

出國報告(出國類別：研究)

目標追蹤與辨識之研究

服務機關：國防大學理工學院電機電子系

姓名職稱：郝樹聲 文職副教授

派赴國家：美國

報告日期：98年9月21日

出國時間：98年6月17日至98年8月27日

摘 要

本心得報告主要內容是本人於 98 年度或國科會補助赴美短期研究之成果及心得之報告，本研究內容主要針對車輛與人物之目標物，進行追蹤與辨識之研究，內容大部份注重在 2D 目標物之追蹤與辨識，考慮未來目標物之追蹤與辨識之發展與趨勢，本報告內容對於 3D 目標物之特性也加以探討，此一研究可以應用在安全監控系統的設計上，以管制人員車輛的出入，另一方面，同樣的理論於國防上可以應用在指揮與管制系統上。本研究的方法採用影像與視訊所具備的各種資訊，例如動作向量、色彩、邊緣資訊、統計特性、時序特性等經由影像處理與統計特性之擷取分析後來加以追蹤，接著分割主體與背景最後加以辨識。追蹤辨識的方法中，在辨識方面，我們會應用統計學上的各種可行的方法從事辨識，而在追蹤上我們應用一些常見的方法，例如卡曼濾波器等法從事追蹤之研究，最後希望可以研發有效、快速的目標追蹤與辨識之演算法，並且希望可以將其應用在安全監控系統上。本計畫之研究場所為美國加州南加大計算機科學系由 Gerard Medioni 教授所主持的 IRIS(Institute for Robotics and Intelligent Systems)實驗室，從事為期三個月之短期研究。

目次

摘要	2
一、目的	6
二、過程	6
三、介紹	6
四、目標追蹤辨識系統	8
(a) 移動偵測模組	8
(b) 陰影去除法	11
(c) 特徵區域選取	12
(d) 追蹤演算法	12
(i) 卡曼濾波器理論分析	13
(ii) 核心函數理論分析	16
(1) 目標模型：	17
(2) 候選模型：	18
五、視窗整合系統	24
(a) 門禁警監系統流程圖	25
(b) 光照均勻化	26
(i) Histogram Equalization	26
(ii) Weighting Block Histogram Equalization	29
(iii) Discrete Cosine Transform	33
(iv) Gamma Intensity Correction	35
(v) Local Normalization	39
(vi) Self Quotient Image	42
(vii) Reflection Ratio	44
(c) 辨識演算法	48
(i) PCA	48

(ii)	LDA	51
(iii)	2DPCA	53
(iv)	2DLDA	55
(v)	特徵比對	57
(1)	曼哈頓距離	58
(2)	相關係數	59
六、	實驗過程與分析	60
(a)	目標追蹤與辨識	60
(i)	車輛追蹤	60
(ii)	室外單人走動	61
(iii)	室外多人走動	62
(iv)	移動人物被遮蔽	63
(v)	追蹤效能分析	65
(vi)	人臉辨識效能分析	66
(b)	門禁警監系統之應用	69
(i)	人物追蹤	69
(1)	更新背景	69
(2)	分割前景	73
(3)	膚色偵測	74
(4)	臉部擷取	75
(5)	光照正規化	76
(ii)	人物辨識	77
(1)	訓練影像資料庫	77
(2)	特徵空間	77
(3)	特徵比對	77
(iii)	實驗結果分析	78
(c)	門禁警監系統之實驗設計	83

(i) 實驗一(受測者正上方有兩盞日光燈)	83
(ii) 實驗二(受測者正上方有一盞日光燈)	85
(iii) 實驗三(受測者正上方日光燈全開)	86
(iv) 實驗四(受測者受側面光照影響)	87
(v) 實驗五(受測者未戴眼鏡)	88
(vi) 實驗六(受測者穿戴接近膚色之衣服)	89
(vii) 門禁警監系統之結果分析	90
七、3D 影像用於人臉追蹤辨識之研究	91
八、結果與討論	96
參考文獻	98
心得	103
建議事項	103

一、目的

本短期研究之目的是赴美國加州南加大與主持目標與追蹤辨識相關理論及計畫在目標追蹤與辨識這個領域著有聲譽的先進學者 Gerard Medion 教授從事研究討論，並且學習這個領域的一些新知，另外也爲了本系 98 年度新成立的『目標與追蹤辨識實驗室』儲備研究能量。在研究期間，所討論的主題有目標追蹤與辨識系統的架構如何訂定，例如硬體的規格與限制，這些限制如何克服，哪些硬體是目前較實用的，以及軟硬體的應用上該注意哪些關鍵點，目前有哪些新的技術在被採用中。研究期間並且與 Gerard Medion 教授討論它所主持之實驗室目前所採用的各項技術，希望學習其經驗並且將這些獲得的新知帶回國內，在本校相關主題的研發上注入新的活水，幫助本人所主持的『目標追蹤辨識實驗室』可以發展更精進的技術與演算法，而且希望藉此更加充實本實驗室之研究能量，幫助將來在添購擴充相關各項軟硬體設備時，在有限的經費上可以採購到最有價值且實用的實驗設備。

二、過程

本研究過程實際在美期間從 98 年 6 月 17 日至 98 年 8 月 27 日爲止，本人前往位於美國加州洛杉磯之南加州大學 Viterbi School 下之計算機科學系，由 Gerard Medion 教授主持的 IRIS (Institute for Robotics and Intelligent Systems) 實驗室，從事短期研究。本計畫是由國科會第四十七屆 (第九十八年度) 補助科學與技術人員國外短期研究所補助之經費，計畫申請編號爲 98BFE09001，計劃補助編號爲 98-2918-I-606-001，補助期限爲三個月。研究期間中主要進行以下幾樣工作，一、與 Gerard Medion 教授互相討論並且請教此次的研究主題，二、與 IRIS 實驗室的同學討論 meeting，三、資料蒐集，四、程式撰寫模擬，五、相關軟硬體資訊以及各項理論新知蒐集及評估，六、研究結果之紀錄撰寫等工作。

研究期間 Gerard Medion 教授提供我一間與一位韓國教授共用之研究室以及網路方便我從事研究，研究期間所用的電腦則是有本人自行從國內帶去使用，研究期間 IRIS 實驗室的同學提供我許多生活與研究上的許多協助，除此之外，南加大計算機科學系與學校之「外國學生中心」之行政同仁也提供我許多在研究外的協助，讓我在美研究期間一切順利。

三、介紹

目標追蹤與辨識最早是應用於軍事與太空應用之中[1, 2]，例如戰鬥機之飛行員依賴雷達提供目標之各種資訊給他，或是海上戰艦之雷達也須提供艦上戰管系統之敵我目標識別資訊，至於地面上之各種載具也都或多或少應用到目標之追蹤與辨識之功能，而至於在太空上的各種應用，如衛星對各種目標的追蹤與辨識之例子更是不勝枚舉。隨著硬體與軟體技術的進步，目標的追蹤與辨識相較以往已越來越精確。近年來因爲恐怖活動的盛行與日益惡化的治安，各種以安全監控查核爲目的所發展的系統也陸續出現以因應各式各樣的需求。以各種應用而言，有車輛[3]與人員[4]的追蹤，而其中又從單一目標進展到多目標[5]的追蹤，至於所辨識的目標有人臉[6]、姿勢[7]、手勢[8]、掌紋[9]、瞳孔[10]、車牌[11]等，而目標的形態又可分爲二維[12]與三維[13, 14]、可見光與不可見光[15, 16]，除此之外，其他的應用還有如年齡之辨識[17]等。

由以上的例子可知，目標物的追蹤與辨識是一項多年以來應用廣泛的關鍵技術，隨著軟硬體的發展，各種演算法也相繼的被提出來，舉例來說，有利用 Color-based Mean Shift (MS)

配合卡曼(Kalman)濾波器[18]來追蹤目標，以及其他的方法如 Particle Swarm Optimization(PSO) [19]，Graph Based Discriminative Learning[20]，Particle Filter(PF)[21]，Hierarchical Active Shape Model(ASM)[22]，Block Based Motion Estimation[23]，或者是利用 Range Data 從事 3D 之 Motion Estimation[24]，以 Geometric Constraints 從事 Incremental Model-Based Estimation[25]，以 Adaptive Block Matching[26]來追蹤目標，以 Voting-based 的方法同時追蹤多個目標[27]，或是以 Wavelet Transform Based 利用 Hierarchical Active Shape Model 同時追蹤多個目標[28]，以 Bounded Hough Transform[29] 來追蹤，多模人臉辨識與追蹤[30]，其他如 Cascade Histograms[31]、Data Driven Markov Chain Monte Carlo(DD-MCMC)[32, 33]、利用 Multiview Geometric Constraints 來偵測動作區域[34]、5D Tensor Voting 執行動作之分割與分析[35, 36]、利用多台照相機從事追蹤[37]等都是目前不同學者所提出的不同演算法。

本校為一軍事院校，除了一般與產業界有關的學術研究之外，另一方面也以服務三軍為主要目標，本研究主題在與產業界有關的研究項目中，以安全監控為主要方向，例如人物車輛的追蹤與辨識就是其一，而在與國防軍事有關的應用舉例來說，有營區的安全查核，船艦車輛的辨識等。因此本研究主題之重要性不容忽視。綜合以上各個參考文獻的研究主題可以發現，所謂的目標已經從以往的 2D 進入到 3D 的領域，這是因為 2D 影像有其先天上的許多限制，利用 3D 的資訊我們在從事辨識時就可以更加精確。而在從事辨識之前的物件分割動作更是關係到整個系統運作的成功與否，若有辦法精確的分割目標物，後續的辨識動作就可以有較大的成功機會。因此物件的分割、辨識與追蹤三者就形成一個不可分開的連結關鍵技術。本人此次所提出的出國短期研究主題將探討不同型態目標物之追蹤、分割與辨識的相關技術與發展。希望以新的技術支援本校在安全查核或是目標追蹤辨識各個主題上相關的計畫。

以下就目前較先進的追蹤辨識系統概念作簡介如圖 1 所示：

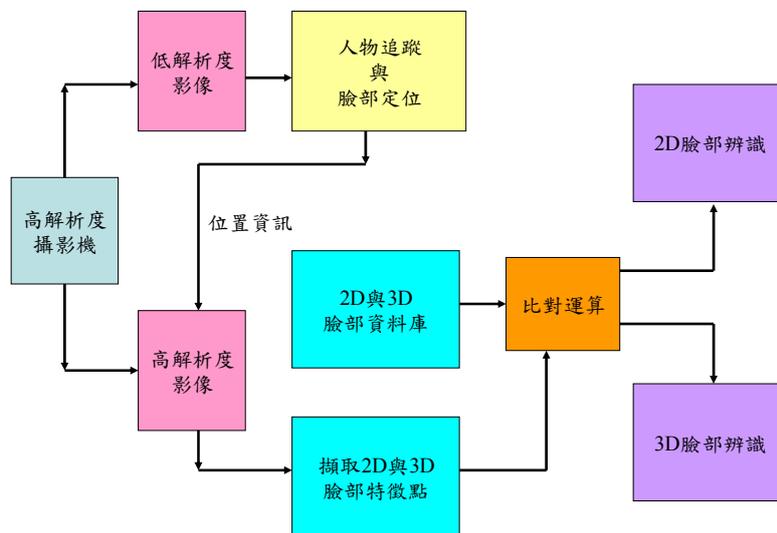


圖 1、追蹤辨識系統概念圖

如圖 1 所示我們可以看出利用一具高解析度的攝影機，我們可以同時獲得高低解析度兩種影像，其中低解析度影像乃是用來快速的偵測到目標物的移動，並且可以快速的定位到人的臉部，接著將位置資訊提供給系統，從高解析度攝影機所擷取影像的同樣位置上，得到 2D 與 3D 臉部的特徵點，至於 2D 與 3D 臉部的資料庫則是事先就已採集好以備比對之用。利用

比對運算的機制從事現場所擷取資料庫的 2D 與 3D 臉部的特徵點進行分析比對，而達到 2D 與 3D 臉部比對之目的。由此一系統圖可看出高解度攝影機是其核心的硬體設備。

以下就針對目標追蹤辨識系統與視窗整合系統兩大系統架構與實驗結果加以介紹，接著對於目標追蹤辨識步驟及實驗結果加以分析與討論。

四、目標追蹤辨識系統

本系統中偵測與追蹤的流程可區分為三個主要系統模組，如圖 2 所示，此一架構包含了移動物體偵測、特徵區域選取、移動向量估測與追蹤等。在移動物體偵測部分，對於攝影機所擷取的連續影像資料從事偵測，將移動物體變化部分自背景中分割、標示出來。偵測到移動物體後，選取適當的特徵區域來描述物體，以便在後續的特徵比對上使用。最後移動估測追蹤就是將目前得到的特徵和上一個時間點所得到的特徵作比較，估測出特徵區域可能的位置，找出相等或是最近似的特徵，修正預測所需參數以達到移動物體追蹤之目的。以下就針對系統內各個重要處理功能方塊作一介紹。

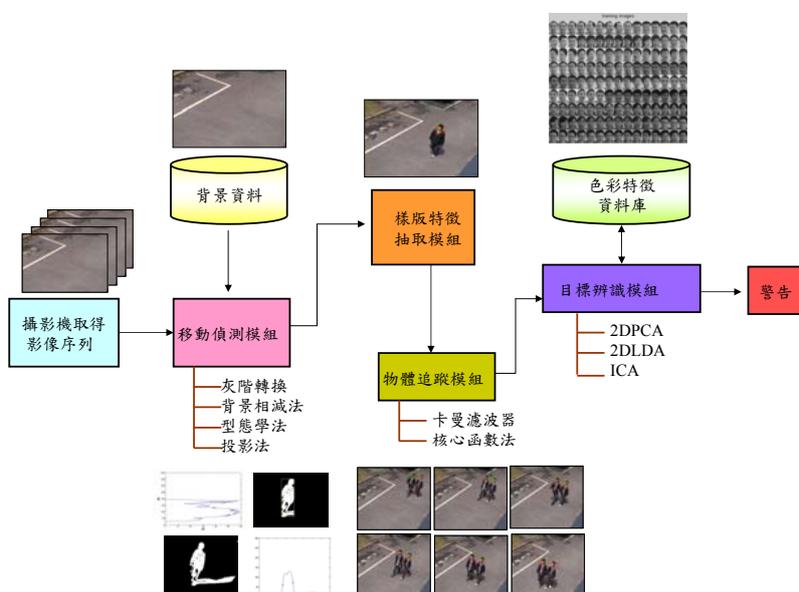


圖 2、目標追蹤辨識系統方塊圖

(a) 移動偵測模組

物體偵測與追蹤的第一步是比較連續影像的差異，將影像序列中連續幾張不相同的區域當作是可能的前景區域，也就是移動的物體，本實驗之流程圖如圖 3 所示，本文是利用背景相減[38]偵測方法偵測出移動物體。背景相減法之流程圖如圖 4 所示。

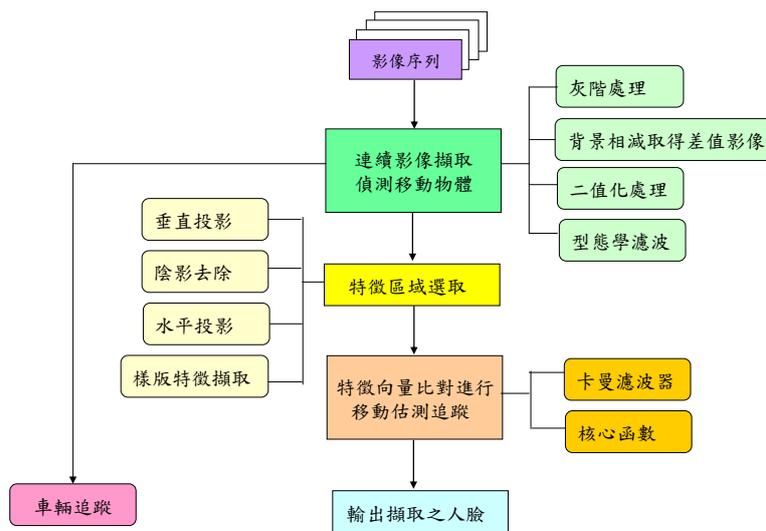


圖 3、目標移動偵測模組方塊圖

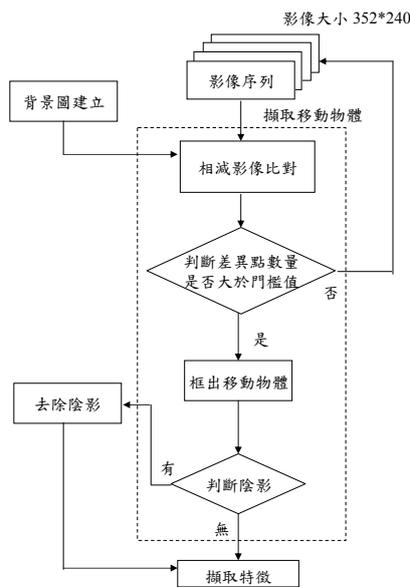


圖 4、背景相減法

此方法採用影像序列中當前畫面和背景畫面，進行比較來獲得移動物體的偵測，在連續影像中，因背景影像相對於前景有較長時間的不變性，我們便可以依此特性將背景分離出來，或是利用統計的方式由多張影像中擷取出平均背景影像，利用這個背景影像與後續進入的影像做比對以找出差異處，若有物體移動時，因背景影像是變化的，只有移動物體會產生位移，故將兩張影像相減即便可獲得移動物體資訊。然而，此方法對於光影的改變十分敏感，光線的變化會導致相減後的影像不精確，因此良好的參考背景影像是應用此法的主要關鍵，本文所介紹的方法是目前最常使用於擷取前景區域的方法。本系統採用連續 10 張影像做像素統計平均，可得到一張平均背景影像做參考，如圖 5 所示。

背景圖建立公式可以表示如下：

$$I_b(x, y) = \frac{1}{N} \sum_{i=1}^N f_i(x, y) \quad (1)$$

其中 $I_b(x, y)$ 為背景圖， $f_i(x, y)$ 為第 i 張影像， N 為採用連續影像數，此處令 $N=10$ 。



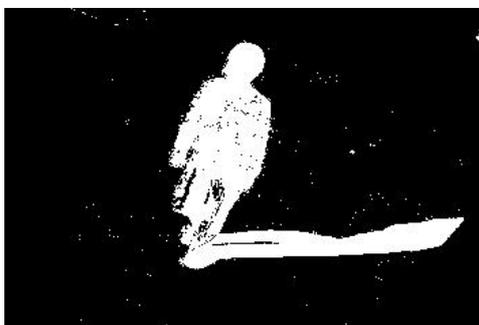
圖 5、連續 10 張背景影像做平均後之結果

連續影像中，當有物體移動時，背景會有位移區域出現，我們將這張影像 I_t 與 I_b 平均背景影像相減後取絕對值，便可得到物體的移動改變量。為了消除亮度變化或其他雜訊干擾影響，我們可加上一個門檻值，當同一位置相減後之絕對值大於門檻值，便可將此點視為移動的像素，因而可定義相減後的二值化影像為

$$D(x, y) = \begin{cases} 1 & , \text{if } |I_t(x, y) - I_b(x, y)| \geq T \\ 0 & , \text{otherwise} \end{cases} \quad (2)$$

其中 $I_t(x, y)$ 、 $I_b(x, y)$ ，代表移動物體與背景影像中 (x, y) 位置之亮度值， T 為門檻值， $D(x, y)$ 則是相減後得到的二值化影像。

在相減所得之二值化影像 $D(x, y)$ 中。值為 1 的部分可視為移動物體所造成亮度變化的地方；值為 0 則代表固定不變的背景，圖 6 所示為不同門檻值所造成移動偵測的結果。



(a) 門檻值為 20



(b) 門檻值為 50

圖 6、不同門檻值所造成移動偵測的結果

(b) 陰影去除法

若追蹤系統的背景是在戶外的環境下，有時會因為光線變化而使物體產生陰影，破壞了物體偵測完整度。為了提升偵測的準確性，必須將陰影去除。圖 7 為本系統去除陰影的流程圖。

判斷確定有物體進入後，先利用垂直投影的方法來計算出物體的垂直投影量，接下來設定門檻值 T （本研究設定為最大垂直投影量值的一半），若垂直投影量大於門檻值 T 可視為移動物體部分，小於門檻值則當成是陰影。這種方法較簡單，但是當兩人走的比較近時，容易誤判為同一物體，造成偵測錯誤，因此本研究利用垂直投影量變化來區分物體的起始點及結束點，在一張投影量曲線圖中，如圖 8(a)，當曲線上某一點的後一個投影量值大於門檻值且前一個投影量值小於門檻值，就可當作是偵測物體的起始點，一般而言，後面接續的投影量值會一直增加，到最大值波峰後開始遞減，一直到當其中一點的前一個投影量值大於門檻值且後一個投影量小於門檻值，則可視為結束點，如圖 8(b)所示，這樣就可以藉由起始點及結束點的標示分割出完整人體部分，並標定出人體的寬度座標。有了此寬度座標後，再計算此範圍內的水平投影量，如圖 8(c)所示，利用水平投影量計算出方框內的投影量，即可標定出人體的高度座標，如圖 8(d)所示。

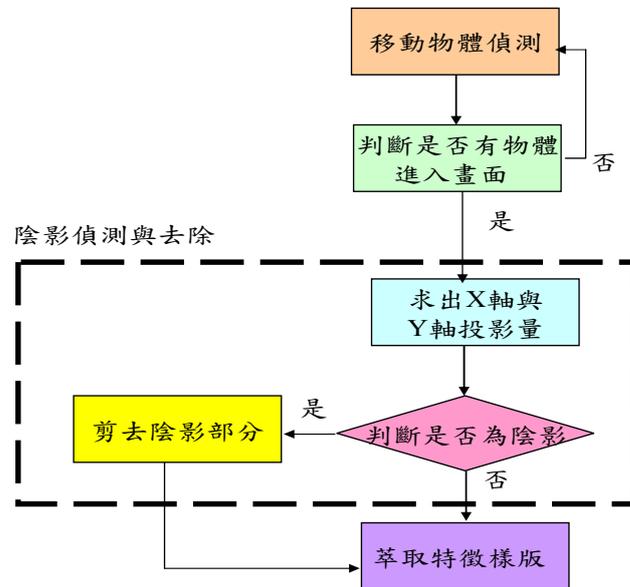
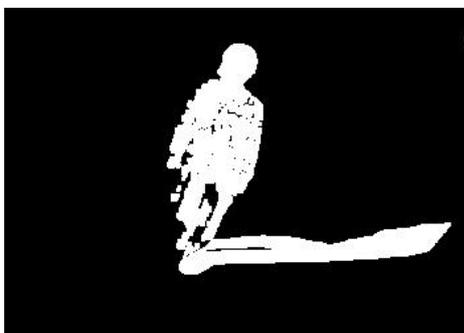
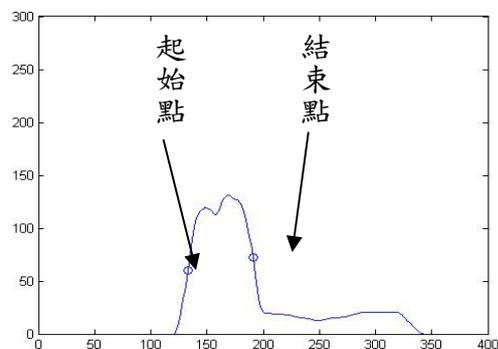


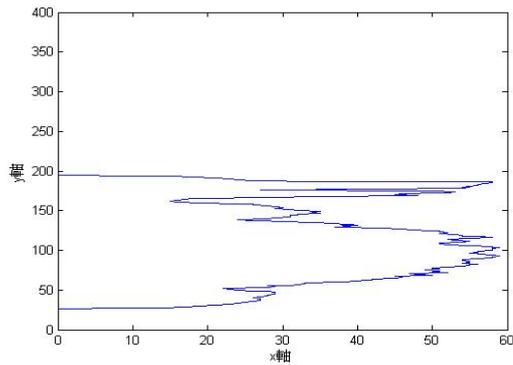
圖 7、陰影去除方塊圖



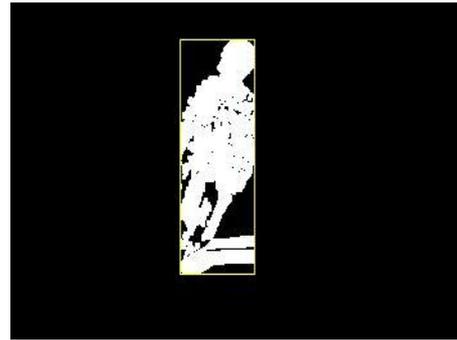
(a)移動物體二值化影像



(b)垂直投影起始點及結束點



(c)移動物體水平投影



(d)框出移動物體

圖8 去除陰影及框出移動物體圖

(c) 特徵區域選取

為了達到人物辨識功能，本報告將特徵區域選定在頭部部位選取，先利用垂直投影量及水平投影量計算出人形方框四個點的座標，再參考人體比率來框出頭部部份，一般而言，人體頭部的垂直位置約佔人體方框上面的 $1/6$ 高處，在兩個場景中的寬度位置分別約介於人體方框中間的 $2/3$ 寬及 $1/3 \sim 2/3$ 寬處，如圖 9 所示。本文將此區域的中心點設定為追蹤點，有了此中心點後，擷取 $m \times m$ 的視窗作為移動物體的樣版影像，可用來作為爾後辨識的參考特徵，本文中 $m \times m$ 值為 39×30 。圖 9 中，方框即為選定的樣版特徵影像。接著就可以在此區域裡尋找臉部特徵做比對。



(a)攝影機於3公尺高
拍攝到的人體



(b)攝影機於平地
拍攝到的人體

圖 9、特徵區域選取示意圖

(d)追蹤演算法

在追蹤方面，我們採用卡曼濾波器(Kalman Filter)[39-40]、核心函數[41-44]演算法來完成目標追蹤之演算，以下將針對這兩個演算法加以描述。

(i) 卡曼濾波器理論分析

卡曼濾波器以遞迴的計算方式，用來估算出一個系統過去、現在和未來不同時間點上的狀態，讓系統在追蹤上更為簡便，卡曼濾波器基於系統之前的狀態序列對下一個狀態序列做最佳化估計，以動態的狀態方程式和量測方程式來描述系統。考慮一個含有處理雜訊(Process Noise)之離散狀態方程式(State Equation)可以完成以下兩個方程式：

$$\text{狀態方程式： } \mathbf{X}_t = \mathbf{A}\mathbf{X}_{t-1} + \mathbf{B}u_{t-1} + \mathbf{W}_{t-1} \quad (3)$$

$$\text{量測方程式： } \mathbf{Z}_t = \mathbf{H}\mathbf{X}_t + \mathbf{V}_t \quad (4)$$

其中 \mathbf{X}_t 表示為時間為 t 時的 $n \times 1$ 維狀態向量(State Vector)；維度 $n \times n$ 的矩陣 \mathbf{A} 為狀態傳輸函數(Transition Matrix)，將過去 $t-1$ 時間狀態和現在的 t 時間狀態連結起來，實際中 \mathbf{A} 可能隨時間變化，但這裡假設為常數；維度 $n \times 1$ 的矩陣 \mathbf{B} 為任意控制輸入函數 u 與狀態 \mathbf{X}_t 的關係矩陣。(4)式中 \mathbf{Z}_t 表示為量測到的 $m \times 1$ 維狀態量測向量；維度 $m \times n$ 的矩陣 \mathbf{H} 表示狀態變量 \mathbf{X}_t 和量測變量 \mathbf{Z}_t 之間的關連矩陣， \mathbf{H} 可隨著時間與每一次測量改變，但在此假設其不變；另外此兩式中的隨機信號 \mathbf{W}_t 和 \mathbf{V}_t 分別表示維度 $n \times 1$ 的狀態雜訊向量和維度 $m \times 1$ 的量測雜訊向量。假設 \mathbf{W}_t 和 \mathbf{V}_t 為相互獨立且均值為零(Zero Mean)的白色高斯雜訊(White Gaussian Noise)，利用 \mathbf{Q}_k 和 \mathbf{R}_k 表示為狀態雜訊 \mathbf{W}_t 及量測雜訊 \mathbf{V}_t 的共變異(Covariance)矩陣：

$$\mathbf{Q}_t = E[\mathbf{W}_t \mathbf{W}_t^T] \quad (5)$$

$$\mathbf{R}_t = E[\mathbf{V}_t \mathbf{V}_t^T] \quad (6)$$

定義 $\hat{\mathbf{X}}_t^-$ 為時間 t 的測量前預估狀態(A Predicted State)， $\hat{\mathbf{X}}_t$ 為時間 t 測量後的更新估測狀態(A Updated State)($-$ 代表先前， $\hat{}$ 代表估測)。由此定義預估誤差(e_t^-)和估測誤差(e_t)如以下兩式所示：

$$e_t^- \equiv \mathbf{X}_t - \hat{\mathbf{X}}_t^- \quad (7)$$

$$e_t \equiv \mathbf{X}_t - \hat{\mathbf{X}}_t \quad (8)$$

而預估誤差與估測誤差的共變異矩陣可寫成：

$$\mathbf{P}_t^- \equiv E[e_t^- e_t^{-T}] \quad (9)$$

$$\mathbf{P}_t \equiv E[\mathbf{e}_t \mathbf{e}_t^T] \quad (10)$$

爲了推導出卡曼濾波器的方程式，目的是計算出時間 t 的估測狀態，因此定義估測狀態 $\hat{\mathbf{X}}_t$ 爲

$$\hat{\mathbf{X}}_t = \hat{\mathbf{X}}_t^- + \mathbf{K}_t (\mathbf{Z}_t - \mathbf{H} \hat{\mathbf{X}}_t^-) \quad (11)$$

其中 $(\mathbf{Z}_t - \mathbf{H} \hat{\mathbf{X}}_t^-)$ 稱爲更新值(Innovation)；而維度爲 $n \times m$ 的矩陣 \mathbf{K}_t 稱爲卡曼增益係數(Kalman Gain)，它會使估測誤差的共變異矩陣 \mathbf{P}_t 最小；爲了達成尋找最佳線性濾波器的目標，將式(11)代入估測誤差 $\mathbf{e}_t = \mathbf{X}_t - \hat{\mathbf{X}}_t$ 和估測誤差共變異矩陣 $\mathbf{P}_t = E[\mathbf{e}_t \mathbf{e}_t^T]$ 中，最後結果如下所示：

$$\begin{aligned} \mathbf{P}_t &= E[\mathbf{e}_t \mathbf{e}_t^T] = E\left[(\mathbf{X}_t - \hat{\mathbf{X}}_t)(\mathbf{X}_t - \hat{\mathbf{X}}_t)^T\right] \\ &= (\mathbf{I} - \mathbf{K}_t \mathbf{H}) \mathbf{P}_t^- (\mathbf{I} - \mathbf{K}_t \mathbf{H})^T + \mathbf{K}_t \mathbf{R}_t \mathbf{K}_t^T \end{aligned} \quad (12)$$

其中 \mathbf{I} 爲單位矩陣，爲了使估測誤差共變異矩陣 \mathbf{P}_t 最小，則需要對 \mathbf{P}_t 矩陣中之跡(Trace)對 \mathbf{K}_t 微分，並令其爲0，即

$$\frac{\partial Tr(\mathbf{P}_t)}{\partial \mathbf{K}_t} = -2 \mathbf{P}_t^- \mathbf{H}^T + 2 \mathbf{K}_t \mathbf{H} \mathbf{P}_t^- \mathbf{H}^T + 2 \mathbf{K}_t \mathbf{R} = 0 \quad (13)$$

因此可以推導出：

$$\begin{aligned} \mathbf{K}_t (\mathbf{H} \mathbf{P}_t^- \mathbf{H}^T + \mathbf{R}) &= \mathbf{P}_t^- \mathbf{H}^T \\ \mathbf{K}_t &= \mathbf{P}_t^- \mathbf{H}^T (\mathbf{H} \mathbf{P}_t^- \mathbf{H}^T + \mathbf{R})^{-1} \end{aligned} \quad (14)$$

再將(14)式代回(12)式，可以得到

$$\mathbf{P}_t = (\mathbf{I} - \mathbf{K}_t \mathbf{H}) \mathbf{P}_t^- \quad (15)$$

卡曼濾波器在執行時主要是以遞迴的方式進行，在某時間上的估測，得到測量的回授。一般來說，卡曼濾波器在執行時可以分爲兩部分：預測機制(Prediction)部分和修正機制(Correction)部分。在預估第 t 時間的系統狀態 $\hat{\mathbf{X}}_t$ 時，我們會先從第 $t-1$ 時間的系統狀態 $\hat{\mathbf{X}}_{t-1}$ 和第 t 時間的測量值 \mathbf{Z}_t 來決定第 t 時間的系統狀態；再將此狀態的估測值用回授方式來修正之前的預測。(11)(14)(15)即構成卡曼濾波器的修正機制。

接著要完成卡曼濾波器的預測機制。我們假定時間 t 的預估狀態為：

$$\hat{\mathbf{X}}_t^- = \mathbf{A}\hat{\mathbf{X}}_{t-1} + \mathbf{B}u_{t-1} \quad (16)$$

代入預估誤差的定義公式(11)中，得到

$$\mathbf{e}_t^- = \mathbf{X}_t - \hat{\mathbf{X}}_t^- = \mathbf{A}(\mathbf{X}_{t-1} - \hat{\mathbf{X}}_{t-1}) + \mathbf{W}_{t-1} \quad (17)$$

再將式(17)代入式(11)中，整理後得到

$$\begin{aligned} \mathbf{P}_t^- &= E[\mathbf{e}_t^- \mathbf{e}_t^{-T}] = E[(\mathbf{X}_t - \hat{\mathbf{X}}_t^-)(\mathbf{X}_t - \hat{\mathbf{X}}_t^-)^T] \\ &= E[(\mathbf{A}(\mathbf{X}_{t-1} - \hat{\mathbf{X}}_{t-1}) + \mathbf{W}_{t-1})(\mathbf{A}(\mathbf{X}_{t-1} - \hat{\mathbf{X}}_{t-1}) + \mathbf{W}_{t-1})^T] \\ &= \mathbf{A} \cdot E[(\mathbf{X}_{t-1} - \hat{\mathbf{X}}_{t-1})(\mathbf{X}_{t-1} - \hat{\mathbf{X}}_{t-1})^T] \cdot \mathbf{A}^T + E[\mathbf{W}_{t-1} \mathbf{W}_{t-1}^T] \\ &= \mathbf{A} \mathbf{P}_{t-1} \mathbf{A}^T + \mathbf{Q} \end{aligned} \quad (18)$$

(16)(18)即為預測機制的方程式。

由於卡曼增益 \mathbf{K}_t 的值決定了調整修正預測狀態的幅度，隨著卡曼濾波器遞迴運作下逐漸收斂，修正次數增加， \mathbf{P}_t 值會逐漸減小，預測精準度不斷提高；並且增益 \mathbf{K}_t 也隨之減小，也就是利用新量測資料改進預測估計值所佔的比重也隨時間減小。所以我們可以依照以上步驟，重複使用上一時態的估測值來預測下一時態的估測值，此回授的特性是以過去的狀態來決定未來的狀態，整個卡曼濾波器的詳細運作流程如圖 10 所示：

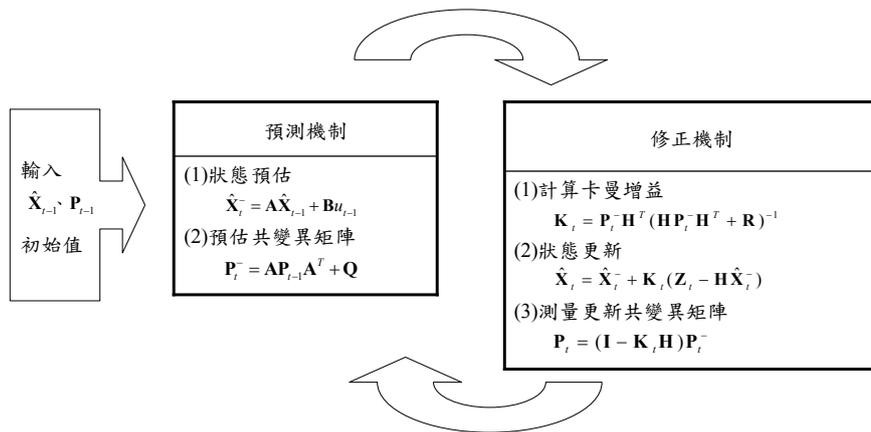


圖 10、卡曼濾波器運作流程

卡曼濾波器算法步驟：

完整的 Kalman Filter 估測演算法如下：

1. 連續兩張影像中正確得到人臉的位置，初始化向量為 $\mathbf{X}(0) = [x(2) \ y(2) \ [x(2)-x(1)]/T \ [y(2)-y(1)]/T]^T$ ，及初始化 $\mathbf{P}_t(0)$ 、 \mathbf{Q}_t 和 \mathbf{R}_t 矩陣。
2. 啟動 Kalman Filter，根據當前時間 t 狀態預估下一個狀態： $\mathbf{X}_t = \mathbf{A}\mathbf{X}_{t-1} + \mathbf{B}u_{t-1} + \mathbf{W}_{t-1}$ 。
3. 計算出下一時間的預估共變異矩陣 $\mathbf{P}_t^- = \mathbf{A}\mathbf{P}_{t-1}\mathbf{A}^T + \mathbf{Q}$ 。
4. 計算時間 t 的卡曼增益 $\mathbf{K}_t = \mathbf{P}_t^- \mathbf{H}^T (\mathbf{H}\mathbf{P}_t^- \mathbf{H}^T + \mathbf{R})^{-1}$ 。
5. 根據量測向量預測出最佳估測值 $\hat{\mathbf{X}}_t = \hat{\mathbf{X}}_t^- + \mathbf{K}_t (\mathbf{Z}_t - \mathbf{H}\hat{\mathbf{X}}_t^-)$ 。
6. 更新下一時間的預測誤差共變異矩陣 $\mathbf{P}_t = (\mathbf{I} - \mathbf{K}_t \mathbf{H})\mathbf{P}_t^-$ 。

隨後，根據時間 t 狀態可以得到最佳估測值，並更新了所有狀態向量，重複 2~6 就可以不斷迭代出新的預測值。

(ii) 核心函數理論分析

本理論是經由移動物體偵測方法，找出移動物體樣版影像的相關資訊後，便可作為移動物體追蹤時的依據。利用取得的樣版區域，計算特徵空間中每個特徵值的機率，當作目標模型的描述。隨後在每張畫面中採用平均位移疊代法(Mean Shift Iteration)0尋找影像中和樣版影像相似度最高的候選影像區塊，經過不斷的疊代計算，最後收斂到目標的真實位置，以達成追蹤的目的。

相似度的比對可利用色彩分佈、輪廓模型、移動向量等資訊的變化來估算，在本論文中是利用樣版影像的色彩分佈密度函數0，使用平均位移疊代法(Mean Shift Iteration)0來判斷樣版影像與候選影像的 Bhattacharya 係數相似度0，達到追蹤的目的。

樣版影像中心點的像素值資訊量較高，越接近邊緣的影像可能被遮蔽或受到背景顏色的影響，其貢獻度較低，核心函數的作用便是針對不同目標內不同位置的像素給予不同的權重分配，在樣版影像中，位置離目標中心點越遠的像素，如邊緣部分，權重逐漸降低；越接近目標中心點、越重要的像素的則權重較高。

核心函數為一中間權重高，四周權重低的函數，定義為：

d 維空間 R^d 上 n 個點的集合為 $\{\mathbf{x}_i\}_{i=1, \dots, n}$ ，多維變量的核心密度估計可寫成：

$$\bar{f}_k(x) = \frac{1}{nh^d} \sum_{i=1}^n k\left(\left\|\frac{x - x_i}{h}\right\|^2\right) \quad (19)$$

h 為核心函數的半徑長度， k 為核心函數並滿足以下條件：

1. k 為負的；
2. k 是非遞增的，若 $a < b$ ，則 $k(a) \geq k(b)$ ；
3. k 為分段連續的，且 $\int_0^{\infty} k(t)dt < \infty$ 。

為了使平均全體誤差(Average Global Error)最小，本文使用 Epanechnikov Kernel Function $K_E(x)$ 做為核心函數估算，定義為：

$$K_E(x) = \begin{cases} \frac{1}{2} C_d^{-1} (d+2) (1 - \|x\|^2) & \text{if } \|x\| < 1 \\ 0 & \text{otherwise} \end{cases} \quad (20)$$

令 $d = 2$ ，則核心函數如圖 11 所示。

C_d 為 d 維度空間中的面積， $d = 2$ 時為單位圓面積，因此 $C_d = \pi \times 1^2 = \pi$ 。

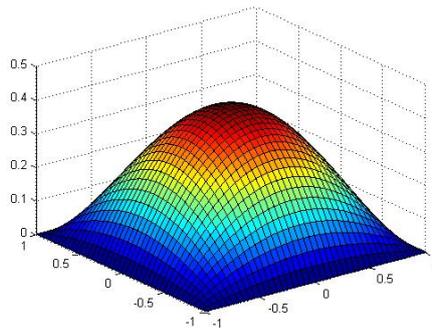


圖 11、Epanechnikov Kernel Function

經過移動物體偵測程序後，可得到移動目標的位置和大小模型，首先要選取該位置的色彩分佈資訊特徵。令人臉樣版影像的顏色特徵 u 具有色彩密度函數 $q(u)$ ，以 y 為中心點的人臉候選區域特徵色彩密度函數為 $p(u, y)$ ，核心函數演算法的目的是要找到位置 y ，使得候選色彩密度 $p(u, y)$ 與目標色彩密度 $q(u)$ 最為相似。

(1)目標模型：

樣版目標模組中某一顏色 u 的色彩分佈機率密度函數 $q(u)$ 可以表示為：

$$q(u) = C \sum_{i=1}^n k(\|x_i^*\|^2) \delta[b(x_i^*) - u] \quad (21)$$

$$C : \text{正規化常數}, C = \frac{1}{\sum_{i=1}^n k(\|x_i^*\|^2)}$$

$\delta(x)$: Kronecker Delta Function

其中， $\{x_i^*\}_{i=1\dots n}$ 表示樣版內所有像素點的集合，其像素總數為 n ，令其中心點為 0 ；函數 $b: R^2 \rightarrow \{1\dots m\}$ 是位置在 x_i^* 的像素值， $b(x_i^*)$ 回傳 x_i^* 相對應的色彩索引 (Color Index)，顏色特徵 $u = 1\dots m$ ； $[1:m]$ 為色彩索引的範圍，在數位攝影機所擷取的影像為 24 位元全彩影像 (R, G, B 各 8 位元)，所以色彩索引的範圍為 $[1:2^{24}]$ ，為了減少運算量，我們降低精確度至 2^{12} (R, G, B 各 4 位元)； C 為正規化常數，使得 $\sum_{u=1}^m q(u) = 1$ 恆成立； $k(\|x_i^*\|^2)$ 為核心函數。色彩分佈密度函數的求取可視為先獲得樣版影像的色彩直方圖，再將其正規化。

(2) 候選模型：

候選影像色彩分佈密度函數的定義 $p(u, y)$ 與 $q(u)$ 類似，差別是前者以 y 為中心；以 $\{x_i\}_{i=1\dots n}$ 來表示候選影像內所有像素位置的集合，並使用相同的核心函數 k ，其定義為：

$$p(u, y) = C_h \sum_{i=1}^n k\left(\left\|\frac{x_i - y}{h}\right\|^2\right) \delta[b(x_i) - u] \quad (22)$$

$$C_h : \text{正規化常數}, C_h = \frac{1}{\sum_{i=1}^{n_h} k\left(\left\|\frac{x_i - y}{h}\right\|^2\right)}$$

$\delta(x)$: Kronecker Delta Function

h : 核心函數之半徑長度

其中，正規化常數 C_h 的功能為使得 $\sum_{u=1}^m p(u, y) = 1$ 恆成立。

定義了移動物體的色彩分佈密度後，便可以目標影像為參考依據，與候選影像做比對，色彩索引越多，精確度越高，而相似度越高代表候選影像某一顏色 u 的色彩分佈機率是與目標影像最相似，在追蹤過程中，目標模型與候選模型的相似性越大，代表候選影像越接近目標真實位置，而相似係數經在下面章節中介紹。

相似性函數用於定義目標模型與候選模型之間的距離，在理想狀況下兩個模型的機率分佈是一樣的。本報告採用 Bhattacharyya 係數 ρ 來量測樣版影像 $q(u)$ 與候選影像 $p(u, y)$ 的相似度。假定對於比對位置的區域中任何位置 y ，目標出現的機率是相等的，其定義如下：

$$\rho(y) \equiv \rho[p(u, y), q(u)] = \int \sqrt{p(u, y)q(u)} du \quad (23)$$

其中 $p(u, y)$ 表示候選影像中第 u 個色彩索引中的量， $q(u)$ 表示樣版影像中第 u 個色彩索引中的量，爲了將 Bhattacharyya 係數運用至色彩分佈密度函數上，其型式可調整如下：

$$\rho(y) \equiv \rho[p(u, y), q(u)] = \sum_{u=1}^m \sqrt{p(u, y)q(u)} \quad (24)$$

由式(3.32)得出 $\rho(y)$ 的值在 $[0,1]$ 之間，且 $\rho(y)$ 的值越大，表示兩個模型越相似，定義兩個色彩分佈密度間的距離(Distance)爲：

$$d(y) = \sqrt{1 - \rho[p(u, y), q(u)]} \quad (25)$$

根據(25)式，Bhattacharyya 係數有下列三個特性：

1. $0 \leq \rho(y) \leq 1$ 。
2. $\rho(y) = 1$ ，則 $p(u, y) = q(u)$ 。
3. $\rho(y) = 0$ ，則 $p(u, y)$ 與 $q(u)$ 爲正交。

根據以上三個特性做推論，我們可以將 Bhattacharyya 係數當作是分類時所造成的錯誤量；當版樣影像的色彩分佈密度和候選影像的色彩分佈密度不相似時，Bhattacharyya 係數將會變得極小(錯誤量極小，導致很好分類)；相反的，當樣版影像的色彩分佈密度和候選影像的色彩分佈密度最近似時，Bhattacharyya 係數將會變得極大(錯誤量極大，所以不好分類)，因此找到 Bhattacharyya 係數的最大值，便可使兩個色彩分佈密度間的距離越小。

將利用泰勒展開式 (Taylor Expansion) $f(x) \approx f(x_0) + f'(x_0)(x - x_0)$ 對 Bhattacharyya 係數展開，以推導出如何尋找與版樣影像色彩分佈密度最相似的候選影像方法，使得 Bhattacharyya 係數有最大值。在泰勒展開式中，越高階的導數項，其重要性越低，於是我們只展開至一階導數項，而忽略其高階項，泰勒展開式對函數 $\rho[p(u, y)]$ 在 $P(u, y_0)$ 的一階導數項展開爲：

$$\rho[p(u, y)] \approx \rho[p(u, y_0)] + \rho'[p(u, y_0)][p(u, y) - p(u, y_0)] \quad (26)$$

式(24)代入(26)可得到以下結果：

$$\begin{aligned} \rho[p(u, y)] &\approx \rho[p(u, y_0)] + \rho'[p(u, y_0)][p(u, y) - p(u, y_0)] \\ &\approx \sum_{n=1}^m [p(u, y_0) \cdot q(u)]^{\frac{1}{2}} + \sum_{n=1}^m \frac{1}{2} [p(u, y_0) \cdot q(u)]^{\frac{1}{2}} \cdot q(u) \cdot [p(u, y) - p(u, y_0)] \\ &\approx \sum_{n=1}^m [p(u, y_0) \cdot q(u)]^{\frac{1}{2}} + \sum_{n=1}^m \left[\frac{1}{2} p(u, y_0)^{\frac{1}{2}} \cdot q(u)^{\frac{1}{2}} \cdot p(u, y_0) - \frac{1}{2} p(u, y_0)^{\frac{1}{2}} \cdot q(u)^{\frac{1}{2}} \right] \\ &\approx \frac{1}{2} \sum_{n=1}^m [p(u, y_0) \cdot q(u)]^{\frac{1}{2}} + \frac{1}{2} p(u, y) \cdot \left[\frac{q(u)}{p(u, y_0)} \right]^{\frac{1}{2}} \end{aligned} \quad (27)$$

將式(22)代入(27)：

$$\begin{aligned} \rho[p(u, y), q(u)] &\approx \sum_{u=1}^m \frac{1}{2} [p(u, y_0) \cdot q(u)]^{\frac{1}{2}} + \frac{1}{2} \sum_{n=1}^m p(u, y) \cdot \left[\frac{q(u)}{p(u, y_0)} \right]^{\frac{1}{2}} \\ &\approx \sum_{u=1}^m \frac{1}{2} [p(u, y_0) \cdot q(u)]^{\frac{1}{2}} + \frac{1}{2} \sum_{n=1}^m \left\{ C_h \sum_{i=1}^n k \left(\left\| \frac{x_i - y}{h} \right\|^2 \right) \cdot \delta[b(x_i) - u] \cdot \left[\frac{q(u)}{p(u, y_0)} \right]^{\frac{1}{2}} \right\} \\ &\approx \sum_{n=1}^m \frac{1}{2} [p(u, y_0) \cdot q(u)]^{\frac{1}{2}} + \frac{C_h}{2} \sum_{i=1}^n \left\{ \sum_{n=1}^m \delta[b(x_i) - u] \cdot \left[\frac{q(u)}{p(u, y_0)} \right]^{\frac{1}{2}} \right\} k \left(\left\| \frac{x_i - y}{h} \right\|^2 \right) \\ &\approx \frac{1}{2} \sum_{n=1}^m [p(u, y_0) \cdot q(u)]^{\frac{1}{2}} + \frac{C_h}{2} \sum_{i=1}^n w_i k \left(\left\| \frac{x_i - y}{h} \right\|^2 \right) \end{aligned} \quad (28)$$

$$\text{其中 } w(x_i) = \sum_{u=1}^m \delta[b(x_i) - u] \sqrt{\frac{q(u)}{p(u, y_0)}} \quad (29)$$

爲了得到最大的相似度，式(28)的 Bhattacharyya 係數必須達到最大值，但是 $\rho(y)$ 的大小是取決於自變數 y 的選擇。其中第一項的值由 y_0 決定，爲一常數。因此，主要影響 Bhattacharyya 係數的是第二項，爲了使 Bhattacharyya 係數達到最大，式(28)中的第二項

$\frac{C_h}{2} \sum_{i=1}^n w_i k \left(\left\| \frac{x_i - y}{h} \right\|^2 \right)$ 需達最大值。接下來我們將推導以平均位移法進行疊代，以求最大值，

並找出相對應的候選影像。

平均位移向量(Mean Shift Vector)是一個函數朝著機率密度梯度方向的函數，它亦是此一函數增加量最大的方向，以下我們便要推導如何用平均位移法求得 Bhattacharyya 係數最大值。

給予一個機率密度函數 $f(x)$ ，在 d 維空間中已知 n 個點 x_i ， $i = 1, \dots, n$ ， $f(x)$ 的核心密度估計如式(19)，對函數進行梯度運算可以得到：

$$\nabla \bar{f}(x) = \frac{2}{nh^{d+2}} \sum_{i=1}^n (x - x_i) k' \left(\left\| \frac{x - x_i}{h} \right\|^2 \right) \quad (30)$$

定義 $g(x)$ 為 $k(x)$ 的負導函數 $g(x) = -k'(x)$ 代入式(30)可以得到：

$$\begin{aligned} \nabla \bar{f}(x) &= \frac{2}{nh^{d+2}} \sum_{i=1}^n (x - x_i) g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right) \\ &= \frac{2}{nh^{d+2}} \times \left[\sum_{i=1}^n g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right) \right] \times \left[\frac{\sum_{i=1}^n x_i g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right)}{\sum_{i=1}^n g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right)} - x \right] \end{aligned} \quad (31)$$

函數 $f(x)$ 之梯度向量的方向便是函數 $f(x)$ 其增加量最大的方向，此梯度向量便是我們所要求之平均位移向量，可將之表示為：

$$M(x) = \frac{\sum_{i=1}^n x_i g \left(\left\| \frac{x_i - x}{h} \right\|^2 \right)}{\sum_{i=1}^n g \left(\left\| \frac{x_i - x}{h} \right\|^2 \right)} - x \quad (32)$$

由式(32)所推導出的平均移動向量，可以用來計算 Bhattacharyya 係數最大值，亦即式(28)

中的第二項 $\frac{C_h}{2} \sum_{i=1}^n w_i k \left(\left\| \frac{x_i - y}{h} \right\|^2 \right)$ 和式(19)函數 $\bar{f}_k(x)$ 具有相同形式的平均位移向量，因此平均位移向量可改寫為：

$$M(x) = \frac{\sum_{i=1}^n x_i w_i g\left(\left\|\frac{x_i - x}{h}\right\|^2\right)}{\sum_{i=1}^n w_i g\left(\left\|\frac{x_i - x}{h}\right\|^2\right)} - x \quad (33)$$

現在我們考慮式(20)核心函數 $K_E(x)$ ，經過變數代換，滿足 $K_E(x) = k(\|x\|^2)$ ，因為我們處理的影像是二維，所以 d 值為 2， C_d 為單位圓的面積 $C_d = \pi \times 1^2 = \pi$ ，代入式(20)得到：

$$k(x) = \begin{cases} \frac{2}{\pi}(1-x) & \text{if } x < 1 \\ 0 & \text{otherwise} \end{cases} \quad (34)$$

令 $g(x)$ 為 $k(x)$ 的負導函數， $g(x) = -k'(x)$ 代入式(34)：

$$g(x) = \begin{cases} \frac{2}{\pi} & \text{if } x < 1 \\ 0 & \text{otherwise} \end{cases} \quad (35)$$

將 $g(x)$ 代入(33)可以得到：

$$M(x) = \begin{cases} \frac{\sum_{i=1}^n x_i w(x_i)}{\sum_{i=1}^n w(x_i)} - x & \text{if } \left\|\frac{x_i - x}{h}\right\|^2 < 1 \\ -x & \text{otherwise} \end{cases} \quad (36)$$

因為 $\left\|\frac{x_i - x}{h}\right\|^2$ 必定小於 1，式(36)可改寫為：

$$M(x) = \frac{\sum_{i=1}^n x_i w(x_i)}{\sum_{i=1}^n w(x_i)} - x \quad (37)$$

使用 Epanechnikov 核心函數找出核心密度的平均位移向量，並計算出 Bhattacharyya 係數的最大值，此步驟可簡化成式(37)，因此可藉由平均位移法快速且有效的計算與目標影像有最大的相似度的候選影像，達到追蹤的效果。

在上一節中定義了目標影像的色彩分佈密度函數 $q(u)$ 與候選影像的色彩分佈密度函數 $p(u, y)$ ，及計算兩者間的 Bhattacharyya 相似係數，並計算出平均位移向量來進行平均位移疊代演算法，進而疊代出和目標影像最為相似且中心點為 y_0 之候選影像，找到中心點為 y_0 的候選影像，即算完成一次移動物體樣版影像追蹤，之後不斷的重複此步驟，便能夠持續的對移動物體的樣版影像進行追蹤。經由移動物體偵測流程得到移動物體的區域及座標，此移動物體的中心點為 y_0 ，我們以 y_0 為平均位移法的起始位置，執行平均位移法的疊代，其流程如下：

1. 由偵測到的移動物體的區域及座標作為目標影像，依據

$q(u) = C \sum_{i=1}^n k(\|x_i^*\|^2) \delta[b(x_i^*) - u]$ 計算出對應的 $\{q(u)\}_{u=1\dots m}$ ，以目標影像的中心點 y_0 為起始位置。

2. 取得起始位置 y_0 後，根據 $p(u, y) = C_h \sum_{i=1}^{n_h} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \delta[b(x_i) - u]$ 計算出候選影像相對應的 $\{p(u, y_0)\}_{u=1\dots m}$ 。

3. 將樣版影像色彩分佈密度 $\{q(u)\}_{u=1\dots m}$ 與候選影像色彩分佈密度 $\{p(u, y_0)\}_{u=1\dots m}$ 代入 Bhattacharyya 係數 $\rho[p(u, y_0), q(u)] = \sum_{u=1}^m \sqrt{p(u, y_0)q(u)}$ ，計算出 Bhattacharyya 係數 $p(y_0)$ 。

4. 根據 $w(x_i) = \sum_{u=1}^m \delta[b(x_i) - u] \sqrt{\frac{q(u)}{p(u, y_0)}}$ 計算出 $\{w(x_i)\}_{i=1\dots n_k}$ 。

5. 根據 $M(y_0) = \frac{\sum_{i=1}^{n_h} x_i g\left(\left\|\frac{x_i - y_0}{h}\right\|^2\right) w(x_i)}{\sum_{i=1}^{n_h} g\left(\left\|\frac{x_i - y_0}{h}\right\|^2\right) w(x_i)} - y_0$ 計算出候選影像的新位置

$$y_1 = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x_i - y_0}{h}\right\|^2\right) w(x_i)}{\sum_{i=1}^n g\left(\left\|\frac{x_i - y_0}{h}\right\|^2\right) w(x_i)}。$$

6. 計算出候選影像的新位置 y_1 後，便可以更新 $\{p(u, y_1)\}_{u=1\dots m}$ 與

$$\rho[p(u, y_1), q(u)] = \sum_{u=1}^m \sqrt{p(u, y_1)q(u)}。$$

7. 當 $\rho[p(u, y_1), q(u)] < \rho[p(u, y_0), q(u)]$ 時，修正 $y_1 = \frac{1}{2}(y_0 + y_1)$ 。

我們必須先定義一個門檻值 ε ，若 y_0 和 y_1 之間的位移量小於此門檻值 ε 時，便停止疊代，也就是當 $\|y_1 - y_0\| < \varepsilon$ 則停止疊代，否則更新 y_0 ，使 $y_0 = y_1$ ，回到步驟 2 中繼續進行疊代演算，如圖 12。

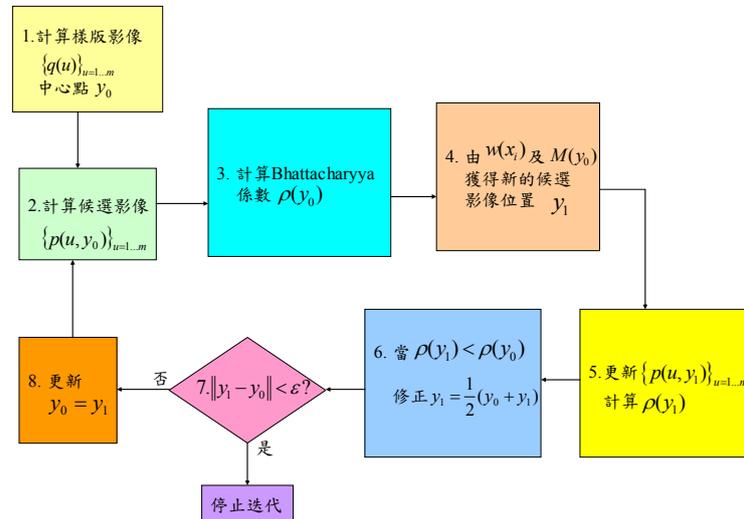


圖12 平均位移法實作流程圖

經由圖 12 的過程得知，平均位移演算法的疊代過程是從樣版影像的中心點 y_0 不斷移動到 y_1 的過程，每一個平均位移向量移動一次，相似性係數的值便會增加，最終達到相似性係數最大值之處。當我們進入移動物體追蹤處理程序後，經由以上步驟的反覆疊代演算，我們可以藉著計算出相似度最高的候選影像，進而完成應用平均位移演算法來追蹤移動物體版樣工作，達到追蹤的效果。

五、視窗整合系統

人臉辨識視窗系統之功能瀏覽，如圖 13 所示。在此系統中包含了基礎影像處理、光照處理、辨識演算法及門禁警監系統。基礎影像處理中有 5 個部分，分別為色彩轉換、邊緣遮罩、型態學遮罩、二值處理及相關係數等。光照處理中有 7 個方法，分別為直方圖等化 (Histogram Equalization, HE)、權重區域直方圖等化 (weighted Block HE, wBHE)、離散餘弦轉換 (Discrete Cosine Transform, DCT)、伽瑪亮度修正 (Gamma Intensity Correction, GIC)、區域正規化 (Local Normalization, LN)、自商數影像 (Self Quotient Image, SQI) 及反射比 (Reflection Ration, RR)。辨識演算法依照矩陣及運算結構的不同，分為三個主要類別，如一維辨識演算法有四種方法，二維單向辨識演算法有四種方法，二維雙向辨識演算法有兩種方法。最後則是介紹我們所出的演算法在門禁警監系統上的應用及性能的分析。

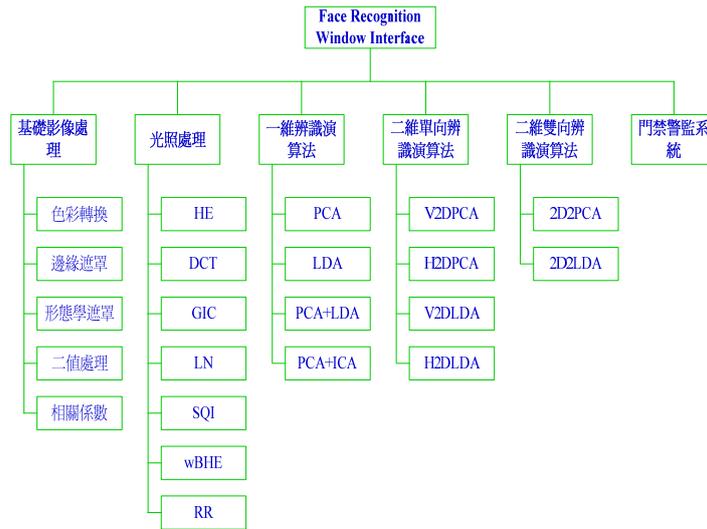


圖 13、視窗整合追蹤辨識介面

(a) 門禁警監系統流程圖

圖 14 為門禁警監系統功能方塊圖，其中包含了三大部份，影像前處理、光照正規化及辨識處理，下面將逐項說明與討論。第一部份影像前處理之輸出為進入畫面之受測者之臉部影像，此一部份包括背景的更新、分割與膚色區域的偵測，最後擷取臉部，第二部份為光照正規化之輸出，主要是針對光影的影響，利用各種光照正規化的演算法，將影像進行光照正規化以減少光照造成辨識錯誤之發生，第三部份辨識過程之輸出，則是利用各種統計學上一維或是二維的辨識演算法，將之用於辨識受測影像與資料庫中取得的受測者身份，驗證兩者是否吻合，因而獲得受測者之辨識結果。以下就針對光照均勻化與各種辨識理論分別加以說明：

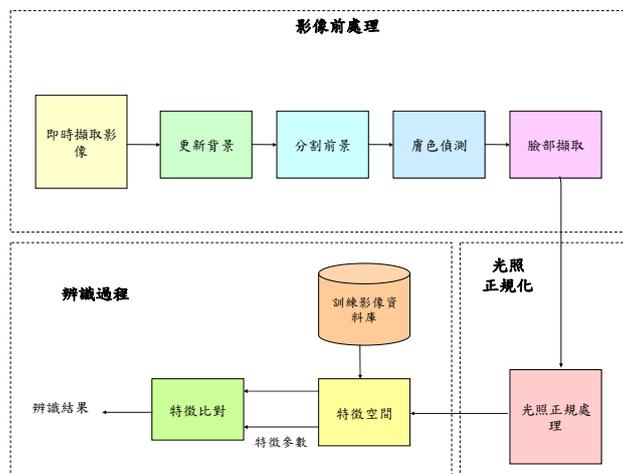


圖 14、門禁警監系統功能方塊圖

(b) 光照均勻化

(i) Histogram Equalization

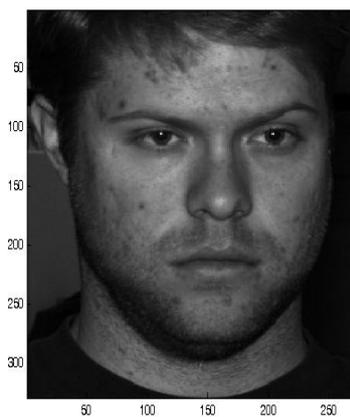
在亮度正規化的方法中，亮度統計值等化(HE)0 之方法是最常被拿來作為比較的均勻化方法之一，任何影像其像素灰階值並不會平均的分布，而亮度統計值等化的想法即將影像上的像素灰階值分佈的機率加以拉近，因此可使用亮度統計值等化，使整張影像之像素灰階值分佈較為平均，此法可增加整體影像的對比，使影像整體較為明顯。其中之參數 N_{Mean} 為所有像素點之灰階值之平均值， r_k 為轉換後的階值。

轉換公式如下：

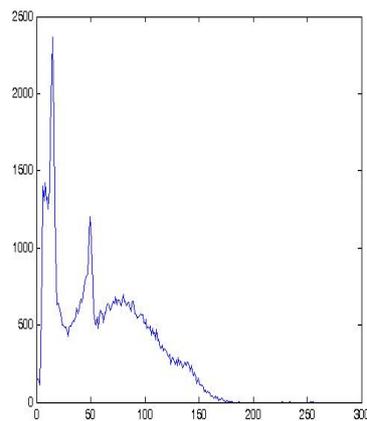
$$N_{Mean} = \frac{N_{all}}{256} \quad (38)$$

$$r_k = T(k) = \left(\frac{\sum_{i=1}^k N_k}{N_{Mean}} \right) \quad (39)$$

其中參數 N_k 為影像中灰階值為 k 時的像素數目， N_{all} 為影像中所有像素總數， $T(k)$ 為轉換函數。



(a)



(b)

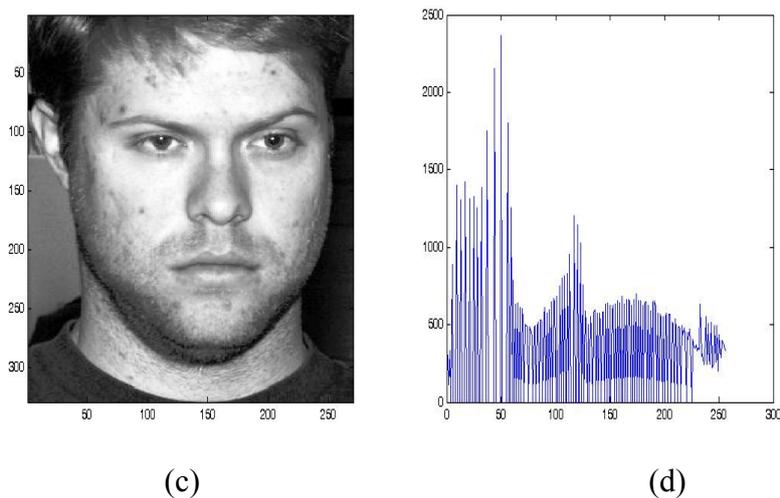


圖15 (a)原始影像(b)原始影像之直方圖
(c)亮度統計值等化後之影像(d)HE處理後之直方圖

由圖 15(b)之直方圖可看出，大部分之灰階值集中於 0-150 而致使整體影像偏暗，經以亮度統計值等化後，得到圖 15(c)，整體之灰階值顯然較為平均分配至 0-255 之間，整體亮度與對比皆較圖 14 明顯。

藉由灰階值前後的轉換，可得到圖 16 之灰階值對應表，由此表可看出灰階值前後轉換的關係。

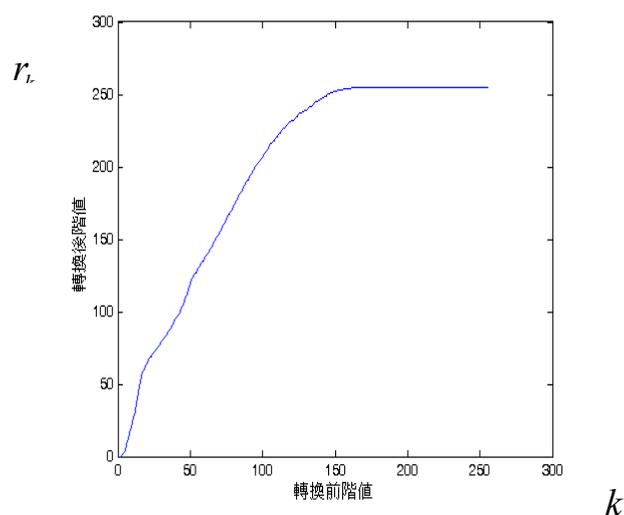


圖16 階值轉換對應圖

視窗介面之直方圖等化參數設定及結果如下：

程式如圖 17，在選擇 Histogram Equalization 後會彈出圖 18 之畫面，此時可設定是否要更改處理的圖片維度，按下 Preview 可看到原始影像，按下 Process 則可看到處理之後的影像，最後點選 OK，則會將 NDU-Media Lab 上的影像從事均勻化運算結果，如圖 19 所示。

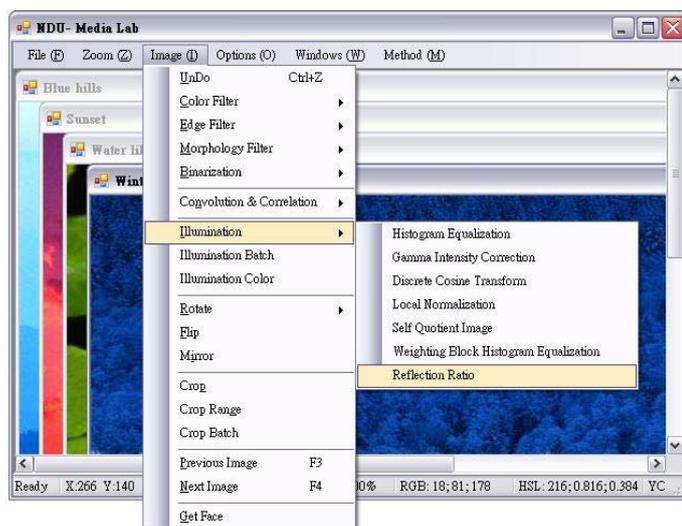


圖 17 光照正規化模擬主畫面

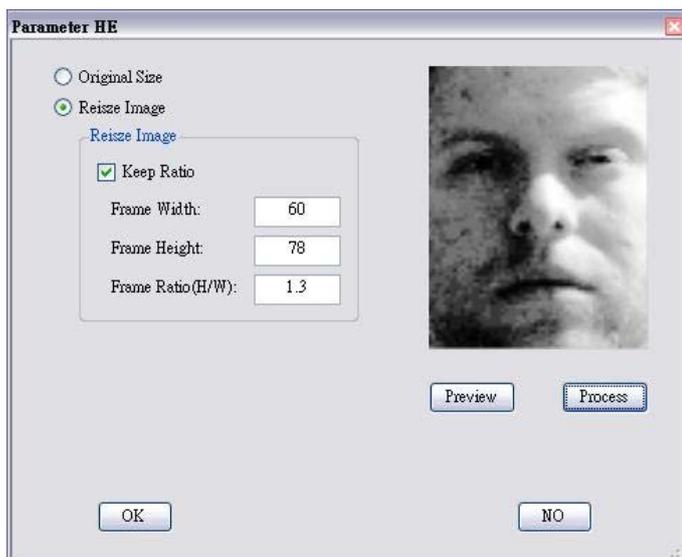


圖18 HE光照處理表單

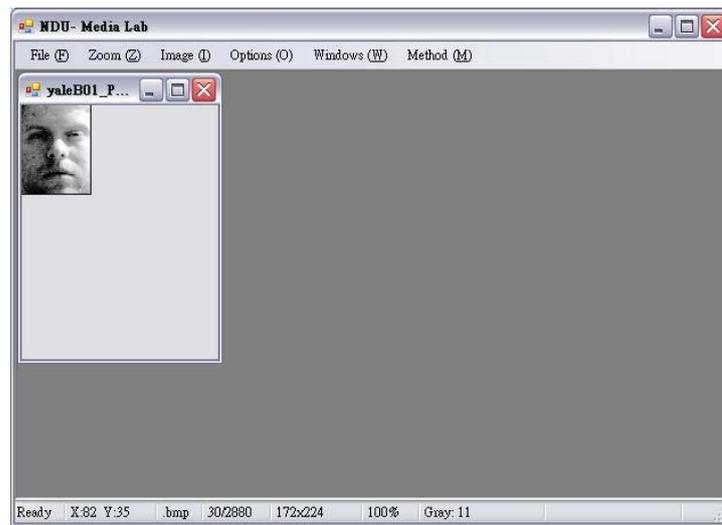


圖19 HE光照處理後的結果

(ii) Weighting Block Histogram Equalization

此節介紹 權重區域直方圖等化 (wBHE)，其做法與 Block-based histogram equalization(BHE)0 類似，不同於其它權重設定，此想法是針對當影像左右或上下亮度不均時，將影像切分成 4 個區域，然後每個區域分別進行亮度統計值等化，最後再將這 4 個影像依照不同的權重如圖 20 所示，將影像合併成一個影像。

假設原影像寬為 $width$ 、高為 $height$ ，區分成 4 個區塊為 B_i ， $i = 1, 2, 3, 4$ ，每個區塊的寬為 SBw 、高為 SBh ，由於 $SBw = \alpha \times width$ 、 $SBh = \alpha \times height$ ，且 $0.5 \leq \alpha < 1$ ，所以每個區塊必定有重疊區域，隨後分別設定每個區塊的權重 w_i ， $0 \leq w_i \leq 1$ ， $i = 1, 2, 3, 4$ ，以消除重疊邊處的邊界效應。

此四區塊再合成之影像為

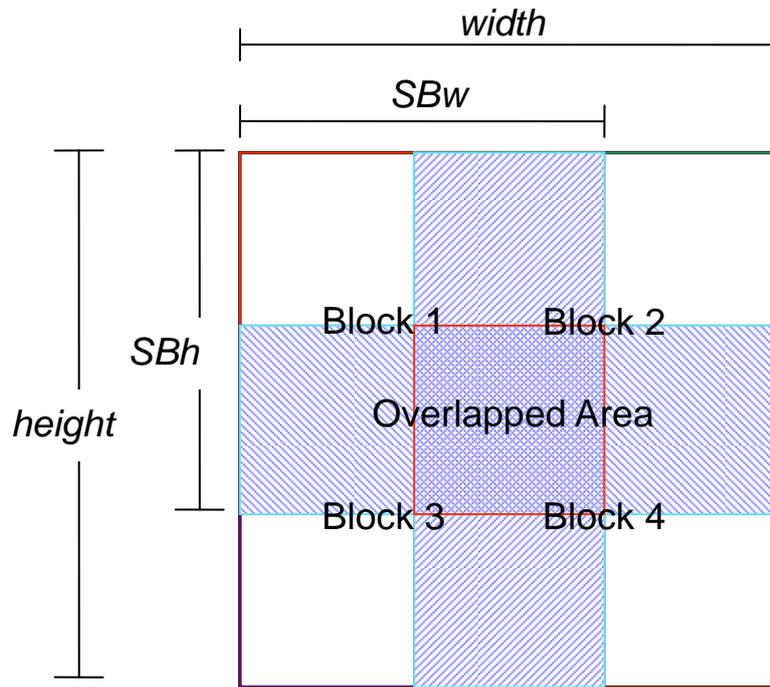


圖20 切分4區塊的示意圖

$$f(x, y) = \sum_{i=1}^4 w_i(x, y) \times B_i(x, y) \quad (40)$$

$$B_i(x, y) = \text{Crop}(\text{Org Image}, SBw, SBh), \quad i = 1, 2, 3, 4 \quad (41)$$

利用切割影像之函式 $\text{Crop}()$ ，將原始影像 Org Image 分割為四個區塊， B_i ， $i = 1, 2, 3, 4$ 。

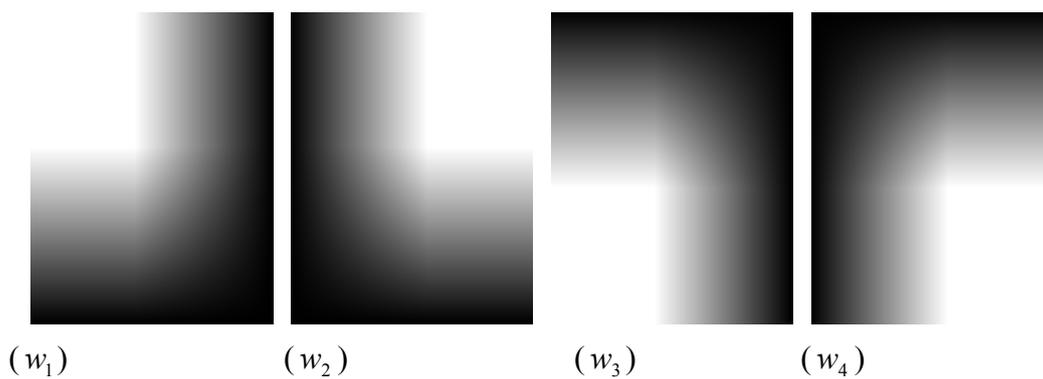


圖21 w_i 之權重影像, $i=1, 2, 3, 4$

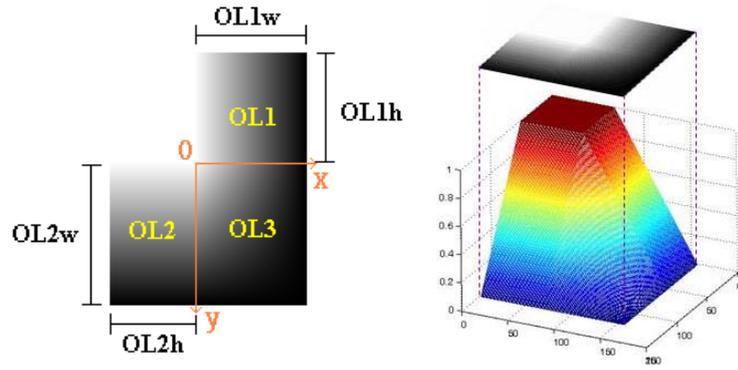


圖22 權重影像示意圖，區塊1之重疊區域分別為OL1-3

圖 21 為區塊 1-4 的權重影像。以區塊 1 來介紹其重疊區塊之 OL1~OL3 的權重為例，如圖 21 所示。

OL1~OL3 之權重分別為 $w_{1-OL1} \sim w_{1-OL3}$ 表示如下：

$$w_{1-OL1}(x, y) = w'_{1-OL1}(x) \times w'_{1-OL1}(y)$$

$$w'_{1-OL1}(x) = 1 - \frac{x}{OL1w} \quad (42)$$

$$w'_{1-OL1}(y) = 1$$

$$w_{1-OL2}(x, y) = w'_{1-OL2}(x) \times w'_{1-OL2}(y)$$

$$w'_{1-OL2}(x) = 1 \quad (43)$$

$$w'_{1-OL2}(y) = 1 - \frac{y}{OL2h}$$

$$w_{1-OL3}(x, y) = w'_{1-OL1}(x) \times w'_{1-OL2}(y) \quad (44)$$

其餘部分為 1(如圖 22 白色部分)。最後透過剛所得之各 w_i 、 B_i 計算出式(39)之 $f(x, y)$ ， $f(x, y)$ 即為權重區域直方圖等化之結果。

圖 23 為介面系統之 wBHE 光照處理的參數設定表單，圖 24 為其處理結果，透過此系統設定不同 α 值來進行討論。

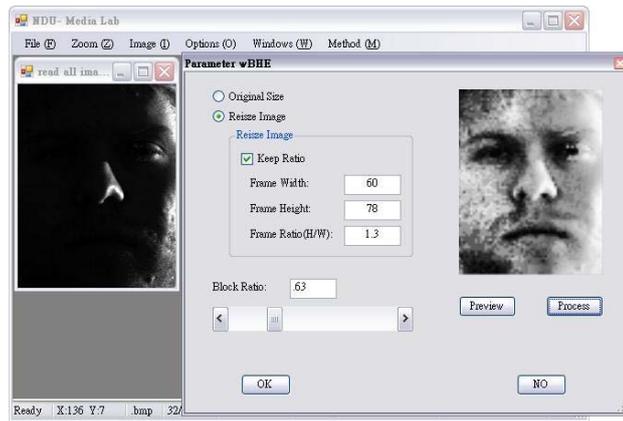


圖23 wBHE光照處理的參數設定表單

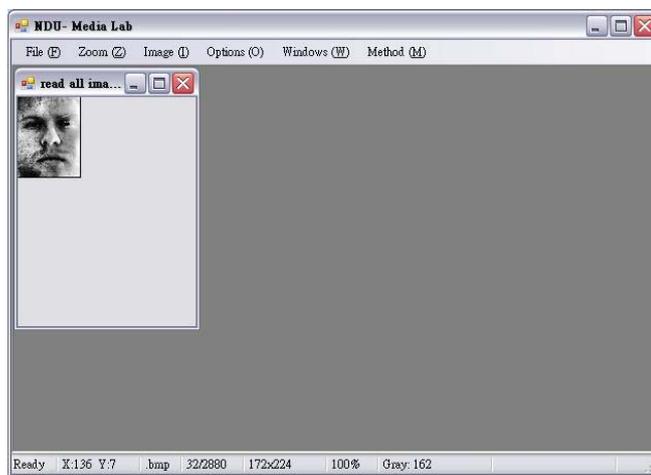


圖24 wBHE光照處理的結果

圖 25 為原始影像，圖 26-27 為 wBHE 後之影像，透過不同的 α ，可看出當 α 為 0.5 時，其邊界有分割線存在，至 0.65 時已較為緩和，當 α 為 0.7 時已可忽略邊界之分割線，且影像左右之亮度也相當接近，有不錯的亮度等化效果，如圖 26、27 所示。



圖25 原始影像



圖26 α 分別為0.5、0.55、0.65



圖27 α 分別為0.7、0.8、0.9

(iii) Discrete Cosine Transform

離散餘弦轉換 θ ，於一般使用中，主要是進行影像壓縮所用，其運算後之結果，可取得 DC 係數及 AC 係數值，DC 係數為一直流值，而 AC 係數則為一高頻係數，如圖 28 所示，此方法則是透過修正 DC 及 AC 係數，藉以調整整張影像之亮度分配。將整張影像進行離散餘弦轉換，為了增加影像整體亮度，故 θ 中其做法修改了 DC 係數為原先之 1.1 倍，為了使整體影像較為平緩部份之灰階值修改至相近，因此針對前 35 個 AC 係數值修改至 50，其數值設定則是以實驗所得，最後再透過反離散餘弦轉換，將頻域轉換回空間域。圖 28 為 DCT 係數的區塊特徵。

考慮一張原始影像，其亮度為 $f(x, y)$ ，則此影像之離散餘弦轉換(DCT)及反離散餘弦轉換(IDCT)分別為

$$C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \times \cos\left[\frac{\pi(2x+1)u}{2M}\right] \cos\left[\frac{\pi(2y+1)v}{2N}\right]$$

其中 $C(0,0) = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)}{M \times N}$ 。

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v)C(u, v) \times \cos\left[\frac{\pi(2x+1)u}{2M}\right] \cos\left[\frac{\pi(2y+1)v}{2N}\right]$$

其中 $\alpha(u) = \begin{cases} \frac{1}{\sqrt{M}} & ; u = 0 \\ \frac{2}{\sqrt{M}} & ; u = 1, 2, \dots, M-1 \end{cases}$, $\alpha(v) = \begin{cases} \frac{1}{\sqrt{N}} & ; v = 0 \\ \frac{2}{\sqrt{N}} & ; v = 1, 2, \dots, N-1 \end{cases}$ 。

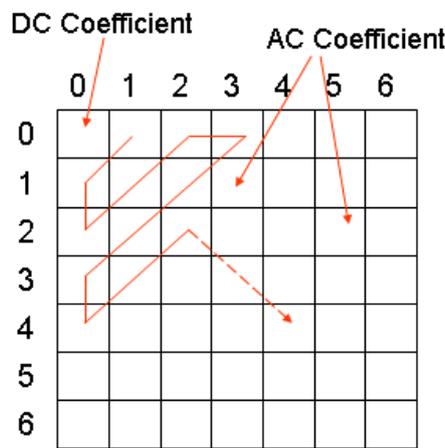


圖28 DCT係數的區塊特徵

圖 29-30 可看出透過本節的做法，由於修改前 35 個 AC 參數，所以部份較低頻的階值，已被改為較為相近，如臉頰處；留下較高頻的邊緣處不變，如眼睛鼻子嘴巴的邊緣特徵。

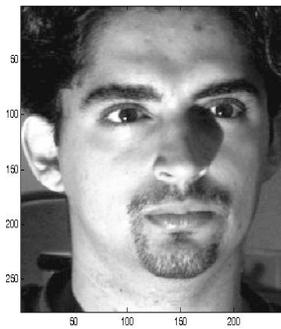


(a)

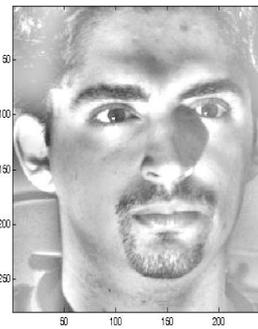


(b)

圖29 (a)已做HE處理影像(b) DCT處理後影像



(a)



(b)

圖30 (a)已做HE處理影像(b) 接著做DCT處理影像

(iv) Gamma Intensity Correction

伽瑪強度修正(GIC)0，其想法是將影像上各點進行開 γ 次方運算，藉以降低明亮之亮度差異，但所處理之結果並無法判斷其好壞，故需要一張正常光源之參考影像(如圖 31)，將 GIC 處理之結果與此張參考影像進行差異平方和之運算，取其最小值視為最佳處理之結果。

假設正常光源影像為 $I_0(x, y)$ ，原始影像為 I_{xy} ，GIC 處理之結果 I'_{xy} 為

$$I'_{xy} = G(I_{xy}, \gamma^*) \quad (45)$$

$$G(I_{xy}; \gamma) = c \times I_{xy}^{\frac{1}{\gamma}} \quad (46)$$

$$\gamma^* = \arg \min_{\gamma} \sum_{x,y} [G(I_{xy}; \gamma) - I_0(x, y)]^2 \quad (47)$$

$I_0(x, y)$ 為正常光源之參考影像， I_{xy} 為光照影響之影像， I'_{xy} 為 GIC 轉換後之最佳影像， γ^* 為最佳 γ 值， c 為灰階延伸係數，其值設定為將 $I_{xy}^{\frac{1}{c}}$ 中之最大值正規至 255。藉由代入不同 γ 至式 (46) 中， γ 的範圍為 1-10，將 I_0 與 $G(I_{xy}; \gamma)$ 點對點相減並平方再相加，其值為最小者時之 $G(I_{xy}; \gamma)$ ，即為轉換後之最佳影像 I'_{xy} 。



圖31 GIC參考影像

圖 33 及圖 35 皆使用圖 31 做為參考影像所處理之結果，可看出使用 GIC 處理後，可縮短影像左右兩側光照的亮度影響，並使影像左側的邊緣特徵強化，即使影像左側相當黑暗，也有不錯的亮度正規效果。



圖32 尚未做GIC之原始影像(與參考影像同人)



圖33 GIC處理後之影像(與參考影像同人)

圖 35 所使用之參考影像同樣為圖 31，可看出參考影像與處理影像並不同人，但也有不錯的亮度結果，除了將黑暗處亮化之外，也保持了人臉上的邊緣特徵。



圖34 尚未做GIC之原始影像(與參考影像不同人)



圖35 GIC處理後之影像(與參考影像不同人)

圖 36 為系統中 GIC 的參考影像輸入表單，由圖 37 設定 γ 值，可由手動輸入或系統透過式(45-47)自動算出 γ 值，參數設定完畢後可由 Preview 按鈕看到原始影像，經由 Process 按鈕則可看到處理之後的影像，最後點選 OK，則會將圖 37 之不均勻光照影像修改如圖 38 之均勻光照影像。

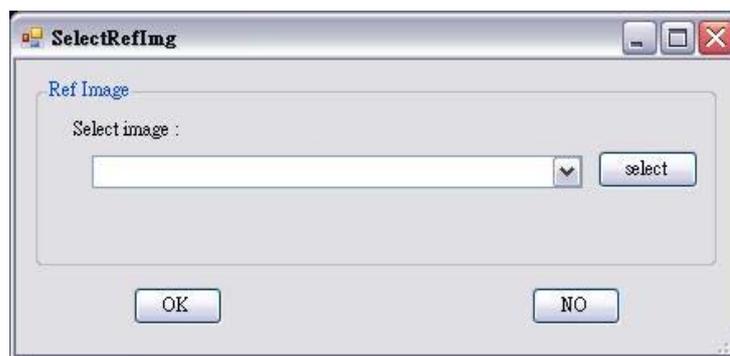


圖36 GIC光照處理的參照輸入表單



圖37 GIC光照處理的參數設定表單



圖38 GIC光照處理的結果

(v) Local Normalization

區域正規化(LN)做法有點類似統計學中常態分佈正規成標準常態分佈的過程，主要以均值 μ 及變異數 σ^2 來做處理，假設正常光源影像為 $f(x, y)$ ，其正規方法，則是將此影像先減去其均值影像再除以變異數影像。

假設 $f'(x, y)$ 為光照影像，定義轉換矩陣 A 為

$$\begin{aligned} A &= f'(x, y) / f(x, y) \\ \Rightarrow f'(x, y) &= A \times f(x, y), \quad (x, y) \in W, \end{aligned} \quad (48)$$

$(x, y) \in W$ ， W 是整張影像點的位置，因此 $f'(x, y) = Af(x, y)$ ，假如考慮雜訊影像 B ，則 $f'(x, y)$ 可表為

$$f'(x, y) = A(x, y) \times f(x, y) + B(x, y) \quad (49)$$

是以影像之平均值與變異量分別為

$$\begin{aligned}
E(f'(x, y)) &= E(A \times f(x, y) + B) \\
&= A \times E(f(x, y)) + B, \quad (x, y) \in W,
\end{aligned} \tag{50}$$

$$\begin{aligned}
Var(f'(x, y)) &= \sqrt{\frac{\sum (f'(x, y) - E(f'(x, y)))^2}{N}} \\
&= A \times \sqrt{\frac{\sum (f(x, y) - E(f(x, y)))^2}{N}} \\
&= A \times Var(f(x, y)), \quad (x, y) \in W,
\end{aligned} \tag{51}$$

N 為影像中的總像素數，則影像之正規化(Normalize)像素(pixel)為

$$f'_p(x, y) = \frac{f'(x, y) - E(f'(x, y))}{Var(f'(x, y))} \tag{52}$$

$$\begin{aligned}
f'_p(x, y) &= \frac{f'(x, y) - E(f'(x, y))}{Var(f'(x, y))} \\
&= \frac{A \times f(x, y) + B - (A \times E(f(x, y)) + B)}{A \times Var(f(x, y))} \\
&= \frac{f(x, y) - E(f(x, y))}{Var(f(x, y))} = f_p(x, y), \quad (x, y) \in W
\end{aligned} \tag{53}$$

其中 $f'_p(x, y)$ 為光照影像 $f'(x, y)$ 之正規影像， $f_p(x, y)$ 為正常光源影像 $f(x, y)$ 之正規影像，經由(53)式，證明 $f'_p(x, y) = f_p(x, y)$ 。

在 LN 過程中的平均影像及變異數影像，為遮罩處理之影像，因此也必須選擇區塊的大小，區塊太小易造成雜訊；區塊太大則處理時間過長，在此區塊大小之選擇為 7x7，是透過後面實驗討論而來。

圖 39 為系統之區域正規化處理的參數設定表單，主要有一個遮罩大小的參數設定，當影像較大時需要有較大的遮罩設定，圖 40 為參數設定好之處理結果。

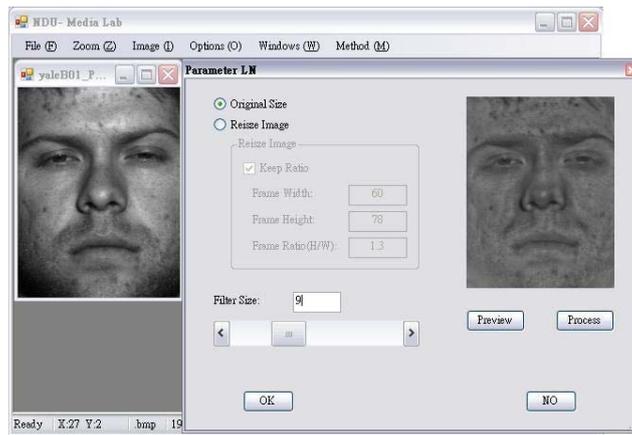


圖39 LN光照處理的參數設定表單

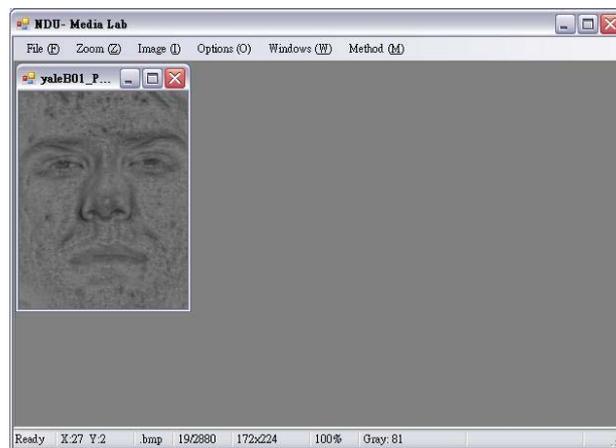


圖40 LN光照處理的結果

以圖示說明當不同的遮罩大小所造成的影響，圖 41 為原始影像及 HE 處理影像，用以強化影像對比。圖 42 為平均影像，當選取之遮罩邊界越大，其影像也越模糊。圖 43 為變異數平均影像。圖 44 看出當遮罩太小時，其影像失去一些邊緣特徵，遮罩大小設為 7 時，除了顯示出臉部邊緣特徵，並已消除影像左右亮暗之差異，因此本做法之遮罩大小為 7x7。



圖41 (左)原始影像，(右)HE後影像



圖42 平均影像遮罩邊界大小依序為3、7、11、15、23



圖43 變異數平均影像遮罩邊界大小依序為3、7、11、15、23



圖44 LN影像遮罩邊界大小依序為3、7、11、15、23

(vi) Self Quotient Image

自商數影像(SQI)[52,53]主要計算商數影像 Q ，商數影像是由光照影像除以光照影像之低通影像所得，其目的在於取得影像之邊緣特徵，藉以達成去除光照影響。

自商數影像對亮化具有不變性，SQI 處理主要針對邊緣特徵，即使影像遭受光照影響而亮化，臉部仍會保留邊緣特徵，因此可由 SQI 處理以進行正規化，定義自商數影像如下：

$$Q = \frac{I}{\hat{I}} = \frac{I}{F * I} \quad (54)$$

其中 \hat{I} 為自影像 I 之低通濾波， F 為低通濾波， Q 為自商數影像。

處理自商數影像的過程中，由式(54)得知只需用到低通濾波，故此處 F 採用高斯濾波器並加上權重來做處理，見圖 45。

此處定義低通濾波如下：

$$F = G \times W \quad (55)$$

$$\frac{1}{N} \sum_{\Omega} WG = 1 \quad (56)$$

$$\tau = \text{Mean}(I_{\Omega}) \quad (57)$$

$$W(i, j) = \begin{cases} 0, I(i, j) \in M_2 \\ 1, I(i, j) \in M_1 \end{cases} \quad (58)$$

$$I' = D(Q) \quad (59)$$

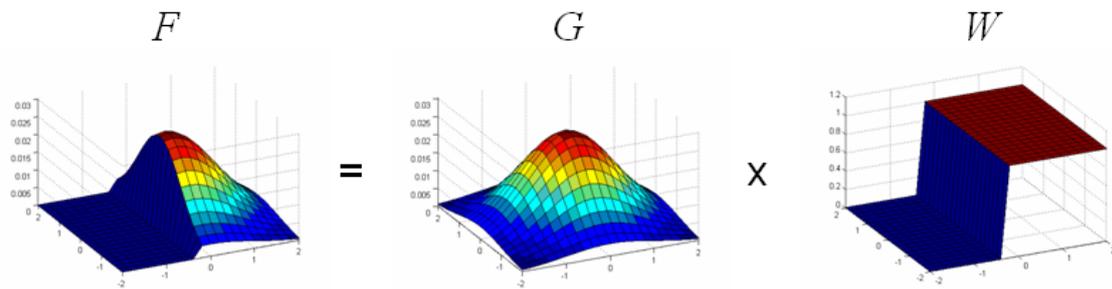


圖45 高斯濾波示意圖

其中式(56)為遮罩係數 W 乘 G 之正規化， W 為遮罩權重，其數值為 $[0,1]$ 矩陣， G 為高斯濾波之遮罩係數， $\frac{1}{N}$ 為正規化常數， Ω 為遮罩內之各點像素。 I_{Ω} 為影像 I 之遮罩待處理之像素， τ 為影像 I 之遮罩內像素的平均值，其為一門檻值，由 τ 將遮罩中矩陣數值 I_{Ω} 分為 2 群， I_{Ω} 指影像 I 之遮罩內像素值，多數的設為 M_1 ，少數則設為 M_2 ，目的為 W 乘 G 遮罩運算後，可減少影像邊界暈開的情況發生，在此我們遮罩大小設定為 7×7 。最後透過式(54)及式(55)取得 Q ，再使用式(59)得正規光照影像 I' ， D 為正規化函數，以本報告中則是將影像灰階值乘一常數值 200。

圖 46 為系統中自商數影像的參數設定表單，其中包含之參數有遮罩大小及灰階延伸係數。圖 47 為自商數影像處理完成的結果，可看出陰影處之臉部特徵皆有明顯改善。

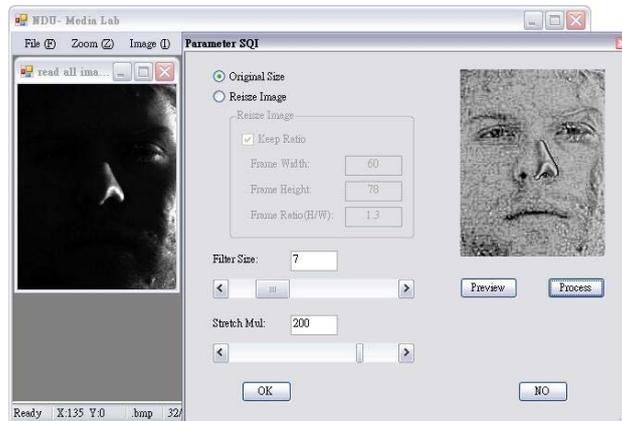


圖46 SQI光照處理的參數設定表單

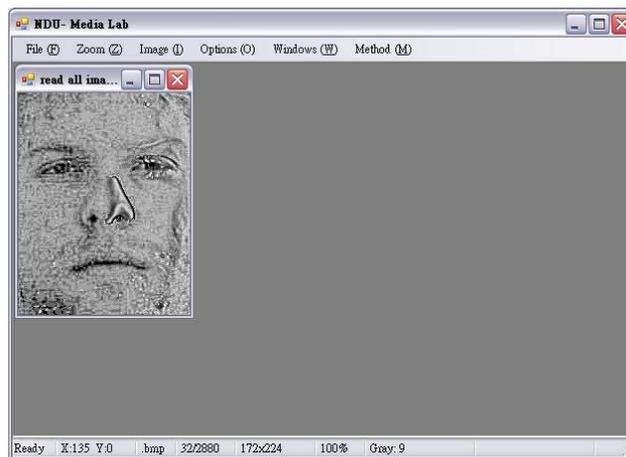


圖47 SQI光照處理的結果

(vii) Reflection Ratio

ReflectionRatio(RR)是參考球面調合的再亮化方法⁰，球面調和的再亮化方法是透過球面調和函數 Y_{lm} 以及入射光 L_{lm} 去尋找最相近原始影像的反射光場 $B(\alpha, \beta)$ ，因此由不同的反射光場，可將影像轉換至任一光線條件，我們基於此想法，以低通濾波方式找出有別於球面調和之反射光場。

定義球面調和之反射光場為

$$B(\alpha, \beta) = \sum_{l=0}^{\infty} \sum_{m=-l}^l B_{lm} Y_{lm}(\alpha, \beta) \quad (60)$$

其中 B_{lm} 為反射光， Y_{lm} 為球面調和函數。人臉可視為 Lambertian surface，透過球面調和表示為

$$\begin{aligned} I(x, y) &= \rho(x, y) \sum_l \sum_m \Lambda_l \hat{\rho}_l L_{lm} Y_{lm}(\theta, \phi) \\ &= \rho(x, y) \sum_l \sum_m B_{lm} Y_{lm}(\theta, \phi) \end{aligned} \quad (61)$$

其中 $\rho(x, y)$ 為影像的反照度， Λ 為正規化因子， L_{lm} 為入射光， $\hat{\rho}_l$ 為 cosine 的轉換函數。

若有另一入射光 $L(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l L_{lm} Y_{lm}(\theta, \phi)$ ，則影像於新光照條件下轉變為

$$\begin{aligned} I'(x, y) &= \rho(x, y) \sum_l \sum_m \Lambda_l \hat{\rho}_l L_{lm} Y_{lm}(\theta, \phi) \\ &= \rho(x, y) \sum_l \sum_m B'_{lm} Y_{lm}(\theta, \phi) \end{aligned} \quad (62)$$

由式(61)-(62)，當光照條件變化後，其反射光也跟著變化，故原始影像可由各自的反射光場來還原，如式(63)。

$$I(x, y) = I'(x, y) \frac{\sum_l \sum_m B_{lm} Y_{lm}(\theta, \phi)}{\sum_l \sum_m B'_{lm} Y_{lm}(\theta, \phi)} \quad (63)$$

介紹我們所發展的 RR 之反射光場的取得，將正常光源影像以及光照條件影像個別做模糊處理，這邊的模糊處理採用低通遮罩 F ，正常光源影像以 13x13 低通處理(即為 $B(x, y)$)，

在此遮罩大小是以實驗所得，當參照影像之低通遮罩太大時，進行正規處理後的影像會發生邊緣暈開的情形，遮罩太小則無法顯示反射光場的影像，會趨近為原圖，因此遮罩以 13x13 來做處理，見圖 48。光照條件影像則以 3x3 低通處理(即為 $B'(x, y)$)，目的為取得邊緣資訊，因此影像使用一般 3x3 低通遮罩即可，無需較大遮罩，見圖 49，而處理後之模糊影像即以此設為反射光場，最後透過下式取得正規後影像。

$$B(x, y) = F_{mean13} * I(x, y) \quad (64)$$

$$B'(x, y) = F_{mean3} * I'(x, y) \quad (65)$$

$$I_{new}(x, y) = I'(x, y) \frac{B(x, y)}{B'(x, y)} \times w_1 + B'(x, y) \times w_2 \quad (66)$$

其中 $I(x, y)$ 為正規之參考影像， $I'(x, y)$ 為光照影像， F_{mean3} 及 F_{mean13} 分別為 3x3 及 13x13 低通濾波遮罩，透過式(66)取得正常光照影像 I_{new} ， w_1 及 w_2 用以調整 RR 處理時，縮小影像明暗邊界處的差異，其值分別為 0.8、0.2。



圖48 正常光照條件影像之模糊處理($B(x, y)$)



圖49 上排為原始影像，中排為模糊處理($B'(x, y)$)，下排為修正後影像



圖50 與上圖皆使用相同之參考影像，上排為原始影像，中排為模糊處理($B'(x,y)$)，下排為修正後影像

圖 51 為 RR 之參數設定表單，可調整參考影像之均化遮罩、光照影像之均化遮罩以及處理後之亮化程度。圖 52 則是 RR 處理結果。

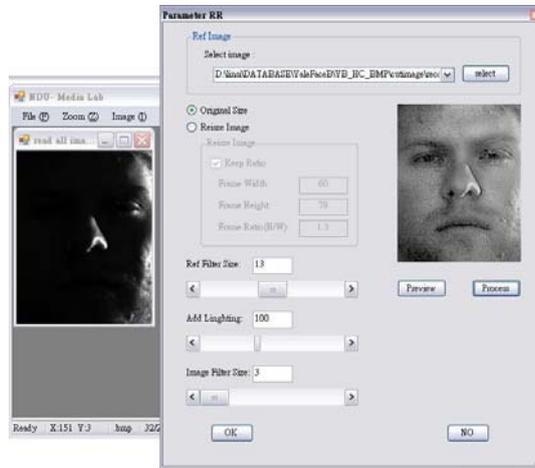


圖51 RR光照處理的參數設定表單



圖52 RR光照處理的結果

(c) 辨識演算法

(i) PCA

主成分分析法(PCA, Principal Components Analysis)[55,56]乃是將資料轉換至特徵空間以便於分析及處理。PCA 為線性的轉換方法，主要是產生一個正交投射基底(Orthogonal Projection Basis)來從事去相關(de-correlation)的動作，以降低資料的維度(Dimension)並作為特徵選取之參考，其特徵空間(Eigenspace)是藉由對整組訓練影像資料執行特徵分析，並由全域共變異矩陣(Global Covariance Matrix)計算出特徵向量(Eigenvector)所獲得，特徵向量的選取是對應非零的特徵值(Eigenvalue)所對應的特徵向量，由大至小排序，再由這些特徵向量產生正規化基底(Orthonormal Basis)，這些基底用來做為我們所需的基底函數，所有的影像都可藉由此基底函數投影至特徵空間，以實施分類與比較。

假設共有 N_T 張影像， $N_T = m \times n$ ，其中有 m 個不同人物，每個人物各有 n 張影像，則第 1 個人物的第 1 張影像為 \mathbf{X}_{11} ，假設影像高為 NY ，寬為 NX ，則 \mathbf{X}_{11} 為 $(NY \times NX) \times 1$ 的一維影像向量， \mathbf{X} 為 N_T 張影像向量之矩陣，其矩陣大小為 $(NY \times NX) \times N_T$ ，如圖 53 所示。 \mathbf{X} 定義為

$$\mathbf{X} = [\mathbf{X}_{11}, \mathbf{X}_{12}, \dots, \mathbf{X}_{1n}, \mathbf{X}_{21}, \mathbf{X}_{22}, \dots, \mathbf{X}_{2n}, \dots, \mathbf{X}_{m1}, \mathbf{X}_{m2}, \dots, \mathbf{X}_{mn}] \quad (67)$$

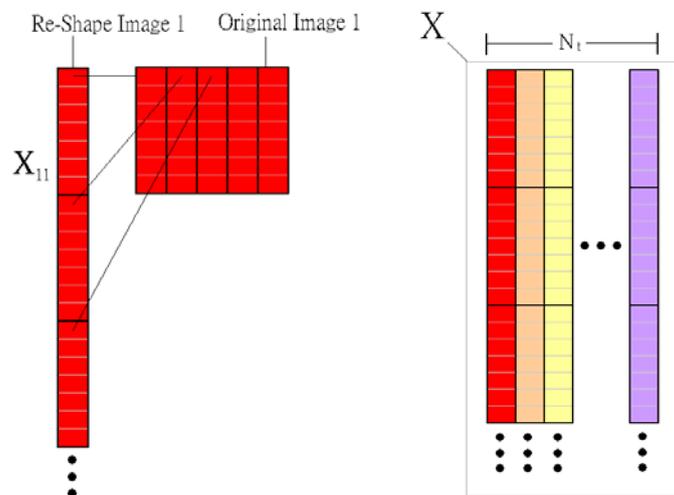


圖53 影像矩陣示意圖

則平均向量 $\bar{\mathbf{X}}$ 為

$$\bar{\mathbf{X}} = \frac{\sum_i^m \sum_j^n \mathbf{X}_{ij}}{N_T} \quad (68)$$

令

$$\mathbf{A}_{ij} = \mathbf{X}_{ij} - \bar{\mathbf{X}} \quad (69)$$

$$\mathbf{A} = [\mathbf{A}_{11}, \mathbf{A}_{12}, \dots, \mathbf{A}_{1n}, \mathbf{A}_{21}, \mathbf{A}_{22}, \dots, \mathbf{A}_{2n}, \dots, \mathbf{A}_{m1}, \mathbf{A}_{m2}, \dots, \mathbf{A}_{mn}] \quad (70)$$

則 \mathbf{A} 之共變異矩陣 \mathbf{C} ，可表為

$$\mathbf{C} = \mathbf{A} \times \mathbf{A}^T \quad (71)$$

其中 T 為轉置矩陣(Transpose Matrix)。由於影像 \mathbf{X}_{ij} 的高具 NY 個畫素，寬具 NX 個畫素，故矩陣 \mathbf{A} 的大小為 $(NY \times NX) \times N_T$ ，共變異矩陣 \mathbf{C} 的大小為 $(NY \times NX) \times (NY \times NX)$ 。

共變異矩陣之特徵值可由 $\det(\mathbf{C} - \lambda \mathbf{I}) = 0$ 解出。設特徵值為 λ ，由 $\mathbf{C}\mathbf{P} = \lambda\mathbf{P}$ 可得特徵向量 \mathbf{P} ，將特徵值由大至小排序，取得自定之前 n 個最大特徵值 $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ 及其所對應之特徵向量 $\mathbf{P} = [\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_n]$ ，則 \mathbf{A} 可投影至特徵空間 \mathbf{W} 。

$$\mathbf{W} = \mathbf{P}^T \times \mathbf{A} \quad (72)$$

圖 54 為 PCA 參數設定表單。圖 55-56 為 PCA 處理後的辨識結果，左方為受測影像，右方為候選影像。圖 57 為處理 PCA 過程中之產出。圖 58(a)為所有 λ 之排序結果(b)為平均臉(即 $\bar{\mathbf{X}}$)(c)為 30 個特徵向量分別轉為二維影像所得之特徵臉。

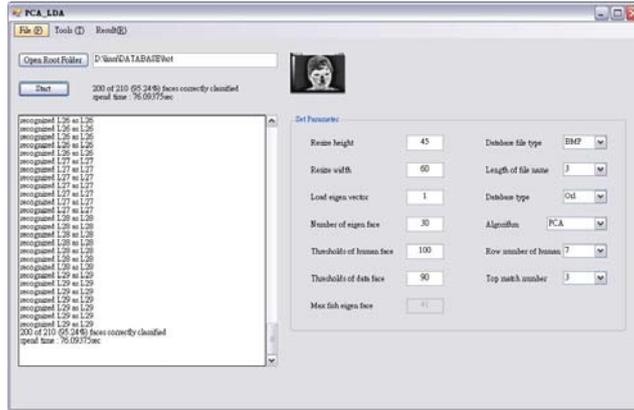


圖54 系統PCA參數設定表單

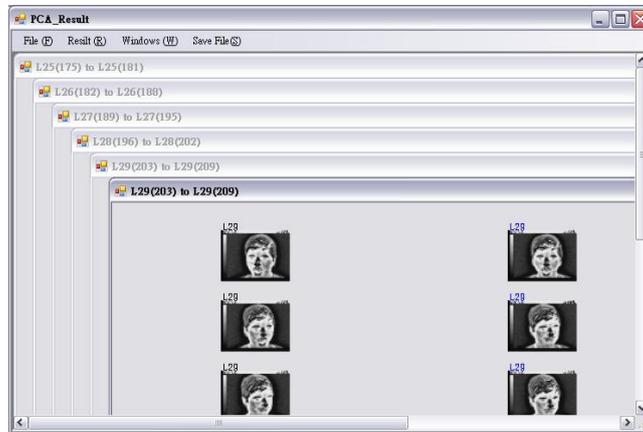


圖55 系統PCA辨識結果

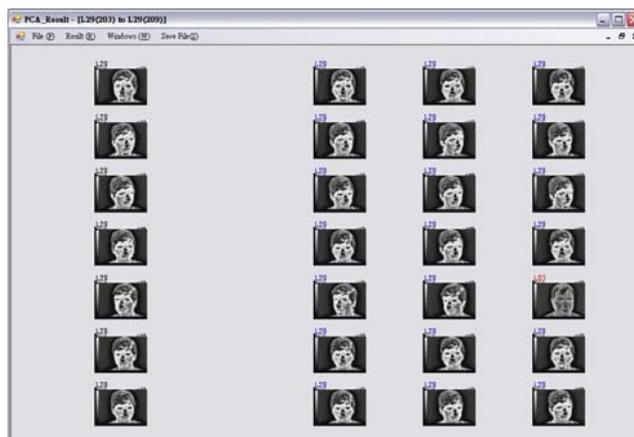


圖56 系統PCA辨識結果

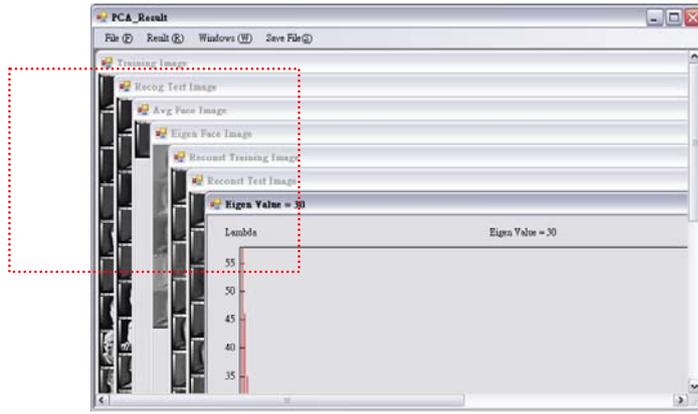
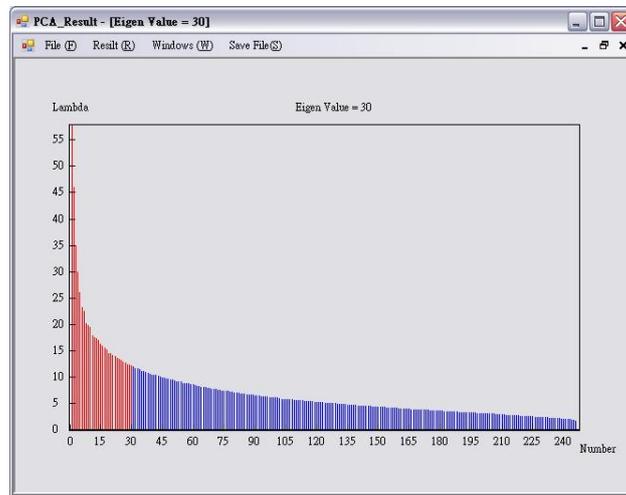


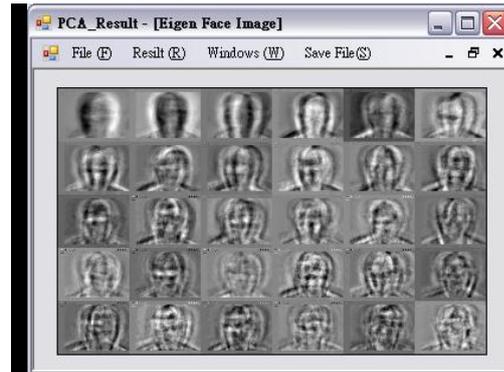
圖57 系統PCA過程中之產出



(a)



(b)



(c)

圖58 系統PCA辨識結果(a) λ 之排序(b)平均臉(c)特徵值 λ 所對應之特徵臉

(ii) LDA

線性鑑別分析法(LDA, Linear Discriminate Analysis)0 的主要概念是尋找一個最佳的投影

空間，將資料進行投影後，使同一類別的資料加以集中，而不同類別的資料加以分開。因此其重點在於拉大不同類別間(between-class)的距離，並縮小類別內(within-class)的距離。

圖 59 為 LDA 最佳投影示意圖，當類別 X_1 及 X_2 投影至 W_2 空間，可看出資料有重疊現象，並無法明確的分類出是否為相同類別，而投影至 W_1 空間，則明顯區別出類別 X_1 及 X_2 ，因此線性鑑別分析法主要為尋求最佳的投影空間 W_1 。

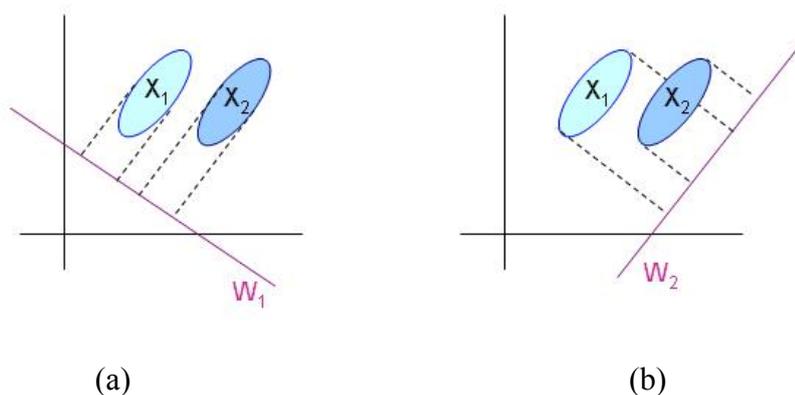


圖59 LDA類別最佳投影示意圖(a)最佳投影(b)錯誤投影

令 m 為可分類數， n 為每類之影像數， N_T 為所有的影像總數。

定義 \mathbf{X} 為

$$\mathbf{X} = [\mathbf{X}_{11}, \mathbf{X}_{12}, \dots, \mathbf{X}_{1n}, \mathbf{X}_{21}, \mathbf{X}_{22}, \dots, \mathbf{X}_{2n}, \dots, \mathbf{X}_{m1}, \mathbf{X}_{m2}, \dots, \mathbf{X}_{mn}] \quad (73)$$

則影像平均向量為

$$\bar{\mathbf{X}} = \frac{\sum_i^m \sum_j^n \mathbf{X}_{ij}}{N_T} \quad (74)$$

每類之平均向量為

$$\bar{\mathbf{X}}_i = \frac{\sum_j \mathbf{X}_{ij}}{n}, \quad i = 1, 2, \dots, m \quad (75)$$

定義類別間散佈矩陣 \mathbf{S}_B (Between-class Scatter Matrix)，以及類別內散佈矩陣 \mathbf{S}_W (Within-class Scatter Matrix)分別為

$$\mathbf{S}_B = \sum_i^m \left[(\bar{\mathbf{X}}_i - \bar{\mathbf{X}})(\bar{\mathbf{X}}_i - \bar{\mathbf{X}})^T \right] \quad (76)$$

$$\mathbf{S}_W = \sum_i^m \sum_j^n \left[(\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)(\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)^T \right] \quad (77)$$

類別間散佈矩陣 \mathbf{S}_B 表示為類別間資料的分散程度，是由每個類別的平均值影像減去全部影像的平均值，再計算其共變異矩陣的總合所得。類別內散佈矩陣 \mathbf{S}_W 表示為類別內資料的分散程度，是由每張影像減去每個類別的平均影像，再計算其共變異矩陣的總合所得。

假設影像 \mathbf{X}_{ij} 的高具 NY 個畫素，寬具 NX 個畫素，則 \mathbf{S}_B 及 \mathbf{S}_W 之維度分別為 $(NY \times NX) \times (NY \times NX)$ 。

與主要成分分析法類似，特徵值之取得，可由 $\det(\mathbf{S}_W^{-1}\mathbf{S}_B - \lambda\mathbf{I}) = 0$ 解出，再以 $\mathbf{C}\mathbf{P} = \lambda\mathbf{P}$ 得出特徵向量 \mathbf{P} 。將特徵值由大至小排序，取得前 n 個最大特徵值 $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ 及其所對應之特徵向量 $\mathbf{P} = [\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_n]$ ，則 \mathbf{A} 可投影至特徵空間 \mathbf{W} 。

$$\mathbf{W} = \mathbf{P}^T \times \mathbf{A} \quad (78)$$

(iii) 2DPCA

二維主成分分析法(2DPCA, Two-Dimensional Principal Components Analysis)[58-59]，其運算時影像矩陣有別於 PCA 作法，係直接以影像之二維資料從事計算，如圖 60 所示。

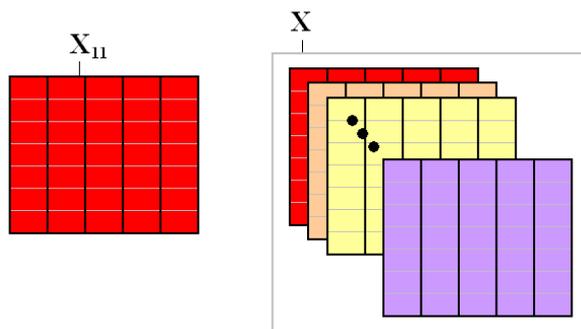


圖60 二維影像矩陣示意圖

令 m 為可分類數， n 為每類之影像數， N_T 為所有的影像總數。定義 N_T 張影像矩陣為

$$\mathbf{X} = [\mathbf{X}_{11}, \mathbf{X}_{12}, \dots, \mathbf{X}_{1n}, \mathbf{X}_{21}, \mathbf{X}_{22}, \dots, \mathbf{X}_{2n}, \dots, \mathbf{X}_{m1}, \mathbf{X}_{m2}, \dots, \mathbf{X}_{mn}] \quad (79)$$

則影像之平均矩陣為

$$\bar{\mathbf{X}} = \frac{\sum_i^m \sum_j^n \mathbf{X}_{ij}}{N_T} \quad (80)$$

令

$$\mathbf{A}_{ij} = \mathbf{X}_{ij} - \bar{\mathbf{X}} \quad (81)$$

$$\mathbf{A} = [\mathbf{A}_{11}, \mathbf{A}_{12}, \dots, \mathbf{A}_{1n}, \mathbf{A}_{21}, \mathbf{A}_{22}, \dots, \mathbf{A}_{2n}, \dots, \mathbf{A}_{m1}, \mathbf{A}_{m2}, \dots, \mathbf{A}_{mn}] \quad (82)$$

則行向的變異矩陣 \mathbf{C}_{ij_cov} 可定義為

$$\mathbf{C}_{ij_cov} = \mathbf{A}_{ij} \times \mathbf{A}_{ij}^T \quad (83)$$

因此 2DPCA 之共變異矩陣 \mathbf{C} 可表為

$$\mathbf{C} = \frac{\sum_i^m \sum_j^n \mathbf{C}_{ij_cov}}{N_T} \quad (84)$$

最後 2DPCA 的共變異矩陣 \mathbf{C} 由 \mathbf{C}_{ij_cov} 之平均所得。假設影像 \mathbf{X}_{ij} 的高具 NY 個畫素，寬具 NX 個畫素，則矩陣 \mathbf{A} 的大小為 $NY \times NX$ ，共變異矩陣 \mathbf{C} 的維度為 $NY \times NY$ ，故運算速度遠快於 PCA。

與主要成分分析法相同，共變異矩陣之特徵值可由 $\det(\mathbf{C} - \lambda \mathbf{I}) = 0$ 解出。設特徵值為 λ ，由 $\mathbf{C}\mathbf{P} = \lambda\mathbf{P}$ 可得特徵向量 \mathbf{P} ，將特徵值由大至小排序，取得自定之前 n 個最大特徵值 $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ 及其所對應之特徵向量 $\mathbf{P} = [\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_n]$ ，則 \mathbf{A} 可投影至特徵空間 \mathbf{W} 。

$$\mathbf{W} = \mathbf{A} \times \mathbf{P} \quad (85)$$

(iv) 2DLDA

二維線性鑑別分析法(2DLDA, Two-Dimensional Linear Discriminant Analysis)0 的主要目的為修改 LDA 過程中產生的 \mathbf{S}_B 及 \mathbf{S}_W 之 $(NY \times NX) \times (NY \times NX)$ 高維矩陣，因此計算 LDA 時，可說是相當消耗時間與空間。2DLDA 作法大致與 LDA 相同，如圖 59 所示，差別在於輸入矩陣與 2DPCA 類似，皆直接處理二維影像，省去轉為一維向量，已達到降低維度的效果，如圖 60 所示。

令 m 為可分類數， n 為每類之影像數， N_T 為所有的影像總數。定義 N_T 張影像矩陣為

$$\mathbf{X} = [\mathbf{X}_{11}, \mathbf{X}_{12}, \dots, \mathbf{X}_{1n}, \mathbf{X}_{21}, \mathbf{X}_{22}, \dots, \mathbf{X}_{2n}, \dots, \mathbf{X}_{m1}, \mathbf{X}_{m2}, \dots, \mathbf{X}_{mn}] \quad (86)$$

則影像平均矩陣為

$$\bar{\mathbf{X}} = \frac{\sum_i^m \sum_j^n \mathbf{X}_{ij}}{N_T} \quad (87)$$

每類之影像平均矩陣為

$$\bar{\mathbf{X}}_i = \frac{\sum_j^n \mathbf{X}_{ij}}{n}, \quad 1 \leq i \leq m \quad (88)$$

定義類別間散佈矩陣 \mathbf{S}_B (Between-class Scatter Matrix)，以及類別內散佈矩陣 \mathbf{S}_W (Within-class Scatter Matrix)，分別為

$$\mathbf{S}_B = \sum_i^m \left[(\bar{\mathbf{X}}_i - \bar{\mathbf{X}})^T (\bar{\mathbf{X}}_i - \bar{\mathbf{X}}) \right] \quad (89)$$

$$\mathbf{S}_W = \sum_i^m \sum_j^n \left[(\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)^T (\mathbf{X}_{ij} - \bar{\mathbf{X}}_i) \right] \quad (90)$$

類別間散佈矩陣 \mathbf{S}_B 表示為類別間資料的分散程度，是由每個類別的平均值影像減去全部影像的平均值，再計算其共變異矩陣的總合所得。類別內散佈矩陣 \mathbf{S}_W 表示為類別內資料的分散程度，是由每張影像減去每個類別的平均影像，再計算其共變異矩陣的總合所得。

假設影像 \mathbf{X}_{ij} 的高具 NY 個畫素，寬具 NX 個畫素，則 \mathbf{S}_B 及 \mathbf{S}_W 之維度為 $NX \times NX$ ，與 LDA 的 $(NY \times NX) \times (NY \times NX)$ 相較之下，故計算速度遠快於 LDA。

與線性鑑別分析法相同，特徵值之取得，可由 $\det(\mathbf{S}_W^{-1} \mathbf{S}_B - \lambda \mathbf{I}) = 0$ 解出，再以 $\mathbf{C}\mathbf{P} = \lambda \mathbf{P}$ 得出特徵向量 \mathbf{P} 。將特徵值由大至小排序，取得前 n 個最大特徵值 $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ 及其所對應之特

徵向量 $\mathbf{P} = [\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_n]$ ，則 \mathbf{A} 可投影至特徵空間 \mathbf{W} 。

$$\mathbf{W} = \mathbf{A} \times \mathbf{P} \quad (91)$$

(v) 特徵比對

在本節利用各種方法獲得訓練影像之特徵向量，將訓練影像透過特徵向量投影至特徵空間，同理，測試資料庫也可透過特徵向量投影至特徵空間，將訓練影像與測試影像之投影向量計算其歐幾里德距離，於一個閾值內以最近的距離視為影像相近之辨識結果，若其值大於閾值則表示為無法辨識此人。

假設訓練影像投影至特徵空間的特徵矩陣 $\mathbf{W}_{train,i}$ 定義為

$$\mathbf{W}_{train,i} = [\mathbf{W}_{train,i,1}, \mathbf{W}_{train,i,2}, \dots, \mathbf{W}_{train,i,L}] = \mathbf{P}^T \times \mathbf{A}_{train,i} \quad (92)$$

其中 $\mathbf{A}_{train,i}$ 為訓練影像 \mathbf{A}_{train} 中的第 i 張影像矩陣， L 為向量長度， \mathbf{P} 為特徵向量矩陣。

所有訓練影像的特徵矩陣表為

$$\mathbf{W}_{train} = [\mathbf{W}_{train,1}, \mathbf{W}_{train,2}, \dots, \mathbf{W}_{train,M}] \quad (93)$$

其中 M 為訓練影像的總張數。

第 j 張測試影像之特徵矩陣為

$$\mathbf{W}_{test,j} = [\mathbf{W}_{test,j,1}, \mathbf{W}_{test,j,2}, \dots, \mathbf{W}_{test,j,L}] = \mathbf{P}^T \times \mathbf{A}_{test,j} \quad (94)$$

其中 $\mathbf{A}_{test,j}$ 為測試影像 \mathbf{A}_{test} 中的第 j 張影像。

所有測試影像的特徵矩陣為

$$\mathbf{W}_{test} = [\mathbf{W}_{test,1}, \mathbf{W}_{test,2}, \dots, \mathbf{W}_{test,N}] \quad (95)$$

則任兩向量間之歐幾里德距離定義為

$$d(\mathbf{W}_{train,i}, \mathbf{W}_{test,j}) = \sum_{k=1}^L \|\mathbf{W}_{train,i,k} - \mathbf{W}_{test,j,k}\| \quad i = 1, 2, \dots, M, j = 1, 2, \dots, N \quad (96)$$

比對結果之定義為

$$R_{recog} = \begin{cases} i & \text{if } (d(\mathbf{W}_{train,i}, \mathbf{W}_{test,j}) < T) \\ \text{UnknownFace} & \text{otherwise} \end{cases} \quad (97)$$

若欲求測試資料庫中的第 j 張影像之比對結果，則求得 $d(\mathbf{W}_{train,i}, \mathbf{W}_{test,j})$ 之最小值時所對應之訓練影像中第 i 張影像，此時之歐幾里德距離若小於辨識閾值 T ，辨識結果即為訓練影像中第 i 張影像之人物，反之則視為辨識不到目標。

其他特徵比對法：

(1)曼哈頓距離(City Block)

我們可以定義曼哈頓距離 0 的正式意義為 L_1 - 距離或城市區塊距離，也就是在歐幾里德空間的固定直角坐標上，兩點所形成的線段對軸產生的投影的距離總和。資料點 X_i 和 X_j 可以利用下列公式(98)計算其曼哈頓距離，圖 61 為資料點之間的曼哈頓距離示意圖。

$$dM(X_i, X_j) = \sum_{d=1}^k |X_{i_d} - X_{j_d}| = \|X_i - X_j\| \quad (98)$$

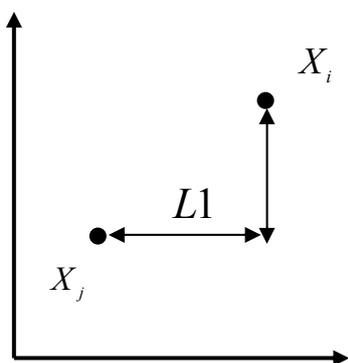


圖61 曼哈頓距離示意圖

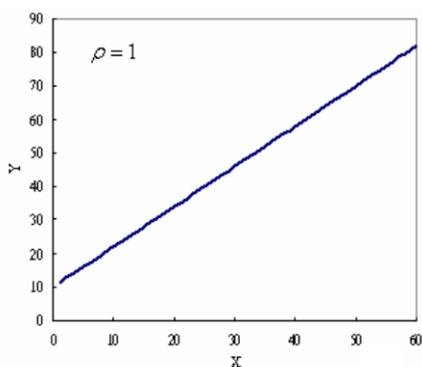
(2) 相關係數(Correlation Coefficient)

相關(Correlation)是用以檢驗兩個變項線性關係的統計技術，以相關係數(Coefficient Of Correlation) ρ 來表示其相關程度。皮爾森相關係數(Pearson's Correlation Coefficient)是一個標準化的關聯係數。其原理是先計算出兩個變項的共變量，再除去兩個變項的標準差，加以標準化，得到的一個標準化分數。相關係數介於-1至1之間。

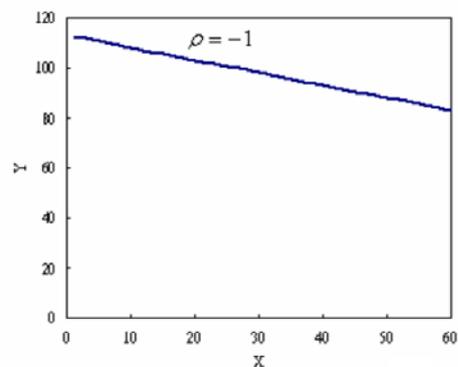
$$\rho_{x,y} = \frac{\text{cov}(X,Y)}{\sigma_x \sigma_y} = \frac{E((X - \mu_x)(Y - \mu_y))}{\sigma_x \sigma_y} \quad (99)$$

其中， $\text{cov}(X,Y)$ 為X與Y的共變異相關， σ_x 為X的標準差， σ_y 為Y的標準差，E為期望值。

當兩個變數的標準差都不為零，相關係數才有定義。當兩個變數的線性關係增強時，相關係數趨於1或-1。當一個變數增加而另一變數也增加時，相關係數大於0。當一個變數的增加而另一變數減少時，相關係數小於0。當兩個變數獨立時，相關係數為0，但反之並不成立。如圖62所示，圖62(a)為完全正相關(所有的點都在一直線上)，圖62(b)為完全負相關，圖62(c)為高度正相關(X與Y的關係接近一直線，Y隨著X的增加而增加)，圖62(d)為沒有線性關係存在(Y不會隨著X改變而改變)。



(a)



(b)

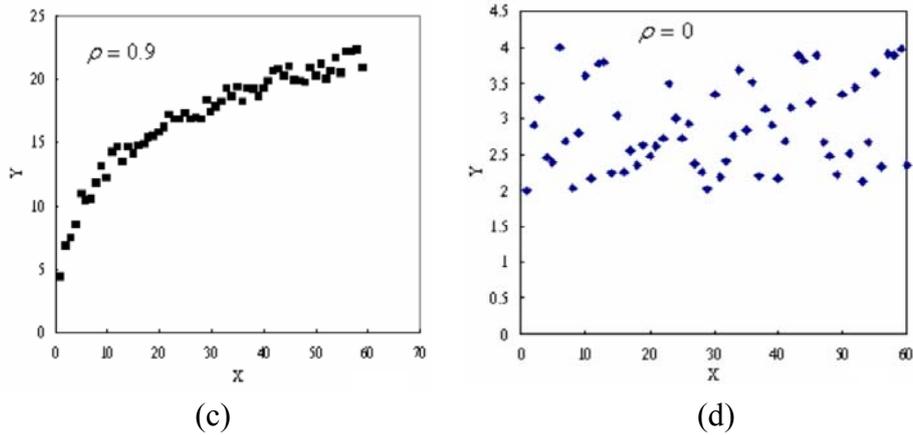


圖 62 不同共變異矩陣數值情況下X與Y的散佈圖

六、實驗過程與分析

(a) 目標追蹤與辨識

本研究的實驗過程區分為兩大部分，第一大部份是做追蹤系統的效能分析，第二大部份則是針對擷取到的人臉做辨識分析。在目標追蹤方面，目標偵測及追蹤步驟找出目標質心位置，本系統判斷是否有目標進入的門檻值，當移動目標二值化影像直方圖中白點面積佔整個畫面的比例大於 0.03 時，則判斷有目標進入。陰影的判斷則是當影像的垂直投影量小於最大垂直投影量值的一半，判斷此資訊為影子並移除之。測試的影像環境包括了室外單人走動、室外二人並排走動、室外多人走動、室外車輛移動、移動人物被遮蔽等狀況，計算各個演算法追蹤時花費的時間及正確率做分析。

本實驗過程中車輛只做追蹤模擬，而人物則做追蹤與辨識，在人臉辨識部分，將經過特徵區域選取後擷取出的人臉影像，與自行拍攝的人臉資料庫做比對。

(i) 車輛追蹤

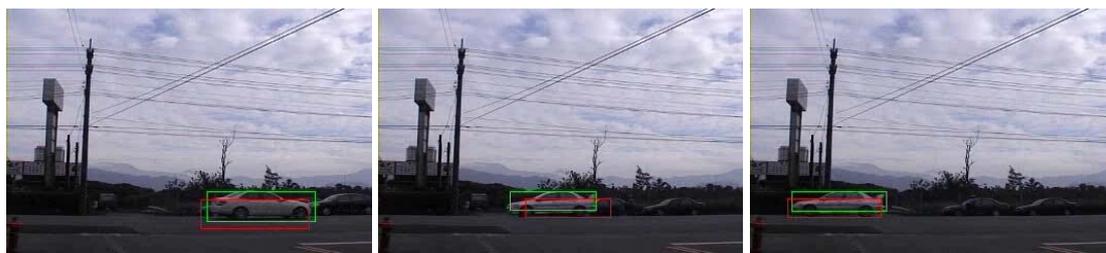
本實驗我們模擬在一室外場景，實施行經車輛的追蹤。車輛行進速度較人物快速，計算時間必須較短暫，我們可以經由提出的影像偵測程序，找出車輛的形狀，進而將外型框出來，達到追蹤的目的，但是畫面中天空、樹葉等環境變化較大，容易產生追蹤錯誤，如圖63。



(a)Frame#42

(b)Frame#45

(c)Frame#53



(d)Frame#63

(e)Frame#69

(f)Frame#75

圖63 使用”卡曼濾波器”追蹤室外車輛移動的實驗結果

(ii) 室外單人走動

本實驗在較複雜的室外環境進行追蹤，背景裡的樹木會隨著風而搖動，也會受到不同光源大小的影響，進而產生人影及較多的雜訊，首先人物由畫面右方進入，因為光源強烈，人體因陽光照射而有影子產生，經過實驗中的去除陰影程序，可以發現人影並未列入偵測範圍，有效的將陰影去除，成功的追蹤到人體頭部部分，如圖64及圖65所示。



(a)Frame#66

(b)Frame#71

(c)Frame#75



(d)Frame#82

(e)Frame#86

(f)Frame#95

圖64 使用”卡曼濾波器”追蹤室外單人走動的實驗結果



(a)Frame#35

(b)Frame#41

(c)Frame#50



(d)Frame#56

(e)Frame#64

(f)Frame#72

圖65 使用”核心函數”追蹤室外單人走動的實驗結果

(iii) 室外多人走動

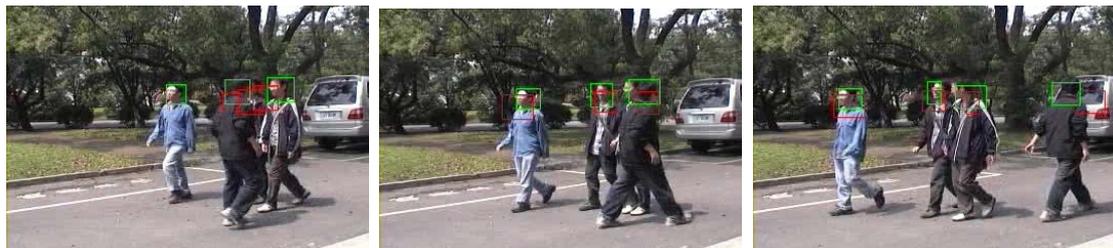
本實驗我們模擬在室外場景，實施多個目標的追蹤如圖66及圖67。在卡曼濾波器追蹤多人交錯的影像序列中，被遮蔽的人會暫時失去追蹤，直到與其他人物分開後才又恢復追蹤，從圖66的Frame#73中發現當多人會合時，會因為人與人之間距離過小，容易被判別為同一物體追蹤，直到彼此離開時在又成功追蹤到每個人；當兩人相遇重疊時，因為樣版色彩資訊相似導致追蹤錯誤，從圖67的Frame#77中發現當兩人會合後開始導致錯誤追蹤而無法回復。



(a)Frame#57

(b)Frame#65

(c)Frame#69



(d)Frame#73

(e)Frame#77

(f)Frame#85

圖66 使用”卡曼濾波器”追蹤室外多人走動的實驗結果



(a)Frame#60

(b)Frame#69

(c)Frame#71



(d)Frame#74

(e)Frame#77

(f)Frame#90

圖 67 使用”核心函數”追蹤室外多人走動的實驗結果

(iv) 移動人物被遮蔽

本實驗我們模擬人物暫時被遮蔽的追蹤，如圖68及圖69。當移動物體部分被其他物體干擾，卡曼濾波器追蹤系統可以容忍部分遮蔽物的阻礙，在干擾不大時仍能執行追蹤功能，而當移動物體被障礙物完全遮蔽時，系統會暫時停止追蹤，待移動物體再次出現時，重新經過影像偵測程序再次獲得移動資訊，成功追蹤到該物體；而核心函數追蹤時會因被遮蔽物遮擋住顏色資訊而導致無法繼續追蹤。



(a)Frame#35

(b)Frame#51

(c)Frame#75



(d)Frame#83

(e)Frame#91

(f)Frame#101

圖 68 使用”卡曼濾波器”追蹤移動人物被遮蔽的實驗結果



(a)Frame#46

(b)Frame#52

(c)Frame#66



(d)Frame#78

(e)Frame#87

(f)Frame#98

圖 69 使用”核心函數”追蹤移動人物被遮蔽的實驗結果

(v) 追蹤效能分析

卡曼濾波器追蹤方法利用估計目標的運動參數，準確的預測出移動目標在下一張畫面出現的位置，縮小搜尋視窗的大小，而後不斷更新目標和參數，一直迭代計算；而核心函數搭配Mean Shift演算法也是一個迭代的過程，先算出當前點的偏移均值，再移動該點到其偏移均值，然後以此點為新的起點，繼續移動，直到滿足一定的迭代條件為止。本實驗模擬了卡曼濾波器及核心函數演算法這兩種追蹤演算法的追蹤結果。當連續影像序列中有物體移動時，可使系統偵測移動物體，以更新物體的資訊，我們實際執行上述五種場景的影像序列，統計移動物體追蹤的執行時間與正確次數，結果如表一所示：

表一 追蹤效能及正確率

實驗演算法(場景編號)	卡曼濾波器					核心函數				
	1	2	3	4	5	1	2	3	4	5
實驗影像序列	1	2	3	4	5	1	2	3	4	5
影片畫面總張數	250	140	130	95	155	250	140	130	108	155
物體出現的畫面張數	120	90	100	72	117	120	90	100	92	117
正確追蹤的畫面張數	115	75	70	51	100	102	72	23	40	53
追蹤正確率	88%	80%	70%	70%	85%	85%	80%	23%	43%	45%
單張畫面平均處理時間(ms)	0.06	0.07	0.08	0.08	0.06	0.09	0.10	0.11	0.10	0.09
平均每秒處理畫面數(fps)	15.1	14.2	13.2	11.3	15.5	10.5	9.2	8.5	9.1	10.1

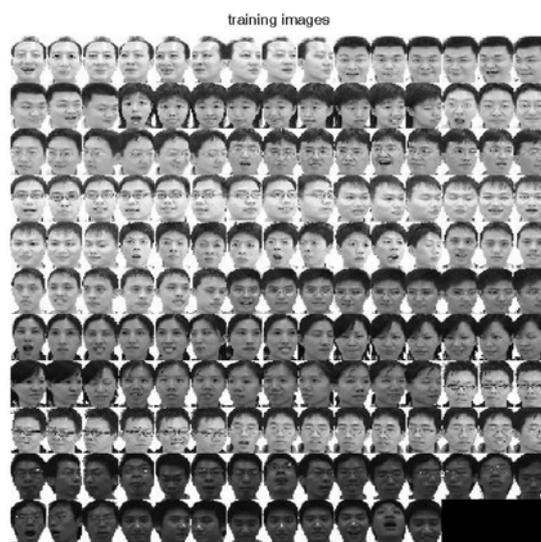
由以上的實驗結果可以發現，在室外的場景，容易受到許多雜訊的影響，如光源的變化、被風吹動的樹葉、背景的改變、在遠處的人、車經過等，而使得移動物體與背景分割不完整或是陰影去除不夠好，影響到移動物體的形狀判斷，導致追蹤目標質心選擇錯誤而追蹤失敗。車輛剛進入畫面時，如圖63的Frame#53，因為車身顏色與背景過於相似，以致誤認為背景而無法分割出來，且兩車交會時會當成同一物體，待分開後又分割成兩輛車，而攝影機受到風吹或雜訊的影響，而發生追蹤錯誤情況。由於單人目標較單純易與背景做分割，因此追蹤效果較為準確。

本報告所提出的追蹤演算法，能有效追蹤多個物體，唯在多人交錯的場景中，在卡曼濾波器演算法中，被遮蔽的人會暫時失去追蹤，直到與其他人分開後又恢復，從圖66的Frame#77開始，可以發現左側移動的人因為與其他人交錯，因此被當作成一個物體追蹤，待分離後即可再次追蹤到個人；而核心函數搭配Mean Shift演算法的目標模型建立在核心直方圖的基礎上，目標周圍出現與目標顏色相近的其他物體，反而會被相似顏色目標吸引而導致追蹤失敗。

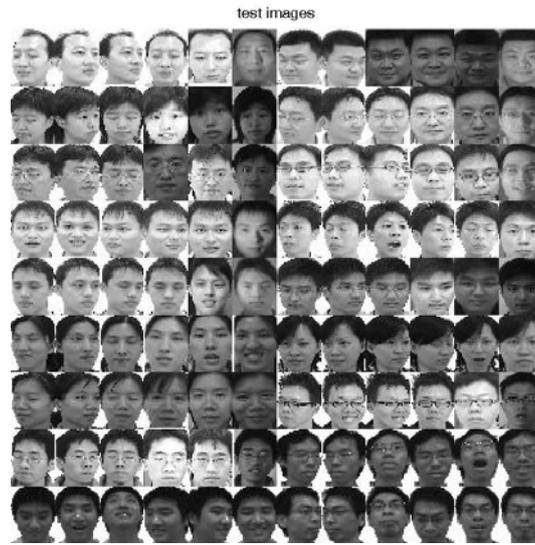
當移動物體被遮蔽時，卡曼濾波器演算法採用軌跡預測和移動趨勢判斷出下一個可能出現的位置，因此可以保持對目標的追蹤；核心函數演算法採用了固定的核心搜尋視窗範圍，可以容忍部分遮蔽或是短時間的阻礙，但是如果目標長時間被遮蔽，會發生目標遺失不能自動更新及定位的缺點。

(vi) 人臉辨識效能分析

本實驗利用數位攝影機拍攝連續影像，經過偵測程序後擷取出人臉作為樣版影像，由於目前本實驗室所配備的攝影機設備等級不佳，擷取的人臉畫面較為模糊，無法有效的辨識，因此我們實驗程序分為兩個獨立步驟，首先由攝影機拍攝可見光人臉影像資料庫做辨識及比對模擬之用，當人物接近時再擷取人物之臉部當做測試影像，預先拍攝所建立的資料庫內有18人，每人15張臉部影像，合計270張影像，影像維度為39x30像素，彩色(RGB)、JPG檔形式。取其9張作為訓練影像，其餘6張作為測試影像，共計162張訓練影像、108張測試影像，圖70為實驗使用之人臉資料庫，圖71為特徵值分佈情形，表二及表三為使用二維主成分分析法(2DPCA)、二維線性識別分析法(2DLDA)及獨立成分分析法(ICA)三種特徵擷取方法之結果比較。

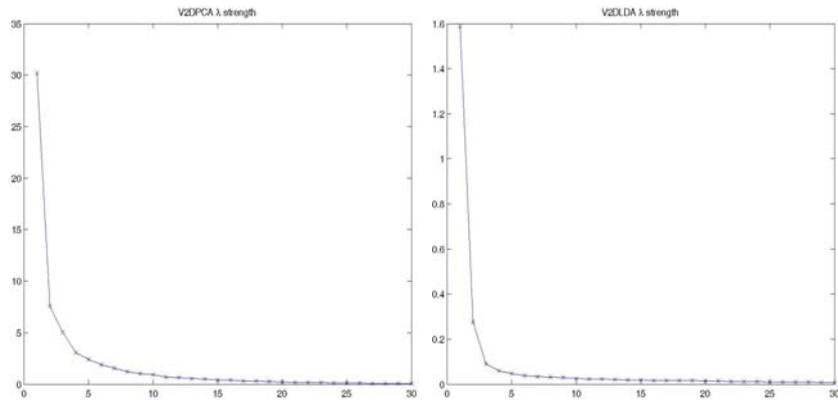


(a)



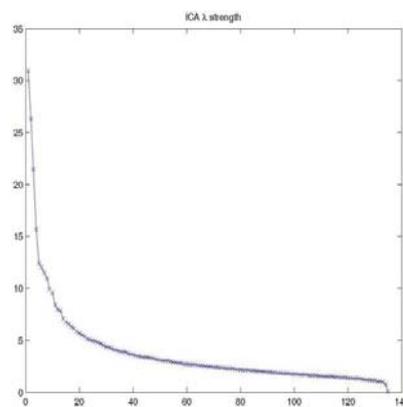
(b)

圖70 可見光資料庫(a)訓練影像(b)測試影像



(a) 2DPCA

(b) 2DLDA



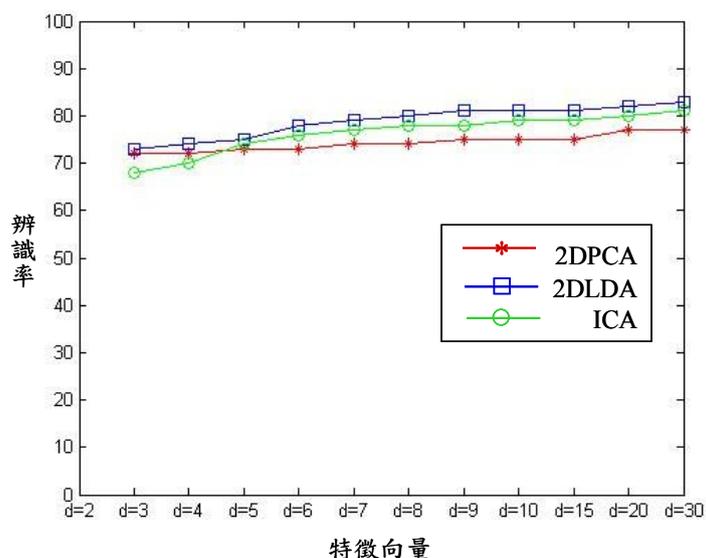
(c) ICA

圖71 特徵值分佈圖

表二 人臉影像執行特徵擷取比對時間及辨識率

特徵擷取比對方法		訓練影像162張、測試影像108張、影像大小39*30											
特徵值		2	3	4	5	6	7	8	9	10	15	20	30
2DPCA	平均訓練時間(秒)	0.34	0.36	0.33	0.36	0.36	0.36	0.38	0.39	0.36	0.42	0.56	0.61
	平均測試時間(秒)	0.28	0.28	0.33	0.33	0.33	0.34	0.36	0.36	0.38	0.42	0.47	0.72
	平均執行時間(秒)	0.63	0.64	0.66	0.69	0.69	0.70	0.73	0.75	0.73	0.84	1.03	1.33
	辨識率(%)	72.41	72.47	73.17	73.71	74.11	74.57	75.16	75.86	75.86	77.01	77.01	77.01
2DLDA	平均訓練時間(秒)	0.16	0.17	0.17	0.17	0.19	0.19	0.20	0.19	0.22	0.23	0.25	0.30
	平均測試時間(秒)	0.28	0.28	0.31	0.33	0.34	0.34	0.34	0.36	0.36	0.41	0.47	0.66
	平均執行時間(秒)	0.44	0.45	0.48	0.50	0.53	0.53	0.54	0.55	0.58	0.64	0.72	0.95
	辨識率(%)	73.56	74.86	75.41	78.16	79.31	80.46	81.61	81.61	81.61	83.91	83.91	83.91
ICA	平均訓練時間(秒)	1.68	1.68	1.70	1.68	1.73	1.68	1.68	1.68	1.71	1.73	1.73	1.73
	平均測試時間(秒)	0.98	0.99	0.96	0.99	1.00	0.99	0.98	0.98	0.98	0.98	0.98	0.99
	平均執行時間(秒)	2.66	2.67	2.66	2.67	2.73	2.67	2.66	2.67	2.69	2.71	2.71	2.72
	辨識率(%)	68.36	70.67	73.71	76.16	77.31	78.33	79.46	79.40	79.60	80.60	81.05	81.75

表三 樣版影像執行三種特徵擷取比對結果圖



由圖71特徵值分佈情形來決定取用特徵向量的數目，當計算特徵值為15之後曲線幾乎沒有變化，表示可達最高辨識率。再由表二、三可得知，執行2DPCA演算法時，當特徵值取15時，辨識率73%，平均每張影像訓練影像時間為0.42秒，平均每張影像測試影像時間為0.42秒，合計每張影像平均執行時間為0.84秒。執行2DLDA演算法時，當特徵值取15時，辨識率便可達到83%，平均每張影像訓練影像時間為0.23秒，平均每張影像測試影像時間為0.41秒，合計

每張影像平均執行時間為0.64秒。執行ICA演算法時，當特徵值取20時，辨識率可達到81%，平均每張影像訓練影像時間為1.73秒，平均每張影像測試影像時間為0.98秒，合計每張影像平均執行時間為2.71秒。三種特徵擷取方法之比較，可以得知在二維演算法中，2DLDA的辨識率可達83%為最高，每張影像平均執行辨識時間0.64秒為最快。

(b) 門禁警監系統之應用：

(i) 人物追蹤

(1) 更新背景

假設只使用一開始之影像視為背景，當影像光影沒有變化時，可藉由背景與當前影像取得前景影像，但隨時間而變，周圍環境之狀態將不會保持固定，後面處理中可能會產生非常多不必要的資訊，如前景擷取錯誤，而降低處理之效率及準確性，因此本系統使用自動更新背景以改善上述狀況。

當程式開始擷取影像時，即將每張影像存入暫存器 I_{temp} 中，暫存器之初始影像為階值 0 之黑色影像替代，暫存器可容納 30 個畫格影像，當存入超過 30 張影像時，則將暫存器中的最早進入的一張影像移除，再存入新影像，一開始的背景影像 B_0 乃是利用前面 30 張影像平均所得。

圖 72 由於是影像的前 30 張平均所得，圖 72(a)為第 10 個 Frame，其暫存器中影像只有儲存 10 張，其它 20 張為階值 0 之黑色影像，因此背景影像偏暗，直至 30 張影像皆讀入 I_{temp} 時，才為一張完整背景，如圖 72(c)所示。

$$B_0 = \sum_{i=1}^{30} I_{temp,t} \quad (100)$$

其中 t 為當前時間， $I_{temp,t}$ 為 t 時之暫存器，見式(100)，不同模擬結果如圖 72 所示。



(a)



圖72 初始背景(a)前10張平均影像(b)前20張平均影像(c)前30張平均影像

但是當有人進入時，這些平均影像就無法視為完整的背景，此時會有鬼影出現，如圖 73。

圖 73 (a)為 Frame 632 以每 3 個 Frame 取一次影像，取至 Frame 722 後，再將此 30 張影像平均所得。圖 73 (b)為 Frame 656 以每 3 個 Frame 取一次影像，取至 Frame 746 後，再將此 30 張影像平均所得。接著加入背景更新方法，來改善鬼影的問題，所使用到的有背景影像、當時影像及平均影像。

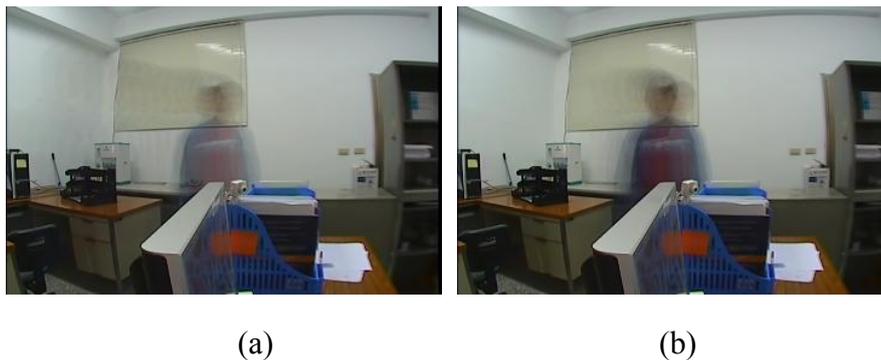


圖73 平均影像(a) Frame 632-722 (b) Frame 656-746

當存入暫存器 I_{temp} 的影像超過 30 張後，就以每間隔 3~8 張才存入暫存器中，此做法是用以控制背景更新速率，存入暫存器之間隔越長，則平均影像 I_{mean} 必須花費越久時間才會接近當前影像 I_t ，本文中透過不同的間隔時間以控制背景的更新速率。

若當前影像 $I_t(x, y)$ 與前一張背景影像 $B_{t-1}(x, y)$ 相減，取其絕對值之後，其值大於閾值 T_1 且背景影像與平均影像差異小於閾值 T_2 時，可將背景影像 $B_{t-1}(x, y)$ 用當前影像 $I_t(x, y)$ 替換掉；若為其它情況，則保持原先之背景影像 $B_{t-1}(x, y)$ ，定義更新後背景影像為

$$B_t(x, y) = \begin{cases} I_t(x, y) & \text{if } |I_t(x, y) - B_{t-1}(x, y)| > T_1 \\ & \& |I_t(x, y) - I_{mean}(x, y)| < T_2 \\ B_{t-1}(x, y) & \text{otherwise} \end{cases} \quad (101)$$

$$I_{mean} = \sum_{i=t-30}^t I_{temp,i} \quad (102)$$

圖 74 (a)(c)為 Frame722 及 Frame722 之原始影像，圖 74 (b)(d)分別為其背景影像，可由圖 74(c)看出，當人位於影像中間，並接近攝影機時，背景影像圖 74(d)仍保持不變，在短的時間中，背景影像不會更新，以保持明確的前景。

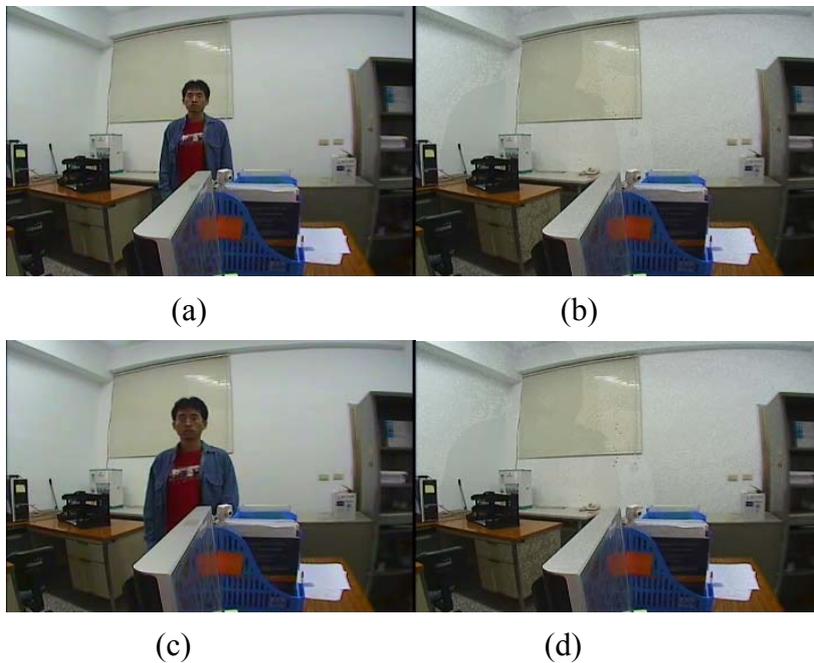


圖74 背景更新(a) Frame 722 (b)背景影像(c) Frame 746 (d)背景影像

圖 75(a)為 Frame 200 之當時畫面，圖 75((b)為 Frame 200 時之背景影像，圖 75 (a)(c)(e)每張影像之間隔為 120 個 Frame，分別為 Frame 200、Frame 320、Frame 440 之當時影像，圖 75(b)(d)(f)每張影像之間隔為 120 個 Frame，則分別為 Frame 200、Frame 320、Frame 440 之背景影像。由圖 75 (a)右方可看出此時有一人將進入畫面，由於此人物進入時期不長，故背景影像此時並無更新，如圖 75 (b)，接著圖 75 (c)看出此人接近攝影機，並持續站著不動至圖 75 (e)，從一開始的人物進入畫面至圖 75(e)，此段時期已過了 240 個 Frame，從圖 75 (d)(f)看出，背景由沒有人影至漸漸顯示出人的外型，Frame 440 之動作則為背景的更新，用於光影變化時，如開關燈時的處理。

圖 76 則是接續圖 75 之影像，圖 75 (a)(c)為 Frame 560 及 Frame 680 時之當前影像，圖 76(b)(d)為 Frame 560 及 Frame 680 時之背景影像。當圖 75 (e)的人從畫面右方走至圖 76 (a)畫面的左方時，再經過一段時間後，原先圖 75 (f)的背景影像中有人影，但卻因為人物的移動，而使該位置漸漸還原成原先的背景影影，在本系統中，不動的物體完全轉換成背景大約要 240 個 Frame 也就是 6 秒鐘左右。

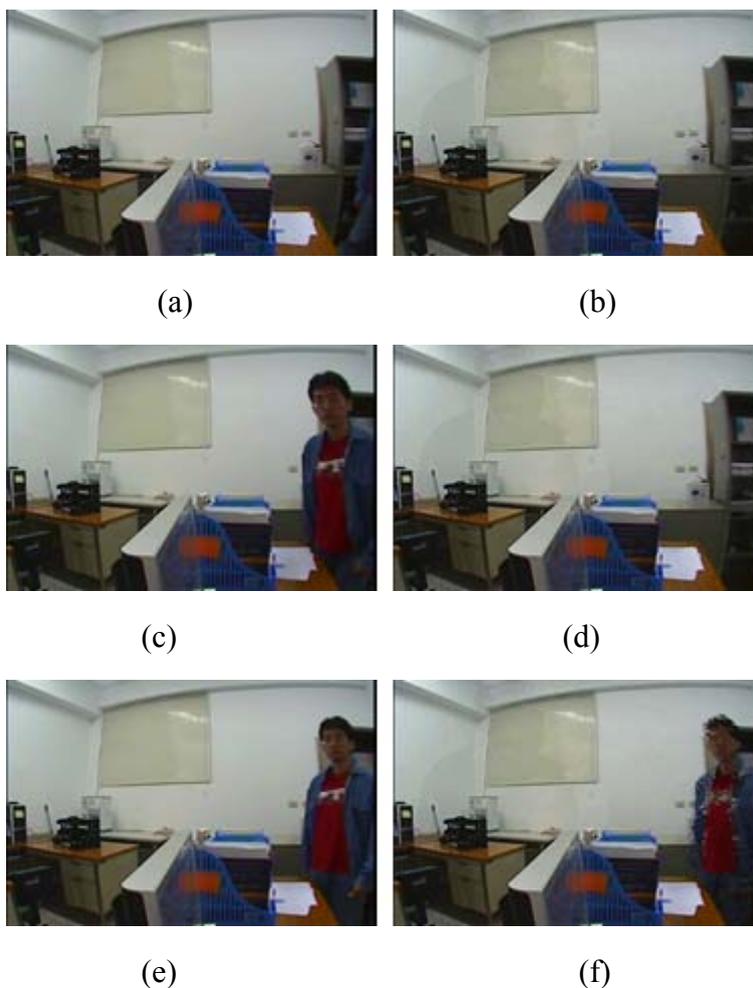


圖75 背景更新(a) Frame 200 (b) Frame 200之背景(c) Frame 320 (d) Frame 320之背景(e) Frame 440 (f) Frame 440之背景





圖76 背景更新(a) Frame 560 (b) Frame 560之背景(c) Frame 680 (d) Frame 680之 背景

(2) 分割前景

為了取的畫面中移動之物體，因此做分割前景之處理。一般最常用的分割前景方法，只將前後兩張影像相減，接著設一閾值，若其值超過此閾值則視為前景影像中的像素，表為

$$F_t(x, y) = \begin{cases} 255 & \text{if } |I_{t-1}(x, y) - I_t(x, y)| > T \\ 0 & \text{otherwise} \end{cases} \quad (103)$$

其中 F_t 為前景影像， I_t 為當前影像， T 為閾值。

在本系統中，利用 5.2.1 節之方法已得到背景影像，因此本系統之前景可表為

$$F_t(x, y) = \begin{cases} I_t(x, y) & \text{if } |I_t(x, y) - B_t(x, y)| > T \\ 0 & \text{otherwise} \end{cases} \quad (104)$$

圖 77 為 Frame 722 之前景分割過程，圖 77(b)為 Frame 722 之背景影像，圖 77 (c)為分割之前景影像。圖 78 為 Frame 746 之前景分割結果，圖 78 (b)為 Frame 746 之背景影像，圖 78 (c)為分割之前景影像。





(c)

圖77 分割前景於Frame 722 (a)原始影像(b)背景影像(c)前景影像



(a)

(b)



(c)

圖78 分割前景於Frame 746 (a)原始影像(b)背景影像(c)前景影像

(3) 膚色偵測

透過動態影像分割前景，接著將前景影像進行膚色偵測，可去除許多非人臉之多餘資訊，如桌面與人臉膚色甚是相近。在此步驟中，利用 YCbCr 轉換，將前景的 RGB 色彩空間轉換至 YCbCr 空間，便於膚色的擷取，定義轉換至 YCbCr 空間之膚色範圍為

$$P_t(x, y) = \begin{cases} 255 & \text{if } [77 \leq C_b \leq 127] \text{ and } [133 \leq C_r \leq 173] \\ 0 & \text{otherwise} \end{cases} \quad (105)$$

當 $77 \leq C_b \leq 127$ 且 $133 \leq C_r \leq 173$ 時，此範圍設為膚色部分， P_t 為膚色二值影像。圖 79 及

圖 80，分別在 Frame 722 及 Frame 746 所取得之前景與其膚色二值影像，可看出在不同時間內，可取得膚色二值影像，由於受測者所穿著之衣服接近膚色，故膚色二值影像包含衣服，但並不包含接近膚色之桌面。



圖79 膚色二值影像於Frame 722 (a)前景影像(b)膚色二值影像

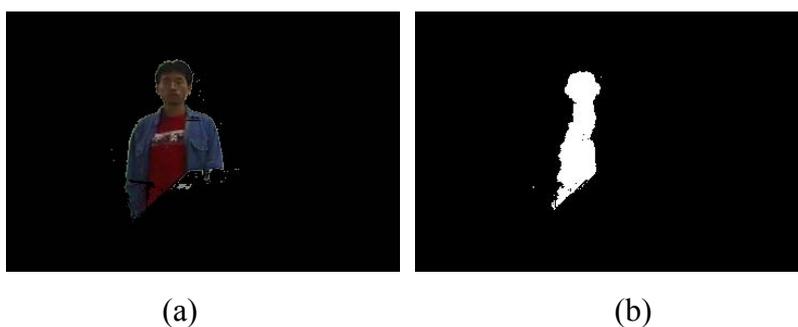
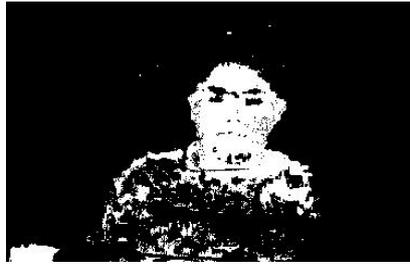


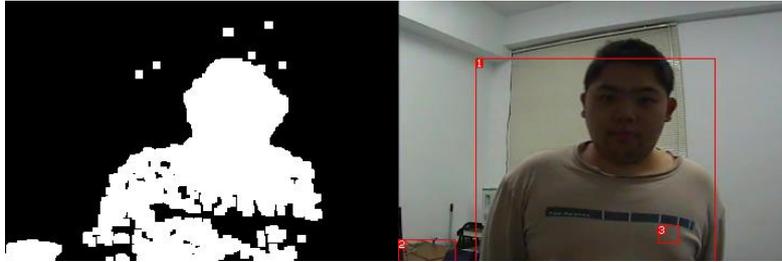
圖80 膚色二值影像於Frame 746 (a)前景影像(b)膚色二值影像

(4) 臉部擷取

進行臉部擷取前，為了消除臉部二值影像可能出現的雜訊，以及使臉部區塊完整，因此做了一次蝕刻和兩次膨脹的動作，接著利用連通法，取得相連的區塊，即可由區塊大小及形狀，取得臉部正確位置。但是若衣服接近膚色，在連通法完成後，則會切割錯誤，如圖 81 所示。為了減少錯誤發生，將針對臉部的可能區塊再次進行水平及垂直投影，再設定一閾值，由投影量來判斷頭部位置，正確的分割結果如圖 82 所示。



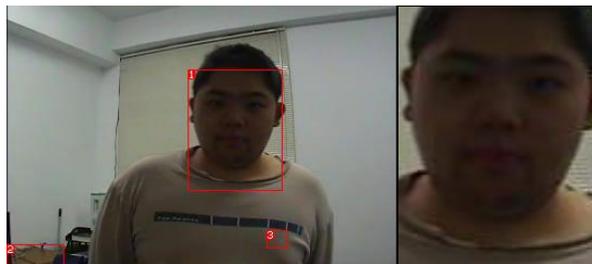
(a)



(b)

(c)

圖81 臉部錯誤擷取(a)膚色二值影像(b)形態學處理(c)框選臉部範圍



(a)

(b)

圖82 臉部擷取修正(a)框選臉部範圍(b)臉部擷取

(5) 光照正規化

做完影像前處理後，已取得臉部影像，爲了降低光影的影響，將所有擷取之臉部影像，使用本報告所介紹的正規化方法，如 HE、wBHE、GIC、LN、SQI 及 RR 來從事光照正規化。

圖 83 及圖 84 則是圖 82(b)正規方法進行處理之結果。



(a)

(b)

(c)

(d)

圖83 光照正規化處理(a)彩色影像(b)灰階影像(c)HE影像(d)wBHE影像

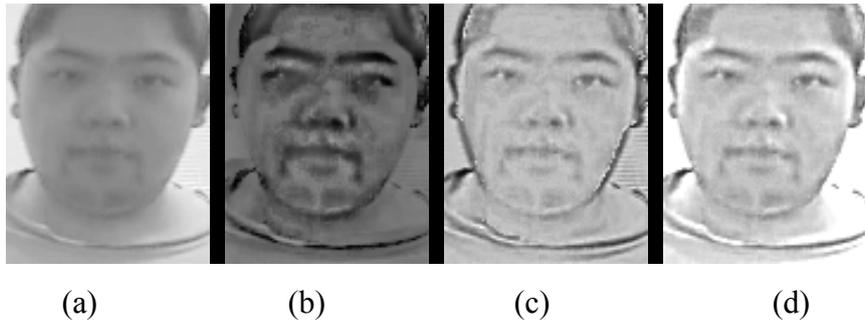


圖84 光照正規化處理(a)GIC影像(b)LN影像(c)SQI影像(d)RR影像

(ii) 人物辨識

(1) 訓練影像資料庫

此資料庫於門禁警監系統執行前，已將訓練影像透過各種光照正規處理，並分別將各方法處理之影像進行各種辨識演算法的訓練，如 PCA、LDA、2DPCA、2DLDA 等，其目的在於取得各訓練資料之特徵向量，故訓練影像資料庫所包含之訓練結果有各種光照處理之訓練影像、各種辨識演算法所訓練之特徵向量、訓練影像透過特徵向量投影至特徵空間之投影向量。以上三種訓練結果於門禁警監系統執行前可先行計算得出，已達到省時之目的。本系統主要是以 2DPCA 做為特徵向量取得之演算法。

(2) 特徵空間

藉由前面章節所得之光照正規臉部影像，可利用資料庫中訓練影像之特徵向量，將此光照正規臉部影像投影至特徵空間，此處產出則為光照正規臉部影像之投影向量。

(3) 特徵比對

利用資料庫中訓練影像之投影向量，以及光照正規臉部影像之投影向量，可將此向量進行歐幾里德距離之特徵比對方法，藉此判斷是否為資料庫人物。在此所設之閾值 T ，經由實驗測試不同影像所得定此閾值為 8.55，此閾值可區別出受測者是否為本實驗室之資料庫人物，此部分之產出為辨識結果。

圖 85 為門禁警監系統之處理結果，影像中標記(a)處為當前影像，(b)為其背景影像，(c)為前景影像，(d)為前景之膚色擷取影像，(e)為所擷取之臉部畫面，(f)為其辨識之結果。



圖85 門禁警監系統辨識處理結果

(iii) 實驗結果分析

實驗的資料庫為 Yale Face Database B，首先介紹資料庫，這資料庫中共有 10 個人如圖 86 所示，9 個不同方向的攝影機如圖 87 所示，以及 66 種光照如圖 88 所示，其中 1 個為正常光照，另一為無光照，在此實驗中以其餘 64 種光照影像，以影像奇、偶順序將資料庫分為各 32 張之訓練影像以及測試影像進行辨識，共 5760 張影像，訓練影像 2880 張，測試影像 2880 張。以下為光照處理後之辨識結果：

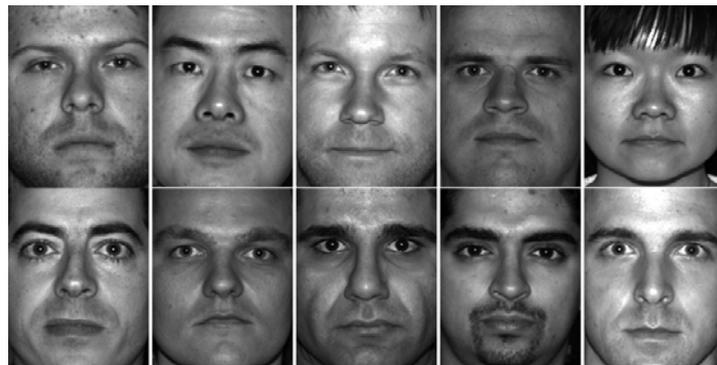


圖86 10個受測者

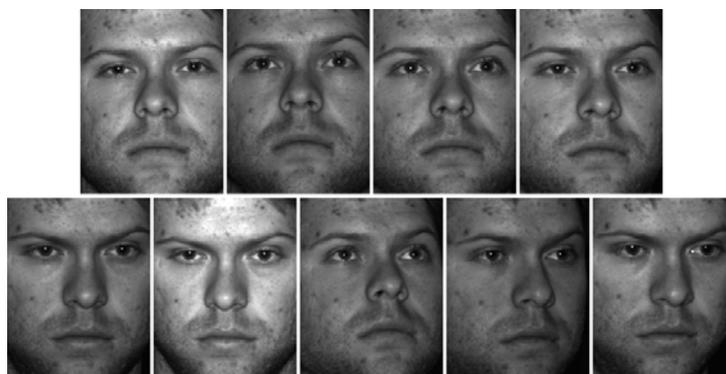


圖87 9個不同方向的攝影方向



圖88 第1人之正向拍攝的66個光照影像

接著為各光照正規化方法之處理影像，除了 DCT 執行時間過長，不適合用於即時處理，不在下面顯示外，其餘結果見圖 89 至圖 94。



圖89 第1人之正向拍攝的HE處理影像



圖90 第1人之正向拍攝的GIC處理影像

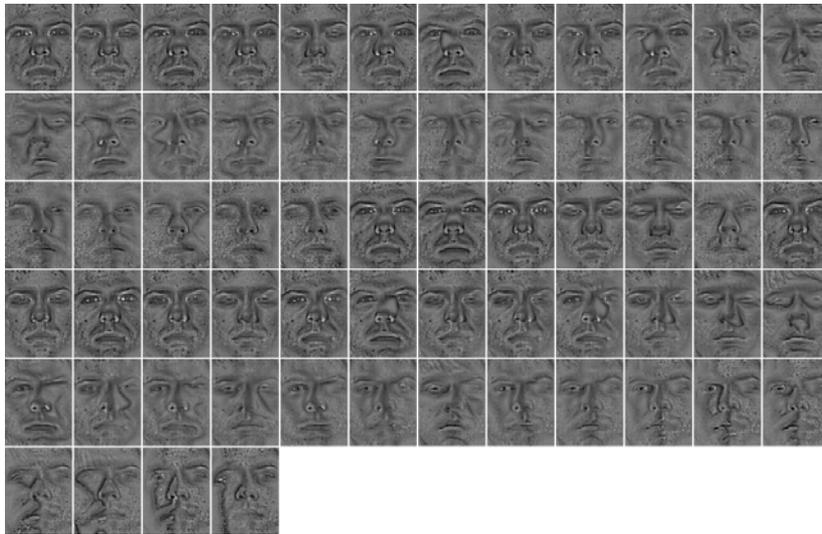


圖91 第1人之正向拍攝的LN處理影像



圖92 第1人之正向拍攝的wBHE處理影像



圖93 第1人之正向拍攝的SQI處理影像



圖 94 第 1 人之正向拍攝的 ReflectionRatio 處理影像

表四為每單張影像，執行各種光照正規之處理時間，其中較快的處理有 HE、GIC、wHE 及 RR，而最慢的處理為 SQI，表五則是說明辨識率及每單張之處理時間，其辨識率來說，wBHE 辨識率達到 95%，而 SQI 及 RR 不管使用 PCA 還是 2DPCA 皆達到 99%，有不錯的光照正規效果。

表四 各光照正規化之處理時間

維度 (60x78)	HE 處理	GIC 處理	LN 處理	SQI 處理	wBHE 處理	RR 處理
處理時間 (秒/單張)	0.0113	0.0469	0.2028	0.4374	0.0143	0.0214

表五 各光照正規化之辨識時間及辨識率

演算法 資料庫	PCA (50x65)			2DPCA (50x65)		
	訓練時間 (sec)	辨識時間 (sec)	辨識率	訓練時間 (sec)	辨識時間 (sec)	辨識率
原始 資料庫	0.544	0.055	84.37%	0.159	0.210	89.16%
HE 處理 資料庫	0.589	0.026	92.99%	0.104	0.177	95.10%
SQI 處理 資料庫	0.584	0.029	99.38%	0.143	0.229	99.38%

GIC 處理 資料庫	0.610	0.051	90.41%	0.111	0.181	94.44%
LN 處理 資料庫	0.554	0.037	88.19%	0.104	0.253	93.71%
wBHE 處理 資料庫	0.561	0.024	95.35%	0.113	0.182	97.81%
RR 處理 資料庫	0.499	0.022	99.38%	0.095	0.159	99.65%

註：以上時間為每張之計算時間

在所設計之門禁警監系統中，將選擇其中一種方法，進行光照正規。雖然 RR 及 SQI 有最高的辨識率，但 RR 需要有參考影像，若參考影像與光照影像中的人物有所偏移，則會影響其光照效果，而 SQI 礙於處理時間較長，不適用於即時的門禁警監系統中，因此選擇辨識率排名第三的 wBHE，做為門禁警監系統所使用之辨識方法進行光照處理。

(d) 門禁警監系統之實驗設計：

在本節實驗中將以受測者正上方有兩盞日光燈、受測者正上方有一盞日光燈、受測者正上方日光燈全關、受測者受側面光照影響、受測者未戴眼鏡及受測者穿戴接近膚色之衣服，六種測試狀況，以驗證本系統。

(i) 實驗一(受測者正上方有兩盞日光燈)

圖 95 至圖 97 為門禁警監系統之處理結果，圖 95(a)為 Frame 299 之處理結果，當有目標進入畫面時，系統會判斷臉部大小，當寬大於 30pixel、高大於 39pixel 時，會將臉部進行光照處理再與資料庫影像進行辨識，由於資料庫影像以正面臉為主，因此可由圖 97 看出此時人物臉部已有所偏移，因此辨識不到目標臉部影像。圖 98 至圖 100 為另一人之處理結果，於圖 98 中可看出有辨識出受測者，圖 99(a)雖然比對結果最接近受測者，但其比對值超過我們所定之閾值，此是為沒辨識到目標。圖 100(a)受測者之臉部已偏離畫面，因此也沒辨識到目標。



(a)



(b)

圖95 兩盞光源之處理結果(a) Frame 299 (b) Frame 309



(a)



(b)

圖96 兩盞光源之處理結果(a) Frame 319 (b) Frame 329

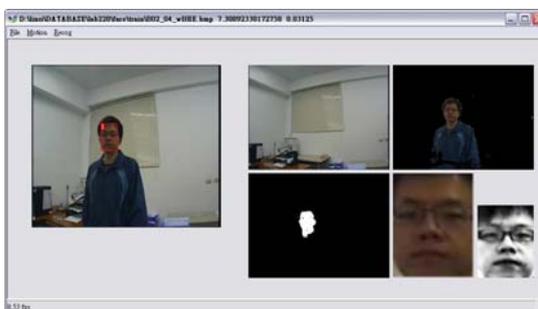


(a)



(b)

圖97 兩盞光源之處理結果(a) Frame 339 (b) Frame 349



(a)



(b)

圖98 兩盞光源之處理結果(a) Frame 170 (b) Frame 180



(a)

(b)

圖99 兩盞光源之處理結果(a) Frame 190 (b) Frame 200



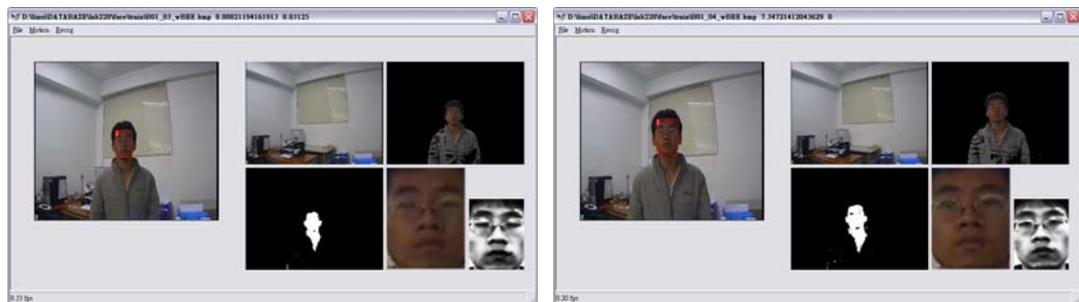
(a)

(b)

圖100 兩盞光源之處理結果(a) Frame 260 (b) Frame 270

(ii) 實驗二(受測者正上方有一盞日光燈)

圖 101(a)(b)(c)(d)(e)(f)分別為 Frame 375、Frame 385、Frame 395、Frame 405、Frame 415 及 Frame 425 之門禁警監辨識結果。其中圖 101 (a)(b)(c)(d)(e)皆有辨識到目標，只有圖 101 (f)之受測者臉部為側臉，因此並沒有辨識到目標。



(a)

(b)



(c)

(d)



(e)

(f)

圖101 一盞光源之處理結果(a) Frame 375 (b) Frame 385 (c) Frame 395 (d) Frame 405 (e) Frame 415 (f) Frame 425

(iii) 實驗三(受測者正上方日光燈全開)

在此實驗中，可看出影像亮度與先前幾個實驗相比較為光亮，這是因為沒有上方光源的影響，使攝影機自動進行反補償的動作，若此臉部影像未做光照正規之處理，則不易辨識到目標，圖 102(a)(b)(c)(d)(e)(f)分別為 Frame 345、Frame 355、Frame 365、Frame 375、Frame 385 及 Frame 395，雖然圖 102 (a)(e)(f) 由於影像有所偏移，因而未辨識到目標，但圖 102 (b)(c)(d) 之臉部影像較為清晰且正面，故即使影像有光照之影響，可用光照正規處理，以取得較好之辨識結果。



(a)

(b)



(c)

(d)



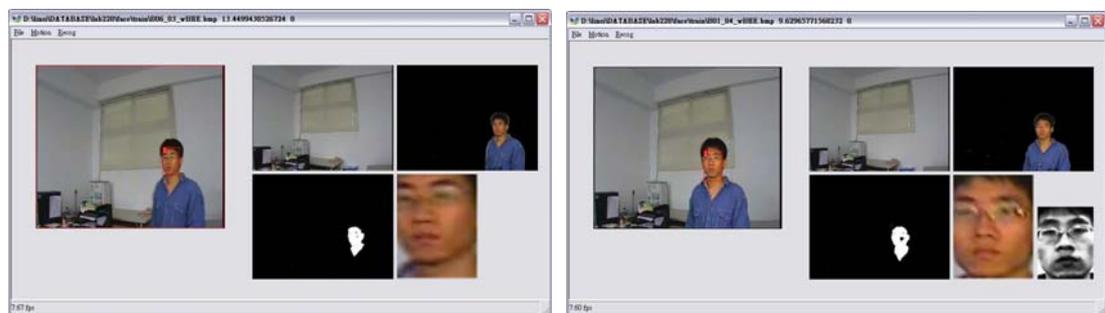
(e)

(f)

圖102 上方光源全關之處理結果(a) Frame 345 (b) Frame 355 (c) Frame 365 (d) Frame 375 (e) Frame 385 (f) Frame395

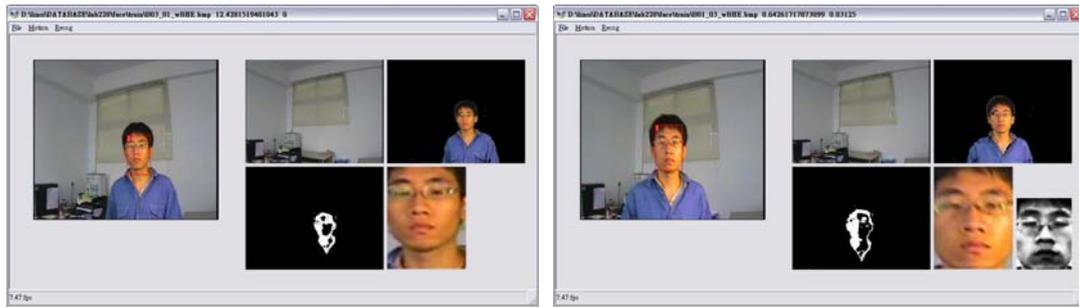
(iv) 實驗四(受測者受側面光照影響)

在此實驗中，使用了側光源照射受測者，圖 103 及圖 104 為門禁警監系統處理側面光照之結果。圖 103 (a)及圖 104 (c)(d)之受測者臉部為側臉，故無法辨識出目標。於圖 103 (b)(d)及圖 104 (a)(b)中，雖然臉部遭受側光源之影響，使膚色有所偏離，造成膚色擷取之不完整，仍可取得正確的臉部位置進行辨識，並透過光照正規處理後識別出正確的受測者身份。於圖 103(c)中可發現，此時所擷取之臉部位置與資料庫影像有所差異，而造成無法辨識出受測者。



(a)

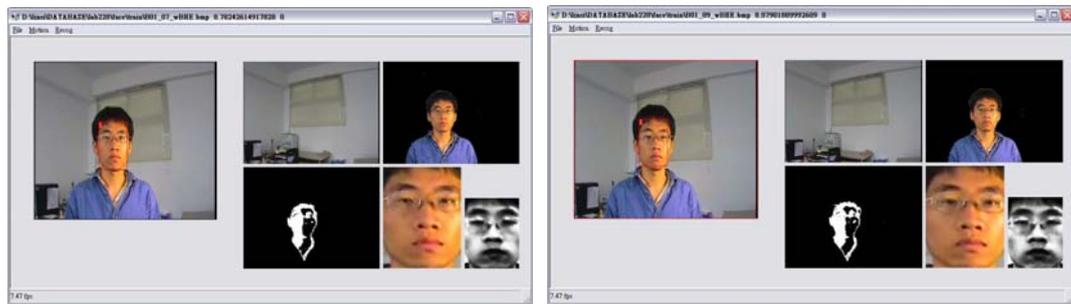
(b)



(c)

(d)

圖103 受測者受側面光照影響之處理結果(a) Frame 100 (b) Frame 110 (c) Frame 120 (d) Frame 130



(a)

(b)



(c)

(d)

圖104 受測者受側面光照影響之處理結果(a) Frame 140 (b) Frame 150 (c) Frame 160 (d) Frame 170

(v) 實驗五(受測者未戴眼鏡)

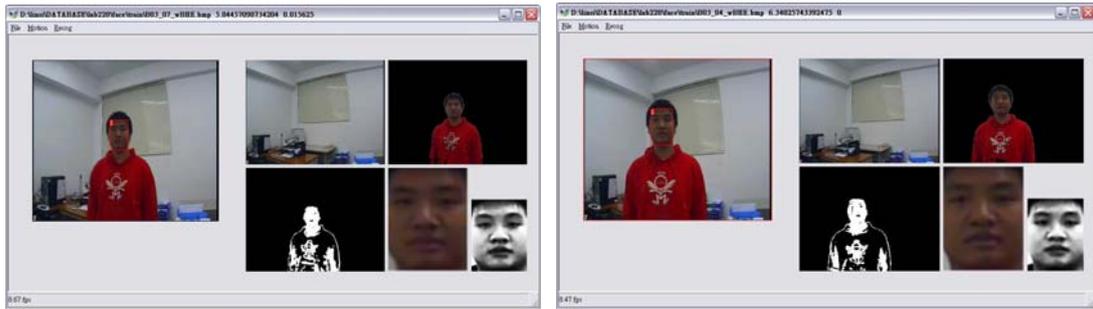
在此實驗中，圖 105(a)(b)(c)(d)(e)分別為 Frame 310、Frame 320、Frame 330、Frame 340 及 Frame 350 之門禁警監辨識結果，除了圖 105 (e)之臉部擷取為側面因而辨識不出目標外，其於圖 105 (a)(b)(c)(d)皆可辨識出受測者。



圖105 受測者未戴眼鏡之處理結果(a) Frame 310 (b) Frame 320 (c) Frame 330 (d) Frame 340 (e) Frame 350

(vi) 實驗六(受測者穿戴接近膚色之衣服)

在此實驗中，受測者穿戴接近膚色之衣服，易使臉部截取錯誤，因而造成辨識不到目標的情況發生，本文之臉部擷取，即針對接近膚色衣服加以修正臉部之擷取，圖 107(a)(b)(c)(d)、圖 106(a)(b)分別為 Frame 197、Frame 217、Frame 237、Frame 257、Frame 277 及 Frame 297 之門禁警監辨識結果，圖 107 (a)(b)(c)(d)及圖 107 (a)中，即使有衣服之影響，仍可正確擷取臉部影像，並且正確的辨識出目標身分，圖 107 (b)之臉部為側臉影像，故辨識不到目標身分。



(a)

(b)



(c)



(d)

圖106 衣服接近膚色之處理結果(a) f-197 (b) f- 217 (c) f- 237 (d) f- 257



(a)

(b)

圖107 衣服接近膚色之處理結果(a) Frame 277 (b) Frame 297

(vii) 門禁警監系統之結果分析：

表六為門禁警監系統透過 wBHE 進行光照正規後所辨識之結果，可由此表看出，測試影像即使受到光照影響，也可順利辨識出受測者身分，除了側面影像無法辨識出受測者，本系統之錯誤率約為 0.35%，這是因為系統比對值之閾值設定為 8.55，較為嚴苛，以取得較精準之結果，相對的正面辨識不到目標的影像數量即會提高。

表六各種影像之辨識結果

測試影像 \ 類型	辨識正確 (張數)	辨識錯誤 (張數)	正面影像辨識 不到目標(張數)	側面影像 (張數)	總張數 (張數)
B01_01 (兩盞日光燈)	41	0	0	33	74
B01_02 (一盞日光燈)	41	0	1	16	58
B01_03 (日光燈全關)	27	0	9	26	62
B01_04 (未戴眼鏡)	32	0	5	18	55
B01_05 (側面光照影響)	28	1	19	50	98
B02_01 (兩盞日光燈)	61	0	31	15	107
B03_01 (衣服接近膚色)	101	1	3	16	121
總和	331	2	68	174	575

註：此為 wBHE 之辨識結果

在大多數正確擷取之影像，即使遭受光照影響，大多能辨識出受測者身份，在實驗四中，許多無法成功辨識之影像，因遭受光照影響，造成臉部擷取不夠準確，因此，臉部擷取若能更為準確，則更能大大提升本系統辨識之準確率。

七、3D 影像用於人臉追蹤辨識之研究

以上各章節所述均為 2D 影像應用於目標物之追蹤與辨識，若是將應用侷限在人臉之辨識則 2D 影像與 3D 影像比較的話，3D 影像佔有極大的優勢，例如 2D 影像及亦受到未裝與光照之影響，如圖 108 所示：



圖 108 人物偽裝與受光照之影響

除此之外，2D 影像易產生假警報，不能解釋 3D 之模板，不能將年齡、體重、頭髮、鬍鬚、戴眼鏡列入辨識之考慮中，攝影機與姿勢嚴重的影響到辨識之正確性，室外的陰影與光照會使得辨識之結果差異相差很大。

實際的 3D 影像需要兩台相距固定距離的攝影機來擷取如圖 109 所示。

Stereo Face Capture for Real 3D

- Pair of cameras known with respect to each other
- Automated calibration and exposure
- Visual features are found in both images and matched
- 3D position of a matched feature is uniquely determined by its position on both focal planes
- Only Passive Stereo Solution Available

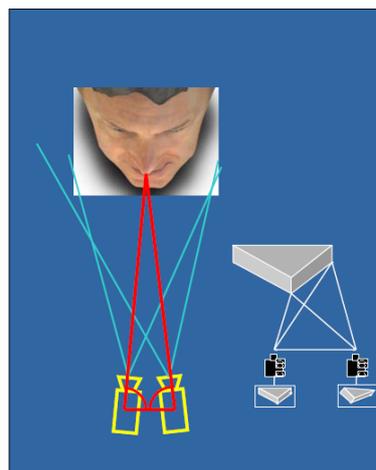


圖 109 3D 影像擷取原理

3D 影像擷取完後會產生一個 3D 的網點模型，利用這個模型來從事辨識就不會受各種限制 2D 辨識的因素影響。圖 110 是一個產生 3D 網點模型的例子。

Enabled by FaceVision™ 3D Regions of Interest for Fusion

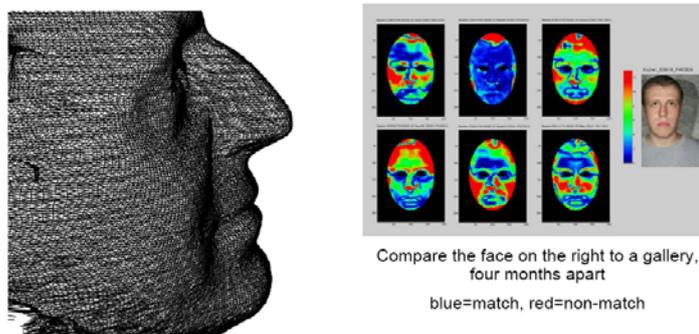


圖 110 3D 網點模型

圖 111 是經由 3D 攝影機擷取經由某段時間後截然不同外型之同一人，由辨識後的結果可知，3D 辨識軟體仍然可以精確無誤的辨識出來。因此 3D 影像的應用可說是日新月異，各個領域都可用到，如圖 112 所示。

18 Months minus 40Lbs minus Full Beard Verified with 3D Face

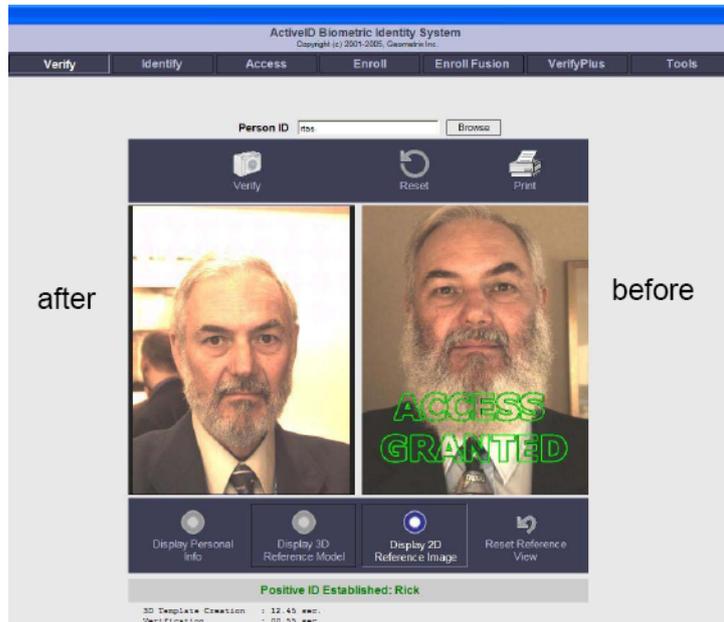


圖 111 經由一段時間後不同外型之同一人辨識結果

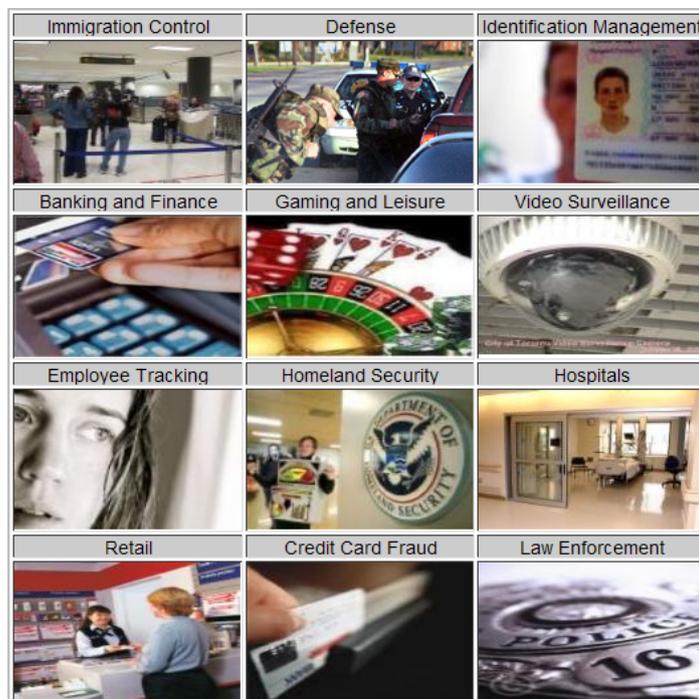


圖 112 3D 影像之各種應用

3D 影像之辨識其實整個流程與 2D 影像之辨識有些雷同，只是加入了 3D 攝影機與 3D 之模型來從事辨識，因此其模擬方塊圖可以大致以圖 113 來描述。

3D Enhances Existing 2D Recognition Leverage Existing 2D Image Databases

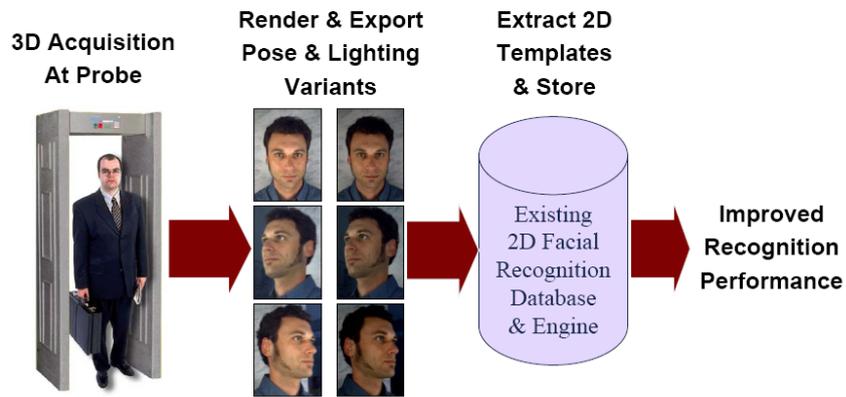


圖 113 3D 影像辨識流程圖

以下各圖就是利用 3D 影像模擬軟體經由 3D 攝影機取像後所產生的各種 3D 資訊圖。



圖 114 3D 影像圖

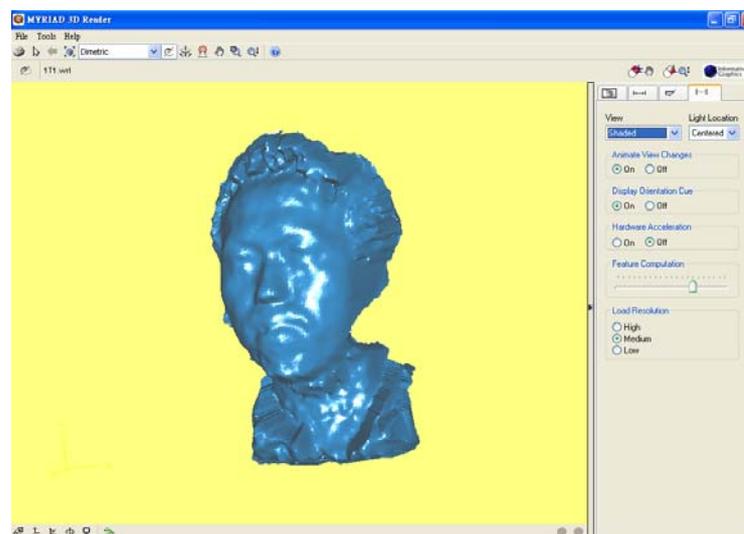


圖 114 3D 陰影圖

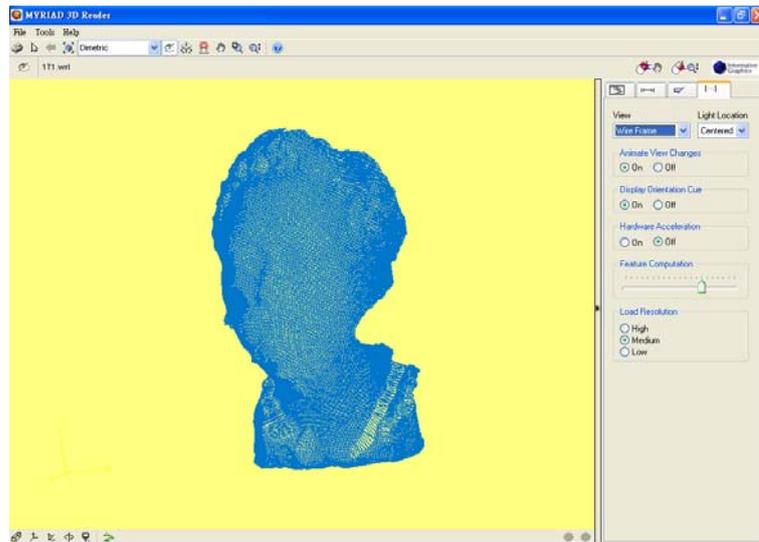


圖 115 3D 網點圖

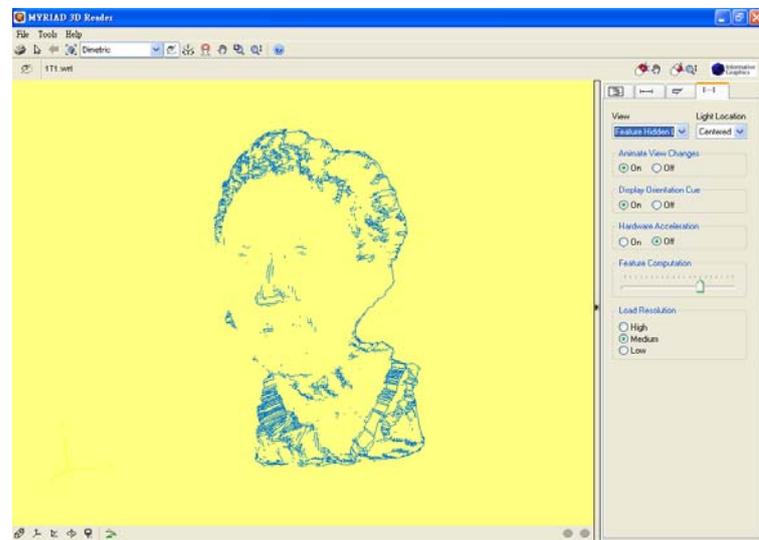


圖 116 3D 特徵圖



圖 117 不同角度 3D 影像圖



圖 118 不同表情 3D 影像圖

由圖 114 至圖 118 可以看出 3D 圖形所包含的大量資訊可以幫助我們從事人臉的辨識，提升辨識之正確率且克服環境因素之影響。

八、結果與討論

本報告目的在研究目標物之追蹤與辨識之方法，利用此研究結果提出了一套入侵人物預警與辨識系統開發系統，此方法先運用固定式攝影機拍攝影像序列，在經過影像處理、追蹤偵測演算法 模擬程序後，持續追蹤移動的物體，並將擷取到的目標特徵與資料庫比對，以判斷該目標是否具有合法身份，具有減少人為誤判、即時回報突發狀況之效果，因應本校之特殊服務需求基地再未來可以開發一套系統，未來可取代國軍營區或機敏庫房衛哨查核人力。

首先的作法是在靜態場景下固定背景中的移動物體偵測，此方式也容易受到光線的變化、背景景物細微變動等影響而降低偵測的效果。在本報告中也提出光照正規化的演算法來解決此一問題，我們先利用背景影像相減法來去除背景，此法比起連續的兩張影像相減法在執行速度上較具優勢，本報告中採用背景相減法，利用事先統計的平均背景資料與畫面做比較，判斷是否有在移動的前景物體，若發現有移動物體時，使用二值化、型態學中的侵蝕膨脹等基本的影像前處理來消除雜訊凸顯移動物體，獲得較乾淨的主體影像及所在位置資訊，再把目標物體交由追蹤模組處理。

而在追蹤系統中，我們比較卡曼濾波器及核心函數兩種演算法的追蹤效果。在移動物體追蹤程序中，可以根據一張畫面測量到的移動物體位置，利用卡曼濾波器對當前畫面中目標

的位置、速度進行估計，並利用此估計值對目標在下一張畫面的位置做出預測，此法優點是減少運算量，能夠適應目標快速移動的情況，也能解決短時間遮蔽的問題。而核心函數演算法是以移動物體偵測系統所得到的移動物體的區域為樣版，因中間畫素的資訊重要度較邊緣高，乘上核心函數賦予不同的權重，在以後的影像中，以平均位移演算法反覆疊代出相似度最高的平均位移向量，此結果便為物體移動後的位置，而達到追蹤的效果，對於目標遭遇遮蔽、相似目標干擾，此一方法則容易發生追蹤失敗情況。

追蹤到的樣版影像，本報告利用二維主成分分析法(2DPCA)、二維線性識別分析法(2DLDA)及獨立成分分析法(ICA)執行特徵擷取，最後再以歐基里德距離進行辨識，由實驗結果可知，維度為 39x30 像素的模擬人臉影像，在辨識運算時間很短符合即時辨識警告需求。

目前本系統在卡曼濾波器追蹤方面的處理速度為 11.3fps，正確率為 70%以上，辨識處理每張平均執行時間為 0.64 秒，辨識率為 80%以上，證明本研究不僅在追蹤及辨識上能符合預警辨識的效能，正確率及辨識率均有不錯的表現。若將自動追蹤及辨識系統做整合，只要有物體進入畫面均可先行實施追蹤，並將擷取的人臉與資料庫內圖片做比對，若非資料庫內人員則實施警告標示，可減少衛哨人力的值勤，達到即時的人物車輛預警與辨識效果。

在目標偵測追蹤過程中，最大的難題是如何準確的分割出每一個物體，每個場景容易受到雜訊干擾、兩個移動目標太過接近、長時間物體遮蔽等影響造成誤判，這是本系統需要克服及改進的地方，在未來可朝改善影像分類演算法，利用特定色彩來輔助物體分割、以及加入背景補償機制等方式，強化偵測效能，增進追蹤速度。而本研究所獲得的成果，現階段可運用於營區警監系統，未來希望能夠利用關鍵技術，推廣應用於三軍目標偵測與識別的相關應用上。

進行人臉識別時，時常因為光影的影響，也就是當光源由側邊發出，或影像偏暗，使臉部光照不平均，因而降低其辨識率，本研究提出光照正規化，透過光照正規化方法，可降低光影對辨識所造成的影響，之後再透過基礎影像處理、光照正規化、辨識演算法及門禁警監系統，以完成整個系統開發之研究程序。

本報告中為了研究發展與使用上的方便性，將我們所開發的許多演算法以模組化的撰寫方式開發一套視窗介面系統，已進行實驗與模擬，此一介面系統其中一個主要模組式光照正規化模組，其中包含 7 種光照正規方法，如直方圖等化(HE)、權重區域直方圖等化(wBHE)、

離散餘弦轉換(DCT)、伽瑪強度修正(GIC)、區域正規化(LN)、自商數影像(SQI)及反射比(RR)，並透過 Yale Face Database B 之資料庫影像以取得光照處理之辨識之測試結果。光照正規方法中，以權重區域直方圖等化(wBHE)、自商數影像(SQI)及反射比(RR)之辨識結果最好。

本介面系統另一個模組包含了各種辨識演算法，如主成分分析法(PCA)、線性鑑別分析法(LDA)、二維主成分分析法(2DPCA)、二維線性鑑別分析法(2DLDA)、二維雙向主成分分析法(2D2PCA)及二維雙向線性鑑別分析法(2D2LDA)等方法，這些方法皆透過其方法計算出訓練影像之特徵向量，再將訓練影像和欲辨識之影像分別透過特徵向量投影至特徵空間，最後再比較其歐幾里德距離，以取得辨識之結果。

利用上述所開發的各個演算法與模組，發展出一套門禁警監系統，本系統之運作乃透過取得自動更新之背景，再透過背景取得移動物體之主體，接著透過膚色以擷取臉部正確位置，當人臉正面靠近攝影機時，即可辨識出是否為資料庫之人物，但若人臉為側臉時，系統依照所擷取之側臉影像，無法於資料庫中找到相近之人臉，因此未來希望加入 3D 之資訊，即使為側臉影像也能準確的分辨出目標人物。

參考文獻

- [1] M. I. Miller, A. Srivastava, and U. Grenander, "Conditional-mean estimation via jump-diffusion processes in multiple target tracking/recognition," *Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on]*, vol. 43, pp. 2678-2690, 1995.
- [2] J. K. E. Tunaley, "Algorithms for ship detection and tracking using satellite imagery," in *Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004 IEEE International*, 2004, pp. 1804-1807 vol.3.
- [3] X. Lei, Z. Guangxi, W. Yuqi, X. Haixiang, and Z. Zhenming, "Real-time vehicles tracking based on Kalman filter in a video-based ITS," in *Communications, Circuits and Systems, 2005. Proceedings. 2005 International Conference on*, 2005, p. 886.
- [4] G. Wei, B. Dong-Liang, and L. Lu, "Human motion tracking based on shape analysis," in *Wavelet Analysis and Pattern Recognition, 2007. ICWAPR '07. International Conference on*, 2007, pp. 1288-1293.
- [5] Z. Tao and R. Nevatia, "Tracking multiple humans in complex situations," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 26, pp. 1208-1221, 2004.
- [6] Z. Haitao and Y. Pong Chi, "Incremental Linear Discriminant Analysis for Face Recognition," *Systems, Man, and Cybernetics, Part B, IEEE Transactions on*, vol. 38, pp.

210-221, 2008.

- [7] S. Mitra and T. Acharya, "Gesture Recognition: A Survey," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 37, pp. 311-324, 2007.
- [8] C. Qing, N. D. Georganas, and E. M. Petriu, "Real-time Vision-based Hand Gesture Recognition Using Haar-like Features," in *Instrumentation and Measurement Technology Conference Proceedings, 2007. IMTC 2007. IEEE*, 2007, pp. 1-6.
- [9] W. Jian-Gang, Y. Wei-Yun, A. Suwandy, and E. Sung, "Fusion of Palmprint and Palm Vein Images for Person Recognition Based on "Laplacianpalm" Feature," in *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, 2007, pp. 1-8.
- [10] M. Li, T. Tieniu, W. Yunhong, and Z. Dexin, "Personal identification based on iris texture analysis," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 25, pp. 1519-1533, 2003.
- [11] W. Cheokman, O. Lei Chan, W. Chan Hon, K. Tong Sio, and N. Kengchung, "A Macao license plate recognition system," in *Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on*, 2005, pp. 4506-4510 Vol. 7.
- [12] D. Gonzalez-Jimenez and J. L. Alba-Castro, "Toward Pose-Invariant 2-D Face Recognition Through Point Distribution Models and Facial Symmetry," *Information Forensics and Security, IEEE Transactions on*, vol. 2, pp. 413-429, 2007.
- [13] I. Mpiperis, S. Malassiotis, and M. G. Strintzis, "3-D Face Recognition With the Geodesic Polar Representation," *Information Forensics and Security, IEEE Transactions on*, vol. 2, pp. 537-547, 2007.
- [14] A. K. Ioannis, P. Georgios, T. George, N. M. Mohammed, L. Yunliang, K. Nikos, and T. Theoharis, "Three-Dimensional Face Recognition in the Presence of Facial Expressions: An Annotated Deformable Model Approach," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, pp. 640-649, 2007.
- [15] Z. L. Stan, C. RuFeng, L. ShengCai, and Z. Lun, "Illumination Invariant Face Recognition Using Near-Infrared Images," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, pp. 627-639, 2007.
- [16] B. Pradeep, T. P. Ioannis, T. Panagiotis, and B. Mike, "Physiology-Based Face Recognition in the Thermal Infrared Spectrum," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, pp. 613-626, 2007.
- [17] G. Xin, Z. Zhi-Hua, and K. Smith-Miles, "Automatic Age Estimation Based on Facial Aging Patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, pp. 2234-2240, 2007.
- [18] L. Zhulin, X. Chao, and L. Yan, "Robust object tracking using mean shift and fast motion estimation," in *Intelligent Signal Processing and Communication Systems, 2007. ISPACS 2007. International Symposium on*, 2007, pp. 734-737.
- [19] Z. Yuhua, and M. Yan, "Adaptive Object Tracking using Particle Swarm Optimization Adaptive Object Tracking using Particle Swarm Optimization," in *Computational Intelligence in Robotics and Automation, 2007. CIRA 2007. International Symposium on*,

2007, pp. 43-48.

- [20] Z. Xiaoqin, H. Weiming, S. Maybank, and A. X. L. Xi Li, "Graph Based Discriminative Learning for Robust and Efficient Object Tracking," in *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, 2007, pp. 1-8.
- [21] Y. Rathi, N. Vaswani, and A. Tannenbaum, "A Generic Framework for Tracking Using Particle Filter With Dynamic Shape Prior," *Image Processing, IEEE Transactions on*, vol. 16, pp. 1370-1382, 2007.
- [22] S. W. Lee, J. Kang, J. Shin, and J. P. A.-J. Paik, "Hierarchical active shape model with motion prediction for real-time tracking of non-rigid objects," *Computer Vision, IET*, vol. 1, pp. 17-24, 2007.
- [23] S. Heon Soo, K. Sung Min, K. Joo Woong, and A. K. H. E. Ki Hwan Eom, "Novel object tracking method using the block-based motion estimation," in *SICE, 2007. Annual Conference*, 2007, pp. 2535-2539.
- [24] G. Hamid, and G. Shaoshuai, "3-D Motion Estimation Using Range Data," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 8, pp. 133-143, 2007.
- [25] C. Sminchisescu, D. Metaxas, and S. Dickinson, "Incremental model-based estimation using geometric constraints," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, pp. 727-738, 2005.
- [26] K. Hariharakrishnan, and D. Schonfeld, "Fast object tracking using adaptive block matching," *Multimedia, IEEE Transactions on*, vol. 7, pp. 853-859, 2005.
- [27] A. Amer and A. Amer, "Voting-based simultaneous tracking of multiple video objects," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 15, pp. 1448-1462, 2005.
- [28] K. Hyunjong, S. Jeonho, and P. Joonki, "Wavelet transform-based hierarchical active shape model for object tracking," in *Intelligent Signal Processing and Communication Systems, 2004. ISPACS 2004. Proceedings of 2004 International Symposium on*, 2004, pp. 256-261.
- [29] M. Greenspan, S. Limin, and P. Jasiobedzki, "Efficient tracking with the Bounded Hough Transform," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, 2004, pp. I-520-I-527 Vol.1.
- [30] P. Vadakkepat, P. Lim, L. C. De Silva, J. Liu, and L. Li Li, "Multimodal Approach to Human-Face Detection and Tracking," *Industrial Electronics, IEEE Transactions on*, vol. 55, pp. 1385-1393, 2008.
- [31] Z. Qiang, Y. Mei-Chen, C. Kwang-Ting, and S. Avidan, "Fast Human Detection Using a Cascade of Histograms of Oriented Gradients," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, 2006, pp. 1491-1498.
- [32] Q. Yu and G. Medioni, "Integrated Detection and Tracking for Multiple Moving Objects using Data-Driven MCMC Data Association," in *Motion and video Computing, 2008. WMVC 2008. IEEE Workshop on*, 2008, pp. 1-8.
- [33] Y. Qian, G. Medioni, and I. Cohen, "Multiple Target Tracking Using Spatio-Temporal Markov Chain Monte Carlo Data Association," in *Computer Vision and Pattern Recognition*,

2007. *CVPR '07. IEEE Conference on*, 2007, pp. 1-8.
- [34] Y. Chang, G. Medioni, K. Jinman, and I. Cohen, "Detecting Motion Regions in the Presence of a Strong Parallax from a Moving Camera by Multiview Geometric Constraints," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, pp. 1627-1641, 2007.
- [35] M. Changki and G. Medioni, "Motion Segmentation by Spatiotemporal Smoothness Using 5D Tensor Voting," in *Computer Vision and Pattern Recognition Workshop, 2006 Conference on*, 2006, pp. 199-199.
- [36] M. Nicolescu and G. Medioni, "A voting-based computational Framework for visual motion analysis and interpretation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, pp. 739-752, 2005.
- [37] J. Kang, I. Cohen, and G. Medioni, "Tracking objects from multiple and moving cameras," in *Intelligent Distributed Surveillance Systems, IEE*, 2004, pp. 31-35.
- [38] Lipton, A.J., Fujiyoshi, H. and Patil, R.S., "Moving target classification and tracking from real-time video," Fourth IEEE Workshop on Applications of Computer Vision Proceedings. WACV '98, pp. 8-14, 1998.
- [39] Salmond, D., "Target tracking: introduction and Kalman tracking filters," IEEE in Target Tracking: Algorithms and Applications (Ref. No. 2001/174), pp. 1-16, 2001.
- [40] Welch, G. and Bishop, G., "An Introduction to Kalman filters," Department of Computer Science University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3175, 2004.
- [41] Comaniciu, D. and Meer P., "Mean shift: a robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp.603-619, 2002.
- [42] Comaniciu, D., Ramesh, V. and Meer P., "Kernel-based object tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp.564-577, May, 2003.
- [43] 鍾宜岑, "應用於動態背景中的移動物體影像之偵測與即時追蹤系統", 碩士論文, 國立交通大學, 新竹, 第 21-34 頁, 2007。
- [44] 簡隆至, "即時移動物體偵測及自動追蹤系統", 碩士論文, 國立臺灣科技大學, 台北, 第 47-76 頁, 2004。
- [45] Comaniciu, D. and Meer, P., "Mean shift analysis and applications," The Proceedings of the Seventh IEEE International Conference on Computer Vision, pp. 1197-1203, 1999.
- [46] Kailath, T., "The Divergence and Bhattacharyya Distance Measures in Signal Selection," *IEEE Transactions on Communication Technology*, pp. 52-60, Feb., 1967.
- [48] Vishwakarma, V. P., Sujata, P. and Gupta, M. N., "A Novel Approach for Face Recognition Using DCT Coefficients Re-scaling for Illumination Normalization," International Conference on Advanced Computing and Communications, pp. 535-539, Dec., 2007.
- [49] Xie, X. and Lam, K.-M., "Face recognition under varying illumination based on a 2D face shape model," *Pattern Recognition*, Vol. 38, No. 2, pp. 221-230, Feb., 2005.
- [50] Shan, S., Gao, W., Cao, B. and Zhao, D., "Illumination normalization for robust face

recognition against varying lighting conditions,” IEEE International Workshop on Analysis and Modeling of Faces and Gestures, pp. 157-164, Oct., 2003.

- [51] Xie, X. and Lam, K.-M., “An efficient illumination normalization method for face recognition,” *Pattern Recognition Letters*, Vol. 27, No. 6, pp. 609-617, Apr., 2002.
- [52] Wang, H., Li, S. Z. and Wang, Y., “Face recognition under varying lighting conditions using self quotient image,” Sixth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 819-824, May, 2004.
- [53] He, L., Liu, J., Zhang, B. and Wang, Z., “Using Quotient Image and Linear Subspace for Face Recognition under Arbitrary Illumination Conditions,” International Asia Conference on Informatics in Control, pp. 231-234, Feb., 2009.
- [54] Jiang, X., Cheng, Y., Xiao, R., Li, Y. and Zhao, R., “Spherical harmonic based linear face de-lighting and compensation,” *Applied Mathematics and Computation*, Vol. 185, No. 2, pp. 857-868. Feb., 2007.
- [55] Ding, S.-F., Shi, Z.-Z., Liang, Y. and Jin, F.-X., “Information Feature Analysis and Improved Algorithm of PCA,” IEEE International Conference on Machine Learning and Cybernetics, Guangzhou, China, Vol. 3, pp. 1756-1761, Aug., 2005.
- [56] Wen, Y. and Shi, P., “Image PCA: A New Approach for Face Recognition,” IEEE International Conference on Acoustics, Speech and Signal Processing, Shanghai, China, Vol. 1, pp. I-1241 - I-1244, April, 2007.
- [57] Martinez, A. M. and Kak, A. C., “PCA versus LDA,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 2, pp. 228-233, Feb., 2001.
- [58] Xu, A., Jin, X., Jiang, Y. and Guo, P., “Complete Two-Dimensional PCA for Face Recognition,” IEEE International Conference on Pattern Recognition, Beijing, China, Vol. 3, pp. 481-484, 2006.
- [59] Yang, J., Zhang, D., Yong, X., Frange, A. F. and Yang, J.-Y., “Two-Dimensional PCA: A New Approach to Appearance-based Face Representation and Recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, No. 1, pp. 131-137, Jan., 2004.
- [60] Kongsontana, S. and Rangsanteri, Y., “Face Recognition Using 2DLDA Algorithm,” IEEE Eighth International Symposium on Signal Processing and Its Applications, Bangkok, Thailand, Vol. 2, pp. 675-678, August, 2005.
- [61] 王精忠, “車牌辨識系統之研究”, 碩士論文, 大同大學通訊工程研究所, 台北, 第 23-26 頁, 2005。

心得

這次本人有機會透過國科會補助赴美短期研究，經過近三個月的研究有以下幾點心得，一、本人求學期間至最高學歷的求學階段都是在國內完成，經由這次難得機會至美國一流大學研究學習，使本人對於國外的教學與研究有一與國內環境完全不同的見識，了解到國外大學與學生無論是在求學或者生活上所遭遇到的各項問題以及他們如何克服解決，尤其是分享體驗留學生在國外求學的各项難得經驗。二、了解到國外學者在研究與教學上與國內的差異性有哪些，他們因為學校政策的關係，在系裡面有許多來至各個國家的訪問學者，這些學者的來訪提供了各項技術交流的機會，使得實驗室裡常常可以激盪出一些新知，活絡了大家的思想。三、因為學校辦學良好經費充足，系裡面有來至各國的留學生，舉目所見其中以大陸、韓國與印度的留學生為最多，反而台灣的留學生以不若以前佔大多數，這些消長的原因值得我們探討深思。四、實驗設備的添購與資訊的獲得較國內方便，許多實驗設備的總公司都設在美國，所以新設備的資訊獲得或者採購都較台灣方便很多。五、整體而言，一個實驗室的規模與產能除了需要有充足的經費設備之外還需要有優秀的學生，這兩者才可以相輔相成，我們國內許多大學的實驗室所具備的研究場所設備並不輸給國外大部分的大學，若是善加利用並且提高研究人員的素質，整體表現應該不會輸給國外的大學。六、國內企業應多提撥經費給研究學者出國研究的經費，不要只是靠國科會的經費補助，與我同研究室的韓國教授就是接受企業的補助提供所有食宿經費赴美從事一年的研究。

建議事項

本人經由這次赴美研究後有以下幾點建議：

- 一、 不論是政府機關教育機構或是鼓勵企業提供並多寬列經費補助國內學者赴國外研究，此舉不只對於很多沒有留學過的老師在各方面都有很大的助益，而且同時可以帶動學術的交流
- 二、 補助老師經費的同時也請考慮是否可以同時補助博士班學生隨指導教授同行如此可以使研究的效果加成
- 三、 政策上是否可以幫助各校多邀請訪問學者來校從事短期研究與教學，這種

作法有助於提升學校之研究與教學水準

- 四、 開放並且鼓勵補助各國研究生來台留學，可以刺激國內學生與老師的思維，並且長期而言對於外交也是一項很大的助益
- 五、 全面的建置無線網路，這是我這次在美期間體驗到無線上網無所不在的經歷，其方便性對於資訊的交流有極大的幫助
- 六、 教學研究設備無論是軟硬體都應盡量加以充實，除了人之外這是教學研究不可或缺的一環，所謂巧婦也難為無米之炊