

以眼口特徵為基礎來評量人像照片優劣之研究

Evaluate the quality of a portrait by analyzing the eyes and the mouth

徐君慧

Chun-Hui Hsu

臺北大學通訊工程研究所

gina11172002@hotmail.com

呂嘉穀

Jia-Guu Leu

臺北大學資訊工程系

leu@mail.ntpu.edu.tw

摘要

一般的生活照是以正面的人像照為主，其臉部特徵資訊幾乎決定了一張人像照片的好壞。有鑑於此，本研究是以人像照片中的臉部特徵資訊為導向，再藉由評分機制來進行評分。根據分數的高低便能判斷該張人像照片的好壞，進而從龐大數量的照片中，篩選出最佳的人像照片。

首先，我們利用顏色資訊進行人臉偵測，並找出雙眼及嘴巴等重要的特徵點位置。由於我們對於一張好的人像照片的認知，人的眼睛必須張大且正視，嘴巴也必須帶有笑意，因此我們定義了三個適用於電腦量化分析的特徵指標。這些特徵指標分別是整個眼睛的高寬比、眼框中心與黑眼珠中心的偏移量，以及嘴巴的彎曲度，並將其數值化為介於 0 和 1 之間的分數。利用量化後的分數設計兩套結合眼睛和嘴巴的綜合評分機制。

由實驗結果證實，將系統的兩種評分結果分別與人工的平均評分比較，皆得到了相關係數為 0.86 的相關性。表示我們系統的兩種評分方式高度符合人類使用者的主觀判斷。

關鍵字：人像照片、人臉偵測、臉部特徵

1. 簡介

我們撇開一些攝影技術層面所探討的條件，真正判斷一張人像照片的優劣，其實是根據欣賞者的角度做出主觀的評論，而欣賞者的注意力是放在照片人物的臉部表情上。一般來說，以單獨人像照來看，能稱為好的照片最直觀的就是能夠清楚地將人物主體表現出來，眼睛睜開正視前方，嘴巴帶有笑意而揚起適度的微笑曲線，這

樣的照片充滿了魅力的眼神和迷人的微笑，會讓人第一眼就印象深刻且相信是美麗的人像照片。



(a)好的人像照片

(b)不好的人像照片

圖 1-1：人像照片的好壞

本研究希望能評量一張人像照片的好壞，從人臉偵測、人臉特徵點擷取到最後照片分析，每一個環節都相當重要且環環相扣，而人臉偵測約略分成以下四大類[5、15]：

- (1) 以知識背景為基礎法 (Knowledge-based methods)：將人臉五官分佈的關係定義成規則，再藉由這些規則去尋找影像中人臉的位置。[14]
- (2) 以特徵為基礎法 (Feature-based methods)：要找出影像中的眼睛、嘴巴等特徵，或是膚色、輪廓、大小等人臉的結構特徵來偵測臉部。[4、5、6]
- (3) 以樣板為基礎法 (Template-based methods)：把輸入的影像與事先建立好的臉部樣板作比較，尋找出符合臉部樣板的區域即為人臉。[5]
- (4) 以影像為基礎法 (Image-based methods)：藉由統計分析及機器學習的方法大量訓練樣本，能針對人臉及非人臉進行正確分類。 [8、13]

關於臉部特徵點擷取，由先前得到的人臉區域影像中擷取獨特的特徵資訊。最明顯且重要的莫過於眼睛和嘴巴，而比較常見的眼睛和嘴巴之擷取方法為色彩資訊 [4]及可變動樣板[10、11]。

Leyvand 等人在[7]將人臉分成 84 個特徵點及位於這些特徵點間 234 組向量的集合，利用支持向量回歸 (Support Vector Regression, SVR) 技術的機器學習方式，讓電腦來辨別人臉的美醜。

2. 人臉偵測與處理



圖 2-1：人臉偵測流程圖

2.1 膚色偵測

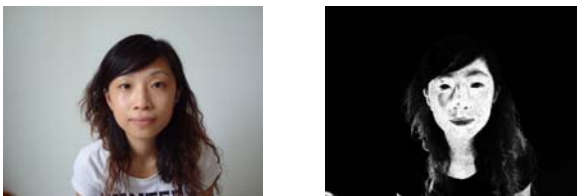
我們先將原影像 RGB 轉換至 YCbCr 色彩空間，以此做為膚色偵測的色彩系統[9]。這個色彩空間的好處就是亮度 Y(luminance)和兩彩度 Cb(blueness)、Cr (redness)之間的關聯性很低。接下來我們任意挑選 8 張不重複的人像照片，分別取這些照片中屬於膚色區域的額頭、鼻子、左臉頰、右臉頰、下巴等五個點的 Cb 和 Cr 值當成膚色的基本樣本(n=40)，紀錄數值並依照公式(1)(2)(3)分別求出 Cb 和 Cr 的整體平均值所組成的平均向量 μ 以及共變異數矩陣 Σ ，再來代入公式(4)形成以機率為特性的混合高斯膚色模型來建立對膚色偵測的機率函數[9][16]，其中的 X 為影像像素的(Cb,Cr)。膚色高斯模型公式如下：

$$\overline{Cb} = \frac{\sum_{i=1}^n Cb_i}{n}, \quad \overline{Cr} = \frac{\sum_{i=1}^n Cr_i}{n} \quad (1)$$

$$\mu = \begin{bmatrix} \overline{Cb} \\ \overline{Cr} \end{bmatrix}, \quad X_i = \begin{bmatrix} Cb_i \\ Cr_i \end{bmatrix} \quad (2)$$

$$\Sigma = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)(X_i - \mu)^T \quad (3)$$

$$P(X) = \frac{1}{(2\pi)^{|\Sigma|^{1/2}}} \exp\left[-\frac{1}{2}(X_i - \mu)^T \Sigma^{-1}(X_i - \mu)\right] \quad (4)$$



(a)原始影像 (b)使用後膚色區域為亮區

圖 2-2：影像代入混合高斯膚色模型

以高斯混合模型和適當的門檻值(threshold)擷取出膚色區塊後，得到一個二值化的影像，我們發現此影像並非十分平滑且會有些許雜訊，所以再對此二值化影像做斷開(open)運算。

2.2 膚色最大區塊 Face Mask

為了避免有些照片經斷開(open)運算處理，剩餘的影像仍無法形成單一完整區塊，這時我們必須將經由形態學運算後的二值化影像再運用連通區域標記(connected component labeling)之技術分割出影像二值化為 1 的最大區塊當做最有可能的 Face Mask 區域。由於位於膚色區域內的眼睛、嘴巴為黑，造成整張人臉呈現空洞狀態，在此我們用影像邏輯運算中的 NOT 將剛剛顯示的二值影像結果反白，針對此影像再做一次連通區域標記，標記區域中值為 1 的最大區塊設為 0，如此便能把剛剛影像中屬於黑色部份的眼睛、嘴巴取出，把眼睛和嘴巴造成的空洞給補起來，獲得較完整的人臉區域 Face Mask。

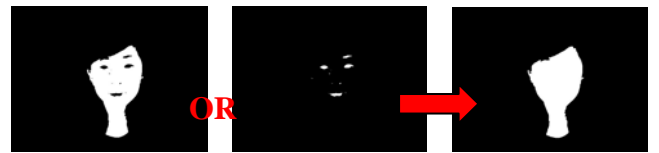


圖 2-3：取出膚色最大區域 Face Mask 過程



圖 2-4：將 Face Mask 套入原彩色圖結果

2.3 重新定義 Face Mask 區域

膚色偵測的結果並非使得最大的膚色區塊為純粹的人臉所在，因此我們將縮小影像的搜尋範圍，只針對感興趣的影像資訊進行處理。在此我們任意挑選 20 張不同人臉的高寬比並統計出人臉的高寬比近似 1.3：1，從膚色偵測所得到最大膚色區塊 Face Mask 找出人臉的寬度，再根據此高寬比的比例來找尋人臉的高度，長方形的邊框框出人臉，此為重新定義的 Face Mask 區域。



圖 2-5：原搜尋框與重新定義 Face Mask 區域範圍

3. 眼睛與嘴巴特徵點擷取與分析

在本章節中，我們將詳細說明如何透過眼睛和嘴巴特徵點發展評分機制，並藉由評分機制產生的分數，進行人像照片優劣的探討。

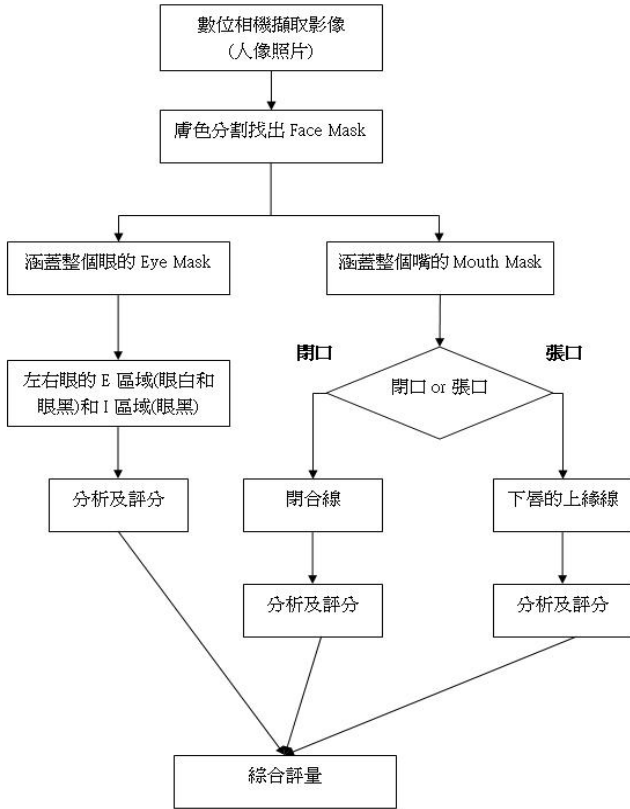


圖 3-1：系統流程圖

3.1 眼睛偵測與分析

首先建立兩種不同的 Eye Maps，第一種是採用彩度的 EyeMapC 來加強 Cb 值、減弱 Cr 值。第二種是採用亮度的 EyeMapL，在 Y 影像中，我們藉由形態學上的運算子膨脹(dilation)和侵蝕(erosion)來強調臉上的明亮元素與黑暗元素。其中我們自訂一個大小類似眼睛形狀的結構元素，而 \oplus 及 \ominus 分別代表對灰階影像做膨脹與侵蝕。最終由這兩種組合而成 Eye Map[4]，經由影像運算建立 Eye Mask 找到可能為眼睛的範圍。

$$\text{EyeMapC} = \frac{1}{3} \{ (C_b^2) + (1 - C_r)^2 + (C_b / C_r) \} \quad (5)$$

$$\text{EyeMapL} = \frac{Y(x, y) \oplus g_\sigma(x, y)}{Y(x, y) \ominus g_\sigma(x, y) + 0.01} \quad (6)$$

$$\text{Eye Map} = (\text{EyeMapC}) \times (\text{EyeMapL}) \quad (7)$$

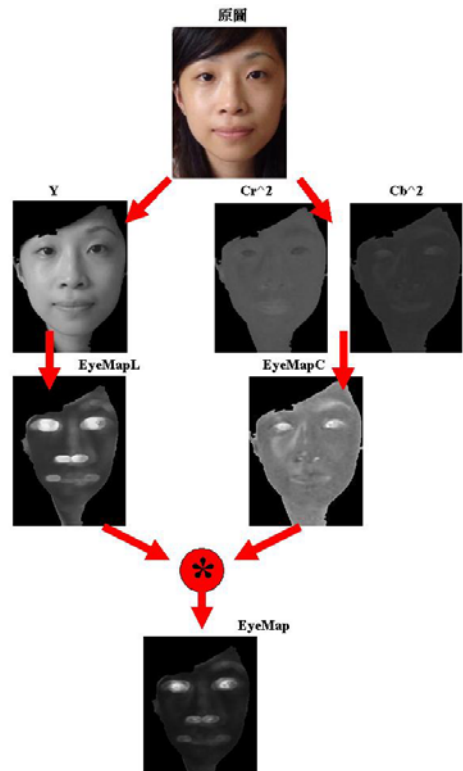


圖 3-2：Eye Map 每階段的變化

為了避免不穩定的結果發生，我們自訂一個眼睛區域搜尋框；由於 Eye Mask 必位於人臉的上半部，且正常眼部的垂直位置介於額頭下方和鼻頭上方，眼部的水平位置比臉寬窄一些。我們針對人臉方框的上下左右邊界進行眼睛區域搜尋框的參數化。

$$\begin{aligned} X_0 &= \text{Left} + \text{Width}/15 \\ X_1 &= \text{Right} - \text{Width}/15 \\ Y_0 &= \text{Top} + \text{Height} \times 9/17 - \text{Height}/3 \\ Y_1 &= \text{Top} + \text{Height} \times 9/17 \end{aligned} \quad (8)$$

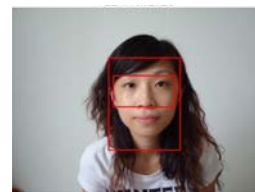


圖 3-3：眼睛的搜尋框範圍

首先我們分別記錄搜尋框內二值化為 1 的區塊像素個數 N 、像素位置 (X_i, Y_i) ，依(9)式計算出各區塊中心點 (X_g, Y_g) 。

$$X_g = \frac{1}{N} \sum_{i=1}^n X_i, \quad Y_g = \frac{1}{N} \sum_{i=1}^n Y_i \quad (9)$$

再把眼睛可能區塊的中心點兩兩一組做判斷：
1. 眼睛搜尋框內沒有或是只有一個中心點，將判定此影像找不到眼睛，此為閉眼的可疑影像。

- 眼睛搜尋框內剛好有兩個中心點 (X_{g1}, Y_{g1}) 和 (X_{g2}, Y_{g2}) ， Y_{g1} 與 Y_{g2} 此兩重心的垂直方向差必須小於或等於眼睛搜尋框高度的 $1/8$ 個像素值， X_{g1} 與 X_{g2} 此兩重心的水平方向差必須大於臉寬的 $3/11$ 個像素值，符合的判定影像為眼睛。
- 眼睛搜尋框內有超過兩個以上的中心點，把這些中心點 Y_g 座標點兩兩相減並記錄其值，去掉兩者間相距太大，保留兩者間相距最小的點，再判斷是否符合第 2 項規則，符合的為眼睛。

接著分別對左右眼進行膨脹運算，如此便可放大左右眼區塊以至於能夠完全包覆整個眼睛的 Eye Mask，以利後續動作的處理。

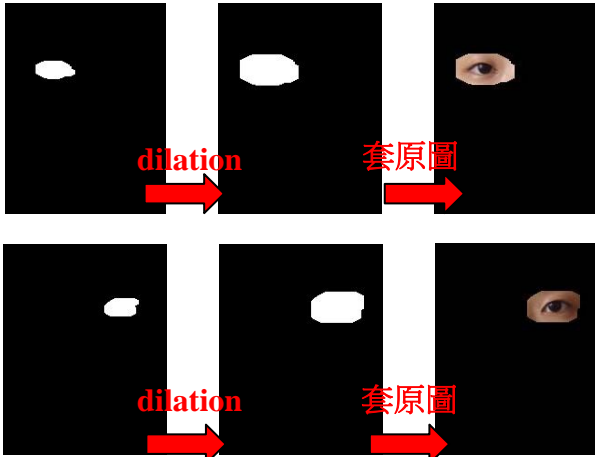


圖 3-4：找出左眼和右眼 Eye Mask 的過程

3.2 眼睛 E 區域及 I 區域特徵點擷取

E 區域(Eye)是包含眼睛的眼白和眼黑，在這邊最主要的是要找出 E 區域的上下左右 4 個特徵點位置以及整個眼框的中心位置，左右兩隻眼睛的 E 區域共有 10 個特徵點。I 區域(Iris)是指較暗的黑眼珠區域，我們要找出 I 區域的中心位置並標示出來，左右兩隻眼睛的 I 區域共有 2 個特徵點。

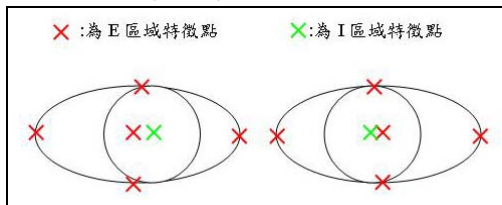


圖 3-5：E 區域及 I 區域特徵點示意圖

取出臉部區域的 Y、Cb、Cr 影像如圖 3-6，觀察到任意一個人在 Y 影像中眼睛區域有很明顯的黑白分明，在 Cb 影像中其眼睛區域的色調值相對於皮膚顏色來說高出許多，而在 Cr 影像

中其眼睛區域的色調值卻是臉部區域最低的。



圖 3-6：脸部區域的 Y、Cb、Cr 影像

1. 找 I 區域並標示其特徵點

根據 Y、Cb、Cr 影像顯示的特性進行影像數學運算，利用 Cb 影像除以 Cr 影像可區別眼睛和膚色部份，此時眼睛區域呈現亮白色，我們必須減弱眼白部分的亮度才能順利取出黑眼珠的部份，取出黑眼珠後的影像再執行線性轉換，將影像元素尺度調整成可顯示的數值，最小值映射成 0，最大值映射成 1。公式中的分母額外再加上一個常數 0.1 的目的是預防分母部份為 0，使得結果達到整個眼睛部份特別明亮，公式如下：

$$C(x, y) = \frac{C_b(x, y)}{C_r(x, y) + 0.1} \quad (10)$$

$$I(x, y) = C(x, y) - Y(x, y) \quad (11)$$



(a)公式 10 結果 (b)公式 11 結果(c)將(b)套入左眼 EyeMask

圖 3-7：I 區域每階段的變化

運算完的影像再套入左眼 Eye Mask，接下來做二值化、4 連通區域標記、形態學運算及計算左眼 I 區域內二值化為 1 的中心點 $I_g(x, y)$ ，同樣利用以上方法亦可找出右眼的 I 區域及特徵點。

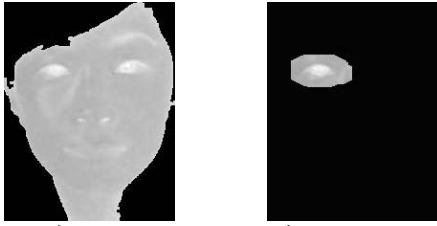


圖 3-8：左右眼 I 區域中心位置

2. 找 E 區域並標示特徵點

我們發現公式(10)的影像結果中，黑眼珠部份呈現較暗的顏色，因此我們必須搭配公式(11)的影像結果加強黑眼珠的亮度，分別給予適當比例能夠使眼黑與眼白的亮度呈現一致，明顯區別整個眼睛顏色和周圍膚色的差異。公式如下：

$$E(x, y) = (9 \times C(x, y) + 1 \times I(x, y)) / 10 \quad (12)$$



(a)公式 12 的結果 (b)套入左眼 Eye Mask

圖 3-9：E 區域每階段的變化

我們以左眼的 E 區域找法來說明，套入左眼 Eye Mask 並定義一個小的結構元素是符合眼睛細長型的構造，以此對灰階影像做膨脹的動作，影像會輕微模糊並稍微放大 E 區域的特徵，如此便能強化眼睛內眼角不明顯的地方。再做二值化及 4 連通的區域標記找出所有連接區域並紀錄下來，只保留最大區域即為 E 區域(Eye)所在位置。接著找整個眼框的中心點 $E_g(x, y)$ ；而 E 區域左右兩邊的特徵點位置就是把最邊界二值化為 1 的點取出， $E_L(x, y)$ 和 $E_R(x, y)$ 當作 E 區域左右特徵點的座標；再來我們是利用 E 區域內每個水平方向往上、往下延伸碰到二值化為 1 的極端點，找出兩個上下的一對座標值所在的垂直方向擁有差異最大的那組， $E_T(x, y)$ 和 $E_B(x, y)$ 當作 E 區域上下特徵點的座標，同樣利用以上的方法亦可找出右眼的 E 區域及特徵點。



圖 3-10：左右眼 E 區域中心位置及上下左右特徵點

3.3 眼睛特徵點距離關係與評分

1. 高寬比

由人像照片中找到眼睛 E 區域的上下左右 4 個特徵點的座標，計算出上特徵點 $E_T(x, y)$ 和下特徵點 $E_B(x, y)$ 間的距離 d_1 ，以及左特徵點 $E_L(x, y)$ 和右特徵點 $E_R(x, y)$ 間的距離 d_2 ，而高寬比值 α 是將 d_1 值除以 d_2 值，藉此數值可以反映出一個人眼睛張開的大小，計算方式如下：

$$\begin{aligned} d_1 &= \sqrt{(x_{E_T} - x_{E_B})^2 + (y_{E_T} - y_{E_B})^2} \\ d_2 &= \sqrt{(x_{E_L} - x_{E_R})^2 + (y_{E_L} - y_{E_R})^2} \\ \alpha &= d_1 / d_2 \end{aligned} \quad (13)$$

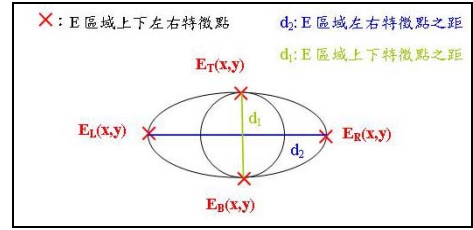


圖 3-11：E 區域特徵點間距離的示意圖

2. 偏移量

紀錄眼睛 E 區域的中心位置坐標 $E_g(x, y)$ 和眼睛 I 區域的中心位置座標 $I_g(x, y)$ ，計算出此兩特徵點的距離 d_3 ，而偏移量值 β 是將 d_3 值除以整個 E 區域寬度 d_2 值，也就是眼睛 E 區域和 I 區域它們的中心點距離相對於整個 E 區域的寬度值，藉此數值可以觀察照片中人物的眼睛是否正視或斜視，計算方式如下：

$$\begin{aligned} d_3 &= \sqrt{(x_{E_g} - x_{I_g})^2 + (y_{E_g} - y_{I_g})^2} \\ \beta &= d_3 / d_2 \end{aligned} \quad (14)$$

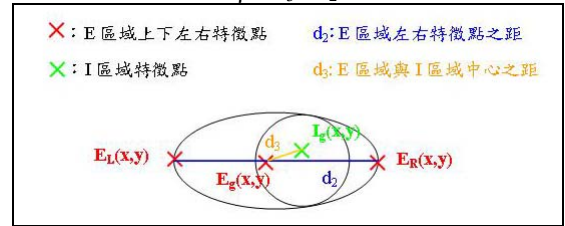


圖 3-12：E 區域與 I 區域中心點間距離的示意圖

以高寬比來說，我們需要所有受測影像的正常狀態下眼睛特徵點的高寬比值取平均，此平均值 α_{normal} 為 0.3，也就是說將任意一張影像中眼睛特徵點的高寬比值相對於正常眼狀態下的高寬比值 0.3，若是拍攝不好眯眼的照片，其比較結果的值必定小於 1，所得到的結果值為指標分數，若比較結果的值等於或大於 1 時，則設定其指標分數為 1，讓指標分數值都分佈在 0 至 1 之間；為了讓指標分數間差異大，我們將上述求得的指標分數平方作為最終的指標分數 $Score_\alpha$ ，此分數可看出一個人眼睛張開的大小。 n 為受測影像中屬於正常眼睛的數量， α_i 為某張影像 E 區域高寬比：

$$\alpha_{normal} = \frac{1}{n} \cdot \sum_{i=1}^n \alpha_i = 0.3 \quad (15)$$

$$Score_{size} = \begin{cases} 1 & \text{if } \alpha_i / \alpha_{normal} > 1 \\ (\alpha_i / \alpha_{normal})^2 & \text{otherwise} \end{cases}$$

以偏移量來說，我們需要所有受測影像的正常狀態下眼睛特徵點的偏移量距離，此平均值

β_{normal} 為 0.05，左眼正常狀況下，眼睛 I 區域的中心位置在眼睛 E 區域的中心位置的右側，屬於正偏移量且兩特徵點相距 0.05；而右眼正常狀況下，眼睛 I 區域的中心位置在眼睛 E 區域的中心位置的左側，屬於正偏移量且兩特徵點相距同樣是 0.05，根據此分析我們得知左右眼的正常偏移量的方向是相反的，因此我們必須先判斷眼睛 I 區域及眼睛 E 區域的中心位置是屬於正偏移或負偏移，再來計算與標準狀態下的偏移量平均值 β_{normal} 相差多少，計算後的偏移量取絕對值再乘以 4 倍，最後用 1 去扣除該結果，結果的值必定大於 0，所得到的結果值為指標分數，若比較結果的值等於或小於 0 時，則設定其指標分數為 0，讓指標分數值都分佈在 0 至 1 之間，此分數可看出一個人眼睛是否正視前方。 n 為受測影像中屬於正常眼睛的數量， β_i 為某張影像的偏移量。

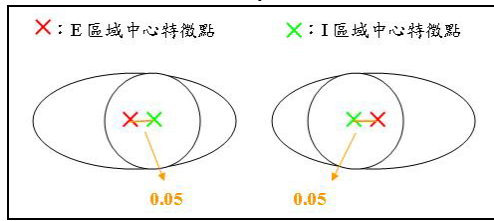


圖 3-13：正常狀況左右兩眼的偏移量

$$\beta_{normal} = \frac{1}{n} \cdot \sum_{i=1}^n \beta_i = 0.05 \quad (16)$$

$$\beta_2 = \begin{cases} \beta - \beta_{normal} & \text{if 正偏移} \\ -(\beta + \beta_{normal}) & \text{if 負偏移} \end{cases}$$

$$Score_{front} = \begin{cases} 0 & \text{if } 1 - |\beta_2 \times 4| < 0 \\ 1 - |\beta_2 \times 4| & \text{otherwise} \end{cases}$$

將上述兩項指標的分數相乘積為單一左眼 Eye Score_L 分數，單一右眼 Eye Score_R 分數同理可得，雙眼的綜合指標分數為左眼綜合指標分數和右眼綜合指標分數相加總取平均，根據雙眼的綜合指標分數，可以決定人像照片中眼睛所拍攝的好壞與否。

Eye Score_L = Score_{size} * Score_{front}, Eye Score_R 同理可得 (17)

$$Eye\ Score = \frac{1}{2} \cdot (Eye\ Score_L + Eye\ Score_R)$$

3.4 嘴巴偵測與分析

嘴巴部分的色彩有著很強烈的紅色元素及很少的藍色元素，運用此特性我們可以強調 Cr 影像降低 Cb 影像的方式建立 Mouth Map 將嘴巴

區域強化出來[4]。公式如下，其中 FG 為擷取的臉部區域位置， n 為該臉部位置的像素個數：

$$Mouth\ Map = C_r^2 \cdot \left(C_r^2 - \eta \cdot \frac{C_r}{C_b} \right)^2 \quad (18)$$

$$\text{其中 } \eta = 0.95 \cdot \frac{\frac{1}{n} \sum_{(x,y) \in FG} C_r(x,y)^2}{\frac{1}{n} \sum_{(x,y) \in FG} C_r(x,y) / C_b(x,y)}}$$

接著我們自訂一個符合嘴型的結構元素做膨脹運算得以放大嘴巴區域，再做二值化及型態學處理，最後結果為涵蓋整個嘴的 Mouth Mask。

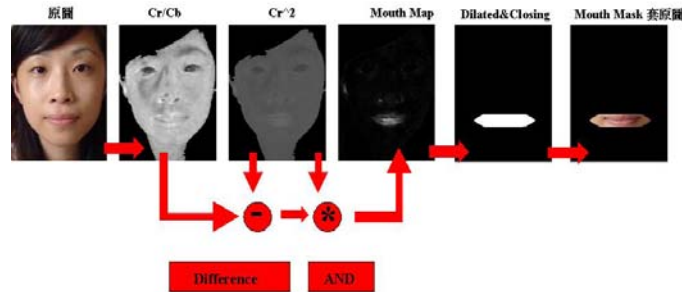


圖 3-14：Mouth Map 各階段與 Mouth Mask 結果

由於 Mouth Mask 必位於人臉的下半部，且嘴巴的垂直位置在鼻頭下方，水平位置比臉寬窄一些。嘴巴區域搜尋框的參數化如下：

$$\begin{aligned} X_2 &= \text{Left} + \text{Width}/5 \\ X_3 &= \text{Right} - \text{Width}/5 \\ Y_2 &= \text{Top} + \text{Height} \times 16/17 - \text{Height} \times 2/5 \\ Y_3 &= \text{Top} + \text{Height} \times 16/17 \end{aligned} \quad (19)$$

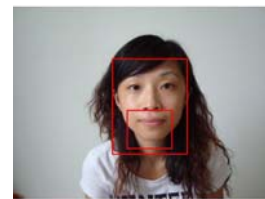


圖 3-15：嘴巴的搜尋框範圍

3.5 判斷開閉並找其嘴巴曲線

當我們開懷大笑到張開嘴巴露出牙齒時，下嘴唇上緣曲線的上揚弧度較明顯，是判別開口時笑容判別的重要依據；當我們嘴巴閉合時，此時嘴巴內的區域則是由上嘴唇下緣曲線與下嘴唇上緣曲線合成一條閉合曲線，此條閉合曲線呈現上揚或是下凹的弧度，是判別開口時笑容判別的重要依據；因此我們必須先判斷嘴巴的開閉狀態進而取得曲線作後續處理。

當嘴巴區域被準確定位並擷取包含上下兩片嘴唇影像後，接下來我們利用 Y 影像及 Cr 影像分別在閉口和開口時的差別，根據公式(20)能將嘴巴開閉的差異明顯的強調出來。

$$\text{Open}(x, y) = \frac{Y(x, y)}{C_r(x, y) + 0.01} \quad (20)$$

1. 擷取閉口時雙唇之閉合線

Y 影像在閉口時呈現上下嘴唇亮度值較亮，中間的閉合線偏暗的特性，設計一個濾波器專門測嘴巴閉合曲線的運算子，以 Y 影像去做運算且運算完後捨棄掉負值再執行線性轉換，將影像元素尺度調整成最小為 0 最大值為 1 的灰階影像，使得影像的垂直方向對比明顯；雖然閉合線並非完全水平，但是它的曲度相當小，故 9x9 遮罩即可突顯閉合曲線在垂直方向上的亮度變化。

3	3	3	3	3	3	3	3	3
2	2	2	2	2	2	2	2	2
1	1	1	1	1	1	1	1	1
-6	-6	-6	-6	-6	-6	-6	-6	-6
-6	-6	-6	-6	-6	-6	-6	-6	-6
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2

圖 3-16：擷取閉合曲線的濾波遮罩

嘴巴閉合曲線會因為嘴巴周圍肌肉的動作，而在兩端嘴角附近產生陰影，將上述步驟所產生影像再做灰階的擴張，讓影像中亮的部份更亮。影像二值化後再使用閉合運算，其目的是獲取更完整的嘴巴閉合曲線。最後作細線化運算。



(a)原圖嘴巴為閉口(b) close 後之影像 (c)閉合線之影像
圖 3-17：擷取閉合曲線的過程

2. 擷取張口時下唇之上緣

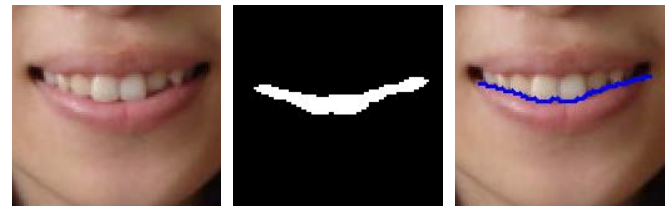
我們想要取出下嘴唇上緣曲線可利用張口有露齒狀況，下嘴唇屬於紅色和牙齒屬於白色的顏色做比對；根據 Cr 影像執行線性轉換後，在閉口時上下嘴唇以及中間閉合曲線的亮度值皆較亮，而開口時因為嘴巴內部有牙齒，顯得中間區域是呈現偏向暗黑色的特性，設計一個濾波器

專門測嘴巴張開時下嘴唇上緣曲線的運算子，以 Cr 影像去做運算且運算完後捨棄掉負值再執行線性轉換，如此便可將強調下嘴唇的上緣曲線。

-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1
0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1

圖 3-18：擷取下嘴唇上緣曲線的濾波遮罩

下嘴唇的上緣曲線會因為張口大笑時的上揚弧度過大，導致在曲線兩端嘴角附近產生不明顯的亮點，我們將上述步驟所產生影像再做一個較窄的灰階的擴張，讓影像中亮的部份更亮。最後作細線化運算。



(a) 原圖嘴巴為張口(b) 下嘴唇上緣曲線(c) 上緣線之影像
圖 3-19：擷取下嘴唇上緣曲線的過程

3.6 嘴巴曲線特徵點擷取

針對此條曲線要找出其兩個極端點 $L_L(x,y)$ 和 $L_R(x,y)$ 、連線的中點位置 $M(x,y)$ 及此線段的中垂線交於嘴巴曲線的 $L_N(x,y)$ 特徵點座標。其中可藉由嘴角左右 2 個特徵點之距離及連線中點的特徵點至嘴巴曲線上特徵點之距離所產生的比例，反映出嘴型變化的彎曲程度。

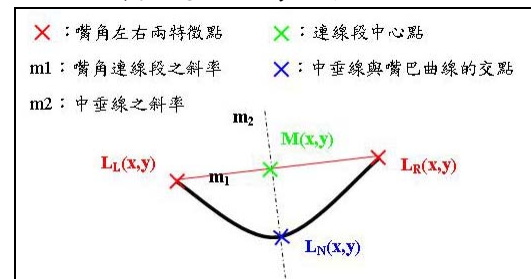


圖 3-20：嘴巴曲線特徵點示意圖

3.7 嘴巴特徵點距離關係與評分

計算各特徵點間的相對距離關係並組成嘴巴的彎曲度特徵值，利用此特徵值來當做判別人

物照片中嘴巴的好壞指標依據。當特徵點 $M(x,y)$ 在 $L_N(x,y)$ 之上方，曲線為往下凹成 V 字形，用正號代表；當特徵點 $M(x,y)$ 在 $L_N(x,y)$ 之下方，曲線為往上凸成 Λ 字形，用負號代表，如此一來可以用正負號辨識嘴巴的輪廓為向上或向下彎曲。再來使用歐氏距離計算出 $M(x,y)$ 和 $L_N(x,y)$ 兩特徵點的距離 d_4 ，以及曲線兩端點 $L_L(x,y)$ 和 $L_R(x,y)$ 的距離 d_5 ，其中 d_4 值包含 $M(x,y)$ 和 $L_N(x,y)$ 間的距離長度值及曲度上下的正負號值，而彎曲度值 γ 是將 d_4 值除以代表嘴巴寬度的 d_5 值，以嘴巴曲線的高度與嘴的寬度之相對比例來當作嘴巴分類的彎曲度特徵向量值，計算方式如下：

$$d_4 = \begin{cases} \sqrt{(x_M - x_{L_N})^2 + (y_M - y_{L_N})^2} & \text{if } y_M < y_{L_N} \\ -\sqrt{(x_M - x_{L_N})^2 + (y_M - y_{L_N})^2} & \text{if } y_M > y_{L_N} \end{cases} \quad (21)$$

$$d_5 = \sqrt{(x_{L_L} - x_{L_R})^2 + (y_{L_L} - y_{L_R})^2}$$

$$\gamma = d_4 / d_5$$

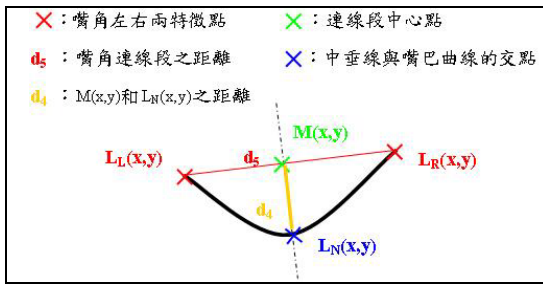


圖 3-21：嘴巴曲線特徵點與距離示意圖

我們從所有受測影像的各種變化發現嘴巴曲線呈現最好的狀態下，嘴巴特徵點的 γ 值取平均是介於 0.03 和 0.09 之間，則設定其指標分數為 1。若 γ 值超過 0.09 的照片，代表照片中人物的嘴巴曲度過大，在視覺感官上並非最完美，指標分數應比最好狀態下的分數略低一點； γ 值介於 0 和 0.03 之間的照片，代表照片中人物的嘴巴曲度小於最好的狀態，指標分數應比最好狀態下的分數低；而 γ 值低於 0 的照片，代表照片中人物的嘴巴曲度的方向向下，具備不好的照片所擁有的特性，指標分數應比最好狀態下的分數低得多。針對以上三種狀況，我們將該張照片的 γ 值與最好的狀態下的 γ 值相減，相減後的值取絕對值再分別乘以 3 倍、3 倍和 9 倍，最後用 1 去扣除該結果，所得到的結果值為指標分數。 γ_i 為某張影像嘴巴曲線彎曲度：

$$\text{Mouth Score} = \begin{cases} 1 & \text{if } 0.03 \leq \gamma_i \leq 0.09 \\ 1 - |\gamma_i - 0.09| \times 3 & \gamma_i > 0.09 \\ 1 - |\gamma_i - 0.03| \times 3 & 0 < \gamma_i < 0.03 \\ 1 - |\gamma_i - 0.03| \times 6 & \gamma_i \leq 0 \end{cases} \quad (22)$$

3.8 結合眼睛和嘴巴的評分機制

最後，透過雙眼綜合分數以及嘴巴綜合分數來完成兩套綜合的評分機制；第一套評分機制是採用雙眼綜合分數與嘴巴分數相乘所得的綜合分數，當判斷眼睛和嘴巴好壞的其中一項條件不好時，即給予否定，屬於較強烈的評分方式。

$$\text{Total Score}_1 = \text{Eye Score} \times \text{Mouth Score} \quad (23)$$

第二套是將雙眼綜合分數與嘴巴分數分別乘上權重 1/2 再相加所得的綜合分數，若測試影像在系統的單項評分分數太低，最後的總分數也只會扣一些分數，屬於較緩和的評分方式，式子如下：

$$\text{Total Score}_2 = \frac{1}{2} \times \text{Eye Score} + \frac{1}{2} \times \text{Mouth Score} \quad (24)$$

4. 實驗結果與討論

4.1. 實驗設計與結果

(一)系統中兩種評分方式

第一套是將眼睛分數與嘴巴分數相乘所得的綜合分數，第二套是將眼睛分數與嘴巴分數分別乘上權重 1/2 再相加所得的綜合分數，以上兩套機制算出的綜合分數皆介於 0 至 1 之間，可作為判斷照片好壞的綜合評分分數。由於影響眼睛分數的變因有兩個，包括眼睛的高寬比和眼睛的偏移量；影響嘴巴分數的變因有一個，也就是這三項因素都會對實驗結果造成影響，因此最後我們從 6 位受測者在眼睛和嘴巴的變化影像中挑選了三項因素都呈現良好狀態，以及三項因素中的其中一項因素呈現不好的狀態，還有三項因素都呈現不好的狀態共分成五種狀況的全臉影像來做實驗。

(二)人工評分方式

我們設計了關於主觀認知照片好壞程度的問卷，並將問卷給予 9 位評分者來填寫。我們要求評分者根據他們對於照片中人物的眼睛和嘴巴特徵點，針對人像照片的好壞程度加以評分，一張照片只可以有一個分數，評分的方式以範圍 1 到 5 分成五個等級，由低至高排列為：1(非常

不好)、2(不好)、3(普通)、4(好)、5(非常好)。接著算出每張照片的平均分數至小數點以下的兩位，最後所得分數就是這 30 張影像的人工評分的平均分數。

(三)系統的兩種評分方式與人工評分間的相關係數及 X-Y 散佈圖

在這次實驗中我們偵測 6 個人，每一個人有 5 張眼睛和嘴巴不同狀況的測試照片，總共 30 張影像。下面的表格 4-8~4-12 是實驗結果的彙整表，列出了各種評分方式對人像照片所進行的評分結果，而人工評分那一欄是取所有評分者給予分數的平均。表格 4-13 為各評分方法間的相關係數及 X-Y 散佈圖，若圖中的點群聚的部份是呈現東北西南走向，則彼此的相關性越高。

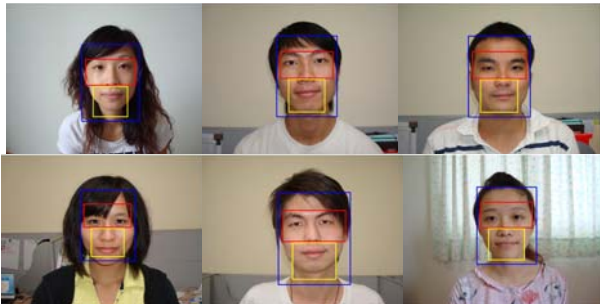


圖 4-1：重新定位 Face Mask 下的眼睛、嘴巴搜尋框

表 4-1：正常眼狀態下各特徵點定位與眼睛分數

照片 \ 分數		左眼			右眼			雙眼總分
		高寬	偏移	總分	高寬	偏移	總分	
	原始數據	0.41	0.11		0.35	0.06		
	分數	1	0.78	0.78	1	0.95	0.95	0.86
	原始數據	0.29	0.07		0.32	0.05		
	分數	0.93	0.92	0.85	1	1	1	0.93
	原始數據	0.21	0.06		0.26	0.09		
	分數	0.49	0.95	0.47	0.74	0.82	0.60	0.54
	原始數據	0.30	0		0.29	0.03		
	分數	1	0.80	0.80	0.95	0.92	0.88	0.84
	原始數據	0.25	0.01		0.24	0.02		
	分數	0.72	0.82	0.59	0.62	0.89	0.55	0.57
	原始數據	0.39	-0.02		0.41	0.07		
	分數	1	0.72	0.72	1	0.91	0.91	0.82

表 4-2：斜視眼狀態下各特徵點定位與眼睛分數

照片 \ 分數		左眼			右眼			雙眼總分
		高寬	偏移	總分	高寬	偏移	總分	
	原始數據	0.45	-0.13		0.31	0.21		
	分數	1	0.30	0.30	1	0.35	0.35	0.32
	原始數據	0.38	-0.10		0.36	0.16		
	分數	1	0.41	0.41	1	0.54	0.54	0.48
	原始數據	0.19	-0.12		0.18	0.22		
	分數	0.40	0.32	0.13	0.36	0.31	0.11	0.12
	原始數據	0.41	-0.18		0.50	0.20		
	分數	1	0.09	0.09	1	0.39	0.39	0.24
	原始數據	0.24	0.22		0.23	-0.22		
	分數	0.65	0.32	0.21	0.61	0	0	0.10
	原始數據	0.35	0.18		0.37	-0.14		
	分數	1	0.48	0.48	1	0.25	0.25	0.36

表 4-3：眯眼狀態下各特徵點定位與眼睛分數

照片 \ 分數		左眼			右眼			雙眼總分
		高寬	偏移	總分	高寬	偏移	總分	
	原始數據	0.23	0.05		0.22	0.10		
	分數	0.61	1	0.61	0.52	0.81	0.42	0.51
	原始數據	0.23	0.06		0.25	0		
	分數	0.61	0.95	0.58	0.69	0.81	0.56	0.57
	原始數據	0.16	0.10		0.19	0.10		
	分數	0.29	0.78	0.23	0.42	0.79	0.33	0.28
	原始數據	0.18	0		0.25	0.06		
	分數	0.37	0.78	0.29	0.69	0.96	0.66	0.48
	原始數據	0.20	0.13		0.13	-0.11		
	分數	0.44	0.70	0.31	0.17	0.36	0.06	0.19
	原始數據	0.18	0.06		0.26	0.09		
	分數	0.34	0.95	0.33	0.78	0.84	0.66	0.49

表 4-4：正常狀態下各特徵點定位與嘴巴分數


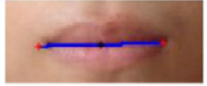

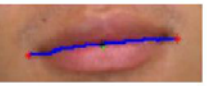

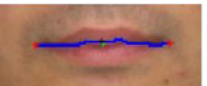






分數		嘴線曲度 原始數據	嘴巴分數
照片(原圖及嘴部)			
		0	0.82
		-0.02	0.70
		-0.02	0.69
		0.01	0.93
		0.01	0.94
		-0.05	0.49

表 4-6：大笑嘴狀態下各特徵點定位與嘴巴分數


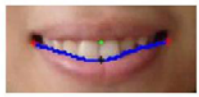










分數		嘴線曲度 原始數據	嘴巴分數
照片(原圖及嘴部)			
		0.13	0.87
		0.12	0.91
		0.10	0.98
		0.12	0.92
		0.09	0.99
		0.08	1

表 4-5：微笑嘴狀態下各特徵點定位與嘴巴分數


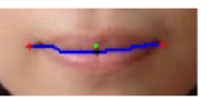

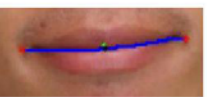



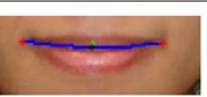

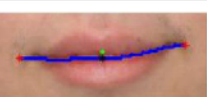

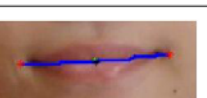
分數		嘴線曲度 原始數據	嘴巴分數
照片(原圖及嘴部)			
		0.05	1.00
		0.02	0.98
		0.01	0.95
		0.03	1
		0.04	1
		0.02	0.97

表 4-7：撇嘴狀態下各特徵點定位與嘴巴分數




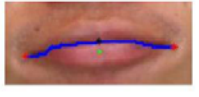

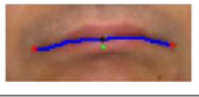

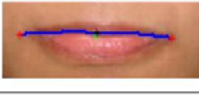



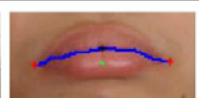
分數		嘴線曲度 原始數據	嘴巴分數
照片(原圖及嘴部)			
		-0.05	0.49
		-0.07	0.41
		-0.05	0.50
		-0.02	0.67
		-0.05	0.50
		-0.11	0.17

表 4-8：眼睛和嘴巴都優的狀態下三種評分分數

照片	系統相乘方法(0-1)	系統相加方法(0-1)	人工評分(1-5)
	0.86	0.93	4.22
	0.91	0.95	3.88
	0.51	0.74	3.55
	0.84	0.92	4.66
	0.57	0.79	4
	0.79	0.89	4

表 4-11：嘴巴不正常的狀態下三種評分分數







照片	系統相乘方法(0-1)	系統相加方法(0-1)	人工評分(1-5)
	0.41	0.66	2.44
	0.37	0.66	2.44
	0.25	0.5	2
	0.31	0.57	2.55
	0.22	0.47	2.55
	0.16	0.55	2.44

表 4-9：眼睛不正常(斜眼)的狀態下三種評分分數







照片	系統相乘方法(0-1)	系統相加方法(0-1)	人工評分(1-5)
	0.3	0.63	2.22
	0.33	0.59	2.22
	0.12	0.56	1.77
	0.22	0.58	3
	0.08	0.46	2
	0.17	0.42	2.22

表 4-12：眼睛和嘴巴皆不正常的狀態下三種評分分數


照片	系統相乘方法(0-1)	系統相加方法(0-1)	人工評分(1-5)
	0.23	0.49	1.55
	0.21	0.46	1.44
	0.07	0.32	1.33
	0.11	0.42	1.44
	0.02	0.29	1.55
	0.17	0.41	1.33

表 4-10：眼睛不正常(眯眼)的狀態下三種評分分數







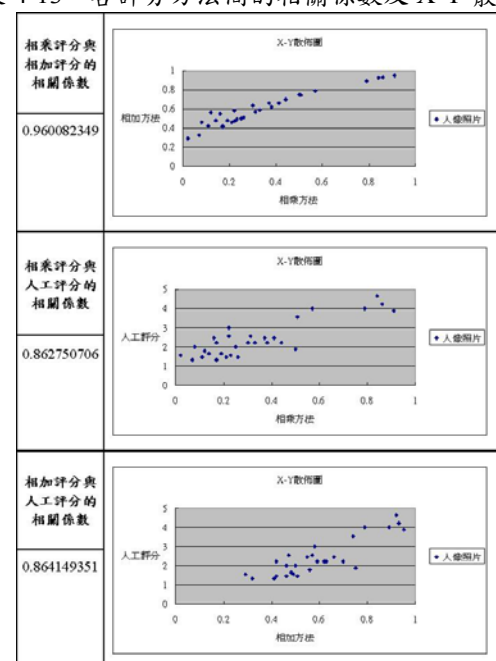
照片	系統相乘方法(0-1)	系統相加方法(0-1)	人工評分(1-5)
	0.5	0.75	1.88
	0.38	0.62	2.22
	0.19	0.48	1.66
	0.44	0.7	2.22
	0.14	0.48	1.66
	0.26	0.51	1.44

表 4-13：各評分方法間的相關係數及 X-Y 散佈圖



4.2. 實驗分析與討論

關於評分方面，人臉長相的特質是影響評分結果的重要因子；例如影像中的人物天生眼睛就較小較細長，而他在正常眼的狀況下所得的評分不會太高；有些人的嘴巴在無表情正常情況下呈現略為下凹嘴型，而他在正常嘴的狀況下所得的評分會較低。在表情綜合評分方面，第一套評分機制中可看出，測試影像若在系統的單項評分分數太低，即給予否定最後加入整體綜合評分時會大大降低總分數，屬於較強烈的評分方式；而第二套評分機制中可看出，測試影像若在系統的單項評分分數太低，最後加入整體綜合評分的總分數也只會扣一些分數，屬於較緩和的評分方式。

我們對系統評分和人工評分做相依性實驗，將系統評分機制所算出來的分數分別和人工評分進行比對，並算出它們之間的相關係數，由實驗結果可以得出，無論是系統的第一套相乘方法或是第二套相加方法，皆取得了與人工評分有相關係數為 0.86 的相關性。

5. 結論

在本篇論文中，我們主要是提出應用影像處理技術與顏色資訊，擷取人臉區域中眼睛 E 區域及 I 區域部份的各特徵點、閉嘴時的嘴巴閉合曲線部份的各特徵點、張嘴時的下嘴唇上緣線部份的各特徵點，透過影像中各特徵參數量化的方式，得到眼睛的高寬比和偏移量以及嘴巴的彎曲度，再針對眼睛的高寬比和偏移量以及嘴巴的彎曲度，設計綜合評分機制進行以眼口特徵為基礎，來評量人像照片優劣的研究。整體進行的工作項目有人臉偵測、眼睛和嘴巴定位、特徵點的擷取、兩種評分機制及人像影像實驗結果顯示。最後，將系統的兩種評分結果與人工評量的方式做比較，實驗結果顯示高達 86% 的相關性，可以驗證我們的系統對於一張人像照片做出優劣判斷，大致與人類主觀的判斷無異。

6. 參考文獻

[1] G. Donato, M. S. Bartlett, J. C. Hager, P. Ekman, and T. J. Sejnowski. "Classifying facial actions". IEEE Transaction on Pattern Analysis and Machine Intelligence, 21(10): 974-989, October 1999.

[2] B. Fasel and J. Luetten, "Automatic facial expression analysis: a survey," Pattern Recognition, Vol. 36, pp. 259-275, 2003.

[3] Rafael C. Gonzalez, Richard E. Woods "Digital Image Processing second edition", Prentice-Hall, 2002.

[4] R.L. Hsu, M. Abdel-Mottaleb, A.K. Jain, "Face detection in color images", IEEE Trans. Pattern Analysis and Machine Intelligence, vol.24, no.5, pp.696-706, May 2002.

[5] E. Hjelmas, and B. K. Low, "Faces Detection: A Survey", Computer Vision and Image Understanding, vol.83, pp.236-274, Apr.2001.

[6] C. H. Lin and K. C. Fan, "Triangle-based approach to the detection of human face," Pattern Recognition, vol. 34, no. 6, pp. 1271-1284, June 2001.

[7] T. Leyvand, D. Cohen-Or, G. Dror, D. Lischinski, "Data-Driven Enhancement of Facial Attractiveness." SIGGRAPH, 2008.

[8] H. Rowley, S. Baluja, and T. Kanade, "Neural Network-Based Face Detection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 1, pp. 23-38, Jan. 1998.

[9] F. Y. Shih, S. Cheng, C. Chuang and P. Wang, "Extracting faces and facial features from color images". Pattern Recognition and Artificial Intelligence, vol. 22, no. 3, pp. 515-534, May 2008.

[10] Y. Tian, T. Kanade, and J. F. Chon "Dual-state parametric eye tracking." In Proceedings of Conference on Computer Vision and Pattern Recognition, 2000.

[11] Y. L. Tian, T. Kanade, and J. F. Cohn, "Multi-State Based Facial Feature Tracking and Detection" Robotics Institute, Carnegie Mellon University, August, 1999

[12] Y. L. Tian, T. Kanade, and J. F. Cohn, "Recognizing Action Units for Facial Expression Analysis," IEEE Trans. Pattern Analysis and Machine Intelligence, vol.23, no.2, 2001

[13] P. Viola and M. J. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," in Proceedings of the IEEE Computer Society International Conference on Computer Vision and Pattern Recognition, vol.1, pp.511-518, Dec. 2001.

[14] G. Yang and T. S. Huang, "Human Face Detection in Complex Background", Pattern Recognition, vol. 27, no. 1, pp. 53-63, 1994.

[15] M.H. Yang, D.J. Kriegman and N. Ahuja, "Detecting Faces in Images: A Survey", IEEE Trans. Pattern Analysis and Machine Intelligence vol.24, no.1, pp.34-58, Jan.2002.

[16] M. H. Yang and N. Ahuja, "Gaussian Mixture Model for Human Skin Color and Its Application in Image and Video Databases," Proc. SPIE: Storage and Retrieval for Image and Video Databases VII, vol. 3656, pp. 458-466, 1999.