

臉部特徵定位與擷取之研究

A Study for Facial Feature Localization and Extraction

研究生 ：郭晉源

指導教授 ：蘇義明 博士

義守大學

電子工程學系

碩士班碩士論文

A Thesis Submitted to
Department of Electronic Engineering
I-Shou University
in
Partial Fulfillment of the Requirements
for the Master Degree
with a
Major in Electronic Engineering
July, 2010
Kaohsiung, Taiwan
Republic of China

中華民國 九十九 年 七 月

義守大學
電子工程學系碩士班
論文口試委員會審定書

本校 電子工程 研究所 郭晉源 君
所提論文 臉部特徵定位與擷取之研究

A Study for Facial Feature Localization and Extraction

合於碩士資格水準，業經本委員會評審認可。

口試委員：洪金車 _____
陳良 _____
蘇義明 _____

指導教授：蘇義明 _____

研究所所長：黃榮生 _____

中華民國九十九年七月六日

義守大學數位論文系統

臉部特徵定位與擷取之研究

研究生：郭晉源 指導教授：蘇義明

義守大學電子工程研究所

摘要

本論文主要建立一個可以正確定位人臉以及臉部特徵擷取，此演算法包含有人臉偵測、嘴唇偵測、膚色檢測、眼睛定位、唇形輪廓、唇型特徵等多個步驟所組成。首先從影片中擷取，含有複雜背景人臉圖片，利用膚色的檢測，可將圖像中的不相關背景物給去除，人臉偵測的部分，則使用矩形特徵 (Rectangle Feature) 和配合自適性布斯特 (AdaBoost) 分類法，定位出人臉的正確位置，藉由人臉位置的確認，在使用臉部的特徵幾何關係，依序從上而下，可畫分出眼睛的候選區域和嘴唇的候選區域，在眼睛候選區域的選定，使用模板比對 (Template Matching) 的方式，取得眼睛位置後，進而縮小嘴唇候選區的範圍，達到更好效果，在嘴唇的選取時，採用邊緣資訊的投影和嘴唇與膚色的差異性，定位出嘴唇的最佳位置；由實驗結果顯示利用膚色檢測的方式，去除背景中不含膚色的區域，不僅可以減少人臉偵測時的運算，對於在複雜的環境有很高的容錯能力。實驗中使用自行錄製短片，對臉部特徵定位與擷取進行測試的動作。在系統對於影片的處理速度可以達到每秒 15 張，整體的偵測率可達到 75%。

關鍵字: 臉部偵測，眼睛偵測，嘴唇偵測，模板比對

A Study for Facial Feature Localization and Extraction

Student : Chin-Yuan Kuo Advisor : Dr. Yih-Ming Su

**Department of Electronic Engineering
I-Shou University**

ABSTRACT

This research is mainly to develop a system for face detection and facial feature extraction. The algorithm consists of face detection, mouth detection, skin color detection and eye detection. Firstly, a face image with complex background is extracted to use skin color detection and remove the background. Then, rectangle feature extraction and AdaBoost classification approaches are used to localize the face region. Furthermore, the eye and lip candidate region are divided using facial geometrize relationship. In the eye candidate region, a template matching approach is used to obtain the position, and then the range of the lip candidate region is narrowed to achieve better result. The best lip region is extracted using edge projection and color difference between the lip and skin color after detecting lip candidate region. The result shows that the use of skin color detection can not only reduce the computation time for face detection, but also have high level of fault tolerant in the complex environment. The experiment results used self-made videos to do the test of facial feature position and selection. The system performed 15 frames/s, and also achieved the overall detection rate of 75 %.

Keywords : face detection, eye detection, lip detection, template matching

義守大學數位論文系統

致謝

首先誠摯的感謝指導教授蘇義明博士，老師悉心的教導使 我得以一窺電腦視覺領域的深奧，不時的討論並指點我正確的方向，使我在這些年中獲益匪淺。老師對學問的嚴謹更是我輩學習的典範。

在研究所這段時間，實驗室裡共同的生活點滴，學術上的討論、言不及義的閒扯、讓人又愛又怕的宵夜、趕作業的革命情感、因為睡太晚而遮遮掩掩閃進實驗室...，感謝眾位學長、學弟的 共同砥礪，你們的陪伴讓我的研究生生活變得絢麗多彩。感謝望秦、冠彰學長們不厭其煩的指出我研究中的缺失，且總能在我迷惘時為我解惑，也感謝同學的幫忙。實驗室的忠延、朝暉、殿凱學弟們當然也不能忘記，你們 的幫忙及搞笑我銘感在心。最後，謹以此文獻給我摯愛的雙親。

目錄	I
中文摘要	II
英文摘要	II
致謝	II
目錄	IV
圖目錄	vi
表目錄	vii
第一章 緒論	1
1-1 研究動機	1
1-2 臉部特徵概述	2
1-3 論文架構	6
第二章 相關文獻討論	7
2-1 色彩空間選擇	7
2-1-1 正規化RGB	8
2-1-2 YUV色彩空間	9
2-1-3 YCbCr色彩空間	10
2-1-4 HSV色彩空間	10
2-2 膚色檢測	12
2-2-1 Normalized RGB 膚色分析	12
2-2-2 YCbCr 膚色分析	13
2-3 人臉偵測	20
2-3-1 積分影像	21
2-3-2 矩形特徵	22
2-3-3 自適性布斯特演算法	23
2-3-4 串列式模型	25
第三章 臉部特徵定位	26
3-1 人臉偵測模組	26
3-1-1 影像前處理	26
3-1-2 膚色檢測	28
3-1-3 人臉偵測	29
3-2 眼睛模組	30
3-2-1 眼睛候選區域切割	30
3-2-2 眼睛偵測	31
3-3 嘴唇模組	32
3-3-1 嘴唇候選區域	32
3-3-2 嘴唇的擷取	34
第四章 實驗結果與討論	36

4-1 系統環境.....	36
4-2 實驗介面.....	38
4-3 實驗結果.....	39
4-3-1 人臉偵測率	39
4-3-2 眼睛特徵偵測率	42
4-3-3 嘴唇特徵偵測率	43
4-4 結果討論.....	44
第五章 結論	45
5-1 結論.....	45
5-2 未來研究方向	45
參考文獻	46



ISU Etheses System

義守大學數位論文系統

圖目錄

圖 1-1 系統流程圖	5
圖 2-1 RGB色彩空間圖	8
圖 2-2 HSV色彩空間	11
圖 2-3 膚色分布	13
圖 2-4 訓練人臉資料庫	14
圖 2-5 膚色在YCbCr色彩空間的分布情形	14
圖 2-6 膚色投影在Cb、Cr的空間	15
圖 2-7 Cb在Y的變化	15
圖 2-8 Cr在Y的變化	16
圖 2-9 YCb'Cr' 膚色的分布情形	18
圖 2-10 投影在Cb'Cr'的膚色模型	18
圖 2-11 矩形特徵	22
圖 2-12 矩形特徵對應在臉部情形	23
圖 2-13 串列模型圖	25
圖 3-1 高斯分布圖	27
圖 3-2 高斯遮罩	27
圖 3-3 平滑處理	28
圖 3-4 膚色檢測	29
圖 3-5 人臉偵測	29
圖 3-6 分割臉部	30
圖 3-7 眼睛候選區域	31
圖 3-8 眼睛模板	32
圖 3-9 眼睛偵測	32
圖 3-10 邊緣偵測	33
圖 3-11 邊緣投影	34
圖 3-12 嘴唇與膚色的分割	35
圖 4-1 系統介面圖	38
圖 4-2 正確臉部區域	41
圖 4-3 偵測錯誤人臉區域	41

義守大學數位論文系統

表目錄

表 4-1 影片資料庫	37
表 4-2 人臉偵測率	41
表 4-3 眼睛偵測率	42
表 4-4 嘴唇偵測率	43



ISU Etheses System

第一章 緒論

1-1 研究動機

在科技越來越發達的時代，電腦系統處理速度也不斷的提昇之下，不論是在休閒娛樂、車用電子和研究實驗上，電腦與日常生活已經是密不可分。雖然電腦可以有效完成複雜性高的和危險性高的工作，但是電腦卻沒有獨自思考與自我判斷的能力，只能依靠程式設計師的命令去完成工作，如何使電腦富有人性化以及自我思考能力，是目前科學家努力積極研究的項目之一，希望能在未來不久的日子，能像電影情節裡一樣，機器人具有自我思考的能力，甚至能夠進步到，了解人類的情感。

人類是透過五官以及身體某些部位的感知功能去搜集外界的資訊，再經由大腦解析之後，進一步的去判斷與分析，來了解環境中的變化，進而判斷下一部動作，現今電腦是無法正確的模擬人類的想法與思維，對於電腦取得影像的方式，得透過一些感測元件，去搜集外界的資訊，資訊的來源可以是多方面，可以從音源、溫度、影像等等，這些資料來源都有各自的專業處理的領域。在電腦視覺（Computer Vision）領域中，取得外界資訊的工具就是攝影機和照相機等一些取像元件，透過這些元件取得色彩資訊來進行分析。其中以「人臉偵測」為熱門的題目，因為人類的情感都是臉部的變化，去表達出情感。然而五官的定位額外顯得重要，在眼睛、鼻子、耳朵、嘴唇和眉毛裡，嘴唇的變化的資訊是最多。在非固定背景環境中，如何正確定位出臉部資訊，對電腦來說是一件困難的事情。如何建構出一個穩定臉部特徵定位與擷取系統，做為日後臉部分析使用，透過人機介面的建立，讓操作起來更佳的容易，成為本論文研究的目的。

義守大學數位論文系統

1-2 臉部特徵概述

電腦視覺一直以來都是人們努力積極研究的課題，人們希望透過攝影機所取得影像，傳送到電腦處理，將電腦端分析出的數據，去做其他相關的應用，例如身分認定、人臉辨識和保全監控等方面應用。這些影像的來源，可分成兩種形式，靜態影像 (Static Image) 和連續影像 (Sequence Images)。則在處理連續影像時，需考慮的硬體的處理速度，是否能夠及時的處理，對臉部特徵的追蹤系統是很大影響因素之一。做臉部特徵定位時，首要的工作，就是如何在取得的影像中，正確又快速定位人臉的資訊，才可以針對臉部特徵的部分去做探討。在人臉偵測的部份，有以下幾種人臉偵測方法，主要成份分析 PCA (Principal Component Analysis) [1][2]、膚色分析 (Skin Color Analysis) [3]-[9]、模板比對 (Template Matching) [11]、矩形特徵 (Rectangle Feature) [11]-[14]、類神經網路 (Neural Network) [15]。

其中主要成份分析 PCA 是一種降低特徵維度的方法。一張圖片(例如：人臉)的原始特徵維度就是這張圖片的長 X 寬，如果圖片是 1024x768 大小，就會有 786432 個特徵值，其特徵維度就是 786432。70 多萬的特徵維度是相當龐大的，會影響圖片辨識的效率，而且這 70 多萬的特徵值可能包含了大量的雜訊，會影響整個圖形辨識的準確率，對於辨識的效率(速度)與準確率都有極大的影響。

在膚色分析上，一般容易被人們接受色彩空間，就是 RGB 色彩空間，是由三原色所組成，經由色彩的調配，可以產生出無數種不同的顏色。不同的生物眼中辨別顏色的細胞並不相同，例如鳥類眼中有四種分別對不同波長光線敏感的細胞，而一般哺乳動物只有兩種，所以對它們來說只有兩種

原色光。再做膚色分析的不單只是RGB色彩空間而已，首先介紹HSV色彩空間，它是由色相（Hue），飽和度（Saturation），明度（Value）三個所組成的。HSV 是RGB色彩空間的一種變形，它的內容與色彩尺度與其出處RGB色彩空間有密切聯繫。HSL(色相（Hue），飽和度（Saturation），亮度（Lightness/Luminance），也稱HLS 或在 HIS色彩空間中 I指(Intensity)與 HSV非常相似，僅用亮度（Lightness）替代了明度（Brightness）。二者區別在於，一種純色的明度等於白色的明度，而純色的亮度等於中度灰的亮度。另外一種色種空間為YCbCr，這種色彩空間最常使用數位電視，而之前提到的RGB色彩空間是最常使用在電腦螢幕顯示上。YCbCr色彩空間主要成份是Y亮度（Luminance）和Cb（Blueness）、Cr（Redness），在作膚色分析時，是不考慮Y分量的變化，只針對Cb和Cr為參考的依據。膚色可以去除掉影像中許多不雜訊，來增加電腦執行速度，但只要人種顏色差異太大的話，則會造成電腦的誤判，單一用膚色的分析是不夠，則需額外增加其他方法來輔助，進而達到更好的效果。

接下來介紹模板比對（template-matching）[11]是最簡單的一種想法。這種想法認為，外界所有的事物，祇要是我們曾經接觸過的，都會在腦子裡留下它們的「烙印」或「模板」，因此我們的腦子裡充滿這類模板。當外界出現一個圖形呈現在我們眼前時，它的形狀與腦子裡儲存的模板一一的比對。應用再電腦上也一樣，將所要偵測目標物，建立一個具有代表性的物件，把所有代表性物件歸類，之後在儲存在資料庫。將所擷取到的數位影像，利用之前所建立的資料庫，一一呼叫出來，再有影像有關地方進行比對的動作。由於影像不只是物件，影像中可能包含前景（Foreground）和背景（Background）的資料，將背景的資料去除之後，經由一些影像處理的技巧，盡可能影像中的資料簡單化，這樣子在作模版比對時，才可減少一些不要的失誤。

近年來，在圖型識別(Pattern Recognition)相關研究的發展，使得在進行目標物偵測(Object Detection)的研究上，已經有許多成功的例子，例如在人臉偵測的領域中，Viola與Jones[12]提出利用矩形特徵(Rectangular Feature)搭配自適性布斯特(Adaboost)演算法，在複雜的背景中快速且正確擷取出人臉資訊。

在介紹完之前所的人臉偵測，接下來是要說明取得人臉資訊後，如何透過幾何的關係，迅速的找到臉部特徵的位置，在進行臉部特徵的定位，可以依照Faisal Shafait[16]所提出來人臉模型，去定義眼睛和嘴唇的位置，可提供嘴唇追蹤和眼睛追蹤使用。本論文旨在發展一套完整臉部特徵定位與擷取的系統，如何在連續的影片來源中，定位人臉的正確位置和擷取臉部出特徵資訊，主要的系統流程如圖 1-1所示。

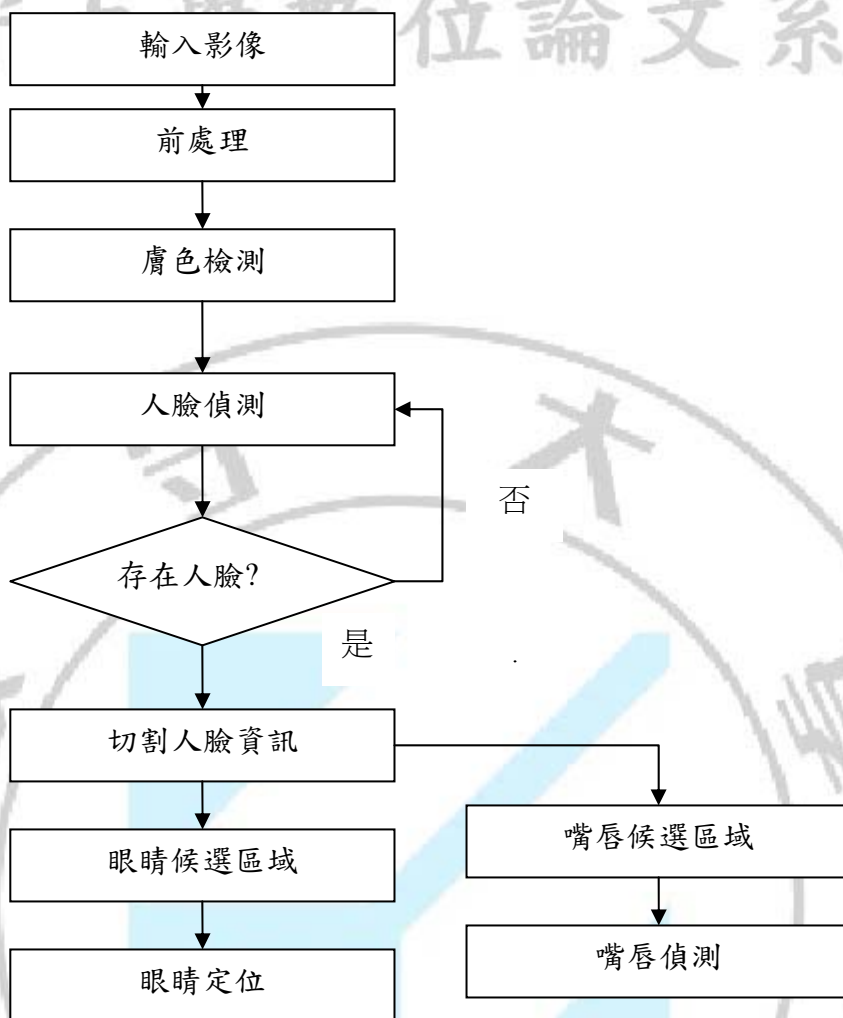


圖 1-1 系統流程圖

1-3 論文架構

本論文之各個章節如下，分為五部份來做詮釋：

第一章 緒論

說明為何探討臉部特徵，研究動機。

第二章 相關文獻討論

探討相關臉部特徵定位研究領域的方法與結果。

第三章 臉部特徵定位

說明本論文所使用方法架構，解釋研究過程中所採用的方法與原理。

第四章 實驗結果和討論

提出實驗的數據，實驗結果及探討，與他種研究方法的比較。

第五章 結論與未來研究方向

討論本論文的成果及不可預期的缺失，並提出未來研究改善方向。

第二章 相關文獻討論

依照影像來源，可以將數位影像或圖片的內容粗分成兩大類，前景（Foreground）和背景（Background）。如何將影像中的前景物件，正確的與背景分離，在影像處理中，是件看似簡單，但要做好並非容易的事情，拜現今的科學所賜，電腦處理的速度越來越快，在繁雜的演算法，透過電腦精密又快速的計算能力，將能夠解決所有的問題，但也藉由許多專學者的努力，提供新的方法結合新技術，使電腦視覺的更加人性化，對於臉部的特徵定位與擷取的方法，可以分成三個部分，色彩空間的選擇、膚色的分析和人臉偵測方法進行探討動作，以下將用三個小節分別對各個主要環節做一個簡略的介紹。

2-1 色彩空間選擇

在數位影像處理對色彩的表示，可以依照不同的參考值，而產生出各種不同的色彩空間，而在一般顯示影像所採取RGB色彩空間，RGB色彩空間是由三原色所組成的，因為RGB色彩空間對於光線比較敏感。由於影像的來源，常常受到光源強弱的影響，進而影響膚色檢測，而產生偵測錯誤的情形，所以在選擇色彩空間時，都會顧忌到亮度對膚色的影響。因此就有許多論文研究對於亮度比較不敏感的色彩空間來做為區隔，如正規化RGB、YUV、YCbCr、HSV等色彩空間[3]。下一個小節將會提到膚色偵測常用到色彩空間優缺點比較。

2-1-1 正規化 RGB

RGB色彩空間，R代表紅色，G代表綠色，B代表藍色，一般紅色、綠色、藍色的灰階度分別以1 Byte來表示，數值的範圍0到255，將三個顏色合成一張圖片時，圖像中的任何一點像素為3 Byte(24Bit)，也相當於 2^{24} 種顏色表示。該色彩空間可以用一個正立方體表示，如圖2-1所示座標值為(255, 0, 0)紅色，(0, 0, 255)藍色，(0, 255, 0)綠色，(255, 0, 255)黃色，(255, 0, 255)洋紅色，(0, 255, 255)黃色。在RGB色彩空間，會因光源強弱，會使RGB的質產生變化，所以RGB色彩空間是需要修正的，改善的方法，只要對RGB的值做正規化的處理，就可以改善光源的影響，其公式表示為：

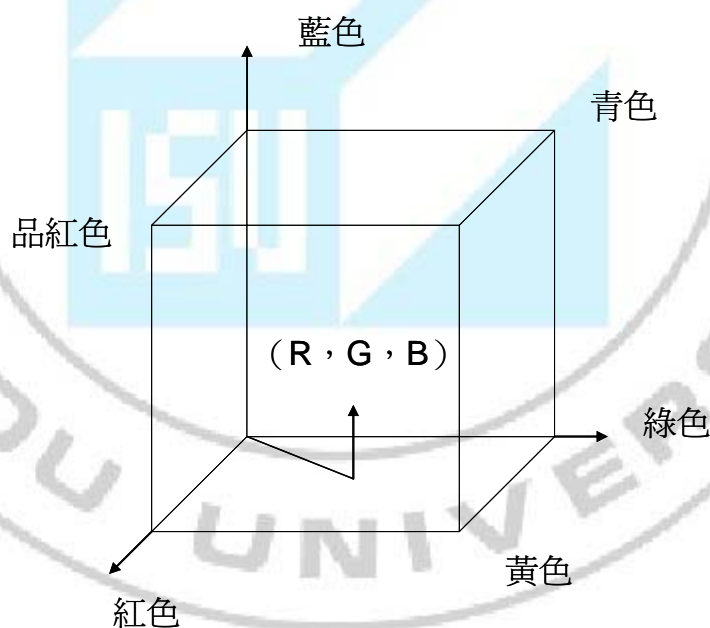


圖 2-1 RGB 色彩空間圖

$$r = \frac{R}{R + G + B} \quad (2-1)$$

$$g = \frac{G}{R + G + B} \quad (2-2)$$

$$b = \frac{B}{R + G + B} \quad (2-3)$$

由於b的值對光源的影響比較小，然而可以忽略。將計算出來的r與g的值，取代原先的值。

2-1-2 YUV 色彩空間

YUV 的發明是由於彩色電視與黑白電視的過渡時期。黑白視訊只有Y (Luma, Luminance) 視訊，也就是灰階值。到了彩色電視規格的制定，是以YUV 的格式來處理彩色電視圖像，把 UV 視作表示彩度的 C (Chrominance 或 Chroma)，如果忽略 C 訊號，那麼剩下的 Y (Luma) 訊號就跟之前的黑白電視訊號相同，這樣一來便解決彩色電視機與黑白電視機的相容問題。其轉換公式如下：

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & -0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2-4)$$

2-1-3 YCbCr 色彩空間

YCbCr 色彩空間從 YUV 色彩空間衍生而來的，YCbCr 色彩空間最常使用在數位視訊中和 JPEG 影像，其中 Y 表示明亮程度，Cr 與 Cb 表示色差的資訊，Cb 成分是藍色與參考值的差距，Cr 成分是紅色與參考值的差距。RGB 和 YCbCr 的關係式如下：事實上，人類的眼睛對色彩的改變比較遲鈍，但對亮度的改變卻是敏感。而在 JPEG 影像的應用上，一般都是只有處理灰階和全彩的影像，全彩的影像是由 Y、Cb、Cr 三個顏色成份（Component）所組成，而灰階的影像由於只有亮度而無色彩可言，因此只有 Y 一個成分。

$$\begin{bmatrix} Y \\ Cr \\ Cb \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \quad (2-5)$$

2-1-4 HSV 色彩空間

HSV 色彩模型和 RGB 色彩空間最大的不同之處是 HSV 模型將亮度 V（Value）變數從色彩變數中抽離出來，也就是說色彩歸色彩，並不帶有亮度的成份，因此混色時不會有加減亮度的情形發生。至於色彩的部份則又分為色彩的種類（Hue，色相）和色彩的飽合度（Saturation）兩個變數來描述色彩。色度 H 是表示與垂直軸的角度，所以圖 2-2 中六角錐中每個頂點的間距夾角是 60 度，因此，紅色在 0 度位置上，黃色在 60 度位置上，綠色在 120 度位置，青色在 180 度位置，藍色在 240 度位置，洋紅色在 300 度的位置。飽和度 S 指的是顏色的濃度，它是從六邊形的中心向邊緣逐漸增

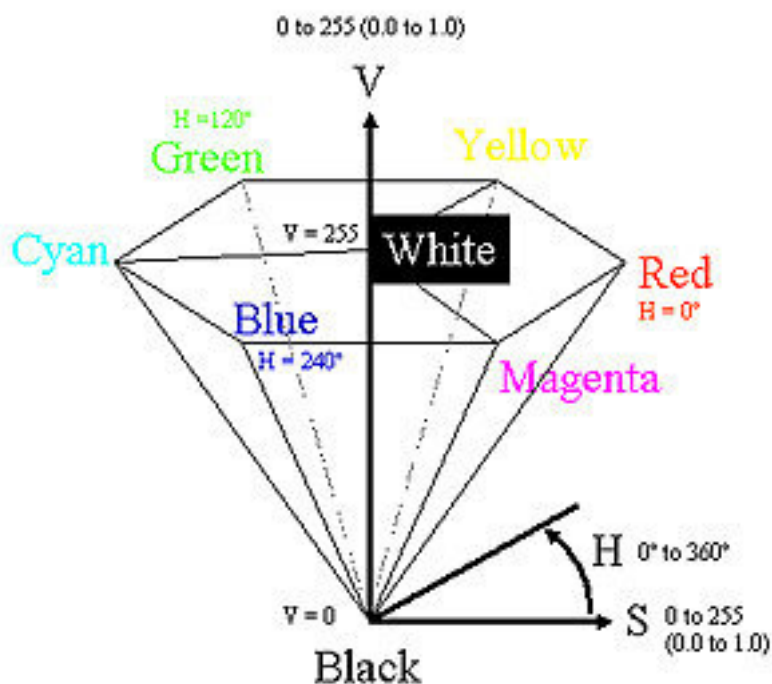


圖 2-2 HSV 色彩空間

加其值，它的值從0 到1 變化代表所選顏色的色飽合度到達最大值，所以當S=0 時，代表它位於六角錐中心點，其色澤是灰階顏色。亮度V指的是顏色的明亮度，它的值從六角錐的尾端0 到頂端1 變化，當V=0 位於底部時代表是黑色，而當V=1 時位在頂端時代表著它有最大的亮度飽合度。當V=1 和S=1 時，代表最純淨的色度（Pure Hues），而白色是V=1 和S=0。HSV的色彩空間是從RGB色彩空間加以轉換，轉換的計算公式如下：

(1) H轉換公式

$$H = \begin{cases} \theta, & \text{if } B \leq G \\ 360^\circ - \theta, & \text{if } B > G \end{cases} \quad (2-6)$$

$$\theta = \cos^{-1} \left\{ \frac{0.5[(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \right\} \quad (2-7)$$

(2) S轉換公式

$$S = \frac{\text{Max}(R, G, B) - \text{Min}(R, G, B)}{\text{Max}(R, G, B)} \quad (2-8)$$

(3) V轉換公式

$$V = \text{Max}(R, G, B) \quad (2-9)$$

2-2 膚色檢測

影響人的膚色的因素有很多種，例如：不同人種膚色的也會所差異或在不同外在環境下的影響，最主要影響的因素是光源，在強光源下，膚色所呈現的會接近白色，在弱光源底下，則膚色會成暗色，這些因素假使都不排除的話，作膚色偵測時將會產生不同的變化。膚色偵測主要的工作是如何從複雜背景環境中找出正確膚色的位置，在比較不同種族的膚色時，可以發現在色彩空間中，不管任何種族膚色會有群聚的效果。下面介紹膚色偵測的方法應用在YCbCr色彩空間和RGB正規劃色彩空間。

2-2-1 Normalized RGB 膚色分析

Normalized RGB 色彩空間是解決傳統 RGB 色彩空間在光源差異度的問題，由 Soriano[6]所提 Normalized RGB 色彩空間，在膚色偵測有很大的幫助，Soriano 所使用 CCD 與 USB web camera 在他設定的十六種不同情況下，當對白種人與黃種人的膚色在 Normalized RGB 色彩空間分布情形。當中 X 軸代表正規化紅色，其分布範圍 0.2 至 0.6，Y 軸代表正規化綠色，其分布範圍 0.2 至 0.4 定義出兩條二次多項式 (Quadratic Funtion)。由此可知，

膚色會有群聚的特性。膚色分布的範圍，可以由兩條的非性線方程式（Nonlinear Function）定義出上、下的邊界。落在這兩條方程式的區域範圍內，認定為膚色的像素（Pixel）。

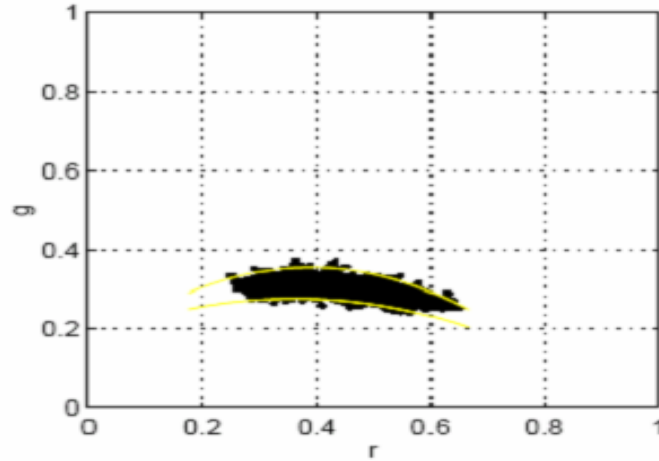


圖 2-3 膚色分布

$$S = \begin{cases} 1 & \text{if } (g < Q_+) \& (g > Q_-) \& (W > 0.004) \\ 0 & \text{otherwise} \end{cases}$$

其中

$$Q_+ = -13767 r^2 + 1.0743 r + 0.1452$$

$$Q_- = -0.776 r^2 + 0.5601 r + 0.1766$$

$$W = (r - 0.33)^2 + (g - 0.33)^2, 0.2 \leq r \leq 0.6$$

(2-10)

2-2-2YCbCr 膚色分析

由2002年Anil K. Jain所提出YCbCr膚色分析[9]，建立一個適合用於YCbCr的膚色模型，訓練資料來源HHI的人臉資料庫，資料庫的內容有多種角度的拍攝及不同光源的下圖形，如圖 2-4所示在資料庫裡，選取出137張圖片，在用人工的方式點選出853571的膚色像素點，將選取的像素投影在YCbCr的色彩空間，如圖2-5所示：



圖 2-4 訓練人臉資料庫

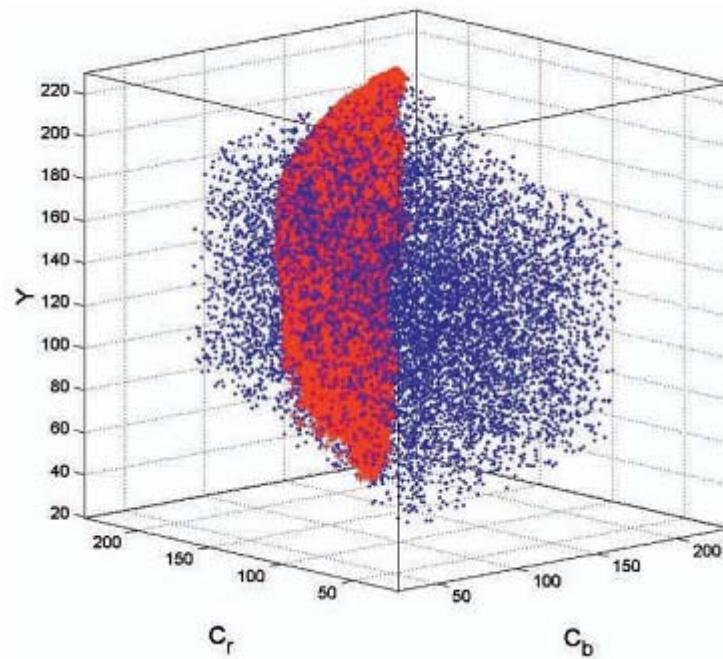


圖 2-5 膚色在 YCbCr 色彩空間的分布情形

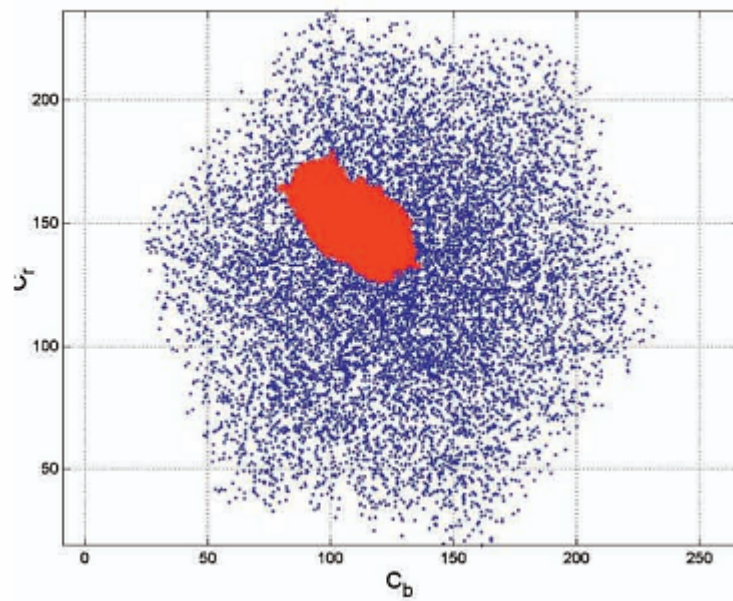


圖 2-6 膚色投影在 C_b 、 C_r 的空間

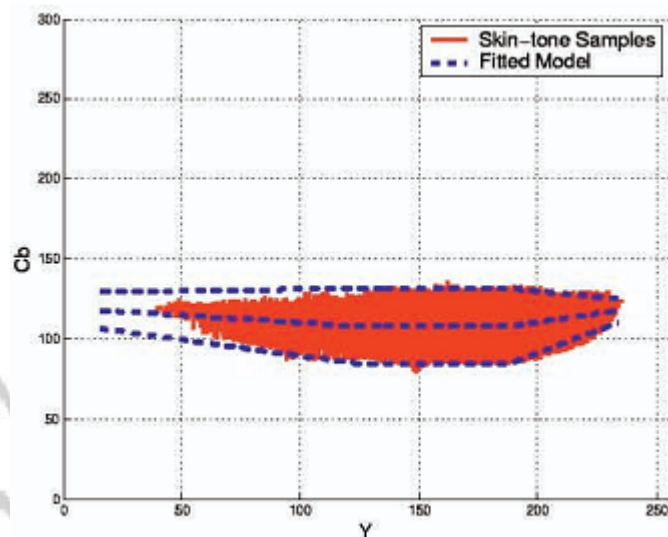


圖 2-7 C_b 在 Y 的變化

從圖2-5可得知在YCbCr色彩空間，Y值的部份在最頂端和最底端呈現縮小的區域，代表Y的大小會影響到膚色分布範圍。由圖2-6得知只要任意取個Y的值投影的情形將會有所改變。依照傳統YCbCr的方式，在作膚色偵測時，是會造某程度上的誤差。

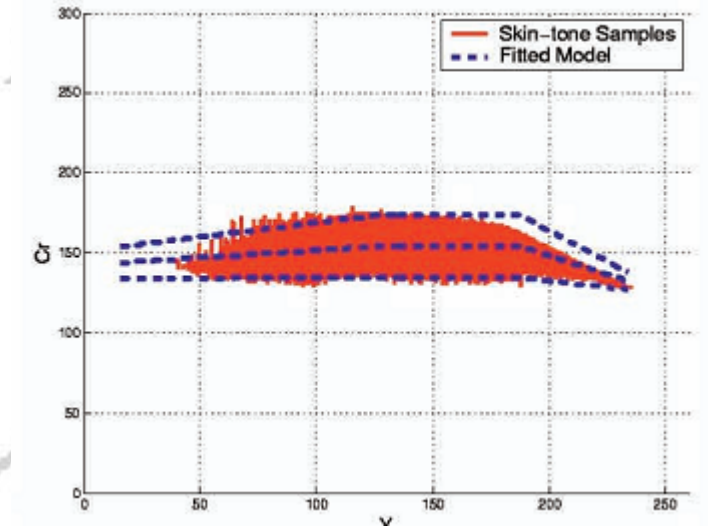


圖 2-8 Cr 在 Y 的變化

如圖2-7和圖2-8所示，可以利用4個邊界條件，來限制住膚色的分布範圍，可以很適應光源過亮和光源不足的區域，來提升YCbCr膚色模型性能，從RGB線性轉換YCbCr，YCbCr非線性的轉換YCb'Cr'，色彩空間推導的公式如下：

(1) 從圖2-7，圖2-8中，將膚色的中軸線分別用

$$\begin{aligned} \overline{C_b}(Y) &= \begin{cases} 108 + \frac{(K_i - Y) \cdot (118 - 108)}{K_i - Y_{\min}} & \text{if } (Y < K_i) \\ 108 + \frac{(Y - K_h) \cdot (118 - 108)}{Y_{\max} - K_h} & \text{if } (K_h < Y) \end{cases} \\ \overline{C_r}(Y) &= \begin{cases} 154 - \frac{(K_i - Y) \cdot (154 - 144)}{K_i - Y_{\min}} & \text{if } (Y < K_i) \\ 154 + \frac{(Y - K_h) \cdot (154 - 144)}{Y_{\max} - K_h} & \text{if } (K_h < Y) \end{cases} \quad (2-11) \end{aligned}$$

其中 K_h 和 K_i 為常數，即非線性區域色彩範圍值，分別為： $K_h=188$ ， $K_i=125$ 。
 Y_{\min} 和 Y_{\max} 為常數，它們是由實驗結果所產生，膚色在 Y 分量的最小值和最大
 值： $Y_{\max}=235$ ， $Y_{\min}=16$ 。

(2) 從圖2-9和圖2-10 中，將膚色寬度分別用

$$Wc_i(Y) = \begin{cases} Wc_i + \frac{(Y - Y_{\min}) \cdot (Wc_i - Wc_i)}{K_i - Y_{\min}} & \text{if}(Y < K_i) \\ Whc_i + \frac{(Y_{\max} - Y) \cdot (Wc_i - Whc_i)}{Y_{\max} - K_h} & \text{if}(K_h < Y) \end{cases} \quad (2-12)$$

其中 i 代表 b 和 r， Y_{\max} 、 Y_{\min} 、 K_h 、 K_i 同上式。 Wc 、 Wc_i 、 Whc 分別為常數，
 它們分別為：

$$\begin{aligned} Wc_b &= 46.97, Wc_i_b = 23, Whc_b = 14 \\ Wc_r &= 38.76, Wc_i_r = 20, Whc_r = 10 \end{aligned} \quad (2-13)$$

(3) 最後根據上面的結果得到如下非線性轉換的公式：

$$C'_i(Y) = \begin{cases} \left(C_i(Y) - \overline{C_i(Y)} \right) \cdot \frac{Wc_i}{Wc_i(Y)} + \overline{C_i(Y)} & \text{if}(Y < K_i) \text{ or } (K_h < Y) \\ C_i(Y) & \text{if}(Y \in [K_i, K_h]) \end{cases} \quad (2-14)$$

經過這樣子的非線性轉換公式後，將轉換後的 Cb' 和 Cr' 投影到 $Cb' - Cr'$ 的二
 維子空間，就可以得到實用的膚色模型。如投影圖 2-9：

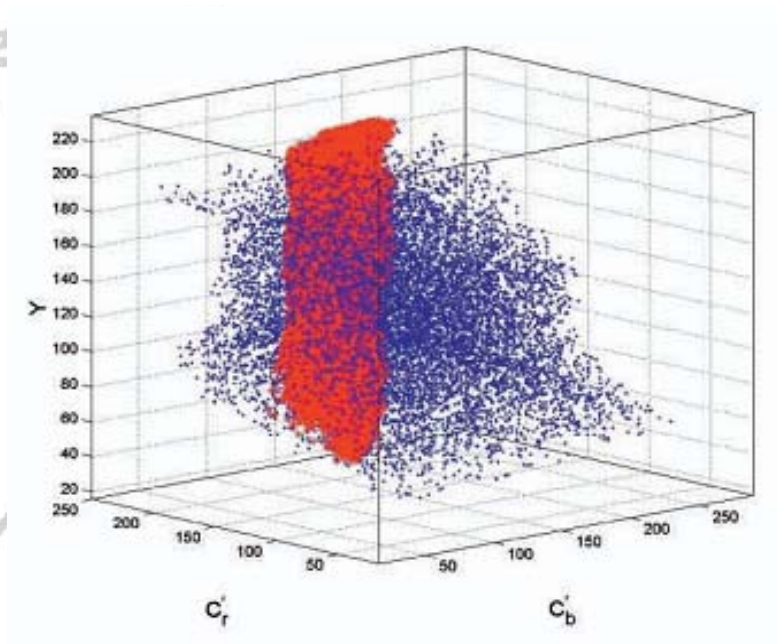


圖 2-9 YCb'Cr' 膚色的分布情形

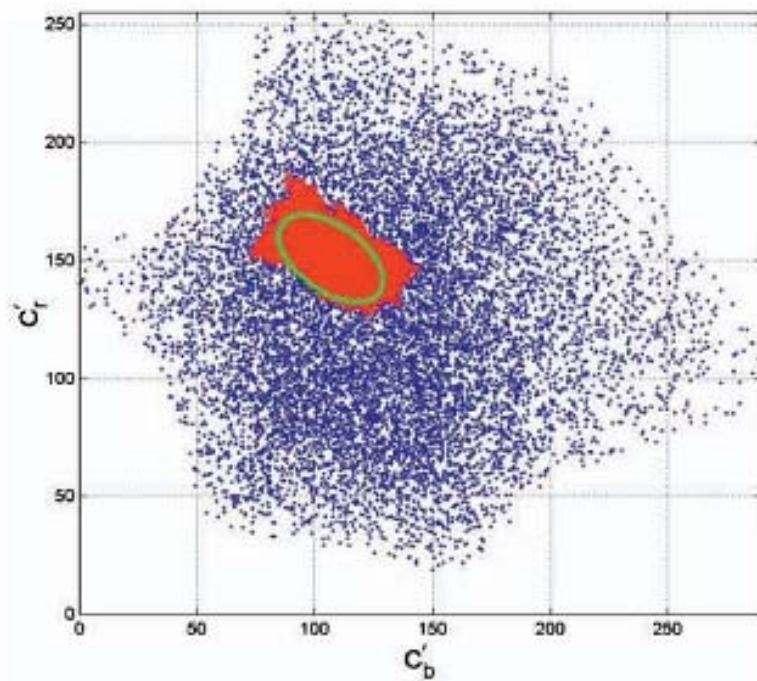


圖 2-10 投影在Cb'Cr'的膚色模型

依照傳統的方法，可以用一個近似橢圓形的方法，來限定膚色的區域，如圖 2-9 中橢圓。可以得到下面算式：圖 2-9 可以利用一個橢圓來近似這個膚色的區域，可以由一個橢圓標準式來表示如下：

$$\frac{(x - ecx)^2}{a^2} + \frac{(y - ecy)^2}{b^2} = 1$$

$$\text{其中, } \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} C'_b - cx \\ C'_r - cy \end{bmatrix} \quad (2-15)$$

其中個別的參數：

$$cx = 109.38,$$

$$cy = 152.02,$$

$$\theta = 2.53 \text{ (弧度)},$$

$$ecx = 1.60, \quad ecy = 2.41, \quad a = 25.39, \quad b = 14.03$$

臉部偵測通常可以運用一些明顯的特徵，如眼睛、鼻子、嘴唇、膚色、臉形、輪廓……等等。當來源為監視器或者攝影機時，這時要完成臉部偵測就必須要求能夠在高偵測率的狀況下，以極快的速度來處理影像。Viola和Jones[12][13]所提出物件偵測應用在人臉偵測的方法，可細分出四個部份來說明，積分影像表示法（Integral Image）、矩形特徵（Rectangle）、自適性布斯特分類法（AdaBoost）及串列式模型（Cadcade Model），將這四部份組成一套快速比對特徵的方法。

由於單一的矩形特徵對複雜目標物的辨識率不高，因此採用AdaBoost[6]作為訓練方法，挑選並結合多個矩形特徵，來找出關鍵的視覺特徵，並產生一個非常有效率的分類器，然後以串列式的決策架構方式組合多個分類器，每個分類器為最小單元分類一個簡單的矩形特徵。此架構可快速地排除影像中與臉部視覺特徵差異較大的區域，減少不必要的計算。因此臉部偵測器主要是使用數個簡單的臉部特徵，稱為簡易分類器（Simple Classifier），串聯成為一個複雜分類器（Complex Classifier）。愈前面的分類器規則愈容易，所需運算量也愈少，但是具有快速判斷影像視窗為人臉或者是背景之效果；愈後面特徵則是愈複雜人臉規則，所以運算量也隨之提高，不過當層數愈多時，人臉偵測的準確性也會提高。論文中利用串聯式的決策架構來判斷輸入的影像視窗是否為臉部特徵，因為結合了膚色的條件，所以在人臉偵測時不必進行全域的掃描，如此可以減少運算量以達到即時性偵測人臉的效果。詳細方法描述如下。

2-3-1 積分影像

對於一個輸入影像 I ，像素 (x, y) 處的積分影像值定義為：

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y') \quad (2-16)$$

表示像素 (x, y) 的灰階值，積分影像值等於陰影部分所有灰階值的和。如此定義，如果要得到一個輸入影像 I 的積分影像時，僅需對原圖逐點掃描一次就可計算出來。計算方式如下：

$$s(x, y) = s(x, y-1) + i(x, y) \quad (2-17)$$

$$ii(x, y) = ii(x-1, y) + s(x, y) \quad (2-18)$$

其中 $s(x, y) = \sum_{y' \leq y} i(x, y')$ 是原影像 (x, y) 點的列累積(Cumulative Row Sum)， $s(x, -1) = 0$ 和 $ii(-1, y) = 0$ 。使用積分影像時，任何影像內部的矩和(Rectangular Sum)可以僅使用四個參考值計算即可取得。在 Viola 的系統中，每個矩形特徵值的計算，最多只需要從積分影像中取9個元素做加減運算。

2-3-2 矩形特徵

矩形特徵也被稱為Haar-Like特徵，最早在1998年提出應用在人臉偵測上，後來被Viola和Jones進一步延伸；常見的矩形特徵如圖2-11所示：

臉部區域亮度的變化會造成影像上的明亮，與灰暗的區塊。因此可以將原本的目標影像，利用多個亮度差異的矩形特徵來近似。以反向來思考，若以知臉部區域灰階變化程度，矩形特徵的資訊做為比對的資料來源，則將可以對未知含有人臉資料的影像與矩形特徵作比對，來判別是否為人臉。

以人臉特徵為例，凹陷的眼窩與突出的鼻子下方通常會造成陰影，因此會較暗。因此如圖 2-12 所示這四個明顯的特徵往往對人臉有著較大的辨識率。

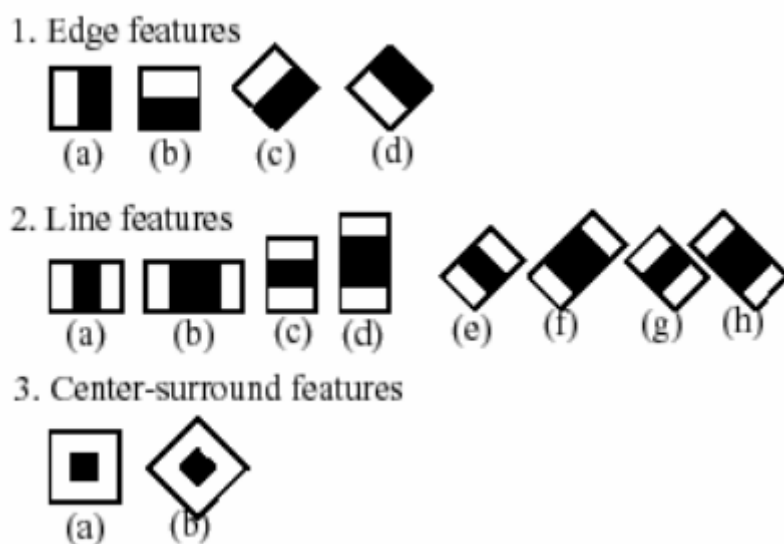


圖 2-11 矩形特徵

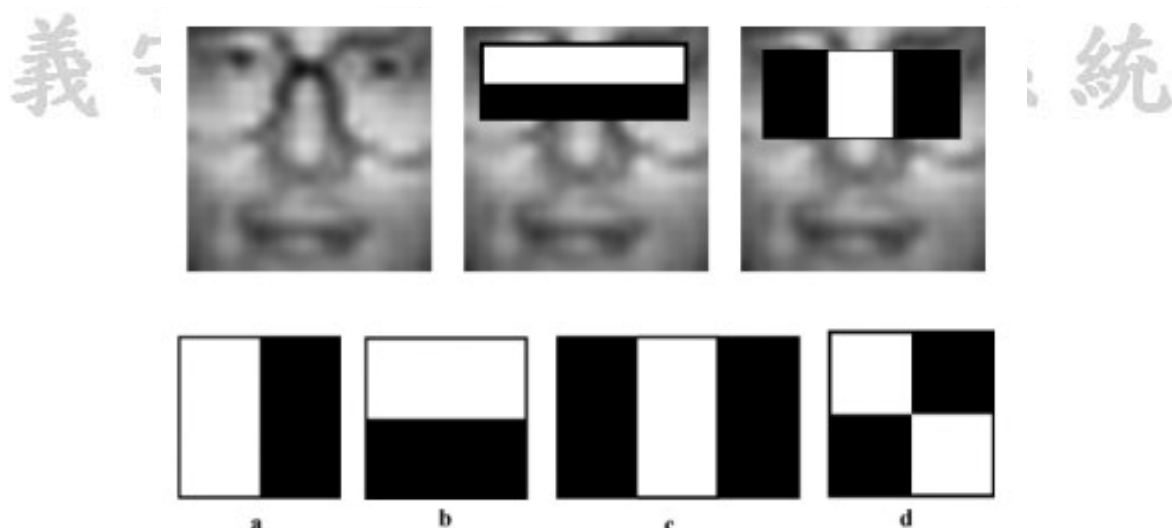


圖 2-12 矩形特徵對應在臉部情形

2-3-3 自適性布斯特演算法

Adaboost 演算法是一種疊代方法，主要的用途是透過從大量的分類器中選取最具有分類意義的那些，組合成一個強分類器，其中弱分器指的是那些性能比隨機分類略好一點的分類器，在 Viola 中也有具體演算法可以參考。在 Viola 的方法中，使用矩形特徵作為分類的依據。矩形特徵的值是指影像上兩個或者多個形狀大小相同的矩形內部所有像素灰階值之盒的差值。在 Viola and Jonse 所提出臉部偵測的方法中，使用了四種矩形特徵，如圖 2-12 所示。

AdaBoost 訓練演算法（計算每一分類器的權重）輸入樣本影像 $(x_1, y_1) \dots (x_n, y_n)$ 其中 $y_i = 0, 1$ 分別代表非目標和目標影像樣本。初始化樣本權重 ω_i ， $i = 1, 2, \dots, n$ for $y_i = 0, 1$ ，其中 m 和 l 是目標影像和非目標影像樣本的數目。

令 $t=1, \dots, T$

1. 正規化權重

$$\omega_{t,i} \leftarrow \frac{\omega_{t,i}}{\sum_{j=1}^n \omega_{t,j}} \quad (2-19)$$

而 ω_t 為一個機率分佈值。

2. 對每一特徵 j ，訓練一個分類器 h_j ，並限制它只使用單一特徵。使用下面的式子評估錯誤率：

$$\omega_t, \varepsilon_j = \sum_i \omega_i |h_j(x_i) - y_i| \quad (2-20)$$

3. 選擇最低錯誤率 ε_t 的分類器 h_t

4. 更新權重

$$\omega_{t+1,i} = \omega_{t,i} \beta_t^{1-e_i} \quad (2-21)$$

其中若樣本影像 x_i 分類正確，則 $e_i = 0$ ，否則 $e_i = 1$ 且

$$\beta_t = \frac{\varepsilon_t}{1 - \varepsilon_t}。$$

最後強分類器為：

$$h(x) = \begin{cases} 1, & \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0, & \text{otherwise} \end{cases} \quad (2-22)$$

其中 $\alpha_t = \log \frac{1}{\beta_t}$

因此對於丟入的一張image，就會由這j個矩形特徵來投票，每一個特徵的投票的權重不太一樣。但只有當加權值大於一半以上的所有分數時，才會認可這一張image為人臉。

2-3-4 串列式模型

串列式分類器是很多強分類器的組合，如圖 2-13 所示。在判斷的過程中以 T 來表示通過某一層特徵的檢測，若為 F 則表示偵測失敗，也就是此影像視窗不是屬於人臉。每一層都是Adaboost 演算法訓練得到的一個強分類器，都經過門檻值調整，使得每一層都能讓幾乎全部人臉樣本通過，而排除了大部分非人臉樣本。每一層分類器根據如下式子判定輸入影像是否是目標影像或非目標影像。

$$\sum_{i=1}^N w_i^M h_M(x_i) > Th_M \quad (2-23)$$

其中 w_i^M 為權重在第 i 個樣本影像 x_i 和第 M 個分類器， $h_M(x_i)$ 和 Th_M 分類器的輸出機率和門檻值。

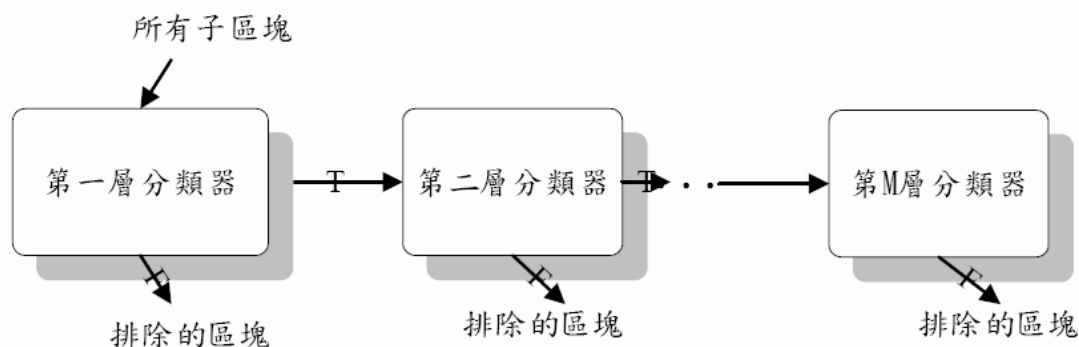


圖 2-13 串列模型圖

第三章 臉部特徵定位

本章節會介紹系統的流程以及架構；依照前面章節所提到的方法，將會用系統的方式呈現出影像的結果。可以將系統分成三個部份為人臉偵測、嘴唇定位與眼睛定位。在人臉偵測的部份，輸入一張含有複雜背景資訊的彩色影像，首先將原始影像經由膚色分析，去除背景資料，再透過 Viola 所提出來人臉偵測方法，去進行判斷是否為人臉。嘴唇特徵的部份，將以擷取好的人臉區域範圍，透過幾何特徵的關係，擷取出嘴唇的範圍。最後嘴唇定位的部份，依據上個步驟設立的嘴唇特徵點的步驟，然後先從連續的圖片去嘗試看系統是否，盡而穩定後，再以即時播放的方式去加以測試。

3-1 人臉偵測模組

3-1-1 影像前處理

由於人臉偵測演算法，主要是透過臉部的紋理的資訊進行偵測，只需考慮到灰階影像的處理。將在執行人臉偵測時，可以利用一些數位影像的處理技巧[1]，例如：直方圖均化（Histogram Equalization）、高斯平滑（Gaussian Smoothing）等處理，利用這些前處理的技巧，不僅可以使影像看起來更柔順，也可以避免一些影像中的雜訊。當影像輸入時，會將單張的彩色影像，利用一個 5*5 的高斯平滑濾波器的來去影像中的雜訊，高斯平滑數學式如下所示：

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (3-1)$$

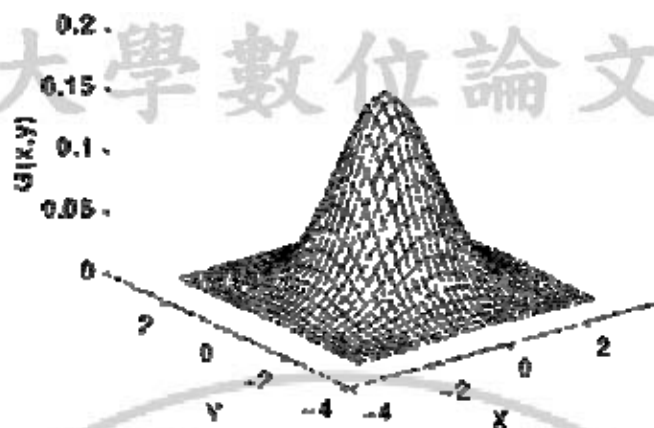


圖 3-1 高斯分布圖

	1	4	7	4	1
	4	16	26	16	4
$\frac{1}{273}$	7	26	41	26	7
	4	16	26	16	4
	1	4	7	4	1

圖 3-2 高斯遮罩

假設高斯分布的平均值為 0，而 σ 是高斯分布的標準差，如果 $\sigma=1$ ，我們可以得到如圖 3-1 所示。例如對於一個 5x5 的高斯濾波器 ($\sigma=1$)，我們可以得到以下的濾波器係數如圖 3-2。高斯濾波器是一個平滑化濾波器，平滑化程度是由標準差 σ 來控制， σ 值越大，平滑程度越高，相對的，影像越模糊。

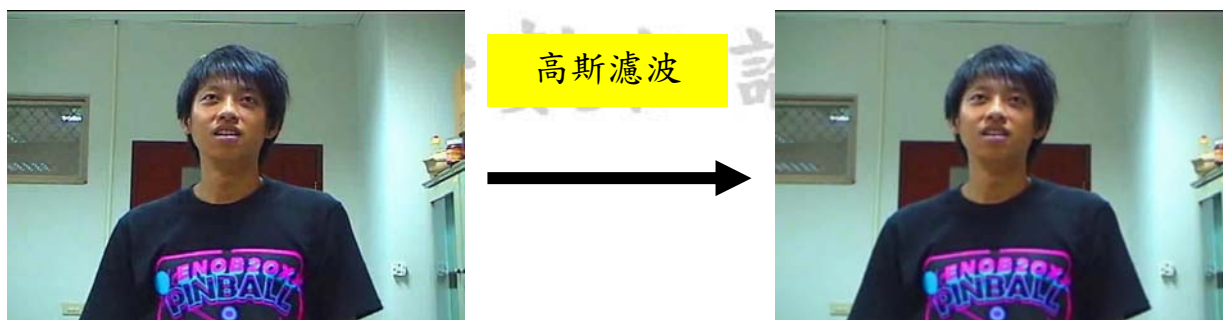


圖 3-3 平滑處理

3-1-2 膚色檢測

利用不同的色彩空間的轉換，膚色檢測也會有不同的效果，依照背景的複雜度的不同，截取出來的膚色也會有所差異。即使膚色的品質再好，在整體上也無法有效完成人臉偵測的工作，作膚色檢測最大的目的，可以將影像中非膚色的區域給去除，膚色檢測的工作做得好的話，不但可以使人脸偵測的成功率提高，和減少運算的時間。為了使膚色的偵測的效果達到更好，通常在選擇色彩空間時也是一大學問，在原始的影像的來源通常是 RGB 為基礎的色彩空間，在之前的章節也都探討過，RGB 色彩空間的光源的敏感度，這個因素對於在作膚色檢測的部分，將不採納以 RGB 色彩空間，作為膚色檢測的工作。在本論文中，A. K. Jain 和 M. Abdel-Mottaleb 和 R. L. Hsu[9]，所提出的提出膚色模型，是在 YCbCr 的色彩空間下做膚色的檢測，僅保留影像中膚色的區域，並且將膚色的灰階影像，位置記錄下來，映射

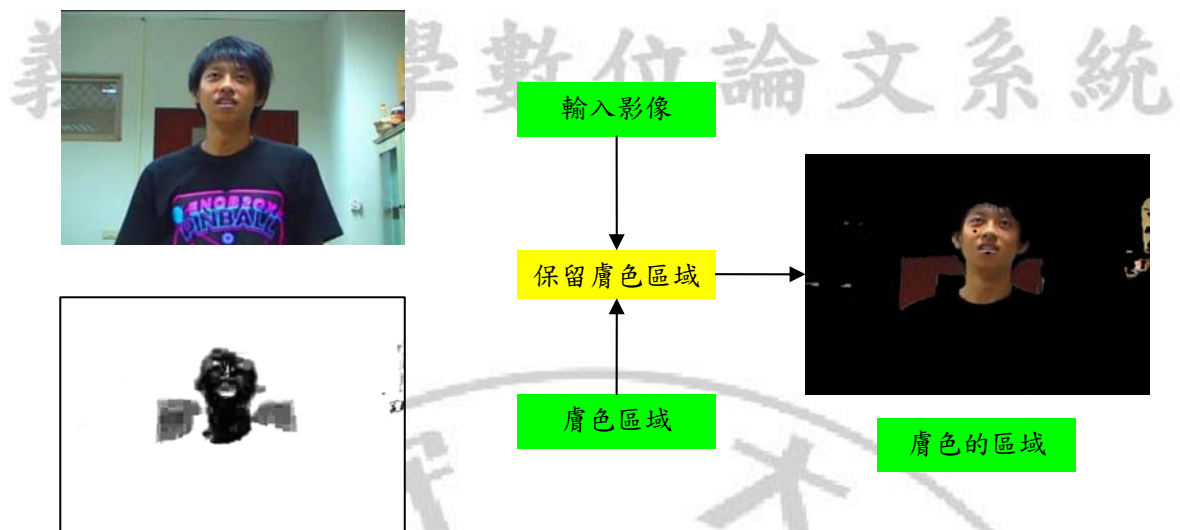


圖 3-4 膚色檢測

3-1-3 人臉偵測

利用 Viola[6]所提到人臉偵測演算法，利用矩形特徵進行臉幾何紋理的偵測，經過膚色檢測後，只須針對影像中剩下膚色區域，進行偵測的動作，然而提取出人臉的位置。在搜尋的人臉時，由於 Viola 所提出的偵測方法，並不考慮的顏色因素，則可以將輸入的影像資料，轉成灰階的圖片，對 Viola 只考慮臉部幾何紋理的因素，並沒有影響到偵測的品質，而又可以加速電腦處理的速度。在取得臉部位置資訊後，利用該資訊對圖片進行切割的動作，則顯示出臉部在圖像正確的位置。在人臉區域內，根據臉部幾何特徵的關係，由上而下分割出眼睛候選區域和嘴唇候選區域。



圖 3-5 人臉偵測

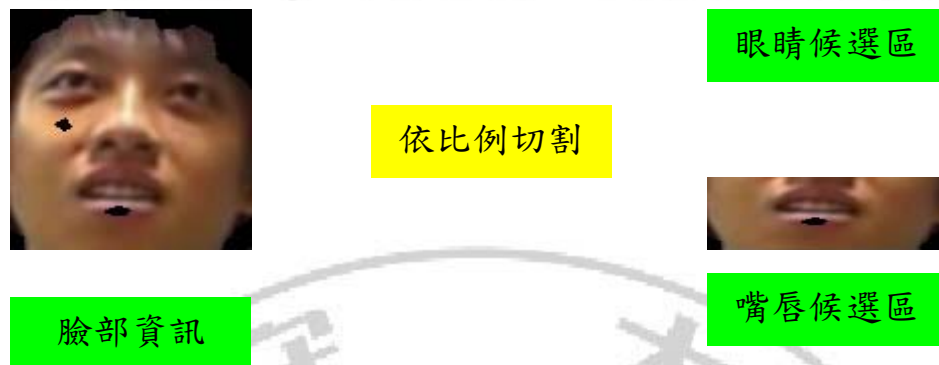


圖 3-6 分割臉部

3-2 眼睛模組

臉部資訊中，不僅眼睛的形狀變化性是最小，眼球的顏色也跟膚色相差甚大，不管是用顏色還是形狀，在眼睛的偵測上都不會造成很大的困難，除了戴墨眼鏡或者遇到頭髮遮蔽，造成眼睛特徵不明顯，才會使眼睛偵測率有下降的情形發生。

3-2-1 眼睛候選區域切割

選擇出定位眼睛候選區域的部份，首先將臉部的區域，依照人臉的幾何關係，將眼睛的候選區域在細分出左右的兩個候選區域範圍。提供給眼睛的定位使用，如圖 3-7 所示：

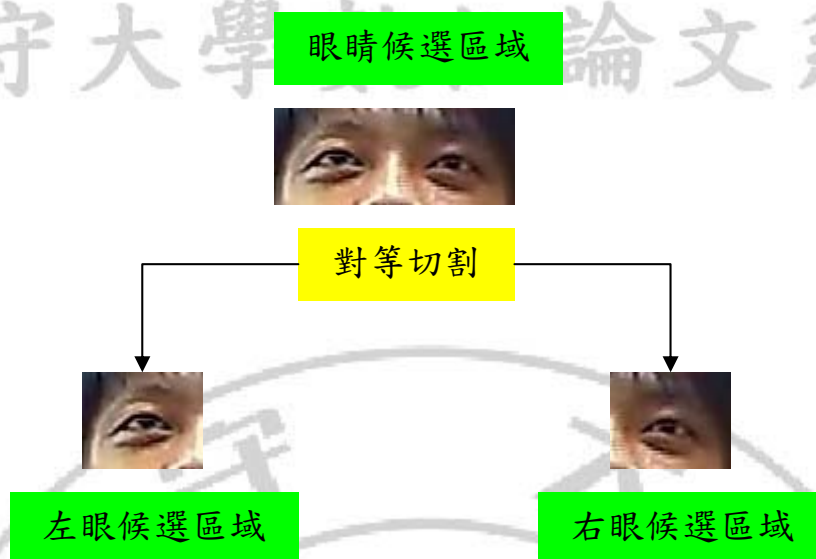


圖 3-7 眼睛候選區域

3-2-2 眼睛偵測

再人臉的五官定位裡，就屬於眼睛的偵測最為簡易，眼睛偵測法，採用形狀的資訊來偵測。首先先建立一個眼睛資料庫，樣板每個都是具有代表性的眼睛如圖 3-8 所示。模板比對 (Template Matching) 將眼睛特徵設定為一個或多個 template，利用標準模型來描述人臉或是部份的人臉特徵，再以 Searching Window 的方式偵測。而此模型的數量依據研究者所使用的演算法而不同。最後偵測結果的判定大多使用輸入影像與樣板間的相關值 (Correlations) 來決定影像中是否含有眼睛部位的存在。此方法的優點是方法簡單容易實作，適用是眼睛定位上，眼睛的變化小，減少運算的時間。結果如圖 3-9 眼睛偵測圖。



圖 3-8 眼睛模板



圖 3-9 眼睛偵測

3-3 嘴唇模組

臉部資訊中，嘴唇的特徵是最具有代表性，在嘴唇偵測領域中，無論是透顏色 (Color)、紋理 (Texture)、邊緣資訊 (Edge) 等方法，但嘴唇偵測成功率，都是不盡理想，會影響嘴唇偵測的因素，可以說是非常的多，考慮牙齒的露出、嘴唇附近鬍鬚的影響 和舌頭吐出等因素。為了排除這些影響嘴唇偵測的因素，唯有不斷縮小嘴唇搜尋範圍，利用局部性的特徵，促使嘴唇偵測率提升，達到更好的效果。

3-3-1 嘴唇候選區域

嘴唇的偵測，由於嘴唇的顏色跟膚色的色彩資訊，有一定的相似度存在，所以在作嘴唇的偵測時，就先不把顏色資訊當做判斷的依據，首先可利用 SOBEL 邊緣偵測的方法，找出水平方向與垂直方向的邊緣，把兩張邊緣偵測做邏輯 OR 的動作如圖 3-10，產生出具有水平跟垂直方向的邊緣利用這些邊緣的資訊，經由投影在一維的平面上，分成水平方向的投影和垂直方向的投影圖 3-11，將垂直投影的資訊，利用唇形的幾何位置，取出中尖峰值區域，認定為唇形寬度，水平投影的資訊，也是利用唇形在幾何位置，決定出嘴唇長度，更小的嘴唇範圍，有利於在作唇形輪廓擷取。將垂直投影的資訊，利用唇形的幾何位置，為了排除臉部輪廓的資訊，把投影量資

訊，前 10%和後 10%的寬度座標，不列入計算的範圍，當垂直投影量少於最大的值的 $1/10$ 時，記錄該點的位置，和第二次小於最大值 $1/10$ 時記錄該點，認定為唇形寬度，由於水平投影的資訊，並不理想，所以在切割嘴唇高度時，就不是利用邊緣資訊來決定嘴唇高度位置。有利於在作唇形輪廓擷取。

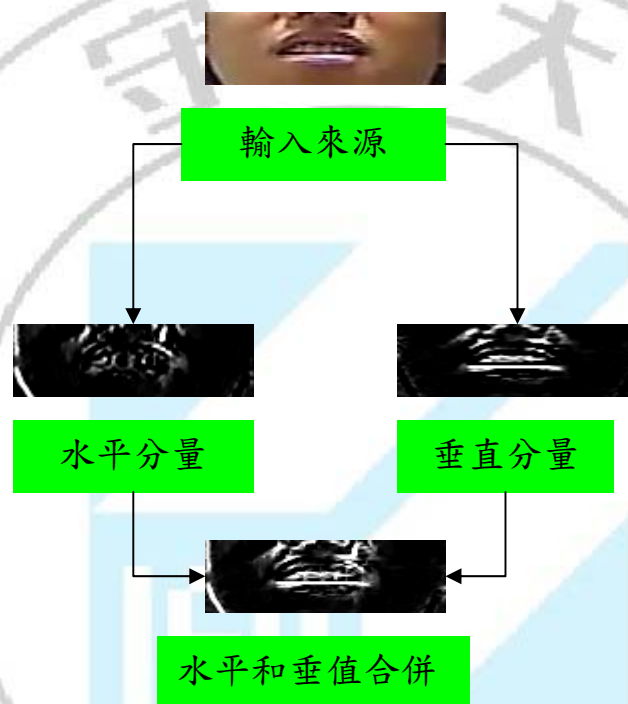


圖 3-10 邊緣偵測

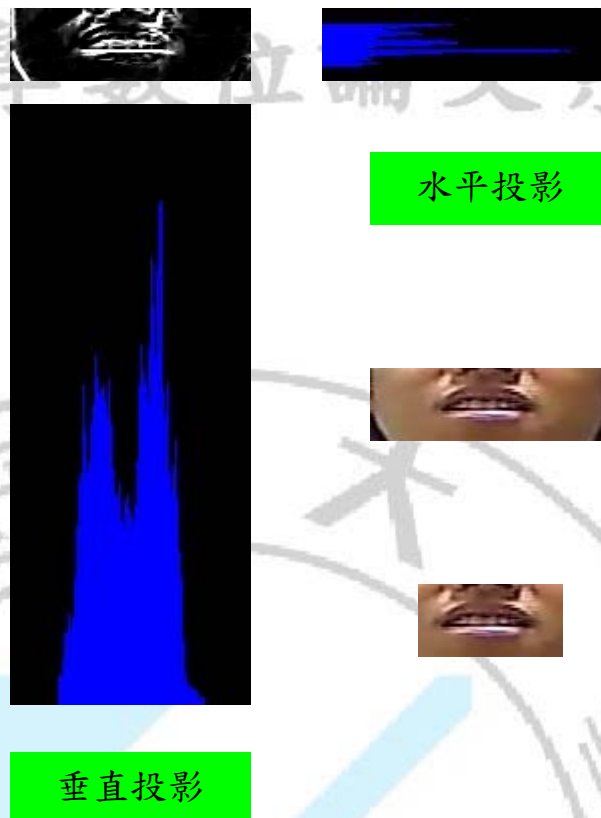


圖 3-11 邊緣投影

3-2-2 嘴唇的擷取

雖然我們可以藉由上述的方法偵測到嘴唇物件的移動，但是在通常的情況下，輸入的嘴唇物件的區域中，經過一次高斯平滑的處理，把顏色的區域給區分出，再利用直方圖均化方法，使影像中的灰階顏色表現更加的明顯，讓嘴唇的顏色與膚色的顏色產生更大的差異度，為了降低更多的雜訊，試圖將灰階影像，做二值化（Threshold）的動作，二值方法最被常來使用，就屬於 OTSU 二值法[17]，方法的使用非常的簡單，不需輸入任何的參數。OTSU 用統計的方法來對其影像之像素分佈來作運算，求取最佳門檻值(Threshold)來二值化，將影像切割變成黑與白，將其背景變成白的，而要取之範圍為前景則為黑色，方便作其它處理。

義守大學數位論文系統

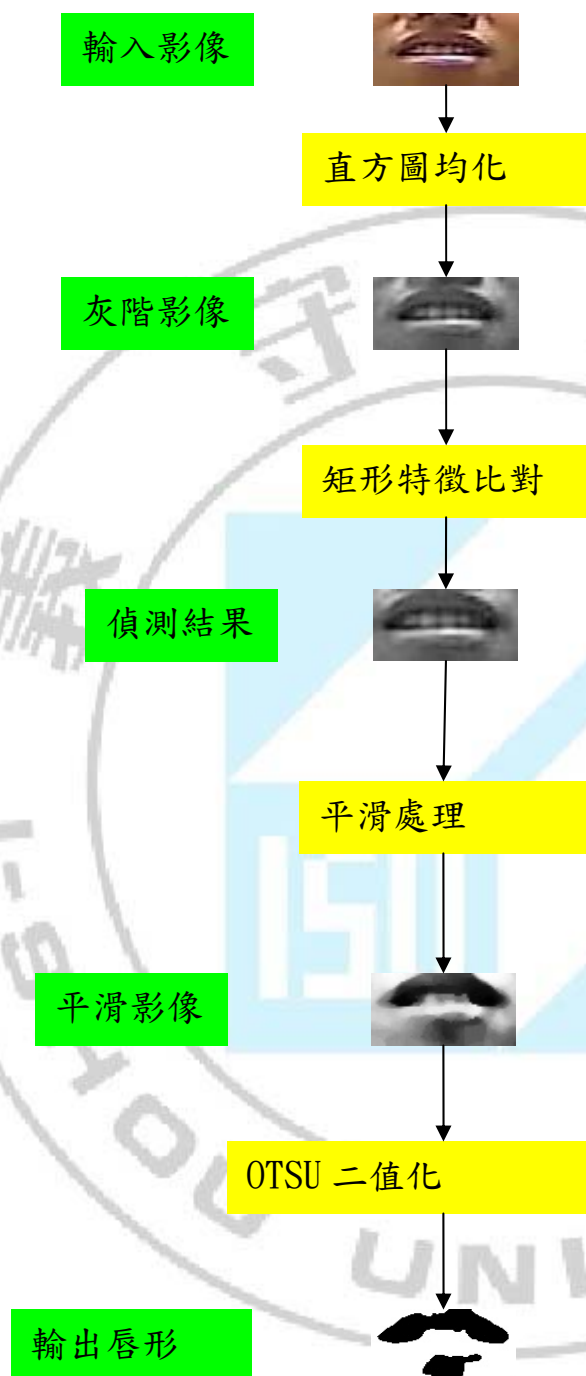


圖 3-12 嘴唇與膚色的分割

第四章 實驗結果與討論

4-1 系統環境

本系統使用的軟硬設施如下：

軟體：

程式平台：Visual c++ 6.0

函式庫：OPENCV

系統：Windows XP

硬體：

CPU：AMD Athlon 64*2 3600

記憶體：1G

資料庫：

自行拍攝的影片，所拍攝的人物，為本科系 95 年級的全體同學，由於女生的人數較少，所以資料庫的內容以男生為大多數，所有的影像都是為正面的人臉影像，每個樣本，所拍攝的時間，大約至 40 秒至 1 分鐘不等的長度，影片的解析度為：720*480，拍攝的環境，在室內，拍攝時無使用補強的燈光。








編號	圖片	解析度	時間	撥放速率
1.		720*480	36.5 秒	每秒 29.7 張
2.		720*480	41.1 秒	每秒 29.7 張
3.		720*480	42.0 秒	每秒 29.7 張
4.		720*480	30.0 秒	每秒 29.7 張
5.		720*480	36.6 秒	每秒 29.7 張
6.		720*480	29.4 秒	每秒 29.7 張
7.		720*480	47.8 秒	每秒 29.7 張

表 4-1 影片資料庫

義守大學數位論文系統

4-2 實驗介面

系統介面設計如下，系統可以開啟影片、視訊裝置和單張圖片的模式，在做偵測之前，可選擇三種模組進行偵測的動作，並在介面顯示各個模組直行的情形。如圖 4-1



圖 4-1 系統介面圖

義守大學數位論文系統

4-3 實驗結果

本實驗的結果，分成三個部分作來探討：第一個部分講探討的人臉偵測在影像中，所偵測的人臉區域，是否為人臉；第二個部分講述眼睛在受到頭髮和眼睛影響和眼鏡的影響；第三個部分是有關嘴唇是否能正確定位和切割候選區域是否有影響到唇形的。

4-3-1 人臉偵測率

在評估系統偵測率，評估標準分別可以用 Recall 和 Precision，定義如下：

$$\begin{aligned} recall &= \frac{Detect}{Detect + miss} \times 100 \% \\ precision &= \frac{Detect}{Detect + false} \times 100 \% \quad (4-1) \end{aligned}$$

其中 *Detect* 代表正確的偵測出臉部區域的數目，*miss* 代表臉部偵測時未被偵測出的數目，*false* 代表在膚色的區域找到多餘人臉數目。因此當 Recall 為 100%，表示所有在膚色區域中，臉部的位置都被找出來，而和 Precision 為 100%，則表示所有在所有膚色區域，找出臉部數目為正確。

測試時，將錄製影片，以每 0.5 秒進行一次偵測的動作 1 張，則影片撥放方式進行為已每秒 29 張的速率的在播放。當中記錄臉部位置與面積的資訊，在進行下一次圖片比對的動作，為了避免分割臉部區域，範圍時發生錯誤，保留圖片所留下人臉資訊。如圖 4-2 偵測間隔中，沒有對該區域進行人臉偵測的部分，此區域是由上一次偵測結果所留下的資料，造成誤錯

的因素，因為目標物快速的移動，以及畫面閃爍，所發生的錯誤情形。








影片編號	總數	Detection	Miss	False	Recall	Precision
1. 	73 張	38 張	19 張	16 張	66%	70%
2. 	82 張	56 張	21 張	5 張	72%	91%
3. 	84 張	49 張	21 張	14 張	70%	77%
4. 	60 張	41 張	7 張	12 張	85%	77%
5. 	73 張	40 張	16 張	17 張	71%	70%
6. 	74 張	36 張	15 張	23 張	70%	61%
7. 	95 張	70 張	18 張	17 張	80%	80%

表 4-2 人臉偵測率



圖 4-2 正確臉部區域



圖 4-3 偵測錯誤人臉區域

義守大學數位論文系統

4-3-2 眼睛特徵偵測率

在眼睛的偵測模組中，只有針對正確臉部區域面積，做處理，將測試分割出眼睛候選區域內，對於大部分眼睛候選區域，會可能發生，只有偵測到單顆眼的問題，或者其中一顆眼睛，產生不穩定，造成左右水平位置差距過大的情形。另有也有可能受到戴眼鏡和眼睛受到頭髮的遮閉的問題，測試的範例，則選擇已經切割好了臉部的區域。依照幾何比例的關係，將候選區域再細分出左右兩邊，進行眼睛偵測的部分。對於眼睛候選區域的分割成功與否，將和人臉偵測後，分割出來的圖片有大的關係，測試的數據使用人臉偵測所儲存資料，去近判斷的。








影片編號	總共張數	左眼偵測張數	右眼偵測張數	左眼偵測率	右眼偵測率
1. 	38 張	27 張	33 張	91%	86%
2. 	56 張	51 張	49 張	91%	87%
3. 	49 張	47 張	43 張	95%	87%
4. 	41 張	40 張	38 張	97%	87%
5. 	40 張	35 張	39 張	87%	97%
6. 	36 張	25 張	18 張	84%	70%
7. 	70 張	64 張	68 張	91%	97%

表 4-3 眼睛偵測率

4-3-3 嘴唇特徵偵測率

嘴唇偵測時，考慮到因素，就遠比在眼睛偵測則要考慮的多，不像眼睛偵測的是可以使用模板比對的方式，而且使用模板比對的方式，只有適於在外觀的形狀，沒有太大的變化時，才能使用。就因為嘴唇的形狀，常常受制皮膚的顏色相似，以及一些嘴唇內的變化，這些可能造成嘴唇的定位已經偵測方法受到挑戰，只有經由不斷的修正縮小嘴唇的候選區域範圍，降低外在因素的影響。

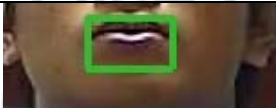
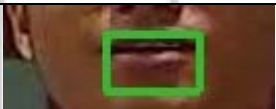
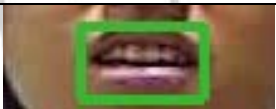
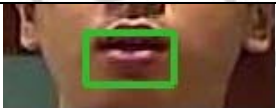
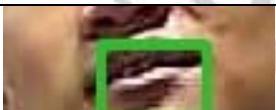
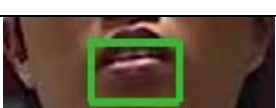

影片編號	偵測張數	嘴唇偵測張數	嘴唇偵測率
1. 	38 張	25 張	65%
2. 	56 張	37 張	66%
3. 	49 張	29 張	59%
4. 	41 張	24 張	58%
5. 	40 張	27 張	67%
6. 	36 張	18 張	50%
7. 	70 張	51 張	72%

表 4-4 嘴唇偵測率

4-4 結果討論

在錄製影片中，雖然是在室內的環境，透過一些影像的分析結果後，都常常受限於硬體的設施，以及實驗燈光的不足，以及拍攝的角度無法統一，導致成果變差環境由於當初都沒有參考，由於國外影片的資料庫，在錄製時，都會考慮到外在因素的變化，且都會詳細說明，列出資料庫中所有外在條件說明和適合在使用在哪一方面，將有可能的影響影片品質因素排除影片在外。在本論文的實驗中，雖然臉部區域範圍，可以輕易從影片擷取出來，透過臉部特徵的幾何關係，切割出眼睛候選區域和嘴唇候選區域的位置，對於眼睛定位可以清楚定位出來，相對在嘴唇的定位，並沒有想像中的順利，無疑影響嘴唇偵測因素，因為在嘴唇產生形狀變化時，都會造成無法成功將嘴唇定位出，另外跟皮膚顏色有一定的相似度，這都會造成嘴唇偵測錯誤的因素。

第五章 結論

5-1 結論

在論文中建構出臉部特徵定位與特徵擷取的系統，在自行錄製的影片中，進行測試。利用矩形特徵做臉部的定位，將成功定位出臉部區域的時，藉由臉部幾何的關係，將眼睛和嘴唇的位置各為劃分出。眼睛的部分，在個別畫出左右兩邊，進行模板比對的方式，在左右眼平均偵測率約 89%，由於左右的偵測率不一定相同。在嘴唇的部分，利用邊緣的投影資訊和顏色的畫分，在垂直投影上，可以縮小嘴唇區域寬度，在水平投影時，都會常被鼻子資訊給干擾，在選擇上水平投影的資訊就不考慮，進而考慮膚色與唇色差異關係，在此取設門檻值，只能區分嘴唇比較亮的部分，只選定亮色的部分當作嘴唇的位置與特徵，嘴唇的偵測率約 52%，就沒有眼睛偵測率來得好。

5-2 未來研究方向

在未來的研究上，希望在人臉偵測的部分，可加入多人臉偵測，何分割出多個嘴唇和眼睛的部分。在眼睛的部分，須加強兩眼的平衡。在嘴唇部分，不但可以加強嘴唇的輪廓描述，以及找出嘴唇和膚色的相關差異的地方，建立一組嘴唇顏色的相關參數，透過邊緣偵測偵測方式，利用邊緣的資訊，將唇形的輪廓，完整的描述出來，進而達到嘴唇追蹤、輪廓的分析。

[1] Yambor WS, Draper BA, Beveridge JR "Analyzing PCA-based Face Recognition Algorithms:Eigenvectors Selection and Distance Measures". 2nd Workshop on Empirical Evaluation in Computer Vision, 2000.

[2] Hatim A Aboalsamh, Hassan I Mathkour, Ghazy M R Assassa, Mona F M Mursi, "Face recognition using principle components and linear discriminant analysis" Mathematics And Computers In Science And Engineering archive Proceedings of the 8th, P276-282,2009.

[3] Leonid Sigal, Stan Sclaroff, Vassilis Athitsos , "Skin color-based video segmentation under time-varying illumination" IEEE Transactions on Pattern Analysis and Machine Intelligence P862 - 877 , 2004.

[4] Zhang Zhengzhen, Shi Yuexiang , "Skin color detecting unite YCbCr color space with YCbCr color space " Image Analysis and Signal Processing , P221 - 225 , 2009.

[5] Seyyed Meysam Hosseini, Hadi Sadoghi Yazdi, Saber Amjadi, S "A Novel Color Space Creating Method Applied to Skin Color Detection", International Conference on Digital Image Processing , P17 - 21 ,2009.

[6] S. Kawato, J. Ohya, "Automatic Skin-color Distribution Extraction for Face Detection and Tracking",Signal Processing Proceedings,P1415 - 1418 , 2000.

[7] Yu-ting Pai,Shanq-jiang Ruan, Mon-chau Shie, chi Liu, "A

Simple and Accurate Color Face Detection Algorithm in Complex Background", IEEE International Conference on Multimedia and Expo , P1545-1548 ,2006.

[8] Cheng-Chin Chiang, Wen-Kai Tai, Mau-Tsuen Yang, Yi-Ting Huang and Chi-Jaung Huang, "A Novel Method For Detecting Lips, Eyes and Faces in Real Time", The Journal of Real Time Image, P277-287, 2003.

[9] R.L. Hsu, M. Abdel-Mottaleb , A.K. Jain, "Face Detection in Color Images" , IEEE Transactions on Pattern Analysis and Machine Intelligence , P696-706 (2002).

[10] Alberto Albiol , Luis Torres , Edward J. Delp, "Optimum color spaces for skin detection", In Proceedings of the International Conference, P122-124, 2001.

[11] Zhefan Jin, Zhen Lou, Jingyu Yu Yang, Quan Sen Sen Sun "Face detection using template matching and skin-color information", Neurocomputing, 2007, P 794-800, 2007.

[12] Viola Paul, J. Jones. Michael, "Rapid Object Detection Using a Boosted Cascade of Simple Features", IEEE CVPR, P511-518, 2001.

[13] Paul Viola , Michael Jones, "Fast and Robust Classification using Asymmetric AdaBoost and a Detector Cascade", Statistical and Computational Theories of Vision, 2001.

[14] Paul Viola , Michael Jones, "Robust Real-Time Face Detection", International Journal of Computer Vision, P137-154. 2004.

[15] Hazem M. El-Bakry , "Face detection using fast neural networks and image decomposition", Neurocomputing , P1039-1046 , 2001.

[16] Faisal Shafait , Ralph Kricke , Islam Shdaifat , Rolf-rainer Grigat, "Real Time Lip Motion Analysis for a Person Authentication System", P1957-1960, 2006.

[17] Nobuyuki Otsu (1979), "A threshold selection method from gray-level histograms", IEEE Trans. Systems, Man, and Cybernetics, P62-66, 1979.

[18] R. C. Gonzalez and R. E. Woods 著，數位影像處理，繆紹綱譯，台灣培生教育出版，台北，民國九十三年。

[19] Face detection by template matching demo

<http://sites.google.com/site/simonegasparini/research/face>