

國立交通大學

電機學院電機產業研發碩士班

碩士論文

基於賈伯小波及區域保留投影之人臉偵測與人臉辨識

Face Detection and Face Recognition Based on Gabor
Wavelets and Locality Preserving Projection



研究生：褚坤龍

指導教授：林昇甫 博士

中華民國九十七年一月

基於賈伯小波及區域保留投影之人臉偵測與人臉辨識

Face Detection and Face Recognition Based on Gabor Wavelets and
Locality Preserving Projection

研究生：褚坤龍

Student: Kun-Long Chu

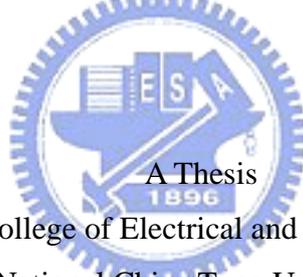
指導教授：林昇甫 博士

Advisor: Dr. Sheng-Fuu Lin

國立交通大學

電機學院電機產業研發碩士班

碩士論文



Submitted to College of Electrical and Computer Engineering

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Industrial Technology R & D Master Program on
Electrical Engineering

January 2008

Hsinchu, Taiwan, Republic of China

中華民國九十七年一月

國立交通大學
電機學院電機產業研發碩士班

論文口試委員會審定書

本校 電機學院電機產業研發碩士班 褚坤龍 君

所提論文：

(中文) 基於賈伯小波及區域保留投影之人臉偵測與人臉辨識

(英文) Face Detection and Face Recognition Based on Gabor
Wavelets and Locality Preserving Projection

合於碩士資格水準、業經本委員會評審認可。

口試委員：



張 翔 林昇甫
潘 晴 財 行政武

指導教授：林昇甫

專班主任：柯明達

中 華 民 國 97 年 / 月 10 日

基於賈伯小波及區域保留投影之人臉偵測與人臉辨識

研究生：褚坤龍

指導教授：林昇甫 博士

國立交通大學電機學院產業研發碩士班

摘 要

人臉辨識在長久以來一直都是很熱門的研究主題，因為可以廣泛的應用在各個不同的領域之中，例如身份識別、基於內容的影像檢索、電腦視覺與人類電腦介面等。而做為必要前處理的人臉偵測也同樣重要，因為要先能把人臉偵測出來，才能夠做接下來的人臉辨識。因此本論文結合了賈伯小波與區域保留投影的優點而提出一個可在數位相片中執行人臉偵測與人臉辨識的方法。先使用膚色偵測來找出可能是人臉的區域，然後使用賈伯小波轉換取得人臉整體特徵，並利用區域保留投影將人臉特徵降至很低的維度，再訓練類神經網路來判別是否為人臉，根據人臉偵測的結果用人工做分類並建立資料庫，透過資料庫中的人臉來訓練類神經網路做辨識，最後在測試相片輸入時，就能夠對於所選定的人臉做辨識，藉此辨識結果就可以將資料庫中含有相同人物的相片挑選出來，達到基於人物身份的影像檢索的目的。本論文的主要貢獻為利用賈伯小波轉換能夠同時在時域和頻域中兼顧足夠的分辨率的特性來取得人臉的整體特徵，再使用區域保留投影能保留局部非線性資訊做降維的優點，將賈伯小波轉換後過於龐大的特徵向量降低至很小的維度，實驗結果顯示本論文的方法在人臉偵測與人臉辨識都有不錯的效能。

Face Detection and Face Recognition Based on Gabor Wavelets and Locality Preserving Projection

Student:Kun-Long Chu

Advisor: Dr. Sheng-Fuu Lin

Industrial Technology R & D Master Program of
Electrical and Computer Engineering College
National Chiao Tung University

Abstract

Face recognition has been a popular research topic for a long time because it can be widely used in many different fields, such as identity identification, content-based image retrieval, computer vision and human computer interaction. However, face detection, which serves as the preprocessing procedure, is equally important since it has to be done first before face recognition is taken. This thesis, therefore, proposes a method which takes the advantages of Gabor transformation and locality preserving projection to implement face recognition and face detection on a digital picture. The first step adopts face detection to find candidate face region. Second, Gabor wavelets transformation is adopted to extract face features of human face, and locality preserving projection is applied to project features of human face into lower dimension space. Afterwards, neural network is trained to decide whether candidate region is human face or not. Then, database is constructed manually according to the result of face detection. Finally, a neural network is trained by the faces is stored in the database. When a test picture is input the proposed method is able to identify the faces of the chosen people. According to the result of identification, pictures of the same person can be chosen from database and implement the identify-based image retrieval. The main contribution of this thesis is to employ the specialty of Gabor wavelets transformation, which is to maintain sufficient recognition rate in both time domain and frequency domain, to obtain face feature; moreover apply the strength of locality preserving projection, which preserves the local structure of the multidimensional structure, the immense feature vectors of Gabor wavelets transformation is lowered to minimum. The experiment results show the proposed method has good performance in both face detection and face recognition.

誌 謝

首先要以最誠摯的謝意，感謝我的指導教授林昇甫博士，感謝他在研究和實驗上的啟迪與指導，以及細心校閱本論文的謬誤之處。還有本實驗室的學長士哲、晉嘉、國育、永吉、逸章、建焜、皇村、培家、啟耀、弦澤、松傑；一起學習成長的同學東邑、瑞良、奎彰、育賢、慶安、宥瑾、俊嘉；由於他們熱心的協助與討論，使得研究過程中的困境能夠迎刃而解。

最後要感謝疼愛我的父母，感謝他們的養育之恩以及在我求學生涯中給我最大的鼓勵與支持，讓我得以順利完成學業。



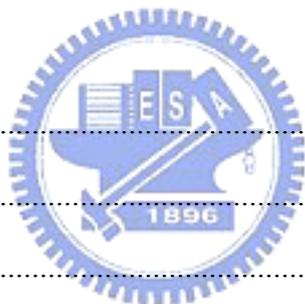
目錄

中文摘要	i
英文摘要	ii
誌謝	iii
目錄	iv
圖目錄	vii
表目錄	ix



第一章 緒論	1
1.1 研究動機.....	2
1.2 相關研究.....	2
1.3 論文架構.....	4
第二章 相關知識與理論	5
2.1 膚色偵測.....	5
2.1.1 色彩空間.....	6
2.1.2 分類方法.....	7
2.2 形態學運算.....	9
2.2.1 基本數學定義.....	9
2.2.2 侵蝕.....	10
2.2.3 膨脹.....	10

2.3 直方圖等化.....	11
2.4 賈伯濾波器.....	13
2.4.1 二維賈伯小波.....	13
2.4.2 二維賈伯濾波器.....	14
2.4.3 二維影像的賈伯小波表示法.....	15
2.5 區域保留投影.....	15
2.5.1 流形學習.....	15
2.5.2 區域保留投影.....	17
2.6 類神經網路.....	19
2.6.1 類神經網路簡介.....	19
2.6.2 倒傳遞類神經網路.....	20
第三章 人臉辨識系統	25
3.1 系統概述.....	25
3.2 特徵抽取.....	27
3.2.1 膚色偵測.....	27
3.2.2 多重解析度視窗搜尋.....	28
3.2.3 直方圖等化.....	30
3.2.4 賈伯小波轉換.....	30
3.2.5 區域保留投影.....	31
3.3 人臉偵測.....	32
3.3.1 訓練類神經網路.....	32
3.3.2 偵測人臉.....	34
3.4 建立人臉資料庫.....	36
3.5 人臉判別.....	37
第四章 實驗結果與分析	38
4.1 實驗機制.....	38



4.1.1 實驗設備	38
4.1.2 效能評估	39
4.2 實驗結果與分析比較.....	40
4.2.1 人臉偵測的實驗結果與分析比較	40
4.2.2 人臉辨識的實驗結果與分析比較	44
4.3 結果討論.....	49
4.3.1 人臉偵測的結果討論	49
4.3.2 人臉辨識的結果討論	51
第五章 結論與未來發展	54
參考文獻	A-I



圖目錄

圖 2.1 色彩空間中的膚色分布，(a)RGB 色彩空間，(b)正規化的 RGB 色彩空間， (c)YCbCr 色彩空間，(d)HSI 色彩空間。.....	8
圖 2.2 使用橢圓邊界模型在 95%偵測正確率的色彩空間邊界圖，(a)R-G 色訊空間， (b)CIE- u^*v^* 色訊空間。[17].....	9
圖 2.3 侵蝕之示意圖，(a)集合 B ，(b)結構元素 S ，(c) B 被 S 侵蝕。.....	10
圖 2.4 膨脹之示意圖，(a)集合 B ，(b)結構元素 S ，(c) B 被 S 膨脹。.....	11
圖 2.5 本論文所使用的賈伯濾波器，有三種尺寸、八個方向，合計 24 種，此為取其 實部的表示法。[21].....	14
圖 2.6 一張人臉影像的賈伯小波表示法，(a)取大小的表示法，(b)取實部的表示法。	16
圖 2.7 倒傳遞類神經網路的架構。.....	22
圖 3.1 人臉辨識系統架構圖。.....	26
圖 3.2 特徵抽取流程圖。.....	27
圖 3.3 膚色偵測的結果，(a)原始相片，(b)經過膚色偵測後的結果，(c)再經過消除雜 訊後的結果。.....	29
圖 3.4 三庭五眼示意圖。.....	30
圖 3.5 經過直方圖等化後的結果，(a)原始影像，(b)原始影像的直方圖，(c)經過直方 圖等化後的影像，(d)直方圖等化後的直方圖。.....	31
圖 3.6 訓練類神經網路流程圖。.....	33
圖 3.7 PIE 資料庫中 13 種不同人臉角度的說明圖。[32].....	33

圖 3.8 倒傳遞神經網路模型。	34
圖 3.9 訓練類神經網路過程的均方誤差。	35
圖 3.10 建立人臉資料庫流程圖。	36
圖 4.1 人臉偵測的部分實驗結果，(a)大多數為正面的人臉，(b)臉部特徵不明顯的正面人臉，(c)有遮擋的正面人臉，(d)各種角度的人臉。	42
圖 4.2 人臉辨識實驗的部分結果，(a)人臉偵測後的結果，(b)人臉辨識的結果，由左至右取出資料庫的人臉做展示。	45
圖 4.3 不同訓練樣本數的人臉辨識實驗結果比較。	48
圖 4.4 不同維度下的人臉辨識實驗結果比較。	48
圖 4.5 人臉偵測的部分結果，(a)因為人臉特徵清楚而偵測成功，(b)因為照明影響造成的偵測失敗，(c)因為角度造成的偵測失敗，(d)因為遮擋造成的偵測位置錯誤，(e)因為遮擋造成的偵測失敗。	51
圖 4.6 人臉辨識的錯誤結果，(a)人臉偵測後的結果，(b)人臉辨識的結果，由左至右取出資料庫的人臉做展示。	53



表目錄

表 4.1 人臉偵測的可能情形。.....	39
表 4.2 三種人臉偵測方法的實驗結果比較。.....	44
表 4.3 不同訓練樣本數的人臉辨識實驗結果。.....	46
表 4.4 不同維度下的人臉辨識實驗結果。.....	47
表 4.5 人臉偵測結果分析。.....	50
表 4.6 人臉辨識結果分析。.....	53



第一章

緒論

近年來，隨著新的資訊科技及儲存媒體的蓬勃發展，人臉辨識(face recognition)在身份識別(identity identification)、基於內容的影像檢索(content-based image retrieval, CBIR)、電腦視覺(computer vision)、人類電腦介面(human computer interaction, HCI)等方面，有越來越多的應用，也吸引了許多學者投入研究這塊領域。而在國家安全、公共安全、金融及軍事應用方面，人臉辨識的技術尤其顯得重要，需求也日益增加。例如應用於門禁系統、相片比對、敏感人物的監控等等。

所謂的人臉辨識，就是指在任意輸入的人臉影像或影片中，把不同大小及姿態的人臉做分析與比對，進而尋找相同的人臉或是區別不同的人臉。因為輸入的人臉影像或影片可能是在各種不同的環境下所取得的，所以光照的影響、拍攝的角度以及成像的條件都會使偵測人臉變得困難；另外，由於人臉的外貌不盡相同、臉部表情變化以及臉部飾品及頭髮所造成的臉部遮擋，也會增加人臉辨識的困難度。為了能夠有較好的辨識成功率，通常都會混合好幾種方法來使用。優點是可以使一種方法的缺點被另一種的優點所彌補，進而形成快速、準確與強健的人臉辨識系統。

在本章中，將在 1.1 節介紹本論文的研究動機；在 1.2 節介紹目前的相關研究；最後在 1.3 節介紹本論文的架構。

1.1 研究動機

現在個人電腦已經變成生活必需品，數位相機也取代了傳統的底片相機，使拍攝相片比以往變得更為方便和容易。以往需要將底片帶到相館沖洗才能得到的照片，變成了在個人電腦上就可以直接觀看的數位相片，一本本厚重的相簿也變成了電腦裡的資料夾與網路上的數位相簿。這些數位相片雖然容易取得及使用，但是當數量越來越多的時候，想要在一大堆的資料夾裡尋找想要的相片，就變成一件不容易的事情。許多人用來整理相片的方法是根據日期或是主題，把相片放到不同的資料夾中，不過當想要尋找某個人的照片時，就缺少一個好的方法來尋找它。想要在相片中尋找人物，就需要人臉辨識，做為其中一項關鍵的人臉偵測，在辨識人臉上扮演了很重要的角色，先將數位相片中的人臉準確的擷取出來，再對這些人臉做分析辨識，才能幫助使用者找尋到所需要的數位相片。因此本論文的重心放在相片裡的人臉辨識這部分的研究，希望能夠發展出一套穩健的相片人臉辨識系統，幫助使用者能夠用來辨識相片中的人臉，以彌補根據日期或是主題來分類相片的不足。

本篇論文是先利用**膚色偵測**(skin detection)把相片中可能不是人臉的部份剔除，再利用**賈伯濾波器**(Gabor filter)取得人臉的整體特徵，然後使用**區域保留投影**(locality preserving projection, LPP)做降維，得到人臉的特徵資訊。然後訓練**類神經網路**(artificial neural networks, ANNs)對輸入影像做**人臉偵測**(face detection)，依照偵測到的人臉建立人臉資料庫，最後在測試相片中辨識選定的人臉。

1.2 相關研究

不管是人臉偵測或是人臉辨識，所使用的方法都具有相同的架構，也就是選取人臉特徵，再使用分類方法來區分人臉。目前不管是在特徵選取上或是分類方法上的研究都很多，因此以下分為特徵選取和分類方法這兩個部分來做說明。首先從人臉特徵的選取來分，大致上可分為三種[1]：

1. 整體的方法(holistic methods)

將整張人臉當做輸入來處理的方法，主要的挑戰是遇到非常小的人臉時要如何得到足夠的人臉資訊。

2. 局部的方法(local methods)

使用局部人臉特徵的方法，如何合併整張臉的資訊到局部的模型會是主要的問題。

3. 混合的方法(hybrid methods)

同時使用整體及局部的方法，通常會比單純使用整體或局部的方法要來的好，但是也會比較複雜。

接下來，從使用的分類方法來分，則可以分成以下四類[2]：

1. 基於知識的方法(knowledge-based methods)

也就是基於規則(rule-based)的方法，基於人類知識對於一個典型臉部的組成，以及臉部特徵的彼此關係的敘述。

2. 不變特徵的方法(feature invariant approaches)

目的在於尋找存在於不同的姿勢、觀察點、或照明情況下的結構特徵，然後用來鎖定臉部。例如臉部特徵、紋理及膚色等等。

3. 符合樣板的方法(template matching methods)

儲存幾種標準的臉部圖樣來描述一整張臉或是單獨的臉部特徵，用計算輸入影像和儲存的圖樣的方法來判斷。

4. 基於外貌的方法(appearance-based methods)

和符合樣板的方法不同，臉部模型(或樣板)是由一群訓練影像中學習出來的，可以掌握臉部外貌的形態的變化。

在上述的幾種方法中，基於外貌的方法目前有較多的學者投入研究。通常是經由靜態分析和機器學習來判別臉部和非臉部特徵之間的差異，然後再使用減少維度來增加運算效率和偵測率。例如 Turk 等人[3]將主成份分析(principal component analysis, PCA) 使用在人臉與非人臉影像樣本，以訓練一個投影的子空間，叫做**特徵臉(eigenface)**，然後

再把輸入影像投影在這個子空間上來做分類。Rowle 等人[4]將三種尺寸的訓練影像樣本分別使用類神經網路來學習，再經過一個網路來仲裁輸出結果。Osuna 等人[5]則使用**支持向量機**(support vector machine)來學習訓練樣本，然後再對輸入影像重新取樣成好幾種尺寸來掃描。Rajagopalan 等人[6]使用子影像樣本來計算**隱馬可夫模型**(hidden Markov model, HMM)的狀態轉換圖，然後使用**k-均值**(k-means)演算法來分類，最後再用隱馬可夫模型來學習分類好的結果。

由於光照變化和人臉角度都會增加人臉偵測與辨識的困難度，因此也有不少學者是針對此類問題做研究，Georghiades 等人[7]使用**照明錐**(illumination cone)建出一個三維臉部模型，可以重建固定視角下各種光源二維臉部的樣貌，因而處理了光線變化的問題。Wiskott 等人[8]則使用**彈性群圖形匹配**(elastic bunch graph matching)的方法，根據人臉上的特徵點建立一個圖形並允許此圖變形，使得在不同視角下臉部的特徵點仍然位於臉上的相對位置，因此能解決人臉角度不同的問題。



1.3 論文架構

本篇論文剩餘章節的安排如下：論文使用到的一些相關知識與理論會在第二章作簡要的描述，針對數位相片來辨識人臉的系統會在第三章提出並且作詳細的說明，實驗結果及分析會詳列在第四章，最後第五章會介紹結論與未來發展。

第二章

相關知識及理論

本章將會介紹本篇論文所使用到的膚色偵測、直方圖等化(histogram equalization)、賈伯濾波器、區域保留投影及類神經網路，並做簡單扼要的介紹。為了能實現一個穩健的人臉辨識系統，本論文先使用膚色偵測將相片中不可能為人臉的部份剔除，再使用直方圖等化來增強影像的對比程度，然後利用賈伯濾波器取得人臉的整體特徵，並利用區域保留投影將人臉特徵降至很低的維度，再訓練類神經網路來判別是否為人臉，根據人臉偵測的結果用人工做分類建立資料庫，透過對資料庫中每個人臉都訓練一個類神經網路來辨識人臉。

在本章中，對於膚色偵測的相關理論會在 2.1 節中做介紹；在 2.2 節介紹形態學運算以及其中的**侵蝕**(erosion)與**膨脹**(dilation)；在 2.3 節會介紹直方圖等化；在 2.4 節說明賈伯濾波器的理論；在 2.5 節介紹區域保留投影；最後在 2.6 節介紹類神經網路。

2.1 膚色偵測

想在相片中尋找人臉的時候，必須對整張相片來尋找，而且要做多重解析度搜尋不同大小的人臉。這樣的運算相當耗時，如果能夠先對相片做膚色偵測，去除掉不含人臉的部分，就可以加快搜尋人臉的速度。由相關的研究結果[9]可以得知，膚色偵測在人臉

偵測這部份的影像處理應用，扮演了很重要的角色，尤其是在人臉偵測、人臉追蹤和手勢分析等等。膚色特徵是隨著照明條件、相機參數、人種及不同的身體部位而有不同的特性。從分類的觀點來看，就是要把膚色與非膚色分開來，因此分類的步驟可以分為：

1. 選擇適合表示膚色影像的色彩空間。
2. 產生用來區分膚色與非膚色的模型。
3. 把做好的模型拿來分類。

2.1.1 色彩空間

選擇色彩空間是做膚色偵測的第一步，依照特性可以分成三種。因為區分膚色與非膚色時是基於色訊(chrominance)的不同，所以一般做法都會儘量的除去亮度部份，以避免因亮度變化而影響膚色偵測的效能。以下介紹幾種常用的色彩空間及優缺點的比較：

1. RGB：用於電腦螢幕顯示及數位影像的色彩空間，其中 R 為紅色，G 為綠色，B 為藍色。為了減少照明因素及人種膚色不同的影響，必須先將 RGB 值正規化，把 R、G、B 除上(R+G+B)，使得 R+G+B=1。正規化後的 RGB 色彩空間在膚色上有低變化的特性，因此就適合用於膚色偵測[10]。
2. HSI：人眼知覺(perceptual)的色彩空間，H 為色調(hue)，S 為飽和度(saturation)，I 為強度(intensity)。色調是色彩從紅色到綠色的變化，飽和度是指有多少比例的白色混入純色中，強度(也就是亮度)是指色彩從黑到白的變化。因為亮度被分離出來，所以也適合用於膚色偵測[11]。以下是用來將 RGB 色彩空間轉換成 HSI 色彩空間的公式：

$$H = \cos^{-1} \frac{\frac{1}{2}((R-G) + (R-B))}{\sqrt{(R-G)^2 + (R-B)(G-B)}}, \quad (2.1)$$

$$S = 1 - (3/(R+G+B)) \times \min(R, G, B), \quad (2.2)$$

$$I = (R+G+B)/3. \quad (2.3)$$

3. YCbCr：正交的色彩空間，Y 為亮度(luminance)，Cb 與 Cr 是藍色與紅色的色訊。因為亮度部分被獨立出來，所以很適合用於膚色偵測，是最多人使用的色彩空

間[12]。以下是將 RGB 色彩空間轉換成 YCbCr 色彩空間的公式：

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix}. \quad (2.4)$$

比較上述的三種色彩模型後可得知，在進行影像分析時，每張不同的相片都會遇到不同的照明環境，因而得到不同的亮度，為了降低亮度對尋找膚色的影響，就必須將亮度作分離，觀察圖 2.1 可知，膚色在除去亮度部分後是較為集中的，有利於膚色的尋找。

2.1.2 分類方法

用於分類膚色與非膚色的方法很多，簡單的方法像是 Tomaz 等人[13]使用修改自 HSI 色彩空間的 TSL，T 是加上白色的色調，L 跟 I 一樣都是亮度，用取 H 與 S 閾值的方法來區分膚色與非膚色。Jones 等人[14]將三維影像直方圖轉換成機率分佈，計算膚色與非膚色直方圖轉換成的機率分佈，使用貝氏分類器來決定是屬於哪一類，藉以把膚色與非膚色分類出來。而需要預先學習的方法像是 Seow[15]等人使用倒傳遞(back propagation)類神經網路，透過膚色與非膚色的樣本來訓練一個三層的類神經網路，然後用訓練過的網路來分類膚色與非膚色。Sebe 等人[16]則使用貝氏網路(Bayesian network)於膚色模型上，特別的是他們使用了隨機結構搜尋(stochastic structure search)演算法在貝氏網路的結構的學習上，這個方法改進了系統的效能。另外，Yang 等人[17]使用的高斯混合模型(Gaussian mixture models)則是一個熱門的方法。因為不同人種的膚色群集在色彩空間是一個小區域，用單個高斯分佈不足以很好的表示出來。再加上不同的照明環境的影響，因此使用多個高斯函數的高斯混合模型就能有更好的準確度。

接下來要介紹的是本論文使用的橢圓邊界模型(elliptical boundary model)，由 Lee 等人[18]所提出。使用了很多種的色彩空間來做實驗及比較：包含 RGB 色彩空間、修改自 RGB 色彩空間的 CIE-l*u*v*、CIE-l*a*b 和 CIE-xyz；YCbCr 及修改自 YCbCr 色彩空間的 YIQ 等等。

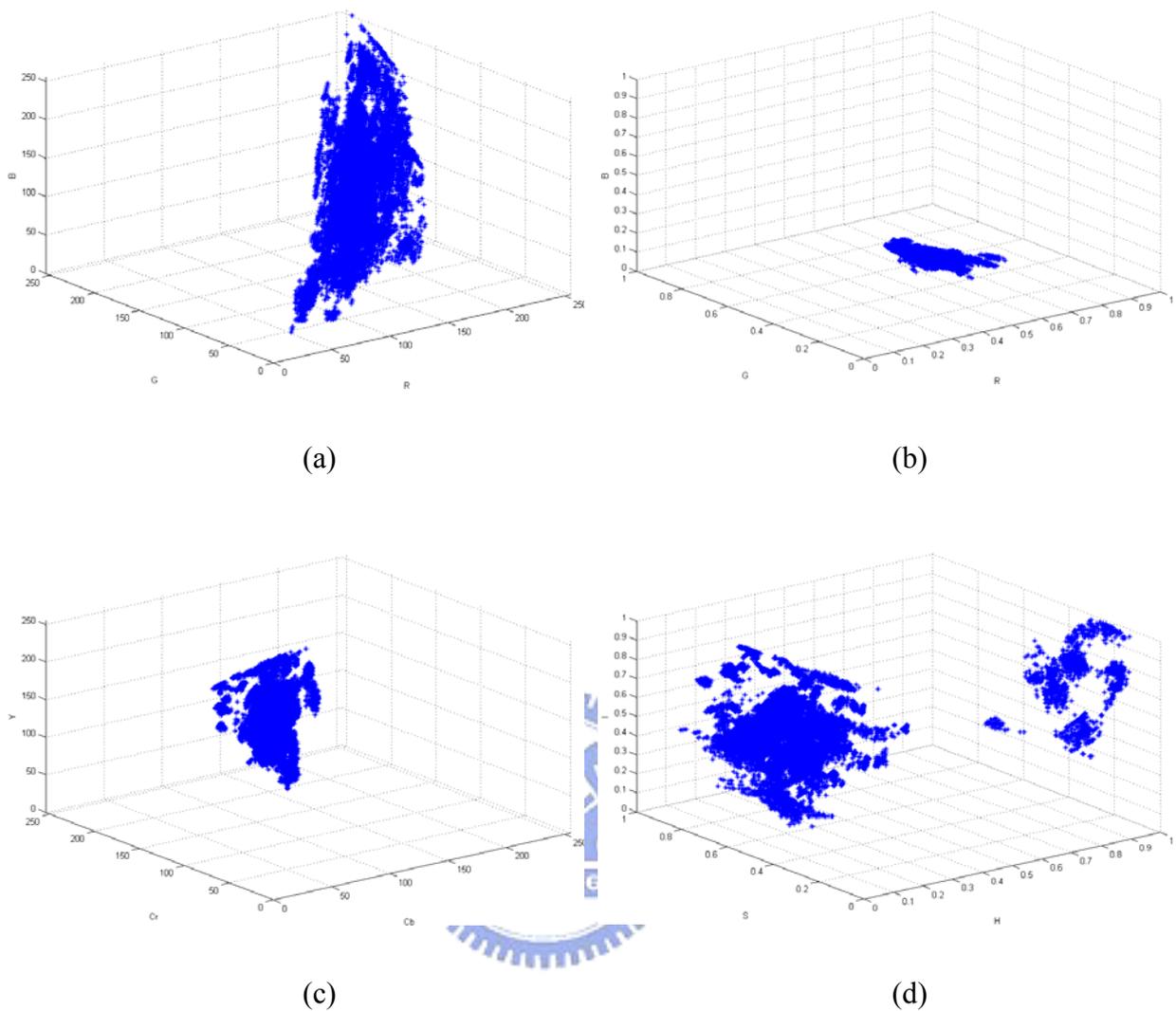


圖 2.1 色彩空間中的膚色分布，(a)RGB 色彩空間，(b)正規化的 RGB 色彩空間，(c)YCbCr 色彩空間，(d)HSI 色彩空間。

與高斯混合模型相比，橢圓邊界模型的運算較為簡單。而且也和單高斯模型(single Gaussian model)一樣，只需要少數的訓練樣本就可以有不錯的效果。橢圓邊界模型的公式為：

$$\Phi(c) = [c - \Psi]^T \Lambda^{-1} [c - \Psi], \quad (2.5)$$

這裡的 c 是色彩向量， Ψ 和 Λ 是模型的參數，定義為：

$$\Psi = \frac{1}{n} \sum_{i=1}^n c_i, \quad \Lambda = \frac{1}{N} \sum_{i=1}^n f_i (c_i - \mu)(c_i - \mu)^T, \quad (2.6)$$

這裡的 N 是訓練樣本的總數， f_i 是 c_i 的色訊樣本總數， μ 是訓練資料集合裡色訊向量的平均值。當 $\Phi(c) < \theta$ ，輸入色訊向量 c 的像素就會被判斷為膚色，這裡的 θ 是由經驗法則所決定的閾值，當 θ 越大時，偵測正確率會增加，但是錯誤率也會隨之增加。圖 2.2 為在 R-G 色訊空間及 CIE- u^*v^* 色訊空間下的膚色分佈圖，出自[17]。

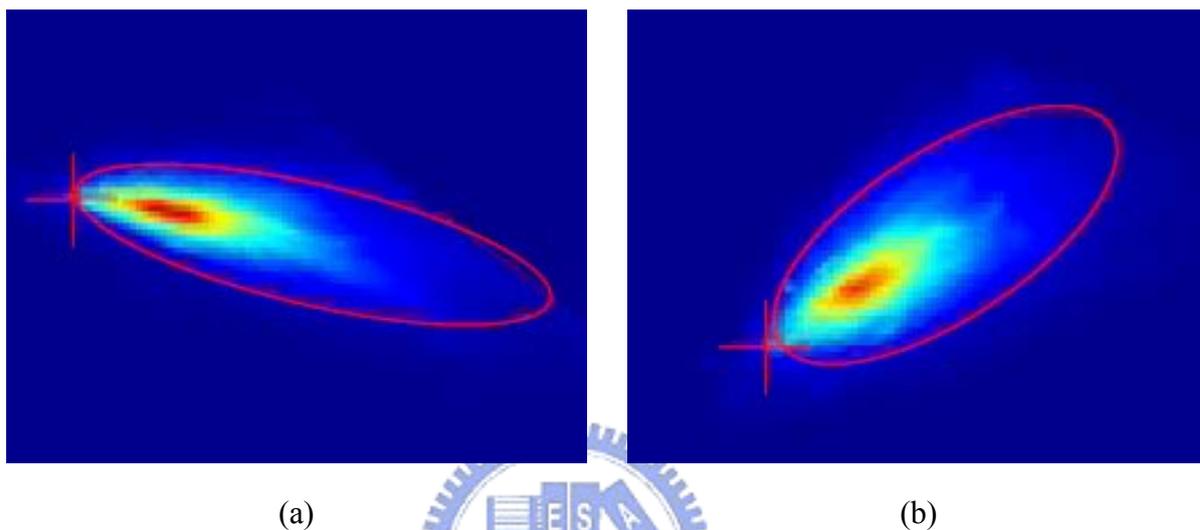


圖 2.2 使用橢圓邊界模型在 95%偵測正確率的色彩空間邊界圖，(a)R-G 色訊空間，(b)CIE- u^*v^* 色訊空間。[17]

2.2 形態學運算

形態學之重點是在於研究物體的形狀或結構，也可說是在探討物體內部之間的相互關係。在數位影像處理之應用領域中又稱其為**數位形態學**(digital morphology)，視為分析影像幾何結構之方法，例如：邊界、輪廓與凸形封包等[19]。

2.2.1 基本數學定義

假設集合 B 與 S 分別含元素 $b = (b_1, b_2)$ 與 $s = (s_1, s_2)$ ，其中 b_i 與 s_i ， $i = 1, 2$ 分別代表整數的座標軸位置。

定義 B 為**水平移動**(translation)， $x = (x_1, x_2)$ 符號可表示成 $((B)_x)$ 表示為：

$$(B)_x = \{a \mid a = b + x, \text{對於 } b \in B\}. \quad (2.7)$$

定義 S 之**反射**(reflection)或稱映像(符號為 \widehat{S})為：

$$\widehat{S} = \{x \mid x = -s, \text{對於 } s \in S\}. \quad (2.8)$$

定義集合 B 之**補集**(complement)(符號為 B^c)為：

$$B^c = \{x \mid x \notin B\}. \quad (2.9)$$

最後定義集合 B 與 S 之**差**(difference)為：

$$B - S = \{x \mid x \in B, x \notin S\}. \quad (2.10)$$

2.2.2 侵蝕

這種方法通常用於物體的邊界上，目的是將物體往內收縮若干像素，使其周圍有種被慢慢吞蝕掉的現象，所以其名為**侵蝕**[19]，數學表示為：

$$E = B \ominus S = \{x \mid (S)_x \subseteq B\}. \quad (2.11)$$

意即被 S 侵蝕之 B 是 S 平行移動 x 之後包含於 B 中所有 x 的集合於圖 2.3 可見。

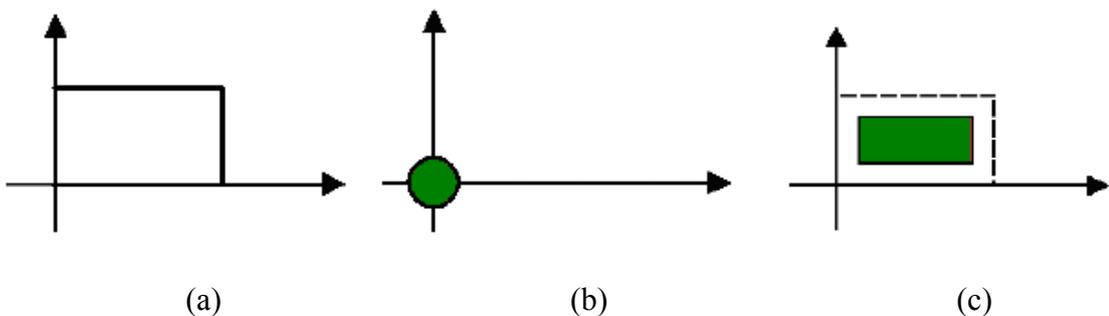


圖 2.3 侵蝕之示意圖，(a)集合 B ，(b)結構元素 S ，(c) B 被 S 侵蝕。

2.2.3 膨脹

由名字可知其意為與侵蝕動作相反之動作[19]，目的在於將物體的邊界往外膨脹若干像素之方法，數學可表示為：

$$D = B \oplus S = \{x | (\hat{S})_x \cap B \neq \emptyset\}. \quad (2.12)$$

意即所有可以使 \hat{S} 在平行移動 x 後與 B 至少重疊一個非零元素的集合如圖 2.4 所示。

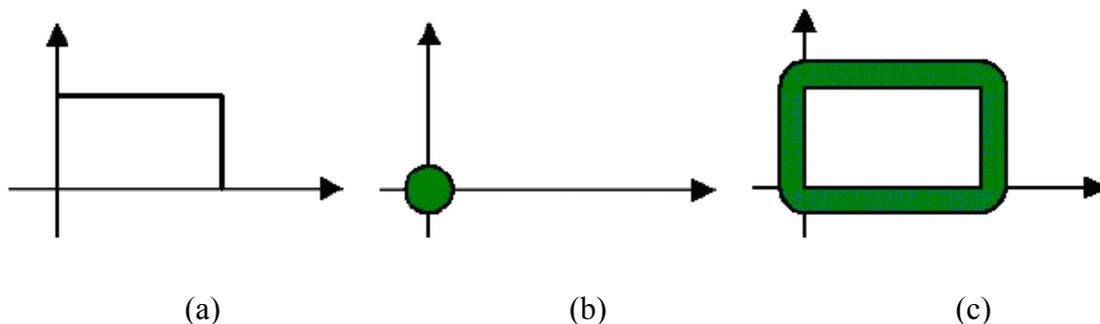


圖 2.4 膨脹之示意圖，(a)集合 B ，(b)結構元素 S ，(c) B 被 S 膨脹。

2.3 直方圖等化

在本論文中，所處理的照片常常會因人臉與背景顏色太過相近，而導致在前處理後還殘有雜訊，因此為了要減少這類雜訊所造成的誤判，使用直方圖等化來增強數位影像的對比程度。

一幅影像中的灰階可以看成區間 $[0,1]$ 上的隨機變數。一個隨機變數的最基本描述子是它的**機率密度函數**(probability density function)，令 $p_r(r)$ 和 $p_s(s)$ 表示隨機變數 r 和 s 的機率密度函數，假如 $p_r(r)$ 和 $T(r)$ 已知，且 $T^{-1}(s)$ 滿足 $T(r)$ 在 $0 \leq r \leq 1$ 區間上是單值且單調遞增的，則轉換後的變數 s 的機率密度函數可以用相當簡單的方法獲得：

$$p_s(s) = p_r(r) \left| \frac{dr}{ds} \right|, \quad (2.13)$$

因此轉換後變數 s 的機率密度函數可以由輸入影像的灰階機率密度函數和經選取的轉換函數所決定。一個在影像處理上特別重要的轉換函數有這樣的型式：

$$s = T(r) = \int_0^r p_r(w) dw, \quad (2.14)$$

其中 w 是虛設的積分變數。(2.14)式的右邊是隨機變數 r 的**累積分布函數**(cumulative

distribution function)。由於機率密度函數永遠是正的，而且函數的積分就是函數的面積，所以其結果是此轉換函數是單值且單調遞增的。同理變數在範圍 $[0,1]$ 內的機率密度函數的積分也同樣在範圍 $[0,1]$ 內。

已知一個轉移函數 $T(r)$ ，應用(2.13)式可求得 $p_s(s)$ 。從基本的微積分學的萊布尼茲(Leibniz)法則可得知，對一個定積分上限微分就是被積分式在該極限的值。也就是說：

$$\frac{ds}{dr} = \frac{dT(r)}{dr} = \frac{d}{dr} \left[\int_0^r p_r(w)dw \right] = p_r(r), \quad (2.15)$$

以此 ds/dr 的結果代入(2.13)式中，並且所有的機率值都為正數，可得：

$$p_s(s) = p_r(r) \left| \frac{dr}{ds} \right| = p_r(r) \left| \frac{1}{p_r(r)} \right| = 1, \quad 0 \leq s \leq 1. \quad (2.16)$$

因為 $p_s(s)$ 是一個機率密度函數，所以在此情況下它在區間 $[0,1]$ 外必須是0，因為對所有 s 值的積分一定要等於1。可以得出(2.16)式之 $p_s(s)$ 的形式是一個均勻的機率密度函數，產生一個特性為均勻機率密度函數的隨機變數 s 。由(2.16)式可看出，不論 $p_r(r)$ 的形式為何，所得的 $p_s(s)$ 永遠是均勻的。

由於數位影像為離散值，所以針對離散值，討論的是機率總和而不是機率密度函數與積分。在一幅影像中的灰階 r_k 發生的機率可近似為：

$$p_r(r_k) = \frac{n_k}{n}, \quad k = 0, 1, \dots, L-1, \quad (2.17)$$

其中 n 是影像中的像素總數， n_k 是有灰階 r_k 之像素的數目， L 是影像中可能灰階的總數。

(2.14)式所給之轉換函數的離散形式為：

$$s_k = T(r_k) = \sum_{j=0}^k p_r(r_j) = \sum_{j=0}^k \frac{n_j}{n}, \quad k = 0, 1, 2, \dots, L-1. \quad (2.18)$$

因此經由(2.18)式將輸入影像中灰階 r_k 的每一像素映射至輸出影像中灰階 s_k 的對應像素中獲得一個處理過的影像， $p_r(r_k)$ 對 r_k 的圖稱為直方圖，經過(2.18)式的轉換後，輸出便為直方圖等化。

2.4 賈伯濾波器

1964年 Gabor[20]提出了一維的賈伯小波，1985年 Daugman[21]提出了二維的賈伯小波。賈伯小波的特性是能夠同時在時域和頻域中兼顧足夠的分辨率來做訊號分析，因此在訊號處理的領域得到很廣泛的應用，而近年來也有不少研究將其應用在人臉特徵抽取(feature extraction)的方面[22]。

2.4.1 二維賈伯小波

在空間域，二維賈伯小波是高斯函數被一個正弦平面波所調變的結果[23]，可以表示為：

$$\varphi(x, y) = \frac{f^2}{\pi\gamma\eta} \exp(-(-\alpha^2 x_r^2 + \beta^2 y_r^2)) \exp(j2\pi f x_r), \quad (2.19)$$

$$x_r = x \cos \theta + y \sin \theta, y_r = -x \sin \theta + y \cos \theta,$$

這裡的 f 是正弦平面波的中央頻率， θ 是橢圓高斯主軸的方向， α 是高斯主軸平行於波的銳利度， β 是高斯最小軸垂直於波的銳利度， $\gamma = \frac{f}{\alpha}$ 和 $\eta = \frac{f}{\beta}$ 定義為保持頻率和銳利度常數的比率。

因此，(2.19)式的傅利葉轉換式為：

$$\Phi(u, v) = \exp(-\pi^2(-\frac{(u_r - f)^2}{\alpha^2} + \frac{v_r^2}{\beta^2})), \quad (2.20)$$

$$u_r = u \cos \theta + v \sin \theta, v_r = -u \sin \theta + v \cos \theta.$$

為了找到不同賈伯小波間的關係，定義一個波向量 $\vec{k} = 2\pi f \exp(j\theta)$ 來表示在頻域的中央頻率部分，這裡假設波向量的方向都和橢圓高斯主軸的方向相同，然後再假設 $\gamma = \eta = \frac{\sigma}{\sqrt{2\pi}}$ ，即 $\alpha = \beta = \frac{\sqrt{2\pi}f}{\sigma}$ 。因此在 $\vec{z} = (x, y)$ 位置的賈伯小波就可以表示為：

$$\varphi(\vec{z}) = \frac{\|\vec{k}\|^2}{2\pi\sigma^2} \exp(-\frac{\|\vec{k}\|^2 \|\vec{z}\|^2}{2\sigma^2}) \exp(j\vec{k} \cdot \vec{z}). \quad (2.21)$$

2.4.2 二維賈伯濾波器

接下來介紹二維的賈伯濾波器，可以把(2.21)式改寫成：

$$\Psi_{s,d}(\vec{z}) = \frac{\|\vec{k}_{s,d}\|^2}{\sigma^2} \exp\left(-\frac{\|\vec{k}_{s,d}\|^2 \cdot \|\vec{z}\|^2}{2\sigma^2}\right) [\exp(j\vec{k}_{s,d} \cdot \vec{z}) - \exp(-\frac{\sigma^2}{2})], \quad (2.22)$$

這裡的 $\vec{z} = (x, y)$ ， $\vec{k}_{s,d} = k_s \exp(j\theta_d)$ ， $k_s = \frac{k_{\max}}{f^s}$ 為頻率， $\theta_d = \frac{\pi d}{8}$ ， $\theta_d \in [0, \pi)$ 為方向， k_{\max}

為最大頻率， f 是賈伯濾波器在頻域的間隔係數(spacing factor)。波向量 $\vec{k}_{s,d}$ 控制賈伯濾波器的空間解析度和方向， s 控制賈伯濾波器的尺寸，也就是決定濾波器的頻域中心；而 d 控制的是賈伯濾波器的方向；尺寸係數 k_s 使得不同空間頻率的濾波器有大致相等的能量； σ 決定了高斯窗(Gaussian window)的寬度與波向量長度的比例。由於(2.22)式的虛部分量的平均值為零，但實部分量的平均值不為零，所以要減去 $\exp(-\frac{\sigma^2}{2})$ ，使得濾波器對於照明的變化敏感度降低[24]。

本論文使用的是三種尺寸($k_{\max} = \frac{5\pi}{4}$, $f = \sqrt{2}$, $s \in 0, \dots, 2$)、八個方向($d \in 0, \dots, 7$)，並設定 $\sigma = 2\pi$ 的賈伯濾波器，如圖 2.5 所示。

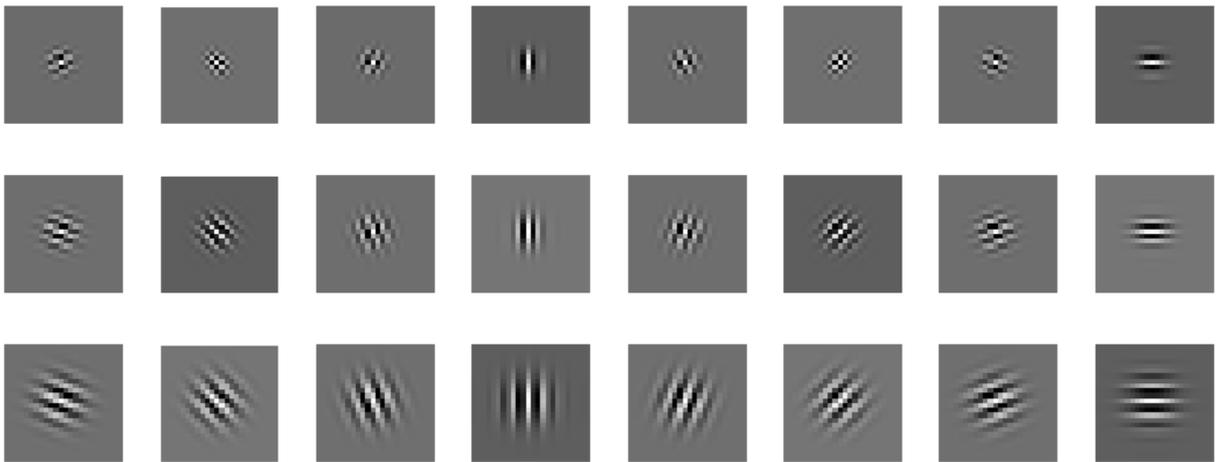


圖 2.5 本論文所使用的賈伯濾波器，有三種尺寸、八個方向，合計 24 種，此為取其實部的表示法。[21]

2.4.3 二維影像的賈伯小波表示法

使用賈伯小波來表示一張二維的影像，就是和賈伯濾波器做摺積(convolution)後的結果：

$$G_{u,v}(\vec{z}) = I * \Psi_{s,d}(\vec{z}), \quad (2.23)$$

這裡的 I 是輸入影像， $\Psi_{s,d}(\vec{z})$ 是(2.22)式的賈伯濾波器，圖 2.6 為一張人臉影像的賈伯小波表示法。

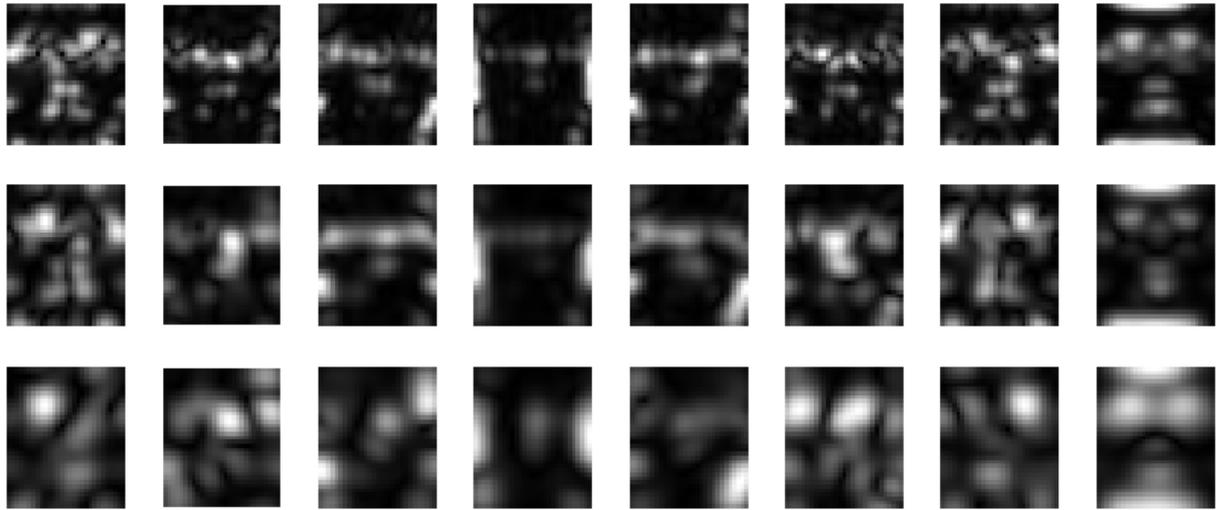
2.5 區域保留投影

在這節將介紹區域保留投影的理論及演算法，是流形學習(manifold learning)中的一種方法。首先在 2.5.1 節對流形學習做簡單的介紹，然後在 2.5.2 節詳細說明區域保留投影的演算過程。

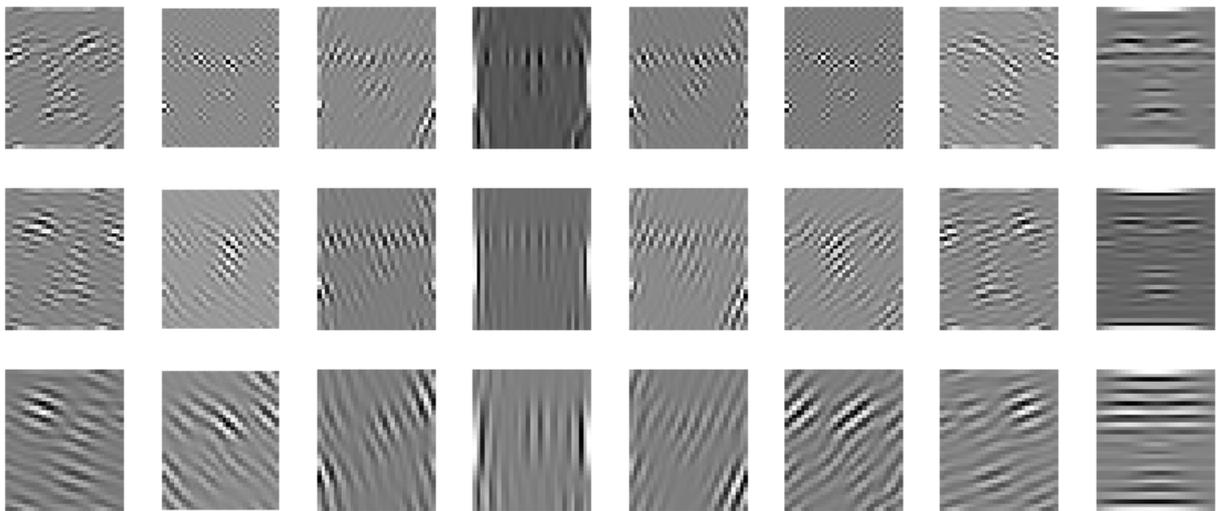


2.5.1 流形學習

流形學習是一種新的非監督式學習(unsupervised learning)方法，可以有效的將高維度非線性的資料做降維，因此近年來逐漸受到重視，目前已經有很多有效的流形學習方法，如 Roweis 等人[25]所提出的局部線性嵌入(locally linear embedding, LLE)，主要是利用在嵌入空間中，每個採樣點可以表示成它的近鄰點的線性組合，因而能在低維空間中保持每個鄰域中的權值不變，重建原數據點使得重建誤差最小；Tenenbaum 等人[26]所提出的等距映射(isometric feature mapping, ISOMAP)是建立在多維尺度變換(multidimensional scaling, MDS)的基礎上，力求保持資料點的內在幾何性質，即保持兩點間的測地距離。



(a)



(b)

圖 2.6 一張人臉影像的賈伯小波表示法，(a)取大小的表示法，(b)取實部的表示法。

上述兩種是目前常用的非線性流形學習方法，優點在於可以保存資料分布中局部的非線性特性；而傳統線性的方法，如主成份分析(principal components analysis, PCA)和線性識別分析(linear discriminant analysis, LDA)則是維持全局的特性，在現實中很多分類的問題都是和局部的非線性特性比較有關，因此 LLE 和 ISOMAP 應用在這些問題上會有較好的效果。而其缺點則是運算時間較長、主要著重於訓練資料的表示而不是分類、當有新的資料時就需要重新學習來降維等。

因此本論文使用 He 等人[28]所提出的區域保留投影，這個方法是基於 Belkin 等人

[27]所提出的拉普拉斯特徵映射(Laplacian eigenmap)。拉普拉斯特徵映射也是非線性的方法，其主要思想是在高維空間中離得很近的点投影到低維空間中也應該離得很近，通過使用兩點間的加權距離作為損失函數，可求得相應的降維結果。而區域保留投影的概念就是對拉普拉斯特徵映射做線性逼近，因此運算速度較快，同時它也找出了一組投影矩陣，在增加新資料時不需要重新學習就可以找到對應的低維空間表示等優點。

2.5.2 區域保留投影

本節將對區域保留投影方法做詳細的說明，給定一組輸入資料 $X = x_1, x_2, \dots, x_m$ 為有 m 筆 n 維的資料，找一個轉換矩陣 A 使輸入資料 X 對應到 $Y = y_1, y_2, \dots, y_m$ 為 m 筆 l 維的輸出資料 ($l < n$)，這樣就可以用 y_i 來表示原本的 x_i ，這裡的 $y_i = A^T x_i$ 。

從拉普拉斯特徵映射法中根據一組輸入資料 X 所建立的權重圖(weighted graph) G 可以判斷任兩個資料點中是否有邊相連，假設 $Y = y_1, y_2, \dots, y_m$ 也跟 X 有相同的權重圖 G ，那麼根據拉普拉斯特徵映射法，最好的映射就是 $\sum_{ij} (y_i - y_j)^2 W_{ij}$ 之最小值，這裡的 W_{ij} 是兩個資料點 y_i 與 y_j 的權重值，也就是若 $\sum_{ij} (y_i - y_j)^2 W_{ij}$ 越小，映射後的 Y 就會越接近原始資料的 X 。再假設這樣的映射是線性的，例如 $y_i = a^T x_i$ ，那麼可以推導(2.24)如下：

$$\begin{aligned}
 & \frac{1}{2} \sum_{ij} (y_i - y_j)^2 W_{ij} \\
 &= \frac{1}{2} \sum_{ij} (a^T x_i - a^T x_j)^2 W_{ij} \\
 &= \sum_i a^T x_i D_{ii} x_i^T a - \sum_{ij} a^T x_i W_{ij} x_j^T a \\
 &= a^T X(D - W)X^T a \\
 &= a^T XLX^T a,
 \end{aligned} \tag{2.24}$$

這裡的 D 是對角矩陣， $D_{ii} = \sum_j W_{ij}$ ， $L = D - W$ 為拉普拉斯矩陣，若 D_{ii} 越大，對應的 y_i 也越重要，因此求 $\sum_{ij} (y_i - y_j)^2 W_{ij}$ 最小值的問題就變成求 $a^T XLX^T a$ 的最小值，再根據拉普拉斯特徵映射法來限制 $a^T XLX^T a = 1$ 就可以移除映射中對於任意的放大因素的問題。

最後轉換拉普拉斯特徵映射法中的拉普拉斯算式 $LY = \lambda DY$ 為 $a^T XLX^T a = \lambda a^T XDX^T a$ ，也就是 $XLX^T a = \lambda XDX^T a$ ，可以看出這裡的 XLX^T 與 XDX^T 都是對稱且正半定義(positive semi-definite)的，根據以上的推導過程就可以將原本拉普拉斯演算法用線性的方式做逼近。接著將解 A 的演算法步驟做說明：

1. 建構相鄰無向圖(adjacency undirected graph)

假設 G 是一個有 m 個點的圖，用來判斷任兩個節點 i 與 j 之間是否有邊相連，即 x_i 和 x_j 是否接近，有以下兩種方式可依照需求選擇：

(1) ϵ -相近(ϵ -neighborhoods)

若 $\|x_i - x_j\|^2 < \epsilon$ ，即兩點間的歐基里德距離(Euclidean distance)小於某個常數 ϵ ，則兩節點 i 與 j 之間有邊相連。

(2) k -相近(k -nearest neighbors)

若兩節點 i 與 j 中，有一節點在另一節點於整組資料中最相近的 k 個點中，則兩點之間有邊相連。

2. 選擇權重(weights)

權重矩陣 W 是一個 $m \times m$ 的稀疏對稱矩陣， W_{ij} 是邊相連的兩個頂點 i 與 j 的權重值，

若兩頂點之間沒有邊相連，則權重值為 0，依照下列兩種方式來選擇：

(1) 熱核心(heat kernel)

如果兩節點 i 與 j 之間有邊相連，則 $W_{ij} = e^{-\frac{\|x_i - x_j\|^2}{t}}$ 。

(2) 簡易指定

如果兩節點 i 與 j 之間有邊相連，則 $W_{ij} = 1$ 。

3. 特徵映射

為了得到一般化的特徵向量問題裡的特徵向量和特徵值，求解以下的算式：

$$XLX^T a = \lambda XDX^T a, \quad (2.24)$$

這裡的 D 是對角矩陣， $D_{ii} = \sum_j W_{ij}$ ， $L = D - W$ 為拉普拉斯矩陣， λ 為特徵值，矩陣 X

的第 i 行元素為 x_i 。假設行向量 a_0, a_1, \dots, a_{l-1} 為(2.24)式的解，根據它們的特徵值做排序， $\lambda_0 < \lambda_1 < \dots < \lambda_{l-1}$ 。因此輸入資料 x_i 就被降維成所想要維度的輸出資料 y_i ，即 $x_i \rightarrow y_i = A^T x_i$ ， $A = a_0, a_1, \dots, a_{l-1}$ ，這裡的 y_i 是一個有 l 維度的向量， A 是一個 $n \times l$ 的矩陣。

2.6 類神經網路

在本論文的演算法中，將會使用到倒傳遞類神經網路架構，所以在 2.4.1 節中，先介紹何謂類神經網路，並在 2.4.2 節說明倒傳遞類神經網路的架構與方法。

2.6.1 類神經網路簡介

類神經網路的理論最早是在 1958 年時 Rosenblatt[29]仿造生物的神經系統，提出了**感知器**(perceptron)的神經元模型開始。後來在 1982 年由 Hopfield[30]提出了神經網路，才使得類神經網路逐漸受到大家的重視。類神經網路的架構是由很多非線性的運算單元，也就是**神經元**(neuron)和位於這些運算元間的眾多連結所組成，其建構的目標在於映射輸入與輸出之間的數學。類神經網路具有高度平行處理、學習、記憶和容錯能力等優點，所以被廣泛的應用在各個領域。

而類神經網路是以電腦的軟硬體來模擬生物神經網路的資訊處理系統，從人類專家解決問題的實際案例中學習，利用非線性函數的轉換，能有效地對大量資料進行分析，而且具備學習能力，可利用來做各種非結構性決策的制定。另外，類神經網路之應用不需要前提假設，只要有充足的歷史資料，即可進行分析，例如有完整之交易資訊的股票市場，或精確且眾多的氣象資料，皆十分適合運用網路模式進行分析，預測的工作。

類神經網路有大量相互連結的處理單元，通常是以平行的方式操作且分散於整個網路結構之中，而整個類神經網路的聚集形式就如同人類的大腦一般，可透過樣本或資料的訓練來展現出學習、回想、歸納推演的能力。類神經網路在處理**圖形辨識**(pattern

recognition)、分類(classification)、函數近似(function approximation)、最佳化(optimization)及資料分類(data clustering)上都有很好的效果。

以學習策略來分類，可以分為**監督式學習**(supervised learning)，非監督式學習及**增強式學習**(reinforcement learning)等三種。監督式學習是在訓練資料裡包含輸入及輸出的資料，讓網路從中學習輸入資料與輸出資料的內在對映規則。這類學習方式就像老師指導學生對於問題做正確的回答，經常應用於圖形辨識及預測。非監督式學習是在缺乏期望輸出值的情況下，自行發掘出資料中哪些特徵是重要的或是可忽略的，然後將資料做分類，輸出值代表的是輸入的資料相似度有多少。增強式學習則是先定義每一種情況的好或壞，也就是給機器一套價值觀，之後讓機器自行去學習。機器觀察很多資料以後，作出不同的反應，但是不同的反應可能會有不同的結果，結果的好壞就成為**回授**(feedback)的資料，讓機器學會下次遇到類似的情形如何反應。

而以網路架構來區分，可以分為**前饋**(feed forward)類神經網路與**回饋**(feedback)類神經網路兩種。前饋類神經網路之架構為單方向的向前傳遞連結，且網路的所有神經元均無後向或是側向的傳遞連結。而回饋類神經網路與前饋類神經網路最大的不同，在於回饋類神經網路至少會有一回饋迴圈。本論文所用到之**倒傳遞**(back propagation)類神經網路屬於監督式的前饋類神經網路。

2.6.2 倒傳遞類神經網路

倒傳遞類神經網路的架構為**多層感知器**(multi layer perceptron, MLP)，一般使用的學習演算法為**誤差倒傳遞演算法**(error back propagation, EBP)，這樣的組合(MLP+EBP)稱之為倒傳遞類神經網路。倒傳遞演算法屬多層前饋式網路，以**監督式**(supervised)學習方式，來處理輸入輸出間之非線性映射關係，由於早期 Rosenblatt[29]發展的感知器僅為單層網路，並無隱藏層的存在，無法解決某些問題，使得類神經網路的發展有很長一段停滯期。直到 Rumelhart 等人[31]發展出倒傳遞學習法，才開始有了突破。

倒傳遞學習法的提出，使得多層感知器的網路架構得以實現，一般常用的多層感知器網路架構在輸入層與輸出層之間有**隱藏層**(hidden layer)，之所以稱為隱藏層，是因為

該層與外界資訊隔離，網路只透過輸入層與輸出層作為對外的介面，有人稱倒傳遞神經網路為三層結構的神經網路，包含輸入層，隱藏層及輸出層，但實際上輸入層如同單層感知器的輸入層，並不對輸入資訊作特別處理，只是當外界輸入訊息的傳遞介面，將輸入項直接往後傳遞，因此實際有作用的神經元只有隱藏層及輸出層兩層。一般倒傳遞類神經網路的架構，輸入層與輸出層是用來表現輸入項與輸出項，兩者神經元數目依問題的型式而定；隱藏層的層數可以依問題複雜度由一層增加到數層。倒傳遞神經網路輸入和輸出的關係可以看作是一種非線性的映射關係，其廣義化特性就可以看作是一種非線性內插，圖 2.8 表示了一個倒傳遞類神經網路的架構，其中參數定義如下：

1. H_l ：在第 l 層中的神經元數目，當 $l=0$ 時，代表的是輸入層；當 $l=1,2,\dots,L-1$ 時，代表的是隱藏層；當 $l=L$ 時，代表的是輸出層。
2. $x_p^{(l)}(k)$ ：在第 l 層中第 p 個神經元的輸出，其中 $l=0,1,\dots,L$ ， $p=0,1,\dots,H_l$ ，參數 k 代表第 k 個輸入向量。
3. $z_q^{(l+1)}(k)$ ：在第 $l+1$ 層中第 q 個神經元的輸入，其中 $l=0,1,\dots,L-1$ 。
4. $w_{pq}^{(l+1)}$ ：在第 l 層中第 p 個神經元到第 $l+1$ 層中第 q 個神經元的權重值，其中 $l=0,1,\dots,L$ ， $q=0,1,\dots,H_l$ 。
5. $w_{0q}^{(l+1)}$ ：在第 $l+1$ 層中第 q 個神經元的**臨界值**(threshold)。
6. $d_n(k)$ 對第 k 個輸入向量而言，在輸出層第 n 個神經元的目標值，其中 $q=0,1,\dots,H_L$ 。

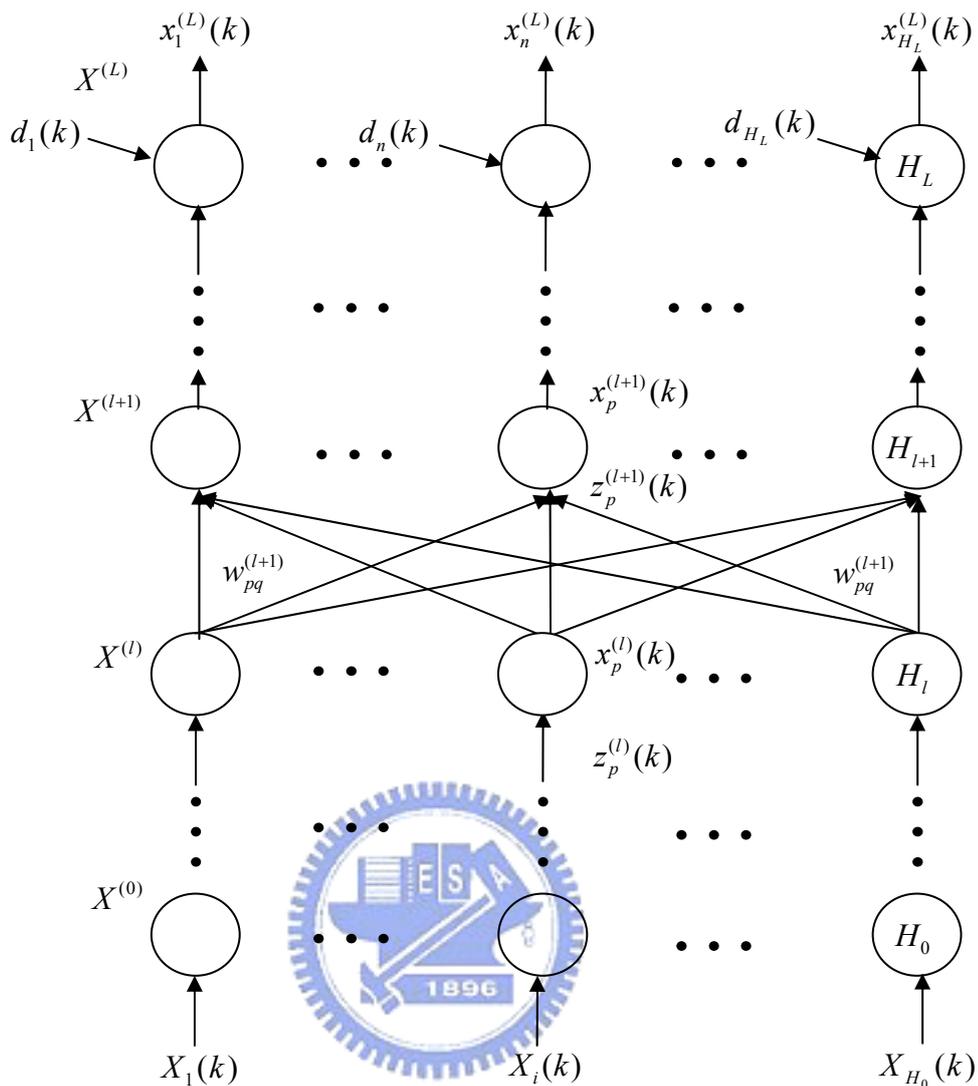


圖 2.7 倒傳遞類神經網路的架構。

網路中靠相關權重連結各層間之神經元，輸入值由輸入層直接傳入隱藏層，經加權累加後再透過作用函數(activation function)運算可得輸出值，同理再傳入輸出層。常使用的作用函數型式如 2.25 式的 S 形函數：

$$f(x) = \frac{1}{1 + \exp(-\alpha x)}. \quad (2.25)$$

由於 S 形函數之活化函數可微分，在學習過程中使用錯誤修正學習法則(error-correlation learning function)修正權重與臨界值，使誤差值逐次減小，達到學習的目的。

倒傳遞演算法的學習過程分為正向傳遞與負向傳遞兩種，在正向傳遞過程中，輸入訊號

從輸入層經由隱藏層加權運算，透過作用函數轉換處理後，再傳向輸出層計算網路輸出值，每一層神經元只影響下一層神經元的狀態，若在輸出層的輸出向量與目標向量差異過大時，則將兩者之間的差異轉為反向傳遞，將誤差訊號沿原來的連接通路回傳，藉此修改各層神經元的權重與臨界值，使得誤差縮小，再重新輸入訊號。如此一直重覆到誤差訊號達到容忍誤差範圍之內而停止。倒傳遞神經網路的多層感知器能透過隱藏層非線性的作用函數來增加神經網路的辨識能力，但是神經網路的學習結果還是會受到許多參數的影響，如隱藏層的神經元個數、隱藏層數目和神經元非線性函數的選擇等等，因此必須藉由試誤的方法來決定適合的參數。

以下就權重與臨界值修改之步驟作說明：

1. 初始值的設定

系統隨機地將初始權重值與臨界值設定為一個很小的數。

2. 決定好每個輸入向量，第 k 個輸入向量：

$$X(k) = [x_1(k), \dots, x_{H_0}(k)]^T, \quad (2.26)$$

其中 $x_i^{(0)}(k) = x_i(k)$ ， $i = 1, \dots, H_0$ 。

3. 順向訊號的傳遞：

$$x_q^{(l+1)}(k) = \text{sgm}(w_{0q}^{(l+1)}(t) + \sum_{p=1}^{H_l} w_{pq}^{(l+1)}(t)x_p^{(l)}(k)), \quad (2.27)$$

其中 $q = 1, 2, \dots, H_{l+1}$ ，和 $l = 0, 1, \dots, L-1$ 。

4. 計算輸出層的 δ 值：

$$\delta_q^{(L)}(k) = (d_q(k) - x_q^{(L)}(k))(1 - x_q^{(L)}(k))x_q^{(L)}(k), \quad (2.28)$$

其中 $q = 1, 2, \dots, H_L$ 。

5. 計算隱藏層的 δ 值：

$$\delta_p^{(l)}(k) = \left(\sum_{q=1}^{H_{l+1}} w_{pq}^{(l+1)}(t)\delta_q^{(l+1)}(k) \right) (1 - x_p^{(l)}(k))x_p^{(l)}(k), \quad (2.29)$$

其中 $p = 1, 2, \dots, H_l$ ， $l = L-1, L-2, \dots, 1$ 。

6. 權重值與臨界值的學習

在 $t+1$ 次疊代之後，權重值變成：

$$w_{pq}^{(l+1)}(t+1) = w_{pq}^{(l+1)}(t) + \alpha \Delta w_{pq}^{(l+1)}(t). \quad (2.30)$$

而臨界值變成：

$$w_{0q}^{(l+1)}(t+1) = w_{0q}^{(l+1)}(t) + \alpha \Delta w_{0q}^{(l+1)}(t), \quad (2.31)$$

其中 α 為學習係數，並滿足 $0 < \alpha < 1$ 。而權重值的變化量為：

$$\Delta w_{pq}^{(l+1)}(t) = \delta_q^{(l+1)}(k) x_p^{(l)}(k). \quad (2.32)$$

臨界值的變化量為：

$$\Delta w_{0q}^{(l+1)}(t) = \delta_q^{(l+1)}(k), \quad (2.33)$$

其中 $p=1,2,\dots,H_l$ ， $q=1,2,\dots,H_{l+1}$ ， $l=L-1,L-2,\dots,0$ 。

7. 回到 2，直到每一個輸入圖樣所產生的誤差，均小於一個事先給定的正實數值 ε ，即：

$$\sum_{n=1}^{H_l} [d_n(k) - x_n^{(L)}(k)]^2 < \varepsilon, \quad \forall k. \quad (2.34)$$

在倒傳遞演算法中，學習係數的選擇，對於網路的性能有著相當大的影響。假若學習係數 α 設得太小，則學習速度將變得非常慢，假若學習係數 α 設得太大，則會引起權重值的振盪。為加速或提升倒傳遞類神經網路之權重調整效率，一種能夠增加學習速度並能減少振盪的方法，就是在原先的學習法則中，加入**動量項**(momentum term):

$$w_{pq}^{(l+1)}(t+1) = w_{pq}^{(l+1)}(t) + \alpha \Delta w_{pq}^{(l+1)}(t) + \beta \Delta w_{pq}^{(l+1)}(t-1), \quad (2.35)$$

其中**動量係數**(momentum coefficient) β 滿足 $0 < \beta < 1$ 。動量項主要的作用，是用來決定前一次的權重值變化量，對於目前權重值移動的方向所造成的影響。

第三章

人臉辨識系統

在有各種複雜背景的相片中自動找出人臉並加以辨識是很困難的，從第一章的文獻探討中可以得知，人臉辨識在很多人臉資訊處理的相關應用中佔有很重要的地位，也有許多學者投入研究在這塊領域上，不管是在人臉特徵的抽取上、或是人臉角度問題的解決、照明環境的補償或是偵測的方法，都已經有許多研究的結果被發表出來。本論文的研究重心是放在從一般的數位相片中辨識出裡面的人臉，目的是想要在數量很多的相片中，尋找出包含指定人物的相片，以彌補相片分類的不足，因此設計了一個人臉辨識系統，在本章將會詳細介紹系統的架構及辨識過程。

在 3.1 節會描述系統流程與架構；在 3.2 節介紹特徵抽取；在 3.3 節介紹人臉偵測；在 3.4 節介紹如何建立人臉資料庫；最後在 3.5 節介紹人臉判別的方法。

3.1 系統概述

圖 3.1 為人臉偵測系統的架構圖，左邊流程圖是系統的分析階段，右邊的流程圖是測試階段。人臉辨識系統需要先從輸入影像樣本來建立人臉資料庫，這個部分為分析階段；對於輸入的測試影像做人臉偵測，選擇待尋人臉並與資料庫做比對，這個部分為測試階段。

在分析階段，第一個部分是特徵抽取，將輸入影像中可能為人臉的區塊轉成特徵向量，以便於接下來的分析；第二個部分是做人臉偵測，使用已訓練好的類神經網路來判斷此區塊是否為人臉；第三個部分是用偵測到的人臉建立人臉資料庫。在測試階段，第一個部分一樣是特徵抽取；第二個部分一樣是人臉偵測；第三個部分則是選擇特定的人臉做搜尋；第四個部分是將選定的人臉與資料庫中的人臉做比對，以找出有包含選定人臉的相片。

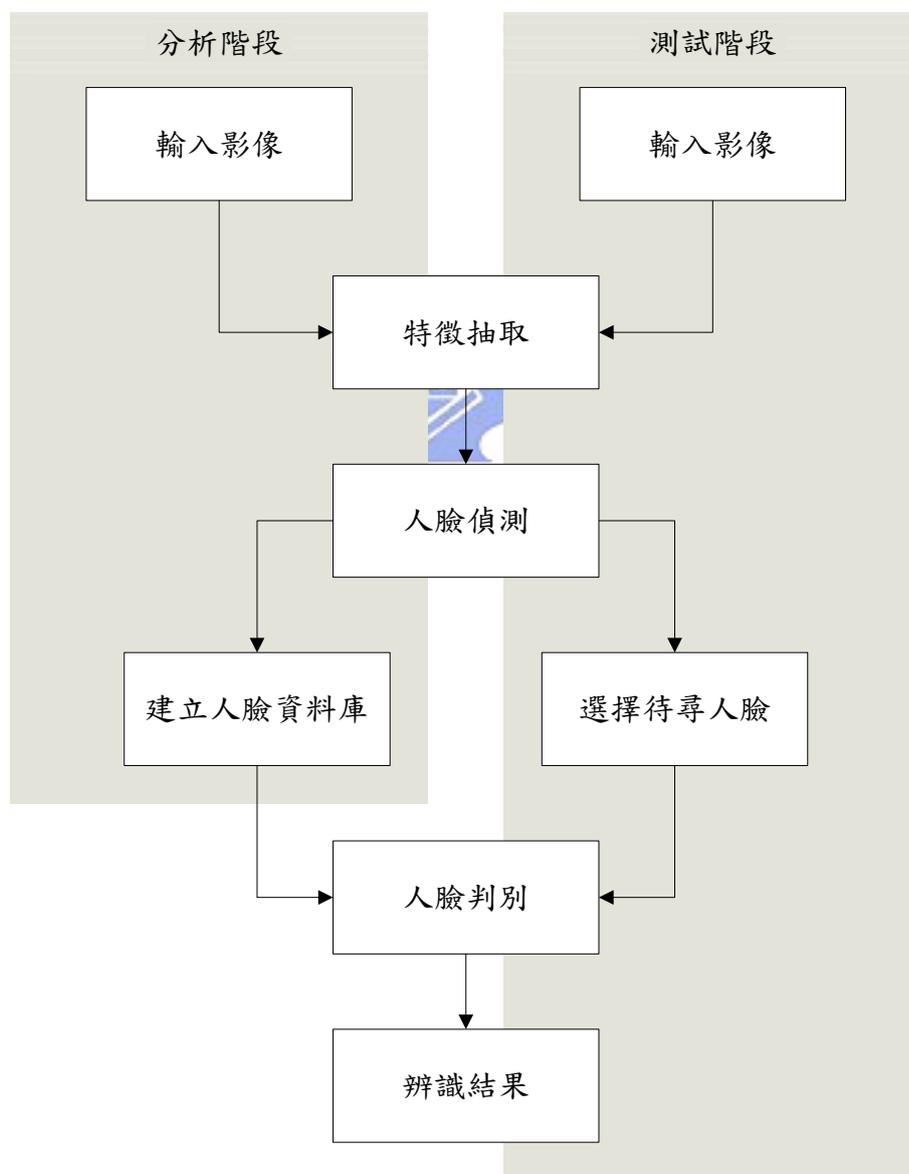


圖 3.1 人臉辨識系統架構圖。

3.2 特徵抽取

圖 3.2 為特徵抽取流程圖，首先將輸入影像做膚色偵測，除去不包含人臉的部份；接下來是多重解析度視窗搜尋(multiresolution window scanning)，將有可能為人臉的部分做各種大小的區塊截圖；然後再將截圖做直方圖等化，以增加影像的對比程度，使得人臉的特徵更為明顯；接著將影像轉換成賈伯小波表示法以產生一個向量；然後再使用區域保留投影對此向量降維，產生影像特徵向量以便於接下來的人臉分析。

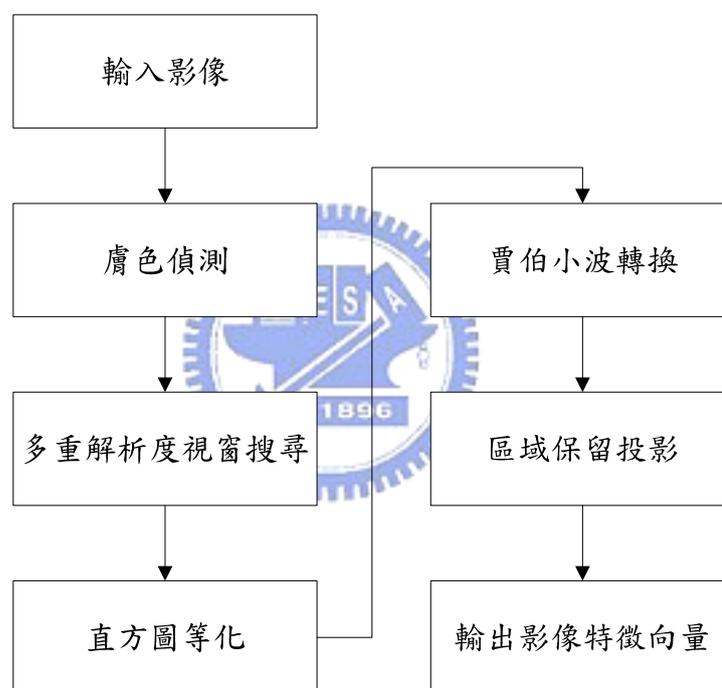


圖 3.2 特徵抽取流程圖。

3.2.1 膚色偵測

首先介紹特徵抽取的第一個步驟，也就是膚色偵測。為了減少在相片中搜尋人臉的時間，先對輸入影像做膚色偵測，以除去不包含人臉的部份，可以有效的增進系統搜尋的效率。在第二章已經介紹過好幾種膚色偵測所選用的色彩空間以及膚色分類的模型，本論文的選用色彩空間是正規化後的 RGB 色彩空間，因為膚色分佈比較集中；使用膚

色分類的模型是橢圓邊界模型[17]，從各種不同人種及不同照明環境的相片中取得了一共三萬筆的膚色資料，用來計算(2.6)式的參數 Ψ 和 λ 。因此利用(2.5)式就可以得到一個橢圓的模型，然後再設定 $\Phi(c) < \theta$ 中的 θ 值就可以用來分類相片中的膚色與非膚色。

為了不讓是人臉的區域因為判斷錯誤使得從接下來的分析中被忽略，因此 θ 值會設的較大一點，雖然會增加判斷錯誤的區域，但是這裡做膚色偵測的目的只是為了減少搜尋人臉的範圍。膚色偵測後的結果會是一個二值影像，把它套回原始圖片上的結果如圖 3.3(b)所示。

做完膚色偵測後會發現結果有許多雜點，所以接下來要濾除這些雜點，使得包含人臉的區塊儘量沒有雜點，讓接下偵測人臉不會受到影響；而不含人臉的區塊則要儘量縮小，就不會浪費時間在這些區塊，而能使整個搜尋的速度變快。在這裡用來濾除雜點的方法是形態學中的膨脹與侵蝕，使用圓形的結構元素，分別經過幾次膨脹與侵蝕後的運算，可以有效的濾除大部份的雜訊，而且可以使膚色偵測的邊緣變得圓滑，便於接下來尋找人臉，結果如圖 3.3(c)所示。



3.2.2 多重解析度視窗搜尋

把相片中的膚色與非膚色區域分出來之後，接下來要對可能為人臉的區域做分析。因為本論文用的賈伯濾波器是針對灰階影像做處理，所以必須先把彩色影像轉換成灰階，然後再根據膚色偵測的結果，在可能是人臉的部分做搜尋。為了能夠偵測到不同大小的人臉，這裡使用多重解析度來搜尋人臉，也就是比例相同但是尺寸大小不同的人臉，在每個尺寸的人臉都搜尋一次。由於賈伯小波對於區塊定位和對齊誤差有相當不錯的容忍度[22]，所以可以用規則分布的取樣點來取代對於每個像素都要做一次搜尋，這樣可以增進系統的執行效率，減少執行所花費的時間。

為了決定比較適合的規則分布取樣點，本論文採用中國古代畫繪畫術語中的「三庭五眼」的觀念來決定取樣點。所謂的「三庭五眼」，是把臉分成三等分的長度及五等分的寬度，如圖 3.4 所示：「三庭」指的是前額髮際線至眉骨、眉骨至鼻底及鼻底至下頰的



(a)



(b)



(c)

圖 3.3 膚色偵測的結果，(a)原始相片，(b)經過膚色偵測後的結果，(c)再經過消除雜訊後的結果。

長度大致相等；「五眼」指的是兩眼之間寬度為一眼及眼外側至髮際間的寬度亦為一眼。

取樣點之間的距離是依照人臉的尺寸而定，寬度為人臉的十分之一，高度為人臉的六分之一。根據實驗結果，這樣不但可以大幅減少執行搜尋的時間，同時也能夠保留足夠分析的資訊。

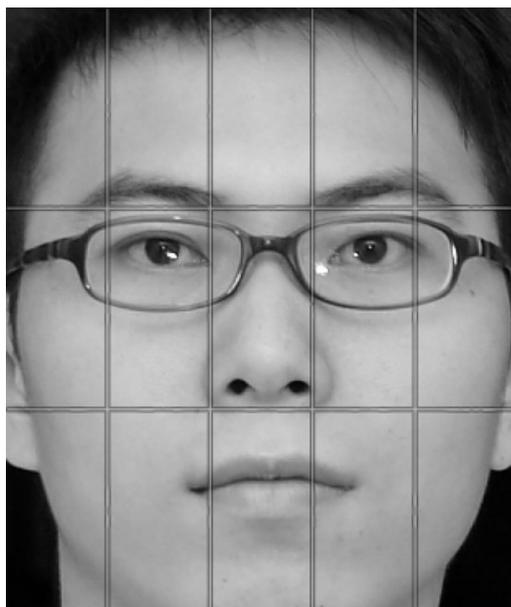


圖 3.4 三庭五眼示意圖。

3.2.3 直方圖等化

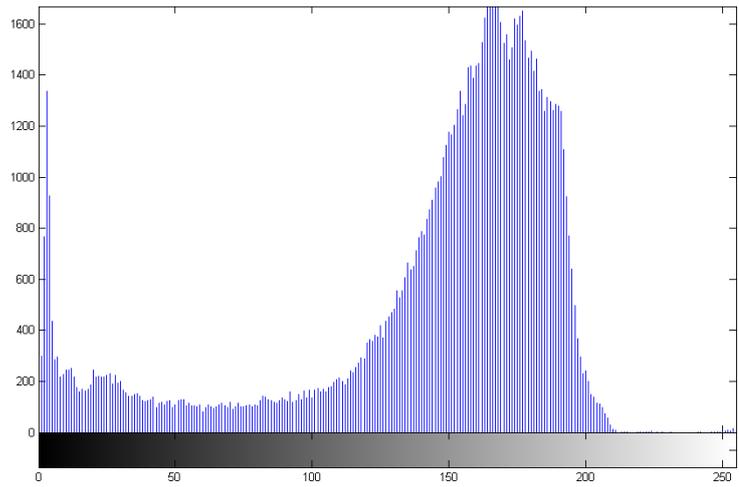
直方圖等化的目的是增強影像的對比，讓臉上的特徵更明顯，可以在一定的程度上減少照明變化的影響。本論文是對灰階影像做直方圖等化，其理論在 2.3 節已經提過，圖 3.5 是一張人臉影像做完直方圖等化後的結果及其直方圖，可以看出做完直方圖等化後的人臉，臉部的特徵變得更為清楚，更能將人臉的特徵表現出來，因而有利於輸入類神經網路做訓練或是分析。

3.2.4 賈伯小波轉換

經過直方圖等化後的影像，如果直接輸入到類神經網路的話，並沒有什麼意義，所以要先經過賈伯小波轉換，使人臉的各部位特徵變得明顯。相關的理論在 2.4 節已經提過，本論文使用三種尺寸($k_{max} = \frac{5\pi}{4}, f = \sqrt{2}, s \in 0, \dots, 4$)、八個方向($d \in 0, \dots, 7$)，並設定 $\sigma = 2\pi$ 的賈伯濾波器來得到人臉影像的特徵。將寬為 29 像素、高為 35 像素的影像經過 24 種賈伯濾波器，會得到總共 24360 個維度的向量，這樣的維度過於龐大，難以做分析，所以先要經過降維，降維的方法是使用區域保留投影，將在下個小節中做介紹。



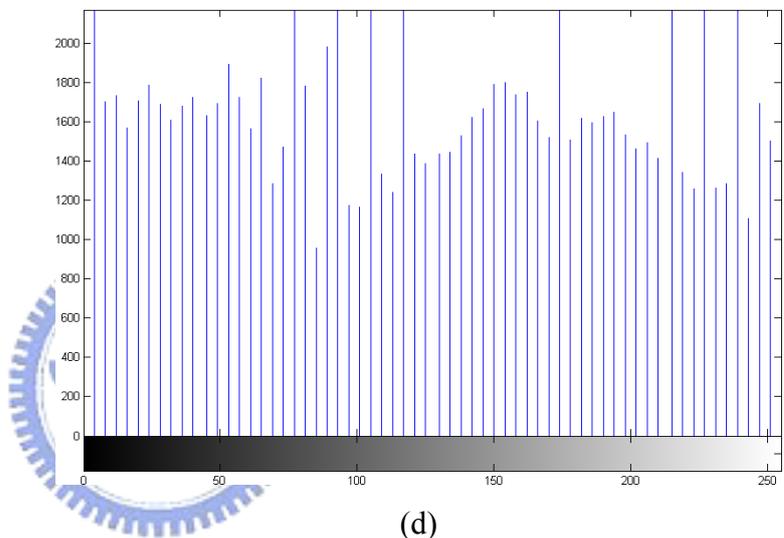
(a)



(b)



(c)



(d)

圖 3.5 經過直方圖等化後的結果，(a)原始影像，(b)原始影像的直方圖，(c)經過直方圖等化後的影像，(d)直方圖等化後的直方圖。

3.2.5 區域保留投影

區域保留投影的主要目的是將輸入資料做降維，以便於輸入到類神經網路中做分析。其理論與優點在 2.5 節已經提過，本論文使用 k-相近來建構出相鄰無向圖，因為在實際應用中，要定義出最好的 ϵ 值比較困難；權重的選擇則是使用熱核心。寬為 29 像素、高為 35 像素的影像就有 1015 個維度，再經過 24 種賈伯濾波器會得到總共 24360 個維度的向量，過於龐大所以要降維。由於影像在經過每種賈伯濾波器後都會變成比較

簡單的圖形，因此本論文對於 24 種賈伯濾波器分別做 24 次的區域保留投影，每次只針對一種賈伯濾波器做降維。這樣就可以保留賈伯小波的特性，同時又能夠將維度降低到很小。

3.3 人臉偵測

本論文是使用類神經網路來做人臉偵測，類神經網路需要先經過訓練的步驟，利用人臉與非人臉的訓練樣本做學習，才能夠用來判斷人臉，以下將介紹類神經網路的訓練過程、本論文使用的人臉資料庫以及人臉偵測的方法。

3.3.1 訓練類神經網路

圖 3.6 為訓練類神經網路流程圖，輸入的訓練影像為灰階影像，先將其做直方圖等化以增加影像的對比程度，使得人臉的特徵更為明顯；接著將影像轉換成賈伯小波表示法以產生一個向量；然後再使用區域保留投影對此向量降維，產生影像特徵向量再送入類神經網路做訓練。

本論文使用的訓練影像的資料庫出自卡內基美隆大學(Carnegie Mellon university, CMU)的機器人學會(robotics institute)所建立的姿勢、照明及表情(pose, illumination, and expression, PIE)資料庫[32][33]，總共包含了有 70 個人、42490 張照片，並分成 13 種不同的姿勢(也就是人臉角度)、43 種不同的照明環境及 3 種不同的表情。圖 3.7 為架設環境的立體空間圖及 13 種不同人臉角度的說明，出自[32]。

圖 3.8 是本論文所使用的倒傳遞神經網路模型， $W(1,1)$ 為 100×120 的輸入權重矩陣，它的輸入為 120×1 的向量， $b(1)$ 為 100×1 的偏權值向量。中間隱藏層的函數為正切雙彎曲函數(3.4)式，其輸入為 100×1 的向量。

$$\frac{e^n - e^{-n}}{e^n + e^{-n}}, \text{ 其中 } n=1,2,\dots,N. \quad (3.1)$$

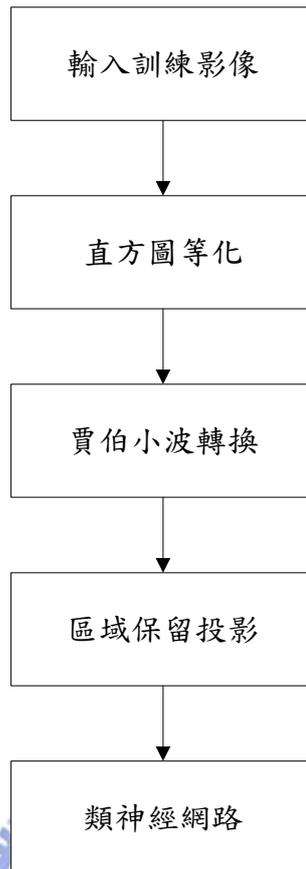


圖 3.6 訓練類神經網路流程圖。



圖 3.7 PIE 資料庫中 13 種不同人臉角度的說明圖。[32]

$LW(2,1)$ 為 100×120 的層權重矩陣，輸入為 100×1 的向量， $b(2)$ 為 120×1 的偏權值向量。

最後的輸出層函數為對數雙彎曲函數(3.2)式，其輸入為 240×1 的向量。

$$\frac{1}{1+e^{-n}}, \text{ 其中 } n=1,2,\dots,N. \quad (3.2)$$

要使得類神經網路能正確的運作，則必須透過訓練的方式，讓類神經網路反覆的學習，直到對於每個輸入都能正確對應到所需要的輸出，因此在類神經網路訓練前，需準備訓練所需的樣本，樣本越多、越正確、差異性越大，類神經網路的能力就越強。

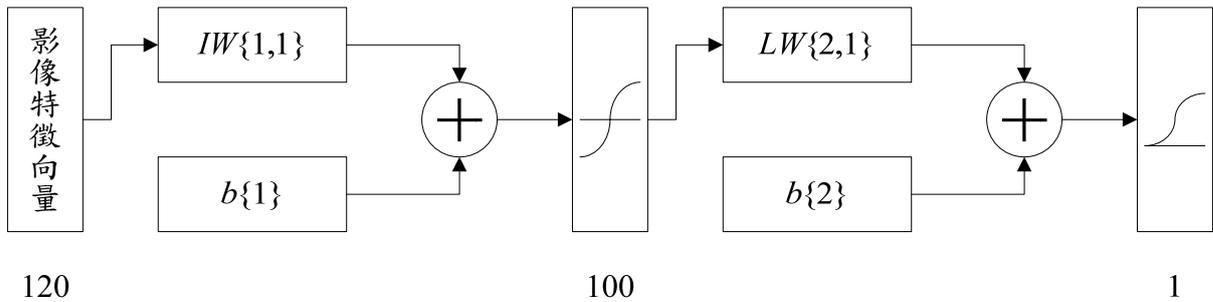


圖 3.8 倒傳遞神經網路模型。

在本論文中，訓練樣本選用了 PIE 資料庫中的 30 個人、5 種不同的姿勢及 8 種照明環境，總共 1200 張的人臉，再加上 2400 張非人臉樣本，總共 3600 張影像樣本來做訓練。對於人臉樣本的目标值設為 1；非人臉樣本的目标值設為-1。根據實驗結果，人臉與非人臉樣本的比例為一比二時，訓練的效果最好。圖 3.9 是網路在訓練過程中的均方誤差(mean square error)，設定目標的均方誤差為 10^{-4} ，是足夠區分人臉與非人臉的差異，從圖中可以看出最後的結果確實到達了 10^{-4} 。訓練完後的神經網路會有最後的權重值。

3.3.2 偵測人臉

輸入影像在經過膚色偵測後會得到可能包含人臉的區域，然後對這些區域中的每個像素點做多重解析度區塊截圖，把每張不同尺寸的截圖經過直方圖等化，轉成賈伯小波表示法，再經由區域保留投影降維，得到影像特徵向量，然後輸入到已訓練好的類神經網路，會得到一個輸出。這個輸出如果越接近訓練時設定的人臉目标值時，就代表是人臉的可能性越高；反之，越接近訓練時設定的非人臉目标值時，就代表不是人臉的可能性越高。

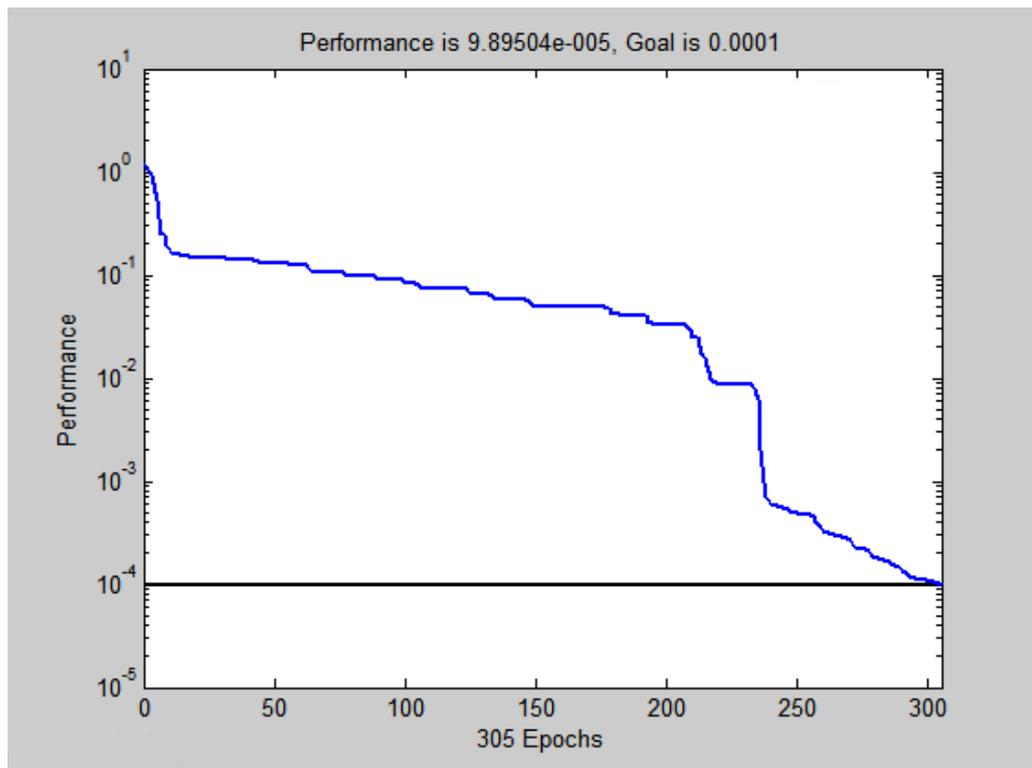


圖 3.9 訓練類神經網路過程的均方誤差。

能否找到最適合的人臉大小，跟之後人臉辨識的準確度有很大的關係，人臉的大小範圍與位置如果偵測的越準確，取得的人臉特徵也就越準確，人臉背景雜訊所產生的干擾也會變小。要如何找到最適合的人臉大小，本論文所使用的判斷方法敘述如下：對可能為人臉的區域中某一點做人臉偵測，會得到相對應的類神經網路輸出值，將此值與訓練類神經網路時所設定的人臉目標值相減再取絕對值後，如果越接近零，就代表人臉的可能性越高。如果在同一點上做不同尺寸的截圖，這個尺寸的截圖比前後大小尺寸的截圖都還要接近零，也就是越可能為人臉時，就代表這張截圖是最適合的人臉大小。

經過類神經網路的計算之後，對於都有人臉的截圖，以與人臉目標值的差為最小的代替，然後記錄這個位置座標、人臉尺寸及差值。等到全部的區域都搜尋完畢後，將有人臉的全部資料點依照其差值降冪排序，然後依照截圖的位置座標及尺寸大小，把在人臉範圍內其他差值較大且的點判斷為同一張人臉。

3.4 建立人臉資料庫

在分析階段，將欲分析的相片輸入，偵測出其中人臉之後，接下來就是建立人臉資料庫。首先由人工指定偵測出的人臉為 a 這個人，人臉影像的維度一樣是寬為 29 像素、高為 35 像素的影像，總共有 1015 個維度。假設同樣為 a 的人臉一共有 n 張，其人臉資料的特徵矩陣為 P_a ，為一個有 $1015 \times n$ 的矩陣。將此矩陣分別經過 24 種賈伯濾波器，並且每次經過一種賈伯濾波器都做一次區域保留投影，降維到 l 個維度 ($l \ll 1015$)，最後會得到一組特徵向量的矩陣 E_a ，每個特徵向量都是 $1015 \times l$ 的矩陣，這樣原始的人臉特徵矩陣，維度為 $24360 \times n$ 就被降維到 $l \times n$ 。然後將全部的人臉特徵矩陣都與 E_a 相乘做降維，再將這些人臉特徵矩陣輸入到類神經網路中做訓練，所使用的類神經網路一樣是倒傳遞類神經網路，與 3.3.1 節相同。對於是 a 的人臉目標值設為 1；不是 a 的人臉目標值設為 -1，將全部的人臉都訓練完之後就得到一個可以用來判斷人物的類神經網路，圖 3.10 為建立人臉資料庫的流程圖。

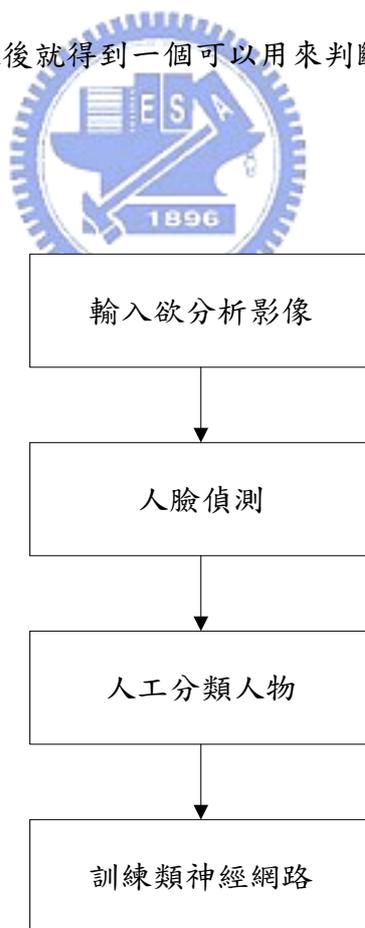


圖 3.10 建立人臉資料庫流程圖。

3.5 人臉判別

在測試階段，將測試影像輸入，做完人臉偵測，然後選擇好待尋的人臉之後，接下來就是辨識出這個人是誰。將待尋人臉先經過特徵抽取的步驟轉成影像特徵向量，然後再輸入到在建立人臉資料庫時，所訓練好的類神經網路，會得出一組輸出結果，這組輸出結果的值就是對應於人臉資料庫中各人臉的相似程度。如果越接近訓練時某人物的人臉目標值，就代表越有可能是此人物，依照相似度的高低來排列，就可以判別待尋人臉為何人。



第四章

實驗結果與分析

在本章中會討論本論文所提出的人臉辨識系統對於測試影像的人臉偵測率、人臉辨識率與系統效能，在 4.1 節會說明用來測試系統效能所採用的實驗機制；在 4.2 節會展示實驗結果及與其他方法的比較；在 4.3 節會對實驗結果做討論。



4.1 實驗機制

在 4.1.1 節中會介紹實驗用的硬體架構以及程式處理效能速度；在 4.1.2 節中會定義用來評估系統效能的方式。

4.1.1 實驗設備

本研究實驗用的設備是中央處理器為 AMD Athlon 64 3000+ 2.0GHz 而且記憶體為 1GB 的個人桌上型電腦，使用 Matlab 2007a 的軟體來編寫程式以實現人臉辨識系統。在人臉偵測的部分，類神經網路訓練時所使用的訓練資料庫在 3.4 節已經介紹過，是卡內基美隆大學的 PIE 資料庫[28][29]，從其中選用了總共包含了有 30 個人、5 種不同的姿勢及 8 種照明環境，總共 1200 張的人臉影像樣本，再加上 2400 張非人臉影像樣本，總共 3600 張影像樣本。在人臉辨識的部分，建立人臉資料庫與測試階段中所使用的測試

相片則包含自己拍攝的以及同學與朋友所提供的照片。

4.1.2 效能評估

人臉辨識系統的主要目的是尋找輸入相片裡是否有人臉存在，所以在系統的效能評估方面分為人臉偵測與人臉辨識的結果兩種。根據人臉偵測的可能情形，可以將結果分為以下四種，如表 4.1 所示：

1. **正確肯定**(true positives, TP)：意思是所偵測到的人臉全都是正確的，也就是成功的偵測到人臉。
2. **正確否定**(true negatives, TN)：意思是沒有把非人臉誤認為人臉的情形，也就是沒有誤判的非人臉。
3. **錯誤肯定**(false positives, FP)：意思是沒有出現人臉卻把某個區域認定為人臉，也就是偵測錯誤的人臉。
4. **錯誤否定**(false negatives, FN)：意思是出現了人臉但沒有偵測到，也就是未被偵測到的人臉。



表 4.1 人臉偵測的可能情形。

		實際情形	
		人臉	非人臉
偵測情形	人臉	TP	FP
	非人臉	FN	TN

有了這四種的偵測可能結果後，接下來定義人臉偵測的正確率、錯誤率及誤判率如下：

$$\text{正確率} = \frac{TP}{TP + FN} \times 100\%, \quad (4.1)$$

$$\text{錯誤率} = \frac{FP}{TP + FP} \times 100\% . \quad (4.2)$$

$$\text{誤判率} = \frac{FP + FN}{TP + FP + FN} \times 100\% . \quad (4.3)$$

當正確率達到 100%時，代表所有出現的人臉都被偵測到；當錯誤率為 0%時，代表所有偵測到的都是人臉，也就是沒有非人臉卻被認為人臉的區域；而誤判率則是正確率和錯誤率兩個的綜合指標，可以直接看出系統的好壞，越低的誤判率代表系統越好，一個理想的人臉偵測必須是誤判率為 0%。

至於人臉辨識的可能情況就比較簡單，只有辨識成功率，定義如下：

$$\text{正確率} = \frac{\text{辨識正確的人臉數目}}{\text{辨識正確人臉的數目} + \text{辨識錯誤的人臉數目}} \times 100\% . \quad (4.3)$$

同樣的，當正確率達到 100%時，代表所有出現的人臉都辨識正確，一個理想的人臉辨識必須是正確率為 100%。

4.2 實驗結果與分析比較

在本節中將展示本論文的實驗結果，分為人臉偵測與人臉辨識兩部分。在 4.2.1 節展示人臉偵測的實驗結果；在 4.2.2 節展示人臉辨識的實驗結果。

4.2.1 人臉偵測的實驗結果與分析比較

在人臉偵測的實驗部分，測試相片是從自己拍攝及同學與朋友所提供的照片中隨機挑選出 100 張來做測試，包含了各種場景及不同的人臉角度與大小，實際的人臉總數為 764 張人臉。當系統偵測到人臉的時候，會用一個方框將人臉框出，根據人臉大小及角度的不同，會自動選擇不同的方框大小。

圖 4.1 為人臉偵測的部分實驗結果。圖 4.1(a)是大多數為正面人臉偵測結果圖。相片中實際人臉總數為 83 張、系統偵測到正確的人臉為 79 張、未被偵測到的人臉為 4 張、偵測到錯誤的人臉為 1 張。因此可以看出當人臉特徵是清楚的時候，就算人臉很小也可以有不錯的偵測效果。圖 4.1(b)則是臉部特徵不明顯的正面人臉偵測結果圖。相片中實際人臉總數為 84 張、系統偵測到正確的人臉為 71 張、未被偵測到的人臉為 13 張、偵

測到錯誤的人臉為 2 張。因此在人臉特徵因為照明關係不太清楚的時候，系統的偵測效能就會受到影響。圖 4.1(c)則是有遮擋的正面人臉偵測結果圖。相片中實際人臉總數為 67 張、系統偵測到正確的人臉為 50 張、未被偵測到的人臉為 17 張、偵測到錯誤的人臉為 1 張。因此在人臉特徵因為受到遮擋變得不清楚的時候，會使系統的偵測效能下降。圖 4.1(d)則是各種角度的人臉偵測結果圖。相片中實際人臉總數為 20 張、系統偵測到正確的人臉為 15 張、未被偵測到的人臉為 5 張、偵測到錯誤的人臉為 0 張。因此在各種角度的人臉下，由於人臉特徵並不完全，所以系統的偵測效能也會受到影響而下降。

接下來將本論文的方法和兩篇相關的論文做比較，這兩篇論文分別是 Huang 等人[34]所提出的方法與 Kadoury[35]等人所提出的方法。Huang 的方法是使用四種賈伯濾波器得到人臉的特徵，再用主成份分析做降維，最後用多項式類神經網路來判別人臉；而 Kadoury 的方法是將二維影像用局部線性嵌入來降維，然後使用支持向量機來判別人臉。選擇這兩個方法來比較是因為 Huang 的方法與本論文都是使用賈伯濾波器來得到人臉的特徵，但 Huang 使用主成份分析來降維而不是本論文使用的區域保留投影；Kadoury 的方法則是直接將人臉影像經過局部線性嵌入來降維，不像本論文是使用賈伯濾波器來得到人臉的特徵。用這兩種方法對本論文所使用的測試相片做人臉偵測的結果如表 4.2 所示。

從表 4.2 中的結果可以看出本論文的方法是較優於另外兩種方法，由於在本論文中，人臉偵測的目的是為了接下來做人臉辨識，因此在設計上是以儘量提高 TP 及降低 FN 為目標，也就是儘量能夠偵測到最多的人臉，因此 FP 的部分會比較高。Huang 的方法雖然也是用賈伯濾波器來得到人臉特徵，但是只用了四種，不像本論文用了 24 種，且 Huang 是使用主成份分析做降維，對於人臉這種局部為非線性的資料就比較不能好好保留原始的資料結構，所以在正確率與錯誤率上都會比較低；Kadoury 的方法是直接將人臉影像經過局部線性嵌入做降維，雖然能保留人臉資訊，但是在複雜的相片背景中效果就會比較差，所以在誤判率上會是最高的。



(a)



(b)

圖 4.1 人臉偵測的部分實驗結果，(a)大多數為正面的人臉，(b)臉部特徵不明顯的正面人臉，(c)有遮擋的正面人臉，(d)各種角度的人臉。



(c)



(d)

圖 4.1(續) 人臉偵測的部分實驗結果。

表 4.2 三種人臉偵測方法的實驗結果比較。

人臉總數：764	TP	FN	FP	正確率	錯誤率	誤判率
本論文的結果	676	88	89	88.48%	11.63%	20.75%
Huang 的結果	641	123	61	83.90%	8.69%	22.30%
Kadoury 的結果	658	106	84	86.13%	11.32%	22.41%

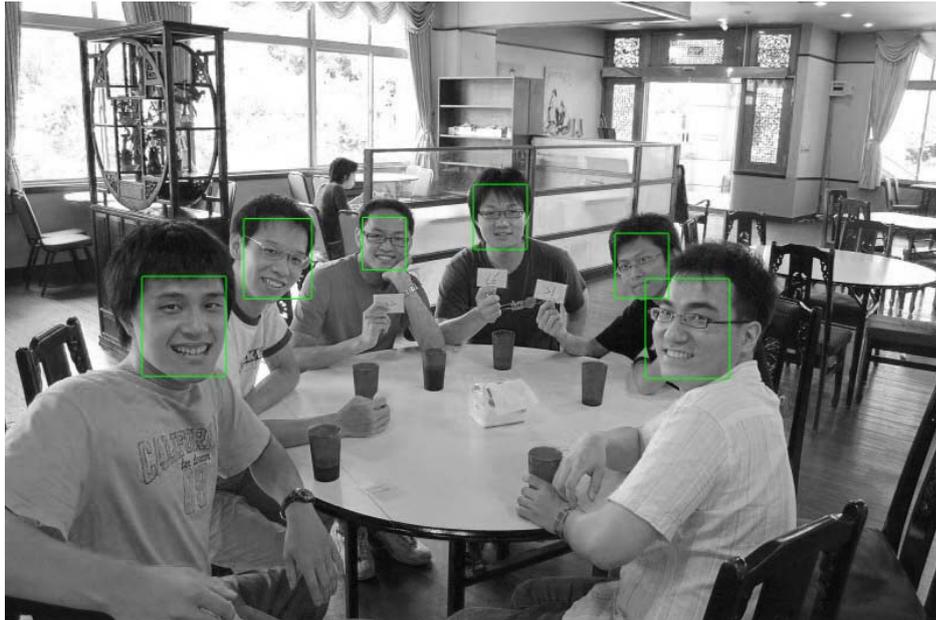
4.2.2 人臉辨識的實驗結果與分析比較

在人臉辨識的實驗結果部分，測試相片是從自己拍攝及同學與朋友所提供的照片中選出 20 個人物，總共 96 張照片來做測試，包含了各種場景及不同人臉角度與大小的照片，一共有 402 張人臉。圖 4.2 是人臉辨識實驗的部分結果。圖 4.2(a)是測試照片做完人臉偵測後的結果，將這些截下來的人臉送入類神經網路做辨識，辨識後的結果對照在人臉資料庫中的人物，將其由左至右依序展示在圖 4.2(b)。

人臉辨識的實驗結果分為兩個部分，第一部分是測試訓練樣本數多寡下的辨識率，隨機抽取不同數量的訓練樣本，這樣可以模擬實際應用情況下，可能會發生訓練樣本數不足的情況，測試結果如表 4.3 所示。第二部分是測試各種維度下的辨識率，這樣可以找出區域保留投影所降低的維度與人臉辨識率之間的關係，由於本論文的方法是對 24 種賈伯小波做降維，因此降低後的維度會以 24 為基數增加來測試辨識率，測試結果如表 4.4 所示。

從表 4.3 中可看出，訓練樣本數越多，辨識率也會隨之增加，但又不是絕對，這是因為隨機選擇訓練樣本並不一定能找到足夠代表性的樣本，也就是說，訓練樣本是否具有代表性才是影響辨識率的關鍵。

從表 4.4 中可看出，維度越高，辨識率同樣的也會增加，但是增加的幅度會越來越小。當維度到達 192 維時有最佳的辨識率與維度的平衡，也就是對於每一個賈伯濾波器，使用區域保留投影做降維，只需要 8 個維度就足以表示了。



(a)



(b)

圖 4.2 人臉辨識實驗的部分結果，(a)人臉偵測後的結果，(b)人臉辨識的結果，由左至右取出資料庫的人臉做展示。

接下來將本論文的方法和兩篇相關的論文做比較，這兩篇論文分別是 Wang 等人[36]所提出的方法與 You[37]等人所提出的方法。Wang 的方法是使用二維賈伯小波視窗的方法去抽取局部特徵向量，用這個向量來表示人臉並且做辨識；而 You 的方法是屬於流形學習中的方法，稱之為鄰域識別投影(neighborhood discriminant projection, NDP)，此方法是從局部線性嵌入的方法修改而來。主要思想是保留分類內部的幾何關係，透過不同類別之間的差異而有識別的能力。選擇這兩個方法來比較是因為 Wang 的方法與本論文都是使用賈伯濾波器來得到人臉的特徵，但 Wang 是使用局部的人臉特徵來辨識人臉，不像本論文使用的是人臉的整體特徵；You 的方法則是直接將人臉影像經過鄰域識別投影來學習，不像本論文是使用賈伯濾波器來得到人臉的特徵。用這兩種方法對本論文所使用的測試相片分別對不同訓練樣本數與不同維度的情形做人臉辨識，結果如圖 4.3 與圖 4.4 所示。

表 4.3 不同訓練樣本數的人臉辨識實驗結果。

訓練樣本數	辨識正確	辨識錯誤	辨識率
2	47	355	11.69%
4	128	274	31.84%
6	182	220	45.27%
8	269	133	66.92%
10	334	68	83.08%
12	361	41	89.80%
14	376	26	93.53%
16	373	29	92.79%
18	375	27	93.28%
20	368	34	91.54%
22	373	29	92.79%
24	369	33	91.79%
26	371	31	92.29%
28	374	28	93.03%
30	367	35	91.29%

由圖 4.3 中同樣可看出本論文的方法是較優於另外兩種方法，Wang 的方法與本論文相近，但是效果仍然沒有像本論文的方法這麼好，雖然特徵與本論文所使用的很類似，但由於 Wang 使用局部的特徵，因此在辨識上對於不同表情與角度的變化，就沒有那麼好的容忍度；另外從圖上可以很明顯的看出來，You 的方法是需要比較多的訓練樣本才能夠達到足夠的辨識率。

表 4.4 不同維度下的人臉辨識實驗結果。

維度	辨識正確	辨識錯誤	辨識率
24	140	262	34.83%
48	217	185	53.98%
72	289	113	71.89%
96	325	77	80.85%
120	341	61	84.83%
144	353	49	87.81%
168	361	41	89.80%
192	366	36	91.04%
216	369	33	91.79%
240	371	31	92.29%
264	372	30	92.54%
288	373	29	92.79%
312	373	29	92.79%
336	374	28	93.03%
360	374	28	93.03%
384	374	28	93.03%
408	375	27	93.28%
432	375	27	93.28%
456	375	27	93.28%
480	376	26	93.53%

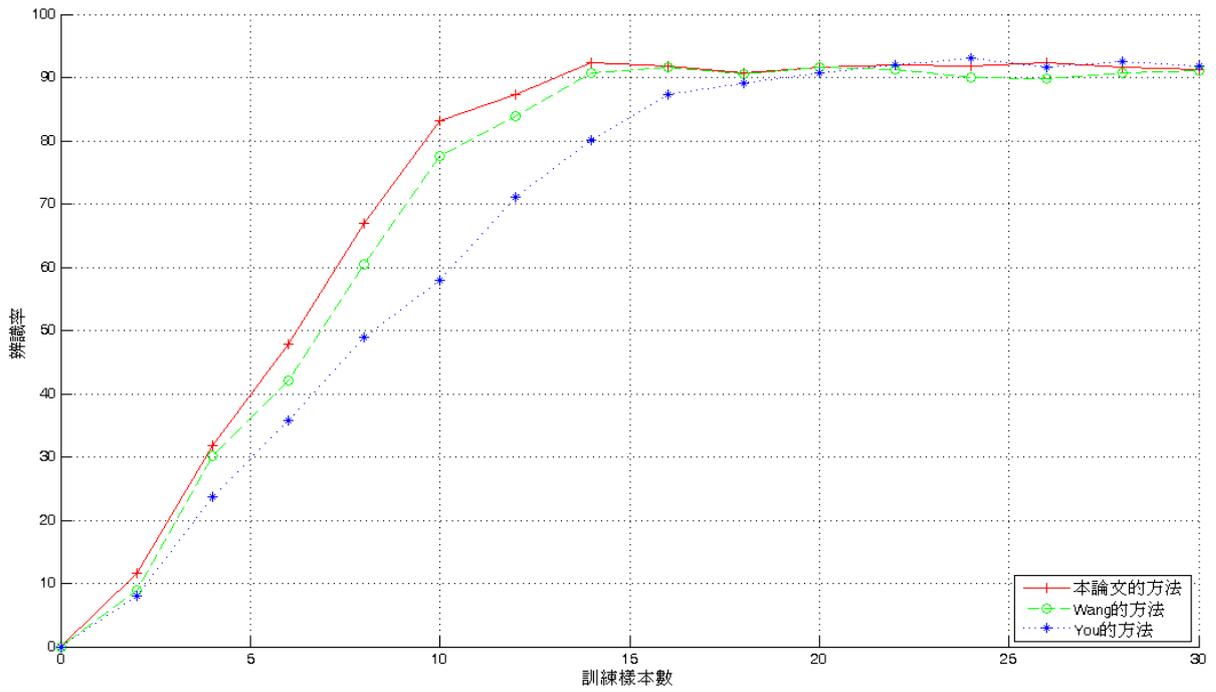


圖 4.3 不同訓練樣本數的人臉辨識實驗結果比較。

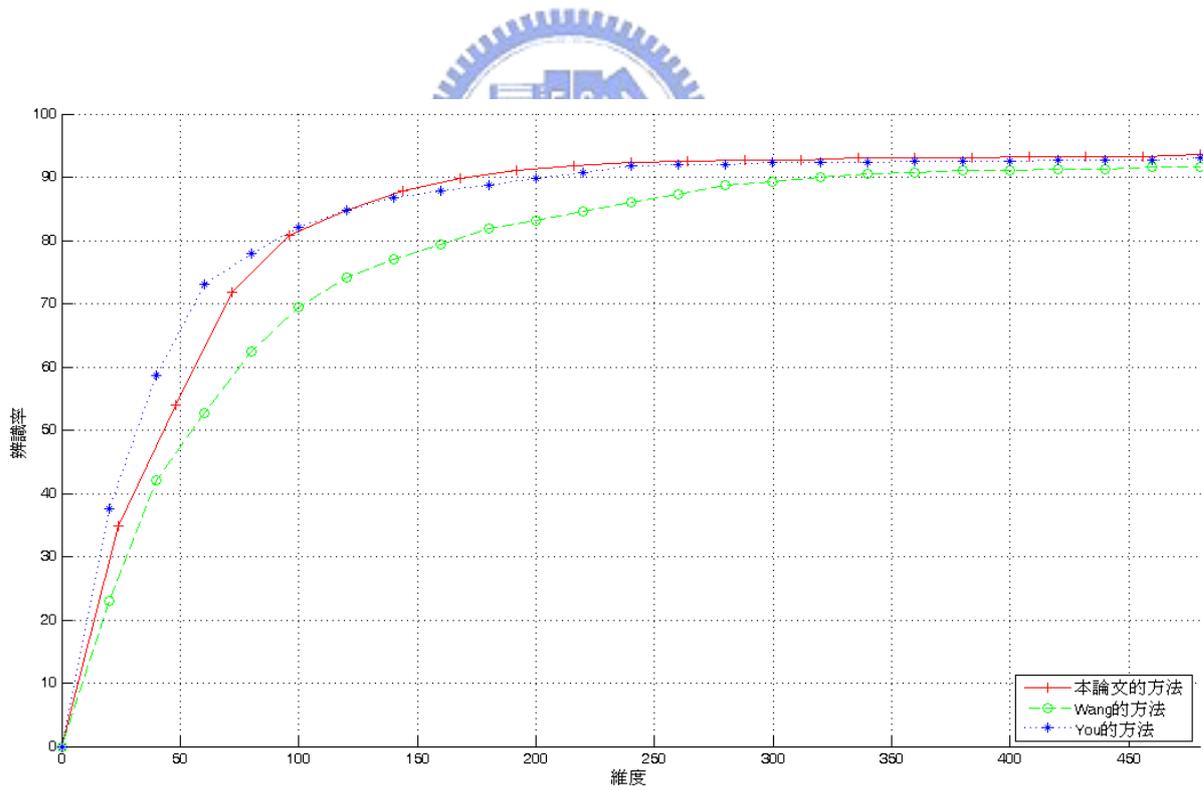


圖 4.4 不同維度下的人臉辨識實驗結果比較。

由圖 4.4 中可看出本論文的方法是較優於 Wang 的方法，但是比 You 的方法稍差。Wang 的方法雖然也是用賈伯濾波器來得到人臉特徵，但 Wang 是使用局部的人臉特徵與主成份分析做降維來辨識人臉，不像本論文使用的是人臉的整體特徵與區域保留投影；雖然 Wang 的方法是將人臉局部特徵分開做主成分分析來降維，但是效果仍然沒有像本論文的方法這麼好；You 的方法的最大優點是只需要比較少的維度就能有不錯的辨識率，但是需要較多的訓練樣本才能達到效能，因此在訓練樣本不足時，辨識率反而會較低。

4.3 結果討論

本論文提出的人臉偵測與人臉辨識的方法，是針對日常生活中隨意拍攝所得到的相片，在對於相片的來源及數量都沒辦法掌握的情況下，已經具有基本的偵測成功率及辨識率。在相片來源的方面，由於偵測成功率的高低在進一步的人臉辨識上有很大的影響，從以上的實驗結果中可得知，本論文的方法在人臉特徵明顯時有較好的偵測成功率，同時也會有較好的辨識率。在相片數量的方面，想要有較好的辨識率，就要有足夠的訓練影像樣本才能夠達到，但是由實驗結果可得知，是否有足夠的代表性訓練樣本才是能否提高辨識率的關鍵。以下分為人臉偵測與人臉辨識兩個部分做結果討論與分析。

4.3.1 人臉偵測的結果討論

在人臉偵測的部分，針對本論文系統的人臉偵測結果做整理，根據影響人臉特徵的因素，可分為照明、角度與遮擋三種，表 4.5 列出了在這三種因素組合下的可能情形。情形(a)指的是在光線足夠環境且人臉上沒有陰影，人臉角度為正面對鏡頭的人臉，沒有受到任何遮擋的情形，此時人臉的特徵最為清楚，因此正確率也達到了 100%；情形(b)為人臉僅單純受到照明影響，人臉角度為正面，沒有受到任何遮擋的情形，在這種情形下人臉的特徵有部分受到照明影響而變的不清楚，正確率為次高的 96.05%；情形(c)為人臉僅有角度變化，沒有受到照明與遮擋影響的情形，這種情形下人臉的特徵會因為某

些角度而消失，正確率為 93.86%；情形(d)為人臉僅有受到遮擋的影響，沒有受到照明影響，人臉角度為正面的情形，從表中可看出，在這幾種影響人臉特徵的因素中，遮擋影響的程度最大，不但會使得人臉特徵消失，同時還可能被誤判為非人臉，因此正確率只有 87.23%；接下來情形(e)為同時包含情形(b)及情形(c)，也就是同時有照明及角度影響，但是沒有遮擋的情形，這時候的正確率是 77.93%；情形(f)為同時包含情形(c)及情形(d)，也就是同時有角度及遮擋影響，但是沒有照明影響的情形，這時候的正確率是 55.32%；情形(g)為同時包含情形(b)及情形(d)，也就是同時有照明及遮擋影響，但是人臉角度為正面對鏡頭的情形，這時候的正確率是 60.42%；最後是情形(h)，同時有照明、角度與遮擋三種因素影響的情形，在這種情形下很難正確偵測到人臉，因此正確率是最低的，僅有 11.11%；而由表中可看出，本論文的系統對於照明有較好的容忍度，其次是角度，最後是遮擋，也就是受到遮擋時容易對系統的偵測正確率造成很大的影響。

表 4.5 人臉偵測結果分析。

人臉特徵情形	TP	FN	正確率
(a) 正面人臉	232	0	100.00%
(b) 有照明影響	146	6	96.05%
(c) 有角度影響	107	7	93.86%
(d) 有遮擋影響	82	12	87.23%
(e) 情形(b)及(c)	53	15	77.94%
(f) 情形(c)及(d)	26	21	55.32%
(g) 情形(b)及(d)	29	19	60.42%
(h) 情形(b)及(c)及(d)	1	8	11.11%
合計	676	88	88.48%

由以上的結論，本論文的系統在人臉特徵明顯時會有比較好的偵測正確率，如圖

4.5(a)所示；因為照明、角度與遮擋等影響使得人臉特徵不明顯時，就會導致人臉偵測失敗。圖 4.5(b)是因為照明關係使偵測人臉失敗；圖 4.5(c)是因為角度關係使偵測人臉失敗；圖 4.5(d)則是因為遮擋而使偵測人臉失敗。由於本論文的系統做人臉偵測的目的是為了下一步驟的人臉辨識，所以設計上是以儘量能找到最多的人臉，也就是儘量提高正確率為目標；就算會將非人臉誤判為人臉也沒關係，因此在錯誤率上會隨著正確率而提高。



圖 4.5 人臉偵測的部分結果，(a)因為人臉特徵清楚而偵測成功，(b)因為照明影響造成的偵測失敗，(c)因為角度造成的偵測失敗，(d)因為遮擋造成的偵測位置錯誤，(e)因為遮擋造成的偵測失敗。

4.3.2 人臉辨識的結果討論

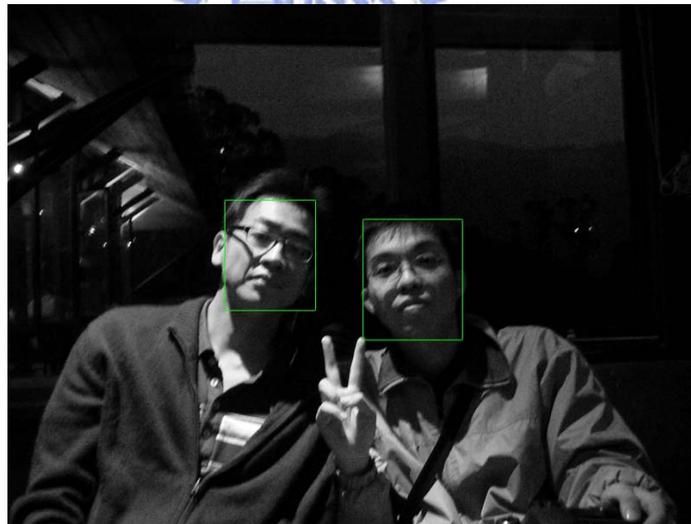
接下來針對人臉辨識的部分做討論，在表 4.6 列出了照明、角度與遮擋三種因素組

合下的可能情形。情形(a)指的是在光線足夠環境且人臉上沒有陰影，人臉角度為正面對鏡頭的人臉，沒有受到任何遮擋的情形，此時人臉的特徵最為清楚，因此辨識率也達到了 98.53%；情形(b)為人臉僅單純受到照明影響，人臉角度為正面，沒有受到任何遮擋的情形，在這種情形下人臉的特徵有部分受到照明影響而變的不清楚，辨識率為次高的 94.62%；情形(c)為人臉僅有角度變化，沒有受到照明與遮擋影響的情形，這種情形下人臉的特徵會因為某些角度而消失，辨識率為 92.21%；情形(d)為人臉僅有受到遮擋的影響，沒有受到照明影響，人臉角度為正面的情形，從表中可看出，在這幾種影響人臉特徵的因素中，遮擋影響的程度最大，不但會使得人臉特徵消失，同時還可能被誤判為非人臉，因此辨識率只有 81.97%；接下來情形(e)為同時包含情形(b)及情形(c)，也就是同時有照明及角度影響，但是沒有遮擋的情形，這時候的辨識率是 69.57%；情形(f)為同時包含情形(c)及情形(d)，也就是同時有角度及遮擋影響，但是沒有照明影響的情形，這時候的辨識率是 24.24%；情形(g)為同時包含情形(b)及情形(d)，也就是同時有照明及遮擋影響，但是人臉角度為正面對鏡頭的情形，這時候的辨識率是 48.65%；最後是情形(h)，同時有照明、角度與遮擋三種因素影響的情形，在這種情形下很難正確辨識人臉，因此辨識率是最低的，僅有 14.29%；從表中可以看出，正面人臉因為人臉的特徵最清楚，所以辨識率也是最高的；其次是照明影響，然後是角度，最後才是遮擋，這個結果與人臉偵測的結果相同，也是可以預期的，因為本論文的系統是以人臉特徵為基礎所建立，所以在特徵清楚時自然會有較好的人臉偵測正確率與人臉辨識率。

針對辨識錯誤的情形，這裡以圖 4.6 為例子並做說明，圖 4.6(a)是測試照片並經過人臉偵測的結果，在其中偵測出兩張待尋人臉，將此兩張人臉做特徵抽取後送入類神經網路做辨識，辨識後的結果如圖 4.6(b)所示，排列的順序是對應於圖 4.6(a)的人臉順序。從圖 4.6(b)中可以看出兩張人臉均辨識錯誤，可能原因為拍攝時受到照明與角度影響，導致人臉特徵與資料庫中的其他人臉相似而造成系統誤判，因此在這種情況下就無法將其區分出來。

表 4.6 人臉辨識結果分析。

人臉特徵情形	辨識正確	辨識錯誤	辨識率
(a) 正面人臉	134	2	98.53%
(b) 有照明影響	88	5	94.62%
(c) 有角度影響	71	6	92.21%
(d) 有遮擋影響	50	11	81.97%
(e) 情形(b)及(c)	32	14	69.57%
(f) 情形(c)及(d)	8	25	24.24%
(g) 情形(b)及(d)	18	19	48.65%
(h) 情形(b)及(c)及(d)	1	6	14.29%
合計	376	26	93.53%



(a)



(b)

圖 4.6 人臉辨識的錯誤結果，(a)人臉偵測後的結果，(b)人臉辨識的結果，由左至右取出資料庫的人臉做展示。

第五章

結論與未來發展

本論文提出了一個可在一般數位相片中執行人臉偵測與人臉辨識的方法。由於數位相片所拍攝的條件相當複雜，會造成偵測人臉及辨識的困難度。因此在人臉偵測的部分，先使用膚色偵測來找出可能是人臉的區域，提升了偵測人臉的效率。然後使用賈伯小波轉換取得人臉整體特徵，並利用區域保留投影將數量龐大的賈伯人臉特徵降至很低的維度，有效增進了類神經網路的效能，使得偵測人臉的所在位置與人臉大小變得更為準確。在人臉辨識的方面，則透過對資料庫中的人臉訓練類神經網路來辨識，在測試相片輸入時，就能夠對於所選定的人臉做辨識，藉此辨識結果就可以將資料庫中含有相同人物的相片挑選出來，達到基於人物身份的影像檢索的目的。

在未來的發展上分為人臉偵測與人臉辨識這兩部分。在人臉偵測的部分，由於偵測是依靠人臉的特徵來進行，當因為人臉角度、照明環境與遮擋導致人臉特徵不明顯時，就會使得偵測率明顯的下降。因此在後續的研究中，首要目標應該是研究改良人臉特徵抽取的方法，避免因為角度與照明導致偵測率下降，同時也能應用在人臉辨識；然後再搭配其他不需要依靠人臉特徵的方法，將人臉成功並準確的偵測出來。在人臉辨識的方面，對於訓練樣本的數量及是否具有代表性的問題，是會直接影響到人臉辨識的辨識率，要如何能從很少的訓練樣本中挑選出具有代表性的樣本來做訓練，將會是後續研究中的首要目標。而從目前的人臉資料庫與本論文所使用的數位相片中來看，可以發現人

臉的樣本都是在相近的時間點所拍攝所得到的。但是人臉會因為隨著年齡增長使臉部特徵產生變化，或著是因為臉部特徵的改變，如眼鏡、化妝等因素，可能會造成辨識的困難，因此如何解決這種問題，是值得繼續研究的。



參考文獻

- [1] X. Tan, S. Chen, Z. H. Zhou, and F. Zhang, "Face Recognition from A Single Image Per Person: A Survey," *Pattern Recognition*, vol. 39, pp. 1725-1745, 2006.
- [2] M. H. Yang, D. Kriegman, and N. Ahuja, "Detecting Faces in Images: A Survey," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, pp. 34-58, Jan. 2002.
- [3] M. Turk and A. Pentland, "Eigenfaces for Recognition," *J. Cognitive Neuroscience*, vol. 3, pp. 71-86, 1991.
- [4] H. Rowley, S. Baluja, and T. Kanade, "Neural Network-Based Face Detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, pp. 23-38, Jan. 1998.
- [5] E. Osuna, R. Freund, and F. Girosi, "Training Support Vector Machines: An Application to Face Detection," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 130-136, San Juan, Puerto Rico, 17-19 Jun. 1997.
- [6] A. Rajagopalan, K. Kumar, J. Karlekar, R. Manivasakan, M. Patil, U. Desai, P. Poonacha, and S. Chaudhuri, "Finding Faces in Photographs," *Proc. IEEE Int'l Conf. Computer Vision*, pp. 640-645, Bombay, India, Jan. 1998.
- [7] A. S. Georghiades, P.N. Belhumeur, and D.J. Kriegman, "From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, pp. 643-660, Jun. 2001.
- [8] L. Wiskott, J. M. Fellous, N. Kruger, and C. v. d. Malsburg, "Face Recognition by Elastic Bunch Graph Matching," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, pp. 775-779, Jul. 1997.

- [9] P. Kakumanu, S. Makrogiannis, and N. Bourbakis, "A Survey of Skin-Color Modeling and Detection Methods," *Pattern Recognition*, vol. 40, pp. 1106-1122, Mar. 2007.
- [10] J. Yang, W. Lu, and A. Waibel, "Skin-Color Modeling and Adaptation," *Proc. Asian Conf. Computer Vision*, pp. 687-694, Hong Kong, Jan. 1998.
- [11] Q. H. Thu, M. Meguro, and M. Kaneko, "Skin-Color Extraction in Images with Complex Background and Varying Illumination," *Proc. IEEE Workshop Applications of Computer Vision*, pp. 280-285, Orlando, FL, USA, Dec. 2002.
- [12] D. Chai and A. Bouzerdoun, "A Bayesian Approach to Skin Color Classification in YCbCr Color Space," *Proc. Intelligent Systems and Technologies for the New Millennium*, pp. 421-424, Kuala Lumpur, Malaysia, Sep. 2000.
- [13] F. Tomaz, T. Candeias, and H. Shahbazkia, "Fast and Accurate Skin Segmentation in Color Images," *Proc. Canadian Conf. Computer and Robot Vision*, pp. 180-187, May 2004.
- [14] M. J. Jones and J. M. Rehg, "Statistical Color Models with Application to Skin Detection," *Int'l J. Computer Vision*, vol. 46, pp. 81-96, 2002.
- [15] M. J. Seow, D. Valaparla, and V. K. Asari, "Neural Network-Based Skin Color Model for Face Detection," *Proc. Workshop Applied Imagery Pattern Recognition*, pp. 141-145, Washington, DC, USA, Oct. 2003.
- [16] N. Sebe, T. Cohen, T. S. Huang, and T. Gevers, "Skin Detection: A Bayesian Network Approach," *Proc. Int'l Conf. Pattern Recognition*, vol. 2, pp. 903-906, Cambridge, UK, Aug. 2004.
- [17] M. H. Yang and N. Ahuja, "Gaussian Mixture Model for Human Skin Color and Its Application in Image and Video Databases," *Proc. SPIE: Conf. Storage and Retrieval for Image and Video Databases*, vol. 3656, pp. 458-466, San Jose, Jan. 1999.
- [18] J. Y. Lee and S. I. Yoo, "An Elliptical Boundary Model for Skin Color Detection," *Proc. Int'l Conf. Imaging Science, Systems and Technology*, pp. 472-479, Las Vegas, USA, Jun.

2002.

- [19] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Upper Saddle River, NJ: Prentice Hall, 2002.
- [20] D. Gabor, "Theory of Communications," *J. Institution of Electrical Engineers*, vol. 93, pp. 429-457, 1946.
- [21] J. G. Daugman, "Uncertainty Relation for Resolution In Space, Spatial- Frequency, and Orientation Optimized by Two-Dimensional Visual Cortical Filters," *J. Optical Society of America A-Optics Image Science and Vision*, vol. 2, pp. 1160-1169, 1985.
- [22] L. L. Shen and L. Bai, "A Review on Gabor Wavelets for Face Recognition," *Pattern Analysis and Applications*, vol. 9, pp. 273-292, Aug. 2006.
- [23] J. G. Daugman, "Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 36, pp. 1169-1179, Jul. 1988.
- [24] M. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba, "Coding Facial Expressions with Gabor Wavelets," *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition*, pp. 200-205, Nara, Japan, Apr. 1998.
- [25] S. T. Roweis and L. K. Saul, "Nonlinear Dimensionality Reduction by Locally Linear Embedding," *Science*, vol. 290, pp. 2323-2326, Dec. 2000.
- [26] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A Global Geometric Framework for Nonlinear Dimensionality Reduction," *Science*, vol. 290, pp. 2319-2323, Dec. 2000.
- [27] M. Belkin and P. Niyogi, "Laplacian Eigenmaps and Spectral Techniques for Embedding and Clustering," *Neural Information Processing Systems*, vol. 14, pp. 585-591, 2002.
- [28] X. F. He and P. Niyogi, "Locality Preserving Projections," *Neural Information Processing Systems*, vol. 16, 2003.
- [29] F. Rosenblatt, "The Perceptron - A Probabilistic Model for Information-Storage and Organization in the Brain," *Psychological Review*, vol. 65, pp. 386-408, 1958.

- [30] J. J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities," *Proc. National Academy of Sciences*, vol. 79, pp. 2554-2558, USA, Apr. 1982.
- [31] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning Representations by Back-Propagating Errors," *Nature*, vol. 323, pp. 533-536, 1986.
- [32] T. Sim, S. Baker, and M. Bsat, "The CMU Pose, Illumination, and Expression (PIE) Database," *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition*, pp. 53-58, Washington, DC, USA, May 2002.
- [33] T. Sim, S. Baker, and M. Bsat, "The CMU Pose, Illumination, and Expression Database," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, pp. 1615-1618, Dec. 2003.
- [34] L. L. Huang, A. Shimizu, and H. Kobatake, "Robust Face Detection Using Gabor Filter Features," *Pattern Recognition Letters*, vol. 26, pp. 1641-1649, Aug. 2005.
- [35] S. Kadoury and M. D. Levine, "Face Detection in Gray Scale Images Using Locally Linear Embeddings," *Computer Vision and Image Understanding*, vol. 105, pp. 1-20, Jan. 2007.
- [36] L. Wang, Y. P. Li, H. Z. Zhang, and C. B. Wang, "A Novel 2D Gabor Wavelets Window Method for Face Recognition," *Multimedia Content Representation, Classification and Security*, vol. 4105, pp. 497-504, 2006.
- [37] Q. B. You, N. N. Zheng, S. Y. Du, and Y. Wu, "Neighborhood Discriminant Projection for Face Recognition," *Pattern Recognition Letters*, vol. 28, pp. 1156-1163, Jul. 2007.