

# 基於人臉影像之人類年齡偵測

Jian-Hao Lai(賴健豪) Don-Lin Li(李東霖) Prof. C.T. Lin(林進燈)

*National Chiao Tung University Taiwan, ROC*

[93502052@cc.ncu.edu.tw](mailto:93502052@cc.ncu.edu.tw) [lazybones.ece95g@nctu.edu.tw](mailto:lazybones.ece95g@nctu.edu.tw) [ctlm@mail.nctu.edu.tw](mailto:ctlm@mail.nctu.edu.tw)

Shih-Chun Chou(周世俊)

*Institute of Information Industry Taiwan, ROC*

[benchou@iii.org.tw](mailto:benchou@iii.org.tw)

## 摘要

近年來，年齡偵測在人臉辨識技術中，漸漸成為一項受重視的領域，加上科技的發達，年齡偵測被認為在多媒體通訊、人機介面、居家照護及安全監控等應用上有相當大的發展潛力。本論文提出一個基於線性識別分析(LDA)技術的自動人臉年齡偵測系統。我們參考數種常見於人臉辨識的特徵抽取技術，觀察它們針對年齡特徵所產生的流形(Manifold)，依據實驗結果，LDA(Linear Discriminant Analysis)最適合用來抽取年齡特徵。強健且容易使用的SVM(Support Vector Machine)則是系統中主要的數據分類法。

現有相關領域論文皆使用較高解析度影像作為實驗資料，而本論文為滿足實際應用需求，特別針對低解析度影像來實驗。在只有40x40像素的極低解析度下，本系統仍能提供不遜於使用高解析度影像的準確率。

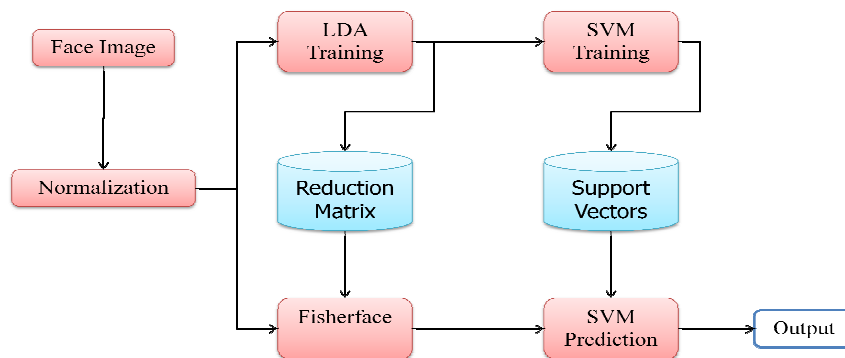
## 一、緒論

人臉影像可以視為一種包含大量資訊的複合訊號，這個複合訊號由身分、性別、年齡跟感情等資訊組成，

在實際人臉影像分析應用上，諸如多媒體通訊、人機介面與安全監控等，扮演重要角色。如何從這龐大的複合訊號中只抽取需要的部分，仍是一門具挑戰性的學問。

目前特徵抽取的方法可以略分為三大類，第一類是基於統計分析之特徵抽取法[1][2][3][4]，第二類為基於頻率分析之特徵抽取法[5]，而第三類則是基於紋理分析之特徵抽取法[6]。由於實際應用上所輸入的大多為低解析度之影像，本論文所設定的實驗環境亦使用低解析度之影像，而基於頻率分析跟基於紋理分析之方法在低解析度下效果較差，故本論文採用基於統計分析之特徵抽取法來抽取年齡特徵。

常用的基於統計分析之特徵抽取法有PCA(Principle Component Analysis)[1][4]、LDA(Linear Discriminant Analysis)[1]、LPP(Locality Preserving Projection)[2]與OLPP(Orthogonal Locality Preserving Projection)[3]四種，藉由觀察它們對年齡特徵產生的流形(Manifold)，依據實驗結果，LDA最適合用來抽取低解析度影像的年齡特徵，並結合SVM，希望建置一個使用低解析度影像的高準確率年齡偵測系統。



圖一、系統架構圖

## 二、系統流程

本節將針對系統流程做詳細說明，如圖一所示，本系統共分為三塊：1) 影像前處理、2) offline 訓練跟 3) online 測試。首先影像前處理的部分，我們用雙線性內插法將原始影像正規化成 40x40 像素的統一大小，方便訓練與測試。Offline 訓練的部分，訓練影像經過正規化，先透過 LDA 訓練出能抽取年齡特徵的降維矩陣，同時將降維後取出的特徵向量放到 SVM 訓練求得支持向量與分割線，根據這一流程完成分類器的訓練。

## 三、年齡特徵抽取

年齡是人臉影像的基本特徵之一，假設年齡影像空間  $X = \{x_i : x_i \in R^D\}_{i=1}^N$  是由  $N$  張人臉影像所構成，其中  $D$  是影像維度，由於人臉影像屬於複合訊號，為了取出跟年齡相關的訊號，本論文透過學習訓練一個低維度的子集合，藉此找到足以表示人臉年齡特徵的子空間  $Y = \{y_i : y_i \in R^d\}_{i=1}^N$ ，其中  $d \ll D$  且與  $X$  為一對一的對應，此對應可視為空

間上的投影，因此年齡特徵抽取的目的在於找到一投影函式  $P$ ，使得  $Y = P(X)$  成立。

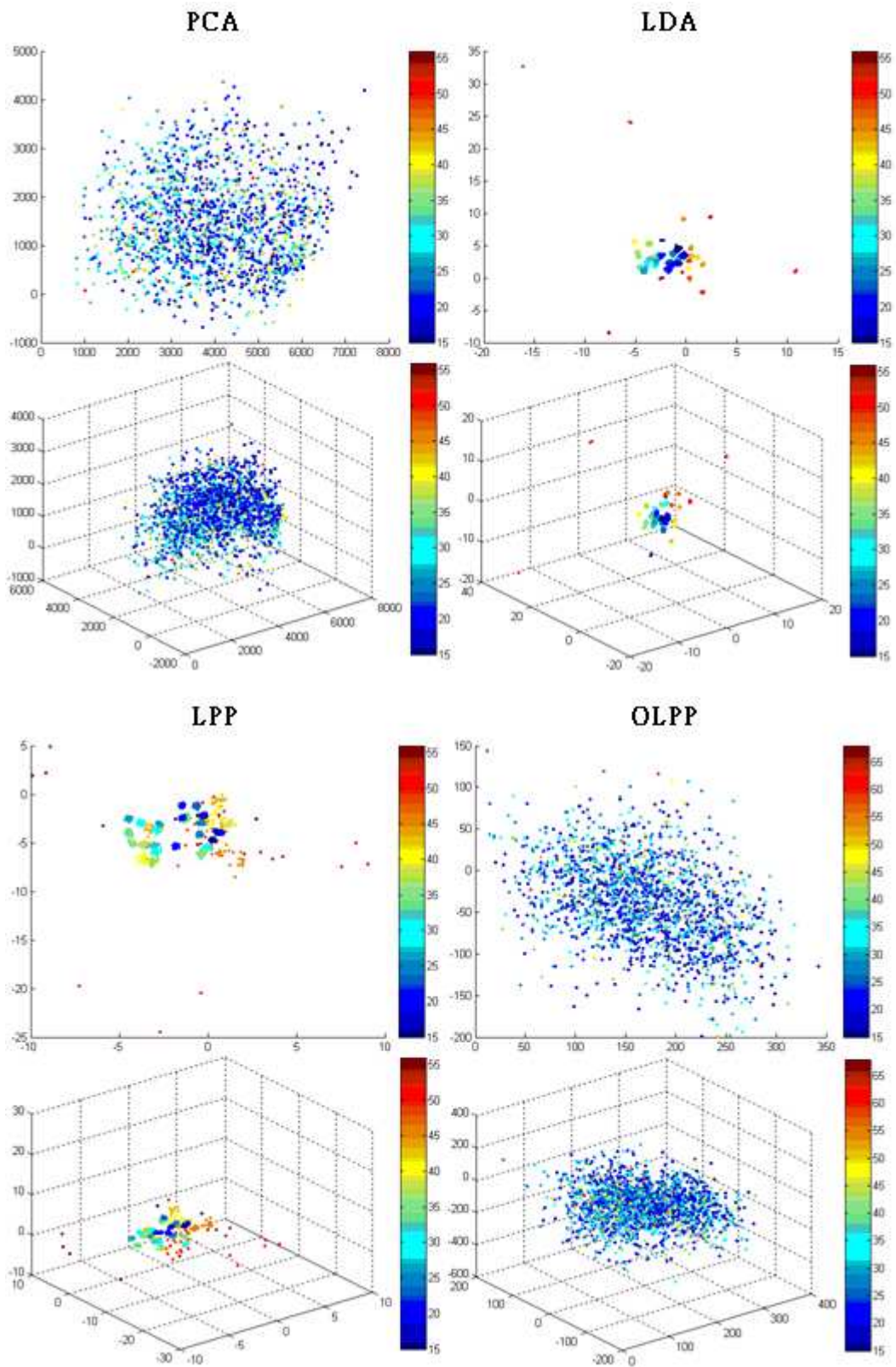
為了此一目的，本節將介紹並比較數種資料降維演算法：

1. *Principal Component Analysis (PCA)*[1][4]：PCA 特徵抽取法是圖形識別領域被廣為使用的經典方法，因此這邊優先提出來實驗比較。本方法的核心在於找到一個能夠最大化投影變異數的子空間投影

$$p = \arg \max_{\|p\|=1} p^T S p, \quad (1)$$

$S = \sum_{i=1}^N (x_i - \mu)(x_i - \mu)^T$  是散射矩陣， $\mu$  是資料平均值， $N$  則是訓練資料總數，簡單來說，就是透過求解訓練資料矩陣的特徵值，並依據特徵值大小排序取出對應的特徵向量作為子空間的投影軸。

2. *Linear Discriminant Analysis (LDA)*：LDA 特徵抽取法的作法與 PCA 雷同，不同的地方在於 LDA 有引入類別資訊，並強化類別內的同質性及類別間的差異性。本方法定義了兩個散射矩陣，一是類別間散射矩陣



圖二、年齡特徵抽取流形比較圖

$S_B = \sum_{i=1}^C n_i (\mu_i - \mu)(\mu_i - \mu)^T$ ， $n_i$  表示第  $i$  類的個數， $\mu_i$  表示第  $i$  類的平均值， $C$  表示總類別數；另一個是類別內散射矩陣  $S_W = \sum_{i=1}^C \sum_{x_k \in X_i} (x_k - \mu_i)(x_k - \mu_i)^T$ ， $x_k$  表示第  $k$  個樣本， $\mu_i$  表示第  $i$  類的平均值藉， $C$  表示總類別數，由定義這兩個散射矩陣來找到最佳投影子空間

$$p = \arg \max_p \frac{|p^T S_B p|}{|p^T S_W p|}。$$

### 3. Locality Preserving Projections (LPP)[2]:

LPP 特徵抽取法具有兩個特性：保留區域結構(local structure)，也就是說保留了資料間的相關性，不同於使用 PCA 或 LDA 等保留全域結構(global structure)的特徵抽取法，LPP 更能有效描述具流形特性的資料；線性投影使得 LPP 可以輕易地新增訓練資料，在訓練樣本數量的需求上也較 PCA 與 LDA 來的少。本方法的核心在於建構相鄰無向圖(adjacency undirected graph)，藉由點的連線與否來表示資料間的關係，再根據距離給定兩點的相關性

$S_{ij} = \exp(-\frac{\|x_i - x_j\|^2}{t})$ ，並定義對角線矩陣  $D$  與拉氏矩陣(Laplacian matrix)

$L$ ，其中  $D_{ii} = \sum_j S_{ij}$ ， $L = D - S$ 。有了這些資訊，便可以找到最佳投影子

$$\text{空間 } p = \arg \min_p \frac{|p^T X L X^T p|}{|p^T X D X^T p|}。$$

4. Orthogonal Locality Preserving Projections(OLPP)[3]: OLPP 特徵抽取法是基於 LPP 特徵抽取法的改良型演

算法，除了 LPP 原本就具有的特性之外，OLPP 還有保留度量結構(metric structure)的功能。基於 LPP 所提出的相鄰無向圖、關連矩陣  $S$ 、對角線矩陣  $D$  與拉氏矩陣  $L$ ，OLPP 的最佳投影子空間定義為

$$p = \arg \min_{p^T X D X^T p = 1} \sum_{i=1}^n \sum_{j=1}^n (p^T x_i - p^T x_j)^2 S_{ij}$$

$n$  表示訓練資料總數， $x_i$  表示第  $i$  個樣本。

本論文採用 MORPH 人臉影像年齡資料庫[7]做為實驗資料庫，為了能夠有效的觀察以上四種特徵抽取法於年齡特徵抽取的效果，MORPH 資料庫內所有人臉年齡影像都當作訓練資料。實驗結果如圖二所示，分別為 PCA、LDA、LPP 與 OLPP 的流形，每一組圖的上半部為最具代表性的兩軸資訊，下半部則是最具代表性的三軸資訊，從圖中不難發現 PCA 抽取出的特徵與年齡的對應是雜亂無章的，而 OLPP 雖然有接近線性的流形，其特徵與年齡的對應依然混亂，至於 LDA 與 LPP 抽取出的特徵，皆能使同年齡的資料聚集在一起，但是 LDA 在鄰近年齡間的投影效果更好，因此本論文選用 LDA 作為年齡特徵抽取的方法，經由 LDA 可將 1600 維的資料降維至 62 維。

## 四、 實驗結果

本節將詳細說明實驗設定、效能評估方式以及實驗結果比較：

### 1. 資料庫

本年齡偵測系統實驗所使用的資料庫是 MORPH 人臉影像年齡資料庫，這個資料庫包含由 515 個人(多種族)



取得的高解析度灰階人臉影像 1690 張，年齡分布介於 15 至 68 歲，圖三是本資料庫的一些樣本。



庫樣本圖

## 2. 實驗設定

本論文採用 leave-one-person-out (LOPO) 的方法進行測試，所謂的 LOPO 就是將同一個人的影像取出來做為測試資料，其餘的部分用來做訓練的一種方法，以本篇使用的資料庫為例，整個實驗流程必須執行 515 次後才能統計結果。本實驗使用的分類器為 SVM，先透過 LDA 將訓練資料的年齡特徵取出，使用 Libsvm[8] 工具訓練出人臉年齡偵測系統的分類器，其中本論文使用 RBF(radial based function) Kernel 作為 SVM 的 Kernel Function，參數設定的部分， $\gamma = 1.2207 \times 10^{-4}$ ，分類誤差懲罰權重  $c = 8.0$ 。

## 3. 結果

年齡偵測系統效能可由下面兩種方法來評估：平均絕對誤差 MAE (mean absolute error) 跟累積分數 CS

(cumulative score)。MAE 的定義是偵測年齡  $\hat{l}_k$  與實際年齡  $l_k$  絕對誤差的平均

值，即  $MAE = \frac{1}{N} \sum_{k=1}^N |l_k - \hat{l}_k|$ ，而 CS 則

定義為  $CS(j) = N_{e \leq j} / N \times 100\%$ ，其中

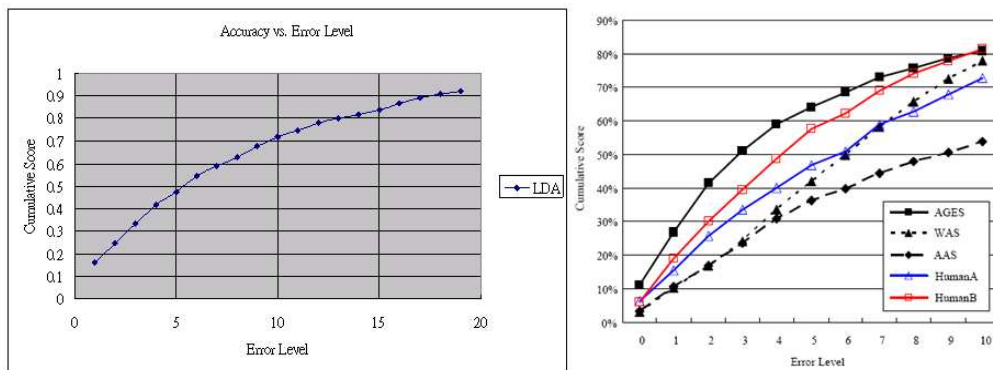
$N_{e \leq j}$  代表測試影像絕對誤差小於  $j$  的數量。

表一是與[9]的 MAE 比較結果，可以看到本論文提出的方法在低解析度下仍優於 WAS 與 ASS 兩個使用高解析度影像的早期經典方法，甚至只略遜於使用高解析度影像的 AGES。

Method	MAE
AGES	6.77
WAS	8.06
ASS	14.83
LDA+SVM	7.95

較表

圖四則是 CS 的比較結果，由圖可知，本論文提出的方法在誤差等級 10 時即可達到 70% 的準確率，誤差等級 13 左右就能達到 80% 的準確率，雖然稍遜於 AGES 與 WAS，但是在低解析度的環境下就能達到如此準確率，實用性較高。



圖四、CS 比較圖

## 五、 結論

人臉年齡偵測仍是人臉辨識中一塊具發展潛力的研究主題，而人臉辨識的核心在於適當的特徵抽取法。為了滿足實際應用的需求，如何在低解析度影像中抽取出適當的年齡特徵就是本系統要達到的目標。使用 LDA 作為特徵抽取法並結合 SVM 來分類可以得到相當不錯的效果，而且具有相當高的實用性。

本系統所使用的 LDA 為全域特徵抽取法的一種，近幾年人臉特徵抽取漸漸有較多的研究傾向使用區域特徵做為辨識的依據，未來本系統將朝著結合區域特徵與全域特徵的優點，藉以建構一個更強健準確的人臉年齡偵測系統的目標邁進。

## 致謝

本研究依經濟部補助財團法人資訊工業策進會「98 年度 互動看板觀眾分析引擎」辦理。

## 參考文獻

- [1] Belhumeur, P. N., J. P. Hespanha, et al. (1997). "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection." *IEEE Transactions on Pattern Analysis and Machine Intelligence* **19**(7): 711-720.
- [2] Xiaofei He, Shuicheng Yan, Yuxiao Hu, Niyogi P., Hong-Jiang Zhang (2005). "Face recognition using Laplacianfaces." *IEEE Transactions on Pattern Analysis and Machine Intelligence* **27**(3): 328-340.
- [3] Deng Cai, Xiaofei He, Jiawei Han, Hong-Jiang Zhang (2006). "Orthogonal Laplacianfaces for Face Recognition." *IEEE Transactions on Image Processing* **15**(11): 3608-3614.
- [4] R. O. Duda, E. H. Peter, and G. S. David, *Pattern Classification*, 2<sup>nd</sup> ed. New York: Wiley Interscience, 2000.
- [5] Yu Su; Shiguang Shan; Xilin Chen; Wen Gao; Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on 14-21 Oct. 2007. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," in *Proc. ECCV, 1998*, vol. 2, pp. 484-498.
- [6] Karl Ricanek, Jr. and Tamirat Tesafaye, "MORPH: A Longitudinal Image Database of Normal Adult Age-Progression", *IEEE 7th International Conference on Automatic Face and Gesture Recognition*, Southampton, UK, April 2006, pp 341-345.
- [7] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," in *Proc. ECCV, 1998*, vol. 2, pp. 484-498.
- [8] C.C. Chang and C.J. Lin, "LIBSVM: a Library for Support Vector Machines," June 14, 2007. Software available at <http://www.csie.ntu.edu.tw/~cjlin/li>

bsvm.

- [9] Xin Geng, Zhi-Hua, *Senior Member, IEEE*, and Kate Smith-Miles, *Senior Member, IEEE* *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, VOL29, NO. 12, DECEMBER 2007.