國立臺灣師範大學應用電子科技研究所

碩士論文

指導教授：高文忠博士

應用區域對比增強於不均勻光源下之人臉辨識

Local Contrast Enhancement for Human Face Recognition in Poor Lighting Conditions

研究生：徐民儕　撰

中 華 民 國 九 十 七 年 六 月

應用區域對比增強於不均勻光源下之人臉辨識

學生：徐民儕　　論文指導教授：高文忠 博士

國立臺灣師範大學應用電子科技研究所碩士班

摘要

近幾年來，由於安全上的需求，所以利用人臉來進行身份辨識的應用越來越廣泛，在許多從事人臉辨識的研究的文獻中，常利用人臉影像擷取出來的特徵，來分辨出不同的人。然而在實際的應用上，常常會因為環境中光源的不均勻照射，使得同一張人臉會有很大的不同，因而導致人臉的辨識率大幅下降，為了提昇辨識效能，我們提出一個區域對比增強的方法，可以有效的解決人臉辨識在不同光源下的改變。

本篇論文提出的人臉辨識的演算法，則是在辨識前對影像做離散餘弦轉換，取出人臉影像的低頻部份，有效降低影像的維度，因此在辨識的時間上也會相對的減少，最後交給支持向量機（SVM），來決定辨識的結果。本論文測試的人臉資料庫為 Yale_B，經使用支持向量機的辨識率可達 99.13%，在已發表的論文中是辨識較好的方法之一。

關鍵字：特徵抽取、人臉辨識、支持向量機

Local Contrast Enhancement for Human Face Recognition in Poor Lighting Conditions

Student：Ming-Chai Hsu Advisor：Dr. Wen-Chung Kao

Institute of Applied Electronics Technology
National Taiwan Normal University

Abstract

In recent years, many face recognition algorithms have been developed for surveillance systems and promising results have been reported in specific environments. The human face recognition highly relies on extracted stable features from input images. In practical application environments, however, the direction of the illuminant is uncontrollable and it will result in unstable feature extraction. For remedying the problems caused by non-uniform light sources, illumination compensation is necessary.

In this thesis, we propose a local contrast enhancement approach to reduce the effect of non-uniform light sources, and integrate it with a face recognition system. Through the process of local contrast enhancement, the facture extraction based on digital cosine transformation (DCT) becomes more reliable. The adopted classification kernel is support vector machines (SVM) which has been shown to be a robust classifier. The well-known human face database Yale_B is used for verifying system performance, and the recognition rate can achieve to 99.13%. As far as we known, the recognition rate is better than all of the published literatures.

Keywords: Feature extract, Face recognition, Support vector machines

致　　謝

首先感謝指導教授高文忠博士，在二年的研究生生涯中給予我課業上的指導和生活上的關懷，特別在面對問題，研究態度和危機處理等方面的能力都是值得我去效法，不僅在學術上的堅持與努力，在對於品德操守也十分的要求，這些都使得我獲益良多。

感謝實驗室的所有成員包括了：上一屆學長宏碩、聯暘、家平及本屆同學與學弟包括瑋琦、奇明、嘉安、志兆、志祥和大學部的學弟敬中，尤其是從業界來幫忙的宏信，提供許多協助，本人難忘和大家一起努力奮鬥的日子，並且給予種種建議和幫助。在這實驗室中充滿了許多酸甜苦辣的回憶。感謝所有幫助過我和關心我的人，謝謝大家我會永遠記得你們。最後祝大家身體健康、萬事如意。

最後要感謝我家人，因為有父母親用心栽培與家人的鼓勵，這樣的關心成為我最大的精神力量，讓我在研究所期間能專注於研究，使得我才能順利完成碩士學業，希望藉由本論文的完成，與他們分享我的榮耀與喜悅，並願他們永遠身體健康。

徐民儕　謹誌

于台灣師範大學系統晶片實驗室

2008年6月30日

目錄

圖　　目　　錄

表　　目　　錄

第一章 緒論

1.1 研究動機

隨著科技的日新月異，人類的經濟規模變得越來越大，人們開始重視個人隱私，因此保護個人隱私就顯得非常重要，早期使用密碼或鑰匙來保護珍貴的資料或物品，但偽造與犯罪的手法不斷翻新，破解密碼或鑰匙的技術也更為提升，因此鑰匙或密碼若為他人取得，珍貴的物品或資訊就有可能被竊取，而且遺忘密碼或遺失鑰匙的情況也時常發生，所以可以隨身攜帶不會遺失而且難以被其他人所偽造的保密方法正好可以解決以上的問題，利用每個人獨有的特徵—「生物特徵」，作為使用者驗證的依據，也因此成為科學界努力研究的方向。其中語音、人臉、DNA、指紋、虹膜等，都是常見甚至在日常生活已被廣泛應用的生物特徵，例如聲控操作、門禁系統、身份證件等，均是生物特徵常見的應用，但是對於使用者而言，一個方便且具有人性化的人機介面是實際應用上必須被考慮到的問題，語音辨識則是容易受到外界噪音干擾，辨識效果常常打折扣。虹膜、DNA、指紋等生物特徵雖然辨識準確率高，但是擷取特徵的儀器成本較高不易取得，且使用時可能需要使用者與儀器直接接觸。

為了能在不同的場合表示自己的身份，所以每個人的身上都要攜帶各種不同場合的證件，而這麼多的證件都只是要來表示一個人的身份。因此造成兩個嚴重的問題，第一個問題是太多的證件會造成攜帶的不方便。第二個問

題在現今各種證件竊取及偽造事件相當頻繁，因此證件的安全性及可靠性也就愈來愈受到質疑。所以近年來，大部分都採用人臉特徵來當作辨識身分的選擇，這是因為人臉特徵和傳統的證件不同，每個人臉上的特徵都不同，因此想要複製或仿冒不是件容易的事，而且取得人臉影像是相當方便，所以如果能自動化地利用人臉特徵來辨識身份，那就不需要攜帶大量的證件，沒有攜帶的麻煩，也不用去擔心證件會被複製或仿冒，這可以增加身份認證的安全性、可靠性和便利性。因此以人臉取得最為方便與最不具侵入式，便成為最常使用的生物特徵，在門禁系統上更是有相當大的需求。雖然人臉辨識可以說是一把不會遺失且難以被仿造的鑰匙，但在處理上卻是相當複雜，往往需要耗費龐大的時間與計算量，並且很容易受到環境光源改變的影響，使生物特徵辨識衍生出了許多值得研究的相關議題，譬如如何解決環境光源變化的影響、提高生物特徵辨識的效率與降低運算量以及增加模型的彈性使生物特徵辨識能在各種環境之下都能正常的運作，每個主題都是相當艱深而且需要長時間的研究。

1.2 相關研究

在人臉辨識方面，近年來有相當多的研究和方法不斷的改善辨識率，而在技術不斷的突破和經驗的累積，其相關的研究也漸漸成熟，所以有非常多的論文都針對人臉辨識會遇到的問題提出說明與他們的解決方法。雖然有這麼多的方法被提出與使用，但是由於人臉辨識需要考慮的因素實在太多，例如：運算複雜度高，因而必須實現在高效能的電腦【18】【19】【20】【21】【22】。人臉辨識會遭遇到的問題如下所示：

(1) 臉部大小：影像中人臉的大小，因為大小的人臉不同，因此會產生

些許的差異，很容易會導致辨識率下降，因此可以利用一些前置處理，來避免此情況的發生。

(2) 姿勢與位置：重疊、遮住等等，都會影響到系統的辨識。

(3) 旋轉角度：人的臉部特徵，在當面對相機作不同角度方向旋轉後會有極大的變化。例如：左轉與右轉後臉上的特徵就會少掉一大半。上仰的話，臉的上半部就會消失不見；反之俯視的話，臉下半部就看不到。最重要的是，不管是什麼角度，在影像中特徵的相對位置都會改變，導致無法辨識。

(4) 光源環境：光源的強度、角度等等，都會影響到臉上色彩分佈與強度的散佈。

(5) 人臉的差異：雖然人的臉大部分就是由眼睛、鼻子、眉毛、嘴唇、耳朵等特徵組成，但是不同人之間還是有相當大的差異。這樣的差異就是人在這些器官上的彈性相當大，如突顯程度與位置，甚至戴眼鏡與髮型遮蓋的問題。

(6) 表情：由於人的五官是彈性的。不同的表情下，人臉會有很大的差異。

(7) 其他：相機的鏡頭特性、CCD 的增益、曝光時間、光圈調整、自動白平衡、色彩校正等都會影響到畫面。

基於上述的幾點就有非常多的文獻去研究以解決上述的問題。Zhao et al.【1】認為人臉辨識的方法，基本上可以分為以下四類：

(1) Feature-based methods：這類的方法主要目標是以人臉特徵點的距離與角度、人臉特徵的形狀或特徵的亮度關係進行辨識。

(2) Template-matching methods：此類方法會事先定義好的代表人臉的 patterns。藉由比對整個輸入影像與樣板間 correlation value 的來做人

臉辨識，當 correlation value 值越高，表示這個輸入 pattern 與人臉 template 之間越像。

(3) Neural Network methods：此類的方法主要是嘗試去模仿人類的神經系統，利用神經元和運算單元間的眾多連結來做平行且分散的方式運算，因此可以同時處理大量的資料。其主要方法是以疊代的方式不斷修正神經網路中的輸入值與期望輸出值之間的誤差，直到誤差小於一定的臨界值。

(4) Statistics-based methods：此類的方法主要是利用數學轉換取得人臉影像中最重要的特徵向量，並利用該重要的特徵向量與資料庫中的向量來進行比較。

1.2.1 Feature-based methods

Li et al.【2】提出了 NFL（Nearest Feature Line）的方法進行人臉辨識，作者認為在同一個人臉部分的影像任取兩個特徵點所連成的特徵線是一致的，因此利用臉部特徵點來針對每一個特徵線做最近距離的計算。實驗中分為二大部份：第一部份利用 NFL 的方法來與傳統的特徵臉（Eigenface）方法來進行比較，使用混合多種常用的人臉資料庫，共計 132 人 1079 張影像，結果顯示 NFL 的方法錯誤率較特徵臉的方法低，錯誤率改善了 43.7~65.4%。第二部份則是利用 NFL 的方法與迴旋類神經網路來進行比較，人臉資料庫採用 ORL 資料庫，每人 10 張影像，共有 40 人 400 張影像，實驗結果 NFL 的方法錯誤率僅 3.125%，較迴旋類神經網路的錯誤率 3.83%來的更低，錯誤率改善了 81.59%。

1.2.2 Template-matching methods

Brunelli et al. 【3】比較特徵式及模板式運用在人臉辨識中的方法，並採用 47 個人的正面人臉影像（包含男性 26 人、女性 21 人，每人 4 張），並利用這兩種方法發展出新的演算法來進行實驗。其中特徵式的方法是計算人臉的幾何特徵值（如眼睛、鼻子及嘴巴）；模板式的方法則是在灰階中進行眼睛、鼻子、嘴巴及臉部等 4 個臉部遮罩的模板來進行比對。經過實驗結果，特徵式的方法辨識率僅達到 90%，而模板式的方法辨識率則達到 100%。

1.2.3 Neural Network methods

Er et al. 【4】對於人臉辨識提出了二個論點：

(1) 使用類神經網路進行人臉辨識前，特徵擷取的方法是相當重要的。
(2) 輻射基底函式（Radial Basis Function）類神經網路僅需要訓練少量的樣本數，就可以達到一定的分類程度。

作者在前置處理中，利用主成份分析（Principal Component Analysis，PCA）的方法來降低影像的維度，再利用 FLD（Fisher's Linear Discriminate）對不同的類別擷取特徵，再將擷取出來的特徵使用輻射基底函式類神經網路來做最後的分類，人臉資料庫採用 ORL 資料庫，每人有 10 張影像，有 40 人共 400 張，200 張為訓練另外 200 張為測試，共實驗了六次，平均辨識率為 98.02%。

1.2.4 Statistics-based methods

Bicego et al.【5】用 Canny 濾鏡來抽取影像中人臉的邊緣並利用多階層 B-spline 曲線來取得人臉的特徵值，以降低人臉影像的維度，最後再給支持向量機完成分類。實驗中採用多種人臉資料庫進行比較辨識率，其中 ORL 資料庫辨識率為 97.25%，Bern 資料庫中辨識率為 95.33%，Yale 資料庫的辨識率為 98.89%。

1.3 本論文提出之方法

在人臉辨識實際應用的環境中，光源的照射常常是不均勻的，而眼睛對亮度相對之間的感覺較亮度之間的絕對差值有較大的靈敏度，與真實的亮度強度絕對值沒有關聯。若光以亮度來看，並不是整張影像都一樣，而是根據亮度強度的變化會群聚成小塊區域來分佈。因此我們提出以區域對比增強（Local Contrast Enhancement）的演算法，來增強區域性的影像對比，而使得影像的主要細節與紋路變得相當清楚，這對後續的影像辨認有很大的助益。

在特徵擷取部份，我們使用 8 × 8 當成每一個 Block，每個 Block 經過離散餘弦轉換（Discrete Cosine Transform，DCT）後的亮度 Y 的直流值與前三個 AC 值當作辨識系統的輸入特徵。

針對人臉辨識方面，我們使用的方法是目前相當流行，也常常被應用在人臉辨識的分類器—支持向量機（Support Vector Machines，SVM）。SVM 是圖形識別中一門相當新穎的技術，傳統的圖形識別是以經驗之風險最小化，

達到對訓練資料有最好的效果，SVM 則是使用結構之風險最小化，而結構之風險最小化是以訓練資料的錯誤總和及 Vapnic-Chervonenkis 維度來決定一般化錯誤的上限，利用對於這個上限做最小化，試圖改善尚未觀察到的資料被分類錯誤的機率，以達到不錯的效果。

SVM 被認為是具有高效能的分類器之一，也可以對大量的資料作處理，在許多方面已被廣泛的利用，但 SVM 在使用上有某些限制，譬如 SVM 是利用對結構風險的上限作最小化達到不錯的結果，但 SVM 若與其他以機率為出發點的辨識器相比卻缺乏能夠衡量決策之不確定性的因素，Kwok 因此將調節輸出應用在 SVM 上【6】，把 SVM 的決策結果改成機率形式，並且建立 SVM 與貝氏理論間的關係，Tipping 則發展出以機率形式為出發點的關聯向量機（relevance vector machine，RVM）。由於不是以機率為出發點，所以 SVM 也無法對已訓練好的參數做漸進式的學習（sequential learning），即使完成超平面參數的訓練後，仍需要把所有資料儲存起來，才能在新資料進入系統時，擁有所有的訓練資料對超平面的參數做重新的訓練，相當花費計算時間與儲存空間。

1.4 論文架構

本章一開始介紹了本論文的研究動機與人臉辨認的相關研究，並提出本論文的研究方法。第二章介紹一些相關文獻。第三章介紹人臉辨識的系統架構。第四章介紹我們如何去抽取特徵並利用統計分析找出有用的特徵值後，丟入辨認核心系統取做人臉辨認，本論文的辨認核心為使用支援向量機(Support Vector Machines, SVMs)。第五章為實驗結果。最後第六章為結論與未來研究工作。

第二章 人臉辨識的相關研究及探討

2.1 相關研究概述

近年來，人臉辨識有相當多的研究和方法，而在技術不斷的突破和經驗的累積，其相關研究也愈來愈多，在這些研究中大致上可分為兩種類型：整體特徵方法及局部特徵方法。整體特徵的方法是直接將整張人臉當特徵來辨識（如圖 2-1），例如：PCA 是不含資料的類別資訊，主要將所有的資料結合在一起做處理或 Linear Discriminant Analysis（LDA）主要是將資料的類別資訊加入，並不是全部的資料混合一起做處理。局部特徵的方法則是先找出臉上的局部特徵，例如眼睛、鼻子、耳朵及嘴巴…等，利用這些局部特徵來辨識。最後在取得特徵做比對的相關研究也相當多，一般來說是用歐基里德的方法，求取最小的距離來當作結果，而不同的比對方法所產生的辨識率也不同，例如：可以使用類神經網路來當分類，利用類神經網路的多層感知機（Multi-layer neural network，MLNN）在辨識上也有不錯的效果。底下我們將分成兩個小節來探討這兩種類別的相關研究。

圖 2-1　整張人臉當特徵

2.2 整體特徵方法

Turk et al.【7】假設所有的人臉都可以用一組基底人臉的線性組合表示，而作者以主成份分析（Principal Components Analysis，PCA）【9】【10】來找出這個基底，其簡介如下：

PCA 主要目的是將許多變項加以減少，使其改變為少數幾個互相獨立的線性組合變項，經由線性組合所得的成分之變異數會變為最大，使得原本的 N 維資料在這些成分上顯出最大的個別差異。

PCA 使用在特徵擷取有兩個理由，第一 PCA 方法能夠快速容易的計算出結果，第二在線性投影中 PCA 方法能夠維持被投影資料最大的資訊。其目的是期望能使用比較少的變數來解釋原先資料中的大部分變異，使得許多相關性很高的變數轉變成彼此互相獨立的新變數，我們從其中選取比原來變數個數少的幾個新變數，就能代表大部份資料中的變異，這就是主成分，其主要特性如以下二點所示：（1）各主成分的方向是互相垂直的。（2）各主成

分是互相獨立的。

而PCA主要的原理簡述如下：

假設原來共有N張訓練用之人臉影像，其原始特徵參數為$\{X_1, X_2, ..., X_n\}$。PCA之目的為找出一個 n×m 線性轉換矩陣P，將原始維度為n之特徵參數X_k轉換成維度為m（m≦n），且更具代表性（即兩兩間的變異數更大）之新參數Z_k，如（2-1）式。

$$Z_k = P^T X_k, k = 1,2,...,n \tag{2-1}$$

令轉換前的平均向量(Mean Vector)為$\overline{X}$，則轉換後的平均向量如（2-2）式：

$$\overline{Z} = \frac{1}{N}\sum_{k=1}^{N} Z_k = \frac{1}{N}\sum_{k=1}^{N} P^T X_k = P^T\left(\frac{1}{N}\sum_{k=1}^{N} X_k\right) = P^T \overline{X} \tag{2-2}$$

以全域散佈矩陣(total scatter matrix）表示所有特徵參數相對於其平均向量的分散程度。令轉換前n × n全域散佈矩陣如（2-3）式。

$$S_{tx} = \sum_{k=1}^{N}\left(X_k - \overline{X}\right)\left(X_k - \overline{X}\right)^T \tag{2-3}$$

則由(2-1)(2-2)(2-3)式可得轉換後 n × n 的全域散佈矩陣如（2-4）式。

$$S_{tz} = \sum_{k=1}^{N} \left(Z_k - \overline{Z}\right)\left(Z_k - \overline{Z}\right)^T = \sum_{k=1}^{N} \left(P^T X_k - P^T \overline{X}\right)\left(P^T X_k - P^T \overline{X}\right)^T = P^T S_{t_x} P \qquad (2\text{-}4)$$

為將轉換後特徵參數與其平均值之間的散佈程度加大，必須找出能使 S_{tx} 最大化之轉換矩陣 P_{opt} ，如（2-5）式。

$$P_{opt} = \arg\max\left(P^T S_{tx} P\right) \qquad (2\text{-}5)$$

根據線性代數理論，可以採用某方陣之跡(Trace）或是行列式(Determinant）表示其內部元素分布的情形。因此（2-5）式可以改寫成以下兩種形式。如（2-6）式或（2-7）式所示。

$$P_{opt} = \arg\max tr\left(P^T S_{tx} P\right) \qquad (2\text{-}6)$$

或

$$P_{opt} = \arg\max\left|P^T S_{tx} P\right| \qquad (2\text{-}7)$$

於（6）式中，為使 $tr\left(P^T S_{tx} P\right)$ 之值有所限制，不至於變成無限大這種無法控制的結果 若令P為n×m轉換矩陣，則需加入 $P^T P = I_m$ 限制條件如（2-8）式：

$$F(P) = P^T S_{tx} P - \lambda(P^T P - 1) \qquad (2\text{-}8)$$

為使F(P)最大化，必須根據P取其一階導數，並將其結果設為零。如此可以得到（2-9）式：

$$\frac{\partial F(P)}{\partial P} = 2S_{tx}P - 2\lambda P = 0 \tag{2-9}$$

再進一步簡化式子如（2-10）式：

$$S_{tx}P = \lambda P$$
$$(S_{tx} - \lambda I)P = 0 \tag{2-10}$$

解(2-10)式之結果，可得 P 恰為 S_{tx} 之特徵向量(eigenvector）所組成之矩陣。

因此，原始的人臉特徵參數經過PCA轉換後，可以得到新的特徵參數，而這組特徵參數除了維度降低之外，各參數間的變異程度也是最大，所以我們可以用較少維度的特徵，來表達出每張不同人臉影像之間的差距，也就是取得最具代表性的特徵。如果我們把每一個特徵向量依照影像形成特徵參數順序來排列成一張影像，會發現其形成的圖形看起來很像一張人臉，所以這種PCA轉換後的人臉特徵，又被稱做特徵臉(Eigenface)【15】，如圖2-2所示。

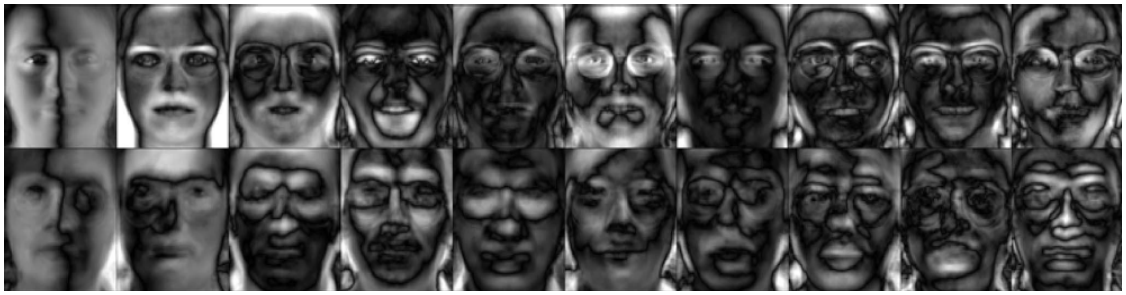

圖 2-2　Yale 人臉資料庫，取 10 個特徵值所對應的特徵臉影像

Yang et al.【16】提出2D-PCA的方法，基本概念是使用較直覺二維矩陣(Matrices)代替特徵臉技術使用之一維向量(Vector)，表示人臉影像像素資料，以進行主要成份分析。首先求出特徵值(eigenvalue)與特徵向量(eigenvector)，它可以把矩陣中的元素成分重組，將重要的資訊集中在較大的特徵值所對到的特徵向量中。其特殊的性質如下：

(1) 方陣中所有非零的特徵值之"積和"，等於該方陣的行列式值如（2-11）式：

$$\prod_{i=1}^{m} \lambda_i = |S_{tz}| \tag{2-11}$$

(2) 方陣中所有特徵值之和，等於該方陣的跡，如（2-12）式：

$$\sum_{i=1}^{m} \lambda_i = tr(S_{tz}) \tag{2-12}$$

第二步為求得我們所需的特徵空間，一般都是先求平均值 -> 求Zero Mean -> 計算Covariance Matrix -> 計算特徵值與特徵向量 -> 最後求得特徵空間。詳細的過程和公式如下：

求平均值：將訓練樣本加總起來除以個數如（2-13）式

$$m = \frac{1}{k}\sum_{i=1}^{k} x^i \quad , x^i = \left[x_1^i, x_2^i, ..., x_N^i\right]^T \tag{2-13}$$

其中 k 為訓練樣本個數，N 為每一樣本的維度

Zero Mean：把所有訓練樣本減掉平均值如（2-14）式

$$\overline{x^i} = x^i - m, i = 1,2,...k \tag{2-14}$$

計算 Covariance Matrix，如（2-15）式：

$$C = \sum_{i=1}^{k} \overline{x^i} \cdot \overline{x^i}^T \tag{2-15}$$

計算特徵值與特徵向量：由 Covariance Matrix 來求得特徵值與特徵向量，如（2-16）式：

$$c\phi_i = \lambda\phi_i \tag{2-16}$$

其中 ϕ_i 為特徵向量，λ 為特徵值

計算特徵空間：依照計算得到的特徵值由大到小做排序，將所對應的特徵向量組合而成特徵空間，而選取的特徵向量則為所對應的特徵值，是個非零的特徵向量如（2-17）式。

$$\phi = [\phi_1, \phi_2, ..., \phi_k] \tag{2-17}$$

其中 $\lambda_i = \phi_i \subset \varphi_i \neq 0$ and $\lambda_i > \lambda_i + 1, for \quad 1 \leq i \leq k$

第三步為將訓練樣本投影到特徵空間，根據上述的步驟可得訓練樣本的

特徵，如（2-18）式。

$$\overline{y^i} = \phi^T \overline{x^i} \Leftrightarrow \overline{Y} = \phi^T \overline{X} \tag{2-18}$$

測試樣本投影到特徵空間：當有一個測試樣本要做辨識時，將其樣本減掉訓練樣本的平均值，接著再投影到特徵向量上，便可以與訓練樣本做比對（如圖2-3所示）

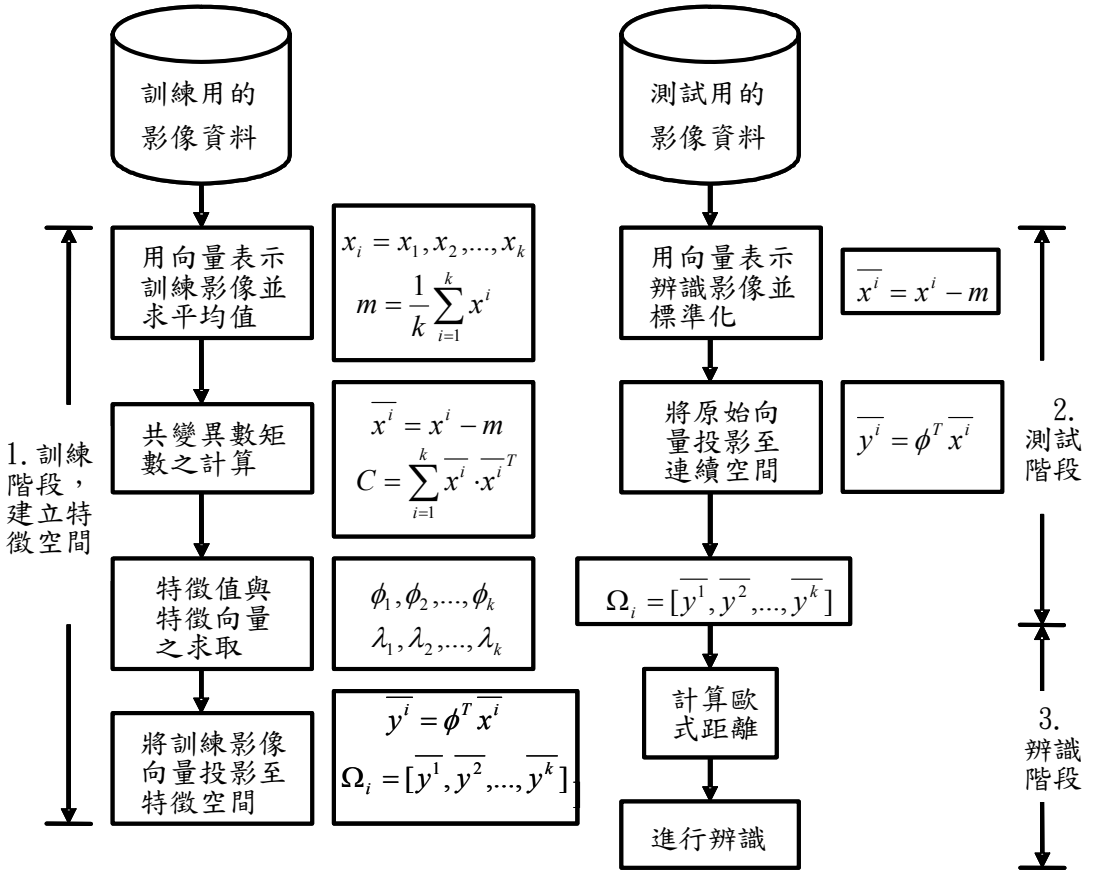

圖 2-3　PCA 分析之研究流程

此種方式提升擷取人臉影像特徵之效率，但在影像儲存方面，因需要較

多之係數表示人臉，故效率較差。而實驗結果顯示，在已知之三個人臉資料庫中，2D-PCA具有較佳之辨識率，但在其他資料庫中，與特徵臉並沒有顯著之差異。

Belhumeur et al.【8】更進一步利用 Linear Discriminant Analysis（LDA）【17】將不同的人臉影像投射到高維空間使其儘量分開，因而提高辨識率。其推導如下，首先定義各類資料的分散程度（S_B）。如（2-19）式：

$$S_B = \sum_{i=1}^{c} n_i (m_i - m)(m_i - m)^T, m = \frac{1}{n}\sum_x x \tag{2-19}$$

當要投影到高維空間後，我們不再是求一個vector basis W，而是要求一組basis，所以多組W 將會寫成一個matrix W來表示，裡面的column vector 就是一個basis。如（2-20）式：

$$\widetilde{S}_B = W^t S_B W, \widetilde{S}_W = W^t S_W W \tag{2-20}$$

所以原來的$J(W)$將變成如（2-21）式：

$$J(W) = \frac{\left|W^t S_B W\right|}{\left|W^t S_W W\right|} \tag{2-21}$$

S_W數字愈大代表同類別之間愈集中，而$J(W)$中的W是一個矩陣，代表一組basis。分子分母因為W是矩陣，所以必須加個determinant 才能變成常數。若要求W為第i個column vector，只要解如（2-22）式，則取得第i大的特徵值

（eigenvalue），即可求出對應的特徵向量（eigenvector）。因此若要投影到k維的空間，只要取前k大的特徵值對應的特徵向量就可以了。

$$S_B W_i = \lambda S_W W_i, \quad i = 1,2,...,n \tag{2-22}$$

如圖2-4所示，為一個PCA及FLD的比較。圖中的 o 跟 ＋ 是兩個不同的類別，用PCA求出降維後的basis，會使所有的資料投影到那basis 產生的error (Euclidean distance)最小。另一方面，Fisher Linear Discriminant (LDA) 所產生的basis 就不一樣了，你可以看得出來 o 跟 ＋ 被投影到LDA basis上時，有明顯被區分成兩群的情況，我們可以在投影過後的空間中，決定出一個點把兩群資料分開，而PCA投影過後的資料沒有辦法決定一個點把兩群資料分開。

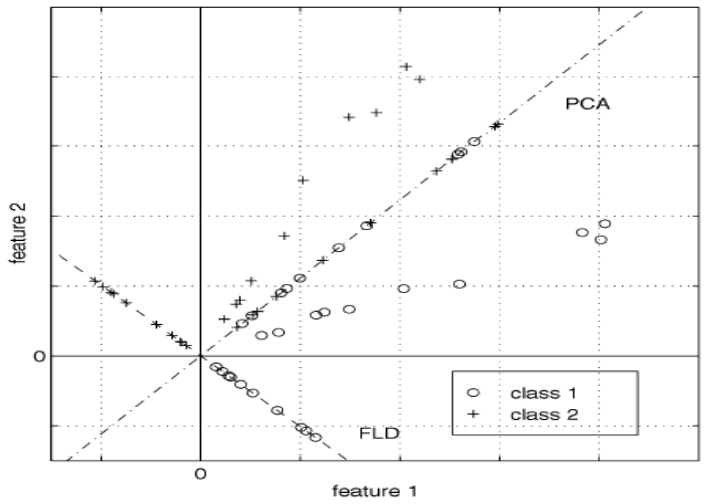

圖 2-4　PCA 及 FLD 的比較

作者在實驗的人臉資料庫中採用了 Harvard 資料庫，此資料庫共分為五個 subset，每個 subset 都是由不同角度的光源照射變化（如圖 2-5），

各種方法的辨識率如圖 2-6，正確數據如表 2-1。

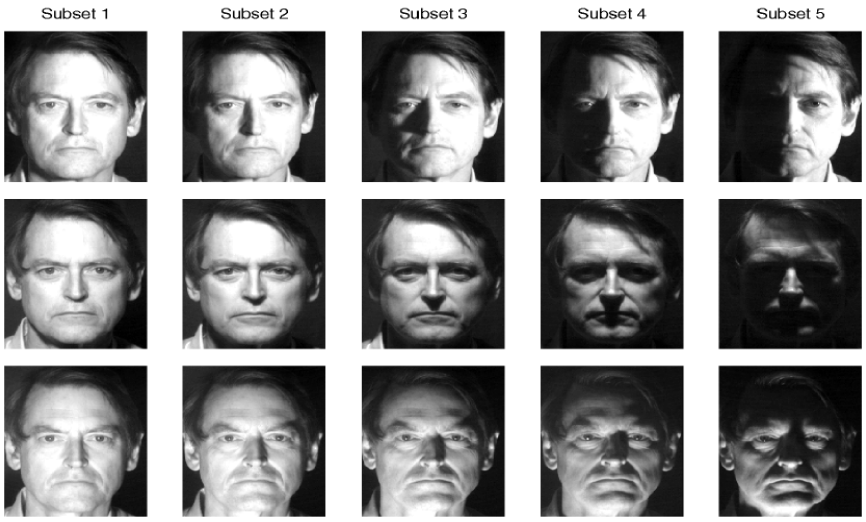

圖 2-5　不同角度的光源照射

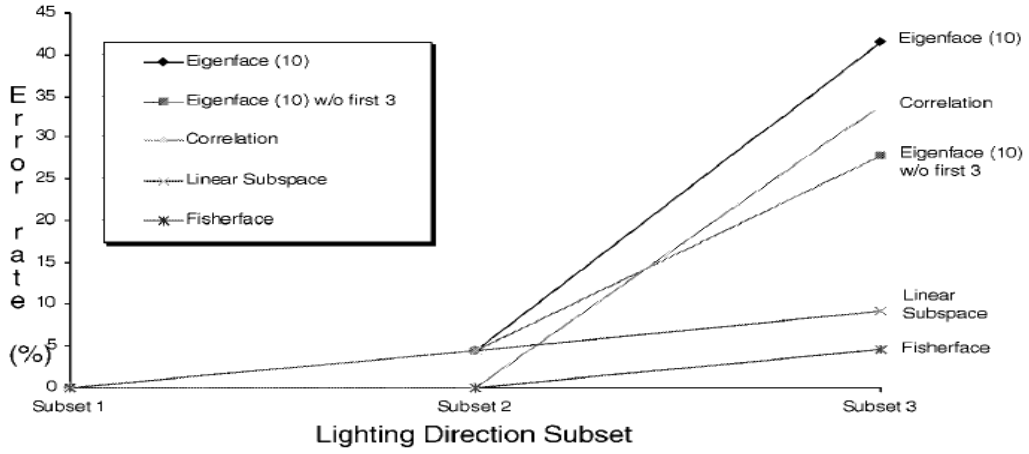

圖 2-6 各種方法的辨識結果

表 2-1　各種方法的辨識數據

Method	Reduced Space	Error Rate (%)		
		Subset 1	Subset 2	Subset 3
Eigenface	4 10	0.0 0.0	31.1 4.4	47.7 41.5
Eigenface w/o 1st 3	4 10	0.0 0.0	13.3 4.4	41.5 27.7
Correlation	29	0.0 0.0	0	33.9
Linear Subspace	15	0.0 0.0	4.4	9.2
Fisherface	4	0.0 0.0	0	4.6

作者在實驗的人臉資料庫中採用了 Yale 資料庫，此資料庫影像各有不同的情緒與光照條件（如圖 2-7），各種方法的辨識率如圖 2-8。

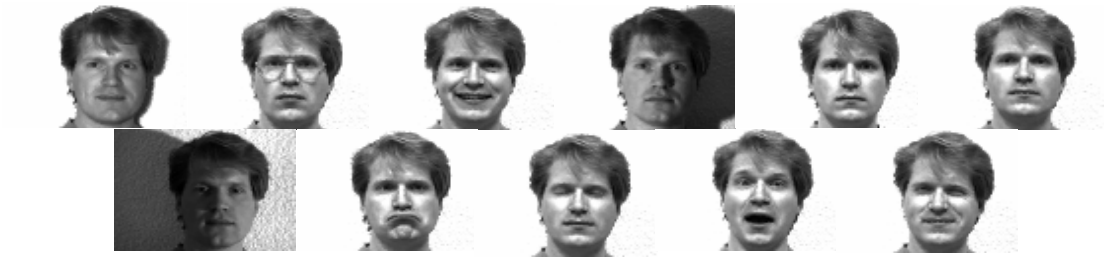

圖 2-7　各種情緒及光照的影像

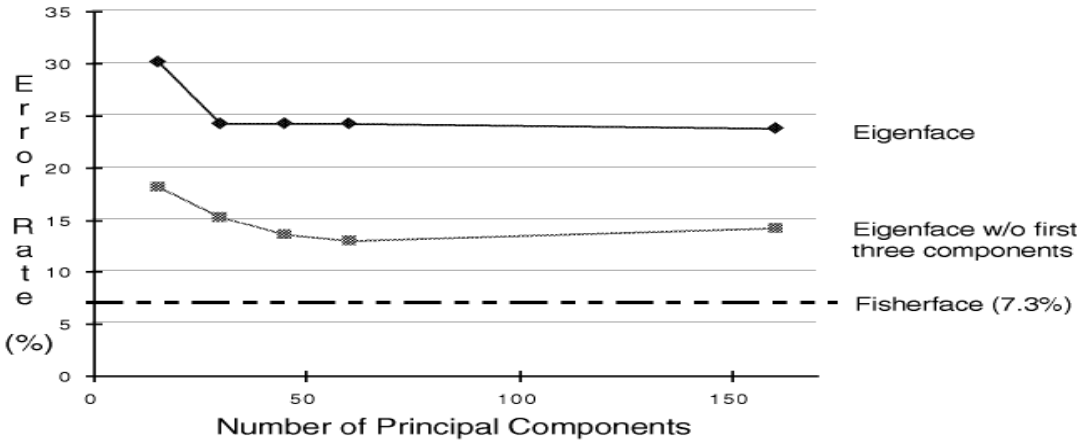

圖 2-8 各種方法的辨識結果

Jing et al.【12】則利用離散餘弦轉換（Discrete Cosine Transform，DCT）將原始影像（如圖 2-9）降維後（如圖 2-10），以降低運算量再加以辨識。此方法與 PCA 及 LDA 不同之處在於，DCT 是抽取影像的低頻部份作為特徵，而且不像 PCA 或 LDA 需要很多樣本來計算轉置矩陣，而是利用（2-23）式直接對每一個樣本降維。而本篇所採用的人臉資料庫為 Yale 及 ORL，均可得到不錯的辨識效果。

$$F(u,v)=\frac{1}{\sqrt{MN}}\alpha(u)\alpha(v)\sum_{x=1}^{M}\sum_{y=1}^{N}f(x,y)\times\cos[\frac{(2x+1)u\pi}{2M}]\cos[\frac{(2y+1)v\pi}{2N}] \quad (2\text{-}23)$$

其中，$\alpha(w)=\begin{cases}\frac{1}{\sqrt{2}}, & w=1\\ 1, & otherwise\end{cases}$

圖 2-9 原始影像

圖 2-10 取重要的頻帶的影像

Er et al.【13】認為一般人將原始影像抽取出特徵後，再利用最近鄰居法（Nearest-Neighbor）來辨識，像 Eigenface 及 Fisherface 都是屬於這種方式，於是本篇作者提出一個比較快速的方式，即原始影像經過 DCT 後轉成一維向量（如圖 2-11），然後再利用類神經網路（Neural Network）來作辨識，作者以 Yale 及 ORL 人臉資料庫做測試，皆得到不錯的辨識率。

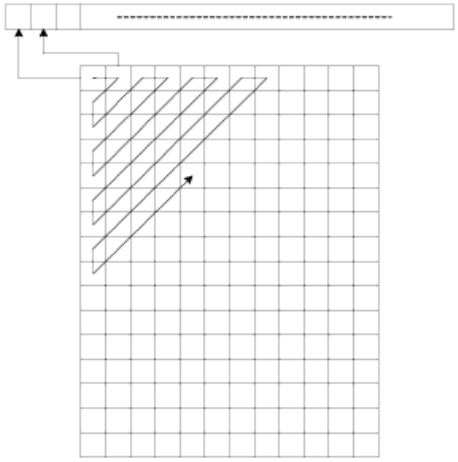

圖 2-11 二維影像經 DCT 轉成一維向量

Chen et al.【14】在處理光源照射不均勻的問題，提出了一個前處理的手法，即將欲辨識的人臉，轉成一個除掉光源資訊的格式，再做人臉辨識。此篇論文提出在對數域（logarithm domain）再進行離散餘弦轉換做處理，作者認為光源的變化主要在低頻，因此只要針對低頻做處理，即可消除光源的影響。圖 2-12 所示為離散餘弦轉換後欲捨棄的順序圖，圖 2-13a 為原始的影像，圖 2-13b 為捨棄 3 個係數的影像，圖 2-13c 為捨棄 6 個係數的影像，圖 2-13d 為捨棄 15 個係數的影像，圖 2-13e 為捨棄 20 個係數的影像，圖 2-13f 為捨棄 35 個係數的影像，圖 2-13g 為捨棄 50 個係數的影像。經實驗結果可得到不錯的辨識率。

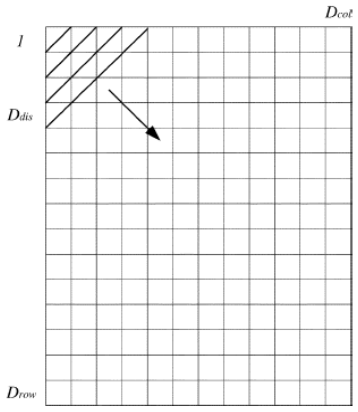

圖 2-12 離散餘弦轉換後欲捨棄的順序圖

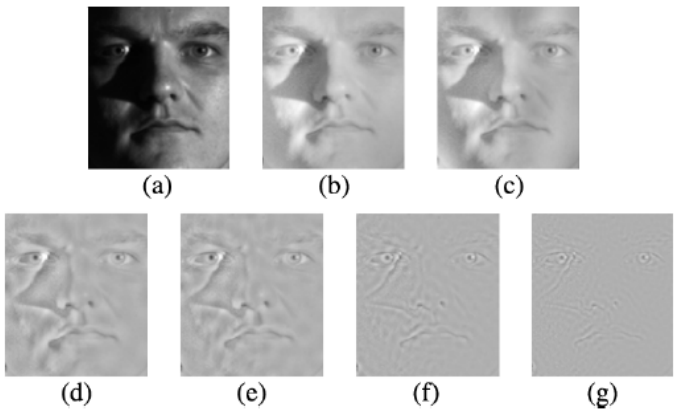

圖 2-13 捨棄係數的影像

2.3 局部特徵方法

Heisele et al【11】提出把偵測到的人臉，分別找出十個重要的特徵（如圖 2-14），然後再根據這些局部特徵做辨識，最後將這些個別的局部特徵的結果結合起來，以得到最後的辨識結果。實驗結果雖然局部特徵的方法比整體特徵的方法有更高的辨識率，但局部特徵的方法對於特徵定位的問題，在實作上會有比較高的困難度。

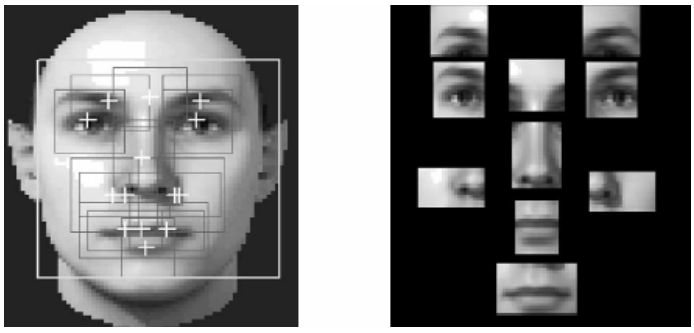

圖 2-14　局部特徵

2.4 問題與探討

在上面的討論中，我們知道特徵臉的作法可以取出一張人臉中，最具代表性的特徵，使不同影像間的差異性加大。然而這個方法也有一些缺點存在。在實做上，如果資料庫中的原始人臉影像維度很大，或是影像張數很多，則求取特徵向量的中間過程相當的複雜，日後要加入或更新資料庫中的影像時，整個特徵空間以及所有影像的特徵參數，也需要重新計算，這是一件相當費時的工作。不過在目前電腦運算速度不斷增加的狀況下，影響還不算是很大。特徵臉應用在人臉辨識上，最大不足的地方，就是當不同影像間差距被拉大的同時，屬於同一個人的影像間，距離也會被拉大。這是因為PCA轉換的過程中，僅考慮個體間的差異性，並沒有針對辨識時真正的需求，也就是類別間的差距來做最佳化的轉換。所以轉換後所得的特徵參數，可能會受到其他外在因素的影響，而非不同的人，像是光線來源所造成的照度不同，或是人臉上的不同表情等。對於PCA轉換的作法而言，這些差異的影響程度，往往比不同人所造成的差異還要大。也因此所得到的特徵參數，僅能區別出各張不同的影像，而不具備辨識出不同人的能力。

第三章 系統架構

3.1 系統簡介

一般生物特徵辨識的工作，我們可以大略將之分為三大部分：前置處理（Preprocessing）、訓練（Training）及辨識（Recognition）。以下就是對這三大部分的說明：

前置處理：這個部分主要是將所有可能影響辨識結果的變因做某個程度的去除。例如正規化，就是把剛收集到的雜亂無章的資料，整理成比較有規則的情況，譬如目標物在不同圖片的位置都不一樣，我們就要找出正確的位置並將其擷取出來；又或者是每張影像被拍攝的時間地點可能不相同，光線與拍攝的角度等也可能不相同，前處理可以強化這些對於辨識有幫助的特徵且能夠在訓練工作執行時更容易把我們所要的特徵取出來，因此使辨識的效果更好。同時因為把有用的部分擷取出來，所以也可以達到資料的降維。

訓練：這個部分主要是取出可以作為辨識依據的特徵，並將所有需要的特徵加以量化以便處理；以人臉為例，需要的特徵可以是每一個像素（pixel）的灰階值（0～255），也可以是人臉的輪廓以及五官相對的位置。由得到的特徵來當作訓練的資料（training data），以得到所有類別（class）的特有資訊，這些資訊即是用參數描述所有類別分佈的情況（例如：假設有 n 個訓練資料，把這些訓練資料訓練為 k 個類別，其中通常 $n > k$ ）。

也因為是用參數來描述每一個類別，在辨識的時候，就可以利用一些較具有物理意義的方法，把要分類的資料分到適合的類別。人腦做辨識的依據就是「抽象的感覺」之間的關係，簡單的說，就是用具體的參數來描述抽象的感覺，使電腦能夠根據這些具體的參數做分析及運算，達到辨識的目的。

辨識：在第二部分時已經得到所有類別各自的參數，當測試資料（testing data）進入辨識系統後，可以利用演算法來估算測試資料是否與某個類別最為接近（最簡單的就是計算測試資料與每一個訓練類別之間的距離），就分到哪個類別。有的辨識系統當遇到較為模糊不清（ambiguous）的測試資料時，由於辨識的結果有較高的錯誤機率，也就是擁有較低的信賴度（confidence），系統則拒絕此次分類的要求，以儘量降低錯誤發生的機率及發生錯誤所造成的影響。

3.2 辨識流程

整個人臉辨識系統的處理流程如圖 3-1 所示，當人臉資料輸入之後，將分為三大部份作處理。

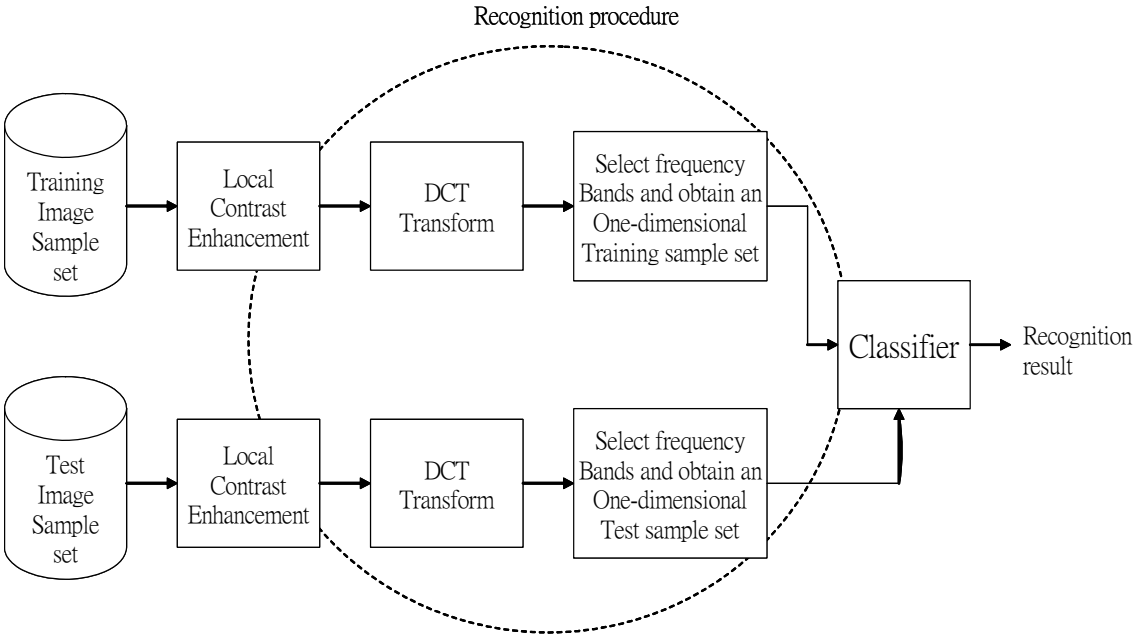

圖 3-1　人臉辨識系統流程圖

本論文所提出的辨識系統中，首先利用區域對比增強（local contrast enhancement）的方式，來處理每張影像被拍攝的光線不均勻的問題，而經過此處理後，可將影像的細節保存下來並將這些可能影響辨識結果的變因做某個程度的去除，之後再利用離散餘弦轉換擷取特徵，以進行下一步的辨識。

在辨識處理的程序中，通常要經過前置處理的步驟，並產生很多相關的數據，但是如何快速的處理這些數據完成判斷，此時就需要進行數據的分類；如果要確認所定的判斷規則或者是所擷取的特徵條件，是否合乎研究所預期的處理結果，一般的做法會先擁有一群用以建立系統的樣本，這些樣本通常稱為訓練樣本(training sample)，訓練樣本進入系統後，會改以各種型態的特徵表現出來，然後根據這些樣本建立某種規則，藉以將來對其它的樣本進行判斷分析，而這樣的過程便稱為分類(classification)，最後根據某些測試樣本(test sample)對這個系統進行測試，利用測試出來的辨識

反覆改進這個分類規則以得到最佳的分類效果，要進行人臉辨識處理程序就必須仰賴分類器(classifier)的幫助，而分類器種類眾多而且用途也不同，本研究採用支持向量機，這個是近年來被廣泛運用在分類問題上的數學工具，其具備參數較易調整、學習收斂速度快速及處理快速等優點。

第四章 人臉辨識演算法

4.1 影像的前置處理

4.1.1 人臉正規化

人臉的五官會隨著體形的高矮胖瘦不同而有所差異，即使是相同身材的人，其臉孔五官往往也有明顯的差異，而這些差異也就是辨識身份的個人重要特徵。因此，為了要達到提高辨識率的目標，除了要有強健的辨識演算法之外，我們可將收集到的雜亂無章的資料，整理成比較有規則的情況，例如目標物在不同圖片的位置都不一樣，我們可找出正確的位置並將其擷取出來，用來減少輸入影像資料間的差異性，以得到標準的影像形式。以 Yale_B 人臉資料庫為例，我們將影像切成 88 × 112 的大小，進行人臉辨識。

4.1.2 色彩空間轉換

由於影像的色彩很容易受到光源強弱的影響而有所變化，所以直接在 RGB 色彩空間做處理時，會產生嚴重的誤差，因此我們採用可對光源處理的 YCbCr 色彩空間。

4.1.2.1 RGB

RGB 為彩色影像的三原色（Primary Color），R 代表紅色光成份、G 代表綠色光成份、B 代表藍色光成份。而一般所謂的全彩色是由 R、G、B 三種不同成份所組成，其全彩影像為 24 位元（若是灰階影像為 8 位元），也就是 R、G、B 分別佔 8 位元，其值域範圍為 0-255。以錐狀體為例，來說明三種顏色對光線的靈敏度，其主要負責眼睛彩色視覺的感應器，人眼中大約有 600-700 萬個錐形體，其中 65%可以感應紅光，33 %可以感應到綠光，2%可以感應藍光；可見光的波長範圍為：400-700nm，紅光的波長範圍為：450-700nm，綠光的波長範圍為：430-680nm，藍光的波長範圍為：310-560nm；在可見光與 R、G、B 波長範圍對人眼的感應，得知人眼對藍光比較不靈敏。

4.1.2.2 YCbCr

YCbCr 色彩空間用於數位視訊中，其中 Y 代表亮度成份，Cb 代表是藍色成份與參考值的差距、Cr 代表是紅色成份與參考值的差距。如（4-1）式為 RGB 與 YCbCr 色彩空間轉換公式。

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.1687 & -0.3313 & 0.5 \\ 0.5 & -0.4187 & -0.0813 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix} \tag{4-1}$$

4.1.3 區域性對比增強

由於每張影像被拍攝的光線常常是不均勻的，而眼睛對亮度相對之間的感覺較亮度之間的絕對差值有較大的靈敏度，與真實的亮度強度絕對值沒有關聯。若光以亮度來看，並不是整張影像都一樣，而是根據亮度強度的變化會群聚成小塊區域來分佈。如圖 4-1 所示，我們對影像以區域性的 contrast enhancement，可以增強區域性的影像對比，若以 5 × 5 像素為一個區塊，計算此區塊的平均亮度（*Lavg*）如（4-2）式，為了避免（4-3）式的分母為 0，若是小於θ則將*value*設為 0。否則可以利用（4-3）式算出平均值與中心點（*Li*）的差異，為了不讓亮暗差異過大，可以選在 logarithm domain 下做處理，最後將結果正規化回來之後，可將影像的細節保存下來。因此前處理可以將這些可能影響辨識結果的變因做某個程度的去除，可以強化對於辨識有幫助的特徵且能夠在訓練工作執行時更容易把我們所要的特徵取出來，因此可以使辨識的效果更好。

$$Lavg = \sum_{i=1}^{5}\sum_{j=1}^{5} L_{ij} \tag{4-2}$$

$$value = \log(Li / Lavg) \tag{4-3}$$

其中 *Li* 為中心點的亮度，*Lavg* 為 5×5 區塊的平均亮度

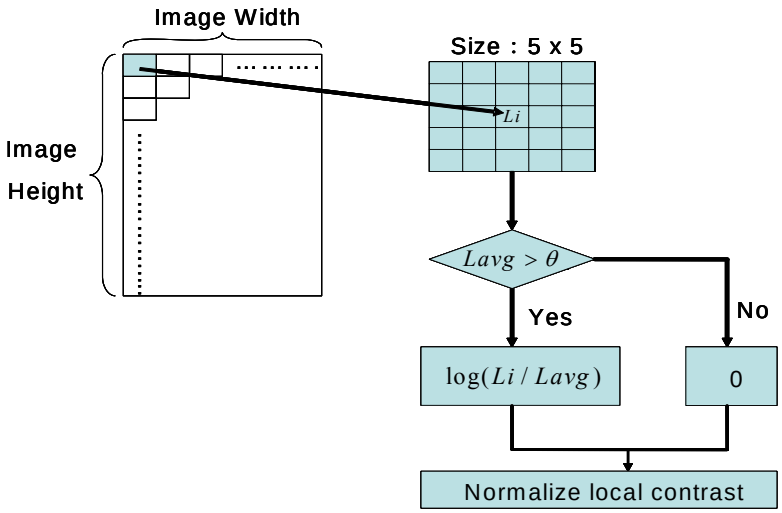

圖 4-1 區域對比增強流程示意圖

4.2 人臉特徵抽取

4.2.1 DCT 簡介

離散餘弦轉換（Discrete Cosine Transform，DCT）是將影像資料經過數學的運算，從空間域轉換成頻率域的表示方式，經過轉換後的訊號能量會比原先的訊號更加集中，而大部份的能量可以集中在某些係數上，因此具有高度緊束能量的特性，因此常被使用在人臉辨識特徵抽取，以減少資料量。而一個 M × N 二維 DCT 轉換公式如（4-4）式所示：

$$F(u,v)=\frac{1}{\sqrt{MN}}\alpha(u)\alpha(v)\sum_{x=1}^{M}\sum_{y=1}^{N}f(x,y)\times\cos[\frac{(2x+1)u\pi}{2M}]\cos[\frac{(2y+1)v\pi}{2N}] \quad (4\text{-}4)$$

其中，$\alpha(w)=\begin{cases}\frac{1}{\sqrt{2}}, & w=1\\ 1, & otherwise\end{cases}$

故 DCT 轉換是將區塊內的資料依空間頻率來分解，再將相同頻率部份

相加，因而形成一個矩陣，矩陣內的係數為其相對位置所代表頻率的振幅，左上角為低頻係數，越往右下角頻率越高，其重要性越低，最左上角由於其水平頻率和垂直頻率皆為 0，故稱為直流項（DC Component），其餘的稱為交流項（AC Component），如圖 4-2 所示。

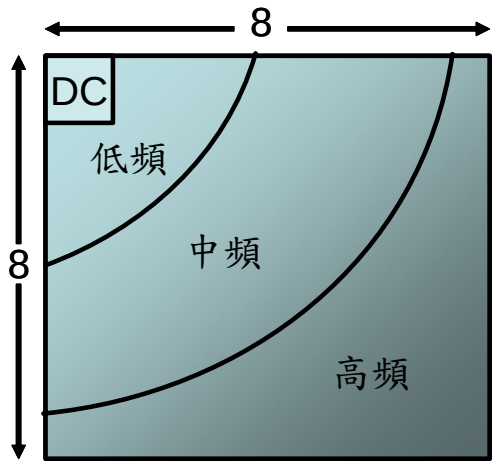

圖 4-2 二維影像頻率分佈圖

4.2.2 特徵抽取與統計分析

我們定義一個辨認特徵抽取的範圍，以 8 × 8 當作一個 Block，對整張影像作特徵抽取，每一個 Block 將會被取出並以 YCbCr 4:2:0 的格式的 block 進行離散餘弦轉換。經過離散餘弦轉換後，我們得到每個的 Block 的 DC 與 AC 係數，如圖 4-3 所示。

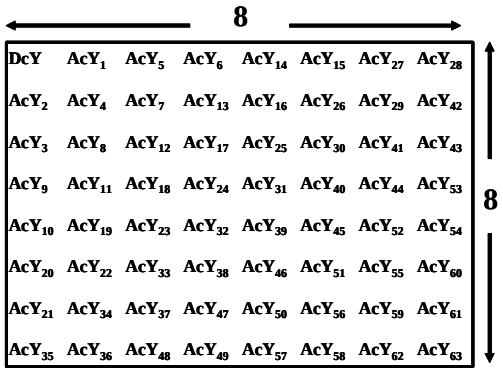

圖 4-3 zigzag scan 示意圖

Zigzag scan 後，系統取出亮度 Y 的 DC 與 AC 值作為辨認的特徵。由於越往後的 AC 值代表的是越高頻的成分，跟前面的 DC 與 AC 值比起來更像是雜訊，不適合用來當人臉辨認的特徵。因此為了降低計算量與辨認效能起見，本系統在一個亮度的 block 取出 DC 值與前三個 AC 值，共四個特徵。假設人臉區域共有 154 個 MB ，則總共可以取出 616 個特徵值。將這些特徵值串成辨認的特徵向量 F，如圖 4-4。其中 f_i 表示特徵值，i 表示維度，一個人臉的最原始被抽出的特徵向量長度為 616 維。

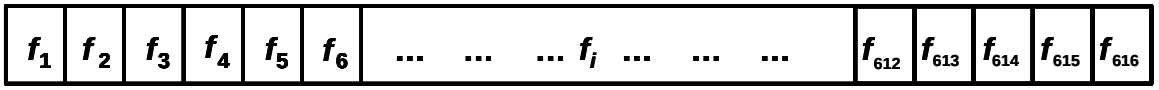

圖 4-4 人臉辨認特徵向量示意圖

接下來是特徵的統計分析，由於 SVM 的長處在於將兩個類別的特徵映射到高維度後，線性切割以辨認出兩個類別。所以我們先計算每個類別每個特徵維度的平均值，共有 10 人。

統計出第 1 類別之第 1 個維度的平均值 μ_1^1 ，統計出第 1 類別之第 2 個維度的平均值 μ_1^2 ，類推到第 1 類第 616 維的平均值 μ_1^{616} 。統計出第 2 類別

之第 1 個維度的平均值 μ_2^1，統計出第 2 類別之第 2 個維度的平均值 μ_2^2，類推到第 2 類第 616 維的平均值 μ_2^{616}。同理類推到第 10 類的第 616 維的平均值 μ_{10}^{616}。其中 μ_j^i 的 μ 代表平均值，i 代表維度，j 代表類別。

計算完平均值後接著計算每個類別之特徵維度的變異數。統計出第 1 類別之第 1 個維度的變異數 ν_1^1，統計出第 1 類別之第 2 個維度的變異數 ν_1^2，類推到第 1 類第 616 維的變異數 ν_1^{616}。統計出第 2 類別之第 1 個維度的變異數 ν_2^1，統計出第 2 類別之第 2 個維度的變異數 ν_2^2，類推到第 2 類第 616 維的變異數 ν_2^{616}。同理類推到第 10 類的第 616 維的變異數 ν_{10}^{616}。其中 ν_j^i 的 ν 代表變異數，i 代表維度，j 代表類別。ν 代表變異數 (variations)。

當我們有了平均值與變異數後，計算同一個維度對不同類別之間的差異。利用算出的平均值來計算差異值。計算出來的差異值越大表示兩類間差別越大。$d_{m,n}^k$ 表示平均值相減的平方。我們會得到(10×9) / 2 = 45 個不同兩類間差異的 d 向量，如圖 3-5 所示，其中 k 表示所屬的特徵維度，下標的兩個數字 m、n 表示所要分類的兩個類別。其中 d 的計算方式由（4-5）式得到：

$$\begin{aligned} &d_{m,n}^k = (\mu_m^k - \mu_n^k)^2, \text{for } k = 1,...616 \\ &1 \le m,n \le 10, m \ne n \end{aligned} \tag{1-5}$$

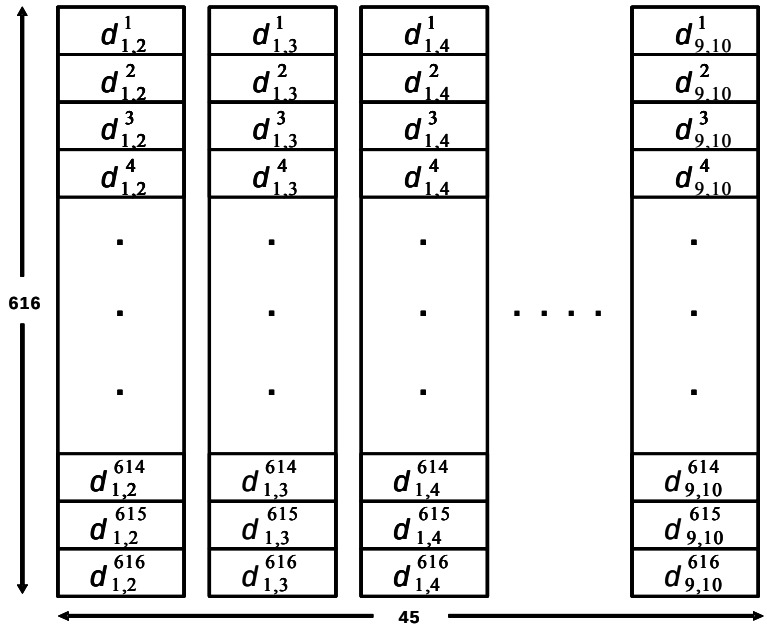

圖 4-5 兩兩之間平均值差異的平方的向量示意圖

計算完$d_{m,n}^{k}$後，就可以利用剛剛產生的$d_{m,n}^{k}$與之前計算的變異數來計算可以用來區別兩類的特徵區別值ω。計算方法如（4-6）式。根據公式來看，若是同類之間的變異數越小，表示兩類之間自己與自己類越相像。而兩類間的$d_{m,n}^{k}$越大就表示這兩類越不像。因此，若是所計算出來的 ω 值越大，表示這個值越可以區分出這兩類，這個特徵越重要。

$$\omega_{m,n}^{k} = d_{m,n}^{k} / (v_{m}^{k} + v_{n}^{k}), \text{for } k = 1,...616 \\ 1 \le m,n \le 10, m \ne n \tag{4-6}$$

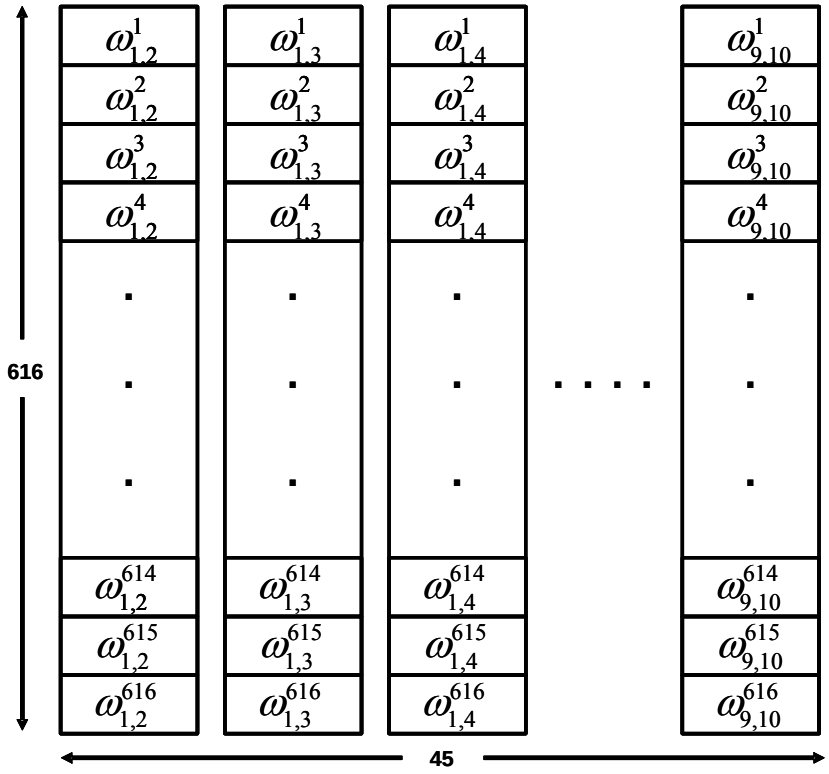

圖 4-6 兩兩之間特徵區別值的向量示意圖

計算完兩類與兩類之間的特徵區別值 ω 後，我們就可以知道針對不同的兩個類別，哪一個特徵維度的特徵值是對於分辨這兩類有幫助的，值越大表示對於分類越有利。知道這個資訊後我們就做 Sort 排序，經過排序後我們就可以知道，使用哪幾個特徵維度可以最佳辨識兩類所對應的兩類。由於是兩類與兩類之間的辨識，針對不同的兩個類別取相對應此兩類的最佳分類特徵維度的做法就會比一般將所有類別一起訓練出一個分類模型的分類效果來的好，而且可以用到更少的特徵。取多少特徵維度來辨認則是取決於當初訓練時最佳的辨識率是落在哪一個維度。若都很好時則取最少維度以減低計算量。下表 4-1 為用來分辨類別所選用的特徵維度，C_{i-j} 表示這幾個特徵是用來分辨第 i 與第 j 類，而後面所顯示的為前 12 大的特徵維度。

表 4-1 分類所選用的維度

C1-2	534	528	440	446	487	530	439	442	444	352	443	128
C1-3	396	528	531	573	440	443	352	485	534	7	441	4
C1-4	534	532	308	440	352	396	535	572	573	575	7	444
C1-5	444	440	614	56	352	54	439	488	68	144	72	84
C1-6	440	441	572	56	397	531	39	532	352	72	573	534
C1-7	572	573	8	446	575	308	52	312	12	444	43	574
C1-8	534	84	573	48	6	40	216	308	532	56	304	572
C1-9	442	164	4	535	573	108	572	534	440	300	8	112
C1-10	573	572	534	575	439	396	484	263	532	308	262	488
C2-3	534	487	443	442	311	440	91	4	72	310	528	444
C2-4	534	532	308	487	528	536	91	352	128	4	490	84
C2-5	91	444	487	530	128	534	88	94	172	54	528	265
C2-6	444	487	6	532	534	528	446	442	443	91	48	440
C2-7	446	528	312	308	487	48	534	4	91	529	530	442
C2-8	534	487	532	530	528	91	443	442	308	490	440	445
C2-9	442	528	4	534	221	48	533	532	487	6	530	8
C2-10	442	128	490	534	443	532	487	172	530	533	446	528
C3-4	440	308	396	352	531	311	443	354	72	24	578	444
C3-5	444	440	72	488	94	56	60	44	148	188	614	531
C3-6	56	72	60	140	64	310	39	443	6	531	68	168
C3-7	12	8	52	116	311	488	120	528	308	2	493	572
C3-8	6	60	72	124	100	84	56	184	48	574	64	24
C3-9	442	164	212	116	112	72	256	108	64	184	104	535
C3-10	488	311	531	572	485	396	575	223	72	60	574	4
C4-5	614	188	144	493	148	94	537	152	536	308	444	84
C4-6	308	140	440	396	56	444	39	580	352	184	6	60
C4-7	536	446	493	48	8	92	492	12	440	52	396	132
C4-8	84	48	396	440	184	6	224	140	355	260	100	124
C4-9	164	184	112	140	442	144	8	224	256	108	212	228
C4-10	308	439	398	443	442	612	440	573	4	352	444	41
C5-6	444	440	39	11	614	148	12	44	36	35	493	45
C5-7	52	614	440	446	44	89	95	572	188	348	48	308
C5-8	440	614	12	5	54	8	89	172	36	444	216	92
C5-9	33	44	614	91	176	45	440	133	125	132	567	493
C5-10	444	78	108	144	88	89	148	116	44	188	160	54
C6-7	12	308	446	8	444	312	19	493	16	39	11	56
C6-8	6	48	308	5	216	260	137	39	3	89	7	0
C6-9	442	220	8	35	39	176	11	19	6	7	3	444
C6-10	39	56	6	308	11	72	3	396	116	120	140	0
C7-8	48	6	92	12	52	8	446	124	16	5	172	260
C7-9	442	164	446	92	132	212	184	176	493	256	133	124
C7-10	116	92	446	52	304	12	536	348	444	308	8	120
C8-9	6	48	442	164	8	5	84	4	40	32	344	212
C8-10	216	84	260	40	128	124	48	172	304	8	100	80
C9-10	164	112	116	108	144	33	4	184	212	104	120	220

4.3 使用 SVM 的人臉辨認系統

4.3.1 SVM 簡介

支持向量機(Support Vector Machine，簡稱 SVM)是 1998 年以統計理論為基礎所提出的機器學習理論，有別於傳統的類神經網路的經驗之風險最小化(Empirical Risk Minimization Principle, ERM)，而是一種結構風險最小化原理(Structural Risk Minimization Principle, SRM) 的統計學習理論，用於分類與迴歸的問題。SRM 使 VC(Vapnik Cherovnenkis) 維度數上限最小化，這使得 SVM 的方法比使用類神經網路具有更好的泛化能力。此外，SVM 的一般化錯誤與其它機器學習的方法不同在於：SVM 的一般化錯誤與資料的維度無關，而是視不同類別資料之間分離的程度而定，因此只要資料的分離程度越好，則 SVM 的分類效果也越好，這也證明了 SVM 在處理大量資料時，能達到很好的分辨能力。

SVM 的分辨效果與資料間的分離程度有很大的關係，而 SVM 是利用最佳分類平面（optimal separating hyperplane）將資料分成兩個部份，因此最佳分類平面可以用（4-7）式來表示，而此時所有的分類就可以用（4-8）式來表示。

$$w \cdot x + b = 0 \tag{4-7}$$

$$y_i(w \cdot x_i + b) \geq 1 \quad i = 1,2,...,N \tag{4-8}$$

其中 w 表示權重向量 weight vector，維度與資料的維度相同，b 表示偏移量

（bias），偏移量的目的是使最佳分類面水平移動後能落在空間中的正確位置，因此偏移量是在權重向量訓練完成後才決定。SVM 應用在分類時，是把最佳分類面視為一個決策函數（decision function）如（4-9）式：

$$f(x) = \text{sgn}(w \cdot x + b) \tag{4-9}$$

當一個未知的資料 x 代入此決策函數後，所得到只有{+1,-1}兩種結果，若以+1 代表資料屬於正的一類，-1 代表資料屬於負的一類，因此所有的資料均可被分為“正”與“負”兩個類別。

分類面之參數是由所有的訓練資料決定，而測試資料是藉由最佳分類面來將區隔分成兩部分，因此可以輕易的被分類。在 SVM 的定義中，一個好的分類面必須使得兩個類別分得越開越好，如圖 4-9 所示。

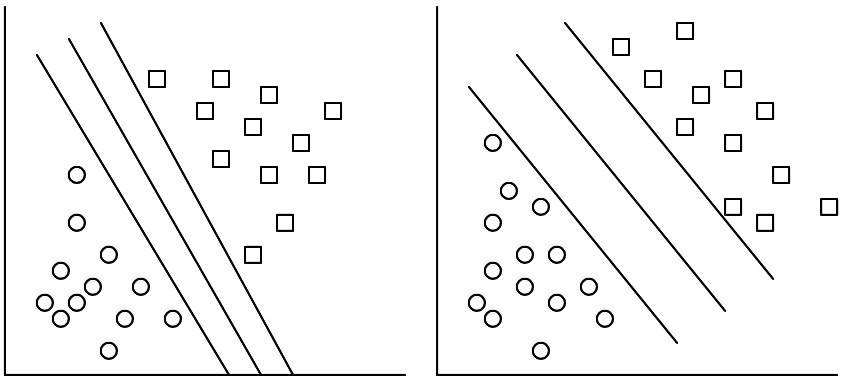

圖 4-9 分類面示意圖

左圖與右圖中的實線都能夠將兩類分開，但因右圖能夠把兩群資料分的最開，因此可說是最佳的分類面，而正負兩類最接近分類面的資料點被稱為支持向量（Support Vectors），因此這些支持向量決定了分類面。正負兩個類別與分類面間的最短矩離稱為間距（Margin），間距的大小相當於不

同類別資料間分離的程度，因此 SVM 是一種間距最大化（Margin Maximization）的分類器。

SVM 使用超平面做分類，但是實際上資料分佈的情況有很多，因此我們大致上把 SVM 分為三種不同的情形來介紹。

- 線性可分離（Linearly Separable）
- 線性不可分離（Linearly Non-Separable ）
- 非線性可分離（Nonlinearly Separable）

4.3.1.1 線性可分離

線性可分離是假設分類面可以將兩類的資料完全分開，即不同類別並沒有互相重疊的部份，此距離可以用（4-10）式來表示：

$$d(x,w,b)=\frac{w\cdot x+b}{\|w\|} \tag{4-10}$$

一般來說，要符合線性可分離的情況，通常是假設所有的訓練資料與分類面的距離不會小於 1，因此要使分類間距最大化，則必須使分母的$\|w\|$最小化，根據 Lagrange 理論引入 Lagrange 乘子 Lagrange multiplier α_i，和 Lagrange 函數$L_d(\alpha)$和 Kuhn-Tucker 條件的補充可得到（4-11）式：

$$L_d(\alpha)=\sum_{i=1}^{N}\alpha_i-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j(x_i\cdot x_j) \tag{4-11}$$

在約束條件（constraint）$\sum_{i=1}^{N}\alpha_i y_i = 0$ 且 $\alpha_i \geq 0$ 之下，滿足條件的輸入向量便稱為支持向量(Support Vector)，經由訓練後便可得到 Lagrange 乘子與偏移量。

4.3.1.2 線性不可分離

線性可分離是描述所有的資料都沒有落在間距內的情形，但在實際上，資料有可能會落在空間上的任何一個位置，因此很難使兩群不同的類別的資料完全分離，所以加入了間隔鬆弛變量（Margin Slack Variable）$\xi_i \quad i = 1,2,...,N$，目的是加入 slack variable 後，能使得所有的資料都能落在間距之外，所以 slack variable 可視為資料的分佈狀況，當 slack variable 的總和越小，代表資料越符合 SVM 的準則；當 slack variable 的總和越大，代表我們需要對資料作越大的調整，才能使所有的資料都符合分離的情況，因此 slack variable 的總和越小越好。以物理意義來說，當資料分佈與正確位置間的錯誤小於 slack variable 時，則原本的式子可表示成（4-12）式，而最佳分類面可表示成（4-13）式。

$$y_i(w \cdot x_i + b) \geq 1 - \xi_i, i = 1,2,...,n \tag{4-12}$$

$$j(w,\xi_i) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\xi_i \tag{4-13}$$

其中 C 是一個衡量訓練錯誤，代表當 C 越大時，SVM 的模型對於錯誤就越敏感，反之當 C 越小時，間距最大化的重要性也就越大，所以 Lagrange

函數為成（4-14）式：

$$L_d(\alpha)=\sum_{i=1}^{N}\alpha_i-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j(x_i\cdot x_j)\quad 0\le\alpha_i\le C,\ \sum_{i=1}^{N}\alpha_i y_i=0 \tag{4-14}$$

求解 Lagrage 函數（Lagrangian）$L_d(\alpha)$之最大值就可以得到最佳分類平面（Optimal Separating Hyperplane）。

4.3.1.3 非線性可分離

然而有很多情況，SVM的線性分離限制對於原始維度的資料而言是過於嚴格的，因此若當兩類資料重疊的情形很嚴重，以致不能很容易用最佳分類面將資料分類，所以在尋找最佳分類面之參數前，可以先將資料利用 ϕ 函數(4-15)式，將原本的空間投影到某個高維的空間：

$$x\rightarrow\phi(x) \tag{4-15}$$

此高維空間稱做特徵空間，所有原始的資料投影到特徵空間後（如圖4-10），能改善原本不同類別的資料分佈重疊的情況，因此能將類別與類別間分離，然後再取得最佳化平面的最佳化參數值。

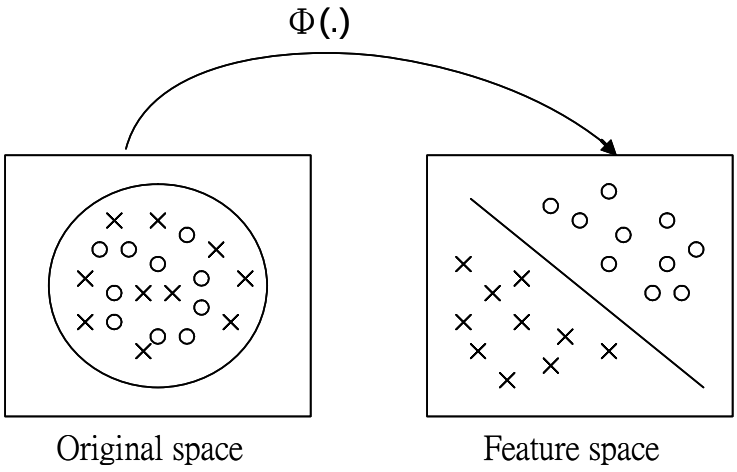

圖 4-10 kernel function 示意圖

而這些核函數比較著名的有下列公式(4-16)、(4-17)、(4-18)、(4-19)：

Linear kernel function： $K(x_i, x_j) = x_i \cdot x_j$ (4-16)

Polynomial kernel function： $K(x_i, x_j) = (\gamma x_i \cdot x_j + coef)^d$ (4-17)

RFB kernel function： $K\left(\overrightarrow{x_i}, \overrightarrow{x_j}\right) = \exp\left(-\gamma\left(\overrightarrow{x_i} - \overrightarrow{x_j}\right)^2\right)$ (4-18)

Signoid kernel function： $K\left(\overrightarrow{x_i}, \overrightarrow{x_j}\right) = \tanh\left(\gamma \overrightarrow{x_i} \cdot \overrightarrow{x_j} + coef\right)$ (4-19)

其中的 d、γ 與 $coef$ 皆為常數參數值。

SVM 理論只考慮高維特徵空間的內積運算 $K\left(\overrightarrow{x_i}, \overrightarrow{x_j}\right) = \Phi\left(\overrightarrow{x_i}\right) \cdot \Phi\left(\overrightarrow{x_j}\right)$，而不直接使用函數 Φ，從而巧妙地解決了因 Φ 未知而 W 無法顯示表達的問題，稱 $K\left(\overrightarrow{x_i}, \overrightarrow{x_j}\right)$ 為核函數。經由文獻証明，只要滿足 Mercer 條件的對稱函

數即可作為核函數。

對於兩類問題分類，存在線性可分和線性不可分的兩種支持向量機。但是在實際中，為了將兩類模式盡可能分類開來，一般都要構造非線性可分的分類器。然而一個複雜的模式識別分類問題，在高維空間比低維空間更容易線性可分，支持向量機就是首先透過核函數把訓練樣本中的低維數據映射到高維特徵空間，然後在高維特徵空間構造一個最佳分類平面。由於構造的核函數滿足 Mercer 條件，所以在訓練中只需考慮核函數 K，而不必明確知道映射函數 Φ 。

從此可以看出：當我們樣本空間通過非線性映射映入特徵空間時，如果只用映射內積，則可以用相對應的核函數來代替，而不需要知道映射的顯示表述式。這是從線性支持向量機到非線性支持向量機的關鍵一步。
在特徵空間 F 中應用線性支援向量機的方法，分類決策函數式變為（4-20）式：

$$y = Sgn(w * \Phi(x) + b) = Sgn(\sum_{sv} \alpha_i y_i (\Phi(x)\Phi(x)) + b) \tag{4-20}$$

這就是非線性支援向量學習機的最終分類決策函數。雖然用到了特徵空間及非線性映射，但實際計算中並不需要知道他們的顯示表述。只需求出支持向量及其支持的 α 和 b 值，通過核函數的計算，即可得到原來樣本空間的非線性輸出值，如（4-21）式與（4-22）式。

$$y(x) = \sum_{i=1}^{\infty} \lambda_i \psi_i \Phi_i(x) + b = \sum_{j=1}^{n} \alpha_j y_j K(x, x_j) + b \tag{4-21}$$

$$\Psi = \sum_{i=1}^{n} \alpha_i y_i \Phi(x_i) = (\psi_1, \psi_2, \cdots, \psi_i, \cdots) \tag{4-22}$$

4.3.2 人臉辨認

最後我們將前一節所抽取出的特徵向量放入支持向量機的分類器中訓練，而預測辨認的部分就根據上一節所介紹的（4-21）式來做。計算（4-21）式後，判斷 y 值的正負就可得知分類結果。

整個人臉辨認的流程如圖 4-11。我們使用的辨認機制為系統輸入未知的向量後，每個未知的向量就根據當初訓練分類的 hyper-plane 使用的特徵維度取用對應此維度的特徵值。例如輸入的未知向量要分辨第一與第二類時，就取第一與第二類的 hyper-plane 辨認以表 4-1 當初訓練出 C1-2 hyper-plane 的所使用的那些特徵維度的值為輸入的特徵值，然後這些值經過之前訓練的模型與支援向量的運算後，最後系統輸出為第一類或第二類。

接著我們使用一個投票機制來做為系統第二級的辨認。例如：系統在經過第一類與第二類的分類後會決定出分類結果，比較第一與第二類系統覺得此未知向量應該是第一類則我們認為系統投給第一類一票，將第一類的堆疊票數加一票。同理比較第一與第三類後就將系統輸出的那個類別的票數加一票，最後則經過 45 個 hyper-plane 的比較後，第一到第十類中都會累積一些 SVM 比較這 45 個 hyper-plane 後所投出的票。在第一到十類中，累計票數最高的那個就是系統最後輸出的辨認結果。

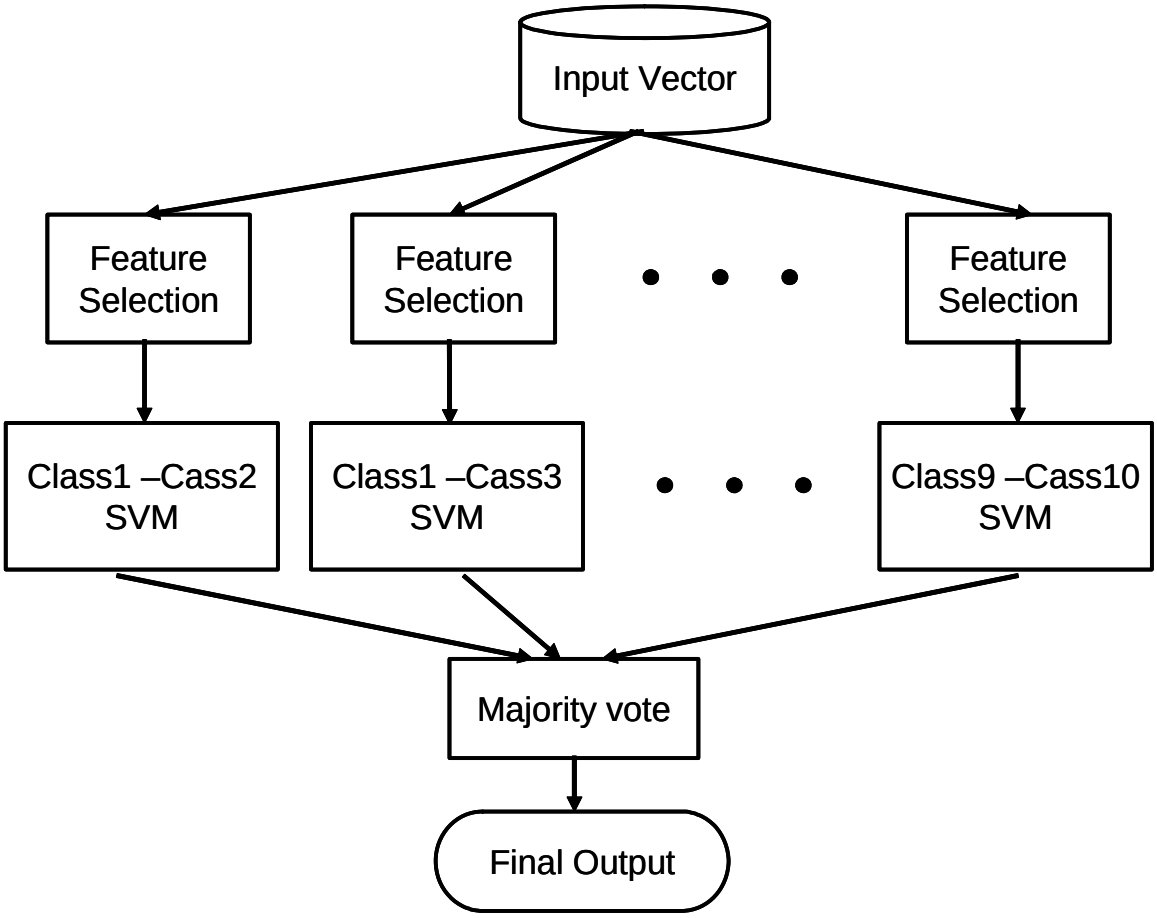

圖 4-11 人臉辨認流程圖

第五章 實驗結果

本章對實驗相關設定及人臉資料庫所基本的介紹，並且就實驗的結果加以探討。

5.1 人臉資料庫

本論文以 SVM 之人臉在不同光源下辨識為研究主題，所使用的實驗資料是廣為被應用且具有高度的公信力之人臉資料庫，以下對人臉資料庫做大略的介紹。

5.1.1 Yale 人臉資料庫

這個人臉資料庫是由耶魯大學所收集的。可免費使用於學術用途上。資料庫中有十五個不同的人，每人十一張照片，共有一百六十五張影像。每個人的十一張影像各有不同的情緒與光照條件，分別有中央打光、戴眼鏡與不戴眼鏡、快樂表情、左邊打光、正常表情、右邊打光、難過表情、想睡覺的表情、驚訝表情與眨眼表情。如果需要有關此人臉資料庫的進一步資料，請參考網頁 http://cvc.yale.edu/projects/yalefaces/yalefaces.html

5.1.2 Yale_B 人臉資料庫

這個資料庫是由耶魯大學所收集的。可免費使用於學術用途上。資料庫中有十個不同的人，每人有九種姿勢，並在不同角度的光源下做變化，所以每人有四百零五張，總共有四千五十張，

5.2 實驗結果

5.2.1 Yale 實驗結果

根據第四章人臉辨識演算法經由 Adaptive Feature Selection 的 SVM 辨識系統的處理，針對 Yale Database 的人臉資料庫做統計分析，此人臉資料庫的樣本數有十五個人共一百六十五張。表 5-1 為每個人選出三張（共 45 張）來訓練，其餘的 120 張影像來測試所得到的結果，下表 5-2 為 Adaptive Feature Selection 的統計圖，表 5-3 為 Adaptive Feature Selection 的支持向量統計圖。如圖 5-1 所示，未經區域對比增強的影像，其辨識率為 93.33%。而經過區對比增強之後的影像（如圖 5-2），其辨識率為 98.33%。

圖 5-1 未經區域對比增強

圖 5-2 經過區域對比增強的結果

表 5-1 人臉辨識的結果

The Number of Feature values	Number of the Misrecognized Images		Recognition Rate (%)	
	Without LCE	With LCE	Without LCE	With LCE
8	21	15	82.5	87.5
10	22	14	81.67	88.33
12	19	9	84.17	92.5
14	19	9	84.17	92.5
16	16	9	86.67	92.5
18	15	11	87.5	90.83
20	15	5	87.5	95.833
22	15	9	87.5	92.5
24	13	8	89.17	93.33
26	13	7	89.17	94.17
28	17	8	85.83	93.33
30	16	6	86.67	95
32	12	8	90	93.33
Adaptive Feature Selection	8	2	93.33	98.33

表 5-2 特徵選取數量的統計圖

The Number of Feature values	The Number of Hyperplanes	
	Without LCE	With LCE
8	67	69
10	11	8
12	10	6
14	5	3
16	1	2
18	4	6
20	3	3
22	1	3
24	0	3
26	0	1
28	1	0
30	1	1
32	1	0
Total Number of Hyperplanes	105	105
The Average Number of Feature values	10.438	10.857

表 5-3 支持向量的統計圖

The Number of Support Vectors	Without LCE	With LCE
1	0	0
2	0	0
3	0	0
4	15	11
5	43	30
6	47	64
7	0	0
8	0	0
9	0	0
10	0	0
11	0	0
Total Number of Hyperplanes	105	105
The Average Number of Support Vecotrs	5.304	5.504

表 5-4 為每個人選出五張（共 75 張）來訓練，其餘的 90 張影像來測試所得到的結果，下表 5-5 為 Adaptive Feature Selection 的統計圖，由此表可看出大部份的 Feature Number 小於 10，而平均值為 8.5，因此可以分析出大部份的 Feature Number 都沒有很大，表示我們選用的 Feature 都不會很大。表 5-6 為 Adaptive Feature Selection 的支持向量統計圖，由此表可看出大部份的 SV Number 小於 4，而平均值為 3.18，因此可以分析出大部份的 SV Number 都沒有很大，表示大部份的 hyper-plane 可以容易分類。未經區域對比增強的影像，其辨識率為 96.67%。而經過區對比增強之後的影像，其辨識率為 100%。由此可知，經過我們建議的方法之後，其辨識率大大的提升，並且有效的解決光源變化的問題。

表 5-4 人臉辨識的結果

The Number of Feature values	Number of the Misrecognized Images		Recognition Rate (%)	
	Without LCE	With LCE	Without LCE	With LCE
8	10	3	88.89	96.67
10	11	4	87.78	95.56
12	10	4	88.89	95.56
14	10	3	88.89	96.67
16	8	2	91.11	97.78
18	8	3	91.11	96.67
20	10	1	88.89	98.89
22	10	2	88.89	97.78
24	9	1	90	98.89
26	9	1	90	98.89
28	9	1	90	98.89
30	7	1	92.22	98.89
32	7	0	92.22	100
Adaptive Feature Selection	3	0	96.67	100

表 5-5 特徵選取數量的統計圖

The Number of Feature values	The Number of Hyperplanes	
	Without LCE	With LCE
8	88	97
10	2	3
12	2	1
14	2	1
16	1	0
18	1	0
20	0	3
22	2	0
24	3	0
26	0	0
28	2	0
30	2	0
32	0	0
Total Number of Hyperplanes	105	105
The Average Number of Feature values	9.92	8.5

表 5-6 支持向量的統計圖

The Number of Support Vectors	Without LCE	With LCE
1	0	0
2	33	66
3	2	6
4	9	4
5	19	12
6	24	13
7	12	2
8	4	2
9	1	0
10	1	0
11	0	0
Total Number of Hyperplanes	105	105
The Average Number of Support Vecotrs	4.59	3.18

5.2.2 Yale_B 實驗結果

根據第四章人臉辨識演算法經由 Adaptive Feature Selection 的 SVM 辨識系統的處理，針對 Yale_B Database 的人臉資料庫做統計分析，此人臉資料庫的樣本數有十個人，每個人有九種姿勢，每個姿勢有四個 subset，分別在不同角度的光源下做變化，所以每人有四百零五張，總共有四千五十張。

首先我們將人臉資料庫經過區域對比增強，並選擇 subset1 當訓練（共 630 張），其餘的 3420 張影像來測試所得到的結果，下表 5-7 為人臉辨識的結果，由表中可知，經過 Adaptive Feature Selection 的辨識率為 98.51%，平均使用的特徵數量為 16.4，平均使用的 SV Number 為 10。表 5-8 為近年來人臉辨識的相關研究與本研究之比較。

表 5-7　人臉的辨識結果

The Number of Feature values	Number of the Misrecognized Images	Recognition Rate (%)
8	289	91.55
10	234	93.16
12	182	94.68
14	166	95.15
16	167	95.12
18	153	95.53
20	131	96.17
22	133	96.11
24	114	96.67
26	94	97.25
28	82	97.6
30	85	97.51
32	75	97.8
Adaptive Feature Selection	51	98.51

表 5-8 相關文獻與本論文之比較

Method	Recognition Rate (%)		
	subset2	subset3	subset4
Eigenface[23]	89.81	47.04	21.9
Direct-LDA[24]	98.15	70.09	30.79
Fisherface[7]	95.14	75.14	34.76
Subspace-LDA[25]	99.84	94.44	49.8
Kernel PCA[26]	92.96	49.44	22.38
KDDA[27]	99.63	89.26	38.89
GDA[28]	99.81	96.94	49.92
Choosing Parameters of Kernel Subspace LDA[29]	99.81	96.99	52.22
Our Approach	99.91	99.07	96.83

表 5-9 為從每個 subset 挑出二張（共 720 張）來訓練，其餘的 3330 張影像來測試所得到的結果，下表 5-10 為 Adaptive Feature Selection 的統計圖，由此表可看出大部份的 Feature Number 小於 12，而平均值為 10.8，因此可以分析出大部份的 Feature Number 都沒有很大，表示我們選用的 Feature 都不會很大。表 5-11 為 Adaptive Feature Selection 的支持向量統計圖，由此表可看出大部份的 SV Number 小於 10，而平均值為 9.578，因此可以分析出大部份的 SV Number 都沒有很大，表示大部份的 hyper-plane 可以容易分類。如圖 5-3 所示，未經區域對比增強的影像，其辨識率為 94.02%。而經過區域對比增強之後的影像（如圖 5-4），其辨識率為 99.13%。由此可知，經過我們建議的方法之後，其辨識率大大的提升，並且有效的解決光源變化的問題。表 5-12 為近年來人臉辨識的相關研究與本研究之比較【30】。

表 5-9 人臉辨識的結果

The Number of Feature values	Number of the Misrecognized Images		Recognition Rate (%)	
	Without LCE	With LCE	Without LCE	With LCE
8	336	101	89.91	96.97
10	300	87	90.99	97.39
12	307	62	90.78	98.14
14	281	50	91.56	98.5
16	273	41	91.8	98.77
18	271	47	91.86	98.59
20	252	49	92.43	98.53
22	253	50	92.4	98.5
24	233	51	93	98.47
26	231	45	93.06	98.65
28	223	34	93.3	98.98
30	207	34	93.78	98.98
32	230	30	93.1	99.1
Adaptive Feature Selection	199	29	94.02	99.13

表 5-10 特徵選取數量的統計圖

The Number of Feature values	The Number of Hyperplanes	
	Without LCE	With LCE
8	22	29
10	6	2
12	2	4
14	3	3
16	0	1
18	1	3
20	0	1
22	2	1
24	4	0
26	1	0
28	3	0
30	0	0
32	1	1
Total Number of Hyperplanes	45	45
The Average Number of Feature values	13.378	10.8

表 5-11 支持向量的統計圖

The Number of Support Vectors	Without LCE	With LCE
1	0	0
2	0	0
3	6	0
4	6	0
5	7	1
6	2	6
7	3	6
8	5	7
9	1	8
10	3	3
11	2	3
12	1	3
13	3	4
14	0	0
15	2	0
16	1	0
17	1	3
18	0	1
19	0	0
20	2	0
Total Number of Hyperplanes	45	45
The Average Number of Support Vecotrs	8.11	9.578

圖 5-3 未經區域對比增強

圖 5-4 經過區域對比增強

表 5-12 相關文獻與本論文之比較

Method	Recognition Rate (%)
Eigenface	31.59
Direct-LDA	47.33
Fisherface	66.97
Kernel Direct LDA	79.94
Regularized Fisher Discriminant	90.56
Our Approach	99.13

第六章結論與未來展望

6.1 結論

由於科技的日新月異，加上人們追求便利，現在有越來越多、各式各樣的安全驗證系統，而這些系統讓使用者不需要帶一大堆的卡片，或是記住一大堆的密碼，而是透過人的眼睛、指紋和聲音等方式，對使用者做身份的驗證，而其中因為人臉辨識系統不會影響到人們正常的活動，由於這個優點，便成為最常使用的一個方式。然而在辨識的過程中，常常受到環境光源改變，使得同一張人臉會有很大的不同，因而導致人臉的辨識率大幅下降。

本論文提出了一個有效解決光源不均勻照射下的影響，由最後的實驗結果顯示，我們人臉辨識結果的正確率非常的高，表示我們的系統是很可靠與所使用的投票機制與 SVM 辨識核心策略都是非常正確的。

6.2 未來展望

本研究雖然成功的解決影像光源的問題，但在人臉辨識的先天限制還是很多，例如：髮絲遮臉、戴眼鏡或是墨鏡、化妝.....等，這些先天的限制，局限了人臉辨識的廣泛應用，如何有效去除上述種種人臉的障礙，且不破

壞人臉的完整性，也是將來極待解決的重要部份。

在硬體實現部份，由於使用支援向量機做為分類器，其軟體實作時都需花較多的時間訓練，因此如果可以實現在硬體上，則計算所需的時間必將減少許多。

參考文獻

[1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, “Face Recognition: A Literature Survey,” *ACM Computing Surveys*, Vol. 35, No. 4, pp. 399-458, Dec. 2003.

[2] S. Z. Li and J. Lu, “Face Recognition Using the Nearest Feature Line Method,” *IEEE Transactions on Neural Networks*, Vol. 10, No. 2, March. 1999.

[3] R. Brunelli and T. Poggio, “Face Recognition: Features versus Templates,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.15, No.10, pp. 1042-1052, Oct. 1993.

[4] M. J. Er, S. Wu, J. Lu, and H. L. Toh, “Face Recognition With Radial Basis Function (RBF) Neural Networks,” *IEEE Transactions on Neural Networks*, Vol.13 No. 3, pp.697-710, May. 2002.

[5] M. Bicego, G. Iacono, and V. Murino, “Face Recognition with Multilevel B-Splines and Support Vector Machines,” *ACM WBMA* March. 2003.

[6] J.T. Kwok, “Moderating the outputs of support vector machine classifiers,” *IEEE Trans. on Neural Networks*, vol. 10, pp. 1018-1031, Sep. 1999.

[7] M.A. Turk and A.P. Pentland, “Face recognition using eigenfaces,” *IEEE Proceeding on Computer Vision and Pattern Recognition*, pp. 586-591, June. 1991.

[8] P.N Belhumeur, J.P. Hespanha, and D.J. Kriegman, “Eigenfaces vs. Fisherfaces: recognition using class specific linear projection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 711-720, July. 1997.

[9] L.I Smith, “A tutorial on Principal Components Analysis,” Feb. 2002.

[10] K. I. Kim, K. Jung and J. Kim, “Face Recognition Using Kernel Principal Component Analysis,” *IEEE Signal Processing Letters*, Vol. 9, pp.40-42, Feb.

2002,

[11] B. Heisele, P. Ho, J. Wu, and T. Poggio, “Face recognition: component-based versus global approaches,” *Computer Vision and Image Understanding*, pp.6–21, Feb. 2003.

[12] X.Y Jing and D. Zhang, “A face and palmprint recognition approach based on discriminant DCT feature extraction,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B,* Vol.34, pp.2405-2415, Dec. 2004.

[13] M.J. Er, W. Chen, and S. Wu, “High-Speed Face Recognition Based on Discrete Cosine Transform and RBF Neural Networks,” *IEEE TRANSACTIONS ON NEURAL NETWORKS*, Vol. 16, pp.679-691, May. 2005.

[14] W. Chen, M.J. Er and S. Wu, “Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B,* Vol.36, pp.458-466, Apr. 2006.

[15] H. Cevikalp, M. Neamtu, M. Wilkes and A. Barkana, “Discriminative common vectors for face recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 27, pp. 4 - 13, Jan. 2005.

[16] J. Yang, D. Zhang, A.F. Frangi, and J.J. Yang, “Two-Dimensional PCA: A New Approach to Representation and Recognition,” *IEEE Transactions on pattern analysis and machine intelligence*, pp.131-137, Jan. 2004.

[17] A. M. Martinez, A. C. Kak, “PCA versus LDA,”, *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 23, no. 2, pp. 228-233, Feb. 2001.

[18] X. He, S. Yan, Y. Hu, P. Niyogi, H. J. Zhang, “Face recognition using laplacianfaces,” *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 27, no. 3, pp. 328-340, March 2005.

[19] B. G. Park, K. M. Lee, S. U. Lee, “Face recognition using face-ARG matching,” *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 27, no. 12, pp. 1982-1988, Dec. 2005.

[20] K. Venkataramani, S. Qidwai, B. V. K. Vijayakumar, “Face authentication from cell phone camera images with illumination and temporal variations,” *IEEE Trans. System, Man, and Cybernetics*, vol. 35, no. 3. pp. 411-418, Aug. 2005.

[21] J. Ruiz-del-Solar, P. Navarrete, “Eigenspace-based face recognition: a comparative study of different approaches,” *IEEE Trans. System, Man, Cybernetics*, vol. 35, no. 3, pp. 315-325, Aug. 2005.

[22] J. Kim, J. Choi, J. Yi, M. Turk, **“**Effective representation using ICA for face recognition robust to local distortion and partial occlusion,” *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 27, no. 12, pp. 1977-1981, Dec. 2005.

[23] M. Turk and A. Pentland, “Eigenfaces for recognition,” J. Cogn. Neur osci.,vol. 3, no. 1, pp. 71–86, 1991.

[24] H. Yu and J. Yang, “A direct LDA algorithm for high-dimensional data—With application to face recognition,” Pattern Recognition, vol. 34, no. 10, pp. 2067–2070, 2001.

[25] J. Huang, P. C. Yuen, W. S. Chen, and J. H. Lai, “Component-based subspace LDA method for face recognition with one training sample,” Opt. Eng., vol. 44, no. 5, p. 057 002, 2005.

[26] B. Schölkopf, A. Smola, and K.Müller, “Nonlinear component analysis as a kernel eigenvalue problem,” MPI fur biologische kybernetik, Tubingen, Germany, Tech. Rep. 44, 1996.

[27] J. W. Lu, K. Plataniotis, and A. N. Venetsanopoulos, “Face recognition using kernel direct discriminant analysis algorithms,” IEEE Trans. Neural Network, vol.

14, no. 1, pp. 117–126, Jan. 2003.

[28] G. Baudat and F. Anouar, “Generalized discriminant analysis using a kernel approach,” Neural Computer, vol. 12, no. 10, pp. 2385–2404, 2000.

[29] J.H. Pong, C. Yuen, W.S. Chen and J.H. Lai, “Choosing Parameters of Kernel Subspace LDA for Recognition of Face Images Under Pose and Illumination Variations,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B,* Vol.37, no. 37, pp.847-862, Aug. 2007.

[30] W.S. Chen, C. Pong, Y.J. Huang and D.Q. Dai, “Kernel Machine-Based One-Parameter Regularized Fisher Discriminant Method for Face Recognition, ” *IEEE Transactions on Systems, Man, and Cybernetics, Part B,* Vol.35, no. 5, pp.659-669, Aug. 2005.