

# 複雜背景下基於統計學習方法之手勢識別

莊頌祥、黃登淵

E-mail: 9806488@mail.dyu.edu.tw

## 摘要

在近幾年當中，基於視覺影像與統計學習方法的人機介面(Human-Computer Interfaces；HCIs)，其相關研究越來越引人注目。由於手勢是一種最直接和最方便的輸入方式，使得它的應用範圍日益廣泛。常見的應用包括有取代傳統的鍵盤和滑鼠輸入、在虛擬實境中作為控制輸入器、或輔助聾啞人士進行交互溝通等。手勢的旋轉角度大小對於識別正確率具有決定性的影響。本研究為了解決大角度手勢旋轉之問題，自行開發「手勢偵測與旋轉校正」模組，以做為影像前處理單元，可將手勢校正回較小的旋轉角度內。除此之外，本研究採用Gabor Feature做為手勢特徵描述的方法，利用PCA做為特徵擷取的方式，最後再應用歐氏距離(Euclidean Distance)、餘弦距離(Cosine Distance)與支持向量機(Support Vector Machine；SVM)等3種不同的分類器來進行手勢識別。透過實驗得知，採用Gabor Feature做為手勢特徵描述的方法，其手勢識別率遠高於僅採用原始手勢影像之識別率；在所採用的3種分類器中，又以支持向量機(SVM)之手勢識別率最高。例如：在混合各種旋轉角度靜態手勢識別的試驗中，Gabor-coded Image + Euclidean Distance的識別率在特徵數量為35時可得到93%，Gabor-coded Image + Cosine Distance的識別率在特徵數量為113時可得到94%，但Gabor-coded Image + SVM的識別率在特徵數量為103時卻可得到96.06%之水準；經比較沒有經過Gabor轉換之原始影像，歐氏距離分類器的識別率在特徵數量25時僅為80%，餘弦距離分類器的識別率在特徵數量35時亦為80%，至於SVM分類器的識別率在特徵數量29時可得到86.36%；由以上結果得知採用Gabor Feature + SVM可以獲得較好的手勢識別結果。在動態手勢偵測方面，本研究所發展的手勢偵測與識別系統亦能有效地識別出各種旋轉角度下之正確手勢。本系統每幀影像處理時間約為250ms。另外，透過一系列動態視訊影像之檢測，當穿著長袖與短袖衣服時，其手勢識別率分別可達到94%與90.8%之水準。

關鍵詞：手勢識別、Gabor Feature、主分量分析法、支持向量機

## 目錄

授權書	iii
中文摘要	iii
英文摘要	iv
誌謝	vi
目錄	viii
圖目錄	ix
表目錄	xiv
第一章 緒論	1.1
1.1 研究背景	1.1.2
1.2 文獻回顧與探討	1.1.2
1.3 研究方法	1.1.3
1.4 研究結果	1.1.4
1.5 本文架構	1.1.5
第二章 即時手勢偵測與辨識系統架構	2.1
2.1 前言	2.1
2.2 影像輸入部分	2.2
2.3 影像偵測與校正部分	2.2.3
2.4 影像偵測與校正部分	2.2.4
2.5 影像辨識部分	2.2.5
2.6 相關軟硬體之規格	2.2.6
第三章 手勢偵測與旋轉校正之前處理	3.1
3.1 前言	3.1
3.2 自適應性膚色模型選擇	3.2
3.2.1 YCbCr色彩空間	3.2.1
3.2.2 Soriano膚色模型	3.2.2
3.2.3 高斯混合模型Gaussian mixture model(GMM)	3.2.3
3.2.4 光線補償—參考白方法(Reference white method)	3.2.4
3.2.5 光線補償—修正參考白方法(Modified reference white method)	3.2.5
3.2.6 光線補償—Gray World	3.2.6
3.2.7 影像二值化	3.2.7
3.2.8 連通區域—4連通法	3.2.8
3.3 Gabor Filter之原理	3.3
3.3.1 運用Gabor Feature之影像旋轉與校正方法	3.3.1

25 3.4 手勢與手臂之分割	25
27 第四章 手勢的特徵擷取與辨識方法 4.1 前言	27
29 4.2 Gabor Feature	29
30 4.3 主分量分析(PCA)理論基礎	30
31 4.3.1 傳統主分量分析方法(PCA)	31
33 4.4 手勢辨識分類器	33
34 4.4.1 歐式距離分類器(Euclidean Distance Classifier)	34
35 4.4.2 餘弦距離分類器(Cosine Distance Classifier)	35
4.4.3 支持向量機分類器(Support Vector Machine Classifier ; SVM Classifier)	36
第五章 支持向量機(SVM) 5.1 前言	37
37 5.2 線性可分離	37
37 5.3 線性不可分離	37
41 5.4 非線性可分離	41
43 5.5 支持向量機之核函數選擇與參數設定	43
46 第六章 手勢辨識系統流程與實驗結果 6.1 前言	46
49 6.2 手勢識別系統流程設計	49
51 6.2.1 靜態手勢識別系統	51
51 6.2.2 靜態手勢識別系統實驗結果	51
52 6.3 動態手勢識別系統	52
59 第七章 結論與未來研究方向 7.1 結論	59
63 7.2 未來研究方向	63
63 參考文獻	63
65 圖目錄 圖2.1 手勢識別系統架構與流程說明	65
6 圖2.2 自建手勢資料庫內的11種手勢	6
8 圖2.3 原始影像(左)與所擷取的手勢影像(右)	8
8 圖2.4 第一列為標準手勢，第二列為變化手勢	8
8 圖3.1 自適應性膚色模型之選擇方法	8
12 圖3.2 輸入之原始圖片	12
13 圖3.3 紅色框為自適應性膚色模型選擇的最佳結果	13
13 圖3.4 經正規化後r, g之分佈	13
15 圖3.5 灰階分佈直方圖	15
22 圖3.6 4-連通示意圖	22
22 圖3.7 4-連通方向搜尋編碼圖	22
22 圖3.8 表示參數 旋轉所代表的角度	22
24 圖3.9 Gabor Filter, 其中 $\sigma=0.785$ ; $\theta=\{0^\circ, 90^\circ, 45^\circ, 72^\circ, -72^\circ, -45^\circ, -36^\circ\}$ ; $k=5$	24
25 圖3.10 手勢影像與Gabor Filter進行迴旋積後所產生的Gabor Feature	25
26 圖3.11 二值化手勢區域分割(歪斜角度 $> 45^\circ$ )	26
27 圖3.12 經過8個旋轉角度、3個尺度大小的Gabor Filter所產生之Gabor影像	27
27 圖3.13 經過校正與分割後的手勢影像	27
27 圖3.14 穿著長袖之狀態	27
28 圖3.15 穿著短袖之狀態	28
28 圖3.16 經過膚色分析與二值化後之結果	28
28 圖3.17 經過手勢與手臂分割後之結果	28
29 圖4.1 手勢特徵擷取與辨識演算法之流程	29
29 圖4.2 使用自建手勢資料庫內第1種手勢的Gabor Feature	29
30 圖4.3 將一張 $n \times m$ 的影像拉成一維陣列	30
31 圖5.1 線性支持向量機	31
38 圖5.2 SVM於資料分類示意圖	38
39 圖5.3 非線性支持向量機	39
44 圖6.1 左方為靜態手勢辨識流程，右方為動態手勢辨識流程	44
49 圖6.2 靜態手勢識別系統實驗平臺	49
50 圖6.3 動態手勢識別系統實驗平臺	50
50 圖6.4 原始影像之辨識結果	50
52 圖6.5 採用Gabor Feature之辨識結果	52
53 圖6.6 原始影像之辨識結果(變化角度 $< 45^\circ$ )	53
54 圖6.7 採用Gabor Feature之辨識結果(變化角度 $< 45^\circ$ )	54
55 圖6.8 原始影像之辨識結果(變化角度 $> 45^\circ$ )	55
56 圖6.9 採用Gabor Feature之辨識結果(變化角度 $> 45^\circ$ )	56
59 圖6.10 穿著長袖衣服時，使用SVM分類器的11種手勢辨識情況	59
59 圖6.11 穿著短袖衣服時，使用SVM分類器的11種手勢辨識情況	59
60 圖6.12 第1 - 6種手勢在各種角度的辨識結果	60
61 圖6.13 第7 - 11種手勢在各種角度的辨識結果	61
62 表目錄 表6.1 有無採用Gabor Feature在不同分類器下的最高辨識率	62

53	表6.2 變化角度 < 45 ° 時，不同分類器下的最高辨識率	57	表6.4 第一部分實驗結果的confusion matrix
55	表6.3 變化角度 > 45 ° 時，不同分類器下的最高辨識率		

## 參考文獻

- [1]J. Triesch and C. von der Malsburg, " Robust classification of hand postures against complex backgrounds, " In: Proceedings of the IEEE Int. Conf. on Automatic Face and Gesture Recognition, pp. 170-175, Killington, Vermont, USA, Oct. 1996.
- [2]J. Triesch and C. Von Der Malsburg, " A system for person-independent hand posture recognition against complex backgrounds, " IEEE transactions on pattern analysis and machine intelligence, Vol. 23, No. 12, pp. 1449-1453, 2001.
- [3]Y. T. Chen, and K. T. Tseng, " Multiple-angle hand gesture recognition by fusing svm classifiers, " In: Proceedings of IEEE conference on Automation Science and Engineering, pp. 527-530, Scottsdale, AZ, USA, Sep. 2007.
- [4]Y. T. Chen, and K. T. Tseng, " Developing a multiple-angle hand gesture recognition system for human machine interactions, " In: Proceedings of the 33rd Annual Conference of the IEEE Industrial Electronics Society (IECON), pp. 489-492, Taipei, Taiwan, Nov. 5-8, 2007,.
- [5]Q. Chen, N. D. Georganas, and E. M. Petriu, " Real-time vision-based hand gesture recognition using haar-like features, " In: Proceedings of IEEE Instrumentation and Measurement Technology Conference Proceedings, pp. 1-6, Warsaw, Poland, May 1-3, 2007 [6]Q.Chen, N. D. Georganas, and E. M. Petriu, " Hand gesture recognition using haar-like features and a stochastic context-free grammar, " IEEE Transaction on Instrumentation and Measurement, Vol. 57, No. 8, Aug 2008.
- [7]F. S. Chen, C. M. Fu, and C. L. Huang, " Hand gesture recognition using a real-time tracking method and hidden markov models, " Image and Vision Computing, Vol. 21, No. 8, pp. 745-758, Aug. 2003.
- [8]M. A. Amin, H. Yan, " Sign language finger alphabet recognition from gabor-pca representation of hand gestures, " In: Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, Hong Kong, August 19-22, 2007 [9]P. Moreno, A. Bernardino, J. S. Victor, " Gabor parameter selection for local feature detection, " In: Proceedings of IbPRIA2005, Lecture Notes in Computer Science 3522, pp. 11-19, 2005.
- [10]H. Y. Chen, C. L. Huang, and C. M. Fu, " Hybrid-boost learning for multi-pose face detection and facial expression recognition, " Pattern Recognition, Vol. 41, pp.1173-1185, 2008.
- [11]A. M. Martnez, and A. C. Kak, " Pca versus lda, " IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 2, pp. 228-233, 2001.
- [12]J. Wang, K. N. Plataniotis, and A. N. Venetsanopoulos, " Selecting discriminate eigenfaces for face recognition, " Pattern Recognition Letters, Vol. 26, pp. 1470-1482, 2005.
- [13]P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, " Eigenfaces vs. fisherfaces: recognition using class specific linear projection, " IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, pp. 711-720, July 1997.
- [14]J. Wang, K. N. Plataniotis, and A. N. Venetsanopoulos, " Selecting discriminate eigenfaces for face recognition, " Pattern Recognition Letters, Vol. 26, pp. 1470-1482, 2005.
- [15]Z. Shaoyan and Q. Hong, " Face recognition with support vector machine, " In: Proceedings of IEEE International Conference on Robotics, Intelligent Systems and Signal Processing, Vol. 2, pp. 726-730, Changsha, China, 2003.
- [16]G. Guodong, S. Z. Li, and C. Kapluk, " Face recognition by support vector machines, " In: Proceedings of IEEE Conference on International Automatic Face and Gesture Recognition, pp. 196-201, Grenoble, France, 2000.
- [17]K. Jonsson, J. Kittler, Y. P. Li, and J. Matas, " Support vector machines for face authentication, " In: Proceedings of British Machine Vision Conference Nottingham, pp. 543-553, Nottingham, 1999.
- [18]J. Qin, and Z. S. He, " A svm face recognition method based on gabor-feature key points, " In: Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, pp. 5144-5149, Guangzhou, China, 2005.
- [19]L. Bing, Z. Yun, and P. Y. Hong, " Face recognition based on wavelet transform and svm, " In: IEEE International Conference on Information Acquisition, pp. 373-377, Hong Kong and Macau, China, 2005.
- [20]M. Safari, M. T. Harandi, and B. N. Araabi, " A Svm-based method for face recognition using a wavelet pca representation of faces, " In: International Conference on Image Processing, Vol.2, pp. 853-856, Singapore, 2004.
- [21]V. N. Vapnik, " Statistical Learning Theory, " John Wiley & Sons, Inc., New York, 1998.
- [22]R. L. Hsu, M. A. Mottaleb, and A. K. Jain, " Face detection in color image, " IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 5, pp. 696-706, 2002.
- [23]C. Garcia, and G. Tziritas, " Face detection using quantized skin color regions merging and wavelet packet analysis, " IEEE Transactions on Multimedia, Vol. 1, No. 3, pp. 264-277, 1999.
- [24]D. Chai, and A. Bouzerdoum, " A bayesian approach to skin color classification in ycbcr color space, " In: Proceedings of TENCON, Vol. 2, pp. 421-424, 2000.

- [25]M. Soriano, B. Martinkauppi, S. Huovinen, and M. Laaksonen, " Skin detection in video under changing illumination conditions, " In: Proceedings of the 15th International Conference on Pattern Recognition, Vol. 1, pp. 839-842, 2000.
- [26]L. M. Bergasa, M. Mazo, A. Gardel, M. A. Sotelo, and L. Boquete,"Unsupervised and adaptive gaussian skin-color model, " Image and Vision Computing, Vol. 18, pp. 987-1003, 2000.
- [27]J. Yang and A. Waibel, " A real-time face tracker, " In: Proceedings of the IEEE Workshop on Applications of Computer Vision, pp. 142-147, Sarasota, Florida, USA, 1996.
- [28]J. Y. Xu, " Face detection and recognition technology research in complex background, " In: Proceedings of M.S. thesis, Shandong university of technology, pp. 22-24, China, 2007.
- [29]E. Y. Lam, " Combining gray world and retinex theory for automatic white balance in digital photography, " In: Proceedings of the Ninth International Symposium on Consumer Electronics, pp.134-139, Macau, 2005.
- [30]H. T. Lin, and C. J. Lin, " A study on sigmoid kernels for svm and the training of non-psd kernels by smo-type methods, " In: Proceedings of Master thesis, Department of Computer Science and Information Engineering, National Taiwan University, 2003.
- [31]A. Caplier, L. Bonnaud, S. Malasiotis, and M. Strintzis, "Comparison of 2d and 3d analysis for automated cued speech gesture recognition," In: Proceedings of the 9th International Workshop on Speech and Computer (SPECOM '04), Saint-Petersburg, Russia, September 2004.