

人臉偵測與定位之 Matlab 應用

陳怡靜

國立高雄海洋科技大學

電訊工程系

eoeeoo@yahoo.com.tw

盧傳傑

國立高雄海洋科技大學

電訊工程系

kwaw2152@yahoo.com.tw

張玉雯

國立高雄海洋科技大學

電訊工程系

may@mail.nkmu.edu.tw

摘要

近年來人臉辨識(face recognition)與人臉偵測(face detection)技術被廣泛地使用在各種安全門禁應用之上。本論文收集上百張影像樣本統計 H、S、V 的平均門檻值，利用門檻值的範圍先過濾出該影像中的膚色區塊，再經由二值化分析、侵蝕和膨脹、標記連通法等方法對影像中的膚色區塊做前置處理動作，接著再藉由水平投影及垂直投影的方式明確地自動定位出影像中人臉所在位置；透過本文方法，可以精確找出該影像中屬於膚色人臉的區域，並且可以針對被框選出來的區域做進一步處理及判別等動作。

關鍵詞：人臉偵測，膚色偵測，矩形定位。

Abstract

The technology of face recognition and face detection has been applied to the various access control card widely these years. This research paper has collected more than one hundred image samples of Hue, Saturation, and Value and got the average. Taking advantage of the range of the average can filter the image of the skin color areas. The skin color areas can be dealt with well beforehand through the analysis of Binary, the steps of erosion and dilation, and connected component labeling. Then precisely, it can be located the position of people's faces itself by the ways of vertical projection and horizontal projection. Making good use of the approach in this study, it can be easily found the image that belongs to the areas of the face skin color explicitly; and it will be focused on the part of the frame elects and go a step further of handling and distinguish properly.

Keywords: Face detection. Skin color detection. Rectangular orientate.

1. 前言

人臉偵測是人類與生俱來的視覺能力，現今社會中大多使用鑰匙、密碼，甚至最近幾年相當熱門的 RFID 感應卡，以上這些開啟方式往往因為破解方法簡單或是容易被複製等因素導致容易被盜取或是破解，若是能再多一層防護措施，可以提高自身的生命財產安全；而由於，人類與生俱來的生物特徵是獨一無二的，也不易被複製甚至取代，這些生物特徵，例如：聲紋、指紋、虹膜、人臉…等等。在指紋辨識方面，有容易被複製盜用的缺點。而虹膜辨識技術對於有配戴眼鏡的使用者有或多或少使用上的不便，另外近年來熱門的瞳孔變色片也會影響到辨識的結果。此外，聲紋辨識[6]技術由於人聲也顧慮到容易被盜取錄音備份，所以危險性更高。至於人臉辨識，無需攜帶鑰匙或是其他開啟工具，可以直接使用人臉作為辨識依據，透過每個人所具有的不同臉部特徵作為辨識項目，例如：眉毛、眼睛、鼻子、嘴巴…等五官特徵作為辨識項目，再經由多重確認特徵以提高安全性，因此該方法是最直接地也是最安全的辨識方法之一。

雖然近年來有關人臉偵測、辨識的應用不斷的被研究及探討，但是仍然無法普及，主要的原因是：在複雜的環境及條件下，(例如：光源位置、背景、拍攝角度…等等因素)，將會影響而無法準確地定位出人臉的位置，這也是缺點之一。因此，如果要提高人臉定位的準確率，就一定要先從影像中偵測出最大膚色區塊，接著，在利用後續的辨識及比對再去做人臉之偵測、定位。因此膚色偵測也成為人臉偵測流程中最重要的處理之一。因此本文就是要在一張影像中找出屬於膚色人臉的區域並做自動定位。

人臉偵測中另一個重要的處理就是將人臉與背景做分割。有關膚色分割的方法有很多種[1-2]，不管是在哪一種色彩空間下進行處理，皆必須先定義出膚色區域的門檻值，然而

影響膚色像素值的因素很多，如光線、天氣、背景、角度、人種…等。如果直接在 RGB 色彩空間上做膚色分割，會容易受到光源強弱的影響導致分割結果有很大的誤差。在簡逸聰[5]與 Du Xingjing, Liu Chunmei and Yu Yongxia[3]中，都是在 YCbCr 的色彩空間上做膚色分割，然而 YCbCr 較適合用在數位視訊中，而本實驗所用的測試影像是以數位相機拍攝完成的數位影像，因此較不適用 YCbCr；再加上影像容易受到光線強弱的影響。

所以，在本文中是利用 HSV 色彩空間[4]的特性，也就是將 H(Hue)做為偵測膚色區域或人臉的主要條件，而且將亮度分離出來，也能減少光源對影像本身的影響，並透過 100 張測試影像取得色調 H(Hue)以及飽和度 S(Saturation)作為膚色門檻值，接著以所得到的膚色門檻值作為膚色分割之判斷條件，並明確地將膚色區域與背景分離；再進行影像前處理，以便將將影像最大膚色區塊透過垂直及水平投影之方法，來自動定位出影像中膚色區塊。

2. 基本原理

2.1 HSV

HSV 是人們從色彩輪或色盤中選擇色彩時所使用的幾個色彩系統之一。HSV 色彩系統是以圓柱形座標為基礎，從 RGB 轉換為 HSV，只是一個發展公式的問題而已，它將 RGB 數值映射至圓柱形座標，此色彩系統比起 RGB 系統更接近於人類所體驗與描繪的色彩感受。其中，色調(Hue; H)是指色彩的外相，涵蓋了所有飽和度、明度為 100%的顏色，色調可以組成一個環狀的色環，所以色調通常是以 $0^\circ \sim 360^\circ$ 來表示的，其中 0° 為紅色、 120° 為綠色、 240° 為藍色。

飽和度(Saturation; S)是指這個顏色的濃度設定值為 0%，代表最不濃或完全稀釋；100%代表最濃或完全不稀釋；越高則越接近純色，越低則越接近白色。

亮度(Value; V)可想像成顏料的深淺，當設定值為 0%表示最暗；100%表示最亮；越高越接近真實的純色，越低則越接近黑色。

其 RGB 與 HSV 的轉換公式如下：

$$H_1 = \cos^{-1} \left\{ \frac{0.5[(R-G) + (R-B)]}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \right\}$$

$$\begin{cases} H = H_1 & \text{if } B \leq G \\ H = 360^\circ - H_1 & \text{if } B > G \end{cases}$$

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)}$$

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

2.2 二值化

二值化影像(Binary image)簡單來說，影像的像素值只會有黑與白的組合，換句話說也就是一個 0 和 1 的陣列，1 代表白色，0 代表黑色。在影像處理中較容易處理、儲存以及分辨，因此二值化影像廣泛用於影像辨識處理及形態學中，以進行影像分割、膨脹侵蝕、細線化、邊緣檢測…等用途。

主要是利用影像資料中像素值的差異特性，以一適當的灰階門檻值(Threshold)當成轉換影像資料的依據，以便將影像資料轉換成二元影像。常用的二值化公式如下：

$$W = \begin{cases} 0, & w \geq F_0 \\ 1, & w < F_0 \end{cases} \quad W = \begin{cases} 0, & x \leq F_0 \leq y \\ 1, & \text{other} \end{cases}$$

其中， w 為像素值， F_0 為灰度門檻值， W 為轉換後的像素值， x 、 y 為門檻的上下限值。

2.3 膨脹與侵蝕

膨脹(Dilation)和侵蝕(Erosion)算是影像形態學裡很重要的基礎。一般來說，膨脹和侵蝕通常會被組合起來使用，一個影像會用相同或是不同的結構元素來進行一連串的膨脹及侵蝕處理。分別說明如下：

2.3.1 膨脹

膨脹(Dilation)是使二值影像中的物體「增大」或是「變厚」，而膨脹的方式和程度是由一種稱為結構元素(Structuring element)的形狀所控制。假設 X 和 Y 為 Z 中的兩個集合， X 藉由 Y 的膨脹(Dilation)記為 $X \oplus Y$ ，定義為：

$$X \oplus Y = \left\{ Z \mid (\hat{Y})_Z \cap X \neq \phi \right\}$$

其中 X 為原始影像， Y 為結構元素 (Structuring element)， $(\hat{Y})$ 為 Y 映射或反射，也就是說將 Y 對它的原點進行反射。

2.3.2 侵蝕

侵蝕(Erosion)使二值影像中的物件「收縮」或是「變薄」，其侵蝕方式和程度也是結構元素(Structuring element)的形狀所控制。對於Z中的集合X與Y，X被Y的侵蝕(Erosion)記為 $X \ominus Y$ ，定義為：

$$X \ominus Y = \{Z \mid [(Y)_z \subseteq X]\}$$

侵蝕的目的是，可除去二元影像中不相干的細節。

2.4 標記聯通法

標記聯通法(Connected component labeling)的作法是將影像中屬於同一物件的相連像素找出並從左上方開始沿著物件邊緣掃描以得到物體的高度、寬和面積等資訊，並藉此標訂出影像中，各個獨立之物件。當遇到像素值為1時，則依序檢查其上、下、左、右是否有同為1的點，如圖1(a)所示，並將其座標紀錄下來，然後把其值設為0，之後依序遞迴檢查剛剛所記錄點之下、左、右、上是否有1的點存在，若有也將其座標紀錄下來，並將其值設為0，持續整個遞迴的檢查，直到所有記錄點之周圍都不再有1的點出現為止，即可停止遞迴檢查之動作，這樣就可以找出第一個群組的物件，依此類推，當整張影像都掃描遞迴後，我們就可以統計出，影像中所有物件之高度、寬和面積等資訊。

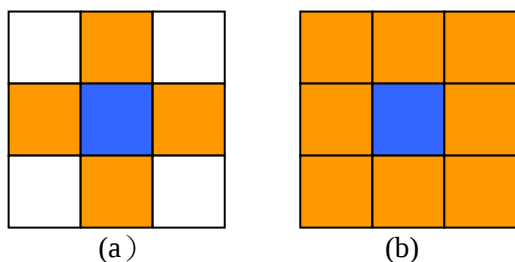


圖 1 (a) 為四聯通 (b) 為八聯通

3. 研究方法

本方法之流程圖如圖2所示。首先將原始影像由RGB色彩空間轉換到HSV色彩空間；另外，利用本系統所建置的測試影像資料庫來統計膚色門檻值、平均值，如表1、2所示。以這些門檻值作為膚色檢測的標準值以便將測試影像中膚色區域及背景做分離。接著，對影像進行一連串前處理的動作，如二值化處理使影像轉成只有黑與白的二元影像，也就是影

像背景區塊會呈現黑色，反之膚色區塊會呈現白色，利用這樣的方式可以使影像更容易做後續處理。接著將上述的二值化影像做侵蝕再膨脹動作，可以消除影像中不相連的細節及平滑影像的輪廓。再利用標記聯通法過濾出最大膚色區塊。最後影像分別利用水平投影以及垂直投影，透過座標軸上所聚集的像素(pixel)多寡來找出人臉的所在座標位置，以達到自動找出影像中膚色區域的座標，並將之對映上，就能使用矩形外框來達到自動定位出影像中人臉所在之區域。

表 1 膚色平均值

	H	S	V
平均值	0.0721	0.4	0.749

表 2 膚色門檻值

	H	S	V
門檻值	0.0221 至 0.5221	0.15 至 0.65	0.419 至 0.879

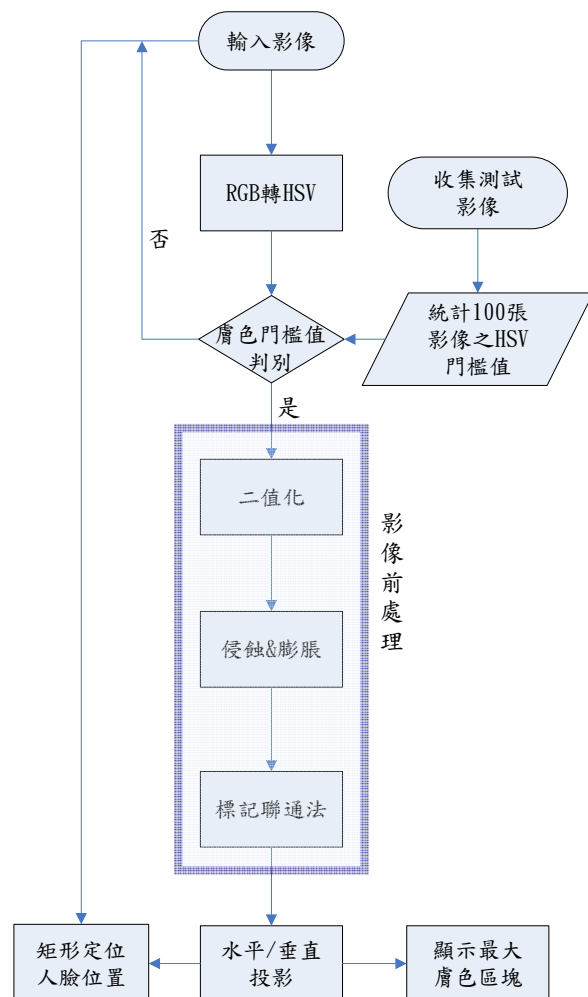


圖 2 人臉偵測流程圖

4.模擬結果與討論

本系統使用的軟硬體條件如下：

軟體：

模擬軟體：Matlab 6.5

作業系統：Windows XP

硬體：

CPU：Intel core 2 Quad Q6600

記憶體：1GB

測試影像：

自行拍攝之相片，所拍攝的影像人物不分男女、拍攝人數、被拍攝者之衣著、拍攝地點（室內或室外），而拍攝方式也不侷限固定腳架、手動拍攝、拍攝角度…等條件；在本實驗中所有測試影像是由 Sony、Nikon 及 Fuji 相機所拍攝，影像解析度 Sony、Fuji 相機解析度為 1280*960，則 Nikon 相機之解析度為 2112*2816 及 2048*1536 兩種。



圖 3 測試影像

由使用者自由選取任意影像以進行測試，如圖 3 所示。再經由表 1 所統計出的 H、S、V 平均值，作為膚色分割的門檻值來進行膚色分割，如圖 4 所示。本研究統計出 H、S、V 門檻值如表 2 所示。



圖 4 膚色切割後影像



圖 5 二值化後影像

利用二值化將灰度介於 0~255 之間的像素，以 0 和 255 來表示，主要目的在使物件與背景能有黑白明顯的對比如圖 5 所示。在本論文中統計水平及垂直投影也是利用二值化的特性來分析像素值多寡。



圖 6 膨脹再侵蝕後影像



圖 7 標記聯通後之影像

侵蝕是將該點像素之灰度值以該點像素與其八鄰接之灰度中之最小值來取代，當影像物體為亮點時，執行此運算將會造成侵蝕縮小之結果。

膨脹則是以該點像素與其八鄰接之灰度中之最大值來取代該點像素值。當影像物體為亮點時，執行此運算將會造成膨脹放大之結果，如圖 6 所示。

透過標記連通法先由左而右，再由上而下掃描影像，在此只考慮四連通分量，結果如圖 7 所示。

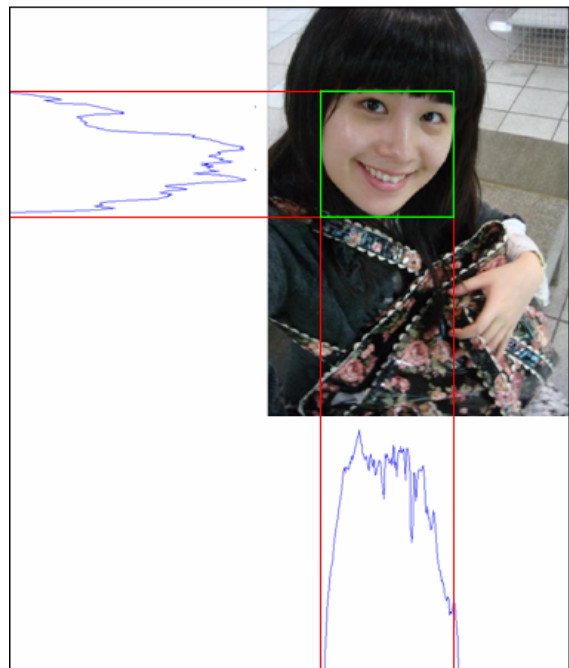


圖 8 水平垂直投影映色後影像

最後，利用做水平、垂直投影法，透過縱座標以及橫座標上像素數目累積多寡以一矩形外框再對映到原始影像上進行自動人臉定位，如圖 8 所示。

在本文中收集的五百張數位影像，所有測試影像皆不侷限任何地點、背景、人數、穿著、天氣、光線、角度甚至被拍攝者有無化妝…等其他的外在參考因素。

綜合上述條件，將影像經由膚色分割後再進一步對影像作前處理，最後成功偵測出最大膚色區塊之人臉位置並且以矩形外框自動定位出人臉位置之影像為 127 張。依據影像測試之條件，光線條件、人數多寡及拍攝距離來進行分類如表 3、4 和 5 所示，進而統計實驗結果如表 6 所示。

表 3 影像測試之光線

		光線充足	光線不充足
光線條件	室外	陽光充足環境下	有遮蔽物、天氣昏暗之條件下
	室內	日光燈照射充足	燈光昏暗之場所

表 4 影像測試之人數

	單人	多人
人數多寡	一人為主	兩人以上包含兩人

表 5 影像測試之拍攝距離

	大頭照	半身照	全身照
拍攝距離	被拍攝者肩膀以上之區塊佔影像約 80% 為主	被拍攝者腰部以上到頭部之區塊佔滿影像約 80% 為主	被拍攝者從頭到腳之區塊皆在影像中

表 6 實驗結果統計

人數	照片種類	光線充足	光線不足
單人	大頭照	26 張	8 張
	半身照	43 張	18 張
	全身照	7 張	3 張
多人	半身照	9 張	6 張
	全身照	4 張	3 張

由表 6 實驗結果 89 張得知，在光線充足、半身和單人之條件其成功的比率高於其他條件；另外在本實驗結果中發現拍攝場景光線條件對於結果成功率有很大的影響。

然而在多人之條件下不論是在光線充足或是拍攝距離遠近之下，所成功偵測出人臉之

機率皆不高，從中判斷多人的情況所影響的外在因素太多容易導致誤判而無法精確辨別出膚色區塊。

5. 結論

在本次的研究中，利用 HSV 門檻值作為最大膚色區塊的辨識方法來選取人臉最大的區塊，再利用矩形外框做臉部的定位，來達到人臉偵測的效果，本研究中沒有限制測試照片的條件，例如：光源、人數、地點、角度、服裝…等外在因素，皆使用日常生活所拍攝的普通生活照片為主，因此準確率與有限制照片條件之準確率相比並不是非常的高；若想要提高準確度，可以再配合利用其他細微的條件來做比對，如對影像進行臉部模板比對。也可透過男女不同的臉部特徵，例如：鬍子、毛孔、頭髮…等條件來判別性別，甚至透過肌膚紋理的特徵來判別年紀，最後再進行五官切割個別比對眼鼻口，當增加越多條件來做比對，其準確度相對會大大提升。

參考文獻

- [1] 繆紹綱，數位影像處理-運用 Matlab，民國 94 年初版。
- [2] 石福僑，“一個使用環場攝影機及多重辨識器的跌倒偵測系統”，中原大學，民國 96 年。
- [3] Du Xingjing^{1,2} , Liu Chunmei² ,and Yu Yongxia² ,“Analysis of Detection and Trackon Partially Occluded Face”,¹School of Information Science and Technology, Beijing Institute of Technology(BIT),Beijing 100081, China,and ²Department of Computer, North China Institute of Science and Technology(NCIST), Beijing East Yanjiao 101601,China,2009。
- [4] Hassan Fahmy Hashem, “Adaptive Technique for Human Face Detection Using HSV Color Space and Neural Networks”, Alexandria High Institute of Technology, March 17-19,2009, Faculty of Engineering, Future Univ., Egypt。
- [5] 簡逸聰，“一個以膚色為基礎之互補人臉偵測策略”，國立中央大學資訊工程系碩士論文，民國 90 年。
- [6] 陳昱達，“語音辨識應用於保全系統之密碼判別研究”，樹德科技大學資訊工程系碩士論文，民國 93 年。