

# 彩色影像之人臉角度分類

林 群 雄

國立臺北大學資訊中心

助理研究員兼作業組組長

(收稿日期：2000 年 10 月 24 日；修訂日期：2001 年 9 月 27 日；接受刊登日期：2001 年 11 月 12 日)

## 摘 要

在本文中，我們建議的系統是由兩個主要部分組成。第一個部分是藉由皮膚顏色分割和等腰三角形為基礎來搜尋潛在臉的區域。我們首先讀取一張 RGB 彩色影像。先判斷此 RGB 彩色影像是否有複雜背景。若無，則我們將跳過「皮膚顏色分割方法」，而直接將原 RGB 彩色影像直接轉變成二值化的影像。若有複雜背景，則將藉由皮膚顏色分割，找出皮膚顏色區域。再將這個皮膚顏色分割後的影像轉變成二值化的影像。再藉由尋找等腰三角形的關係去得到潛在臉的區域。第二部分是要完成臉部角度分類的任務。我們首先將每一個潛在臉的區域，都做了尺寸標準化的處理。然後，藉由人臉權值面具函數獲得每一個人臉的正確位置。其次，再藉由方向權值面具函數判斷人臉的正確方向。最後，再藉由角度權值面具函數決定人臉轉的角度。實驗結果顯示約百分之九十九的成功比率，並且相對錯誤比率很低。

**關鍵字：**人臉偵測、人臉角度分類、皮膚顏色分割、權值面具函數。

# POSE CLASSIFICATION OF HUMAN FACES IN COLOR IMAGES

CHIUNHSIUN LIN

Assistant Researcher, Information Center  
National Taipei University, Taipei, Taiwan, 104, R.O.C.

Received : Oct. 24 2000 ; Accepted : Nov. 12 2001

## Abstract

In this paper, we introduce a novel approach for automatic estimation of the poses/degrees of human faces embedded in complicated environments. The proposed system consists of two primary parts. The first part is to search the potential face regions. First, if the input image contains complex background, then the potential face regions are gotten from skin-color-segmentation and the isosceles-triangle criterion that is based on the rules of "the combination of two eyes and one mouth". If the input image contains complex background, then we will use the input RGB color image to perform the human-skin color-segmentation task to remove the complicated surroundings. Then the result of the input image that is removed the complicated surroundings will be converted to a binary image. If the input image doesn't contain complex backgrounds, then we will skip the human-skin color-segmentation task. The input image will be directly converted to a binary image. Secondly, label all 4-connected components and detect any 3 centers of 3 different blocks that form an isosceles triangle. Then, clip the regions that satisfy the

isosceles triangle criteria as the potential face regions. The second part of the proposed system is to perform the task of pose verification. In the second part, each face candidate obtained from the previous process is normalized to a standard size (60\*60 pixels). Then, each of these normalized potential face regions is fed to the face weighting mask function to obtain the location of the face region. Next, the face region is fed to the direction weighting mask function to determine which direction the matching face region looks at. Last, the face region is fed to the pose weighting mask function to decide the poses/degrees of the human faces.

The proposed face poses/degrees classification system can determine the poses of multiple faces embedded in complicated backgrounds. Experimental results demonstrate that an approximately 99% success rate is achieved and the relative false estimation rate is very low.

**Key Words:** Face detection, face poses classification, skin color segmentation, weighting mask function

## 壹、前言

自動人臉偵測、臉部角度分類與人臉辨識是目前圖形識別和電腦視覺的主要研究課題之一。通常，在人臉辨識過程中，我們首先需要自動偵測出人臉的位置。其次，需要自動分類每一個臉部之角度。最後，才做人臉辨識。以前大多數的人臉辨識系統往往忽略了自動人臉偵測與臉部角度分類的重要性。因此，當人臉在複雜背景中或人臉的角度轉變較大時。大多數的人臉辨識系統便不能正確地完成人臉辨識了。所以，發展一個高效率的自動人臉偵測與臉部角度分類系統是幫助人臉辨識系統正確地完成任務所不可或缺的。在許多文獻與研究論文中已經報導了一些不錯的系統，如 Juell 等人（1996, pp. 781-787）提出層次化神經網路（hierarchical neural network）用以偵測灰階影像人臉。他們設計邊界強化前處理機制，且在一個層次化架構中安排四個反遺傳神經網路（four backpropagation neural networks）以偵測多樣的人臉。Gong 等人（1996, pp. 265 -270）介紹以 Gabor 子波轉換（Gabor wavelet transform）和經由主要構成要素分析而得的角度特徵空間（pose eigenspaces based on PCA – Principal Component Analysis）所合成的人臉代表結構（face representation scheme）去偵測人臉角度變化。C. H. Lee 等人（1996, pp.1877-1889）採用了人臉移動和顏色的資訊。他們的系統有兩個主要步驟。在第一個步驟中，他們用連續影像作輸入且假設連續影像中最大的運動物體就是人臉。因而，他們用移動的資訊作視覺暗示以分離複雜背景。而在第二個主要步驟中，他們使用經由第一個步驟取得的眼睛、眉毛和嘴的區塊，再加上顏色資訊作視覺提示，以自動偵測出人臉的位置。爲了把顏色用作視覺提示，他們分析韓國人臉的顏色資訊，並提出使用他們的色調資訊（hue information）之分割演算法。然而，萬一「假設連續影像中最大的運動物體就是人臉」並不成立。可想而知，他們的系統將會失效。Sobottka 等人（1996, pp. 421-425）發表了從彩色影像自動分類臉部角度之方法。他們運用 HSV 顏色空間（Hue-Saturation-Value color space）與橢圓形去自動分類臉部角度。Dai 等人（1996, pp. 1007-1017）介紹了結合紋理（texture）與顏色資訊的一個架構模型。他們的系統基礎在

## 彩色影像之人臉角度分類

於「空間灰階倚賴性」(space gray-level dependence – SGLD) 矩陣。由實驗基礎取得了由一套不相等條件 (a set of condition inequalities) 創作的人臉紋理架構模型。此外，他們先將 RGB 顏色空間轉變成 YIQ 顏色空間，然後使用 I 部分，以提取皮膚顏色地區以改進性能。然而，他們用一個窗戶區域來掃描影像的所有部分，且計算這個平均色飽和度 (the average intensity) 以決定一個窗戶區域是否含有人臉區域。他們的系統是一個很消耗時間的系統 (因為低效能的搜查)。

由此觀之，以前大多數的自動人臉偵測或臉部角度分類系統之研究，皆需強加許多限制，如：他們不允許不同人臉表情、不同光度及影像模糊與雜訊之問題。為了消除這些限制，我們發展了一個健全的「彩色影像自動分類人臉角度系統」，其架構如圖 1。

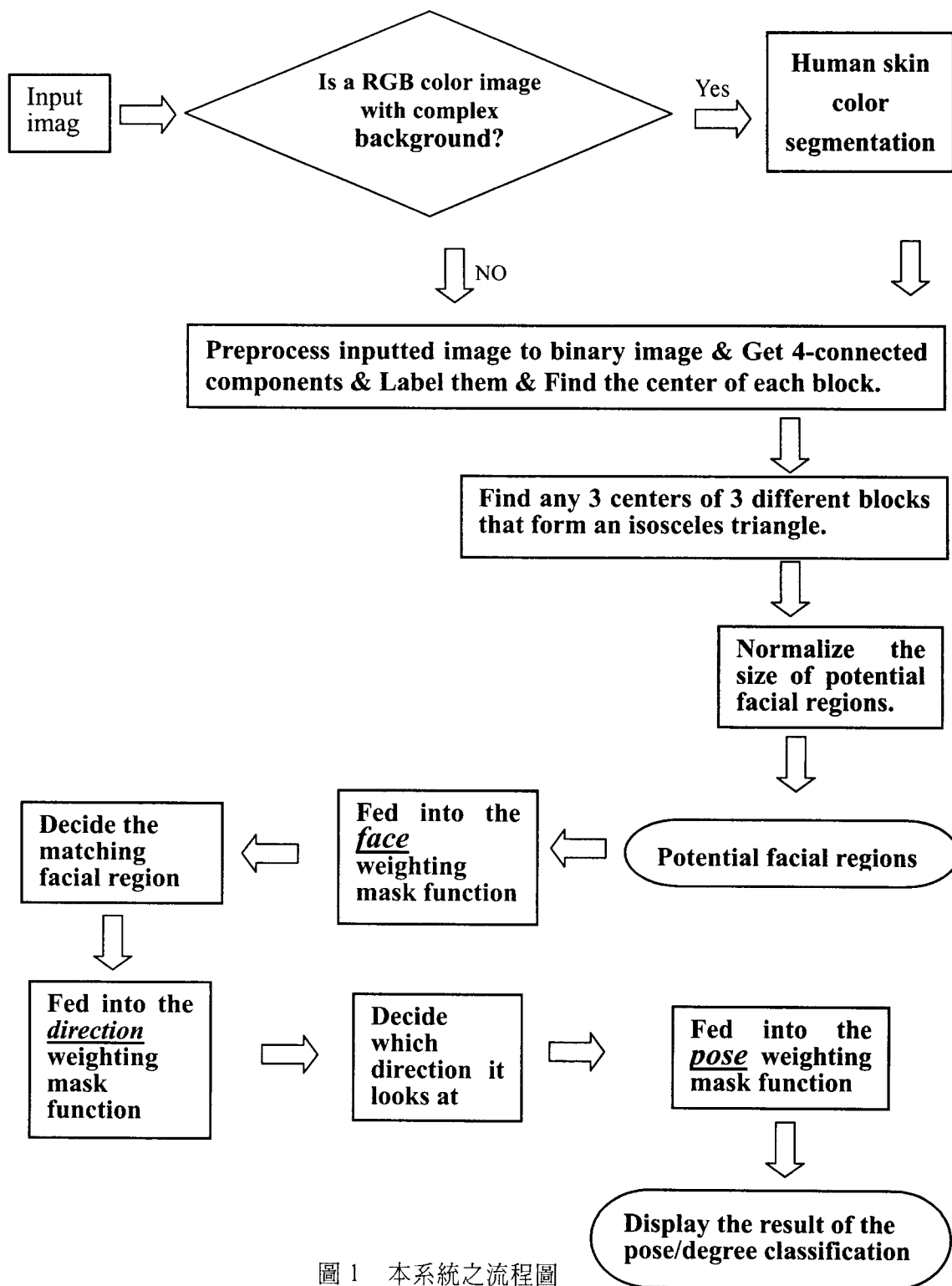


圖 1 本系統之流程圖

## 貳、潛在人臉之偵測

潛在人臉之偵測主要分為四個步驟。首先，我們讀取一張彩色影像。然後，應用「簡單總體的門檻方法」(Simple global thresholding method)並取門檻值(threshold) 100，去二值化影像。如果二值化影像中的區塊數目比25大，我們就認為(假設)原始輸入RGB彩色影像有複雜背景。則我們將藉由「皮膚顏色分割方法」，找出人臉皮膚顏色區域。再將這個皮膚顏色分割後的影像轉變成二值化的影像。反之，則我們將跳過「皮膚顏色分割方法」，而直接將原始輸入RGB彩色影像轉變成二值化的影像。最後，再藉由尋找等腰三角形的關係去得到潛在臉的區域。

### 一、人臉皮膚顏色區域之分割

首先，我們讀取一張彩色影像。先用「簡單總體的門檻方法」(Simple global thresholding method)並取門檻值(threshold) 100，去二值化影像。如果二值化影像中的區塊數目比25大，我們就認為原始輸入RGB彩色影像有複雜背景。則我們將藉由「皮膚顏色分割方法」，找出人臉皮膚顏色區域。再將這個皮膚顏色分割後的影像轉變成二值化的影像。反之(如果二值化影像中的區塊數目比25小)。則我們將跳過「皮膚顏色分割方法」，而直接將原始輸入RGB彩色影像轉變成二值化的影像。我們使用的RGB顏色空間是一個三度的向量空間，而每一個像素，是由  $(r(i), g(i), b(i))$  所組成，它們分別代表紅色光的強度、綠色光的強度與藍色光的強度。

經由許多實驗、分析與觀察，(我們取1,000張人臉做樣品，每一張人臉隨機取樣20點皮膚顏色之像素，總計20,000個樣品，經由實驗—跑程式、分析實驗結果與細心觀察)，我們歸納分析後獲得如下三個規則：

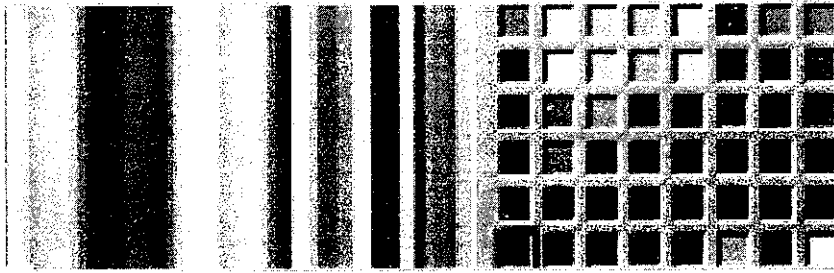
1.  $r(i) > \alpha$ ：意味著「紅色光的強度」應該比  $\alpha$  更大。

2.  $0 < (r(i) - g(i)) < \beta$ ：意味著((紅色光的強度) - (綠色光的強度))應該在 0 和  $\beta$  之間。

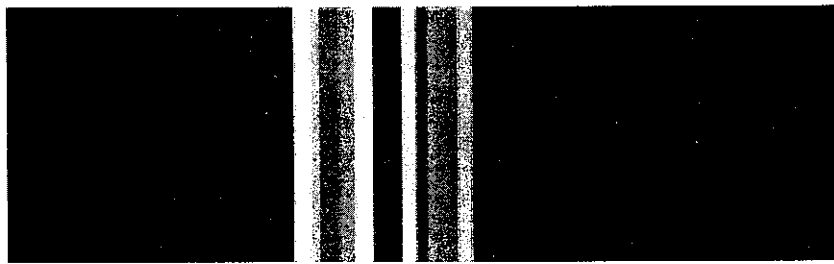
$3.0 < (r(i) - b(i)) < \gamma$ ：意味著（（紅色光的強度）－（藍色光的強度））應該在 0 和  $\gamma$  之間。

從測試大量的皮膚顏色樣品（我們取1,000張人臉做樣品），我們得到  $\alpha = 80$ 、 $\gamma = 56$ 、 $\beta = 98$  有令人滿意的結果。換句話說，如果輸入的RGB影像之任一像素能滿足以上這三個規則，則我們便把這一像素看作是「近似皮膚顏色」。否則，便把這一像素看作是「非近似皮膚的顏色」。例如：圖1(a)表示原始的RGB影像(b)表示經由以上三個規則過濾後的結果—「近似皮膚的顏色」仍被保持原來顏色，而「非近似皮膚的顏色」則被處理成「純黑色」。如此一來，我們便可將「非近似皮膚的顏色」與「近似皮膚的顏色」完完全全的區分開來。然而，本「皮膚顏色分割方法」卻不適用於黑色人種，原因並非種族歧視，而是因為在本系統下一步驟中，我們將尋找眼睛與嘴巴，而所用的方法又是「簡單總體的門檻方法」並取門檻值100，去二值化影像。如果將黑色人種之膚色亦納入「近似皮膚的顏色」，則想再找眼睛與嘴巴，難矣。另外值得一提的是化妝與光學鏡片的影響，若化妝與光學鏡片並未使「近似皮膚的顏色」改變太大，如仍是「近似皮膚的顏色」，如圖2 (b)中間未被轉換成黑色部份之部份，那將不會有任何影響。若已變為藍色或綠色或其它「非近似皮膚的顏色」，當然就不適用於本「皮膚顏色分割方法」。





(a)



(b)

圖2 (a)表示原始的RGB影像；(b)表示原始的RGB影像經由「皮膚顏色分割方法」過濾後的結果，中間未被轉換成黑色部份則是我們所採集的皮肤樣品之一部份。

圖3(a)、(c)與(e)表示含有真實人臉的原始RGB影像；(b)、(d)與(f)表示經由「皮膚顏色分割方法」過濾後的結果—「近似皮膚的顏色」（如：人臉的部份）仍被保持原來顏色，在此「近似皮膚的顏色」採從寬政策一如圖3(f)，而「非近似皮膚的顏色」（如：不是人臉的部份）則被處理成「純黑色」。如此一來，我們便可將真實人臉的原始RGB影像之「人臉的部份」（近似皮膚的顏色）與「不是人臉的部份」（非近似皮膚的顏色）區分開來。

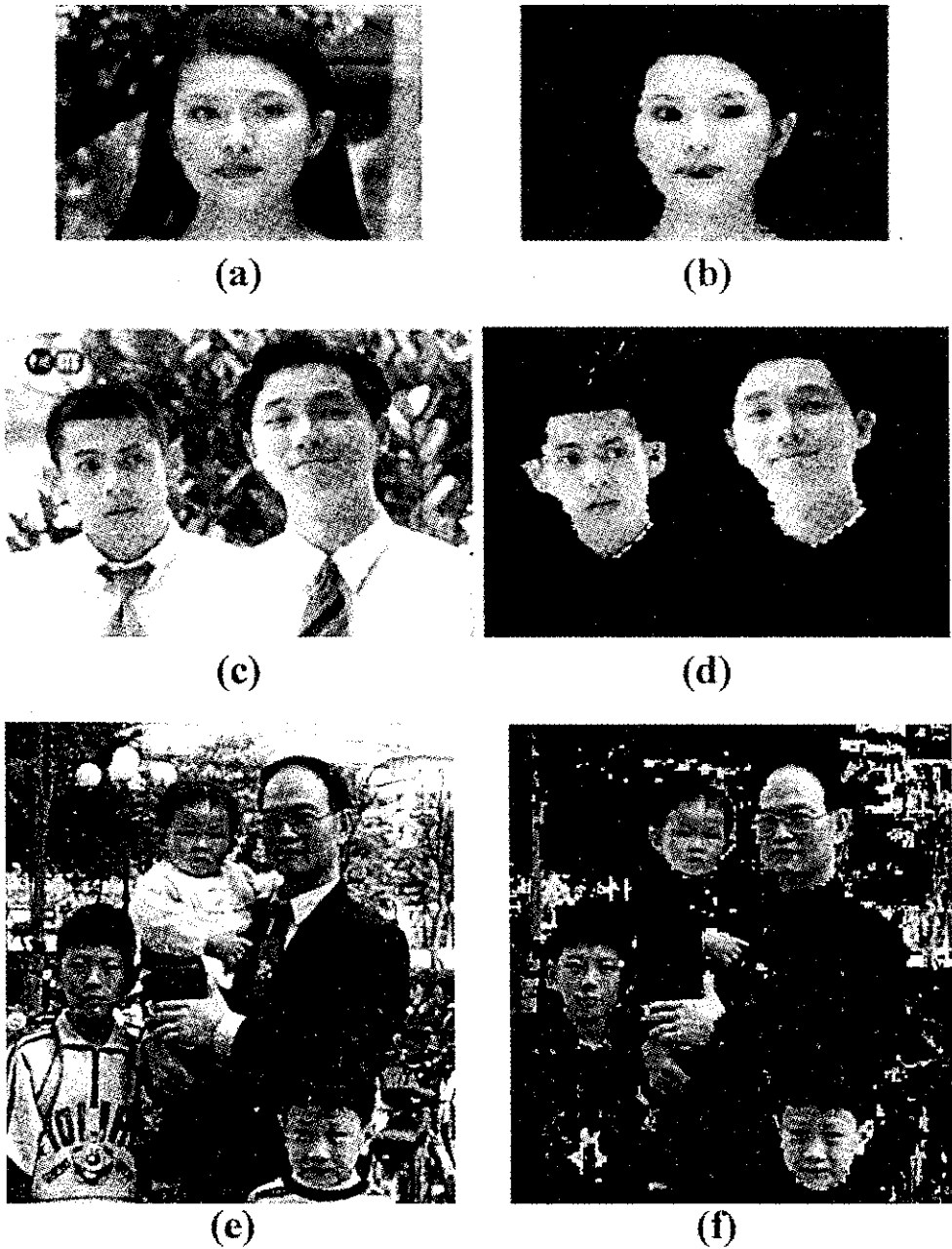


圖3 (a)、(c)與(e)示含有真實人臉的原始RGB影像；(b)、(d)與(f)表示含有真實人臉的原始RGB影像，經由「皮膚顏色分割方法」過濾後的結果。

圖4展示一個例子說明我們的「皮膚顏色分割方法」，對不同照明條件仍然表現優異。圖4(a)是原始輸入的RGB彩色影像。圖4(b)是原始輸入的RGB彩色影像經由「皮膚顏色分割」的結果。圖4(c)是光線較亮的RGB彩色影像。圖4(d)是光線較亮的RGB彩色影像經由「皮膚顏色分割」的結果。圖4(e)是光線較暗的RGB彩色影像。圖4(f)是光線較暗的RGB彩色影像經由「皮膚顏色分割」的結果。

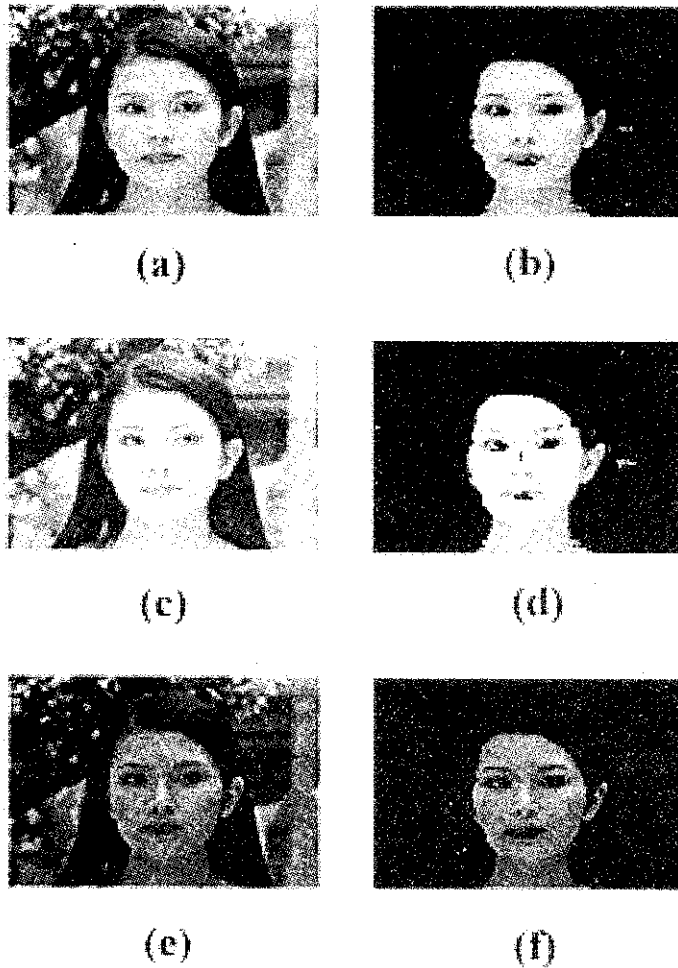


圖4 說明光線對「皮膚顏色分割方法」的影響

圖5展示一個例子說明我們的「皮膚顏色分割方法」，對室內影像仍然表現優異。圖5(a)與(c)是原始輸入的RGB彩色影像。圖5(b)與(d)是原始輸入的RGB彩色影像經由「皮膚顏色分割」的結果。

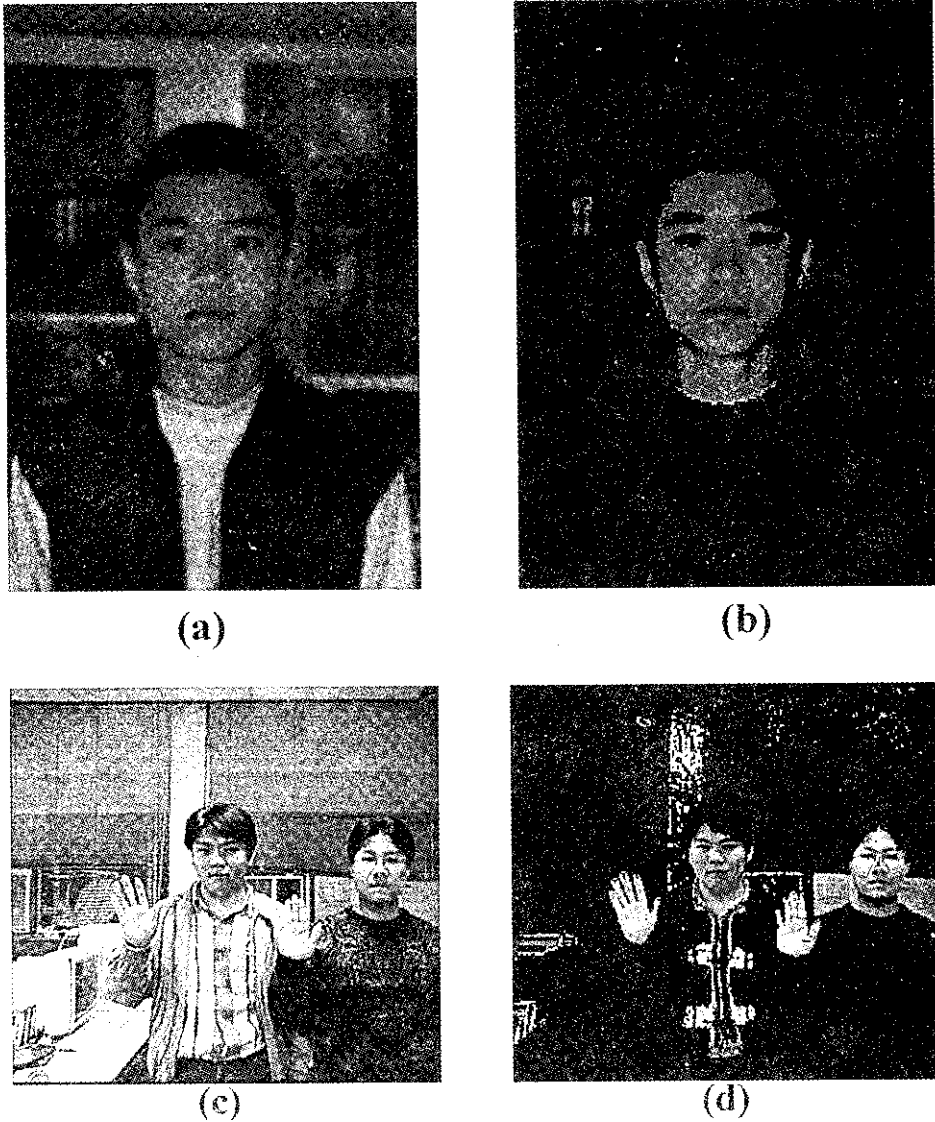


圖5 說明「皮膚顏色分割方法」對室內影像仍然表現優異

圖6展示一個例子說明我們的「皮膚顏色分割方法」，對不同對比(反差—contrast)與不同光線角度 (the direction of light) 影像仍然表現優異。圖6(a)是原始輸入的RGB彩色影像；(b)是高對比原始輸入的RGB彩色影像；(c) 是低對比原始輸入的RGB彩色影像。圖6(d)是(a)經由「皮膚顏色分割」的結果；圖6(e)是(b)經由「皮膚顏色分割」的結果；圖6(f)是(c)經由「皮膚顏色分割」的結果。

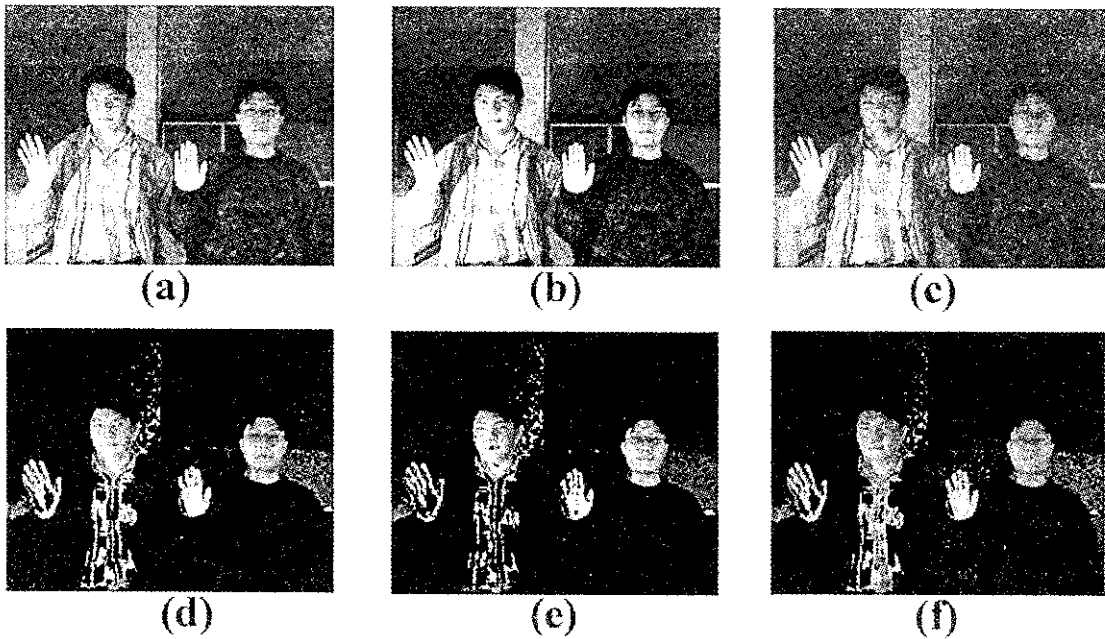


圖6 說明「皮膚顏色分割方法」對不同對比(contrast)影像仍然表現優異

## 二、彩色影像之二值化

在前一段，我們介紹了如何讀取一張彩色影像。然後，應用「簡單總體的門檻方法」(Simple global thresholding method)並取門檻值(threshold) 100，去二值化影像。如果二值化影像中的區塊數目比25大，我們就認為(假設)原始輸入RGB彩色影像有複雜背景。然後，藉由皮膚顏色分割方法，找出人臉皮膚顏色區域。在這一段，我們將介紹如何將這個經由「皮膚顏色分割後的影像」轉變成二值化的影像。

由於我們有興趣的物體(兩個眼睛與一個嘴巴)通常會比背景暗。所以，我們應用「簡單總體的門檻方法」(Simple global thresholding method)並取門檻值(threshold) 100。或許有人會質疑：用「簡單總體的門檻方法」且訂定門檻值100，真的可以適用於所有的情況嗎？它不會因不同的光線而使二值化結果失敗嗎？或許以下的例子——圖7，可以解答您的疑問。藉由圖7的結果，我們可以得知：我們的方法的確可以在一個極大的光線變化下，仍能成功地二值化影像。

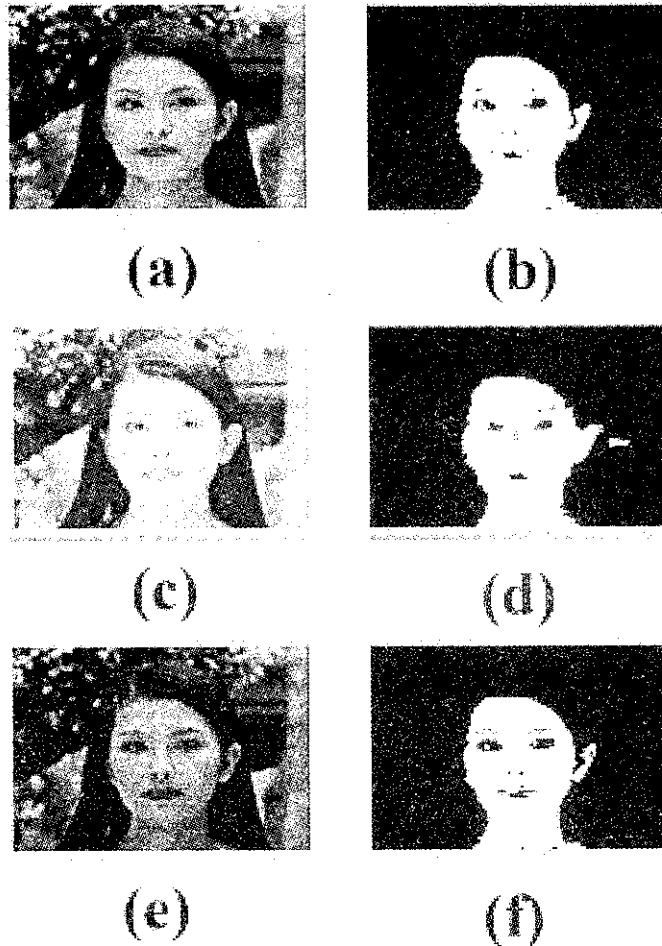


圖7 說明光線對經由「皮膚顏色分割方法」後，再將這個「皮膚顏色分割後的影像」轉變成「二值化影像」之影響。

### 三、尋找潛在人臉的區域

經過前兩個步驟後，我們將再藉由「尋找等腰三角形的關係」去得到潛在人臉的區域。這裡，我們使用光柵掃描 (raster scanning) (先從左而右，再從上而下) 以獲得各個連接的部分，再將各個連接的部分做標籤 (標記)。然後找到每一個區塊的中心。至於光柵掃描的細節有興趣者可於參考文獻 (R. C. Gonzalez, 1992) 查詢。經由許多小心觀察、實驗與分析，我們發現人臉有「兩個眼睛和一個嘴」恰好形成一個等腰三角形。而這就是我們尋找潛在人臉地區的絕佳索引。因此，我們從二值化的影像去搜查「兩個眼睛和一個嘴」的組合以尋找等腰三角形。再經由等腰三角形的三個頂點去推算潛在人臉的區域。至於如何找到等腰三角形的細節各位可以參考我們以前的研究論文 (C. H. Lin, accepted and to appear in Pattern Recognition)。在我們以前的研究論文中，在找到這個等腰三角形後，我們使用「左、右眼睛之間的歐幾里德距離」 (Euclidean distance) 去計算形成這個潛在人臉區域的四個角落點之座標。然而，我們發現當人臉左右轉動角度愈來愈大時，左、右眼睛之間的歐幾里德距離將變得越來越小。另一方面，我們發現右眼與嘴巴之間的歐幾里德距離卻仍舊保持幾乎相同的距離。因為我們認為人臉區域應該要包含眼睛、鼻子與嘴巴。因此，在此研究論文中我們改用「右眼與嘴巴之間的歐幾里德距離」去計算潛在人臉區域的四個角落點之座標。假設  $(X_i, Y_i)$ 、 $(X_j, Y_j)$ 、 $(X_k, Y_k)$  分別代表左、右眼睛與嘴巴三個區塊之三個中心點座標，如圖8(a)。而  $(X_1, Y_1)$ 、 $(X_2, Y_2)$ 、 $(X_3, Y_3)$  和  $(X_4, Y_4)$  分別代表人臉區域的四個角落點之座標，如圖8(b)。 $X_1$ 和 $X_4$ 位於的相同的座標  $(X_i - 1/4 * D(i, j))$ ； $X_2$ 和 $X_3$ 位於的相同的座標  $(X_k + 1/4 * D(i, j))$ ； $Y_1$ 和 $Y_2$ 位於的相同的座標  $(Y_i + 1/4 * D(i, j))$ ； $Y_3$ 和 $Y_4$ 位於的相同的座標  $(Y_j - 1/4 * D(i, j))$ 。在這裡  $D(i, j)$ 代表右眼與嘴巴之間的歐幾里德距離。

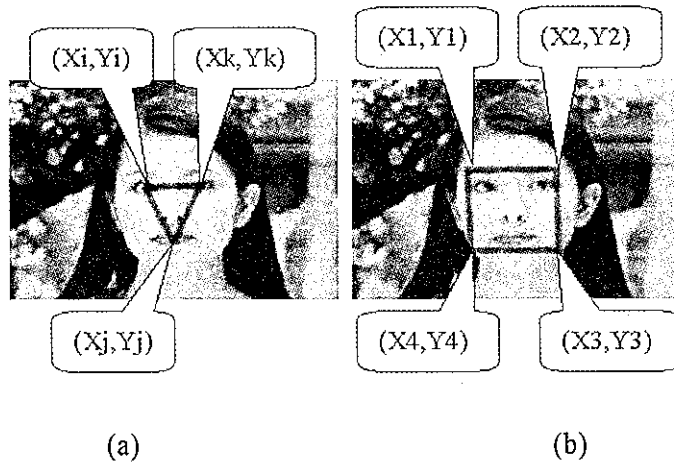


圖8 (a)  $(X_i, Y_i)$ 、 $(X_j, Y_j)$ 、 $(X_k, Y_k)$  分別代表左、右眼睛與嘴巴三個區塊之三個中心點座標；(b)  $(X_1, Y_1)$ 、 $(X_2, Y_2)$ 、 $(X_3, Y_3)$  和  $(X_4, Y_4)$  分別代表人臉區域的四個角落點之座標。

### 參、人臉角度分類

本系統的第二部分是要完成人臉角度分類的任務。在前段中，我們偵測出影像中的潛在人臉區域。而在這個章節中，我們創新提出運用三個不同權值面具函數來做人臉角度分類。這部分主要分兩步驟。第一步驟是要使所有潛在人臉區域的尺寸標準化。第二步驟是要把每個標準化的潛在人臉區域，都運用三種不同權值面具函數來做人臉角度分類。

#### 一、潛在人臉區域尺寸標準化

因為潛在人臉區域的尺寸都被將標準化（如 $60*60$ 像素），而尺寸標準化可去除人臉大小的影響（如相機和人臉的距離變大而使人臉變小）。因此，我們的系統可以偵測不同大小尺寸的人臉。我們所用來標準化的方法是「bicubic插補技術」（bicubic interpolation technique）。對於「bicubic插補技術」的細節有興趣者可於參考文獻（R. C. Gonzalez, 1992）中查詢。



## 二、人臉角度分類

把潛在人臉區域的尺寸標準化後，我們將每個標準化的潛在人臉區域都運用三種不同權值面具函數來做人臉角度分類。我們首先將每一個潛在臉的區域，都做了尺寸標準化的處理。然後，藉由人臉權值面具函數獲得每一個人臉的正確位置。其次，再藉由方向權值面具函數判斷人臉的正確方向。最後，再藉由角度權值面具函數決定人臉轉的角度。

### (一)人臉權值面具函數

在我們的系統中，我們的人臉權值面具是隨機取樣且經由加總十個二值化的人臉區域(面具一...面具十)，如圖9，這十個面具是經由影像處理軟體做手工切割而得到。

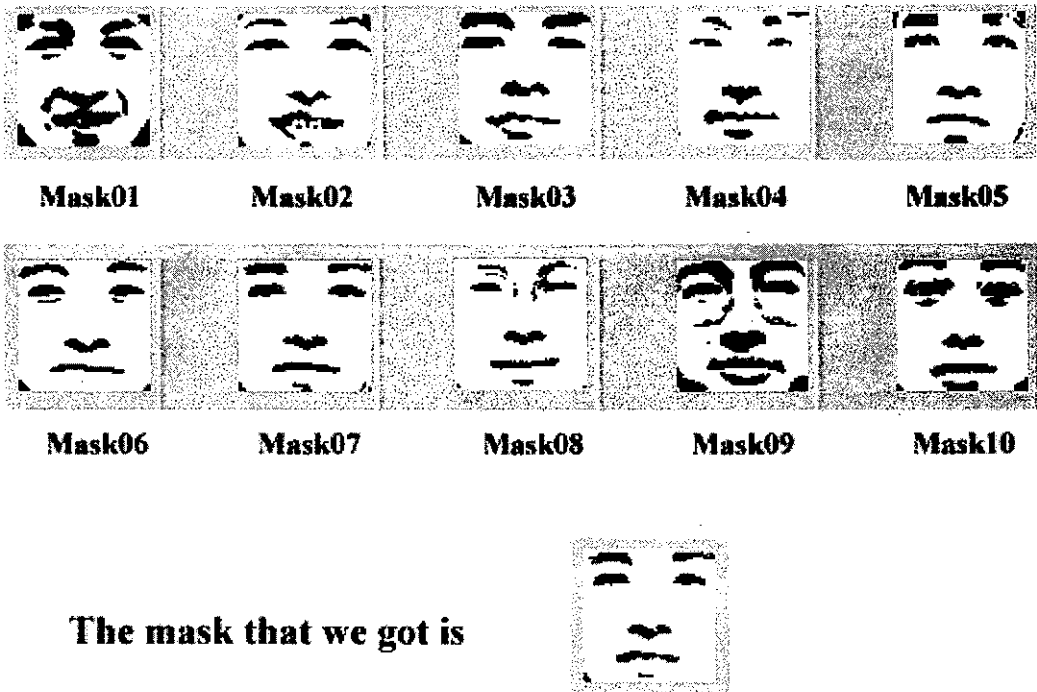


圖9 十個由影像處理軟體做手工切割而得之面具

$$\begin{array}{|c|c|c|} \hline 0 & 0 & 0 \\ \hline 0 & 0 & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline 0 & 0 & 1 \\ \hline 0 & 0 & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline 0 & 1 & 1 \\ \hline 0 & 0 & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 0 & 0 & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 0 & 0 & 1 \\ \hline 0 & 0 & 0 \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 0 & 1 & 1 \\ \hline 0 & 0 & 0 \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 0 & 0 & 0 \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 0 & 0 \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 0 \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array}$$

$$= \begin{array}{|c|c|c|} \hline 7 & 8 & 9 \\ \hline 4 & 5 & 6 \\ \hline 3 & 2 & 1 \\ \hline \end{array} \text{ Take threshold} = 5, \text{ then} = \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 0 & 1 & 1 \\ \hline 0 & 0 & 0 \\ \hline \end{array}$$

圖10 加總與得到人臉權值面具的方法

加總的方法如圖10，加總後再經由門檻值五（ $T = 5$ ）而得到人臉權值面具。例如：第一行第一列的加總值為7，而7大於等於5。因此，第一行第一列的最終值被定為1（代表該像素為白色）。在得到人臉權值面具後，我們即將尺寸標準化的潛在人臉區域與人臉權值面具互相比對。如果潛在人臉區域與人臉權值面具的兩隻眼睛、鼻子或者嘴皆位於相同位置（如：他們兩者皆是黑色像素），則我們將把權值加  $\alpha$  ( $\text{weight} = \text{weight} + \alpha$ )。如果潛在人臉區域與人臉權值面具的皮膚部分皆位於相同位置（如：他們兩者皆是白色像素），則我們將把權值加  $\beta$  ( $\text{weight} = \text{weight} + \beta$ )。如果潛在人臉區域的像素是黑色而人臉權值面具的像素是白色。那麼我們將把權值減  $\gamma$  ( $\text{weight} = \text{weight} - \gamma$ )。如果潛在人臉區域的像素是白色而人臉權值面具的像素是黑色。那麼我們將把權值減  $\delta$  ( $\text{weight} = \text{weight} - \delta$ )。經由實驗—跑程式、分析實驗結果與細心觀察，我們發現 $\alpha = 6$ 、 $\beta = 2$ 、 $\gamma = 4$ 與 $\delta = 2$ 最為恰當。因為如此一來，人臉區域的權值最為穩定。換句話說，人臉區域的權值範圍最小。

## (二)方向權值面具函數

藉由人臉權值面具函數，我們獲得每一個人臉的正確位置後，我們將藉由方向權值面具函數來判斷人臉是向前看、向左看或向右看。我們使用類似人臉權值面具的方法，但是不同的 $\alpha$ 、 $\beta$ 、 $\gamma$ 與 $\delta$ 值。而 $\alpha$ 、 $\beta$ 、 $\gamma$ 與 $\delta$ 的值乃是經由許多實驗與測試而得。經由實驗—跑程式、分析實驗結果與細心觀察，我們發現 $\alpha = 8$ 、 $\beta = 2$ 、 $\gamma = 4$ 與 $\delta = 2$ 最為恰當。（向前、向左及向右看的方向權值面具函數使用相同的 $\alpha$ 、 $\beta$ 、 $\gamma$ 與 $\delta$ 值）。

在方向權值面具函數中，我們使用了三類面具：(1)向前看的面具(2)向左看的面具(3)向右看的面具。因為我們相信如果潛在人臉區域是向前看的，則應該對向前看之面具[由10個二值化的向前看的面具（隨機取樣）所形成的]，有著較高的相似度。另一方面，如果潛在人臉區域是向左看的，則應該對向左看之面具[由10個二值化的向左看的面具（隨機取樣）所形成的]，有著較高的相似度。同樣地，如果潛在人臉區域是向右看的，則應該與向右看之面具[由10個二值化的向右看的面具（隨機取樣）所形成的]，有著較高的相似度。因此，決定人臉是看那個方向則可根據權值的大小而判斷之。而這三十個面具（向前、向左、向右各十個）不再是經由影像處理軟體去手工切割而得到。取而代之的是由前述的偵測「潛在人臉區域」的方法做自動切割而來。

### (三)角度權值面具函數

經由前述幾個步驟後，我們將再藉由角度權值面具函數去分類人臉轉的角度。我們亦使用類似人臉權值面具的方法，但是不同的 $\alpha$ 、 $\beta$ 、 $\gamma$ 與 $\delta$ 值。而 $\alpha$ 、 $\beta$ 、 $\gamma$ 與 $\delta$ 的值亦是經由許多實驗與測試而得。經由實驗 — 跑程式、分析實驗結果與細心觀察，我們發現 $\alpha = 18$ 、 $\beta = 2$ 、 $\gamma = 4$ 與 $\delta = 2$ 最為恰當。因為如此一來，人臉角度權值面具函數產生之權值差異較大。而這十個面具（向前看的）亦不再是經由影像處理軟體做手工切割而得。取而代之的亦是由前述的偵測「潛在人臉區域」的方法去自動切割而來。因為我們相信如果潛在人臉區域是向前看的，則應該對向前看之面具，有著較高的相似度。而向左、向右看的人臉，亦將隨著人臉向左、向右看的角度變大，而相似度愈變愈小。換言之，隨著人臉向左、向右看的角度變大，其權值亦將愈變愈小。因此，我們根據權值的大小去分類人臉轉的角度大小。

我們的系統將人臉轉的角度大小分成五類(1)向前看（如果權值比7,500大）；(2)在向前看和向左看20度之間（權值在7,499和 6,300之間）；(3)在向前看和向右看20度之間（權值在7,499和 6,300之間）；(4)向左看大於20度（權值比6,299更小）；(5)向右看大於20度（權值比6,299更小）。

茲舉例如圖11，圖11(a) 是一張100\*132像素的彩色影像。應用「簡單總體的門檻方

法」(Simple global thresholding method)並取門檻值(threshold) 100,去二值化影像。結果二值化影像中的區塊數目是12(比25小),我們就認為原始輸入RGB彩色影像沒有複雜背景。因而我們將跳過「皮膚顏色分割方法」,而直接將原始輸入RGB彩色影像變成二值化的影像如圖11(b)。經由第二章第三節尋找潛在人臉的區域之方法,我們從二值化的影像去搜查「兩個眼睛和一個嘴」的組合找等腰三角形如圖11(c)。再經由等腰三角形的三個頂點去推算二值化的潛在人臉的區域如圖11(d)。再將潛在人臉區域的尺寸標準化為60\*60像素如圖11(e)。之後,再將圖11(e)標準化之潛在人臉區域套入「人臉權值面具函數」。因為我們相信如果潛在人臉區域是向前看的,則應該對向前看之面具[由10個二值化的向前看的面具(隨機取樣)所形成的],有著較高的相似度。由於此例中,向前看的「人臉權值面具函數」之權值比其他兩個「人臉權值面具函數」(向左看的與向右看的)為高。因此,我們判定此潛在人臉區域是向前看的。最後,再套入「角度權值面具函數」去分類人臉轉的角度,由「角度權值面具函數」得到之權值為7988(大於7500),亦佐證此潛在人臉區域是向前看的。如圖11(f)則是再經由等腰三角形的三個頂點去框出RGB彩色影像的潛在人臉的區域。最後,如圖11(g)則是在原RGB彩色影像中框出人臉偵測之結果。

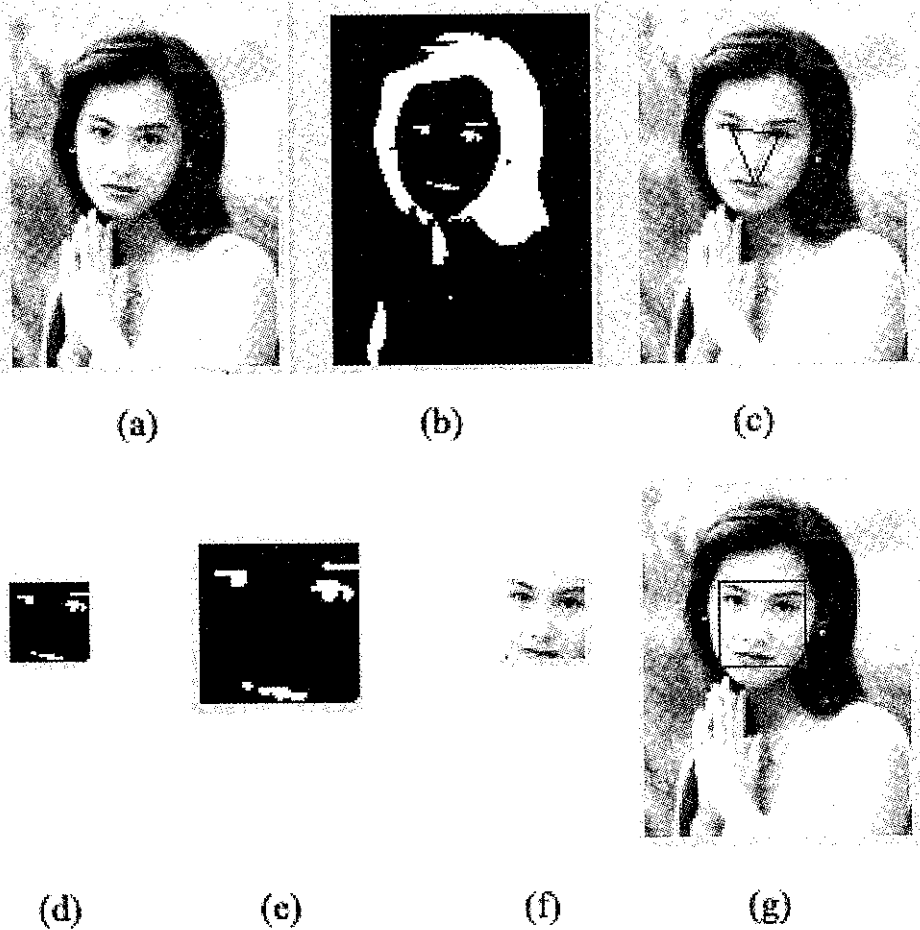


圖11 自動分類臉部角度之細部步驟

## 肆、實驗結果和討論

我們使用 650 個影像（每人 13 個影像）用以測驗我們的系統。這些測驗影像的尺寸是 352\*240 個像素。在這些測驗影像中包含在各種環境中捕獲的人臉影像。結果顯示沒有被正確地分類者僅僅 6 個。實驗結果得到一個近似百分之九十九（644/650 =

99.08 %) 的成功率。而相關錯誤警報比率 (the relative false alarm rate) 低於百分之一 ( $6 / 650 = 0.92$  百分比)。這些影像是使用彩色攝影機 (V8) 與彩色數位攝影機 (DV) 拍攝而得。我們的系統能將人臉轉的角度大小正確地分成五類(1)向前看 (如果權值比 7,500大) ; (2)在向前看和向左看20度之間 (權值在7,499和 6,300之間) ; (3)在向前看和向右看20度之間 (權值在7,499和6,300之間) ; (4)向左看大於20度 (權值比6,299更小) ; (5)向右看大於20度 (權值比6,299更小) 。

在這個研究中，要把人臉轉的角度大小正確地分類，所需的時間取決於影像的尺寸和複雜度。圖12(a) 是一張352\*240像素的彩色影像，倘若使用PII 233的個人電腦只需要 2.5 秒即可正確地分類人臉轉的角度大小。圖12(a)代表向前看的人臉與其權值 (比7,500大) ; 圖12(b)與(c)代表在向前看和向左看20度之間 (權值在7,499和 6,300之間) ; 圖12(d)、(e)、(f)與(g)代表向左看大於20度 (權值比6,299更小) ; 圖12(h)與(i)代表在向前看和向右看20度之間 (權值在7,499和 6,300之間) ; 圖12(j)、(k)、(l)與(m)代表向右看大於20度 (權值比6,299更小) 。圖13(a)代表在向前看和向右看20度之間 (權值6,770在 7,499和 6,300之間) ; 圖13(b)代表在向前看和向左看20度之間 (權值6,448在7,499和 6,300之間) ; 圖13(c)代表向前看的人臉與其權值 (7,656比7,500大) ; 圖13(d)代表向左看大於20度 (權值5,206比6,299更小) ; 圖13 (e)代表在向前看和向右看20度之間 (權值 6,754在7,499和6,300之間) 。

彩色影像之人臉角度分類



圖12 自動分類臉部角度之結果



(a) weight = 6770 (Looking right < 20 & > 0)



(b) weight = 6448 (Looking left < 20 & > 0)



(c) weight = 7656 (Looking front)



(d) weight = 5206 (Looking left > 20degrees);  
(e) weight = 6754 (Looking right < 20 & > 0)

圖13 自動分類臉部角度之結果

## 伍、結 論

在本文中，我們介紹一個健全且有效的人臉角度分類系統。我們創新提出利用三種面具函數：(一)人臉權值面具函數(二)方向權值面具函數(三)角度權值面具函數去自動分類人臉角度。且本系統的訓練(只需要向前看、向左看及向右看之面具各10個即可)，不需要如類神經網路的訓練般需花費許多時間。如Juell等人(1996, pp. 781-787)的系統



訓練需花費許多時間。另外，我們對複雜背景的RGB彩色影像利用「皮膚顏色分割方法」找出人臉皮膚顏色區域，並將「人臉的部份」（近似皮膚的顏色）保留，且將「不是人臉的部份」（非近似皮膚的顏色）處理成「純黑色」，去除了大部份的複雜背景。在此需特別聲明：我們利用「皮膚顏色分割方法」的主要目的是在於去除了大部份的複雜背景。如此一來，更使得整個系統對複雜背景的RGB彩色影像的執行速度加快許多。而且，在我們的系統中，不需要任何彩色空間之轉換。即可對RGB彩色影像直接做皮膚顏色分割。在我們的系統中，我們能夠處理複雜環境的彩色影像、不同尺寸的人臉及容許不同的光線與表情。實驗結果，證實了我們的新方法在效率和精確度上，遠比傳統方法略勝一籌。

## 參考文獻

- R. Brunelli, (1997), "Estimation of Pose and Illuminant Direction for Face Processing" Image and Vision Computing, Vol. 15, pp. 741-748.
- Y. Dai and Y. Nakano, (1996), "Face-texture model based on SGLD and its application in face detection in a color scene", Pattern Recognition, vol. 29, no. 6, pp. 1007-1017.
- S. Gong, S. McKenna, J.J. Collins, (1996), "An investigation into face pose distributions", Proceedings of IEEE the Second International Conference on Automatic Face and Gesture Recognition, 1996, Page(s): 265 -270
- S. Gong, E.J. Ong and P. J. Loft, (1997), "Appearance-based face recognition under large head rotations in depth", (The paper is gotten from <http://www.dcs.qmw.ac.uk/~sgg/msc/hbt/>)
- R. C. Gonzalez and R. E. Woods, (1992), "Digital Image Processing", copyright © 1992 by Addison-Wesley Publishing Company, Inc.
- P. Juell and R. Marsh, (1996), "A hierarchical neural network for human face detection", Pattern Recognition, vol. 29, no. 5, pp. 781-787.

- C. H. Lee, J. S. Kim and K. H. Park, (1996), "Automatic Human Face Location in a Complex Background Using Motion and Color Information," *Pattern Recognition*, Vol. 29, No. 11, pp.1877 –1889.
- C. H. Lin, K.C. Fan, (2001), "Triangle-based Approach to the Detection of Human Face", accepted and to appear in *Pattern Recognition*
- H. A. Rowley, S. Baluja, and T. Kanade, (1995), "Human face detection in visual scenes", Tech. Rep. CMU-CS-95-158R, Carnegie Mellon University, (The paper is gotten from <http://www.cmu.edu/~har/faces.html>)
- J. Sherrah and S. Gong, (1999), "Fusion of 2D Face Alignment and 3D Head Pose Estimation for Robust and Real-Time Performance" *Proceedings of IEEE International Workshop on Recognition, Analysis and Tracking of Faces and Gestures in Real-Time Systems*, Corfu, Greece September 1999, pp. 26-27)
- K. Sobottka and I. Pitas, (1996), "Extraction of facial regions and features using color and shape information", in *Proc. 13th International Conference on Pattern Recognition*, Vienna, Austria, Aug. 1996, pp. 421-425.
- S. Waldherr, Sebastian Thrun, Roseli Romero, and Dimitris Margaritis, (1998), "Template-Based Recognition of Pose and Motion Gestures on a Mobile Robot", Computer Science Department, Carnegie Mellon University, Pittsburgh, PA, 1998
- L. Wiskott, J. M. Fellous, N. Krijger, and C. von der Malsburg, (1997), "Face Recognition by Elastic Bunch Graph Matching", *IEEE transactions on pattern analysis and machine intelligence*, vol. 19, no. 7. July 1997, pp.769-775.
- H. Wu, Q. Chen, and M. Yachida, (1996), "A fuzzy-theory-based face detector", in *Proc. 13th International Conference on Pattern Recognition*, Vienna, Austria, Aug. 1996.
- R. P. Würtz, "Object recognition robust under translations, deformations, and changes in background", *IEEE transactions on pattern analysis and machine intelligence*, vol. 19, no. 7. July 1997, pp. 775-779.