

行政院國家科學委員會專題研究計畫 成果報告

多辨識器組合之強健式臉部表情辨識技術 研究成果報告(精簡版)

計畫類別：個別型
計畫編號：NSC 99-2221-E-216-050-
執行期間：99年08月01日至100年07月31日
執行單位：中華大學資訊工程研究所

計畫主持人：黃雅軒

計畫參與人員：碩士班研究生-兼任助理人員：莊順旭
碩士班研究生-兼任助理人員：陳冠豪
碩士班研究生-兼任助理人員：歐志鴻
碩士班研究生-兼任助理人員：林啟賢
碩士班研究生-兼任助理人員：吳東懋
碩士班研究生-兼任助理人員：李允善
碩士班研究生-兼任助理人員：張倞禕

報告附件：出席國際會議研究心得報告及發表論文

公開資訊：本計畫可公開查詢

中華民國 100 年 10 月 31 日

中文摘要： 本論文提出一種混合權重式區域方向特徵(WLDP)和二元化形態(LBP)的表情辨識方法。一開始 WLDP 和 LBP 會分別對人臉影像進行特徵抽取，接著利用 PCA 分別對 WLDP 和 LBP 所抽取出來的特徵進行特徵降維處理，最後將兩種特徵進行混合，產生出一種對人臉具有分辨能力的混合特徵，並使用 SVM 分類器來進行表情辨識。

實驗的資料庫是使用著名的 Cohn-Kanade 表情資料庫，該資料庫由於具有完整的表情資料，因此是表情辨識領域的研究學者們常用來進行研究的一套表情資料庫。本論文所提出的方法，對 Cohn-Kanade 表情資料庫進行 7 類表情辨識，並使用 10-fold person-independent cross-validation 架構，可以得到高達 91.1% 的辨識率，這顯示本論文所提出的權重式區域方向特徵，和辨識方法具有一定的效果。

英文摘要： A method of combining Weighted Local Directional Pattern (WLDP) and Local Binary Pattern (LBP) for facial expression recognition is proposed. First, WLDP and LBP are applied to extract human facial features. Second, principle component analysis (PCA) is used to reduce their feature dimensions respectively. Third, both reduced facial features are merged to form the final feature vector. Fourth, support vector machine (SVM) is used to recognize facial expressions. Experiment on the well known Cohn-Kanade expression database, a high accuracy rate up to 91.1% for recognizing seven expressions can be achieved with a person-independent 10-fold cross-validation scheme.

1. 前言

在人與人的交流行爲中，情緒表達佔據了相當重要的位置，可幫助人們了解對方的內心感受並讓自己能做出相對的反應；在各種的情緒表達行爲上，包含肢體動作(臉部表情與身體姿勢)、說話語調與用字遣詞等。實際上，人在互相交流的時候，口語，僅佔了 7%，而肢體動作則佔了 93%，其中又以臉部表情所佔最多，高達 55%，由此可見臉部表情對於人類在溝通時的重要性。因此若能讓機器人辨識使用者的情緒變化並做出反應，對於提升人機互動的親合力將有相當大的提升。

2. 研究目的

爲了提升人機的親和力，本計畫提出「多辨識器組合之強健式臉部表情辨識技術」的主題，我們將提出一套即時的自動化臉部表情辨識系統，包含有人臉偵測技術、臉部特徵點抽取技術、臉部特徵點追蹤技術與臉部表情辨識技術。在本計畫中，將對自然表情(即無表情)及六種基本情緒表情(高興、傷心、生氣、驚訝、厭惡與害怕)進行分析與辨識，正確的偵測出使用者的情緒表達，以做爲人機互動的重要資訊。

3. 文獻探討

爲了能使電腦認識人類的表情變化，已有許多學者投入大量的心力在電腦視覺研究上，並致力於能讓電腦理解人類的表情。在近幾年來，表情辨識已經成爲一個熱門的課題，因此有許多學者提出各種方法，來對表情進行辨識。表情辨識方法主要可以分成兩種，其中一種是使用Facial Action Coding System(FACS)[1][2][3]，另一種則是以特徵爲基礎(Feature-based)[24][5][6][7][8][9][10]的辨識方法。

FACS是Ekman和Friesen於1978年所提出用於人臉表情描述的編碼系統，在這套系統中，會依據人臉肌肉的分佈，以及一些肌肉群的運動狀況，定義出動作單元(Action Units)，每個動作單元表示臉部的特定區域的移動狀況，如眉毛上升、嘴角上揚等，共定義了44種動作單元(如圖一所示)，透過動作單元的組合，來進行表情判斷。Tian[2]等人發展出一套自動臉部分析系統(Automatic Face Analysis, AFA)，能依照人臉上永久或暫時性的特徵，對人臉正面影像序列進行分析，辨識出每種單獨的動作單元。Donato[3]等人發現使用GaborWavelet來進行特徵擷取，再進行上半部和下半部人臉的FAUs分類，比傳統的幾何方法可以達到更好的效果。

除了基於動作單元的表情辨識方法之外，也有基於紋理特徵等方法的表情辨識研究。Bartlett和Littlewort[4]等人將輸入的影像序列，偵測出正面的人臉位置，並經過Gabor Wavelet 擷取出紋理特徵，最後再使用一連串的SVM分類器來分類出7種不同的主要表情(包含自然、生氣、猶豫、恐懼、快樂、悲傷和驚訝)。Ma和Khorasni[5]則使用離散餘弦轉換(Discrete Cosine Transform)對整張影像進行特徵偵測和抽取，並使用前饋式類神經網路(Feedforward Neural Networks)來進行辨識。Dubuisson和Davoine[6]等人則先利用主成分分析法(Principal Component Analysis)和LDA(Linear Discriminant Analysis)進行前處理將影像降低維度後，再進行辨識。有些基於2D或3D(Model Template-Based)的表情辨識方法[7][8][9][10]，這些方法計算於3D模型中，特徵點的幾何變化或是對應於2D的紋理特徵變化，最後再經過辨識器，進行表情辨識。

參考文獻

1. 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.
2. Y-L. Tian, T. Kanade, J.F. Cohn, “Recognition action units for facial expression analysis”, IEEE Trans. Pattern Anal. Mach. Intell. 23(2) (2001) 87-115.
 3. G. Donato, M.S. Bartlett, J.C. Hager, P. Ekman, T.J. Sejnowski, “Classifying facial actions”, IEEE Trans.. Pattern Anal. Mach. Intell. 21(10)(1999) 974-985.
 4. M.S. Bartlett, G. Littlewort, I. Fasel, and J.R. Movellan, “Real time face detection and facial expression recognition: Development and applications to human computer interaction,” in Proc. Conf. Computer Vision and Pattern Recognition Workshop, Madison, WI, Jun. 16-22, 2003, vol. 5, pp. 53-58.
 5. L. Ma and K. Khorasani, “Facial expression recognition using constructive feedforward neural networks,” IEEE Trans. Syst., Man, Cybern. B, Cybern., vol. 34, no. 3, pp. 1588–1595, Jun. 2004.
 6. S. Dubuisson, F. Davoine, and M. Masson, “A solution for facial expression representation and recognition,” Signal Process.: Image Commun., vol. 17, no. 9, pp. 657–673, Oct. 2002.
 7. I. A. Essa and A. P. Pentland, “Facial expression recognition using a dynamic model and motion energy,” presented at the Int. Conf. Computer Vision, Cambridge, MA, Jun. 20–23, 1995.
 8. M. Pantic and L. J. M. Rothkrantz, “Expert system for automatic analysis of facial expressions,” Image Vis. Comput., vol. 18, no. 11, pp. 881–905, Aug. 2000.
 9. Irfan A. Essa, “Coding, analysis, interpretation, and recognition of facial expressions,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 7, pp.757–763, Jul. 1997.
 10. M. S. Bartlett, G. Littlewort, B. Braathen, T. J. Sejnowski, and J. R. Movellan, “An approach to automatic analysis of spontaneous facial expressions,” presented at the 5th IEEE Int. Conf. Automatic Face and Gesture Recognition, Washington, DC, 2002.

4. 研究方法

本計畫針對人臉影像進行表情辨識，期許能有良好的辨識效果。主要的研發技術包含臉部特徵點定位和臉部表情辨識等。本計畫所開發的主要技術成果有二：

(1) 臉部特徵點定位

動態形狀模型(Active Shape Model, ASM)已經成功的被應用在臉部特徵點定位上，然而當人們臉部有著誇張的臉部表情變化時，如：驚訝、大笑和挑眉等，其特徵點定位的結果仍有明顯的誤差率。爲了克服這個問題，我們提出串接式多階段的臉部特徵點定位方法。第一階段，我們使用 Adaboosting 學習演算法來定位臉部特徵點中較有明顯鑑別度之特徵點，而這些有明顯鑑別度的特徵點都是角點類型的地標點，如左右內眼角、左右外眼角、左右內眉角、左右外眉角和左右嘴角等共 10 個點。並於第一階段做此 10 個角點的動態形狀模型之重建與定位，來獲得較符合臉部幾何架構之臉部特徵點位置。第二階段，我們根據第一階段所定位的臉部角點位置，重新初始化臉部五官模型位置，再根據地標點變動的分布情形，制定出每個地標點的搜尋範圍，然後進行第二次的動態形狀模型的重建與定位。我們採用 BIOID 和 Cohn Kanade 二種人臉資料庫來進行測試實驗，傳統的定位誤差是 7.93%，而我們所提出的方法是 4.97%。明顯地，我們所提出串接式多階段的臉部特徵點定位方法具有更好的處理效果。下圖爲人臉特徵點定位處理結果的一些範例，可見得此技術可適用於多人種。經實驗分析，我們方法具有下列四項優點：

a. 地標點分類：

把地標點分成角點形態地標點與邊緣形態地標點並且使用不同的地標點搜尋方式。角點形態地標點，使用矩形的搜尋範圍，可以避免因正確地標點不在法線上所造成的誤差。

b. 使用 Adaboost 偵測器來偵測角點形態地標點：

Adaboost 偵測器對於有特殊紋理結構的物件能夠有效的偵測，再搭配所提出的候選點選擇方式，可以成功的定位角點位置。

c. 使用角點位置重新初始化五官形狀位置：

此方式可以大幅減少不同的臉部結構及臉部表情使用同一個平均形狀樣板所造成的初始位置誤差，因此可以大幅提升定位的準確度。

d. 不同的五官地標點使用不同的搜尋長度：

因為各地標點的分散程度不同，所以根據各五官地標點的分佈情形給予相對應的搜尋範圍，可以減少因搜尋範圍短產生的局部最佳解及搜尋範圍長所增加的誤判機率。

(2) 表情辨識

我們提出一種混合權重式區域方向特徵(WLDP)和二元化形態(LBP)的表情辨識方法。一開始 WLDP 和 LBP 會分別對人臉影像進行特徵抽取，接著利用 PCA 分別對 WLDP 和 LBP 所抽取出來的特徵進行特徵降維處理，最後將兩種特徵進行混合，產生出一種對人臉具有分辨能力的混合特徵，並使用 SVM 分類器來進行表情辨識。實驗的資料庫是使用著名的 Cohn-Kanade 表情資料庫，該資料庫由於具有完整的表情資料，因此是表情辨識領域的研究學者們常用來進行研究的一套表情資料庫。本論文所提出的方法，對 Cohn-Kanade 表情資料庫進行 7 類表情辨識，並使用 10-fold person-independent cross-validation 架構，可以得到高達 91.1%的辨識率，效果優於目前多種主要的表情辨識演算法。這顯示我們所提出的權重式區域方向特徵和辨識方法較先前的演算法具有進步性，已達到本計畫的研究目標。下圖為整合系統的測試結果畫面，所拍攝的影像為畫面的主體，在左上角顯示所偵測和抓取的人臉影像，而在臉部的右上方則以圖畫方式顯示所辨識出來的表情結果，如下列圖形所示。





5. 結果與討論

本計畫發展完成的內容包含有(1)完成「臉部表情資料庫」之六種表情(高興、悲傷、驚訝、痛苦、沮喪和自人表情)的收集與所有臉部特徵點之標示和整理；(2)完成「臉部特徵點定位」演算法的開發；(3)完成「表情辨識」演算法的開發；和(4)完成「表情辨識」整合系統。計畫進行中，雖然遇到多次的困難，但均一一克服，已順利達成計畫目標。

本計畫的成果已於 International Conference on Machine Learning and Cybernetics (ICMLC) 國際會議中發表論文，題目為「An Adaboost-Based Facial Expression Recognition Method」，此會議於 2011 年 7 月於中國廣西桂林舉行，論文內容如下

AN ADABOOST-BASED FACIAL EXPRESSION RECOGNITION METHOD

YEA-SHUAN HUANG, SHUN-HSU CHUANG, FANG-HSUAN CHENG

Dept. of Computer Science & Information Engineering
Chung Hua University, Hsinchu, Taiwan
E-mail: yeashuan@chu.edu.tw

ABSTRACT:

A method of combining Weighted Local Directional Pattern (WLDP) and Local Binary Pattern (LBP) for facial expression recognition is proposed. First, WLDP and LBP are applied to extract human facial features. Second, principle component analysis (PCA) is used to reduce their feature dimensions respectively. Third, both reduced facial features are merged to form the final feature vector. Fourth, support vector machine (SVM) is used to recognize facial expressions. Experiment on the well known Cohn-Kanade expression database, a high accuracy rate up to 91.1% for recognizing seven expressions can be achieved with a person-independent 10-fold cross-validation scheme.

Keywords:

Facial Expression Recognition; Local Binary Pattern; Weighted Local Directional Pattern; Principal Component Analysis; Support Vector Machine;

1. Introduction

In recent years, Human-Computer Interaction (HCI) becomes more and more popular. If a computer can know a person's feeling or emotions and react immediately, it will be regarded as very friendly. In general, there are many ways for a computer to communicate with human beings, such as by speaking, by gesturing and by making facial expressions. Researchers demonstrate that information passed by language accounts for 7 percent, information passed by tone of voice accounts for 38 percent, however, information passed by speaking people expressions accounts for 55 percent during human intercourses. Therefore, a lot of research effort has been devoted to facial expression recognition. However, there is no matured product or technology today so that facial expression recognition is still an active research topic.

In this section, we will briefly introduce the existed methods for facial expression recognition. In 1978, Ekman and Friesen defined six universal facial expressions [1], including happiness, sadness, anger, surprise, fear and disgust. They proclaimed these expressions are stable not only across different races, but also across different cultures. In order to make recognition procedure more standardized, a set of

muscle of movements, known as Action Units, was created by psychologists, thus forming the so-called Facial Action Coding System (FACS) [2]-[3], and these action units can be transformed into six universal expressions by rules proposed in [4].

To recognize these six universal expressions, we can in general classify these algorithms into two categories; one is feature-based methods and the other is appearance-based methods. Recently, Valstar et al. [5]-[6] have demonstrated that geometric feature-based methods provide better or equal performance than appearance-based approaches in Action Unit recognition. In [7], Kotsia et al. use shape and texture information to recognize six expressions and got superior performance on Cohn-Kanade. However, feature-based methods usually require robust facial feature localization in order to get good recognition results. For appearance-based methods, image filter such as Gabor-wavelet, is in general applied to the whole or a partial face image to extract the facial texture frequencies [8]-[12]. Although, the appearance-based method by using Gabor-wavelet gets excellent performance, but due to convolve face images with banks of Gabor filters, it is hard to be applied on real-time facial expression recognition systems. There are some methods [13] which use LBP to extract effective appearance features and get similar performance as Gabor-wavelet method, but it is still suffer much from non-monotonic illumination variation, random noise, and change in pose, age, expression. Some subspace methods such as 2DLPP [14] are also used to extract appearance information for facial expression recognition but it still have some drawback at illumination changes. Hence, we try to combine another feature to use the edge information to enhance the recognition performance. This paper proposes a novel algorithm by using the combined features of Boosting LBP and Boosting WLDP to recognize human facial expressions. The proposed method can enhance the performance than only use LBP or LDP method.

The rest of this paper is organized as follows. Section 2 describes the preprocessing step and LBP and WLDP feature extraction method. Section 3 illustrates the boosting algorithm and fusion feature. Section 4

introduces our experiments and the corresponding results. Finally, conclusions are drawn in Section 5.

2. Algorithm Description

The proposed algorithm can be separate into several steps. The first is the preprocessing step and the second step is the extract the proposed fusion feature, and finally uses SVM to recognize the expression. The detail preprocessing step and proposed fusion feature is described in this section. The Fig. 1 demonstrates the proposed algorithm flowchart.

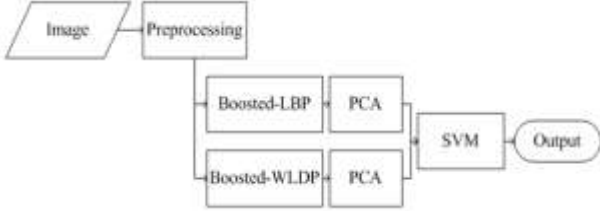


Fig. 1: The flowchart of the proposed algorithm

2.1. Preprocessing Procedure

Before recognizing a human face, an appropriate pre-processing procedure to obtain both slant and size normalized face images is very important because it can make recognition easier and more accurate. In this section, we illustrate how the preprocessing procedure performs.

For reducing the face structure variation resulting from different human races, all the face images are cropped first and then normalized. To obtain a normalized face region, we not only use the two eye positions but also with two mouth-corner positions. The main concept of using the two mouth-corner positions is that some people's faces are longer than others, but a complete face image of either longer face or shorter face can still be well extracted according to two eyes and two mouth corners.

At first, we rotate the face image to align the horizontal eye coordinates. The left boundary and the right boundary of the face region are extracted from the centre of the two eyes, and the bottom boundary of the face region is extracted by the centre of the mouth corners. The centre of the two eyes here we define as (x_{ec}, y_{ec}) . The top-left corner can be defined as $(x_{ec}-0.9d, y_{ec}-0.6d)$ and the top-right corner can be defined as $(x_{ec}+0.9d, y_{ec}-0.6d)$. The bottom-left and bottom-right corner here, we use the centre of the two mouth corners, and the coordinate of the centre point we define as (x_{mc}, y_{mc}) . The bottom-left corner can be defined as $(x_{mc}-0.9d, y_{mc}+0.3d)$ and the bottom-right corner can be defined as $(x_{mc}+0.9d, y_{mc}+0.3d)$. The proportion of facial region is shown clearly in Fig. 2.

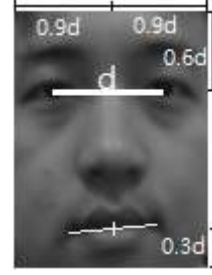


Fig. 2: The definition of facial boundary model

When a face image is extracted, we normalize the extracted image to size of 128x96 pixels. Fig. 3 shows several image samples of the same person with different head rotations and facial expressions, and the normalized extracted face image is shown in the right bottom of each image sample. Obviously, the normalized face images look quite similar by using the specified preprocessing procedure. This makes successive expression recognition plausible.



Fig. 3: Several Image samples and their extracted images with different head rotations and expressions.

2.2 Local Binary Pattern (LBP)

The LBP operator is introduced by Ojala et al. [15] which label each image pixel by comparing its value with a 3x3 neighborhood pixels individually and converting the individual comparison result together to form an integrated feature array. The LBP operator can be described as follow. The (x_c, y_c) means the centre of the LBP mask and $s(x)$ is a thresholding function.

$$LBP(x_c, y_c) = \sum_{p=0}^8 s(g_p - g_c) 2^p \quad (1)$$

where

$$s(x) = \begin{cases} 1, & \text{if } x \geq 0; \\ 0, & \text{if } x < 0. \end{cases} \quad (2)$$

Fig. 4 illustrates an example of obtaining a LBP micro-pattern LBP.

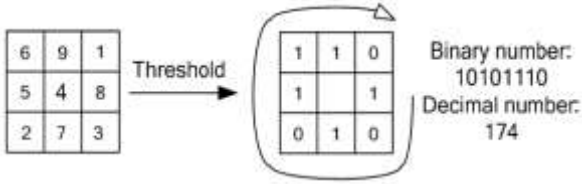


Fig. 4: The basic LBP operator

One variation of the LBP code is known as uniform pattern which is in general represented as LBPu2. The LBPu2 means the LBP binary string contains at most two bitwise transitions from 0 to 1 or vice versa. For example, 00110000, 00000000 and 11000000 are uniform patterns, but 11010100 and 01010101 are not uniform patterns. In general, LBPu2 has better or similar recognition performance than LBP. Therefore, LBPu2 is used as feature in our experiments and in the following discussion LBP indeed represent LBPu2 for simplicity.

2.3 Weighted Local Directional Pattern (WLDP)

The Local Directional Pattern [16] (LDP) is an eight-bit binary code assigned to each image pixel by comparing its 8-directional edge responses. The eight Kirsch masks ($m_0 \sim m_7$) shown in Fig. 5 of different directions are used to extract the 8-direction edge response values of each image pixel.

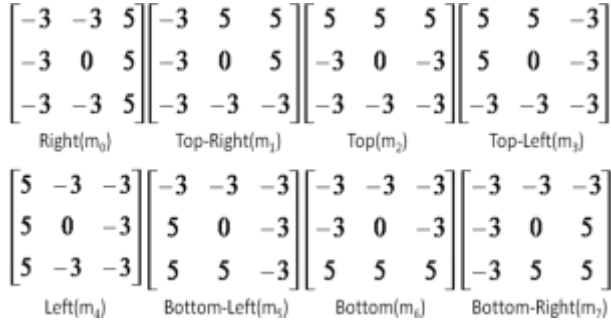


Fig. 5: Kirsch masks in eight directions.

After applying the eight Kirsch masks on an image pixel, eight edge response values m_0, m_1, \dots, m_7 , can be calculated and each edge response value represents the edge strength of one specific direction, which m_0, m_1, \dots, m_7 means the $0, 45, \dots, 275$ degree of the edge directions. Among the eight edge response values, the k largest absolute values are selected and their corresponding bits are set to 1. Of course, the other $8-k$ bits not corresponding to the k largest absolute values are set to 0. Accordingly, each image pixel has an eight-bit LDP code. The $LDP(x_c, y_c)$ means LDP code at coordinate (x_c, y_c) , $s(x)$ means a thresholding function and m_i means the i -th Kirsh mask is used. The LDP operator can be defined as follow.

$$LDP(x_c, y_c) = \sum_{i=0}^7 s(m_i) 2^i \quad (3)$$

where

$$s(x) = \begin{cases} 1 & , \text{if } |x| \geq k^{th}(M) \\ 0 & , \text{if } |x| < k^{th}(M) \end{cases} \quad (4)$$

$$M = \{m_0, m_1, \dots, m_7\} \quad (5)$$

Fig. 6 illustrates an example of obtaining LDP code with parameter $k=3$.

Index	M_7	M_6	M_5	M_4	M_3	M_2	M_1	M_0
Value	-320	240	113	126	410	-128	96	-30
Rank	2	3	6	5	1	4	7	8
Bit	1	1	0	0	1	0	0	0
LDP($k=3$) binary string 11001000								
LDP code is 200								

Fig. 6: The example of LDP code calculation.

The LDP operator by selecting the k largest absolute edge response values to encode the direction of the edge, but if the target image has no obviously edges such as white wall or the facial skin. The LDP algorithm still forced to encode the target image and get LDP code. Hence, we propose a novel weighting method to avoid the situation of the same importance to the facial skin and to the facial part such as eye, eyebrow, nose and mouth. The proposed WLDP method is illustrated as follow.

For each image pixel at coordinate (x, y) , beside computing its LDP code, a corresponding weighting value $W_{LDP}(x, y)$ is also calculated by (3).

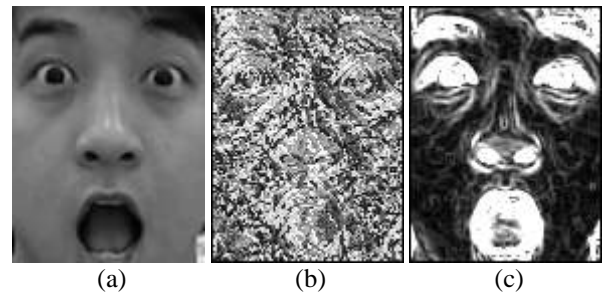
$$W_{LDP}(x, y) = \frac{1}{v(x, y)} \times \sum_{r=1}^k |r^{th}(M)| \quad (6)$$

where

$$v(x, y) = \sum_{i=-1}^1 \sum_{j=-1}^1 I(x+i, y+j) \quad (7)$$

$$M = \{m_0, m_1, \dots, m_7\} \quad (8)$$

The motivation of the weighting function is that some of the facial regions such as eyebrows, eyes, noses and mouth have darker pixel value then others and has stronger edge response value. For this proposes, the left term of the weighting function is designed by summing all pixel values from its neighbors of the center coordinate at (x, y) and the right term is designed by summing the top k absolute edge response values. Fig. 7 shows an example of the original image, the LDP encoding image and the weighting image.



(a) (b) (c)

Fig. 7: (a) the original image, (b) the LDP encoding image and (c) the weighting image.

3. Facial Expression Recognition

In this section, the whole proposed algorithm is introduced. Suppose the eye positions and the mouth corners positions have been obtained by either detection algorithms or manually marking.

At first, we use an AdaBoost.M2 algorithm to select the most discriminative sub-regions from the training samples. Then according to these selected sub-regions, the individual histogram of LBP and WLDP are respectively computed from each normalized face image. Usually, each of the two features has more than one thousand feature dimension, but the high feature dimension will easily degrade the system performance because of curse of dimensionality. Therefore, we further apply PCA to reduce the dimension of each feature. After feature reduction, both reduced features are then concatenated to form an enhanced LBP-WLDP feature vector. In the testing stage, the corresponding enhanced LBP-WLDP feature vector is computed for each testing sample. The classification result is obtained by using SVM with RBF kernel. The AdaBoost.M2 and the histograms of LBP and WLDP are described in this section.

3.1 The histograms of LBP and WLDP

After applying the LBP and WLDP operator to an image, the result image ILBP and ILDP can be obtained. Then, the features of the image can be obtained by calculating the histogram of LBP and WLDP. Hence, the LBP histogram HLBP is a 59 bin histogram with uniform pattern and the LDP histogram HLDP is a 56 bin histogram ($C_k^8 = 56, k = 3$). The calculation of the LBP histogram can be defined by (9).

$$H_{LBP_i} = \sum_{x,y} P\{I_{LBP}(x, y) = c_i\} \quad (9)$$

where

$$P\{A\} = \begin{cases} 1, & \text{if } A \text{ is true;} \\ 0, & \text{if } A \text{ is false.} \end{cases} \quad (10)$$

The WLDP histogram can be defined by (11).

$$H_{WLDP_i} = \sum_{x,y} P_{x,y}\{I_{LDP}(x, y) = c_i\} \quad (11)$$

where

$$P_{x,y}\{A\} = \begin{cases} WLDP(x, y), & \text{if } A \text{ is true;} \\ 0, & \text{if } A \text{ is false.} \end{cases} \quad (12)$$

3.2 AdaBoost

For obtaining better description of face image, we apply AdaBoost.M2 to find the most discriminative LBP and WLDP histograms. Here we represent the sub-region of LBP and WLDP as weak. The size of the weak is from 10 to 25 with 5 pixel difference for each scale and each weak is shifted 4 pixels to the whole image. In total, there are 7,119 weak features. At the each iteration, a weak classifier with the minimum weighted error is selected, and distribution is updated to increase the weights of misclassified samples. For a weak classifier, we adopt template matching as follow. In training, the LBP and LDP histograms in a given class are averaged to generate a histogram for this class. In recognition, the LBP and LDP histograms of the test pattern is use to match the closest reference template by the nearest-neighbor classification. The Chi-Square Statistics (χ^2) is used to measure the histogram dissimilarity.

$$\chi^2(S, M) = \sum_i \frac{(S_i - M_i)^2}{S_i + M_i} \quad (13)$$

where S and M are two LBP or WLDP histograms.

4. Experimental Results

In order to evaluate the performance of different expression recognition algorithms, we conducted experiments on the well known Cohn-Kanade database, which is one of the most comprehensive databases in the facial-expression research community. The Cohn-Kanade facial expression database consists of 95 university students aged from 18 to 30, of which 65% was female, 15% were African-American, and 3% were Asian or Latino. Subjects were instructed to perform a series of 23 facial displays, in which six were based on descriptions of prototypic emotions (i.e., anger, disgust, fear, happiness, sadness, and surprise). Image sequences from neutral to target displays were digitized into 640x490 pixel arrays.

For our experiments, we selected 332 image sequences which consist of 93 subjects from the Cohn-Kanade database. The only selection criterion was that a sequence could be labeled as one of the six basic expressions. For each sequences, the neutral face and three peak frames were used for prototype expression recognition, resulting in 1327 images (105 Anger, 120 Disgust, 129 Fear, 270 Happiness, 153 Sadness, 219 Surprise and 331 Neutral).

For more practical evaluation, the 10-fold person-independent leave-one-out cross-validation scheme was adopted. The person-independent means the same subjects will not appear in both training and testing stages and for testing the generalization abilities, the 10-fold cross-validation is adopted. The 10-fold cross-validation means the one of the 10 groups was selected as the testing data and the remaining 9 groups were as

the training data each time. To avoid over-fitting, we further used the first group of the training data as the validation data to test the generalization ability of the trained classifier. This above process was repeated ten times and each group is in turn selected to serve the testing data which is omitted from the training process. Finally, the average recognition result on the 10 testing dataset is reported.

Table 1 shows the recognition performance of the proposed algorithm and other existed methods. Due to the different environmental setup such as image pre-processing, and different image sequences selected from the Cohn-Kanade database, we cannot compare their recognition rates directly, but this table still has the valuable for a research reference. Table 1 illustrates that WLDP gets better performance than the original LDP and the proposed algorithm with concatenating both Boosting LBP-PCA and Boosting WLDP-PCA gets almost the best performance.

Table 1: Recognition Performance

<i>Method</i>	<i>Recognition Rate(%)</i>
Block-Based LBP[11]	88.9
Boosting LBP[11]	91.4
Block-Based LBP	85.90
Block-Based LDP	86.73
Block-Based WLDP	87.71
Boosting LBP-PCA	88.77
Boosting LDP-PCA	87.33
Boosting WLDP-PCA	88.31
Proposed method (Boosting LBP-PCA+Boosting WLDP-PCA)	91.10

The confusion matrix of the proposed method in seven expressions recognition (AN=Anger, DI=Disgust, FE=Fear, HA=Happiness, NE=Neutral, SA=Sadness, SU=Surprise) is illustrated in Table II.

Output Input	AN	DI	FE	HA	NE	SA	SU
AN	72.4%		1.0%		21.0%	5.7%	
DI		94.2%	0.8%	1.7%	3.3%		
FE	1.6%		80.6%	8.5%	4.7%		4.7%
HA	0.4%		1.9%	94.8%	3.0%		
NE	0.9%	0.3%	0.3%		97.3%		0.3%
SA	2.6%		2.0%		17.0%	78.4%	
SU			0.5%				99.5%

Fig. 8: The confusion matrix of the proposed method.

In Table 1, the proposed method gets best performance than only use Boosting LBP-PCA or Boosting WLDP-PCA method individually. We also

implement the block-based method for comparing LBP, LDP and WLDP performance in [33]. In our experiments, the Block-Based WLDP gets best performance in the three methods, but still lower than [33] proposed. One of the possible reasons is that most of our samples in harder recognize expressions (105 Anger, 129 Fear, 331 Neutral and 153 Sadness images) in Table II is more than [33] (108 Anger, 99 Fear, 320 Neutral and 126 Sadness images) used and remaining obviously expressions in ours (120 Disgust, 270 Happiness and 219 Surprise images) is less than [33] (120 Disgust, 282 Happiness, 225 Surprise images) used.

5. Conclusion

This paper proposed a novel weighting method for LDP operator and fusion feature of Boosted-LBP and Boosted-WLDP are also proposed. We achieve the successful recognition rate 91.10% on Cohn-Kanade database with proposed fusion feature than only use Boosted-LBP and Boosted-WLDP.

Facial expression recognition is a challenge task due to each expression has different emotional representations. Until now, the research on facial expression recognition still focuses on six universal expressions and there is still a long way to go in cognizing facial expression from the psychology and physiology.

ACKNOWLEDGEMENT

The authors would like to thank the National Science Council, R.O.C. for the funding support of the project NSC 99-2221-E-216-050.

REFERENCES

- [1] P. Ekman, W.V. Friesen, "Facial Action Coding System: A Technique for the Measurement of Facial Movement," Consulting Psychologists Press, Palo Alto, 1978.
- [2] P. Ekman and W.Friesen, "Manual for the facial action coding system," Consulting Psychologists Press, 1977
- [3] J. C. T. Kanade and Y. Tian, "Comprehensive database for facial expression analysis," in Proceeding of IEEE International Conference on Face and Gesture Recognition, pp. 46-53, March 2000.
- [4] M. Pantic and L. Rothkrantz, "Automatic analysis of facial expressions: The state of the art," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, pp. 1424-1445, December 2000.
- [5] M. Valstart, I.Patras, M. Pantic, "Facial action unit detection using probabilistic actively learned support vector machines on tracked facial point data," IEEE Conference on Computer Vision and Pattern Recognition Workshop, vol. 3, pp. 76-84, 2005.
- [6] M. Valstar, M. Pantic, "Fully automatic facial action unit detection and temporal analysis," IEEE Conference on Computer Vision and Pattern Recognition Workshop, pp. 149, 2006.

- [7] Kotsia, S. Zafeiriou, I. Pitas, "Texture and shape information fusion for facial expression and facial action unit recognition," *Pattern Recognition*, vol. 41, issue 3, pp. 833-851, 2008.
- [8] Z. Zhang, M. J. Lyons, M. Schuster, S. Akamatsu, "Comparison between geometry-based and gabor-wavelets-based facial expression recognition using multi-layer perceptron," *IEEE International Conference on Automatic Face & Gesture Recognition (FG)*, 1998.
- [9] M. J. Lyons, J. Budynek, S. Akamatsu, "Automatic classification of single facial images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, issue 12, pp. 1357-1362, 1999.
- [10] G. Donato, M. Bartlett, J. Hager, P. Ekman, T. Sejnowski, "Classifying facial actions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, issue 10, pp. 974-989, 1999.
- [11] Y. Tian, "Evaluation of face resolution for expression analysis," *CVPR Workshop on Face Processing in Video*, 2004.
- [12] M. S. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, J. Movellan, "Recognizing facial expression: Machine learning and application to spontaneous behavior," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005.
- [13] C. Shan, S. Gong, P. W. McOwan, "Facial Expression Recognition based on Local Binary Patterns: A comprehensive study," *Image and Vision Computing*, vol. 27, issue 6, pp. 609-839, 2009.
- [14] R. Zhi, Q. Ruan, "Facial expression recognition based on two-dimensional discriminant locality preserving projections," *Neurocomputing*, vol. 71, pp 1730-1734, 2009.
- [15] T. Ojala, M. Pietikäine, D. Harwood, "A comparative study of texture measures with classification based on featured distribution," *Pattern Recognition*, vol. 29, issue 1, pp. 51-59, 1996.
- [16] T. Jabid, M.H. Kabir, C. Oksam, "Local Directional Pattern (LDP) for face recognition," *International Conference on Consumer Electronics (ICCE)*, pp. 329-330, 2010.

Invitation Letter

Dear PCM 2010 Authors,

We are very pleased to inform you that your paper:

Facial landmark detection by combining object detection and active shape model

authored by:

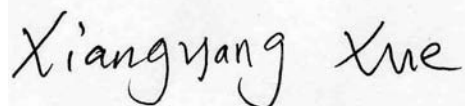
Hsu, Ting-Chia; Huang, Yea-Shuan; Cheng, Fang-Hsuan

has been selected for presentation at the IEEE 11th Pacific-Rim Conference on Multimedia (PCM 2010), which will be held in Shanghai, China, from the 21st to 24th September 2010.

As the hosting organization of this great event, it is our honor to invite you to come to Shanghai, China and attend the conference. Please check out the conference website, <http://pcm2010.fudan.edu.cn/>, for registration, travel and accommodation information. You can also contact our Local Organization Committee in case you need any help: http://pcm2010.fudan.edu.cn/org_com.htm

We look forward to seeing you at the conference.

Best regards,



(Xiangyang XUE)

General Co-Chair
Professor

PCM 2010
School of Computer Science,
Fudan University, China

行政院國家科學委員會補助國內專家學者出席國際學術會議報告

99 年 9 月 29 日

附件三

報告人姓名	黃雅軒	服務機構 及職稱	中華大學資工系 副教授
時間 會議 地點	2010. 09. 20~2010. 09. 25 中國 上海	本會核定 補助文號	NSC 97-2221-E-216-040 NSC 98-2221-E-216-029
會議 名稱	Pacific Rim Conference on Multimedia (PCM)		
發表 論文 題目	A novel ASM-based two-stage facial landmark detection method		

報告內容應包括下列各項：

一、參加會議經過

- 此次會議共有 4 天，第一天有 2 個 Tutorial Talk：
 - ◇ Histogram in tersection kernel learning for multimedia applications
 - ◇ MPEG activities for 3D video coding
 而第二天到第四天均為論文發表，其中每天都安排一個 Keynote Speech，題目分別是
 1. (9/22) A New Machine Learning Framework with Application to Image Annotation, Prof. Zhi-Hua ZHOU
 2. (9/23) The Evolution of Image Search, Prof. Kyros Kutulakos, Prof. Yong RUI;
 3. (5/24) Learning to Rank: Pushing the Frontier of Web Search, Prof. Tie-Yan LIU.
- 此會議總共包含 15 sections，大部分時段都有 poster section，口頭報告時段的主題包含四大類，分別為
 1. Multimedia Analysis and Retrieval
 2. Multimedia Systems and Applications
 3. Multimedia Compression, Communication and Optimization
 4. Multimedia Security and Right Management
- 我的論文發表於第二天下午第一時段(13:30~15:30)，題目為” A novel ASM-based two-stage facial landmark detection method”，由於上一篇論文作者沒到場發表，故我有充

分時間講解論文，在場聽眾事後表示本論文發表非常清楚和完整，充分達到研究心得交流目的。下二圖為 PCM 會議和論文發表時的照片



圖一 PCM 會場與看板



圖一 PCM 論文發表

- 此會議參加人數比預期的少，故可以有較多的時間彼此討論。由於會議專注於多媒體(特別是影像和電腦視覺)的技術和應用，而所有與會人員都是從事於此領域研究，所以有很好的經驗交談和學習的機會。

二、與會心得

- 此會議共有 261 篇論文投稿，最後只選出 75 篇口頭報告和 56 篇海報論文，可見此會議是一個具有較高水準的會議。
- 然而，個人對主辦單位感到失望，整體會議並沒有經過精心的規畫和佈置，不但沒有見到醒目的會議旗幟，第二天開幕時也只見到少數著名學者出席，似乎主辦單位(復旦大學)沒有投入足夠的人力和邀請來舉辦此會議。
- 有很多大陸學生參與此會議，感覺他們很有勇氣，積極的發問，有強烈的企圖心，相信他們在此會議中會有很好的學習。
- 幾乎看不到台灣學生的參加，比起大陸學生的積極參與，個人對台灣年輕一代的研究感到憂慮，若不努力，恐怕會落後大陸越來越多。

三、建議

- 本校老師多參與國際性會議，除了介紹研究成果，增加學校的知名度以外，也能快速擴展視野，建立合作管道，對未來研究和教學有很大的幫助。
- 鼓勵學校的研究生參與這種研究與應用結合的國際性會議，讓他們更了解研究的實用價值，以激發學習和研究的熱誠。

四、攜回資料名稱及內容

會議論文集上、下二冊

International Conference on Machine Learning and Cybernetics 2011

Dear Author,

Congratulations. Your paper has been accepted for publication in the proceedings of the International Conference on Machine Learning and Cybernetics (ICMLC) 2011.

Please verify the following items to ensure their accuracies:

- (1) Please confirm the following:

Paper ID: **3747**

Title: **An Adaboost-Based Facial Expression Recognition Method**

Authors: **Yea-Shuan Huang, Chung-Hua University, Taiwan**
SHUN-HSU CHUANG, Chung-Hua University, Taiwan
Fang-Hsuan CHENG, Chung-Hua University, Taiwan

Please be reminded your paper *will be rejected* if the title, names of author(s) or the order of author list of your final version is *different* from the original version.

- (2) If your paper exceeds six pages, you must pay **HKD 540** (USD 70 #) per extra page.
- (3) Your paper will **NOT** be published in the conference proceedings unless you COMPLETE EVERY STEP of the following tasks by **7-May-2011**
 - a. Register on or before **7-May-2011**

The registration fee:

Non-IEEE Member:	USD 550 (equivalent to HKD 4270 #)
IEEE Member*:	USD 500 (equivalent to HKD 3890 #)
Student*:	USD 450 (equivalent to HKD 3500 #)

Please pay in Hong Kong Currency.

* Identification(s) may be required. Please bring your identification(s) to the conference.

As of 20-April-2011, the exchange rate is HKD 7.78 for one USD.

The registration fee of each participant should be settled by credit card payment. For detail, please see the other attachment (creditcardpayment.doc)

Please be reminded when you settle the registration fee, *attach the Paper ID, Paper Title and the Registrant Name as a remark.*

- b. We will need the *LaTeX or WORD version of your paper (WORD document is preferable)* for English grammar and/or usage editing. Please revise your paper according to the reviewer comments (shown on the ICMLC Submission System). Pay special attention to your paper format and make sure your paper follows the Formatting Guidelines as stated in the ICMLC website. The LaTeX or WORD version of the paper should be submitted via the ICMLC Submission System.
- c. Fill in the *Online Registration Form*. The Registration Form can be found in the ICMLC Submission System. The Oral or Poster presentation will be assigned according to your preference. Please state your preference for your paper presentation. You will be notified whether your paper will be presented in an Oral Session or Poster Session in June 2011.
- d. *Fill in the IEEE Copyright Form* The IEEE Copyright Form can be signed via the ICMLC Submission System.

Steps *b*, *c* and *d* will need to be completed when you logon to the *ICMLC submission system*. (<http://www.icmlc.com/icmlcSystem/author/>)

- (4) After the completion of your registration, you will receive an email attached with *a registration confirmation letter. You need to print it and bring it to the conference as proof*. If you cannot attend the conference and would like to ask your colleague to pick up the conference proceeding CDROM and one volume of the proceeding, your colleague must present your registration receipt to the conference organizer.
- (5) You will be informed whether your paper will be invited to participate in the *Conference Awards* and/or *Ph.D. Colloquia Series* in *June 2011*.
- (6) We will confirm your *attendance* in ICMLC 2011 in *June 2011*.
- (7) If you have any enquiry, please contact the Conference Secretary, Patrick Chan, at patrickchan@ieee.org. Please *indicate your paper ID*.

Thank you again for your contribution to the ICMLC, and we look forward to seeing you in Guilin.

Best regards,
ICMLC 2011 Program Committee

行政院國家科學委員會補助國內專家學者出席國際學術會議報告

100 年 7 月 17 日

附件三

報告人姓名	黃雅軒	服務機構 及職稱	中華大學資工系 副教授
時間 會議 地點	2011. 07. 10~2011. 07. 15 中國 桂林	本會核定 補助文號	NSC 99-2221-E-216-050 099-B02-007(產學計畫)
會議 名稱	International Conference on Machine Learning and Cybernetics (ICMLC)		
發表 論文 題目	An Adaboost-Based Facial Expression Recognition Method		

報告內容應包括下列各項：

一、參加會議經過

- 此次會議共有 4 天，第一天是 Tutorial，題目為：Essentials of research methodology and effective dissemination of research results。第二天到第四天則是會議的口頭發表和海報展示時間，大會還安排二個 Keynote speeches 和一個 Panel Discussion，他們的題目分別為
 - ✧ Keynote 1: (7/11) Machine Learning Challenges for Human Brain Decoding by Prof. Seong-Whan Lee.
 - ✧ Keynote 2: (7/12) Fuzzy Forecasting Based on High-Order Fuzzy Time Series and Genetic Algorithms by Prof. Shyi-Ming Chen.
 - ✧ Panel Discussion: (7/12) The Genesis of an Innovative Research Topic by Dr. T.K. Ho, Prof. H. Yan, Prof. V. Marik, and Prof S. Kwong.
- 我的論文安排於第二天(7/11)下午時段(15:10~16:50) 發表，題目為” An Adaboost-Based Facial Expression Recognition Method” ，此篇論文結合了二種紋理特徵(WLDP 和 LBP)，來進行臉部表情辨識。下圖為此會議的大會場景照片。



圖一 ICMLC 會場與看板

二、與會心得

- 此會議的 Panel Discussion 邀請了四位在業界和學界具有豐富的專家 Dr. T.K. Ho (Bell Lab), Prof. H. Yan (City University of Hong Kong.), Prof. V. Marik (Czech Technical University), 和 Prof S. Kwong (City University of Hong Kong,)來談論如何開啟或做出具創新性的研界成果，每位專家從他們當初如何選擇博士研究題目開始談起，進而論及如何從原來的研究議題中繼續或擴大地找出新議題，並作出有意義的研究成果。在這次的專題討論中，我印象最深刻的內容有二：
 - Dr. T.K.Ho 建議一個好的研究人員應該具有三種研究態度，即”真誠(Sincere)”、“好奇(Curiosity)”和”分享(Sharing)”。
 - Prof. S. Kwong 以遺傳演算法來說明如何進行創新性的研究，包含”適應(Adaption)”、“交換(Crossover)”、“突變(Mutation)”和”持續努力(Iteration)”等四種生物界常用的法則，可以產生更好的物競天擇的更佳結果。
- 此次會議中有些台灣學生參加，它們當中有些是博士生，有些是碩士生，甚至有些是大學生，他們大部份表現得很好，頗有大將之風。他們表示參加此會議擴展了自己視野，對未來工作和研究都有正面的影響。我真為這些台灣年輕一代的學生喝采，覺得台灣學生是很優秀也很有希望的。

三、建議

- 本校老師多參與國際性會議，除了介紹研究成果，增加學校的知名度以外，也能快速擴展視野，建立合作管道，對未來研究和教學有很大的幫助。
- 鼓勵學校的研究生參與這種研究與應用結合的國際性會議，讓他們更了解研究的實用價值，以激發學習和研究的熱誠。

四、攜回資料名稱及內容

會議論文集一冊和一片光碟

國科會補助計畫衍生研發成果推廣資料表

日期:2011/10/31

國科會補助計畫	計畫名稱: 多辨識器組合之強健式臉部表情辨識技術
	計畫主持人: 黃雅軒
	計畫編號: 99-2221-E-216-050- 學門領域: 圖形辨識
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：黃雅軒		計畫編號：99-2221-E-216-050-					
計畫名稱：多辨識器組合之強健式臉部表情辨識技術							
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	0	0	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	0	1	100%		
		專書	0	0	100%		
	專利	申請中件數	0	0	100%	件	
		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（本國籍）	碩士生	4	7	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%		
國外	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	0	2	100%		
		專書	0	0	100%		章/本
	專利	申請中件數	0	0	100%	件	
		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%		

<p>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。)</p>	<p>無</p>
--	----------

	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）

已發表二篇 EI 會議論文：(1) An Adaboost-Based Facial Expression Recognition Method 和 (2) A novel ASM-based two-stage facial landmark detection method，分別發表於 International Conference on Machine Learning and Cybernetics (ICMLC) 和 Pacific Rim Conference on Multimedia (PCM) 國際會議中。

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）

本計畫提出一種混合權重式區域方向特徵(WLDP)和二元化形態(LBP)的表情辨識方法。一開始 WLDP 和 LBP 會分別對人臉影像進行特徵抽取，接著利用 PCA 分別對 WLDP 和 LBP 所抽取出來的特徵進行特徵降維處理，最後將兩種特徵進行混合，產生出一種對人臉具有分辨能力的混合特徵，並使用 SVM 分類器來進行表情辨識。實驗的資料庫是使用著名的 Cohn-Kanade 表情資料庫，該資料庫由於具有完整的表情資料，因此是表情辨識領域的研究學者們常用來進行研究的一套表情資料庫。本論文所提出的方法，對 Cohn-Kanade 表情資料庫進行 7 類表情辨識，並使用 10-fold person-independent cross-validation 架構，可以得到高達 91.1% 的辨識率，效果優於一般常見的表情辨識演算法。這顯示我們所提出的權重式區域方向特徵和辨識方法具有良好的表情辨識效果，具有研究和應用的價值。

另外，雖然動態形狀模型(Active Shape Model, ASM)已經成功的被應用在臉部特徵點定位上，然而當人們臉部有著誇張的臉部表情變化時，如：驚訝、大笑和挑眉等，其特徵點定位的結果仍有明顯的誤差率。為了克服這個問題，本計畫提出串接式多階段的臉部特徵點定位方法。第一階段，我們使用 Adaboosting 學習演算法來定位臉部特徵點中較有明顯鑑別度之特徵點，而這些有明顯鑑別度的特徵點都是角點類型的地標點，如左右內眼角、

左右外眼角、左右內眉角、左右外眉角和左右嘴角等共 10 個點。並於第一階段做此 10 個角點的動態形狀模型之重建與定位，來獲得較符合臉部幾何架構之臉部特徵點位置。第二階段，我們根據第一階段所定位的臉部角點位置，重新初始化臉部五官模型位置，再根據地標點變動的分佈情形，制定出每個地標點的搜尋範圍，然後進行第二次的動態形狀模型的重建與定位。經由實驗的結果證明，我們所提出的臉部特徵點定位方法，的確可以達到較好的臉部特徵點定位效果。此部分的技術已經技轉給凌通科技，可證明此技術的實用性。