

第四章 臉部表情識別系統

臉部表情做為人與人溝通重要資訊，在多數的情況下往往能夠代表一個人情緒的直接反應，可能是高興或是悲傷、難過等情緒，皆可經由臉部表情變化來傳達。有研究指出人在開車時情緒較容易激動，這是相對於走在路上時，即便是與他人產生碰撞，兩人均能直接透過表情傳達抱歉之意，使得對方則較不易動怒；而在駕駛座上，則因無法看到對方的表情，所以較容易引起不滿情緒。大體而言，人是群居的動物，常需要與他人互動，因此表情的溝通有時候勝過千言萬語，常言道一個微笑便可以化解許多不滿情緒，因此如何讓電腦也懂得人類表情所代表的涵義是項有意義的研究。

在本章中，我們希望從影像中已被偵測、識別出的人臉區塊，進行更細部的表情特徵分析，藉由引用相關的心理學研究從而更智慧化的體認出目標的情緒變化狀態；這些表情辨識的基礎除了在識別情緒的應用之外，從數位影像合成的角度來看，目前應用於網路上的虛擬人物，也能經由這些被識別的特徵賦與其情緒上的表情變化，使其在互動式介面中增加生動感。此外，亦可應用在相關的新一代玩工具業上的開發，試想同時具有識別情緒及遠端監控等功能的玩具來陪伴小孩，能為父母帶來於幼兒照護上多大的輔助及便利性，此等多元價值乃是方今表情分析之價值所在。

4.1 臉部表情辨識相關研究

早在 1971 年 Ekman 和 Friesen[14]對於臉部表情有相當深入的研究，將人的表情主要分成七大類，有高興(happiness)、悲傷(sadness)、害怕(fear)、討厭(disgust)、驚訝

(surprise)、生氣(anger)，以及中性表情 (neutral)，之後相關研究為區分七個不同類別的表情。在表情識別問題中，用來分析資訊來源可分為在動態影片[15]以及與靜態單張照片而有所不同，在動態的環境下，可使用的資訊比較多，因此有較高的準確度，但缺點是應用環境較容易受到限制，而本研究則著重在於靜態單張照片的研究。由於可用的參考資訊較少，因此單張照片的處理較為困難，目前在這方面的研究也比較多，其中所用到的分類技術有 Neural Networks[16][17][18]、SVM[21][22][23][24]、Conditional Random Fields[25]、AdaBoost[26][27]、Hidden Markov model(HMM)[28][31]、Nearest-Neighbor [29] 等相關技術，在臉部特徵萃取則用到 Gabor Wavelet[20][27]、Active Appearance Models (AAM)[19][30][31][32][33][34]等方法，在這麼多相關的研究當中，多數在處理如何定位臉部五官的特徵位置，再設計一個有效的分類器來區分這些定位出來的特徵位置，在此我們對這相關技術做一些探討。

4.1.1 Active Appearance Models (AAM)特徵定位的方法

起初的做法是 Active Shape Model (ASM)，ASM 用一些點來表示這個模型的形狀，這些點需由手動方式標記，標記這些點相關位置是很重要的，點的座標用 \bar{x} 來表示，再透過 Principal Component Analysis(PCA)做基底的轉換（式子 4.1），建立起 shape model，再依據 model 來找尋最佳對應表情特徵位置，而找尋的過程有位移、大小、旋轉等步驟。

$$x = \bar{x} + P_s b_s \quad (4.1)$$

式 4.1 \bar{x} 代表 mean shape， P_s 代表 shape 正交變異量的集合， b_s 代表改變 shape 的參數。

$$g = \bar{g} + P_g b_g \quad (4.2)$$

式 4.2 \bar{g} 代表平均灰階值， P_g 代表灰階值正交變異量的集合， b_g 代表改變灰階值 shape 的參數。

$$b = \begin{pmatrix} W_s b_s \\ b_g \end{pmatrix} = \begin{pmatrix} W_s P_s^T (x - \bar{x}) \\ P_g^T (g - \bar{g}) \end{pmatrix} \quad (4.3)$$

式 4.3 W_s 是一個對角矩陣為 shape 參數的權重， b 是代表將兩個矩陣合成一個矩陣來表示。

$$b = Qc \quad (4.4)$$

式 4.4 Q 是代表 eigenvectors[35]， c 代表控制參數的向量包括有 shape 與灰階值的參數。

$$\text{當 } Q = \begin{pmatrix} Q_s \\ Q_g \end{pmatrix} \quad (4.5)$$

$$x = \bar{x} + P_s W_s Q_s c \quad (4.6)$$

$$g = \bar{g} + P_g Q_g c$$

將式 4.1、4.2、4.3、4.4 合併即可得式 4.6

$$g = (g_{im} - \beta 1) / \alpha \quad (4.7)$$

式 4.7 中 g_{im} 代表灰階資訊，在灰階值的參數部份可分為 α 與 β 兩個部份， α 是代表 scaling， β 則是代表 offset，透過正規化的過程以調整光線的影響。

圖 4-1 即為使用 AAM 將五官特定位置予以標示出來，在經過 shape 與 appearance 兩個 model 的訓練之後，當有新的照片就會找尋最佳 shape 與 appearance 的組合，但需

假設原本的 model 與 fit model 的形狀不能夠有太大的差異或者太多的位移。

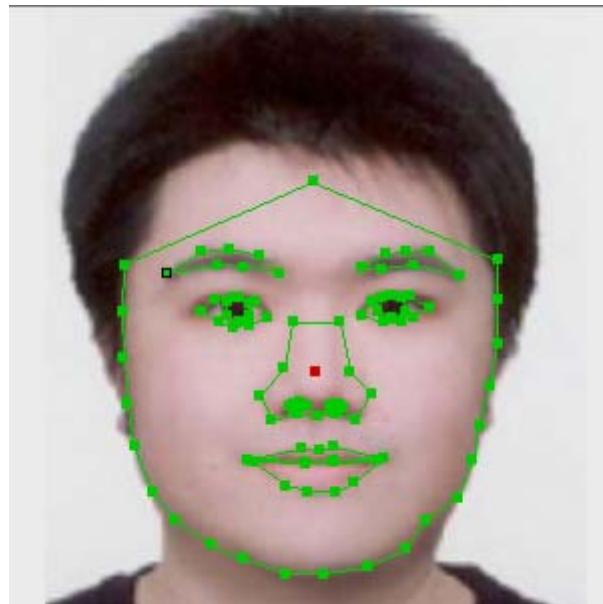


圖 4-1 為用 AAM 找出臉部特徵對應位置

4.1.2 Facial Action Coding System

人的表情是透過特定幾條肌肉控制而有不同變化，而此方法並不是針對單一條肌肉的變化來判別表情，測量的單位是透過 Action Units(動作單元)來做識別，Action Units 將各種表情加以細分成 56 種，表 4-1 左邊是取部份表情圖片照片，右邊是說明相對應的 Action Units[13]。

	內眉毛向上
	外眉毛向上
	眉毛向內

	額頭向上
	臉頰向上
	強烈擠眉毛
	嘴脣向上
	緊閉嘴脣
	嘴角向上

表 4-1 各種表情單元[13]

4.1.3 臉部表情辨識相關研究比較表

在自動表情識別相關研究中其結果差異很大，Kotsia[23]使用 Multi-Class SVM 作為分類方法，其正確率高達 98.2%，而 Tang [24]也是使用 SVM 作為分類方法，但實驗結果卻只有 65.74%，大部份的研究結果是介於 65% 至 90% 之間，其中差異如此大，主要問題在測試資料庫或訓練樣本上，當然所有測試樣本當訓練樣本其結果必然不差，但實質意義並不大，較客觀標準應該是訓練與測試資料是分開的，在這此相關研究我們整理如表 4-2。

Author	Features	Classification	Recognition rate [%]	Database
S.C. Tai [16]	IODB	Neural Networks	88.18%	JAFFE DB
M.P. Loh[17]	Gabor Wavelet	Neural Network	83.75%	EFEC DB
J.Y. Chang[18]	Facial Characteristic points FCP	Neural Networks	92.1%	-
F. Tang[19]	Local Facial Features AAM	Neural Networks	Nearly 88%	CMU-Pittsburgh AU-Coded Face expression image DB
W.F. Liu[20]	Multiple Gabor Features	Neural Network	79.71	JAFFE DB
C.F. Shan [21]	Local Binary Patterns	SVM	87%	Cohn-Kanade DB
Q. Xu[22]	-	Confusion Support vector machine tree	96.31%	Cohn-Kanade DB

		SVMT		
I. Kotsia[23]	Deformed Grid	Multi-Class SVM	98.2%	Cohn-Kanade DB
J. Tang [24]	Intensity Difference Image	SVM	65.74%	Cohn-Kanade DB
A. Kanaujia[25]	Local Non-negativ e Matrix	Conditional Random Fields	93%	-
U.k. Jung[26]	Rectangle features	AdaBoost	-	AR face DB
H. Deng[27]	Gabor Feature Extraction	AdaBoost	AR 89.16% JAFFE 97.33%	JAFFE DB AR DB
Petar S. [28]	Facial definition parameters(FDPs)	HMM	88.73%	Cohn-Kanade DB
X. X. QI [29]	Wavelet Energy Feature FLD	Nearest-Neigh bor	90.36%~94.6 4%	JAFFE DB
K.S. Cho[30]	AAM	EFM	87.2%	
F. De la Torre[31]	AAM	HMM	68%	DS 107

X. Feng[32]	AAM Local Binary Patterns	Linear Programming	83.6%	JAFFE DB
G. Zhou[[33]]	AAM	Rule base	90%	Cohn-Kanade DB
F. Dornaika[34]	AAM	Maximum Likelihood Estimator	80%	CMU DB

表 4-2 表情識別相關研究

4.2 表情識別系統架構與方法

上表列舉的許多研究，其臉部表情特徵定位是使用 AAM 方法，經過實測過後光線對 Appearance model 的影響很大，且測試的樣本需要經過訓練才能夠比較精準的定位，所以本論文中則使用簡單的邊緣來做位置定位，我們主要處理正面表情及非特殊眼鏡邊框表情特徵定位。

當偵測臉部位置，本系統分成兩個決策的模組(如圖 4-2)，一個為 Action-Unit model 的方式，我們定義一些固定表情 pattern 透過訓練的方式而區分不同類別，產生的結果並不是單一類別，而是每個類别的信心指數，第二個決策模組則是用 Component-based 方法，透過偵測表情特徵定位，以偵測出眉毛向上、嘴角上揚等，系統會依對應表情類別提高信心指數，最後信心指數則為各種不同表情類別輸出之結果。

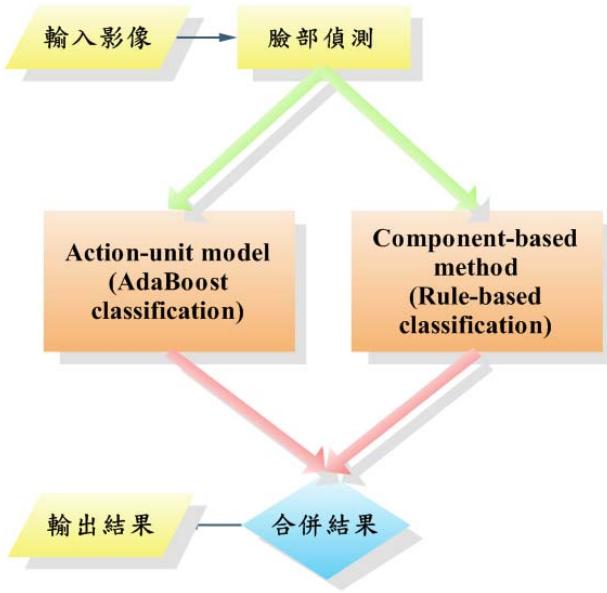


圖 4-2 臉部識別系統架構圖

4.3 表情資料庫的收集

臉部表情資料庫部份，除了使用 JAFFE 與 Yale 做為我們所實驗的資料，而更進一步的實驗方式則透過網路個人相簿所收集到照片，網路上所收集到的照片大多是生活照，透過人臉偵測技術將是臉部的區域儲存，再經由多位同學來做表情的標示，雖然網路上可以收集大量資料，但大部份所呈現表情都只有微笑與一般自然的表情，研究上還是會受到一些限制。

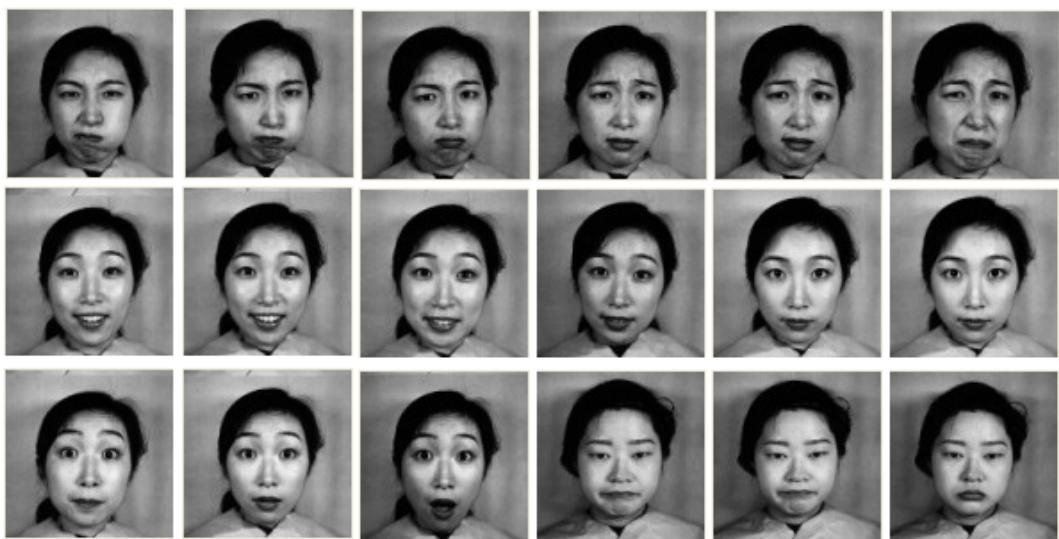


圖 4-3 JAFFE 表情資料庫



圖 4-4 Yale 表情資料庫



圖 4-5 網路收集臉部資料圖

4.4 Component-based 表情識別方法

先前多數的研究在定義表情特徵大都採用 AAM 的方法，在經過我們測試過發覺使用 AAM 的方法在於有形狀定義（shape）與臉部外貌特徵（appearance）都有不錯的效果，但其缺點在於 形狀擬合（shape fitting）過程中一但產生誤判就很難再回到對的模型，且臉部外貌特徵（appearance）需要經過訓練，在沒經過訓練的情況之下擬合的效果並不佳，因此一個很有效的模型的訓練過程將需要耗費大量時間和嘗試，所以我們希望透過較精確臉部定位，先找出眼睛位置，再運用以 DEM 為基礎來定位臉部眼睛與眉毛等表情特徵。

4.4.1 依比例定位表情特徵位置

首先我們先依照比例換算五官的位置， $W_m = W_f * 0.4$ ， W_f 為臉部的偵測臉部寬， W_m 為可能是嘴巴區域寬， $H_m = H_f * 0.45$ ， H_f 為臉部的偵測臉部區域的高， H_m 為可能是嘴巴區域的高，如圖 4-6 綠色方框所示， $W_e = W_f * 0.5$ ， w_e 為可能是眼睛區域的寬，

$H_e = H_f * 0.5$, H_e 為可能是眼睛區域的寬，如圖 4-6 紅色方框所示。

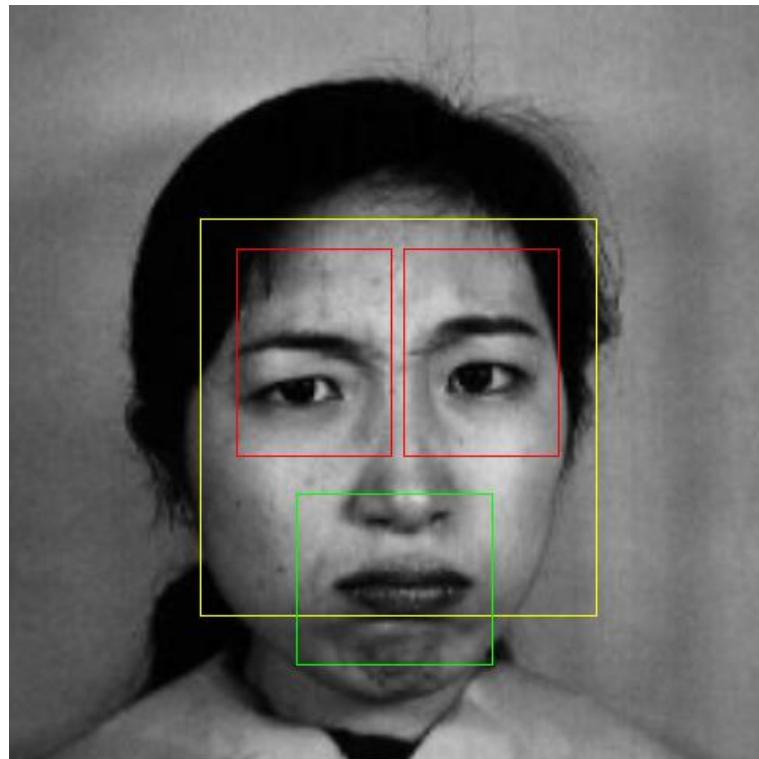


圖 4-6 依比例算出可能五官位置

4.4.2 DEM 五官位置定位

在定位出可能的五官位置時，再由邊緣的投影來找出最佳邊界，但由於傳統 Sobel edge(如圖 4-7)所處理的邊緣是將所有方向全部合在一起而產生一個結果，而我們的作法是將這些資訊分開處理，如此就可以得到更多不同方向資訊，且也比較不會受其他不同方向的邊緣影響。

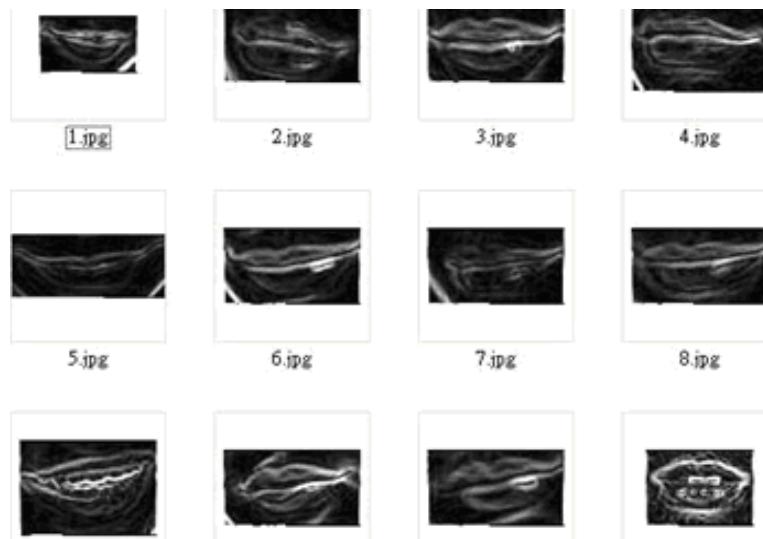


圖 4-7 Sobel 處理結果

4.4.3 嘴巴細部定位

圖 4-8 是我們經過方向性邊緣偵測處理，主要只有顯示水平方向的邊緣值，利用方
向邊緣偵測的結果來做投影，並找出最佳四個邊界點如圖 4-9，在找到四個邊界之後，
再用百分比的方式，取原本定位嘴巴位置 $1/2$ 做為找尋的範圍，寬度則依原始比例 $1/10$
做為找尋的寬度，再依每個不同的區塊進行更細微找尋，以更加精準的定位，其演算法
與上述定邊界的方式相同，其結果如圖 4-10 所示。

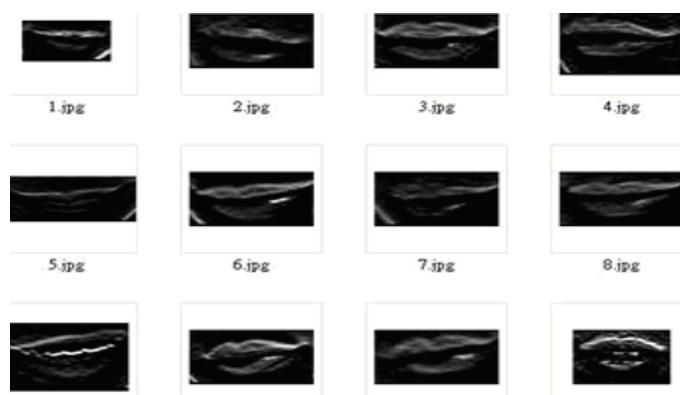


圖 4-8 方向邊算偵較不受其他方向影響



圖 4-9 嘴巴四邊定位之結果



圖 4-10 區域更加精準定位之結果

4.4.4 DEM 眼睛細部定位

眼睛可能的區域依照上述方法先定位出四個邊界點，即為圖 4-11 綠色框出的位置，當找到這四條線時再更進一步的找出眼睛的上邊界，找出眉毛的下邊界即可定位出眉毛及眼睛最佳可能的區塊。



圖 4-11 找出左眼眉毛上下邊界與眼睛定位之結果

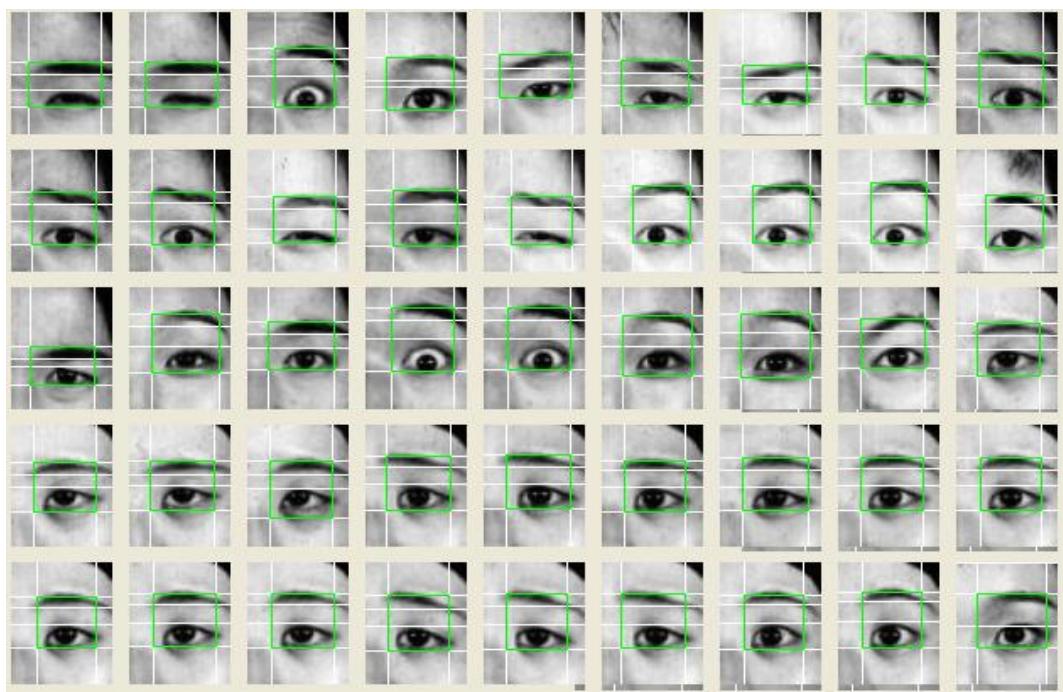


圖 4-12 找出右眼及眉毛上下邊界與眼睛定位之結果

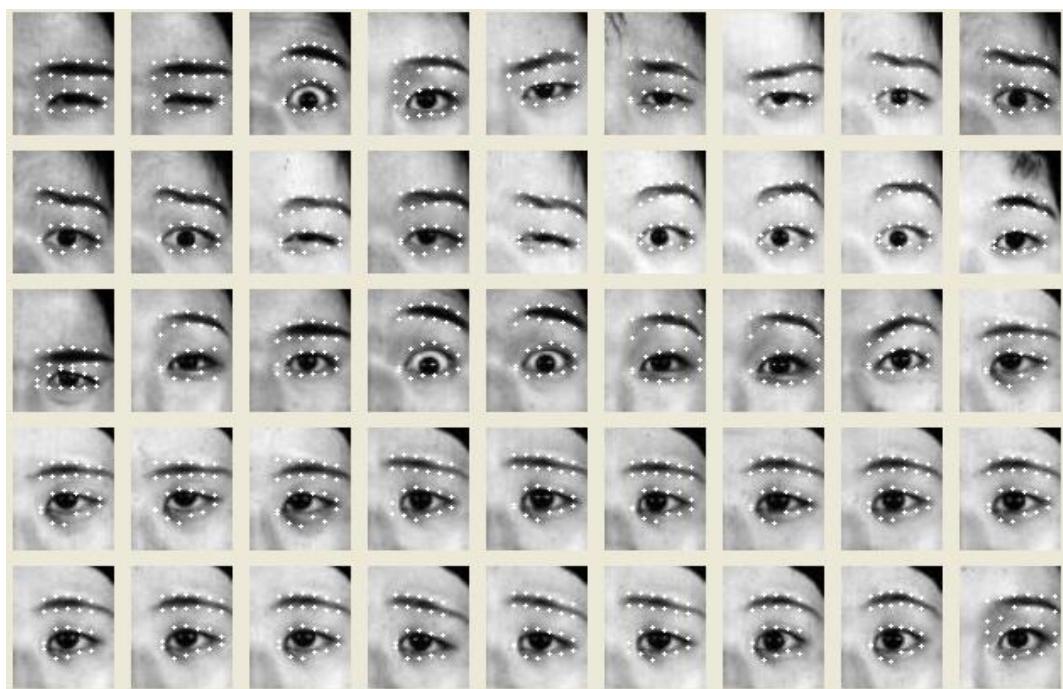


圖 4-13 依照比例的方式進行更細部的位置定位。

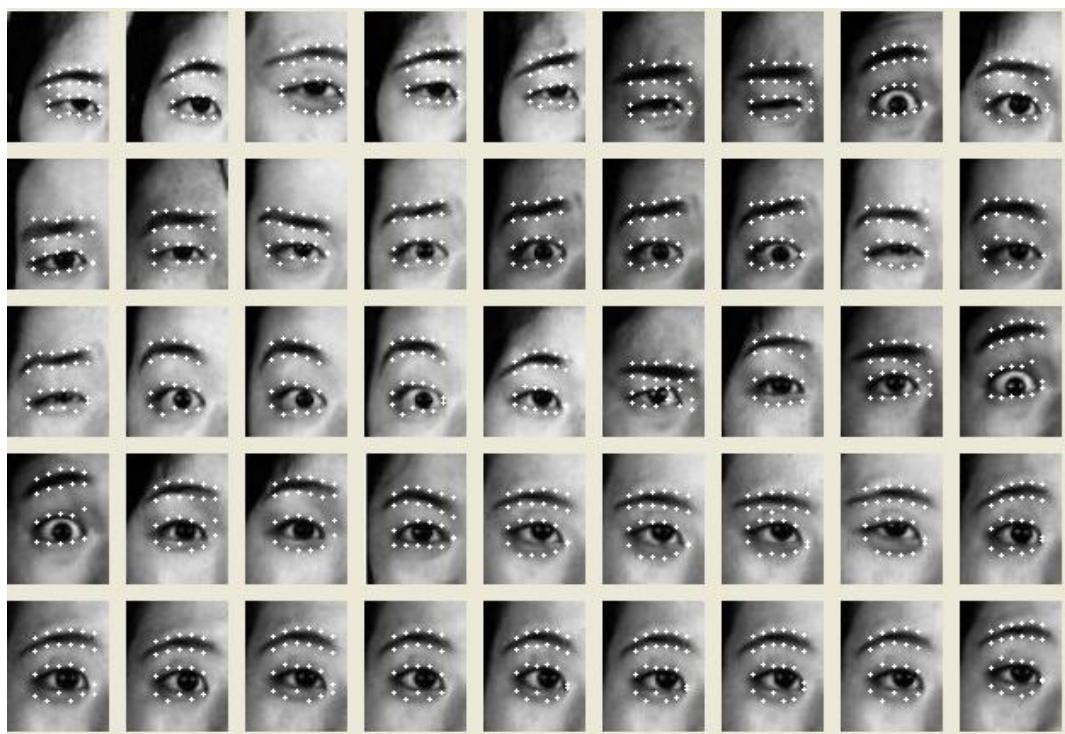


圖 4-14 為左眼睛所定位結果

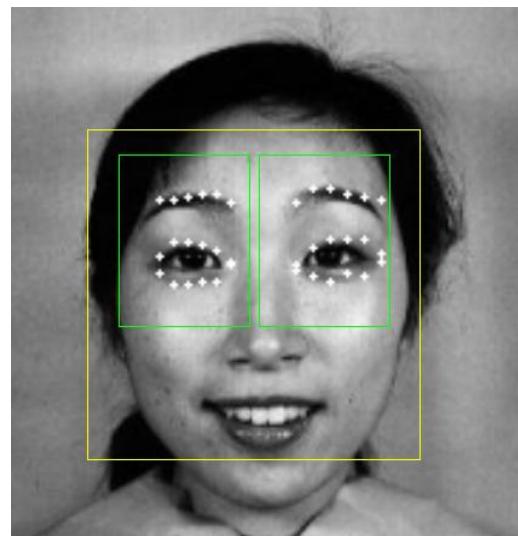


圖 4-15 依照可能位置依照比例定位結果

在表情五官特徵的定位實作方法上，首先我們假設額頭眉毛無遮蔽情形與不戴粗框眼鏡，在於這兩個基本假設成立下，如此才能夠先定位出四個邊界，進一步定位。在取

得五官的特徵位置後，經由這些座標相關資訊，建立一些簡單的規則來計算不同表情種類可信度。

4.5 精確臉部定位方法

經過臉部定位演算法找出臉部位置，在於正面情況下可以精確偵測到表情特徵位置(如圖 4-15 a)，但可能會因為臉部不同角度而影響定位表情特徵的結果(如圖 4-15 b)，因此我們需要一個能夠更加準確定位臉部位置方法，在先前的研究當中有利用臉部區域資訊做為不同光源膚色識別[7]，而我們則應用這概念做為臉部區域精確定位方法。



圖 4-16 (a) 精確定位表情特徵位置，(b) 臉部非正面表特徵產生偏移

4.5.1 轉換成二值化

先取出部份臉部資訊如圖 4-17 紅色區域所示，當此區塊為彩色時我們則轉換到 HSV 色彩空間，再使用色相 (Hue) 直方圖做為取樣標準，當此區塊為灰階則使用灰階直方圖做為取樣標準，當直方圖單一值累積大於總和值的 5% 則視為有效取樣值，再經過二值化其結果為圖 4-18。

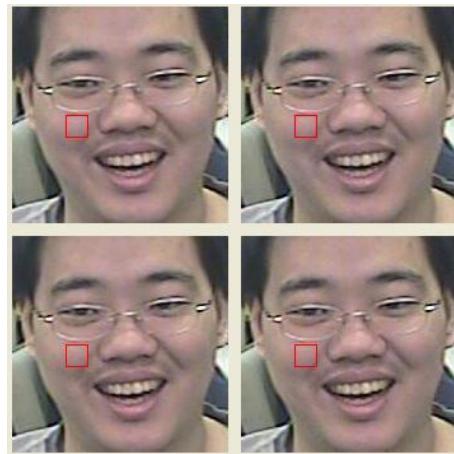


圖 4-17 取出紅色區域做為臉部區域定位的資訊

由於受到日光燈影響，CCD 攝影機在擷取影像時會有雜訊產生，所以要定位邊界並不容易，即使透過原本臉部的資訊來決定範圍，但還是有許多不相連情況與臉部之外錯誤的點，如圖 4-18 所示，需經過 Dilation 及 Erosion 處理過，再找相連元件的最大區塊其結果如圖 4-19 所示。

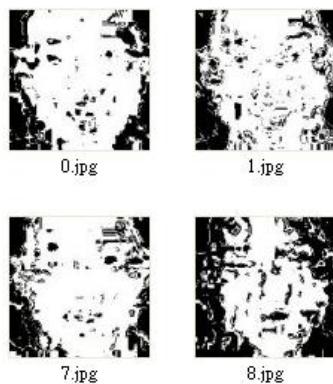


圖 4-18 彩色二值化結果

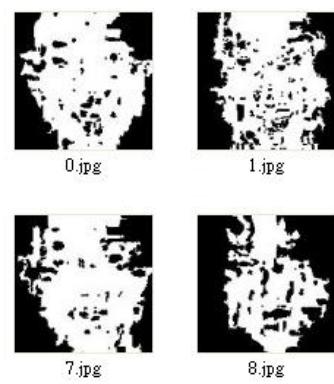


圖 4-19 找相連元件最大區塊結果

4.5.2 定位臉部邊界

在經過二值化後我們先找出所有邊界點，再取 30% 邊界點平均值做為臉部區域邊界值(如圖 4-20)，以白色線條為重新定位邊界結果，圖 4-21 紅色區塊為所要找區域參考

值，藍色線條為臉部區域邊界。

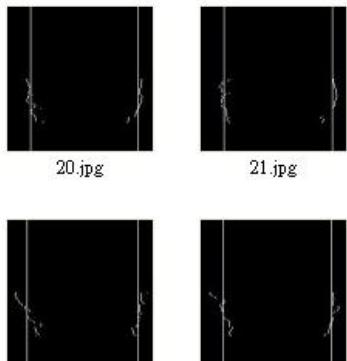


圖 4-20 兩個平均值為邊界點

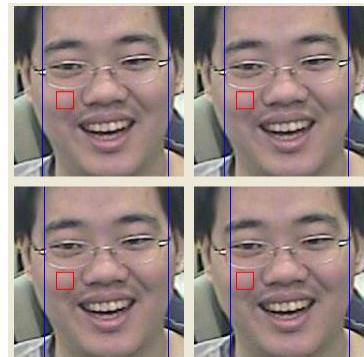


圖 4-21 藍色線條為臉部區域邊界

4.5.3 精準定位臉部位置結果

臉部區域經過較精準定位後，找臉部特徵位置則比較正確，圖 4-22 黃色點代表更精確定位表情特徵點。

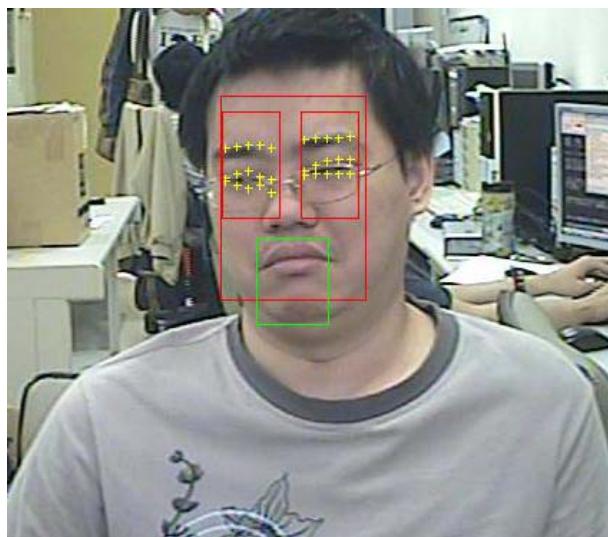


圖 4-22 黃色點代表更精確定位表情特徵點

4.5.4 臉部特徵表情識別規則

經由上述繁複處理程序，其最終目的就是找出五官特徵的對應位置，依這些相對位置建立表情識別的規則，在此我們定義五個規則(如表 4-3)，當這些規則成立時，則代表特定表情類別信心度會增加。

	表情識別規則	信心偏向趨勢
規則一	偵測眼睛是否閉上	悲傷
規則二	偵測眉毛上揚	驚訝
規則三	偵測嘴角上揚	高興
規則四	偵測嘴角向下	悲傷及生氣
規則五	偵測嘴巴是否張開	高興及驚訝

表 4-3 表情識別規則

4.6 Action-Unit model 實作方法

FACS (Facial Action Coding System)系統將表情的動作元件（Action Units）定義達 56 種之多，其中還有包括頭的角度，將可能包含的情緒分類很詳細，但若分類的過細，在各種表情的區別較不明確，造成系統設計上的困難，所以應該區分成較明顯的單元，如此可以讓系統有較高的準確度，我們先收集固定較容易識別的表情來做測試。

4.6.1 Action-Unit model 簡化

由於過多的 Action-Unit 會超出訓練與測試的負擔，不同類別之間差異性很小，對於實際應用上幫助不大，所以我們將對不同的表情找出最有代表性的 Action-Unit 做為訓練的 model。

4.6.2 Action-Unit 分類

我們將比較有代表性分成七個 model，且這些類別也是比較容易可以做出有效的區分，如表 4-4。

	額頭的部位，主要偵測上額是否有皺紋。
	下巴的部位，主要偵測下巴是否有皺紋。
	臉部兩側是否有明顯的線條。
	嘴巴的部份變化，依資料庫類別的標示為微笑表情。
	嘴巴的部份變化，依資料庫類別的標示為生氣表情。
	嘴巴的部份變化，依資料庫類別的標示為驚訝表情。
	嘴巴的部份變化。依資料庫類別的標示為悲傷表情。

表 4-4 Action-Unit Model

4.6.3 Action-Unit model 正規化

由於每個人的五官特徵的寬高比不盡相同，所以我們在訓練之前會將所有類別物件進行相同寬高比正規化，正規化的步驟為，先重新更精確定位四個邊界，再計算平均寬高比值，再依相同的寬高比值進行正規化，而每個物件類別的寬高可是不相同大小。

4.7 Action-Unit model 影像處理增加樣本數

由於所收集到的表情資料庫非常有限，為了增加訓練的樣本數，我們則透過影像處理的方法來增加，所用的方法分別是向左旋轉 1~5 度、向右旋轉 1~5 度、Dilation、Erosion、不同寬高比的縮放、高斯模糊等影像處理運算，透過這些影像的處理可以讓訓練的樣本快速增加，以達到可以識別更多類似的樣本。

4.7.1 影像處理結果

經過上述增加訓練樣本的方法，每單一個訓練樣本可以增加 20 倍的樣本數，其部份處理的結果如表 4-5。

	原圖
	Blur 處理結果
	Dilation 處理結果
	Erosion 處理結果
	Rotation -5 度處理結果
	Rotation 5 度處理結果
	高度調大
	高度調小

表 4-5 影像處理結果

4.8 表情單元實驗訓練樣本

收集到的 Action-Unit 影像再經由正規化與變型後，再透過機器學習的方式來分不同的類別，主要的四種表情單元模型為微笑、生氣、驚訝、悲傷等，及三種輔助表情單元模型為上額是否有皺紋、下巴是否有皺紋、臉部兩側是否有明顯的線條等，每種 Action-Unit 有自己的一個模型來做識別，再由得到的信心度來決定是那一個類別，可以取分數最高的方式，圖 4-23 驚訝表情樣本，圖 4-24 為微笑表情樣本，圖 4-25 為生氣表情訓練樣本，圖 4-26 為傷心表情訓練樣本。

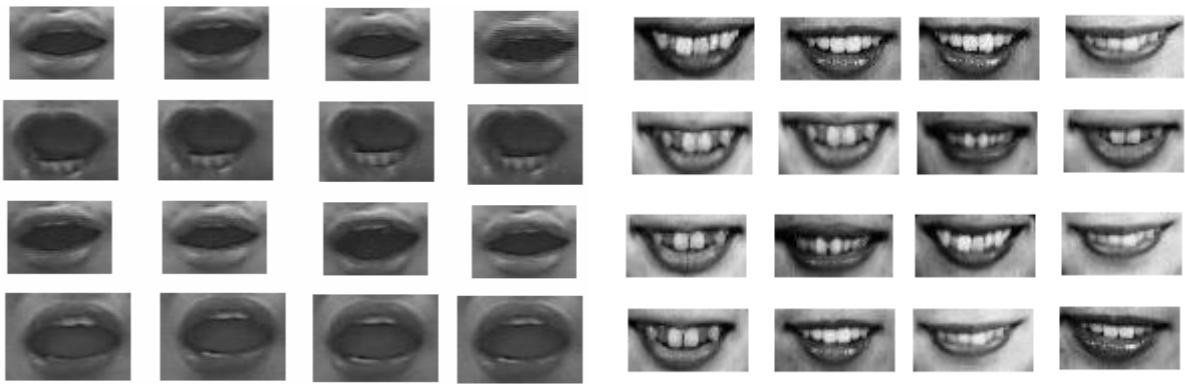


圖 4-23 為驚訝表情訓練樣本的照片

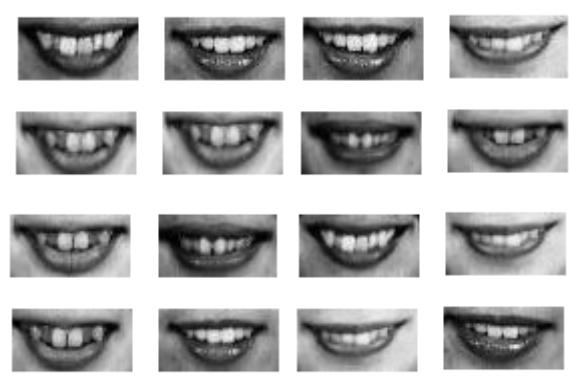


圖 4-24 為微笑表情訓練樣本的照片

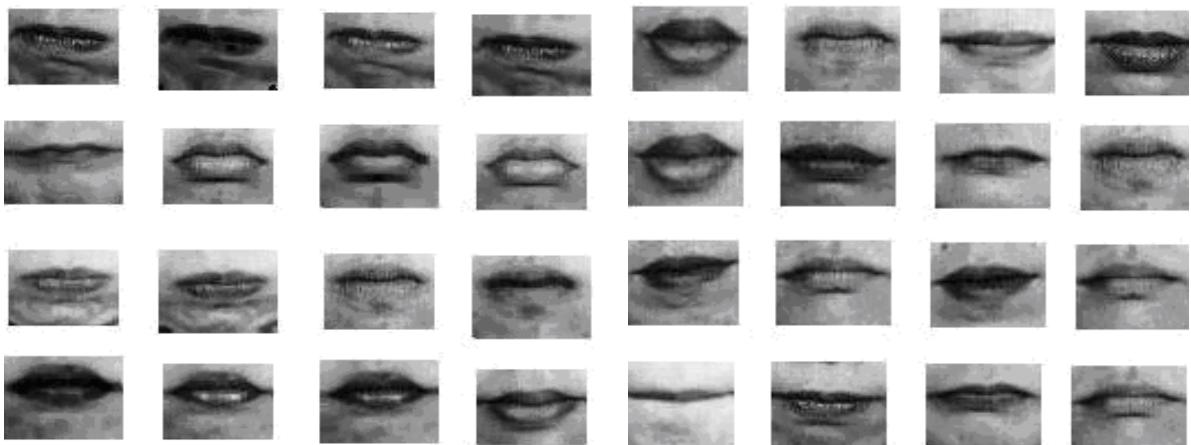


圖 4-25 為生氣表情訓練樣本的照片

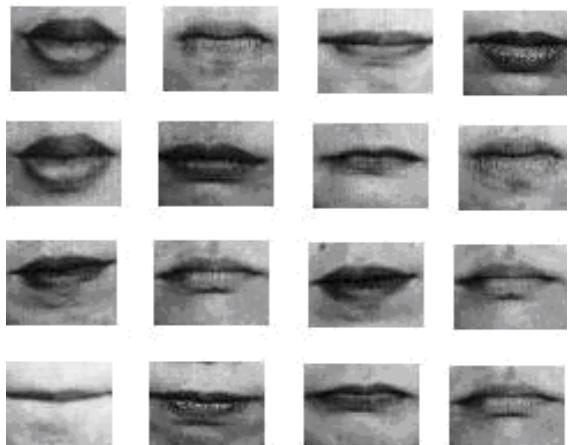


圖 4-26 為傷心表情訓練樣本的照片

4.9 表情識別實驗結果與分析

本章節主要是表情識別實驗方法與辨識率，在辨識率的部份我們採用通用分類，而不是為特定資料庫而設計分類器，所以我們是從網路收集很多訓練樣本，在測試數據部份我們則使用 JAFFE 與 Yale 兩個表情資料庫做為我們測試依據。

4.9.1 JAFFE 資料庫測試結果

JAFFE 資料庫有 10 個人員，共收集 213 張照片，分成六類表情：高興(happiness)、

悲傷(sadness)、害怕(fear)、討厭(disgust)、驚訝(surprise)、生氣(anger)等類別，每一個類別約有 30 張照片，其標示類別是採用投票(voting)方式，有些標示類別並非是很明確與絕對，是界於兩個類別中間如圖 4-28 (a) 看起來比較像是高興的類別但卻分在驚訝的類別，圖 4-28 (b)看來比較像是中性但還是分類在驚訝的類別，所以我們將分類出結果轉換成信心指數，在準確性評估時是以最高分數為主，而我們評估標準則分為 5 種類別，其結果如表 4-6。

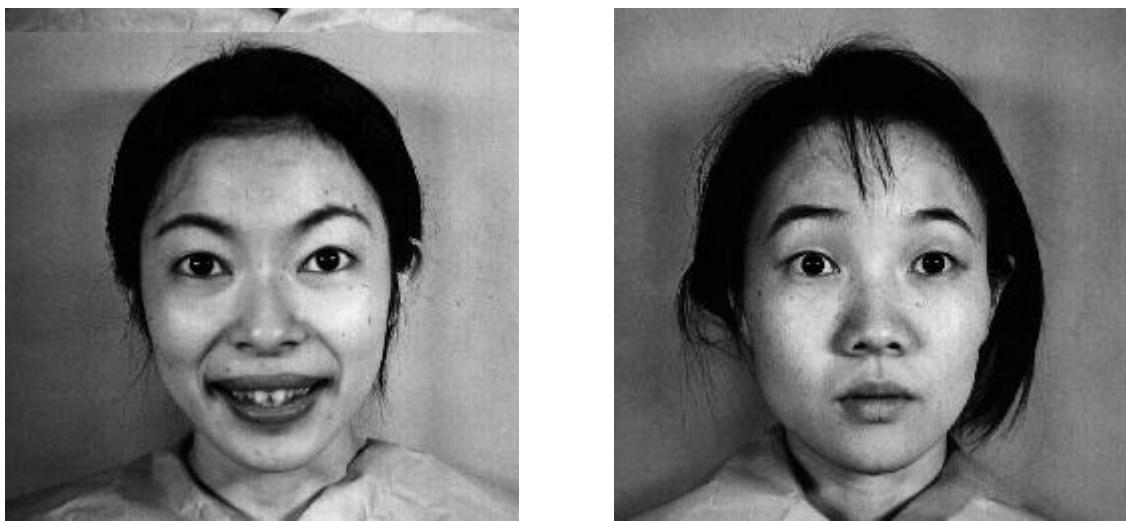


圖 4-27 (a)為驚訝表情類別，(b)為驚訝表情類別

類別 結果	高興 (happiness)	悲傷 (sadness)	驚訝 (surprise)	生氣 (anger)
高興(happiness)	73.08	9.68	6.67	3.23
悲傷(sadness)	11.54	58.06	3.33	16.13
驚訝(surprise)	0	6.44	73.33	3.23
生氣(anger)	15.38	25.81	16.67	77.42
中性(neutral)	0	0	0	0

表 4-6 原始解析度(256x256)下所測試結果

表情識別通常需要有較高的解析度才能有良好辨識率，在此我們針對較差的解析度進行識別的實驗，首先將原圖的解析度 256 x 256 像素，降成 230 x 230 像素，其結果並不會影響太大，如表 4-7，之後再將解析度降到 208 x 208 像素，其部份辨識率有降低的情況，如表 4-8，當解析度降成 128 x 128 像素，因臉部區域僅 52 x 52 像素，表情特徵並不明顯，以致於辨識率下降，在實際應用中使用者與攝影機的距離是很重要，我們使用 8 mm 的攝影鏡頭，在 320 x 240 的解析度下，使用者有效辨識距離為 1.5 公尺，當超過這個距離時系統辨識率則開始下降。

類別 結果	高興 (happiness)	悲傷 (sadness)	驚訝 (surprise)	生氣 (anger)
高興(happiness)	73.08	3.23	11.67	6.45
悲傷(sadness)	11.54	71.61	8.33	22.58
驚訝(surprise)	0	3.23	63.33	3.23
生氣(anger)	15.38	21.94	16.67	67.74
中性(neutral)	0	3.23	0	0

表 4-7 較差解析度(230x230)下所測試結果

類別 結果 %	高興 (happiness)	悲傷 (sadness)	驚訝 (surprise)	生氣 (anger)
高興(happiness)	73.08	6.45	26.67	6.45
悲傷(sadness)	7.69	61.29	6.63	19.35
驚訝(surprise)	0	6.45	60	3.23
生氣(anger)	19.23	22.58	6.7	64.52
中性(neutral)	0	3.23	0	6.45

表 4-8 較差解析度(208x208)下所測試結果



圖 4-28 資料庫為生氣類別，辨識出類別為生氣與傷心，信心度最高為生氣



圖 4-29 辨識出類別為高興

4.9.2 Yale 資料庫測試結果

Yale 資料庫有 15 個人員，共收集 165 張照片，分為高興(happiness)、悲傷(sadness)、驚訝(surprise)等三類表情，每一個類別約有 15 張照片，其標示類別是採用直接分類方式，其資料庫的光源變化比較大，因此樣本較具代表性，包括男生與女性及不同年齡層人員，在準確性評估時是以最高分信心度為分類類別，我們其結果分成 3 種類別，如表 4-9。

類別 結果 %	高興 (happiness)	悲傷 (sadness)	驚訝 (surprise)
高興(happiness)	66.67	0	0
悲傷(sadness)	0	60	0
驚訝(surprise)	6.67	10	73.33
生氣(anger)	13.33	16.67	13.33
中性(neutral)	13.33	13.33	13.33

表 4-9 Yale 資料庫所測試結果

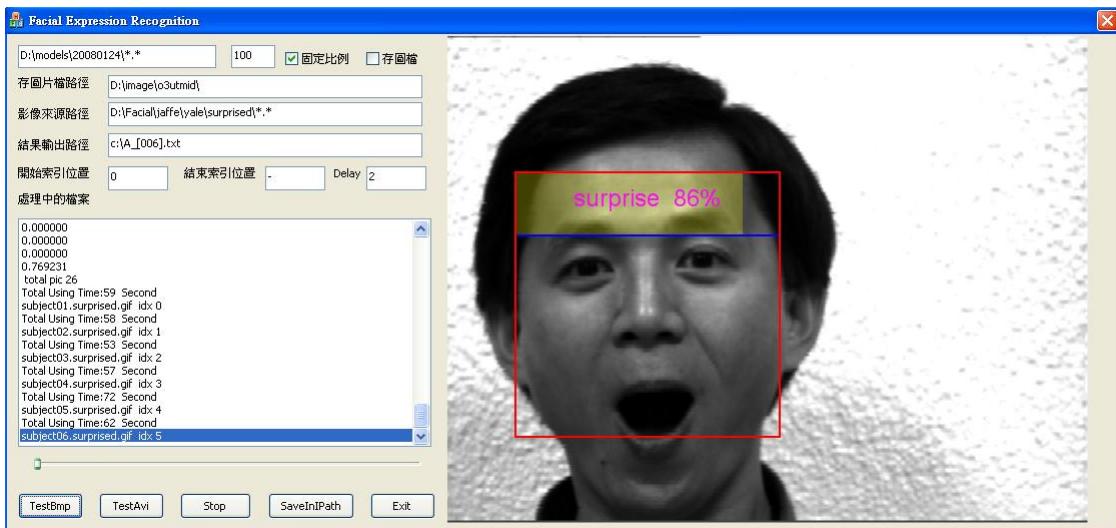


圖 4-30 其辨識出類別為驚訝

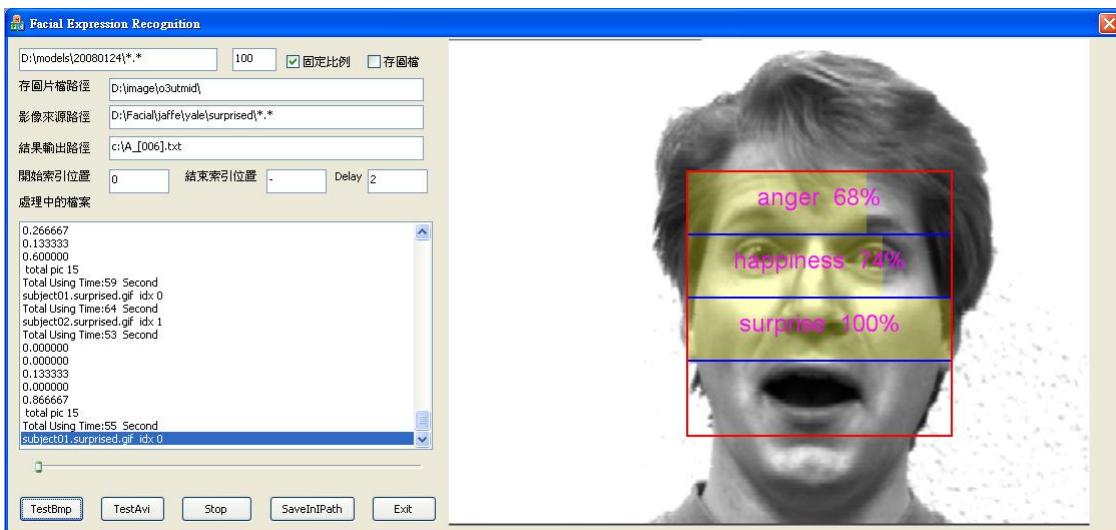


圖 4-31 其辨識出有三種類別，信心度最高為驚訝。

4.9.3 自行收集資料庫測試結果

自行收集資料庫有 10 個人員，共收集 440 張照片，分為高興(happiness)、悲傷(sadness)、驚訝(surprise)、生氣(anger)等四類表情，每一個類別有 110 張照片，每個測試人員收集 11 張圖片，其標示類別是採用直接分類方式，在我們收集臉部資料庫有較

佳辨識率，其結果如表 4-10。

類別 結果 %	高興 (happiness)	悲傷 (sadness)	驚訝 (surprise)	生氣 (anger)
高興(happiness)	83.6	4.5	3.6	3.6
悲傷(sadness)	5.5	72.7	8.2	11.8
驚訝(surprise)	2.7	12.7	80	3.6
生氣(anger)	6.4	4.5	6.4	78.1
中性(neutral)	1.8	5.5	1.8	2.7

表 4-10 本論文收集資料庫所辨識結果

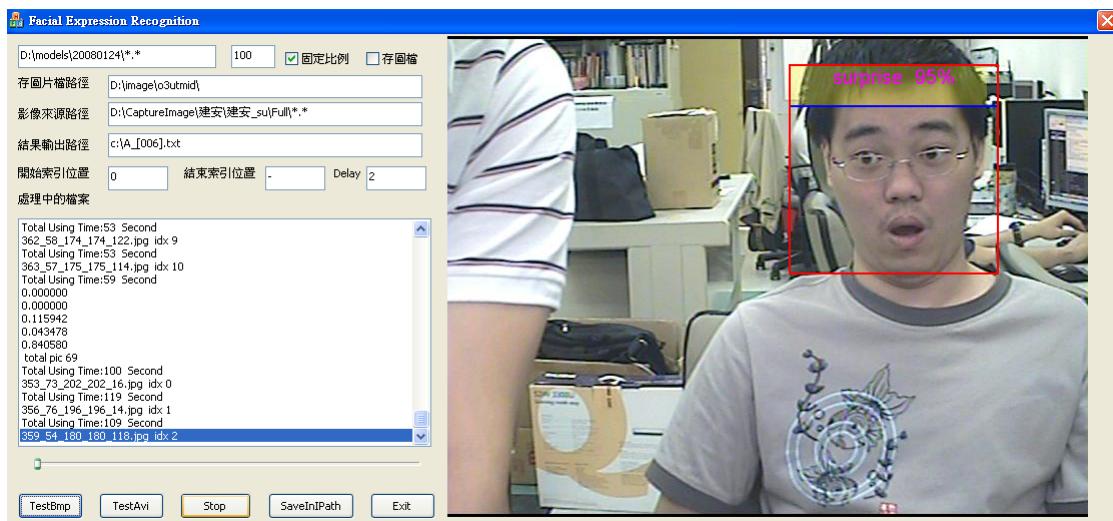


圖 4-32 所辨識出類別為驚訝，其信心度 95%

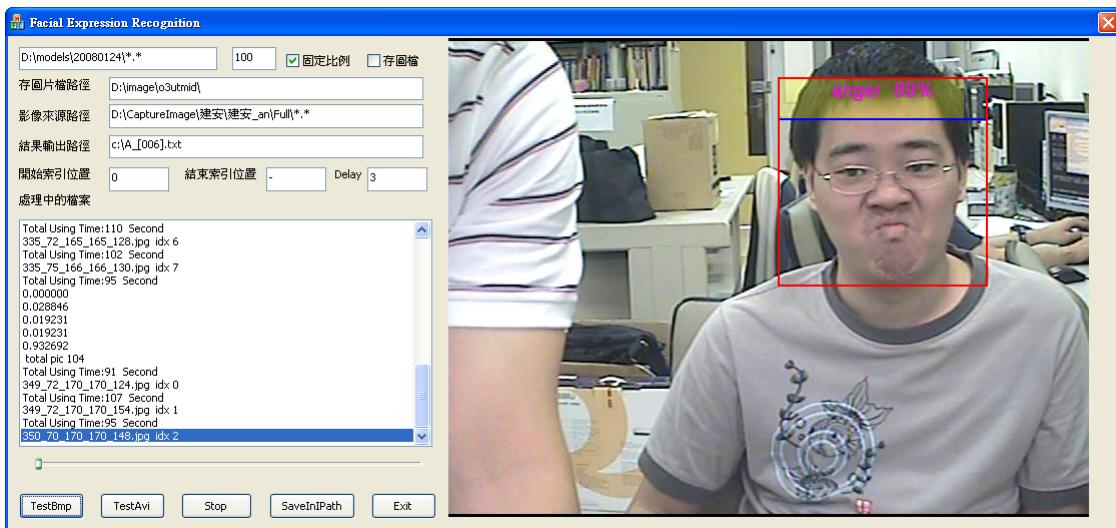


圖 4-33 所辨識出類別為生氣，其信心度 89%

4.9.4 多人表情識別測試結果

在單張照片中我們試著以多人同時做表情識別，由於臉部區域比較小所能顯示資訊較受限制，但還可以清楚描述不同表情類別，有部份測試人員結果多種表情類別，在未出現任何表情時，則代表中性類別，其結果如圖 4-34。



圖 4-34 多人表情識別的結果

4.9.5 從電影蒐集影像測試結果

一般在日常生活當中所會出現的表情大部都是中性(neutral)與高興(happiness)兩種表情，相對傷心與生氣類別比較少見到，而在所以我們就透過影片分析，取出有臉部的圖片，來測試臉部表情識別的效果，我們取電影「愛情三選一」中活靈活現的小女孩，結果如圖 4-35。

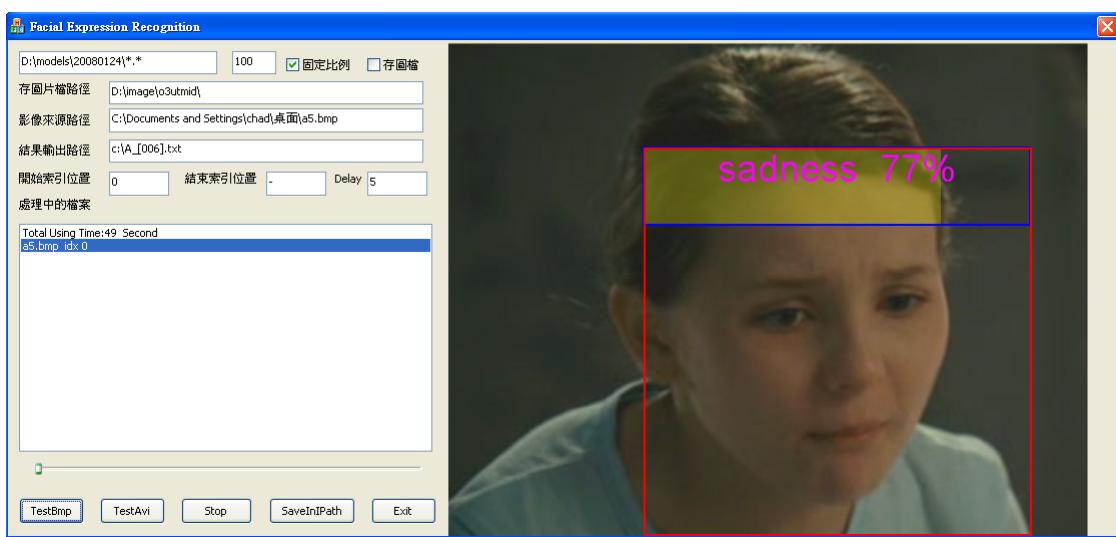


圖 4-35 電影「愛情三選一」中的小女孩

4.9.6 即時影像測試結果

在單張靜態的照片已有不錯的辨識結果，於是我們將系統移植到動態即時表情辨識，在動態視訊中多了時間資訊，透過累計分類結果，再決定是那一個類別表情，如此做法可以增加準確度，雖然會影響辨識時間，本論文所實作的結果可在一秒內進行不同表情分類，並不會有太多延遲的情形，測試結果如圖 4-36 至 4-39。

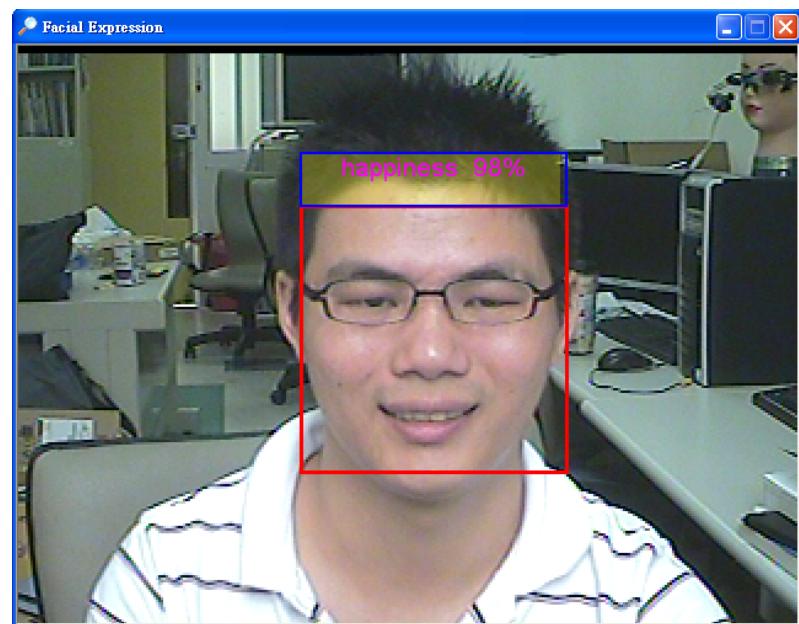


圖 4-36 系統自動辨識為高興的表情

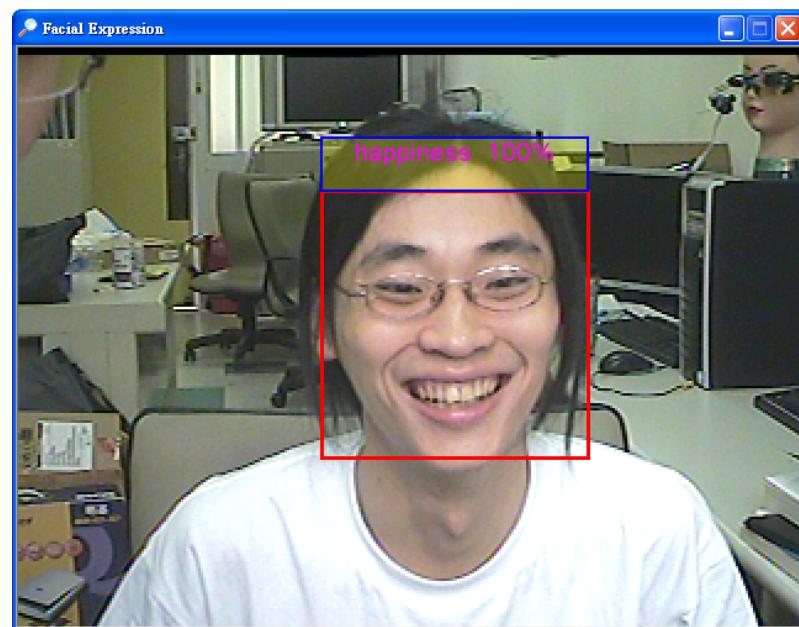


圖 4-37 系統自動辨識為高興的表情。

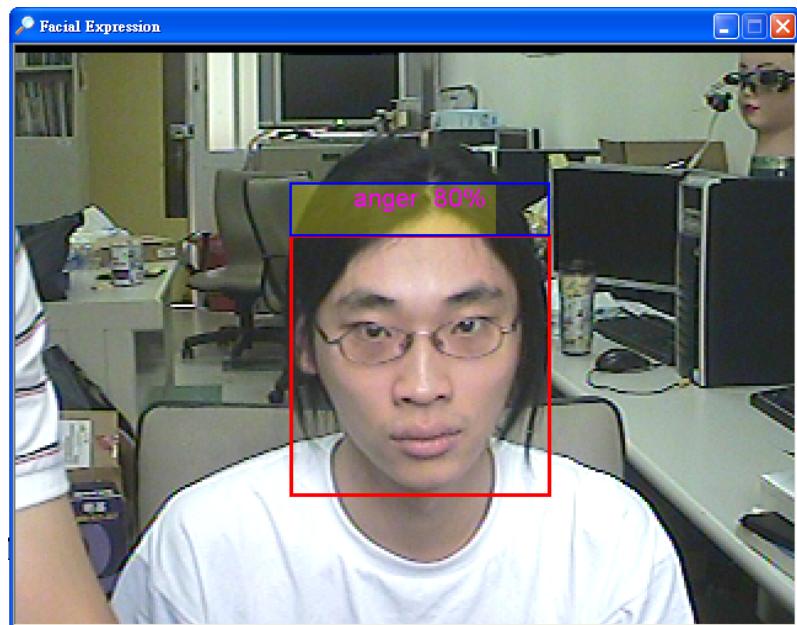


圖 4-38 自動辨識為驚訝表情。

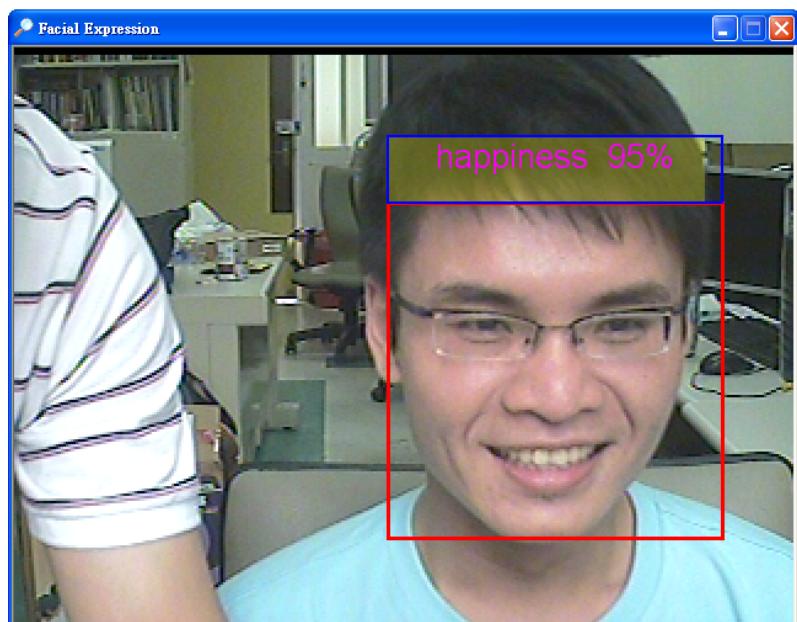


圖 4-39 自動辨識為高興表情