

第一章 緒論

第一節 研究背景與動機

在宵小盛行的現在，監視系統守護著你我的安全，但發生事故去查看監視器時，往往令人失望。假使監視系統可以拍攝到清晰的照片，我們可以加入人臉辨識系統，那麼此等狀況便會大幅獲得改善。建立社區資料庫，非本社區的人員，系統便會知曉。個人財產若不幸遭到破壞，我們便可清楚的看到犯人，以利警方辦案，加速破案效率。

事實上現在市面買的相機，就有許多內建人臉對焦功能的，使許多不懂攝影原理的人，也可以輕鬆的拍出好照片，保留美好回憶。科技始終來自於人性，兩者的基本原理並無不同，重點皆是人臉偵測，如何正確無誤的判斷，自然就是我們努力研究的目標。

人臉偵測受到膚色、環境等影響很大，一個好的軟體，自然正確性、穩定性要高。利用色彩模型、形態學等技術再進一步幫助我們擷取更加完整的人臉。其中可以學習的東西有很多，如果我們可以從中學到些什麼？對我們自己能力上的進步，我想這對我們才是最大的幫助。

第二節 研究重點

從照片中把人臉找出來，以人類的眼睛來觀察，自然是很容易的。我們可以很輕鬆的辨識出，這是誰，天氣如何，風景等。但是要把照片讓電腦辨識就不是那麼簡單了！程式會受影響的因素有很多，如光影、眼鏡、膚色、背景複雜程度、臉的方向等問題，都會造成不同程度的誤判，要如何成功的辨識出整張照片中的人臉，這當然便是我們研究的重點。臉部偵測目前主要研究有下列四種方法[1][2][3]：

(1) Knowledge based methods (認知為基礎的方法)

人有兩個眼睛，一個鼻子，嘴巴，此方法便是運用人臉五官特點的相對關係去訂出一些規則，如眼睛與嘴巴呈現一個倒三角形、人臉的長寬比例等等，利用這些淺而易見的規則來搜尋人臉。

(2) Feature invariant approaches (特徵為基礎的方法)

人類臉部區域大致上的特定器官，如眼睛、鼻子、嘴巴、鼻子及頭髮等等，這部分，會因解析度及表情變動受到一定程度之影響。利用這個方法可以解決亮度與人臉的方向性的干擾。

(3) Template matching methods (樣板為基礎的方法)

以樣板為基礎的方法必須要事先定義好標準樣板，然後計算輸入影像與樣板之間的相識度，作為判斷影像是否有人臉的依據。此方法

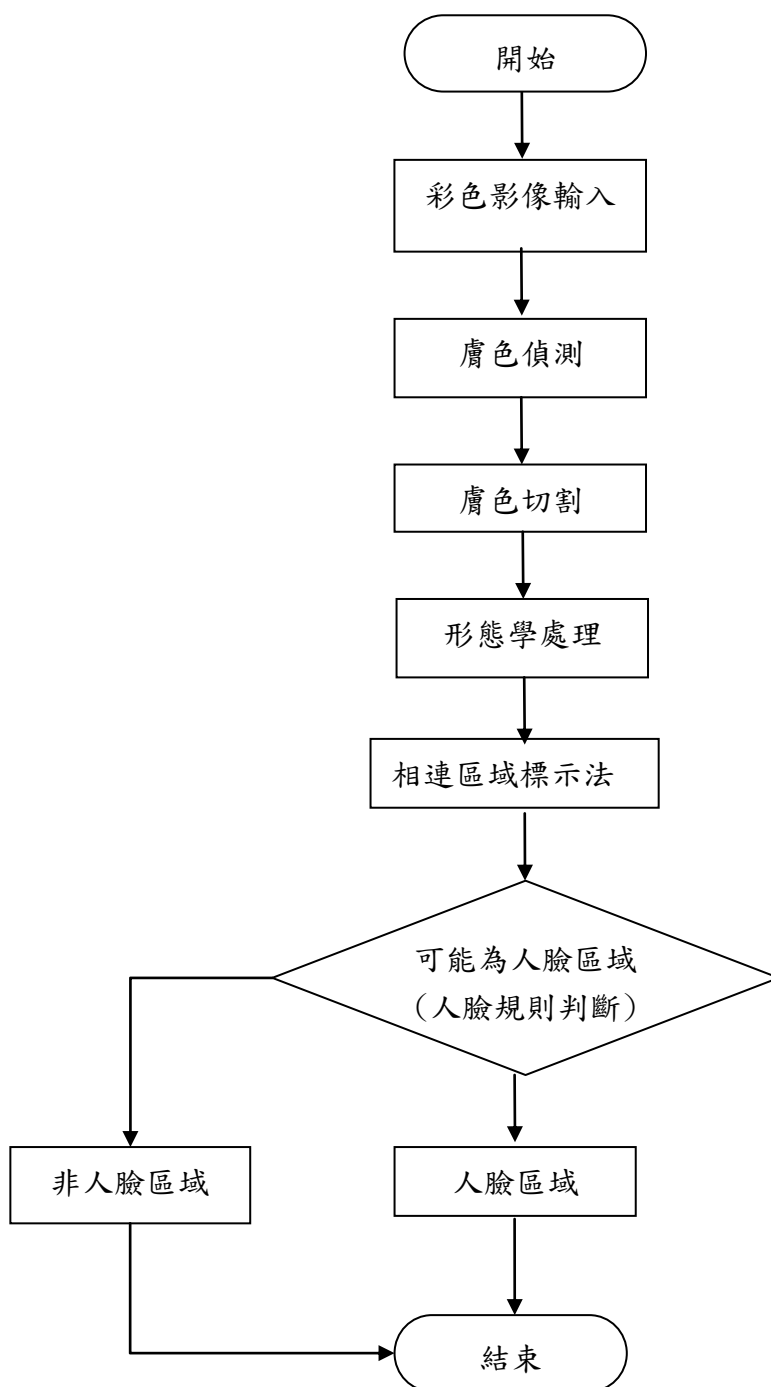
常對整張照片做地毯式掃描，利用已設定好的樣板來判斷圖形是否為人臉。

(4) Appearance based methods (顯性特徵為基礎的方法) [4]

主要是建立、訓練一組臉部資料庫，學習不同人臉的特徵，利用統計分析和機器學習，來判別人臉的存在。這類方法的效能會受到樣本訓練好壞的影響。

第三節 系統流程

本系統以處理彩色影像為主，由於彩色影像中，膚色資訊往往是判斷人臉的基礎偵測方式。在多人且複雜的影像中，人臉的區域位置是較為難偵測的，如何正確定位出人臉位置與判斷人臉的方法，以下系統流程圖的一些研究方法加以說明。



第二章 人臉偵測之相關研究

第一節 色彩模型

2.1-1 RGB 色彩空間：

人眼所見各種色彩是因為光線有不同波長所造成的，經實驗發現，人類肉眼對其中三種波長的感受特別強烈，只要適當調整這三種光線的強度，就可以讓人類感受到『幾乎』所有的顏色。這三個顏色稱為光的三原色(RGB)，就是紅(red)，綠(green)，藍(blue)。所有的彩色電視機、螢幕都具備產生這三種基本光線的發光裝置。因為這三種光線的混合幾乎可以表示出所有的顏色，因此電腦裡頭就用 RGB 三個數值的大小來標示顏色，每個顏色用 8bits 來記錄，可以有 0~255，共 256 種亮度的變化，三種乘起來就有一千六百多萬種變化，這也是我們常聽到的 24 bit 全彩。R.G.B.色彩系統為一直角座標系統，也是最常用的顏色系統，它的三個分量是選擇了可見光中三原色光，為了使色光的定義能夠統一，C.I.E.協會定出三原色光的波長紅 (Red)、綠 (Green) 和藍 (Blue) 為此色彩系統。R.G.B.色彩系統主要為加色法的三度色彩空間，可使用不同強度的紅、綠和藍色光強度來組成各式各樣的色彩。示意圖如圖 1：

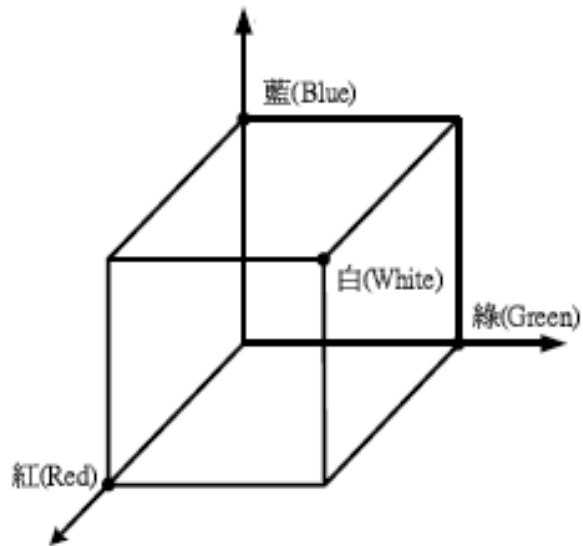


圖 1. RGB 色彩模型

2.1-2 YcbCr 色彩空間[5]：

Y 即是由彩色轉換成灰階影像的灰階值或稱之為亮度值，而其轉換公式主要是依據人類的眼睛對於紅、綠、藍三原色的不同敏感度而來，其中係數值越大則表示人的眼睛對於該顏色較為敏感，所以三種顏色之敏感度依序為綠色(0.587)、紅色(0.299)、藍色(0.114)。

RGB 和 YCbCr 的關係如公式(1)：

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

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 0.00456621 & 0 & 0.00625893 \\ 0.00456621 & -0.00153632 & -0.00318811 \\ 0.00456621 & 0.00791071 & 0 \end{bmatrix} \begin{bmatrix} Y - 16 \\ Cb - 128 \\ Cr - 128 \end{bmatrix}$$

(1)

我們將值簡化，RGB 與 YCbCr 之間的轉換如公式 (2)：

$$Y = 0.2990R + 0.5870G + 0.1140B$$

$$Cb = -0.1687R - 0.3313G + 0.5000B + 128$$

$$Cr = 0.5000R - 0.4187G - 0.0813B + 128$$

$$R = Y + 1.40200 (Cr - 128)$$

$$G = Y - 0.34414 (Cb - 128) - 0.71414 (Cr - 128)$$

$$B = Y + 1.77200 (Cb - 128)$$

(2)

RGB 系統轉換至 YCbCr 空間下將 CbCr 平面切割成四個區域，如圖 2 所示。因此依據此圖判定紅色(red)的位置坐落在區域”I ”與區域”II ”之間；黃色(yellow)的位置坐落在區域”II ”與區域”III ”之間；綠色(green)的位置坐落在區域”III ”上；青色(cyan)的位置坐落在區域”III ”與區域”IV ”之間；藍色(blue)的位置坐落在區域”IV ”與區域”I ”之間。因此僅須找出各個區域下，不同亮度 Y 與色度 CbCr 之間的關係式即可。

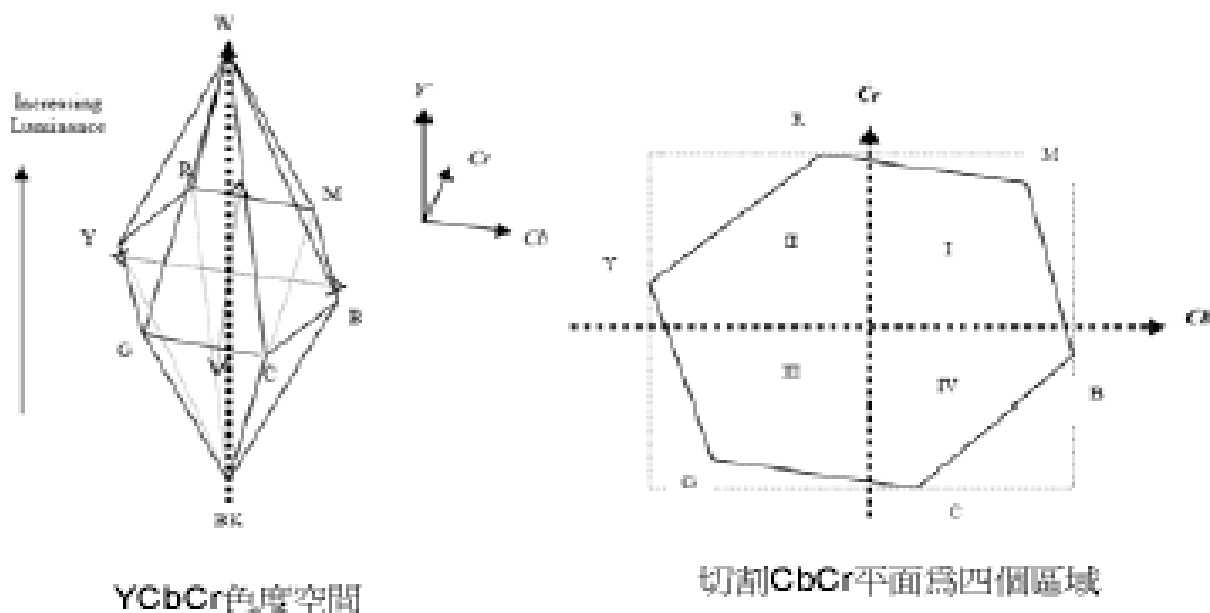


圖 2. CbCr 平面切割成四個區域

2.1-3 HSV 色彩空間[6]：

HSV 利用色度 (Hue)、飽和度 (Saturation)、亮度 (Value) 來區分膚色及非膚色區域，相較於YCbCr 色彩空間對於各種人種膚色偵測上有較準確的偵測結果，不過HSV 的轉換上並不是像YCbCr 是線性轉換，HSV 轉換比較消耗運算時間。

將 RGB 色彩空間轉為 HSV 色彩空間，轉換公式如下：

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

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

$$V = \frac{\text{Max}(R, G, B)}{255} \quad (3)$$

膚色區域將會落在下列的區間：

$$0 \leq H \leq 0.1388$$

$$0.23 \leq S \leq 0.68$$

$$0.35 \leq V \leq 1$$

對影像內像素由RGB 色彩空間轉換成HSV 色彩空間，比對像素中HSV 數值是否落在設定的膚色區間，若符合則判定為初步膚色區域，其餘則為非膚色區域。

圖3為原始影像，該圖經膚色偵測後，結果為已能除去大部分非膚色區域，非膚色區域以白色像素來表示，膚色區域以原始數值顯示。但仍可觀察到初步膚色區域仍不十分準確，所以有必要再針對初步的膚色區域再作分析，取出更準確的膚色區域。



a.原圖



b.初步膚色區域

圖3. 膚色區域偵測

2.1-4 YUV 色彩空間[7]：

YUV 色彩空間用於歐洲的電視系統上，屬於 PAL (Phase Alternation Line) 系統，其 Y 表示為顏色的明亮度 (Luminance)，U 和 V 是指色調(Chrominance) 表示顏色的差異訊號，包括黃色藍色和紅色的差異，YUV 很適合用於電視工業 因為他有效的應用頻寬，是因為人類對於亮度的感覺比顏色的些微改變來的敏感，所以減少顏色的頻寬而增加亮度的頻寬。而 YUV 與 RGB 之間的對應關係 為：

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.148 & -0.289 & 0.437 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0.114 \\ 1 & -0.395 & -0.581 \\ 1 & 2.032 & 0 \end{bmatrix} \begin{bmatrix} Y \\ U \\ V \end{bmatrix}$$

(4)

第二節 人臉偵測四種方法[1][2][3]

2.2-1 Knowledge based methods (認知為基礎的方法)

在這個方法，運用人臉特徵的既定認知，來判斷人臉的存在。例：兩眼的對稱性，利用兩眼中間判斷出中線，接著尋著中線，便可找到鼻子及嘴巴。利用這些規則，便可迅速找出人臉及相對位置。

不過在這個方法下有個很大的問題，因為將人臉描述的非常清楚，可能幾乎所有人臉皆不能通過此規則；描述的模糊，則有可能判斷失誤，出現非人臉的區塊。在不同的姿勢、光影、環境下，有許多無法預測的問題，因此這個方法較適合運用在朝向正面的人臉。

2.2-2 Feature invariant approaches (特徵為基礎的方法)

人臉的特徵，除了五官的特徵外，整體的人臉、光線及色彩的部分也是特徵之一。雖然膚色是人臉偵測方法中算是最有效的一個特徵，不過仍然需搭配標準的演算法，考量處理不同的狀況，將一連串的數據透過轉換及聯結，得到最後結果，以提高人臉搜尋的正確率。

許多文獻提出不同的色彩空間來切割膚色，最常使用的則是 RGB 色彩空間，其好處則是運算快速，無須計算座標轉換，其缺點為偵測結果會受到環境、光源的影響；為了解決此種影響學者提出色彩空間的轉換，雖然膚色在色彩空間中佔有很大範圍，但其色調及飽和度上的變化不如亮度，因此 HSV、YCbCr 及 YUV 便廣泛應用在膚色偵測上。

2.2-3 Template matching methods (樣板為基礎的方法)

這個方法是用人臉進行分析過後做成不同的標準樣板，然後將輸入的影像與樣板計算之間的關係值(Correlation Value)，作為判斷是否有人臉或部分臉部的依據，然後再將關係值作為偵測及定位用途。

然後藉著人工定義的一組標準臉部樣板或是就藉由函式產生樣板參數，輸入的影像與標準樣板針對臉部外形(Contour)、眼睛、鼻子、嘴巴等分別計算出關聯值(Correlation Value)，藉由關聯值來決定臉部的存在。

這種方式的優點就是很容易完成，不過該方法已經被證明不足夠去偵測變異度臉的存在性[8]，因為他不能很有效控制變量，比如說 比例(Scale)、形狀、姿勢。

2.2-4 Appearance based methods (外表為基礎的方法) [1][4]

以外表為基礎的方法通常是建立起一組人臉資料庫，利用機器學習和統計分析等方法，訓練程式來分析人臉及非人臉。

統計分析所使用的方法，主要可區分為兩個部份，一個是特徵擷取的統計方法，如：主成份分析。另一個是判別分離的算法上採用統計模型[1]，如：SVM、隱藏馬可夫模型、支援向量機等。

第三節 相關研究論文探討

2.3-1 馬炳義—人臉偵測技術研究[8]

在人臉偵測技術研究文中，膚色偵測採用 YCbCr 色彩模型，著眼點在於運算及處理上較簡單。經過直方圖統計後，發現膚色範圍值分別落在 $R_{Cb}[96,131]$ ， $R_{Cr}[131,173]$ 。

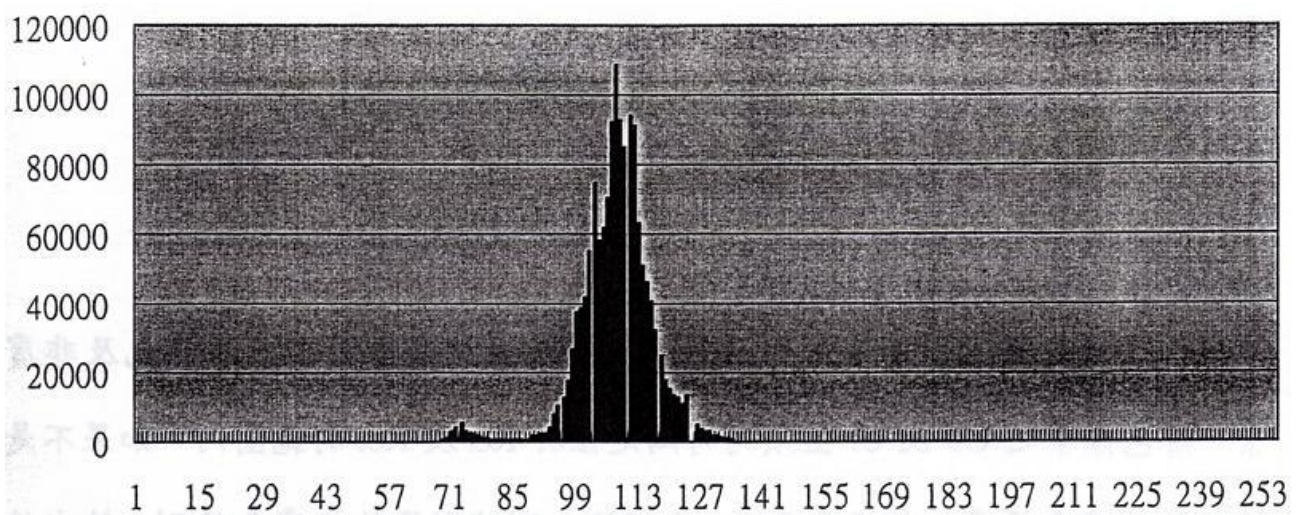


圖 4-1. Cb 統計直方圖

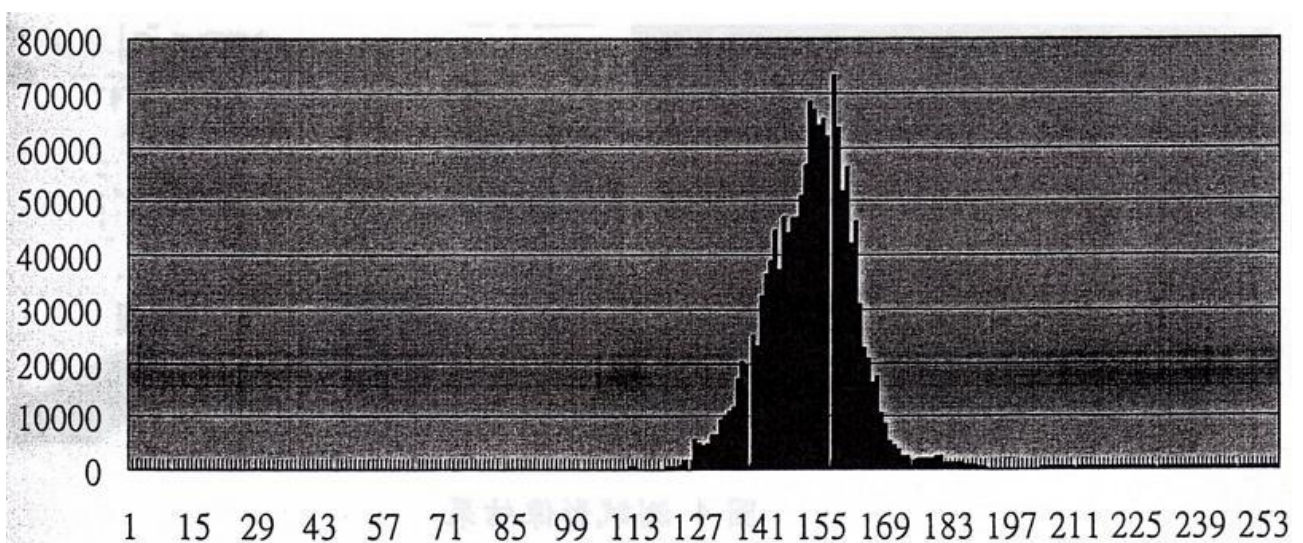


圖 4-2. Cr 統計直方圖

統計完膚色分佈後，接下來依據分佈圖對影像進行色彩分離的步驟，僅針對影像彩度成份處理，演算法如下：

$$O(x, y) = \begin{cases} 1, & \text{if } [Cr(x, y) \in R_{Cr}] \cap [Cb(x, y) \in R_{Cb}] \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

$O(x,y)$ 為像素 (x,y) 輸出的影像，經轉換後分類為膚色及非膚色像素，膚色像素之 Cb 及 Cr 值須同時滿足 R_{Cb} 及 R_{Cr} 的範圍內，如果不是則歸類為非膚色像素，並將值設定為 0，輸出結果如圖 5 所示。



圖 5. 色彩分離

輸出之影像圖，除了人臉區域外，還包括了雜訊，可能出現在臉部非膚色區塊或背景影像中。我們可以利用膨脹的方式填補在臉部區域非膚色的雜訊，或是利用侵蝕的方式移除背景雜訊。當然，這種方式不可能完全的去除雜訊，不過可以大幅降低雜訊的數量及大小。

我們將影像區分為人臉及非人臉區域，利用密度計算的方式來消除雜訊。膚色像素出現在背景時，不是以大的聚集方式呈現，要不然就是小的獨立物件方式。而出現在臉部區域時，則是以較大的聚集方式呈現，因此利用這種特性來計算膚色密度的分佈狀況，演算法描述如下：

$$D(x, y) = \sum_0^3 \sum_0^3 O_1(x, y) \quad (6)$$

其中 $x=0, y=0$ ， $O_1(x, y)$ 不重疊的 4×4 像素， M 及 N 代表影像的寬度及高度。在這個 4×4 區塊中，每個區塊內所累加的膚色像素點數，我們在陣列中將對應三個類別：

第一個類別為 0 類別 ($D=0$)，這代表沒有臉部區域存在，在測試中我們以黑色為代表；第二個類別中間過度類別 ($0 < D < 16$)，這代表該區域可能屬於雜訊，測試以紅色為代表；第三個稱為完全類別 ($D=16$)，在這個類別有極高的機率屬於臉部區域，測試結果中以白色為代表。

接下來我們將非完全類別做二值化處理並將影像像素侵蝕，演算法如下：

$$O_2(x, y) \begin{cases} 1, & \text{if } D(x, y) = 16 \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

在區塊中有可能為人臉的部份，我們將他填入白色做置換，而其他沒有人臉或雜訊的部份便把他填入黑色，好方便我們做後續的處理。

在人臉偵測技術研究本文中提出，採用 Entropic thresholding 的技術可以很容易發現適合影像內容的最佳門檻值「T 值」，藉著 T 值的選擇可以執行非邊緣及邊緣像素類別計算 Entropy「熵」總和，滿足最大的資訊量。演算法說明如下：

$$H(T) = \max_{T=0,1,2,\dots,M} \{Hn(T) + He(T)\} \quad (8)$$

首先假設邊緣像素強度值它的範圍在 $[0, M]$ ，輸入的影像共有 f_i 個像素，邊緣像素強度值 $i \in [0, M]$ ，假定門檻值 T 代表邊緣像素與非邊緣像素分佈，機率的分佈視為獨立，非邊緣像素機率 $Pn(i)$ 定義如下：

$$Pn(i) = \frac{f_i}{\sum_{h=0}^M f_h}, 0 \leq i \leq T \quad (9)$$

$\sum_{h=0}^T f_h$ 代表所有區域最大邊緣強度特徵落於邊緣強度落於 $0 \leq i \leq T$ 之間所有的像素數量。同理，我們設 $Pe(i)$ 為邊緣的機率，定義如下：

$$Pe(i) = \frac{f_i}{\sum_{h=T+1}^M f_h}, T + 1 \leq i \leq M \quad (10)$$

$\sum_{h=T+1}^M f_h$ 代表所有區域最大邊緣強度特徵落於邊緣強度落於 $T + 1 \leq i \leq M$ 之間所有的像素數量。對於這兩個類別的 entropies 定義如下：

$$Hn(T) = -\sum_{i=0}^T Pn(i) \log Pn(i)$$

$$He(T) = - \sum_{i=T+1}^M Pn(i) \log Pn(i) \quad (11)$$

我們將 $Hn(T)$ 與 $He(T)$ 所得值帶回公式 (8)，藉由函式整體最大值結果如圖所示。

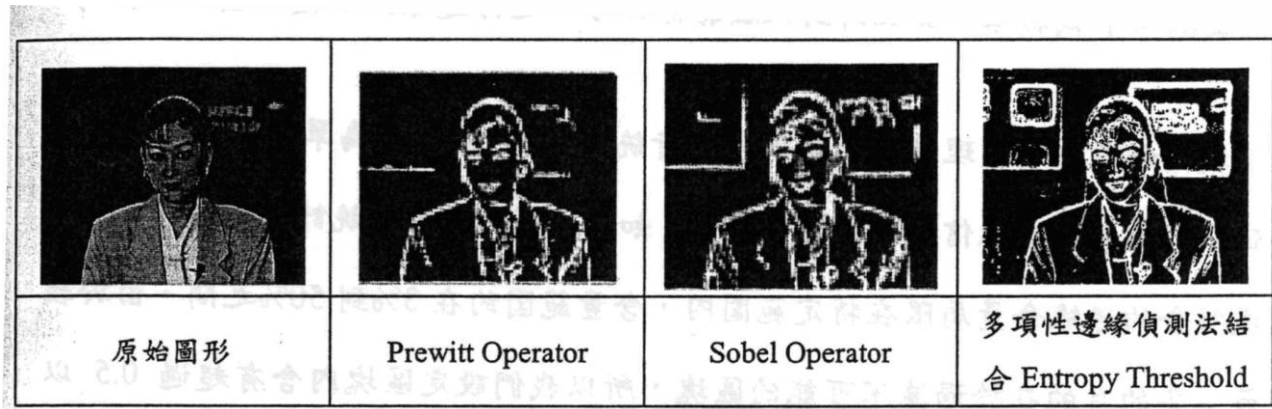


圖 6. 利用 Entropy 取 T 值

2.3-2 吳宣諭—運用膚色過濾的人臉偵測法[1]

本論文所採用的人臉偵測技術為 YCbCr 色彩空間模型，著重於其公式簡單(完全矩陣運算)、運算簡單(有助於硬體化的實現)、廣泛使用(應用於 JPEG 影像壓縮、MPEG、數位相機規格等)，在 YCbCr 中膚色分布有不錯的集中效果，且各色人種的膚色分布差異並不會相差太多。

探討 YCbCr 相關優點後，本論文還利用形態學(Morphology)的基本運算相關技術，類似於影像處理的空間域濾波器，其相關運算及定義如下：

◇ 膨脹(Dilation)：

A 透過 B 的膨脹運算，記為 $A \oplus B$ ，定義為

$$A \oplus B = \left\{ Z \mid \left[(\hat{B})_z \subseteq A \right] \right\} \quad (12)$$

Z 表示空間的位置，A 藉由 B 的膨脹運算是表示 Z 的所有位移的集合。圖形有破碎邊緣或是掃描下來的文字有些許不清楚時，可以使用膨脹運算將其斷開部分接回來。

◇ 侵蝕(Erosion)：

A 透過 B 的侵蝕運算，記為 $A \ominus B$ ，定義為

$$A \ominus B = \left\{ Z \mid (B)_z \subseteq A \right\} \quad (13)$$

侵蝕的應用是去除二值化影像中不相干的細節，簡單說，就是一種消除雜訊演算法，將影像中的雜訊點過濾掉，留下較大的區塊。

◇ 斷開(Opening)：

集合 A 被結構元素斷開，記為 $A \circ B$ ，定義為：

$$A \circ B = (A \ominus B) \oplus B \quad (14)$$

集合 A 被 B 的斷開運算就是 A 先被 B 侵蝕，再將結果對 B 膨脹。

◇ 閉合(Closing)：

集合 A 被結構元素斷開，記為 $A \cdot B$ ，定義為：

$$A \cdot B = (A \oplus B) \ominus B \quad (15)$$

集合 A 被 B 的閉合運算就是 A 先被 B 膨脹，再將結果對 B 侵蝕。

◇ 標記化

像素間的聯結性是許多數位影像中的邊界及區域的一種簡化概念，利用此種方式，判斷兩個像素是否相連，本論文是以 3X3 的範圍考慮圖形的關係，如圖，X 代表其中一個 3X3 的遮罩，而 $X_0 \sim X_7$ 代表周圍和 X 的位置關係。

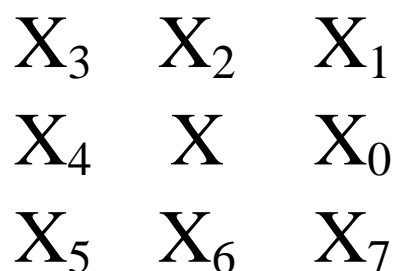


圖 7. 3X3 的遮罩示意圖

在座標(x,y)中有四個垂直和水平方向的相鄰像素點，座標分別為(x+1,y)、(x-1,y)、(x,y+1)、(x,y-1)，這稱為像素四相鄰的點集合。

對於一張二值化影像連通成分是很有用的拓撲性質，當知道二值化圖形的相鄰關係後，希望將區域的關係標示出來，此時必須使用到連通元件標記化(Connected Component Labeling)的方法，其示意圖如下：

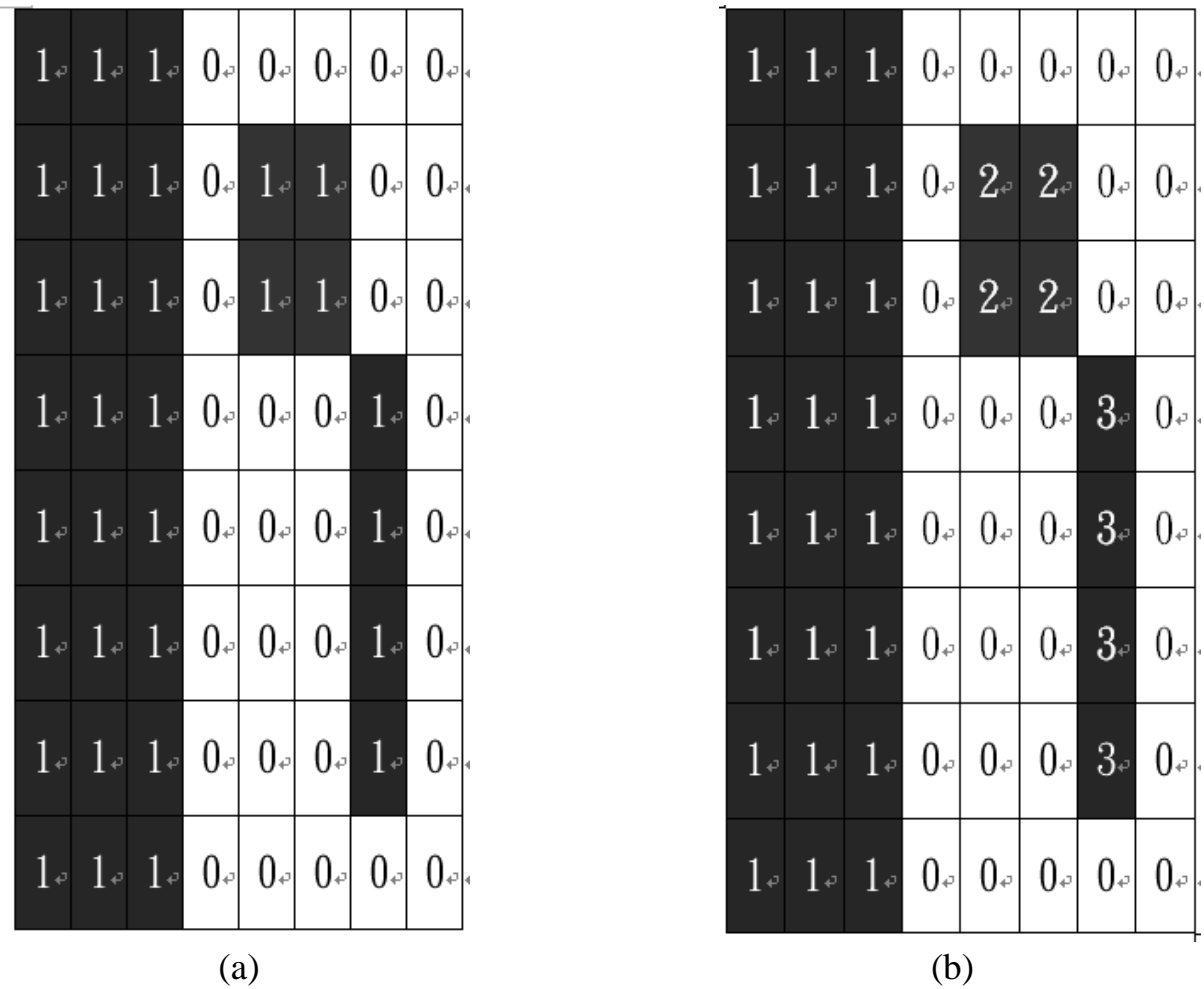


圖 8. 區域標記化示意圖

由圖我們可發現，原本二值化影像，在經過標記化（採四連通）後分成了三個不同的區塊，各區塊中標上了號碼，表示為其各分屬的區塊，在分析過程中便可以針對各個區塊作運算。

本論文利用了集合膨脹(dilation)、取補集 (complex) 以及交集之區域填充的演算法，將二值化圖形區域中，含有空洞的區塊填滿，此方法稱之為區域填充(region fill)其數學的表示方法如下：

$$X_K = (X_{K-1} \oplus B) \cap A^C, K = 1, 2, 3, \dots$$

其方法是設定一個對稱結構元素 B ，假設 $X_K = X_{K-1}$ ，則在疊代到第 K 次時終止， X 會被 B 集合向四連通方向膨脹，不過因為有和 A^c 取交集，所以不至於擴大至整個區域。其演算法示意圖如下：

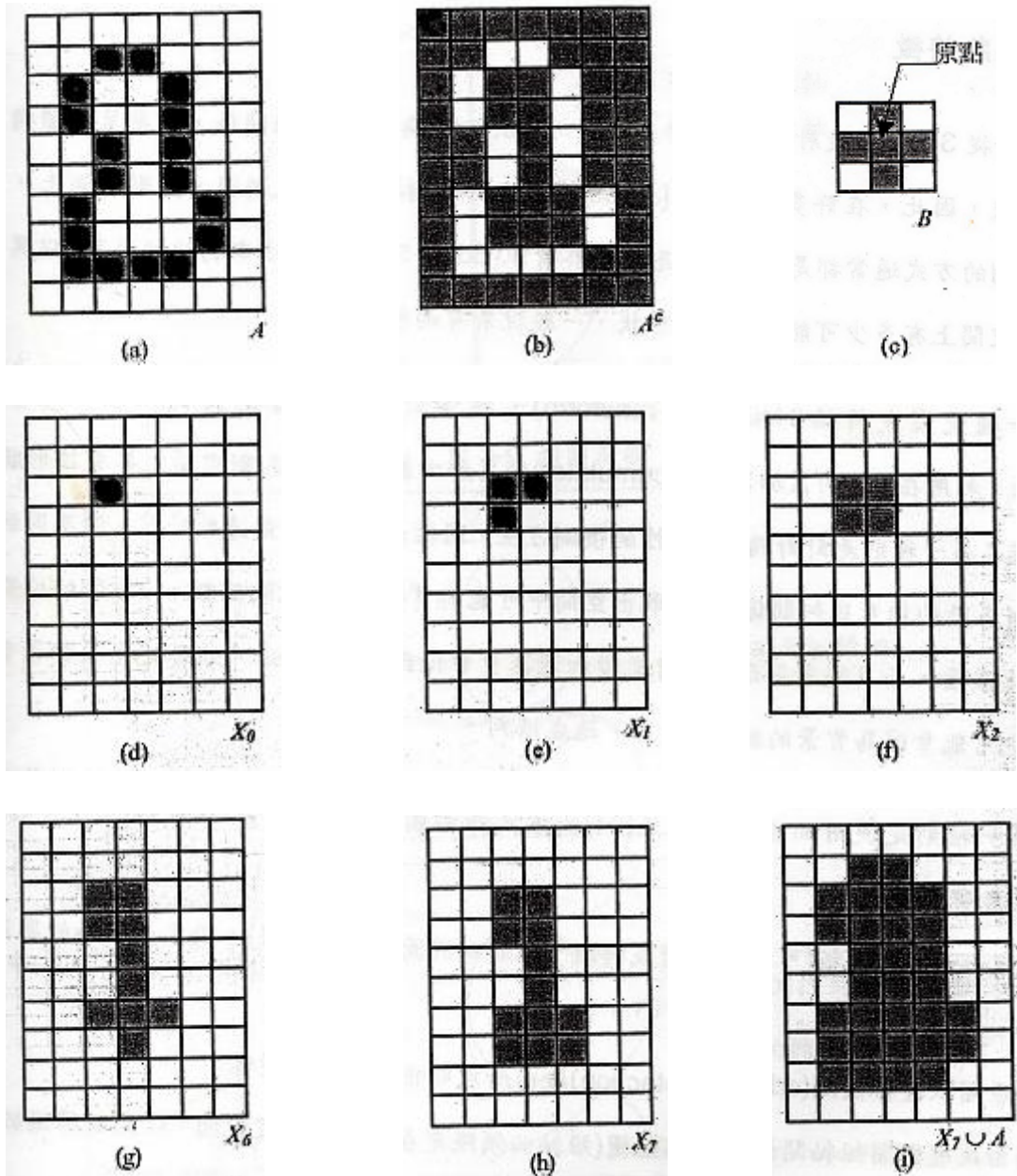


圖 9. 區域填充演算法示意圖

2.3-3 王淑儀-以膚色分割及類神經網路為基礎之人臉偵測[9]

在『以膚色分割及類神經網路為基礎之人臉偵測』文中，膚色偵測採用 Normalized RGB 及 HSV 兩種色彩空間來偵測膚色，選擇這兩種色彩空間在於這兩種色彩空間可以減少亮度對膚色所造成的影響。

利用 Normalized RGB 及 HSV 兩種色彩空間來偵測膚色，膚色樣本分佈情形如下：

$$\left\{ \begin{array}{l} R > G \ \& \ |R - G| \geq 11, \\ 0.33 \leq r \leq 0.6, \\ 0.25 \leq g \leq 0.37, \\ 340 \leq H \leq 359 \text{ or } 0 \leq H \leq 50, \\ 0.12 \leq S \leq 0.7, \\ 0.3 \leq V \leq 1.0 \end{array} \right. \quad (16)$$

由於上式的偵測結果過濾之後還含有類似膚色的雜點與一些相似膚色的背景區塊，因此他們還做了去除雜訊的處理。

接下來第二個步驟對 Skin map 作 3 X 3 中間值濾鏡去除較小的雜點，接著又進行著 Opening 運算(先侵蝕再擴張)，侵蝕主要是為了去除背景中較小的膚色區塊；擴張則是為了填補臉上的小洞再利用八鄰居 Connected component 分割出相連的膚色區塊。

接著第三個步驟過濾出可能的人臉區塊，他們用了三種方法：

第一是膚色區塊面積大小：

設定偵測的最小之人臉為 30x30 因為過小的人臉，其五官特徵並不明顯，臉部特徵很容易在處理影像的期間被抹除掉，如此會增加後續工作的困難度，因此我們對膚色區塊面積大小做個限制。

第二是膚色區塊外觀比：

人臉在直立狀態下，通常包含脖子的部份，所以高度會比寬度來得大，然而考慮了人臉可以旋轉的情況，我們放寬人臉的高寬比例限制，如下式：

$$0.8 < H/W < 2.6$$

以找出符合人臉比例的膚色區塊。

第三是 區塊中膚色所佔比例：

若此區塊內含有人臉的話，則人臉膚色在此區塊中應佔有一定比例 T ，為了避免包含有人臉的膚色區塊被去除，而影響到後續的處理，在此我們放鬆膚色所佔比例的條件限制，設定 T 為 40%。

經由以上三個條件的限制，過濾掉背景中的雜訊及類似膚色的區塊後，我們得到了一些可能包含人臉的區塊，即「可能的人臉區塊」。

在目前類神經網路學習模式，最具代表性且最多人使用的就是倒傳遞類神經網路了。倒傳遞類神經網路屬於監督式學習網路，基本組成單位為「處理單元」，由處理單元組成「層」，在由層組合成「網路」，網路架構由圖所示：

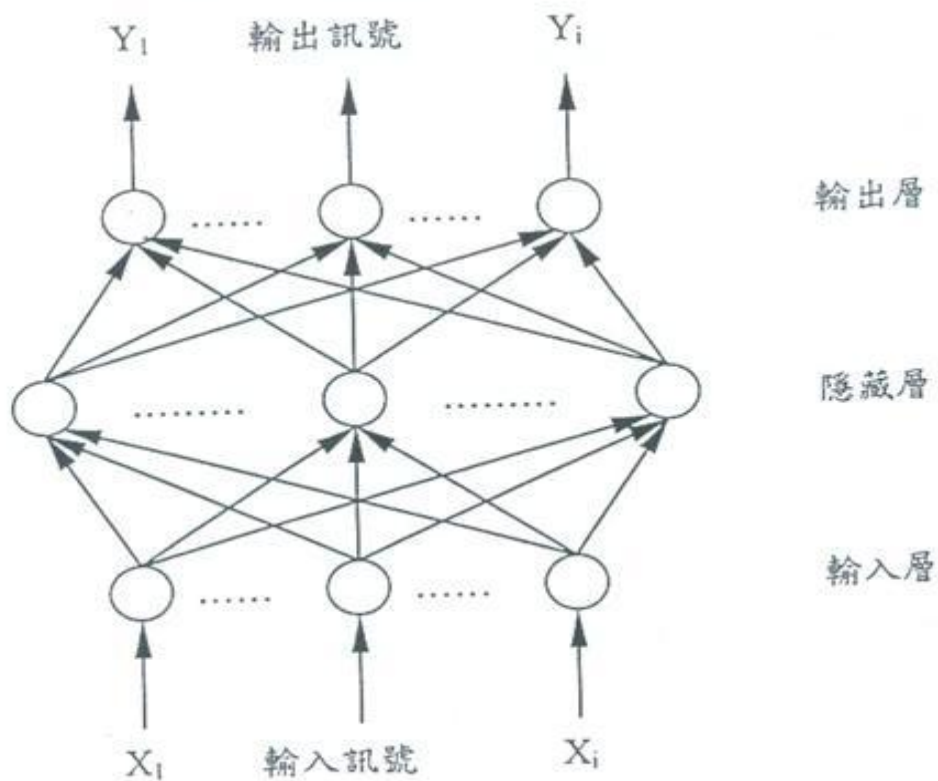


圖 10. 倒傳遞神經網路架構

- ◇ 輸入層：用來表現輸入變數，其處理單元數目依問題而定。
- ◇ 隱藏層：用以表現輸入層處理單元間的交互影響。
- ◇ 輸出層：表現網路的輸出變數。

以下簡述倒傳遞類神經網路的學習過程：

1. 設定網路架構，決定網路層數與各層的處理單元數目。
2. 隨機亂數設定鏈結加權值與閾值。
3. 輸入訓練樣本的輸入向量與目標輸出向量。
4. 計算推論輸出向量。
5. 計算差距量。
6. 計算鏈結加權值與閾值的修正量。

7. 調整鏈結加權值與閾值。
8. 重複步驟 3-7，直到誤差不再有明顯變化或執行一定次數的學習循環為止。

第三章 人臉偵測

在前面兩個章節我們提到了許多影像處理的基本概念，在系統中選擇 YUV 模型來作為我們將影像輸入所需要的數值範圍。緊接著擷取膚色，對圖片做二值化，形態學處理，人臉條件值設定，在這個章節我們將討論這些方法。

第一節 YUV 色彩模型[7]

系統讀入一張照片，我們第一個步驟便是要尋找出人臉可能存在的區域，為了要快速找出人臉區域，我們先找出膚色可能的範圍，利用程式將符合這段區域的顏色保留下來，供我們後續繼續使用。

在我們實作的系統中我們使用 YUV 色彩模型做膚色偵測。根據我們反覆實驗的結果膚色區域會存在於式(17)的區間內。非區間內的顏色即為非膚色部份我們便進一步加以捨棄，將原本的顏色全部變成白色，只保留我們所需要的膚色部份。

$$0 \leq Y \leq 0.1388$$

$$105 \leq U \leq 120$$

$$130 \leq V \leq 180 \tag{17}$$

我們利用YUV模型做膚色偵測，所保留的顏色大致上都接近膚色，如圖(11)(b)所示，顯示YUV色彩模型的效果非常好，我們所需要的人臉部分保留得大致完整。處理過後的膚色及非膚色部份，我們將繼續用來做後續的二值化處理。



圖 11. 過濾人臉膚色偵測結果

第二節 二值化處理

在做完膚色偵測後，我們從影像中偵測到近似膚色的部份，在這一章節，我們使用二值化處理主要的目的是將影像中膚色及非膚色部份，經過二值化處理後，將所有非膚色區域以白色(255)像素來表示，膚色區域以黑色像素(0)來表示，做此項處理的實質意義，是以利於之後須做更深入的影像處理。顯示結果如圖(12)(b)。

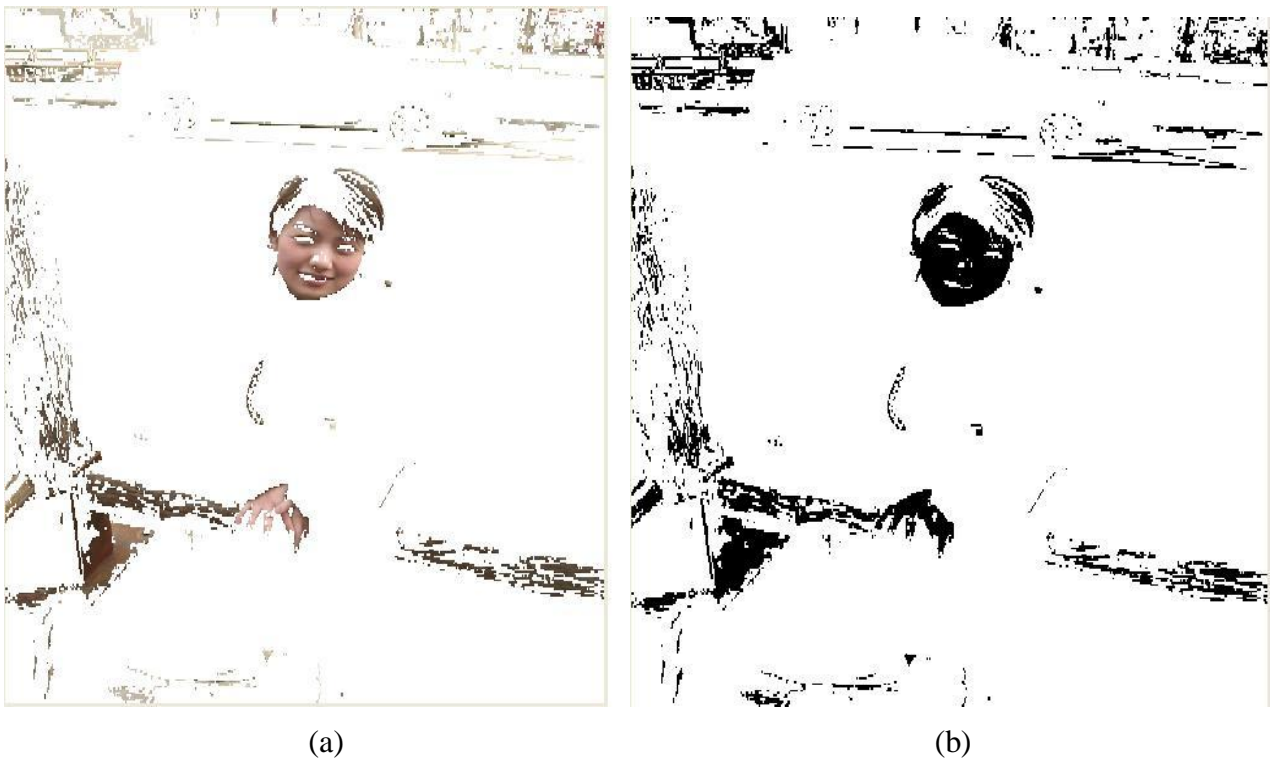


圖 12. 原始圖像(a)，針對膚色二值化結果(b)

第三節 形態學處理[1]

形態學是專門用來處理形狀的方法論，我們以數學上的形態學來當作擷取影像的工具，用來處理邊界、區域消除、粗化及細化等，是非常重要的處理工具。

當然並不是每張照片都這麼完美，此外像素間的連接性是眾多影像中邊界與區域的一種基本概念，利用此種方法判斷兩像素是否相連，以決定是否互相接近，我們第一次做標記化，便可以得到我們所想要的結果，因此我們就有必要做進一步的處理。這時就可能要使用 opening(斷開)，如圖 14(a)(b)與 closing(閉合)，如圖 15(a)(b)對圖片做處理。

公式 (14) 有提到：斷開就是 A 先被 B 侵蝕，再將結果對 B 膨脹。膚色與雜訊所連接比較薄弱、細小的部份，我們就有可能把雜訊連接的部份給割開。這時我們就可能可以偵測到之前與雜訊連接，系統無法確認的區域，如圖 13(a)(b)。

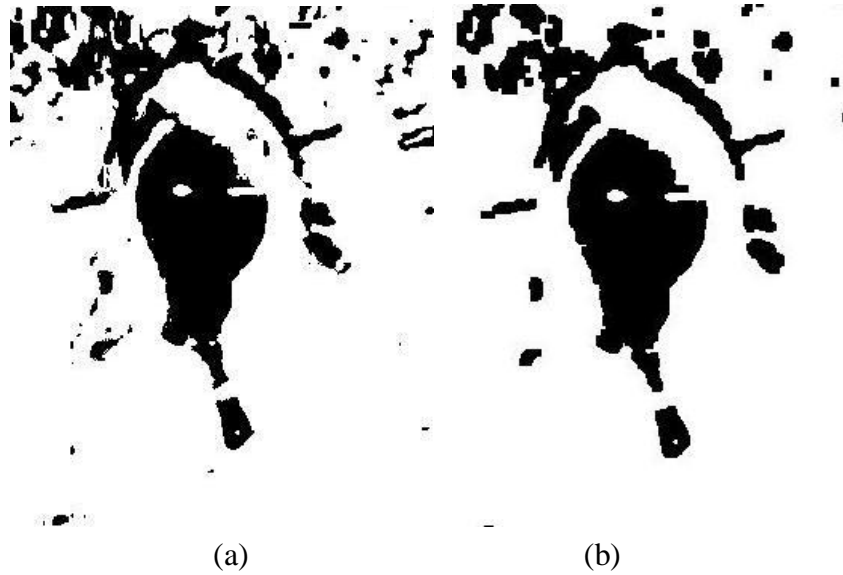


圖 13. 將額頭可能為雜訊的部份斷開

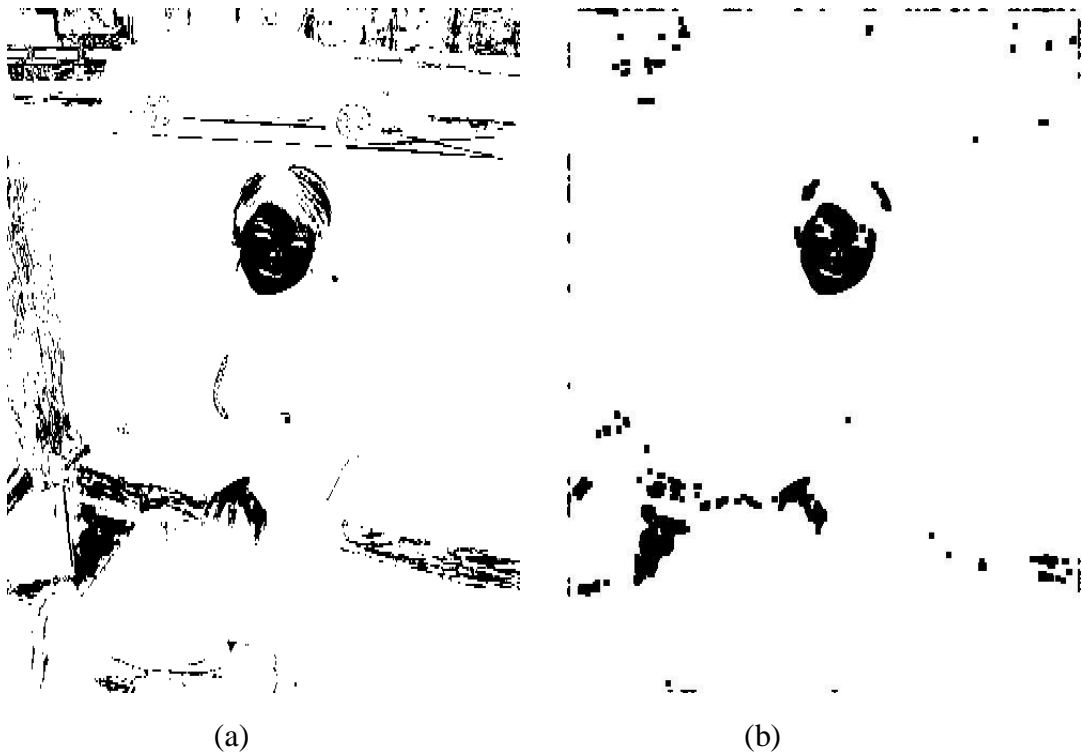


圖 14. opening 前(a)，opening 後(b)

相反地公式 (15)，所說的閉合是：A 先被 B 膨脹，再將結果對 B 侵蝕。便是將可能被視為雜訊所過濾掉的部份，膨脹回來，再將與其他雜訊所連接的部份給斷開。

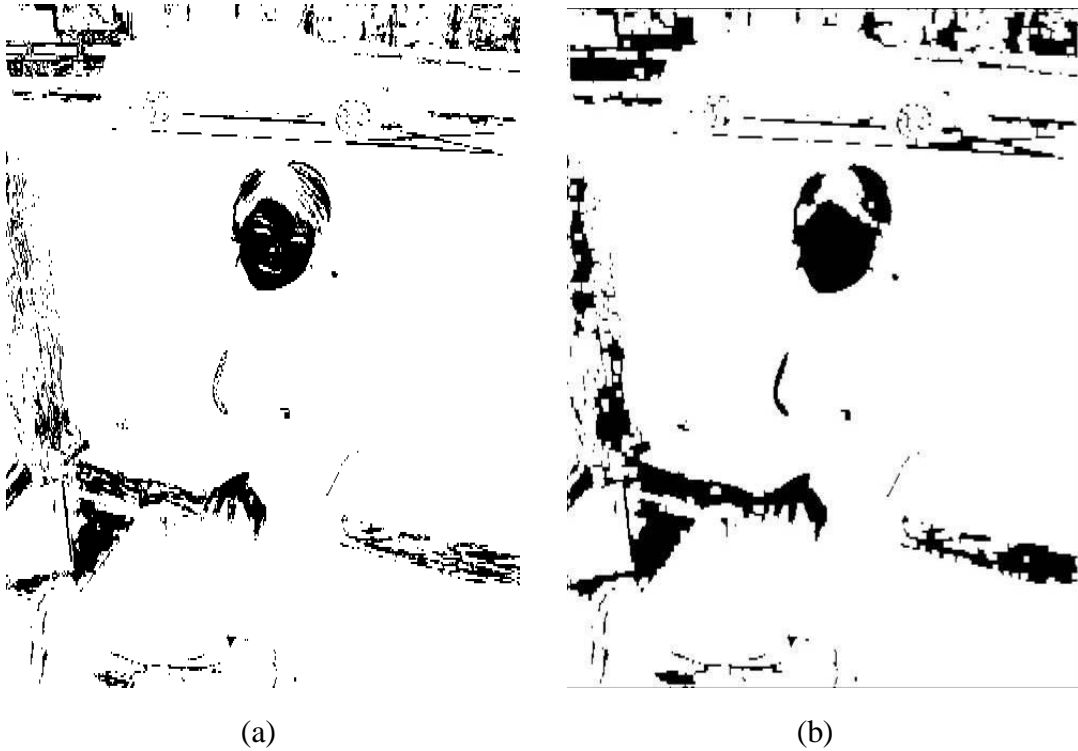


圖 15. closing 前(a)，closing 後(b)

因此我們使用斷開消除細的雜訊連接部份，用閉合消除小洞，填補輪廓上的缺口。藉此得到完整的圖形，以利判斷是否為人臉。

第四節 相連區域標示法[1]

圖片經過二值化的處理後，便會形成黑白分明的照片。我們想要在這張照片中取得人臉的部份，要做的便是 Connected Component Labeling（相連區域標示）。系統中我們採用的是 8-connected（8 連通），將膚色區域被二值化所變成黑色的部份做上標記。在 2.3-2 討論到四連通標記化，如圖便是四連通與八連通之差異。

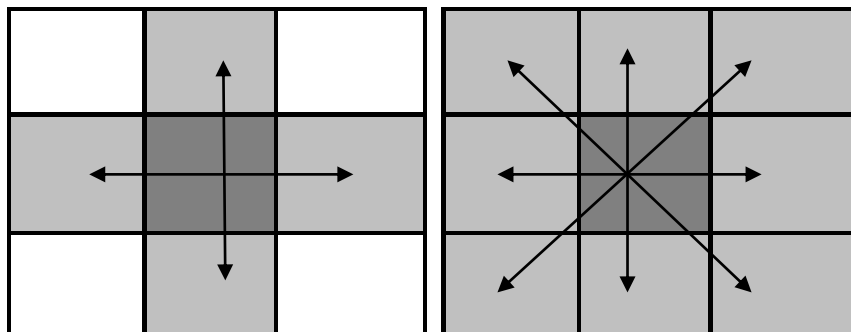


圖 16. 圖(a)四連通，圖(b)八連通

運用標記化我們可以將不同區域的圖形做有效的分離，在不同的區域，隨著掃描的順序，依序給予不同的編號。如此，我們便根據該數字找到不同的區域，做進一步的分析處理，如圖 17(a)(b)。

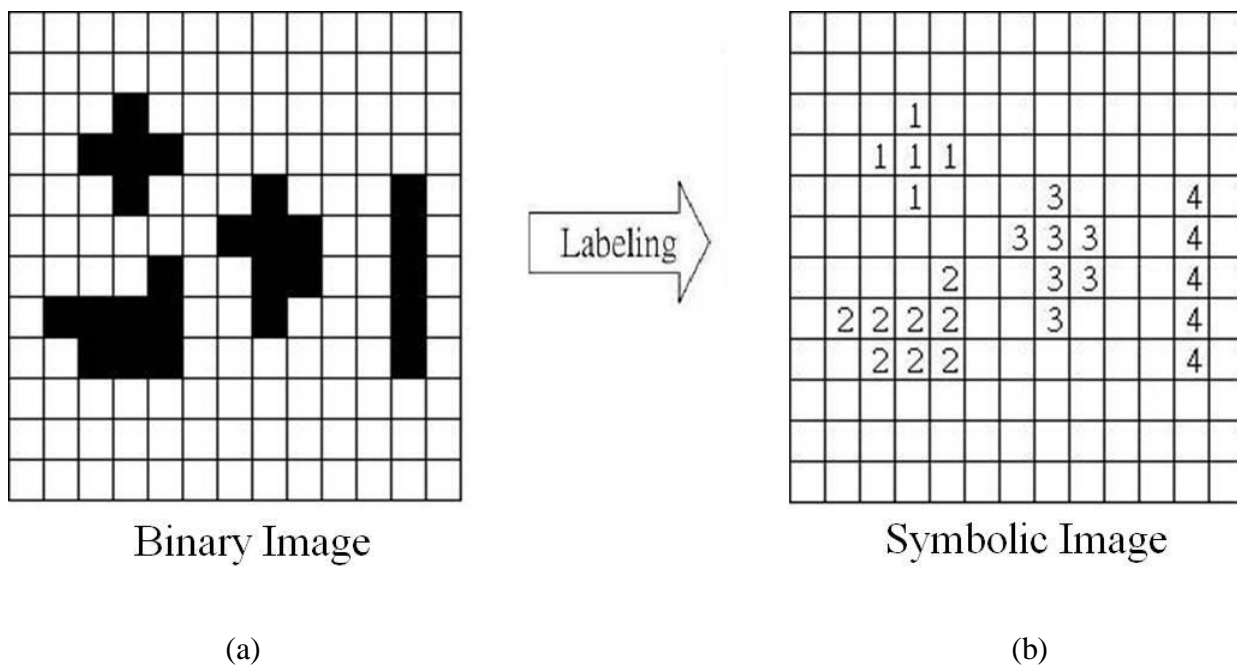


圖 17. 標記化

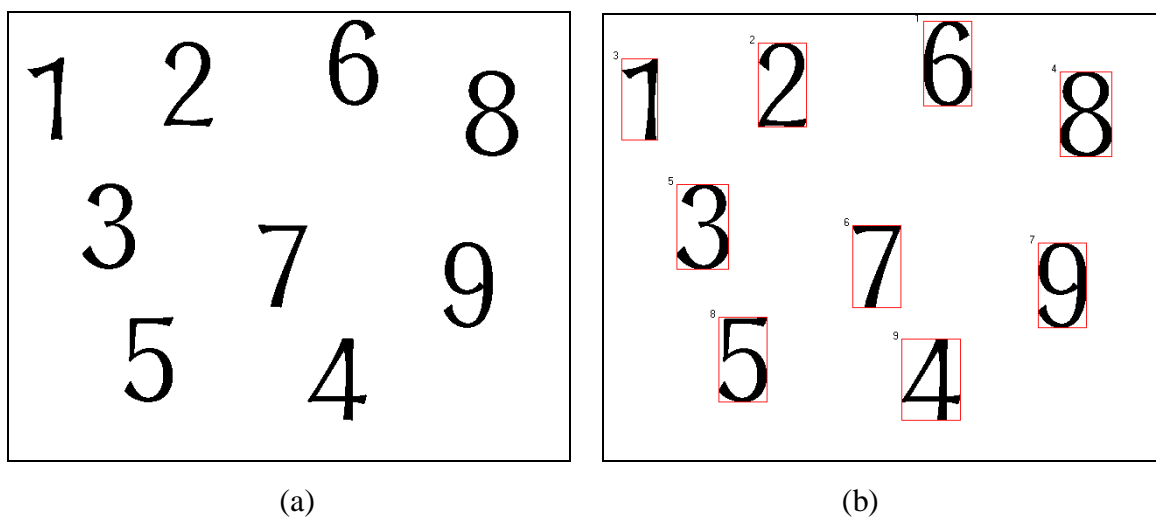


圖 18. 標記化前(a)，標記化後(b)

事實上，最後做完標記化，我們只需要可能為人臉的區域，其餘不是人臉的部份，我們統統視為雜訊將其捨棄。最後並在人臉的週圍畫上一個紅色的方框，明確顯示我們是否得到正確的人臉區域，如圖 18(a)(b)。

第五節 人臉條件值設定 [3] [6] [10]

當我們做完二值化與形態學處理後，保留下許多膚色區塊，像是脖子、胸口、手、腳等與，當然會被系統給擷取下來，再者沙子、夕陽、樹葉等偏向黃色的物件，也是有很高的可能性被擷取下來。所以系統中我們運用了二個方法來判斷是否可能為人臉部份：分別為長、寬比例、膚色佔區域面積。

人臉的比例，大致上有長寬比為三比二的比例在，因此我們便很容易判斷出來，如腳、樹枝等長條形的比例就不對，系統就自動過濾掉。在這個條件下許多的物件都會被我們過濾掉，當然還是會有符合此像條件的雜訊，我們就要在進行膚色多寡判斷。

我們規定在做標記化時，框框內的膚色部份必須超過整個區塊面積之 1/2。這是因為如果該區塊為人臉的話，其膚色區域便會群聚在一起，呈現一大片都是膚色的狀態，但是如果是樹葉等被系統視為可能為膚色的物件，在此條件下便可能會因為其膚色數值不夠多，便會被系統給判斷出來，進而確認。

如圖 19(a)臉部很完美的抓取下來，但是在經過 closing 的處理過後(圖 19(b))眼睛部位的小洞，便會被填補起來。系統因為其長寬比有達到三比二與方框內的膚色比例超過 1/2 便會保留下來，成功的做到了我們所需要的人臉部份。

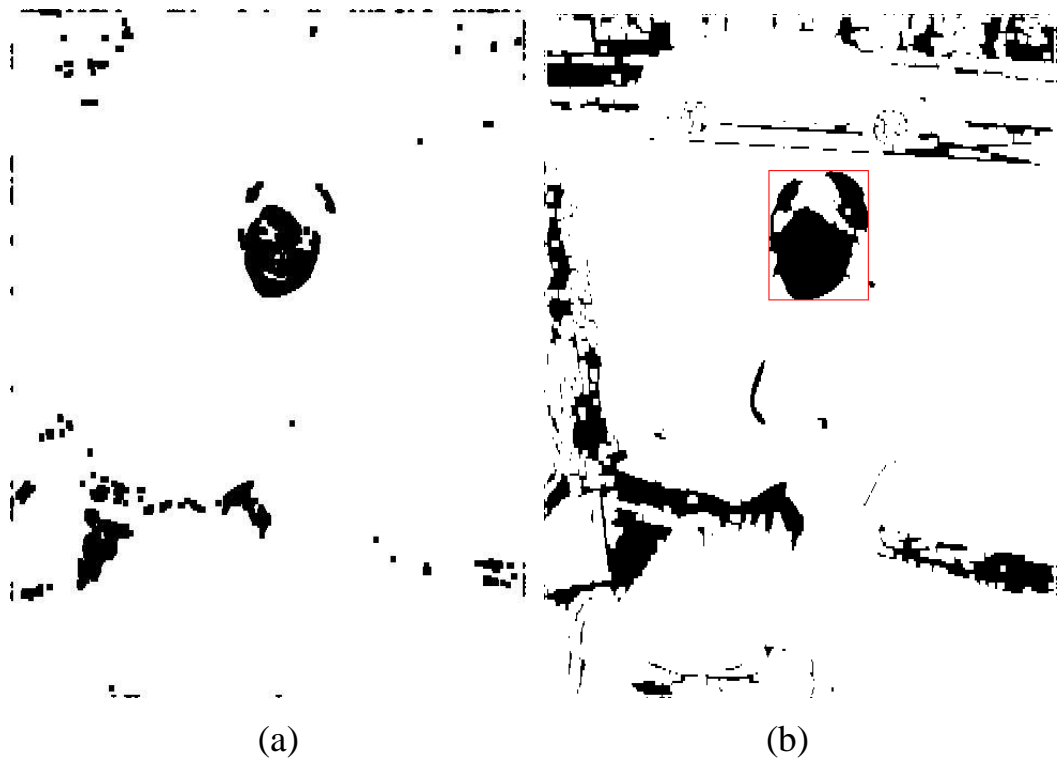


圖 19. 人臉條件值所造成的差異

第四章 人臉偵測系統

在這個章節，我們將實做上一章所提出的方法，方法雖然簡單，但會影響結果的因素有很多，如光線、背景、眼鏡、衣服，基本上所有的外在因素都有可能會影響，因此我們發現原先所沒有想到的缺點一一浮現，經由將系統改善修正，最後提出我們所得到的數據結果。在系統測試上，我們利用網路隨機取得戶外照片 97 張，測試結果是否為我們所需的人臉部份。

圖 20 是系統介面，整體介面上的設計相當簡單，希望可以做到人人都可以輕鬆使用，而不是一個看起來相當複雜的系統，需要學習過才能夠使用。但是，假使照片必須經過些微的處理，才能避免誤判的情形發生，我們也可以在主選單下的 color 做調整，方便我們得到正確的人臉區塊。



圖 20. 系統介面

第一節 人臉偵測系統效能

人臉偵測系統要分辨系統效能好壞很簡單，能夠正確的抓取到人臉便是成功，發生誤判抓取不是人臉部份或根本就沒有抓取到任何的部份那便是失敗了。圖 21 正確的抓取到兩人人臉部份，旁人及手臂、眼鏡沒有干擾到我們的判斷。圖 22 便抓取失敗，雖然有正確的抓到兩人的的人臉部份，但右邊的手因為比例、膚色相素與人臉相近，故也會被系統當成人臉區域抓取。當遇到此類的錯誤判斷，我們便需要反覆測試、修改程式內的人臉規則，使照片中的人臉能夠被正確的抓取。

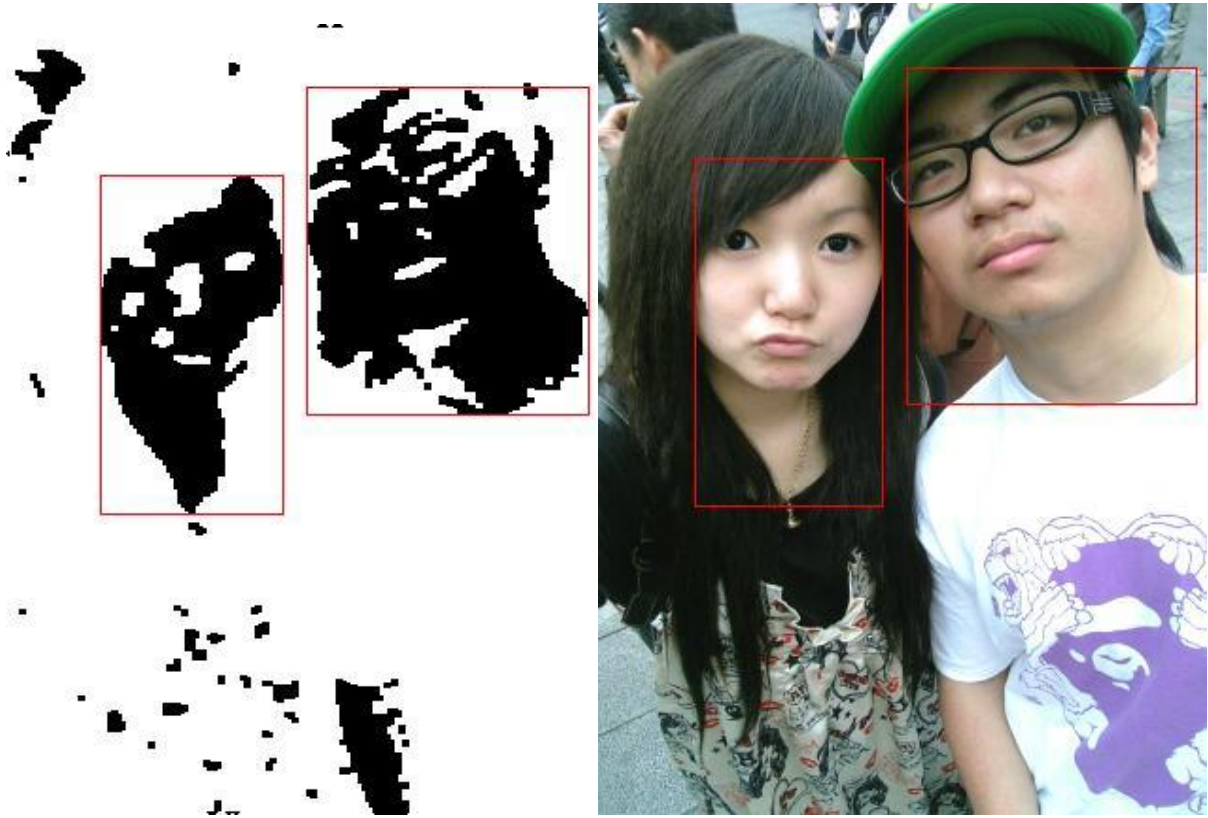


圖 21. 成功的偵測



圖 22. 錯誤的偵測

第二節 效能分析

我們在檢討程式抓取照片時所出現的錯誤，不斷反覆測試照片的條件值並修改程式的人臉規則，希望程式可以做到盡善盡美，使人臉偵測發生誤判出現的機率降低。在這一節我們將最後發生的誤判整理出來，「O」代表成功的抓取，「X」代表抓取失敗或發生誤判的情形發生。

圖片編號	人臉正確抓取與否	錯誤改進
1	O	
2	O	
3	O	
4	O	
5	O	
6	O	
7	O	
8	O	
9	O	
10	O	
11	O	
12	O	
13	O	
14	O	
15	O	
16	X	膚色相素不夠多，需 Dilation
17	O	
18	X	三人中有一人抓不到
19	O	
20	O	
21	O	
22	X	
23	O	
24	O	
25	O	
26	O	
27	O	

28	O	
29	O	
30	O	
31	O	
32	O	
33	X	也會偵測到手臂部份
34	O	
35	O	
36	X	
37	X	也會偵測到手臂部份
38	X	也會偵測到手臂部份
39	O	
40	O	
41	O	
42	X	手部部份比例符合比例
43	O	
44	O	
45	O	
46	O	
47	O	
48	X	膚色相素不夠多，需 Dilation
49	O	
50	O	
51	X	也會偵測到手臂部份
52	O	
53	X	有一人沒有抓到
54	X	有一人沒有抓到
55	X	有一人沒有抓到
56	X	只抓的到一人
57	X	只抓的到一人
58	X	只抓的到一人
59	X	也會偵測到手臂部份
60	X	只抓的到一人
61	X	只抓的到一人
62	X	只抓的到一人
63	O	
64	X	也會偵測到手臂部份
65	O	
66	O	

67	O	
68	X	其中一人膚色相素不夠多，需 Dilation
69	O	
70	O	
71	O	
72	O	
73	O	
74	O	
75	O	
76	O	
77	O	
78	O	
79	X	也會偵測到手臂部份
80	X	也會偵測到手臂部份
81	O	
82	O	
83	O	
84	O	
85	O	
86	O	
87	O	
88	O	
89	O	
90	O	
91	O	
92	O	
93	O	
94	O	
95	O	
96	O	
97	O	

所有照片共 97 張，完整抓取 75 張，失敗 22 張，成功率 77.3%。

第五章 結論

我們系統運用了許多學理上的技術，希望能做出一個戶外人臉自動偵測系統，系統在充足的光源之下，正確率與光線不足的照片相比好出許多。多人團體照方面，雖然偵測結果未能盡如人意，但是在正常的情況條件下，我們也還是能夠抓取到人臉部份，只是會連同非人臉部份一併被標示出的情況發生。

人臉偵測系統有『準確度』與『速度』兩大難題，應該選擇執行迅速或者應該準確的判斷，事實上這是人臉偵測系統中的難題。需要執行速度快，便需要犧牲一些細節的部份，這樣做便不夠的精準；若需準確的判斷，系統的演算法便會複雜且運算的時間就會需要比較長。因此，我們會依其應用來做選擇，若是人臉辨識資料庫，所需的人臉特徵比較多，以免後續辨識系統出錯，發生不必要的困擾，這時所擷取的人臉影像，便需要包含整個臉部特徵，我們就會採用精準的判斷防止失敗。如果是即時性高的人臉移動，因為不需要太多的擷取人臉特徵，只需要正確的抓到人臉，演算法的速度就可以加快，對系統的速度有所要求，我們便會使用類似這類處理迅速的運算。

在這兩學期的專題製作中，我們將人臉偵測系統的前端部份實做出來，其中在不斷反覆測試修改程式，最後得到了人臉偵測成功率 77%。雖然這個結果未能盡如人意，部份的照片還是會有其他因素，造成偵測失敗的結果。這一部份的失敗改進便是需要透過新的人臉/非人臉判斷條件，才能夠獲得解決，也是我

們需要精進的地方。日後若需建立人臉辨識資料庫，也可以輕易的修改成所需的系統，達到完整應用的目的。

參考文獻

- [1] 吳宣諭，95 年 6 月，運用膚色過濾的人臉偵測法，台灣科技大學電機工程研究所碩士班。
- [2] 鄭凱方，94 年 6 月，人臉可辨識度計算用於監控系統中人臉正面最佳影像判定，中央大學資訊工程學系碩士班。
- [3] 曹培彥，94 年 6 月，利用灰階影像直方圖偵測人眼，義守大學資訊工程研究所碩士班。
- [4] 陳秋旺、葉賜旭，2006 年，生活伴侶機器人技術簡介，機械工業雜誌，第 281 期，第 14~22 頁。
- [5] 吳鴻鈞，95 年 6 月，一個以半徑對稱性加強人臉偵測可靠度的方法，元智大學資訊工程學系碩士班。
- [6] 張銘元，90 年 6 月，植基於條件式判別法則及直覺式特徵點辨識法之人臉自動辨識管理系統，靜宜大學資訊管理學系碩士班。
- [7] 梁皓雲，95 年 6 月，利用區塊人臉特徵為基礎之混合式之人臉辨識系統，中央大學資訊工程學系碩士班。
- [8] 馬炳義，95 年 6 月，人臉偵測技術研究，義守大學資訊工程研究所碩士班。
- [9] 王淑儀，94 年 6 月，以膚色分割及類神經網路為基礎之人臉偵測，淡江大學資訊工程學系碩士班。

[10] 廖維中，94 年 6 月，以色彩分析為基礎之臉部遮蔽偵測應用於 ATM 監視系統，淡江大學資訊工程研究所碩士班。