

國立交通大學

電控工程研究所

碩士論文

室內環境之人物辨識

Person Identification in an In-door Environment

研究生：廖天健

指導教授：張志永

中華民國九十九年七月

室內環境之人物辨識

Person Identification in an In-door Environment

學 生：廖天健

Student : Tien-Chien Liao

指導教授：張志永

Advisor : Jyh-Yeong Chang

國立交通大學

電機工程學系



Submitted to Department of Electrical Engineering

College of Electrical Engineering

National Chiao-Tung University

in Partial Fulfillment of the Requirements

for the Degree of Master in

Electrical and Control Engineering

July 2010

Hsinchu, Taiwan, Republic of China

中華民國九十九年七月

室內環境之人物辨識

學生：廖天健

指導教授：張志永博士

國立交通大學 電控工程研究所

摘要

本篇論文以開發智慧型居家照護系統為目標，對於居住者之居住環境提供監控之服務，而人物辨識則為本系統中的一個重要部份，藉由人物辨識技術，得以監視環境內每個人物的身分。在本篇論文中，我們開發出一個人物辨識系統，由兩個子系統所構成，分別為人臉辨識系統以及色彩模型辨識系統。兩個系統皆使用獨立的攝影機搭配特定的監視角度，並且透過網際網路分享資訊。

人臉辨識系統用以擷取進入者臉部影像進行身分的辨識，其使用一基於特徵空間轉換與標準空轉換的方法進行識別。而在實際的應用裡，並非所有時間的影像皆適合進行人臉辨識，接著我們使用衣著色彩模型辨識系統進行人物之辨識。當人臉辨識系統辨識進入者身分的同時，亦建立進入者之色彩模型，再將色彩模型以及人物身分資訊傳送至色彩模型辨識系統，最後透過色彩模型辨識系統對於房間內每個人物的身分進行識別。

Person Identification in an In-door Environment

STUDENT: Tien-Chien Liao

ADVISOR: Dr. Jyh-Yeong Chang

Institute of Electrical Control Engineering
National Chiao-Tung University

ABSTRACT

We aim to implement an intelligent home-care system for nursing home purpose, in which person recognition is an important part in the system. In this thesis, we develop a person recognition system which monitors the room by the two sub-systems: face recognition system and dressing color model recognition system. Each sub-system utilizes a camera with the specific angle of view, but the two sub-systems share information with another system by the internet.

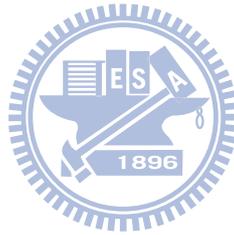
The face recognition system uses the capturing face image and then recognizes the person when one enters the room. The face recognition technique is based on eigenspace transform and canonical space transform. In fact, not all of the image frames are suitable for face recognition, hence, we also propose the color model recognition technique to recognize a person by the color of one's dress. Whenever the face recognition system is possible and done, we build this person's dress on immediately. Then we pass this color model to the color model recognition system. Finally, each person will be identified by the dress color models generated in the system.

致謝

本論文承蒙指導教授 張志永博士的悉心指導與督促，在每一次的討論當中，皆讓我獲益匪淺，使得論文中的方法與實驗皆能夠順利完成。

我也非常感謝實驗室的同學：懷顥、品宏、嘉臨、泳發、竣超以及家杰學長，有你們的配合，讓我能順利建立出一套人臉影像資料庫，使得實驗能順利的進行。

最後僅將此論文獻給我摯愛的家人，你們對我的支持是我奮發向上的原動力，給予我鼓勵、關心、支持與照顧，讓我可以無憂無慮的完成學業，在此希望能將我的這份榮耀與你們一同分享。



目次

摘要	i
ABSTRACT	ii
致謝	iii
目次	iv
圖次	vii
表次	x
第一章 緒論	1
1.1 研究動機及背景	1
1.2 人臉辨識	2
1.3 色彩模型辨識	5
1.4 論文架構	6
第二章 人臉辨識	7
2.1 人臉偵測	7
2.1.1 OpenCV	7
2.1.2 Haar Cascade Classifier	7
2.2 膚色偵測	11



2.3	人臉辨識	13
2.3.1	特徵空間轉換	14
2.3.2	標準空間轉換	16
第三章	人物辨識系統	19
3.1	人臉辨識系統	19
3.1.1	人臉辨識	19
3.1.2	色彩模型建立	22
3.2	色彩模型辨識系統	24
3.2.1	建立背景模型與上半身區域分離	26
3.2.2	直方圖交集	30
3.2.3	直方圖激發	31
第四章	實驗結果	32
4.1	實驗環境	32
4.2	人臉辨識實驗	33
4.2.1	人臉偵測與人臉正規化	33
4.2.2	人臉辨識	34
4.2.3	即時人臉辨識	38



4.3 色彩模型辨識實驗	41
4.3.1 上半身區域擷取	41
4.3.2 色彩模型選用	42
第五章 結論與未來展望	53
參考文獻	54



圖次

圖 1.1. 系統方塊圖	2
圖 1.2. 人臉偵測方法分類	4
圖 2.1. 矩形特徵圖	8
圖 2.2. 矩形特徵偵測人臉特徵示意圖	8
圖 2.3. 積分影像示意圖	9
圖 2.4. 計算矩形特徵示意圖	10
圖 2.5. 層疊分類器處理流程圖	11
圖 2.6. 膚色在 $r-g$ 平面的分佈	12
圖 3.1. 人臉辨識方塊流程圖	19
圖 3.2. 人臉正直或傾斜與嘴唇、鼻尖連線關係圖(a)未傾斜的人臉 (正臉), (b) 傾斜的人臉	20
圖 3.3. 人臉旋轉與嘴唇、鼻尖連線關係圖 (a)未旋轉的人臉, (b)旋轉的人 臉	20
圖 3.4. 人臉特徵範圍圖, 鼻尖位於(25, 30), 如同人臉的中心	20
圖 3.5. 人臉邊界範圍示意圖, (a)為偵測出之人臉範圍, 此範圍無法與人臉邊 界貼齊, (b)為修正過之人臉範圍, 其與人臉邊界貼 齊	22

圖 3.6.	擷取之上半身範圍與人臉區域範圍之比例圖	22
圖 3.7.	色彩模型辨識系統將與人臉辨識系統的資訊進行連結並存入資料庫...	25
圖 3.8.	色彩模型辨識系統將畫面內之前景影像依序比對色彩模型資料庫，以辨識人物身份	26
圖 3.9.	使用前景之二值化影像投影至 X、Y 軸，藉以求出前景物件之外切矩形參數	28
圖 3.10.	前景影像之外切矩型與上半身區域範圍關係圖，(a)為前景外切矩形與上半身區域，(b)為前景外切矩形與上半身區域之比例關係圖	29
圖 3.11.	上半身區域與取樣區域示意圖，(a)人物影像，(b)從(a)上半身擷取之五個影像方塊	30
圖 4.1.	實驗環境，(a)人臉辨識監視之畫面，(b)色彩模型辨識之監視之畫面 ...	32
圖 4.2.	人臉偵測圖，(a)人臉重複偵測的結果，(b)消除人臉重複偵測的結果	33
圖 4.3.	鼻尖與嘴唇的偵測，圖中綠色矩形為被鼻尖與嘴唇中間位置，藍色線為鼻尖與嘴唇的連線，白線為水平線。(a)未傾斜的人臉，(b)傾斜的人臉	34
圖 4.4.	人臉旋轉的判定，(a)-(b)為鼻尖與嘴唇偵測位置圖，(c)-(d)為鼻尖與嘴唇於人臉膚色邊界位置關係	34
圖 4.5.	人臉資料庫影像範例	35

圖 4.6. 前八個特徵臉	35
圖 4.7. 特徵向量之能量分佈	36
圖 4.8. 特徵向量之累積能量分佈	36
圖 4.9. 人臉影像於特徵空間之分佈	37
圖 4.10. 人臉影像於標準空間之分佈	37
圖 4.11. 即時人臉辨識之辨識流程	38
圖 4.12. 即時人臉辨識系統介面	38
圖 4.13. 家庭成員與非家庭成員之人臉影像於標準空間之分佈	39
圖 4.14. 不同 k 值對於非家庭成員之排除率	40
圖 4.15. 不同 k 值對於家庭成員之辨識率	40
圖 4.16. 不同 k 值排除率與辨識率之總和	41
圖 4.17. 人物之外切矩形偵測，(a)畫面影像，(b)前景分佈，(c)前景分佈投影至 X 軸，(d)前景分佈投影至 Y 軸，(e)擷取出人物之外切矩 形	43
圖 4.18. 比較 YC_bC_r 與 rgb 之正確率	51

表次

表 4.1. 比較 EST 以及 EST+CST 之辨識效果	37
表 4.2. 身體區域之偵測結果	44
表 4.3. 色彩模型於 RGB 之分佈	45
表 4.4. 色彩模型於 rgb 模型之分佈	46
表 4.5. 比較淺色系與深色系可區分的程度於 rgb 模型	47
表 4.6. 色彩模型於 $YCbCr$ 之分佈	48
表 4.7. 比較同為淺色系或深色系可區分的程度於 $YCbCr$ 空間	48
表 4.8. 比較淺色系與深色系間可區分的程度於 rgb 模型	49
表 4.9. 比較淺色系與深色系間可區分的程度於 $YCbCr$ 空間	49
表 4.10. 色彩模型於 HSV 之分佈	50



第一章 緒論

1.1 研究動機及背景

在日常生活中，多媒體已經被廣泛的應用在教育、居家保全、娛樂以及生醫等方面，並對客戶端提供多元的服務。在近年來，智慧型居家照護系統是個熱門的研究學門，其廣泛地應用電腦視覺技術以建構出一良好之人機互動環境。

本篇論文以發展智慧型居家照護系統為目標，對於老人之居住環境提供監控服務，藉著監控老人的生活作息形態，及觀察生活作息的變化關係，可作為老人健康狀態的一項指標。此外，也能監視老人們異常的動作，例如老人摔倒時，系統將即時通報相關單位，在第一時間予以處理。在應用面上，此系統亦可擴展成提供居家保全之服務，監視居家環境的進入者，以避免異常入侵者的進入，影響居家安全。為了解決這個問題，我們在居家環境中以兩部攝影機進行監視，而依照攝影機不同的工作需求，裝設在不同位置，負責判斷進入者身分的攝影機將裝設於近門口處；負責判斷居家環境內人物動作的攝影機，則裝設於離門口較遠處。我們結合了人臉影像以及穿著服裝顏色作為特徵，實現一即時身分辨識與追蹤及其動作辨識系統，以掌握居家環境中每個人之身分，接續之整合動作辨識技術，隨時監控居家環境內人們的動作以及生活作息。

在此篇論文中，我們以人臉偵測、人臉辨識以及色彩模型辨識等技術，希望在特定環境中掌握人們的身分以及所在的位置。完整的系統架構如圖 1.1 所示，本系統主要以兩部攝影機進行監視工作，一部裝設靠近門邊，負責監視居家環境進入者的身分；另一部裝設於離門口較遠處，用以判斷整個居家環境中每個人的位置與動作。我們的系統主要可分為人臉辨識與色彩模型辨識兩個部分，於人臉辨識部分，我們應用 Viola 等人[1]所發表之人臉偵測技術，偵測房間環境中進入者之臉部影像，隨後使用結合了特徵空間轉換 (Eigenspace Transformation, EST) 及標準空間轉換 (Canonical Space Transformation, CST) 之人臉識別技術[2]辨識來訪者之身分。於人物追蹤的部份，當來訪者身分辨識完成後，隨即將其所穿著

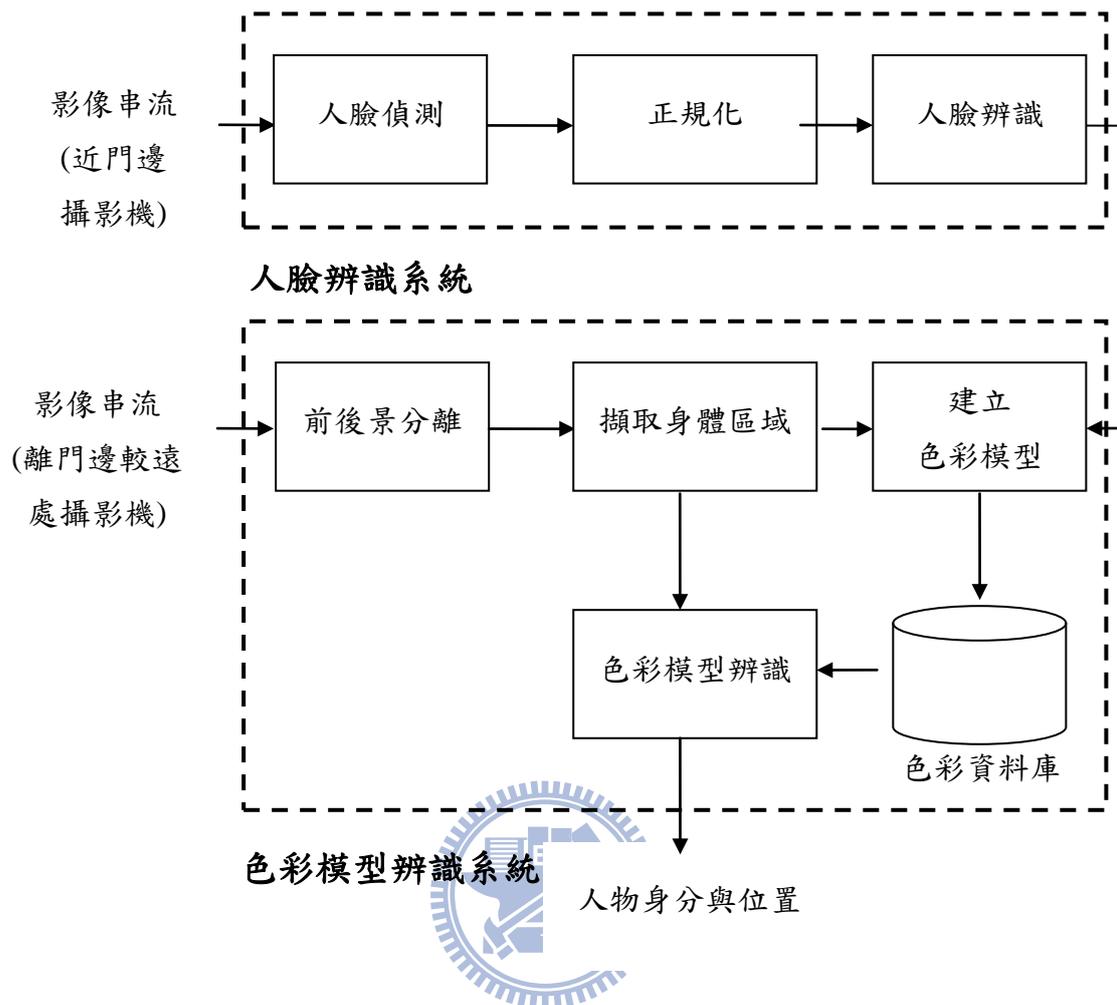


圖 1.1. 系統方塊圖

之服裝顏色建立成色彩模型。另一方面，我們使用一基於前後景影像比值而建立之背景統計模型將環境內之人物抽取出來，透過先前所建立之穿著服裝色彩模型分辨環境內各個人物所在之位置予以監控。

1.2 人臉辨識

人臉辨識屬於生物認證 (Biometrics) 的一種識別技術，利用每個人特有之臉部特徵進行身分的識別。除了人臉辨識以外，生物認證用於辨識身分也包含了常見的指紋、掌紋、視網膜以及語音等識別技術。人類透過眼睛觀察人臉上的特徵，經由對於人臉特徵的記憶，即可判定不同人的身分。而在電腦視覺裡的人臉辨識技術則是透過攝影機擷取人臉影像，經由比對預先建立的人臉辨識模型，藉此判

定出影像中人物的身分。相較於其他的生物認證技術，人臉辨識為一種更貼近自然的辨識方法，其近似於人類辨識人臉之原理。此外人臉辨識有別於指紋以及視網膜等生物認證技術，人臉辨識屬於一種非接觸式的辨識方式，辨識過程中被辨識者不需要按壓指紋或是貼近設備，甚至可以在不被人們察覺的狀態下完成辨識，也因此可避免人們對於身分辨識的偽裝與反感。

在電腦視覺裡，人臉辨識技術屬於一種整合型的技術領域，其中最重要的兩個部份分別為人臉偵測與人臉辨識。人臉偵測的目標在於找出畫面中人臉的位置，使人臉能從背景中分離出來，一般人臉辨識技術可分為：以特徵為基礎 (Feature-based) 以及以影像為基礎 (Image-based) 兩大類[3]，其細節可參考圖 1.2。

以特徵為基礎的人臉偵測藉由分析人類臉部結構上之特徵，例如：輪廓、顏色、形狀以及五官的幾何關係作為人臉判定之依據，其可分為 Low-level analysis、Feature analysis 以及 Active shape models 三類。

以影像為基礎的人臉偵測使用圖形識別的方式進行偵測，一般來說都會先以人臉影像以及非人臉影像經由一個學習的機制，依人臉影像與非人臉影像兩個群體間的差異作為判定人臉的依據，其可分為 Linear subspace methods、Neural networks 以及 Statistical approaches 三類。Chiang 等人[4]使用膚色進行人臉偵測，並輔以臉部眼睛與嘴唇的幾何關係排除非人臉的影像以提昇偵測效能。Rowley 等人[5]使用神經網路的方法進行人臉偵測，其將人臉影像區分為數個不同形狀、大小以及尺度的子區域影像進行學習，而將待偵測之影像縮放為不同尺度進行偵測，並且提供了一系列的處理方式解決臉部影像亮度不均勻；消除人臉在不同尺度重複偵測。對於背景影像對偵測效能的影響，其定義一橢圓遮罩加以排除。Wu 等人[6]統計膚色及髮色之分佈模型，再透過預先建立之人臉膚色與髮色分佈位置之樣板以 Fuzzy pattern matching 的方式進行人臉偵測，其效能更優於單獨使用膚色進行人臉偵測。

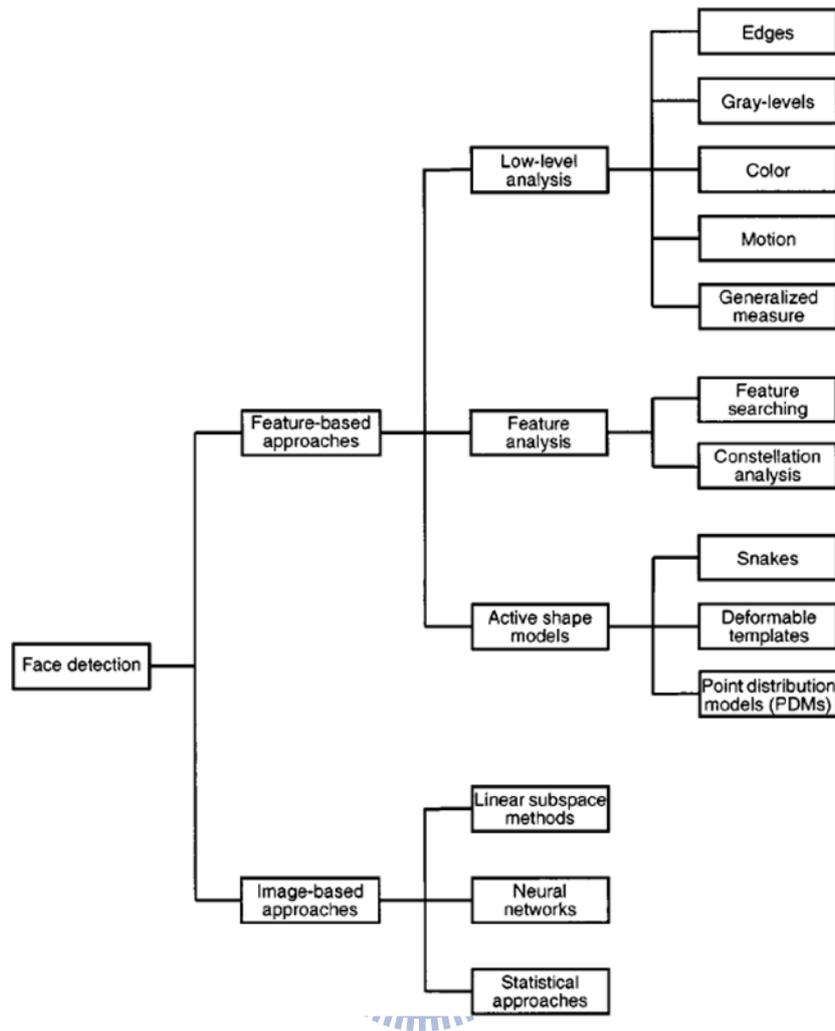


圖 1.2. 人臉偵測方法分類[3]

在人臉辨識上，Turk 等人[7] 透過主成份分析(Principal Component Analysis, PCA) 的方式抽取臉部特徵進行人臉辨識，即為著名的特徵臉 (Eigenfaces)。Etemad 等人[8]應用線性鑑別分析 (Linear Discriminant Analysis, LDA) 在人臉辨識以及性別辨識，比起主成份分析，線性鑑別分析能拉近相同群體間的距離且分散相異群體間的距離以提昇辨識效能。Moghaddam 等人[9]使用支持向量機 (Support Vector Machine, SVM) 進行人臉影像之性別辨識，並與放射狀基礎函數 (Radial Basis Function, RBF) 網路進行性別辨識之效能進行比較。

本研究採用 Viola 等人[1]所提之人臉偵測技術進行人臉偵測，其為包含了積分影像 (Integral image)、矩形特徵 (Haar-like feature) 以及 Adaboost 演算法的一種人臉偵測技術。在完成了人臉偵測後，我們結合了 Soriano 等人[10]的膚色偵

測技術，用以檢測並正規化待辨識之人臉影像，而後利用 Huang 等人[2]的人臉辨識方法，其將人臉影像透過特徵空間轉換以及標準空間轉換兩種方式，將人臉影像投影至標準空間進行人臉辨識。基於主成份分析的特徵空間轉換能夠擷取出臉部影像最具代表性的主要成份，對於縮減資料量是個強而有力的工具，在縮減資料量的同時，也同時能保留資料的最大代表性。基於標準分析的標準空間轉換能同時增大類別間的分散度與同類間聚集度，進而增強分類器表現，結合了這些特性使得人臉辨識能夠有良好的效能。

1.3 色彩模型辨識

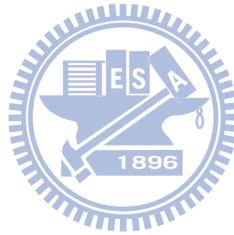
色彩模型辨識之目的在於使系統能夠掌握畫面內各個人物的身分與動向，進而判定人們日常生活中的行為以及所遇到的危險，例如判定一個人唸書的時間以及使用電腦的時間，要是當人不小心跌倒時也可以由系統判定出來，即時進行適當的應變措施。

在攝影機所能監視的影像範圍中，並非所有畫面內的位置都適合使用人臉辨識來進行身分的判定，因此勢必將取得臉部影像以外的特徵作為身份判定的依據。Swain 等人[11]介紹了利用色彩的物件辨識方式，其談到了直方圖交集 (Histogram Intersection) 的方式對數種物件進行辨識，並且分析了不同色彩空間辨識之效能。

我們提出了一種搭配人臉辨識而運作的人物辨識方法，我們在實驗的環境內架設兩部攝影機，一部設置於入口附近負責人臉辨識，辨識入口處進入者的身份。另一部則負責色彩模型辨識，其架設在較高處以監視整個環境且追蹤各個人物的動向與行為。當進入者由人臉辨識系統完成人臉辨識確認身份後，隨即由色彩模型辨識系統依其所穿著之服裝顏色建立色彩模型，而後藉由此色彩模型追蹤畫面內人物的行為及動向。

1.4 論文架構

本論文共分五章，在第一章中介紹研究相關的背景與動機，說明系統流程以及回顧相關的研究。第二章介紹人臉辨識的相關技術，包含人臉偵測、特徵空間轉換、標準空間轉換以及膚色偵測。在第三章中我們會談到人臉辨識系統與色彩模型辨識系統的整合，並且說明在實驗環境中利用色彩模型辨識各個人物身分所用之方法。第四章則是介紹我們的實驗環境以及實際測試的相關數據。第五章為結論的部分，統整我們的研究並探討未來展望。



第二章 人臉辨識

在這個章節裡，將介紹我們應用的人臉偵測與人臉辨識方法，人臉及人臉特徵的偵測使用 OpenCV[12]之 Haar cascade classifier，其原理是建構於 Viola 等人[1]的物件偵測技術之上，再以膚色輔以人臉特徵做人臉影像正規化。在人臉辨識部分，使用特徵空間轉換與標準空間轉換將待辨識的影像轉換至標準空間，計算其與預先建立之人臉辨識模型之歐基里德距離，用以判斷人臉的身份。

2.1 人臉偵測

2.1.1 OpenCV

OpenCV 全名為 Open Source Computer Vision Library[12]，由英特爾公司於 1999 年發起之專案，以開發即時電腦視覺系統為目標，所開發之跨平台電腦視覺程式庫。於 2000 年 OpenCV 第一版首次於 IEEE Conference on Computer Vision and Pattern Recognition 發表，目前版本為 2.1 版。



2.1.2 Haar Cascade Classifier

我們使用 OpenCV[12]所提供之 Haar cascade classifier 進行人臉偵測的工作，其演算架構是基於 Viola 等人[1]所提之物件偵測方法，能夠快速地進行即時物件偵測。此種方法需使用大量的正樣本(Positive samples)以及負樣本(Negative samples)進行訓練，在我們的應用裡正負樣本分別為人臉與非人臉，而我們使用 [13]以及[14]的訓練結果進行人臉、嘴唇與鼻尖的偵測。

Viola 等人[1]所提之人臉辨識方法採用矩形特徵，其用以辨識人臉的灰階影像。如圖 2.1 所示，矩形特徵是由數個白色以及灰色的矩形區域所組成，其以兩種區域各自所涵蓋之像素灰階值總和之差作為在影像之矩形特徵值。如圖 2.2 所示，透過不同的矩形特徵可以表現出人臉區域內各個部分灰階值的變化，例如圖

2.1(b) 之矩形特徵可用於偵測圖 2.2(b) 處之臉部特徵；圖 2.1(c) 之矩形特徵可用於偵測圖 2.2(c) 之臉部特徵。

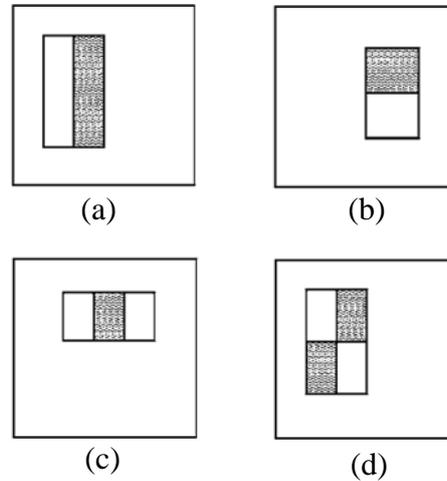


圖 2.1. 矩形特徵圖[1]

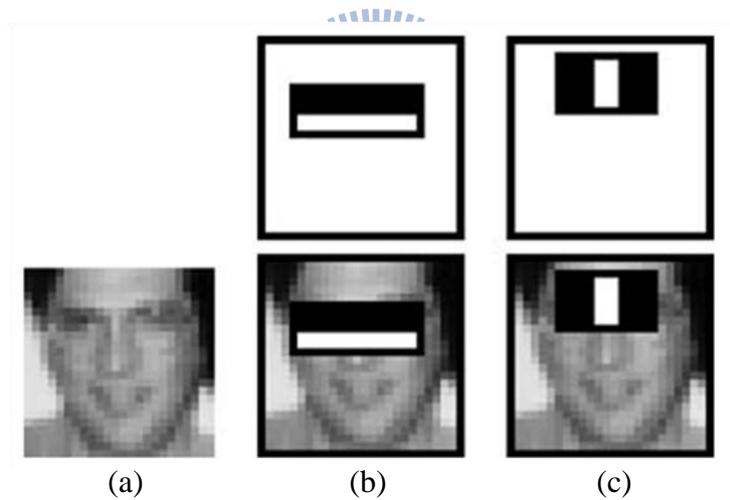


圖 2.2. 矩形特徵偵測人臉特徵示意圖[1]

積分影像 (Integral image) 是一個能加速計算矩形特徵的方式，在計算矩形特徵前先將原始影像轉換成積分影像，即可透過積分影像的特性，不需經由造訪每個像素點而求得矩形特徵。

考慮原始灰階影像 $i(x, y)$ ，積分影像 $ii(x, y)$ 之值定義為

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y') \quad (2.1)$$

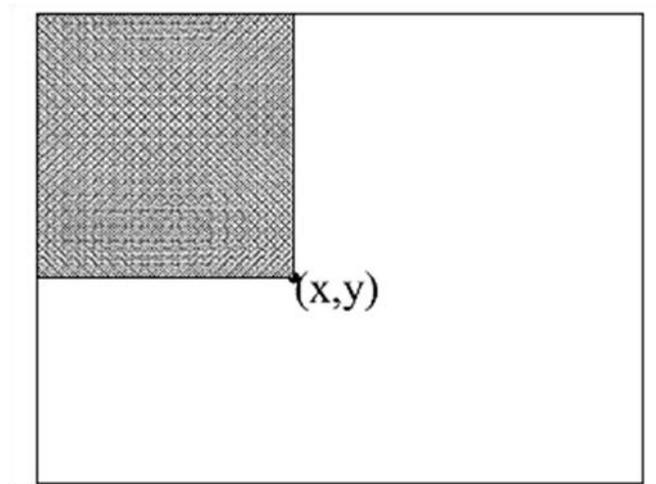


圖 2.3. 積分影像示意圖[1]

如圖 2.3 所示，積分影像值即為灰色區域所覆蓋之原始影像灰階值總和，我們能夠使用下列兩組式子疊代求得積分影像

$$s(x, y) = s(x, y-1) + i(x, y) \quad (2.2)$$

$$ii(x, y) = ii(x, y-1) + s(x, y) \quad (2.3)$$

其中 $s(x, y)$ 為原始影像於座標 (x, y) 的列累積總和

$$s(x, y) = \sum_{y' \leq y} i(x, y') \quad (2.4)$$

且 $s(x, -1) = 0$ 以及 $ii(-1, y) = 0$ ，因此僅需掃描過原始影像一次即可求得積分影像。

在求得積分影像後，計算每個矩形特徵可透過積分影像的特性減少運算量，如圖 2.4 所示，圖中位置 2 於積分影像的數值代表著原始影像 A、B 區域灰階值的總和，位置 3 代表著原始影像 A、C 區域之總和，因此當我們要計算 D 區域總和時，可以藉由積分影像數值 $4 + 1 - (2 + 3)$ 求得，其大幅減少運算矩形特徵

時所需之運算量，這也就是 Viola 等人[1]所提之人臉偵測方法其能夠快速的即時運算之原因。

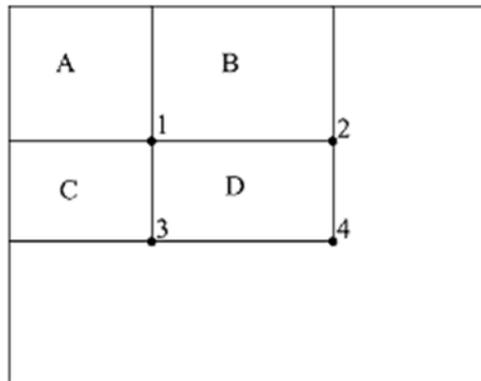


圖 2.4. 計算矩形特徵示意圖[1]

在學習的過程使用一個類似 Adaboost 方法來選擇適當的矩形特徵以及訓練分類器。Adaboost 是一種疊代的演算法，其應用於人臉偵測的目標在於從大量的弱分類器 (Weak classifier) 中挑選出一個最具分類意義的組合，形成一個強分類器 (Strong classifier)。在人臉偵測上是使用矩形特徵且搭配門檻值而形成弱分類器，選取一些具有較好分類效果的矩形特徵，使其搭配適當的權重組合成一個強分類器，其中每個弱分類器不需要極高的辨識率，一般來說只要其辨識率大於 50% 即可算是一個較好的弱分類器。最後串起數個強分類器組合成一個疊疊分類器 (Cascade classifier)。

於人臉偵測階段，從原始影像依照不同尺度依序掃描取出影像裡的子區塊影像，再將其通過數個層疊分類器進行偵測，當子區塊影像被其中一個分類器辨識為非人臉時該影像即被排除；而當一個子影像區塊能夠通過越多分類器時，即代表其屬於人臉的可能性越大，如圖 2.5 所示。

由於偵測的結果可能會是單個人臉被重複偵測出數次，我們選擇將重疊的人臉區塊範圍取平均表示成單一個人臉。在即時的實做過程裡可能會因為雜訊的影響造成某些非人臉的區塊在連續畫面中的某幾個影像產生誤判，我們使用一種簡單的主要投票法 (Majority vote) 來排除這些非人臉，於連續影像中人臉在相鄰

的兩個影像間所偵測出的範圍通常會是重疊的，假若某個偵測出的人臉，其範圍附近於前三個影像亦有被偵測出人臉的紀錄，我們才將其人臉影像資料傳給後級做處理。

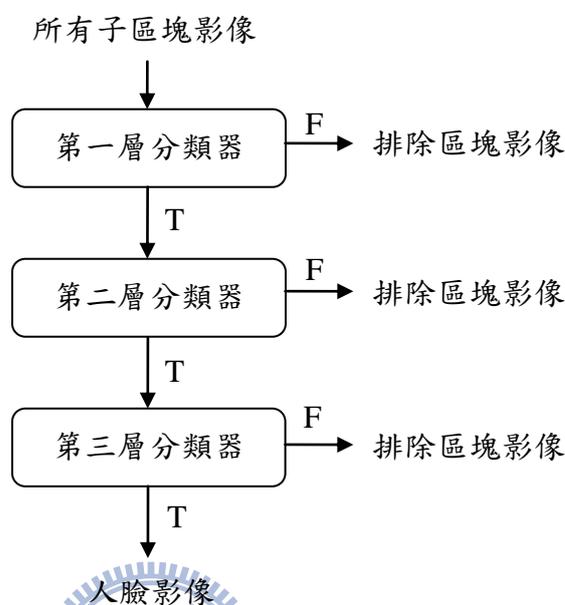


圖 2.5. 層疊分類器處理流程圖[1]

2.2 膚色偵測

為從彩色影像中偵測膚色及唇色像素，本篇專題參考 Soriano 等人[10]所提出在正規化 RGB 色彩空間(Normalized RGB)中的膚色偵測範圍。在 RGB 色彩空間所顯示的物體顏色，會受到光線亮度的強弱影響，經過光線量度正規化之後則可以減少物體顏色對於光線亮度的敏感度。其膚色偵測則是運用正規化 RGB 空間中的 r 、 g 進行分析，正規化公式如下：

$$\begin{aligned}
 r &= \frac{R}{R+G+B} \\
 g &= \frac{G}{R+G+B}
 \end{aligned}
 \tag{2.5}$$

圖 2.6 顯示膚色在 $r-g$ 平面可以用一對二次多項式來定義其上、下邊界，此條件式被描述為

$$f_{\text{lower}}(r) = -0.776r^2 + 0.5601r + 0.1766 \quad (2.6)$$

$$f_{\text{upper}}(r) = -1.3767r^2 + 1.0743r + 0.1452 \quad (2.7)$$

$$R1: g > f_{\text{lower}}(r) \text{ and } g > f_{\text{upper}}(r) \quad (2.8)$$

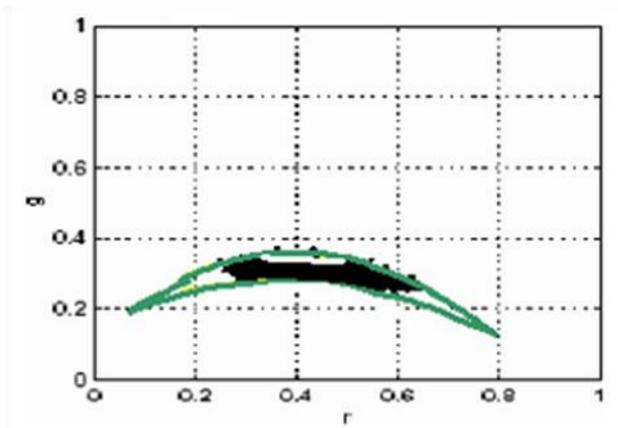


圖 2.6. 膚色在 $r-g$ 平面的分佈[10]

然而白色點 ($r = g = 0.33$) 可能落在兩函數包圍的區域之內，故作以白色點為中心、0.0004 為半徑的圓，要求像素點需落在圓外，以避免白色或接近白色的像素被偵測為膚色，此條件式被描述為：

$$R2: W = (r - 0.33)^2 + (g - 0.33)^2 \geq 0.0004 \quad (2.9)$$

由於 R1、R2 兩個條件並不能準確的濾除非膚色區域，因此再以下列兩個條件式來濾除黃綠色、藍色以及橘色區域：

$$R3: R > G > B \quad (2.10)$$

$$R4: R - G \geq 45 \quad (2.11)$$

總結膚色的條件式如下：

$$S = \begin{cases} 1, & \text{if all } R_1, R_2, R_3 \text{ and } R_4 \text{ are true,} \\ 0, & \text{otherwise.} \end{cases} \quad (2.12)$$

如果 $S=1$ 則被偵測為膚色像素。

2.3 人臉辨識

在我們的研究裡以人臉的灰階影像進行訓練與辨識，其轉換公式為

$$I = 0.2989 \cdot R + 0.5866 \cdot G + 0.1145 \cdot B \quad (2.13)$$

假設將有 c 類的資料要進行訓練，各個類別代表著不同的人，而訓練的資料即為每個人的人臉影像， $\mathbf{x}'_{i,j}$ 可以表示成第 i 個人的第 j 張人臉影像，而 N_i 表示第 i 個人所訓練的人臉影像數量，而所有訓練的人臉影像數可以表示成 $N_T = N_1 + N_2 + \dots + N_c$ ，其代表第 1 至 c 個人其訓練影像數的總和，而所有訓練的人臉影像所成的集合可以表示為

$$[\mathbf{x}'_{1,1}, \dots, \mathbf{x}'_{1,N_1}, \dots, \mathbf{x}'_{2,1}, \dots, \mathbf{x}'_{c,N_c}] \quad (2.14)$$

其中 $\mathbf{x}'_{i,j}$ 都是具有 n 個像素的影像。

首先我們將每張待訓練的人臉影像亮度進行正規化

$$\mathbf{x}_{i,j} = \frac{\mathbf{x}'_{i,j}}{\|\mathbf{x}'_{i,j}\|} \quad (2.15)$$

接下來計算所有影像的平均灰階值

$$\mathbf{m}_x = \frac{1}{N_T} \sum_{i=1}^c \sum_{j=1}^{N_i} \mathbf{x}_{i,j} \quad (2.16)$$

而訓練影像集合可以寫成一個維度為 $n \times N_T$ 的矩陣 \mathbf{X} ，其由各個訓練影像 $\mathbf{x}_{i,j}$ 減去平均灰階值 \mathbf{m}_x 所構成

$$\mathbf{X} = [\mathbf{x}_{1,1} - \mathbf{m}_x, \dots, \mathbf{x}_{1,N_1} - \mathbf{m}_x, \dots, \mathbf{x}_{c,N_c} - \mathbf{m}_x] \quad (2.17)$$

2.3.1 特徵空間轉換

特徵空間轉換已經被廣泛的應用在資料壓縮上，其原理是將資料從彼此相關性高的高維空間映射到相關性較低的低維空間，且此過程將最小化其均方誤差以保留資料的完整性。在這個過程中，將會尋找對原始資料具有最大變異的方向、第二大變異的方向...等前幾大變異的方向，以這幾個方向組成一組新的座標基底，使得原始資料在這個子空間中，能夠將資料彼此間的差異性表現得最大。在運算上以原始資料所構成之共變異數矩陣 (Covariance matrix) 的特徵值 (Eigenvalue) 與特徵向量 (Eigenvector) 作為依據，旋轉座標軸使得原始資料具有最大的變異度。

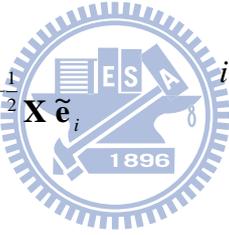
若矩陣 $\mathbf{X}\mathbf{X}^T$ 的秩數 (Rank) 為 K ，則其具有 K 個非零的特徵值分別為 $\lambda_1, \lambda_2, \dots, \lambda_K$ ，且與其相對應之特徵向量 $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_K$ 滿足

$$\lambda_i \mathbf{e}_i = \mathbf{R} \mathbf{e}_i, \quad i = 1, 2, \dots, K \quad (2.18)$$

其中 $\mathbf{R} = \mathbf{X}\mathbf{X}^T$ 為對稱矩陣。為了求解 Eq.(2.18)，我們需要求解維度為 $n \times n$ 的矩陣 $\mathbf{X}\mathbf{X}^T$ 之特徵值與特徵向量，但是由於 $\mathbf{X}\mathbf{X}^T$ 的維度太大難以計算，基於奇異值分解定理 (Singular value decomposition theory)，而以 $\tilde{\mathbf{R}}$ 進行取代

$$\tilde{\mathbf{R}} = \mathbf{X}^T \mathbf{X} \quad (2.19)$$

此時 $\tilde{\mathbf{R}}$ 的維度為 $N_T \times N_T$ ，大幅的減低了運算的困難度。假設 $\tilde{\mathbf{R}}$ 具有 K 個非零的特徵值 $\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_K$ ，且其相對應的特徵向量為 $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \dots, \tilde{\mathbf{e}}_K$ ，則其與 \mathbf{R} 的關係滿足

$$\begin{cases} \lambda_i = \tilde{\lambda}_i \\ \mathbf{e}_i = (\tilde{\lambda}_i)^{-\frac{1}{2}} \mathbf{X} \tilde{\mathbf{e}}_i \end{cases} \quad i = 1, 2, \dots, K \quad (2.20)$$


因此我們可透過這 K 個特徵向量形成一組正交基底，其可展延(Span)出一個新的向量空間，而每張訓練之人臉影像可以被投影至這個新的向量空間以一個點表示。根據主成份分析理論，每張訓練影像可以由較大的前 k 個特徵值 $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_k|$ ， $k \leq K$ 及其相對應之特徵向量 $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k$ 近似。由這 k 個特徵向量所組成的部份集合可以展延出一個特徵空間 (Eigenspace)，而每張影像 $\mathbf{x}_{i,j}$ 投影至這個特徵空間以點 $\mathbf{y}_{i,j}$ 表示

$$\mathbf{y}_{i,j} = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k]^T \mathbf{x}_{i,j} \quad i = 1, 2, \dots, c; j = 1, 2, \dots, N_c \quad (2.21)$$

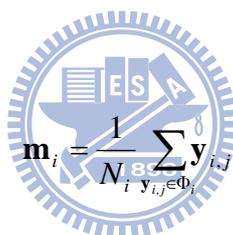
我們稱 $[\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k]^T$ 為特徵空間轉換矩陣 (Eigenspace transformation matrix)，經由了特徵空間轉換，每張影像 $\mathbf{x}_{i,j}$ 可以由這 k 個特徵向量的線性組合近似，形成特徵空間中的一個點 $\mathbf{y}_{i,j}$ ，亦即 $\mathbf{y}_{i,j}$ 為一個具有 k 個係數的一維向量。

2.3.2 標準空間轉換

基於標準分析理論，假設 $\mathbf{y}_{i,j}$ 為第 i 類的第 j 個向量，所有類別之向量平均可以表示為

$$\mathbf{m}_y = \frac{1}{N_T} \sum_i \sum_j \mathbf{y}_{i,j} \quad i=1, 2, \dots, c; j=1, 2, \dots, N_i \quad (2.22)$$

而僅考慮第 i 類的向量平均為



$$\mathbf{m}_i = \frac{1}{N_i} \sum_{\mathbf{y}_{i,j} \in \Phi_i} \mathbf{y}_{i,j} \quad (2.23)$$

令 \mathbf{S}_w 為類別內散佈矩陣 (Within-class matrix)； \mathbf{S}_b 為類別間散佈矩陣 (Between-class matrix)，其分別表示為

$$\mathbf{S}_w = \frac{1}{N_T} \sum_{i=1}^c \sum_{\mathbf{y}_{i,j} \in \Phi_i} (\mathbf{y}_{i,j} - \mathbf{m}_i)(\mathbf{y}_{i,j} - \mathbf{m}_i)^T \quad (2.24)$$

$$\mathbf{S}_b = \frac{1}{N_T} \sum_{i=1}^c N_i (\mathbf{m}_i - \mathbf{m}_y)(\mathbf{m}_i - \mathbf{m}_y)^T \quad (2.25)$$

其中 \mathbf{S}_w 代表類別內向量距離的平均， \mathbf{S}_b 代表各個類別間距離的平均，而標準空間轉換的目標在於最小化 \mathbf{S}_w ，使得類別內的向量距離最小；最大化 \mathbf{S}_b ，使得類

別間的距離最大。透過去最佳化目標函數 $\mathbf{J}(\mathbf{W})$ (generalized Fisher linear discriminant function) 我們可得其解

$$\mathbf{J}(\mathbf{W}) = \frac{\mathbf{W}^T \mathbf{S}_b \mathbf{W}}{\mathbf{W}^T \mathbf{S}_w \mathbf{W}} \quad (2.26)$$

在此我們要選取適當的 \mathbf{W} 使得 $\mathbf{J}(\mathbf{W})$ 為極大值，藉由

$$\frac{\partial \mathbf{J}}{\partial \mathbf{W}} = 0 \quad (2.27)$$

假設 \mathbf{W}^* 為最佳解，且其中的每個向量 \mathbf{w}_i^* 即為相對於第 i 大特徵值 λ_i 的特徵向量。根據[2]，eq.(2.28)可以表示成



$$\mathbf{S}_b \mathbf{w}_i^* = \lambda_i \mathbf{S}_w \mathbf{w}_i^* \quad (2.28)$$

於求解後選用 $c - 1$ 個非零特徵值及其相對應之特徵向量 $[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_c]$ ，藉由這組特徵向量可以產生另一組新的正交基底，即為標準空間 (Canonical space)，而於特徵空間中的每個向量 $\mathbf{y}_{i,j}$ 能夠被投影至標準空間，藉由

$$\mathbf{z}_{i,j} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{c-1}]^T \mathbf{y}_{i,j} \quad (2.29)$$

其中 $\mathbf{z}_{i,j}$ 表示投影至標準空間的各個點，而 $[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{c-1}]^T$ 表示於標準空間的各個分量，其被稱為標準空間轉換矩陣 (Canonical space transformation matrix)，基於上述分析，每個訓練的人臉類別可以被明顯的分離。

藉由合併 eq.(2.21)以及 eq.(2.29)，每張人臉影像可以直接投影至維度為 $c - 1$ 的標準空間

$$z_{i,j} = \mathbf{H} \cdot \mathbf{x}_{i,j} \quad (2.30)$$

其中 $\mathbf{H} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{c-1}]^T [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k]^T$ 。各張影像於標準空間中，其相同群體的將聚集在一起，相異群體將被分離開來。



第三章 人物辨識系統

我們所實做之人物辨識系統是由兩個子系統所組成，分別為人臉辨識系統以及色彩模型辨識系統。人臉辨識系統辨識環境內進入者的身分，在我們的研究中僅考慮正面的人臉作為辨識對象，辨識完成後隨即建立進入者上衣色彩模型，並連同進入者的身分傳送給進入者上衣色彩模型辨識系統。色彩模型辨識系統同時也建立色彩模型，並且連結人臉辨識系統所傳來的資訊後，將包含進入者身分的色彩模型存入資料庫，而後即時在環境中監視其動向。

3.1 人臉辨識系統

3.1.1 人臉辨識

人臉辨識是本研究中的第一個主要部分，負責辨識環境內進入者的身分，以提供後級系統作為辨識的參考依據，其方塊流程圖如圖 3.1 所示。

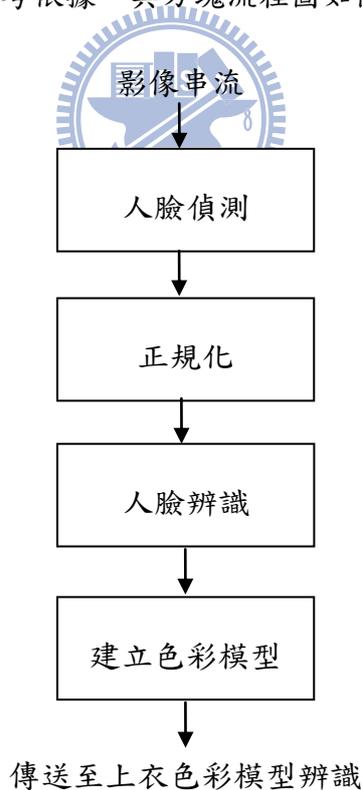


圖 3.1. 人臉辨識方塊流程圖

我們使用第二章所介紹之 Haar cascade classifier 進行人臉、嘴唇以及鼻尖的偵測。首先將攝影機擷取的影像先進行人臉偵測，偵測出來人臉的位置後，再對人臉影像進行正臉的判斷。正臉的判斷是透過偵測嘴唇、鼻尖以及膚色，藉由嘴唇與鼻尖所連成之直線的歪斜度，可以反應出人臉傾斜的程度，當嘴唇與鼻尖所連成之直線為垂直時，代表人臉無傾斜，如圖 3.2 所示。

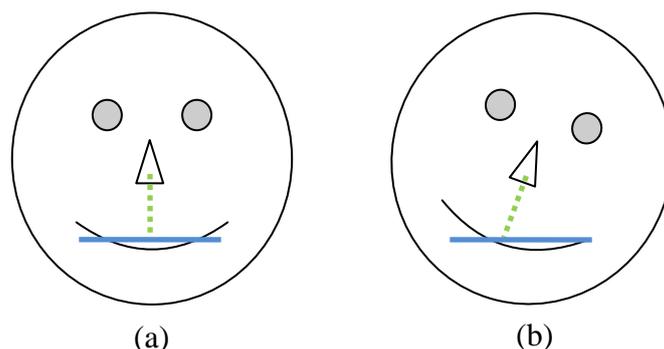


圖 3.2. 人臉正直或傾斜與嘴唇、鼻尖連線關係圖
(a)未傾斜的人臉 (正臉)，(b)傾斜的人臉。

膚色的偵測是為了判斷人臉旋轉的程度，如圖 3.3 所示，我們透過先前對嘴唇的偵測位置作為依據，判斷其與兩側膚色邊界距離的關係，以未傾斜且未旋轉的臉來說，嘴唇與兩側膚色邊界的距離大約是相等的，而我們將此區域區分為五個等份，當嘴唇位置落於中間的位置時，則判定其人臉影像為未旋轉的正臉。

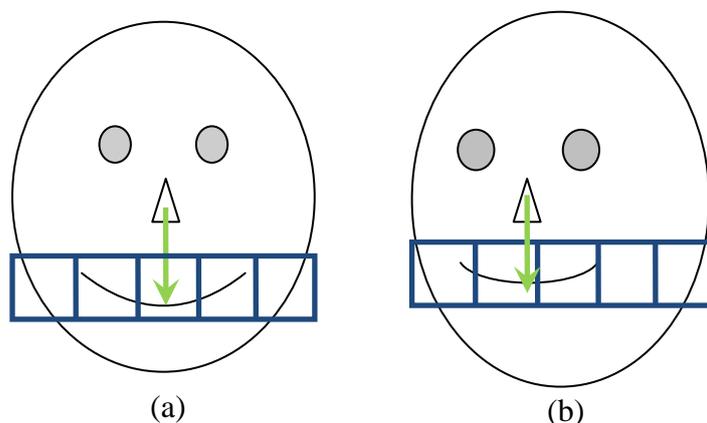


圖 3.3. 人臉旋轉與嘴唇、鼻尖連線關係圖
(a)未旋轉的人臉，(b)旋轉的人臉。

在確認此人臉影像為適合進行辨識的正臉後，由於偵測出之人臉範圍無法與人臉邊界貼齊，我們使用預先定義之人臉特徵距離作為裁切人臉的參考，由於人臉長寬比通常為 1.1 : 1 至 1.3 : 1，並且考慮攝影機與人臉辨識區域之距離後，我們將定義人臉特徵範圍圖，如圖 3.4 所示，人臉的長寬分別被定義為 60、50 個像素。

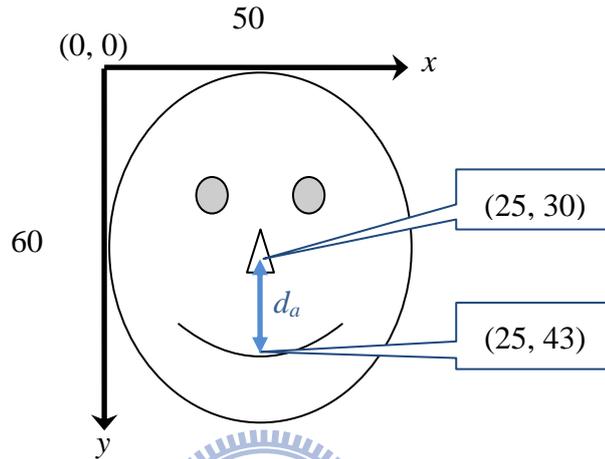


圖 3.4. 人臉特徵範圍圖，鼻尖位於(25, 30)，如同人臉的中心。

首先我們計算偵測出的人臉與預先定義的人臉內嘴唇與鼻尖距離的比值，如圖 3.4 所示， d_a 為預先定義之人臉其嘴唇與鼻尖的距離，其值為 13；如圖 3.5(a)， d_b 為偵測出之人臉其嘴唇與鼻尖的距離。則兩距離之比值為 $R_{ab} = \frac{d_b}{d_a}$ ，透過這個比值可以重新換算偵測出之人臉應由原始影像中裁切出來的邊界。假設人臉影像經由換算過之邊界參數為 $x_0, y_0, x_{width}, y_{height}$ ，如圖 3.5(b)所示，則各個參數將藉由比值的關係計算

$$\begin{aligned}
 x_{width} &= 50 \cdot R_{ab} \\
 y_{height} &= 60 \cdot R_{ab} \\
 x_0 &= x_{nose} - 25 \cdot R_{ab} \\
 y_0 &= y_{nose} - 30 \cdot R_{ab}
 \end{aligned} \tag{3.1}$$

其中 (x_{nose}, y_{nose}) 為偵測出鼻尖的座標位置。計算完人臉裁切的範圍後我們取出其人臉範圍影像，並將其尺寸正規化為 50×60 ，接著進行人臉辨識。

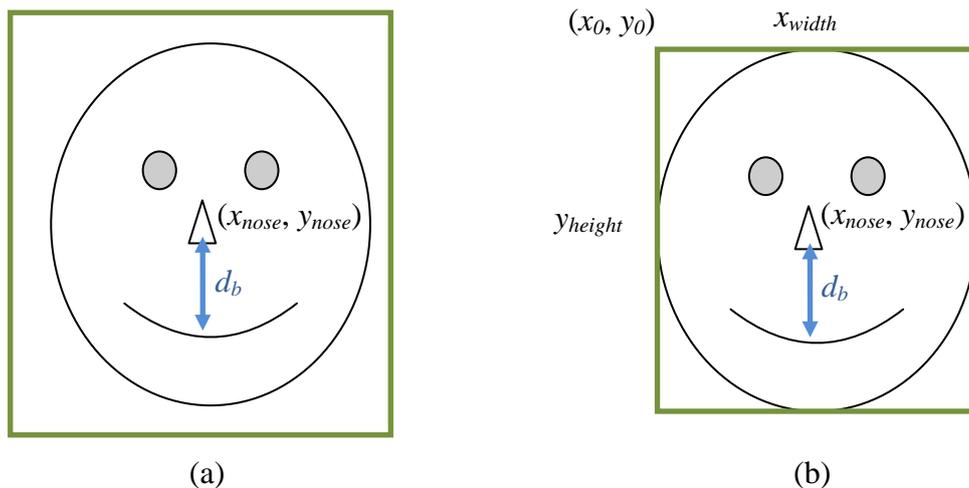


圖 3.5. 人臉邊界範圍示意圖，(a)為偵測出之人臉範圍，此範圍無法與人臉邊界貼齊，(b)為修正過之人臉範圍，其與人臉邊界貼齊。

人臉辨識使用我們預先建立之人臉辨識模型，於辨識階段我們將計算待辨識之人臉影像與人臉辨識模型中每個人於標準空間中的群中心的距離，我們使用歐基里德距離 (Euclidean distance) 進行量測，選擇距離最近的人作為辨識結果。

3.1.2 色彩模型建立

人臉辨識系統完成人臉辨識工作後將於畫面中以人臉下緣取兩倍人臉寬度為寬，三倍人臉高度為高，劃分出人體上半身區域以作為建立上半身衣服色彩模型之用，爾後簡稱色彩模型，如圖 3.6 所示。

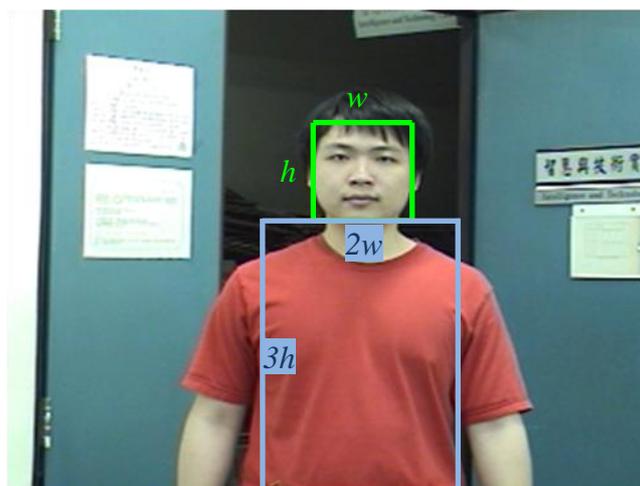


圖 3.6. 擷取之上半身範圍與人臉區域範圍之比例圖

建立色彩模型所使用的色彩空間我們做過了許多的嘗試，其中 YC_bC_r 與 HSV 這兩種色彩空間皆具有亮度與彩度分離的特性，而 YC_bC_r 在轉換速度比 HSV 快，且 HSV 可能會有色調不穩定的現象，因此我們選擇以 YC_bC_r 作為建立色彩模型之色彩空間。

YC_bC_r 色彩空間以 Y 表示亮度，範圍由 16 至 235 共 220 個區間， C_bC_r 分別表示相對於藍色及紅色的偏差值，其分佈範圍皆是 16 至 240 共 225 個區間，其與 RGB 之轉換關係為

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \quad (3.2)$$

透過將上半身區域影像由 RGB 轉換至 YC_bC_r ，再以三個通道獨立建立色彩模型，其資料大小為 $220 + 225 + 225 = 670$ ，若以三個通道一起建立模型，以每個通道量化成 10 個單位為例，其資料量為 $10 \times 10 \times 10 = 1000$ ，比起三個通道一起建立模型，獨立建立模型的優點在於能縮減色彩模型的資料量，不論對於資料庫的容量以及比對的時間都是有益的。

我們選用色彩之直方圖作為色彩模型，色彩之直方圖為統計影像內各個不同大小的顏色出現之次數，而以色彩直方圖作為模型具有下面數項優點：

- 對於觀察視角變換 (Translation) 以及旋轉 (Rotation)，直方圖皆有一定的穩健程度。例如以正面觀察一個物件以及 $\pm 45^\circ$ 觀察同個物件，對於色彩直方圖僅有些微的影響，其色彩分佈仍是於物件主要顏色處。
- 使用色彩直方圖描述物件的顏色能減少背景顏色的影響。在我們的研究中，建立色彩模型之影像範圍內大部分是人體上半身之影像，僅有少部份會參雜分割錯誤的背景，以直方圖建立色彩模型，背景顏色所佔之比重將遠小於上半身的顏色。

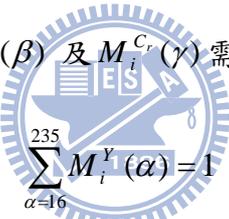
• 使用色彩直方圖進行辨識，較能容忍物件的遠近造成不同的大小的影響。我們使用離散機率密度模型描述色彩的分佈，即使擷取之畫面大小變動時，仍能正確的進行偵測。

我們分別統計上半身區域之 YC_bC_r 值且繪製其直方圖，並以總像素點數進行正規化以建立色彩之離散機率密度模型。假設人物人數為 n ，色彩模型為一三維向量 $(M_i^Y(\alpha), M_i^{C_b}(\beta), M_i^{C_r}(\gamma))$ ，建立於下述

$$\begin{aligned} M_i^Y(\alpha), & \quad 16 \leq \alpha \leq 235 \\ M_i^{C_b}(\beta), & \quad 16 \leq \beta \leq 240 \\ M_i^{C_r}(\gamma), & \quad 16 \leq \gamma \leq 240 \end{aligned} \quad (3.3)$$

其中 $i=1, 2, \dots, n$ ，以 $M_i^Y(\alpha)$ 來說代表著當像素點之 $Y=\alpha$ 時，該像素點屬於第 i

個人的機率，因此 $M_i^Y(\alpha)$, $M_i^{C_b}(\beta)$ 及 $M_i^{C_r}(\gamma)$ 需滿足



$$\begin{aligned} \sum_{\alpha=16}^{235} M_i^Y(\alpha) &= 1 \\ \sum_{\beta=16}^{240} M_i^{C_b}(\beta) &= 1 \\ \sum_{\gamma=16}^{240} M_i^{C_r}(\gamma) &= 1 \end{aligned} \quad (3.4)$$

於模型建立完成後，依此模型可作為辨識房間內人員使用。

3.2 色彩模型辨識系統

在房間環境下，上衣色彩模型人物辨識系統使用人物的衣服顏色作為特徵，接著辨識每個人的動向。色彩模型辨識系統的運作分為兩個部份，第一個部份為建立色彩模型，其工作在於比較人臉辨識系統所傳來之某人身分，並建立其之色彩模型，以完成色彩與人物身份之連結，並存入資料庫；第二個部份是人物的辨識，由於在監視整個環境的攝影畫面內較難對畫面內每個人進行人臉辨識以辨識

身份，我們將畫面內人物之服裝顏色依序比對資料庫的色彩模型，以在不需要偵測臉部影像的情況下，以上衣色彩模型進行身份的辨識。

如圖 3.7 所示，人臉辨識系統於完成進入者的身份辨識後，隨即建立衣服顏色辨識模型並傳送至色彩模型辨識系統，此時色彩模型辨識系統將畫面內的前景分離出來，並依人體特定的比例關係擷取出上半身的局部影像，同時取其影像建立衣服顏色辨識模型，並將建立之模型與人臉辨識系統所傳來之模型進行比對，以完成人物身份與色彩模型之連結，而後色彩模型辨識系統將已完成身份與色彩連結的色彩模型存入資料庫，以待辨識時使用。

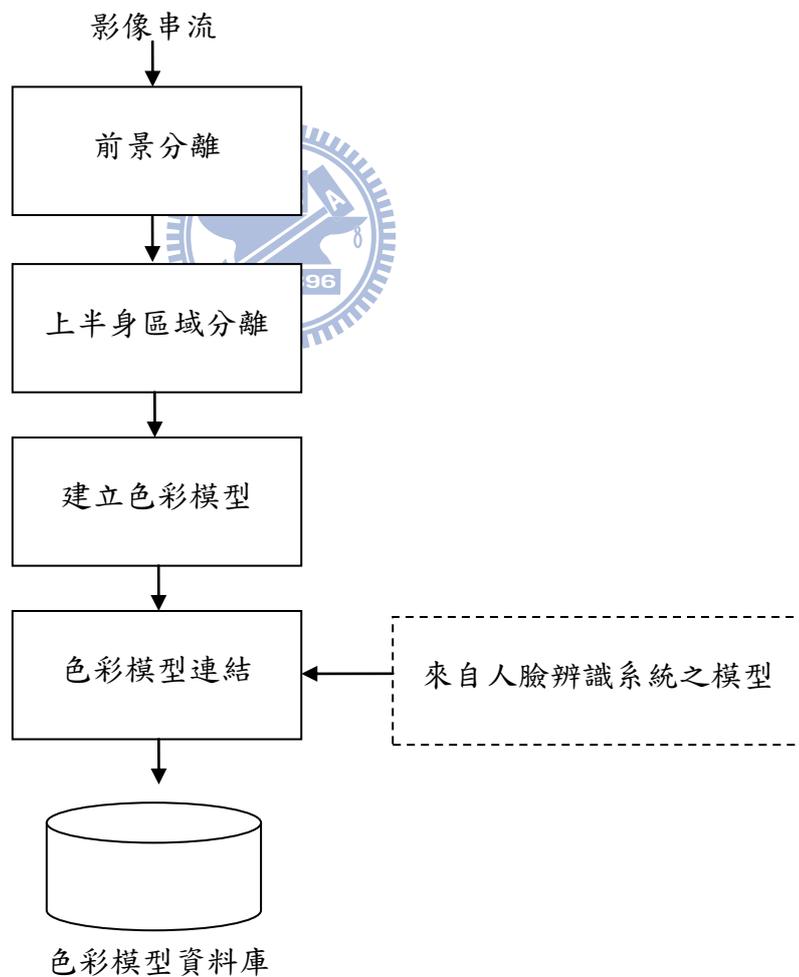


圖 3.7. 色彩模型辨識系統將與人臉辨識系統的資訊進行連結並存入資料庫

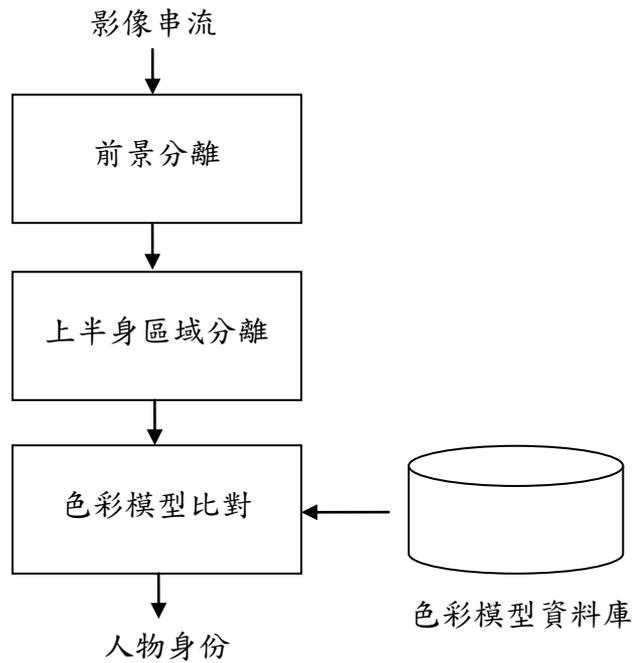


圖 3.8. 色彩模型辨識系統將畫面內之前景影像依序比對色彩模型資料庫，以辨識人物身份。

於辨識階段我們將分離出之前景影像擷取出上半身區域，並將影像轉換至 YC_bC_r 色彩空間，再依序比對色彩模型內的所有模型以辨識人物之身分。在色彩模型比對上我們介紹直方圖交集 (Histogram intersection) 以及直方圖激發 (Histogram activation) 兩種方式，詳細說明將在 3.2.2 與 3.2.3 節。

3.2.1 建立背景模型與上半身區域分離

假設由攝影機拍攝之影像序列所擷取的第 i 個畫面為

$$I_i(x, y) = S_i(x, y)r_i(x, y) \quad (3.5)$$

其中 (x, y) 為畫面中像素點的位置； I_i 為擷取出的畫面亮度； S_i 為光源亮度的分佈； r_i 為畫面內反射比的分佈，擷取出的畫面亮度為原始畫面內光源的亮度經由反射比的影響所呈現出來的結果。假設攝影機位置固定，且畫面內並無會移動的物件，則背景的反射比在每個時間點大致是相同的

$$r_i(x, y) = r(x, y) \quad (3.6)$$

即使反射比是不變的，但前後景之間亮度的關係仍是存在的。Frame ratio 為介於兩個畫面間亮度的比值，其於兩個連續畫面間的關係可以被寫成

$$\begin{aligned}\log\left(\frac{I_i(x,y)}{I_{i-1}(x,y)}\right) &= \log\left(\frac{S_i(x,y)r(x,y)}{S_{i-1}(x,y)r(x,y)}\right) \\ &= \log\left(\frac{S_i(x,y)}{S_{i-1}(x,y)}\right) \\ &= \log(S_i(x,y)) - \log(S_{i-1}(x,y))\end{aligned}\quad (3.7)$$

我們使用 Frame ratio 來建立背景模型，統計連續且無移動物件之背景影像之最小亮度值 $n(x, y)$ ；最大亮度值 $m(x, y)$ ；最大畫面間之 Frame ratio $d(x, y)$ ，藉此描述背景模型，如同

$$\begin{bmatrix} m(x, y) \\ n(x, y) \\ d(x, y) \end{bmatrix} = \begin{cases} \begin{bmatrix} \max_i \{I_i(x, y)\} \\ \min_i \{I_i(x, y)\} \\ \max_i \{I_i(x, y)/I_{i-1}(x, y)\} \end{bmatrix}, & \text{if } I_i(x, y)/I_{i-1}(x, y) \geq 1 \\ \begin{bmatrix} \max_i \{I_{i-1}(x, y)\} \\ \min_i \{I_i(x, y)\} \\ \max_i \{I_{i-1}(x, y)/I_i(x, y)\} \end{bmatrix}, & \text{otherwise} \end{cases}\quad (3.8)$$

其中連續的背景畫面總數為 N ，且 $i=1, 2, \dots, N$ 。

透過先前所建立之背景模型，我們能夠將每個畫面中的前景與背景區分出來

$$B(x, y) = \begin{cases} 0, \text{ a background pixel} & \text{if } \begin{cases} I_i(x, y)/m(x, y) < kd(x, y) \\ \text{or} \\ I_i(x, y)/n(x, y) < kd(x, y) \end{cases} \\ 255, \text{ a foreground pixel} & \text{otherwise} \end{cases}\quad (3.9)$$

其中 B 為前景之影像，以白色表示； k 為一個門檻值，此門檻值將影響到各個像素點是否被歸為前景，而 k 值的決定通常由不同的環境特性來判斷。

如圖 3.9 所示，在前景的像素點被判斷出來後，我們將前景影像 B 分別投影至 X、Y 軸，藉由其投影之直方圖求出前景物件之外切矩形參數。

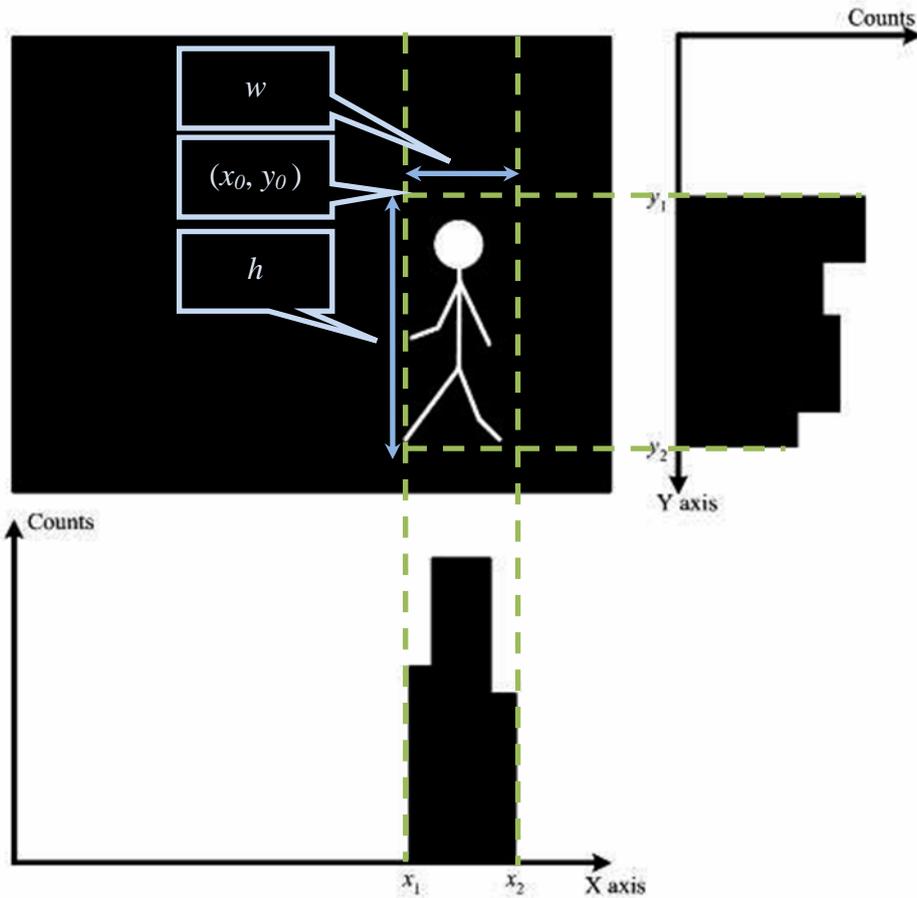


圖 3.9. 使用前景之二值化影像投影至 X、Y 軸，藉以求出前景物件之外切矩形參數。

其中 (x_0, y_0) 為前景人物外切長方形左上角的點， w 和 h 分別為為外切長方形的寬度及高度。透過以上的參數，我們進一步的定義人體上半身的範圍以切割出上半身之區域，如圖 3.10 所示。此上半身區域影像將做為色彩模型辨識系統建立色彩模型以及色彩模型比對之用。由圖 3.10 我們可看出此上半身區域內之影像可能會包含少許的背景成份，但由於我們採用色彩之直方圖來建立色彩模型，少許的背景成份將不太影響實際的辨識結果。

此處上半身區域影像擷取方式與人臉辨識系統不同，人臉辨識系統以人臉的長寬比例關係擷取出上半身區域；色彩模型辨識系統以前景之外切矩形之比例擷取出上半身區域。而我們透過這兩個系統所建立之色彩模型進行比對，藉此將較遠處攝影機與周邊攝影機完成連結。

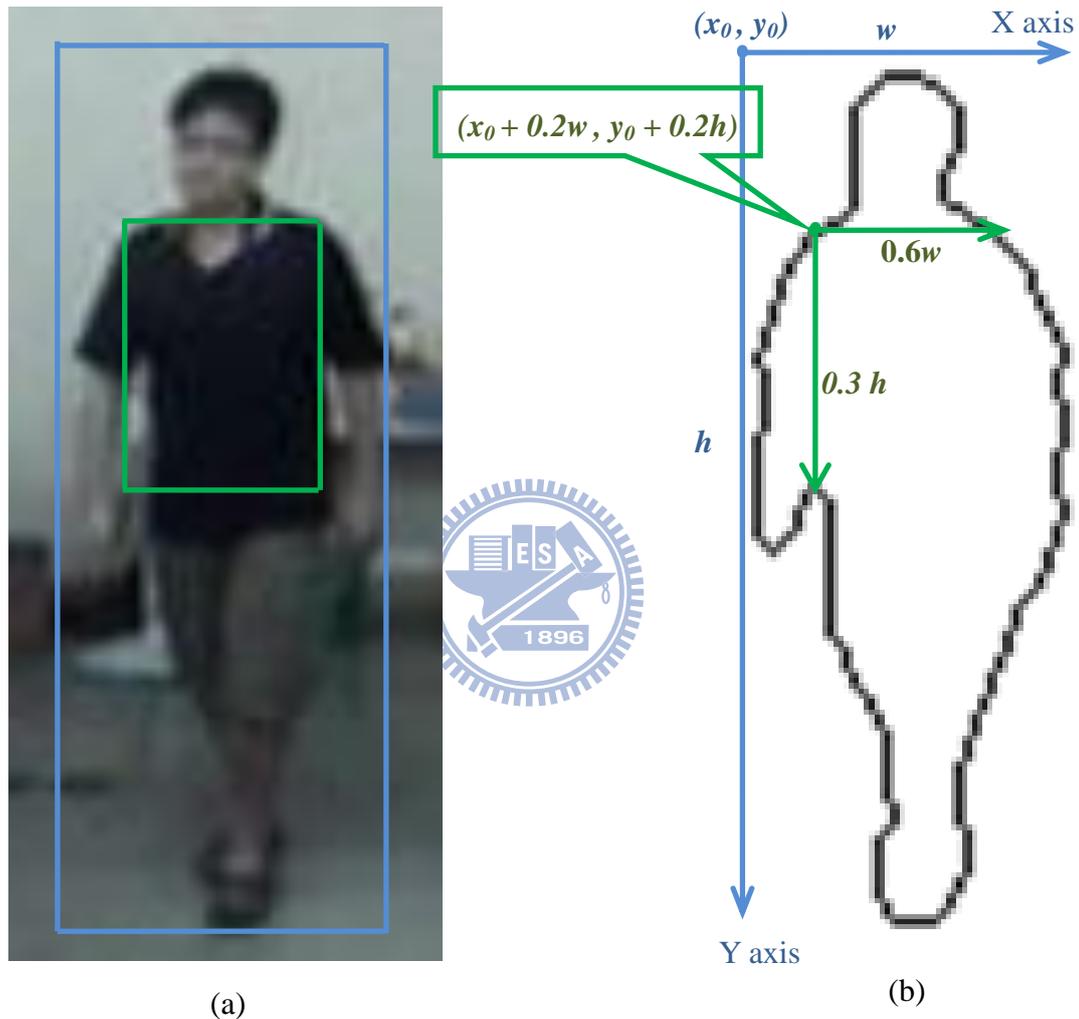


圖 3.10. 前景影像之外切矩型與上半身區域範圍關係圖，(a)為前景外切矩形與上半身區域，(b)為前景外切矩形與上半身區域之比例關係圖。

3.2.2 直方圖交集

直方圖交集方法參考至 Swain 等人[11]所建議之方法，假設 I 和 M 分別為兩組直方圖，且各自含個 n 單位，其交集定義為

$$\sum_{j=1}^n \min(I_j, M_j) \quad (3.10)$$

其結果為兩個影像中顏色相同的像素點數，而在兩個影像比對的數值可以正規化為 0 至 1 間藉由除去其中一個模型的總和

$$H(I, M) = \frac{\sum_{j=1}^n \min(I_j, M_j)}{\sum_{j=1}^n M_j} \quad (3.11)$$

在實做上則是直接待辨識之上半身區域依序與資料庫中的色彩模型進行交集運算，其結果以交集之最大值的某人衣服模型作為辨識其為某人身分。

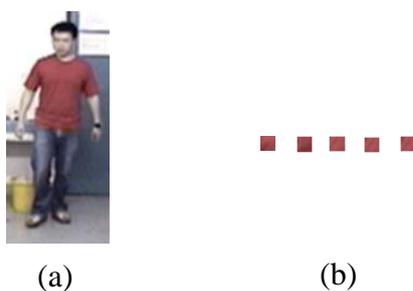


圖 3.11. 上半身區域與取樣區域示意圖
(a)人物影像，(b)從(a)上半身擷取之伍個影像方塊

3.2.3 直方圖激發

參考圖 3.10 與 3.11，我們在已標定之上半身區域內隨機取樣伍個 7×7 之局部影像，如圖 3.11 所示，而後將這些局部影像由 RGB 轉換至 $YCbCr$ 色彩空間，並與資料庫內的色彩模型逐點進行直方圖激發，以辨識人物之身分，假設局部影像內位於 (x, y) 處之 $YCbCr$ 分別表示為 $Y(x, y)$ ； $C_b(x, y)$ ； $C_r(x, y)$ ，而直方圖激發值定義為

$$M_i^{Act}(x, y) = M_i^Y(Y(x, y)) + M_i^{C_b}(C_b(x, y)) + M_i^{C_r}(C_r(x, y)) \quad (3.12)$$

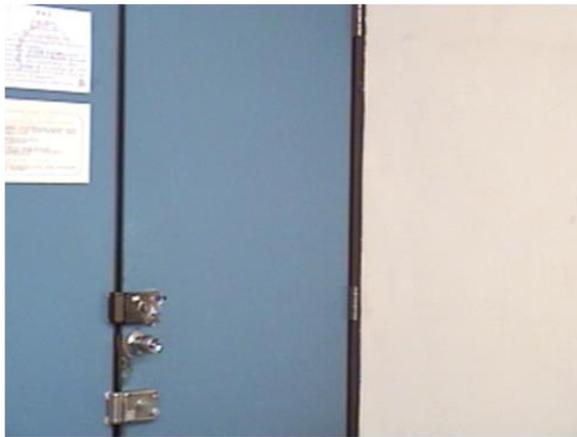
最後再將局部影像內所有的直方圖激發值加總起來，選擇有最大激發量的人物做為辨識結果。



第四章 實驗結果

4.1 實驗環境

我們於國立交通大學，工程五館之智慧與技術實驗室進行系統的開發與測試。人臉辨識系統使用 Sony Evi-D70 攝影機；色彩模型辨識系統使用 Sony Evi-D100 攝影機，使用光源為一般室內日光燈，如圖 4.1 所示，分別為兩個不同攝影機所擷取的實驗環境內畫面。圖 4.1(a)為人臉辨識系統的監視區域，鏡頭畫面為實驗室的入口處，當人們開門進入後即可被攝影機拍攝進行人臉辨識；圖



(a)



(b)

圖 4.1. 實驗環境，(a)人臉辨識監視之畫面，(b)色彩模型辨識之監視之畫面。

4.1(b)為色彩模型辨識的監視區域，得以辨識在此區域內活動之人物身分。人臉辨識系統與色彩模型辨識系統之開發環境硬體設備皆為 Intel(R) Core(TM)2 Quad CPU Q9400 2.66GHz, 2GB RAM。

4.2 人臉辨識實驗

4.2.1 人臉偵測與人臉正規化

圖 4.2 為使用人臉偵測的結果，其包含了重複偵測的人臉，我們使用第二章所介紹過的方法，將重疊的人臉邊界參數取平均合併，用以消除重複偵測的現象。

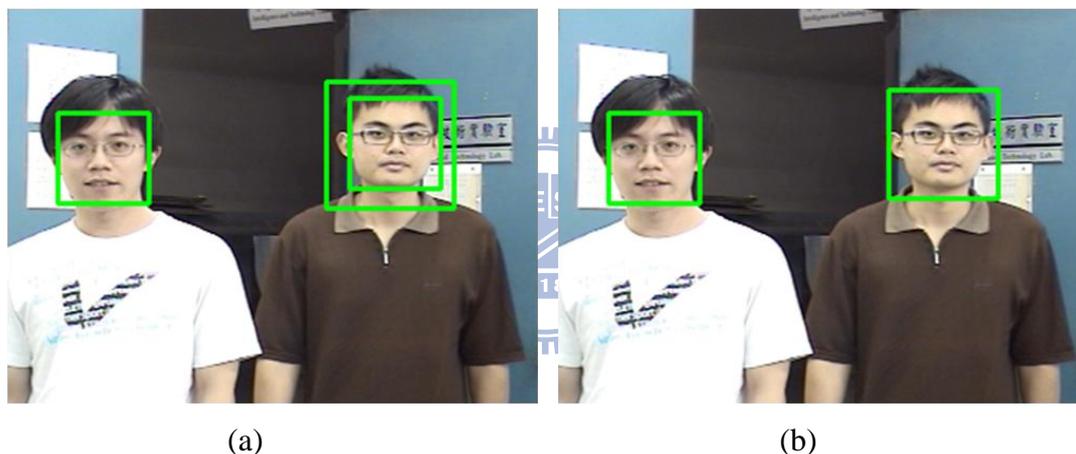


圖 4.2. 人臉偵測圖，(a)人臉重複偵測的結果，(b)消除人臉重複偵測的結果。

於偵測出人臉影像之後，我們將篩選適合辨識的人臉，並且將其進行人臉正規化。如圖 4.3 所示，首先我們偵測鼻尖與嘴唇用綠色矩形表示，並依照鼻尖與嘴唇所偵測出的位置中心將其連成一線，以藍色線表示，此線即可反應出人臉傾斜的程度。如圖 4.4 所示，我們透過膚色偵測找出人臉膚色之左右邊界，並依照此左右邊界將膚色區域之寬度分為五個等分，當嘴唇與鼻尖的位置皆位於中間的位置時，則判定為未旋轉之人臉。圖 4.4(a)、(c)為未旋轉之人臉之人臉影像與膚色影像，(b)、(d)為旋轉之人臉。透過上述方法，我們可以濾除不適合進行人臉

辨識之人臉影像，以確保人臉辨識能夠正確。

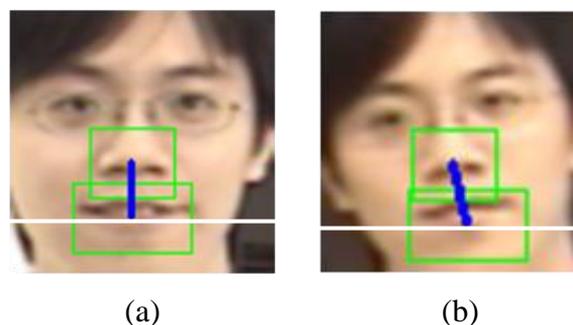


圖 4.3. 鼻尖與嘴唇的偵測，圖中綠色矩形為被鼻尖與嘴唇中間位置，藍色線為鼻尖與嘴唇的連線，白線為水平線。(a)未傾斜的人臉，(b)傾斜的人臉。

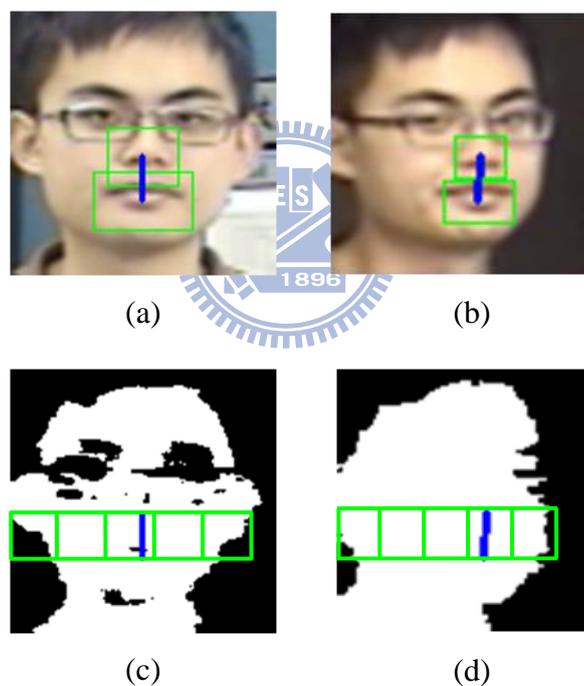


圖 4.4. 人臉旋轉的判定，(a)-(b)為鼻尖與嘴唇偵測位置圖，(c)-(d)為鼻尖與嘴唇於人臉膚色邊界位置關係。

4.2.2 人臉辨識模型建立

建立人臉辨識模型之影像使用我們實驗室所建立之人臉影像資料庫，其中包含八個人物，為期四個月所蒐集之人臉影像。我們以每個人分別選取 100 張在不同日期所拍攝之人臉影像使用特徵空間轉換以及標準空間轉換進行人臉辨識模

型之建立，圖 4.5、4.6 為人臉影像以及特徵臉，透過特徵臉可以展延出一個特徵空間使得每張人臉影像可投影至此空間，且滿足彼此間最大的變異程度。各個特徵向量之能量分佈及其累積分佈如圖 4.7、4.8 所示，我們可看出當選取之特徵向量數為 50 個時所包含的能量已達 90%。而人臉於特徵空間之分佈如圖 4.9 所示，我們以相同顏色代表同個人物的各張人臉影像，在特徵空間中除了咖啡色以及黑色所代表之群體以外，其他群體彼此之間分佈情形相當接近，這將不利於人臉辨識，這也是需要進行標準空間轉換的原因。圖 4.10 為人臉於標準空間之分佈，透過標準空間轉換能同時增大類別間的分散度與同類間聚集度，進而增強分類器表現。

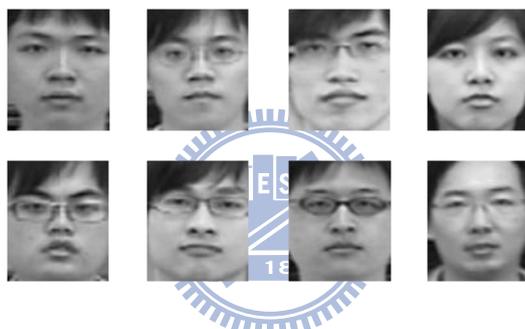


圖 4.5. 人臉資料庫影像範例

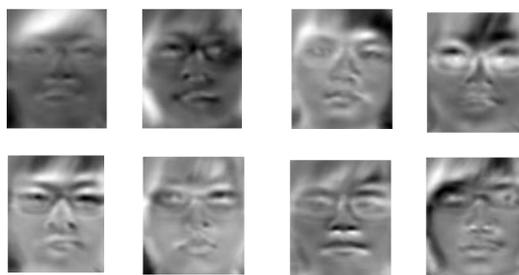


圖 4.6. 前八個特徵臉

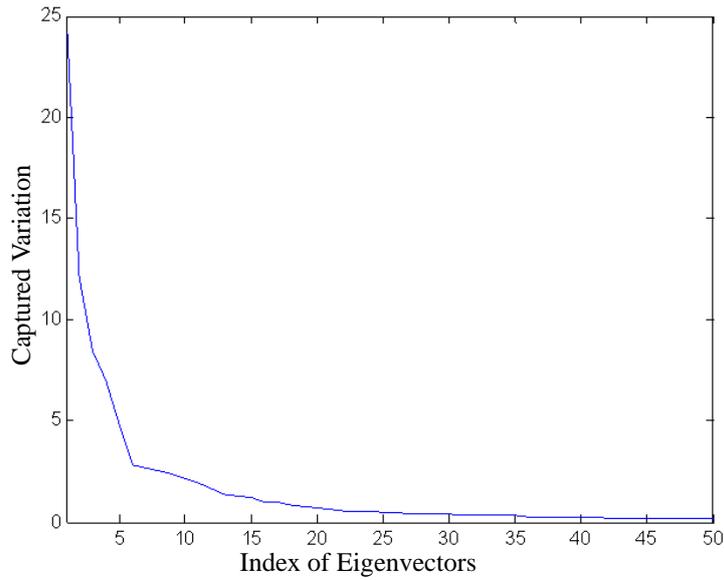


圖 4.7. 特徵向量之能量分佈

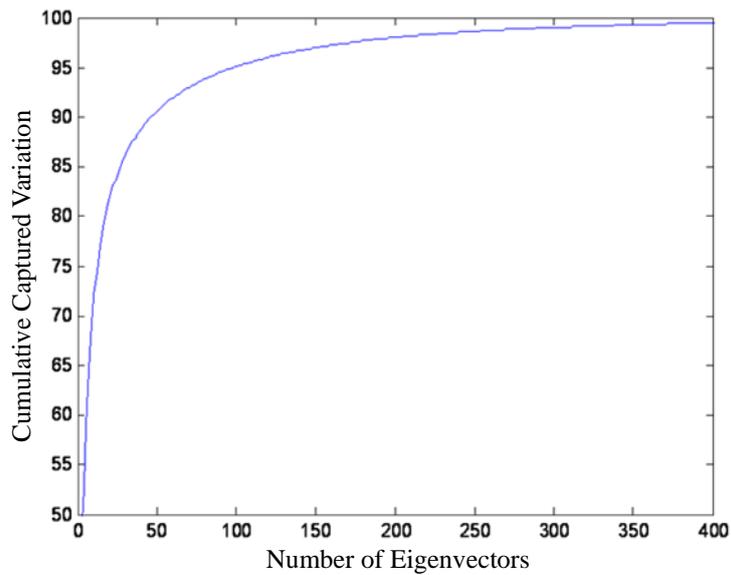


圖 4.8. 特徵向量之累積能量分佈

我們將單獨使用特徵空間轉換以及使用特徵空間轉換加標準空間轉換兩種方法以資料庫內影像，每個人物取 100 張(共 800 張)人臉影像進行 leave-one-out 驗證，其正確率分別為 97.5%、98.1%，如表 4.1 所示，由於標準空間轉換具備增大類別間的分散度與同類間聚集度之特性，使辨識率較僅使用特徵空間轉換為佳。

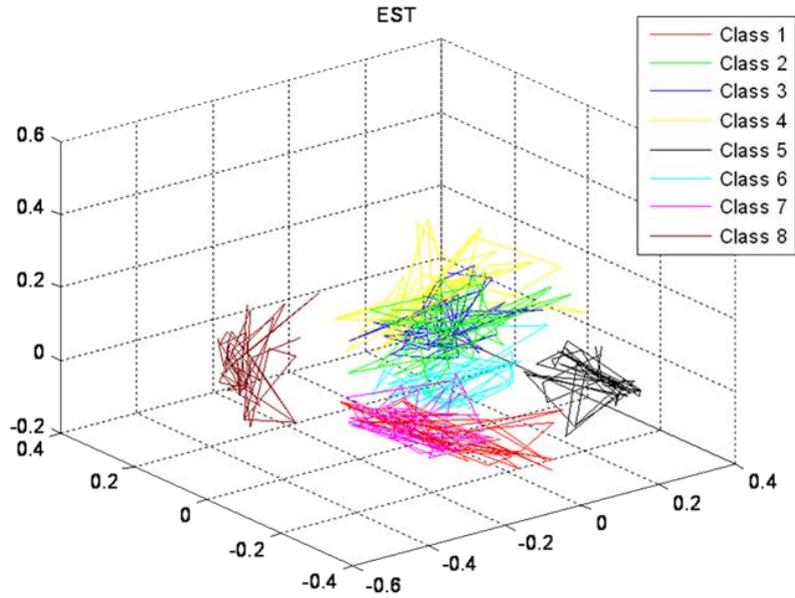


圖 4.9. 人臉影像於特徵空間之分佈

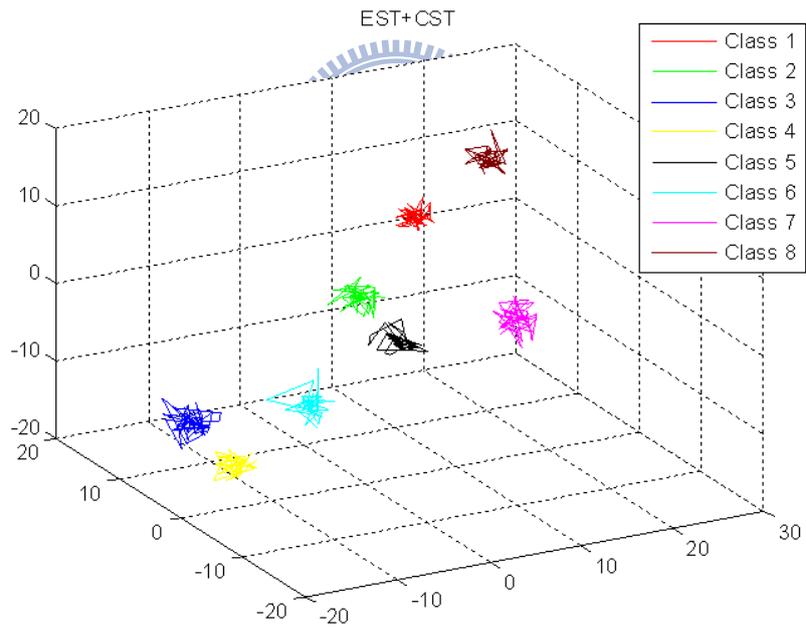


圖 4.10. 人臉影像於標準空間之分佈

表 4.1. 比較 EST 以及 EST+CST 之辨識效果

使用方法	辨識率	類別間分離效果
EST	97.5%	較為重疊
EST+CST	98.1%	較為分離

4.2.3 即時人臉辨識

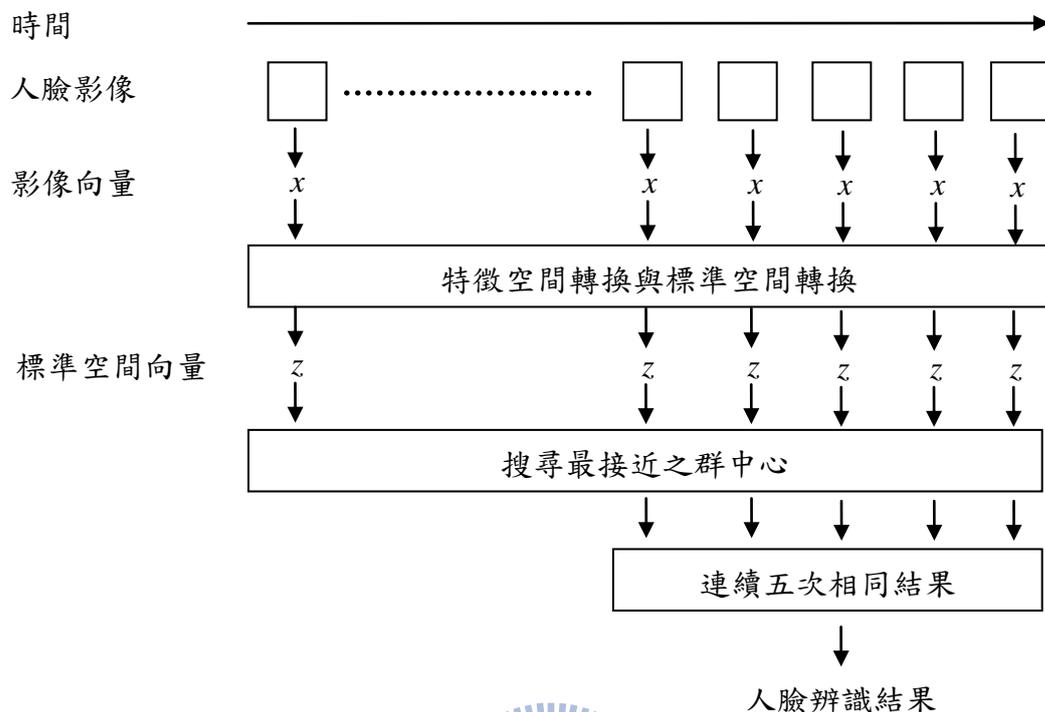


圖 4.11. 即時人臉辨識之辨識流程

即時人臉辨識之辨識流程如圖 4.11 所示，我們將影像轉換至標準空間，並依序計算影像距離每個人於標準空間的群中心，選擇最近的群體作為辨識輸出，再比較最近五次之辨識結果，假若五次皆為同個人，則我們即可將此辨識結果作為人物身份最終的判定，圖 4.12 為人臉辨識系統介面，每張人臉影像之處理時間約 40 毫秒。

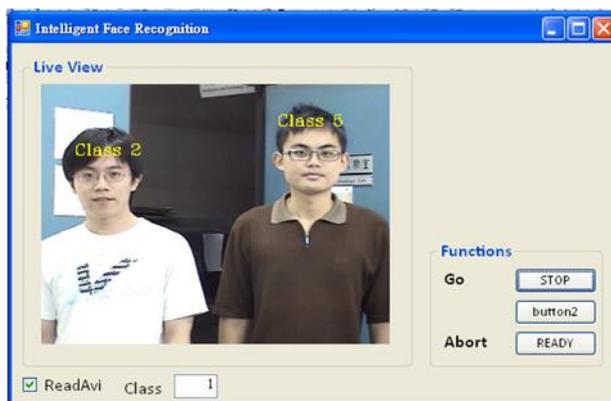


圖 4.12. 即時人臉辨識系統介面

為了能區別進入者是否為家庭的成員，因此我們藉由分析家庭成員與非家庭成員之人臉影像於標準空間之分佈，如圖 4.13 所示，家庭成員之人臉影像使用彩色的線段連接，而非家庭成員之人臉影像以深綠色之星號表示。

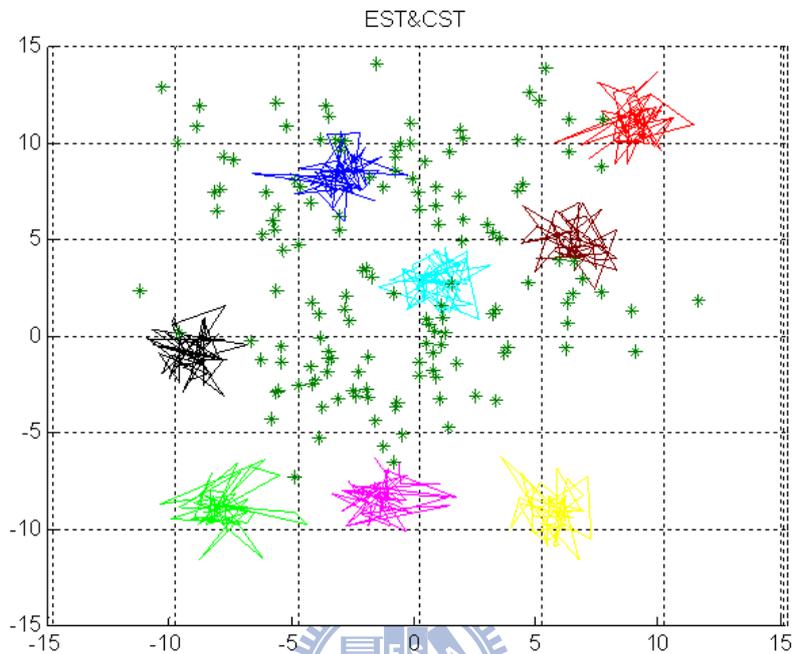


圖 4.13. 家庭成員與非家庭成員之人臉影像於標準空間之分佈

其中屬於家庭成員之人臉影像於標準空間中皆具有群聚之現象，相同群體間將凝聚在一起，而非家庭成員之人臉影像散佈於標準空間中。因此藉由計算家庭成員之人臉影像與其相對應之群中心之距離的變異數，作為判斷進入者是否為家庭成員，其判斷方式為

$$\text{Class} = \begin{cases} i, & \text{if } \frac{\|z^{\text{test}} - z_i^{\mu}\|}{z^{\sigma}} \leq k \\ \text{unknown,} & \text{otherwise} \end{cases} \quad (4.1)$$

其中 z^{test} 為進入者之人臉影像於標準空間之向量， z_i^{μ} 為第 i 個人之所有訓練影像與其相對應之群中心間距離的平均， z^{σ} 為所以訓練影像與其群中心距離之變異數，其值為 1.05， k 為一個閾值，藉由此閾值判斷進入者之人臉影像需小於多少個變異數才能作為正確之人臉辨識。我們取 198 張非家庭成員之影像進行試驗，

圖 4.14 表示不同的 k 值對於非家庭成員之人臉影像的排除率，其計算方式為

$$\text{排除率} = \frac{\text{成功被排除之影像數}}{\text{非家庭人員之人臉影像數}} \times 100\% \quad (4.2)$$

當 k 值越小時代表其排除的門檻越為嚴苛，使得大部分的非家庭成員之人臉影像皆能被排除。

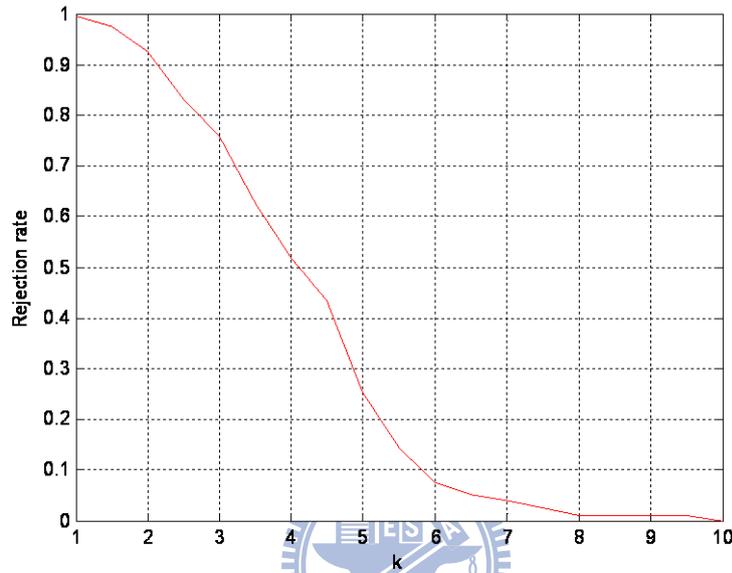


圖 4.14. 不同 k 值對於非家庭成員之排除率

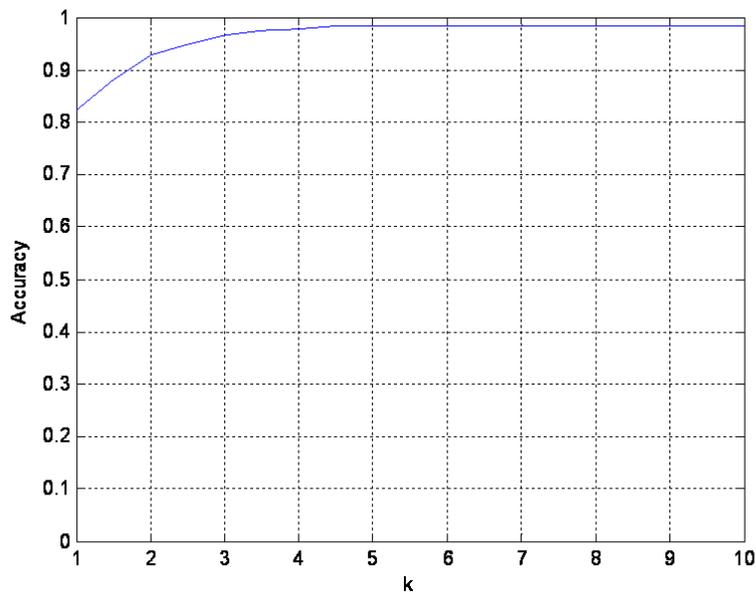


圖 4.15. 不同 k 值對於家庭成員之辨識率

圖 4.15 為資料庫內所蒐集的人物影片進行辨識效能測試，影片內容為人物由實

驗室門口進入的畫面，我們擷取出人臉之影像進行辨識率之測試，計算辨識率的公式為

$$\text{辨識率} = \frac{\text{辨識正確之影像數}}{\text{被偵測的人臉影像數}} \times 100\% \quad (4.3)$$

與計算排除率相同，當 k 值越小時，判斷的標準也越為嚴苛，這使得有些實際為辨識正確的人臉也被排除，應此我們參考正確率與排除率間的總和做為選用 k 值的依據，如圖 4.16 所示。

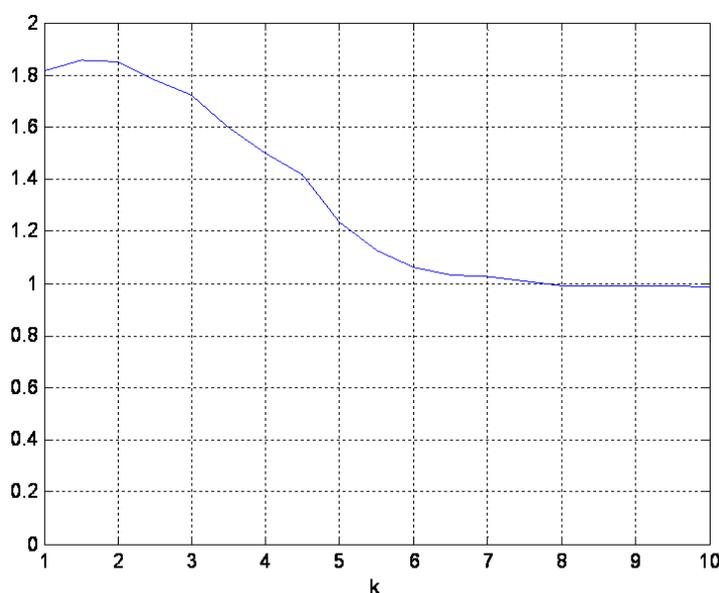


圖 4.16. 不同 k 值排除率與辨識率之總和

其中我們選擇 k 值為 2，其排除率與辨識率分別為 92.42%、92.75%。

4.3 色彩模型辨識實驗

4.3.1 上半身區域擷取

於色彩模型辨識階段，如圖 4.17 所示，使用前後景分離的方式擷取出影像前景的分佈，圖 4.17(a)為原始影像畫面，其解析度為 320×240 像素，透過預先建立之背景模型，將可分離出前景分佈，如圖 4.17(b)。在取得前景分佈後，我們將前景分佈分別投影至 X、Y 軸，如圖 4.17(c)–(d)，依照其投影至此兩軸所累

積的像素點數，得以判斷出前景之外切矩型之範圍，如圖 4.17(e)。

我們以 3.2.1 節所談到之擷取人物上半身區域的方法，將擷取出前景人物之上半身區域，參考表 4.2，其中包含了由左側身到右側身一共 7 個偵測結果，其中以黃色矩型表示偵測出之前景外切矩型；以綠色矩型表示偵測出之人物上半身區域。透過上半身區域的影像我們得以建立人物之色彩辨識模型並即時辨識每個人物的動向。

4.3.2 色彩模型選用

於色彩模型辨識階段嘗試了使用 RGB、正規化 RGB、 $YCbCr$ 以及 HSV 等色彩空間建立色彩模型，我們計算一個近似於統計學中 Z 分數 (Z-values) 的評量參數，用以判斷各個色彩空間對於不同人物的色彩模型能夠被區分的程度，其假設各個色彩模型分佈為常態分佈，選取兩個欲判斷之色彩模型，以標準差較大的模型作為基準，判斷標準差較小之模型平均數距離標準差較大之模型的標準差數，再參考模型的平均以及標準差，方能判斷各個色彩空間是否適合於進行色彩模型辨識。考慮 i 、 j 兩個色彩模型， μ 與 σ 分別為平均值與標準差，計算公式定義為

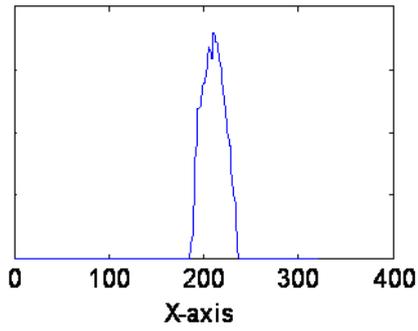
$$z_{i,j} = \frac{|\mu_i - \mu_j|}{\max(\sigma_i, \sigma_j)} \quad (4.4)$$



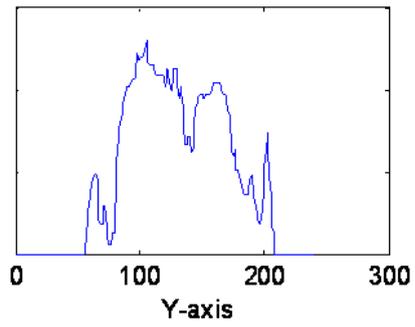
(a)



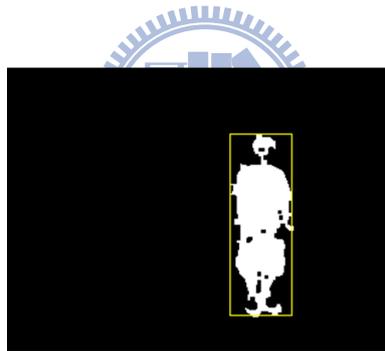
(b)



(c)



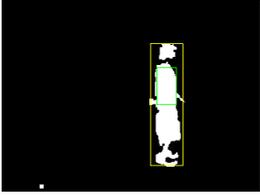
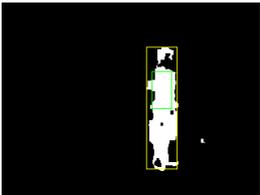
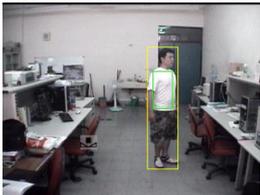
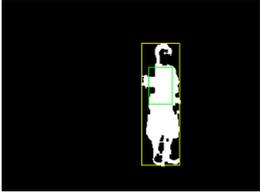
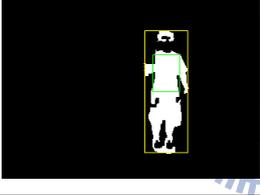
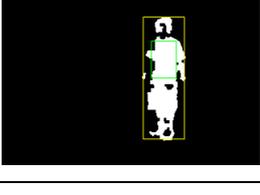
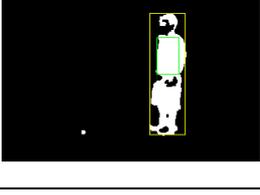
(d)



(e)

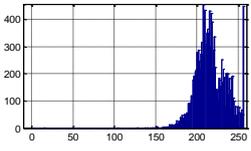
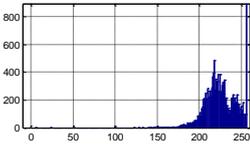
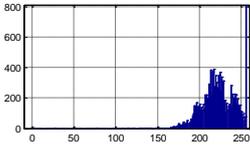
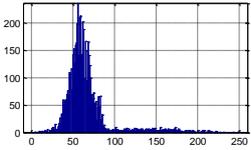
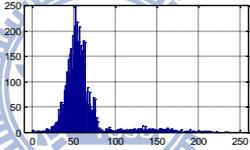
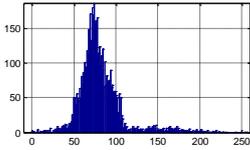
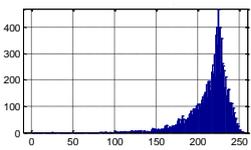
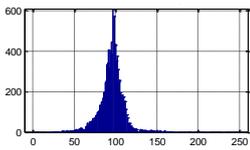
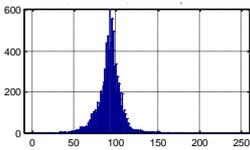
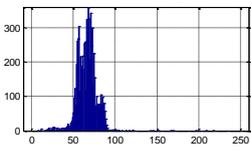
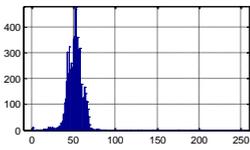
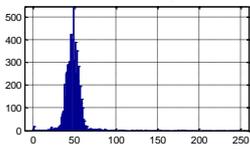
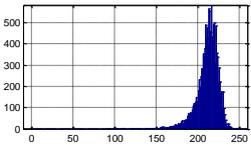
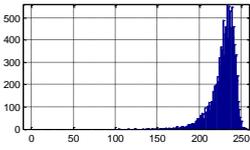
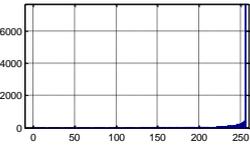
圖 4.17. 人物之外切矩形偵測，(a)畫面影像，(b)前景分佈，(c)前景分佈投影至 X 軸，(d)前景分佈投影至 Y 軸，(e)擷取出人物之外切矩形。

表 4.2. 身體區域之偵測結果

No.	前景分佈	影像畫面
1		
2		
3		
4		
5		
6		
7		

使用 RGB 色彩空間進行分析，如表 4.3 所示，各個空間皆與亮度有相當大的關聯，如第 1、5 類淺色系衣服亮度較大，其於 RGB 之分佈偏向數值較高的一側，相反地第 2、4 類深色系衣服亮度較低，分佈將偏向較低的一側。相較於之後所提及之色彩空間，色彩模型於 RGB 分佈變異程度普遍較大，也因此模型之分佈也較為寬，使得相異類別間容易混淆。因此 RGB 空間仍不適合做為建立色彩模型之用。

表 4.3. 色彩模型於 RGB 之分佈

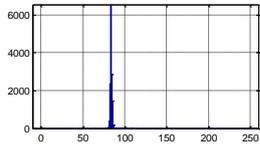
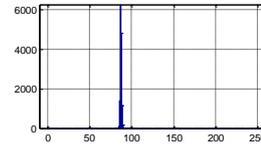
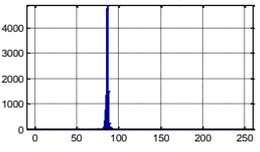
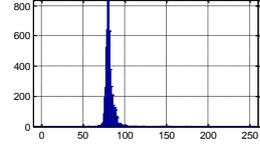
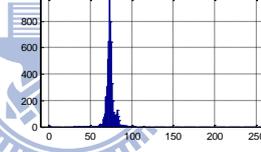
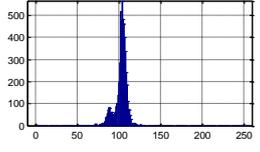
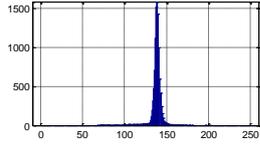
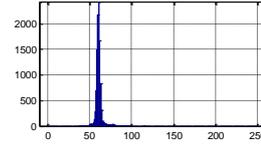
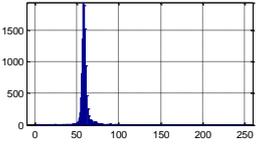
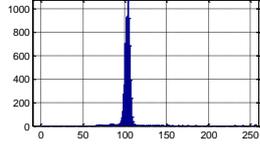
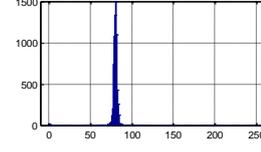
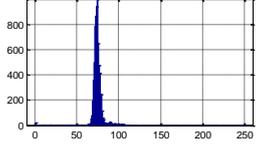
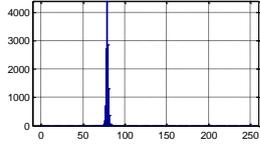
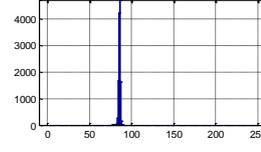
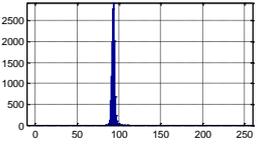
	R	G	B
 1	 $\mu: 215, \sigma: 18.51$	 $\mu: 224, \sigma: 18.18$	 $\mu: 221, \sigma: 19.92$
 2	 $\mu: 64, \sigma: 28.03$	 $\mu: 58, \sigma: 27.18$	 $\mu: 79, \sigma: 26.43$
 3	 $\mu: 213, \sigma: 22.23$	 $\mu: 94, \sigma: 13.67$	 $\mu: 91, \sigma: 13.8$
 4	 $\mu: 65, \sigma: 11.89$	 $\mu: 50, \sigma: 9.86$	 $\mu: 46, \sigma: 9.27$
 5	 $\mu: 210, \sigma: 13.19$	 $\mu: 227, \sigma: 14.12$	 $\mu: 249, \sigma: 11.71$

正規化 RGB 空間之特性在於消除光線對色彩的影響，其轉換公式為

$$\begin{bmatrix} r \\ g \\ b \end{bmatrix} = \frac{1}{R+G+B} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4.5)$$

其分佈於 0 至 1 間，我們以後把正規化 RGB 稱為 rgb 模型。為了能與其他色彩空間進行比較，我們調整其尺度為 0 至 255 間，如表 4.4 所示。再次比較淺色系

表 4.4. 色彩模型於 rgb 模型之分佈

	<i>r</i>	<i>g</i>	<i>b</i>
 1	 $\mu: 83, \sigma: 1.39$	 $\mu: 86, \sigma: 1.8$	 $\mu: 85, \sigma: 1.7$
 2	 $\mu: 80, \sigma: 6.28$	 $\mu: 72, \sigma: 5.04$	 $\mu: 101, \sigma: 8.14$
 3	 $\mu: 136, \sigma: 8.2$	 $\mu: 60, \sigma: 4.11$	 $\mu: 58, \sigma: 4.99$
 4	 $\mu: 102, \sigma: 8.68$	 $\mu: 79, \sigma: 4.32$	 $\mu: 73, \sigma: 6.66$
 5	 $\mu: 78, \sigma: 1.4$	 $\mu: 84, \sigma: 1.05$	 $\mu: 92, \sigma: 1.87$

的第 1、5 類及深色系的第 2、4 類，由直方圖上觀察可得，各個模型的分佈並沒有如同 RGB 般淺色系偏右深色系偏左的現象，確實減低了其分佈與亮度之間的影响，且各個分佈的變異數也比於 RGB 的分佈較為小，其分佈明顯較為窄。我們藉由計算 z 值比較深色系的兩類以及淺色系的兩類是否容易區分，如表 4.5 所示，深色系與淺色系皆在 r 、 g 空間有較好的區分效果。

表 4.5. 比較淺色系與深色系可區分的程度於 rgb 模型

	r	g	b
$z_{1,5}$ (淺色系)	3.57	1.11	3.74
$z_{2,4}$ (深色系)	2.53	1.39	3.44

我們再以 YC_bC_r 空間進行分析，如表 4.6 所示，其轉換方式如 3.1.2 節所介紹。 YC_bC_r 將亮度與彩度進行分離，從亮度觀察，可以明顯區分出深色系與淺色系的差異，淺色系的 1、5 類 Y 值平均偏大，深色系的 2、4 類 Y 值平均偏小，而這可能造成同為淺色系或深色系的模型難以區分，我們再計算深色系與淺系之 z 值，如表 4.7 所示。同為深色系與淺色系的色彩模型在亮度上確實難以區分，但在 C_b 、 C_r 上仍有一定的區分效果，在這些試驗的樣本中又以 C_b 的可區分程度為佳。

雖然同為深色系或淺色系的模型無法透過 Y 進行區分，但其可有效的區分深色系與淺色系的模型，表 4.8–4.9 比較相異亮度的色系間能區分的程度，比起 rgb 模型， YC_bC_r 藉由 Y 使得相異亮度的色彩模型能明顯區分，其 z 值至少也有 5.85，而在 C_r 上 z 值至少也有 2.96 的水準。考慮 rgb 三個通道來看，最佳之 z 值出現在 $z_{5,4}^b$ ，其值為 2.85，且 z 值全部皆小於 3，因此 rgb 模型在分辨相異亮度間的色彩模型之效能仍不及 YC_bC_r 。經由上述討論，其反應出 YC_bC_r 比 rgb 模型

更適合作為色彩模型辨識。

表 4.6. 色彩模型於 YC_bC_r 之分佈

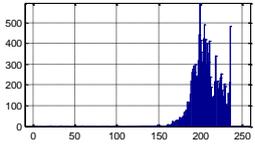
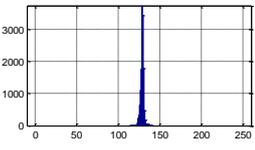
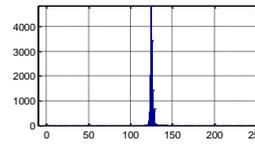
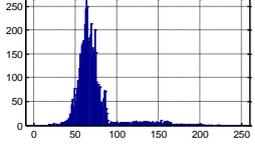
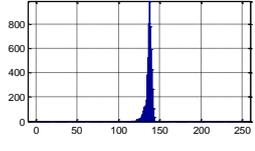
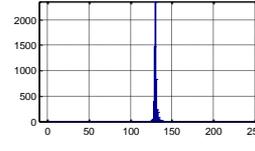
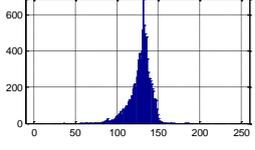
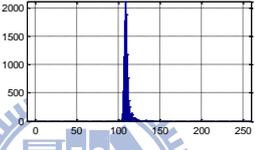
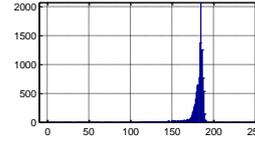
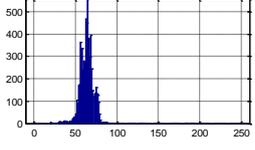
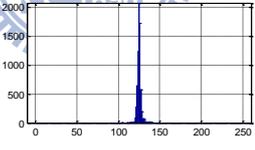
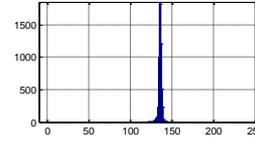
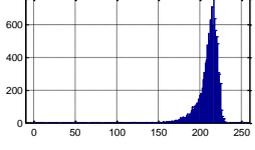
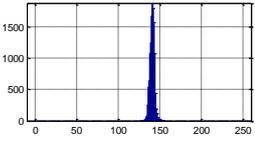
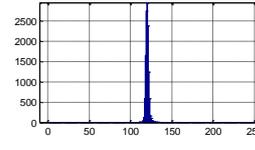
	Y	C_b	C_r
 1	 $\mu: 206, \sigma: 15.76$	 $\mu: 128, \sigma: 2.02$	 $\mu: 124, \sigma: 1.69$
 2	 $\mu: 69, \sigma: 23.41$	 $\mu: 136, \sigma: 3.5$	 $\mu: 129, \sigma: 1.62$
 3	 $\mu: 127, \sigma: 12.36$	 $\mu: 109, \sigma: 3.74$	 $\mu: 180, \sigma: 8.05$
 4	 $\mu: 62, \sigma: 8.74$	 $\mu: 124, \sigma: 2.13$	 $\mu: 134, \sigma: 1.93$
 5	 $\mu: 209, \sigma: 11.46$	 $\mu: 139, \sigma: 2.85$	 $\mu: 119, \sigma: 1.88$

表 4.7. 比較同為淺色系或深色系可區分的程度於 YC_bC_r 空間

	Y	C_b	C_r
$z_{1,5}$ (淺色系)	0.19	3.85	2.65
$z_{2,4}$ (深色系)	0.3	3.42	2.59

表 4.8. 比較淺色系與深色系間可區分的程度於 rgb 模型

	r	g	b
$z_{1,2}$	0.48	2.78	1.97
$z_{1,4}$	2.19	1.62	1.8
$z_{5,2}$	0.32	2.38	1.11
$z_{5,4}$	2.77	1.16	2.85

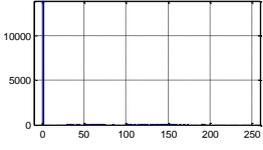
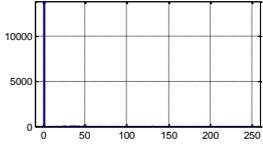
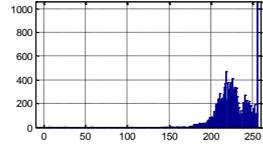
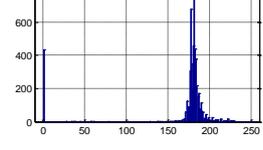
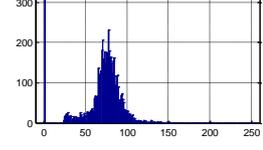
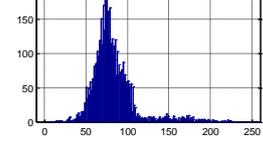
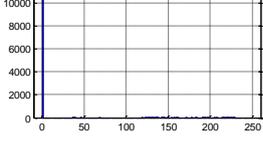
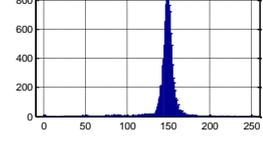
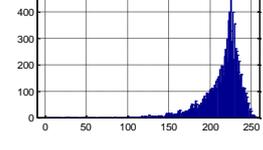
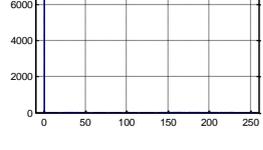
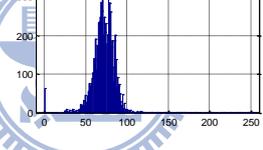
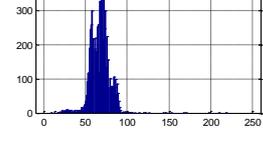
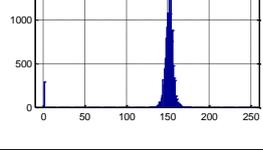
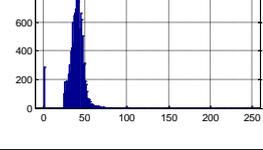
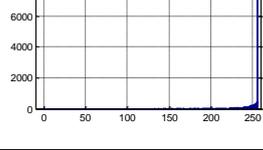
表 4.9. 比較淺色系與深色系間可區分的程度於 $YCbCr$ 空間

	Y	C_b	C_r
$z_{1,2}$	5.85	2.28	2.96
$z_{1,4}$	9.13	1.87	5.18
$z_{5,2}$	5.98	0.86	5.32
$z_{5,4}$	12.83	5.26	7.77

HSV 色彩空間在處理上將會遇到一些問題，當亮度 (Value) 過低時色度 (Hue) 將失去意義；當飽和度 (Saturation) 過低時，色度將不穩定。為了避免上述的問題我們在建立 HSV 色彩模型時加入了門檻值。HSV 之門檻值使用 $H_T = 36^\circ$ 、 $S_T = 22.95$ 、 $V_T = 25.5$ ， H 介於 $\pm 36^\circ$ 間則歸於 0° ，以消除 H 值 0° 與 359° 度間不連續的現象。 S 小於 22.95 代表含有過多的白色成份，顏色將不易區分，因此 H 值也歸為 0° ， S 值也歸 0 。 V 值低於 25.5 表示色彩亮度偏暗，亦不易識別其顏色，因此 H 也歸為 0° 。

表 4.10 中，受到 H 之特性關係，僅以出現最大之 H 值表示，而 S 、 V 仍

表 4.10. 色彩模型於 HSV 之分佈

	H	S	V
 1			
	H = 0	$\mu: 0, \sigma: 3.74$	$\mu: 225, \sigma: 18.1$
 2			
	H = 181	$\mu: 70, \sigma: 23.8$	$\mu: 81, \sigma: 26.68$
 3			
	H = 0	$\mu: 146, \sigma: 16.1$	$\mu: 214, \sigma: 21.82$
 4			
	H = 0	$\mu: 71, \sigma: 15.65$	$\mu: 66, \sigma: 11.88$
 5			
	H = 151	$\mu: 38, \sigma: 9.08$	$\mu: 250, \sigma: 11.71$

以平均值與標準差表示。第 1 類的白色由於 S 值低於門檻，使得 H 與 S 值皆為 0。這將使得原始 H 為 0 的紅色（第三類）以及經由門檻值使得 H 歸 0 的第一類在 H 值有極高的相關性，其中 $z_{5,4}^V$ 為 0.5，使得 1、3 類僅能藉由 S 進行區分。HSV 在 H 的分佈易受到 S、V 的影響，且 S、V 之分佈變異程度僅次於 RGB，因此我們在進一步的討論裡，將選擇 $YCbCr$ 以及 rgb 模型進行色彩模型辨識的比較。

我們選用四個人物的影片，當人物從入口處進入時使用 $YCbCr$ 以及 rgb 模型

建立色彩模型，再擷取人物於畫面中與攝影機距離由遠到近各個影像，以直方圖激發的方式進行試驗，其正確率計算方式為

$$\text{正確率} = \frac{\text{辨識正確之影像數}}{\text{人物畫面影像數}} \times 100\% \quad (4.6)$$

圖 4.18 比較上半身區域取樣數與兩個色彩空間對正確率之影響，我們可得 YCbCr 的正確率皆大於 rgb 模型。

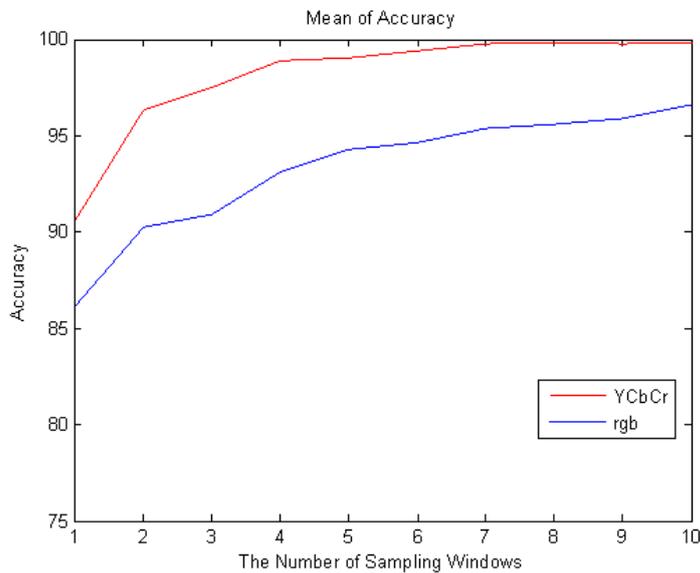


圖 4.18. 比較 YCbCr 與 rgb 之正確率

使用直方圖交集[11]以及直方圖激發於 YCbCr 進行比較，直方圖交集取用所有上半身區域之像素點建立模型，並與預先建立的色彩模型做交集運算，其考量範圍最為完整，試驗之正確率為 100%，但由於其需經由建立模型、交集運算兩個處理程序，在運算效率上較直方圖激發來得慢。我們統計所有試驗樣本之上半身區域像素點數，其平均為 735 個像素，在使用直方圖交集時平均每次處理皆要運算如此的點數，且需經過兩個處理程序。使用直方圖激發則依照取樣區域數而定，每個取樣區域為 7×7 之遮罩，共 49 個像素點，而由圖 4.18 觀察可得，在取樣數為 7 時正確率即以逼近 100%，而使用像素點僅 343 個像素。經由上述討論，

直方圖激發比起直方圖交集有更好的處理效能，而我們在處理的影像為連續的影像序列，因此透過數個影像間進行主要投票，則可再次縮減直方圖激發的取樣數至 3 個取樣樣本，並且維持辨識水準。



第五章 結論與未來展望

本研究整合了數種電腦視覺技術，實現一即時人物辨識系統。首先，使用 OpenCV 之 Haar cascade classifier 進行人臉、鼻尖以及嘴唇之偵測，偵測進入環境內的人物臉部影像，其演算原理乃基於 Viola 等人之人臉偵測技術，能夠快速的偵測物件且具有良好的偵測率，再搭配我們的演算法則消除人臉的誤判以及重疊偵測。使用膚色、鼻尖、嘴唇等臉部特徵，在連續的影像序列中，選取出適合進行人臉辨識之人臉影像，且透過預先定義之人臉範圍擷取出人臉。在辨識上結合特徵空間轉換以及標準空間轉換進行人臉辨識，透過特徵空間轉換，縮減人臉影像資料維度，擷取出人臉中最具代表性之特徵，再以標準空間轉換增加相異群體間的距離，並且縮減相同群體間分散程度，以提供更佳的辨識效能。我們並自行建立人臉資料庫進行測試，此人臉資料庫蒐集了四個月內實驗室成員之人臉影像，使用人臉資料庫測試之正確率為 98.1%，而以即時影片並考慮非家庭成員侵入的情形進行測試之正確率為 92.75%。

人物辨識系統結合了人臉辨識以及色彩模型辨識兩種技術，以即時監視環境內各個人物之動向，使用前後景分離技術擷取出環境內人體區域，並以其色彩模型進行辨識。在色彩空間的選用，我們使用 $YCbCr$ 色彩空間，並以直方圖激發作為色彩模型辨識方式。

在未來我們可以結合動作辨識等相關技術，記錄人們的生活作息形態以及動向，藉此判定居家中人們的身體健康狀態，在出現異常時也能即時將警訊傳達給人們；當人們發生危險時也能即時進行通報，以保護居住者之健康照護與安全。

參考文獻

- [1] P. Viola, M. Jones, “Robust Real-Time Face Detection,” *International Journal of Computer Vision*, 2004.
- [2] P. S. Huang, C. J. Harris and M. S. Nixon, “Canonical Space Representation for Recognizing Humans by Gait and Face,” in *Proc. of IEEE Southwest Symposium on Image Analysis and Interpretation*, pp. 180–185.
- [3] E. Hjelmås and B. K. Low, “Face Detection, a Survey,” *Computer Vision and Image Understanding*, Vol. 83, pp. 236–274, 2001.
- [4] C.C. Chiang, W. K. Tai, M.T. Yang, Y. T. Huang, C. J. Huang, “A Novel Method For Detecting Lips, Eyes and Faces in Real Time,” *The Journal of Real Time Image*, vol.9, pp.277–287, 2003.
- [5] H. A. Rowley, S. Baluja, and T. Kanade, “Neural Network-Based Face Detection,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 1, pp. 23–38, 1998.
- [6] Q. Chen, H. Wu, and M. Yachida, “Face Detection by Fuzzy Pattern Matching,” in *Proc. Int’l Conf. Computer Vision*, pp. 591–596, 1995.
- [7] M. Turk, A. Pentland, “Eigenfaces for Recognition,” *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, pp. 71–86, 1991.
- [8] K. Etemad, R. Chellappa, “Discriminant Analysis for Recognition of Human Face Images,” *Journal of the Optical Society of America A*, Vol. 14, No. 8, pp. 1724–1733, 1997.
- [9] B. Moghaddam and M. H. Yang, “Gender classification with support vector machines,” in *Proc. IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, pp. 306–313, 2000.

- [10] M. Soriano, B. Martinkauppi, S. Huovinen and M. Laaksonen, “Using the skin locus to cope with changing illumination conditions in color-based face tracking,” in *Proc. IEEE Nordic Signal Proc. Symp.*, pp.383–386, 2000.
- [11] M. J. Swain and Dana H. Ballard, “Color indexing,” *International Journal of Computer Vision*, 7(11), pp. 11–32, 1991.
- [12] “OpenCV 1.0, Open Source Computer Vision Library,”
<http://www.intel.com/technology/computing/opencv/>, 2006.
- [13] Rainer Lienhart, <http://www.lienhart.de/index.html>.
- [14] Modesto Castrillon-Santana,
<http://mozart.dis.ulpgc.es/Gias/modesto.html?lang=0>.

