

# 結合人臉紋理與邊緣資訊之性別分類

賴智錦 吳信融 林致圻 林伯豪

國立高雄大學電機工程學系

cclai@nuk.edu.tw, {wux4030, ssn4010302, borhaur}@gmail.com

## 摘要

隨著資訊科技地進步，人類的生活型態跟著產生變化，以人臉特徵進行相關辨識的技術也隨著孕育而生。由於相關技術具有無限的發展潛力與應用層面，使得人類逐漸實現了智慧型的生活型態。本論文提出一個植基於臉部紋理與邊緣資訊的人臉性別辨識方法，該法是找出臉部適合表示性別的特徵，再計算出影像中各像素點的特徵值，最後利用支持向量機進行這些特徵的分類，藉此判定該影像之所屬性別。實驗結果顯示，本論文的方法可以獲得不錯的辨識效果。

**關鍵詞：**人臉性別分類；臉部特徵資訊；支持向量機。

## 1. 前言

在日常生活中，很多人都曾經接觸過人臉辨識的技術，這項技術被廣泛的運用在公共安全、商業分析、安全監控、身分認證、與通行控制等方面。最近，隨著恐怖主義的崛起，世界各國機場的出入境處則相繼導入人臉辨識系統，希望藉由自動化的人臉辨識系統提高反恐的安全檢查能力。

與人臉辨識相關的研究橫跨許多面向，例如：年齡的估測、性別與表情的辨識、與人種分類等。由於人臉相關的辨識只需透過攝影照相設備，即可獲取偵測需要的資料，使得這項技術成為目前最方便的辨識方法之一。利用人臉影像進行分析並擷取出特徵，以判定其性別歸屬，則是電腦視覺領域中熱門的研究主題。

人臉性別辨識技術，依據所取特徵的不同，大致可以分成兩大類：基於幾何特徵或外表特徵。Alexandre[1]提出不同的看法，其將特徵擷取分成基於形狀特徵及基於紋理特徵兩部分，在形狀特徵部分使用類似梯度方向直方圖之邊緣方向直方圖，輸入的影像是沒有經過標準化直方圖處理之灰階影像，在紋理特徵部分則使用區域二元圖樣(local binary pattern,

LBP)來編碼，並採用決策融合(decision fusion)的方式融合在單一尺度下不同分類器訓練後的決定或者是在不同尺度下所獲取的特徵。Lian and Lu[2]考慮多視角並結合形狀和紋理特徵資訊的方法來改進以 LBP 代表臉部的特徵資訊，他們認為此種方法可以描述三種局部性的層級：在像素層級時直方圖的標籤代表影像的資訊，接下來標籤會被加總以產生區域層級的資訊，最後將區域影像直方圖串接起來以產生整個臉部的資訊。Ylioinas 等人[3]則認為 LBP 是一個只有提取臉部影像紋理特徵而忽略灰階對比度的演算法，他們提出在影像資訊中的灰階度是獨立且可以旋轉，而對比度則反之，因此該兩種資訊在 LBP 中可形成互補以改進其效能。

Makinen and Raisamo [4]系統性的透過實驗方式，對於四種性別辨識的方法進行效能的比較；此外，對於臉部影像進行縮放、旋轉、與轉換等處理時，分析性別辨識所使用的分類器之效能。Makinen and Raisamo [5]透過分析 FERET 及 WWW 兩個資料庫中的人臉影像是否有頭髮，進而評估某些性別辨識方法之辨識率；同時他們也指出此種方法並不適用於 WWW 資料庫，因為人臉影像正規化是相當重要的一環，包括位置和旋轉，而資料庫的影像背景差異性也會使辨識率下降。Shen 等人[6]提出使用矩形特徵向量(rectangle feature vector, RFV)辨別人臉性別影像，RFV 具有計算速度快，與可在人臉影像的局部區域變化進行有效編碼的雙重優點。Verschae 等人[7]提出以 AdaBoost 與 domain-partitioning based 分類器為基準的性別分類架構，該架構可在動態的環境中，同時達成高速的處理與訓練速度。Shan [8]提出以 LBP 作為臉部特徵之描述，並以 AdaBoost 挑選出具有辨識能力之特徵集合，最後交由支持向量機進行性別的辨識。

針對以臉部影像進行性別辨識之研究，本論文提出一個植基於臉部紋理與邊緣資訊的人臉性別辨識方法，該法是以雙向區域樣板圖樣(bidirectional local template patterns, B-LTP)

為基準，藉以找出適合辨識的臉部特徵。我們的方法用於計算影像中各個像素點的特徵值，並將前述特徵描述子所表示之直方圖進行串接，藉此表示為人臉性別影像之特徵向量，最後將此特徵向量交由支持向量機(support vector machine, SVM)進行辨識。實驗結果顯示，相較於其他類似的方法，本論文的方法可獲得不錯的性別辨識效能。

本論文架構如下：第二節回顧本論文所使用之相關技術，第三節介紹本論文所提出的方法，第四節是相關實驗結果與討論，第五節為結論與未來研究方向。

## 2. 背景介紹

### 2.1 雙向區域模板圖樣

受到樣板直方圖(histogram of template, HOT)與中心對稱區域二元圖樣(center-symmetric local binary patterns, CS-LBP)的觀念啟發，Xu 等人[9]提出了雙向區域樣板圖樣(bidirectional local template patterns, B-LTP)的觀念，並已成功地應用在行人偵測的研究主題。

B-LTP 主要由四個基本樣板所構成，這些樣板主要是考慮在某一區域裡，中心點像素和對角像素點之間的紋理資訊和梯度資訊的差異；因此，可視為是影像中某種區域結構特徵的表現。此外，這些樣板亦可反應影像裡的邊緣資訊，與突出區域的紋理資訊。有此四個基本樣板，使用者便可以定義出諸多的數學公式，藉此分別用於計算影像中的梯度與紋理資訊。此四個樣板的範例，如圖 1 所示。

針對影像中紋理的部分，B-LTP 可以下述公式予以定義：

$$B-LTP_1 = S(g(P_i), g(P), g(P_{i+4})) \times 2^i, \quad (1)$$

$$S(g(P_i), g(P), g(P_{i+4})) = \begin{cases} 1, & (g(P_i) \geq g(P) \cap g(P) \geq g(P_{i+4})) \cup \\ & (g(P_{i+4}) \geq g(P) \cap g(P) \geq g(P_i)) \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

$$B-LTP_2 = g(P) \geq g(P_i) \cap g(P) \geq g(P_{i+4}), \quad (3)$$

$$B-LTP_3 = \arg \max_k \{g(P)_k + g(P_i)_k + g(P_{i+4})_k\}, \quad (4)$$

其中， $g(P)$ 表示像素  $P$  的灰階值。至於影像中梯度大小的資訊，則以下述公式考慮之：

$$B-LTP_4 = S(Mag(P_i), Mag(P), Mag(P_{i+4})) \times 2^i, \quad (5)$$

$$B-LTP_5 = Mag(P) \geq Mag(P_i) \cap g(P) \geq g(P_{i+4}), \quad (6)$$

$$B-LTP_6 = \arg \max_k \{Mag(P)_k + Mag(P_i)_k + Mag(P_{i+4})_k\}, \quad (7)$$

其中， $Mag(P)$ 表示像素  $P$  的梯度大小。

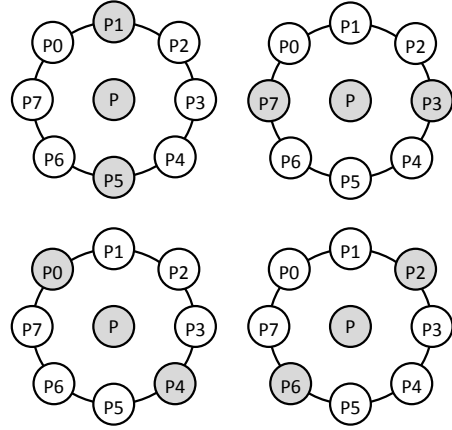


圖 1. B-LTP 的四個基本樣板[9]

以此方式，對於影像中紋理資訊的部分，我們可以得到一個維度大小是 24 直方圖；其中，前 16 個數值是由 B-LTP1 的數值決定，接著 4 個數值則是由 B-LTP2 的條件決定，最後的 4 個數值則是由 B-LTP3 決定。梯度大小的計算方式與紋理的部分雷同，所以亦可產生維度同為 24 的直方圖。最後將梯度大小的部分和紋理部分的直方圖串接在一起，可得到維度是 48 的直方圖。然而，在文獻[9]中，作者並未對梯度大小的計算提供詳細的說明或採行的方法；因此，在本論文中，我們將以 Canny 運算子[10]作為計算影像中邊緣資訊的依據。

### 2.2 Canny 運算子

Canny 演算子[10]是基於下述三個準則以進行影像中的邊緣偵測：

1. 具有良好的偵測性：能夠找到所有邊緣，且對雜訊具有良好的抵抗力。
2. 邊緣點具有好的局部性：所偵測到的像素邊緣間的距離最小，真實邊緣的像素也最小。
3. 單一邊緣響應：每個邊緣只會有一個偵測結果。

依據前述準則，Canny 的邊緣偵測步驟如下：

#### 2.2.1 高斯濾波

首先利用高斯濾波器抑制雜訊。公式(8)是

標準差為 1.4 的高斯平滑濾波器。

$$G = \frac{1}{159} \begin{bmatrix} 2 & 4 & 5 & 4 & 2 \\ 4 & 9 & 12 & 9 & 4 \\ 5 & 12 & 15 & 12 & 5 \\ 4 & 9 & 12 & 9 & 4 \\ 2 & 4 & 5 & 4 & 2 \end{bmatrix} \quad (8)$$

### 2.2.2 尋找影像的梯度值

使用類似 Sobel 方法或其他類似的遮罩，求取影像中各像素的梯度強度與方向。

### 2.2.3 非最大值抑制

此步驟係將前一步驟所得影像中較為模糊的邊緣予以銳利化，其作法是將梯度影像中的區域最大值予以保留，並刪除其他不相關的邊緣資訊。

### 2.2.4 雙門檻值

經過非最大值抑制作用後，所遺留下來的點大部分已經是真正的邊緣點；然而，仍有少許是雜訊被誤判的邊緣資訊。為了避免此種情形發生，Canny 採用了雙門檻值，用以區隔強邊緣點或示弱邊緣點。

### 2.2.5 決定正確的邊緣軌跡

將邊緣點進行區分後，強邊緣點可立即標示為最終的邊緣影像；至於，弱邊緣點僅供於連接至強邊緣點的參考。

## 2.3 區域紋理特徵的擷取

為了有效地表示整張人臉影像中細微紋理的組成，本論文採取區塊式方法建構區域紋理特徵。所謂區塊式特徵的擷取與建構是先將人臉影像分割成固定大小且不重疊的子區塊，針對每個子區塊進行 B-LTP 特徵的計算與相關直方圖的建構，最後將全部子區塊的直方圖串接在一起，形成單一且代表該張臉部表情影像的特徵直方圖，其觀念如圖 2 所示。

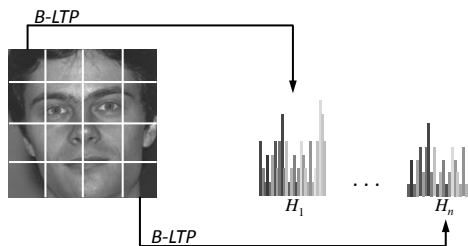


圖 2. 局部紋理特徵擷取示意圖

## 3. 研究方法

本論文所提的臉部性別辨識方法，乃是利用上述理論擷取臉部具有代表性之紋理與邊緣特徵，再將這些特徵送交 SVM，以進行男女性別的辨識。首先我們需將臉部影像進行區塊式 B-LTP 的運算。為了求得區域代表性之臉部紋理，在進行 B-LTP 的運算時，我們會把影像切割成  $N \times N$  個不重疊的子區塊以進行特徵直方圖的運算。最後將各區塊所求取之 B-LTP 特徵直方圖串接一起，形成一個延伸特徵直方圖，我們以此作為輸入臉部表情影像之特徵向量。

圖 3 是本論文所提方法之系統流程圖，可分為訓練與測試兩部分。在訓練部分，每一張影像經上述方法提取相關特徵後，利用 SVM 將欲訓練之性別樣本進行分類訓練，可獲得一個訓練好的 SVM 模型。在測試部分，將欲測試之影像經上述方法提取特徵後送入訓練好的 SVM 模型，進行預測並獲得辨識結果。

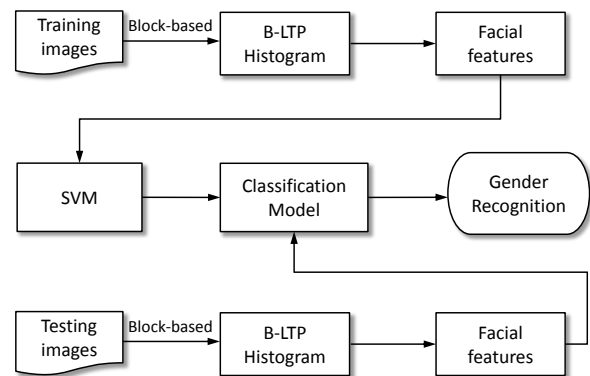


圖 3. 本文所提人臉性別辨識系統流程圖

## 4. 實驗結果

為了驗證本論文所提方法之效能，我們以 FERET[11]和 BioID[12]兩個臉部資料庫作為測試的對象。在 FERET 資料庫的影像部分，我們選取了男性影像 1704 張，女性影像 987 張；在 BioID 資料庫部分，則考慮男性影像 945 張，女性影像 538 張。在這兩個資料庫中的所有影像都經過臉部偵測及提取後，轉成大小皆為  $128 \times 128$  像素的灰階影像，供後續處理與辨識之用。本論文試驗的方法為 10 次交叉驗證法(10-fold cross-validation)。

所謂 10 次交叉驗證法是將資料集隨機分成十等分，每次選取其中九個資料集當作訓練

樣本，剩餘一個資料集當作測試樣本，並且重複上述動作 10 次，直到每個資料集都曾被選為測試樣本後結束。我們將所有性別影像樣本隨機分成十等分，每等分皆有兩類樣本，以 10 倍交叉驗證法進行訓練及測試，所求取的平均數值為該法的辨識率。

本論文之實驗，是以 LIBSVM[13]作為 SVM 分類運算之工具。由於 SVM 的原始版本是以 CPU 進行運算，所以在訓練階段會花費大量的時間；所以 LIBSVM 的作者將原有版本改以 GPU 進行運算，如此可以大幅減少訓練階段所需的時間。因此，本論文所有實驗皆以 GPU 版本的 LIBSVM[14]進行。

#### 4.1 實驗一

為了瞭解使用區域紋理與梯度特徵擷取時，區塊數目對於辨識結果的影響，我們先以 FERET 資料庫作為測試對象，實驗不同的區塊數目並觀察其辨識效能。相關數據如表 1 所示。

從表 1 的數據中我們可以發現，隨著區塊數目的增加，可以獲得更多的紋理資訊，進而提升辨識的效果；當區塊數目為  $6 \times 6$  時，可以得到最高的辨識率。雖然區塊數目的增加，提供後續辨識更多的特徵資訊，然而，特徵資訊的維度亦隨之增加，進行樣本訓練的時間也隨之拉長。因此，區塊數目的適當數量與訓練時間的長短，將是設計一個辨識系統必須考量的兩項因素。

表 1. 不同區塊數目之辨識率(%) -- 以 FERET 資料庫為例

區塊數目	2×2	3×3	4×4	5×5	6×6
辨識率	95.57	97.51	97.99	98.32	98.92

#### 4.2 實驗二

為了得知我們的方法與其他方法在效能上的差異，我們實作文獻上以區域紋理為基準或變形之方法，並以 FERET 資料庫進行比較，結果如表 2 所示。實驗數據顯示，我們的方法相較於其他現有的方法，不論是在男女個別的辨識，或是整體的辨識率，都可以獲得更令人滿意的辨識效果。

表 2. 本論文與其他方法之辨識率比較

方法	辨識率(%)		
	Overall	Male	Female
LBP + SVM[15]	93.38	95.73	89.78
PPPTF [16]	91.76	96.14	79.84
LDP [17]	95.05	94.81	95.33
Ours	98.92	99.00	98.78

#### 4.3 實驗三

除了以 FERET 為實驗對象之外，我們另外考慮 BioID 為測試效能的對象，以驗證當影像的取相對象或方式不同時，我們的方法是否依舊可以獲得不錯的辨識效果。如同在實驗一的目的，我們先驗證不同的區塊數目，對於辨識效果的影響性，實驗數據如表 3 所示。

從表 3 的數據中我們可以發現，雖然隨著區塊數目的增加，可以獲得更多的紋理資訊，進而提升辨識的效果。然而，在此性別資料庫，當我們的方法在區塊數目為  $7 \times 7$  時，已經得到最高的辨識率；如果再繼續增加區塊數目，將無法提升辨識率。因此，不同的資料庫適用的區塊數目有著明顯的不同。

表 3. 不同區塊數目之辨識率(%) -- 以 BioID 資料庫為例

區塊數目	2×2	3×3	4×4	5×5	7×7
辨識率	99.39	99.25	99.46	99.66	99.80

#### 4.4 實驗四

為了得知我們的方法與其他方法在 BioID 資料庫效能上的差異，我們參考了文獻中以 BioID 資料庫作為測試的方法，結果如表 4 所示。雖然比較的方法在影像張數、環境的設定都略有不同，直接比較相關數據較為不妥；但是，實驗的結果仍可反應該方法的效能。從實驗數據顯示，我們的方法比其它的方法，仍可以獲得較令人滿意的辨識效果

表 4: 本論文與其他方法之辨識率比較

方法	辨識率(%)		
	Overall	Male	Female
MSLBP + EDA [18]	95.79	98.50	92.08
SICoNNe [19]	86.40	89.20	81.40
Ours	99.80	100.00	99.44

## 5. 結論

影響人臉性別辨識系統效能的重要關鍵，在於影像特徵表示法的優劣；該特徵表示必須能夠同時滿足特徵維度低，且具有高辨識率的雙重特性。本論文提出的方法是將影像透過區塊式雙向區域模板圖樣，計算出各個像素點的紋理與邊緣特徵，並將其所表示之直方圖當作擷取出來的特徵，以此作為該張影像之特徵資訊。為了取得更詳細的影像結構資訊，我們藉由將影像分割成許多不重疊的區塊，個別進行相關區塊的特徵擷取。由於本文所提之方法，是以整張人臉進行考慮，某些較不具辨識效果的區塊亦會一併進行計算，因此，朝向人臉中較具有鑑別度的部分來研究，是我們未來可以努力的方向。例如：眼形、鼻頭、嘴角或眉毛等區域，再進行特徵值擷取的動作，以達到提高辨識效果的目的。

## 致謝

本研究承蒙行政院科技部計畫補助，特此致謝。研究計畫編號：  
MOST 104-2221-E-390-018-MY2。

## 參考文獻

- [1] L. A. Alexandre, "Gender recognition: A multiscale decision fusion approach," *Pattern Recognition Letters*, vol. 31, no. 11, pp. 1422-1427, 2010.
- [2] H.-C. Lian and B.-L. Lu, "Multi-view gender classification using local binary patterns and support vector machines," in J. Wang, Z. Yi, J. M. Zurada, B.-L. Lu, and H. Yin (Eds): *ISNN 2006, LNCS 3972*, pp. 202-209, 2006.
- [3] J. Ylioinas, A. Hadid, and M. Pietikainen, "Combining contrast information and local binary patterns for gender classification," in A. Heyden, and F. Kahl (Eds): *SCIA 2011, LNCS 6688*, pp. 676-686, 2011.
- [4] E. Makinen and R. Raisamo, "Evaluation of gender classification methods with automatically detected and aligned faces," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 30, no. 3, pp. 541-547, 2008.
- [5] E. Makinen and R. Raisamo, "An experimental comparison of gender classification methods," *Pattern Recognition Letters*, vol. 29, no. 10, pp. 1544-1556, 2008.
- [6] B.-C. Shen, C.-S. Chen, and H.-H. Hsu, "Fast gender recognition by using a shared-integral-image approach," in *Proc. IEEE Intel Conf. Acoustics, Speech, and Signal Processing*, Taiwan, 2009, pp. 521-524.
- [7] R. Verschae, J. Ruiz-del-Solar, and M. Correa, "Gender classification of faces using Adaboost," in *Proc. 11th Iberoamerican Congress in Pattern Recognition, CIARP 2006, LNCS 4225*, Mexico, 2006, pp. 68-78.
- [8] C. Shan, "Learning local binary patterns for gender classification on real world face images," *Pattern Recognition Letters*, vol. 33, no. 4, pp. 431-437, 2012.
- [9] J. Xu, N. Jiang, and S. Goto, "Pedestrian detection based on bidirectional local template patterns," in *Proc. 20th European Signal Processing Conference*, 2012, pp. 400-404.
- [10] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679-698, 1986.
- [11] The **Facial Recognition Technology (FERET) Database**. Available: [http://www.itl.nist.gov/iad/humanid/feret/feret\\_master.html](http://www.itl.nist.gov/iad/humanid/feret/feret_master.html)
- [12] The **BioID Face Database**. Available: <https://www.bioid.com/About/BioID-Face-Database>
- [13] C.-C. Chang and C.-J. Lin, LIBSVM: A Library for Support Vector Machines. Available: <http://www.csie.ntu.edu.tw/~cjlin/libsvm>, 2001.
- [14] GPU-accelerated LIBSVM package. Available: <http://mklab.itl.gr/project/GPU-LIBSVM>
- [15] C. Shan, "Gender classification on real-life faces," in *Proc. 12th Intel Conf. Advanced Concepts for Intelligent Vision Systems, LNCS 6475*, Sydney, 2010, pp. 323-331.
- [16] H. Lu, Y. Huang, Y. Chen, and D. Yang, "Automatic gender recognition based on pixel-pattern-based texture feature," *J. Real-Time Image Proc.*, vol. 3, issue 1, pp. 109-116, 2008.
- [17] T. Jabid, M. H. Kabir, and O. Chae, "Gender classification using local directional pattern (LDP)," in *Proc. 2010 Intel Conf. Pattern Recognition*, Istanbul, Turkey, 2010, pp.

2162-2165.

- [18] R. Azarmehr, R. Laganieri, W.-S. Lee, C. Xu, and D. Laroche, "Real-time embedded age and gender classification in unconstrained video," in *Proc. The Eleventh IEEE Embedded Vision Workshop*, Boston, MA, USA, 2015, pp. 57-65.
- [19] F. H. C. Tivive and A. Bouzerdoum, "A gender recognition system using shunting inhibitory convolutional neural networks," in *Proc. Intel Joint Conf. Neural Networks*, 2006, pp. 5336-5341.