

智慧型校園大樓門禁監控系統

An Entrance Guard Surveillance System for Intelligent Campus Buildings

蔡俊明

臺北市立教育大學資訊科學系

cmtsai@tmue.edu.tw

葉榮木

國立台灣師範大學機電科技學系

zongmu@ntnu.edu.tw

摘要

傳統的校園大樓門禁系統是以磁卡、鑰匙或密碼為主，其缺點是必須隨身攜帶，容易造成遺失或忘記，人員管理上，也會有漏洞。由於科技的進步，影像辨識技術可以應用於門禁監控系統，減少上述問題的發生。所以，本文提出以智慧型的影像辨識技術為基礎，來做校園大樓門禁監控系統。本系統主要包括人臉偵測和人臉辨識，其中人臉偵測包括膚色偵測和人臉確認，在膚色偵測中，因人的膚色容易受亮度的影響，所以，首先在不同亮度下，分別訓練出不同亮度的CbCr值，來做膚色偵測；接著，近距離的人臉，利用人臉中有眼睛和嘴唇以及上方有頭髮三種特徵，來確認出人臉區域，遠距離的人臉，利用變異度和頭髮，來驗證出人臉；最後，在人臉辨識中，利用主成份分析法和灰關聯分析法，做人臉辨識。經由實驗結果顯示，動態人臉偵測率和人臉辨識率分別達97.0%和94.3%。

The traditional entrance guard system of campus buildings relies mainly on magnetic stripe cards, keys or password. Its shortcoming must be hand-carried and it is apt to cause losing and forgetting. Also, the

personnel management will have loophole. Because of the progress of science and technology, the face recognition can be applied to entrance guard's surveillance systems to reduce above described problems. Thus, this paper proposes an entrance guard surveillance system based on intelligent image recognition method. This system includes face detection and face recognition. The face detection includes skin detection and face identification. In the skin detection, the human skin is easily affected by luminance. So, we train different CbCr values for different class's luminance to segment the skin color. Next, eye, lip, hair and variance features are used to identify the face area. Finally, the Principle Component Analysis (PCA) and Grey Relational Analysis (GRA) are used to recognize the faces. Experiment results show that the proposed method has good performance. The dynamic face detection rate and dynamic face recognition rate are up to 97.0% and 94.3%, respectively.

關鍵詞：膚色偵測、人臉偵測、人臉辨識、主成份分析法、灰關聯分析法，Skin

detection, Face detection, Face recognition, Principle component analysis, Grey relational analysis。

一、緒論

現在的學校，為了學校財產和學童的安全，都裝設了視訊監控系統，當學校遭受小偷入侵，甚至財產失竊時，可以事後調閱錄影帶，提供警方辦案線索。但是，這是屬於事後再調查，視訊監控系統無法提供『即時資訊』，立即通知警衛和警方馬上處理。另外，校園警衛人數不多，要照顧到全校安全、巡邏校園、門口管制人員和車輛的進出等工作，非常繁雜，所以，學校常被宵小入侵，偷竊財產的事，時有所聞。另外，雖然有裝設視訊監控系統，但是因為錄到的影像品質不佳，或是因為光線問題，或是因為監控位置設計不佳，錄影到的影像不是很理想，加上需要額外人力去看影片，造成相關單位許多工作負擔。

傳統的校園大樓門禁系統是以磁卡、鑰匙或密碼為主，其缺點是必須隨身攜帶，容易造成遺失或忘記，人員管理上，也會有漏洞。以影像為基礎的生物辨識系統，是近幾年熱烈研究的題目，它是一個最方便的生物認證方式。各級學校的大樓，如果都能使用人臉辨識系統，那門禁進出管理會變得很方便很自然，也能節省大樓警衛人力。所以，本文提出一以膚色偵測、人臉確認和人臉辨識為基礎的智慧型校園大樓門禁監控系統，希望來解決上述傳統的校園門禁管理問題。

圖 1.1 是本文所提出的智慧型校園大樓門禁監控系統流程圖，主要包括兩大部分：系統使用前的訓練和線上系統測試。在使用前的訓練階段，要先將現有編制人員或新進人員的影像擷取到電腦中，再利用膚色偵測和人臉確認組成的人臉偵測演算法將人臉偵測出來，最後，抽取特徵存入人臉特徵資料庫中。在線上系統測試階段，利用 CCD 擷取影像，同樣，利用人臉

人臉特徵建制系統 人員監控系統

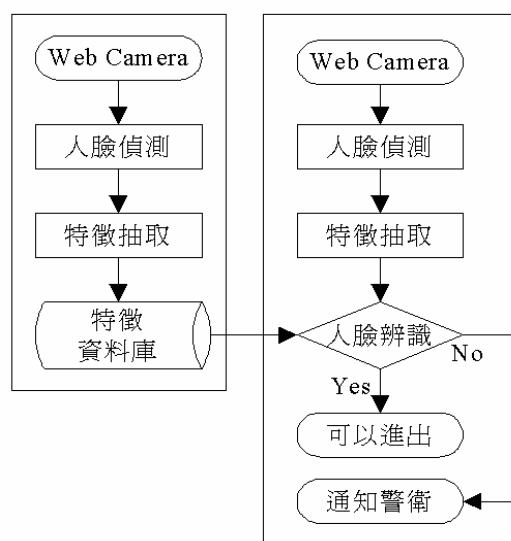


圖1.1『智慧型校園大樓門禁監控系統』流程圖。

偵測模組來偵測人臉，再利用人臉辨識模組抽取特徵與人臉資料庫比對，若是校方人員，解除門禁，讓其進出，並記錄進出時間，若不是校方人員，則通知警衛處理，以下章節分別介紹之。

二、文獻回顧

(一) 人臉偵測的相關研究

人臉偵測的研究，大致上可以分為三大類[1]：基於特徵法(feature-based methods)，模板比對法(template matching methods)，影像分析法(image-based methods)。基於特徵法[2-4]是根據人臉結構上的特徵，例如眼睛、嘴巴、鼻子、眉毛等特徵，先將其定位出來，再利用統計或特徵在人臉結構上的比例，來確定人臉的位置。此類做法容易受到外在因素的影響，例如，燈光的強度和背景的複雜度，會影響到偵測的結果。模板比對法[5-6]是利用事先定義好的人臉 patterns 或人臉上各個器官的 patterns，和輸入影像做比對，來偵測出人臉所在區域。此類方法可以找出人臉區域並且判斷此區域是否為人臉，這方法對於戴眼鏡或是落腮鬚的人不適用，而且針對不同角度的人臉，也要事先建立不同人

臉模板，才可以處理。影像分析法[7-9]藉由統計分析 (statistical analysis) 與機器學習 (machine learning) 來找出人臉影像有意義的樣式，然後藉由不同大小的視窗，將輸入灰階影像與人臉樣式比對，進而判斷是否有人臉。由於此類方法是以各種大小的視窗掃描影像，來偵測不同大小、位置的人臉，因此，計算量相當的大。

(二) 人臉辨識的相關研究

人臉辨識的研究，大約可以分為四大類 [10]：以整體外觀為主的方法 (appearance-based approaches)，以樣板比對為主的方法 (template-matching approaches)，以階層式影像分析為主的方法 (range-image approaches)，以幾何分析為主的方法 (geometric-analysis approaches)。以整體外觀為主的方法 [11-15]，主要是收集同一類別，或者是同一對像的多筆影像樣本，並且以這些資料來代表此類別。也就是利用每張影像的像素 (pixel) 值做為數據並透過處理與計算，獲得每一類別最具代表性特徵，進而利用這些特徵來達到辨識的結果。以樣板為主的方法 [6] 是將人臉標準化為數個區域，例如，眼睛、鼻子、嘴巴、臉型 … 等，並將這些區域予以特徵化，用來驗證或辨別人臉。以階層式影像分析為主的方法 [16-17]，不同於上述的方法，此法是以 3D 的方式進行分析，此方法比 2D 圖片多了深度的資訊，利用在三度空間中人臉動作的模擬，使系統更貼近人類，以達到更準確和更有效的結果。以幾何分析方式 [18] 主要是利用人臉上五官之間的距離當作鑑別的依據，例如，眼睛、鼻子、嘴脣之間相對的距離，事先定義人臉上的特徵點，然後來做比對，因為人臉上的特徵點並無強制性，可以由系統建構者所決定，因此，此方法的效果會與臉上設定的特徵點有非常高度的相關性，所以這個方法並不是非常準確和可靠 [19]。

三、人臉偵測

在動態的人臉偵測應用中，需要有一具有強健性和即時性的偵測系統來偵測人臉。在本文所提出的人臉偵測步驟有：人臉候選區偵測，人臉特徵抽取，人臉確認。在人臉候選區偵測中，提出在不同亮度下的膚色偵測，雜訊去除，連通成分抽取，不適當連通成分去除。在人臉特徵抽取中，提出自適性閾值的臉部特徵抽取方法，抽取出頭髮，嘴巴和眼睛特徵。在人臉確認中，利用抽取出來的人臉特徵，確認出人臉的區域，以下分別說明這些方法。

(一) 人臉候選區偵測

在使用攝影機或是 CCD 對人臉取像時，發現人臉和攝影機之間的距離，會影響膚色的呈現，也就是當人臉靠近攝影機時，膚色的亮度會提高，遠離攝影機時，膚色會顯的比較暗淡，因為拍攝的遠近是我們無法控制的，若要對人臉先做光線補償，再去做人臉偵測，對於即時的系統而言，會造成負擔，甚至偵測的速度下降，所以，我們一開始就先統計好不同亮度的膚色範圍，這樣輸入的影像就可以依照每一個 pixel 的亮度去選擇他的膚色範圍，不用經過前處理，這樣可以提高系統的速度。

本文的膚色偵測是以色彩為基礎，根據各色彩成分的像素值大小，來進行膚色的判斷。在文獻 [20] 的方法中，發現在不同亮度下膚色所群聚的範圍不一樣，也就是說不同的亮度膚色，其 CbCr 的範圍是不一樣的，所以，我們重新統計不同亮度的膚色影像，來得到不同亮度下膚色的 CbCr 值，其結果如表 1 所示。另外，因為非色彩 (黑色，灰色和白色) 的 CbCr 也在 127~128 之間，所以，我們利用 $R_n = R / (R+B+G)$ 來避免非色彩被偵測為膚色，這是因為非色彩的 RGB 值差不多相等其 R_n 接近 1/3。另外，雖然膚色的 R 值很高，但是當一個 pixel 的 R 值大於 R+G+B 的一半時，這個 pixel 的顏色就接近紅色而不是膚色。所以，利用之前統計膚色範圍的人臉，來計算 R_n 的值，得到膚色的 R_n 範圍介於 0.36~0.6 之間。所以，所提

表 3.1 不同亮度下膚色群聚表

n	亮度值(Y)	Cb 膚色範圍	Cr 膚色範圍
0	$Y \leq 50$	$108 \leq Cb \leq 131$	$120 \leq Cr \leq 145$
1	$51 \leq Y \leq 110$	$83 \leq Cb \leq 132$	$123 \leq Cr \leq 175$
2	$111 \leq Y \leq 180$	$75 \leq Cb \leq 135$	$125 \leq Cr \leq 185$
3	$181 \leq Y \leq 205$	$68 \leq Cb \leq 138$	$115 \leq Cr \leq 160$
4	$206 \leq Y \leq 219$	$65 \leq Cb \leq 140$	$115 \leq Cr \leq 158$
5	$Y \geq 220$	$65 \leq Cb \leq 141$	$115 \leq Cr \leq 153$

出的膚色偵測方法，是利用訓練出的不同亮度下的膚色範圍，加上 Rn 範圍，可以偵測出正確的膚色範圍，其式子如 3.1 所示。

$$\text{Face}(x, y) = \begin{cases} 1 & \text{if } [(0.36 \leq Rn \leq 0.6) \text{ and } \text{Map}_n(a, b) = 1], \\ 0 & \text{otherwise,} \end{cases}$$

$$n = \begin{cases} 0 & \text{if } [Y(x, y) \leq 50], \\ 1 & \text{if } [51 \leq Y(x, y) \leq 110], \\ 2 & \text{if } [111 \leq Y(x, y) \leq 180], \\ 3 & \text{if } [181 \leq Y(x, y) \leq 205], \\ 4 & \text{if } [206 \leq Y(x, y) \leq 219], \\ 5 & \text{if } [Y(x, y) \geq 220], \end{cases}$$

$$Rn = R/(R + B + G),$$

(3.1)

其中 a 和 b 分別指 Cb、Cr 的膚色範圍，Y 表示在輸入影像(x, y)位置中的亮度值，n 是指不同的亮度等級，式子 3.1 的意思是輸入一個 pixel 並計算它的 Rn 值，假如其 Rn 值的範圍介於 0.36~0.6 之間，再計算它的亮度值，並判斷它是屬於 0~5 那一個 map 的範圍，再判斷它的 Cb、Cr 範圍是不是屬於這個 map，若是的話就判定為膚色。

在偵測完膚色後，因為在背景中會有一些 pixels 的顏色屬於膚色的範圍，這會產生雜訊。另外，也會有一些膚色沒有事先被訓練到，而造成膚色破碎。所以，利用形態學(morphology)的膨脹(Dilation)和侵蝕(Erosion)來處理這些問題。接下來，利用連通成分，找出膚色的連通區域，

因為在背景中可能出現和膚色相接近的原木色桌子或是類似膚色的壁紙，這些與膚色相近的背景，會產生大的連通成分。依照人體計測資料庫[21]的說明，人臉的長寬比是固定的，所以，利用長寬比，可以排除非人臉的膚色連通成分，最後，就得到人臉候選區域。

(二) 人臉特徵抽取

因為人臉距離攝影機有遠有近，近距離人臉的五官特徵明顯，所以抽取頭髮，嘴巴和眼睛特徵，以便確認人臉。遠距離時，人的五官不明顯，採取計算變異度和抽取頭髮。如何決定選用那一類的特徵擷取呢？這裡採用面積的方法來判斷，當人臉候選區的面積大於 A 時，抽取五官特徵，否則計算變異度和抽取頭髮特徵。

在人臉中五官的顏色比膚色的顏色較陰暗，例如，眼睛比膚色暗，嘴巴比膚色紅，頭髮的灰階值遠小於膚色灰階值，以下分別說明如何抽取特徵。首先抽取頭髮特徵，其原因是頭髮所占的區域較大，而且，頭髮顏色比較一致性和亮度值較低，一致性使用標準差來表示，理想上為 0，實際上，經由實驗得到為 10 以下，亮度值用平均值表示，經由實驗設為 60。抽取頭髮的演算法如下：在人臉候選區塊上，往左、上、右三個方向做固定延伸，得到頭髮候選區，先計算頭髮亮度平均值，若亮度平均值大於 60 則不是頭髮，若是小於等於 60，再計算標準差，若標準差小於 10，則判斷為頭髮，否則為非頭髮。

在嘴唇的偵測中，利用色彩分析發現，皮膚 R 值(紅色的成分)比嘴唇的 R 值高，而且嘴唇的 R 值不高，但是比 G、B 值(綠色、藍色的成分)高出許多，這也是嘴唇的顏色看起來像紅色。我們統計不同亮度下嘴唇的 R-G 和 R-B 值，發現當亮度提高後 RGB 值也會跟著提高，嘴唇顏色也隨著亮度的提高，其 R-G 和 R-B 的值，也會等比例的提高，經由實驗的知，在不同的亮度情況下，R-G 的值大約是 R 值的 0.25 倍以上，R-B 的值大約是 R 值的 0.3 倍以上。另外，在 YCbCr 空間，發現嘴唇的 Cr

比較高，偵測嘴唇的公式如式子 3.2 所示。

$$\text{Mouth}(x,y) = \begin{cases} 1 & \text{if } [Cr \leq T_{\text{mouth}}] \wedge [R-B > 0.25 \times R] \wedge [R-G > 0.3 \times R], \\ 0 & \text{otherwise,} \end{cases} \quad (3.2)$$

其中， T_{mouth} 是嘴唇二值化門檻值。還有人的嘴唇面積占整張人臉的 5%左右[21]，所以，利用式子 3.2 和面積，提出自適性嘴唇偵測方法。演算法開始時先設定嘴唇二值化門檻值($T_{\text{mouth}}=130$)，當輸入人臉候選區，利用式子 3.2 來偵測嘴唇，若輸入影像其像素的Cr值小於 T_{mouth} 時，統計嘴唇 pixel 個數，若是小於原始影像的 5%，表示嘴唇還沒有完全偵測到，更新 T_{mouth} 值(加 1)，重新用式子 3.2 來偵測嘴唇，直到嘴唇的像素點大於整個人臉候選區像素的 5%，表示偵測嘴唇已經偵測到，若是偵測到的嘴唇大於 1，因為嘴唇是屬於長方形，利用長寬比可以得到真正的嘴唇。

在眼睛的偵測中，因為不同人種的眼睛顏色(黑色、褐色、深藍色)不同，為了解決這個問題，先將眼睛做黑色增強，其方法如式子 3.3 所示：

$$g(x,y) = \begin{cases} \min & \text{if } [f(x,y) - \min < \max - f(x,y)], \\ \max & \text{otherwise,} \end{cases} \quad (3.3)$$

其中 $f(x,y)$ 代表 3x5 遮罩中心的灰階值，max和min分別為 $f(x,y)$ 相鄰區域灰階值中的最大和最小值，此公式可以增強黑色區域並改善影像中灰階值的對比程度。再來，因為眼睛比較暗，所以，可以設定一門檻值(T_{eye})，來偵測眼睛，但是，在不同亮度下和不同人種下，固定門檻值會出現問題，所以提出自適性的閾值，針對不同亮度和人種影像，給予不同的閾值。另外，因為眼睛的面積占整張臉的面積大約 2%左右[21]，所以，所提出的自適性眼睛偵測方法如下：先加強黑色影像，再把初始閾值(T_{eye})設為 0，計算輸入候選區的 pixels 數，影像從左到右從上到下計算每個 pixel 的灰階值，假如小於閾值就當作是眼睛的 pixels，當眼睛的 pixels 數目等於 2%候選區的 pixels 數目，這時的閾值就是

最適當的閾值，反之，當眼睛的 pixels 數目小於 2%候選區的 pixels 數目，我們就把閾值加 1 再重新計算一次，直到當眼睛的 pixels 數目大於 2%候選區的 pixels 數目，如此就可以偵測出眼睛後選區。因為眉毛和鼻孔的灰階度也低，容易被誤認為眼睛，所以，本文利用幾何規則，來確認出真正眼睛所在，所用的幾何規則如下：兩眼距離是臉寬的 0.2~0.8 倍，兩眼間的斜率必須小於 0.5(27°)，嘴唇位置在兩眼間，兩眼中心點到嘴唇的距離大於 0.3 倍臉長。

上述方法是當人臉候選區面積大時，因為人臉上的特徵比較清楚，所以可以抽出頭髮，嘴唇和眼睛，但是，當人臉候選區太小時，人臉上的特徵不明顯，則無法抽出特徵，對於這種情況，在此提出以人臉變異度為特徵的抽取方法。首先，我們觀察人臉無論是在遠方或是近處，或是有旋轉角度，它都具有旋轉不變(Rotation-invariant)[25]的特徵存在，所以，可以應用此特徵來做人臉驗證工作。由實驗結果發現，人臉候選區保有一定的變異度，所以，可以事先以訓練的方式找出人臉候選區的變異度範圍，在動態測試的時候，就可以直接算變異度的範圍，且並判定待測人臉候選區的變異度是否在範圍之內，我們提出的人臉變異度抽取方法介紹如下。

首先，輸入 800 張人臉，並將每一張人臉影像轉成灰階，再來，對灰階影像做 Sobel 邊緣偵測，再來，利用低通濾波器來降低高頻雜訊，這樣處理之後，人臉五官的邊緣特徵會特別明顯，接著，計算和統計每一張人臉的平均值和標準差，其分不如圖 3.1 所示，在此圖中可以看出，由 800 張影像所描繪出的資料點，其分佈像是兩條拋物線所限制的範圍，所以，利用拋物線的方程式來逼近分佈圖的上下線，將資料點帶入拋物線標準方程式，可以推導如公式(3.4)與(3.5)的結果。所以，這兩條拋物線，就是 800 張人臉影像訓練出的變異度範圍。

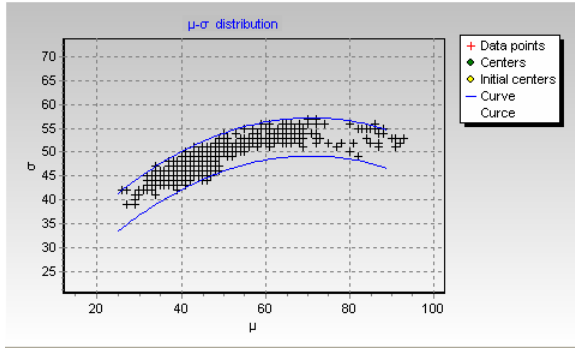


圖3.1 人臉平均值和標準差的二維分佈圖。

$$\sigma_1 = -0.0076\mu_1^2 + 1.0744\mu_1 + 16.2326 \quad (3.4)$$

$$\sigma_2 = -0.0076\mu_2^2 + 1.0744\mu_2 - 1.7674 \quad (3.5)$$

其中 σ_1 , μ_1 和 σ_2 , μ_2 分別表示上下拋物線的標準差和平均值。經過上述的訓練後，得到人臉變異度的範圍，所以，往後測試時，就可以直接做計算，並判定是否符合在兩條拋物線之間的範圍。

(三) 人臉確認

有了上述特徵的抽取，本文的人臉驗證流程如下：首先判定人臉候選區是否夠大（在此設為 50×50 ），若是夠大，只要有找到頭髮、眼睛和嘴唇三個特徵中兩個，並經過幾何判斷，若有符合條件時，就可以判定為人臉。若是不夠大，則計算變異度，代入拋物線方程式(3.4)和(3.5)，測試是否為人臉，若是為人臉，在測試其上方是否有頭髮，若有頭髮，就可以判定為人臉，否則為非人臉。

四、人臉辨識

在人臉辨識的階段，首先，我們將前面所偵測到的人臉，正規化為 112×92 的大小，接著，為了減少系統執行時間，使用Haar小波轉換[22]，對輸入影像做資料降低維度處理，再來，使用主成份分析法找出特徵臉，最後，使用灰關聯分析法來辨識出正確的人臉，以下分別介紹這些方法。

(一) 離散小波轉換

輸入的人臉影像大小為 112×92 ，如果不先降低影像維度，要辨識大量人臉，容易造成整個系統負荷過大，所以，使用小波轉換來降低影像的維度。這裡做兩次小波轉換，把影像的維度降低到 28×23 。二維小波轉換的方法如下：假設原影像大小為 $m \times n$ 像素，先沿著原始影像列向量做一維小波分解，分解成高頻和低頻兩部分，再將此兩部份各自沿著行向量分解成高頻與低頻兩部分，最後得到LL、LH、HL、HH四個子影像，這四個子影像的維度大小各為 $m/2 \times n/2$ 。做了兩次小波轉換後，利用LL重要的低頻影像，進行辨識特徵處理。

(二) 主成份分析法

主成份分析法[11-12](Principal Component Analysis, PCA)是使用較少的資訊來表示龐大的資料，這些少量資訊中所包含的特徵，對於原始資料而言是較具有代表性。以影像為例，一張 $N \times N$ 的影像可以轉成一維向量 $1 \times N^2$ 來代表一筆資料(I)。假如給定的資料有 M 筆(I_1, I_2, \dots, I_M)，主成份分析法是要找到一個轉換矩陣，利用這個轉換矩陣，將這些向量投影到一個較低維度的空間，這樣可以使用較少並且較具有代表性的資料來表示這筆資料。

取得用來訓練用的輸入人臉資料後，可以得到人臉訓練資料的平均向量(ψ)，如式子(4.1)所示，再來利用平均臉和每筆資料的差，求得差異向量(Φ)，如式子(4.2)所示。這裡平均向量代表平均臉，平均臉是用資料庫中所有人臉影像所計算出來的。

$$\psi = \frac{1}{M} \sum_{i=1}^M I_i \quad (4.1)$$

$$\Phi_i = I_i - \psi \quad (4.2)$$

利用上面差異向量，可以得到所有資料的共變異矩陣(Covariance matrix, C)，如式子(4.3)所示。

$$C = \frac{1}{M} \sum_{i=1}^M \Phi_i \Phi_i^T = AA^T \quad (4.3)$$

其中 $A = [\Phi_1 \Phi_2 \dots \Phi_M]$ 。計算完共變異矩陣可以得到特徵向量，本文選取前十個特徵向量來表示事先訓練的人臉資料庫影像，在此，特徵向量就是特徵臉。為了得到較小的資料量與差異性較大的特徵，必須找一個讓共變異矩陣 C 最大化的轉換矩陣 W_{opt} ，此轉換矩陣如式子(4.4)所示。

$$W_{opt} = \arg \max_w |W^T C W| \quad (4.4)$$

由線性代數的理論中得知，轉換矩陣 W 恰為共變異矩陣 C 的特徵向量所構成之矩陣，而其所相對應的特徵值，為其向量所代表的能量。但是共變異矩陣 C 是一個非常巨大的矩陣，很難計算其特徵向量，因此在[11]中使用 AA^T 矩陣來求特徵向量，這是因為 AA^T 中的 M 個特徵向量恰好是 AA^T 的 N^p 個特徵向量中，最大的前面 M 個特徵向量。所以，可以在這 M 個特徵向量中，選取 K 個特徵向量 $\{u_1, u_2, \dots, u_k\}$ ，來構成較低維度的子空間，每個差異向量便可以由式子(4.5)來表示：

$$\hat{\Phi} - \psi = \sum_{j=1}^K w_j u_j \quad (4.5)$$

其中 $w_j = u_j^T \Phi_i$ 稱為主成份 (principal component)。

(三) 灰關聯分析法

傳統的辨識比對法則大都是採用歐式距離，本文採用灰關聯分析法[23-24]做為辨識時的決策法則。灰色系統理論 (Grey System Theory)，主要是在不完整或沒有足夠資訊下，用來進行系統分析的理論。灰關聯分析 (Grey Relational Analysis) 是在灰色系統理論中，用來分析資料序列的工具，在給定一參考序列和一比較序列，透過灰關聯分析法就可以知道這些序列之間的關係。假設給定一參考序列集合為 $S_r = \{S_{r1}, S_{r2}, \dots, S_{rn}\}$ ，其中 $n \geq 1$ ，其序列為 $S_{ri} = \{S_{ri}(1), S_{ri}(2), \dots, S_{ri}(p)\}$ ；給定比較序列集合為 $S_c = \{S_{c1}, S_{c2}, \dots, S_{cm}\}$ ，其中 $m \geq 1$ ，且其序列為 $S_{cj} = \{S_{cj}(1),$

$S_{cj}(2), \dots, S_{cj}(p)\}$ 。上述 p 都是有限的正整數，比較序列與參考序列之間的關係稱為灰關聯度，其值域為 0 到 1，定義如式子(4.6)所示：

$$\gamma(S_{ri}, S_{cj}) = \sum_{k=1}^p \beta_k \gamma_{ij}(k) \quad (4.6)$$

其中 β_k 為一權重值，定義為 $\sum_{k=1}^p \beta_k = 1$ 。式子(4.6)常簡化為式子(4.7)式：

$$\gamma(S_{ri}, S_{cj}) = \frac{1}{p} \sum_{k=1}^p \gamma_{ij}(k) \quad (4.7)$$

其中 γ_{ij} 稱為灰關聯係數 (Grey Relational Coefficients)，定義為式子(4.8)[23]：

$$\gamma(S_{ri}(k), S_{cj}(k)) = \frac{\Delta_{\min} + \zeta \Delta_{\max}}{\Delta_{ij}(k) + \zeta \Delta_{\max}} \quad (4.8)$$

其中

$$\Delta_{ij}(k) = |S_{ri}(k) - S_{cj}(k)|,$$

$$\Delta_{\max} = \max_{j \in m} \max_{\forall} \Delta_{ij}(k),$$

$$\Delta_{\min} = \min_{j \in m} \min_{\forall} \Delta_{ij}(k),$$

ζ 為區別係數，其值介於 0 到 1 之間。但是式子(4.8)的分佈不夠線性，在此情況下，計算灰關聯度時，會因為加總的動作而扭曲了最後的結果，所以在[24]中提出了新的灰關聯係數的計算：

$$\gamma(S_{ri}(k), S_{cj}(k)) = \left[\frac{\Delta_{\max} - \Delta_{ij}(k)}{\Delta_{\max} - \Delta_{\min}} \right]^{\zeta} \quad (4.9)$$

其中區別係數 $\zeta \in (0, \infty]$ ， $\gamma(S_{ri}(k), S_{cj}(k)) \in (0, 1]$ 。雖然式子(4.9)也不是線性的轉換公式，但是當區別係數 $\zeta = 1$ 時會產生線性曲線，所以本文採用式子(4.9)，並且取 $\zeta = 1$ 的條件下，計算灰關聯係數，可以獲得較正確的辨識率。

五、實驗結果與討論

本文研究的電腦硬體配備為 CPU P4 2.4G/256MB RAM，應用介面是使用 Borland C++ Builder 6 撰寫，實驗影像擷取的攝影器材包括 FUJIFILM V10 數位相機，Logitech QuickCam Pro 以及 PCB-P100WebCam，每一個影片和擷取的影

表 5.1 不同拍攝距離的人臉偵測結果

距離	1 公尺	3 公尺	5 公尺
張數	1500	1500	1500
正確	1455	1410	1350
錯誤	45	90	150
正確率	97.0%	92.0%	90.0%
FPS	6	8	10

像大小是 320×240。本文的實驗分為動態人臉偵測和動態人臉辨識，以下分別說明之。

(一) 動態的人臉偵測

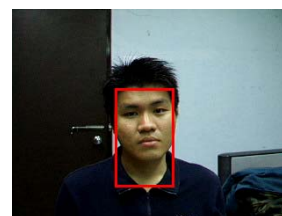
為了驗證本系統在不同的狀況下的偵測效果，設定了以下幾個實驗：(1)對正面而且不同拍攝距離的人臉偵測率。(2)對移動中的人做人臉偵測實驗。(3)對不同亮度下的人做人臉偵測實驗。(4)對不同角度拍攝的人臉做人臉偵測實驗。(5)對多人做人臉偵測實驗。

在對正面而且不同拍攝距離的人臉偵測實驗中，人與相機距離分別是 1 公尺、3 公尺及 5 公尺，實驗的影像分別對每次實驗者錄製 15 秒鐘的影像，在這 15 秒鐘實驗者可以輕微的擺動，在 15 秒鐘內大約會有 150 個 frame，每位實驗者在 15 秒內都取 50 張影像，在每次實驗中有 30 位實驗者，所以影像張數 1500 張，計算處理速度是使用 FPS 表示法，FPS 是 frame per second 的簡寫，其意義是每秒可以處理的畫面數，FPS=10 代表每秒鐘可以處理十張畫面，用此來驗證所提出的系統速度。

表 5.1 是動態的人臉偵測實驗的結果。在此實驗中可以發現當拍攝距離約 1 公尺時，人臉偵測率高達 97%，這是因為(1)使用事先訓練在不同亮度下膚色的 CbCr 值，相較於不分亮度的膚色訓練，可



(a)



(b)



(c)

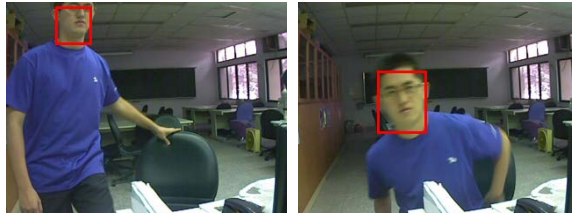
圖 5.1 不同拍攝距離的人臉偵測結果(a)1 公尺(b)3公尺(c)5公尺。

以提高膚色在不同亮度下的偵測，(2)使用自適性的閾值來偵測人臉特徵，對於每張不同的影像我們都分別給予不同的閾值來抽取眼睛和嘴巴，相較於使用固定的閾值，可以得到較高的偵測。對於沒有偵測到人臉影像的原因，是因為實驗者有閉眼的情況導致特徵點遺失。另外，因為拍攝距離較遠，偵測到的人臉越來越小，偵測率會下降，這是因為使用色彩分析來偵測人臉特徵，人臉影像越來越小時人臉特徵相對變小，當人臉特徵太小時就會被當成雜訊去除了，所以我們的系統在近距離時有較高的偵測率，但是遠距離時，系統就不能有效偵測人臉特徵，造成偵測率無法達到理想的效果。另外，對於處理的速度是根據連通成分的數量以及大小來決定，連通成分數量越多，後面處理步驟要花費較多時間，連通成分面積越大代表要處理的影像也變大，所以速度會降下來，所以可以看到人臉在近距離時，FPS 比較慢，但是隨著人臉在距離的增加，FPS



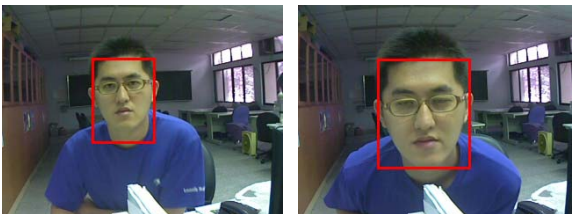
(a)

(b)



(c)

(d)



(e)

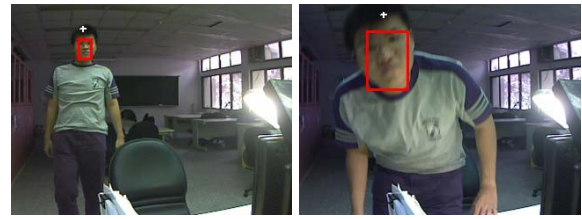
(f)

圖5.2 移動中的人做人臉偵測結果。(a)到(f)是人由遠到近。

會有提高的趨勢，所以，如何在偵測率與處理速度之間取得平衡，是一項直的研究的問題。圖 5.1 是在不同拍攝距離(1、3、5 公尺)的人臉偵測結果。

再來是對移動中的人做人臉偵測的實驗，其結果如圖 5.2 所示，在此結果中可以發現，所提出的系統也有追蹤的效果。在偵測過程中，由於左邊的木製書櫃非常接近膚色，在膚色偵測階段，也被偵測為人臉候選區，但是沒有抽出人臉特徵，所以，沒有被驗證為人臉，但是系統處理時間增加不少，在實驗過程中發現，其 FPS 會下降。

在對不同亮度下的人做人臉偵測實驗，其結果如圖 5.3 所示，在圖 5.3 中影像左邊偏暗，影像右邊有一盞檯燈，當人臉靠近時會因為檯燈的照射而變亮，實驗結果得到本系統在光線變化的情形，也可以偵測到人臉。



(a)

(b)



(c)

(d)



(e)

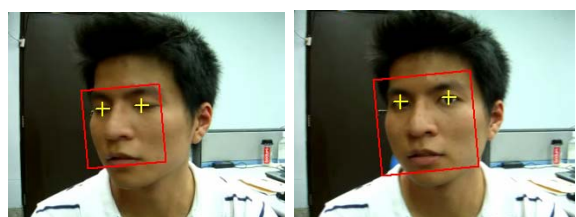
(f)

圖5.3 不同亮度下的人做人臉偵測結果。

在對不同角度拍攝的人臉做人臉偵測實驗中，因為本系統是根據臉部膚色和特徵來偵測和驗證人臉，可以偵測的角度在 45 度到 -45 度間，在這個角度範圍內，兩眼和頭髮的特徵都可以偵測到，當大於這個角度時，眼睛特徵就會偵測不到，所以只討論 45 度到 -45 度的人臉偵測。對不同角度拍攝的人臉做人臉偵測，其結果如圖 5.4 所示，頭由左轉到右的偵測結果，最左邊角度 45 度，最右邊角度也是 45 度

在對多人做人臉偵測實驗時，多人影像也是可以偵測出來，其結果如圖 5.5 所示，在此圖中發現，有些人臉無法偵測出來，如圖 5.5(b)所示，右上方的人臉傾斜，嘴唇張開，表情高興，所以無法偵測到。另外，在圖 5.5(c)中，左下方的人臉被上方的手影響到，發生重疊現象，導致也偵測不到人臉。所以，未來要針對表情和重疊的人臉問題，進行研究和改善。

(五) 動態的人臉辨識



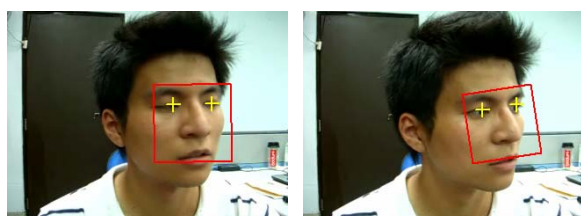
(a)

(b)



(c)

(d)



(e)

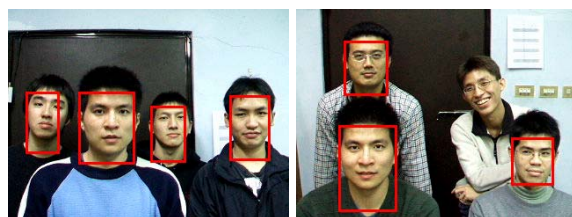
(f)

圖5.4 不同角度的人臉做人臉偵測結果。

(a)到(f)是臉部由-45度轉到45度。

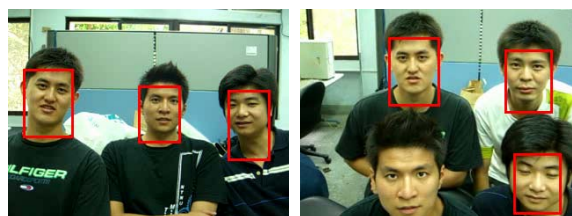
在動態人臉辨識實驗中，是將動態人臉偵測的結果儲存下來當作人臉資料庫，本實驗資料庫包含了30位實驗者，每位實驗者有50張圖片，總共1500張影像，其大小為112x92。為了測試本系統在不同的狀況下的動態辨識效果，做了以下實驗：(1)針對正面的人臉和不同拍攝距離來做人臉辨識實驗；(2)針對不同角度拍攝的人臉做人臉辨識實驗(人臉左右偏轉正負45度間)。在動態辨識實驗中，依據以上不同的實驗條件，本文的人臉資料庫是可以被更新的，以上兩種的實驗，都會有不同的人臉資料庫來當作訓練影像。

在對正面的人臉和不同拍攝距離來做人臉辨識實驗中，有三個不同的距離來拍攝人臉影像，其拍攝距離分別是1公尺、2公尺和3公尺，其辨識結果如表5.2所示，最好的辨識率是在拍攝距離1公尺時，拍攝距離越遠，辨識效果會下降，這是因為



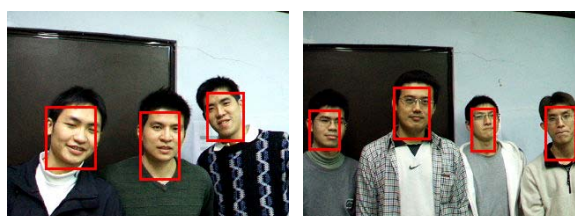
(a)

(b)



(c)

(d)



(e)

(f)

圖5.5 多人的臉偵測結果。

距離越遠，人臉的特徵愈不明顯，所以，在人臉動態辨識中，人臉與攝影機的距離不可以太遠。

在針對不同角度拍攝的人臉做人臉辨識實驗中(人臉左右偏轉正負45度間)，本實驗也是把動態人臉偵測的結果儲存下來，當作本實驗的訓練人臉影像其辨識結果如表5.3所示，最好的辨識率也是在近

表5.2 不同拍攝距離的人臉辨識結果

距離	1公尺	2公尺	3公尺
張數	1500	1500	1500
正確	1415	1305	1250
錯誤	85	195	250
正確率	94.3%	87.0%	83.3%
FPS	5	5	5

表 5.3 不同角度的人臉做人臉辨識結果

距離	1 公尺	2 公尺	3 公尺
張數	1500	1500	1500
正確	1275	1155	1050
錯誤	225	345	450
正確率	85.0%	77.0%	70.0 %
FPS	5	5	5

距離時，距離愈遠，辨識率會下降。

另外，從表 5.2 和表 5.3 可以知道，人臉有偏轉時，辨識率也會下降，所以，要應用本系統於實際大樓門禁監控時，要將攝影機裝設於適當距離和方向，以便取得比較正面而且適當大小的影像，這樣的人臉偵測和辨識結果才會提高。

六、結論

本文提出一個應用於智慧型大樓門禁監控的雛形系統，所使用的方法包括先統計在不同亮度下的膚色範圍，進而找出膚色的區域，當膚色的區域大時，用自適性的閾值方法抽出頭髮，嘴唇和眼睛特徵，最後，再用幾何判斷的方法驗證真正的人臉，當膚色的區域小時，利用變異度和頭髮特徵來驗證出人臉。人臉偵測方法簡單又有效，人臉膚色和特徵，在不同度亮度下，大都可以偵測出來，在最好的實驗條件下，動態人臉偵測率可達 97%。另外，在動態人臉辨識中，在近距離以及比較正面取像情形下，最好的辨識率 94.3%。

本篇主要是利用人臉偵測和辨識技術，應用於門禁監控系統，來解決傳統門禁的缺點。此系統只是一個雛形，未來的工作如下：(1)將系統架設於實際環境中，(2)取得更多和更自然的人臉來做實驗，(3)最佳化動態人臉偵測程式，(4)改進動

態人臉辨識的方法，(5)研究各種人臉表情的偵測與辨識，(6)研究解決人臉重疊的問題。

誌謝

This work was supported by the Ministry of Economic, R.O.C., under Grants MOEA-96-EC-17-A-02-S1-032 and National Science Council, R.O.C., under Grants NSC 96-2221-E-133-001-.

參考文獻

- [1] M.H. Yang, D.J. Kriegman, and N. Ahuja, "Detecting faces in images: a survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 1, pp. 34-58, 2002.
- [2] Y. Nara, J. M. Yang, and Y. Suematsu, "Face detection using the shape of face with both color and edge," *Cybernetics and Intelligent Systems*, 2004 IEEE Conference on, Vol.1, pp. 147-152, 2004.
- [3] R.L. Hsu, M. Abdel-Mottaleb, and A.K. Jain, "Face detection in color images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 24, No. 5, pp. 696-706, 2002.
- [4] C.H. Lin and K C. Fan, "Triangle-based approach to the detection of human face," *Pattern Recognition*, Vol. 34, pp. 1271-1284, 2001.
- [5] H. Wu, Q. Chen, and M. Yachida, "Face detection from color images using a fuzzy pattern matching method," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.21, No.6, 1999.
- [6] K. Anderson and P.W. McOwan, "A real-time automated system for the recognition of human facial expressions," *Systems, Man and Cybernetics, Part B*, *IEEE Transactions on*, Vol. 36, pp. 96-105, 2006.
- [7] H.A. Rowley, S. Baluja, and T. Kanade, "Rotation invariant neural network-based face detection," in *Proceedings of Computer Vision and Pattern Recognition*,

- pp. 38-44, 1998.
- [8] C.A. Waring, X.W. Liu, "Face detection using spectral histograms and SVMs," *Systems, Man and Cybernetics, Part B, IEEE Transactions*, Vol. 35, Issue 3, pp.467-476, 2005.
- [9] A.N. Rajagopalan, R. Chellappa, and N.T. Koterba, "Background learning for robust face recognition with PCA in the presence of clutter," *IEEE Transactions on Image Processing*, Vol. 14, Issue 6, pp.832-843, 2005.
- [10] R. Chellappa, C.L. Wilson, and S. Sirohey, "Human and machine recognition of faces: a Survey," *Proceedings of the IEEE*, Vol. 83, Issue 5, pp. 705-741, 1995.
- [11] M.A. Turk, A.P. Pentland, "Face recognition using eigenfaces," in *Proceedings of Computer Vision and Pattern Recognition*, pp. 586-591, 1991.
- [12] P.N. Belhumeur, J.P. Hespanha and D.J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine*, Vol. 19, No. 7, pp. 711-720, 1997.
- [13] J.P. Ye and Qi Li, "A two-stage linear discriminant analysis via QR-decomposition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 6, pp. 929-941, 2005.
- [14] B.L. Zhang, H.H. Zhang, and S.S. Ge, "Face recognition by applying wavelet subband representation and kernel associative memory," *IEEE Transactions on Neural Networks*, Vol. 15, Issue 1, pp.166-177, 2004.
- [15] S. Krishna and S. Panchanathan, "A methodology for improving recognition rate of linear discriminant analysis in video-based face recognition using support vector machines," *IEEE International Conference on Multimedia and Expo*, pp. 1054-1057, 2005.
- [16] Y.X. Hu, D.L. Jiang, S.C. Yan, L. Zhang, and H.J. Zhang; "Automatic 3D reconstruction for face recognition," *Automatic Face and Gesture Recognition*, 2004. *Proceedings. Sixth IEEE International Conference on*, pp. 843- 848, 2004.
- [17] B. Heisele and T. Koshizen, "Components for face recognition," *Automatic Face and Gesture Recognition*, 2004. *Proceedings. Sixth IEEE International Conference on*, pp. 153-158, 2004.
- [18] F. Zuo, and P.H.N. de With, "Real-time embedded face recognition for smart home," *IEEE Transactions on Consumer Electronics*, Vol. 51, No. 1, pp.183-190, 2005.
- [19] L.N. Chu, "A novel partitioned gradient Fisherface algorithm for robust face recognition," *National Tsing Hua University, Department of Computer Science, Master thesis*, July 2004.
- [20] K.W. Wong, K.M. Lam, and W.C. Siu, "A robust scheme for live detection of human face in color images," *Signal Processing: Image Communication*, Vol. 18, No.2, pp.103-114, 2003.
- [21] 人體計測資料庫，
<http://www.iosh.gov.tw/ergo.htm>
- [22] R.M. Rao and A.S. Bopardikar, *Wavelet transforms: introduction to theory & applications*. Addison Wesley Longman, July 1998.
- [23] C.C. Wong and H.R. Lai, "A new grey relational measurement," *The journal of Grey System*, Vol. 12, No. 4, pp.341-346, 2000.
- [24] K.C. Chang and M.F. Yeh, "Grey relational analysis based approach for data clustering," *Vision, Image and Signal Processing, IEE Processings*, Vol. 152, No. 2, pp. 165-172, 2005.
- [25] H.J. Kim, S.H. Lee and N.I. Cho, "Rotation-invariant face detection using angular projections," *Electronics Letters*, Vol. 40 No. 12, pp.726-727, 2004.