

人臉偵測及 辨識方法探究

曾婉菁

一、介紹

生物辨識 (Biometric) 是透過人體指紋、臉部、聲音或虹膜等身體器官組織的獨特性來辨識使用者身分的一項技術。主要運用人體身上的特徵做為識別的密碼，因此在技術的開發上必須選擇準確度高、容易使用的辨識特徵以利使用，主要可分為生理上的 (如臉形、指紋、虹膜)，或是獨特的行為模式 (如聲音、簽名、密碼)。以準確度來說，「生理特徵」在唯一性及安全性上明顯優於「行為特徵」。生物辨識系統是利用每個人獨一無二的生理或行為特徵來辨識使用者的身分。運用生物辨識技術，你的身體就是密碼，不需要記憶一長串不易記住的數字，也不怕遺失，複製不易，更不用擔心遭人盜用，因為你是獨一無二且「隨身攜帶」。目前生物辨識核心技術的發展，指紋辨識佔技術比率的 54%，簽名辨識佔技術比率約 21%，臉部辨識則佔技術比率 16%，虹膜辨識佔技術比率 9%。市面上以「指紋辨識」技術較成熟，市場占有率最高，其次則為成長速度最快的「臉部辨識」技術。而「虹膜辨識」的準確度最高，但是由於使用上必需以紅外線掃描眼球，在價格及安全性的考慮下，並不容易發展成為大眾化的產品，相對的市場占有率也就無法迅速拓展。其餘生物辨識科技則仍受一般消費者的使用習慣、可接受度，以及經濟價格因素影響，成長較緩慢。

近年來以人臉為特徵的偵測與辨識技術發展十分迅速，作為人臉自動辨識系統的第一步，人臉偵測技術有著十分重要的作用，它為後續的人臉分類提供了待辨識人臉的具體資訊。人臉偵測常常作為人臉追蹤與辨識的前置作業，是一個複雜且困難的研究課題，其結果足以影響整個系統的效能，人臉偵測要走向實際應用，準確和速度是很重要的兩個關鍵因素。

二、人臉偵測方法

國內外相關研究在人臉偵測與辨識已有許多的相關研究〔1～6〕，目前人臉偵測主要方法有

1. 樣板比對法 (Template matching method) :

Template matching method 是目前最普遍的人臉辨識方法，首先制定一個標準的人臉樣板，這個人臉樣板通常都是正面的 (frontal)，利用這個人臉樣板來與輸入的影像作比對，以其結果來做進一步的判斷，然而，一張人臉的維度通常都很大，若直接將一筆影像資料轉成向量進行辨識，其計算量將相當驚人，所以採用此方法進行辨識之前，通常會先以統計的方法將原始影像投射至另一維度較低的特徵空間，此特徵空間依然保有大部份的影像資

訊，以期降低計算量。

2. 條列式規則區分法 (Rule-based method) 又稱為 (knowledge-based method) :

條列式規則區分法主要就是考慮到人臉上應有一定的規則，並以此規則在人臉候選區域進行搜尋，此方法的缺點是很難明確定義描述人臉特徵的規則，如果規則定義太過鬆散，則可能將非人臉的區域誤判為人臉，若定義太過嚴謹，又很容易偵測失敗。

3. 以訓練影像資料為基礎的方法 (Appearance-based method) :

(1) EigenFaces (五官特徵) :

Appearance-based method 與 Knowledge-base method 不同之處在於，系統建立者不需定義規則來描述人臉的五官特徵 (EigenFaces)，而是輸入一連串的人臉影像如圖示 (1.1, 1.2)，讓偵測系統透過一些統計或學習方法自行定義人臉影像的模型，並用此定義的模型進行人臉區域的搜尋。George Bebis 建構的人臉偵測與辨識系統，以 Principle Component Analysis (PCA) 的方法來定義人臉影像模型，並使用基因演算法搜尋某個特定對象的人臉區域。George Bebis 並未使用任何影像處理的技術，而是直接使用基因搜尋演算法搜尋整

表 1 各種生物辨識技術比較表

	原理	優點	缺點
指紋辨識	利用每個人手指指紋圖像特徵不同，來辨識其身分	1. 指紋每個人都不一樣。 2. 資料庫系統建立較早	1. 指紋容易被他人取得。 2. 為接觸式辨識，容易產生衛生問題。 3. 手汗症、或是手指、磨損受傷可能會影響辨識結果。
虹膜辨識	虹膜為附著在瞳孔上的生理薄膜，顏色和結構特徵獨特；利用眼睛掃描進行辨識	1. 虹膜特徵紋理結構複雜不易改變。 2. 不易被複製或取得。 3. 精準度高。	1. 掃描範圍必須只控制在眼睛範圍，速度可能較慢。 2. 透過紅外線掃描可能會有傷害眼睛的風險。 3. 辨識機器較昂貴。
聲音（紋）辨識	利用每個人聲帶和聲道的生理結構不同，分析說話者特性來做語音辨識	每個人說話習慣所成現的聲波幅度不同。	聲音有可能因為各種因素面臨變聲，影響辨識準確度。例如：男孩青春期的、感冒沙啞、聲帶受損等。
人臉辨識	利用人臉五官輪廓的角度、距離來建立 3D 結構模型，進行身分辨識。	1. 為非接觸式的辨識，不會有衛生問題。 2. 簡單方便，速度快。 3. 人臉辨識無法以照片作假；不會受帽子、眼鏡其他外物影響。 4. 不需受辨識者配合，可直接擷取影像	可能受環境光線影像。

個影像畫面，找出特定對象的人臉區域，再將此區域所涵蓋的影像投影到 PCA 特徵空間中進行偵測與辨識。

(2) 類神經網路 (Neural Network) :

以數學統計學習模型 (SVM)，對每張影像取出 20x20 像素 (pixels) 的小影像，經過前置處理，如：histogram equalization 等，再輸入到預先訓練好的類



圖示 1.1 : PCA 的訓練習影像 (a)



圖示 1.2：PCA 的訓練習影像 (b)

神經網路模型或數學統計學習模型來判斷，而因為人臉大小不一，對原始的影像要一步一步地縮小來做偵測，在訓練類神經網路模型或數學統計學習模型，我們輸入的訓練資料分為兩類，一類是人臉影像，另一類是非人臉影像，這裡的困難之處在於非人臉影像是難以用一堆非人臉訓練資料來代表，所以訓練多個模型來加以判斷是必要的，不過偵測的速度就變慢了。

三、應用

1. 車內駕駛監視預警系統

提昇行車安全為交通運輸產業的重要議題，大多數的肇事原因歸究於駕駛本身的駕駛行為，例如疲勞或酒醉駕駛，因此，車內監視預警系統即可透過攝影機觀看駕駛人的精神狀態，進而降低交通事故的發生。

利用人臉辨識系統，估算出人臉特徵座標，再依特徵擷取出人臉五官，計算左右兩眼真實距離遠近，來估算駕駛頭部擺



圖示 2：人臉辨識—駕駛人頭部擺動方向之追蹤

動方向。左眼距離近，右眼距離遠，代表擺動方向是右邊；右眼距離近，左眼距離遠，代表擺動方向是左邊。此頭部擺動方向即可作為預警的參考之用，如圖示 (2)。

2. 人機介面：

使用者在觀看電視時，利用預先建立的人臉資料庫及喜愛頻道，音樂，節目類別等資料，與現有的節目系統作一結合，當使用者打開電視，藉由 WebCam 或相機取得使用者的影像，與預先建立的人臉資料庫一一比對，從中辨識觀看者的喜好，再從節目表單中列出可能的喜好的節目，列出清單以供選擇，省去觀看者要一切換頻道的困擾。

3. 虛擬擴增實境：

最有名的莫過於是 Microsoft 體感裝置，不需要任何控制器的 Kinect，Kinect 一次可擷取三種東西，分別是彩色影像、3D 深度影像、以及聲音訊號。中間的鏡頭是



圖示 3：圖為遊戲《Your Shape Fitness Evolved》畫面

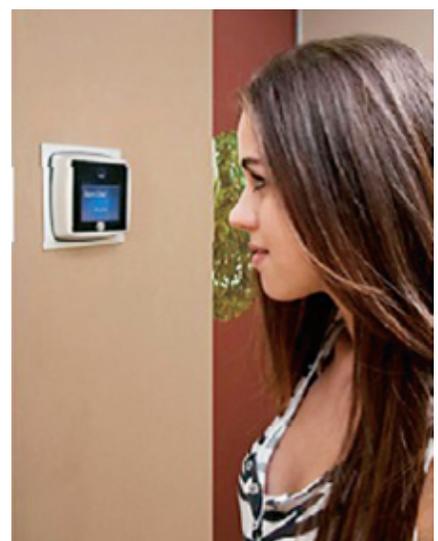
一般常見的 RGB 彩色攝影機，左右兩邊鏡頭則分別為紅外線發射器和紅外線 CMOS 攝影機所構成的 3D 深度感應器，中間視訊鏡頭則是用來辨識玩家身分（靠著人臉辨識和身體特徵）、以及辨識基本的臉部表情，圖示（3）為遊戲” Your Shape Fitness Evolved”。

4. 課堂自動點名系統

主要因素為鬆散的自我學習或自主學習，其成效不易彰顯。有鑑於此，利用低價位網路攝影機（Web-Cam），配合自行開發之一套偵測程式，當學生啟動課程時，同時也啟動偵測程式。程式運用網路攝影機擷取影像，找尋影像中的人臉，來判斷



圖示 4：人臉辨識—課堂點名系統



圖示 5：人臉辨識—大樓門禁系統

課程播放期間，學習者是否坐於電腦螢幕前。配合其它學習紀錄如圖示（4），可提供學習成效分析，提升自我學習的成效。

5. 門禁系統

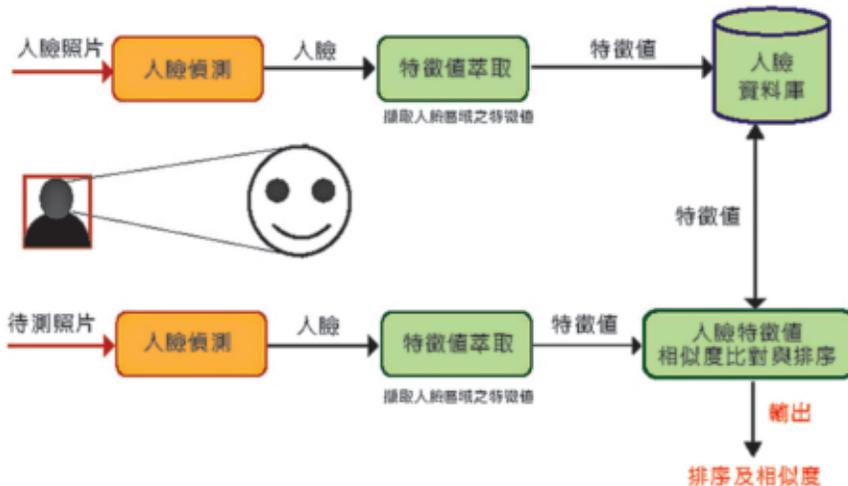
在大樓入口或機密重地，利用人臉辨識確認進入者的身份，可作為輔助識別之用，相較於傳統的方式，除了更有效外，可省去感應卡的管理建置費用，也能補足員工忘記帶卡的困擾，如圖示（5）。相對於其他生物辨識方式，人臉辨識這種非侵入、非接觸式的方式是較為先進、文明且易於為大多數人接受的，在雙手無暇刷卡或按密碼時，特別能感受到人臉辨識技術的好處。

至於其他可能影響辨識率，例如光線強弱，臉部受創或者是否戴帽、髮型的變化或染髮、帶眼鏡或鬍鬚的有無變化，只

要主要的人臉特徵能清楚分辨的情況下，辨識結果可在控制之內。常見代打卡的困擾也能因使用人臉辨識系統而獲得解決，以人臉代替感應卡時，只有本尊能完成打卡動作。一旦發現陌生人也能立即警示，通知相關人員處理，但此類的安全性也可能遭到質疑其準確度，故經常搭配其他的生物特徵來多重確認，以達到安全管控。

6. 犯罪偵查系統：

目前台灣警政署在各地道路皆裝設監視系統，同時鄰里社區也有自己的監視器，平時錄影，一旦有犯罪情事發生，利用人臉辨識的技術，將犯罪嫌疑人的人臉特徵輸入到人臉辨識系統，利用各地的錄影系統，可快速的將犯罪嫌疑人的行跡一一揭露，當然，這也會有抵觸人權的問題產生。端看監督的機制是否完善。



圖示 6：人臉辨識系統 -- 警政署建置的警政治安相片比對系統

內政部警政署建置的警政治安相片比對系統如圖示（6），是如何協助警察機關執行特定人物比對呢？首先，員警將待比對樣本相片進行基本前置處理後，將影像上傳系統，啟動人臉偵測功能，擷取人臉區塊的特徵值，與人臉資料庫進行相似度比對排序，最後再輸出結果。也可經過設定各種條件，協助員警縮小偵查範圍，加速破案；亦可輸入身分證字號，驗證待比對樣本相片身分，系統會自動計算相片的相似度，協助承辦人員判斷兩者是否為同一人。

不明人士的身分查證，是警察機關每天都會遇到的工作。未來可運用臉部辨識比對技術，必要時由勤務處所或現場人員以相機拍攝不明人士的正面臉部影像，上傳進行影像比對，大幅加速身分查證時間，亦可於確認人別後，深入調查，以發現不法。此外，針對路倒人士、失智老人，也可比對過濾後之聯絡清單，短時間內就能找到待協尋人口的家屬，讓一家團聚，家人早日放下焦急的心。

7. 入出境管理系統：

可作為入出境旅客之身份識別輔助系統之用，大部份的國家在旅客入出境時，皆會拍照及按捺指紋，兩者合一，可將誤判率降低，而人臉辨識為一種非接觸式的生物特徵，可快速的在公共場所中，進行



圖示 7：人臉辨識系統 -- 高雄國際機場

追緝識別。

移民署針對入出境查驗作業，已修改成「三方比對」，高雄國際機場採用的人臉辨識系統為「人臉辨識 + 雙指紋」的客制化系統如圖示（7），外國人從高雄國際機場入境時需按捺雙手指紋及採集臉部特徵資料，出境時則以「護照 + 臉部特徵 + 雙手指紋」做為身份認證依據，國境安全再往前邁進一大步。

8. 相簿搜尋：

利用姓名及對應的照片，來搜尋相片，可幫助使用者快速分類及搜索，而不需要使用者一個個打開相片尋找，節省相當多的時間作搜索及分類。

Google 以獨特的人臉辨識技術，為 Picasa 網路相簿建立了一個新功能—「名稱標記」如圖示（8），它能幫你自動搜尋相片中的人臉，讓你能迅速標記他們的名稱，就能快速地整理、尋找相片中的人物。

使用「名稱標記」，Picasa 會將找到的



圖示 8：人臉辨識—Picasa 之名稱標記

臉孔加以分組，並顯示畫面上方的橫向捲軸中，捲軸中每張縮圖就代表一組相似的臉孔。點選縮圖後下方會列出相似的照片，直接在下方的「輸入名稱」處鍵入人物的名稱，就能幫該照片加上名稱標記。如果有誤判的情形，可以將圖片下方的「~」取消即可。

9. 廣告系統：

利用人臉偵測系統根據人臉位置、注意力、眨眼、搖頭、點頭、停留的時間，來判定廣告的吸引力及喜好度，計量出人流及廣告效益評估。

在 7-11、全家、麥當勞或星巴克，我們總能看到牆上的大型數位電子看板如圖示 (9) 輪播著一個又一個的廣告畫面。殊

不知當我們目光停留於這些數位看板時，看板上的感應器會在背景分析人臉生物特徵，統計每個人的觀看行為，並透過統計資料讓廣告商更精準地投放內容。

這類的智慧電子看板會偵測螢幕前方 5



圖示 9：高雄捷運上的數位看板應用人臉偵測 -- 人群對廣告看板之效益評估

公尺內的動態活動，一旦有人注視螢幕 2 秒，系統會將其認定為有效注視者（正臉對螢幕），並開始分析每一秒的特徵狀態（稱為非結構性資料），包含目光位置、年齡、性別等資訊。透過這龐大的資料庫與廣告內容交叉分析，便能了解廣告是否有效收視；從長線來看，店家能掌握店面的客群組成，或是在特定氣候因素下消費者在意的廣告內容，精準掌握銷售族群。

10. 網路安全付款機制：

隨著網際網路應用普及化、電子交易頻繁化，傳統以密碼和帳號的認證方式，已難以確保認證的安全性，然而，利用個人獨特的生物特徵所辨識之「生理密碼」，像是藉由人臉、指紋、個人簽名與虹膜辨識等，由於具有難以複製或遭竊的特性，即能有效且便利地提高資料安全防護等級。

利用人臉辨識技術，僅需使用一般攝影機，即可提供自動擷取、主動辨識、身份確認、影像儲存等強大功能，人臉辨識除可作為密碼之外的另一道安全防線，亦能增加產品附加功能，加上人臉特徵不需記憶且隨身攜帶，對於使用者來說，有其絕對的便利性及私密性。

目前，是一家芬蘭的初創公司 Uniqul，他們已經對申請了專利的 unqu 支付系統完成測試，免去了消費者在商店內進行支付時的很多安全隱患。在 Uniqul 的



圖示 10：人臉辨識 -- 之付費系統

系統如圖示（10）中，你的臉就是你的密碼。

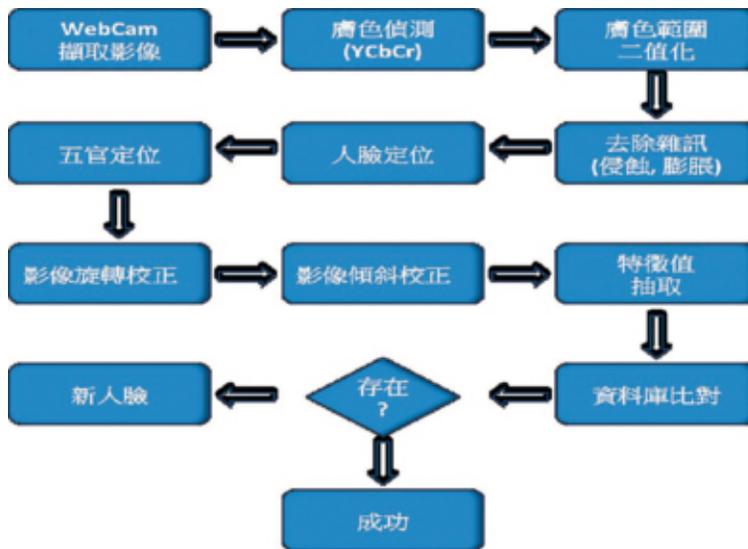
這家公司準備儘快在赫爾辛基地區鋪設它的終端。事實上它的運行機制聽起來很簡單：在對交易進行確認支付的時刻，使用者僅需要把臉對著攝影機，等著他們的 ID 顯示出來，然後在平板顯示幕上點擊「OK」便可完成支付確認。

整個過程不需要信用卡、錢包、手機——這表明該系統會把你的 ID 以及你的支付方式等細節儲存到伺服器上。Uniqul 說你可以用任何支付系統進行註冊，從 PayPal 到傳統銀行卡，而且這些資料會使用軍事等級的加密保護技術加密。

四、人臉偵測及辨識流程：

本文以樣板比對法來探討人臉偵測及辨識，然而條列式規則區分法及以訓練影像資料為基礎的方法不在探討之列。

以人臉偵測而言，從攝影機取得之彩



圖示 11：人臉偵測與辨識流程圖

色影像經過膚色偵測處理，去除了影像中大部份的非人臉的區域，其餘為人臉之可能區域，而要判斷該區域是否為人臉，最好的方法即是偵測該區域是否存在人臉之五官特徵，一旦該區域偵測到五官成分，即可判定該區域存在人臉，反之則無。

再者，一旦偵測人臉的五官成分，即可從中取出特徵值（如眼嘴鼻的比例、度、距離等），作為辨識資料之用，進而比對資料庫中所預存之人臉資料，符合此類特徵者，則判定符合，否則為新的人臉，人臉，偵測及辨識流程如圖示（11）所示。

五、人臉偵測及辨識的背景知識：

1. 色彩空間：

HSV：

HSV 色彩空間如圖示（13）對人類來說較容易理解，它將亮度（Value）與彩度（Chrominance）從色彩中分離出來，彩度又分為色相（Hue）與飽和度（Saturation），HSV 的優點是分離了影像中色彩與亮度的部分，對於顏色分割有很好的效果，因此許多人臉偵測的系統都使用此色彩空間。由於亮度已經被獨立出來，彼此之間的關聯性很小，故適合做影像前處理，但由於電腦絕大部分都是以 RGB 來表示，所以必須經由 RGB 轉換到 HSV，轉換的過程為複雜，其公式如圖示（12）：

$$H1 = \cos^{-1} \left\{ \frac{0.5[(R-G) + (R-B)]}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \right\}$$

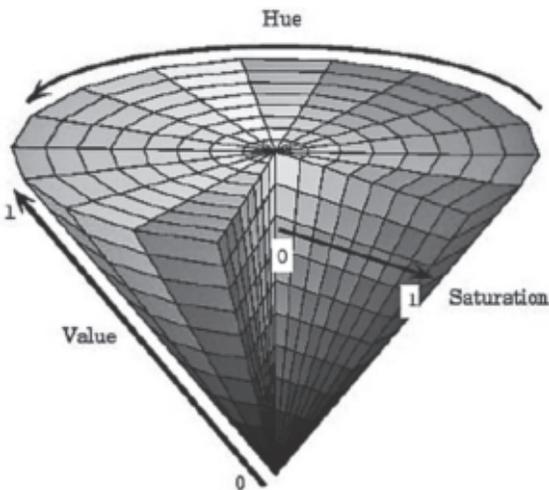
$$H = \begin{cases} H1 & \text{if } B \leq G \\ 360^\circ - H1 & \text{if } B > G \end{cases}$$

$$S = \frac{\text{Max}(R, G, B) - \text{Min}(R, G, B)}{\text{Max}(R, G, B)}$$

$$V = \frac{\text{Max}(R, G, B)}{255}$$

圖示 12：RGB 與 HSV 轉換公式

其中色調 H (Hue) 為偵測人臉的依據，主要原因是 H 比較不受光線強弱的影響；S 代表顏色中的飽和度 (Saturation)，其值介於 0 ~ 1 之間；V 代表顏色的明暗度 (Value)，也是介於 0 ~ 1 之間。



圖示 13：HSV 之色彩空間圖

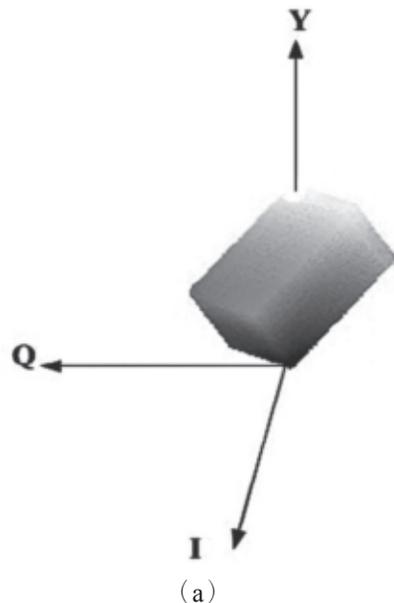
YIQ：

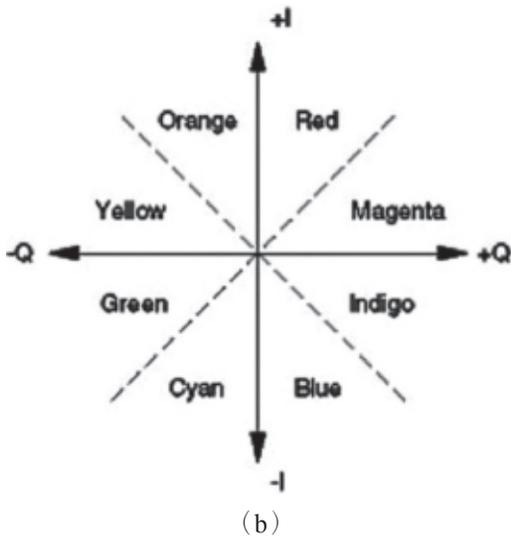
YIQ 是 NTSC 的色彩系統，定義了三種影像資料成分：照度 (Y：luminance)，色調 (I：hue) 和飽和度 (Q：Saturation)，其中選擇使用英文字母 YIQ 為約定俗成，照度表示灰階資訊，而其他兩個成分具有電視訊號的彩色資訊，YIQ 與 RGB 轉換公式如圖示 (14)：

$$\begin{Bmatrix} Y \\ I \\ Q \end{Bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.320 \\ 0.212 & -0.532 & 0.311 \end{bmatrix} \begin{Bmatrix} R \\ G \\ B \end{Bmatrix}$$

圖示 14：YIQ 與 RGB 轉換公式

其中 RGB 與 YIQ 的範圍均在 0 ~ 255 之間，而在色彩空間的關係如圖示 (15)：





圖示 15：RGB 與 YIQ 的色彩空間關係圖

YIQ 也經常被用於描述膚色的一種色彩空間，使用 Y 與 I 兩種色彩元素將膚色區域自背影中分離出來。

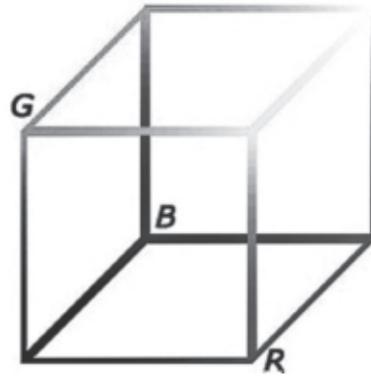
YCbCr：

YCbCr 又稱為 YCC，是 REC601-1 (CCIR 601-1) 所制定的，MPEG 視訊與 JPEG 影像的標準色彩空間，被廣泛運用在數位視訊中，其中亮度資料為 Y 表示，而色彩資料以 Cb (blueness) 和 Cr (Redness) 成分來表示，對亮度的分離性高，方便與彩度分開操作，適合影像處理使用，YCbCr 與 RGB 的轉換公式如下如圖示 (16)：

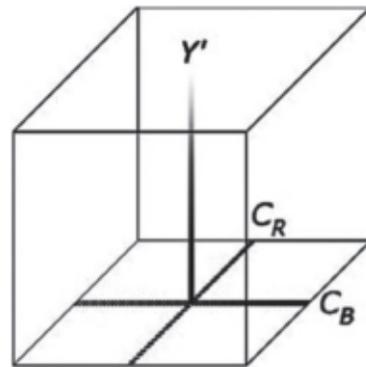
$$\begin{Bmatrix} Y \\ Cb \\ Cr \end{Bmatrix} = \begin{Bmatrix} 0 \\ 128 \\ 128 \end{Bmatrix} + \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.5 \\ 0.5 & -0.418 & -0.081 \end{bmatrix} \begin{Bmatrix} R \\ G \\ B \end{Bmatrix}$$

圖示 16：YCbCr 與 RGB 的轉換公式

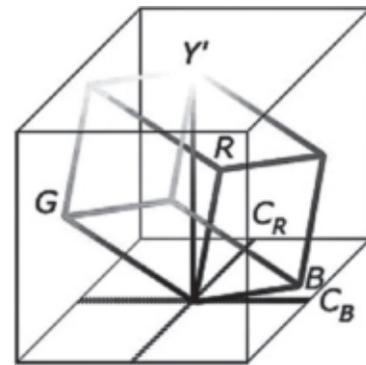
其中 RGB 與 YCbCr 的值皆介於 0 ~ 255，兩者色彩空間關係圖如圖示 (17)：



(a)



(b)



(c)

圖示 17：RGB 與 YCbCr 之色彩空間關係圖

2. 型態學的方法

(1) 四連通成份搜尋 (4-connected component)

連通成份搜尋主要是用來計算影像中屬性相同且相連之像素所組成的物體面積，依連通成份的定義不同，可分為 4 連通成分，8 連通成分等，其所計算的結果也不同，如圖示 (18.a) 為 4 連通成份的結構元素，假設一個成分相連的區域 Y，並指定 Y 中的任一元素 P 作為搜尋的起始點，如圖示 (18.b) 所示，再依據以下的定義，以疊代的方式進行搜尋，找出所有屬於 Y 的元素個數，如圖示 (18.c-e)

$$X_k = (X_{k-1} \oplus B) \cap A \quad k=1,2,3\dots$$

其 $X_0 = P$, B 是結構元素，若 $X_k = X_{k-1}$

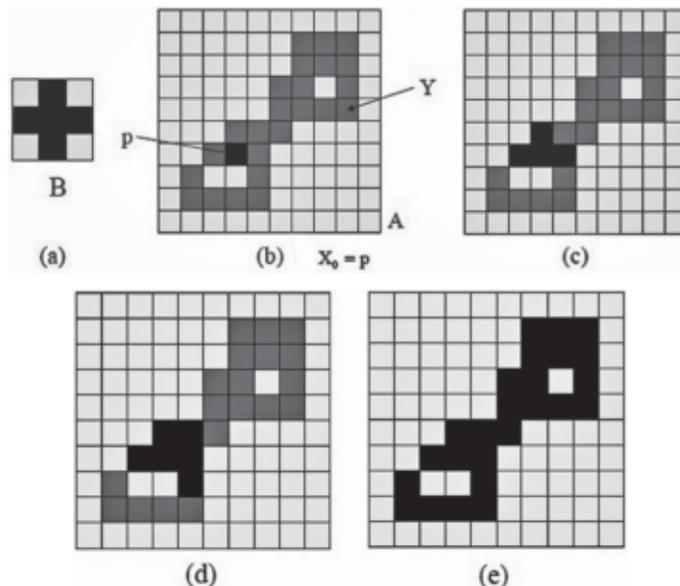
則代表演算法收斂，並令 $Y=X_k$

(2) 膨脹

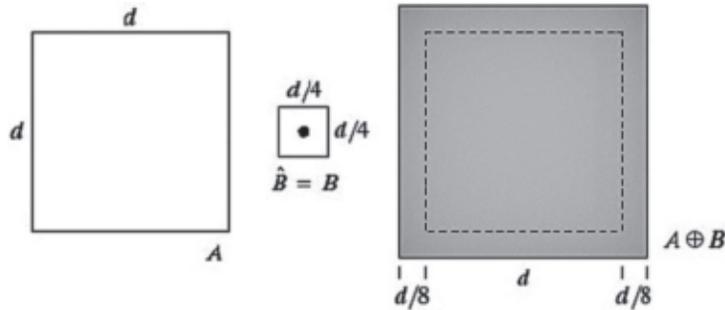
膨脹運算主要的作用是讓二值化影像中的物體增大與增厚，其運算方式也是由結構元來決定，不同的二元素可得不同的膨脹效果，數學式以 $A \oplus B$ 來表示 A 被結構元素 B 膨脹，定義為：

$$A \oplus B = \{z | (\hat{B})_z \cap A \neq \emptyset\}$$

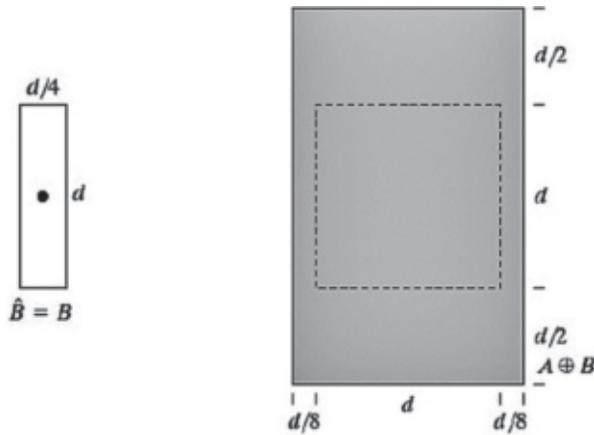
其中 \emptyset 為空集合，B 為結構元素， (\hat{B}) 為 B 的反射 (reflection)，而 $(\hat{B})_z$ 代表 (\hat{B}) 平移 (translation) z 單位，實際的作法是讓結構元素的原點在整個二元影像中移動，當結構元素的原點與標示為 1 的影像重疊時，將所有與結構元素有交集的像素設為 1。如圖示 (19)：



圖示 18：4 連通成分之搜尋

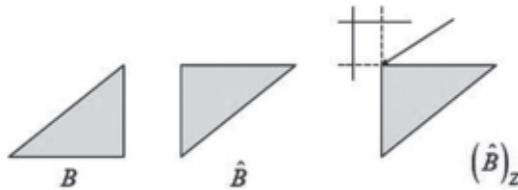


(a)



(b)

圖示 19-1：不同之結構元素所得之不同膨脹效果



圖示 19-2：型態學之平移與反射

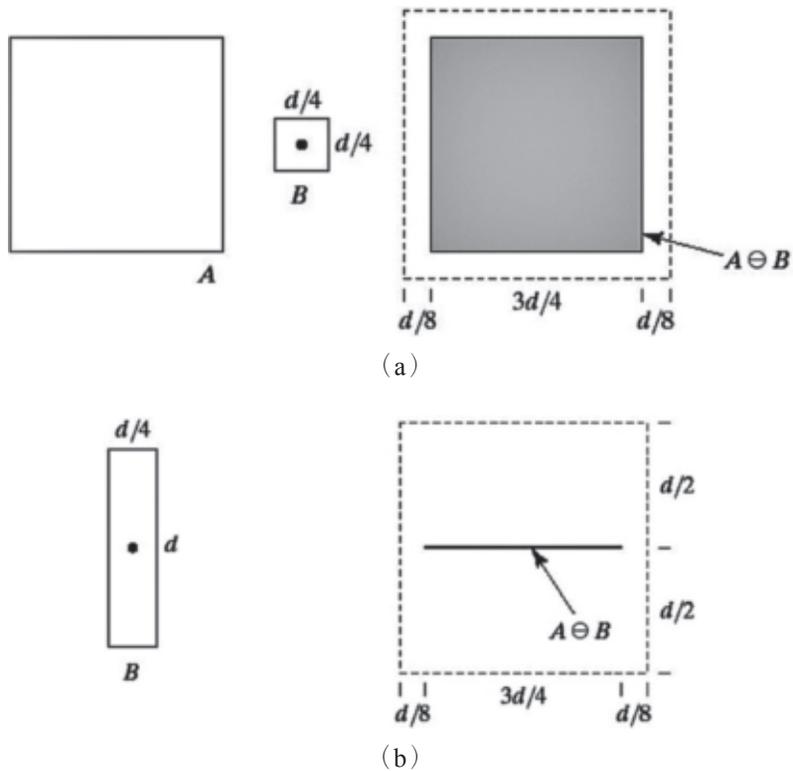
(3) 侵蝕

侵蝕運算主要作用是讓二元影像中的物體收縮或變薄，其收縮的方式與程度可由結構元素的形式加以控制，不同的結構元素，可得不同的侵蝕效果如圖示 (20)，以 $A \ominus B$ 來表示 A 被結構元素 B 所侵蝕，其

定義為：

$$A \ominus B = \{ z | (B)_z \cap A^c \neq \emptyset \}$$

其中 \emptyset 為空集合， B 為結構元素， A^c 為 A 的補集合，



圖示 20：不同的結構元素可得不同之侵蝕效果

3. 膚色分割 (Color Segmentation)

人類的眼睛對低頻的資料比對高頻的資料具有更高之敏感度，事實上，人類的眼睛對亮度的改變也比對色彩的改變要來得敏感許多。在 JPEG 的應用上，一般說來都是只有處理灰階及全彩的影像，全彩的影像是由 Y、Cb、Cr 三個顏色成份

(component) 組合而成，YCbCr 色彩空間中分別代表的意思是，亮度 Y (luminance) 和彩度 Cb (blueness)、Cr (redness)，也因為其明顯對於亮度和彩度有高分離性，所以廣泛的被應用在影像處理中，其 YCbCr 與 RGB 轉換公式如圖示 (21)：

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix}$$

圖示 21：YCbCr 與 RGB 轉換公式

以下為膚色判斷公式：

$$\begin{cases} Y = 0.2989 \times R + 0.5866 \times G + 0.1145 \times B \\ C_b = 0.5647 \times (B - Y) \\ C_r = 0.7132 \times (R - Y) \end{cases}$$

(a)

$$\begin{cases} \text{if}(Y > 128) \\ \theta_1 = -2 + \frac{256 - Y}{16} \\ \theta_2 = 20 - \frac{256 - Y}{16} \\ \theta_3 = 6 \\ \theta_4 = -8 \\ \text{if}(Y \leq 128) \\ \theta_1 = 6 \\ \theta_2 = 12 \\ \theta_3 = 2 + \frac{Y}{32} \\ \theta_4 = -16 + \frac{Y}{16} \end{cases}$$

(b)

$$\begin{cases} C_r \geq -2(C_b + 24); & C_r \geq -(C_b + 17); \\ C_r \geq -4(C_b + 32); & C_r \geq 2.5(C_b + \theta_1); \\ C_r \geq \theta_3; & C_r \geq 0.5(\theta_4 - C_b); \\ C_r \leq \frac{220 - C_b}{6}; & C_r \leq \frac{4}{3}(\theta_2 - C_b); \end{cases}$$

(c)

圖示 22：膚色判斷公式

其中如圖示 (22.b) 及圖示 (22.c) 中的膚色判斷公式為根據圖示 (22.a) 中包含於膚色區域的八條邊界之判斷公式，另

外，由於膚色分布區域在 Y 值大於或小於 128 時有顯著之不同，因此以 Y 值等於 128 為界分別作不同的判斷處理。

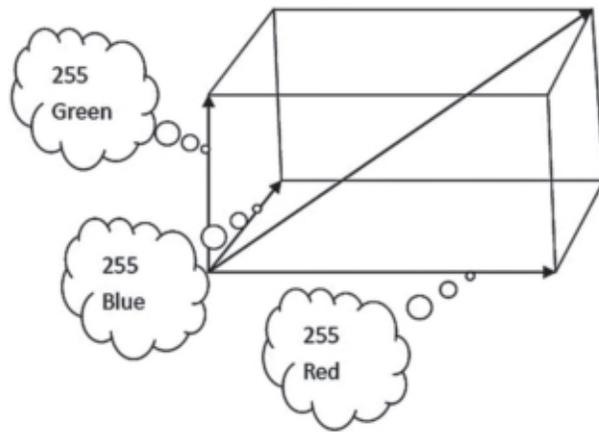
4. 二值化原理

二值化是在處理影像時最常用的一種方式，也是由灰階化影像經過處理而成。在二值化處理圖像的時候把大於某個臨界灰度值的像素灰度設為灰度極大值，把小於這個值的像素灰度值設為灰度最小值，從而實現二值化，在我們把彩色影像圖檔經過二值化處理之後，圖檔較容易儲存、處理與辨認，所以在本專題中二值化佔有非常重要的地位。

5. 影像灰階化

灰階是一種 8 位元的影像，以 8bit 來表示，共有 256 種深淺不一的灰色組合，當 Webcam 擷取彩色影像圖檔之後，若不做灰階化處理，會讓背景與人物的顏色太過相近，使得辨識很容易產生誤差，也會讓人物的定位產生非常大的誤差，所以在擷取彩色圖檔之後就要使用灰階化，把背景跟人物的差別區分出來，這樣能把辨識臉部特徵如 (眼、鼻子、嘴巴。) 的誤差值降到最低，也就是能更順利的把臉部特徵值給找出來。

經過上面所述，灰階就是液晶螢幕上人們肉眼所見的一個點，即一個像素，都



圖示 23：RGB 與灰階示意圖

是由紅、綠、藍、(RGB) 三個子像素所組成的。每一個子像素，其背後的光源都可以顯現出不同的亮度級別，經由這三種顏色組成起來，最後會形成不同的色彩點。而灰階代表了由最暗與最亮之間不同亮度的層次級別。(圖示 (23) 為 R、G、B 與灰階示意圖)。

N 設定為 3，即使用 3*3 矩陣進行侵蝕的處理影像。

侵蝕的條件：

P1	P2	P3
P8	P	P4
P7	P6	P5

圖示 24：3x3 矩陣

6. 雜訊移除 (Noise Filtering)

將之前膚色偵測之後的範圍取二值化，然後發現圖片周圍有許多微小的雜訊，為了避免這些雜質影響往後的辨識工作，對於這些雜訊，經常使用了影像型態學中的斷開運算。

斷開運算包含了之前所提的侵蝕與膨脹，二值化後的影像先運用侵蝕來縮小區域，接著再利用膨脹來擴張該區域，經過這樣的方法就可以將雜訊去除。本專題使用遮罩 (N * N) 的矩陣如圖示 (24)，將

首先判斷判斷 P 點是否為 1，若為 1，且周圍的 P1 ~ P8 皆為 1。則 P=1。

反之，則都設定為 0，其運算公式如圖示 (25)。

$$P = P1 \cap P2 \cap P3 \cap P4 \cap P5 \cap P6 \cap P7 \cap P8$$

圖示 25：斷開運算公式

7. 人臉定位：

將去除雜訊的膚色偵測二值化所留下的白色面積，運用標記連通的方法取出最

大面積，利用白色面積的長寬比，來作為人臉定位的依據，如圖示（26）：



圖示 26：人臉定位

8. 五官定位

在經過膚色偵測，去除雜訊之後，最後就是對於五官的定位。利用二值化後的影像與去除雜訊之後的影像做比對，如圖示（27）（28）。

9. 人臉特徵值：

以下列出人臉的特徵，可能有哪些特徵：

（1）比例值：

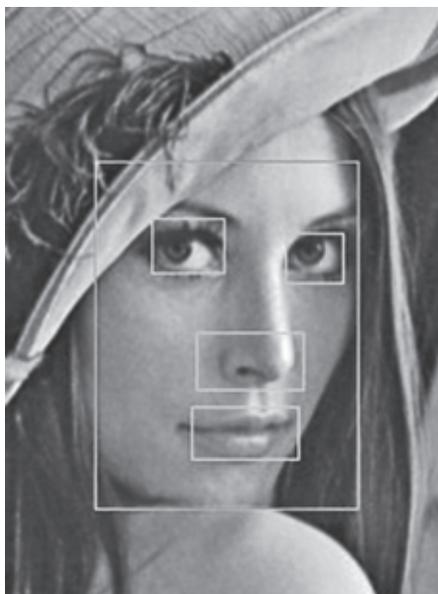
眼臉比例（臉寬 / 兩眼間距）*255。

眼嘴比例 1（兩眼間距 / 嘴的長度）

*255。



圖示 27：五官切割



圖示 28：五官定位

眼嘴比例 2（眼至嘴的距離 / 兩眼間距）

*255。

- (2) RGB 均值：
- 眼的 RGB 均值。
- 嘴的 RGB 均值。
- 鼻的 RGB 均值。

(3) 四組 R-G R-B G-B (左眼、右眼、嘴、鼻)

- (4) 兩眼距離 =X、臉部寬度 =Y、
- 特徵值 =Z
- $(Y/X) \times 255 = Z$

其中乘以 255 是因為要把特徵值的差異性放大，因為 Y/X 的值每個人都有所不同，但是這個值又沒有很大，所以要乘以 255 放大之後，能更有效的辨識出照片中的人物。

10. 影像旋轉校正

利用人臉影像的投影圖來判斷人臉旋轉方向，以藉此找出連續影像中最接近正面的人臉影像，利用人臉影像的投影圖來判斷人臉旋轉角度，而我們利用臉部特徵集中線 A 與臉部方框的中心線 B，相減求出距離 d 如圖示 (30)，再把距離與臉部邊界寬度相除，求出旋轉率，最後利用門檻值過濾轉動過大的影像，公式如下所示：

轉動公式如圖示 (29)。

$$turning_rate = \frac{d}{box_width}$$

其中，*turning_rate* 為旋轉率，*d* 是距離，*box_width* 是臉部邊界寬度

圖示 29：影像旋轉轉動公式



圖示 30：臉部特徵與中心線圖

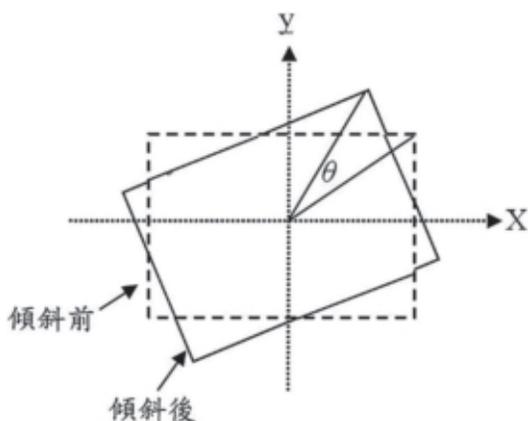
11. 影像傾斜校正

影像傾斜為一種基本的影像調整方法，用以修正各個影像中不同角度的差異，假設影像中一個像素點位於 (x, y) 座標點，在經過 θ 度傾斜後，像素點位移至新的座標點 (x', y')，兩個座標點之間的映射關係如圖示 (31) (32)：

$$x' = x \cdot \cos \theta + y \cdot \sin \theta$$

$$y' = -x \cdot \sin \theta + y \cdot \cos \theta$$

圖示 31：影像傾斜座標轉換公式



圖示 32：影像傾斜座標映射關係

六、參考書目

1. <http://www.techbang.com/posts/14116-to-see-the-painted-faces-cover>
2. <http://www.techbang.com/posts/16531-digital-sign>
3. http://www.delightpress.com.tw/bookRead/sknn00001_read.pdf
4. <http://tw.shuttle.com/main/productsNewsDetail?newsId=1412&productId=1738>
5. <http://www.techbang.com/posts/2936-get-to-know-how-it-works-kinect>
6. http://www.moi.gov.tw/chi/chi_ipmoi_note/ipmoi_note_detail.aspx?sn=112
7. www.asmag.com.tw/product/product_detail.aspx?pid=6465
8. <http://www.staff.chu.edu.tw/focus/focus548.htm>
9. 蔡安智、林達德、徐子建、陳加增，“基於雙眼視覺之駕駛人臉部特徵擷取演算法”，中華民國第十四屆車輛工程學術研討會，2009年10月30
10. 黃登淵、莊國楨、楊晏和、陳南樺、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，民國九十六年
11. 黃金鷗、蔡孟伸，“即時人臉偵測及辨識系統的開發：民國九十六年
12. 姚志佳、謝林宏、陳柏志、劉辰哲、廖崇哲、周慶哲，“人臉辨識系統”
13. 呂芳懌、余少棠、許閔雄、林晁立，“影像資料庫在人臉辨識上的應用—前科犯資料庫”，東海科學第一卷：103-122，民國88年7月
14. 李建興、林應璞、游凱倫，“即時人臉偵測與辨識”，技術學刊第二十四卷，第二期，民國九十八年
15. 薛傑仁，“生物辨識之人臉辨的方法”，民國99年7月
16. 吳明芳、李振興、王炳聰、詹慧珊、黃建邦，“多人臉影像視覺辨識技術”，Journal of Kun Shan University, No.8, pp.75 ~ 95 (June, 2011)
17. 蘇祖澤、蘇信宏、洪志宗，“數位學習之人臉偵測與辨識”，北台灣學報，第31期，民國97年3月。

曾婉菁 / 中央印製廠技研室管理師