

中 華 大 學

碩 士 論 文

一種以特徵點為基礎的新穎性人臉辨識技術
A Novel Feature-point Based Face Recognition
Technology

系 所 別：資訊工程學系碩士班

學號姓名：M09702033 彭國達

指導教授：黃雅軒 博士

中華民國 101 年 2 月

摘要

本論文提出一種搭配統計式降維演算法的特徵點雙向辨識之人臉辨識技術，藉由人臉上有明顯紋理的特徵點(例如：角點)資訊，來提升統計式演算法的辨識效能。首先，利用廣義鑑別分析法(Generalized discriminant analysis, GDA)來淘汰與輸入影像之人臉整體結構較不相同的人，接著以特徵點雙向辨識演算法，從剩餘的候選者當中，根據特徵點的區塊對應程度與彼此之間的分佈關係，判斷出最後的辨識結果。我們針對三套人臉資料庫(Banca 人臉資料庫、Feret 人臉資料庫和 Cas-Peal 人臉表情資料庫)進行效能驗證，只單用 GDA 的辨識率分別為 Banca：97.88%、Feret：83.65%和 Cas-peal：87.82%，而加入特徵點雙向辨識演算法後辨識率則提升為 Banca：98.65%、Feret：93.26%和 Cas-peal：91.66%。這顯示出本論文所提出的特徵點雙向辨識演算法具有良好的地區性特徵辨識結果，當再與整體性 GDA 辨識演算法結合，則可產生具有互補性的辨識效果，因此能達到高準度的辨識效能。



關鍵詞：人臉辨識、特徵點、廣義鑑別分析法、區塊比對

Abstract

A method of feature point based two way recognition algorithm with statistics type dimension reduction algorithm for face recognition is proposed. Using the significant texture of feature points to enhance the recognition performance of statistics type dimension reduction algorithm. First, Generalized discriminant analysis (GDA) is used to eliminate the person whose overall structure of face less the same with input image. Then, according to the block matching score and the distribution with feature points determine the final recognition result. We focused on three sets of face database (Banca face database, Feret face database and the Cas-Peal face database) for performance verification. The recognition rate with only use GDA in three sets of database are Banca : 97.88%, Feret : 83.65%, and Cas-peal : 87.82%. Both GDA and our propose method are used, the three of the recognition rate are Banca : 98.65%, Feret : 93.26%, and Cas-peal : 91.66%. This shows that our method have complementary effects of other algorithm to achieve high accuracy of recognition performance.

Keywords: face recognition; feature point; Generalized discriminant analysis (GDA); block matching

致謝

感謝指導教授 黃雅軒教授，在我研究所的學習生涯中，對我的耐心指導與教誨。在課業研究上給予我相當多的指導與鼓勵，經常的討論並指點我正確的方向，使我在這些年中獲益匪淺。老師對於學術研究的認真與做事態度的執著，是我所要學習的最重要精神。

感謝實驗室的同學陳禹仲、李允善、周泫廷與廖辛隆，在一同研究時互相幫助和砥礪，並從旁提供了我寶貴的建議，給予我在研究過程中很多幫助。還有感謝學長劉偉成、張倞禕、顏華慶、莊順旭、吳羿蓁與許廷嘉在多方面的指導與幫忙，另外還有感謝學弟陳冠豪、歐志鴻、林啟賢與陳宣有的協助，讓我能更快速的完成研究。

最後感謝我最愛的家人與親戚朋友，感謝他們給我的關心、支持與鼓勵。特別感謝我的父母，讓我在就學期間不用擔憂生活，能專心致力於研究工作。僅以此文獻給我摯愛的親人朋友。



目錄

摘要.....	i
Abstract.....	ii
致謝.....	iii
目錄.....	iv
圖目錄.....	vii
表目錄.....	vi
第一章 緒論.....	1
1.1 研究背景與動機.....	1
1.2 相關研究.....	1
1.3 論文架構.....	4
第二章 人臉辨識模組架構.....	6
第三章 第一階段辨識模組.....	7
3.1 影像前處理 A.....	7
3.1.1 正規化處理.....	7
3.1.2 光線補償.....	8
3.2 廣義鑑別分析法.....	9
第四章 第二階段辨識模組.....	13
4.1 影像前處理 B.....	15
4.2 特徵影像轉換和特徵點偵測.....	16
4.2.1 特徵影像轉換.....	17
4.2.2 特徵點偵測.....	19
4.3 區塊比對演算法.....	22
4.4 特徵點雙向辨識演算法.....	24
4.4.1 最長環狀排列模型.....	26

4.4.2	全域性模型.....	29
第五章	系統實驗與分析.....	32
5.1	Banca 人臉資料庫.....	32
5.2	Feret 人臉資料庫	33
5.3	Cas-Peal 人臉資料庫	34
5.4	分析結果.....	35
第六章	結論與未來展望.....	36
參考文獻	37



表目錄

表 5-1、BANCA 資料庫測試結果	33
表 5-2、FERET 資料庫測試結果	34
表 5-3、CAS-PEAL 資料庫測試結果	35



圖目錄

圖 1-1、長向相似卻不相關的二人	錯誤! 尚未定義書籤。
圖 2-1、人臉辨識演算法流程圖	6
圖 3-1、人臉區域定義 A	8
圖 3-2、光線補償處理架構圖	9
圖 3-3、光線補償輸出影像	9
圖 3-4、KERNEL METHOD 基本概念例圖	10
圖 4-1、特徵點雙向辨識演算法流程圖	13
圖 4-2、人臉特徵點資訊影像	14
圖 4-3、雙向辨識器流程圖	14
圖 4-4、辨識模組流程圖	15
圖 4-5、人臉區域定義 B.....	16
圖 4-6、正規化後的人臉影像	16
圖 4-7、(A)(B)(C) 為 GAUSSIAN SECOND ORDER PARTIAL DERIVATIVE;(D)(E)(F) 為趨近 GAUSSIAN SECOND ORDER PARTIAL DERIVATIVE 的 BOX FILTER. 圖中灰色區域權重值為 0.....	17
圖 4-8、LBP 計算方法	18
圖 4-9、LBP 特徵影像	19
圖 4-10、三個影像尺度之 BLOB RESPONSE MAP 範例	20
圖 4-11、P 點與周遭 16 個點之範例圖	20
圖 4-12、P 點與周遭 16 個點的比較結果	21
圖 4-13、統計圖 13 連續 T 個數的結果	21
圖 4-14、漢明距離範例	22
圖 4-15、人臉特徵點比對示意圖	23
圖 4-16、特徵點種類圖	24

圖 4-17、特徵點匹配對應圖	26
圖 4-18、隨機環狀式排列幾何模型缺點案例	27
圖 4-19、最長環狀排列模型	27
圖 4-20、特徵點示意圖	30
圖 5-1、BANCA 人臉資料庫中 3 種不同環境下的影像，第 1 列為 CONTROLLED、 第 2 列為 DEGRADED、第 3 列為 ADVERSE。	32
圖 5-2、FERET 資料庫的部分影像	33
圖 5-3、CAS-PEAL 人臉資料庫的部分影像	34



第一章 緒論

1.1 研究背景與動機

隨著時代的推進，科技的進步日新月異，人們對於生活品質的要求也日益提升。在傳統的個人身分認證措施上，需要使用到一些物品(例如:身分證件、晶片卡或是 RFID 卡) 或記憶一組冗長的密碼，若是遺失物品、遺忘密碼或是密碼洩漏，就會造成相當大的不便及損失。因此，為了改善這些可能的缺失，研究者提出了生物識別技術 (Biometric Identification Technology)，利用人體所具有的獨特生物特徵(例如指紋、虹膜及人臉等)作為身份辨識的依據，此類特徵因是與生具有且不易仿造，又可隨身攜帶，可大幅減低遺失、毀損與被竊取等各種狀況的發生。故此，各種針對人體特徵辨識的研究相繼出爐，而人臉辨識更是其中的佼佼者。

人臉辨識技術的應用層面也是非常地廣泛，例如：身份證明、人機介面、安全監控、關卡控制、嫌犯追蹤等用途上，加上目前主觀條件及客觀環境已有所改善，因此人臉辨識的應用預期將有更大的發展。近十年來，擁有人臉辨識的商品慢慢出現在市面上，早期應用於筆記型電腦的登入系統，使用者必須通過身分識別才能使用此電腦；近期人臉辨識則應用於智慧型手機上，手機持有者透過此系統就能輕鬆杜絕他人窺探手機裡的隱私達到手機加密的效果。但也由於其應用層面皆與個人隱私或人身安全有關，所以其嚴重性就顯得格外重大。本論文的研究希望能提昇人臉辨識的學理與技術水準，增進人臉辨識的準確率及可用性。

1.2 相關研究

人臉辨識技術的相關領域，近年來有相當多的研究與方法被提出，研究的議題大多在於人臉的角度變化、光線變化、表情變化和影像解析度等因素影響下如何有效地提升技術的容忍度及辨識率。早期人臉辨識的關鍵在於人臉特徵的擷取，由於人臉影像資料屬於高維度的特徵，不只運算量大，過高的維度特徵也會干擾

資訊的篩選，甚至癱瘓資料分析，因此如何適當地縮減資料儲存量就變得相當重要。在傳統常見的方法中，使用統計式方法的有主成分分析法(Principal Component Analysis, PCA)[1]、線性鑑別分析法(Linear Discriminant Analysis, LDA)[2][2]、局部保持投影法(Locality Preserving Projections, LPP)[3] 和正交局部保持投影法(Orthogonal Local Preserving Projection, OLPP)[4]等。PCA 的做法是找出人臉影像中共同存在的主要成分影像，也就是找到造成資料投影後變異量相對較大的基底向量，再利用這些主要成分影像來簡化原始影像的表現方式。LDA 是另一種常見的方法，其作法是增加不同類別之間特徵的差異性(Between-Class Variation)，並同時降低同一類別之不同影像間特徵的差異性(Within-Class Variation)，藉此可提升識別的精確度。比起 PCA 和 LDA 在資料結構上使用全域的歐基里德結構(Global Euclidean Structure)，LPP 建構每個資料點的最近鄰居圖(Nearest-Neighbor Graph)，保留了影像資料間的多方面結構(Manifold Structure)資訊；OLPP 所建立出的特徵空間基底，具有彼此正交的性質，更完整的保持住原始資料的多方面結構性。除了上述所說的線性降維技術以外，另外還有一些屬於非線性的降維演算法，例如基於核(Kernel)[5][6]的方法或是局部線性嵌入演算法(Locally Linear Embedding, LLE)[7]等。核方法的原理在於將原始資料映射到高維度空間，使得資料在此空間中能夠呈現線性可分割的狀態，但並不是每個核函數對每種資料都適用，所以核方法的關鍵在於如何選擇適當的轉換函數；LLE 是一種局部的映射方法，它保持資料局部幾何特性，此微小的區域上仍具有線性的特徵，透過 LLE 就能保持鄰近資料點間的結構，並達到降維的效果。保持結構的降維方法(例如：LPP、LLE)皆面臨到相同的問題，那就是無法有效地將同類別的資料做有效的分類。為了有效的達到資料維度的縮減，同時在不同類別之間的資料有明顯的差異性，Baudat G.提出了廣義鑑別分析法(Generalized Discriminant Analysis, GDA)[8]，GDA 是一種使用核函數的非線性區別分析，它將原始資料向量映射到高維度特徵空間，而在這個高維度的特徵空間上資料分佈盡量形成線性可分割，所以可獲得更好的分類效果。

原始影像中參雜著多種環境因素的干擾，如何從影像中擷取出有鑑別力的特徵影像也是一個重要的課題。其中，線段邊緣圖(Line Edge Map, LEM)[9]是將輸入的人臉影像經過邊緣偵測，如：Sobel 演算法、Prewitt 演算法等，再利用多邊線段擬合(Polygonal Line Fitting)，產生出人臉線段邊緣圖，最後，使用線段豪斯多夫距離(Line Segment Hausdorff Distance)來做判定兩張影像之間的相似度。LEM 找出影像上的邊緣線，所以其對光線變化較不敏感，但若當人臉發生大幅度的表情變化則，邊緣線段也會相對應的大幅度改變，同時使的辨識率降低；局部二元圖形(Local Binary Pattern, LBP)[10]近年來被廣泛的使用，其中也被應用於人臉辨識當中[LBP facerecognition]，LBP 以遮罩運算的方式，計算其中心與八鄰域的值的相關性，藉由產生一串二位元代碼，將這 8-bit 二位元代碼轉換成十進制即得到 LBP 影像特徵。此影像特徵可以搭配常見的分類器如上述介紹的 PCA、LDA，進而萃取出高辨識度之特徵。LBP 是一種區域的亮度關係，對光線變化有一定程度的抑制能力，這意味著在相同場景不同光照條件下擷取不同的影像，所得到的 LBP 特徵影像是不太有變化的。

近年來，物件辨識所使用的影像特徵點資訊也慢慢的被應用於人臉辨識當中例如：尺度不變特徵轉換(Scale-invariant feature transform, SIFT)[11]和 Speed-Up Robust Feature(SURF)[12]。SIFT 於 2004 年被 Lowe[11]提出且成功的應用於物件辨識方面，其認為影像在經過高斯模糊之後，可以表示影像在不同尺度下放大所小的結果，來改進更早之前的 Harris 角點偵測法本身非尺度不變的缺點，而 SIFT 特徵有一定的穩定性和強健性但由於其特徵點維度過大，比對時時間複雜度過高，加上攝影機畫素提升影像解析度越高 SIFT 所需處理的時間相對不理想。為解決 SIFT 缺點，Brown 和 Lowe[13]提出藉由霍夫轉換(Hough transform)濾除錯誤的特徵點，來減少不必要的比對，但效果仍有限。SURF 特徵成功的改進 SIFT 不理想的部分，利用積分影像(Integral images)來計算來逼近高斯的一階和二階導數數值，達到降低複雜度的目的；而 SIFT 將影像放大縮小的方式被改進成調整 Filter 大小以達到相同的目的；在比對方面，加入了 Hessian matrix 跡數的前處理，減

少比對時間。其實，在人臉影像找特徵點的理念與 Unsang Park 和 Anil K. Jain 所提出的使用人臉上的軟性生物特徵(Soft Biometrics)例如：胎記、痣、疤[14]大致上是相似的，其理念皆在於如何找出人臉上的局部紋理資訊。不同於 SIFT 或 SURF 的特徵點比對，[14]利用高斯差異影像(difference of Gaussians)配合不同的門檻值找出人臉上的軟性生物特徵，再以統計式的方式，統計人臉各個區域的特徵數量，將每個區域數量串接起來做為特徵使用。

統計式的人臉辨識固然重要，但仍有不足之處，例如孿生兄弟將無法有效的分辨出來。但從局部特徵的概念出發找出孿生兄弟的差異處，即可有效的分辨出結果。基於這樣的理解，本論文將以人臉局部特徵點的理念出發，搭配統計式降維演算法，發揮出截長補短的作用提升人臉辨識的效能。

1.3 論文架構



本論文共分為六章，各章節內容安排如下：

第一章說明本論文的研究動機與人臉辨識的相關研究。

第二張介紹整體的演算法流程，並簡單說明各個處理步驟。

第三章介紹第一階段人臉辨識模組，說明此階段如何選出與測試者相似的 N 名候選者。

第四章為本論文所提出的以特徵點為基礎的人臉辨識技術，特徵點雙向辨識演算法

第五章為實驗結果，本論文使用三套人臉資料庫(Banca 人臉資料庫、Color-Feret 人臉資料庫、Cas-Peal 人臉資料庫)進行效能驗證。

第六章為結論和未來展望，它總結本論文的研究成果，並對未來可繼續研究的方向做一個說明。



第二章 人臉辨識模組架構

此章節介紹本論文人臉辨識技術的整體架構，並針對每個模組進行簡單說明，圖2-1為本論文所提出的演算法流程圖。一開始當影像輸入進來之後，影像經過人臉及眼睛點偵測模組找到人臉的初步位置以及眼睛瞳孔的座標點，隨即進入第一階段的辨識流程。首先，利用瞳孔的座標點，我們在前處理A的模組中進行人臉的轉正、抓取以及光線補償動作，接著透過GDA模組將補償後的人臉影像做特徵降維，利用降維後的特徵向量進行辨識找出與輸入影像相似的M名候選人。找出候選人之後，對此M人進行第二階段的辨識流程，為因應特徵點雙向辨識模組，在前處理B中重新抓取人臉區域影像並對此影像進行特徵點偵測。緊接著利用這些特徵點的幾何分佈資訊於特徵點雙向辨識模組中使用區塊比對，在比對影像裡找到輸入的參考特徵點所對應的匹配特徵點，比較參考特徵點與匹配特徵點的分佈差異作為相似判斷的依據，得出最後的辨識結果。

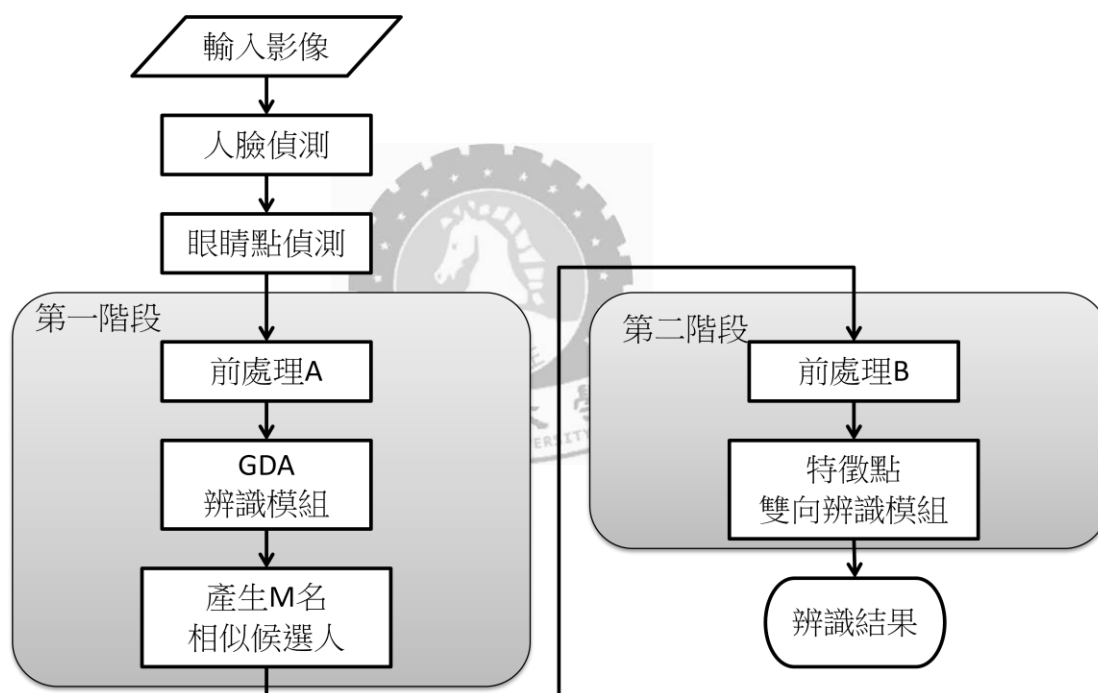


圖 2-1、人臉辨識演算法流程圖

第三章 第一階段辨識模組

本技術在偵測人臉部分使用傳統 AdaBoost 所訓練的分類器來對輸入影像進行人臉偵測。使用 AdaBoost 偵測出可能的人臉後，接著會經過輻射型對稱性轉換(Radial Symmetry Transform)產生眼睛瞳孔後選點；再從這些候選點當中，使用支持向量機(Support Vector Machine, SVM)來過濾出可能的眼睛點位置，若是符合眼點偵測的條件，則將 AdaBoost 偵測出來的結果判斷為人臉，並輸出眼睛點位置。利用得到的眼睛點位置重新定義出人臉區域，並對此人臉區域進行光線補償動作，最後，補償後的輸入人臉影像經由 GDA 模組，找出相似的 N 名候選者。3.1 節將會介紹人臉的正規化處理以及光線補償，3.2 節介紹廣義鑑別分析法。

3.1 影像前處理 A

3.1.1 正規化處理



首先，影像經由瞳孔點所連成的直線與影像水平線的夾角 θ 來進行影像轉正處理。接著由於雙眼瞳孔為較穩定之臉部特徵點，所以將雙眼瞳孔位置間之距離 d 定義為人臉區域的比例尺。人臉區域之寬度及高度皆定義為 1.6 倍的眼距，瞳孔距離左右邊界為 0.3 倍眼距，而瞳孔高度為 0.3 倍的眼距，如圖 3-1 所示。由於頭髮及嘴巴部分並不是穩定的臉部特徵，因此本研究定義人臉區域的上方邊界大約到眉毛部分，而下方邊界則是大約取到上嘴唇部分，最後將影像正規化至 36×36 的影像大小。

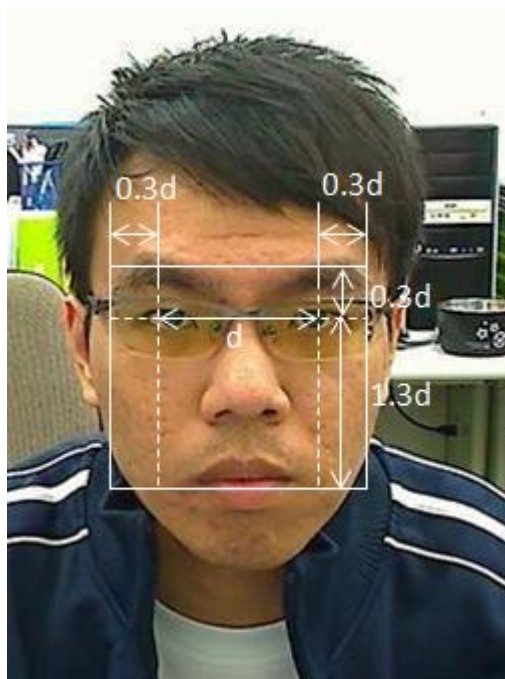


圖 3-1、人臉區域定義 A

3.1.2 光線補償

由於光線所產生的明暗變化會嚴重影響人臉特徵的抽取效果，因此如何對光線變化進行補償，也是一項重要的處理步驟。[15]論文中，提出了一套具有良好的光線補償效果。作法將於本段落中簡單的介紹，如圖 3-2。首先，同態濾波的作用在於對圖像灰度範圍進行調整，它能增強暗區的圖像細節，同時又不損失亮區的圖像細節。當影像經由比值影像運算處理後會將低頻的訊號強度大幅降低，因此可有效保存在辨識方面最重要的高頻紋理資訊，但同時高頻的雜訊部分亦將會被保留起來。非等方性平滑化的演算法不但可以達到降低高頻雜訊的目的，而且可以保持甚至強化紋理的特徵。最後利用直方圖等化從而調整像素值的區間，並得到更穩定的臉部分析與辨識結果。圖 3-3 為影像經由光線補償之後的輸出結果，第一列至第五列分別為：輸入樣本、濾波影像、比值影像、非等方影像、辨識輸出影像。

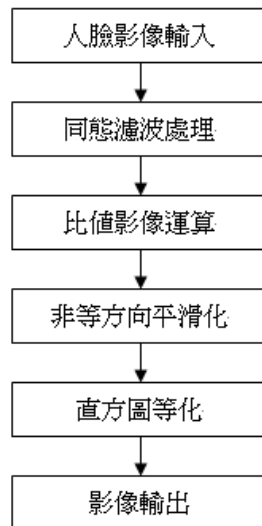


圖 3-2、光線補償處理架構圖



圖 3-3、光線補償輸出影像

3.2 廣義鑑別分析法

廣義鑑別分析法可說是 Kernel 化的 LDA，主要概念是透過映射函數將原始資料映射至高維度特徵空間，使其分佈盡量能以線性分割，提升了分類的效率，再降維取出具有最大分辨能力的低維次特徵。圖 3-4 以簡單的例子表示 Kernel Method 的基本概念，其中左圖代表兩種類型樣本在 R^2 空間上的原始資料分佈情

形，可以清楚的看到在原来的空間若要把這兩類樣本完全地分開，則需要一個橢圓。如果能夠先做一個動作將資料對應到三維空間（如圖 3-4 右圖所示），則只需要一個平面就可以達到目的。原本的橢圓方程式 $\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1$ 進行變數轉化（ $z_1 = x_1^2$ ， $z_2 = \sqrt{2}x_1x_2$ 和 $z_3 = x_2^2$ ）以後，可以化為 $\frac{1}{a^2}z_1 + 0 \cdot z_2 + \frac{1}{b^2}z_3 = 1$ ，所以用一個平面就可以將兩類樣本完全分開。因此如果映射函數(Mapping Function)選擇的恰當，本來比較複雜的關係就會變得比較簡單而容易分辨。

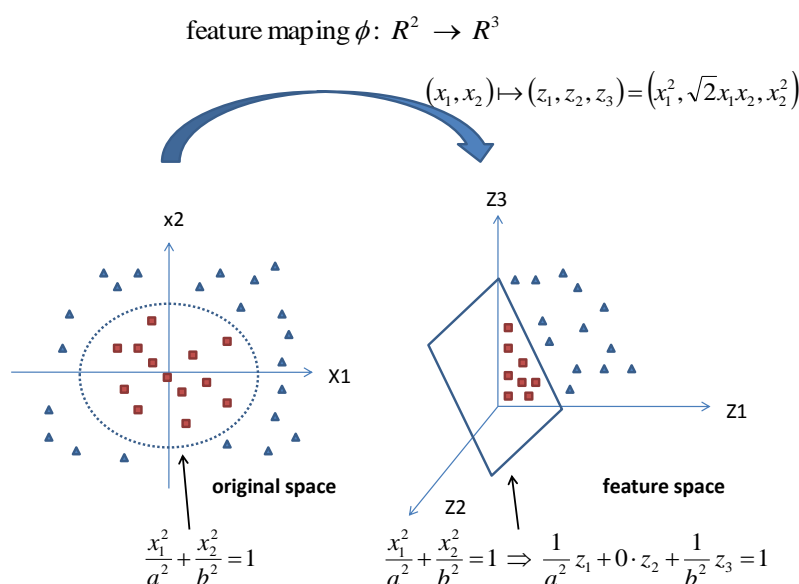


圖 3-4、 Kernel Method 基本概念例圖

在說明 Kernel Method 之前，在此先定義所用到的符號： $\phi(x)$ 表示將資料 x 對應到高維度空間， $\langle x, x' \rangle$ 為 x 和 x' 進行內積運算， $\|x - x'\|$ 為 x 和 x' 的距離。由於 Kernel Method 運用到複雜的觀念而比較難理解，在此先使用上面的例子來說明。若已知 Feature Mapping $\phi: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ 為

$$(x_1, x_2) \mapsto (z_1, z_2, z_3) = (x_1^2, \sqrt{2}x_1x_2, x_2^2)$$

假設在 \mathbb{R}^2 空間下 $x = (x_1, x_2)$ ； $x' = (x'_1, x'_2)$ ，則 \mathbb{R}^3 空間上的內積可推導如下：

$$\langle \phi(x_1, x_2) \cdot \phi(x'_1, x'_2) \rangle = x_1^2 x_1'^2 + 2x_1 x_2 x'_1 x'_2 + x_2^2 x_2'^2$$

$$= (x_1x'_1 + x_2x'_2)^2 = (\langle x, x' \rangle)^2$$

由推導結果得知，高維空間上的內積 $\langle \phi(x_1, x_2) \cdot \phi(x'_1, x'_2) \rangle$ 相等於低維空間上內積的平方 $(\langle x, x' \rangle)^2$ 。所以要計算高維空間二個特徵點間的距離或夾角時，並不需要真的進行高維次特徵的計算，而只需要做原本低維次的特徵計算即可。事實上，低維次的計算結果仍需要經由適當的轉換函數才能得到和高維次特徵計算相同的結果。這個轉換函數稱為 Kernel Function κ (kappa)。以上個例子來說， κ 可被定義為：

$$\kappa(x, x') = (\langle x, x' \rangle)^2$$

所以若定義出適合資料分類的 Feature Mapping ϕ ，即可使複雜的關係變得簡單而容易分辨，並且可透過 Kernel Function 的技巧減少運算的複雜度。因此，不需要在高維空間上計算，只要將原始的低維資料帶入 Kernel Function，就可以計算出在高維空間上的內積值。但是，如何選擇分類資料的適用 Mapping Function 是很困難的工作，而且每次都要根據 Mapping Function 來計算其對應的 Kernel Function，也是相當繁複的步驟。在本論文中，我們選擇已經過證明能符合 Kernel Function 條件且對一般分類問題具有良好成效的 RBF (Radial Basis Function) Kernel 作為我們的 Kernel Function κ ，透過調整變數 σ 可以找出最適合樣本資料的 Feature Space，其方程式如下：

$$\kappa(x, z) = \exp\left(\frac{-\|x - z\|^2}{2\sigma^2}\right), \sigma \in R - \{0\}$$

接下來，我們會介紹如何將 LDA 和 Kernel Method 結合起來，這樣結合的概念稱為 GDA (Generalized Discriminant Analysis)。為了方便 GDA 後續的說明，在此先定義一些變數： L 為所有類別的數目， N_i 為第 i 類別的樣本數目， N 為所有類別樣本的總數， M 為相似候選人個數， $\phi(x_j^{(i)})$ 為第 i 類中第 j 個樣本的高維空間之對應特徵， $X_i^T = [\phi(x_1^{(i)}), \dots, \phi(x_{N_i}^{(i)})]$ ， $X^T = [X_1^T, \dots, X_L^T]$ 。首先假設所有樣本在高維度空間中都經過中心點平移運算（即平均值 m_0 為 0），則其類別間散佈矩陣(Between-Class Scatter Matrix) S_b^{GDA} 與類別內散佈矩陣(Within-Class Scatter

Matrix) S_b^{GDA} 的定義如下：

$$S_b^{GDA} = \sum_{i=1}^L \frac{N_i}{N} (m_i - m_0)(m_i - m_0)^T = \sum_{i=1}^L \frac{N_i}{N} m_i m_i^T$$

$$S_w^{GDA} = \frac{1}{N} \sum_{i=1}^L \sum_{j=1}^{N_i} \phi(x_j^{(i)}) \phi(x_j^{(i)})^T$$

GDA 的目的是找出向量 v 使得類別間散佈矩陣(S_b^{GDA})與類別內散佈矩陣(S_w^{GDA})之比值為最大，其數學表示法為

$$v = \arg \max \frac{v^T S_b^{GDA} v}{v^T S_w^{GDA} v}$$

這相當於求出 $S_b^{GDA} v = \lambda S_w^{GDA} v$ 的特徵根問題(Eigen Problem)。由線性代數的理論可知，特徵向量 v 可由所收集的樣本在高維空間的對應向量做線性組合而得。

在辨識的過程中，採用最小距離分類法則(Nearest Distance Classification)，輸入影像 I_p 經過 GDA 轉換出代表每張影像的特徵參數 I_p^{GDA} ，資料庫中所註冊的第 k 個使用者的 n 張影像樣本同樣經過 GDA 轉換後的特徵參數 $R_{k,q}^{GDA}$ ($q = 1, \dots, n$)，則兩者間的距離為

$$d(I_p^{GDA}, R_{k,q}^{GDA}) = \|I_p^{GDA} - R_{k,q}^{GDA}\|$$

而測試影像 I_p 與資料庫中第 k 個使用者($k = 1, \dots, L$)的距離為

$$sim_{distance}^k = \min[d(I_p^{GDA}, R_{k,1}^{GDA}), d(I_p^{GDA}, R_{k,2}^{GDA}), \dots, d(I_p^{GDA}, R_{k,n}^{GDA})]$$

將此 L 個相似分數由小至大排列，取其前 M 個相似分數則所相對應類別即為相似候選者。

第四章 第二階段辨識模組

在第二階段辨識模組中，針對第一階段的辨識結果進行特徵點雙向辨識，產生最後的辨識結果。圖 4-1 為演算法的主要流程圖。首先，我們將原始灰階影像經由前處理 B 模組抓取出不同大小的人臉範圍後，將影像中有用的特徵資訊擷取出來，在此使 SURF(Speed Up Robust Feature)和 LBP(Local Binary Pattern)兩種特徵。在做完特徵轉換之後，藉由特徵點偵測找出人臉上具有代表性的局部特徵，如圖 4-2，圖中紅點即為偵測之後的結果，最後，雙向辨識器將判斷最後的辨識結果。

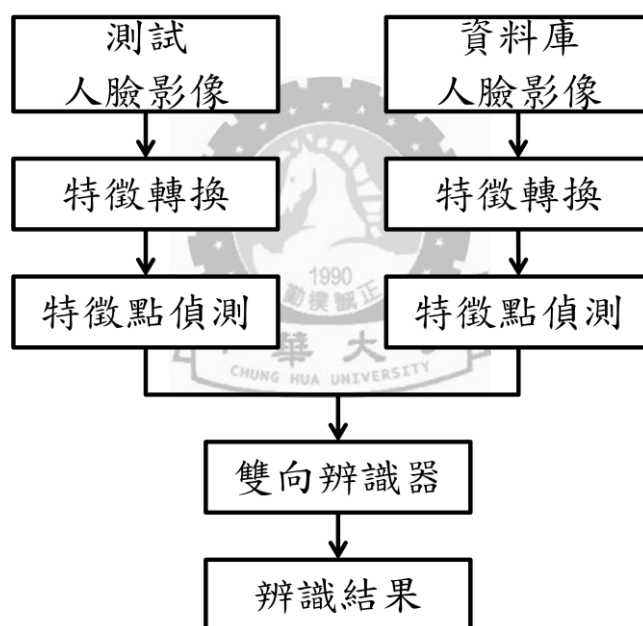


圖 4-1、特徵點雙向辨識演算法流程圖



圖 4-2、人臉特徵點資訊影像

下圖圖 4-3 為雙向辨識器之流程圖，配合此流程圖簡單的說明雙向辨識器的主體架構。

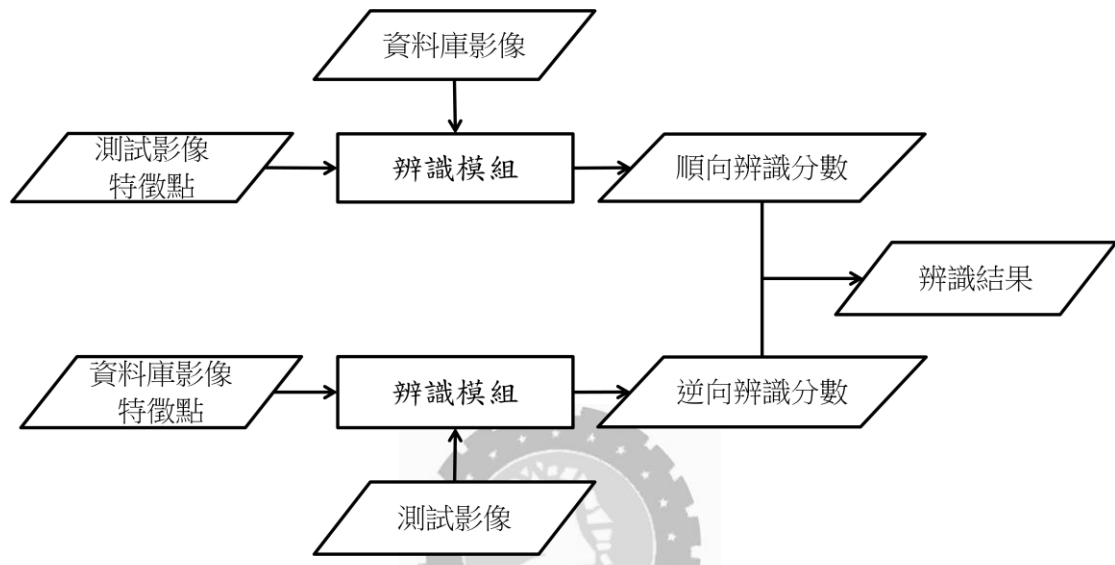


圖 4-3、雙向辨識器流程圖

順向辨識的部分，將測試影像所偵測出的特徵點，與資料庫影像經由辨識模組運算處理後，可得到一個順向辨識分數；而逆向辨識分數即是，將資料庫影像所偵測出的特徵點與測試影像，經由辨識模組運算後，所得到的辨識分數。最後，這二個辨識分數以加總運算後，即成為測試影像與資料庫影像之間的綜合分數。辨識系統在正和逆等雙方向辨識的運算設計下，可以對影像受到外部環境(例如光線變化或是影像上的雜訊)的影響具有較大的容忍性。下圖 4-4 為辨識模組的處理流程圖。

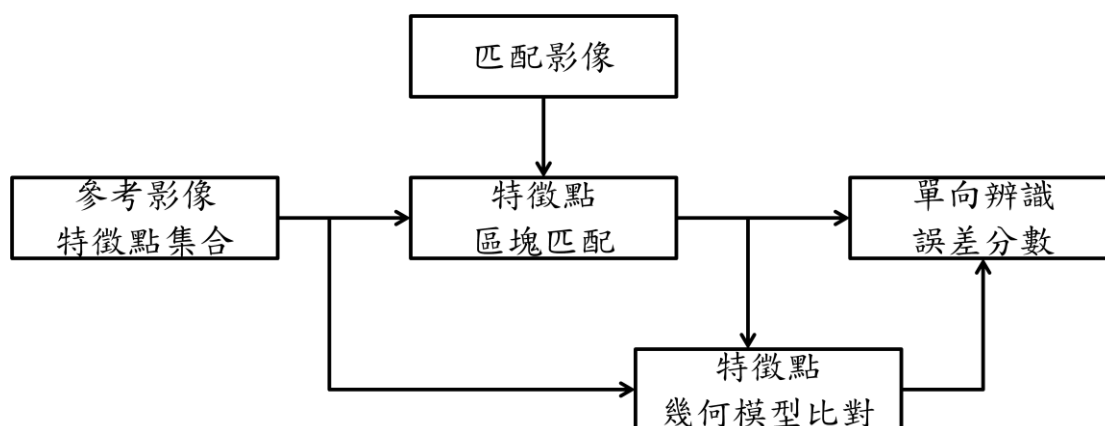


圖 4-4、辨識模組流程圖

若參考影像為輸入影像，則匹配影像為資料庫影像，此時可得到順向辨識分數；反之，若參考影像為資料庫影像，則匹配影像為測試影像，此時可得到逆向辨識分數。將參考影像特徵點，利用特徵點區塊匹配技術，於匹配影像中找尋最相似的區塊(即匹配強度能量最強的影響區塊)，即可得到與其對應的匹配特徵點。接著將參考影像特徵點集合與所得到的匹配特徵點集合進行幾何模型比對，產生出幾何模型誤差分數。最後，將幾何模型誤差分數與匹配強度能量值經由函數結合，即得到一個單向辨識結果的分數。

在 4.1 節，將會介紹適用於特徵點雙向辨識演算法的人臉區域定義。而 4.2 節，說明特徵影像轉換和特徵點偵測。4.3 節敘述區塊比對演算法。最後，特徵點雙向辨識演算法將於 4.4 節做詳細介紹。

4.1 影像前處理 B

考慮到特徵點的位置有機率出現在人臉的輪廓位置，則需要保留較大範圍的人臉資訊提供區塊比對使用。在此利用瞳孔典和嘴角點來定義出人臉的區域範圍，以雙眼瞳孔位置之距離 d 定義為人臉比例尺。瞳孔距離左右邊界為 0.6 倍的眼距，兩眼中心距離上方邊界為 0.7 被眼距，兩嘴角中心距離下方邊界為 0.6 被眼距，如圖 4-5 所示。接著，影像正規化的部分依據瞳孔的距離，將人臉影像所放置兩

眼距離為 70 像素點的距離如圖 4-6 所示。



圖 4-5、人臉區域定義 B



圖 4-6、正規化後的人臉影像

4.2 特徵影像轉換和特徵點偵測

本研究使用 SURF 與 LBP 特徵點二種特性的特徵點進行研究，4.2.1 節介紹特徵影像的轉換，而 4.2.2 節介紹特徵點偵測。

4.2.1 特徵影像轉換

4.2.1.1 SURF 特徵轉換

SURF 的特徵轉換[11]是使用 Hessian matrix 來表示，如下(3-1)式所示， $H(x, \sigma)$ 定義為影像 x 點在高斯 scale 大小為 σ 的 Hessian matrix，其中 L_{xx} 、 L_{yy} 與 L_{xy} 分別為水平、垂直與 45 度方向的 Gaussian second order derivative，如圖 4-7 的 a、b、c 子圖所示，再利用運算與它們近似的 box filter (如圖 3-2 的 d、e、f 子圖所示)進行 convolution 運算，得到 D_{xx} 、 D_{yy} 與 D_{xy} 的運算結果，來達到 Laplacian of Gaussian approximations 的目的。

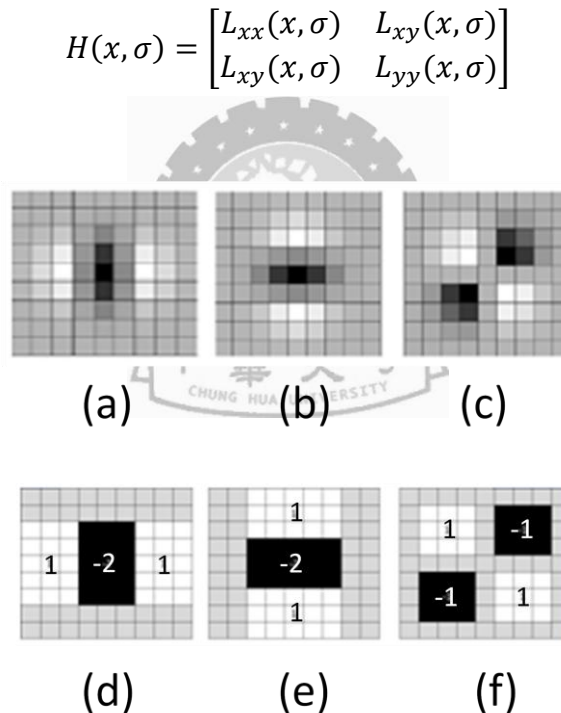


圖 4-7、(a)(b)(c)為 Gaussian second order partial derivative;(d)(e)(f)為趨近 Gaussian second order partial derivative 的 box filter. 圖中灰色區域權重值為 0.

因為這樣的 box filter 實作方法，需要大量累加影像上矩形區域的灰階值後再進行相減，為了增加效能，所以再搭配使用 Viola and Jones 所提出的 Integral image 方法[11]錯誤! 找不到參照來源。來減少計算量；再針對近似後的 Hessian

matrix 進行 determinant 計算，如下所示：

$$\det(H_{approx}) = D_{xx}D_{yy} - w(D_{xy})^2$$

經過 determinant 計算後就能得到每個像素點的響應強度，且利用不同高斯尺度大小的 box filter，來計算出固定大小影像中每個像素點的響應強度，且將每個像素點計算後的結果儲存在 blob response map 中，在 4.2.2.1 節將會利用 blob response map 來進行 SURF 特徵點的偵測。

4.2.1.2 LBP 特徵轉換

Local Binary Pattern (LBP)在許多影像紋理特徵資訊的擷取應用上已有相當不錯的成果，其計算的方法是從灰階影像上進行 3x3 的遮罩運算所產生的結果，而 3x3 的遮罩運算方法是比較遮罩中心點位置亮度值與周遭 8 個像素點亮度值的大小關係， (x_c, y_c) 為 LBP mask 的中心點， $s(x)$ 為二元判斷函式。

$$LBP(x_c, y_c) = \sum_{p=0}^7 s(g_p - g_c) 2^p$$

又

$$s(x) = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{if } x \leq 0 \end{cases}$$

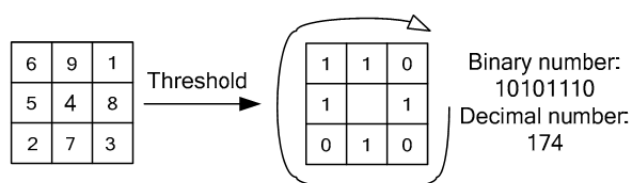


圖 4-8、LBP 計算方法

圖 4-8 為一個 LBP 特徵影像點的計算方法，最後計算出每個灰階影像像素點的 LBP 特徵值，其結果如圖 4-9 所示。

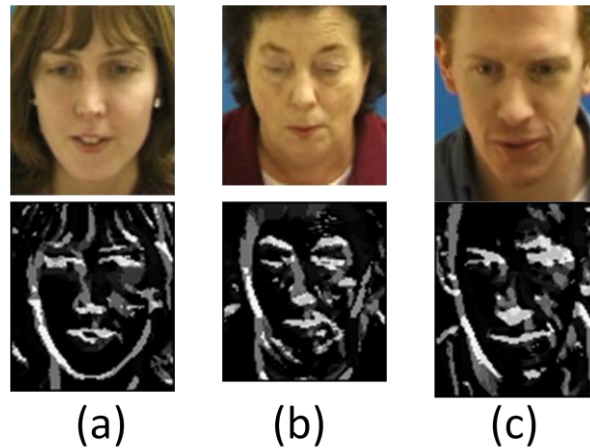


圖 4-9、LBP 特徵影像

4.2.2 特徵點偵測

影像經由特徵轉換後，利用所抽取出來的特徵資訊，並且使用特徵點偵測方法來將特徵點的位置偵測出來。4.2.2.1 節將說明 SURF 特徵點偵測演算法、4.2.2.2 節說明 LBP 特徵點偵測演算法

4.2.2.1 SURF 特徵點偵測

將原始灰階影像經過 SURF 特徵轉換後，針對 Hessian matrix 進行 determinant 計算後可得到每個像素點的響應值，再將許多不同影像尺度所計算的響應值儲存在 blob response map 中。圖 4-10 表示為使用 3×3 大小的 block 針對連續 3 個不同尺度大小的影像進行 non-maximum suppression 方法的判斷[]，當某一個位置 x 在影像尺度為 0.6 的 blob response map 之強度響應值比周遭 8 個鄰近位置的強度響應值都大，且也比影像尺度為 0.4 與 0.8 的 blob response map 之 18 個鄰近位置強度響應值都大時，則此位置即被偵測為一個 SURF 特徵點。然後使用 Brown 與 Lowe[]所提出的 Invariant Features from Interest Point Groups 方法，將該尺度所偵測到的特徵點位置與大小轉換到原始的 image space 中，如此一來就可以在多個影像尺度中都能偵測到 SURF 特徵點。

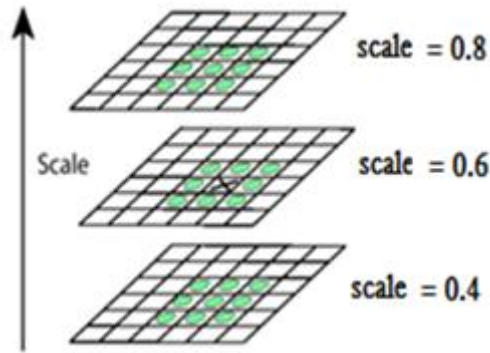


圖 4-10、三個影像尺度之 blob response map 範例

4.2.2.2 LBP 特徵點偵測

針對每一張 LBP 特徵影像來找出 LBP 特徵點，在 7×7 的範圍內，如下圖 4-11 的右圖所示，在中心點位置 p ，計算 p 點位置的 LBP 值與周遭 16 個位置點的 LBP 值之漢明距離，設定一臨界值 t ， $H_{p \rightarrow x}$ 為 p 點到 x 位置的 LBP 值之漢明距離，且 $x \in \{1 \dots 16\}$ ；在 x 集合裡，當 $H_{p \rightarrow x} > t$ 連續有 n 個 $p \rightarrow x$ 成立，且 n 大於臨界值 B ，則定義此 p 點為我們所要擷取的紋理特徵點，因為當 $H_{p \rightarrow x} > t$ 表示 x 位置與中心點 P 差異較大，而當影像為平滑區域時其 $H_{p \rightarrow x} < t$ ，又 $H_{p \rightarrow x} > t$ 的連續個數 n 表示此特徵點紋理的幾何角度程度，若 n 越大其角度越小，在這裡利用此參數來區分與剔除比對時容易比對錯誤的直線邊緣點。

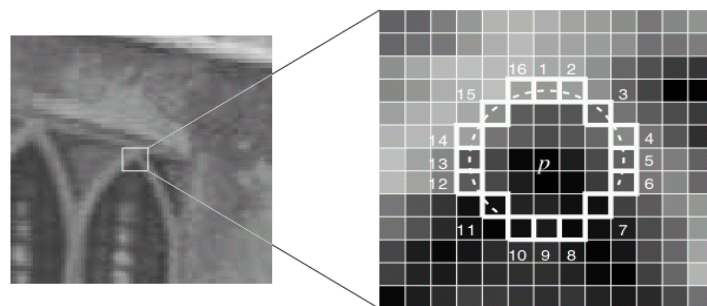


圖 4-11、 p 點與周遭 16 個點之範例圖

由於在中心點 p 的周遭可能伴隨著雜訊，如果我們只有使用 $H_{p \rightarrow i} > t$ 的個數來判斷這一點是否為特徵點將會是很不穩定的，所以我們加入了”連續成立”的條件，當 $H_{p \rightarrow i}$ 在 $1 \leq i \leq 16$ 需要連續成立，且次數要大於 B ，如此一來 p 點才會被認定為我們所偵測的 LBP 特徵點。下圖 4-12 為 p 點與周遭 16 個點計算後的結果，在每一次比對我們判斷 $H_{p \rightarrow i}$ 是否大於 t ，如果大於 t 則計算結果標記為’T’，否則標記為’F’

T	T	T	T	T	T	T	T	T	F	F	T	T	F	F	T	T
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

圖 4-12、 p 點與周遭 16 個點的比較結果

圖 4-13(a)是針對圖 3-13 統計出連續’T’個數的結果(i.e. 8, 2 and 2)。如果當比較的開頭與節尾都為’T’時，則將開頭與節尾連續 T 的個數進行相加，如圖 4-13(b). 最後我們得到了 10 與 2 的連續 T 個數. 一般的情況，我們會將 B 臨界值設定為 8，所以在這個例子當中，此 p 點將會被判定為 LBP 特徵點。經由我們所提出的 LBP 特徵點偵測計算後，可以得到如圖 3-15 的結果。

8	0	2	0	2
---	---	---	---	---

(a) 原始統計連續 T 個數的結果

10	0	2	0
----	---	---	---

(b) 最後統計連續 T 個數結果

圖 4-13、統計圖 13 連續 T 個數的結果

本論文所提出的 LBP 特徵點偵測方法對於光線變化具有相當程度的容忍性，且能將影像中鑑別度較高的紋理特徵點偵測出來，若使用一般常見的特徵點偵測

方法，如角點偵測，當特徵點中心與周遭像素點的亮度對比不夠大時，響應的能量值將會很低，則此特徵點將不會被偵測出來，而我們所提出的 LBP 特徵點擷取方法，由於是利用灰階影像在不同方向的亮度對比結構作為特徵，並進行能量值的計算，所以與一般的角點偵測方法相較之下，能偵測到更多的特徵點。在實際應用上，當影像的亮度值對比不強時，如果使用角點偵測方法來濾除位於影像平滑區域的 SURF 特徵點，將可能會發生整張影像中所擷取到的 SURF 特徵點數量過少的情形，造成後續的應用資料量不足的問題。

4.3 區塊比對演算法

區塊比對演算法主要應用於影像的追蹤系統中，用來判斷追蹤目標上一時刻的位置與目前時間的位置相對關係。演算法使用 LBP 特徵影像並利用漢明距離來計算兩個特徵點的相似度，相似度最高的特徵候選點即為目前追蹤物件的座標位置。在 LBP 影像中，一個像素點所代表的是 8 個灰階影像像素點與區塊中心像素點的亮度大小關係，利用互斥(XOR)邏輯運算將兩個要比對的 LBP 像素點進行二進制的 8 位元與 8 位元比對，其中只有 1 對 1 與 0 對 0 才會有 0 的結果，否則皆為 1，最後再計算 XOR 之位元值為 1 的個數，也就是所謂的漢明距離(Hamming Distance)，如以下範例圖 4-14 所示，其漢明距離為 3，而當 1 的個數較多時，則表示漢明距離越大，也就是差異度越大。

漢明距離範例EX

	LBP Pixel value1	10110111
XOR)	LBP Pixel value2	10111010
		00001101
		(漢明距離為3)

圖 4-14、漢明距離範例

當我們得到參考影像及批配影像的 LBP 特徵和參考影像的特徵點位置之後，

緊接著就進行 LBP 特徵區塊比對來找出參考影像的特徵點在匹配影像中可能的位置，已達到追蹤目的，如下圖 4-15 所示。

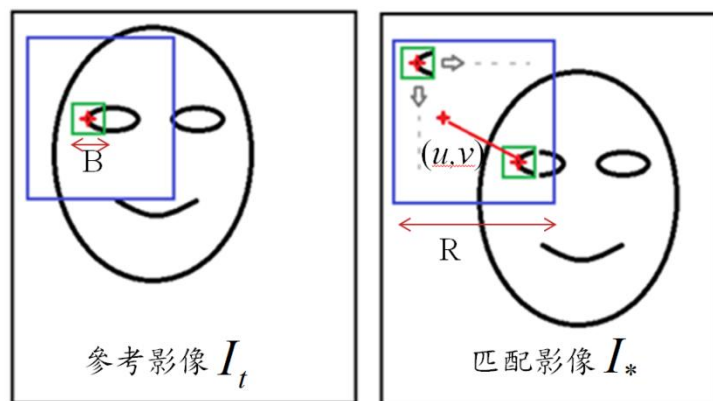


圖 4-15、人臉特徵點比對示意圖

在圖 4-15 中，我們以參考特徵點的位置為中心(如左圖影像的紅色特徵點)，在搜尋範圍內(如藍色方框區域)取適當大小的 LBP 特徵區塊(如綠色方框區域)來進行比對，統計出區塊內漢明距離的總和，做為兩個特徵點區塊相似程度的依據，而在 LBP 計算漢明距離的實作上，將兩個 LBP 值進行 XOR 邏輯運算後產生一個 8 位元數值 X，且進行 BIT[X]運算，此運算是利用查表的方式來統計總共有多少個 1，來轉換出比對後的相似度；而查表的方法是先建立一個 8 位元的表格，將所有 8 位元可能的 0 與 1 的排列組合都列出來，其資料表大小所佔的記憶體空間總共為 $8 \times 256 = 2,048$ 位元。假設 I_t 為參考影像、 I_* 為匹配影像、 (x,y) 為參考特徵點位置、 (u,v) 為特徵點在匹配影像中可能的位移向量，而 B 為以特徵點為中心所建立的特徵點區塊(Template Block)半徑，我們將參考特徵點位置 (x,y) 之 LBP 特徵點區塊與匹配影像中 $(x+u, y+v)$ 位置之 LBP 特徵點區塊進行漢明距離的計算，其數學表示式如：

$$LBP_{(x,y)}(u, v) = \sum_{j=-B}^B \sum_{i=-B}^B BIT[XOR(I_t(x+i, y+j), I_*(x+u+i, y+v+j))]$$

$LBP_{(x,y)}(u, v)$ 表示二個區塊之間的影像相似度。假設 R 為以特徵點區塊比對的搜

尋半徑，則參考影像特徵點 (x,y) 會在匹配影像中的搜尋區域內各影像點所對應的區塊逐一計算它們的 LBP 相似度，然後找出在搜尋區域內有最大相似度的一組位移量 (u,v) ，來產生最佳的位移量 (u^*,v^*) 即

$$LBP_{(x,y)}(u^*,v^*) = \arg \max_{\substack{-R \leq u \leq R \\ -R \leq v \leq R}} LBP_{(x,y)}(u,v)$$

在本研究中，將區塊比對方法應用於人臉辨識當中，用來判斷參考影像上的特徵點，將會坐落於匹配影像上的何處，若參考影像與匹配影像屬於同類別影像則所形成的偏移向量的幅度將會很微小，反之，若影像不屬於同一類別則偏移向量將會產生明顯的偏移幅度。

4.4 特徵點雙向辨識演算法

在本節中將會詳細的說明個模組之間的運作方式，在事前訓練的部分，資料庫人臉影像經由特徵點偵測處理，以先將所偵測到的特徵點資訊以及人臉影像儲存於一儲存裝置內，回復供後續測試端進行辨識。在測試方面，測試人臉影像經由進行特徵點偵測，找出測試影像上具有鑑別性的特徵點。如圖 4-16 所示，測試影像與資料庫影像各別找尋出具有鑑別性的特徵點。特徵點根據其位於人臉的位置將給予不同的權重值，若特徵點位於人臉輪廓處或是嘴巴附近則給予較低的權重值，因為這些地方容易受到邊緣環境以及人臉的情緒表情所影響，如圖 4-16 所示，圖中紅點為正常的特徵點，白點為容易受到表情或背景影響的特徵點。



圖 4-16、特徵點種類圖

在特徵點偵測完成之後，進入雙向辨識系統，進行測試比對。以順向辨識為例，將測試影像的特徵點資訊於資料庫影像中進行區塊匹配運算，來尋找於資料庫影像中有最大相似度的區塊，然後自此區塊即可找到匹配像素點。區塊比對法(Block Matching)為一種常用的相似度比較方法，其中 I_t 為參考影像， I_* 為匹配影像， (x,y) 為參考影像特徵點的位置， (u,v) 為位移量，而 B 為以特徵點為中心所建立的樣板區塊(Template Block)半徑。

$$LBP_{(x,y)}(u,v) = \sum_{j=-B}^B \sum_{i=-B}^B BIT[XOR(I_t(x+i, y+j), I_*(x+u+i, y+v+j))]$$

$LBP_{(x,y)}(u,v)$ 表示在參考影像上以 (x,y) 為中心的區塊和匹配影像上以 $(x+u, y+v)$ 為中心的區塊之間的影像差異量，當此差異量越小時，表示所對應的區塊影像相似度越高；反之差異量越大時，則表示對應的區塊影像之相似度越低。假設 R 為以特徵點區塊比對的搜尋半徑，則參考影像特徵點 (x,y) 會在區域影像中的搜尋區域內各影像點所對應的區塊逐一計算它們的 LBP 誤差量，然後找出在搜尋區域內有最小誤差量的一組位移量 (u,v) ，來產生最佳的位移量 (u^*, v^*) 即

$$LBP_{(x,y)}(u^*, v^*) = \arg \min_{\substack{-R \leq u \leq R \\ -R \leq v \leq R}} LBP_{(x,y)}(u, v)$$

當找出最佳的位移量以後，即可得知最佳的匹配位置 P 和其匹配能量強度 MD ，即

$$P = (x + u^*, y + v^*)$$

$$MD = LBP_{(x,y)}(u^*, v^*)$$

假設參考影像中總共有 K 個特徵點，而 MD_i 表示第 i 個特徵點的匹配強度，則其均匹配純量強度(AMD)為

$$AMD = \frac{\sum_{i=1}^K w_i \times MD_i}{\sum_{i=1}^K w_i}$$

圖 4-17 為特徵點匹配的結果，其上半部影像所標記的紅點為參考影像中所偵測的特徵點位置，而下半部影像中的紅點是參考影像特徵點在匹配影像中所匹配到的最佳對應點位置。



圖 4-17、特徵點匹配對應圖

當所有的參考影像特徵點都以區塊比對運算得到了它們在匹配影像上的最佳對應點位置之後，我們會根據這二組特徵點位置來建構出二個幾何模型。由參考影像特徵點所建構的幾何模型稱為參考幾何模型(RG)，而由匹配影像特徵點所建構的幾何模型稱為匹配幾何模型(MG)。然後藉由這二個幾何模型(RG和MG)，我們進行幾何模型匹配運算，以求出這二個幾何模型之間的差異程度。幾何模型是用來表示特徵點的結構關係，主要包含兩點之間的距離、方向以及三點所形成的夾角等，以下將會介紹二種幾何模型，於4.4.1小節中說明最長環狀排列模型，4.4.2小節說明全域性模型。在說明模型前，在此先假設參考影像總共有 K 個特徵點，而 x_i 、 y_i 和 w_i 分別記錄第 i 個特徵點的水平座標、垂直座標和權重值。

4.4.1 最長環狀排列模型

此模型將利用特徵點二點之間的距離以及三點所形成的角度做為幾何模型

的基底。我們簡單的以其編號的順序來連接此 K 個特徵點位置，形成一種所謂環狀排列的幾何模型，如下圖 4-18(a)所示。但是任意的排列存在一個嚴重的問題，今天 n_{i-1} 特徵點往右移動些許距離如圖 4-18(b)所示，雖然 $\overline{n_{i-1}n_i}$ 並沒有明顯的變化但是 n_{i-1} 、 n_i 和 n_{i+1} 所形成的夾角卻產生大幅度的改變，這將會造成後續辨識產生錯誤的比對。

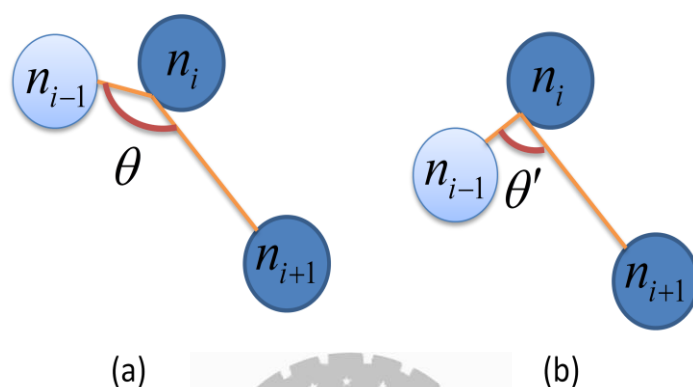


圖 4-18、隨機環狀式排列幾何模型缺點案例

為了改善上述問題，我們提出最長環狀排列模型如下圖 4-19，在不失其一般性的情形下，選擇 n_0 作為起始點，接著挑選與其最遠的點為 n_4 ，再以 n_4 為出發點找到最遠的點 n_1 ，直到找到最後一個點 n_5 後，將 n_5 與 n_0 做銜接形成一個環狀排列，如此一來即可避免角度不穩定的現象。

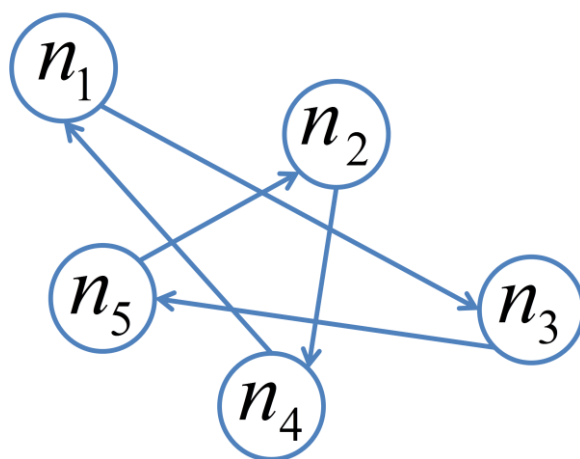


圖 4-19、最長環狀排列模型

在最長環狀排列模型中，任何連續 3 個特徵點 $n_s(x_s, y_s)$ 、 $n_i(x_i, y_i)$ 和 $n_t(x_t, y_t)$ 即構成一個夾角(θ_i)和其二邊線段($\overrightarrow{n_i n_s}, \overrightarrow{n_i n_t}$)，

$$\overrightarrow{n_i n_s} = [x_s - x_i, y_s - y_i]^T$$

$$\overrightarrow{n_i n_t} = [x_t - x_i, y_t - y_i]^T$$

$$\theta_i = \cos^{-1} \frac{\overrightarrow{n_i n_t} \cdot \overrightarrow{n_i n_s}}{\|\overrightarrow{n_i n_t}\| \|\overrightarrow{n_i n_s}\|}$$

由於每個特徵點各自擁有一權重值，所以每個線段以及夾角都有其相對應的權重值。假設 w_{is} 、 w_{it} 和 w_{ist} 代表 $\overrightarrow{n_i n_s}$ 、 $\overrightarrow{n_i n_t}$ 和 θ_i 的權重值，則

$$w_{is} = w_i \times w_s$$

$$w_{it} = w_i \times w_t$$

$$w\theta_i = w_{is} \times \|\overrightarrow{n_i n_s}\| \times w_{it} \times \|\overrightarrow{n_i n_t}\|$$

因此，在環狀排列的情況下，K 個特徵點可產生一組具有 2K 個數值的幾何模型 G，即

$$G = (d_1, d_2, \dots, d_K, \theta_1, \theta_2, \dots, \theta_K) \in R^{2K}$$

其中 $d_i = \|\overrightarrow{n_{i-1} n_i}\|$ ，也就是 d_i 為第 i-1 特徵點和第 i 特徵點之間的距離。而 RG 和 MG 的差異量(GD_{ring})計算方式如下：

$$GD_{ring} = \text{diff}(RG_i - MG_i)$$

其中，

$$RG_i = (rd_1, rd_2, \dots, rd_K, r\theta_1, r\theta_2, \dots, r\theta_K)$$

$$MG_i = (md_1, md_2, \dots, md_K, m\theta_1, m\theta_2, \dots, m\theta_K)$$

而相對應 RG_i 的權重值為

$$(w_1, w_2, \dots, w_K, w\theta_1, w\theta_2, \dots, w\theta_K)$$

GD_{ring} 可改寫為如下

$$GD_{ring} = \frac{\sqrt{\sum_{i=1}^K w_i \times (rd_i - md_i)^2}}{\sum_{i=1}^K w_i} \times \frac{\sqrt{\sum_{i=1}^K w\theta_i \times (r\theta_i - m\theta_i)^2}}{\sum_{i=1}^K w\theta_i}$$

4.4.2 全域性模型

本小節裡將有效有規則性的從 K 個特徵點中選取 3 個點並記錄此三點形成的線段長以及夾角，作為幾何模型中的一個組件。但為了簡化說明，我們將 K 個特徵點以 $n_i (i = 1, \dots, K)$ 表示，如下圖 4-20 所示。以特徵點 n_i 為參考點為例，屏除點 n_i ，從剩餘 $K-1$ 個點中隨機取出兩個特徵點，即會產生 $C_2^{(K-1)}$ 種組合，在多種組合當中，假設其中一組選擇了 n_s 以及 n_t 兩個特徵點。接著以點 n_i 為中心點，計算此三點所構成的夾角 θ_{ist} 和其二邊線段 $(\overrightarrow{n_i n_s}, \overrightarrow{n_i n_t})$ ，而

$$\overrightarrow{n_i n_s} = [x_s - x_i, y_s - y_i]^T$$

$$\overrightarrow{n_i n_t} = [x_t - x_i, y_t - y_i]^T$$

$$\theta_{ist} = \cos^{-1} \frac{\overrightarrow{n_i n_t} \cdot \overrightarrow{n_i n_s}}{\|\overrightarrow{n_i n_t}\| \|\overrightarrow{n_i n_s}\|}$$

其中

$$i = 1, \dots, K$$

$$s = 1, \dots, K \text{ and } s \neq i, s \neq t$$

$$t = 1, \dots, K \text{ and } t \neq i, t \neq s$$

由於每個特徵點各自擁有一權重值，所以每個線段以及夾角都有其相對應的權重值。假設 w_{is} 、 w_{it} 和 w_{ist} 代表 $\overrightarrow{n_i n_s}$ 、 $\overrightarrow{n_i n_t}$ 和 θ_{ist} 的權重值，則

$$w_{is} = w_i \times w_s$$

$$w_{it} = w_i \times w_t$$

$$w_{ist} = w_{is} \times \|\overrightarrow{n_i n_s}\| \times w_{it} \times \|\overrightarrow{n_i n_t}\|$$

以圖 4-18 為例， n_0 為參考點， n_1 以及 n_2 為選曲組當中的一組則 n_0 、 n_1 和 n_2 三個點所夾角 θ_{012} 以及形成的線段 $(\overrightarrow{n_0 n_1}, \overrightarrow{n_0 n_2})$ 為一組模型組件。

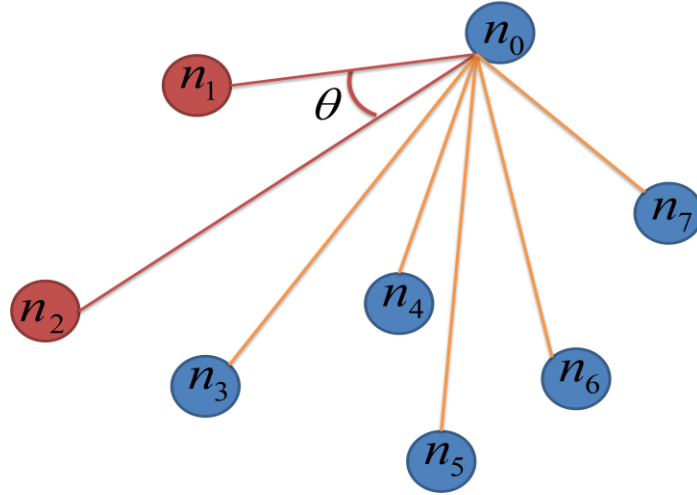


圖 4-20、特徵點示意圖

因此， K 個特徵點可產生出一組具有 $K \times \binom{K-1}{2}$ 個組件的幾何模型 G ，即

$$G_i = (d_{i1}, d_{i2}, \theta_i), i = 1, \dots, K', \quad K' = K \binom{K-1}{2}$$

其中 d_{i1} 以及 d_{i2} 代表模型組件的兩個線段長， θ_i 表示 3 個特徵點所夾的夾角。而 RG 和 MG 的差異量(GD)計算方式如下：

$$GD = \text{diff}(RG_i - MG_i)$$

其中，

$$RG_i = (rd_{i1}, rd_{i2}, r\theta_i)$$

$$MG_i = (md_{i1}, md_{i2}, m\theta_i)$$

而相對應 RG_i 的權重值組件為

$$(wd_{i1}, wd_{i2}, w\theta_i)$$

故，差異量 GD_{star} 可以改寫成如下：

$$GD_{\text{star}} = \frac{\sum_{i=1}^{K'} [wd_{i1} \times wd_{i2} \times \sqrt{(rd_{i1} - md_{i1})^2 + (rd_{i2} - md_{i2})^2} \times w\theta_i \times |r\theta_i - m\theta_i|]}{\sum_{i=1}^{K'} wd_{i1} \times wd_{i2} \times w\theta_i}$$

4.4.3 辨識分數整合

藉由組合特徵點的平均匹配純量強度值 AMD 以及幾何模型的變化量 GD，

來得出最後的單向辨識分數 S ，如下函式：

$$S = \frac{AMD}{GD}$$

假設 S_1 為順向辨識分數，而 S_2 為逆向辨識分數。將它們加總即得到最後的綜合辨識分數 FS 。

$$FS = S_1 + S_2$$

而擁有最大綜合辨識分數者即為最後的判斷結果。若相似候選人個數為 M ，每個人的影像張數以 M_i 表示則 FS_i 為第 i 類樣本的辨識分數，

$$FS_i = \max_{t=1, \dots, M_i} FS_t$$

針對每一類的辨識分數，選取擁有最大辨識分數的類別做為最後的辨識結果，

$$i^* = \arg \max_{i=1, \dots, M} FS_i$$

i^* 即代表測試影像的辨識結果。



第五章 系統實驗與分析

在本論文中提出了特徵點雙向辨識系統，以人臉的局部資訊配合使用區塊比對技術成功的達到人臉辨識的目的。而在此章節中將針對 PCA 和 GDA 二種降維演算法配合三種幾何模型(隨機環狀排列模型、最長環狀排列模型、全域模型)而相似候選人在實驗當中選取 3 名進行驗證，實驗對象使用三套資料庫分別為 Banca 人臉資料庫[16]、Feret 人臉資料庫[17]和 Cas-Peal 人臉資料庫[18]。

5.1 Banca 人臉資料庫

本論文中進行的實驗使用了 Banca 人臉資料庫[15]，其包含了 52 個不同人的影像，其中有 26 個為男人 26 個為女人，並且在不同的攝影機及三種場景下拍攝(Controlled、Degraded 和 Adverse)。每個人在三種場景下各於不同時間拍攝 4 組影像，每次拍攝 10 張影像，所以每一組有 520 張人臉影像，12 組一共 6240 張人臉影像。拍攝的過程中被拍攝者的表情並不受限制，每張人臉影像的大小為 360×288 像素。每一組皆於不同時間拍攝，時間前後達三個月，可以說是一個非常齊全，十分適合拿來進行人臉辨識與驗證實驗的資料庫。圖 5-1 為其中 3 人在 3 種不同的環境所拍攝的影像。



圖 5-1、BANCA 人臉資料庫中 3 種不同環境下的影像，第 1 列為 Controlled、第 2 列為 Degraded、第 3 列為 Adverse。

實驗中所使用的實驗樣本為第一種場景(Controlled)下拍攝的 4 組人臉影像，資料庫中共有 52 個人，拿其中 2 組作為訓練影像、一組為驗證影像而另一組為測試

影像，詳細的辨識結果如下表 5-1。

Banca	原始效能	隨機環狀 排列模型	最長環狀 排列模型	全域模型
PCA	89.42%	90.57%	91.15%	93.46%
GDA	97.88%	96.9%	98.07%	98.65%

表 5-1、Banca 資料庫測試結果

5.2 Feret 人臉資料庫

Feret 人臉資料庫為公開資料庫中較為龐大的，其擁有 1208 個人，每個人有多個角度的影像以及不同時間點的影像，下圖圖 5-2 為 Feret 資料庫的部分影像。本研究從其中挑選出 52 人，每個人 4 張的正面影像進行人臉辨識，2 張作為訓練影像、2 張作為測試影像，下表 5-2 為 Feret 的辨識結果。



圖 5-2、Feret 資料庫的部分影像

Feret	原始效能	隨機環狀 排列模型	最長環狀 排列模型	全域模型
PCA	64.42%	66.34%	67.30%	70.19%
GDA	83.65%	85.57%	85.57%	93.26%

表 5-2、Ferret 資料庫測試結果

5.3 Cas-Peal 人臉資料庫

Cas-Peal 資料庫為現今公開的資料庫中少數以亞洲人為主體的人臉資料庫，同時也建立了多種情境的人臉影像，包括有人臉的多角度影像、不同的光線角度影像、攜帶配件的影像如眼鏡、帽子、和不同的表情影像等等。資料庫包含了 1040 個人，其中男性有 595 人女性有 445 人，總共有 99594 張影像，下圖 5-3 為 Cas-Peal 人臉資料庫的部分影像。

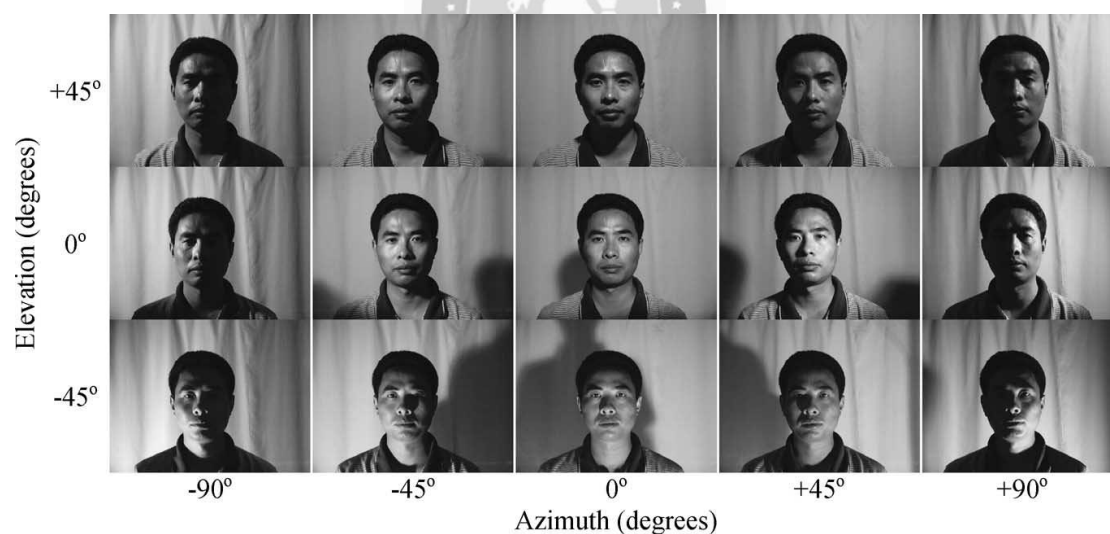


圖 5-3、Cas-Peal 人臉資料庫的部分影像

本論文在此使用光線變化的人臉影像進行實驗，從資料庫中挑選出 52 人擁有相同條件下的光線變化影像，1 個人 3 張訓練影像、3 張測試影像，辨識結果如表 5-3

Cas-Peal	原始效能	隨機環狀 排列模型	最長環狀 排列模型	全域模型
PCA	76.92%	78.84%	78.84%	80.76%
GDA	87.82%	88.46%	89.74%	91.66%

表 5-3、Cas-Peal 資料庫測試結果

5.4 分析結果

本論所提出的特徵點雙向辨識演算法搭配傳統降維式的演算法，皆成功達到效能提升的效果，利用特徵點的幾何關係改善指針對整體人臉結構的演算方法。其中，第一階段的演算法的選取扮演相當重要的角色，由 Feret 資料庫即可看出結果，在相同候選人個數的情形下，將 GDA 的辨識率提升 10%，反觀 PCA 卻只有 6% 的辨識率，這意味著對 PCA 來說，3 名相似候選人的空間不足以找到正確的結果。在幾何模型的選取方面以全域結構搭配權重值的設計，充分的反映出特徵點之間的群聚關係，且不容易受到幾何上的誤差干擾。

整體的實驗效果皆對傳統方法產生明顯的效能提升，因此本論文所提出的特徵點雙向辨識技術搭配傳統統計式降維演算法，對人臉辨識而言，具有一定的效果。

第六章 結論與未來展望

由於現今科技進步快速，人們為了有更好的生活品質和更好的效率，人機介面的重視已不可同日而語，使用者只需要面向攝影機的鏡頭，就可以對使用者做出身分識別，改善過去要依靠密碼或是卡片所帶來的不便。在本論文中，提出一種以特徵點為基礎的雙向辨識演算法，利用人臉局部紋理的特性使用特徵點區塊比對搭配特徵點之間的幾何關係，成功改善統計式降維演算法的效能。

由於區塊比對所花費的時間過長，造成辨識需要花較長時間的等待，可能會給使用者帶來不便。而區塊比對所需的時間與特徵點的數目有絕對的相關，雖然降低特徵點數量可以立即改善此問題，但任意的刪減往往容易適得其反，所以，如何有效的過濾特徵點以及找出穩定的特徵點來達到系統的加速以及穩定將是未來改進的主要方向。



參考文獻

- [1] M. Turk and A. Pentland, “Eigenfaces for recognition”, *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, pp. 71-86, 1991.
- [2] V. Belhumeur, J. Hespanha, and D. Kriegman. “Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, July 1997.
- [3] X. He, S. Yan, Y. Ho, P. Niyogi and H. J. Zhang, “Face recognition using Laplacianfaces,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 3, pp. 328-340, 2005.
- [4] D. Cai, X. He, J. Han, and H. Zhang, “Orthogonal Laplacianfaces for face recognition,” *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3608-3614, 2006.
- [5] J. Shawe-Taylor and N. Cristianini. *Kernel Methods for Pattern Analysis*. Cambridge University Press, 2004.
- [6] W. Liu, J. Principe and S. Haykin. *Kernel Adaptive Filtering: A Comprehensive Introduction*. Wiley, 2010.
- [7] Hong Chang & Dit-Yan Yeung, “Robust Locally Linear Embedding,” Technical Report HKUST-CS05-12, 2005.
- [8] G. Baudat and F. Anouar, “Generalized discriminant analysis using a kernel approach,” *Neural Comput.*, vol. 12, pp. 2385–2404, 2000.
- [9] Yongsheng Gao and Maylor K.H. Leung, “Face Recognition Using Line Edge Map,” *IEEE Pattern Analysis and Machine Intelligence* 2002.
- [10] T. Ojala, M. Pietikäinen and T. Mäenpää, “Multi-resolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 24, pp. 971-987, 2002.

- [11] David G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints”, International Journal of Computer Vision, 60, 2, pp. 91-110, 2004.
- [12] H. Bay, T. Tuytelaars, and L. V. Gool, “SURF: Speeded up robust features,” in Proc. of the 9th European Conference on Computer Vision (ECCV’06), ser. Lecture Notes in Computer Science, A. Leonardis.
- [13] Brown, M.; Lowe, David G. (2003). "Recognising Panoramas". *Proceedings of the ninth IEEE International Conference on Computer Vision*. **2**. pp. 1218–1225. doi:10.1109/ICCV.2003.1238630
- [14] Unsang Park; Jain, A.K., “Face Matching and Retrieval Using Soft Biometrics,” IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 5, NO. 3, SEPTEMBER 2010
- [15] 劉偉成，連續性人臉辨識技術，中華大學資訊工程研究所碩士論文，2009.
- [16] S. Bengio, F. Bimbot, J. Mari ´ethoz, V. Popovici, F. Por ´ee, E. Bailly- Bailli`ere, G. Matas, and B. Ruiz, “Experimental protocol on the BANCA database”, IDIAP-RR 05, IDIAP, 2002.
- [17] P.J. Phillips, H. Wechsler, J.S. Huang, and P.J. Rauss, “The FERET Database and Evaluation Procedure for Face-Recognition Algorithms,” Image and Vision Computing, vol. 16, no. 5, pp. 295-306, 1998.
- [18] Wen Gao, Senio, Bo Cao, Shiguang Shan, Xilin Chen, Delong Zhou, Xiaohua Zhang, and Debin Zhao, “The CAS-PEAL Large-Scale Chinese Face Database and Baseline Evaluations” IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART A: SYSTEMS AND HUMANS, VOL. 38, NO. 1,

JANUARY 2008.

