

第三章 人臉偵測方法

本文主要分為人臉偵測、人臉辨識兩大部分。在人臉偵測階段，本文分為人臉候選區的尋找、候選區中眼睛和嘴巴的定位、最後利用嘴唇和眼睛的幾何判斷條件去定位出正確的人臉位置。人臉偵測的流程如圖3.1所示。

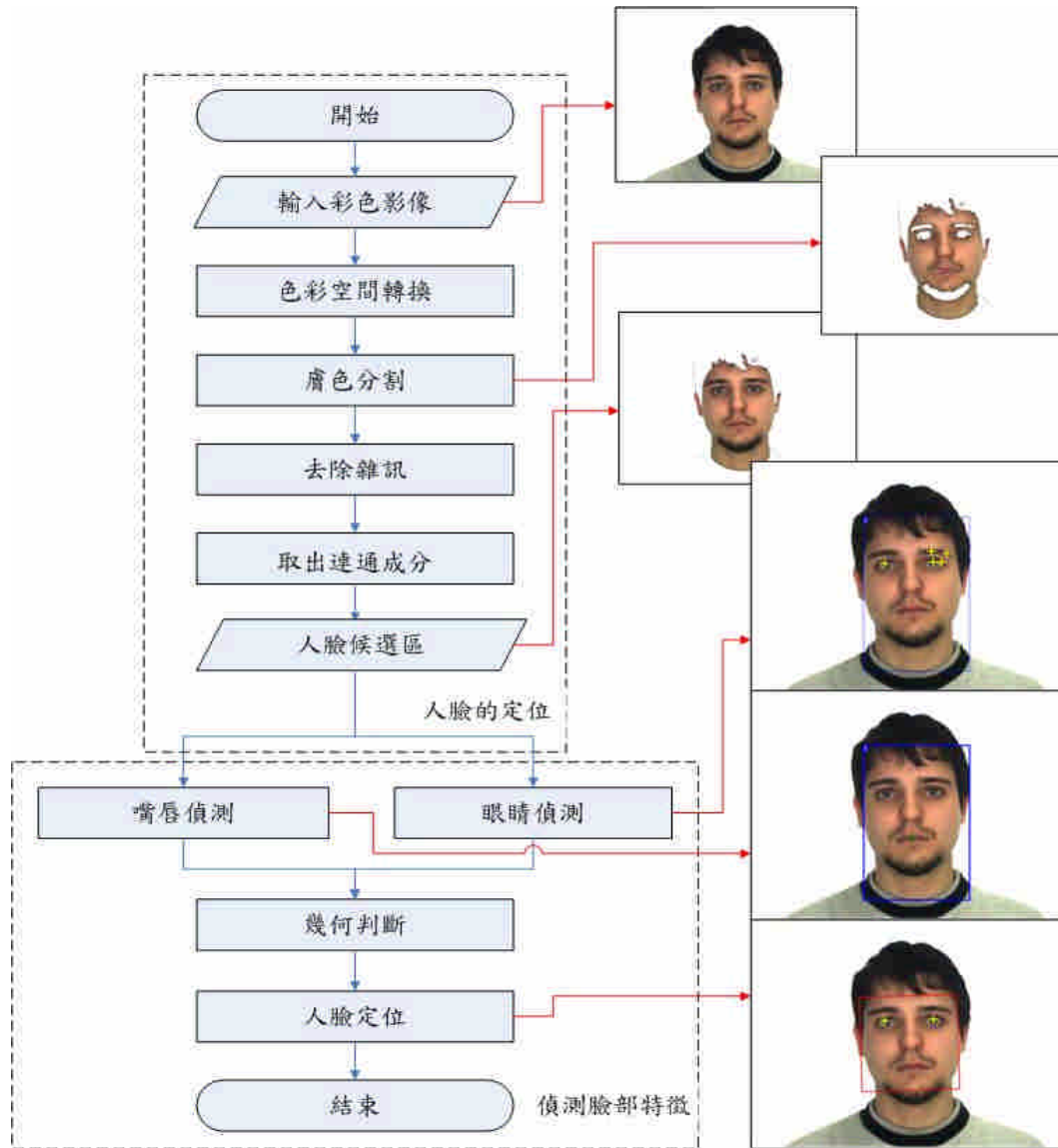


圖3.1 人臉偵測流程圖[35]

由於人的膚色分佈在彩色空間上具有一致性、區域性的關係，所以有很多文章在彩色空間上，提出了膚色分佈的數據，有的是用Y、Cb、Cr來

定義膚色的分佈[30][31]，是較常應用在硬體導向的彩色模型上使用，如數位照相機。有的是用H、S、V來定義膚色的分佈[32]，其目的是強調此H、S較能抵抗因外界光線照射的干擾，且較容易讓人類視覺接受的彩色模型。有的是用R、G、B來定義膚色的分佈[33][34]，目的是要加快運算速度，不需要做彩色空間的轉換計算。以事先訓練好的人臉膚色的分佈範圍，再從影像中篩選出這些屬於膚色的像點，再用膨脹原理將這些不連續的膚色像點連接起來。最後再用長方形或橢圓形人臉近似的方法，來判斷是否為人臉的位置所在。這種方法容易受不同的光線照射、複雜的背景及環境顏色影響。

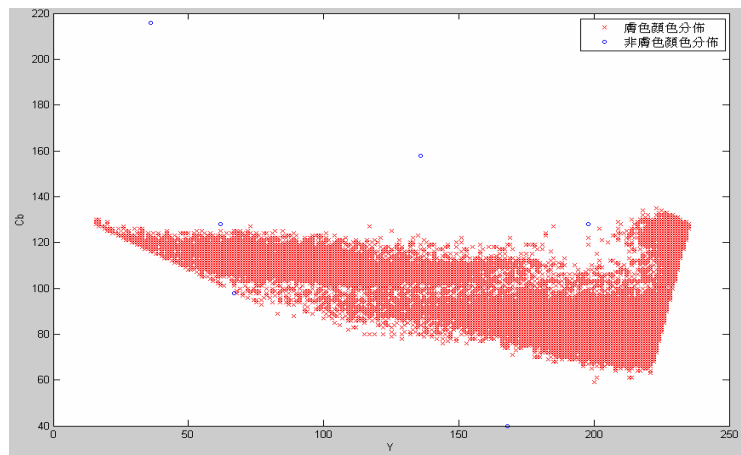
到目前為止，膚色偵測在不同的應用中，都會碰到一些不同的問題。同時對於何種色彩座標較佳，似乎也沒有一致的看法。我們歸納膚色對於影像形成的條件，有極大的關係。所以只用膚色特徵並不是最好的辦法。雖然其優點是快速且容易實現人臉偵測的方法，但缺點是容易受外界環境、光線照射及人臉上的表情等因素影響。

3-1 膚色分割

人類的皮膚顏色在影像中，若以RGB三原色色彩空間來看，對人的眼睛或許沒問題可以很快的辨識出膚色區域，但是以機器視覺或電腦來說並不是件容易的事情，且近年來的研究當中利用人類膚色的特性來判別人類的臉部區域越來越受到重視。但是就以RGB色彩空間來說容易受到光源影響，會隨著光源的變化而有很明顯的變動，會造成在RGB色彩空間上所設定的閾值做顏色的分割會有很大的誤差，所以並不適用於光源變動大的環境中進行色彩的分割動作，如此必須將色彩空間做一個調整。本文是使用YCbCr的色彩空間搭配上形態學運算和幾何關係達到人臉和人臉特徵的偵測，因為YCbCr色彩空間較不易受到光源變化的影響，有好的亮度和彩度分離的特性。在YCbCr色彩空間中Y是指亮度、Cb是指藍色

的成分、Cr 是指紅色的成分。

本文的膚色分割[35]是以色彩為基礎，也就是根據各色彩成分之像素值大小來進行膚色的判斷，在文獻[30]的方法中發現在不同的亮度中膚色所群聚的範圍不一樣，也就是說不同的亮度膚色的Cb、Cr的範圍是不一樣的，所以我們統計不同亮度的膚色影像，統計出來的圖如圖3.2所示，其中Y值是指亮度值。3.2(a)中x軸是指亮度的範圍由0~255，0是黑色、255是白色中間部分以此類推，y軸是指Cb的範圍，圖中紅色的部份就是膚色在Y-Cb空間所群聚的範圍，3.2(b)是代表膚色在Y-Cr空間的群聚範圍。



(a)

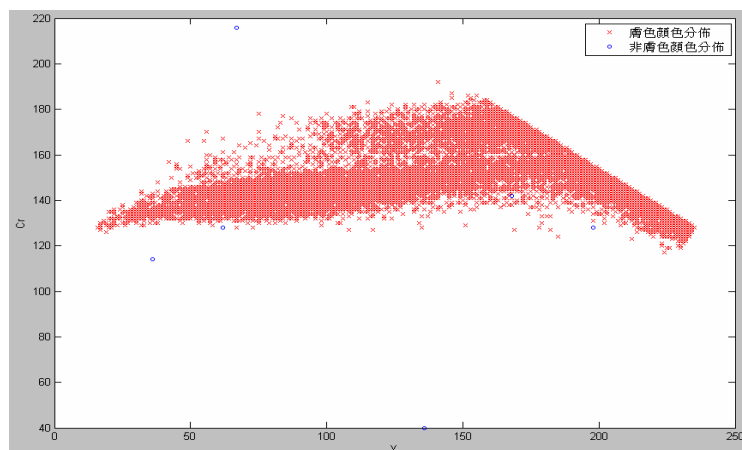


圖 3.2 膚色群聚的範圍在不同的亮度下(a) 膚色在 Y-Cb 空間的群聚範圍；(b) 膚色在 Y-Cr 空間的群聚範圍[35]

基於圖 3.2 的色彩空間我們可以定義可能是膚色的範圍如下表 3.1 所示，因為黑色和白色的 Cb、Cr 也在 127~128 之間，所以我們多了一個判別式可以避免白色和黑色會被當作膚色來偵測， $R_n=R/(R+B+G)$ 因為黑色和白色的 R、B、G 值差不多所以 R_n 接近 1/3 時，那個 pixel 可能是黑色或白色的機率很高，且膚色的 R 值雖然很高但是當一個 pixel 的 R 值大於 $R+G+B$ 的一半時時這個 pixel 的顏色就接近紅色而不是膚色。我們使用之前統計膚色範圍的人臉計算 R_n 的值，發現膚色的 R_n 範圍介於 0.343~0.702 之間大部份的膚色 R_n 值是介於 0.4~0.5 之間的，由之前的條件和膚色 R_n 的範圍我們定義膚色 R_n 值的範圍為 0.36~0.6，這樣可以保留下膚色的部份又可以濾除掉黑色白色和太紅的顏色。使用 R_n 的範圍再加上我們歸納出的不同亮度下的膚色範圍，我們就可以利用 3.1 式分割出正確的膚色範圍。

表 3.1 不同亮度值下膚色群聚範圍表[35]

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	$Y \geq 220$	$65 \leq Cb \leq 141$	$115 \leq Cr \leq 153$

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

$$n = \begin{cases} 0 & \text{if } [Y(x, y) \leq 50], \\ 1 & \text{if } [51 \leq Y(x, y) \leq 110], \\ 2 & \text{if } [111 \leq Y(x, y) \leq 180], \\ 3 & \text{if } [181 \leq Y(x, y) \leq 205], \\ 4 & \text{if } [206 \leq Y(x, y) \leq 219], \\ 5 & \text{if } [Y(x, y) \geq 220], \end{cases} \quad (3.1)$$

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

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，上列公式的意思是輸入一個pixel計算他的Rn值假如Rn值的範圍介如0.36~0.6之間的話，再計算它的亮度值判斷它是0~5那個map的範圍，定義是哪一個map之後再判斷它的Cb、Cr的範圍是不是屬於這個map假如是的話就判定為膚色。經由膚色分割後我們可以得到一個二值影像，由於我們使用不同亮度用不同的Cb、Cr值來判斷，所以得到的效果比一般的方法還好。

3-2 去除雜訊

在我們得到膚色分割的二值影像，可以看出裡面還有許多雜訊或是一些破碎的情況產生，這些雜訊的產生是因為在背景中剛好有些 pixels 的顏色屬於膚色的範圍，通常這些 pixels 的數量不大且都是分散開來的點雜訊所以我們使用中值濾波器去除雜訊。我們所使用的中值濾波器的遮罩是3×3的大小。

首先把遮罩內所含蓋的像素灰階值由小到大排列，中值是指排序在中間的那一個值，此值即為濾波器的輸出。此法特別適合用在有很強的胡椒粉式或脈衝式的雜訊時，因為這些灰階值的雜訊值與其鄰近像素的灰階值有很大的差異，因此經排序後取中值的結果是強迫將此雜訊點變成與其鄰

近的某些像素的灰階值一樣，達到去除雜訊的效果。由中值濾波器的原理不難看出中值濾波器是一個非線性的動作。

在決定每一點的值之前，要對區塊內的每個亮度值排序一番，其計算量稍大於簡單的加權運算。然而以原有值取代，便不會產生影像中不曾出現過的新值，且雜訊也不會有什麼影響力。如圖3.3所示，一個3×3遮罩中的數字由小排到到可以得知第五個數字是70，所以就用70去取代原本遮罩中間的值，這是用於灰階影像時的狀況。在二值影像中，只有黑與白也就是0和255，中值濾波器的效果就是比較3×3遮罩中白點還是黑點多，假如白點多就用白點取代原本遮罩中間的值。

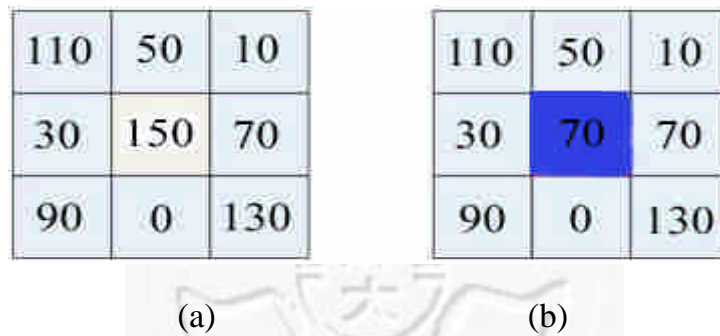


圖 3.3 中值濾波器 (a)原本的遮罩；(b)取代後的遮罩

在尋找人臉候選區的時候，會有許多破碎的產生，所以我們提出以形態學(morphology)演算法來改善這個問題。接下來會先介紹一些形態學演算法的基本概念，然後再說明使用形態學演算法來處理雜訊結果。

形態學上之基本定義設A和B是二維空間中的兩個集合，它們的元素分別為 $a=(a_1,a_2)$ 和 $b=(b_1,b_2)$ 。A平移 $x=(x_1,x_2)$ 後用 $(A)_x$ 表示之，定義為：

$$(A)_x = \{c \mid c = a + x, a \in A\}, \quad (3.2)$$

A的補集是：

$$A^c = \{x \mid x \notin A\}, \quad (3.3)$$

A和B的差集，表示為 $A-B$ ，定義為：

$$\begin{aligned}
 A - B &= \{x \mid x \in A, x \notin B\}, \\
 A - B &= A \cap B^c,
 \end{aligned}
 \tag{3.4}$$

擴張(Dilation)，A被B擴張，記為 $A \oplus B$ ，定義為：

$$A \oplus B = \bigcup_{b \in B} (A)_b,
 \tag{3.5}$$

意思是將A集合向外平移B集合中每個元素的距離，然後把這些集合聯集起來，所以擴張會使A集合擴大，且擴大的範圍是由B集合之結構來決定。

侵蝕(Erosion)A被B侵蝕，記為 $A \ominus B$ ，定義為：

$$A \ominus B = \bigcap_{b \in B} (A)_{-b},
 \tag{3.6}$$

意思是將A集合向內平移B集合中每個元素的距離，然後取這些集合的交集，所以侵蝕會使A集合縮小，且縮小的範圍是由B集合之結構來決定。圖3.4為擴張和侵蝕的結果其中B集合為一個3×3的遮罩。在圖3.4(b)中橘色的部份是擴張出來的結果，在圖3.4(c)中淡掉的部份是被侵蝕的地方。

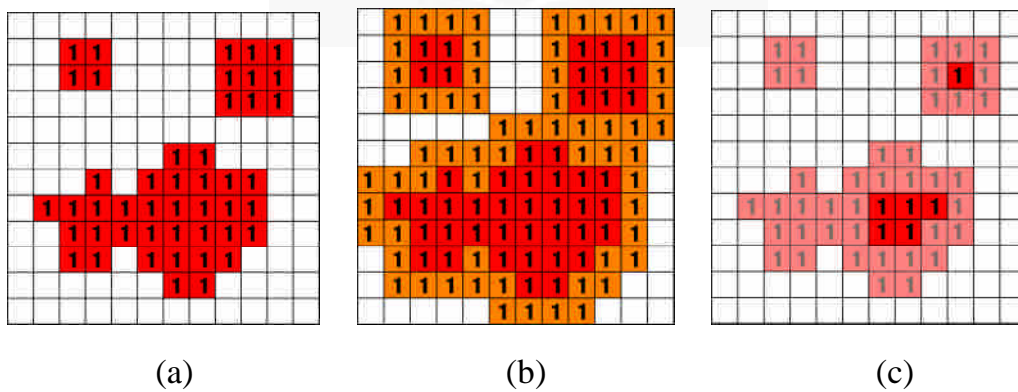


圖 3.4 擴張和侵蝕的結果(a)輸入影像 (b)擴張的結果(c)侵蝕的結果[35]

中值濾波器是利用一個 3×3 的遮罩，將遮罩內所包含的九個亮度值做大小的排序後，排序在中間的值(中值)作為該濾波器的輸出值，以此替代

掉遮罩內位於中間的亮度值。這樣的方法特別對於處理孤立的點雜訊有效。在經過遮罩的作用之後，孤立的點或是面積較小的點都會被黑色所取代，視為雜訊而濾除。形態學運算是透過結構元素的中心點 X 對應到二值圖上，取出該點以及相對映所參照到的所有色階值執行邏輯 AND 運算，若是其中一個點的值為 0，則該點也設定為 0，否則不變。經過侵蝕後可以把體積稍大的雜訊點消去掉。我們把膚色分割後的二值圖形做中值濾波器和做侵蝕擴張運算後，可以得到如圖 3.5 的結果。

在使用中值濾波器和型態學運算之後可以去除掉比較小的點雜訊，不過對於在背景中出現的原木色桌子或是類似膚色的壁紙，這些屬於比較大塊的雜訊無法使用中值濾波器和型態學運算來去除所以我們就先用長寬比等條件去除一些比較大塊的雜訊，詳細內容在下一節提出。

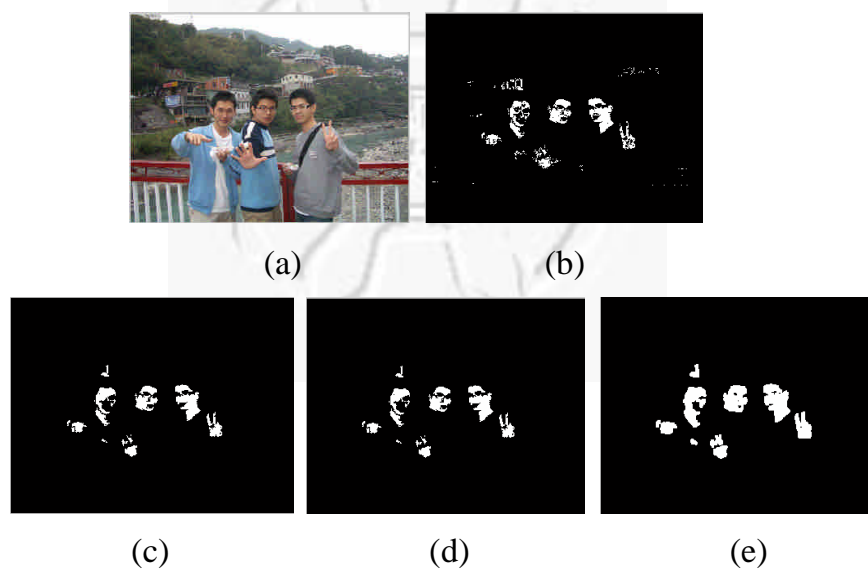


圖 3.5 (a)原始影像 (b)膚色分割影像
(c)中值濾波器影像 (d)侵蝕影像
(e)擴張影像

3-3 取得連通成分和標記人臉候選區

在我們去除雜訊之後，我們得到圖 3.5(e)的影像，但是因為臉部特徵的顏色也不屬於膚色，所以導致眼睛、嘴巴等特徵也被當作是雜訊，這樣在接下來的處理會找不到眼睛、嘴巴的所在位置，所以我們使用區域填充的方法保留臉部的特徵以便於在之後的處理。

以圖3.6[36]為例，A代表一個集合，它含有一個子集合，該子集合是依照邊界抽取演算法所求得的邊界，從邊界內側任選一點p就可以開始區域填充的程序，方法如下：

$$X_k = (X_{k-1} \oplus B) \cap A^c, \quad k=1,2,3,\dots, \quad (3.7)$$

其中 $X_0=p$ ，B集合如圖3.6(c)。第k次疊代時 X_{k-1} 會被B集合向4連通(4-connected)方向膨脹，不過因為有跟 A^c 取交集，不至於擴大到整個區域。如果不再變化則演算法終止，此時令 $X_{conv}=X_k$ ，A的區域填充 $R(A)$ 為：

$$R(A) = X_{conv} \cup A, \quad (3.8)$$

儘管這個例子只有一個子集合，但是演算法仍然可以用於任意有限數目的子集合，只要對每個子集合分別求出邊界與給定邊界內側的一點，就可以遞迴式的區域填充完各個子集合。

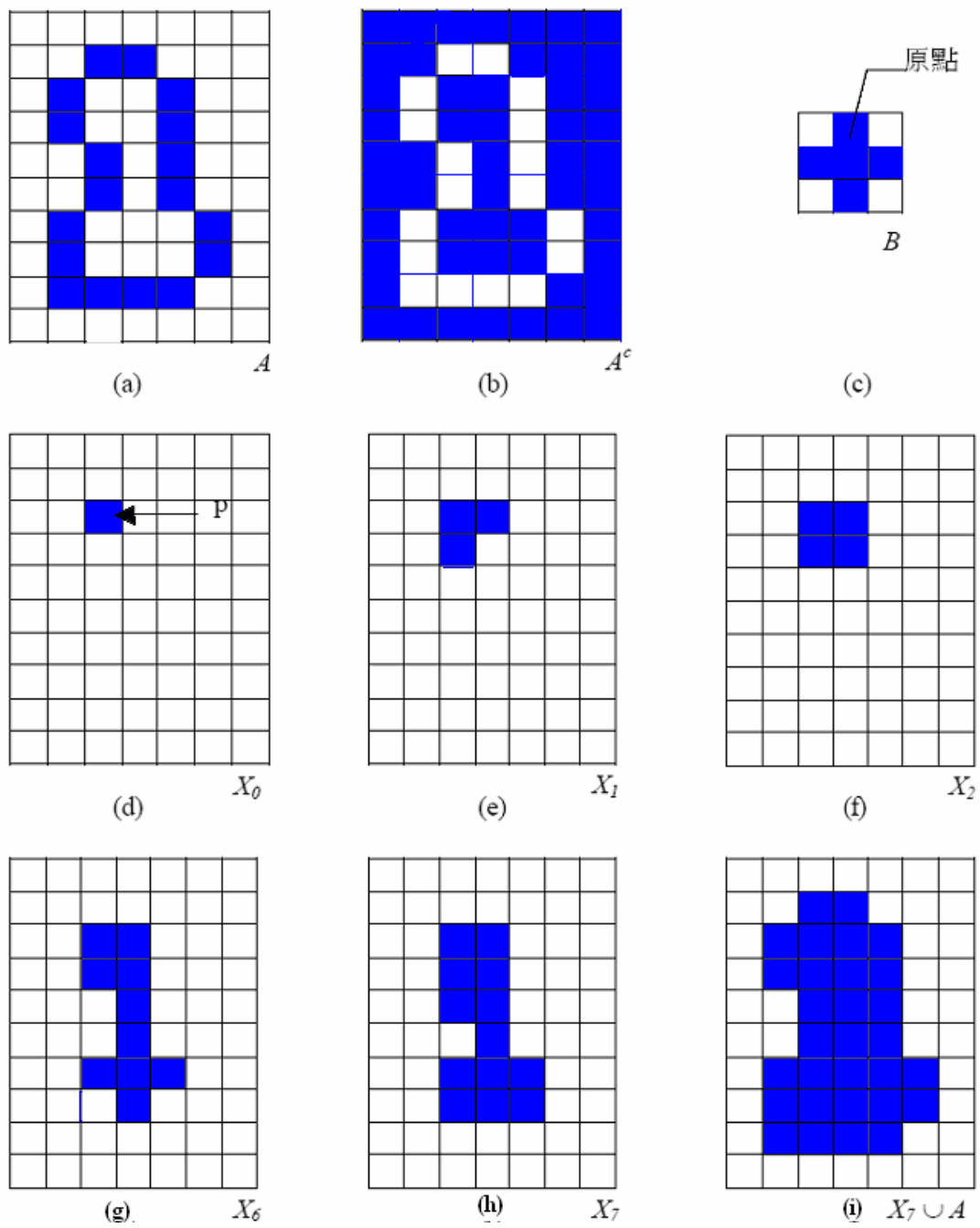


圖3.6 區域填充(a)含有一個邊界子集合的集合A；(b)A的補集；(c)結構集合B；(d)邊界內側的起始點p；(e)~(h)分別為各步驟疊代結果；(i)由(a)和(h)的聯集得到的最後結果[36]

使用區域填充的方法可以將影像中白色部份中的黑色雜訊疊代成為白色的臉部區域，這樣代表在人臉中那些眼睛、嘴巴原本被認為是雜訊的部份，可以經由區域填充的方法保留下以便我們之後的處理。圖 3.8 所示的為經過區域填充後的影像。

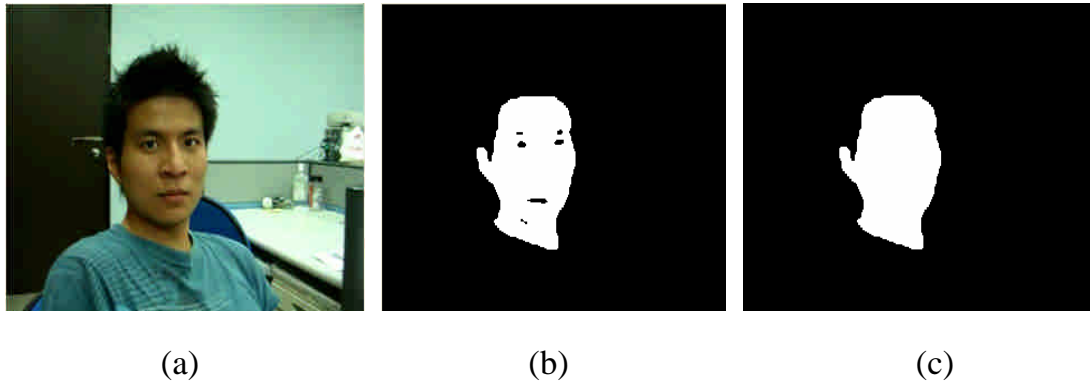


圖3.7 區域填充後的結果 (a)原始影像 (b)經過去除雜訊的膚色分割影像
(c) 區域填充後的影像

在尋找影像中臉部狀態前，必須先找到人臉在影像中的位置。經過雜訊去除的步驟後，我們僅僅是在視覺上達到動態物件的萃取，為了要能得到物體實際在影像中的二維座標位置，我們可以抽取二值影像中的連通成分去取得該連通成分的位置。

以圖3.8[36]為例，假設Y表示包含於集合A的一個連通成分，並假定已知Y的一個點p。則下面疊代表示式產生Y的所有元素：

$$X_k = (X_{k-1} \oplus B) \cap A, k=1,2,3,\dots, \quad (3.9)$$

其中 $X_0=p$ ，而B是一個合適的結構元素，假如 $X_k=X_{k-1}$ 則此演算法收斂，並另 $Y=X_k$ 。方程式(3.9)在形式上類似於方程式(3.8)，唯一的差別是這裡用的是A而不是他的補集。這種差別的發生是因為待尋找的元素(及連通成分的元素)都被標記為1。每疊代一次與A的交集消去了以標記為0之元素為中心上的擴張。

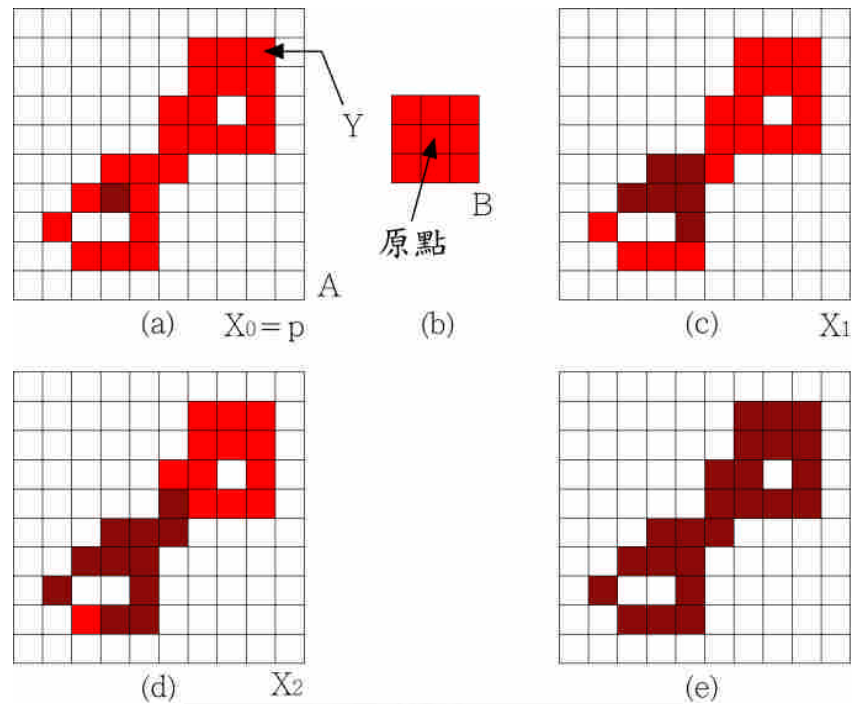


圖3.8 連通成分(a)含有子集合Y的集合A， p 為起始點；(b)結構元素B；(c)第一次疊代結果；(d)第二次疊代結果；(e)最後結果[36]

畫面中的膚色範圍可能會有一個以上，為了能夠分辨不同的膚色範圍，我們建立了一個與畫面同樣大小的矩陣，將二值化圖形轉換成一個二元圖(Binary map)，接著透過上述連通成分的判斷，在與結構元素作用時分別給予不同的編號值。之後則可以利用Binary Map判斷出各個連通物件的座標，依照這個與畫面同樣大小的矩陣索引值找出在x軸與y軸方向上最大與最小的座標值，並以此座標值在畫面中畫出可框出其人臉範圍。如圖3.9所示。



圖3.9 標記人臉候選區(a)單人的結果；(b)複數人的結果

3-4 眼睛和嘴唇的偵測

在定出人臉區域以後，我們要先抓取眼睛和嘴唇的特徵點，首先利用眼睛在臉部的的位置，根據人體計測資料庫，在人臉長方長區塊內定出在高度的15%到55%區間內為瞳孔大概所在區域，然而在此區域內找出數個可能為眼睛的座標，所以我們還要找出嘴唇的座標來定義最正確的臉部特徵，因為頭會旋轉所以嘴唇的位置會隨著x軸移動由左移到右或由右移到左，所以嘴唇會出現的位置就在高度的40%到80%之間，使用眼睛和嘴唇在臉部出現的範圍來偵測眼睛、嘴唇可以降低運算量，也不會在不可能出現眼睛的位置去找眼睛在不可能出現的嘴唇的位置去尋找嘴唇。我們用表3.2的公式可以找出眼睛和嘴唇可能出現的範圍，再使用之後所介紹的方法在我們定義出眼睛和嘴唇的範圍去尋找眼睛和嘴唇。其中X1-X0表示人臉候選區的寬，Y1-Y0表示人臉候選區的長。在圖3.10中紅色的框表示人臉候選區，黃色的框表示尋找眼睛的範圍，綠色的框表示尋找嘴唇的範圍，在圖3.10[35]中可以發現我們偵測眼睛會嘴唇的位置重疊了，一個正常的人臉眼睛會嘴唇的位置不會重疊，我們是考量測膚色時因為實驗者衣著的關係脖子的部份可能會露出來可能不會，這樣會造成人臉寬高比不固定，所以為了避免定義尋找眼睛嘴唇範圍的錯誤，所以我們就加寬偵測的範圍以對應各種不同的情況。

表3.2 眼睛和嘴唇的範圍座標表[35]

	偵測眼睛的範圍	偵測嘴唇的範圍
左上的X座標	$X2=X0$	$X4=X0$
左上的Y座標	$Y2=Y0+0.15 \times (Y1-Y0)$	$Y4=Y0+0.4 \times (Y1-Y0)$
右下的X座標	$X3=X1$	$X5=X1$
右下的Y座標	$Y3=Y0+0.55 \times (Y1-Y0)$	$Y5=Y0+0.8 \times (Y1-Y0)$

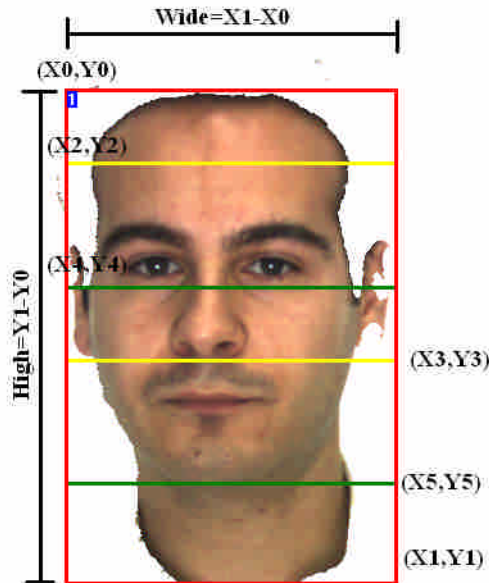


圖 3.10 偵測臉部特徵的範圍[35]

3-4-1 自適性的眼睛候選區偵測

不論人種，眼睛區域相對於人臉膚色在色調上往往都會較為陰暗。因此可以利用眼睛與膚色的差異，讓人眼從臉部區域中分離出來。但若直接用彩色影像進行處理，則要判斷不同人種的眼睛顏色(黑色、褐色、深藍色等)將會是一項繁雜的程序。有鑑於此，我們就希望在眼睛顏色判斷上，減少不同人種眼睛顏色差異的影響。我們發現將彩色影像轉為灰階影像之後，不但可大幅減少色調(Hue)和飽和度(Saturation)的依賴。而且發現眼睛在人臉之中，不管那一人種，將依然呈現出與周圍皮膚灰階度上的明顯差異。因此，我們只需要再訂定出一個門檻值，即可將眼睛從人臉區域中區隔出來。灰階值大約在一個門檻值(65~80)以下，但門檻值以下的灰階值，不一定是人的瞳孔。而且在眼睛部分所形成的黑色區域形狀，應該是長方形、方形或圓形。所以先保留這些形狀的黑色區域，並消除一些不是直線長條形或是垂直方向形狀的黑色區域。所以眼睛是可以經過形狀來分析篩選的。然後再使用增加黑色區域的方法，做法是先將彩色影像轉成灰階的影像，以一個 5×3 的遮罩掃描整幅影像，並改變影像的灰階值。

$$g(x,y)=\begin{cases} \min & \text{if } [f(x,y)-\min < \max-f(x,y)], \\ \max & \text{otherwise,} \end{cases} \quad (3.10)$$

$f(x,y)$ 代表 3×5 遮罩中心的灰階值， \max 、 \min 分別為 $f(x,y)$ 相鄰區域灰階值中的最大最小值，用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，下圖3.11[35]是用公式(3.10)所得到的結果。



圖3.11黑色增強的結果 (a)原本的灰階圖(b)是用公式(3.10)後的灰階圖[35]

由上圖可以明顯看出對比和黑色的程度都增強了，然後在人臉可能區域中依照我們之前定義的眼睛範圍，作為所要搜尋眼睛特徵之範圍，接下來將原圖中相對之區域由彩色影像先轉為灰階圖形，再取一門檻值如公式(3.11)，將小於門檻值的像素定為1，大於門檻值的像素定為0。接下來使用斷開運算，將雜訊點移除。並再次運用連通成份標定程序，求得每一物件的範圍和面積，最後並將所有找出的連通成份區域座標位置記錄下來。經由此眼睛偵測的步驟後即可獲得。

$$\text{Eye}(x,y)=\begin{cases} 1 & \text{if } [g(x,y) < T_{\text{eye}}], \\ 0 & \text{otherwise,} \end{cases} \quad (3.11)$$

我們知道眼睛的面積占整張臉的面積大約2%左右(資料來源：人體計測資料庫)，先把初始閾值 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 T_{eye} 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 T_{eye} 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 T_{eye} 加 1 再計算一次，直到當眼睛的 pixels 數目大於 2%候選區的 pixels 數目。這流程就如圖 3.12 所示[35]。

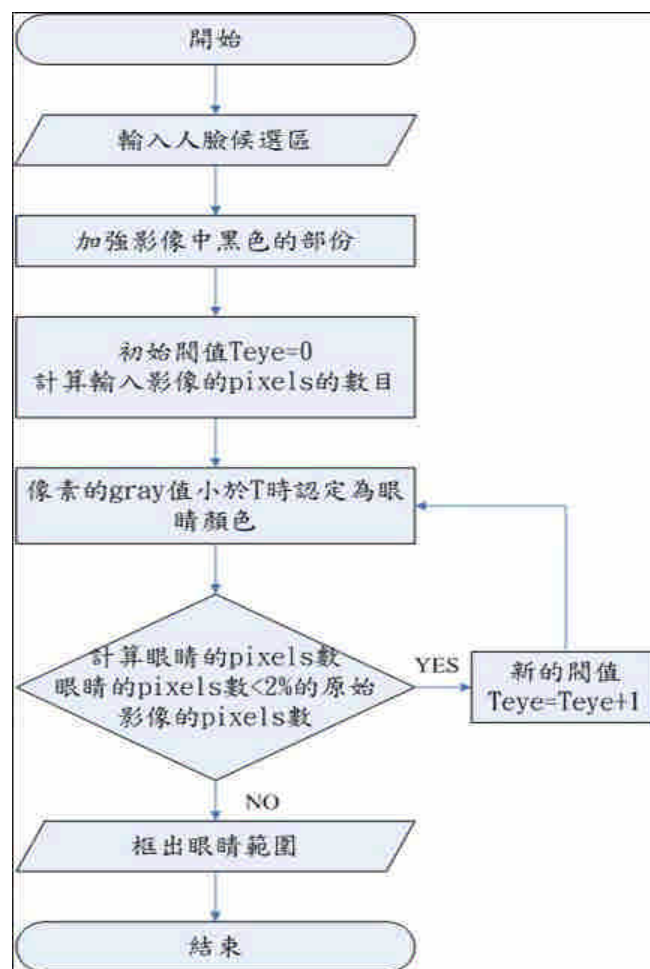


圖 3.12 自適性眼睛偵測流程圖[35]

3-4-2 自適性的嘴唇候選區偵測

同樣使用之前偵測到的人臉可能區域中依照我們之前定義的嘴唇範圍，我們使用色彩分析法發現，皮膚R值(紅色的成分)比嘴唇的R值高，雖然嘴唇的R值不高但相對的比G、B值(綠色、藍色的成分)高出許多，所以我們看嘴唇的顏色才會像紅色。我們統計不同亮度下嘴唇的R、G、B值，統計出來的結果如下表3.3。

表3.3 不同亮度下唇色的RGB值[35]

Y 值範圍	平均 R 值	平均 G 值	平均 B 值	R-G	R-B
51~70	89.5	53.25	34.75	36.25	54.75
71~90	111.33	69.33	50.83	42	60.5
91~110	137.33	83.8	71.13	53.2	66.33
111~130	165.25	103.5	91.13	61.75	74.25
131~150	184.75	122.75	109.88	62	74.57
151~170	207.14	144.14	132.57	63	74.87
171~190	224.17	160.83	151.67	66.5	66.5

我們知道當亮度提高後 RGB 值也會跟著提高，在唇色的部份隨著亮度的提高 R-G 和 R-B 的值也會等比例的提高有上表 3.3 可以得知，在不同的亮度情況 R-G 的值大約都是 R 值的 0.25 倍以上而 R-B 的值大約都是 R 值的 0.3 倍以上，這樣可以先使用 RGB 之間的關係先濾除掉一些在臉部中非唇色的部份，在利用自適性閾值分割正確的嘴唇部分，如(3.12)式。

$$\text{Mouth}(x,y) = \begin{cases} 1 & \text{if } [Cr \leq T_{\text{mouth}}] \text{ and } [R - B > 0.25 \times R] \text{ and } [R - G > 0.3 \times R], \\ 0 & \text{otherwise,} \end{cases} \quad (3.12)$$

如同偵測眼睛的流程一樣根據嘴唇占人臉的比例為 5%(資料來源：人體計測資料庫)，因此我們取 5% 的 pixels 數為嘴唇可能的大小，來決定閾值 T_{mouth} ，當嘴唇的 pixels 數目大於 5% 候選區的 pixels 數目，我們就把 T_{mouth} 加 1 再計算一次，直到符合我們的條件，不同的閾值得到的嘴唇 pixels 數也不一樣。圖 3.13 所示為偵測嘴唇的流程圖[35]。

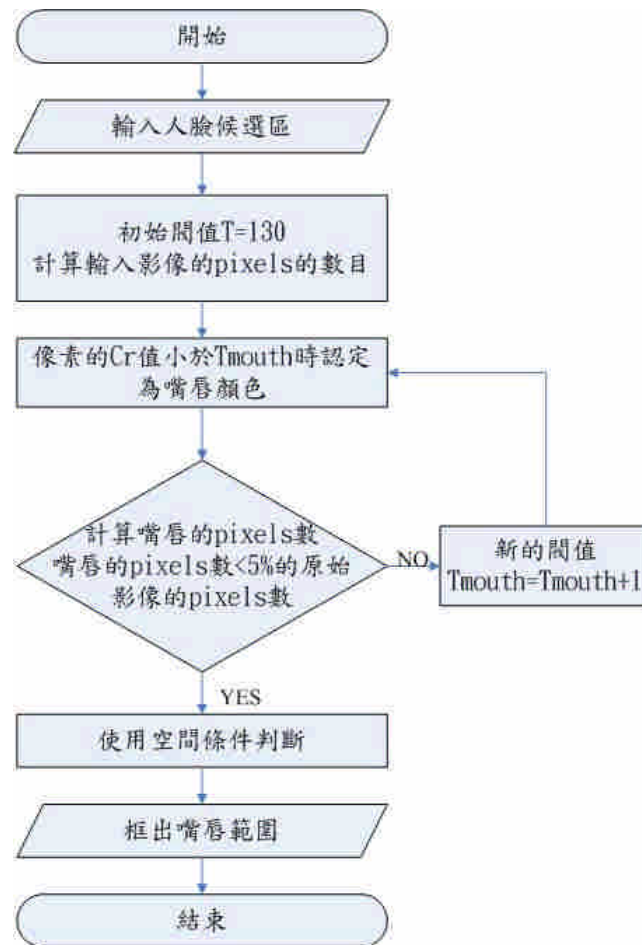


圖 3.13 嘴唇偵測的流程圖[35]

我們在偵測嘴唇候選區時，因為有時候耳朵和臉頰的部份會呈現粉紅色，所以我們在使用連通成份取得嘴唇之前先使用嘴唇的寬高比，因為嘴唇是呈現長方形的樣子，所以我們設定寬高比要大於 1.5 才可以定義為嘴唇，再判斷連通成份最大的一個且符合寬高比大於 1.5 我們才會判斷為嘴唇。

3-5 幾何判斷

找到臉部特徵，利用眼睛、嘴唇之間的幾何關係，由人體計測資料庫發現到眼睛嘴唇之間呈現等腰三角形，以及眼睛出現的位置，我們定出下列規則：

規則一：兩眼之間距離在臉寬的 0.2 倍到 0.8 倍之間。

規則二：兩眼之間的斜率限制於 0.5 以下大約 27 度左右。

規則三：嘴唇的位置必須在兩眼之間。

規則四：兩眼中心點到嘴唇的距離必須大於 0.3 倍的臉長距離。

因為頭部會旋轉所以眼睛的距離在正面人臉的時候是臉寬的 0.8 倍，但是當臉轉動的角度載 0~45 度或是 0~-45 度時兩眼的距離會越來越小，一直到 45 度或是-45 度時兩眼的距離是臉寬的 0.2 倍，所以我們取兩眼的距離必須是臉寬的 0.2~0.8 倍。在我們偵測眼睛時是使用灰階度來偵測眼睛，所以有時會把眉毛或是鼻孔也偵測進去，我們加入斜率的條件因為兩眼會呈現水平的狀態，使用斜率就可以避免錯誤的眼睛對。還有根據嘴唇必須在兩眼之中的條件，也可以濾除掉一些被誤測為眼睛的雜訊。

基於以上的規則，我們可以先慮除掉一些非眼睛的雜訊點，接下來我們用下來公式來定義出正確的人臉。

$$D = d1 + |(d2/d3) - 1| \times \frac{\text{Height}}{2} + d4 \quad (3.14)$$

其中Height是指人臉候選區的高度，d1是指兩眼之間的距離，d2、d3分別指左眼和右眼到嘴唇中心點的距離，d4是代表兩眼距離的中點到嘴唇中心點的距離，如圖3.14所示[35]。因為在人臉候選區中幾使使用上列規則保留下來的可能會有一個以上的眼睛對，所以對於每組眼睛對都會去計算它的D值計算出來D值最小的眼睛對我們就把它定義為正確的眼睛對，我們知道臉部特徵眼睛和嘴唇是呈現等腰三角形，所以d2和d3的距離會相

等，為了避免眉毛對會被偵測為眼睛對我們知道眉毛的位置在眼睛的上方，所以眉毛對的 d_2 ， d_3 ， d_4 都會比眼睛對的值還要高，所求出的D值越小是正確人臉的機率越高，在某些時候因為鼻孔的位置也會呈現低灰階度的情況，因為我們把眼睛的距離設定在臉寬的0.2~0.8倍之間，而鼻孔對和眼睛對都是呈現水平狀態，所以我們加入嘴唇到兩眼中心的位置必須大於臉寬的0.3倍這也是由人體計測資料庫中統計出來的距離。所以我們是取計算出來的D值最小為最正確的人臉，我們所偵測出來的結果會在下一章呈現出來。

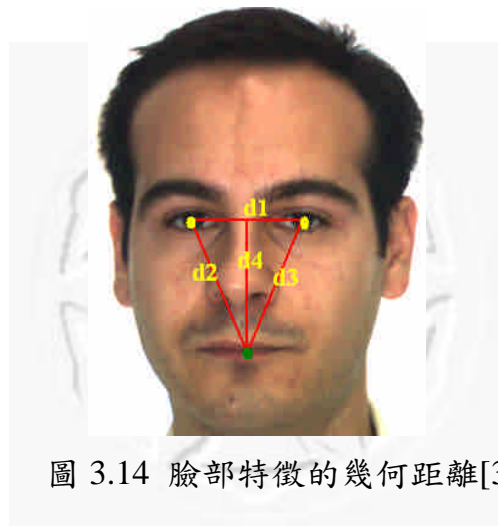


圖 3.14 臉部特徵的幾何距離[35]