

第五章

轉描法及降色原理

在本章節中，我們將介紹如何運用另一種卡通化轉換—轉描法的實作原理及細節，另外還有如何針對輸入圖片，在前處理之後除了臉部區域之外的背景影像的降色機制。於第二章的相關研究之中，我們介紹過其他非真實影像繪圖的原理，此處轉描法實作，大體上也是承續同樣的原則對於整個的影像區域來進行一種特別的影像過濾器實作，然而我們亦同時擁有有關臉部特徵點的定義相關資訊，因此如何使用這些特徵點來加強真實影像中的抽象化結果，便是本章主要闡述的重點。

5.1 邊緣偵測 (Edge Detection)

邊緣偵測核心是影像處理的基礎[25]中常運用來抽象化的一種工具，其設計的方式也因應不同的影像內容而定，常用如 Sobel 及 Laplacian 及 Canny 演算法等，這些典型的邊緣偵測方法也都內建在常用的影像處理平台或函式庫之中，可以進行簡單的測試，且與其他影像處理核心進行多種過濾器的組合實驗，來取得適合的參數及結果影像。就人臉的影像資訊特色來說，邊緣偵測往往也是可以用來尋找雙眼及嘴唇位置的參考，將邊緣偵測的結果以 x 軸投影的邊緣強度分布可以偵測嘴巴的大致位置，而對 y 軸做投影的強度分布則大致指出了雙眼的相對位置。

以下圖 5.1 是我們使用邊緣偵測方法來進行影像內容的簡化的結果：



圖 5.1、使用邊緣偵測進行影像線條簡化

典型的邊緣偵測核心在影像抽象化過程中，大體都是運用影像中的像素在水平或垂直方向的梯度變化，來進行實作過濾器，而人臉在下巴及脖子部份由於都是接近的膚色，因此是人臉影像簡化中最難處理的部份，而我們可以由特徵點定義中的臉部中段所形成的開放輪廓來補強這一個部份的弱點（見圖 5.2）。



圖 5.2、使用特徵點定義資訊補足封閉形狀

然而僅使用這些邊緣仍未足準確產生進行轉描效果的卡通肖像，在轉描的過程中，可以分為剪影及填色兩大主要過程，第一步的剪影，首先需要是選取一個封閉的形狀區間，除了利用邊緣偵測核心之外，另外我們也可以調整不同的核心來對真實影像進行更

多保留細節的抽象化。

經過調整的核心相對於典型的邊緣偵測，除了保留更多影像內容中的細節，而且也具有簡略的內容區塊分割，運用這個簡單的 3x3 Convolution 核心來進行計算，可以獲得影像內容中的梯度 (gradient) 資訊，對梯度資訊作排序，約保留原始影像大小的三分之一的像數量，我們可以將真實的影像轉換成類似黑白漫畫的效果 (見圖 5.3)。

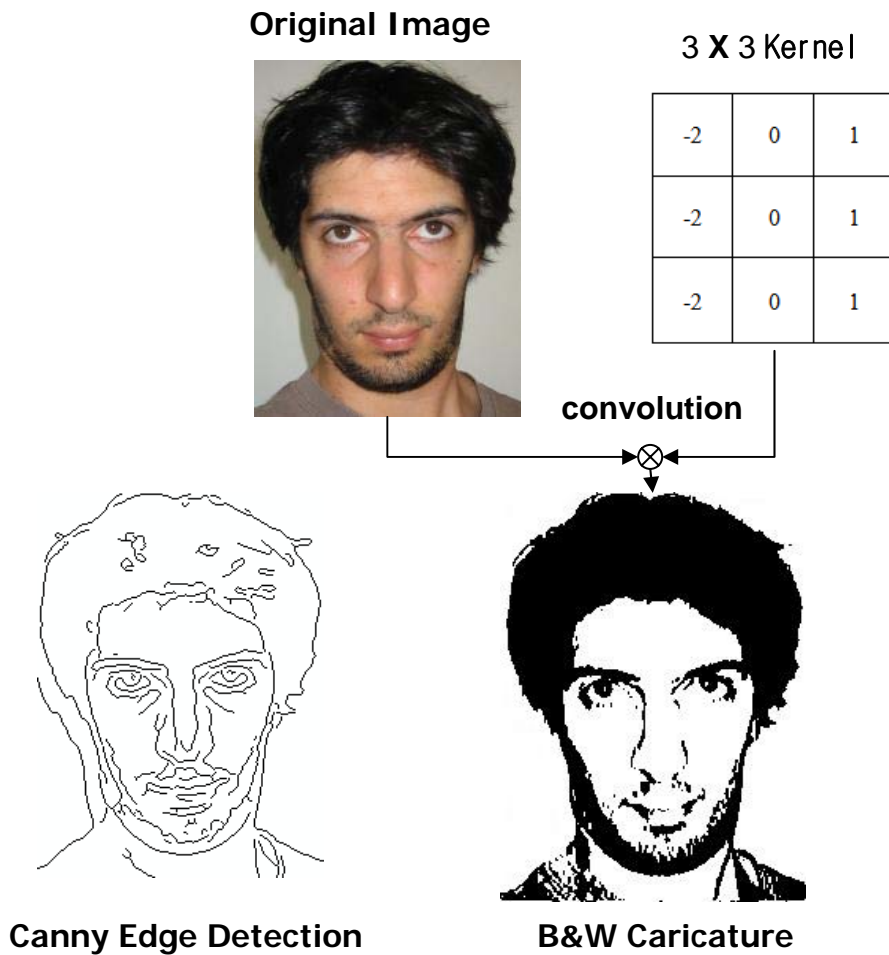


圖 5.3、運用 Canny、梯度濾波器及的影像抽象化

以人臉影像中主要的形狀區間就是我們的五官，這些五官就運用已知特徵點來標記

成封閉區間。接下來的填色主要目的為選取一個代表顏色來填滿此封閉區間。我們將此封閉區間的影像資料，可以選定以各種不同的顏色空間（Color Space），如 RGB、LUV 等來表示，並且將此區域的顏色分布視為一個機率分布，則顏色的選擇策略可以從以下兩種來決定：

- 平均色（average color）：將區間的顏色資訊加總再除以區間面積來取的平均的顏色。（見圖 5.4a）
- 接近自選色（approximates to manually selected color）：依特定應用情境，將由使用者自定輸出裝置所支援的色盤資訊，區塊顏色則由色盤中最接近的顏色替補。（見圖 5.4b）



圖 5.4、運用五官特徵點的封閉形狀的填色結果，左圖為平均色，右圖為使用者自定四種顏色（黑、白、綠、紅）。

5.2 Mean Shift 協助內容分割（Segmentation by Mean Shift）

Mean Shift[20][21]是一種非參數式的特徵分析演算法，其利用反覆計算空間中的群集技巧，可以運用在尋找各種複雜的形狀群集之中，在不同的群集間可以運用 Mean Shift

的技巧估計出其邊界來分割，一開始的提出這個技巧是將其應用是在電腦視覺中的物體追蹤辨識之上，而同樣的技巧運用在影像處理的領域中也常使用在影像的內容分割及降色之上，以下將簡單介紹其技術的原理及我們如何使用在人臉影像中的抽象化及內容分割。

利用通用數學公式，我們可以將影像內容的分割問題轉變為處理特定維度的空間中所散布的資料點，這些資料點的分布成特定的群集，則我們可運用求出這個資料分布中的梯度變化估計出所謂的 Mean Shift 向量；將特定維度假設為在 d 維度的空間 \mathbf{R}^d 中，資料分布則用 \mathbf{x}_i 表示，其中 $i = 1, \dots, n$ 共有 n 筆資料點。再運用多變量核 $K(\mathbf{x})$ 及固定窗口的大小為 h ，可以估計資料分布在多變量的空間下的密度分布函數寫為：

$$f(\mathbf{x}) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right) \quad (5.1)$$

$$K(\mathbf{x}) = c_{k,d} k(\|\mathbf{x}\|^2) \quad (5.2)$$

由於 $K(\mathbf{x})$ 核心的幅射對稱 (radially symmetric) 特性，因此我們也可以將 $K(x)$ 標註為 (5.2)，其中的正規化常數 $c_{k,d}$ ，是保證 $K(\mathbf{x})$ 的分布在 0 至 1 區間；且在梯度 (gradient) 變化為 $\nabla f(\mathbf{x}) = 0$ 時將滿足密度分布函數 $f(\mathbf{x}) = 0$ 的特性。因此運用梯度變化的定義及我們將公式 (5.1) 中增加一個變量空間，我們也可以將之改寫成下面的形式：

$$\nabla f(\mathbf{x}) = \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^n (\mathbf{x}_i - \mathbf{x}) g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right) \quad (5.3)$$

此處我們定義 $g(s) = -k'(s)$ ， $G(\mathbf{x}) = c_{g,d} g(\|\mathbf{x}\|^2)$ 為新的多變量核依據定義，梯度變化 $\nabla f(\mathbf{x})$ 可以再展開為公式 (5.3)，這裡我們便將最後一項視為 $m_h(\mathbf{x})$ ，即 Mean Shift 向量。

$$\nabla f(\mathbf{x}) = \frac{2c_{k,d}}{nh^{d+2}} \left[\sum_{i=1}^n g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right) \right] \left[\frac{\sum_{i=1}^n \mathbf{x}_i g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right)} - \mathbf{x} \right] \quad (5.4)$$

由以上定義，可知在 Mean Shift 的計算過程中可以分為以下兩部份：

- 計算 Mean shift vector $m_h(\mathbf{x}^t)$
- 運用滑動視窗的移動反覆計算： $\mathbf{x}^{t+1} = \mathbf{x}^t + m_h(\mathbf{x}^t)$

由前所述資料分布的密度函數 (5.1) 具有在梯度函數 (5.4) 為 0 時收斂的特色，整個 Mean Shift 的運作方式可以參考圖 5.5，由此可以體現 Mean Shift 可以尋找不同區塊的中心所在 (梯度變化為 0 處)，相當適合使用在影像的內容分割，分割結果參見圖 5.6。

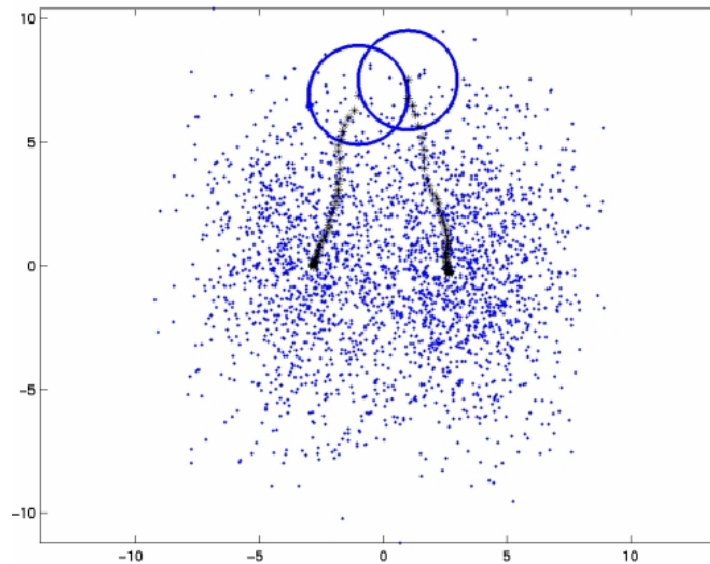


圖 5.5、Mean Shift 邊界尋找[20]

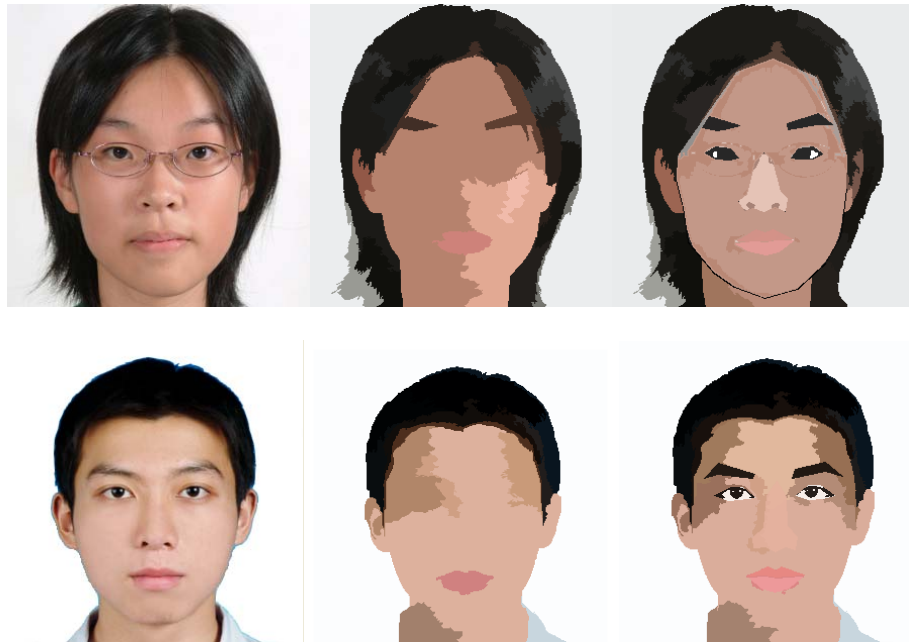


圖 5.6、使用 Mean-Shift 進行顏色抽象化及轉描法生成卡通肖像

5.3 降色與影像內容抽象化 (Color Quantization and Abstraction)

影像內容的抽象化，大體可以應用於邊緣線條與顏色區塊兩大部份，在邊緣線條中我們可以運用前述的 5.1 節中所提的濾波器來抽象化產生描邊式的卡通化肖像。而顏色區塊的抽象化，則是運用降色[5]的演算法來將原來的全彩影像，簡化成僅運用少數代表顏色的方式來替代，如此抽象化過程中可以粗略的形成少數顏色區塊，間接構成簡單的影像內容分割。

在前一節中，我們已提及如何在影像像素空間中使用 Mean Shift 來達成影像內容的區塊分割，在本節之中，為達成影像內容的降色，我們亦可以將 Mean Shift 的群聚機制加入顏色空間的參數，整合為一個以顏色區塊來尋找邊界的影像分割方式。新的多變量核加入顏色空間後，可以重新寫成以下的形式：

$$K_{h_s, h_r}(\mathbf{x}) = \frac{C}{h_s^2 h_r^p} k\left(\left\|\frac{\mathbf{x}^s}{h_s}\right\|^2\right) k\left(\left\|\frac{\mathbf{x}^r}{h_r}\right\|^2\right) \quad (5.5)$$

在新的多變量核之中，運用不同的 h_r 表示 p 個維度的不同的顏色空間，例如：LUV 顏色空間僅運用 L 一個維度則 $p = 1$ 或 RGB 顏色空間模型中，則 $p = 3$ 。為取得較佳的影像分割，一般常取用 LUV 中的 lightness 來做為 h_r 的分割限制，雖然可以有不錯的降色和分割結果，但是在大部份以 RGB 的輸出裝置中，結果影像的色盤無法僅限於限定的顏色種類，Mean Shift 降色及內容分割結果如圖 5.7 所示。

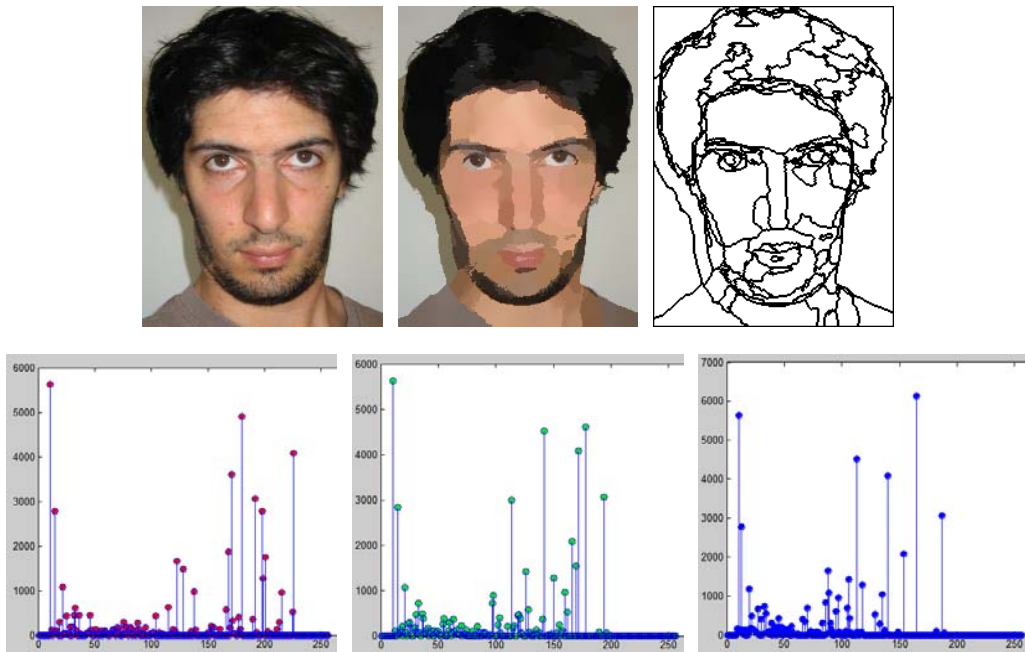


圖 5.7、Mean Shift 進行 LUV 影像分割 (h_s, h_r) = (7, 6.5)，
下圖分別為 R, G, B 個別顏色空間 histogram

除了使用 Mean Shift 以外，也仍有其它的降色演算法可以運用，包括有 Median-Cut[27]、Agglomerative Clustering[28]、Octree[29]等，這些演算法的運作過程大體分為以下三大步驟：

- 抽取代表色盤過程：此過程中將運用上述相關演算法來決定新的色盤，在其用

來所顯現的降色影像與原始影像間，我們希望能在像素扭曲錯誤率（Pixel Distortion Error）上可符合最小化的條件。

- 遞色（Dithering）過程：遞色的問題是根據人類視覺的生理基礎所研究的一個重點，在訊號處理的理論中，前面的代表色盤相當於對訊號來重新取樣，然而重新取樣後的訊號，由於取樣頻率的降低，將產生有許多不平順或不連續的訊號，就經由遞色過程來將這些訊號的斷裂處加入一些不規律的雜訊來進行欺騙感知系統，常用的演算法如 error diffusion[30]及 Floyd dithering[31]。
- 剔除（Pruning）過程：選擇性的將較小的顏色區塊剔除或和附近的主要區塊合併，以尋求較佳的分割結果。

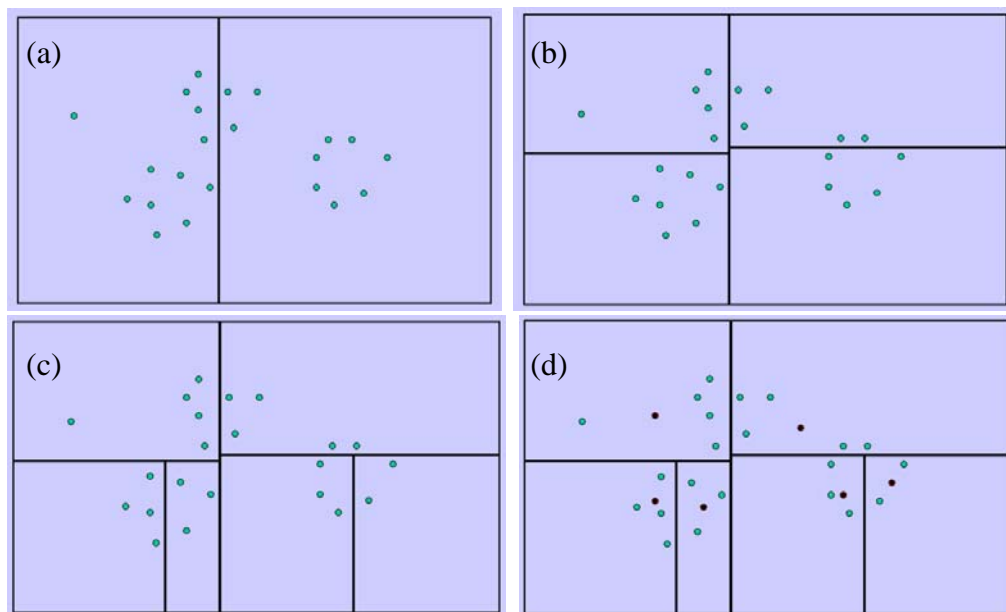


圖 5.8、Median Cut 運作過程示意圖，由(a)至(d)

在前述諸多降色演算法之中 Median-Cut[27]是最常被提及和使用的一種，其原理是將影像在 RGB 的個別顏色空間中將範圍最長的一個維度切割成兩半，然後對半作遞迴，直到達到欲數目滿足新色盤的總數為止，則每一個被切半的區塊 median 作為色盤

中的代表顏色（運作過程見圖 5.8，d 中的黑色實心即為最後挑選的代表顏色）。

Median-Cut 的方式具有易於實作的優點，但其降色的結果大體還是遵循由個別像素的循序遞色而產生，因此在影像分割的效果上顏色區塊仍略顯瑣碎，但其優勢是在限定顏色的情境中，可以分毫不差的將影像降色為所需的色階，因此在針對限定顏色的應用情境下，將之搭配 5.4 節中的 Mean Shift 來配合使用，以生成符合輸出裝置需求的卡通化肖像的結果影像（如圖 5.9）。

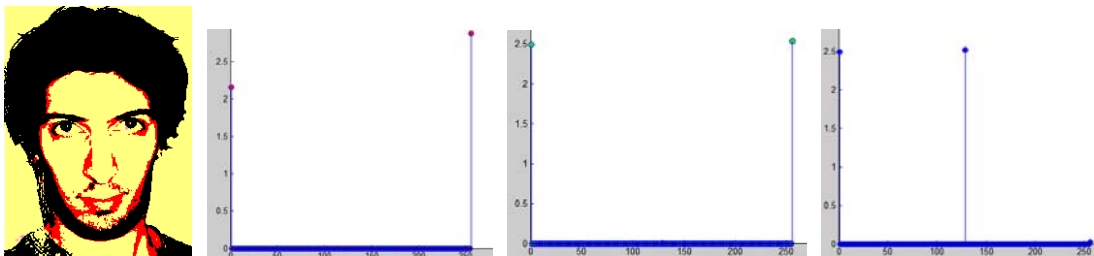


圖 5.9、使用 Median Cut 進行限定顏色（黑、紅、黃）的降色取樣
下方三圖分別為：R, G, B 個別顏色空間 histogram

5.4 以直方圖規格化輔助色彩抽象化 (Histogram Specification)

在前面，我們運用 Mean-Shift 以及 Median-Cut 等演算法來進行影像降色取樣，以及討論了不同的降色演算法產生的結果影像的在 RGB 顏色空間中的顏色直方圖分布，在前面提出的顏色抽象化方法，都是以輸入的真實影像本身的內容為依據，進行具代表性顏色抽取，最後達到降色的效果。

但在我們的應用中，亦使用到的卡通影像資料庫裡，仍擁有許多不同的用色風格的卡通藝術影像，因此我們的思考方式也有了另一種的選擇：將真實影像的色彩轉換至參考的卡通影像顏色風格；直方圖規格化[33] (Histogram Specification) 的技術是達成此

種色彩風格的轉換最直接的選項。藉由原始影像的顏色直方圖分布資訊，此處我們亦可考慮使用在顏色直方圖空間作進行顏色的重新對應（轉換結果參考圖 5.10）的轉換技巧來解決卡通化後結果的膚色一致性和背景顏色契合度的問題。

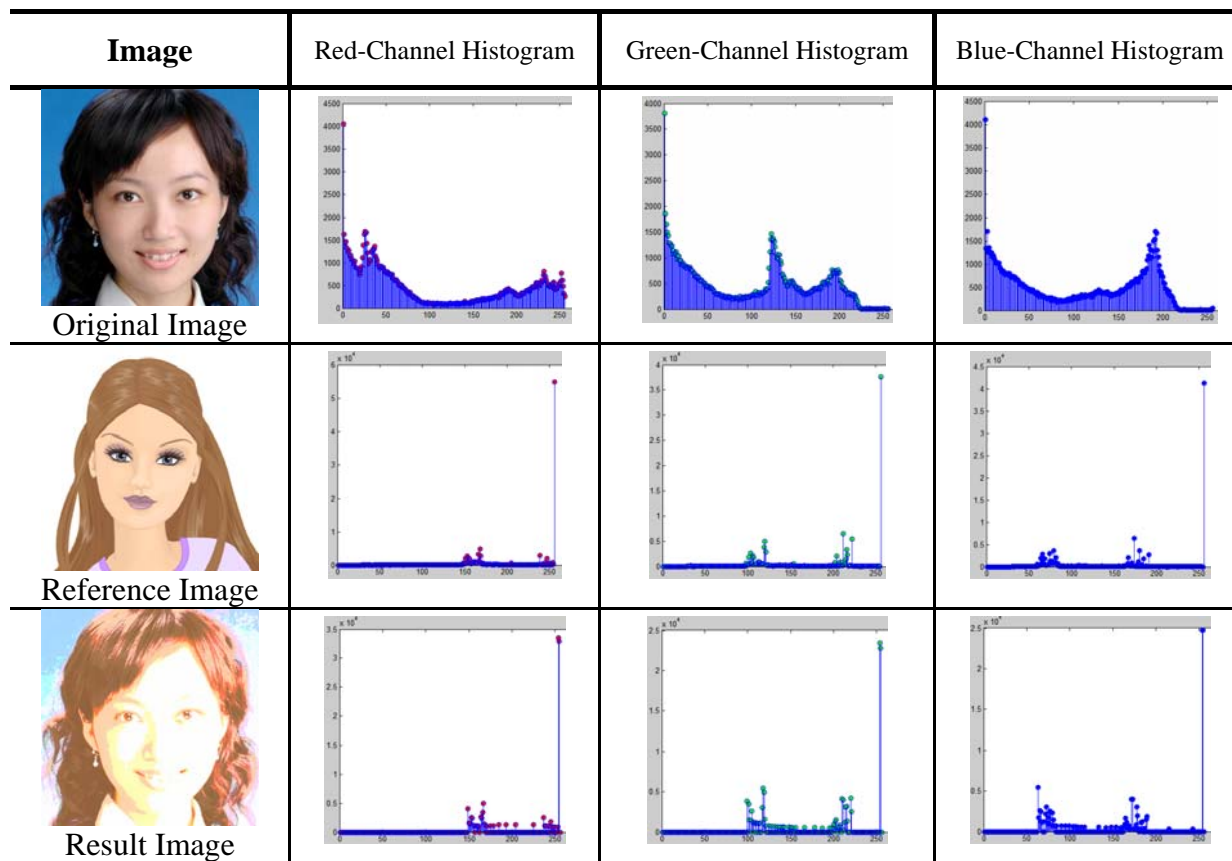


圖 5.10、運用直方圖規格化校正輸入影像顏色分布

顧名思義，直方圖規格化（Histogram Specification），就是將影像的顏色直方圖視為一種機率分布（PDF）方式，藉由被參考影像的分布方式，尋求一個目標轉換的函數對應，能將原始影像轉換成具有參考影像的顏色分布特性的一種演算法[38]。進行規格化之前，由於兩個影像的大小維度不同，必需先針對兩者的顏色直方圖的累積分布機率作均化（Histogram Accumulation Equalization）[25]，意即將依序每個顏色出現的次數作累積，並且除以全部樣本總數（總像素）的次數，讓整個顏色直方圖累積分布的機率的和

可以正規化為 1 (見公式 5.1)。

$$S_k = T(r_k) = \sum_{j=0}^k p_r(r_j) = \sum_{j=0}^k \frac{n_j}{n}, \quad k = 0, 1, 2, \dots, L-1 \quad (5.1)$$

欲從原始影像參考另一影像的顏色分布來進行分布同化，我們可以把原始影像 x 及參考影像 y ，在連續性的假設下將累積機率分布 h_x 、 h_y 以下列式子表示：

$$h_x = f(x) = \int_0^x p_r(u)du, \quad h_y = g(y) = \int_0^y p_r(v)dv \quad (5.2)$$

依規格化定義我們欲得到的結果影像 x' 將會與參考影像 y 擁有一樣累積機率分布，我們假定 $h_x = h_y$ 在此條件下，因此可以由公式 5.2 的原始的累積機率分布 h_x 代入參考累積分布函數的逆轉換函數，重新定義規格化函數為：

$$y = g^{-1}(h_y) = g^{-1}(h_x) = g^{-1}(f(x)) = h_y(x) = x' \quad (5.3)$$

從公式 5.2 我們可以得知在連續性的分布下的規格化定義為：

$$g^{-1}(f(x)) = h_y(x) \quad (5.4)$$

然而實際的顏色分布是以離散的直方圖形式，因此在離散分布中，規格化的過程即是最小化兩者於累積機率分布 $g(y)$ 及 $f(x)$ 上的差距：

$$\left| F_x(l) - G_y(j) \right| = \text{MIN}_k \left(\left| F_x(l) - G_y(k) \right| \right) \quad (5.5)$$

從公式 5.5 可知，在離散的顏色直方圖規格化中，重新將每一個原始影像建立起一個離散函數對應，當滿足 $F_x = G_y$ 條件時， $h_y(x) = g^{-1}(f(x)) = g^{-1}(g(y)) = y$ 即為所求。

雖然直方圖規格化在尋找對應色有其優勢，但使用此方法缺點是僅在顏色分布上取得最接近的對應，較欠缺其他空間資訊上的運用，為了能達到最佳的抽象化效果，仍建議與其它降色演算法來相互配合。利用規格化的方法我們可以將使用者輸入的真實照片，在顏色對應上轉換成所喜好的卡通樣版 (見圖 5.11)。考慮到單純使用變形法所產生的卡通肖象，對臉部特徵的保留有所不足，如：膚色便是相當具有代表性的特色，除

了不同人種有別之外，膚色也會根據所受到年紀、性別、外在環境燈光等因素的影響，因此我們除了產生在直方圖規格化後的背景樣版時，在臉部區塊和部份還可以運用混色技巧，將轉描法與變形法妥善地合成最後的卡通化肖像，不但能保有較佳的膚色特徵，同時在臉部區域貼上卡通化材質後也能與背景影像有比較高的契合度（見圖 5.12）。

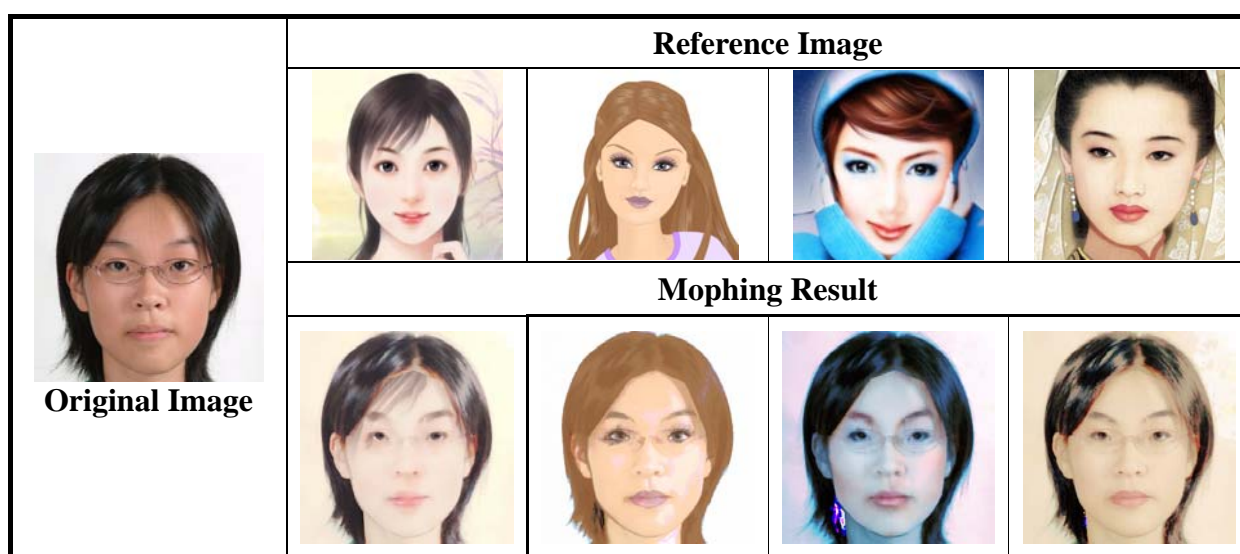


圖 5.11、直方圖規格化後變形法卡通結果

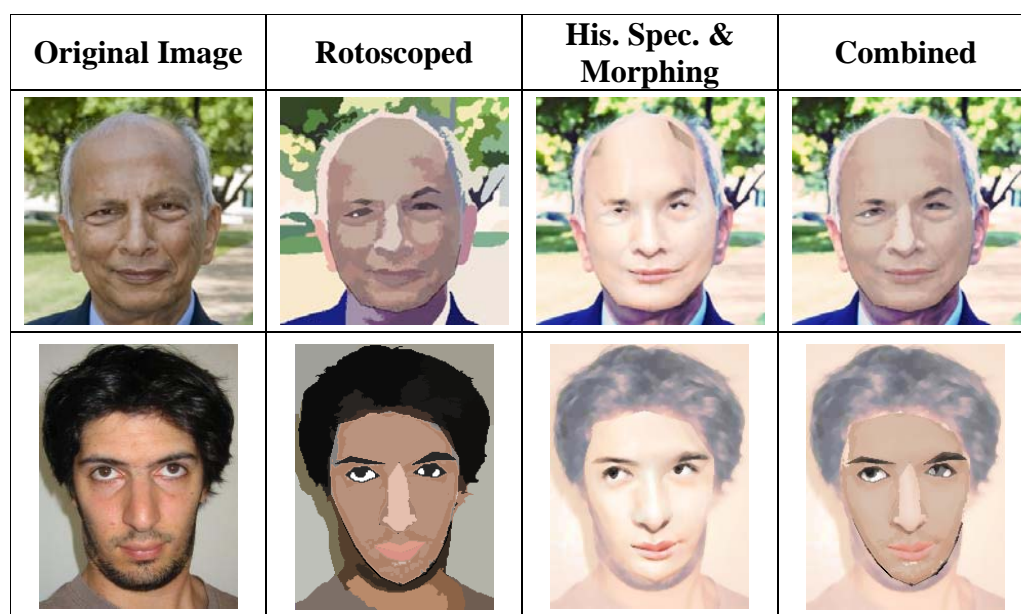


圖 5.12、運用混色技巧整合轉描及變形法，保留膚色及臉部特徵