

# 以改良式區域方向圖樣為基準之臉部表情辨識

賴智錦、邱建璋、謝鴻瑋  
國立高雄大學電機工程學系

**摘要** — 近年來，臉部表情辨識是電腦視覺領域中許多專家學者投入研究的主题之一。脸部表情辨識主要透過人脸影像进行脸部特征提取，再将取出的特征经由分类器进行表情的分类。本论文提出一个以改良式区域方向图样为基准之特征提取方法，再将取得的脸部特征以支持向量机进行分类。实验结果显示我们的方法相较于其他的表情辨识方法，有更好的辨识效能。

**关键词**：脸部表情辨识、改良式区域方向图样、支持向量机

## 一、前言

人的脸部表情非常多样，使得人们可以藉由表情与他人进行互动与交流。在生活中，表情透露出一個人內心的情緒，舉凡：開心、生氣、傷心、或驚訝等，都能從人脸辨识出來。為了減少機器與使用者之間的隔閡，使機器更具有人性，多種人性化的功能相繼而生，脸部表情辨识即是其中一門重要課題。藉由脸部表情辨识技術的研發，可以讓機器學會判讀人的表情，並且強化人機之間的互動。

經過幾十年的發展，以機器辨識人脸表情有了很大的進步，各種辨識方法相繼提出，這些方法大致分為兩類：基於幾何特征的方法與基於外觀特征的方法[1]。幾何特征是指人脸的局部特征訊息，可以簡潔地表示人的脸部結構；換言之，基於幾何特征的方法就是利用人脸的幾何特性，例如：眼睛、嘴巴、鼻子等形狀與距離變化進行分析與辨識。人脸動作編碼系統 (Facial Action Coding System, FACS) [2]是由 Ekman 及 Friesen 所提出，藉由分析臉部的動作，並定義出 46 種動作單元 (Action Unit, AU)，再透過各單元間的組合以識別各種脸部表情。Valstar 等人[3]則提出了 AU 檢測的方法，藉由計算與基準點最相近的 AU 以辨識表情，但該法需要準確的定位及追蹤五官位置才有較好的效能。如今此類的辨識系統大多會加上 AdaBoost 演算法偵測運動單元的位置[4]，藉以提升辨識的準確性。Hamm 等人[5]則從動態的影像序列中取出幾個特定的 AU 特征，再使用 Gabor 小波來表示這些 AU 特征，再以 AdaBoost 分類器進行表情的判別。

基於外觀特征的方法泛指使用全臉影像的特征，依據紋理或像素值的差異進行分析；例如：主成分分析法 (Principal Component Analysis, PCA)[6]、獨立成分分析 (Independent Component Analysis, ICA)[7]、區域特征分析 (Local Feature Analysis, LFA)、線性判別分析 (Linear Discriminant Analysis, LDA)[8]與 Gabor 小波分析[9]。文獻[10]中則探討了一般常用的外觀特征提取方式對於臉

部表情辨識的優劣，結果顯示 Gabor 小波表示法及 ICA 有較佳的效果。Zhang 等人[11]與 Tian[12]也證明了 Gabor 小波相較於幾何特征的方法有更好的辨識效能。

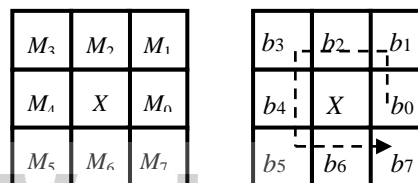
為解決 Gabor 小波表示法耗費大量運算時間及記憶體儲存空間的問題[13]，區域二元圖樣 (Local Binary Patterns, LBP) 逐漸廣用於表達圖像的資訊[14]。Viola 等人[15]證實了 LBP 對光照的變化及對低解析度的影像較具強韌性，辨識效果也較 Gabor 小波佳。然而 LBP 依然存在著特征維度過大與部分特征遺失的問題；因此，Heikkilä 等人提出以 LBP 為基礎，並加入中心點對稱處理的概念[16]。Jabid 等人[17]則是將像素梯度大小變化及方向性加入 LBP 中，提出了所謂的區域方向圖樣 (Local Directional Pattern, LDP)。

在本論文中，我們提出一個方法將表情影像透過改良式區域方向圖樣 (Improved Local Directional Pattern, ILDP) 計算出各個像素点的特征值，藉此表示成人脸表情影像之特征，再以支持向量机 (Support Vector Machine, SVM) [18] 進行脸部表情特征之分类，以達成脸部表情的辨識。

## 二、背景介紹

### 2.1 區域方向圖樣

由於 LBP 僅考慮像素之間灰階值的比較，並沒有引入權重的概念，此舉可能造成部分特征的遺失[19]。為了改善 LBP 之缺點，Jabid 等人[17]提出 LDP 描述方法；此方法的概念是將影像中的每一個像素進行 LDP 運算，藉以擷取出影像的區域紋理特征，再將這些特征以八位元二進制編碼方式來表示。LDP 主要藉由計算不同方位的邊緣響應 (Edge Response)，對每一個像素點進行方向屬性編碼，以像素點右邊位置為 0，依逆時針方向依序編號置位置 7，如圖一所示。邊緣響應值則透過圖二的 Kirsch edge masks 計算後，可得到八個方位的邊緣響應值。



圖一：(a)八個方向的頻率響應位置，(b)LDP 二位元編碼位置

$\begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix}$	$\begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix}$	$\begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$	$\begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$
East $M_0$	North East $M_1$	North $M_2$	North West $M_3$
$\begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix}$	$\begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix}$	$\begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix}$	$\begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix}$
West $M_4$	South West $M_5$	South $M_6$	South East $M_7$

圖二：八方位 Kirsch edge masks 示意圖

85	32	26
53	50	10
60	38	45

圖三：某像素點與其八鄰點之像素值

圖三為某影像中  $3 \times 3$  的區塊內像素值的分佈，經過八個方向的邊際遮罩運算後，可分別得到八種不同方向的邊緣響應值。將響應值取其絕對值，再取出前  $K$  個最大的數，將其所對應之位置設置為 1，其餘位置設置為 0，依序進行編碼即為此像素點的 LDP 編碼，如圖四所示。

Mask index	$M_7$	$M_6$	$M_5$	$M_4$	$M_3$	$M_2$	$M_1$	$M_0$
Mask value	161	97	161	537	313	97	-503	-393
Rank	6	7	5	1	4	8	2	3
Code bit	0	0	0	1	0	0	1	1
LDP code	19							

圖四：LDP 編碼範例 ( $K=3$ )

LDP 編碼可經由公式(1)與(2)的計算而得：

$$LDP_k = \sum_{i=0}^7 b_i(m_i - m_k) \times 2^i \quad (1)$$

$$b_i(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (2)$$

其中  $m_i, i = 0, 1, \dots, 7$  表示 Kirsch edge masks 的八個方向響應位置， $m_k$  為第  $k$  個最具代表性的值。

## 2.2 中心對稱區域二元圖樣

中心對稱區域二元圖樣(center-symmetric local binary pattern, CS-LBP)[16]的觀念是以 LBP 為基礎，引入中心點對稱處理的方式以進行改良。CS-LBP 重新定義了運算比較的規則，將一像素點的八個鄰點，依序進行兩兩對稱式的運算。CS-LBP 運算方式定義如公式(3)與(4)。

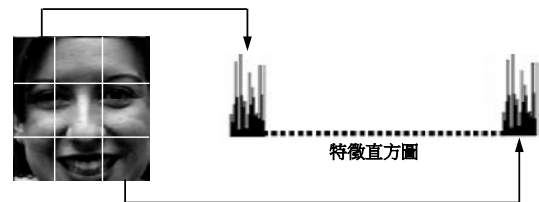
$$CS-LBP_{R,N,T}(x,y) = \sum_{i=0}^{\frac{N}{2}-1} b_i(g_0 - g_{i+(N/2)}) \times 2^i \quad (3)$$

$$b_i(x) = \begin{cases} 1, & x \geq T \\ 0, & x < T \end{cases} \quad (4)$$

其中  $R$  是圓周半徑， $N$  是圓周上的像素點個數， $T$  是門檻值。CS-LBP 運算子是以中心像素值為對稱點，將其鄰點進行對稱式的計算，得到一長度為  $N/2$  的二進制編碼，將其轉為十進制即為該中心點的 CS-LBP 值。

## 2.3 局部紋理特徵擷取

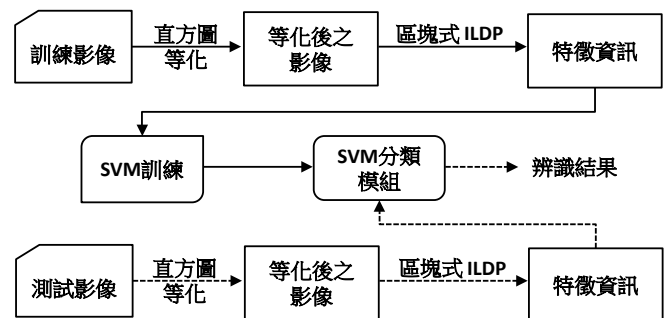
區塊式的特徵建構可以更有效的表示出整張臉部表情影像的細微紋理特徵。區塊式紋理特徵的取得是將臉部表情影像分割成固定大小的子區域，並對每一個子區域進行特徵值擷取；再將每一區域所擷取出的特徵直方圖進行串接，即可成為代表該臉部表情影像的特徵直方圖，其觀念如圖五所示。



圖五：局部紋理特徵擷取示意圖

## 三、所提方法

由於 LDP 的方法，隨著  $K$  值的提高，編碼組合的數目會大幅增加[17]。如果再以區塊式的方式建構特徵向量，特徵維度將再大幅成長，如此龐大的資料對運算時間有著極大影響。為了改良上述缺點，我們提出以改良式區域方向圖樣(ILDP)為基準進行臉部表情辨識，基本觀念是：在取得臉部特徵之前置處理中加入了臉部影像的直方圖等化(histogram equalization)動作，接著進行區塊式 ILDP 運算；如此，將可大幅降低特徵值的維度，進而提升運算上的效能。圖六是我們所提方法的流程。



圖六：所提方法之流程圖

### 3.1 直方圖等化

我們在前置處理中加入了直方圖等化的動作，目的是讓每張影像的灰階值分佈更為平衡，以減少影像因光影的關係影響到整體的辨識效能。直方圖等化是將原始影像中所有像素的灰階值加以平均並分配到各個像素點，使得直方圖盡可能呈現均勻分佈，以提高影像的對比度，得到較清晰的影像。

將影像中的每一可能灰階值視為一隨機變數，則直方圖可視為一離散的機率密度函數(probability density function)，直方圖等化的計算方法如公式(5)所示，

$$S_k = T(r_k) \\ = \sum_{j=0}^k \frac{n_j}{n} = \sum_{j=0}^k \Pr(r_j), 0 \leq r_k \leq 1, k = 0, \dots, L-1 \quad (5)$$

其中  $L$  是影像中使用的灰階總數， $r_k$  為原始直方圖分布中第  $k$  種灰階的隨機變數， $n_j$  是第  $j$  種灰階的像素數目， $n$  是整張影像像素點的總數， $\Pr(r_k)$  是隨機變數  $r_k$  的離散機率密度函數， $S_k$  是第  $k$  種灰階的隨機變數  $r$  等化後的隨機變數值。

### 3.2 改良式 LDP

為了改良 LDP 特徵維度過於龐大的缺點，我們引入了以中心點對稱處理的概念對影像進行編碼，此舉可大幅度的將特徵維度進行縮減，有利於降低運算上所耗費的時間。ILDP 運算方式定義如公式(6)與(7)，

$$ILDP = \sum_{i=0}^{N/2-1} s(mv_i - mv_{i+(N/2)}) \times 2^i \quad (6)$$

$$s(x) = \begin{cases} 1, & x \geq Th \\ 0, & x < Th \end{cases} \quad (7)$$

其中  $mv_i, i = 0, 1, \dots, 7$  表示經過 Kirsch edge masks 計算後得到的八個方向響應值， $N$  為響應值個數， $Th$  為門檻值。ILDP 運算子是以某像素點為中心，將其八個鄰點的響應值取其絕對值後，接著進行對稱式的計算，得到一長度為  $N/2$  的二進制編碼，再將其轉換為十進制即為該中心點的 ILDP 值。圖七是圖三範例之 ILDP 編碼範例。

Mask index	$M_7$	$M_6$	$M_5$	$M_4$	$M_3$	$M_2$	$M_1$	$M_0$
Mask value	161	97	161	537	313	97	-503	-393
Code bit	-	-	-	-	1	1	1	0
ILDP code	14							

圖七：ILDP 編碼範例( $Th=0$ )

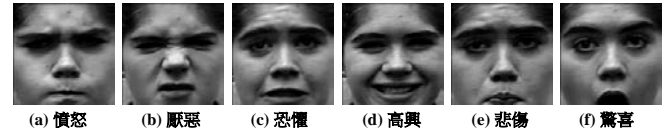
## 四、實驗結果

在驗證所提方法的效能上，我們以 Cohn-Kanade 資料庫[1]實際進行前述方法的測試。資料庫中有七種表情類別，分別是：憤怒(Anger)、厭惡(Disgust)、恐懼(Fear)、高興(Joy)、悲傷(Sadness)、驚喜(Surprise)、與無表情(Neutral)，其中六類表情範例如圖八所示。我們共選取了 1280 張影像，各類別表情的樣本數如表 I 所示。我們採取與文獻[14]相同之 10 倍交叉驗證法(10-fold cross-validation)，進行所提方法的辨識率之計算。

表 I：七種表情類別之樣本數

Anger	Disgust	Fear	Joy	Sadness	Surprise	Neutral
108	120	99	282	126	225	320

為了瞭解直方圖等化對辨識率之影響及其穩定性，在實驗中，我們測度了在不同的門檻值其辨識效果之變化，表 II 為實驗結果。該實驗是以區塊切割數  $2 \times 2$  進行測度，從實驗數據可觀察出一個合適的門檻值，將有助於辨識度的提升；並且加入直方圖等化的處理後，可以讓後續的相關處理較不易受到門檻值的影響。從表 II 中顯示，門檻值訂為 120 時可獲取最佳的辨識效果；因此，在區塊分割數對辨識率之影響實驗中，我們便將門檻值設定為 120 以進行實驗。從表 III 的數據中顯示在合適的區塊分割數內，將有助於辨識率的提升，但區塊切割數大至一定程度時，將造成某些區塊內的特徵資訊不足，反而造成辨識率下降的情形。



圖八：六類表情範例圖

表 II：門檻值與直方圖等化對辨識率之影響

門檻值	未等化處理之辨識率(%)	經等化處理之辨識率(%)
50	93.98	93.59
60	94.76	94.06
70	93.75	94.45
80	93.67	94.53
90	94.76	94.92
100	94.45	95.07
110	92.34	95.22
120	93.59	95.39
130	92.96	94.92
140	92.81	94.53
150	91.79	94.29

表 III：區塊分割數對辨識率之影響

區塊數	經等化處理之辨識率(%)
$2 \times 2$	95.39
$3 \times 3$	97.50
$4 \times 4$	98.12
$6 \times 6$	98.75
$8 \times 8$	99.06
$9 \times 9$	99.21
$10 \times 10$	98.90

表 IV 為經本論文所提方法所獲取之混淆矩陣，從中可觀察出恐懼、高興、和驚喜這些較具有鑑別度的表情，辨識結果較佳。然而，厭惡和悲傷因表情特徵較不明顯，則有 2% 的機率被分到無表情類別。此外，我們將無表情去除後的六類表情進行分類的測試，結果如表 V 所示。實驗結果顯示憤怒和悲傷可能因為兩者的特徵較為相似，所以有混淆的情形發生。表 VI 則是我們的方法在考慮不同的區塊數之下，六類表情與七類表情之效能比較。數據顯示在去除無表情後可讓辨識率些許提升，此外，我們仍可以觀察出，在合適的區塊分割數內，將有助於提升辨識率。

表 IV：本論文方法所得之七種表情混淆矩陣(%)

	Anger	Disgust	Fear	Joy	Sadness	Surprise	Neutral
Anger	98.67	0	0	0	0	0	1.33
Disgust	0	97.62	0	0	0	0	2.38
Fear	0	0	99.07	0	0	0	0.93
Joy	0	0	0	100	0	0	0
Sadness	0	0	0	0	97.62	0	2.38
Surprise	0	0	0	0	0	99.07	0.93
Neutral	0.31	0	0.63	0.31	0.94	0	97.81

表 V：本論文方法所得之六種表情混淆矩陣(%)

	Anger	Disgust	Fear	Joy	Sadness	Surprise
Anger	99.07	0	0	0	0.93	0
Disgust	0	100	0	0	0	0
Fear	0	0	100	0	0	0
Joy	0	0	0	100	0	0
Sadness	0.79	0	0	0	99.21	0
Surprise	0	0	0	0	0	100

表 VI：六類表情與七類表情之效能比較

區塊數	六類表情辨識率(%)	七類表情辨識率(%)
2×2	98.33	95.39
3×3	99.27	97.50
4×4	99.47	97.96
6×6	99.65	98.75
8×8	99.65	99.06
9×9	99.75	99.21
10×10	99.65	98.90

為了證明我們所提的研究方法具有一定的準確性，我們與文獻[14][17][20]所提方法進行效能比較，表 VII 為相關實驗數據。文獻[14]是使用區塊式 LBP 進行臉部紋理特徵擷取，再使用 AdaBoost 演算法訓練出較具辨識度的區域。文獻[17]則是採用區塊式 LDP 進行臉部紋理特徵擷取，再對重點特徵區塊給予較高權重值的方式。文獻[20]則以 AdaBoost 演算法對各類表情訓練出對分類結果較有幫助的區域後，再使用 LBP 及權重式區域方向圖樣特徵(WLDP)對這些區域進行特徵值的擷取，最後將兩種特徵結合以彌補辨識上的不足。

表 VII：本論文方法與其他方法之比較

方法	辨識率(%)
Boosted-LBP [14]	91.4
LDP [17]	93.4
Boosted-LBP +WLDP [20]	91.1
Ours	99.2

## 結論

在人臉表情辨識系統中，一個優異的特徵擷取方法是影響系統效能的關鍵。本論文中，我們改善了 LDP 特徵維度過大的問題，並在改善的同時將其效能進一步提升。本論文採用對整張臉部影像進行特徵擷取的方式，一些較不具辨識度的部分也一併進行擷取，進而導致重點特徵資訊有機會被其較不具辨識度的資訊所混淆。未

來我們將朝向對影像特徵進行篩選，也就是只取出較具辨識度的部分進行特徵值的擷取，或是在五官的重點部分，例如：眼睛、嘴角、眉毛等部分區塊給予較高的權重，應該有助於效能的提升。

## 參考文獻

- [1] Y. Tian, T. Kanade, and J. F. Cohn, *Handbook of Face Recognition*, 2nd Ed., Springer, 2005, pp. 487-520.
- [2] P. Ekman and W. V. Friesen, "Facial action coding system," *Consulting Psychologists Press*, College Avenue, Palo Alto, CA, 1978.
- [3] M. Valstar, I. Patras, and M. Pantic, "Facial action unit detection using probabilistic actively learned support vector machines on tracked facial point data," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2005, vol. 3, pp. 76-84.
- [4] M. Valstar and M. Pantic, "Fully automatic facial action unit detection and temporal analysis," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2006, pp. 149-156.
- [5] J. Hamm, C. G. Kohler, R. C. Gur, and R. Verma, "Automated facial action coding system for dynamic analysis of facial expressions in neuropsychiatric disorders," *Journal of Neuroscience Methods*, vol. 200, no. 2, pp. 237-256, 2011.
- [6] M. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 1991, pp. 586-591.
- [7] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, "Face recognition by independent component analysis," *IEEE Trans. Neural Network*, vol. 6, no. 13, pp. 1450-1464, 2002.
- [8] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, 1997.
- [9] M. J. Lyons, J. Budynek, and S. Akamatsu, "Automatic classification of single facial images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 21, no. 12, pp. 1357-1362, 1999.
- [10] G. Donato, M. Bartlett, J. Hager, P. Ekman, and T. Sejnowski, "Classifying facial actions," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 21, no. 10, pp. 974-989, 1999.
- [11] Z. Zhang, M. J. Lyons, M. Schuster, and S. Akamatsu, "Comparison between geometry-based and gabor-wavelets-based facial expression recognition using multi-layer perceptron," in *Proc. IEEE International Conference on Automatic Face & Gesture Recognition*, Apr. 1998, pp. 454-459.
- [12] Y. Tian, "Evaluation of face resolution for expression analysis," in *Proc. CVPR Workshop on Face Processing in Video*, 2004, pp. 82-88.
- [13] M. S. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J. Movellan, "Recognizing facial expression: machine learning and application to spontaneous behavior," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2005, vol. 2, pp. 568-573.
- [14] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," *Image and Vision Computing*, vol. 27, no. 6, pp. 803-816, 2009.
- [15] P. Viola and M. J. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137-154, 2004.
- [16] M. Heikkilä, M. Pietikäinen, and C. Schmid, "Description of interest regions with local binary patterns," *Pattern Recognition*, vol. 42, no. 3, pp. 425-436, 2009.
- [17] T. Jabid, M.H. Kabir, and O. Chae, "Robust facial expression recognition based on local directional pattern," *ETRI Journal*, vol. 32, no. 5, pp.784-794, 2010.
- [18] C. C. Chang and C. J. Lin, LIBSVM: a library for support vector machines. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>, 2001.
- [19] X. Feng, A. Hadid, and M. Pietikäinen, "A coarse-to-fine classification scheme for facial expression recognition," in *Proc. International Conference on Image Analysis and Recognition*, 2004, vol. 3212, pp. 668-675.
- [20] 莊順旭，基於權重式區域方向特徵和區域二元化形態之表情辨識方法。中華大學資訊工程學系碩士論文，99年8月。