

以區域為基礎之表情辨識系統

葉純孝 林建宏 李承諺 劉遠楨

國立臺北教育大學理學院資訊科學系

E-mail: liu@tea.ntue.edu.tw

摘要

本文提出了一個在 Android 上，整合人臉辨識與表情辨識的學習管理系統，內容分為兩個部分，人臉辨識及表情辨識。在人臉辨識的部分，本文的辨識方式，需要使用者先將自己的影像存放至人臉資料庫中，再透過 PCA 的方式，先行訓練資料庫中的資料，得到前 K 個影像的權向量，再與輸入圖片比較，找出較接近的人臉，辨識正確率可達 85% 左右。在表情辨識的階段，先將偵測到的人臉區域依照比例找出雙眼及嘴巴的位置，接著透過本文的演算法將特徵點定位出來，再根據這些特徵點之間的關係，判斷出表情為高興、悲傷、閉眼、一般或是驚訝，辨識正確率約為 80.2% 左右。

關鍵詞：Android、人臉辨識、表情辨識、PCA。

Abstract

In this thesis, we proposed a learning management system with face recognition and facial expression recognition on Android. It is divided into two parts, face recognition and facial expression recognition. In face recognition, PCA is our primary method. The users should put their face images to the face database first. Then, training all data in the database. The weight vectors were obtained. The vectors compared with the input image. Finding the smallest distance, and we can get this input face belongs to whom. Recognition rate can reach 85%. In facial expression recognition, we detected the face first. According to face proportions, we can locate eyes and mouth. Next, found the feature points by our algorithm. There are 12 points at face. Then according to the relationship between the feature points, we can get a facial. Recognition rate can reach about 80.2%.

Keywords: Android, Face recognition, Facial expression recognition, PCA.

1. 研究動機與目的

由於科技的快速進步，平板與手機等等的行動裝置已經成為人們生活上不可或缺的物品之一，不論是影片、聲音、影像或文字等等…的資訊，無時無刻都可以快速地流通，為人們的生活造成了極大的改變，就連教育的部分也不例外，近幾年來，不論是國中或國小，逐漸由一般的學

習方式，改變為數位學習，也就是利用平板或著是其他科技產品的輔助，使課程能夠更佳的豐富且生動，再加上 Wi-Fi 及行動網路的普及，上課不一定要侷限在教室內，透過網路也能夠讓學生與老師之間，達到即時的互動，達到無時無刻皆能學習的目的。

數位學習的興起，有些教師已經從傳統的教學方式改變為使用科技輔具的教學方式，學生也不再是拿著傳統的書本，而是拿著行動裝置的方式學習，因此本文藉由人臉辨識的方式取代傳統的點名方式，希望減少點名時所需要的時間，教師能夠將較多的時間放在教學課程上，則表情辨識的部分，將學生在上課時的情緒變化收集，提供教師在課程上的作為參考，比方說，學生可能在某些時候會因為課程較為無趣而有疑惑或是想睡覺的動作，又或著是因為課程有趣而有高興的情緒產生，這些情緒都透過電腦分析而得到的數據，供教師作為參考的依據，使教師在設計課程上能夠更加偏向學生的喜好，使學生的專注力提升，進而提高學習的效率。

2. 文獻蒐集與探討

PCA 全名為 Principal components analysis，也就是主成分分析，通常在進行人臉辨識時，不會使用原始的資料作分析，因為原始影像的維度較高，運算時會造成負擔，因此 PCA 的概念就是以 m 維來表示 n 維的空間($m < n$)，也就是利用少數的維度來代表全部的維度，以上為 PCA 的基本概念，但要將 PCA 用運用於辨識上需要做一些變化。

Mallipeddi 等人[8]在 2012 年提出之文獻內提到，PCA 提取高維度特徵時，會有 NP 困難(NP-Hard)的產生，樣本量也會因為數量過少的問題，使得強分類器不易建構，因此它們透過差異進化演算法(differential evolution)，使訓練的數據之間的距離最大化以獲得各子集的不同特徵向量的組合，再利用這些數據來訓練分類器。

Ebied[9]於 2012 年提出結合線性與非線性的方法進行特徵的提取，線性的部分使用 PCA 做特徵的提取，非線性的部分則使用 Kernel-PCA，Kernel-PCA 是由 PCA 發展而來的，它代表一個高維特徵空間的非線性映射，從實驗結果可以證明，與單獨使用 PCA 的識別方法比較，有著更高的識別率。

表情辨識為利用電腦視覺得到影像後，再利用演算法將臉部分類，找出其中的特徵點，模擬出人所能看到的情緒反應，一般分為三個階段，人臉偵測、特徵點定位及表情辨識。

特徵點定位大致可分為整體分析與部分分析，整體分析是，先選定特徵點或區塊，利用連續性的影像，得到變形或是位移，進而判斷出結果，而部分分析，是將靜態影像分為幾個區域，根據這幾個區域中得到的資訊去做判斷的動作，Irene 等人[10]在 2006 年提出利用網格模型去定義人的表情變化，網格由 62 個節點組成，並利用 Kanade-Lucas-Tomasi 的追蹤技術，追蹤各節點的變化，再根據這些變化將表情辨識出來。

Ekman 與 Friesen[11]在 1978 年提出 FACS，其概念是將臉部分為多個小區塊，根據區塊的形狀，定義不同的參數，一個區塊就代表一個臉部動作，總共定義了 44 個動作，基本的表情包括生氣、高興、厭惡、害怕、驚訝及傷心六種表情。另一種方法是直接透過已知的特徵點來判斷表情，可分為整體特徵和局部特徵，在整體特徵，整張臉就是一個最大的特徵，局部特徵是先將人臉的眼睛、鼻子或嘴巴等區域找出來，接著根據這些局部區域中所得到的數據，得出最終的辨識結果，如 Tian 等人[12]提出的方法。

Tang 等人[13]在 2008 年提出一個基於自動選取特徵的 3D 表情辨識方法，利用 MPEG-4 standard 的 FAPUs 作正規化，再使用機率分布的方法找出特徵值，並取出前 K 個最大值，再以 AdaBoost 分類器分類，此方法較不會受到光線的影响。

有一些表情辨識是以模糊規則當作辨識依據，比如說，生氣時嘴巴會閉合、驚訝時會張嘴、高興嘴角會上揚等，在[14]中，使用 Mamdani' s inference method 進行模糊推論，再使用 Center of gravity defuzzification 解模糊，得到各表情的權重值，選出最大的權重值，將其對應的表情選擇出來，[15]也是依據此種做法。由於每個人的臉部狀態會有些許的不同，因此在條件的定義上如果不是很完備，可能就會有辨識錯誤的情形產生。

表情辨識也可以使用支持向量機(SVM)，SVM 是一種機器學習演算法，透過資料的訓練，建立一組資料庫，較不會因為人的表情變化因素而造成誤判，如[16]的方法中，使用雙模的方法，根據兩種資訊得到辨識數據，權重會因為可靠度的不同而改變，將訓練好的 SVM 分類器依照權重值進行多階層的分類，最後輸出可能的表情。

3. 表情辨識演算法

表情辨識最重要的步驟就是定位臉部的特徵點，有了這些特徵點才能夠做後續的表情判斷，因為此演算法需要在行動裝置上運行，為了效能

上的考量，本文抓取雙眼及嘴巴的特徵點作為辨識時的依據，每個嘴巴及眼睛各抓取四個特徵點，也就是會得到總計 12 個特徵點，再依據這些特徵點之間的相互關係，判斷出相應的表情，首先，對輸入影像做人臉偵測，確定人臉是否存在，若有偵測到人臉，才繼續後面的步驟，偵測到人臉後，接著大致框選出雙眼及嘴巴的位置，利用人臉偵測時所得到的資訊，即雙眼中心點與瞳孔之間的距離，接下來我們將分析的步驟分為三個部分，定位眼睛特徵點、定位嘴巴特徵點及辨識階段。

3.1 定位眼睛特徵點

步驟 1 框選到雙眼及嘴巴的位置以後，先處理眼睛的部分，首先將框選到的眼睛區域轉換為灰階形式。

步驟 2 採用 Valley-like 偵測器[17]來判定雙眼的輪廓，Valley-like 偵測器是一個 5x3 的遮罩，也可以看成是一個 5x3 的矩陣，將轉換後的眼睛灰階影像套用此遮罩，會得到一組影像矩陣。

步驟 3 將這組影像矩陣透過定義的閾值，將其二值化，此閾值的定義是根據反覆試驗後所得到的結果，會得到圖 3-1 的結果。



圖 3-1. 二值化後的眼睛影像

步驟 4 經過二值化後，很明顯看得出有非常多的雜訊，因此本文採取先膨脹再侵蝕，也就是閉合運算，將部分較弱的區域先連接，再將獨立的雜訊點去除，輸出結果如圖 3-2。



圖 3-2. 經閉合運算後的眼睛影像

步驟 5 經過步驟 4 的處理，雜訊已經降低不少，但仍然無法正確判斷出眼睛的位置，為了找出眼睛位置，本文使用八連通演算法，將各個連通區域做區分，找出前二大的區域作比較分析，根據長寬比來作初步的分析，若大於 1 者，就將其視為為眼睛區域，小於 1 則捨棄，如果兩個區域的長寬比皆大於 1，則有可能包含到眉毛的區域，根據眉毛及眼睛的相對位置，眼睛位於眉毛的下方，則判斷下方區域為眼睛區域。

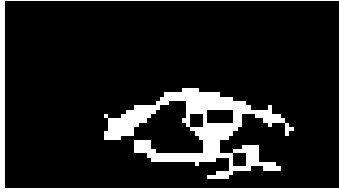


圖 3-3. 經分析後的眼睛區域

步驟 6 找到眼睛的實際位置後，先將最靠左右兩邊的特徵點定位出來，之後由這兩個點的 X 座標，藉此找出眼睛的中心點，以眼睛點為中心由上至下掃描，定位出眼睛的上界與下界。

完成上述六個步驟之後，即可定位出眼睛的特徵點，因為人有兩隻眼睛，所以需要再重複以上步驟一次，將左眼及右眼的特徵點定位出來。

3.2 定位嘴唇特徵點

步驟 1 將嘴巴區域的影像轉換至 YCbCr 色彩空間。

步驟 2 將影像二值化，得到 M1 影像。

步驟 3 再將嘴巴區域的影像轉換至 NCC(Normalize Color Coordinates)色彩空間，得到的影像稱為 M2。

步驟 4 將 M1 與 M2 做 AND 運算，得到影像 M，結果如圖 3-4。



圖 3-4. 嘴唇區域膚色偵測結果

步驟 5 為了使嘴唇區域能夠突顯出來，因此使用形態學上的膨脹，使嘴唇區域的輪廓更加明顯。

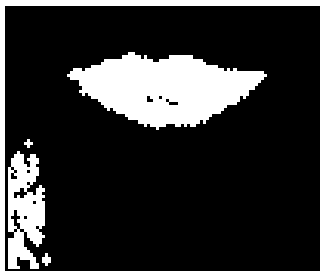


圖 3-5. 經膨脹後的結果

步驟 6 為了避免上唇與下唇會有分離的情形，使用垂直填補，使分離的區域結合在一起。

步驟 7 用八連通演算法，找出區域中，最大的

區塊，將此區塊判斷為嘴唇區域。

步驟 8 判斷出的區域，先將最左邊及最右邊的特徵點定位出來，再來根據這兩個特徵點的 X 軸座標，找出中心點，並由上至下掃描，找出嘴唇上方與下方的特徵點。

經過上述的步驟以後，雙眼及嘴唇的特徵點即可被定位出來，其結果如圖 3-6 所示。



圖 3-6. 特徵點定位結果

3.3 表情辨識階段

本文總共定義五種表情，包括一般、高興、悲傷、閉眼及驚訝，雙眼及嘴唇的特徵點已經被定為出來，每個位置各 4 個點，總計 12 個點，先將這 12 個點做一些基本的定義，首先將雙眼及嘴唇的斜率求出，如圖 3-16 所示，接下來要我們先簡單定義各個表情的可能條件，根據這些條件，將它們數據規則化，例如高興時，嘴唇會有上揚的情形，閉眼時眼睛上端與下端會較為接近…等等。

本文希望以不需要建立表情資料庫的方式，也能夠將表情辨識出來，因此依照定義，訂定出多條的規則，將不同表情的影像輸入，依照這些規則做訓練，每種類別的表情會各自得到一組數據，再經由分析這些數據，將各規則與表情之間的權重比定義出來。

由於規則較多的關係，在此不一一將其列出來，輸入的影像通過這些規則，如果符合條件，就加上該表情的權重分數，經由一連串的條件下來，每一種表情都有各自的數值，取出數值最大的作為辨識的結果，若在最終結果，數值有相同的情況，會依照悲傷、閉眼、驚訝、高興極一般的順序作為判斷的依據，越前面的表情優先權越大，圖 3-7 為表情辨識的完整流程圖。

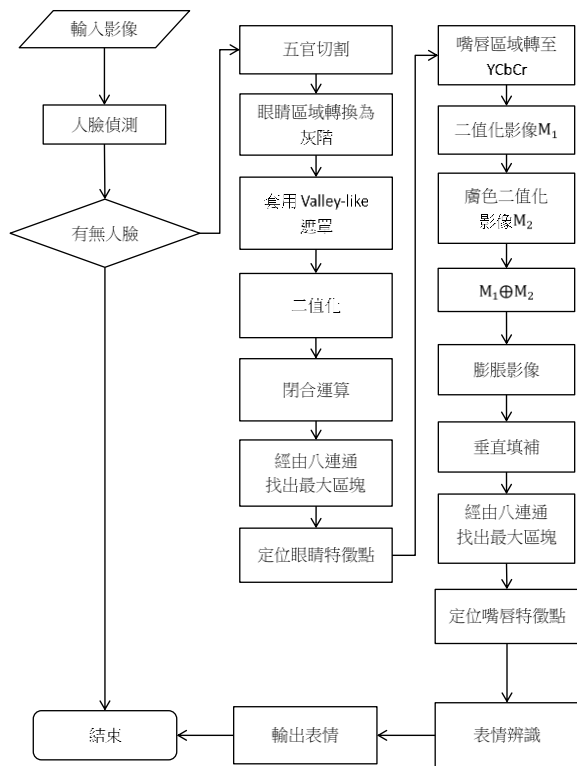


圖 3-7. 表情辨識整體流程圖

4. 實驗結果

本文實作一個整合人臉辨識及人臉表情辨識的行動裝置應用程式，本章節將針對人臉辨識及表情辨識的各項實驗結果進行分析與探討。

4.1 測試環境與設備

本文之應用程式實做於 Android 之上，本文選 ASUS Transformer Pad 做為測試之設備。

4.2 人臉資料庫

本文實驗使用兩種人臉資料庫為實驗的樣本，分別為 Frontal face dataset 及 Yale 人臉資料庫，由於本文表情辨識演算法需要再全彩的影像下才可運作，Yale 人臉資料庫僅提供灰階影像，因此選用兩種不同的人臉資料庫，作為實驗時的樣本資料。

Frontal face dataset 此資料庫是由 Markus Weber 匯集而成，當中包含 450 張人臉影像，是由 27 位不同的人所組成，且每張圖片皆有不同的光照、表情及背景，每張格式皆為 896x592 像素的 Jpeg 圖檔。

Yale 資料庫總共有 165 張影像，其中包含 15 個人，每個人各 11 張影像，這 11 張影像包括有光線的變化、臉部的遮蔽物(如眼鏡)及表情的變化，每一張影像的大小皆為 320x243 像素。

4.3 人臉辨識實驗結果與分析

本文將人臉辨識實驗分為三個部分，分別為資料庫內的人數、訓練的人臉數及選取的特徵向量數，根據以上實驗的結果加以分析。

圖 4-1 是為了測試資料庫的人臉數量是否會影響辨識的結果，此實驗以 Yale 資料庫做為樣本資料，使用資料庫內前 3 張影像做為訓練之用，後 3 張做為測試用，並且保留前 90% 的特徵向量，實驗結果如圖所示，在資料庫中只有 5 個人時，辨識率可達到 100%，但是隨著人數的增加，可以發現辨識率有著顯著的下降，在人數為 13 的時候，辨識率降低到 84% 左右，證明辨識率與資料庫中人數有著明顯的關係。

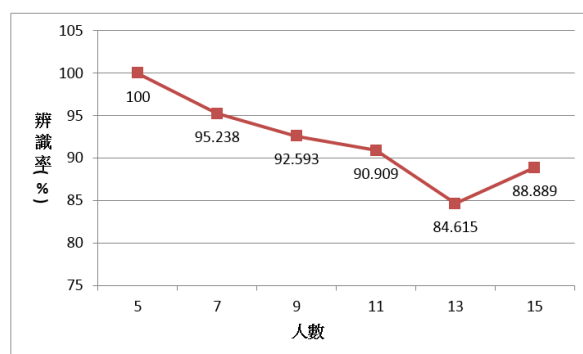


圖 4-1. 資料庫人數與辨識率之相關圖

圖 4-2 實驗以 15 個不同的人，並選用前 90% 的特徵向量，如圖 4 所示，當訓練張數只有 1 張時，辨識率僅僅只有 23.33%，但隨著訓練的人臉數越多，辨識率也會隨之增加，大約到達 3 張人臉的時候，辨識率就沒有顯著的提升

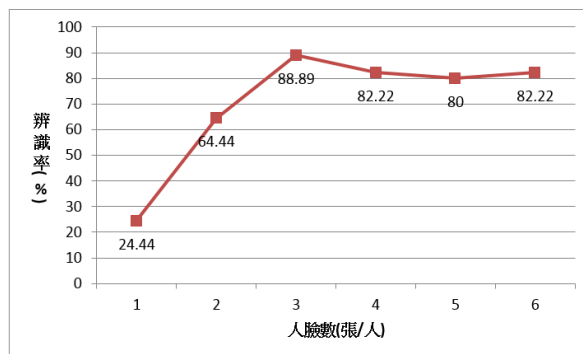


圖 4-2. 人臉數與辨識率之相關圖

圖 4-3 實驗藉由改變特徵向量，來檢測辨識率是否有所變化，以 15 個不同的人及各 3 張人臉影，總共 45 張人臉影像作資料庫的訓練，實驗結果如圖 4-5 所示，從實驗結果選取的特徵向量愈多，得到的辨識率也會是愈高，但是相對的，選取愈多的特徵向量，所以需要運算的時間也會相對地變長。

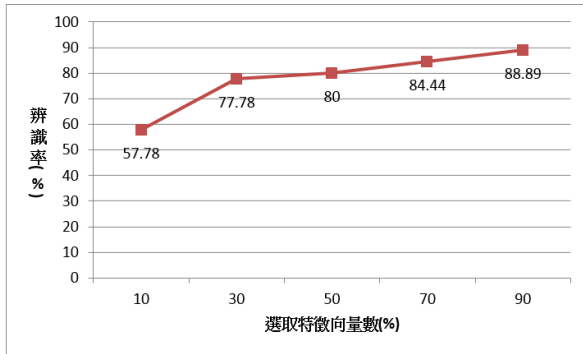


圖 4-3. 徵向量與辨識率之相關圖

4.4 表情辨識實驗結果與分析

以本文提出之表情辨識演算法進行實驗，實驗結果如下如圖 4-4 所示：

輸出 輸入	一般	高興	驚訝	閉眼	悲傷	辨識率 (%)
一般	43	2	3	6	0	79.6
高興	3	23	2	2	0	76.7
驚訝	3	4	14	2	0	60.9
閉眼	3	2	0	17	0	77.3
悲傷	3	1	0	2	11	64.7

圖 4-4. 期辨識結果

結果中可見，整體辨識率僅 71.8% 左右，經過分析每一張辨識錯誤的影像，是在嘴唇與眼睛特徵點定位錯誤，而造成錯誤，以其中一張錯誤影像為例，圖 4-5(A) 原始演算法所偵測出的結果，眼睛上面與下面的邊緣是分離的，因此經由八連通運算後，紅色的區域就是偵測到的區域，與預期的結果不相符，這些錯誤影像，多半是在眼角及嘴角連接處發生問題，眼角及嘴角容易受到光線的影響而造成陰影，所以本文在原始的偵測結果之後垂直填補演算法，將原本分離的角落的區域連結，圖 4-5(B) 加入填補演算法後的結果，眼睛區域已經正確地被偵測出來，經過修正後的演算法的實驗結果，如圖 4-6。



(A)

(B)

圖 4-5. 眼偵測結果 (A) 原始演算法 (B) 加入垂直填補演算法

輸出 輸入	一般	高興	驚訝	閉眼	悲傷	辨識率 (%)
一般	47	4	2	1	0	87
高興	3	25	2	0	0	83.3
驚訝	3	1	19	0	0	82.6
閉眼	2	3	0	17	0	77.3
悲傷	3	1	0	1	12	70.6

圖 4-6. 用垂直填補之表情辨識數據

從實驗結果可以得知，整體的辨識率可以到達 80.2%，但是在閉眼及悲傷的表情時的辨識率低於平均辨識率，所以我們針對這兩個表情深入分析，閉眼的表情因為判斷的條件較少，眼睛是否閉上是主要的判斷依據，分數權重也最重，但在嘴巴部分，條件容易與其他表情混在一起，在判斷上容易有誤判的產生，而悲傷的部分，造成誤判則是在嘴巴部分，在定位嘴巴特徵點時，會因為光線的問題造成嘴角的陰影較深，由於基於膚色偵測的緣故，陰影較深的區域會在偵測的範圍外，定位出的位置會跟一般的嘴型相似，進而將其判斷為一般表情。

本文提出一個整合人臉辨識與表情偵測的學習管理系統，系統登入時使用 PCA 方法進行人臉辨識，由實驗數據可見，辨識率達到 85% 上下，以此辨識得到的結果進行登入的動作，並於登入時同步進行點名，之後利用行動裝置的前置攝影機，定時抓取影像，找出這張影像中的人臉區域，並將區域內的五官做切割，根據本文提出的演算法，定位出五官的特徵點，一共 12 個點，依據這 12 點之間的相對位置及斜率算出個表情的權重比，根據這個權重比判斷出相對應的表情，可達到 80.2% 的辨識率。

未來的研究方向，可以著重在以下幾個重點，首先是在辨識率的方面，單獨使用 PCA 進行辨識，資料庫內人數如果愈來愈多，辨識率會有下降的趨勢，如果加入其他的演算法輔助，例如類神經網路、SVM 等等...，在辨識率方面應該能有顯著的提升，在表情辨識的部分，本系統僅能辨識出五種情緒反應，但是人類的情緒是很複雜的，不只有這五種表情，未來可以從特徵點下手，將特徵點的數量增加，有了更多的特徵點，定義的條件就能夠增加，能夠辨識的表情也就能隨之增加，最後，由於本系統是建築在 Android 的行動裝置上，就現階段而言，行動裝置的效能還是與一般家用型電腦有差距，因此，如何將演算法優化，也是未來重要的課題，能夠在行動裝置上運行越複雜的演算法，也就代表能夠處理更多的數據，對辨識率也能夠有效地提升。

參考文獻

- [1] Canny, John, "A Computational Approach to Edge Detection," Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol.PAMI-8, no.6, pp.679,698, Nov. 1986.
- [2] Ming-Hsuan Yang; Kriegman, D.; Ahuja, N., "Detecting faces in images: a survey," Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol.24, no.1, pp.34,58, Jan 2002. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271-350.
- [3] H.P. Graf, T. Chen, E. Petajan, and E. Cosatto, "Locating Faces and Facial Parts," Proc. First Int'l Workshop Automatic Face and Gesture Recognition, pp. 41-46, 1995.
- [4] Augustejjn, M.F.; Skufca, T.L., "Identification of human faces through texture-based feature recognition and neural network technology," Neural Networks, 1993., IEEE International Conference on , pp.392,398 vol.1, 1993.
- [5] Viola, P.; Jones, M., "Rapid object detection using a boosted cascade of simple features," Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on , vol.1, no., pp.1-511,1-518 vol.1, 2001.
- [6] Lienhart, R.; Maydt, J., "An extended set of Haar-like features for rapid object detection," Image Processing. 2002. Proceedings. 2002 International Conference on , vol.1, no., pp.1-900,1-903 vol.1, 2002.
- [7] 薛傑仁, "生物辨識之人臉辨識的方法", 亞洲大學生物與醫學資訊研究所碩士論文, 2010年7月。
- [8] Mallipeddi, R.; Minh Lee, "Ensemble based face recognition using discriminant PCA Features," Evolutionary Computation (CEC), 2012 IEEE Congress on , pp.1,7, 10-15 June 2012.
- [9] Ebied, H.M., "Feature extraction using PCA and Kernel-PCA for face recognition," Informatics and Systems (INFOS), 2012 8th International Conference on , pp.MM-72,MM-77, 14-16 May 2012.
- [10] Kotsia, I.; Pitas, I., "Facial Expression Recognition in Image Sequences Using Geometric Deformation Features and Support Vector Machines," Image Processing, IEEE Transactions on , vol.16, no.1, pp.172,187, Jan. 2007.
- [11] P. Ekman and W. V. Friesen, "The facial action coding system: a technique for the measurement of facial movement," San Francisco: Consulting Psychologists Press, 1978.
- [12] Ying-Li Tian; Kanade, T.; Cohn, J.F., "Evaluation of Gabor-wavelet-based facial action unit recognition in image sequences of increasing complexity," Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on , vol., no., pp.229,234, 21-21 May 2002.
- [13] Hao Tang; Huang, T.S., "3D facial expression recognition based on automatically selected features," Computer Vision and Pattern Recognition Workshops, 2008. CVPRW '08. IEEE Computer Society Conference on , vol., no., pp.1,8, 23-28 June 2008.
- [14] Khanam, A.; Shafiq, M.Z.; Akram, M.U., "Fuzzy Based Facial Expression Recognition," Image and Signal Processing, 2008. CISP '08. Congress on , vol.1, no., pp.598,602, 27-30 May 2008.
- [15] 鐘仁厚, "基於模糊邏輯之臉部表情辨識", 碩士論文, 國立中央大學, 2008。
- [16] 許晉懷, "結合影像及語音之雙模情緒辨識系統", 碩士論文, 國立交通大學電機與控制工程學系, 2006。
- [17] 蘇芳生, "人臉表情辨識系統", 碩士論文, 國立中正大學通訊工程研究所, 2004。
- [18] Li Song Jin. "A method of face detection using geometrical structure of human face", Kuahakwontongbo(Bulletin of natural science of Academy of Science of D. P. R of Korea) 51(1) 21-26.
- [19] Jianke Li; Baojun Zhao; Hui Zhang, "Face Recognition Based on PCA and LDA Combination Feature Extraction," Information Science and Engineering (ICISE), 2009 1st International Conference on , pp.1240,1243, 26-28 Dec. 2009
- [20] Andrea F, Abate, Michele Nappi, Daniel Riccio, Gabriele Sabatino, "2D and 3D face recognition: A survey", ScienceDirect Pattern Recognition Letters , vol. 28, 2007 pp. 1885-1906.
- [21] Jelsovka, D.; Hudec, R.; Breznan, M., "Face recognition on FERET face database using LDA and CCA methods," Telecommunications and Signal Processing (TSP), 2011 34th International Conference on , pp.570,574, 18-20 Aug. 2011.
- [22] Xiao-Ning Zhang, Z.; Jue Jiang; Zhi-Hu Liang; Chun-Liang Liu, L., "Skin color enhancement based on favorite skin color in HSV color space," Consumer Electronics, IEEE Transactions on , vol.56, no.3, pp.1789,1793, Aug. 2010.