)

在RGB空間裡表示的影像，會因光的亮度不同而使本來實物相同顏色的地方在影像中卻變得不同，所以不適合直接以RGB做顏色分割，但只要把R，G予以正規化(normalized)，就可使R，G對光的相關性減少，公式如下式(2.1)：

$$R_n = \frac{R}{R+G+B}, G_n = \frac{G}{R+B+G}, \quad (2.1)$$

在這裡只要對 R_n 跟 G_n 作適當的閾值選取，就可以把膚色跟影像中的顏色分開，由於大部份的電腦擷取系統都是以RGB作為基本顏色的表示方法，而且正規化RGB的計算簡單快速及顏色對光的依賴性減少。

2-1-2 HSV色彩空間

HSV色彩空間對人類來說較容易理解，它將亮度value(V)與彩度chrominance從色彩中分離開來，彩度又分為色相hue(H)與飽和度saturation(S)，由於元素之間關聯性小，不會互相影響，所以非常適合影像處理的使用。HSV色彩空間與RGB色彩空間的轉換關係式(2.2)：

$$H = \cos^{-1} \left\{ \frac{\frac{1}{2} [(R-G) + (R-B)]}{\left[(R-G)^2 + (R-B)(G-B) \right]^{\frac{1}{2}}} \right\},$$

$$S = 1 - \frac{3}{R+G+B} [\min(R, G, B)],$$

$$V = \frac{\max(R, G, B)}{255}, \quad (2.2)$$

以軟體的方法在較慢的電腦下把RGB轉成HSV較耗時而不適合即時(real-time)的訴求，但若硬體系統上本身就有HSV的轉換或不是即時的場合，HSV是最多人臉偵測系統用到的顏色空間，因為在HSV做顏色分割會有較好的效果。

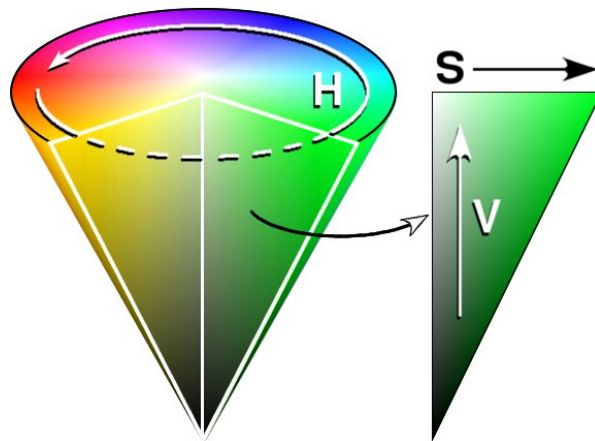


圖2.2 HSV的色彩空間(資料來源：COULEUR.ORG)

2-1-3 YCbCr色彩空間

利用RGB格式處理圖像時，每個圖素由3個分別對應於三原色的8位元或10位元字確定，因而這不是一種最高效的處理方法。這種格式下，必須在所有的紅色、綠色和藍色通道上對圖素進行作業，所需的儲存空間和數據頻寬毫無疑問將比其它可供選擇的色彩格式更大。為了解決這個問題，許多廣播標準(如歐洲的PAL和北美的NTSC電視系統)均採用亮度和色度視頻信號。因此，不同的色彩格式之間需要一種能互相轉換的機制。最通用的一種轉換是由RGB格式轉換至YCbCr格式(反之則從YCbCr格式轉換至RGB格式)。其中Y表示亮度(而Cb與Cr代表另外二個色度(chrominance))。研究顯示，人眼察覺到的光亮度資訊(Y)的60%至70%來自綠色光。紅色和藍色通道實際上只是亮度資訊的複制，因此這些重覆資訊完全可以去除掉。最終的結果是圖像可用表徵色度和亮度的信號來表示。在這種格式下，8位元系統規定亮度的取值範圍介於16至235之間，而Cb和Cr信號的取值範圍介於16至240之間，並規定128表示亮度為0或255。

YCbCr色彩空間是修改YUV色彩空間所得到的，分為亮度元素luminance(Y)與blueness(Cb)、redness(Cr)兩個彩度元素，對亮度的分離性高，

方便與彩度分開操作，適合影像處理使用。它也是REC601-1(CCIR 601-1)所制訂MPEG視訊與JPEG影像的標準色彩空間。YCbCr色彩空間與RGB色彩空間的轉換關係式(2.3)：

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.29900 & 0.58700 & 0.11400 \\ -0.16874 & -0.33126 & 0.50000 \\ 0.50000 & -0.41869 & -0.08131 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \quad (2.3)$$

顏色分割的目的就是要把臉部跟背景分辨出來，又或是從臉部判斷眼睛，嘴唇等位置。但在複雜的背景中，有時也會出現其他人或物跟目標的顏色相同的現象，所以在複雜的背景中單靠顏色分割作判斷有時候也有可能造成誤判。所以一般顏色分割後的影像會跟其它資訊結合運用。本論文之膚色色彩系統是使用YCbCr色彩空間。

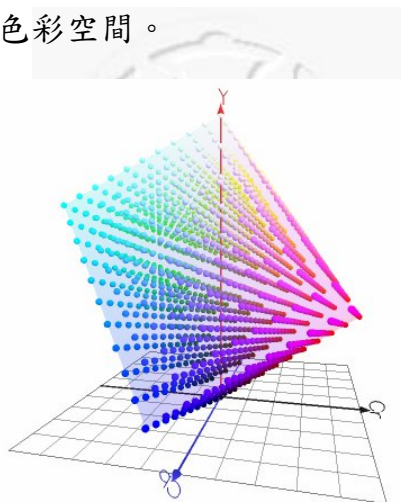


圖2.3 YCbCr的色彩空間(資料來源：COULEUR.ORG)

2-2 相關研究

人臉偵測的研究已經風行了多年，從一開始的灰階影像到現在的彩色影像，人臉偵測即是判斷影像中是否有人臉圖案如果有的話，則標出人臉的位置、大小與數目等資料。此問題長久以來一直都是一個困難且熱門的研究課題，有相當多的研究人員投入這個領域當中，也獲得了相當豐碩的成果。使用的方法有許多種，也有結合多種方法的人臉偵測，大致上可以分為三大類如圖2.4所示：

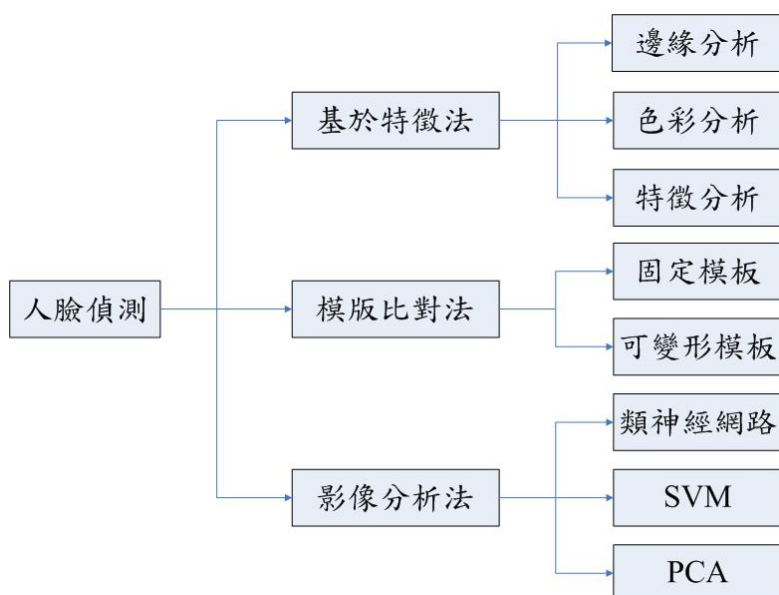


圖 2.4 人臉偵測的分類

2-2-1 基於特徵法 (feature-based methods)

此類方法是根據人臉結構上之特徵，諸如眼睛、嘴巴、鼻子、眉毛等等，先將其定位出來，再利用統計或是上述特徵與人臉結構上之比例來定出人臉的位置。此類的做法容易受到外在因素的影響，燈光的強度以及顏色，背景的複雜度，都會影響到偵測出來的結果。很多研究人員會針對膚色下手，先將顏色空間轉換，再建立膚色模型用以偵測人臉區域，最後結合人臉特徵來確認人臉區域是否為人臉。

例如邊緣分析[2,3,4]，尋找影像中物件的邊緣，利用人臉的邊緣屬於圓形或是橢圓來定位人臉，不過在複雜背景之下會有誤判的產生。Nara Yusuke 和 Jianming Yang[3]使用邊緣偵測和膚色偵測兩個模型來偵測人臉，在一個複雜背景的影像中現做sobel的邊緣偵測，再對同一張圖做膚色偵測尋找有膚色又有橢圓邊緣的地方，定位為人臉所在，不過這方法在光線的影響下雖然可以偵測出人臉邊緣，不過無法偵測正確的人臉膚色範圍。Stan Brichfield [4]提出一種利用橢圓為基礎的臉部輪廓搜尋模型，並且檢測橢圓內的色彩分佈來追蹤人臉。利用此兩種模型的交互使用，能在複雜的環境中達到穩健的追蹤效果。此系統對於物體的轉動、暫時遮蔽、接近與離開皆有穩定的追蹤效果。不過橢圓偵測的缺點在於容易受到光源及背景影響，並且需要大量的運算量而難以達到即時性人臉追蹤。

色彩分析[6-10]，使用色彩的資訊統計膚色的範圍達到人臉定位，不過色彩資訊對於光線很敏感在過暗或過亮的狀況下會影響偵測率。Rein-Lien Hus[7]提出使用光線補償後再統計出膚色範圍建立出膚色模型，使用膚色模型找出人臉範圍，再使用型態學運算找出臉特徵眼睛和嘴唇，最後使用霍氏轉換找出人臉的橢圓輪廓。這方法是目前人臉偵測在色彩分析中引用次數較多的方法，不過此方法不適用於即時的人臉偵測。S.H. Kim and H.G. Kim[10]利用動態偵測及膚色以實現人臉偵測。首先，利用兩張連續影像相減，並且設定一閾值，當差值大於此閾值的話則為1，反之則為0。得到差值影像之後，再利用膚色特徵在群集（Grouping）的範圍裡面判斷此區域是否為人臉。雖然此方法結合了動態及膚色可以快速的找出人臉的位置，不過作者是在理想的光源之下偵測膚色，如果膚色受到光源的影響而造成色偏，則此方法恐怕會失敗。為了得到差值影像，所以此系統必須使用固定式攝影機對著不變的背景，因此，此方法比較適用於特定場合，當有動態的物體，例如人體，出現在影像當中即可將移動的物體和背景分離。

特徵分析[11-15]，此類方法試圖找出人臉影像結構上的一些特徵來偵測人臉區域。此類方法主要用在找出人臉所在的區域。這方法多侷限在正面的人臉，在側面以及臉部特徵遮蔽的狀況效果不盡理想。Carlos[14]提出利用人的瞳孔做為人臉追蹤時的特徵。他們使用主動式紅外線（Infra-ray）照射特定的區域，當此區域出現人的臉時，人眼的瞳孔受到紅外線的照射而產生反射，所以當攝影機照到人臉時，瞳孔相對之下會變亮許多，他們也就是利用此特性並且加上啟發式規則來偵測及追蹤人臉。此方法的缺點只能偵測正面的人臉，因為當眼睛瞳孔特徵不存在時，則此系統就無法進行人臉追蹤；另外，當背景光源很亮時此方法恐亦不適用。C. C. Han[15]提出一種利用眼睛來偵測影像點中是否有人臉的方法。眼睛為一穩定的特徵，即使受到光線的影響，眼睛灰階值會小於鄰近的灰階值，利用形態學（Morphology）中開啟（Open）與閉合（Close）運算元的特性，找出圖形中可能為眼睛的區域，再配合臉部特徵的幾何關係與類神經網路完成最後的人臉偵測。此方法可快速偵測人臉位置，但只使用臉部特徵(眼、口、鼻...等)作為判別是否為人臉，容易受到擁有相似特徵的物體影響而產生誤判。

這些方法都不需要事前訓練，都是利用臉部的資訊或是型態學的運算來找出人臉的位置，直接使用膚色和臉部特徵的幾何特性以及人臉的輪廓呈現橢圓形來做分析，因為運算簡單且不需要建立龐大的資料庫所以常常被使用在即時的系統上，雖然運算的速度快不過其可靠度卻也比需要事前訓練的系統差。本文人臉偵測也是屬於此類的。

2-2-2 模板比對法 (template matching methods)

模板比對[16-19]，此類方法用一些事先定義好代表人臉的patterns或人臉各個器官的patterns與輸入影像間的關係來偵測人臉區域。此類方法可以找出人臉區域並且判斷此區域是否為人臉，這方法對於戴眼鏡或是落腮鬚的人不

適用，且針對不同角度的人臉要建立不同人臉模板。Sakai et al. [17] 使用眼睛、鼻子、嘴巴與臉的輪廓的 patterns 來描述人臉，每個 pattern 是由一條條的線條區段所組成。將灰階影像轉成由線條所組成的圖形，然後找出與 patterns 相似的區域。Young Ho Kwon[18]建立了五個不同的人臉模版分別是人臉由左到右的五個角度，在使用Fuzzy的方法建立膚色和髮色的模型，用這兩個模型先尋找膚色和髮色的所在位置，找到之後再跟先前建立的五個人臉模版做Fuzzy的比對，這方法對於光頭和有戴帽子或是頭巾的人無法偵測，且髮色的定義是針對哪種髮色也未加說明，這方法對於太小的人臉也會有偵測上的困難。

此類方法必須先事先定義好人臉 (face patter) 做成模板，而這些人臉大部分皆是正面，藉由這些模板去對整張灰階影像做全搜尋，以判斷被掃描的區域是否為人臉。此類方法的優點就是較容易實作。不過對於非正面的人臉用模板比對的方法可行性就降低許多，而且對於整張影做搜尋對於大張的影像就會降低偵測速度，且影響中的人臉大小不固定也會對偵測造成影響。除了使用事先定義好的template外，也有許多研究人員利用大小與外型 (deformable template) 可以變動模型，成功地偵測人臉。

2-2-3 影像分析法 (image-based methods)

與模板比對法不同的是影像分析法藉由統計分析 (statistical analysis) 與機器學習 (machine learning) 來找出人臉影像有意義的樣式，然後藉由不同大小的視窗對整個灰階影像作地毯式的掃描一一與人臉樣式比對，以判斷被視窗掃描到的區域是否為人臉。由於此類方法須以各種大小的視窗掃描影像以便偵測不同大小、位置的人臉，因此計算量相當的龐大。

例如類神經網路[20,21]，是將模式的統計特性隱含在ANN的結構與參數之中，其方法是基於經驗風險最小化的原理，它提供一個方便的工具來做影

像分類和人臉偵測，不過需要一個龐大的訓練資料庫來做訓練。H. A. Rowley[20] 作者提出一個以neural network為主的方法偵測正面的人臉。連結神經網路檢查一張影像中的每一個小方框和決定是否每一個小方框包含一張人臉。這個系統可以視察多個網路之間可以改善單一網路的性能。作者用一連串正面的人臉做training。為了提高偵測率作者收集反面事例，並使用引導算法，把錯誤的人臉偵測增加進訓練裝置作為訓練發展。使用臉很少在圖像內重疊的事實，能更進一步改進準確。

SVM[22,23]，是基於結構風險最小化的原理之統計學學習理論，常用於分類與回歸的問題，SVM使VC(Vapnik Chervonenkis)維數的上限最小化，也是用分類的方式來偵測人臉。Hyungkeun Jee[22] 在這篇文章裡作者使用顏色，邊緣的資訊偵測人臉後選區，先使用sobel偵測邊緣分離出前景背景，在對分離出對位置使用YCbCr的色彩空間分割出膚色範圍，然後使用邊緣和二值化的資訊來訊找眼睛對。最後作者在使用SVM驗證是否為正確的人臉。

PCA[24]，將影像看成一個有 $n \times n$ element的vector，該影像可以看成落於 $n \times n$ 空間中的一點。對一群人臉的影像中，我們做PCA，可以得到一個影像中最重要的元件而這些最重要的元件，它們組合成的空間稱為principal space（人臉稱為eigenface）所有的人臉影像投影在這個principal space都會相當接近對於一張影像，要判斷它是否為一張人臉，只需將它投影在principal space，然後看他和人臉影像所座落位置的距離，如果此影像是人臉的話那麼他的距離會非常小，於是我們可以將影像投影至principal space看它和其它人臉投影的距離來判斷它是否也為一張人臉。Sung 與 Poggio [25]利用兩個距離評量去測度出輸入影像與群聚的距離，他們利用改良式k-mean 的分類方法，去訓練出包括 6 類人臉及 6類非人臉等共計 12 類的族群，再將含有 12 個值的特徵向量輸入至multi-layer perceptron (MPL) 做確認的工作。

這些方法不需要透過特徵分析，不過都需建立資料庫用來作人臉的比對，雖然偵測的效果較強都需要經過複雜的運算，在時間上就大打則扣了。可以將人臉區域的圖像視為一高維度線性空間集合，於是可以透過統計或者類神經網路來判斷某一子視窗是否為人臉。利用此方法必須搜尋大量的人臉及非人臉樣本以訓練此系統，並且準確性會受到樣本的影響。在下醫張表中我們將簡單描述各方法之間的優缺點。

表 2.1 相關文獻綜合比較表

類別	方法	優點	缺點
基於特徵法	邊緣分析[3]	1.處理速度快 2.可用於非正面人臉 3.不受眼睛、鬍鬚的影響	1.不適用於複雜背景 2.遮蔽破壞輪廓 3.易受光線影響
	色彩分析[7]	1.有光線補償 2.可偵測不同角度 3.有偵測人臉特徵	1.難以實現即時系統 2.對於人臉大小有限制
	特徵分析 [15]	1.處理速度快 2.有偵測人臉特徵	1.易受擁有相同特徵物體影響 2.特徵遺失問題
模板比對法	固定模板 [18]	1.容易實做 1.可偵測比較小的人臉 2.可針對五個角度	1.搜尋整個影像 2.對於頭上有戴帽子或頭巾無法偵測
	可變形模板 [19]	1.容易實做	1.對於形狀以及姿勢不一處理 2.只能偵測正面人臉

影像分析法	類神經網路 [19]	1. 通常人臉的位置一定可以偵測到 2. 可搜集反面資料提高偵測率	1. 偵測的好壞與 Training Set 的好壞非常相關 2. Network 的參數很難調整到最佳狀況 3. 不利於即時系統
	SVM[21]	1. 只適用於正面人臉 2. 有偵測臉部特徵	1. 需要事前訓練 2. 輸入影像和訓練影像不可差異過大
	PCA[23]	1. 不需預先進行人臉偏轉角度的估計	1. 需要事前訓練 2. 易受外界環境、光線等因素影響

2-2-4 綜合討論

在即時的人臉偵測的應用當中，需要有一具有強健性及即時性的追蹤系統可以幫助臉部特徵的擷取，為了可以在偵測到人臉以後可以做接下來的處理，例如人臉辨識、表情及情緒的分析與唇形的辨識等等。我們可以將即時的人臉偵測分成以下幾個方法並且討論其可行性。

1. 動態偵測

利用兩張連續影像相減以偵測出影像當中移動的物體。此方法雖然可以快速找出移動的物體，不過容易受到身體及其他物體的影響其缺點在於攝影機必須是固定不動的。而且對於在畫面中靜止不動的人臉就無法判斷，對於智慧家庭系統使用人臉辨識當做依據通常是只有輕微的移動，此方法只適用於監視系統。

2. 臉部模型法

利用此方法的研究者認為臉部3D影像的資訊不會受到光源的影響，所以每一個輸入的影像即被投影到另一個人臉模型以進行後續的處理。此方法

的缺點在於有許多模型參數必須設定，例如大小、正規化、旋轉、變形及光源之變化，不僅要很多的運算量，而且也很難將現實所有的情況考慮進去。

3. 外形搜尋法

利用橢圓為基礎來搜尋影像當中符合此條件的物體，利用此方法的研究是根據頭部的形狀輪廓為橢圓形。此方法對於物體的轉動、暫時遮蔽、接近與離開皆有穩定的追蹤效果。不過橢圓偵測容易受而光源及複雜背景影響而造成人臉追蹤失敗，而且也需要大量的運算量而難以達到即時人臉偵測。

4. 色彩搜尋法

在相關的研究領域中有許多研究都認為，利用膚色可以將人臉與背景有效地分離出來，且在解決臉部旋轉及遮蔽性有一定的效果。但是色彩容易受光源的影響，造成人臉偏暗或偏亮，進而影響人臉偵測的效果。雖然此派研究者提出多種色彩轉換空間，例如正規化RGB、HSV及YCrCb等等，要將亮度及色度分離開來以減少亮度對於色度的影響，不過解決的程度有限。以上為人臉追蹤常使用的方法。另外，尚有利用紅外線偵測人臉的瞳孔等等，不過都必須在特定的情況之下才可以成立，無法在實際的情況當中實現。

我們主要所用的方法是色彩搜尋法，使用 YCbCr 的色彩空間，為了避免光線的影響所以統計不同亮度的 Cb、Cr 範圍。接著我們使用色彩的資訊尋找臉部特徵，不過是因為使用色彩資訊所以一樣對於光線的效果過於敏感，光線的暗或亮都會對於臉部特徵的偵測造成影響，所以我們使用自適性的閾值來偵測臉部特徵，對於臉部的明暗程度不同使用不同的閾值來偵測，這樣就可以減少光線的影響。接著對偵測出來的特徵點加一些條件限制使其可以找到正確的人臉特徵位置。詳細方法將在下一章提出。

第三章 自適性閾值的臉部特徵偵測

在本文的方法中主要分成三個部份：人臉候選區的尋找、候選區中眼睛以及嘴巴的定位、最後利用眼睛嘴唇的幾何關係去定位出正確的人臉。流程如圖 3.1 所示。

由於人的膚色分佈在彩色空間上具有一致性、區域性的關係，所以有很多文章在彩色空間上，提出了膚色分佈的數據，有的是用 Y、Cb、Cr 來定義膚色的分佈[6,7]，是較常應用在硬體導向的彩色模型上使用，如數位照相機。有的是用 H、S、V 來定義膚色的分佈[8]，其目的是強調此 H、S 較能抵抗因外界光線照射的干擾，且較容易讓人類視覺接受的彩色模型。有的是用 R、G、B 來定義膚色的分佈[9,10]，目的是要加快運算速度，不需要做彩色空間的轉換計算。以事先訓練好的人臉膚色的分佈範圍，再從影像中篩選出這些屬於膚色的像點，再用膨脹原理將這些不連續的膚色像點連接起來。最後再用長方形或橢圓形人臉近似的方法，來判斷是否為人臉的位置所在。這種方法容易受不同的光線照射、複雜的背景及環境顏色影響。

到目前為止，膚色偵測在不同的應用中，都會碰到一些不同的問題。同時對於何種色彩座標較佳，似乎也沒有一致的看法。我們歸納膚色對於影像形成的條件，有極大的關係。所以只用膚色特徵並不是最好的辦法。雖然其優點是快速且容易實現人臉偵測的方法，但缺點是容易受外界環境、光線照射及人臉上的表情等因素影響。

本文是使用 YCbCr 的色彩空間搭配上形態學運算和幾何關係達到人臉和人臉特徵的偵測，在 YCbCr 色彩空間中 Y 是指亮度、Cb 是指藍色的成分、Cr 是指紅色的成分。

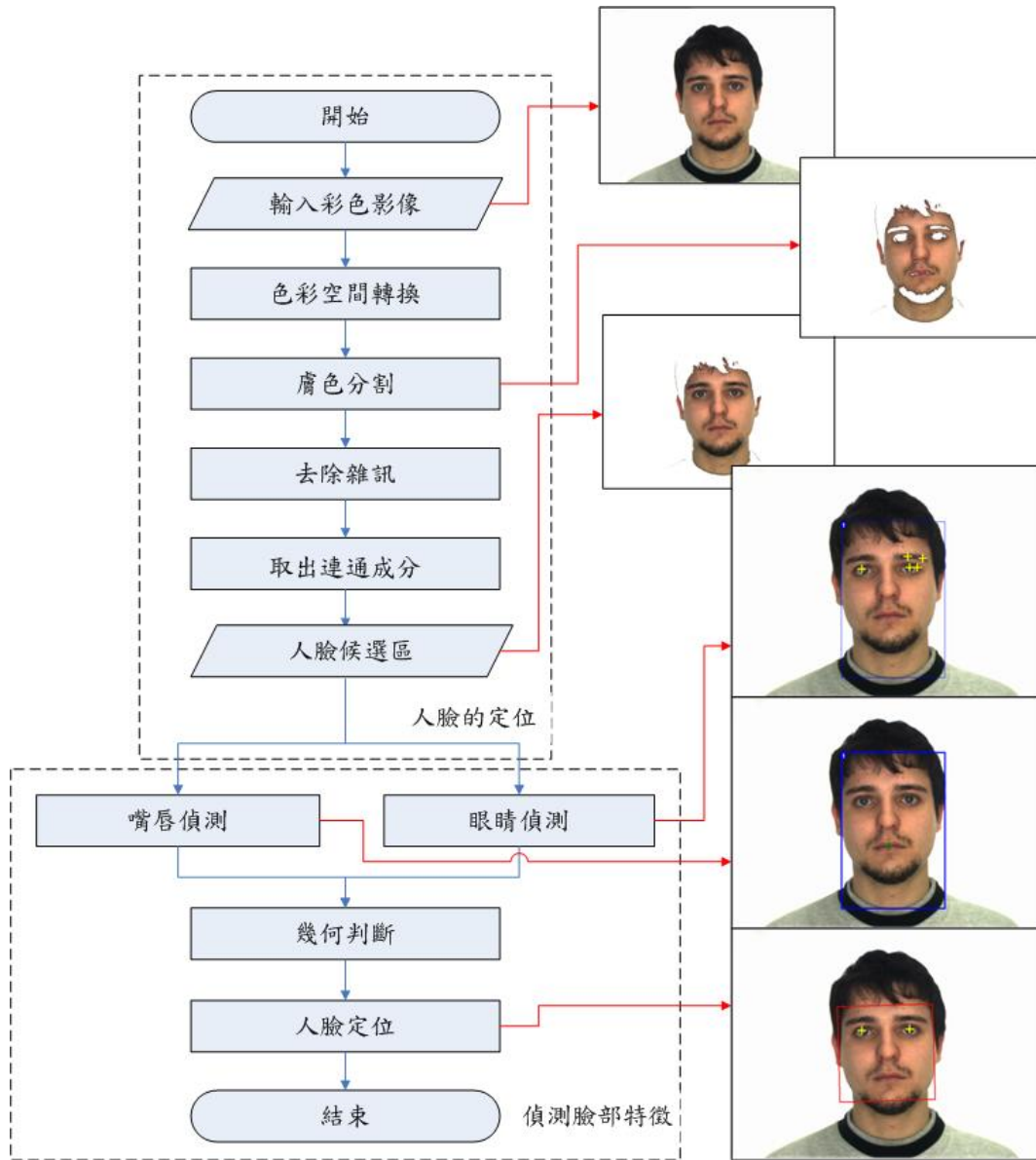
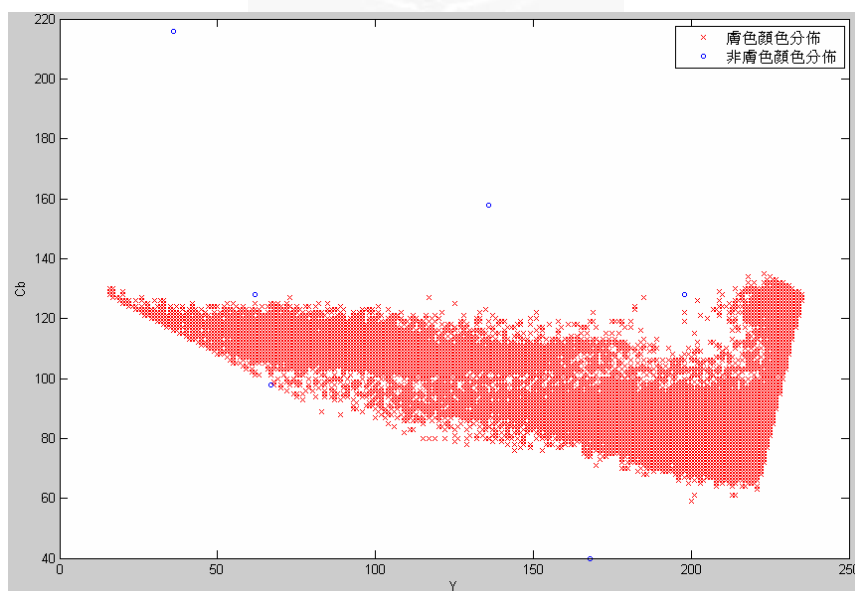


圖 3.1 人臉偵測流程圖

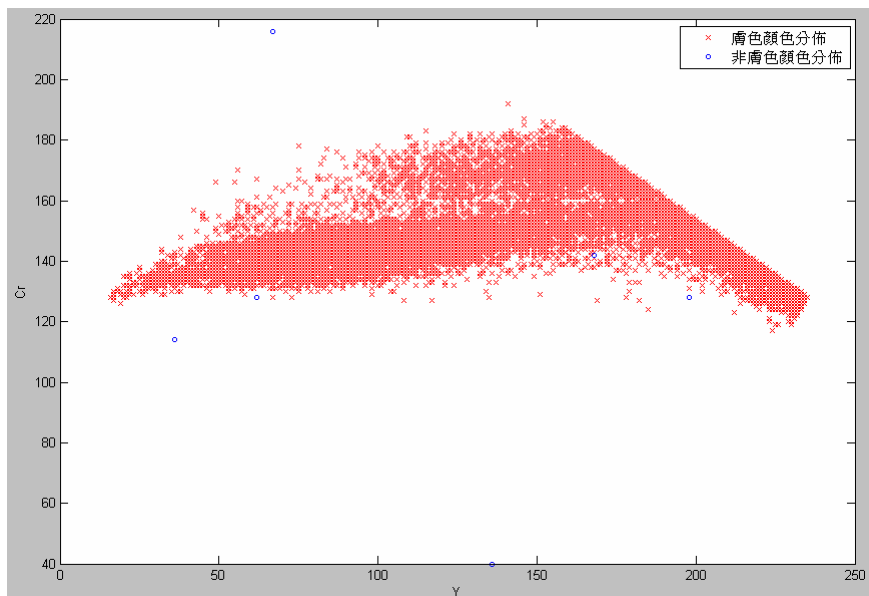
3-1 膚色分割

到目前為止，膚色偵測在不同的應用中，都會碰到一些不同的問題。同時對於何種色彩座標較佳，似乎也沒有一致的看法。我們歸納膚色對於影像形成的條件，有極大的關係。所以只用膚色特徵並不是最好的辦法。雖然其優點是快速且容易實現人臉偵測的方法，但缺點是容易受外界環境、光線照射及人臉上的表情等因素影響。

本文的膚色分割是以色彩為基礎，也就是根據各色彩成分之像素值大小來進行膚色的判斷，在文獻[6]的方法中發現在不同的亮度中膚色所群聚的範圍不一樣，也就是說不同的亮度膚色的Cb、Cr的範圍是不一樣的，所以我們統計不同亮度的膚色影像，統計出來的圖如圖 3.2 所示，其中 Y 值是指亮度值。3.2(a)中 x 軸是指亮度的範圍由 0~255，0 是黑色、255 是白色中間部分以此類推，y 軸是指 Cb 的範圍，圖中紅色的部份就是膚色在 Y-Cb 空間所群聚的範圍，3.2(b)是代表膚色在 Y-Cr 空間的群聚範圍。



(a)



(b)

圖 3.2 膚色群聚的範圍在不同的亮度下(a) 膚色在 Y-Cb 空間的群聚範圍;(b) 膚色在 Y-Cr 空間的群聚範圍

基於圖 3.2 的色彩空間我們可以定義可能是膚色的範圍如下表 3.1 所示，因為黑色和白色的 Cb、Cr 也在 127~128 之間，所以我們多了一個判別式可以避免白色和黑色會被當作膚色來偵測， $R_n = R / (R+B+G)$ 因為黑色和白色的 R、B、G 值差不多所以 R_n 接近 1/3 時，那個 pixel 可能是黑色或白色的機率很高，且膚色的 R 值雖然很高但是當一個 pixel 的 R 值大於 $R+G+B$ 的一半時時這個 pixel 的顏色就接近紅色而不是膚色。我們使用之前統計膚色範圍的人臉計算 R_n 的值，發現膚色的 R_n 範圍介於 0.343~0.702 之間大部份的膚色 R_n 值是介於 0.4~0.5 之間的，由之前的條件和膚色 R_n 的範圍我們定義膚色 R_n 值的範圍為 0.36~0.6，這樣可以保留下膚色的部份又可以濾除掉黑色白色和太紅的顏色。使用 R_n 的範圍再加上我們歸納出的不同亮度下的膚色範圍，我們就可以利用 3.1 式分割出正確的膚色範圍。

表 3.1 不同亮度值下膚色群聚範圍表

n	亮度值(Y)	Cb 膚色範圍	Cr 膚色範圍
0	$Y \leq 50$	$108 \leq Cb \leq 131$	$120 \leq Cr \leq 145$
1	$51 \leq Y \leq 110$	$83 \leq Cb \leq 132$	$123 \leq Cr \leq 175$
2	$111 \leq Y \leq 180$	$75 \leq Cb \leq 135$	$125 \leq Cr \leq 185$
3	$181 \leq Y \leq 205$	$68 \leq Cb \leq 138$	$115 \leq Cr \leq 160$
4	$206 \leq Y \leq 219$	$65 \leq Cb \leq 140$	$115 \leq Cr \leq 158$
5	$Y \geq 220$	$65 \leq Cb \leq 141$	$115 \leq Cr \leq 153$

$$\text{Face}(x,y) = \begin{cases} 1 & \text{if } [(0.36 \leq R_n \leq 0.6) \text{ and } \text{Map}_n(a,b) = 1], \\ 0 & \text{otherwise,} \end{cases}$$

$$n = \begin{cases} 0 & \text{if } [Y(x,y) \leq 50], \\ 1 & \text{if } [51 \leq Y(x,y) \leq 110], \\ 2 & \text{if } [111 \leq Y(x,y) \leq 180], \\ 3 & \text{if } [181 \leq Y(x,y) \leq 205], \\ 4 & \text{if } [206 \leq Y(x,y) \leq 219], \\ 5 & \text{if } [Y(x,y) \geq 220], \end{cases} \quad (3.1)$$

$$R_n = R/(R + B + G),$$

a 和 b 分別指 Cb、Cr 的膚色範圍，Y 表示在輸入影像(x,y)位置中的亮度值，n 是指不同的亮度等級，上列公式的意思是輸入一個 pixel 計算他的 Rn 值假如 Rn 值的範圍介如 0.36~0.6 之間的話，再計算它的亮度值判斷它是 0~5 那個 map 的範圍，定義是哪一個 map 之後再判斷它的 Cb、Cr 的範圍是不是屬於這個 map 假如是的話就判定為膚色。經由膚色分割後我們可以得到一個二值影像，由於我們使用不同亮度用不同的 Cb、Cr 值來判斷，所以得到的效果比一般的方法還好。在使用攝影機或是 CCD 時對如人臉的部份

會有一個特性，就是當人臉靠近攝影機時膚色的亮度會提高，遠離攝影機時膚色會顯的比較暗淡，因為拍攝的遠近是我們無法控制的，假如要對影像先做光線補償之後再去做後續的處理，對於即時的系統在做人臉偵測前就要先做光線補償的前處理，會造成系統的負擔造成偵測的速度下降，所以我們一開始就先統計好不同亮度的膚色範圍，這樣輸入的影像就可以依照每一個 pixel 的亮度去選擇他的膚色範圍，不用經過前處理這樣可以提高系統的速度。如圖 3.3 所示。圖 3.3(b)中因為左半臉有強光照射，所以使用一般的方法左半臉有許多地方無法偵測到，圖 3.3(c)使用文獻[7]的方法，可以發現跟我們的效果差不多但是處理的時間文獻[7]耗時許多。

使用不同亮度的 Cb、Cr 範圍可以有效的偵測即使影像過亮或過暗都有不錯的效果，不過對於複雜背景中的 pixels 的色彩資訊剛好坐落在我們的膚色範圍內，就會被當作膚色即產生雜訊，我們在下一節中使用中值濾波器和型態學運算的方法去除小塊的雜訊。

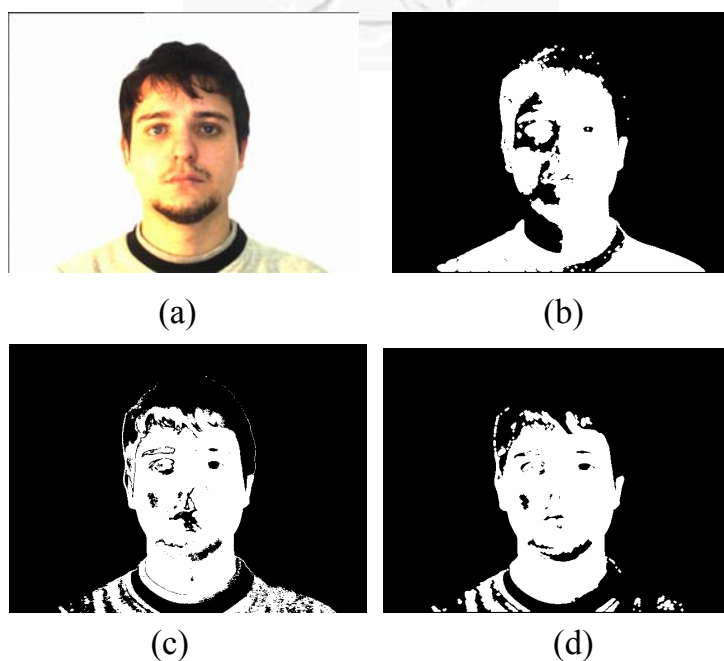


圖 3.3 膚色分割 (a)原始影像；(b)使用一般的方法；(c)使用文獻[7]光線補償之後的效果；(d)使用不同亮度取不同 Cb、Cr 值的方法

3-2 去除雜訊

在我們得到膚色分割的二值影像，可以看出裡面還有許多雜訊或是一些破碎的情況產生，這些雜訊的產生是因為在背景中剛好有些 pixels 的顏色屬於膚色的範圍，通常這些 pixels 的數量不大且都是分散開來的點雜訊所以我們使用中值濾波器去除雜訊。我們所使用的中值濾波器的遮罩是 3×3 的大小。

首先把遮罩內所含蓋的像素灰階值由小到大排列，中值是指排序在中間的那一個值，此值即為濾波器的輸出。此法特別適合用在有很強的胡椒粉式或脈衝式的雜訊時，因為這些灰階值的雜訊值與其鄰近像素的灰階值有很大的差異，因此經排序後取中值的結果是強迫將此雜訊點變成與其鄰近的某些像素的灰階值一樣，達到去除雜訊的效果。由中值濾波器的原理不難看出中值濾波器是一個非線性的動作。

在決定每一點的值之前，要對區塊內的每個亮度值排序一番，其計算量稍大於簡單的加權運算。然而以原有值取代，便不會產生影像中不曾出現過的新值，且雜訊也不會有什麼影響力。如圖3.4所示，一個3×3遮罩中的數字由小排到到可以得知第五個數字是80，所以就用80去取代原本遮罩中間的值，這是用於灰階影像時的狀況。在二值影像中，只有黑與白也就是0和255，中值濾波器的效果就是比較3×3遮罩中白點還是黑點多，假如白點多就用白點取代原本遮罩中間的值。

120	60	20
40	160	80
100	0	140

(a)

120	60	20
40	80	80
100	0	140

(b)

圖 3.4 中值濾波器(a)原本的遮罩；(b)取代後的遮罩

在尋找人臉候選區的時候，會有許多破碎的產生，所以我們提出以形態學(morphology)演算法來改善這個問題。接下來會先介紹一些形態學演算法的基本概念，然後再說明使用形態學演算法來處理雜訊結果。

形態學上之基本定義設A和B是二維空間中的兩個集合，它們的元素分別為 $a=(a_1,a_2)$ 和 $b=(b_1,b_2)$ 。A平移 $x=(x_1,x_2)$ 後用 $(A)_x$ 表示之，定義為：

$$(A)_x = \{c \mid c = a + x, a \in A\}, \quad (3.2)$$

A的補集是：

$$A^c = \{x \mid x \notin A\}, \quad (3.3)$$

A和B的差集，表示為 $A-B$ ，定義為：

$$\begin{aligned} A - B &= \{x \mid x \in A, x \notin B\}, \\ A - B &= A \cap B^c, \end{aligned} \quad (3.4)$$

膨脹(Dilation)，A被B膨脹，記為 $A \oplus B$ ，定義為：

$$A \oplus B = \bigcup_{b \in B} (A)_b, \quad (3.5)$$

意思是將A集合向外平移B集合中每個元素的距離，然後把這些集合聯集起來，所以膨脹會使A集合擴大，且擴大的範圍是由B集合之結構來決定。

侵蝕(Erosion)A被B侵蝕，記為 $A \ominus B$ ，定義為：

$$A \ominus B = \bigcap_{b \in B} (A)_{-b}, \quad (3.6)$$

意思是將A集合向內平移B集合中每個元素的距離，然後取這些集合的交集，所以侵蝕會使A集合縮小，且縮小的範圍是由B集合之結構來決定。

圖3.5為膨脹和侵蝕的結果其中B集合為一個 3×3 的遮罩。在圖3.5(b)中橘色的部份是膨脹出來的結果，在圖3.6(c)中淡掉的部份是被侵蝕的地方。

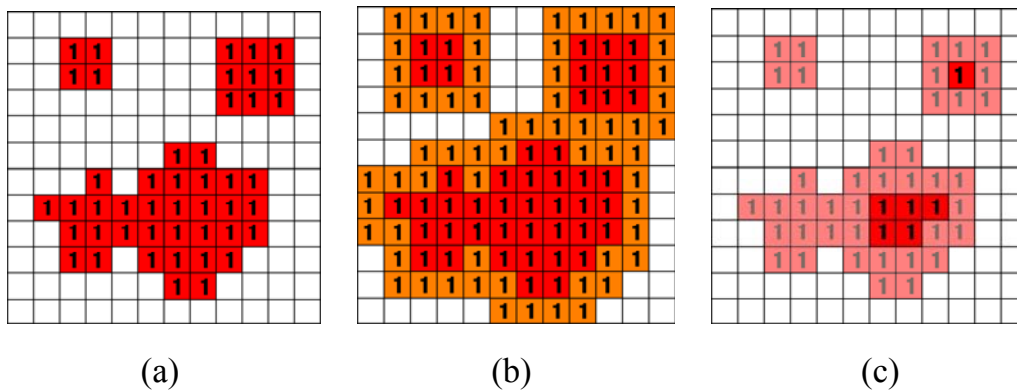
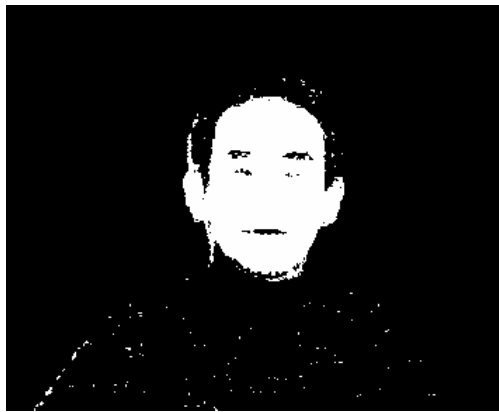


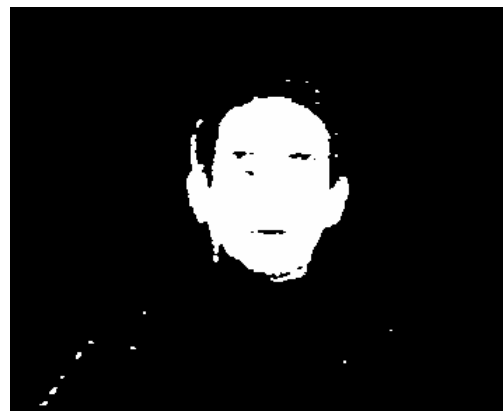
圖 3.5 膨脹和侵蝕的結果(a)輸入影像；(b)膨脹的結果；(c)侵蝕的結果

中值濾波器是利用一個 3×3 的遮罩，將遮罩內所包含的九個亮度值做大小的排序後，排序在中間的值(中值)作為該濾波器的輸出值，以此替代掉遮罩內位於中間的亮度值。這樣的方法特別對於處理孤立的點雜訊有效。在經過遮罩的作用之後，孤立的點或是面積較小的點都會被黑色所取代，視為雜訊而濾除。形態學運算是透過結構元素的中心點 X 對應到二值圖上，取出該點以及相對映所參照到的所有色階值執行邏輯 AND 運算，若是其中一個點的值為 0，則該點也設定為 0，否則不變。經過侵蝕後可以把體積稍大的雜訊點消彌掉。我們把二值圖形做中值濾波器和做閉合運算後，可以得到如圖 3.6 的結果。

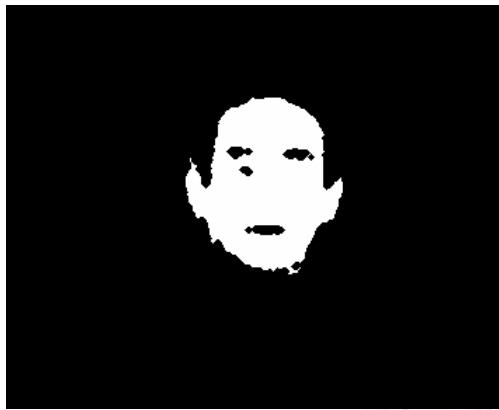
在使用中值濾波器和型態學運算之後可以去除掉比較小的點雜訊，不過對於在背景中出現的原木色桌子或是類似膚色的壁紙，這些屬於比較大塊的雜訊無法使用中值濾波器和型態學運算來去除所以我們就先用長寬比等條件去除一些比較大塊的雜訊，詳細內容在下一節提出。



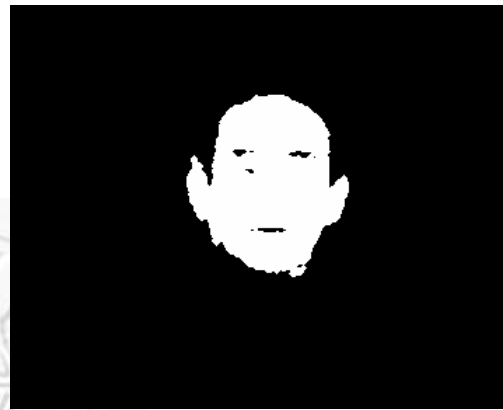
(a)



(b)



(c)



(d)

圖 3.6 經過中值濾波器和型態學運算的結果(a)經過膚色分割的影像；(b)經過中值濾波器的影像；(c)經過侵蝕運算的結果；(d)經過擴張運算的結果

3-3 取得連通成分和人臉候選區

在我們去除雜訊之後，我們得到圖 3.6(d)的影像，但是因為臉部特徵的顏色也不屬於膚色，所以導致眼睛、嘴巴等特徵也被當作是雜訊，這樣在接下來的處理會找不到眼睛、嘴巴的所在位置，所以我們使用區域填充的方法保留臉部的特徵以便於在之後的處理。

以圖3.7為例，A代表一個集合，它含有一個子集合，該子集合是依照邊界抽取演算法所求得的邊界，從邊界內側任選一點p就可以開始區域填充的程序，方法如下：

$$X_k = (X_{k-1} \oplus B) \cap A^c, \quad k=1,2,3,\dots, \quad (3.7)$$

其中 $X_0=p$ ，B集合如圖3.7(c)。第k次疊代時 X_{k-1} 會被B集合向4連通(4-connected)方向膨脹，不過因為有跟 A^c 取交集，不至於擴大到整個區域。如果不再變化則演算法終止，此時令 $X_{conv}=X_k$ ，A的區域填充 $R(A)$ 為：

$$R(A) = X_{conv} \cup A, \quad (3.8)$$

儘管這個例子只有一個子集合，但是演算法仍然可以用於任意有限數目的子集合，只要對每個子集合分別求出邊界與給定邊界內側的一點，就可以遞迴式的區域填充完各個子集合。

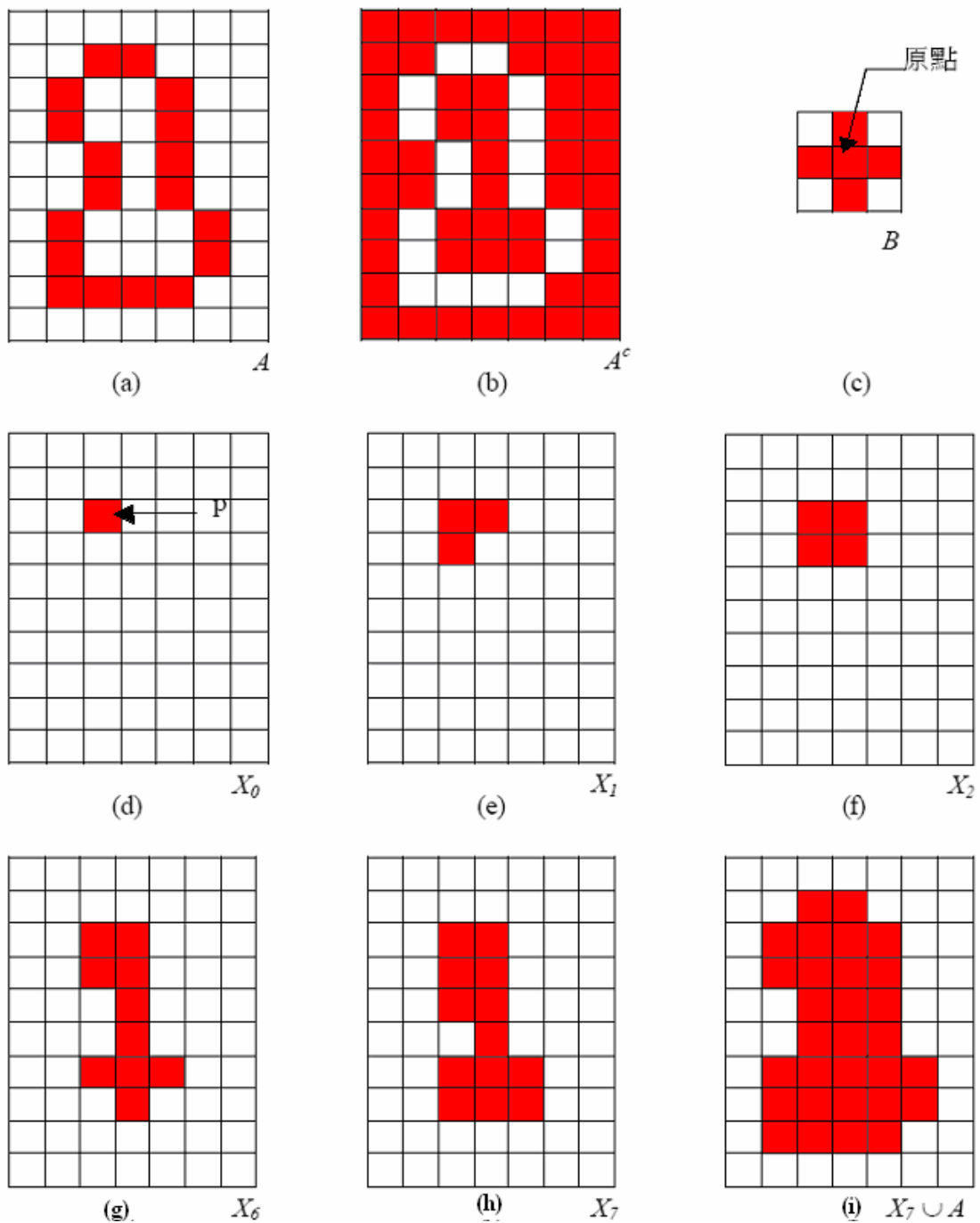


圖3.7 區域填充(a)含有一個邊界子集合的集合A；(b)A的補集；(c)結構集合B；(d)邊界內側的起始點p；(e)~(h)分別為各步驟疊代結果；(i)由(a)和(h)的聯集得到的最後結果。(資料來源：Digital Image Processing)

使用區域填充的方法可以將影像中白色部份中的黑色雜訊疊代成為白色的臉部區域，這樣代表在人臉中那些眼睛、嘴巴原本被認為是雜訊的部份，可以經由區域填充的方法保留下以便我們之後的處理。圖 3.8 所示的為經過區域填充結果。

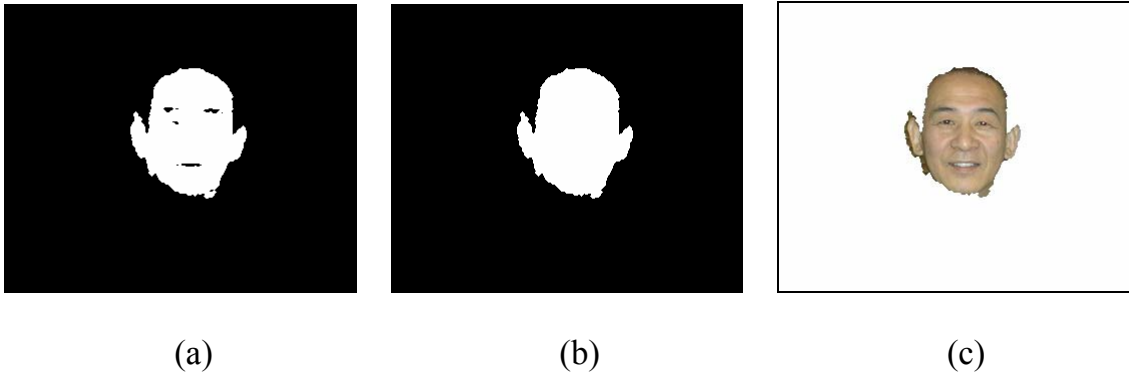


圖 3.8 區域填充的結果(a)經過去除雜訊的影像；(b)使用區域填充的結果；(c)

保留臉部的影像

在尋找影像中臉部狀態前，必須先找到人臉在影像中的位置。經過雜訊去除的步驟後，我們僅僅是在視覺上達到動態物件的萃取，為了要能得到物體實際在影像中的二維座標位置，我們可以抽取二值影像中的連通成分去取得該連通成分的位置。

以圖3.9為例，假設Y表示包含於集合A的一個連通成分，並假定已知Y的一個點p。則下面疊代表示式產生Y的所有元素：

$$X_k = (X_{k-1} \oplus B) \cap A, k=1,2,3,\dots, \quad (3.9)$$

其中 $X_0=p$ ，而B是一個合適的結構元素，假如 $X_k=X_{k-1}$ 則此演算法收斂，並另 $Y=X_k$ 。方程式(3.9)在形式上類似於方程式(3.8)，唯一的差別是這裡用的是A而不是他的補集。這種差別的發生是因為待尋找的元素(及連通成分的元素)都被標記為1。每疊代一次與A的交集消去了以標記為0之元素為中心上的膨脹。

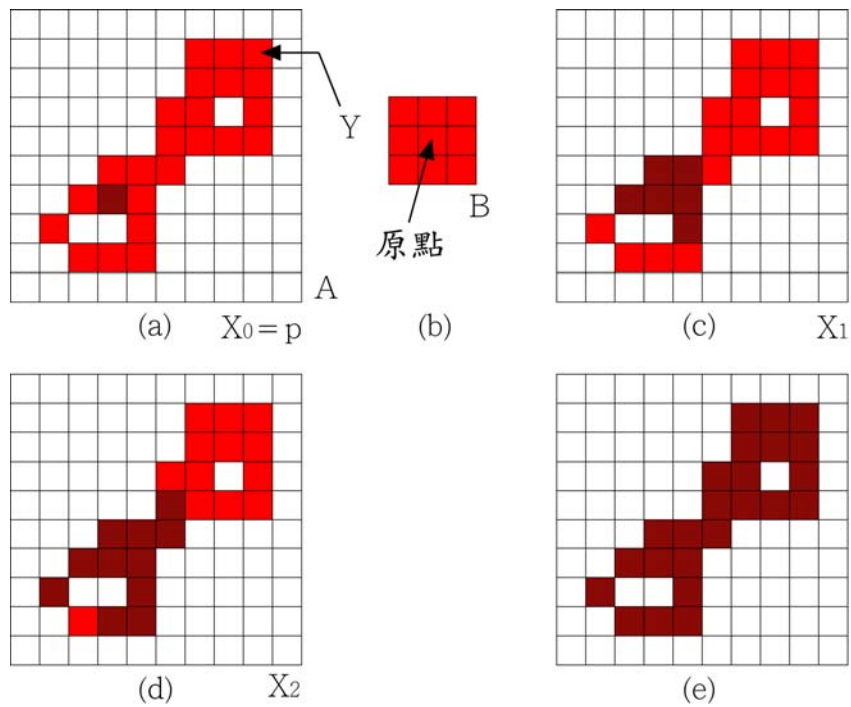


圖 3.9 連通成分(a)含有子集合 Y 的集合 A，p 為起始點；(b)結構元素 B；(c)第一次疊代結果；(d)第二次疊代結果；(e)最後結果。(資料來源：Digital Image Processing)

畫面中的膚色範圍可能會有一個以上，為了能夠分辨不同的膚色範圍，我們建立了一個與畫面同樣大小的矩陣，將二值化圖形轉換成一個二元圖 (Binary map)，接著透過上述連通成分的判斷，在與結構元素作用時分別給予不同的編號值。之後則可以利用 Binary Map 判斷出各個連通物件的座標，依照這個與畫面同樣大小的矩陣索引值找出在 x 軸與 y 軸方向上最大與最小的座標值，並以此座標值在畫面中畫出可框住其人臉範圍。如圖 3.10 所示。



圖 3.10 人臉候選區(a)單一人的狀況；(b)複數人的狀況

使用連通成分去框出人臉候選區，在框出人臉候選區之前我們先加入一些條件，因為使用中值濾波器和型態運算只能去除掉點雜訊，對於一些比較大塊的雜訊就無法達到去除的效果，所以我們第一個條件是候選區的 pixels 數不可以太小，對於一個 320×240 的影像中人臉的大小大約在 30×30 以上，對於小於 20×20 的區域大多是雜訊或是小到可以當作背景的人臉，所以我們計算每個區域的 pixels 的數量小於影像的 0.4% 我們就不會把他當作人臉候選區也不會去做接下來的處理。還有框出來長方形的寬高比必須要在一定範圍內，由人體計測資料庫中人臉的寬高比介於 0.7~1 之間，但是有些人臉的膚色範圍是有包含脖子的部份以及避免誤差所以我們在另外統計人臉的寬高比大約為 0.5~1.2 之間，所以當一個區域過長或是過寬的時候，就可能是人的手臂或是類似膚色的木頭桌子，所以就用長寬比這個條件去濾除一些大塊面積卻不是人臉的區域。符合上述兩個條件的候選區才會被框起來做之後臉部特徵擷取的步驟，如圖 3.11 所示，在(b)圖中多了兩個連通成分，使用我們的條件後就有兩個非人臉的候選區被移除，沒有被選中的候選區就不會進行之後的運算，這樣可以減輕運算量。其中框出來的候選區有些並不是真的人臉會在之後的步驟再做判斷。

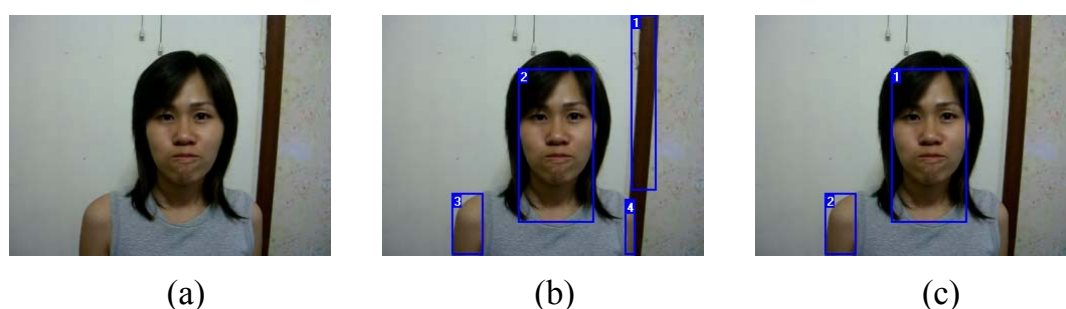


圖 3.11 連通成分的結果(a)原始影像；(b)沒有使用條件；(c)使用條件。

3-4 眼睛和嘴唇的偵測

在定出人臉區域以後，我們要先抓取眼睛和嘴唇的特徵點，首先利用眼睛在臉部的位置，根據人體計測資料庫，在人臉長方長區塊內定出在高度的20%到55%區間內為瞳孔大概所在區域，然而在此區域內找出數個可能為眼睛的座標，所以我們還要找出嘴唇的座標來定義最正確的臉部特徵，因為頭會旋轉所以嘴唇的位置會隨著x軸移動由左移到右或由右移到左，所以嘴唇會出現的位置就在高度的40%到80%之間，使用眼睛和嘴唇在臉部出現的範圍來偵測眼睛、嘴唇可以降低運算量，也不會在不可能出現眼睛的位置去找眼睛在不可能出現的嘴唇的位置去尋找嘴唇。我們用表3.2的公式可以找出眼睛和嘴唇可能出現的範圍，再使用之後所介紹的方法在我們定義出眼睛和嘴唇的範圍去尋找眼睛和嘴唇。其中 $X1-X0$ 表示人臉候選區的寬， $Y1-Y0$ 表示人臉候選區的長。在圖3.12中紅色的框表示人臉候選區，黃色的框表示尋找眼睛的範圍，綠色的框表示尋找嘴唇的範圍，在圖3.12中可以發現我們偵測眼睛會嘴唇的位置重疊了，一個正常的人臉眼睛會嘴唇的位置不會重疊，我們是考量測膚色時因為實驗者衣著的關係脖子的部份可能會露出來可能不會，這樣會造成人臉寬高比不固定，所以為了避免定義尋找眼睛嘴唇範圍的錯誤，所以我們就加寬偵測的範圍以對應各種不同的情況。

表3.2 眼睛和嘴唇的範圍座標表

	偵測眼睛的範圍	偵測嘴唇的範圍
左上的X座標	$X2=X0$	$X4=X0$
左上的Y座標	$Y2=Y0+0.15 \times (Y1-Y0)$	$Y4=Y0+0.4 \times (Y1-Y0)$
右下的X座標	$X3=X1$	$X5=X1$
右下的Y座標	$Y3=Y0+0.55 \times (Y1-Y0)$	$Y5=Y0+0.8 \times (Y1-Y0)$

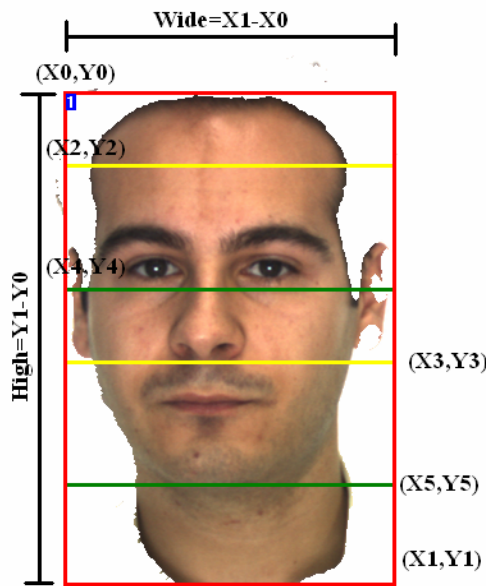


圖 3.12 偵測臉部特徵的範圍

3-4-1 自適性的眼睛候選區偵測

不論人種，眼睛區域相對於人臉膚色在色調上往往都會較為陰暗。因此可以利用眼睛與膚色的差異，讓人眼從臉部區域中分離出來。但若直接用彩色影像進行處理，則要判斷不同人種的眼睛顏色(黑色、褐色、深藍色等)將會是一項繁雜的程序。有鑑於此，我們就希望在眼睛顏色判斷上，減少不同人種眼睛顏色差異的影響。我們發現將彩色影像轉為灰階影像之後，不但可大幅減少色調(Hue)和飽和度(Saturation)的依賴。而且發現眼睛在人臉之中，不管那一人種，將依然呈現出與周圍皮膚灰階度上的明顯差異。因此，我們只需要再訂定出一個門檻值，即可將眼睛從人臉區域中區隔出來。灰階值大約在一個門檻值(65~80)以下，但門檻值以下的灰階值，不一定是人的瞳孔。而且在眼睛部分所形成的黑色區域形狀，應該是長方形、方形或圓形。所以先保留這些形狀的黑色區域，並消除一些不是直線長條形或是垂直方向形狀的黑色區域。所以眼睛是可以經過形狀來分析篩選的。然後再使用增加黑色區域的方法，做法是先將彩色影像轉成灰階的影像，以一個5×3的遮罩掃描整幅影像，並改變影像的灰階值，公式如下[2]。

$$g(x,y)=\begin{cases} \min & \text{if } [f(x,y)-\min < \max-f(x,y)], \\ \max & \text{otherwise,} \end{cases} \quad (3.10)$$

$f(x,y)$ 代表 3×5 遮罩中心的灰階值， \max 、 \min 分別為 $f(x,y)$ 相鄰區域灰階值中的最大最小值，用此公式可以增強黑色區域並改善影像中灰階值的對比程度，下圖 3.13 是用公式(3.10)所得到的結果。

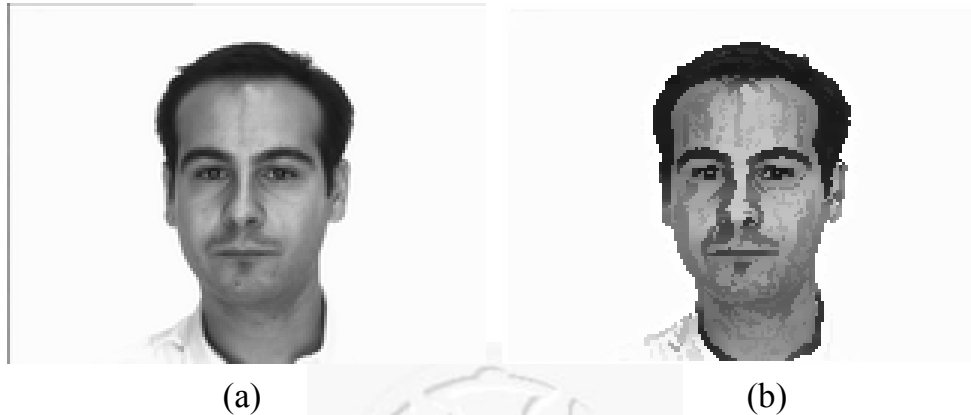


圖 3.13 黑色增強的結果 (a)原本的灰階圖；(b)是用公式(3.10)後的灰階圖

由上圖可以明顯看出對比和黑色的程度都增強了，然後在人臉可能區域中依照我們之前定義的眼睛範圍，作為所要搜尋眼睛特徵之範圍，接下來將原圖中相對之區域由彩色影像先轉為灰階圖形，再取一門檻值如公式(3.11)，將小於門檻值的像素定為1，大於門檻值的像素定為0。接下來使用斷開運算，將雜訊點移除。並再次運用連通成份標定程序，求得每一物件的範圍和面積，最後並將所有找出的連通成份區域座標位置記錄下來。經由此眼睛偵測的步驟後即可獲得。

$$\text{Eye}(x,y)=\begin{cases} 1 & \text{if } [g(x,y) < T_{\text{eye}}], \\ 0 & \text{otherwise,} \end{cases} \quad (3.11)$$

但是由於光線以及種族的的不同，所以固定閾值的話會造成有些影像某種程度上的誤判，所以我們使用自適性的閾值來判斷。圖 3.14 表示在不同亮度下的眼睛候選區，我們使用的閾值固定為 90，其中白色的部份被當作眼

睛的候選區，可以發現在(a)圖中因為臉部過暗使用閾值為 90 的話可以明顯看到整張臉的大部分都被判定為眼睛，在(c)圖中因為設定的閾值太低導致眼睛的候選區太小，可能會在之後去除雜訊的動作被當成雜訊去掉使得我們無法偵測到眼睛候選區，所以我們知道使用固定的閾值來分割眼睛，在不同亮度下會可能會無法偵測，所以我們使用自適性的閾值，針對不同影像給予不同的閾值。

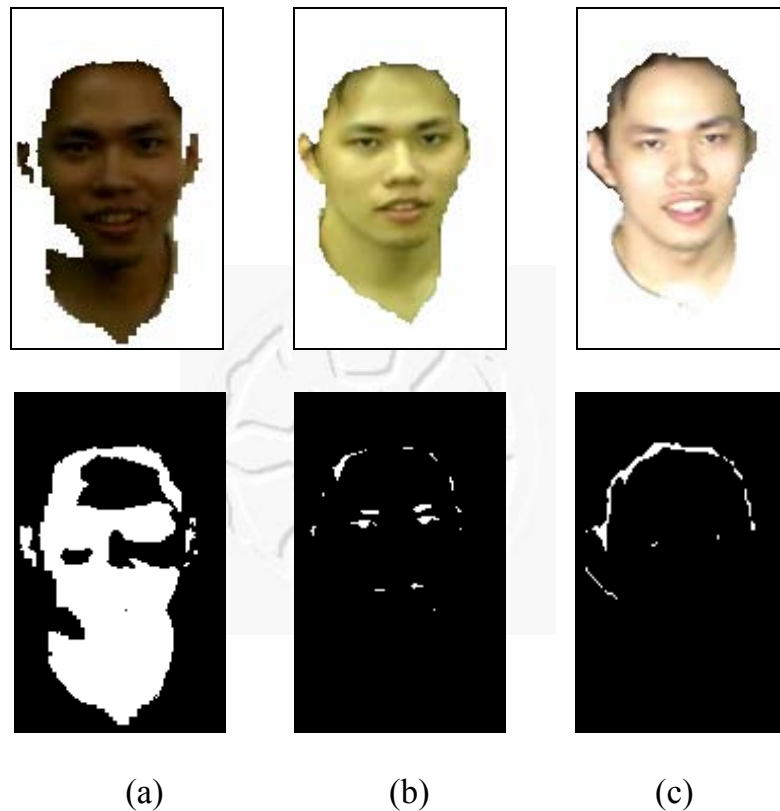


圖 3.14 不同亮度下的眼睛候選區(使用的閾值 T_{eye} 為 90)(a)臉部平均亮度大約 50；(b)臉部平均亮度大約 145；(c)臉部平均亮度大約 200

我們知道眼睛的面積占整張臉的面積大約 2% 左右(資料來源：人體計測資料庫)，先把初始閾值 T_{eye} 設為 0，計算輸入候選區的 pixels 數，影像從左到右從上到下計算每個 pixel 的灰階值，假如小於 T_{eye} 就當作是眼睛的

pixels，當眼睛的 pixels 數目等於 2%候選區的 pixels 數目，這時的 T_{eye} 就是最適當的閾值，反之當眼睛的 pixels 數目小於 2%候選區的 pixels 數目，我們就把 T_{eye} 加 1 再計算一次，直到當眼睛的 pixels 數目大於 2%候選區的 pixels 數目。這流程就如圖 3.15 所示。

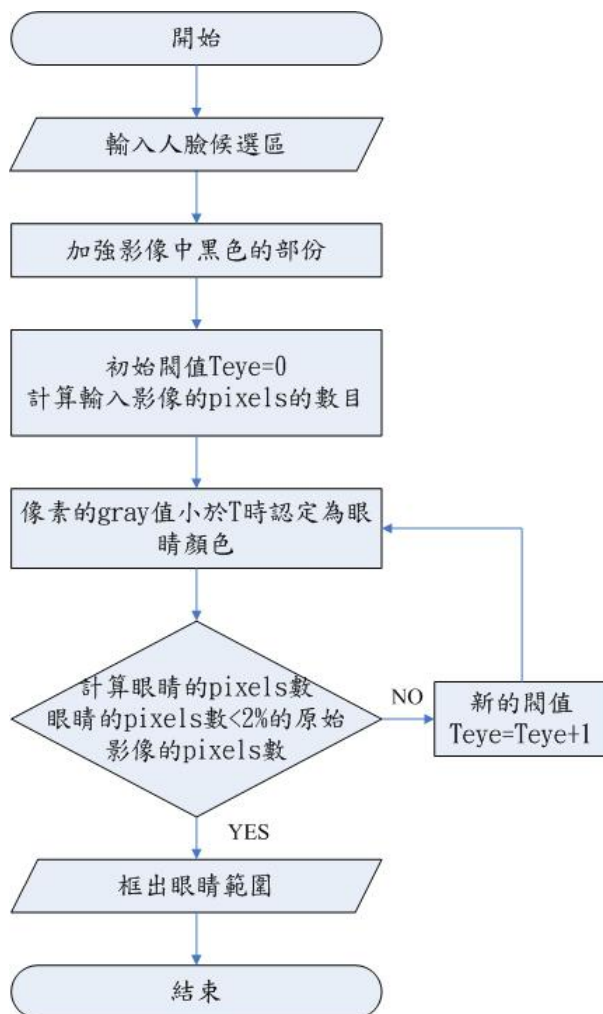


圖 3.15 自適性眼睛偵測流程圖

使用這個方法可以用正確的閾值去分割每張影像中的眼睛，但是只使用灰階值來判斷的話，會有其他的雜訊產生例如眉毛、鼻孔、眼鏡的邊框和反光等等…，如圖 3.16(a)中眉毛也被當成眼睛的候選區，如圖 3.16(b)鏡框和鼻孔都造成錯誤的雜訊。所以我們還要進一步的偵測出嘴唇，利用眼睛和嘴唇的幾何關係來判斷哪些是正確的眼睛。



圖 3.16 眼睛偵測的結果

3-4-2 自適性的嘴唇候選區偵測

同樣使用之前偵測到的人臉可能區域中依照我們之前定義的嘴唇範圍，我們使用色彩分析法發現，皮膚 R 值(紅色的成分)比嘴唇的 R 值高，雖然嘴唇的 R 值不高但相對的比 G、B 值(綠色、藍色的成分)高出許多，所以我們看嘴唇的顏色才會像紅色。我們統計不同亮度下嘴唇的 R、G、B 值，統計出來的結果如下表 3.3。

表 3.3 不同亮度下唇色的 RGB 值

Y 值範圍	平均 R 值	平均 G 值	平均 B 值	R-G	R-B
51~70	89.5	53.25	34.75	36.25	54.75
71~90	111.33	69.33	50.83	42	60.5
91~110	137.33	83.8	71.13	53.2	66.33
111~130	165.25	103.5	91.13	61.75	74.25
131~150	184.75	122.75	109.88	62	74.57
151~170	207.14	144.14	132.57	63	74.87
171~190	224.17	160.83	151.67	66.5	66.5

我們知道當亮度提高後 RGB 值也會跟著提高，在唇色的部份隨著亮度的提高 R-G 和 R-B 的值也會等比例的提高有上表 3.3 可以得知，在不同的亮度情況 R-G 的值大約都是 R 值的 0.25 倍以上而 R-B 的值大約都是 R 值的 0.3 倍以上，這樣可以先使用 RGB 之間的關係先濾除掉一些在臉部中非唇色的部份，在利用自適性閾值分割正確的嘴唇部分，如(3.12)式。

$$\text{Mouth}(x, y) = \begin{cases} 1 & \text{if } [Cr \leq T_{\text{mouth}}] \cap [R - B > 0.25 \times R] \cap [R - G > 0.3 \times R], \\ 0 & \text{otherwise,} \end{cases} \quad (3.12)$$

如圖 3.17 所示，其中白色的點為判定位唇色的點，在(a)圖中因為閾值設的太小所以在臉部很多地方都被認為是嘴唇的顏色，(c)圖也是沒有給予適當的閾值使的偵測的效果沒有很好。

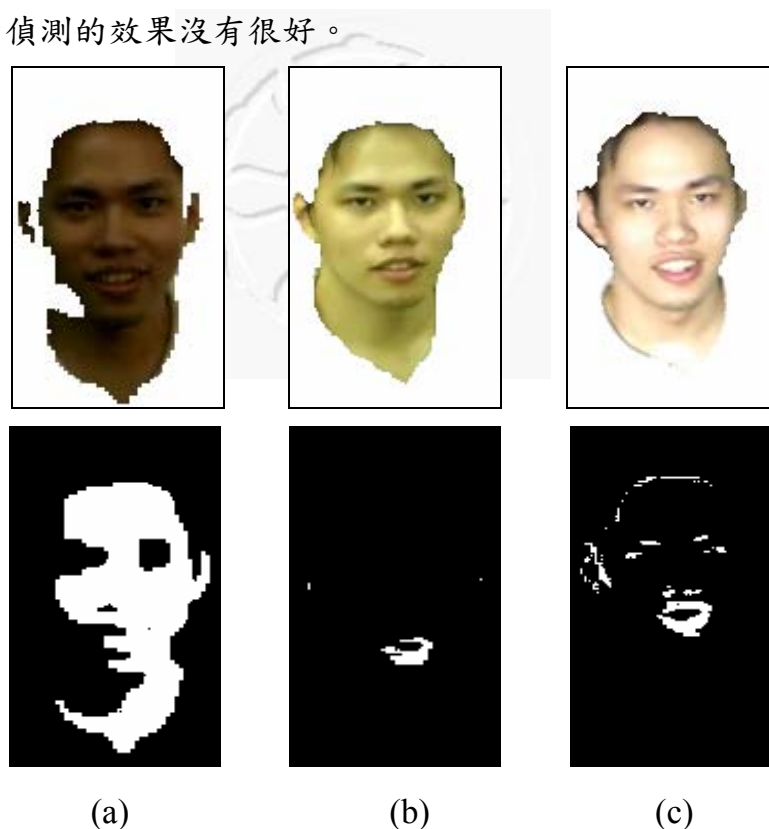


圖 3.17 不同亮度下的嘴唇候選區(使用的閾值 T_{mouth} 為 140)(a)臉部平均亮度

大約 50；(b)臉部平均亮度大約 145；(c)臉部平均亮度大約 200

如同偵測眼睛的流程一樣根據嘴唇占人臉的比例為 5%(資料來源：人體

計測資料庫)，因此我們取 5% 的 pixels 數為嘴唇可能的大小，來決定閾值 T_{mouth} ，當嘴唇的 pixels 數目大於 5% 候選區的 pixels 數目，我們就把 T_{mouth} 加 1 再計算一次，直到符合我們的條件，不同的閾值得到的嘴唇 pixels 數也不一樣。圖 3.18 所示為偵測嘴唇的流程圖。

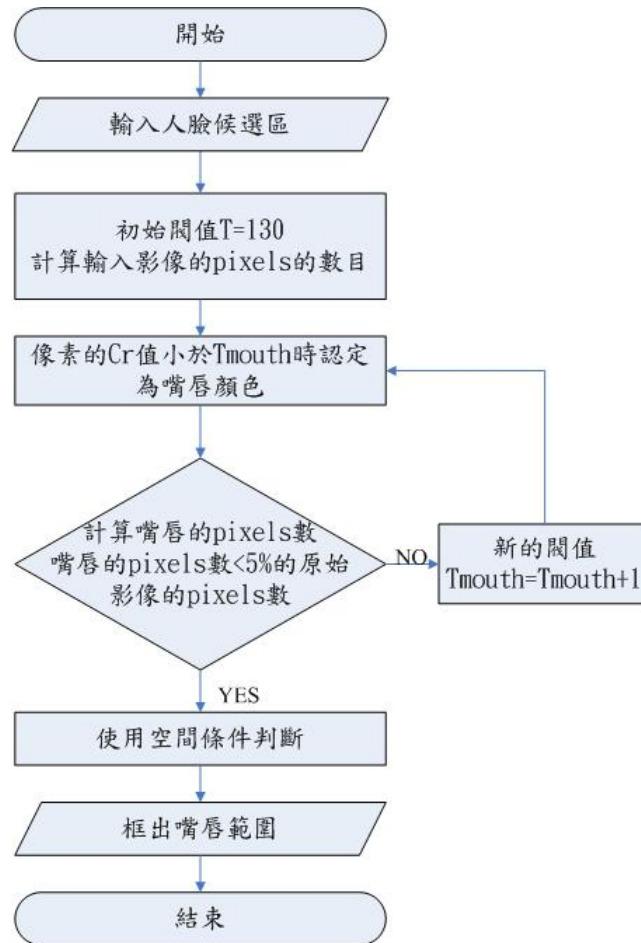


圖 3.18 嘴唇偵測的流程圖

我們在偵測嘴唇候選區時，因為有時候耳朵和臉頰的部份會呈現粉紅色，所以我們在使用連通成份取得嘴唇之前先使用嘴唇的寬高比，因為嘴唇是呈現長方形的樣子，所以我們設定寬高比要大於 1.5 才可以定義為嘴唇，再判斷連通成份最大的一個且符合寬高比大於 1.5 我們才會判斷為嘴唇。由於有些影像嘴巴是張開的，這樣會偵測到上下兩個嘴唇，所以我們使用(3.13)

式去找嘴唇區域的權重位置， n 代表嘴唇區塊 B_i 中的 pixels 數， x_i 、 y_i 分別代表嘴唇區塊中每一個 pixels 的 x 、 y 座標，即使嘴巴張開我們也可以找到整張嘴的中心位置 (x_c, y_c) 。圖 3.19 呈現嘴唇偵測的結果綠色十字的地方就是我們偵測出嘴唇的位置，圖 3.19(a) 是張開嘴唇的實驗者因為我們是取得中心位置，所以綠色十字的位置是在上下嘴唇的中間。

$$\begin{aligned} x_c &= \sum_{x_i \in B_i} x_i / n, i=1,2,3,\dots, \\ y_c &= \sum_{y_i \in B_i} y_i / n, i=1,2,3,\dots, \end{aligned} \quad (3.13)$$

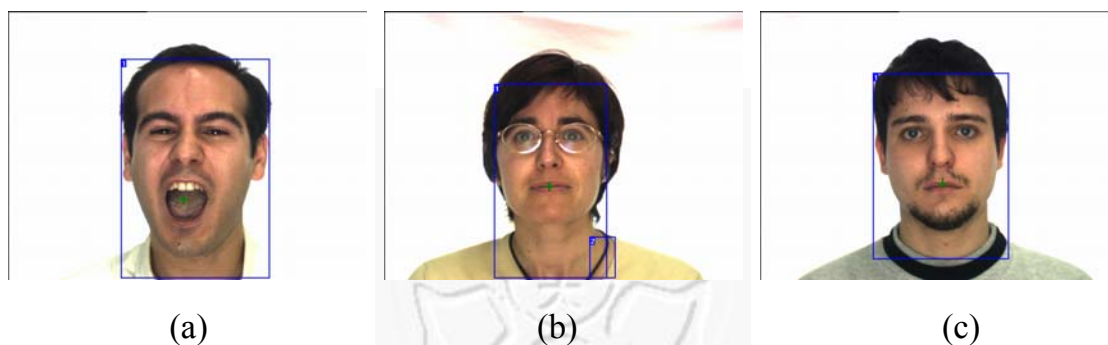


圖 3.19 嘴唇偵測的結果

3-5 幾何判斷

找到臉部特徵，利用眼睛、嘴唇之間的幾何關係，由人體計測資料庫發現到眼睛嘴唇之間呈現等腰三角形，以及眼睛出現的位置，我們定出下列規則：

規則一：兩眼之間距離在臉寬的 0.2 倍到 0.8 倍之間。

規則二：兩眼之間的斜率限制於 0.5 以下大約 27 度左右。

規則三：嘴唇的位置必須在兩眼之間。

規則四：兩眼中心點到嘴唇的距離必須大於 0.3 倍的臉長距離。

因為頭部會旋轉所以眼睛的距離在正面人臉的時候是臉寬的 0.8 倍，但是當臉轉動的角度載 0~45 度或是 0~-45 度時兩眼的距離會越來越小，一直到 45 度或是-45 度時兩眼的距離是臉寬的 0.2 倍，所以我們取兩眼的距離必須是臉寬的 0.2~0.8 倍。在我們偵測眼睛時是使用灰階度來偵測眼睛，所以有時會把眉毛或是鼻孔也偵測進去，我們加入斜率的條件因為兩眼會呈現水平的狀態，使用斜率就可以避免錯誤的眼睛對。還有根據嘴唇必須在兩眼之中的條件，也可以濾除掉一些被誤測為眼睛的雜訊。

基於以上的規則，我們可以先慮除掉一些非眼睛的雜訊點，接下來我們用下來公式來定義出正確的人臉。

$$D = d1 + |(d2/d3) - 1| \times \frac{\text{Height}}{2} + d4 \quad (3.14)$$

其中 Height 是指人臉候選區的高度，d1 是指兩眼之間的距離，d2、d3 分別指左眼和右眼到嘴唇中心點的距離，d4 是代表兩眼距離的中點到嘴唇中心點的距離，如圖 3.20 所示。因為在人臉候選區中幾使使用上列規則保留下來的可能會有一個以上的眼睛對，所以對於每組眼睛對都會去計算它的 D 值計算出來 D 值最小的眼睛對我們就把它定義為正確的眼睛對，我們知道臉部特徵眼睛和嘴唇是呈現等腰三角形，所以 d2 和 d3 的距離會相等，為了

避免眉毛對會被偵測為眼睛對我們知道眉毛的位置在眼睛的上方，所以眉毛對的 d_2 ， d_3 ， d_4 都會比眼睛對的值還要高，所求出的 D 值越小是正確人臉的機率越高，在某些時候因為鼻孔的位置也會呈現低灰階度的情況，因為我們把眼睛的距離設定在臉寬的 $0.2\sim 0.8$ 倍之間，而鼻孔對和眼睛對都是呈現水平狀態，所以我們加入嘴唇到兩眼中心的位置必須大於臉寬的 0.3 倍這也是由人體計測資料庫中統計出來的距離。所以我們是取計算出來的 D 值最小為最正確的人臉，我們所偵測出來的結果會在下一章呈現出來。

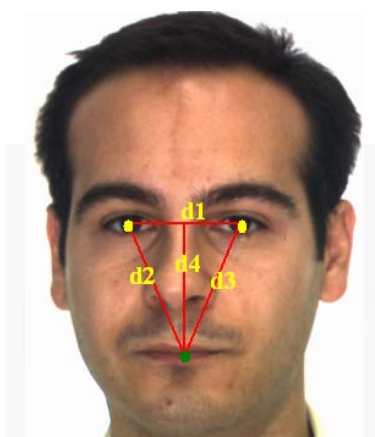


圖 3.20 臉部特徵的幾何距離

第四章 實驗與分析

本研究中利用 DirectX 9 Software Development Kit 將與電腦連結的一般攝影機所攝得的畫面即時的傳送至系統中，而系統在接收到畫面後即對影像進行處理，這讓我們僅需要將網路攝影機與電腦連結即可偵測即時的畫面，或是在系統中載入預先錄下的影片檔案也可在事後進行影像的分析處理，表 4.1 為實驗用的硬體規格表。

表 4.1 實驗中所使用的硬體規格

CPU規格	Pentium 4 2.4GHz	
記憶體	256MB	
攝影器材	CASIO EX-Z750	Logitech QuickCam Pro
影像解析度	320×240	

為了驗證本系統在不同的狀況下的偵測效果，所以我們設定了以下幾個實驗：

1. 針對不同大小的人臉來驗證系統的偵測率。
2. 針對不同亮度的影像來驗證系統的偵測率。
3. 對於其他情況驗證系統的偵測率，這裡的其他情況是指頭旋轉的角度以及戴眼鏡的實驗者和多人的實驗。
4. 針對我們的方法和別人的方法比較。

第一個實驗中我們使用三個不同的人臉大小來做偵測，分別是 100×100、75×75、50×50 三個大小，這邊大小是指最後被框出來的正確人臉的大小，在實驗的影像我們分別對每個實驗者錄製 10 秒鐘的影像，在這 10 秒鐘實驗者可以輕微的擺動，在 10 秒鐘會有大約 80~110 個 frame，我們每個實驗者都會取 30 張影像在 10 秒之內，在每個實驗中我們都有五個實驗者所以我們影像張數都有 150 張，我們計算處理速度是使用 FPS 的方法，FPS

是 frame per second 的簡寫代表意義為每秒可以處理的畫面數，FPS 10 代表每秒鐘可以處理十張畫面，用此來驗證我們的系統速度。下表 4.2 是我們實驗結果。

表 4.2 不同大小人臉的實驗結果

實驗項目	100×100	75×75	50×50
影像張數	150	150	150
正確人臉	145	136	121
誤判人臉	5	14	29
正確率	96.66%	90.66%	80.66%
FPS	8	11	13

在第一個實驗項目中我們可以看到當我們人臉大小在 100×100 時我們有極高的偵測率，因為我們使用自適性的閾值來偵測人臉特徵，對於每張不同的影像我們都分別給予不同的閾值來分割眼睛和嘴巴，相較於使用固定的閾值可能有些影像可以偵測到但是影像一有變動就會產生誤測。對於我們沒有偵測到的影像是因為實驗者有閉眼的情況導致特徵點遺失。但是當我們的人臉影像越來越小時我們偵測率就會下降，因為我們是使用色彩分析來偵測人臉特徵，人臉影像越來越小時人臉特徵也會相對變小，當人臉特徵太小時就會被當成雜訊去除了，所以我們的系統在人臉大小約 75×75 時也有良好的偵測率，但是人臉大小小於 50×50 我們的系統就不能有效的偵測人臉特徵，造成我們的偵測率無法達到理想的效果。對於處理的速度是根據連通成分的數量以及大小，因為連通成分數量越多之後步驟的處理數量也比較多，而連通成分面積越大代表要處理的影像也變大所以速度會降低下來，所以可以看到人臉大小在 100×100 的時候速度是處於比較慢的情況，隨著人臉範圍的變小處理的速度就有提高的趨勢。下圖 4.1 是由近拍到遠的實驗結果，圖 4.2 是由遠拍到近的實驗結果。在圖中的#符號是代表第幾張處理的畫面。

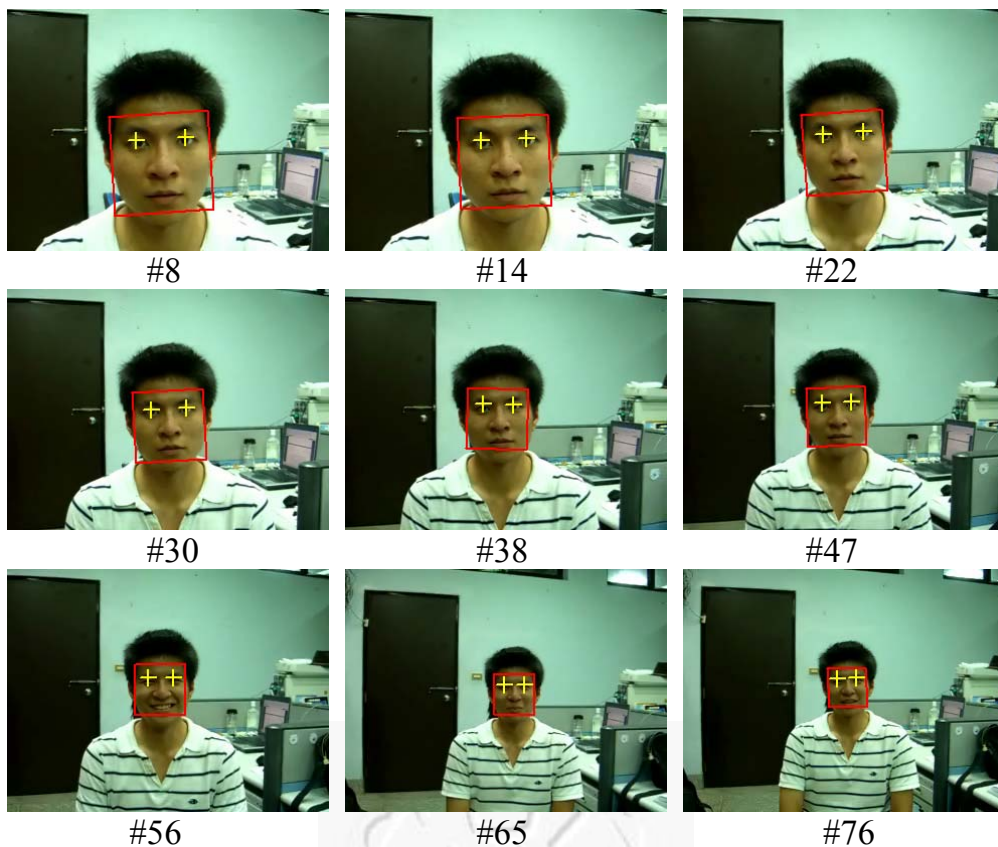


圖 4.1 由近拍到遠的實驗結果

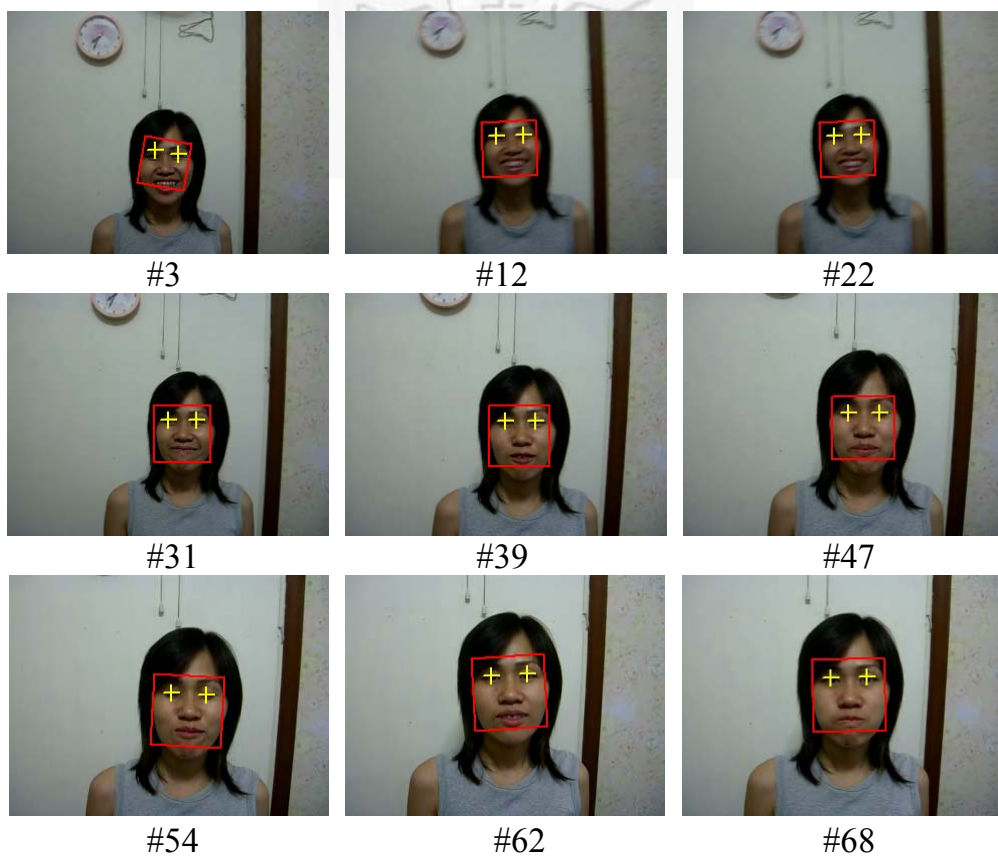


圖 4.2 由遠拍到近的實驗結果

第二個實驗中我們使用兩種不同的亮度情況來偵測，第一個情況是在只有微弱光線的陰暗室內，達到我們想要偵測在即使光線不足的情況下我們的系統依然有良好的偵測率，第二個情況我們實驗者的臉部打上白光，使實驗者的臉部膚色和特徵呈現比較高的亮度，來驗證我們的系統，在實驗的影像我們一樣對兩種亮度的情況分別取五個實驗者個別錄製 10 秒鐘，每一個情況分別得到 150 張影像。下表 4.3 是我們實驗結果。

表 4.3 不同亮度情況的偵測結果

實驗項目	較暗的情況	較亮的情況
影像張數	150	150
正確人臉	126	123
誤判人臉	24	27
正確率	84.00%	82.00%
FPS	10	10

對於較暗或是較亮的影像因為我們在統計膚色範圍的時候，就會統計包含較暗或是較亮的人臉影像，雖然統計的數量比一般光線的人臉影像還少，但是在實際實驗的時候對於光線不足和打光的情況下，我們對不同亮度的 pixels 給於不同的膚色範圍，所以在不同光線的情況下的膚色偵測效果也是不錯，除了比較極端的亮度過暗的膚色接近黑色以及打光太強造成有些膚色的 pixels 呈現接近白色的情況。在我們的實驗中那些錯誤的例子大多是臉部膚色影像在特別的情況下造成臉部膚色的破碎情況，臉部膚色破碎會造成區域填充的錯誤，我們使用區域填充是因為偵測膚色時，臉部特徵不屬於膚色所以使用區域填充把臉部特徵保留下來，當區域填充出錯時臉部特徵就會被當成背景，在接下來的步驟就不會去背景的地方偵測臉部特徵，所以照成在沒有臉部特徵的地方偵測臉部特徵，自然無法偵測到臉部特徵也就無法判定

為人臉。對於處理的速度因為我們在做不同亮度的偵測時我們都有固定實驗者的人臉大小，都是固定在 75x75 的大小，不過因為亮度的關係會使得原本非膚色的背景也會因為光線的關係變得接近膚色，雖然在之後的步驟會判定這些類似膚色的背景為非人臉，不過因為連通成分的增加所以會造成處理速度的下降，所以在不同亮度的人臉偵測的速度會比同樣的大小的人臉在一般光線的情況下還要慢。下圖 4.3 是較亮情況的實驗結果，在影像中人臉的平均亮度大約都在 180~200 左右，圖 4.4 是較暗情況的實驗結果，在影像中人臉的平均亮度大約都在 50~60 左右。



圖 4.3 較亮情況的實驗結果

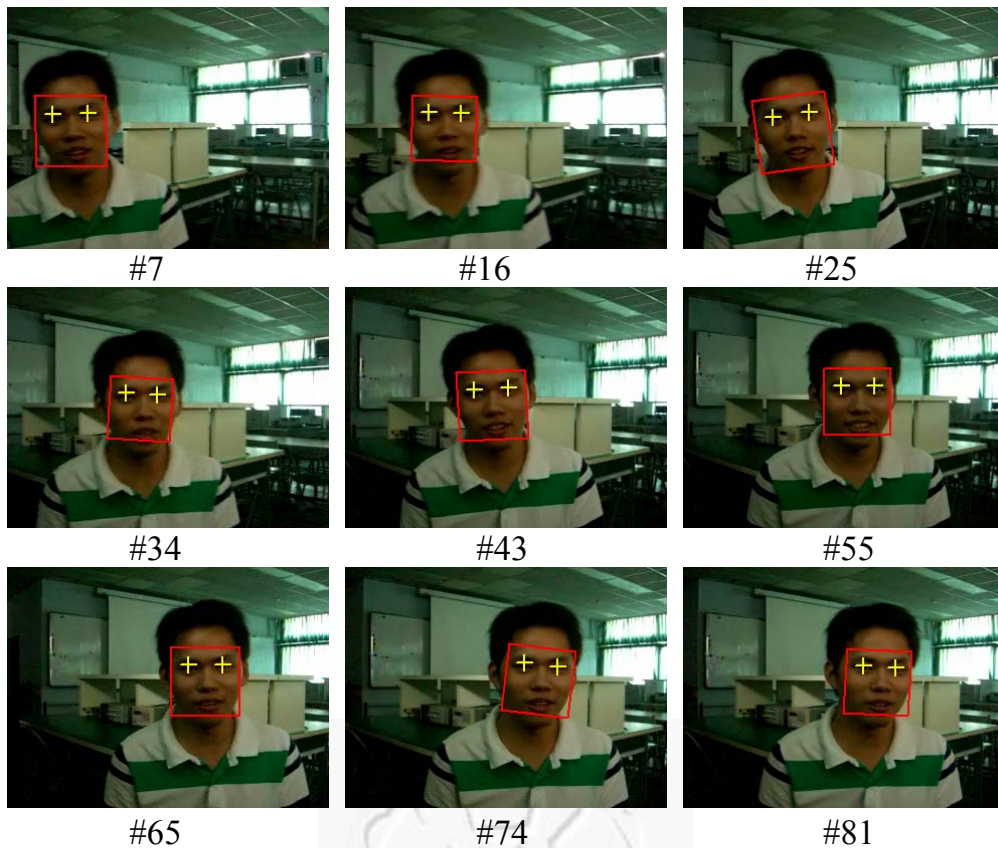


圖 4.4 較暗情況的實驗結果

第三個實驗中我們針對不同的情況偵測人臉，第一個情況是針對不同角度的人臉，因為我們是根據臉部特徵來偵測人臉所以我們可以偵測的角度在45度到-45度之間，在這個角度範圍兩眼的特徵都還存在所以我們的系統可以偵測，當大於這個角度眼睛特徵就會不見所以無法偵測，所以我們只討論45度到-45度的人臉偵測率，第二個情況我們實驗者為戴眼鏡的人，這個實驗是考驗我們系統對於偵測眼睛特徵的強健性，在實驗中因為我們是使用影像增強的方法再用灰階值分割眼睛特徵，所以戴有色鏡片會影響灰階值所以在這項實驗我們的實驗者都不會戴有色鏡片。第三個情況我們多人的偵測，偵測人數為2~4人，在實驗的影像我們一樣對每一種情況分別取五個實驗者個別錄製10秒鐘，每一個情況分別得到150張影像，第三個情況是分別做2人、3人、4人分別10秒鐘然後我們分別取30張影像共90張，90張影像

中共有 270 個人臉，下表 4.4 是我們實驗結果。

表 4.4 其他其況下的人臉偵測率

實驗項目	不同人臉角度	戴眼鏡的實驗	多人偵測
人臉數目	150	150	270
正確人臉	128	130	226
誤判人臉	22	20	44
正確率	85.33%	86.66%	83.70%
FPS	10	11	8

在第一個情況中因為有時候耳朵會呈現粉紅色，所以在 45 度和-45 度的時候會把耳朵誤測為嘴唇的顏色，這樣根據我們的幾何判斷就會把不是為正確的眼睛判定為正確的眼睛，造成人臉偵測的錯誤，以及眼睛太靠近邊緣會造成眼睛的特徵被當作背景使得眼部特徵遺失，使得我們系統無法偵測到。第二種情況對於眼鏡被強光打到造成反光，這樣使得眼睛的特徵點遺失，我們的系統就無法做正確的偵測，和對於戴黑色粗框的實驗者在膚色偵測的時候因為黑色不屬於膚色然後又是粗框，這樣會使得一張人臉被分成兩個部份，這樣在膚色的地方就先出錯之後臉部特徵自然無法偵測，就會造成人臉偵測錯誤，第三個情況中因為是多人偵測有時會有人臉來在一起或是遮蔽的現象產生，會造成特徵點的遺失，就會有錯誤的偵測。除了以上的問題會造成誤測，基本上這三個情況的偵測率是依據人臉的大小，除了第三個情況多人偵測不易控制人臉的大小，而我們在其他兩個狀況的實驗都是以 75x75 的人臉大小來做實驗。因為我們的實驗者在人臉的大小都控制在 75x75 的大小所以在速度上大多相同，只有在多人偵測時的連通成份數量增加所以速度有明顯的下降。下圖 4.5 為頭由左轉到右的實驗結果，圖 4.6 是帶眼鏡的實驗結果，圖 4.7 是多人的實驗結果。



圖 4.5 頭由左轉到右的實驗結果



圖 4.6 帶眼鏡的實驗結果



圖 4.7 多人的實驗結果

在第四個實驗中我們從網路上下載178張包含216人的影像，這些影像為 CCD所擷取的影像，包含了不同的環境條件。有各種不同的照明、不同的背景，有不同膚色的、戴眼鏡的、有正面的、有頭轉動小於 30 度的、有胖的、有瘦的人臉。我們使用這些影像用文獻[26,27]這兩篇的方法跟我們的方法做比較。Froba與Kublbeck [26]使用 edge orientation matching 的方法快速地偵測人臉。其 frame rate 在 P4 2.4G 的個人電腦偵測 320×240 的影像可達到 20 fps。在偵測率的表現雖然有很高的偵測率，但同樣地也有很多的錯誤偵測。這裡的錯誤偵測是指在沒有人臉的地方卻判定為人臉。此系統錯誤偵測太多，對於安全需求較高的系統比較不適合。網站上可以下載 demo 的執行檔。BuFaLo (Biometric base Unit for Face Localization) 系統實作了 Viola 與 Jones [27] 的方法。BuFaLo 不但有高偵測率、低錯誤偵測，而且速度上也相當

快。在P4 2.4G的個人電腦上，偵測320×240的影像，其frame rate可達6 fps。
(上面兩篇執行檔的網址在其他參考資料中)

表4.5 不同系統的比較

實驗項目	Froba 與 Kublbeck [26]	BuFaLo	我們的方法
人臉數目	216	216	216
正確人臉	205	192	187
錯誤偵測	115	8	11
正確率	94.91%	88.89%	86.57%
FPS	20	6	11

在上表中可以發現Froba與Kublbeck [26]的方法雖然有很高的偵測率，但是他同時也會在背景中有許多的錯誤，在樣在一個影像中只有一個人臉但他們的方法會偵測出很多張人臉，這樣不利於之後做人臉辨識的部份或是門禁系統。而BuFaLo這個系統他們也有不錯的偵測率不過他並沒有偵測出臉部特徵這樣無法做接下來的表情辨識，而相較之下我們的方法執行的速度快了將近兩倍更適合用在動態的影像偵測。我們的系統仍有很高的偵測率與少量的錯誤偵測。故在監控系統中，即使外在環境改變有太亮或是太暗，人臉有各種表情或是人們頭稍微傾斜等。

第五章 結論

本文提供一個簡單又有效的方法偵測人臉，利用人臉器官的比例在不同情況下可以自適性的調整閾值，先使用統計不同亮度的膚色範圍的方法找出膚色的範圍，再用自適性的閾值去分割人臉特徵，最後再用幾何判斷的方法找出適當的眼睛嘴唇定位出人臉。用我們的方法在有強光的情況用可以找出不同亮度的膚色再用自適性的閾值去針對不同亮度的眼睛和嘴唇給予不同的閾值然後用幾何判斷定義出正確的眼睛嘴唇，在手臂以及腳或是類似膚色的背景中雖然是膚色的區域但是他們不會擁有臉部的特徵，所以系統不會找出臉部特徵，也可以正確的判定不是人臉，在複雜背景的情況下，因為我們是使用色彩分析，而不是使用邊緣偵測的方法，所以可以分割出正確的膚色區域，即使背景有膚色的情況，我們也可以在下面的步驟判斷沒有臉部特徵，進而判斷為非人臉區域。在速度方面因為我們使用自適性的閾值來尋找眼睛和嘴唇，在決定閾值的時候會一直跑迴圈造成速度的下降，我們使用自適性的閾值來偵測人臉 FPS 大約是 10~11，假如使用固定閾值來偵測人臉特徵的 FPS 大約 12~13，可以發現我們的速度慢大約兩個 frame，在動態偵測的條件下還是可以接受每秒處理 10 張畫面的速度，不過在不同亮度的情況中使用固定閾值會造成無法偵測的情況，相較之下可以發現我們的方法雖然比較慢但是還在可以接受的範圍內，不過對於不同情況的偵測效果卻比使用固定閾值的方法高出許多。

在我們實驗測試結束後，我們歸納及分析了我們的方法。發現了此演算法之優點及缺點，以下將一一加以列舉。

優點：

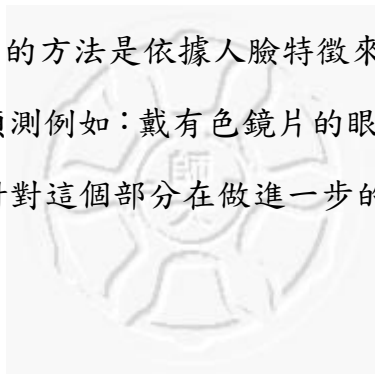
1. 方法簡單易於實作，且所耗成本低。
2. 能抓取出人臉特徵之位置，這裡所指的是眼睛及嘴巴。
3. 即使嘴巴形狀特別亦可偵測出來，如張大嘴巴或歪嘴等。

4. 可用於複雜的背景情況之中。
5. 可用於不同的光線情況之中。
6. 可以偵測戴眼鏡的人臉。
5. 可用於單人或多人偵測上。

缺點：

1. 當眼鏡鏡片受強光反射，將造成偵測錯誤。亦或是配戴過粗之暗色系鏡框，也將造成偵測錯誤，因為在膚色判斷上，將會把人臉從眼睛地方切為兩半。
2. 被測者頭部擺動過大，將造成偵測錯誤。

我們的方法在多人的情況下，也有良好的偵測率。未來可以繼續再做人臉識別的部分，不過我們的方法是依據人臉特徵來定位人臉，所以在某些狀況下我們無法做有效得偵測例如：戴有色鏡片的眼鏡或是墨鏡以及戴口罩的人，所以未來我們必須針對這個部分在做進一步的探討。



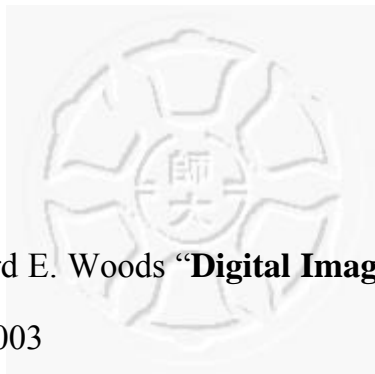
參考資料

- [1] Ming-Hsuan Yang, Kriegman, D.J. and Ahuja N., “Detecting faces in images: a survey,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.24, NO. 1, pp. 34-58, Jan.-2002.
- [2] Karin Sobottka, Ioannis Pitas,”Extraction of Facial Regions and Features Using Color and Shape Information”, *IEEE, Proceedings of ICPR ‘96*
- [3] Nara Yusuke, Jianming Yang, and Yoshikazu Suematsu, “Face detection using the shape of face with both color and edge” *Cybernetics and Intelligent Systems, 2004 IEEE Conference on*, Vol.1, pp. 1-3 Dec-2004.
- [4] S. Birchfield, “An elliptical head tracker,” *Proc. Conference Record of the Thirty-First Asilomar Conference on Signals, Systems & Computers*, pp. 1710-1714, 1997.
- [5] Jianping Fan, David K.Y. Yau, Elmagarmid, A.K., Aref, Walid G., “Automatic image segmentation by integrating color-edge extraction and seeded region growing,” *IEEE Transactions on Image Processing*, Vol.10, NO 10, pp.1454 - 1466, Oct-2001.
- [6] Kwok-Wai Wong, Kin-Man Lam and Wan-Chi Siu, “A Robust Scheme for Live Detection of Human Face in Color Images,” *Signal Processing: Image Communication, The Netherlands*, Vol. 18, No.2, pp.103-114, February-2003.
- [7] Rein-Lien Hus, and Anil K. Jain “Face detection in color images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.24, NO.5, pp. 696-706, May-2002.
- [8] Yanjiant Wang, and Baozong Yuan, “A novel approach for human face detection from color images under complex background,” *Pattern Recognition*, Vol. 34, NO.3, pp. 1983-1992, February-2000.

- [9] Chiunhsiun Lin, Kuo-Chin Fan, "Human face detection using geometric triangle relationship", *Pattern Recognition, 2000. Proceedings. 15th International Conference*, vol.2, pp.941 - 944 Sept-2000.
- [10] S.H. Kim and H.G. Kim, "Face Detection using Multi-model Information," *Proc. Fourth IEEE International Conference Automatic Face and Gesture Recognition*, pp.14-19,2000.
- [11] Alessandro Rizzi, Carlo Gatta, Daniele Marini, "A new algorithm for unsupervised global and local color correction", *Pattern Recognition Letters* 24, pp.1663-1677,2003.
- [12] S. Karungaru, M. Fukumi, and Akamatsu, "Feature extraction for face detection and recognition" *Robot and Human Interactive Communication, 2004. ROMAN 2004. 13th IEEE International Workshop on*, pp. 20-22, Sept-2004.
- [13] Chiunhsiun Lin, Kuo-Chin Fan, "Triangle-based approach to the detection of human face," *Pattern Recognition* Vol.34, NO.6, pp.1271-1284,2001.
- [14] C.H. Morimoto and M. Flickner, "Real-time Face Detection Using Active Illumination," *Proceedings. Fourth IEEE International Conference*, pp. 8-13, 28-30, March-2000.
- [15] C.C. Han, H.Y. Liao, G.J. Yu, and L.H. Chen, "Fast Face Detection via Morphology-based Pre-processing," *Pattern Recognition*, vol.33, pp. 1707-1712, 2000.
- [16] Heng Liu, Shengye Yan, Xilin Chen, and Wen Gao, "Rotated face detection in color images using radial template (RT)" *Multimedia and Expo, 2003. ICME '03. Proceedings. 2003 International Conference on*, Vol.3, pp. 6-9 July-2003.

- [17] T. Sakai, M. Nagao, and S. Fujibayashi, "Line Extraction and Pattern Detection in a Photograph," *Pattern Recognition*, vol. 1, pp. 233-248, 1969.
- [18] Haiyuan Wu, Qian Chen, and Masahiko Yachida, "Face Detection From Color Images Using a Fuzzy Pattern Matching Method", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, NO. 6, JUNE-1999.
- [19] Anderson K. and McOwan P.W., "A real-time automated system for the recognition of human facial expressions," *Systems, Man and Cybernetics, Part B, IEEE Transactions on*, vol. 36, pp. 96-105, Feb-2006.
- [20] Henry A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection" *Computer Vision and Pattern Recognition, 1998. Proceedings. 1998 IEEE Computer Society Conference on*, pp. 18-20, June-1996.
- [21] Henry A. Rowley, S. Baluja, and T. Kanade, "Rotation invariant neural network-based face detection" *Computer Vision and Pattern Recognition, 1998. Proceedings. 1998 IEEE Computer Society Conference on*, pp. 23-25, June-1998.
- [22] Hyungkeun Jee, Kyunghee Lee, and Sungbum Pan, "Eye and face detection using SVM" *Intelligent Sensors, Sensor Networks and Information Processing Conference, 2004. Proceedings of the 2004*, pp. 14-17, Dec-2004.
- [23] Waring C.A., Xiuwen Liu, "Face detection using spectral histograms and SVMs," *Systems, Man and Cybernetics, Part B, IEEE Transactions*, Volume 35, Issue 3, pp.467-476, June-2005.
- [24] Rajagopalan A.N, Chellappa R, Koterba NT, "Background learning for robust face recognition with PCA in the presence of clutter", *Image Processing IEEE Transactions*, Vol. 14, Issue 6, pp.832-843, June-2005.

- [25] K.-K. Sung and T. Poggio, “Example-based learning for view-based human face detection,” *IEEE Trans. on Pattern Analysis and Machine Intelligence* Vol. 20, 39-51, 1998.
- [26] B. Froba and C. Kublbeck, “Face Detection and Tracking using Edge Orientation Information,” *In SPIE Visual Communications and Image Processing*, pp. 583-594, Jan 2001
- [27] P. Viola and M. Jones, “Rapid Object Detection using a Boosted Cascade of Simple Features,” *IEEE Conf. on Computer Vision and Pattern Recognition*, 2001.



其他參考資料

Rafael C. Gonzalez, Richard E. Woods “**Digital Image Processing**” 台灣培生教育出版股份有限公司，2003

COULEUR.ORG，<http://www.couleur.org/>。

人體計測資料庫，<http://www.iosh.gov.tw/ergo.htm>。

文獻[26]的執行檔，<http://www.iis.fhg.de/bv/biometrie/download/index.html>。

文獻[27]的執行檔，<http://www.geocities.com/fritzfra2001/>。