

第二章 人臉追蹤演算法

本章中我們將發展一人臉追蹤演算法，以期將其整合在實驗室發展之家用機器人影像平台上，完成在光源變動之下也可達到持續追蹤人臉之效果。在追蹤方面，加入預測及時間性以減少運算量及增加追蹤的穩定性；利用色彩分佈的改變，在不同時刻下得到最佳膚色之色彩分割，藉此得到此刻之人臉位置及投影長寬，再將此資訊傳遞到下一時刻，做為人臉追蹤時所需之資訊；最後，將影像中人臉的位置傳到工業電腦，並由工業電腦控制機器人的頭部及腳部，即使人臉在移動，此頭部平台會將人臉固定在影像平面的中心位置。

2.1 人臉追蹤系統

假設影像中已偵測到人臉（此部分將在第三章更詳細的描述），接下來會將此人臉的位置(MA_X_0, MA_Y_0)及大小(MA_W_0 與 MA_H_0)傳遞到人臉追蹤系統。追蹤系統會根據這些資訊產生出感興趣區域 (Region Of Interest, ROI)，此部分的用意在於縮小搜尋範圍及減少運算量。在感興趣區域裡面，我們將提出適應性膚色搜尋 (Adaptive Skin Searching, ASS) 找出最佳的膚色色彩分佈，並且將人臉的範圍框選出來，接著若符合以下兩個條件就把此刻的人臉位置及大小傳遞到下一時刻輸入影像中，以設定下一時刻感興趣區域：第一，利用人臉投影長寬而產生感興趣區域，故人臉區域投影出來的面積 (Mapping Area, MA) 會占感興趣區域面積一定的比例，若投影的面積大於或小於此感興趣區域面積某一程度的話，即表示前一時刻選擇的人臉位置及大小不是最佳的，導致此刻人臉的位置沒有在設定之感興趣區域中心附近，即表示追蹤將失敗；第二，假設人臉的移動在一定的速度以內，且利用本論文發展之嵌入式影像平台進行即時性影像擷取，我們認為人臉投影寬度的變化是很微小的，所以我們將人臉追蹤時前後兩張影像所得到之人臉投影寬度差值，設定為第二個判斷是否追蹤成功的條件，此條件為：當 $t-1$ 時刻的投影寬度和 t 時刻的投影寬度差值之絕對值小於某一閾值，即表示人臉追蹤成

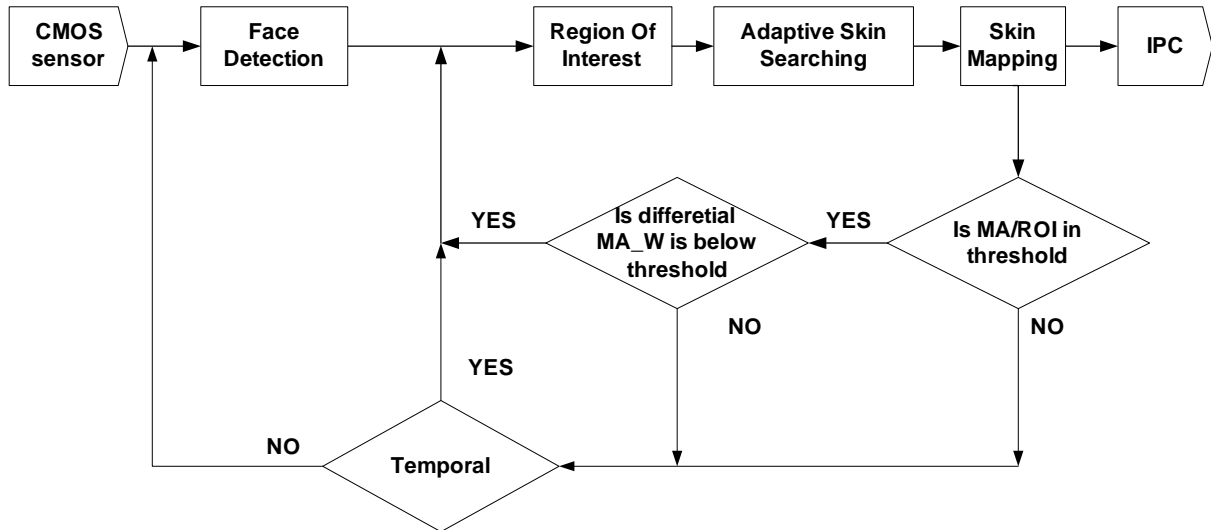


圖 2-1 人臉追蹤之流程圖

功。在人臉追蹤穩定性方面，我們採用時間性（Temporal）來改善追蹤之穩定性。當 t 時刻之投影面積及前後臉部寬度差值此兩個條件都失敗的話，我們會利用 $t-1$ 時刻有效的人臉位置及大小再進行 $t+1$ 時刻輸入影像之人臉追蹤。圖 2-1 為人臉追蹤之流程圖。



2.2 適應性膚色搜尋法

在移動方面，藉由時間性及預測下一次人臉出現的位置及大小，縮小處理的範圍，並且省略不必要的膚色搜尋，以達到即時追蹤及穩定性的效果；利用色彩空間模型 YCrCb 三維色彩分佈進行人臉分割，雖然膚色之色彩分佈是隨時間及光源變動，但是我們藉由前一時刻的人臉位置及大小，找出此刻最佳之膚色色彩分佈，並且在感興趣區域內進行色彩分割及投影，再將此刻的人臉位置及大小再傳遞下一時刻輸入的影像。詳細的設計將以下各小節說明及描述。

2.2.1 感興趣區域

在膚色區域投影中，最常發生的問題是閾值的選定及判斷數個投影區域何者才是正確的人臉區域，以下就這兩部分別討論其解決的方法。

第一，臉部的大小也會因為人與機器人之間的距離而改變，所以在閾值的設定方

面，如果人和機器人之間的距離遠於某個距離而造成人臉區域長寬之數值小於設定之閾值，則此膚色區域則會被忽略；反之，若閾值設定太小的話，則一些類似膚色的區域有可能會被選擇而造成判斷失誤。我們提出利用感興趣區域來解決設定閾值的問題。我們利用 $t-1$ 時刻之人臉位置在 t 時刻人臉的附近產生一區域，此區域我們稱之為感興趣區域。我們定義投影的長度及寬度分為 MA_H 及 MA_W ，感興趣區域之長寬分別為 ROI_H 及 ROI_W 。公式(2-1)及公式(2-2)表示感興趣區域的長寬是依據前一次人臉投影的長寬。

$$ROI_W_t = MA_W_{t-1} \times 1.5 \quad (2-1)$$

$$ROI_H_t = MA_H_{t-1} \times 1.5 \quad (2-2)$$

我們假設此區域的大小在人臉移動的時候，都可以將人臉放置在此區域中心的附近位置，藉此方法我們可以設定投影的閾值，如公式(2-3)及公式(2-4)所示。在此區域內有一適應性膚色視窗 (Adaptive Skin Windows, ASW)，其功用在統計此視窗內之膚色色彩分佈，此部分將在 2.2.3 節有更詳細的說明。因為 ROI 的長寬為臉部長寬的 1.5，所以投影之後臉部最長及最寬的距離大約是 ROI 最長及最寬長度的 $2/3$ ，若考量其他不確定因素，

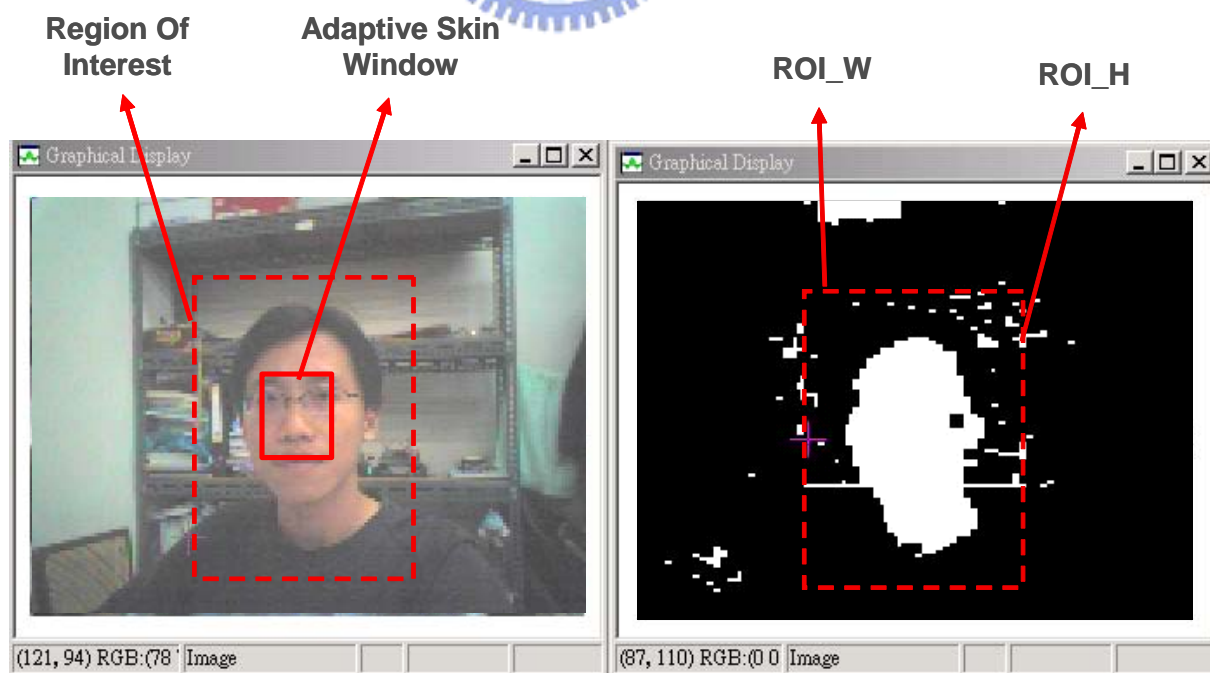


圖 2-2 感興趣區域之設定

例如光源對眼鏡造成反射而在眼睛的附近形成非膚色之色彩分佈，故我們設定當投影的長寬大於感興趣區域長寬一半時即表示有膚色區域的存在。

$$\text{Face_width_threshold} > \text{ROI_W} / 3 \quad (2-3)$$

$$\text{Face_height_threshold} > \text{ROI_H} / 3 \quad (2-4)$$

第二，在判斷數個投影的區域何者才是正確的區域方面，經由適應性視窗統計膚色的色彩分佈，再利用此色彩分佈在整張的影像中進行色彩分割，如圖 2-3 所示，在實際的環境當中會找到其他類似膚色的像素，若利用整張影像做膚色區域投影的話，不僅會將不是人臉區域框選出來而降低準確性，而且也會浪費運算量，所以透過感興趣區域的選取可以減少不必要的運算及判斷，而且可以更集中在處理此區域內之膚色的色彩分佈。

在判斷人臉追蹤是否成功方面，因為投影面積和 ROI 的面積比率為百分之 40 左右，由公式(2-5)可知，所以如果投影面積比率介於合理的範圍內，則表示追蹤成功；反之，若面積比率大於或者小於某一程度的話，即表示此追蹤條件失敗。

$$\text{MA} / \text{ROI} * 100\% = 40\% \quad (2-5)$$

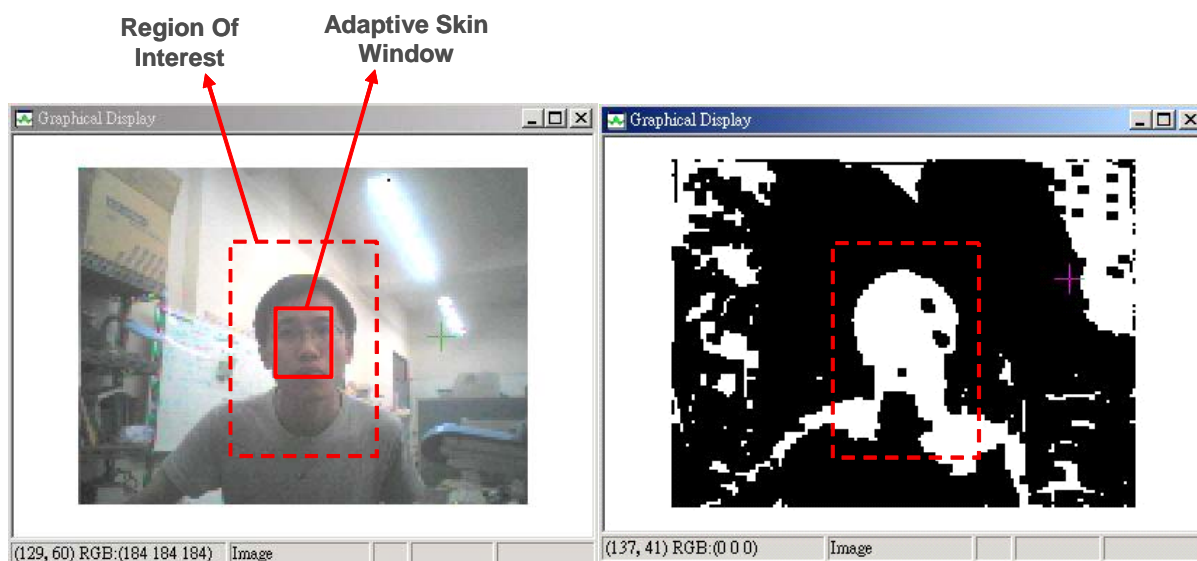


圖 2-3 膚色區域搜尋及追蹤條件設定

2.2.2 色彩轉換與分割

膚色常被用於作為人臉偵測的主要特徵之一，由感測器攝影機所擷取得到的影像資訊多為 RGB (紅、綠、藍) 三原色所構成，但 RGB 三原色所構成的色彩空間模型 (Color Space Model) 容易受到光線的影響，會隨著光源的變化而有很大的變動，所以 RGB 色彩空間模型並不適用於亮度變動大的環境中進行膚色與非膚色的分割。藉由色彩空間模型轉換，可以將其轉換到對亮度變化敏靈度低的色彩空間模型，以避免色彩資訊隨著亮度變動而產生劇烈的變化。目前除了 RGB 色彩空間之外，常見的色彩空間還有 Normalized RGB、HSV、HSI、YCrCb 等等，為了使系統能即時處理及適應大部分的環境，通常會選擇色彩空間轉換公式簡單且對光的靈每度低的色彩空間。

在相關的研究當中，文獻[20]利用在 YCrCb 及 HSV 的色彩空間中作膚色分割的探討，並提供了在這兩個色彩空間中分別找到的判別式。我們經由實驗室發展之嵌入式影像系統進行影像分割的測試，如圖 2-4(a)為包含人臉資訊的測試影像，經膚色分割處理，得到結果如圖 2-4(b) 所示。YCrCb 是一種色彩空間，其中 Y 的定義是亮度資訊，而色度資訊是由 C_r 及 C_b 所組成，其中 C_r 代表了光源中的紅色分量， C_b 代表了光源中的藍色分量。



(a) 原始影像

(b) 膚色分割

圖 2-4 膚色分割測試及結果

2.2.3 色彩分佈閾值

由觀測的結果顯示亮度對會色度產生影響，若只有使用色度的資訊，在色彩分割的準確性而言可能會稍嫌不足。本論文主要是採用 YCrCb 此色彩空間，所以除了使用 Cr 及 Cb 色度資訊之外，也要考慮亮度 Y 對色度產生的影響。圖 2-5 為人臉平均亮度對平均色度之分佈圖，主要是根據圖 1-2 所統計之色彩分佈，此色彩分佈是利用第四章本論文發展之嵌入式影像平台所擷取之人臉區域分析而得到之資料。由圖 2-5 可以得知，當人臉平均亮度愈高時，對應之平均 Cr 值愈高且平均 Cb 值愈低，Cr 值及 Cb 值分別代表為紅色分量及藍色分量，所以色度之紅色分量 Cr 值也會愈高，藍色之分量 Cb 值也就愈低；反之，當人臉平均亮度愈低時，則平均 Cr 值降低且平均 Cb 值增高，並且兩數值會愈逼近數值 128。

色彩分割就是找出膚色收斂的範圍並且設定膚色之色彩分佈上下閾值，若設定的範圍太大，則會將不屬於膚色的背景判定為膚色像素；若設定之閾值小於膚色分佈的範圍的話，當人臉受到光源照射或者臉部過暗時，則無法將人臉膚色的區域與背景分離出來，

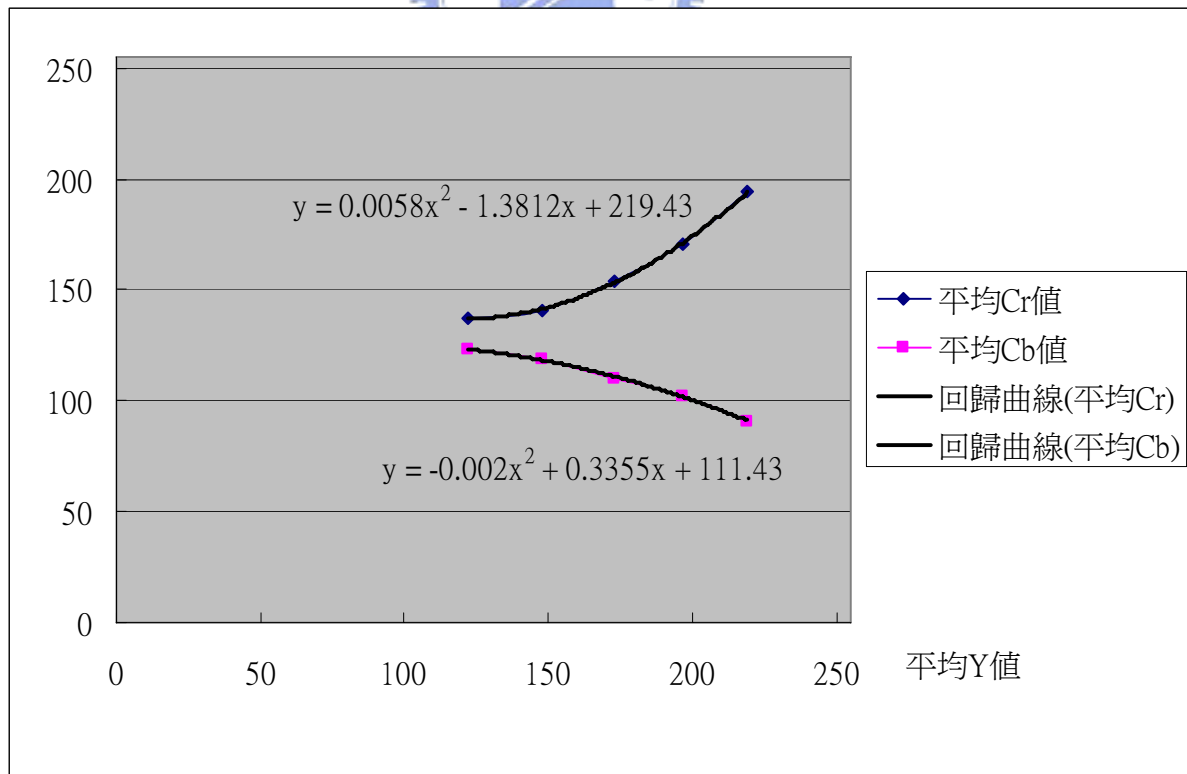


圖 2-5 平均亮度對平均色彩之分佈圖

所以若能在某一時刻設定最佳之膚色色彩分佈閾值，就能將屬於膚色之像素正確地和背景分離出來。

本論文利用統計方法得到人臉膚色在不同的平均亮度 Y 之下，以得到色彩分割所需之色彩分佈閾值。我們利用圖 1-2 五組不同的色彩分佈為基礎，並且使用回歸分析以得到連續平均亮度對應 Y、Cr 與 Cb 之上下閾值。在亮度之上下閾值方面，定義 Y 的下閾值為 Y1，上閾值為 Y2，並且根據圖 2-6 之五組不同的亮度色彩分佈，分別利用公式(2-6)近似 Y2，公式(2-7)近似 Y1，

$$Y2 = -0.0002y^3 + 0.0816y^2 - 11.709y + 682.61 \quad (2-6)$$

$$Y1 = 2E-05y^3 - 0.0042y^2 + 0.4557y + 55.38 \quad (2-7)$$

另外，再利用圖 1-2 的五組色彩分佈得到平均亮度對應 Cr 與 Cb 的上下閾值範圍。我們定義 Cr 的上閾值為 Cr1，下閾值為 Cr2，並且根據圖 2-7 之色彩分佈的情況，以得到平均亮度對應 Cr 之上下閾值。在平均亮度對應 Cr 之上下閾值，主要是利用公式(2-8)近似 Cr2 和公式(2-9)近似 Cr1，

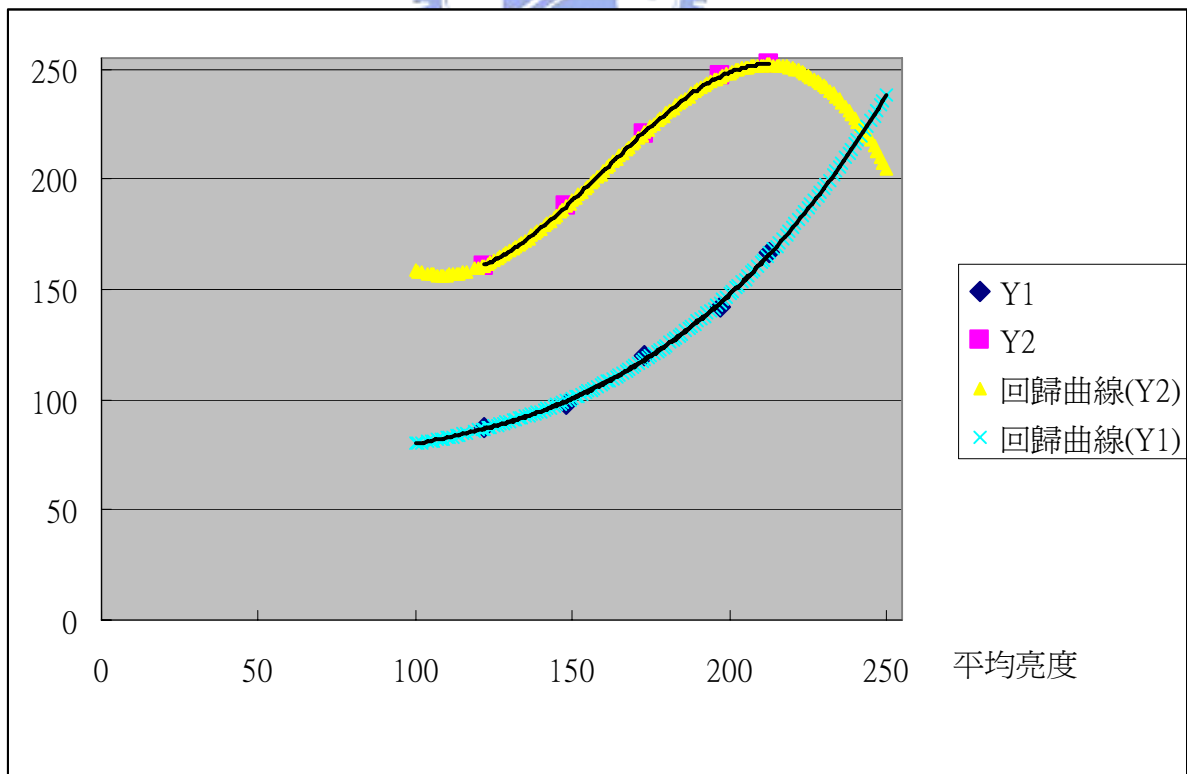


圖 2-6 平均亮度對上下 Y 閾值之分佈圖

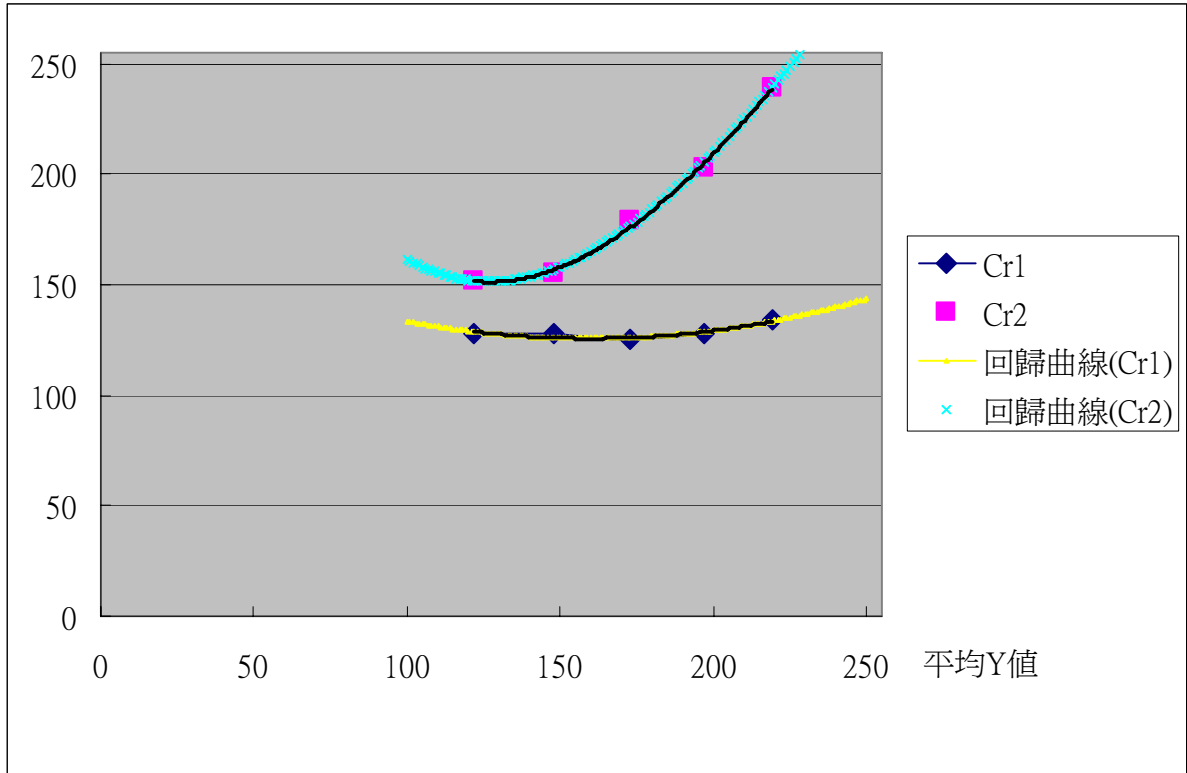


圖 2-7 平均亮度對上下 Cr 閾值之分佈圖

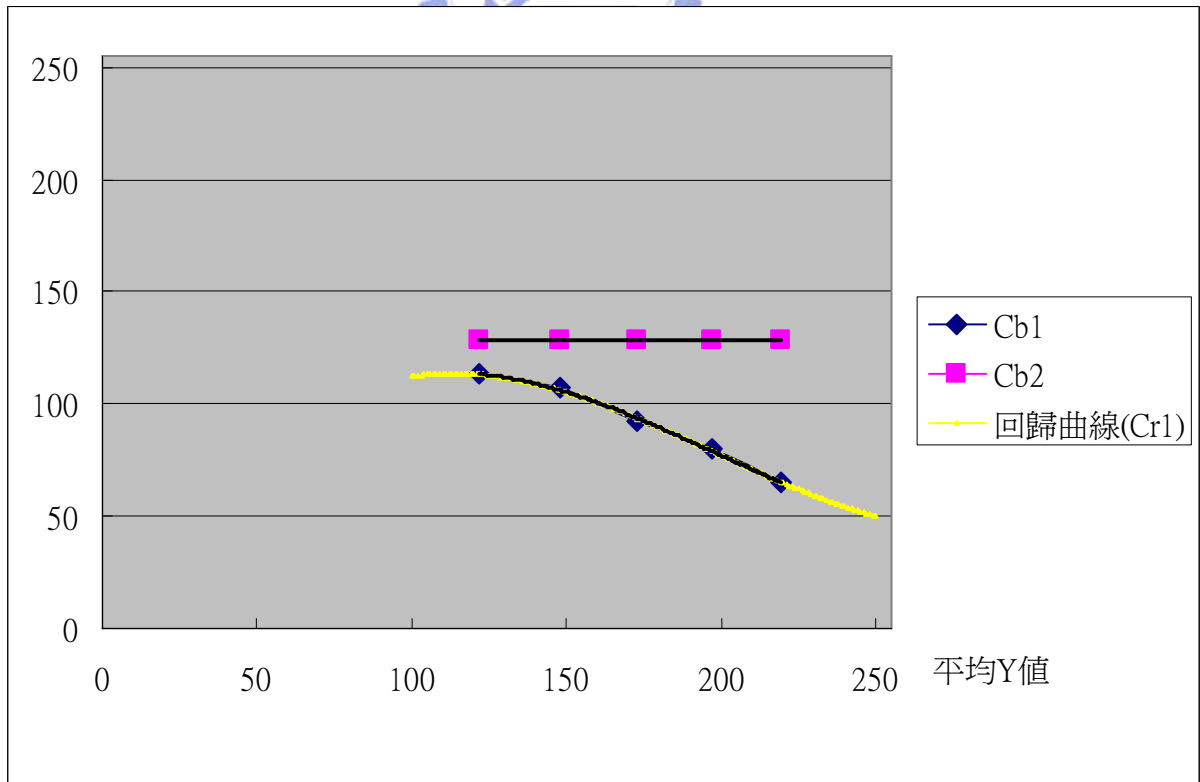


圖 2-8 平均亮度對上下 Cb 閾值之分佈圖

$$Cr2 = -3.11(E - 5) \times y^3 + 2.50(E - 2) \times y^2 - 4.84 \times y + 4.27 \quad (2-8)$$

$$Cr1 = 2.20(E - 3) \times y^2 - 7.02 \times y + 181.72 \quad (2-9)$$

最後，在得到 Cb 的上下閾值範圍和 Cr 的上下閾值範圍的方法一致。我們定義 Cb 的上閾值為 Cb1，下閾值為 Cb2，並且根據圖 2-8 之 Cb 之色彩分佈情況，再利用公式(2-10)表示 Cb2 和公式(2-11)近似 Cb1。

$$Cb2 = 128 \quad (2-10)$$

$$Cb1 = 2.91(E - 5) \times y^3 - 1.73(E - 2) \times y^2 + 2.80 \times y - 23.6 \quad (2-11)$$

當完成以上之步驟後，我們即可事前得到平均亮度 100 到 220 所對應之 Y、Cr 與 Cb 之上下閾值，並且使用查表法 (Lookup Table, LUT) 的方式，來完成在特定膚色平均亮度之下，得到其對應之色彩分佈上下閾值，因此當人臉追蹤系統在運作時，只要利用查表的方式即可得到膚色色彩分佈之閾值，不須要再花額外的運算量求得各個閾值。透過以上的方法，可以精確地將人臉膚色與背景分離出來。



2.2.4 形態學

形態學[21] (Morphology) 為一非線性的影像分析技術，開始之初是應用於圖形辨識 (Pattern Recognition)，後來演進成一門集合理論、非線性濾波及應用系統的概念，此方法為有效率之影像處理演算法、容易在硬體上實現及應用於形狀方位問題等優點。

形態學上的濾波器 (Morphology filter)，主要的運算有侵蝕 (Erode)、膨脹 (Dilate)、閉合 (Closing)、及開啟 (Opening)，下面就各別的運算方式及效果分述如下：

膨脹與侵蝕運算為形態學之基本運算。膨脹主要是利用遮罩在輸入的影像當中移動，並把求得到的結果存於遮罩的中心位置。假設遮罩 B 為一 3x3 之二元化結構 (Binary Structure)，如公式(2-12)所示，輸入影像 A 為二值化影像，膨脹的演算法如公式(2-13)所示：

$$B = \{1 \ 1 \ 1, 1 \ 1 \ 1, 1 \ 1 \ 1\} \quad (2-12)$$

$$A \oplus B = \{X \mid (B)_X \cap A \neq \Phi\} \quad (2-13)$$

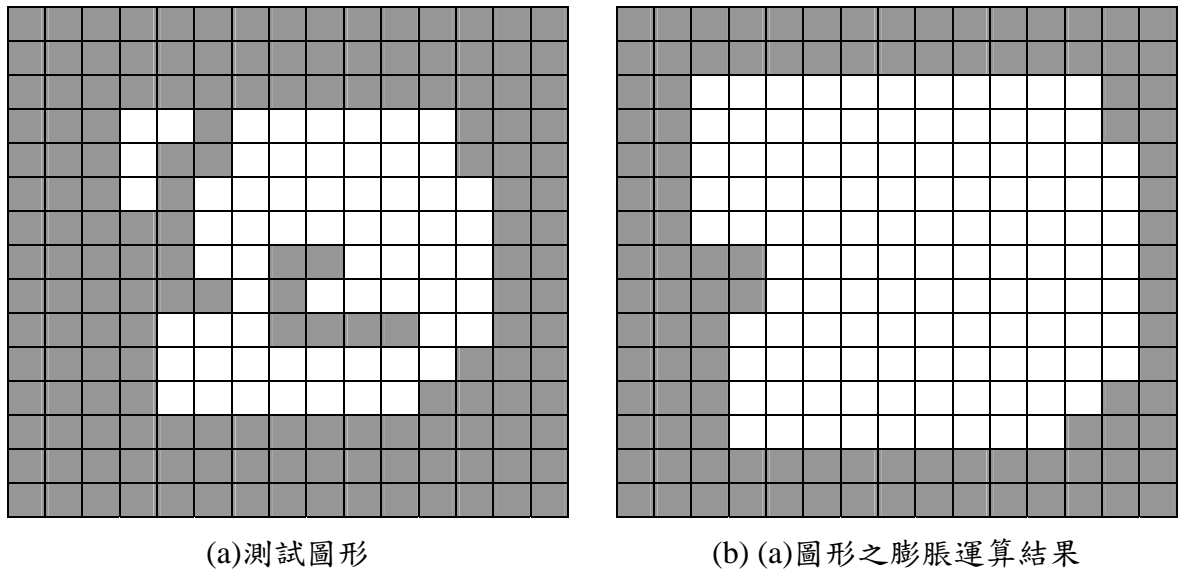


圖 2-9 膨脹過程示意圖

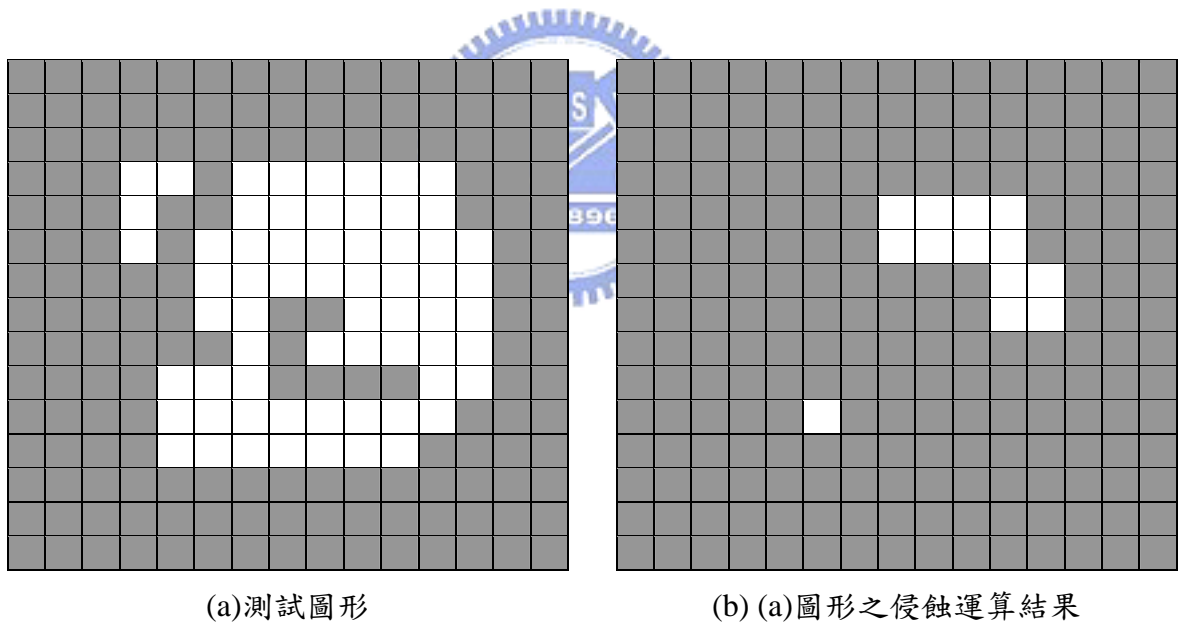


圖 2-10 侵蝕過程示意圖

由結果可以得知，經由膨脹運算之後會擴大原始的影像，而擴大的範圍則由遮罩 B 決定，B 的維度愈高則擴大的面積也愈大。另外，經由閉合運算得到之結果如圖 2-9(b)所示，原本在試測影像 2-9(a)當中有空洞之處也會被填滿。

在侵蝕運算方面，假設遮罩 B 為一 3x3 之二元化結構 (Binary Structure)，輸入影像

A 為二值化影像，侵蝕的演算法如公式(2-14)所示：

$$A \ominus B = \left\{ \left\{ X \mid (B)_x \subseteq A \right\} \right\} \quad (2-14)$$

由結果可得知，經由侵蝕運算之後會縮小原始的影像，而縮小的範圍則由遮罩 B 決定，B 的維度愈高則縮小的面積也愈大。如圖 2-10 所示，侵蝕運算的效果主要是在將原始影像中細小之處，例如影像當中的雜訊，過濾掉以免不必要的干擾。

侵蝕與膨脹可以分別加以組合以得到不同的處理效果。若在二值影像先進行膨脹運算之後再進行侵蝕運算，此方法稱為閉合 (Closing)：

$$A \bullet B = (A \oplus B) \ominus B \quad (2-15)$$

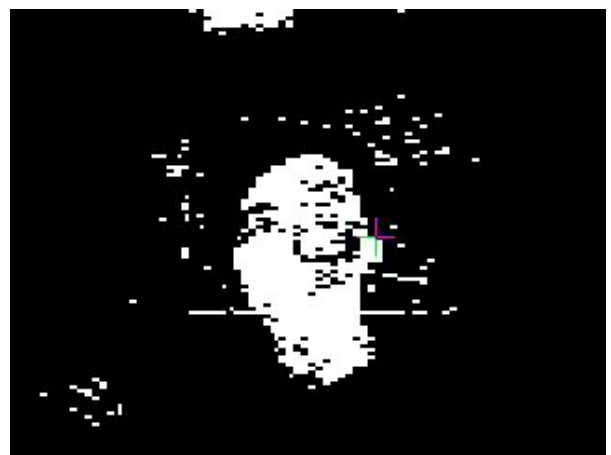
及之，若進行侵蝕運算之後再進行膨脹運算，則稱為開啟 (Opening)：

$$A \circ B = (A \ominus B) \oplus B \quad (2-16)$$

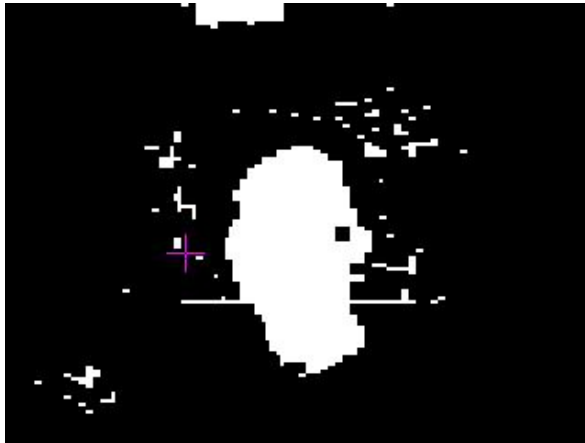
閉合及開啟此兩個方法在形態學中常被使用。圖 2-11(a)為原始測試圖，圖 2-11(b)為經過色彩分割之後所得到之二值化影像，圖 2-11(c)與(d)分別顯示這兩個方法處理過後的結果。閉合運算雖然無法將雜訊過濾，不過卻可以將色彩分佈有洞空之處填滿。開啟運算雖然無法填滿空洞處，但由結果看來卻可以過濾掉大部份的雜訊。因此，此二者可根據不同的需求重覆使用，以增加影像處理之效果。



(a)原始影像



(b)色彩分割



(c)閉合運算



(d)開啟運算

圖 2-11 閉合運算及開啟運算測試

2.2.5 膚色區域投影

利用前面所介紹的色彩分割及形態學的閉合運算之後，著我們希望從這些二值化的影像中找出可能為人臉的區域，以縮小搜尋的範圍。

在此我們參考實驗室陳文俊學長[22]論文所使用之色彩分割方法，利用先前所求得的色彩分割後的資訊 I_{skin} 進行可能人臉區域的判別。我們利用色彩分割後的資訊作水平



(a) 測試原始影像



(b) 膚色分割且經過閉合運算之結果

圖 2-12 膚色分割測試

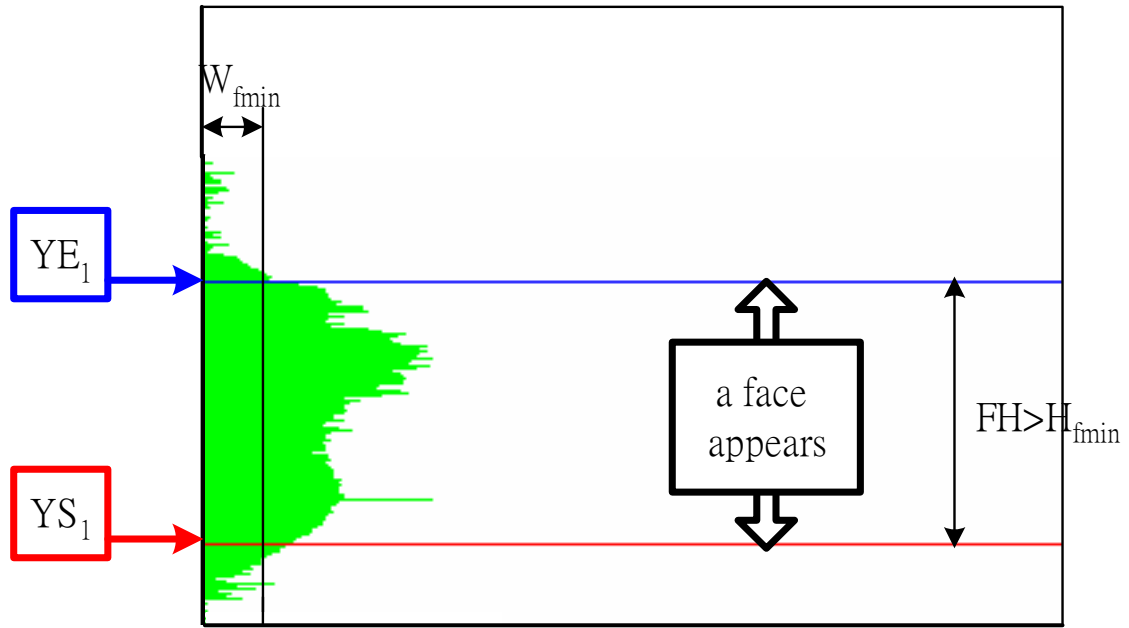


圖 2-13 膚色色彩於 Y 軸的第一次投影

軸與垂直軸上的投影以找出符合要求的區域。前面我們已求得的特定色彩資訊 I_{skin} 為一 $M \times N$ 的矩陣，其中 M 為影像高， N 為影像寬， $I_{skin}(x, y) = 1$ 表示影像中 (x, y) 座標上的像素屬於膚色的色彩分佈，反之，則為非膚色。我們將 I_{skin} 在 Y 軸上進行投影求得 I_{skin} 在 Y 軸上的投影量 $H(y)$

$$H(y) = \sum_{x=0}^M I_{skin}(x, y) \quad (2-17)$$

並設定兩閾值對 H 進行檢測，一閾值為最小可偵測的人臉寬 W_{fmin} ，另一閾值為最小可偵測的人臉長 H_{fmin} ，利用式(2-18)找出可能有人臉的 K 個 Y 方向的區間，如圖 2-13 所示。圖 2-13 為圖 2-12(b)對 Y 軸投影後的結果。

$$\begin{cases} YE_i - YS_i = \left(\sum_{y=YS_i}^{YE_i} \begin{cases} 1, & \text{if } H(y) \geq W_{fmin} \\ 0, & \text{otherwise} \end{cases} \right) \geq H_{fmin} \\ YS_{i+1} > YE_i \end{cases}, i=1 \sim K \quad (2-18)$$

其中 YS_i 及 YE_i 為第 i 個區域的啟始位置及結束位置，且每個區域之間互不重疊。

接著我們再對這 K 個 Y 區間內的膚色資訊利用式(2-19)各別進行 X 軸上的投影，其

投影量各別為 $W_i(x)$ ， i 為 1 到 K

$$W_i(x) = \sum_{y=YS_i}^{YE_i} I_{skin}(x, y), \quad i=1 \sim K \quad (2-19)$$

設定與前面相同的兩個閾值，對 $W_i(x)$ 進行分割，得到

$$\begin{cases} XE_{ij} - XS_{ij} = \left(\sum_{x=XS_{ij}}^{XE_{ij}} \begin{cases} 1, \text{if } W_i(x) \geq H_{fmin} \\ 0, \text{otherwise} \end{cases} \right) \geq W_{fmin}, \quad i=1 \sim K, j=1 \sim L_i \\ XS_{i(j+1)} > XE_{ij} \end{cases} \quad (2-20)$$

其中 L_i 為第 i 個 Y 區間中， X 方向可能有人臉的區間數。 XE_{ij} 與 XS_{ij} 分別為第 i 個 Y 軸區間內的第 j 個 X 軸區間的啟始位置與結束位置。至此，我們就可以概略的找出可能有人臉的矩形區域，如圖 2-14 為圖 2-11(b) 經過上述兩個步驟後得到的結果。

但是有可能會出現圖 2-15 的情況，只進行兩次投影後得到如圖 2-16 的結果，顯示出區域分割的並不完全，所以我們再對各個區間進行一次 Y 軸的投影，以縮小可能有人臉的區域範圍。

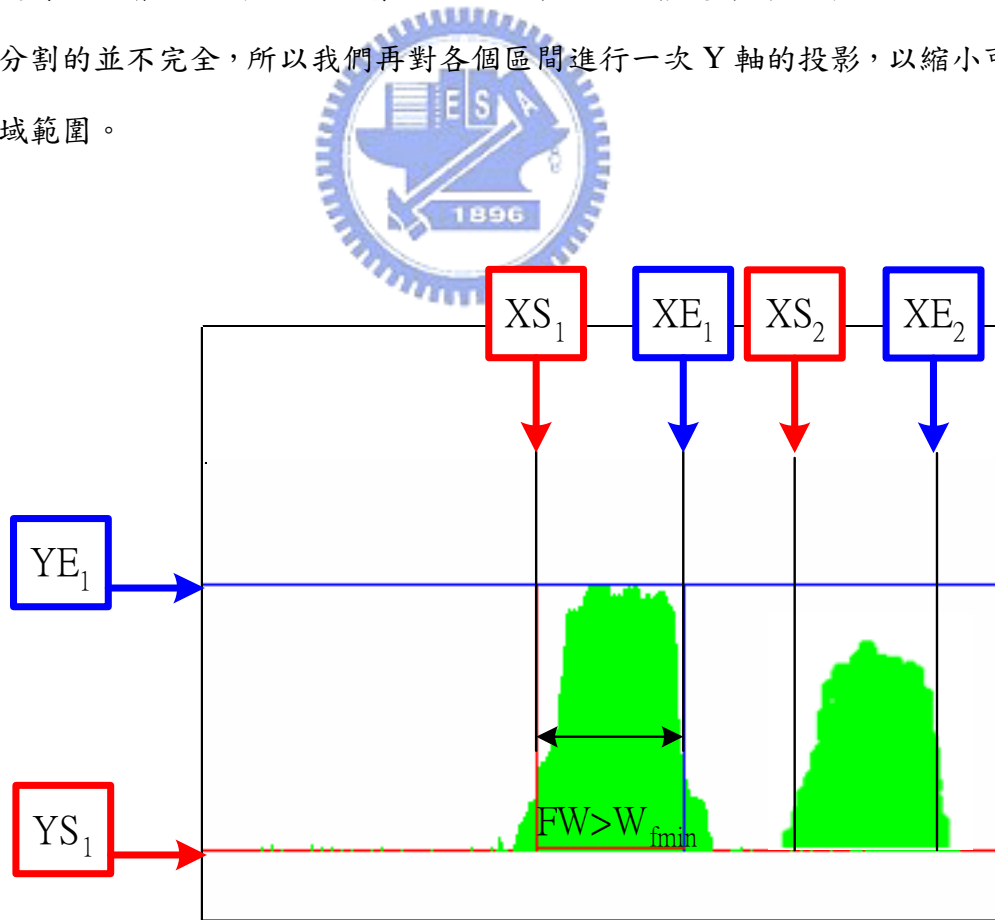


圖 2-14 特定色彩於 X 軸的第一次投影

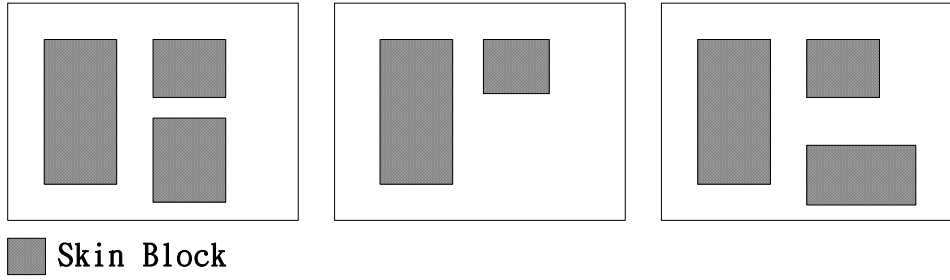


圖 2-15 特殊的特定色彩區塊排列

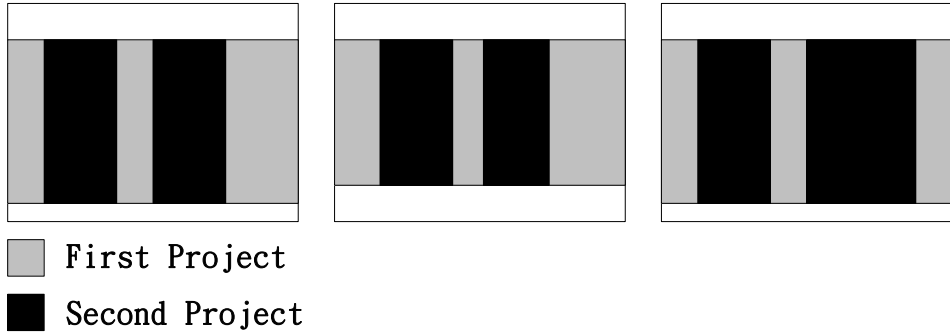


圖 2-16 特殊特定色彩區塊經兩次投影後的結果

$$WH_{ij}(y) = \sum_{y=YS_i}^{YE_i} \sum_{x=XS_{ij}}^{XE_{ij}} I_{skin}(x, y), \quad i=1 \sim K, j=1 \sim L_i \quad (2-21)$$

設定與前面相同的兩個閾值，對 $WH_i(x)$ 進行分割，得到

$$ZE_{ijk} - ZS_{ijk} = \sum_{y=YS_i}^{YE_i} \begin{cases} 1 & ,if \quad WH_{ij}(y) > W_{f \min} \\ 0 & otherwise \end{cases}, \quad k=1 \sim Q_{ij} \quad (2-22)$$

如此我們就可以求得 $\sum_{i=1}^K \sum_{j=1}^{L_i} Q_{ij}$ 個可能含有人臉的膚色區域。各個可能有人臉的矩形

區域之左上角與右下角座標分別為：

左上角： (XS_{ij}, ZS_{ijk})

右下角： (XE_{ij}, ZE_{ijk})

其中 $i=1 \sim K, j=1 \sim L_i, k=1 \sim Q_{ij}$ 。

經由色彩投影之後我們即可知道影像中可能出現人臉的數量及其位置，接著只需再

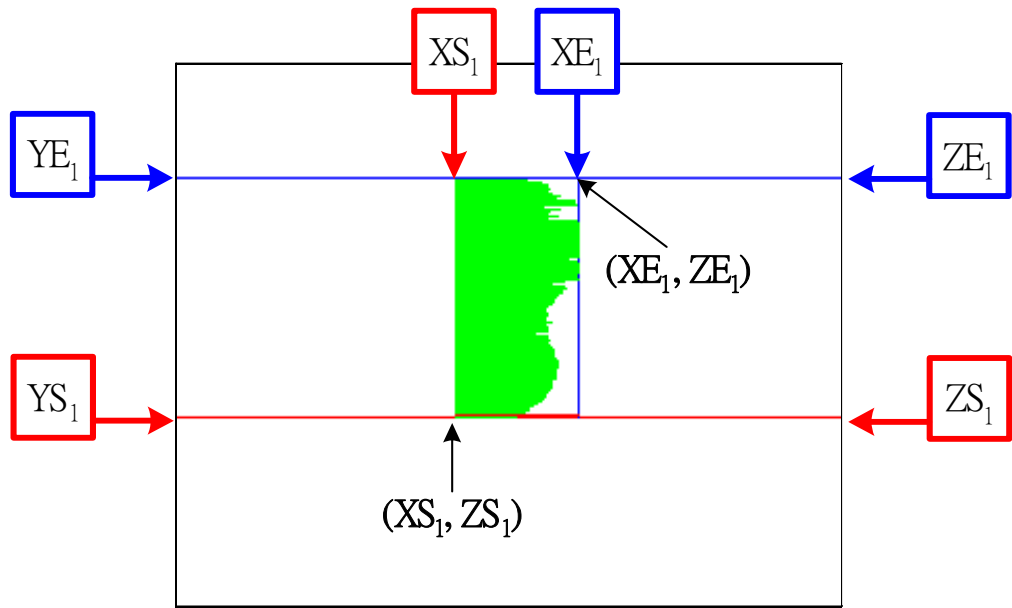
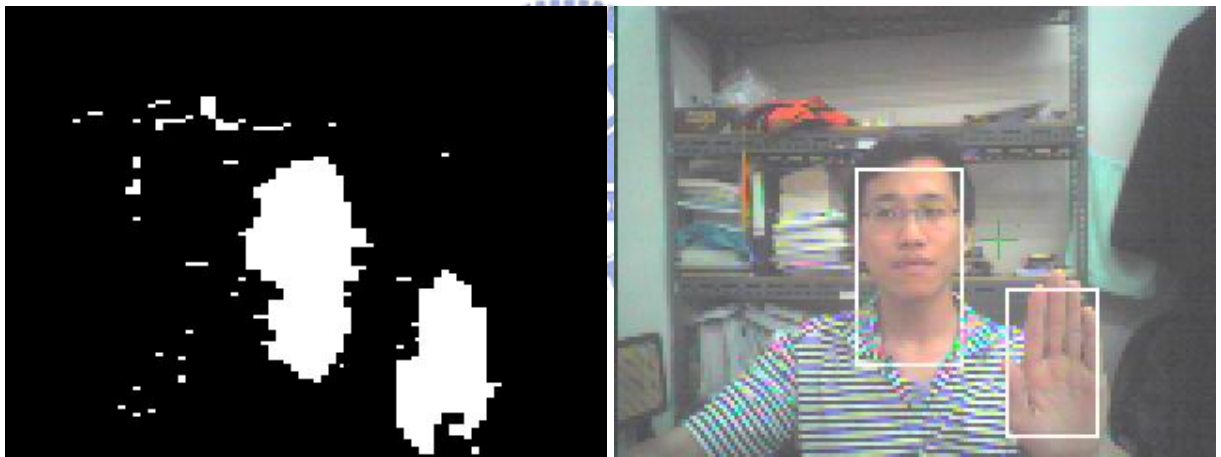


圖 2-17 特定色彩於 Y 軸的第二次投影



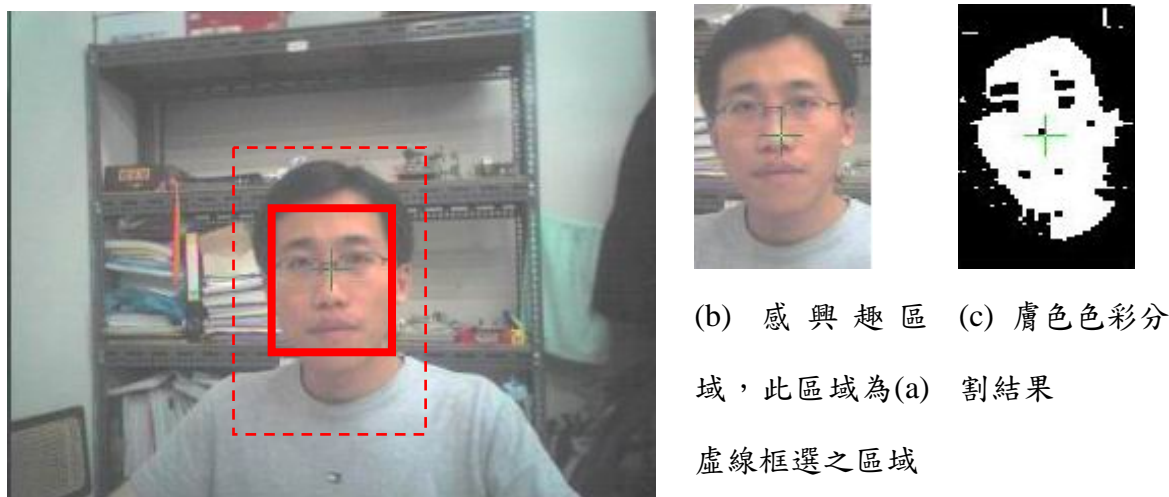
(a) 色彩區域搜尋測試圖

(b) 色彩區域搜尋結果(白色框為搜尋結果)

圖 2-18 特定色彩區域搜尋測試

對這些區域進行臉部特徵搜尋，即可判定此區域是否為人臉，也可得到人臉的位置。圖 2-18(b)為圖 2-18(a) 經過三次投影後所得到的最後結果。

在實際的應用上，只有在初始化狀態，也就是人臉偵測時才需要進行三次投影以找出可能為人臉的區域，在人臉追蹤時只需要進行一次即可。由圖 2-19 所示，因為搜尋



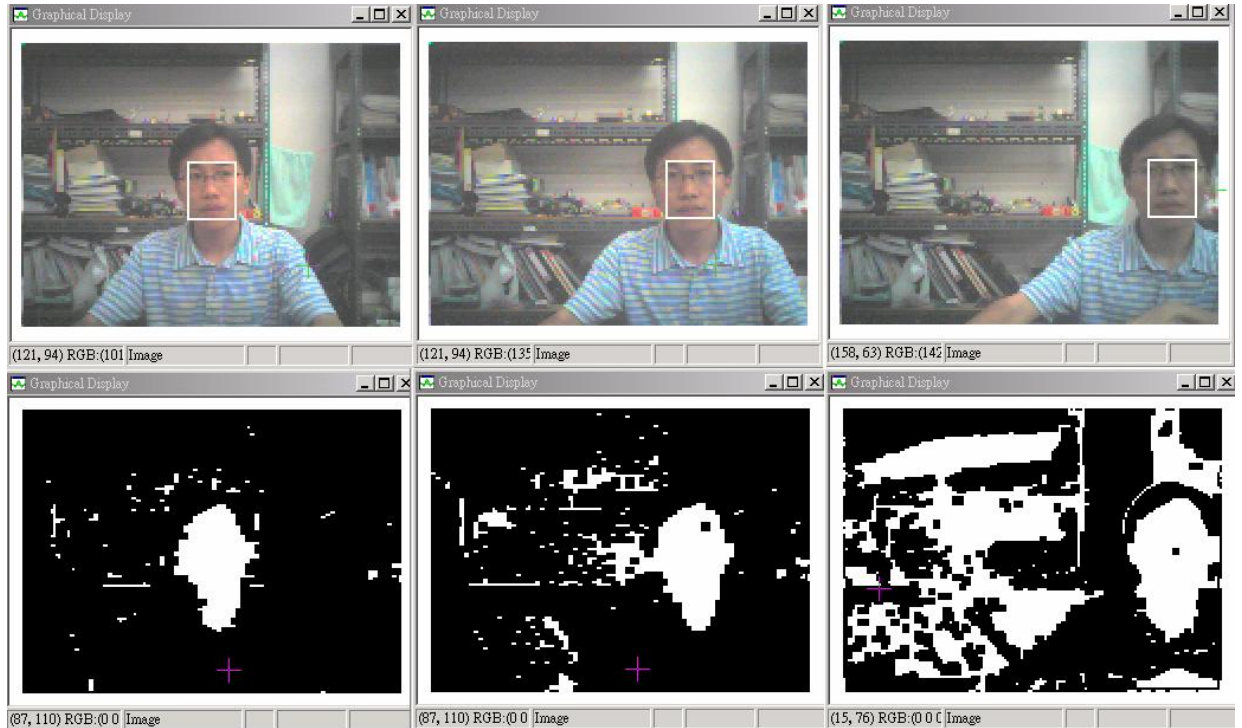
(a) 測試圖，實線區域為(c)之色彩投影結果

圖 2-19 感興趣區域及色彩分割及投影示意圖

的範圍已經縮小為感興趣區域，我們假設此區域內屬於膚色的像素一大半部分是屬於人臉的像素，即使有其他非人臉的像素被判定為膚色，不過所佔之區域有限，除非人臉附近的背景和膚色都很相近，否則人臉投影區域應該佔感興趣區域內最大的面積，所以，在人臉追蹤時只要經過一次色彩投影即可將人臉區域框選出來。

2.2.6 適應性膚色搜尋

此節我們將說明如何利用之前所提及的感興趣區域選取、膚色色彩分佈閾值、色彩分割、形態學與膚色區域投影這些方法，以得到每一時刻人臉所在的位置及大小，並且將 t 時刻之人臉位置及臉部投影長寬傳遞到 $t+1$ 時刻影像平面，此用意在於搜尋 $t+1$ 時刻人臉所在的位置。圖 2-20 為測試影像及膚色色彩分割之結果，上半部影像為測試圖，下半部為經由統計膚色分佈閾值之後的結果。在影像當中有一塊白色框的視窗，此視窗稱之為適應性膚色視窗 (Adaptive Skin Window, ASW)，此視窗為統計膚色分佈閾值之區域，而此視窗內色彩分佈閾值是利用 2.2.3 節之色彩分佈閾值得到的，主要是先計算此視窗內之平均亮度並且根據平均亮度以得到 YCrCb 三維色彩分佈之上下閾值，然後



(a)

(b)

(c)

圖 2-20 適應性膚色視窗搜尋結果。上半部為測試圖，白色框為適應性膚色視窗，下半部為對應之色彩分割結果

再利用這些膚色閾值進行感興趣區域內之色彩分割。由圖 2-20(a)、(b)與(c)可以得知，只要將適應性膚色視窗落於人臉的區域，並且依照此適應性膚色視窗統計出來的色彩分佈閾值進行色彩分割，人臉區域大部分都可以和背景分離出來，即使在光源的影響，如圖 2-20(c)如示，人臉膚色明顯地偏暗，但是透過適應性視窗統計之後的色彩分佈閾值進行色彩分割，還是可以將膚色分割出來，不過相對圖 2-20(a)分割的結果，圖 2-20(c)分割的結果會將許多不是人臉區域的背景也判定為膚色。圖 2-20(b)為 2-20(a)及 2-20(c)膚色色彩分佈中間的狀態，由色彩分割的結果，可再次證明適應性膚色視窗是適用於統計膚色之色彩分佈。由以上的討論為，若此適應性膚色視窗落於人臉區域內的話，則可以有效地將適應性膚色視窗所定義之膚色與背景分割出來。

假設我們已得到前一張影像得到人臉所在的位置及臉部投影的長寬，並依照 2.2.1 節將感興趣區域框選出來，如圖 2-21 所示，則人臉的位置應該會在此區域中心的附近。

接下來，我們依照公式(2-23)定義四個適應性視窗之位置及大小，如圖 2-22 所示，此四個適應性視窗分別為：人臉左上方、右上方、左下方及右下方四個角落。

圖 2-21 為適應性膚色視窗之參數之設定。適應性膚色視窗長度 (Block_H)、寬度 (Block_W)、適應性膚色視窗起始位置 (ASW_X, ASW_Y) 及移動的距離都是根據感興趣區域長寬而設定的，而感興趣區域的長寬則是根據上一時刻之人臉投影長寬而求得到。設定此四個視窗之長寬不可過大或過小，太小的話則有可能無法得到最好的色彩分佈，太大的話則有可能因為人臉移動使得視窗落於人臉外面，而得到錯誤的色彩分佈。在此視窗的起始位置及移動向量，則是根據人臉在感興趣區域的位置以及人臉長寬比 1.2 的條件而設定。

$$\begin{cases} \text{Block_W}_t = \text{MA_W}_{t-1} \times 3/4 \\ \text{BlockH}_t = \text{MA_H}_{t-1} \times 3/4 \\ \text{ASW_X}_t = \text{ROI_W}_t / 4 \\ \text{ASW_Y}_t = \text{ROI_H}_t / 4 \\ \text{Shift_right} = \text{MA_W}_{t-1} / 6 \\ \text{Shift_down} = \text{MA_H}_{t-1} / 5 \end{cases} \quad (2-23)$$

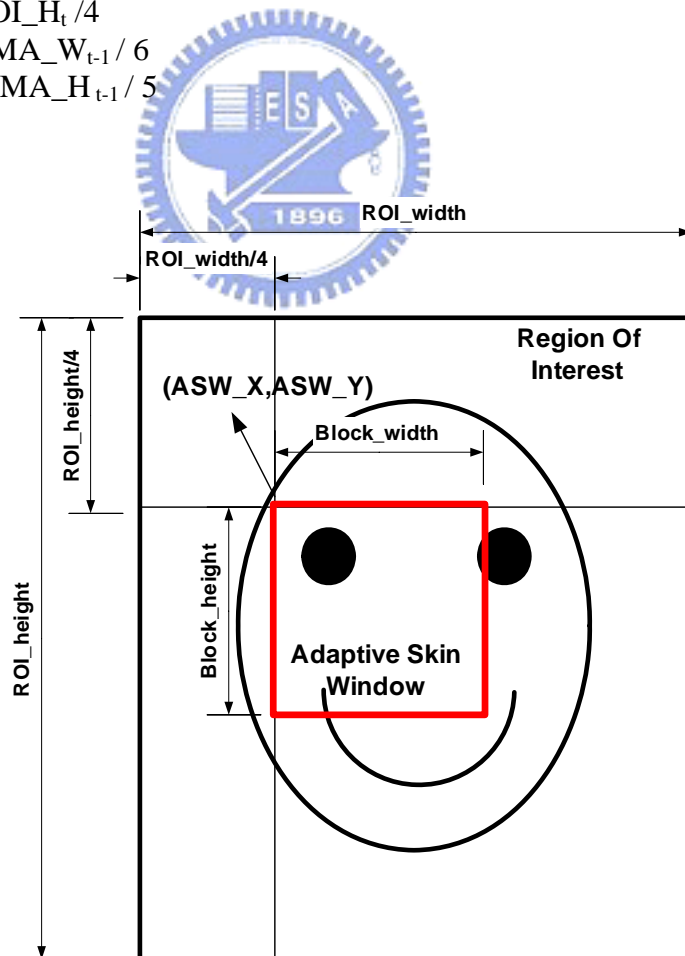


圖 2-21 適應性膚色視窗之設定

圖 2-22 中顯示感興趣區域內之四個適應性膚色視窗，這四個適應性膚色視窗是位於感興趣區域內靠近中心區域附近，分別為：適應性膚色視窗 1 在人臉左上角，適應性膚色視窗 2 在人臉的右上角，適應性膚色視窗 3 在人臉的左下角及適應性膚色視窗 4 在人臉的右下角。選擇這四個角落的用意在於若人臉在某時刻是置中的，則此四個視窗大部分都能落於人臉的範圍內，若人臉因移動而造成人臉不在感興趣區域內的中心位置，則此四個視窗可能有某一個能落於人臉的範圍以內。

我們將先前提及之感興趣區域選取、膚色色彩分佈閾值、色彩分割、形態學與膚色區域投影之處理流程稱為適應性膚色搜尋。如圖 2-23 所示。利用這四次的投影面積來

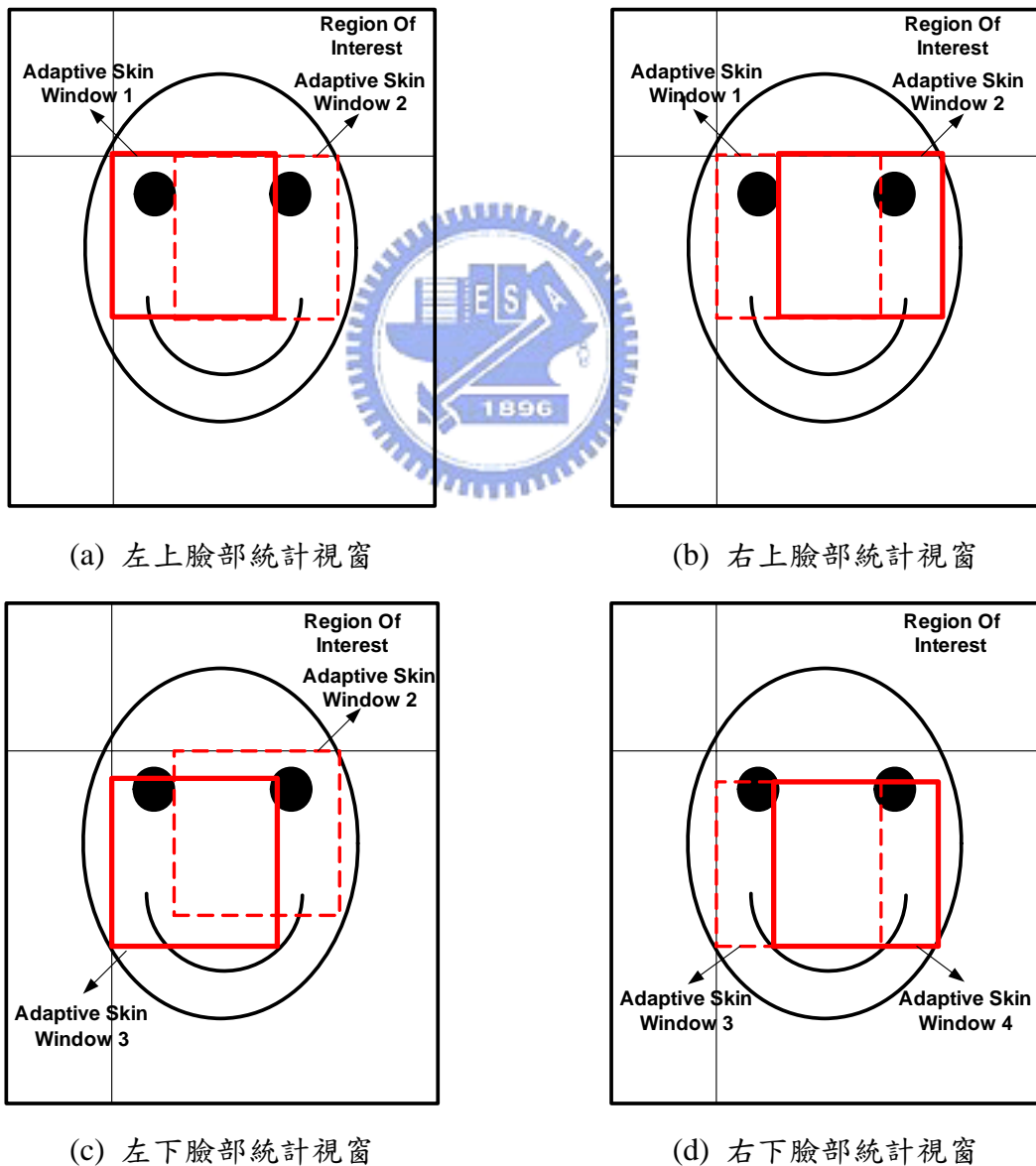


圖 2-22 適應性視窗之選取示意圖

決定何者最符合此刻人臉所在的位置及臉部長寬，並且再根據前後影像臉寬差值的條件來判斷是否追蹤成功，若以上條件成立的話，則將投影區域的位置及長寬等資訊傳遞到下一張輸入的影像，以搜尋下一時刻人臉所在的位置。以下說明為適應性膚色搜尋法之流程，其順序為：

- 1) 當人臉偵測系統偵測到影像平面出現人臉之後，將人臉的投影位置（此時的投影位置為相對於全域影像原點位置）及投影長寬傳遞到人臉追蹤系統，以設定下一刻影像平面感興趣區域及適應性膚色搜尋法之相關參數。
- 2) 利用 $t-1$ 時刻之感興趣區域位置(ROI_X_{t-1} , ROI_Y_{t-1})、臉部投影位置(MA_X_{t-1} , MA_Y_{t-1} ，此時投影位置為相對於感興趣區域原點位置)及投影長與寬 MA_W_{t-1} 與 MA_H_{t-1} ，以得到 t 時刻之 ROI_X_t 、 ROI_Y_t 、 ROI_W_t 與 ROI_H_t ，其對應關係為公式(2-24)所示，此用意在於將人臉置於感興趣區域中心附近的位置。

$$\begin{cases} ROI_W_t = MA_W_{t-1} \times 1.5 \\ ROI_H_t = MA_H_{t-1} \times 1.5 \\ ROI_X_t = ROI_X_{t-1} + MA_X_{t-1} - ROI_W_t/2 \\ ROI_Y_t = ROI_Y_{t-1} + MA_Y_{t-1} - ROI_H_t/2 \end{cases}, ROI_X_0=ROI_Y_0=0 \quad (2-24)$$

- 3) 利用適應性膚色視窗分別計算四個視窗之平均亮度，並且根據 2.2.3 節之色彩分佈閾值分別選擇對應之 $Y1$ 、 $Y2$ 、 $Cb1$ 、 $Cb2$ 、 $Cr1$ 與 $Cr2$ ，再利用此四組色彩分佈閾值在感興趣區域內完成色彩分割及色彩投影。
- 4) 經過四次的適應性膚色視窗搜尋之後，選擇投影面積 $MA_X_t \times MA_Y_t$ 最大且投影面積介於 $ROI_t/3$ 及 $ROI_t/2$ 之間，並且又滿足 $|MA_X_t - MA_X_{t-1}| < MT$ 之條件，若以上兩點都成立的話，則將 t 時刻之感興趣區域之位置(ROI_X_t , ROI_Y_t)、臉部投影位置(MA_X_t 、 MA_Y_t)及投影長與寬 MA_W_t 、 MA_H_t 傳遞到 $t+1$ 時刻做為搜尋 $T+1$ 時刻人臉所在位置，即表示追蹤成功，回到步驟 2。
- 5) 若步驟 4 不成立的話，則利用 $t-1$ 時刻之感興趣區域之位置及投影長寬來搜尋 $t+1$ 時刻輸入之影像平面人臉的位置。若符合以下兩點：第一，投影面積與感興趣

區域面積之比值在合理的範圍內；第二，前後影像所得到之投影寬度之差值小於某一閾值，則表示追蹤成功，並且將人臉追蹤所需之參數更新為 t+1 時刻之人臉位置及大小，回到步驟 2；反之，則表示追蹤失敗，回到步驟 1。

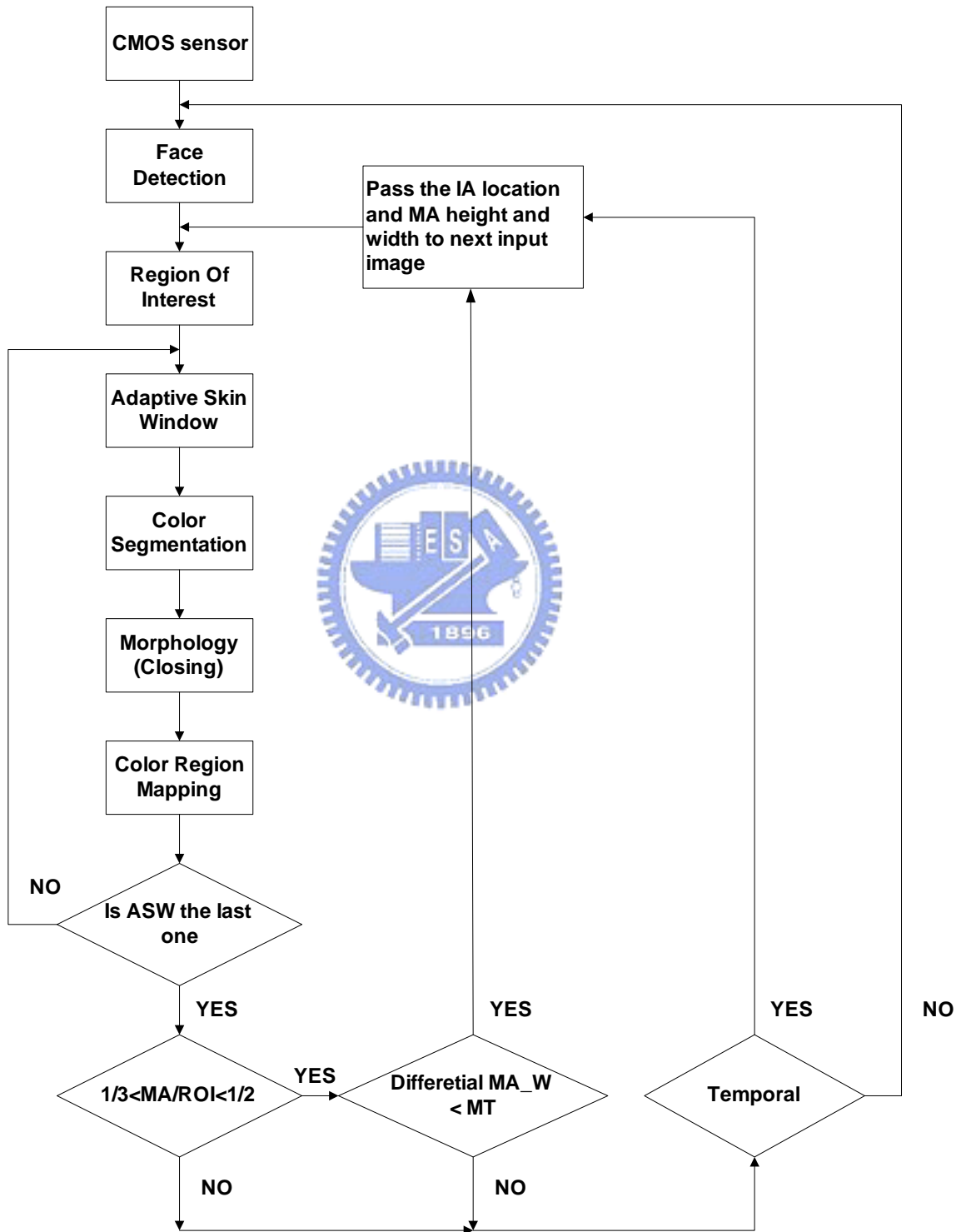


圖 2-23 適應性膚色搜尋法之流程圖

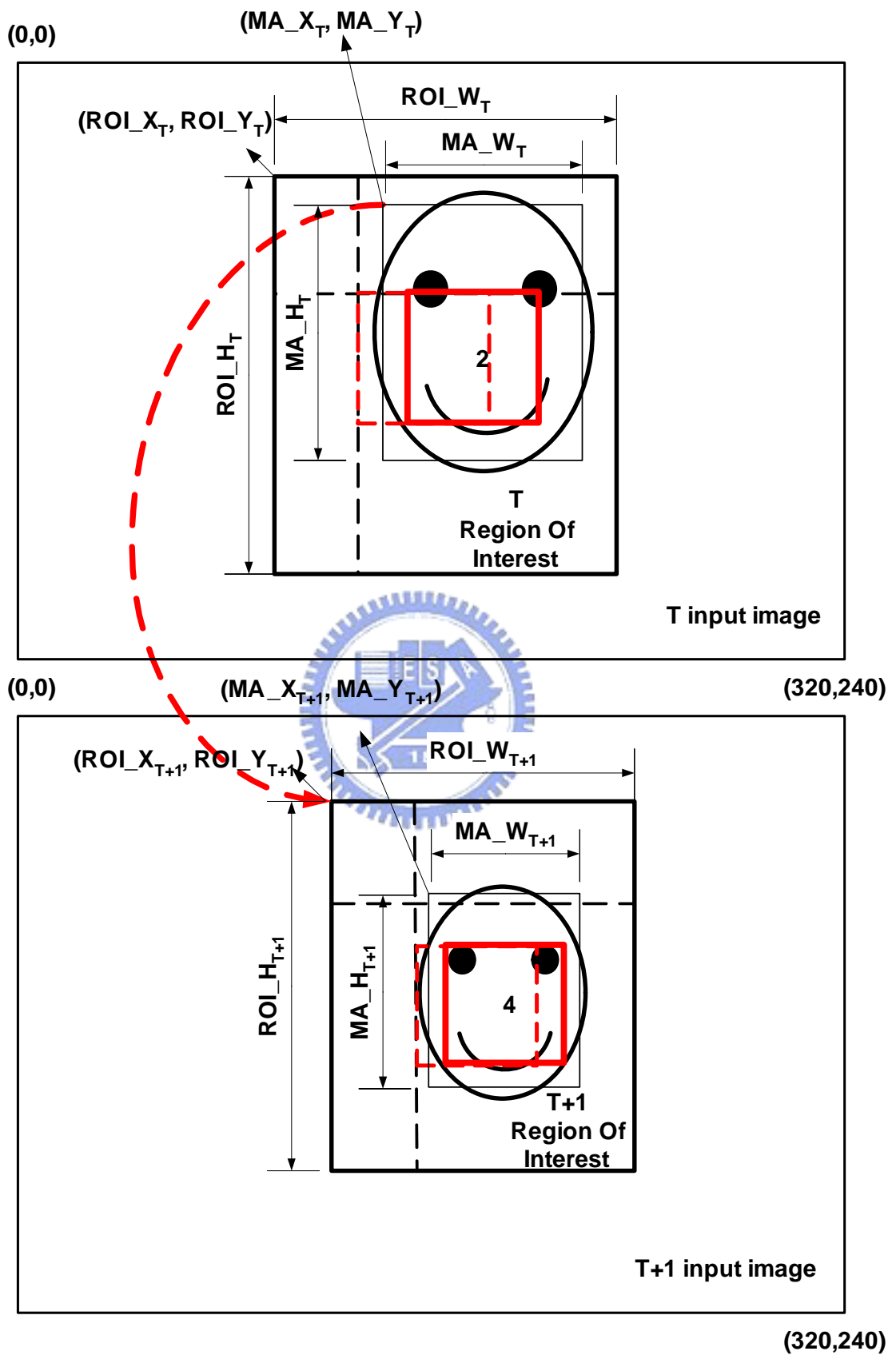


圖 2-24 適應性膚色搜尋法之示意圖

圖 2-24 為適應性膚色搜尋法之示意圖。在圖 2-24 上半部，也就是 t 時刻之影像平台，我們可以看到即使利用 t-1 時刻所傳遞過來的參數設定 t 時刻之感興趣區域，但是人臉卻因為移動而在此區域偏右上方，不過由圖可以得知，適應性膚色視窗 2 落於此人臉區域，即表示有機會得到 t 時刻之人臉投影位置及大小。利用此方法再將 t 時刻之人臉投影位置及大小傳遞到 t+1 時刻以進行人臉追蹤，反覆此動作，直到人臉追蹤失敗為止。

2.3 人臉追蹤之行為控制

完成人臉追蹤系統及偵測系統之後，我們將利用本論文發展之嵌入式影像平台以實現人臉追蹤系統及偵測系統兩個系統，並整合到家用機器人平台，藉由控制頭部運動及機器人本體移動以完成家用機器人之人臉追蹤功能。

在追蹤狀態我們是採用一頭部影像追蹤系統[22]，來達到控制機器人頭部之轉動以進行人臉追蹤，使機器人頭部架設之影像感測器所擷取到影像的中心區域能移到目標物所在的位置上。此系統是基於影像的伺服控制系統，其中包含兩種回授迴圈：1.控制系統影像位置回授迴圈，或稱為外迴路控制。2.頭部機構伺服控制迴圈，或稱為內迴路控制。利用此章人臉追蹤時所得到的人臉位置，再配合以上兩種回授迴圈完成以影像特徵為基礎之閉迴路控制。圖 2-26 為影像追蹤控制器系統架構圖。影像追蹤的目的就是為了使目標能夠保持在影像的中心區域內，所以我們設定其期望位置(Desired Position) (X_d, Y_d) 為 A，如公式(2-25)所示，人臉中心位置為 B，如公式(2-26)所示。

我們定義影像平面中心點為(0,0)，如圖 2-25 所示，假設 A 屬於中心區域內集合的一點，B 為人臉投影之中心點 (MA 之中心點)。

$$A = \{(X_d, Y_d) \mid -2 < X < 2, -2 < Y < 2\} \quad (2-25)$$

$$B = (X_B, Y_B) \quad (2-26)$$

$$\overrightarrow{BA} = ARG(MIN \|\overrightarrow{BA}\|) \quad (2-27)$$

當公式(2-27)成立時，即表示人臉投影中心點 B 到影像平面中心區域 A 為最近的距離，

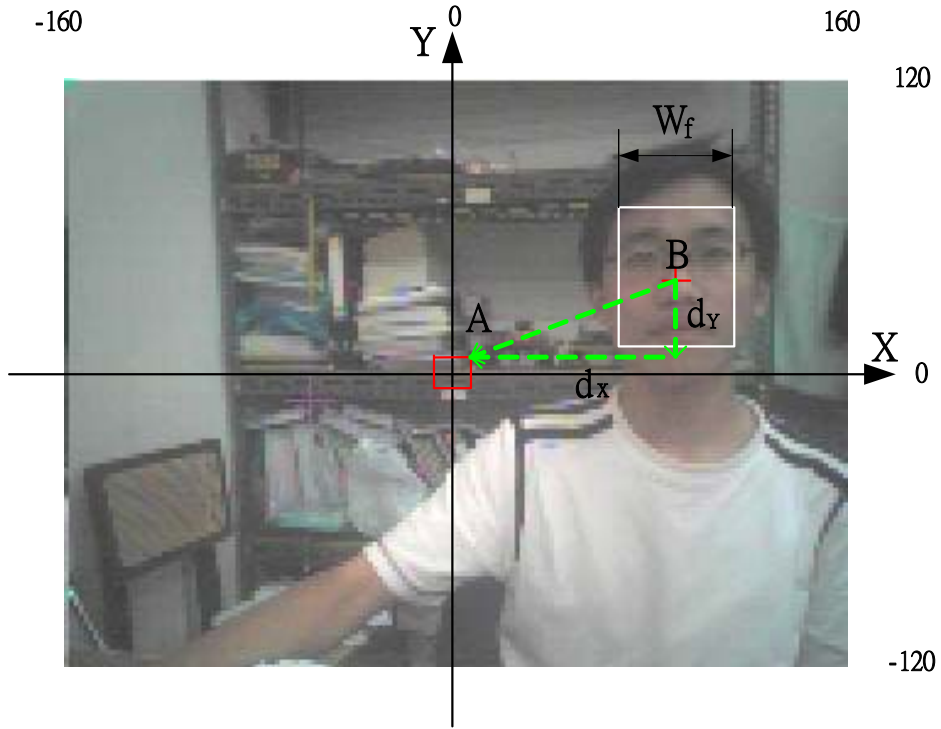


圖 2-25 影像平面及人臉位置坐標

並且利用公式(2-27)求得位置誤差 \vec{d}_x 與 \vec{d}_y ，

$$\vec{d}_x = A_x - B_x \quad (2-28)$$

$$\vec{d}_y = A_y - B_y \quad (2-29)$$

外迴路控制的回授量為人臉在影像中的位置，經由人臉追蹤系統將人臉的所在中心位置 (X_B, Y_B) 標示出來後，與設定之中心區域 (X_d, Y_d) 比較後，得到位置誤差 (dx, dy) ，其單位元為像素。頭部機構的旋轉是以速度為主，透過移動速度轉換將 \vec{d}_x 與 \vec{d}_y 分量轉換成 θ_{pd} 及 θ_{id} 速度控制命令，如公式(2-30)及公式(2-31)所示，gain1 及 gain2 分別為垂直及水平增益係數，此係數愈大，則馬達輸出速度愈慢。透過 PID 控制器完成頭部機構轉動。在系統中內迴路控制，主要在確保馬達能移動到所指定的位置。

$$\theta_{id} = \vec{d}_y / \text{gain1} \quad (2-30)$$

$$\theta_{pd} = \vec{d}_x / \text{gain2} \quad (2-31)$$

由架構圖，如圖 2-26 所示，可知頭部運動控制卡的功能，主要為把位移像素轉換成相對應之移動脈波數，產生水平與垂直之移動軌跡，伺服馬達的 PID 伺服控制，同時藉由移動限制功能，防止頭部有不當的轉動情況，如防止水平轉動超過 180 度，垂直轉

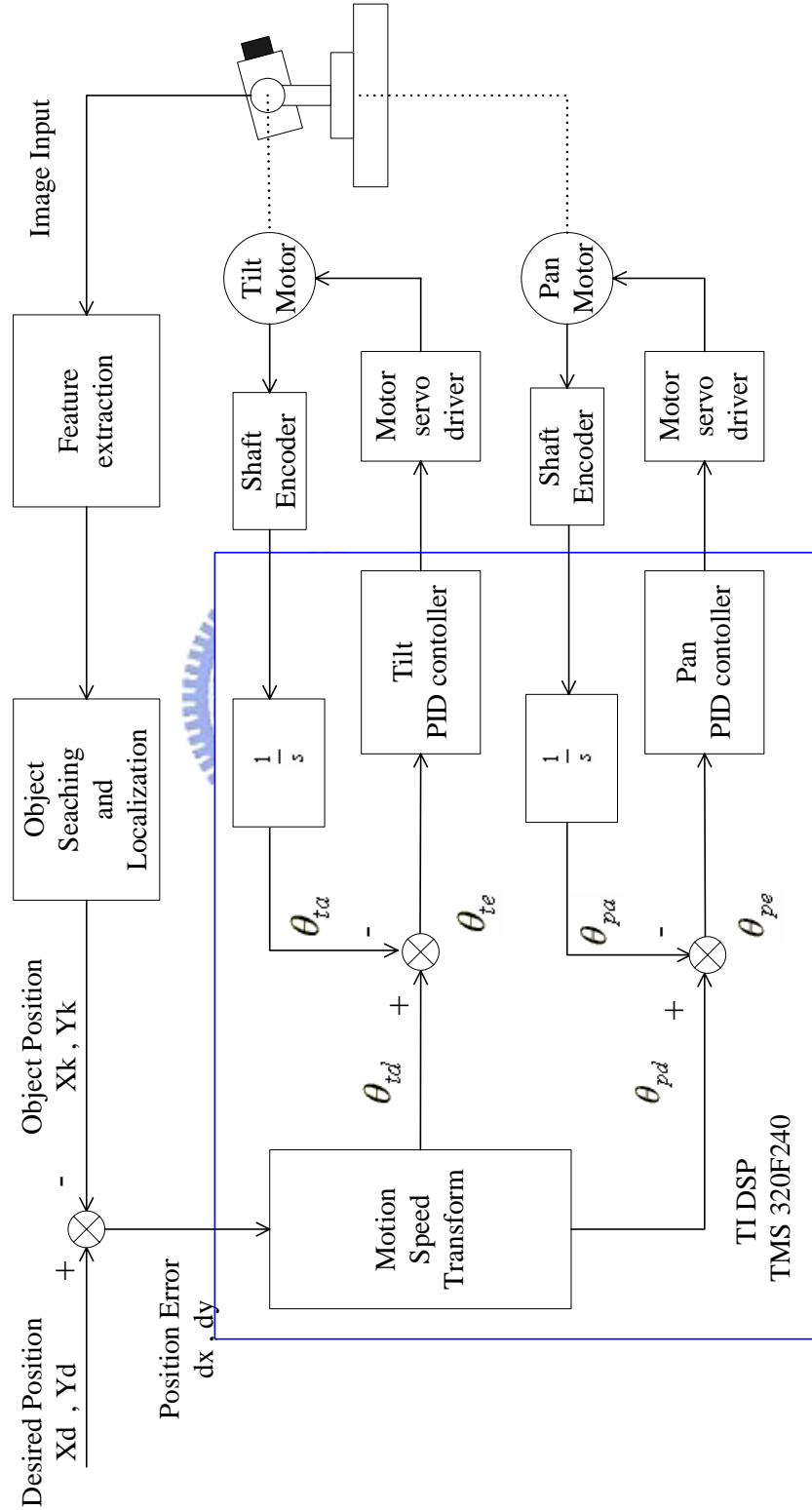


圖 2-26 影像追蹤控制器系統架構圖[22]

動超過 90 度，轉動時若超過所設定的角度範圍將造機構損壞。控制卡中 PID 伺服控制的取樣時間設定為 1ms，使控制卡能有良好的反應。

先前實驗室陳文俊學長已完成一家用機器人的追蹤行為控制[22]，雖然有考慮到人與機器人之間的互動及機器人的行為法則，但是沒有考慮到人與機器人在互動時機器人的靈敏性。當在執行人臉追蹤的時候，此機器人會先轉動其頭部，若頭部轉動角度左右小於 45 度的話，則機器人本體不進行本體的修正，一旦左右大於 45 度則機器人才開始進行本體的修正；機器人會根據人臉投影的大小及頭部仰角的資訊，控制機器人本體前進或後退以與人保持一定的距離。根據以上的行為法則則會發生一個問題，即當人員向機器人前進或者後退時，若機器人的頭部若和本體不在同一平面，則必須等到頭部和本體修正成同一平面之後才可以前進及後退，如此一來，機器人的反應就會比較遲頓，故我們將重新設計機器人頭部及本體的行為決策系統。

在家用機器人之本體移動及頭部控制考量方面，首先，由本體先進行左右轉的人臉追蹤，當在某一速度以內可以追蹤到人臉時，利用機器人本體修正即可，若本體左右轉速度大於某個上限時，再利用頭部追蹤加強人臉追蹤的效果。頭部在進行人臉追蹤時，機器人同時會根據頭部及本體的角度（Pan angle）進行本體的修正，此行為可以使機器人隨時保持面向人的姿勢，如此機器人可以根據人臉的位置及大小同時進行人臉追蹤及機器人本體之修正，機器人之反應速度也可以隨之提昇。機器人頭部 tilt 是隨時根據人臉的位置而修正。圖 2-27 為本論文之機器人行為決策架構圖。

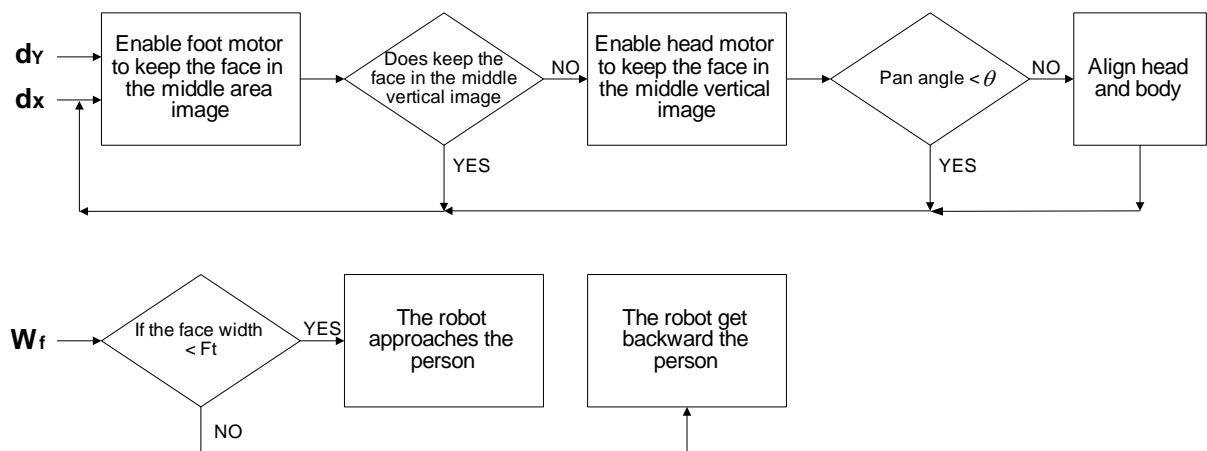


圖 2-27 機器人行為決策架構圖

2.4 結論與討論

本章節主要完成的部分為人臉追蹤系統，由第五章實驗的結果來看，即使在光源的影響下，也可以有效地及即時性地追蹤人臉。在移動方面藉由感興趣區域的選取以預測下一次人臉出現的位置及大小，不但可以縮小處理的畫面以免不必要的搜尋，在色彩投影方面也可以由多次投影減少到只要做一次投影即可；利用時間性加強追蹤系統的穩定性；在膚色之色彩分佈方面，利用適應性膚色色彩空間模型 YCrCb 三維色彩分佈以完成膚色之色彩分割以克服光源的影響。目前在人臉追蹤處理速度方面，本論文利用 L2 快取記憶體進行資料的處理[23]，所以嵌入式影像平台只有使用工作頻率 300Mhz 處理的速度而已，處理一張 320x240 解析度彩色影像大約花費 20 毫秒。利用本論文使用之追蹤架構，可以使用少量的運算量達到相同的結果，如此一來可以將其他應用，例如人臉辨識或者唇形辨識，能在同一平台一起實現對於將來即時性的要求有一定的助益

在家用機器人頭部控制方面，利用以下兩點改善追蹤時不穩定的情況：第一，以速度控制取代角度控制；第二，人臉中心位置距離影像平面中心某一區域即表示人臉置中。先前實驗室學長利用人臉在影像平面上出現的位置與影像中心點的距離做角度控制，當人臉距離愈遠時即移動角度愈多，反之則減少移動的角度，不過因為種種不穩定的因素，例如光源影響，而造成人臉的位置會有少許的改變，所以在控制頭部運動時會有不穩的狀況。用以上兩點來改善頭部追蹤，在實測方面的確得到相當不錯的結果。

雖然本論文利用適應性膚色搜尋來對人臉做持續性的追蹤，由實驗的結果的確有一定的成效，不過實際環境之背景是複雜的，如果人臉位置附近的背景與膚色類似的話，則有可能造成人臉追蹤失敗，所以若能再加入其他追蹤的條件，例如輪廓及動態偵測，一定能提高人臉追蹤的效果及穩定度。此追蹤系統是根據特定人員的色彩分佈而進行實驗，所以有可能因為不同場合、色溫或者不同的人員而造成膚色的色彩分佈發生改變進而追蹤失敗，本論文提出另一解決方法，若隨時利用色彩的高斯分佈並且選擇低於色彩分佈最大值及高於最小值某一程度，做為色彩分佈閾值的話，則此方法有一定的機會不會因為不同人員及色溫而造成影響，這對於人臉追蹤的強健性及實用性有一定的幫助。