行政院國家科學委員會專題研究計畫 成果報告

人臉辨識及於情緒辨識之分析設計應用
研究成果報告(精簡版)

計 畫 類 別 ： 個別型
計 畫 編 號 ： NSC 99-2221-E-216-022-
執 行 期 間 ： 99 年 08 月 01 日至 100 年 07 月 31 日
執 行 單 位 ： 中華大學電機工程研究所

計 畫 主 持 人 ： 駱樂

計畫參與人員： 碩士班研究生-兼任助理人員：黃湧益

處 理 方 式 ： 本計畫涉及專利或其他智慧財產權，2 年後可公開查詢

中 華 民 國 100 年 10 月 26 日

行政院國家科學委員會補助專題研究計畫 ☑ 成 果 報 告 □期中進度報告

人臉辨識及於情緒辨識之分析設計應用

計畫類別：☑個別型計畫　□整合型計畫

計畫編號：NSC　99－2221－E－216－022

執行期間： 99年08月01日至100年07月31日

執行機構及系所：中華大學

計畫主持人：　駱　樂

共同主持人：

計畫參與人員：黃湧益

成果報告類型(依經費核定清單規定繳交)：☑精簡報告　□完整報告

本計畫除繳交成果報告外，另須繳交以下出國心得報告：

□赴國外出差或研習心得報告

□赴大陸地區出差或研習心得報告

□出席國際學術會議心得報告

□國際合作研究計畫國外研究報告

處理方式：除列管計畫及下列情形者外，得立即公開查詢

☑涉及專利或其他智慧財產權，□一年□二年後可公開查詢

中　華　民　國 100 年 07 月 31

人臉辨識及於情緒辨識之分析設計應用

主持人： 駱 樂

執行機構： 中華大學電機工程學系

執行期間： 民國 99 年 08 月 01 日至 100 年 07 月 31 日

國科會計畫編號：NSC 99-2221-E-216-022

摘 要

人臉表情與情緒的自動分析，近幾年已廣泛於心理學、電腦科學、語言學、神經科學等領域展開相關研究與探討。本計劃研發出嶄新的理論方法，以利辨識人臉六種基本靜態影像之情緒或表情。計畫中首先精進修改現有的膚色偵測法，使能於複雜背景中擷取人臉。同時根據中國面相術與心理學觀念，創新於人臉中定義關鍵的表情特徵點，以取代現今被廣泛使用耗時的 FACS 特徵擷取方式。再藉由特徵點變動偵測來量化情緒狀態的改變。然後於情緒分類部分，創新採用高斯隨機混合模型 Gaussian Mixture Model (GMM)建立人臉情緒數學模型，以獲得足夠抗影像誤差與雜訊的強健性。本計劃所研發出的技術，將同時以實驗室資料庫以及著名的 JAFFE 資料庫來驗證。經分析與實驗，人臉情緒平均辨識率可達 83%，同時整套人臉情緒辨識系統以 2007a Matlab 軟體完成，在硬體 Pentium 2.33Ghz 平台，可 0.01 秒內瞬間完成情緒辨識。由實驗結果與分析，証明本計劃所研發出的技術與策略，可有效辨識人臉六種基本情緒，獲得較優於一般情緒辨識相關的方法。

關鍵字：情緒辨識、面相術、高斯隨機混合模型

Facial recognition and its applications on emotion analysis and design

Leh Luoh

Department of Electrical Engineering,
Chung Hua University

2010/08/01~2011/07/31
NSC 99-2221-E-216-022

Abstract

1. Automated analysis of human affective behavior (or facial expressions) has recently been extensively researched in psychology, computer science, linguistics, neuroscience, and other related disciplines. This paper presents a novel approach to recognizing six universal human facial expressions from static images. First, a modified skin color detection process is adopted to pick up the facial region from a complicated background. Then, fourteen crucial facial features, based on face reading in psychology, are further defined, instead of relying on relatively time-consuming FACS representation in extracting facial features. Feature motion detection follows to quantize the characteristic values of the emotion states. In further emotion classification, the statistical Gaussian Mixture Model (GMM) is especially introduced with an attempt to gain robustness to noise. Finally, the proposed systems are verified by our comprehensive databases along with the famous Japanese female facial expression (JAFFE) : the average recognition rate is as high as 83% and the discriminative time is less than 0.01 sec by 2007a Matlab software (Pentium 2.33Ghz). The experiment has shown that the proposed strategy is effective, and generating superior results as compared to some other approaches.

***Index Terms*—Emotion recognition, Face reading, Gaussian Mixture Model**,

1. Introduction

Automated analysis of human affective behavior (or facial expressions) has attracted increasing attention from researchers in psychology, computer science, linguistics, neuroscience, and related disciplines. Moreover, it would be highly beneficial for fields as diverse as behavioral science, medicine, security, law enforcement, education, psychiatry, telecommunications (such as lip reading, visual and speech synthesis, videoconferencing, and human-machine interfaces, *etc*).

The review of facial expression analysis should go back to the nineteenth century. Darwin demonstrated already in 1872 the universality of facial expressions and their continuity in man and animals. In 1971, Ekman and Friesen postulated six primary emotions in which each possess distinct content. It relates to ethnicities and cultures, and it comprises happiness, sadness, fear, disgust, surprise and anger [35]. In 1978, Suwa *et al*. has presented a first investigation on automatic facial expression analysis with an image in sequence. Thereafter, automatic facial expression analysis has gained much attention after the pioneering work of Pentland and Mase [21],[32]. In a human society, relationships with friends, colleagues and family (Ekman, 2004) are carefully maintained by the ability to understand, interpret and react to emotions. Emotions positively affect intelligent functions such as decision making, perception and empathic understanding (Bechara, Damasio, & Damasio, 2000; Isen, 2000). The automated facial expression recognition system can be divided roughly into three stages, *i.e.*, face detection, features extraction, and facial expression classification. The image of emotion recognition can also be categorized to static and sequence image. Usually, the method of recognizing a single image can also be used to recognize the dynamic image. A continuous image sequence can join those temporal characteristics while recognizing the relation between the front and the back of single image.

In face detection stage, generally, it mainly uses template matching, feature based, Principal Component Analysis (PCA), color analysis, neural network and evolutionary computation, *etc*. The method using template matching and characteristic [2-6] need a great deal of computing. The template and feature method may diverge for different people, and the corresponding model is not easy to build up. The color analysis uses different color space, like HSV, YCbCr, *etc.* Some use specific neural network to separate skin color from background. However, it also needs a great deal of training data [8][9][10][15]. In face features extraction stage, it can be divided into the method of geometric feature-based and the wavelet [16-19], [20-26]. Geometric feature-based approach adopt facial feature model (*e.g.*, eyes); or by using of the relation of locations to extract characteristics, such as eyebrow *vs.* eyes, *etc.* For holistic approach, a 3D facial representation may even be adopted with the corresponding texture and a spatial-temporal facial motion in the study. It also can use the wavelet to transform the image into the frequency domain, and then find out the characteristic, *e.g.*, Gabor Wavelet [23-26]. However, even it gains fruitful benefits, the process might be a little bit complex. Several factors make this task more difficult, such as the presence of facial hair, glasses, *etc*.; another problem is the variation of the image size and orientation of the face.

In expression classification stage, Ekman and Friesen have constructed the Action Unit (AU) with Facial Action Coding System (FACS) [35]. Among them, each AU represents a face department of an action. For instance, the rising in the seamy side of eyebrow is an action unit, the rising in right corner of mouth is another action unit, *etc*. [35] has defined 44 action units to evaluate a facial expression. FACS is probably the most well-known study on facial expression activity. In past few years, many questions arose around this study[35]. However, the automatic emotional classification is difficult to handle for number of reasons. Firstly, the Ekman's description of six prototypic of emotion is linguistic, *i.e.*, ambiguous. There is no unique definition or exact description either in terms of facial actions or in terms of some other defined facial codes. Hence, the verification of the classification scheme is a difficult and crucial work. Second, the classification of facial expressions should be feasible. Meanwhile, for more credible, the psychological scrutiny should be involved in this topic.

This project provides a systematic approach to analyze Ekman and Friesen's six primary facial expression of human in colorful background, including happiness, sadness, surprise, anger, disgust and fear. Instead of relying on relative time-consuming FACS representation of facial features, emotion detection is derived only from fourteen feature points. The fourteen crucial feature points are defined based on the psychology's face reading. Through monitoring the change of these feature points, the characteristic values of the emotion states are quantitatively computed. For gaining robust to noise, a statistical Gaussian Mixture Model (GMM) is built for further emotion classification. The main contributions of this project are as follows: (1) A fourteen crucial feature points is defined according to psychology's face reading, (2) A robust emotion classification system based on GMM is proposed to recognize the emotion states.

2. Preliminary

It is known that the background of human face recognition system may be complicated. Generally, the input image can be acquired from a CCD or a digital camera. A typical recognition system would comprise of the three essential stages, *i.e.*, face detection, features extraction, and expression classification. The process of extracting the facial expression information is referred to localize the face and its features in the scene. A model should be also given to describe the characteristics of a specific emotion. Our proposed emotion recognition system is shown as illustrated in Fig. 1.

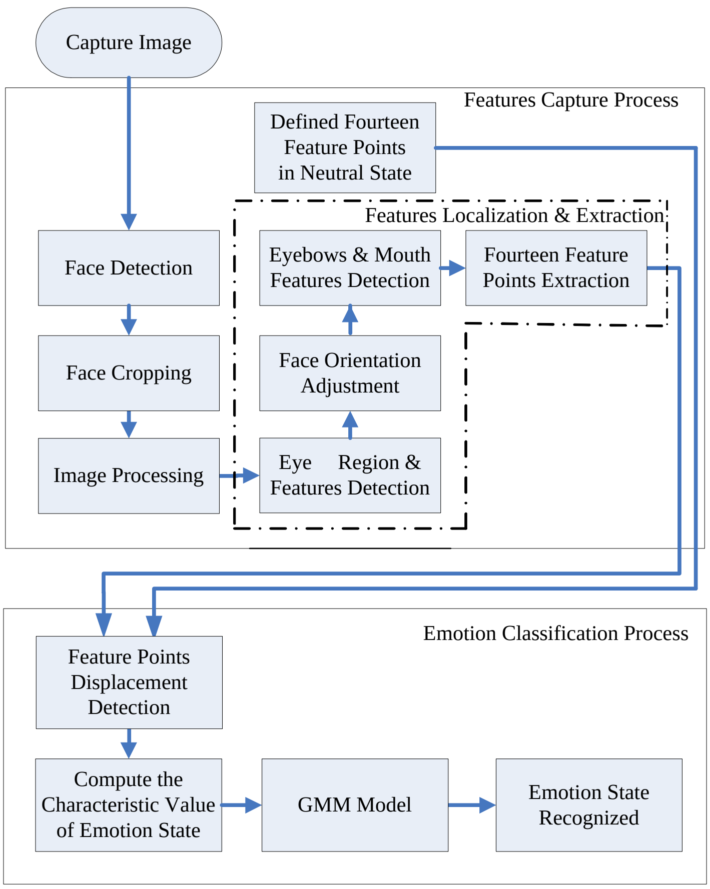

Fig.1 The proposed face emotion recognition system

2.1.Face Detection

For works in automatic facial expression analysis as Fig.1, the conditions under which a facial image is obtained are controlled. Usually, the image has the face in frontal view. Hence, the presence of a face in the scene is ensured and some global location of the face in the scene is known in priori. The face detection process is proposed as Fig. 2.

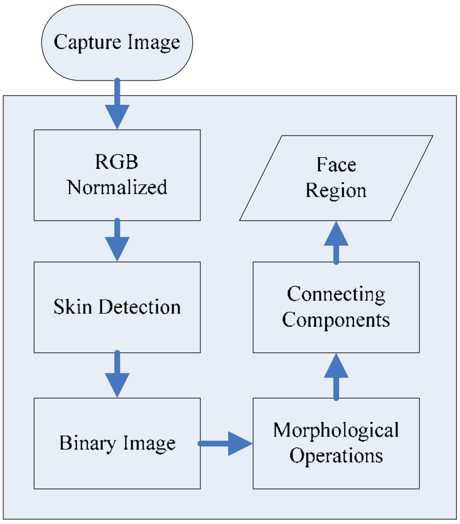

Fig.2 The process of face detection

For skin color detection, Soriano's RGB normalization method [27] not only considers the camera property, but the light of environment. The normalized color coordinates (NCC) is used to reduce the color's dependence on brightness. That is

$$r = \frac{R}{R+G+B} \quad , \qquad g = \frac{G}{R+G+B} \tag{1}$$

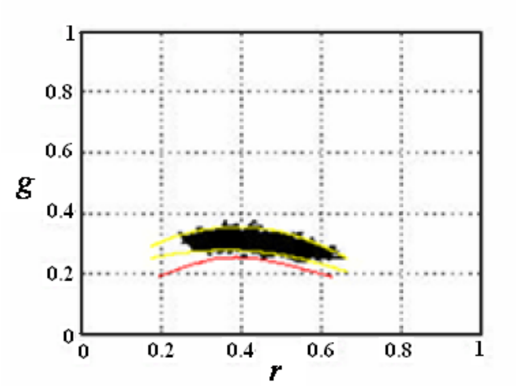

Fig.3 Normalized color coordinates

In NCC (Fig.3), r and g represent X-axis and Y-axis, respectively. Since the skin color is distributed on the range of the 0.2~0.6 of X-axis and 0.2~0.4 of Y-axis, respectively. Two quadratic equations F_1 and F_2 are tuned as boundary.

$$F_1(r) = -1.376r^2 + 1.0743r + 0.2 \tag{2}$$

$$F_2(r) = -0.776r^2 + 0.5601r + 0.18 \tag{3}$$

Owing to white color(r=0.33 and g=0.33) also is included in skin range, criteria (5) is chosen in similar way to remove the region of white color, *i.e.*,

$$w = (r - 0.33)^2 + (g - 0.33)^2 > 0.001 \tag{4}$$

and the range of the skin color is defined as

$$skin = \begin{cases} 1 & \text{if } (F_1(r) < g < F_2(r) \,\&\, w > 0.001 \,\&\, R > G > B \,\&\, (R - G) \geq 5) \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where conditions $R > G > B$ and $R - G > 5$ are used to detect skin color and exclude yellow or green noise of non-skin part, respectively. After converting to binary image, morphology skill is generally used to remove the noise [14]. Moreover, for excluding wrong skin parts (arms, *etc.*), the connected component labeling is in sequel to take out the face region [2].

2.1.Feature points definition and extraction

A. Feature points definition

After the face has been detected, the next step is to extract the emotion information of the face. The famous FACS rule has provided a linguistic description of all detectable change of face in terms of 44 so-called action units (AU)[35]. Various analysis of emotion has been developed. An effective ways for automatic emotion recognition is still anxious to be explored. This paper first gives some modified emotion descriptions originated from Ekman. Especially, for future embedded system or intelligent context-aware space implementation possible, this paper defines directly some crucial feature points related to emotion states from face reading in psychological view. Ekman has defined six categories as the basic emotions *i.e.*, happiness, sadness, surprise, fear, anger, and disgust. He described each basic emotion in terms of a facial expression that uniquely characterizes the emotion [24,26,31,35]. Instead of rule of FACS in terms of 44 action units, based on Table 1, this paper defines crucial fourteen feature points as Fig.4. Among them, two points for each eyebrow, three points for each eye, and four points for mouth. The detail features localization and extraction process is depicted in following subsection.

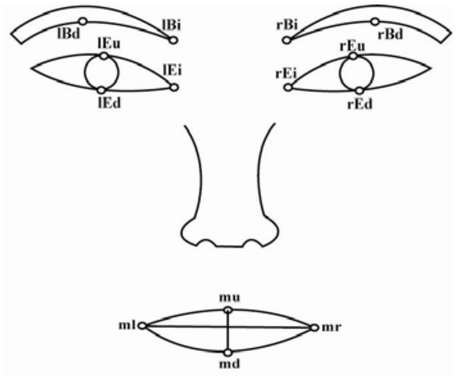

Fig.4 The proposed crucial fourteen feature points

B. Face orientation adjustment

As illustrated in Fig.1, in this paper, the emotion states are extracted according to the change of relative position of the feature points. The orientation of face needs to be adjusted appropriately before getting emotional information. Since the feature point **lEi** and **rEi** are always in horizontal (see Fig.5). The face orientation can be calculated easily by $\theta = \tan^{-1}(y/x)$, where θ is the orientation angle to be modified.

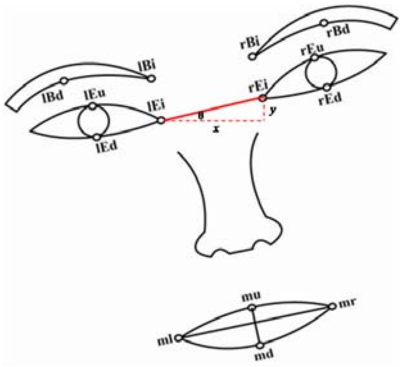

Fig.5 The face orientation adjustment

C. Features localization & extraction

It is well-known that determining the exact features location of face is a complex problem. However, the features characterization and their geometrical relationship are useful for feature points localization and extraction.

-Eyes Detection

As defined in prior, there are three feature points of eyes for each. The extraction process of eyes is shown as Fig. 6.

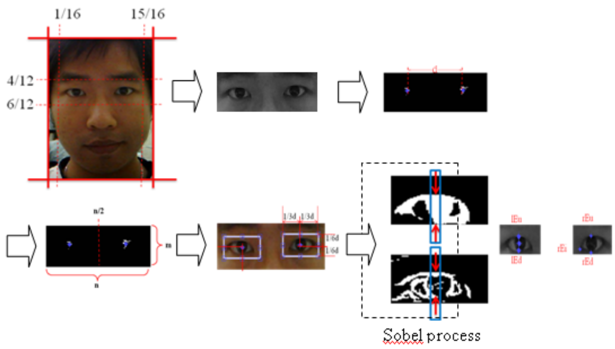

Fig. 6 The features extraction process of eyes

3. Facial emotion classification

3.1 The definition of characteristic value

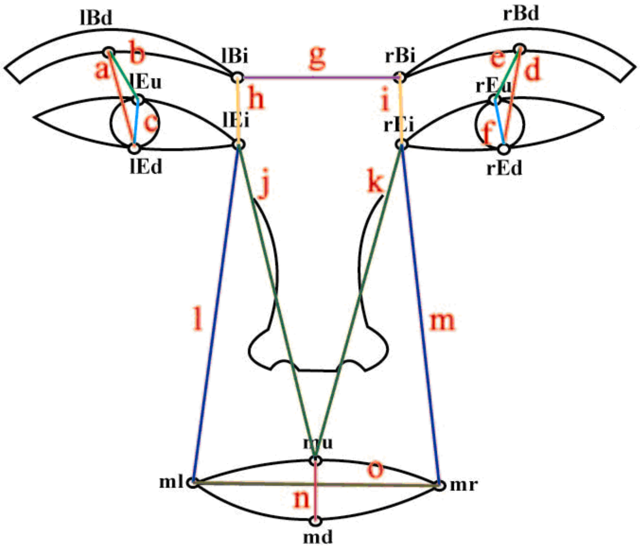

Fig. 7 The defined distance of fifteen features

According to those defined characteristic points in Fig. 4, we can compute the corresponding characteristic value (Fig.7), which represents the changing of emotion. Let the x-y coordination of each characteristic point be denoted as Suffix. For instance, the x-y coordination of lBu is (lBu_x, lBu_y). Accordingly, those defined characteristic values $(a \sim o)$ that represent the changing of emotion for the corresponding characteristic point can be computed be using the distance measure.

For instance, $a = \sqrt{\left|lBu_x - lEd_x\right|^2 + \left|lBu_y - lEd_y\right|^2}$. With the similar computation, the others can be also derived.

Owing to a the similar photo may take from different position of camera, the dimension of face may is not be the same. Therefore, a normalized process is needed here. Moreover, we have to use a invariant distance that will not fluctuate in the process. The distance g between point lEi and rEi seems a good choose. Therefore, those 15 characteristic values can be normalized by invariant distance g. For example, $\bar{a} = a/g$, the others $\bar{b} \sim \bar{o}$ are the same. Accordingly, the changing of facial expression can be estimated. Firstly, let us select a base photo with $\bar{a}_{base}, \cdots \bar{o}_{base}$ that hasn't any facial expression. Now, suppose a facial expression with $\bar{a}, \cdots, \bar{o}$ is captured in an time instance. Then the 15 characteristic values can be calculated. That is, subtract those captured values $\bar{a}, \cdots, \bar{o}$ from the value $a'_{base}, \ldots o'_{base}$ of base image, respectively. For instance, $\vec{F}_a = \bar{a} - \bar{a}_{base}$. Accordingly, the other normalized characteristic values $\overline{F}_b, \overline{F}_c, \cdots \overline{F}_o$ can be estimated.

3.2 GMM based facial affective states classification

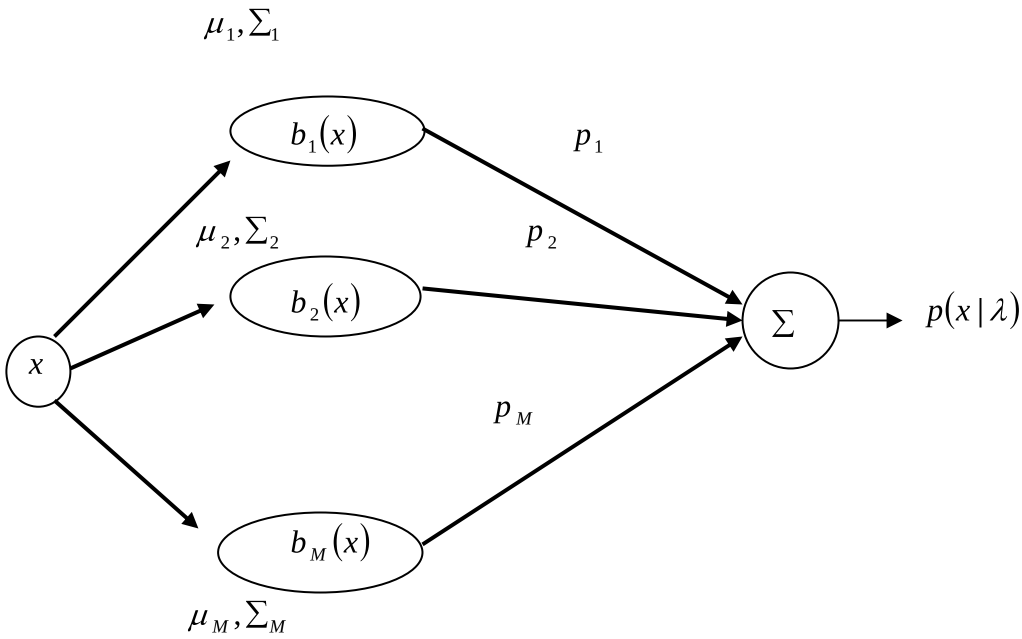

Fig. 8 Gaussian mixture density diagram

This section describes the form of the Gaussian mixture model (GMM) and motivates its use as a representation of human affective states for emotion identification. Gaussian mixture model (GMM) is the extension of single gauss probability density function. It is successfully using in some classifications problem recently. The Gaussian mixture model and its parameterization are described. The use of the Gaussian mixture density for emotion identification is then motivated by two interpretations. First, the individual component gaussians in an emotion-dependent GMM are interpreted to represent some broad facial expression classes. Second, a gaussian mixture density is shown to provide a smooth approximation to the specific facial expression by a given person. Finally, the maximum-likelihood parameter estimation and emotion identification procedures are first described.

A. Model description

A Gaussian mixture density is a weighted sum of M component densities, as depicted in Fig. 8 and given by the equation Gaussion mixture density:

$$p(\vec{x} \mid \lambda) = \sum_{i=1}^{M} p_i b_i(\vec{x}) \tag{6}$$

where $\vec{x}$ is a D-dimensional random vector, $b_i(\vec{x})$, $i = 1,\ldots,M$, are the component probability densities function and p_i, $i = 1,\ldots,M$, are the mixture weights. Each component density is a derivate Gaussian function of the form

$$b_i(\vec{x}) = \frac{1}{(2\pi)^{D/2} |\Sigma_i|^{1/2}} \exp\left\{-\frac{1}{2}(\vec{x} - \vec{\mu}_i)' \Sigma_i^{-1} (\vec{x} - \vec{\mu}_i)\right\} \tag{7}$$

with mean vector $\vec{\mu}_i$ and covariance matrix Σ_i; The mixture weights satisfy the constraint that $\sum p_i = 1$. The complete Gaussian mixture density is parameterized by the mean vectors, covariance matrices and mixture weights from all component densities. These parameters are collectively represented by the notation

$$\lambda = \{p_i, \vec{\mu}_i, \Sigma_i\} \quad i = 1, \ldots, M. \tag{8}$$

For six facial affective states identification, each facial affective state is represented by a GMM and is referred to by his/her model λ. The linear combination of diagonal covariance Gaussians is capable of modeling the correlations between feature vector elements. That is, given training facial emotion from a person, the goal of facial emotion model training is to estimate the parameters of the GMM, λ, which in some sense best matches the distribution of the training feature vectors. There are several techniques available for estimating the parameters of a GMM [24]. By far the most popular and well-established method is maximum likelihood (ML) estimation. The aim of ML estimation is to find the model parameters which maximize the likelihood of the GMM, given the training data. For a sequence of T training vectors $X = \{\vec{x}_1 \ldots, \vec{x}_T\}$, the GMM likelihood can be written as

$$p(X \mid \lambda) = \prod_{t=1}^{T} p(\vec{x}_t \mid \lambda) \tag{4}$$

5. Experiment Result

There is two database that used in this project, the JAFFE(the Japanese Female Facial Expression) database([23][25]), and the database from our Lab. We use thirty different person for experiment, and each person has twenty face images in different background.

	高興	悲傷	驚訝	生氣	嫌惡	害怕	正確率
高興	17	3	0	0	0	0	85%
悲傷	1	19	0	0	0	0	95%
驚訝	0	0	17	0	0	3	85%
生氣	0	2	0	13	3	2	65%
嫌惡	0	0	0	2	16	2	80%
害怕	1	2	0	0	0	17	85%
Average							83%

6. Conclusion

This paper has provided a system method to recognize a facial expression automatically. The experiment result also shows the facial emotion recognition rate for the well-known JEFFE database can be attained to 82.5 % for the database of LAB.

7. Acknowledgment

The authors express especially the gratitude to the foundation support by NSC 99－2221－E－216－022.

8. Reference

[1]. M. Bartlett, J. Hager, P.Ekman, and T. Sejnowski. “Measuring facial expressions by computer image analysis,” *Psychophysiology*, 36:253–264, 1999.

[2]. M. J. Black, and Y. Yacoob, “Recognizing facial expressions in image sequences using local parameterized models of image motion,” *International Journal of Computer Vision*, 25(1):23–48, October 1997.

[3]. R.Brunelli and T.Poggio , “ Face recognition : features vs. template,” *IEEE Trans. Pattern Analysis and Machine Intelligence, vol.1 ,no. 10*, pp. 1042-1052,Oct.1993.

[4]. 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.

[5]. C. Garcia, G. Simandiris, and G. Tzirita, “A feature-based face detector using wavelet frames,” *Proc of Intern , workshop on very low bit coding*, pp. 71-76, Athens, October, 2001.

[6]. C. Garcia, and G. Tzirita, “Face detection using quantized skin color region merging and wavelet packet analysis,” *IEEE trans. Multimedia*, 1(3), pp. 264-277, 1999.

[7]. C. Garcia and Manolis Delakis, “A neural architecture for fast and robust face detection,” *Pattern Recognition, 2002. Proceedings. 16th International Conference on , Volume: 2 , 11-15* , pp. 44-47 ,Aug. 2002.

[8]. Rafael C. Gonzalez, Richard E. Woods, and Steven L. Eddins, “Digital Image Processing Using Matlab. ”

[9]. 黄志長，“以影像處理技術為基礎的人臉情緒辨識” ，畢業碩士論文.

.

附件二

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一综合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

☑ 達成目標

☐ 未達成目標（請說明，以 100 字為限）

　☐ 實驗失敗

　☐ 因故實驗中斷

　☐ 其他原因

說明：本計劃研發出嶄新的理論方法，以利辨識人臉六種基本靜態影像之情緒或表情。根據中國面相術與心理學觀念，以取代現今被廣泛使用耗時的 FACS 特徵擷取方式。同時創新採用高斯隨機混合模型 Gaussian Mixture Model (GMM)建立人臉情緒數學模型，所研發出的技術，將同時以實驗室資料庫以及著名的 JAFFE 資料庫來驗證。經分析與實驗，人臉情緒平均辨識率可達 83%

2. 研究成果在學術期刊發表或申請專利等情形：

論文：☐已發表 ☐未發表之文稿 ☑撰寫中 ☐無

專利：☐已獲得 ☑申請中 ☐無

技轉：☐已技轉 ☐洽談中 ☐無

其他：（以 100 字為限）

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以 500 字為限）

人臉表情與情緒的自動分析，近幾年已廣泛於心理學、電腦科學、語言學、神經科學等領域展開相關研究與探討。本計劃研發出嶄新的理論方法，以利辨識人臉六種基本靜態影像之情緒或表情。計畫中首先精進修改現有的膚色偵測法，使能於複雜背景中擷取人臉。同時根據中國面相術與心理學觀念，創新於人臉中定義關鍵的表情特徵點，以取代現今被廣泛使用耗時的 FACS 特徵擷取方式。再藉由特徵點變動偵測來量化情緒狀態的改變。然後於情緒分類部分，創新採用高斯隨機混合模型 Gaussian Mixture Model (GMM)建立人臉情緒數學模型，以獲得足夠抗影像誤差與雜訊的強健性。本計劃所研發出的技術，將同時以實驗室資料庫以及著名的 JAFFE 資料庫來驗證。經分析與實驗，人臉情緒平均辨識率可達 83%，同時整套人臉情緒辨識系統以 2007a Matlab 軟體完成，在硬體 Pentium 2.33Ghz 平台，可 0.01 秒內瞬間完成情緒辨識。由實驗結果與分析，証明本計劃所研發出的技術與策略，可有效辨識人臉六種基本情緒，獲得較優於一般情緒辨識相關的方法。此研發技術成果極有學術與實用價值，未來可應用於「智慧型機器人」與「智慧型空間系統」等相關議題與產業，包含人機互動、陪伴居家照顧、心理諮商、精神鑑別、測謊、教育、教學等許多人文與科技之廣泛層面。

國科會補助計畫衍生研發成果推廣資料表

日期:2011/10/26

國科會補助計畫	計畫名稱: 人臉辨識及於情緒辨識之分析設計應用
	計畫主持人: 駱樂
	計畫編號: 99-2221-E-216-022- 學門領域: 智慧型機器人

無研發成果推廣資料

99 年度專題研究計畫研究成果彙整表

計畫主持人：駱樂			計畫編號：99-2221-E-216-022-				
計畫名稱：人臉辨識及於情緒辨識之分析設計應用							
成果項目			量化			單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）
			實際已達成數（被接受或已發表）	預期總達成數(含實際已達成數)	本計畫實際貢獻百分比		
國內	論文著作	期刊論文	1	1	100%	篇	中華大學理工學刊 2010
		研究報告/技術報告	1	1	100%		國科會補助專題研究計畫成果報告
		研討會論文	1	0	100%		NSSSE 2010
		專書	0	0	100%		
	專利	申請中件數	0	1	100%	件	中華民國專利
		已獲得件數	0	0	100%		
	技術移轉	件數	0	1	100%	件	盛暘科技
		權利金	0	0	100%	千元	
	參與計畫人力（本國籍）	碩士生	1	1	100%	人次	碩士生
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	0	2	100%	篇	1. IEEE Trans on Image Processing 2. International Journal of Innovative computing, Information and Control
		研究報告/技術報告	0	0	100%		
		研討會論文	2	2	100%		1. 2010 International Conference on System Science and Engineering 2. 2010 International Conference on System Science and Engineering 兩篇
		專書	0	0	100%	章/本	
	專利	申請中件數	0	1	100%	件	美國
		已獲得件數	0	0	100%		

	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（外國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)		正與盛暘科技公司洽談產業技術發展與產品研製推廣合作					

	成果項目	量化	名稱或內容性質簡述
科教處計畫加填項目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與（閱聽）人數	0	

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

■達成目標

□未達成目標（請說明，以 100 字為限）

□實驗失敗

□因故實驗中斷

□其他原因

說明：

2. 研究成果在學術期刊發表或申請專利等情形：

論文：□已發表 □未發表之文稿 ■撰寫中 □無

專利：□已獲得 ■申請中 □無

技轉：□已技轉 ■洽談中 □無

其他：（以 100 字為限）

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以 500 字為限）

人臉表情與情緒的自動分析，近幾年已廣泛於心理學、電腦科學、語言學、神經科學等領域展開相關研究與探討。本計劃研發出嶄新的理論方法，以利辨識人臉六種基本靜態影像之情緒或表情。計畫中首先精進修改現有的膚色偵測法，使能於複雜背景中擷取人臉。同時根據中國面相術與心理學觀念，創新於人臉中定義關鍵的表情特徵點，以取代現今被廣泛使用耗時的 FACS 特徵擷取方式。再藉由特徵點變動偵測來量化情緒狀態的改變。然後於情緒分類部分，創新採用高斯隨機混合模型 Gaussian Mixture Model (GMM)建立人臉情緒數學模型，以獲得足夠抗影像誤差與雜訊的強健性。本計劃所研發出的技術，將同時以實驗室資料庫以及著名的 JAFFE 資料庫來驗證。經分析與實驗，人臉情緒平均辨識率可達 83%，同時整套人臉情緒辨識系統以 2007a Matlab 軟體完成，在硬體 Pentium 2.33Ghz 平台，可 0.01 秒內瞬間完成情緒辨識。由實驗結果與分析，証明本計劃所研發出的技術與策略，可有效辨識人臉六種基本情緒，獲得較優於一般情緒辨識相關的方法。此研發技術成果極有學術與實用價值，未來可應用於「智慧型機器人」與「智慧型空間系統」等相關議題與產業，包含人機互動、陪伴居家照顧、心理諮商、精神鑑別、測謊、教育、教學等許多人文與科技之廣泛層面。