

第二章 文獻探討

2-1 人臉檢測的問題

人臉檢測的問題，其所包含的內容相當廣泛，由不同的角度分析，大致可分成幾個大類，如表1.1：

表1.1 人臉檢測之問題分類

問題分類	內容	
影像的來源	靜態影像：其主要的問題在於演算法的適應性與強健性。	動態影像[17]：其主要問題，在於檢測影像的速度有很高的要求。
影像的色彩	彩色影像	灰階影像
人臉的姿勢	正面：影像中限制人臉為正面的方向[18,19]。	側面：影像中的人臉是否俯仰或旋轉一個角度。
影像中人臉之數目	單一：限制在影像中只能有一個人臉[19]。	未知：需判斷影像中，是否存在影像與人臉的數目。
影像背景的複雜度	簡單的背景：在影像中，限制背景的特徵[19]。	複雜的背景：不限制影像的背景種類或特徵。[8][20-22]

由以上人臉檢測問題的分類可知，針對不同的檢測問題，其所需解決問題的演算法與特徵要求皆不相同，因此，在設計人臉檢測演算法之前，須針對應用場合，分析各種狀況與問題，設計合適的演算法。本文主要是討論靜

態彩色影像、不限制人臉的姿勢與數目，且不限定影像背景的複雜度，因此，本文依上述情況下，設計合適的人臉檢測系統。

2-2 人臉檢測方法的綜述

人臉影像所包含的特徵相當豐富，如何在人臉影像中決定人臉的特徵，是人臉檢測中最關鍵的問題，但是人臉的特徵具有複雜且細緻的變化，因此，需要採用綜合模式的處理方法，取得人臉特徵。

依據人臉特徵的特性，發展出許多不同的特徵處理方法，其可以分類如 1.2 表所示：

表 1.2 人臉的特徵與處理方法

人臉特徵	處理方式
顏色	膚色分佈[15-17][23,24]
直方圖	直方圖的分佈與距離[25,26]
結構	對稱性與投影特性[19][25,26]
輪廓	橢圓輪廓[27]
模板	均方差距離[25][28]
變換域	特徵臉[29]與小波特徵[30,31]

人臉檢測的方法在近幾年中，已有相當的發展，利用方法論可概要分類出幾種主要的人臉檢測方法[3]，當然，此分類的界線並不是絕對的，其結合多個方法一起使用的檢測方法亦相當普遍的。

2-2-1 知識為基礎之方法[19,21]

這些方法，是基於我們對一些典型人臉的認識而總結出一些知識，且將這些

特徵之間的關係歸納出一些檢測的法則，例如，正面的人臉具有兩個對稱的眉毛與眼睛，如此尋找出一系列的法則後，將其應用在檢測影像中，檢測出符合規則之人臉，這些方法主要是應用人臉的定位。

2-2-2 人臉特徵為基礎之方法[32,33]

這些方法是利用特定組織不變的特徵，尋找結構性的特徵偵測人臉，首先，針對輸入影像中，擷取出某些特徵，再與人臉特徵之間進行比較，判斷此區域是否為人臉，例如可利用影像測邊的方法，尋找邊界判斷正面人臉的邊線而構成正面人臉的輪廓，或從影像中擷取出色彩與紋理再與人臉的色彩與紋理模型進行判斷比較，得到可能是人臉的區域，若運用組合相關的人臉特徵，更可以提高檢測演算法的效能。

2-2-3 模板匹配為基礎之方法[34,35]

此方法，是先建立且儲存一些人臉的標準模板，例如建立正面的人臉、眼睛、鼻子、嘴巴與眉毛之模板，再利用相關的演算法，來比對輸入影像中的區域與標準的模板其相似程度與相關性，進而判斷此區域是否為人臉。此方法需耗相當大的系統計算量，而且，為了增加檢測率，事先需建立收集龐大的比對樣本資料庫。此方法常用於人臉的定位與檢測。

2-2-4 人臉外觀為基礎之方法

此方法是透過大量的人臉與非人臉的樣本做為訓練集，利用相關技術訓練出一分類器，用來檢測人臉，例如類神經網路或支持向量機 (SVM)。此方法與其他方法的不同，在於不用在事先利用人工針對人臉進行分析或建立模板，而是利用大量整體外觀，用結構化的方法訓練出人臉檢測的分類

器。這些方法主要是用於人臉的檢測。

2-2-4-1 支持向量機[12,13]

支持向量機(Support Vector Machines, SVM) 是 Vapnik 等人提出的，基於結構風險最小化原理(Structural Risk Minimization Principle, SRM)之統計學學習理論[12]，常用於分類與回歸的問題。SRM 使 VC (Vapnik Chervonnikis) 維數的上限最小化。

2-2-4-2 類神經網路[9-11]

類神經網路(ANN)的方法，是將模式的統計特性，隱含在ANN的結構與參數之中，其方法是基於經驗風險最小化(Empirical Risk Minimization Principle, ERM)的原理。如HideakiIshii[11]運用類經網路，先做人臉膚色之搜尋，再用另一組類神經網路做人臉區域之確認。與Anifantis[12]用一個類神經網路直接搜尋人臉的方式。這類方法將受限於所訓練的樣本是否足夠，及其所訓練的類神經網路是否適當的影響。尤其是在即時系統中，對於訓練樣本之收集，需要較多之資料量，加以分類。

2-2-4-3 特徵臉[29]

特徵臉技術之基本原理是由統計的觀點，尋找人臉影像分佈的基本元素，亦即尋找人臉的特徵向量，近似表示為該人臉，此特徵向量稱為特徵臉(Eigenface)，實際上，特徵臉反映了隱含在人臉樣本集合內部的資訊和人臉的結構關係。

另外，近幾年，亦有許多的專家學者利用小波當基底，解析其訊號，運用於人臉定位與人臉辨識[30,31]。

以上各種方法都有其優缺點與適用的領域，許多學者專家亦提出了許多相關的檢測演算法，但是，這些演算法都存在效率與性能的矛盾，在檢測正確率佳的演算法，通常需要花費相當的時間，效率與性能在本質上是對立的。例如：使用模板與特徵的方法需大量的計算，並且會因為人的不同而產生差異，而模板建立亦不容易；類神經網路的方法則是利用獲得的特徵值，然後再輸入類神經網路來分類，但此方法必須有大量的訓練資料，才会有較好的結果。因此，我們在設計檢測系統時，須考慮如何使檢測演算法在整體性能上表現為最佳？

2-3 色彩空間的選擇

依據色彩理論，對同一顏色而言，可以有不同的表示方法，因此，產生了許多不同的色彩空間。色彩空間是表示所有可能顏色的一個模型，可應用在許多的輸出裝置上，例如監視器、印刷機、掃描器和照相機等。數位彩色影像通常利用RGB色彩空間，R（紅色），G（綠色）和B（藍色）通常稱為三原色，利用RGB值，我們可使用線性或非線性的轉換，獲得其他的色彩空間，每一種色彩空間都有其產生的背景與應用領域，主要的色彩空間與其優缺點敘述如下：

2-3-1 RGB 色彩空間（紅、綠、藍三原色模型）

彩色影像為多光譜影像的一種現象，對應於人類視覺三原色為紅、綠、藍三個波長，近似人眼的光譜量化性質，因此，利用 RGB 三原色分量來表達色彩是很自然的一種格式，許多影像擷取設備是直接感知色彩的 RGB 分量，因此，其為十分重要的色彩空間。RGB 彩色系統為一直角座標系統，其中三個基量為可見光中的三原色光，為使色光的定義能夠統一，CIE 定出三原色的波長為紅(700nm)，綠(546.1nm)，和藍(435.8nm)。RGB 色彩空間表示為一 3D 色彩立方體如圖 2.1 所示。

一般色彩系統的形式可稱為 Additive 線性 RGB 色彩系統如圖 2.2 所示，在各種色光混合中，均有一對色光可以混合成白色光，稱為補色光，對紅、綠、藍而言，其補色光如下：

紅 --- 青(Cyan)

綠 --- 紫(Magenta)

藍 --- 黃(Yellow)

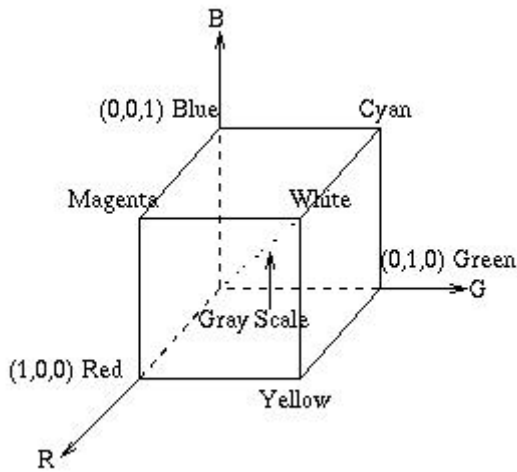


圖2.1 RGB色彩立方體

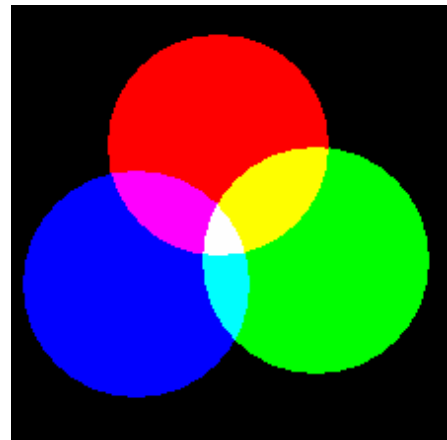


圖2.2 Additive RGB色彩系統

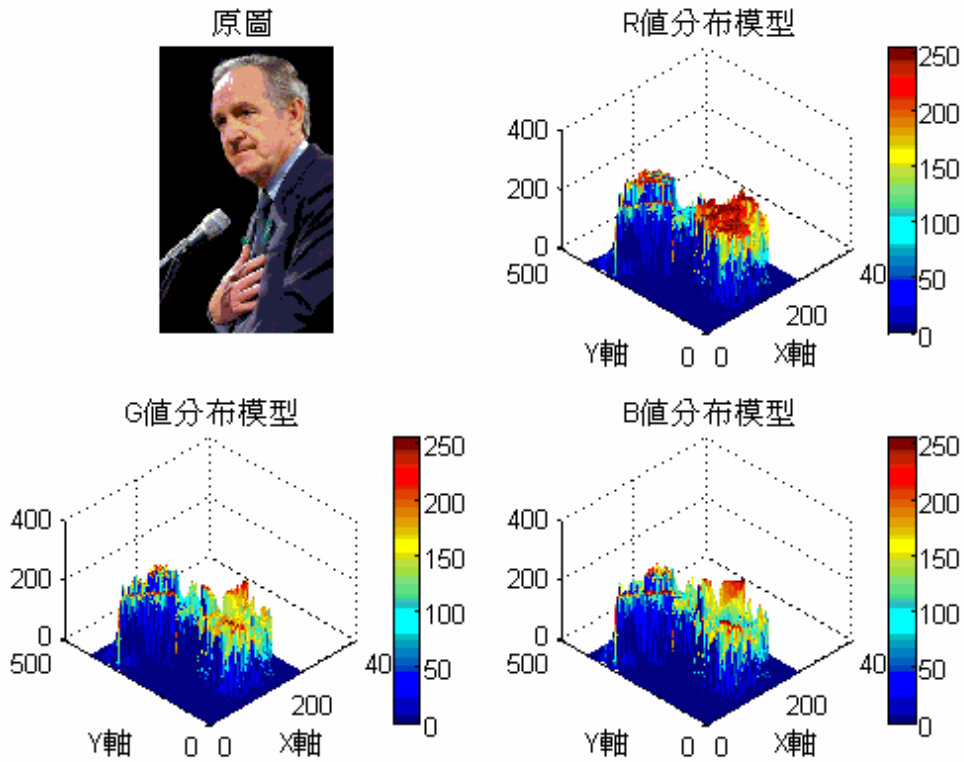


圖2.3 RGB三原色分佈

圖 2.3 所示，為觀察人的膚色在 RGB 三個獨立分量是否具有聚類性？是否可以做為膚色偵測的特徵？由圖 2.3，可發現膚色在 R 分量有較佳的聚類特徵。

由於 RGB 色彩空間主要的缺點是其每一個分量之間有很高的關聯性及對於光線反應較為靈敏，為了減少其關聯性與降低光線的影響可將 RGB 色彩空間作正規化轉換。

其轉換方式如 2.1 式所示：

$$r = \frac{R}{R+G+B} \quad g = \frac{G}{R+G+B} \quad b = \frac{B}{R+G+B} \quad (2.1)$$

2-3-2 HSI 色彩空間(色度、飽和度、亮度模型)

HSI 色彩空間模型常為藝術家所使用，此模型反應人類觀察色彩的方式。在色彩資訊的應用上分辨不同色彩，主要是利用色彩三要素：色度(Hue)、飽和度(Saturation)與亮度(Intensity)。我們若分析某一類物質之色彩的特性時，則需分析其在色彩空間的聚類特性，而此一聚類特性通常反映在色彩的本質上，而且又常受到光的明暗干擾而影響此特性。光的明暗對物質的直接影響為亮度的分量，因此若能將亮度由色彩中移出，而只利用色彩本質的色度與飽和度來作聚類分析，可得到較佳的效果。

在應用上，HSI 色彩系統的優點是其分量彼此之間是不相關的，可以將顏色分別分離出亮度與反應色彩本質的色度與飽和度，而缺點是要將 RGB 三原色轉換為 HSI 的計算成本。

色度(Hue)平常是用來區分顏色的主要成分，亦即為不同波長的光在眼睛所感覺的顏色，如紅、橙、黃、綠、藍等。飽和度(Saturation)表示顏色飽和的程度，亦即顏色中滲入白色的程度，高飽和度的色彩表示此顏色滲入白色愈少，如粉紅色比紅色的飽和度低。

色彩學(Chromaticity)是指色度與飽和度，CIE 協會定義析色圖(chromaticity diagram)統一色光的定義，所有光譜的色光，分佈在析色圖的

周圍，而色光所混和的所有顏色皆在析色圖上。CIE 所制訂的 RGB 色彩空間和所有色光之關係如圖 2.4 所示。

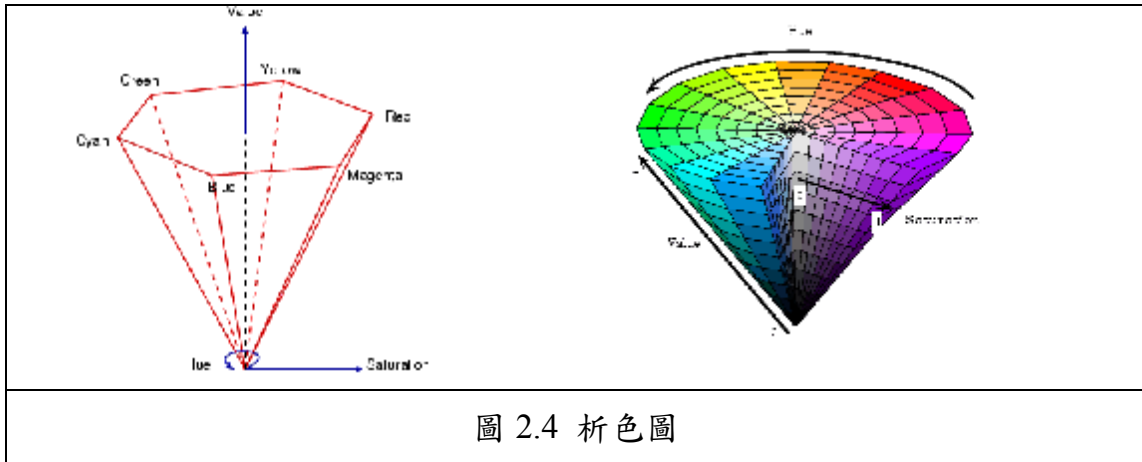


圖 2.4 析色圖

其 RGB 色彩空間轉換至 HSI 色彩空間的轉換方程式如 2.2 式所示：

$$\begin{aligned}
 H &= \arccos \frac{\frac{1}{2}((R-G)+(R-B))}{\sqrt{((R-G)^2 + (R-B)(G-B))}} \\
 S &= 1 - 3 \frac{\min(R, G, B)}{R+G+B} \\
 I &= \frac{1}{3}(R+G+B)
 \end{aligned}
 \tag{2.2}$$

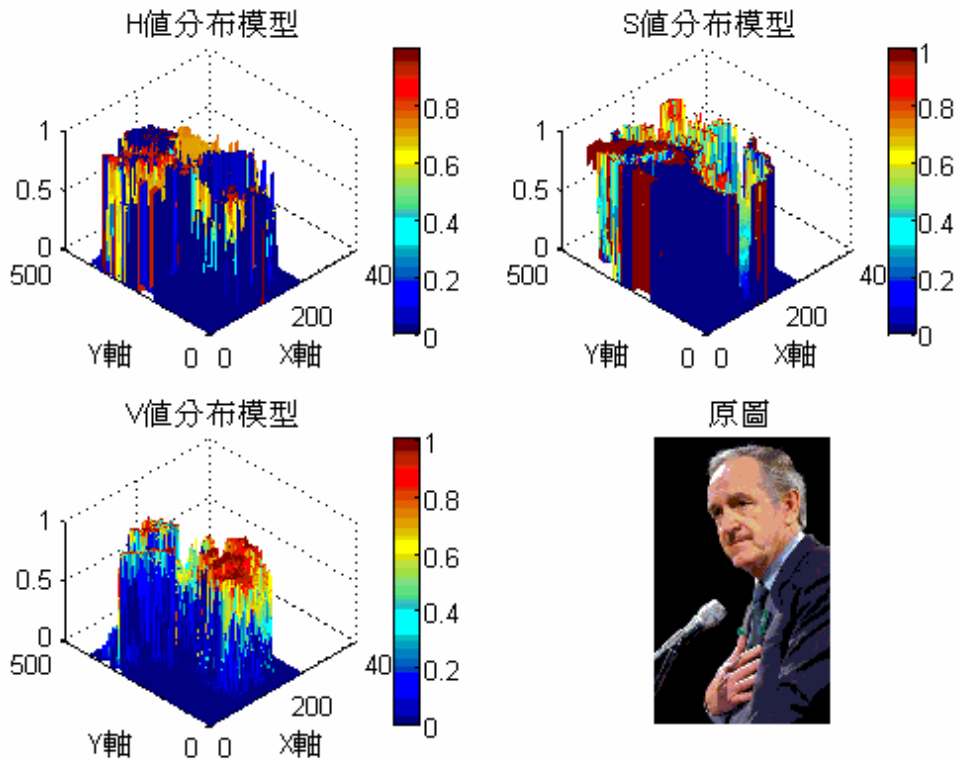


圖2.5 HSI 色彩空間分佈

圖 2.5 所示為 HSI 色彩空間之聚類特徵。

2-3-3 CMY 色彩空間

彩色印刷是利用顏料的吸光特性顯示色調，通常是採用青色(Cyan)、紫色(Magenta)和黃色(Yellow)為三原色，其分別吸收各自的補色光，此三種顏料依不同的比例打點在白紙上，在白光照射下，反射出不同比例的紅、綠、藍光，而呈現各種的色彩。

RGB 色彩空間與 CMY 色彩空間之間的轉換關係如 2.3 式：

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1-R \\ 1-G \\ 1-B \end{bmatrix} \quad (2.3)$$

2-3-4 YCbCr 色彩空間

此色彩空間是應用於電視之色彩顯示格式，為了要讓彩色電視信號能夠利用原有的黑白電視傳播系統傳送信號，而針對視覺的特性，所發展出來節省傳輸頻寬的方法。

YCbCr 色彩空間，其 Y 代表亮度分量，Cb 與 Cr 代表色度分量，Cb 與 Cr 所在的平面又稱為色度平面。

YCbCr 色彩空間具有將亮度分量分離的優點，但是因它是利用 RGB 做線性變換而得到，因此，較少直接使用在物體色彩之聚類分析，其大多需做一些修正之後，才能應用在物體的聚類分析。

其基本的轉換方程式如 2.4 式所示：

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.5 \\ 0.5 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.4)$$

Y 式中的加權值代表眼睛對相同的 RGB 值的不同反應，Cb 和 Cr 代表去除亮度後的色度值，對於白色光，亦即當 R=G=B 時，Cb 和 Cr 值皆為 0，其代表此時無色差。

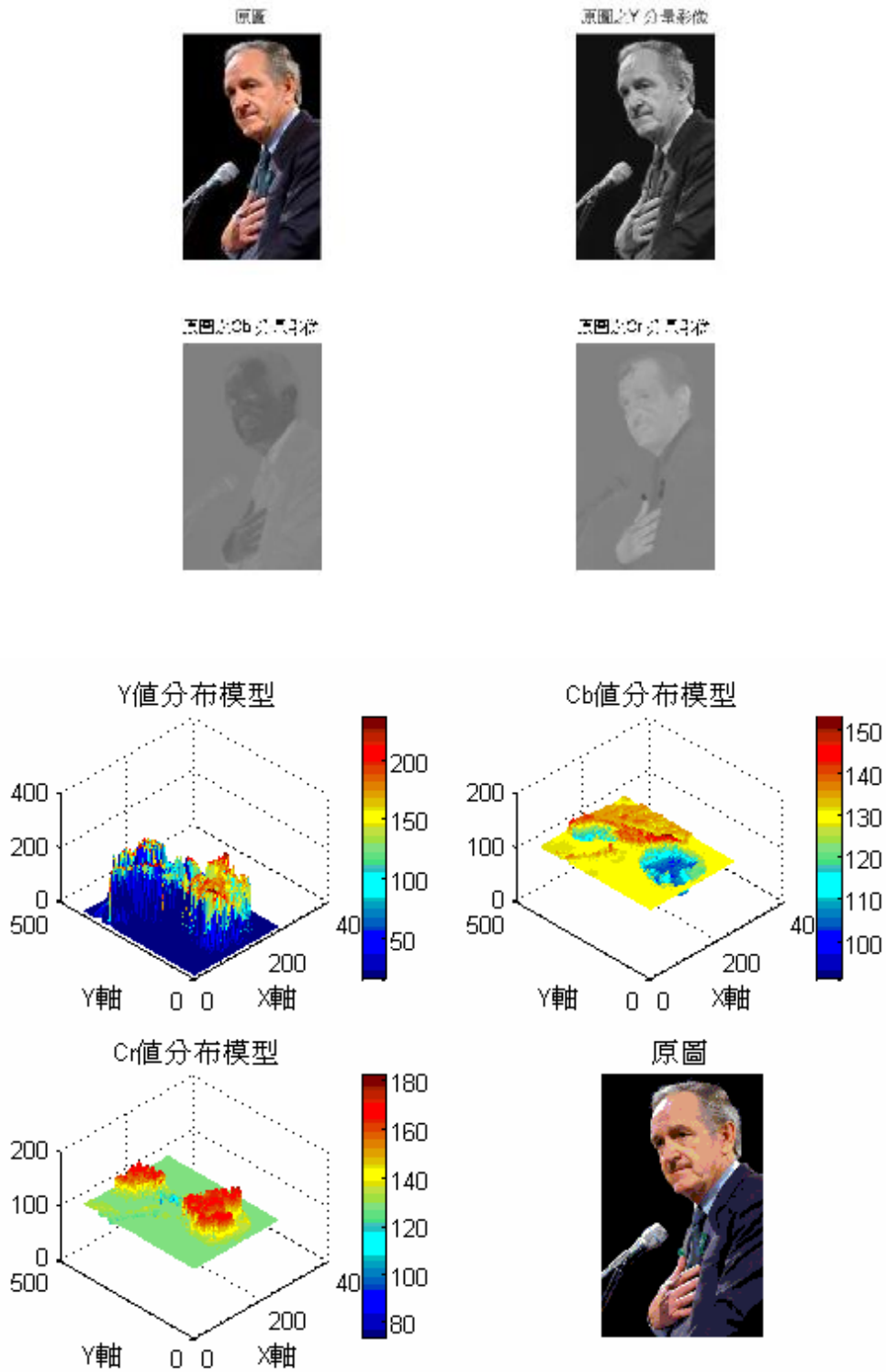


圖 2.6 YCbCr 之分量分析

由圖 2.6 所示，可發現膚色在 Cb 與 Cr 的分量，具有很好的聚類特徵。

2-4 膚色模型

建構膚色模型的主要目的是建立膚色的判斷法則，用於區別像素為膚色或非膚色，建立膚色模型的方法常用有以下幾種[36]：

1. 明確定義色彩空間的聚類範圍[37,38]

此模型是藉由色彩空間的變換，使膚色在色彩空間中具有良好的聚類特性，然後將此一聚類在色彩空間定義出明確的聚類範圍，例如若像素之色度值滿足 H 值落在 0 至 85 之間、Cr 值落在 92 至 107 之間及 Cb 值落在 102 至 117 之間，則此像素被分類為膚色，其中 H 為 HSI 色彩空間的 hue 分量，其範圍正規化為 0~255，Cr 為 YCbCr 色彩空間之 Cr 分量，Cb 為 YCbCr 色彩空間之 Cb 分量[39]。此方法可利用簡單的數學方程式來判斷，其優點是在檢測膚色時較簡單且可快速的分類。但是此方法若要達到高的正確檢測率，其需要選擇好的色彩空間與合適的判斷法則，例如利用機械學習 (machine learning) 演算法尋找出合適的色彩空間與簡單的判斷法則[40]。

2. 參數式膚色分佈模型[41,42]

參數式膚色分佈模型，常見有高斯膚色分佈模型與高斯混合膚色分佈模型，其是利用統計學的原理，主要的概念是認為膚色之隨機樣本，在色彩空間的分佈應符合高斯分佈，而高斯分佈其形式較為簡單，在統計學中應用較為深入，因此，應用在建構膚色模型其效能應該不錯。

利用高斯分佈模型建立膚色模型的方法亦可分為兩種：一種是可以利用多個人臉樣本建立一個膚色模型[41]，另一種是針對每一單個人臉樣本，分別建立相應的膚色高斯模型，然後在疊加成一個高斯混合模型[42]。

常見的參數式膚色分佈模型，大多是利用色彩空間中的色調平面，而忽略亮度分量資訊。

3. 非參數式膚色分佈模型[28,23]

非參數式膚色分佈模型，其方法主要是從訓練樣本中，事先不需作特徵歸納取得明確的膚色模型，用於評估膚色的分佈，常見方法有 Bayes classifier [23]及 Self Organizing Map (SOM) [18]。

非參數式膚色分佈模型的優點是在訓練與處理速度上非常快而且其方法在理論上與膚色在色彩空間分佈的形狀無關，缺點是其需要較多的訓練樣本與儲存的記憶體空間。

2-5 光線補償

在影像中檢測人臉是相當有挑戰性的，因為在影像中，很難預知在影像中人臉的尺寸大小、位置與數量，而且人臉檢測在不同的光線情況下，會增加人臉檢測的困難度；一般的人臉檢測演算法，大多是假定影像是在均勻光線下取得，但實際上，光線通常是不均勻的，因為不均勻的光線甚至有強光與陰影的影響，使得人臉的檢測率大為降低，因此，我們有必要針對不均勻的光線做適度的補償，但在目前實用的影像增強技術，需要額外人為調整參數，才有較好的效果，因此必須依賴人來調整相關參數，否則易導致影像的過度強化或不足，使得影像增強技術及其演算法並不適用於即時人臉檢測系統。

近年來，以色彩資訊為基礎的方法，成為一新技術且具有較佳之效能，因為色彩資訊是我們視覺重要的資訊來源，而且，在複雜的背景下利用色彩資訊的分佈來檢測分割物體可減少檢測的時間。因此可將其應用在相關的系統，所以我們利用人臉膚色色彩的資訊作為物體檢測技術之基礎，利用色彩資訊在某些色彩空間的分佈，可以建立一人臉膚色之色彩模型，做為人臉檢測之過濾器，因此可應用人類膚色資訊在人臉檢測系統中，做為檢測系統之前置處理，例如 Yanjiang Wang 等[20]使用正規化 RGB 色彩空間與 HSI 色彩空間聚類出膚色之範圍。Douglas Chai 等[23] 使用 YCbCr 色彩空間聚類出膚色之範圍。但是若使用膚色之色彩資訊作處理時，即會存在上述之光線明暗的影響，使檢測膚色的強健性不佳。

模糊控制器可應用於一些控制領域[43][44]，例如無法建構出數學模式、無法給予明確定義或難以用傳統量化技巧分析之系統，而光線對於影像的影響目前尚無較佳的數學模式，因此適合使用模糊理論來解決光線補償的問題。

光線補償藉此改善物體的檢測率之研究，最基本簡單的方式是利用影

像之直方圖均化原理，其作法是將影像中之像素灰階值直方圖以累積機率密度函數重新分配灰階值的分佈，因此處理後的直方圖灰階值較均化，並能增強影像的對比，但是利用影像直方圖均化方法可能會將數個灰階值較少的像素合併成一新的灰階值造成過度強化，雖然有許多針對直方圖均化的改進方法，例如 K. Wongsritong 等[45]使用多峰直方圖均化法以增強影像對比值，其方法雖然可避免影像過度強化，但是卻造成影像對比的不明顯，限制增強影像的效果。另外其他改進的方法雖可改善過度強化的影響，但其亦相對產生一些問題如：處理成本過大或明亮對比過大，且應用在彩色影像中則需轉換至其他色彩空間，例如利用 HSI 色彩空間，使用直方圖均化來調整影像亮度之分量，在處理後將圖形轉回 RGB 色彩空間，修正原影像。R.L. Hsu, M. Abdel-Mottaleb, A.K. Jain [46]提出利用參考白(reference white)之方式，改善彩色影像因偏光的影響與提出一個非線性色彩轉換之演算法，補償光線的影響。Hayit Greenspan 等人[47]提出一個混合式高斯(mixture of Gaussians)分佈建構陰暗之人臉膚色模型。Shimizu, Kondo, Kohashi, Tsuruta and Komuro [48]提出使用模糊邏輯與 HIST 產生歸屬函數，進而得到補償量，達到影像之光線補償的效果。Murakami and Honda[49]提出自動曝光系統，其是使用影像像素之色調(hue)與色度(chroma)取得背景的資訊及藉由模糊推論取得補償量。上述的演算法其因為需額外作色彩空間的轉換或作非線性轉換，使其計算成本較高且較繁雜，因此，為了避免上述的問題，在本文中設計一簡單的分散式模糊推論系統，推論光線的補償量。