

# 行政院國家科學委員會專題研究計畫 期中進度報告

## 子計畫三：利用人臉表情及唇形動態資訊進行身分確認之研究(1/3)

計畫類別：整合型計畫

計畫編號：NSC92-2213-E-002-082-

執行期間：92年08月01日至93年07月31日

執行單位：國立臺灣大學資訊工程學系暨研究所

計畫主持人：洪一平

計畫參與人員：王嘉銘、劉憲璋、江岳軒

報告類型：精簡報告

處理方式：本計畫可公開查詢

中 華 民 國 93 年 5 月 31 日

# 行政院國家科學委員會補助專題研究計畫期中精簡報告

## 結合音訊與視訊之多模組身分確認之研究 — 子計畫三： 利用人臉表情及唇形動態資訊進行身分確認之研究(1/3)

計畫類別：個別型計畫      整合型計畫

計畫編號：NSC -92-2213-E-002-082

執行期間：92年8月1日至93年7月31日

計畫主持人：洪一平

本成果報告包括以下應繳交之附件：

- 赴國外出差或研習心得報告一份
- 赴大陸地區出差或研習心得報告一份
- 出席國際學術會議心得報告及發表之論文各一份
- 國際合作研究計畫國外研究報告書一份

執行單位：國立台灣大學資訊工程學系暨研究所

中 華 民 國 93 年 5 月 30 日

# 行政院國家科學委員會專題研究計畫期中精簡報告

## 結合音訊與視訊之多模組身分確認之研究 — 子計畫三： 利用人臉表情及唇形動態資訊進行身分確認之研究(1/3)

### Biometric Person Authentication Using Dynamic Information of Facial Expression and Lip Tracking

計畫編號：NSC-92-2213-E-002-082

執行期限：92年8月1日至93年7月31日

主持人：洪一平

計畫參與人員：王嘉銘、劉憲璋、江岳軒

#### 一、中文摘要

本子計畫的目標在於結合人臉表情與唇形追蹤等動態資訊進行身分確認。在傳統利用靜態資訊的方法中，由於影像中人臉的大小、方位，以及取像時的背景、光線均存在極大的變異性，增加辨識的困難度。因此，在本子計畫中，我們嘗試利用使用者的臉部動態表情中所蘊含的資訊來進行身分確認。在第一年的研究中，我們設計了一個基本的系統，將系統運作分成「系統設定階段」和「系統使用階段」。首先，在系統設定階段，使用者臉部先做出某一種特定的表情，讓攝影機得以擷取一段用來訓練系統的視訊。接著，使用主成份分析法(PCA)可以找到一個參考用的固有子空間。在系統使用階段中，使用者如果做出與系統設定階段一樣的特定表情，我們即可使用相互子空間法，計算此表情影像所形成的子空間與宣稱對象之固有子空間的夾角。如果角度小於某個門檻值時，即可通過身分確認。未來兩年，我們將再結合唇形的動態資訊，進一步改進此系統之效能。

**關鍵詞：**身分確認、生物測定學、人臉表情、唇形追蹤、光流場、相互子空間法

#### Abstract

The goal of this project is to combine dy-

amic information of human face, such as facial expression and lip tracking, for person authentication. In the traditional methods that only utilize static information, the large variations in face size, face pose, lighting and background increase the difficulty of face verification. Therefore, in this research project, we will use the dynamic information contained in face expressions for person authentication. Our system we designed in the first year contains two stages: the setup stage and the usage stage. In the setup stage, the user first shows a specific expression, and the corresponding image set will form a subspace that can be obtained by using principle component analysis (PCA). Second, in the usage stage, the user simply shows this specific facial expression, and the system will calculate its subspace. Once the subspace is computed, we use the mutually subspace method (MSM) to calculate the angle between these two subspaces. If the angle is less than a threshold, our system will let the user pass the verification process. In the next two years, we will combine the dynamic information obtained from lip tracking, and try to further improve the system performance.

**Keywords:** Person Authentication, Biometrics, Facial Expression, Lip Tracking, Optical Flow Field, Mutual Subspace Method.

## 二、緣由與目的

隨著科技的進步，自動身分確認已成為一個愈來愈重要的問題。基於憑證(token-based)或基於知識(knowledge-based)的方法已經愈來愈不敷安全及便利的需求。因此，利用生物特徵的身分確認系統在近年來是一個非常熱門的課題。在各種生物特徵中，人臉是最明顯的外露特徵。就人類視覺而言，在相當遠的距離時我們便可藉由人臉來分辨出對方的身分，因此人臉辨識與確認一直是電腦視覺領域持續關注探討的問題。

所謂的「人臉身分確認」(face authentication)，與「人臉識別」(face recognition)並不是完全相同的工作，前者是要針對處理對象所宣稱的身分做出確認的動作，後者則是要判斷處理對象的身分是誰。這兩種工作在決策方式與評估方法不盡相同。相關的研究大致可以根據人臉特徵的資訊來源分成兩大類，一類使用人臉的靜態資訊，這一類的方法會希望所處理的人臉盡量不要有表情變化；另一類則是利用人臉的動態資訊，這一類的方法希望所處理的人臉最好能有一些表情或唇形變化。

在現有利用靜態臉部資訊的方法中，利用固有特徵臉(eigenface)的方法[1]、使用線性判別分析(LDA)子空間的統計方法[2]，以及使用彈性串圖形比對的類神經網路方法[3]是已知比較好的三種方法。利用動態資訊的部分，CMU 的陳祖瀚等人提出了利用固有特徵光流(eigenflow)的方法來從事人臉身分確認[4]。他們先是計算處理對象的人臉影像與訓練用的人臉影像之間的光流場與相對應的光流補償殘像誤差，然後計算在事前訓練階段所求得的固有特徵子空間下的固有特徵光流誤差，最後再結合光流補償殘像誤差及固有特徵光流誤差，以判斷兩者是否具有同一身分。在此一子計畫中，我們採用不同的方法，以期能獲得較佳的人臉確認正確率。

## 三、研究方法

為了表示不同的人說出同一段特定語彙所造成的不同特徵，我們利用「主成分分析法」(Principle Component Analysis, PCA)來對於每個人的特定表情的影像集合計算個別的子空間[5]。

一般的人臉身分確認大都採用單張影像比對單張影像。在這裡，為了比較兩個影像集合之相似度，我們採用 Yamaguchi 等人所提出的相互子空間方法(mutual subspace method, MSM)[6]。這方法定義相似度為兩個子空間的夾角。一個子空間是由受測對象的表情的一序列影像所定義出來的，另一個則是由宣稱身分的表情的一序列影像所定義出來的。假設子空間 D 和 G 代表這兩個序列影像所定義出來的子空間，而 P 和 Q 分別代表子空間 D 和 G 的正交投影矩陣。角度  $\theta$  被定義為

$$\cos^2 \theta = \sup_{d \in D, g \in G, \|d\| \neq 0, \|g\| \neq 0} \frac{|(d, g)|^2}{\|d\|^2 \|g\|^2}$$

上式有個最大值，也就是 PQP 這矩陣所計算出來最大的固有值[7][8]。如下式所示：

$$\cos^2 \theta = \lambda_{\max}$$

比較此一夾角  $\theta$  與事前設定的某個門檻值，即可判斷此人身分的真假。

## 四、結果與討論

在本年度的計畫中，我們把焦點放在特定表情的研究上。我們的系統分為「系統設定階段」和「系統使用階段」兩個部分，分別說明如下。在「系統設定階段」部分，我們先建立人臉資料庫，對每個使用者錄製一段其說出特定語彙(例如：“Open Sesame”)的表情。在我們初步建立的資料庫中共有 7 個受測者。每個受測者分別選擇三個不同的日子。在每個選定的日子中錄製兩次唸特定語彙的表情。其中一次拿來當做訓練影像集合，另一次拿來當做測試影像集合。圖一顯示某一受測者在唸出這段語彙時的部分影像序列。



圖一：使用者唸“Open Sesame”的一段影像序列

我們將每一個受測者的一連串動態影像經過主成分分析法後，取固有值最大的前八個主軸，形成一個固有子空間，如圖二。



圖二：受測者的影像序列經由主成分分析所得到的八個主軸。

給定一個使用者的測試影像集合，我們可以將其拿來跟資料庫中的七個受測者的訓練影像集合作比對。利用上一節所說明的 MSM 方法，對於每一次的比對，我們都可以計算出八個夾角。表格一顯示其中一名使用者的測試結果。如果使用者宣稱的身分確為其人，則兩個子空間之間的夾角通常都會比其他的夾角要來的小。

由於僅對動態影像本身使用主成分分析法，並不能表示時間軸方面的前後順序之資訊，因此加入運動估測可能還是有必要的。光流估測是一種常見的運動估測方法。我們可以利用光流資訊來找出一序列影像時間前後的關聯性。利用光流資訊來進行身分確認的部分目前還在研究階段，在過去，常被使用的光流演算法是 Lucas-Kanade 方法[10]，但這個方法僅僅利用前後兩張影像，準確性較差。在我們未來的做法中，將會利用一序列影像來估測光流，以達到更準確的效果。Irani 在過去

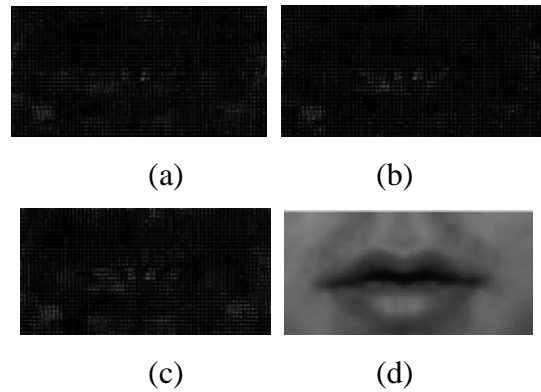
的論文中曾經發表利用多張影像估測光流的方法[9]，但這個方法必須用到整張影像所有資訊來做估測。在本年度，我們也設計了以特徵點(例如：嘴角、眼睛)為基礎的多圖框光流估測技術。假設  $(u_{ij}, v_{ij})$  是某張影像第  $i$  個點對第  $j$  張影像的光流，並且我們在該張影像前後各取  $K$  張影像。假設我們以階層式的方法來表示這組影像序列，在第  $n$  個迴圈，我們可以定義能量函數為

$$E(x, y; p) = \sum_{j=-K}^K \sum_{k \in \Omega_i} (u_i \cdot I_{xk} + v_i \cdot I_{yk} + I_t^{(n)})^2.$$

其中  $\Omega_i$  是點  $i$  的鄰域， $p$  表示假設的不同運動模型。假定在等速度運動模型之下，將此式對  $p$  微分，我們可以得到光流估測如

$$p = [u_i \quad v_i]^T = - \begin{bmatrix} \sum_{j=-T}^T \sum_{k \in \Omega_i} (j \cdot I_{xk})^2 & \sum_{j=-T}^T \sum_{k \in \Omega_i} j^2 (I_{xk} \cdot I_{yk}) \\ \sum_{j=-T}^T \sum_{k \in \Omega_i} j^2 (I_{xk} \cdot I_{yk}) & \sum_{j=-T}^T \sum_{k \in \Omega_i} (j \cdot I_{yk})^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{j=-T}^T \sum_{k \in \Omega_i} j \cdot I_{xk} I_t^{(n)} \\ \sum_{j=-T}^T \sum_{k \in \Omega_i} j \cdot I_{yk} I_t^{(n)} \end{bmatrix}.$$

圖三顯示 Lucas-Kanade 演算法、Irani 演算法以及我們設計的演算法對於某特定表情的光流場。未來我們將進一步分析三種演算法對於辨識效果的影響。



圖三：使用者說出特定語彙時嘴唇所產生之光流場。(a)Lucas-Kanade 演算法，(b)Irani 演算法，(c)我們的演算法，(d)原始序列的某一張。

## 五、計畫成果自評

在本年度的計畫中，我們主要集中心力在研究如何利用特定表情的人臉動態資

訊來進行身分確認工作。目前，我們採用的是主成分分析法，針對每一段包含人臉動態資訊的影像序列，建立一個子空間，並且利用相互子空間法來計算受測者與對照者的夾角（夾角愈小表示愈相似），來判斷其提出之身份是否確為此人。初步的實驗結果顯示此一方法頗具潛力。由於這個方法還沒有真正利用到不同圖框影像之間的時間先後關係，因此透過適當的修改，此一方法的整體效能仍有相當大的改善空間。在未來兩年中，我們將再結合唇形追蹤的動態資訊，進一步改進此系統之整體效能。

## 六、參考文獻

- [1] M. Turk and A. Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, vol. 3, pp. 71-96, 1991.
- [2] W. Zhao, R. Chellappa, and A. Krishnaswamy, "Discriminant Analysis of Principal Components for Face Recognition," in *Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 336-341, 1998.
- [3] L. Wiskott, J.M. Fellous, and C. von der Malsburg, "Face Recognition by Elastic Bunch Graph Matching," *IEEE Trans. On Pattern Analysis and Machine Intelligence*, Vol. 19. pp. 775-779, 1997.
- [4] X. Liu, T. Chen, and B.V.K.V. Kumar, "On modeling variations for face authentication," in *Proceedings of Fifth IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 369-374, May 2002.
- [5] R.O. Duda, P.E. Hart, D.G. Stork, *Pattern classification*, Second edition. John Wiley & Sons. Inc., New York, 2001.
- [6] O. Yamaguchi, K. Fukui, and Ken-ichi Maeda, "Face recognition using temporal image sequence," *Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on*, 14-16 Apr 1998, pp 318-323.
- [7] K. Maeda and S. Watanabe, "A Pattern Matching Method with Local Structure", *Trans. IEICE(D)*, vol.J68-D, No.3, pp. 345-352(1985)
- [8] A. Bjorck and G. H. Golub, "Numerical Methods for computing Angles Between Linear Subspaces", *Mathematics of Computation*, Vol.27, pp. 579-594(1973).
- [9] M. Irani, "Multi-frame Optical Flow Estimation Using Subspace Constraints", *International Conference on Computer Vision*, pp.626-633, 1999.
- [10] T. Kanade and M. Okutomi, "A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, No.9, pp.920-932, 1994.

受測者	資料庫	夾角 1	夾角 2	夾角 3	夾角 4	夾角 5	夾角 6	夾角 7	夾角 8
<b>Session 1</b>									
Yatsai	Allen	15.1261	56.5892	68.6528	70.7796	75.8874	86.0846	87.013	89.9223
Yatsai	Crazy	15.1746	61.3103	72.0037	76.5791	77.6392	82.9625	86.3912	89.9008
Yatsai	Erock	14.7892	52.7614	70.2518	73.5445	77.2063	82.3426	84.7797	89.4103
Yatsai	Godfly	14.3104	62.7899	69.6326	74.3964	76.5328	82.4891	84.848	89.7677
Yatsai	Octa	17.363	49.8645	57.0503	73.2373	81.0328	82.2257	85.1206	89.5638
Yatsai	Yatsai	3.56226	40.5787	48.1972	60.8913	73.1239	77.9121	79.4314	88.4604
Yatsai	Zenki	15.212	56.6714	66.8188	69.9788	75.1852	79.9702	87.7398	89.2436
<b>Session 2</b>									
Yatsai	Allen	15.1808	41.7848	52.1925	67.8493	72.5617	81.0085	85.2585	89.8671
Yatsai	Crazy	14.295	50.7199	59.4723	65.4031	74.4513	77.2478	83.2443	84.856
Yatsai	Erock	13.586	56.1336	63.1037	72.9202	77.5626	81.5932	83.6202	89.7543

Yatsai	Godfly	15.2753	60.6161	63.15	69.5188	76.5818	81.255	84.7983	86.9649
Yatsai	Octa	16.2679	47.5306	52.3611	72.5689	73.9557	80.3949	85.9563	89.159
Yatsai	Yatsai	2.0674	22.1726	38.9787	61.2333	69.8247	77.0793	80.7894	85.6871
Yatsai	Zenki	15.6169	45.2804	58.3738	58.8305	68.3537	74.9051	80.8163	85.5431
<b>Session 3</b>									
Yatsai	Allen	13.889	45.2082	54.8808	61.9026	72.2734	74.243	84.2365	88.9229
Yatsai	Crazy	12.2855	46.0714	56.0498	62.1921	77.8984	80.6854	86.2238	88.8117
Yatsai	Erock	13.1428	50.232	64.5085	70.7815	77.2661	80.2433	85.6941	88.9322
Yatsai	Godfly	11.7334	51.5942	52.4547	64.9881	72.9804	79.479	81.2379	88.3049
Yatsai	Octa	16.5847	51.2569	58.6044	71.1942	78.7897	81.8155	88.3809	89.6836
Yatsai	Yatsai	1.87008	24.8471	35.0366	51.398	67.7377	80.9306	86.5121	88.7081
Yatsai	Zenki	12.9935	35.7316	47.5683	54.5927	69.3047	73.2854	80.747	84.4227

表一：使用者Yatsai在不同天的資料庫所測試的結果。我們可以發現由於主成分分析的原因，愈前面的夾角愈能辨識出該使用者。