

即時人臉偵測與辨識

李建興 林應璞 游凱倫

國立成功大學 系統及船舶機電工程學系

摘 要

本文以 AT&T Laboratories Cambridge, Georgia Institute of Technology, California Institute of Technology 與作者自建人臉影像資料庫為對象，進行即時人臉偵測與辨識，並討論訓練樣本數對人臉辨識之影響。首先，以 Matlab/Simulink 進行即時人臉偵測，用以降低人臉偵測所需時間，再使用 Haar 小波、主成份分析法與改良式主成份分析法擷取人臉特徵，最後以歐氏距離決策法、最接近特徵線決策法與線性鑑別式分析法作為人臉辨識之決策法則。

關鍵詞：主成份分析法，改良式主成份分析法，線性鑑別式分析，最接近特徵線法。

REAL TIME FACE DETECTION AND RECOGNITION

Chien-Hsing Lee Ying-Pu Lin Kai-Lun You

*Department of Systems and Naval Mechatronic Engineering
National Cheng Kung University
Tainan, Taiwan 701, R.O.C.*

Key Words: principal component analysis, improved principal component analysis, linear discriminant analysis, nearest feature line method.

ABSTRACT

This paper presents a real-time face detection and recognition system using the database of AT&T Laboratories Cambridge, Georgia Institute of Technology, California Institute of Technology and a database created by the authors. In addition, the effect of training samples on face recognition is investigated in this paper. The face detection algorithm was implemented in a Matlab/Simulink environment to reduce the running time to detect a face. Then, the facial feature extraction part uses Haar wavelets, principal component analysis and improved principal component analysis. Finally, the facial recognition rule is based on Euclidean distance, the nearest feature line method and linear discriminant analysis.

一、前 言

生物辨識包括人臉辨識、聲紋辨識、虹膜比對、指紋辨識等[1-4]，其中以人臉辨識之非接觸性較廣為所接受。在實際上，人臉辨識易受光線、亮度、臉部表情變化等因素影響而增加辨識困難度。為強健辨識系統，一般皆以大

量具代表性影像來訓練，而將增加運算量與辨識時間。

人臉辨識研究始於 1991 年之特徵臉演算法，而 Turk 與 Pentland 以特徵臉演算法[5, 6]為基礎延伸出主成份分析法 (PCA)。但因此法所擷取之特徵屬人臉最佳特徵，並不是最佳人臉分類，在整體人臉辨識表現上是有限的。於 2001 年，Linda [7]等人基於人臉各類別間之差異性且於投

影空間中可大幅拉開各類別間距，而提出線性鑑別式分析法[8, 9]並應用於人臉辨識，以大幅提升整體辨識率。但線性鑑別式分析法需要足夠樣本才可運作，否則會因樣本不足，而產生不具反矩陣的奇異矩陣。Yusuke [10]等人以主成份分析法為基礎提出改良式主成份分析法 (IPCA)，其辨識率表現上較優於主成份分析法，並可有效縮短辨識所需時間。表一所示為相關文獻使用之人臉偵測與特徵擷取方法的比較[5, 10-14]。

本文以 Haar 小波[15, 16]、主成份分析法、改良式主成份分析法擷取人臉特徵，並配合歐氏距離[17]、線性鑑別式分析法與最接近特徵線法[18]組合出 12 種人臉辨識方式，再以 AT&T [19]、GIT [20]與作者自建資料庫評估人臉辨識率與辨識時間，最後以 CIT 資料庫[21]來驗證。在人臉偵測部份，吾人以 Matlab/Simulink 進行即時人臉偵測，用以有效減少非即時人臉偵測所需之時間。在人臉辨識部份，吾人著重在探討辨識率與訓練樣本數之關連性。圖 1 所示為本文所提人臉辨識流程圖，而圖 2 則為即時人臉偵測流程圖。

二、即時人臉偵測步驟

1. 膚色分割 (Color Segmentation)

人臉膚色常被視為人臉偵測的主要特徵之一[22-25]，因一般視訊裝置所擷取影像皆由光的三原色 RGB (紅色、綠色、藍色) 所組成，但由 RGB 所構成之色彩空間易受光線影響而產生大幅度的波動，使其不適合進行膚色與非膚色之色彩分割。為避免光的影響，於膚色分割前大都將 RGB 色彩空間轉換至其他如 YIQ、YUV、HSV 與 YCbCr 等色彩空間，其中 YCbCr 對膚色的分割已於文獻[23]被證明其效果不錯。本文亦以此色彩空間來進行膚色分割，其 Simulink 模型如圖 3 所示。有關 YCbCr 及 RGB 色彩空間的轉換，其公式為[26]

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.533 & 24.966 \\ -37.797 & -74.203 & 112 \\ 112 & -93.786 & -84.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

式 (1) 主要是依人類眼睛對 RGB 三原色不同敏感度而來。另外，吾人採用[23]所提供 YCbCr 色彩空間膚色之判斷式，再根據膚色區域的邊界來判定，且由於膚色分佈區域在 Y 值大於或小於 128 時有顯著的不同，而以 Y = 128 為門檻值分別做不同判斷與處理。

2. 雜訊移除 (Noise Filtering)

針對膚色分割後的區域，吾人將非膚色區轉成二值影像，發現在人臉區塊附近或多或少會出現類似膚色之雜訊，可加以移除來降低偵測失敗率。雜訊移除方法主要以形態學 (Morphology) 中的二值斷開運算 (Binary Opening

表一 相關文獻回顧表

相關文獻	偵測方法				特徵擷取方法		
	Color	Edge	Motion	Real time	Haar	PCA	IPCA
Turk [5]			x			x	
Yusuke [10]	x			x			
Yang [11]							x
Sirovich [12]						x	
Huang [13]		x					
王科翔 [14]	x				x	x	

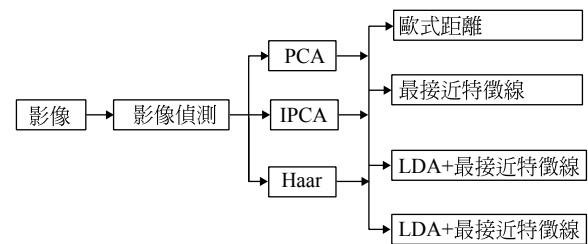


圖 1 人臉辨識流程圖

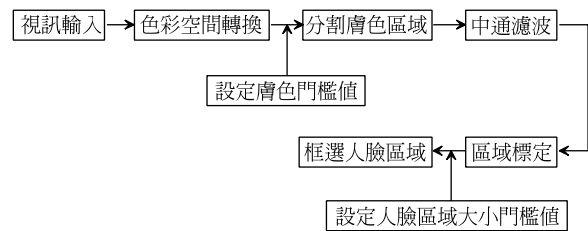


圖 2 即時人臉偵測流程

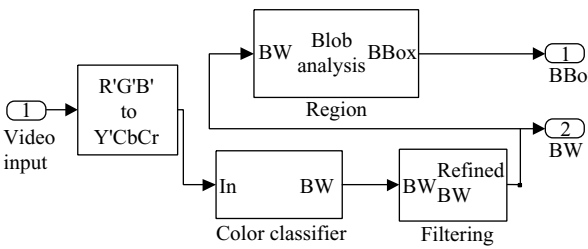


圖 3 膚色分割 Simulink 模型圖

Operation) 包括侵蝕 (Erosion) 和膨脹 (Dilation)，即圖 3 中的 Filtering 方塊。由膚色分割後的影像，可發現人臉周圍因在複雜環境中會有類似膚色的雜訊點，如圖 4(a)所示，此雜訊區域有大有小，而需運用形態學將雜訊移除。首先，以侵蝕來消除雜訊，其結果如圖 4(b)所示，再使用膨脹來擴大膚色範圍區域，其目的是希望留下相同結構元素尺寸之幾何特性，其餘則歸屬於雜訊部分，可再將它移除，其結果如圖 4(c)所示。



(a) 膚色範圍二值化影像 (b) 侵蝕後影像 (c) 膨脹後影像

圖 4 人臉偵測之雜訊移除影像

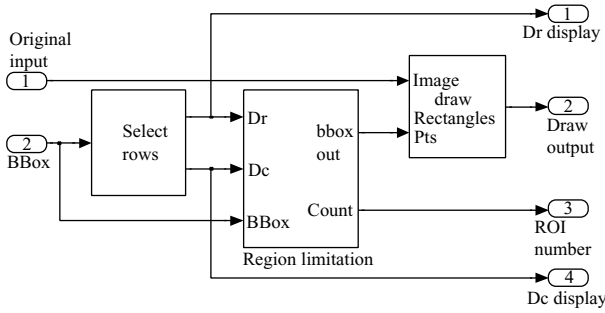


圖 5 框選目標物 Simulink 模型圖

3. 區域標定 (Region Marker)

經膚色分割與雜訊移除後，吾人可得人臉膚色區塊，再應用連接元區域標定程序 (Connected Component Labeling)，將影像中屬於同一物件之相同像素找出，以便獲得該物件之面積、高度、寬度等資訊，即圖 3 中的 Blob Analysis 方塊。由於某些雜訊點經斷開運算後，可能還是無法消除，吾人便可利用簡單的比例或面積大小之關係，即可將所感興趣的人臉區塊 (Region of Interested, ROI) 標定出來。

4. 框選目標物 (Object Framing)

經前述步驟後，吾人可得僅包含目標物的二值化影像，其中像素值為 0 是背景，而像素值為 1 即是目標物。框選目標物之目的是要取得目標物的上、下、左、右四個邊界，再利用最小矩形把目標物框選，並將目標物座標資訊回傳至程式中，以進行後續辨識步驟。圖 5 所示為框選目標物區域之 Simulink 模型。

三、人臉特徵擷取演算法

1. Haar 小波轉換法

本文以 Haar 二維小波轉換將原始人臉影像先經高通及低通濾波器，從行方向分解成高頻及低頻兩部份，再將各自列方向分解成高頻及低頻兩部份，如圖 6 所示。其中 $h_0(x)$ 與 $h_1(x)$ 分別為高通與低通濾波器、 $\downarrow 2$ 代表次取樣 (down sampling)、 $f_1^0(x, y)$ 為一張原始訓練影像如人臉影像，經二維 Haar 小波轉換後，原始影像會被分解為 $f_1^1(x, y)$ 、 $f_2^1(x, y)$ 、 $f_3^1(x, y)$ 、 $f_4^1(x, y)$ 四個子影像，其中

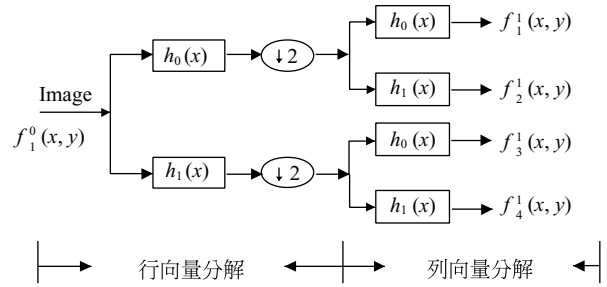


圖 6 二維小波轉換示意圖

$f_4^1(x, y)$ 低頻部份除可大幅降低影像維度外，也保留大部分影像資訊。例如：一張原始影像，經行及列向量分解後為一次小波分解；而維度方面，每執行一次小波轉換，只取出低頻部份其將會變成原本的 1/4 影像，由於選擇二維小波轉換所獲得的人臉影像資料將變為原本人臉影像的 1/16 之大小維度，如此可降低鑰匙影像維度且同時在低頻 $f_4^1(x, y)$ 中仍保留重要的影像資料，此乃小波轉換最主要的貢獻。

2. 主成份分析法

此法為電腦視覺演進中較早期的方法，主要是希望找出影像基底用以線性組合其他影像來構成影像特徵空間，再經特徵參數所排序的特徵向量堆疊出影像。如依大至小的特徵值排序，人臉的主要特徵與形狀等訊息均可由前方的特徵向量來表示，而後方的特徵向量則只存在細微的人臉訊息。所以，一般均取前幾個特徵向量來組成人臉影像的特徵空間。

假設有 N 張影像，將其轉換為原始特徵樣本 $\{x_1, x_2, \dots, x_k, \dots, x_N\}$ ，其所有原始樣本相對於平均樣本 \bar{x} 的分散程度可以全域散佈矩陣 (total scatter matrix) 來表示，即轉換後的特徵參數平均為 $\bar{z} = W^T \bar{x}$ ， W 為原始樣本轉換前後的全域散佈轉換矩陣。假設全域散佈矩陣 Σ_{T_x} 於轉換後可表示為

$$\Sigma_{T_y} = W^T \Sigma_{T_x} W \quad (2)$$

其中， $\Sigma_{T_x} = \sum_{k=1}^N (x_k - \bar{x})(x_k - \bar{x})^T$ 。用線性代數跡 (trace) 的觀念來表示 Σ_{T_y} 內部分佈情形，可得最佳全域散佈轉換矩陣 W_{opt} 為

$$W_{opt} = \arg \max_W \text{tr}(W^T \Sigma_{T_x} W) \quad (3)$$

$$\Sigma_{T_x} W = \lambda W_x \quad (4)$$

其中， λ 為原始樣本矩陣之特徵值序列， W_x 為原始樣本所組成之全域散佈矩陣， $\arg \max$ 為透過全域散佈轉換矩陣所得之最大數值。由此可知，取式 (4) 特徵向量所組成之矩陣即為欲求之基底。

3. 改良式主成份分析

此法主要是在求特徵向量中之全域散佈矩陣，用以加大轉換後之特徵參數與其平均值間之散佈程度。但在吾人討論之元素中，屬於人臉特徵向量元素是很少的，而非人臉特徵向量元素卻是多的，如雜訊。換言之，人臉特徵向量之全域散佈矩陣是小的，而非人臉全域散佈矩陣卻是大的。所以，主成份分析法需有足夠的訓練樣本用以避免在整個特徵擷取中，儲存太多非特徵之元素。而改良式主成份分析法雖取出較少特徵向量，但卻較具代表性。假設 x 為原始特徵樣本，因一般人臉資料庫中至少有上百張影像， x 會是一大型矩陣，而矩陣元素受雜訊的影響會隨矩陣大小而改變，即矩陣越大，雜訊的影響也越大。所以，吾人可藉由式 (5) 來降低矩陣維度，如下所示：

$$u = x / \bar{x} \quad (5)$$

其中， \bar{x} 為平均樣本， u 為降低維度後的原始特徵樣本。同樣地，假設有 N 張訓練用影像，其原始特徵參數為 $\{x_1, x_2, \dots, x_k, \dots, x_N\}$ 且維度為 n ，而轉換後之特徵參數為 $z = \{z_1, z_2, \dots, z_k, \dots, z_m\}$ 且維度為 m ，可表示為

$$z = W^T u \quad (6)$$

其中， W 為原始樣本轉換後的全域散佈矩陣，而 W^T 為 W 之轉置矩陣。相較於主成份分析法主要是要求最大全域散佈矩陣，而改良式主成份分析法是要求出 z 的最小全域散佈矩陣之轉換矩陣。假設全域散佈矩陣轉換前可表示為

$$\Sigma_{T_x} = W^T \Sigma_{T_x} W \quad (7)$$

$$\Sigma_{T_x} = \frac{1}{n} \sum_{k=1}^N (u - \bar{u})(u - \bar{u})^T = \sum_{x \in X} \left(\frac{x}{\bar{x}} - 1 \right) \left(\frac{x}{\bar{x}} - 1 \right)^T \quad (8)$$

其中， \bar{u} 為 u 之平均值。藉由線性代數跡的觀念來表示 Σ_{T_x} 內部分佈情形，可得最佳全域散佈矩陣 W_{opt}

$$W_{opt} = \arg \max_W \text{tr}(W^T \Sigma_{T_x} W) \quad (9)$$

$$W^T \Sigma_{T_x} W = \lambda \quad (10)$$

所以，由改良式主成份分析之轉換結果，吾人可於訓練樣本中分析出一特徵空間，再將每張人臉影像投影到此空間就可得一組新的特徵參數，而這些新特徵參數的散佈程度會最大，即會是最具代表性的特徵。

四、人臉辨識決策法則

1. 線性鑑別式分析法

線性鑑別式分析之觀念是先建構特徵空間，並將影像訓練資料庫中所有特徵樣本投影於其中，再將每一類別內

的影像樣本投影點之距離縮短，以及伸長每一類別間的影像樣本投影點之距離，以減少在比對時的辨識誤差。意即可視每一類別的影像樣本為一群組，用以拉開相異群組間之差異度並拉近相同群組的距離[7-9]。

假設訓練資料庫之影像樣本分成 d 類不同人臉影像且每類中均有 I 張影像，則原始影像資料庫中，第 d 類第 i 張的原始樣本為 x_i^d ，此類中所有原始樣本的平均值以 \bar{x}^d 表示。如以 \bar{x} 代表所有原始特徵樣本的平均值，則類別內原始樣本間的散佈程度為

$$\Sigma_{Q_x} = \sum_{d=1}^D \sum_{i=1}^I \left(x_i^d - \bar{x}^d \right) \left(x_i^d - \bar{x}^d \right)^T \quad (11)$$

而其類別間的散佈程度為

$$\Sigma_{B_x} = \sum_{d=1}^D \left(\bar{x}^d - \bar{x} \right) \left(\bar{x}^d - \bar{x} \right)^T \quad (12)$$

當原始樣本 x_i^d 轉換後，可以特徵樣本 z_i^d 來表示如下：

$$z_i^d = Q^T x_i^d \quad (13)$$

再透過轉換可得類別內矩陣 Σ_{Q_x} 與類別間矩陣 Σ_{B_x} 分別為

$$\Sigma_{Q_x} = Q^T \Sigma_{T_x} Q \quad (14)$$

$$\Sigma_{B_x} = Q^T \Sigma_{B_x} Q \quad (15)$$

如利用費雪條件式 (Fisher Criterion)，吾人可得 Q_{opt} 為

$$Q_{opt} = \arg \max_Q \frac{\text{Tr}(Q^T \Sigma_{B_x} Q)}{\text{Tr}(Q^T \Sigma_{Q_x} Q)} \quad (16)$$

由式(15)之運算可得矩陣中 $(\Sigma_{Q_x})^{-1}$ 與 Σ_{B_x} 兩矩陣乘積，再另求其特徵向量矩陣即為所求；因特徵空間 Q (由特徵向量組成之矩陣) 的秩 (rank) 為 $C-1$ 。但此法仍有一大缺失須克服，即需要有足夠的訓練樣本方可達最佳分類。換言之，在類別內散佈矩陣之運算，因訓練樣本不足會產生不具反矩陣的奇異矩陣，以致無法得到轉換矩陣。常見的解決方式是設法取得更多訓練樣本，或先將原始訓練樣本的維度降低，再做線性鑑別式分析。為了避免產生不具反矩陣之奇異矩陣，可移動原始影像樣本像素的方向，以增加訓練影像樣本空間或使用常見的 Waveletface [15]與 Fisherface [8]等方法。

2. 歐氏距離法

歐氏距離為目前較普遍的分類法，其乃是將測試影像樣本用來比對資料庫中每一訓練影像樣本，以找出最短距離即為吻合之影像，等同於計算測試影像的向量與訓練資料庫中每一向量間的最短距離[17]。若 y 為測試影像的特

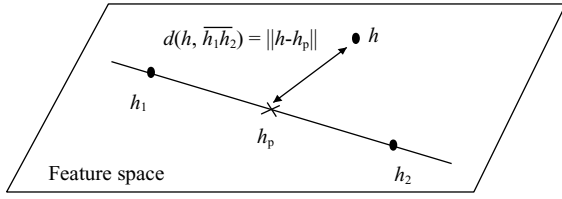


圖 7 特徵線示意圖

徵樣本，資料庫中共有 C 類，每類有 N 張影像且已訓練好，其特徵參數為 $\{z_i^1, z_i^2, \dots, z_i^k\}$ ， $1 \leq k \leq C$ 且 $1 \leq i \leq N$ ，則測試影像 y 到資料庫影像 z_i^k 的歐氏距離為

$$d(y, y_i^k) = \|y - z_i^k\| \quad (17)$$

而測試影像的特徵參數與類別 k 的最短距離為

$$d(y, k) = \min(d(y, z_1^k), d(y, z_2^k), \dots, d(y, z_N^k)) \quad (18)$$

所以，測試影像 y 的身分為

$$ID_y = \arg \min_k d(y, k) \quad (19)$$

其中， ID_y 為系統判斷出的測試影像中被拍攝者的身分。

3. 最接近特徵線法

當特徵樣本的維度為 m 時，可將特徵樣本 m 度空間中視為一個點，其類別內任意兩點可形成一條特徵線 (feature line) [18]。假設每類別中有 N 個特徵樣本，即意謂能形成 $I(I-1)/2$ 條特徵線。當比對測試影像與影像資料庫時，需找出測試影像之特徵點與每一類別中每一條特徵線最接近的距離，方可確認屬於影像資料庫中哪一幅人臉影像。

假設一特徵線 h_1h_2 由 h_1 與 h_2 特徵點形成，測試影像特徵點為 h 如圖 7 所示，其在 h_1h_2 上的特徵點 h_p 為

$$h_p = h_1 + \gamma(h_2 - h_1) \quad (20)$$

因 h_1h_2 與 hh_p 相互垂直，其內積 (inner product) 為 0，即

$$\overrightarrow{hh_p} \cdot \overrightarrow{h_1h_2} = (h_p - h) \cdot (h_2 - h_1) = 0 \quad (21)$$

而長度比例 γ 為

$$\gamma = \frac{(h - h_1) \cdot (h_2 - h_1)}{(h_2 - h_1) \cdot (h_2 - h_1)} \quad (22)$$

可由此得到特徵線上特徵點 h_p ，其距離位於測試樣本的特徵點到特徵線之間，可以歐氏距離計算即

$$d(h, h_1h_2) = \|h - h_p\| \quad (23)$$

注意：訓練影像中所產生的拍攝角度可作為建構影像資料庫時之依據，運用此法將可減少直接比對特徵點的誤差，再線性組合出更接近測試影像，以提高辨識準確率。

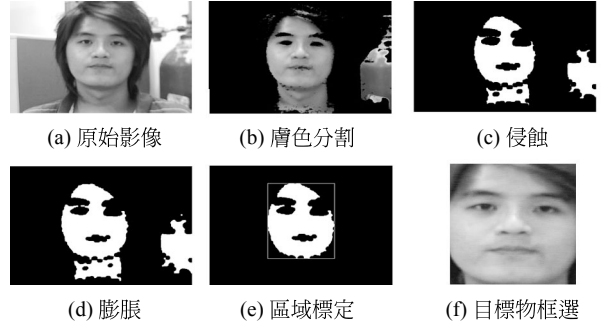


圖 8 人臉偵測模擬結果

五、模擬結果

1. 即時人臉偵測部份

本文以 Matlab/Simulink 程式語言中的 Video and Image Processing Blockset，以及 Image Acquisition Toolbox 來實現即時人臉偵測。首先，由視訊攝影機連續擷取 100 張影像，再進行膚色偵測，並比較即時與非即時處理之偵測時間和偵測率。即時偵測時間為 9.1 秒，而非即時偵測需要 29.5 秒；前者僅需後者的 30% 時間，即可完成所有偵測程序，並提供足夠辨識的人臉區塊資訊；且即時偵測的偵測率為 97%，較非即時偵測高出 7% 左右。所以，藉由 Simulink 即時人臉偵測不僅可有效減少偵測所需之時間並可達較高偵測率。

於非即時人臉偵測中，藉由膚色分割來獲取人臉影像之膚色區域如圖 8(a)與(b)所示，並轉換為較易分辨之二值影像。因經膚色分割後之影像，會發現在人臉區域附近可能有些許之類似膚色之雜訊點，須經由二值運算來消除雜訊點以降低偵測失敗率，其結果分別如圖 8(c)與(d)所示。再者，經膚色分割和雜訊移除後，可應用連接元區域標定程序來獲得膚色所在的區域如圖 8(e)所示，最後將區域標定之區塊從影像中取出如圖 8(f)，即可進行特徵擷取及辨識程序。

於即時人臉偵測中，同樣使用膚色分割、雜訊移除、區域標定、目標物框選等影像前處理，但不同的是即時人臉偵測需要設定目標物大小門檻值。當視訊攝影機擷取動態影像時，會依門檻值大小來判別是否將所擷取之影像存入記憶體中，以作為辨識時所需的資訊，再將此影像以連接元區域標定程序將膚色區塊標出，來獲得所要之人臉區塊。吾人根據資料庫影像維度大小來設定門檻值，若門檻值設太小，其所擷取影像的維度和解析度會不足而降低辨識率。圖 9 所示為門檻值設定之 Simulink 模型。如圖 10(a) 中所示，左邊同學並未符合所設定之門檻值，故僅有符合門檻值的右邊同學人臉區域會被框出，再將人臉區域相關的座標資訊存入記憶體。而圖 10(b)中，當左邊同學逐漸靠近視訊攝影機且符合門檻值時，系統將框選出人臉區域，並記錄兩個人臉區塊的座標資訊。當系統偵測到有框選出

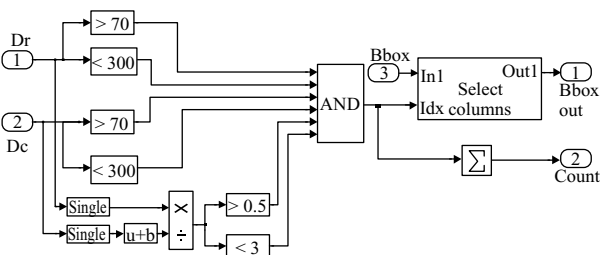


圖 9 門檻值設定之 Simulink 模型圖

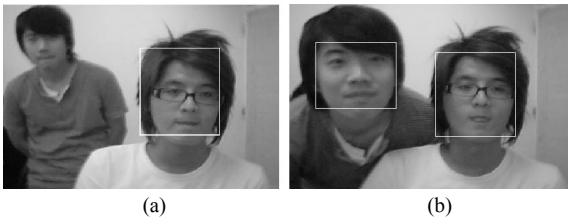


圖 10 人臉偵測模擬圖

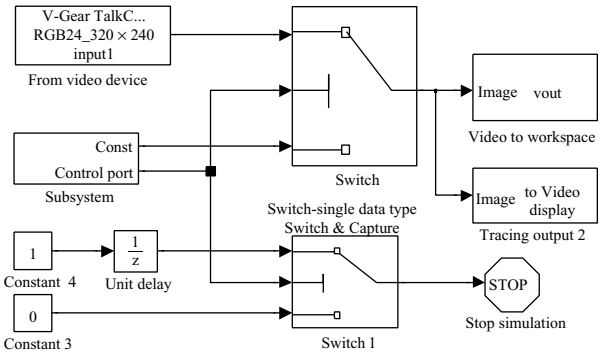


圖 11 Switch & Delay Simulink 模型圖

符合門檻值的人臉區域時，可藉由 Simulink 中 Switch 功能，開啓紀錄動態影像功能，將所擷取的人臉區域資訊紀錄和傳回 Matlab 程式中，並在 Simulink 中 Unit Delay 功能設定延遲時間，以擷取足夠的影像作為辨識用途。當到達指定延遲時間時，則自動停止程式運作，並切換至 Matlab 程式，準備進行辨識動作。圖 11 所示為 Switch & Delay Simulink 模型圖。

2. 人臉辨識部份

本文比較 AT&T、GIT 與作者自建資料庫 (如表二所示) 於人臉辨識之差異。AT&T 資料庫共 40 人，每人 10 張影像包含睜眼、閉眼、微笑、扁嘴等不同表情，共 400 張。其拍攝背景固定為黑色，且大部分為正面影像並允許影像有小角度之偏斜。GIT 資料庫共 50 人，每人 15 張包含不同表情之正面及傾斜影像，共 750 張。但拍攝環境為彩色背景且皆在不同光線狀況下所拍攝。作者自建資料庫

表二 四種人臉資料庫其中一人之不同表情影像

人臉資料庫	其中一人之不同表情影像				
AT&T					
GIT					
作者自建					
CIT					

表三 十二種人臉辨識之方式

決策法	特徵擷取演算法		
	Haar 小波轉換	主成份 分析法	改良式主成份 分析法
歐氏距離法	1	2	3
最接近特徵線法	4	5	6
線性鑑別式分析法結合 歐氏距離法	7	8	9
線性鑑別式分析法結合 最接近特徵線法	10	11	12

共 8 人，每人 20 張包含正面、上、下、左、右以及不同表情和小角度傾斜之影像，共 160 張。而拍攝背景固定為白色，以消除背景對辨識結果之影響。而人臉辨識則由上述特徵擷取演算法與決策法所線性組合出 12 種辨識方式以代號區分如表三所示。

表四 不同人臉資料庫之辨識率比較

辨識方式	辨識率 (%)			辨識時間 (sec)		
	AT&T	GIT	作者自建	AT&T	GIT	作者自建
1	85.50	68.40	83.75	17.15	41.44	31.13
2	77.50	52.00	81.25	24.40	62.17	11.06
3	92.00	78.40	87.50	24.52	78.92	10.04
4	82.00	62.40	78.00	29.45	79.97	24.63
5	70.00	41.00	72.50	31.50	88.65	21.20
6	87.50	70.00	81.25	33.28	95.00	25.55
7	87.00	65.00	81.20	30.38	90.41	10.62
8	70.00	51.00	77.50	31.58	71.38	39.85
9	89.00	74.00	86.25	46.28	111.34	57.66
10	84.00	63.00	80.25	54.10	124.47	31.20
11	72.00	45.00	75.50	55.30	135.51	32.43
12	88.50	72.00	83.00	67.55	150.45	33.12

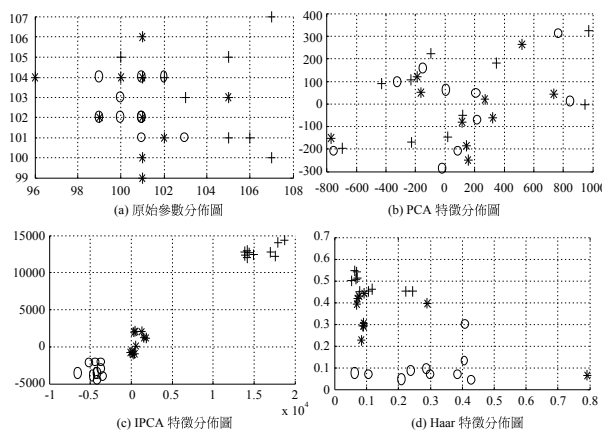


圖 12 原始參數與三種特徵擷取演算法的特徵向量分佈

表四所示為 12 種辨識方式於三種人臉資料庫之辨識結果，其中以辨識方式 3 之結果最佳。以 AT&T 資料庫為例，其辨識率高達 92%。而如以線性鑑別式分析法為決策法則時，因此法之目的是拉開類別間與縮小類別內分散程度，並非適用於任意的特徵擷取演算法，且在使用此法時，會執行分類的動作而耗費較多時間。另外，最接近特徵線和歐式距離法兩原理極為相似，歐式距離法是找出測試點到訓練點的最短距離，而最接近特徵線是找出測試點到兩個訓練點所連的最短距離，於辨識時需要比較訓練點中可連成特徵線的所有情況，故會增加許多時間。

另外，圖 12 所示為三種特徵擷取演算法所得特徵向量之分佈，吾人隨機選取 ORL 人臉資料庫中三類影像，並將此三類影像的特徵向量各別投影到該類特徵空間中，其中 X 軸與 Y 軸分別代表二維影像的 X 方向與 Y 方向之特徵值。首先，將圖 12(a)與圖 12(b)、12(c)、12(d)做比較，吾人發現在原始影像分佈圖中，三類影像的原始參數點

(○、*、+) 彼此間之距離相當接近，然經特徵擷取演算法轉換後，此三類影像類別內的距離縮短而類別間距擴大。一般在辨識過程中，同一類別內影像特徵點之距離要越小，而不同類別間影像特徵點之距離要越大，才能確實減少誤判以提升辨識率。所以，吾人發現相較於圖 12(b)與 12(d)，圖 12(c)中三個類別影像的特徵點已明顯區分出來，且同一類別影像特徵點彼此間的距離也明顯縮小。然而，改良式主成份分析法比另二種特徵擷取演算法較為強健，亦可由表四得知，改良式主成份分析法的辨識結果比另二種特徵擷取演算法較優異。

再者，於人臉辨識時，如選用過多訓練樣本數可能會增加系統的運算量，但過少訓練樣本數卻會因樣本數不足，而無法得到較高辨識率。所以，足夠訓練樣本數不但可減少系統整體運算量，且可維持一定程度的辨識結果。吾人以受外在因素影響較少的 AT&T 資料庫與外在因素較為複雜的 GIT 資料庫為對象，比較不同訓練樣本數對其辨識率之影響，再以作者自建資料庫加以驗證，如表五所示。有關訓練和測試兩個流程的辨識時間，皆以辨識方式 3 來進行。觀察表五中之 AT&T 資料庫結果，當訓練樣本數為 7 時，其最高辨識率為 94.67%；當訓練樣本數為 8 時，辨識率有下滑傾向，此乃因訓練樣本數不足所致，且無法明顯看出辨識率曲線是否有收斂現象。所以，吾人以每人 15 張影像的 GIT 資料庫進行更多訓練樣本數的模擬分析。觀察表五中之 GIT 資料庫結果，當訓練樣本大於 9 張時，其辨識率較佳；然而，當訓練樣本數增加至 13 或 14 時，其最佳辨識率僅維持在 86.00% 左右，此主因乃是 GIT 資料庫中人臉表情變化、臉部傾斜角度與光線影響皆比 AT&T 資料庫中人臉影像差異較大所致。最後，再以作者自建資料庫加以驗證，當訓練樣本數為 17 時，辨識率最佳為 91.68%，且辨識一張影像需 0.597 秒。當訓練樣本數介於 12 至 19 間，辨識率有上下小幅震盪的情形。因此，增加訓練樣本數，並未能顯著提升此資料庫之辨識率。

圖 13 所示為 AT&T、GIT 與作者自建人臉資料庫使用三種特徵擷取法搭配歐氏距離於不同訓練樣本數之辨識結果。由圖得知，改良式主成份分析法不論在何種訓練樣本數，結果都較主成份分析法和 Haar 小波轉換法優異，且辨識曲線的震盪收斂情形皆相似於主成份分析法和 Haar 小波轉換之辨識曲線。所以，本文訓練樣本數之觀點應適用其他以內部特徵擷取演算法為基礎之辨識。且由圖推論，當訓練樣本數為 13 至 17 時，已有相當程度之辨識結果，即使增加訓練樣本數並無法再大幅提升整體辨識率。若要再提升辨識率僅能從特徵擷取演算法和決策法兩方面進行改進，如僅增加訓練樣本數是無法有效改善的。由 AT&T 與 GIT 資料庫模擬得知，改良式主成份分析法應用於人臉辨識有較佳結果，而此兩種資料庫雖包含許多角度和表情之變化，仍有略顯不足之處，且模擬結果和實際應用仍有差異。所以，吾人以自建人臉資料庫進行比較，所得結果與上述兩種資料庫皆十分相似。

表五 不同訓練樣本數目對不同人臉資料庫之辨識率比較

訓練樣本數	辨識率 (%)			辨識時間 (sec)			單張辨識時間 (sec)		
	AT&T	GIT	作者自建	AT&T	GIT	作者人員	AT&T	GIT	作者自建
1	67.77	37.86	76.97	3.563	3.50	7.54	0.010	0.004	0.049
2	80.00	46.77	75.69	6.308	8.35	7.60	0.020	0.012	0.052
3	84.64	48.67	81.62	22.719	14.45	7.64	0.081	0.024	0.056
4	87.50	52.73	78.91	24.573	21.25	7.99	0.102	0.039	0.063
5	92.00	56.60	80.84	24.524	28.88	8.48	0.123	0.058	0.071
6	93.63	67.55	80.36	35.886	37.43	8.47	0.224	0.083	0.076
7	94.67	70.75	80.77	76.346	49.32	8.85	0.633	0.123	0.085
8	94.25	73.14	79.17	88.795	60.01	9.40	1.109	0.171	0.098
9	93.12	72.67	87.50	97.988	80.97	9.65	2.450	0.270	0.109
10		78.40	87.50		78.92	10.04		0.316	0.125
11		82.50	88.89		108.63	10.53		0.543	0.146
12		82.67	85.94		122.05	10.91		0.813	0.171
13		86.00	87.93		150.28	11.42		1.502	0.203
14		86.00	84.08		160.95	12.87		3.218	0.268
15			87.50			13.06			0.326
16			88.38			13.89			0.434
17			91.68			14.32			0.597
18			88.50			15.03			0.939
19			90.00			15.81			1.977

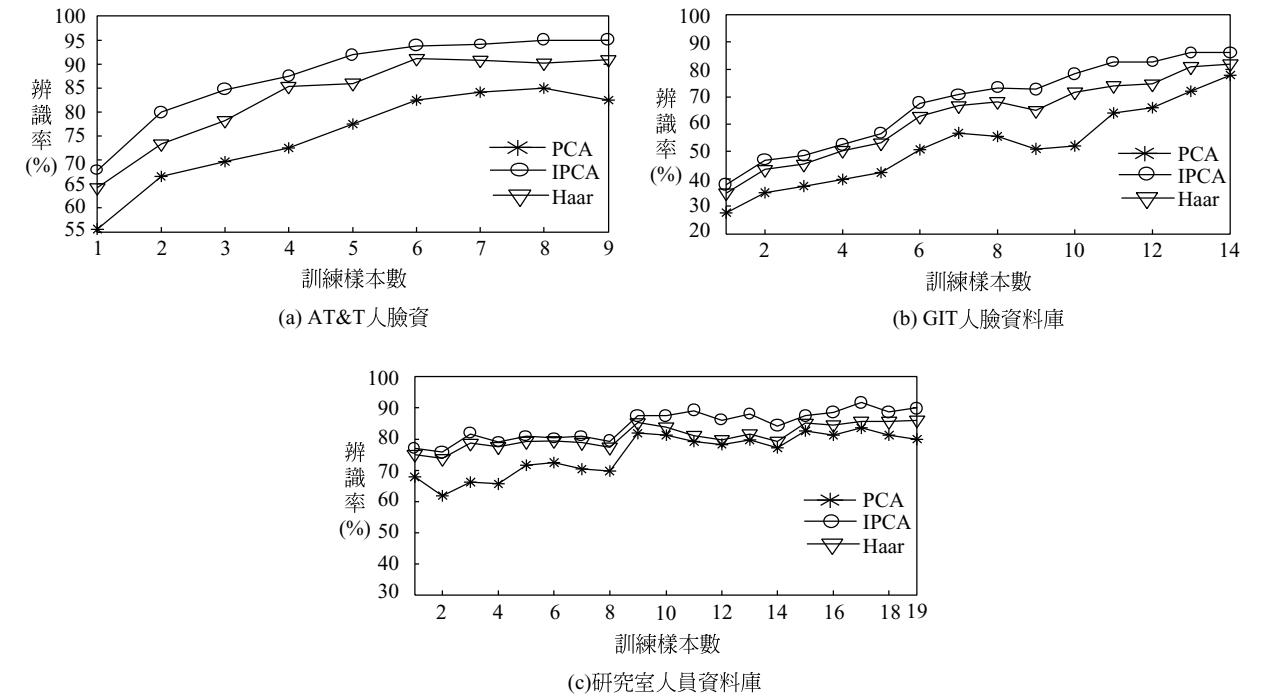


圖 13 三種人臉資料庫使用主成份分析法 (PCA)、改良式主成份分析法 (IPCA) 與 Haar 小波轉換 (Haar) 搭配歐氏距離法對不同訓練樣本數之人臉辨識結果

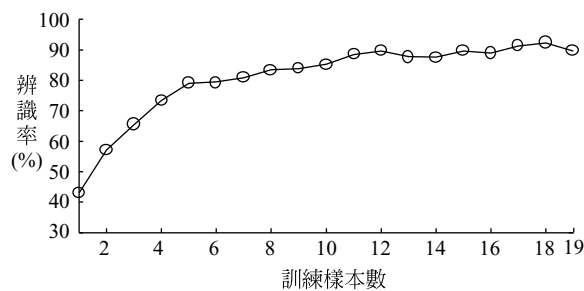


圖 14 CIT 資料庫以改良式主成份分析法搭配歐式距離法對不同訓練樣本數之人臉辨識結果

表六 不同訓練樣本數目對 CIT 人臉資料庫之辨識結果

訓練樣本數	辨識率 (%)	辨識時間 (sec)
1	42.936	2.618
2	57.018	3.508
3	65.325	4.840
4	73.355	6.368
5	78.947	8.001
6	79.323	9.635
7	80.567	12.335
8	83.333	14.091
9	83.731	17.266
10	85.263	19.966
11	88.304	23.263
12	89.474	25.156
13	87.722	29.116
14	87.482	32.358
15	89.474	38.095
16	88.858	41.427
17	91.288	44.418
18	92.105	49.143
19	89.477	53.052

接下來，吾人選用 California Institute of Technology (CIT) 人臉資料庫進行以改良式主成份分析法搭配歐式距離來驗證上述推論，結果如圖 14 與表六所示。CIT 資料庫共 27 人，每人包含 10 到 20 張不同光線、角度、表情背景等變化如表二所示，共 450 張。因每人有 10 至 20 張不同影像，故選取含有完整 20 張影像的 19 人作為驗證的資料庫。由表六得知，當訓練樣本數為 18 時，最高辨識率為 92.105%，而訓練樣本數為 19 時，辨識率略下降 3% 左右。由此得知，如再增加訓練樣本數將無法明顯提升辨識率，反而增加辨識時間。由圖 14 亦發現，訓練樣本數為 13 到 18 時，辨識曲線已趨向平穩，並無大幅震盪，由此可驗證圖 13 的推論。

表七 以改良式主成份分析法搭配歐式距離進行混合資料庫之人臉辨識結果

資料庫	AT&T 資料庫	GIT 資料庫	CIT 資料庫	作者 自建	混合 資料庫
辨識率	87.50%	75.00%	81.57%	75.00%	80.33%

最後，吾人分別使用 AT&T、GIT、CIT 與作者自建資料庫每人 8 張、13 張、18 張及 18 張訓練影像加以混合，共 1456 張影像作為混合資料庫，而剩餘的 234 張影像則作為測試庫，並以改良式主成份分析法搭配歐式距離進行人臉辨識，其辨識結果如表七所示。在 AT&T、GIT、CIT 與作者自建資料庫的 80 張、100 張、38 張及 16 張測試影像中，其辨識成功率分別為 87.50%、81.57%、75.00% 與 75.00%，而整體辨識率為 80.33%。因本文中的三種特徵擷取演算法皆為圖像辨識的範疇，且四種人臉資料庫原始影像的檔案格式和維度皆不同，為使系統正常運作，則需進行資料維度大小的調整。然在維度大小正規化的同時，會造成原始影像資料結構的破壞，而降低解析度與辨識率。

六、結 論

本文以 Matlab/Simulink 進行即時人臉偵測，用以取代非即時人臉偵測，並探討訓練樣本數對人臉辨識之影響。在人臉偵測部份，即時人臉偵測僅需 9.1 秒，而非即人臉偵測需 29.5 秒才可完成偵測步驟。在人臉辨識部份，以三種特徵擷取演算法與四種決策法則組成 12 種辨識方式，再探討不同訓練樣本數對人臉辨識之影響，其目的是了解訓練樣本數對辨識結果與辨識時間之差異。由結果得知，以改良式主成份分析法搭配歐氏距離法的辨識方法所得之辨識較為優異。另外，亦可推論得知，作者自建資料庫之訓練樣本數至少需 13 張訓練影像可得較佳辨識率。然而，如增加訓練樣本數至 18 以上，只能些微提升辨識結果，但卻會增加辨識時間。綜合言之，足夠的訓練樣本數將有助於提升辨識率，但過多的訓練樣本數將會增加辨識時間。所以，如何在辨識率與辨識時間做出取舍，是件值得探討的問題。

符號索引

$\arg\max$	透過全域散佈矩陣所得之最大數值
$\arg\min$	透過全域散佈矩陣所得之最小數值
B	RGB 色彩空間之藍色成份像素點
Cb	YCbCr 色彩空間之藍色彩度
Cr	YCbCr 色彩空間之紅色彩度
$f_1^1(x, y)$	小波轉換後對角區域資訊
$f_2^1(x, y)$	小波轉換後垂直區域資訊
$f_3^1(x, y)$	小波轉換後水平區域資訊

$f_4^1(x, y)$	小波轉換後低頻區域資訊
G	RGB 色彩空間之綠色成份像素點
$h_0(x)$	高通濾波器
$h_1(x)$	低通濾波器
h_1, h_2	特徵點
$\overline{h_1 h_2}$	特徵線
h_p	測試圖像投影於 $\overline{h_1 h_2}$ 的特徵點
Q_{opt}	由 Σ_{B_i} 和 Σ_{Q_y} 組成的轉換矩陣
R	RGB 色彩空間之紅色成份像素點
u	降低維度後原始影像樣本
\bar{u}	降低維度後原始影像樣本平均值
W	全域散佈矩陣
W_{opt}	最佳全域散佈矩陣
W_x	原始樣本所組成之全域散佈矩陣
x	原始人臉樣本
\bar{x}	原始人臉樣本平均值
x_i^d	第 d 類第 i 張原始人臉影像
Y	YCbCr 色彩空間之亮度
y	測試影像
z	人臉特徵樣本
\bar{z}	人臉特徵樣本平均值
z_i^d	第 d 類第 i 張轉換後人臉影像
λ	特徵值序列
Σ_{T_x}	轉換前全域散佈矩陣
Σ_{T_y}	轉換後全域散佈矩陣
Σ_{B_x}	類別間矩陣
Σ_{Q_y}	類別內矩陣

參考文獻

- Zhao, W., Chellappa, R., Phillips, P. J., and Rosenfeld, A., "Face Recognition: A Literature Survey," *ACM Computing Surveys*, Vol. 35, No. 4, pp. 399-458 (2003).
- Brunelli, R. and Poggio, T., "Face Recognition: Features Versus Templates," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 10, pp. 1042-1052 (1993).
- Chellappa, R., Wilson, C. L., and Sirohey, S., "Human and Machine Recognition of Faces: A Survey," *Proceeding of the IEEE*, Vol. 83, No. 5, pp. 705-741 (1995).
- Hong, L. and Jain, A., "Integrating Faces and Fingerprints for Personal Identification," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 12, pp. 1295-1307 (1998).
- Turk, M. A. and Pentland, A. P., "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, pp. 71-86 (1991).
- Turk, M. A. and Pentland, A. P., "Face Recognition Using Eigenfaces," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 586-591, Maui, HI, USA (1991).
- Shapiro, L. G. and Stockman, G. C., *Computer Vision*, Upper Saddle River, Prentice Hall, NJ, USA (2001).
- Belhumeur, P. N., Hespanha, J. P., and Kriegman, D. J., "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 711-720 (1997).
- Yin, H., Fu, P., and Meng, S., "Sampled Two-Dimensional LDA for Face Recognition with One Training Image Per Person," *ICICIC First International Conference on Innovative Computing, Information and Control*, Vol. 2, No. 30-01, pp. 113-116 (2006).
- Nara, Y., Yang, J., and Suematsu, Y., "Face Recognition Using Improved Principal Component Analysis," *IEEE International Symposium on Micromechatronics and Human Science*, pp. 77-82 (2003).
- Yang, J. and Waibel, A., "A Real-Time Face Tracker," *Proceedings of the 3rd IEEE Workshop on Applications of Computer Vision*, Vol. 26, No. 11, pp. 1408-1423 (2004).
- Sirovich, L. and Kirby, M., "Low-Dimensional Procedure for the Characterization of Human Faces," *Journal Optical Society America*, Vol. 4, No 3, pp. 519-524 (1987).
- Huang, J., Gutta, S., and Wechsler, H., "Detection of Human Faces Using Decision Trees," *IEEE Proceedings of 2nd. International Conference on Automatic Face and Gesture Recognition*, Killington, VT, USA, pp. 248-252 (1996).
- Wang, K. -S., "Multiple Human Faces Detection and Identification System," Master Thesis, National Cheng Kung University, Tainan, Jul. (2005).
- Chien, J. -T. and Wu, C. -C., "Discriminant Waveletface and Nearest Feature Classifier for Face Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 12, pp. 1644-1649 (2002).
- Wang, H., Yang, S., and Liao, W., "An Improved PCA Face Recognition Algorithm Based on the Discrete Wavelet Transform and the Support Vector Machines," *International Conference on Computational intelligence and Security Workshops*, pp. 308-311 (2007).
- Wang, L., Zhang, Y., and Feng, J., "On the Euclidean Distance of Images," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 8, pp. 1334-1339 (2005).

18. Li, S. Z. and Lu, J., "Face Recognition Using the Nearest Feature Line Method," *IEEE Transactions on Neural Networks*, Vol. 10, No. 2, pp. 439-443(1999).
19. <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>, (01/08/2009 retrieved).
20. http://www.anefian.com/face_reco.htm, (01/08/2009 retrieved).
21. <http://www.vision.caltech.edu/html-files/archive.html>, (01/08/2009 retrieved).
22. Hsu, R. -L., Abdel-Mottaleb, M., and Jain, A. K., "Face Detecting in Color Images," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 5, pp. 696-706 (2002).
23. Garcia, C. and Tziritas, G., "Face Detection Using Quantized Skin Color Regions Merging and Wavelet Packet Analysis," *IEEE Transaction Multimedia*, Vol. 1, No. 3, pp. 264-277 (1999).
24. 黃登淵、莊國禎、楊晏和、陳南樺和王嘉宏，「複雜背景下多重人臉偵測演算法之研究」，科學與工程技術期刊，第三卷，第三期，第 35-43 頁(2007)。
25. Gargsha, M. and Panchanathan, S., "Face Detection from Color Images by Iterative Thresholding on Skin Probability Maps," *IEEE International Symposium on Circuit and Systems*, Vol. 5, No. 5, pp. 673-676 (2002).
26. Gonzalez, R. C. and Woods, R. E., *Digital Image Processing*, Prentice Hall, NJ, USA (2002).

2009 年 02 月 26 日 收稿
 2009 年 03 月 05 日 初審
 2009 年 04 月 15 日 接受