

以手勢識別為基礎之視覺化人機互動技術

蔡耀弘

玄奘大學資訊管理研究所 助理教授
tyh@hcu.edu.tw

林思妤

玄奘大學資訊管理研究所 學生
susamlin2001@yahoo.com.tw

摘要

在近年來，智慧型機器人這個科技產業越來越受到關注，智慧型機器人不僅可應用於國防用途、工業用途，它也漸漸的走進了人類的日常生活中，可應用於休閒娛樂以及服務用途；其中主要的關鍵是利用人機介面的技術讓智慧型機器人與人類之間傳達訊息和動作指令以產生互動。

本論文則以發展即時手勢識別技術當作人機介面與智慧型機器人互動，主要提出一種方法能應用在不同環境下，並且達到即時識別效果的人機互動技術。我們系統目的是利用手臂的肢體動作作為輸入，並使用 Hough Transform 方法在手臂特徵的影像中找出線段的特徵，藉以辨識手臂的姿勢，再將該手臂動作轉換成訊息傳達給機器人，重要的是在執行時能夠達到即時的正確識別，讓人和機器人能產生互動。本系統在不同環境下經過多次測試後，在實驗結果顯示手勢辨識的成功率可達 75%。

關鍵詞：手勢識別、智慧型機器人、人機互動。

Abstract

In recent years, the technology of the intelligent robot is increasingly concerned. In this paper, we develop real-time gesture recognition technology, which is used as an interactive human-computer interaction for intelligent robots. We provide a method that can be applied in different circumstances. The capture image is from the webcam, then the system separates the arms and the background image from the input images by using image processing methods, like erosion, Hough Transform methods. Finally we show the identifying result of gesture of the application interface and can interact with the robot.

Keywords: Gesture recognition, intelligent robots, human-computer interaction, Hough Transform.

1.簡介

1.1 機器人技術

以目前來說，日本與美國在智慧型機器人這個產業上的技術最為領先，而日本也因此有「機器人王國」之稱[1]，但兩國卻是各有各的市場，日本是以智慧化服務型機器人為主，所以在控制方式及靈活度較大，而美國則是以人工智慧和研發技術為主。

現在機器人的技術可分為軟機器人技術、虛擬機器人技術、多智慧體控制技術、微型和微小機器人技術[2]及倣人和倣生技術[3,4]，而倣人和倣生技術是目前最高指標，讓機器人能完成像人一樣的動作。

1.2 人機互動介面

人機互動 (human-computer interaction, HCI)，就是所謂人類與機器之間的互動以及溝通方法；在人機介面上，以往都是以電腦為中心，需要透過其他媒介轉成機器的語言才能操控機器，例如需要利用滑鼠、鍵盤、手寫板或搖控器…等輸入介面來做為輔助工具才可達到互動，但現在卻是轉變為以人為中心，讓機器以我們人類習以為常的肢體動作、語言或情感來進行與機器的溝通及互動。

人機互動的模式有語音識別[5]、手寫識別、手勢識別、人臉識別、步態識別及虛擬環境…等技術來搭起人類與機器的溝通橋樑[7]，而人機介面裝置的範圍非常的廣，有運用在螢幕上使操作者方便使用的多點觸控技術，也有在互動上不受距離及環境限制的肢體上的語言[6]，也就是以人最直覺的肢體動作反應達到互動的直覺式控制介面，還有使本身具有人工智慧和自動化的機器人技術…等。

1.3 手勢辨識

人與人之間的互動方式有很多種，包括：語言、神情、手勢、眼神、文字，相對的人與機器人之間也可以利用這些方式作為互動的方式；但是對於人類來說，最自然的表達方式就是利用手勢的肢體動作來進行溝通，所以本

論文則選擇以手勢的部份作為與機器的溝通橋樑。

手勢辨識(Gesture Recognition)主要是透過電腦視覺的方法，利用手指或手臂的肢體動作來傳達訊息給機器，讓機器將攝影機所擷取的手指或手臂的動作指令做辨識，進一步的分析手勢的動作指令所代表的意義為何[8]；手勢辨識在應用上的範圍也非常的廣，目前已運用到軍事、娛樂、居家…等各領域上的控制，利用手勢識別去操控機器人也有相當多的應用，像是居家機器人，或是在娛樂方面有互動的機器；而目前在手勢識別技術上有雷射追蹤擷取手勢技術[9]、手勢輸入擬人化、虛擬鍵盤、指尖追蹤[10]、LED 手套手勢輸入[11]、手勢控制虛擬 3D 物件…等各種應用，有些技術是戴上資料手套去提取手指的動作，有些則是利用各種光的感應器作為輔助，除此之外還有一種最直接的方式就是不需要用手套及光的感應器的輔助就能自由的動作，直接經由人類本身的手勢肢體動作給機器一個動作，透過攝影機以視覺化影像作為輸入，讓它轉為指令，則可達到目的，而上述的手勢技術大部份都是使用輔助器具當媒介來達到控制的作用，但輔助的設備儀器既昂貴又會受到環境上的影響，所以本論文則是利用手臂的肢體動作作為輸入，透過 Webcam 以視覺化影像做後續影像處理來達到控制的作用，而成本既不高也不受到環境上的限制及影響。

本論文主要是以手臂的肢體動作來作為與機器人互動的依據；目前系統定義前進、後退、向左、向右及停止，這五個手勢肢體動作，而方法是首先利用背景相減法得到手勢的影像，再將此張影像做形態學的侵蝕運算，利用此運算的方法清除不必要的雜訊，再利用 Hough Transform 的方法偵測出手臂的所在位置，之後將偵測到的結果，再做手臂肢體動作的判斷。

本論文主要內容分為五章：第一章為簡介，內容為簡述手勢辨識及人機互動運用上的基本概念；第二章為針對相關的研究討論；第三章為本篇主要使用的方法；第四章為本篇方法所實驗的結果；第五章則為對本論文的研究心得及結論。

2.文獻探討

如何能夠即時性的讓機器人辨識出人類的肢體動作所代表的意思是非常重要的，並且又必須能夠應用在不同的環境下，這是有點難度

的，而目前大部份的手勢識別都是在控制的環境下去進行識別，Manar Maraqa 使用手套的輔助再做識別，並且在手套的五指及手腕上加上各種不同的顏色，以利用顏色去做手勢上的辨識，簡易區分出手的位置[12]；Chin-Chen Chang 則是使用者除了身穿黑色衣服外，連背景也都是用黑色，以此可降低周圍環境的影響，如此可輕易得到所要的手部特徵，在此篇方法則是利用曲率尺度空間(Curvature Scale Space)作為主要的手勢識別方法[13]；Kai Nickel 則是使用多假設追蹤(Multi-Hypothesis Tracking)尋找出頭部和手部的位位置，還結合類神經網路(ANN)及隱藏式馬可夫模型(HMM)做手勢的識別[14]。

3.建置方法

本系統主要是將手勢的肢體動作透過 Webcam 以視覺化影像做後續的影像處理，而建置方法如圖 1。

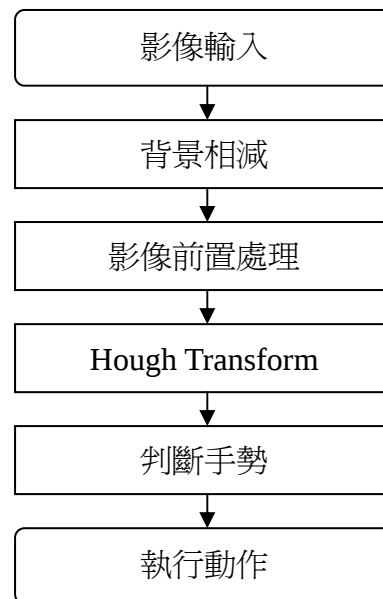


圖 1 建置方法流程圖

3.1 影像前置處理

在影像前置處理的部份分為背景相減法及影像前處理兩部份；首先先利用一般的 Webcam 取得連續影像前後畫面，經由背景相減法找出畫面移動的區塊部份，相減後就會根據大幅移動的區塊得到一張二值化只剩下手勢部份較明顯的影像，如圖 2；再將得到的影像做雜訊濾除，因為相減後的影像會因光線的變化而對於影像背景會產生一些雜訊，而為了

清除雜訊，則使用形態學中的侵蝕運算，此運算主要是利用在遮罩中所有像素點灰度值大小排序為最小的灰度值，執行侵蝕運算則可去除影像中的雜訊小點，如圖 3。



圖 2 背景相減結果



圖 3 形態學之侵蝕運算

3.2 特徵擷取

本論文則是使用 Hough Transform 的方法做特徵的提取；將在上述前置處理過後的影像中的特徵點利用 Hough Transform 的方法去尋找手勢肢體動作的直線，而 Hough Transform 主要是利用影像中點的位置找出直線的參數值，首先將影像空間映射到參數空間，再累計每一個點所產生的參數值，而最後參數點出現頻率最高的則為影像空間上的直線，從此處可得出手臂正確位置，進而獲得手勢的資訊。

3.3 手勢識別

判斷手勢識別的部份，本論文定義出五個手勢指令，(a)前進：雙手各向上四十五度；(b)退後：雙手各向下四十五度；(c)向左：左手向上四十五度，右手則向右平舉；(d)向右：右手向上四十五度，左手則向左平舉；(e)停止：雙手平舉，依序如圖 4(a)-4(e)。

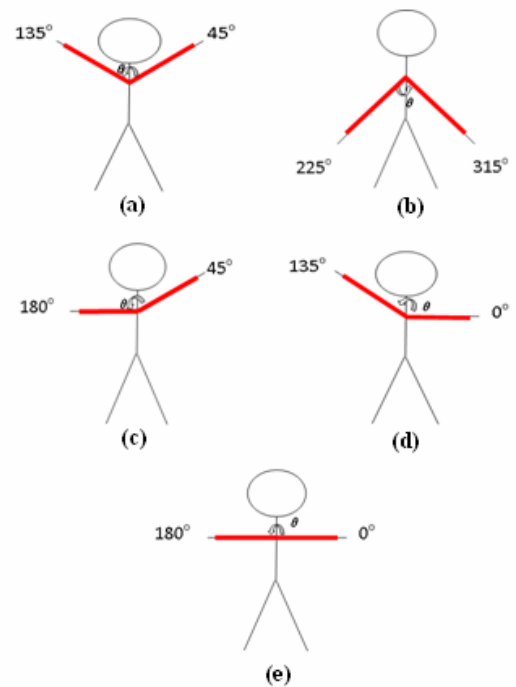


圖 4 手勢指令示意圖

以定義的五個手勢指令為依據，利用上述 Hough Transform 的方法所得出的影像中的二條直線做前進、後退、向左、向右及停止指令的判斷，從所得到的二條直線可得到不同角度的結果，例如在停止的指令則可得知角度為一百八十度時，機器人則執行停止的動作，以此類推。圖 5 則是以後退為例，經由 Hough Transform 方法所得出的結果，而灰色線段為所找出的直線線段。

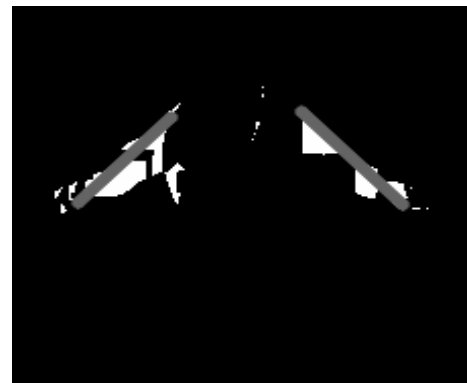


圖 5 利用 Hough Transform 找出線段

在上述方法已經得到二條直線，將直線利用向量內積的原理來求出兩向量之間的夾角，得知角度即可判斷此手勢的方向為何，而向量內積公式如公式 1，例如：以前進為例，

參考圖 6，首先以 0° 線為基準線，計算出 0° 與 $\bar{A}$ 的夾角為 θ_1 ，再計算 $\bar{A}$ 與 $\bar{B}$ 之間的夾角為 θ_2 ，之後得到這兩個角度後，再進行比對角度是否小於預設值 ε_1 、 ε_2 ，如公式 2。

$$\theta = \cos^{-1} \frac{\bar{A} \bullet \bar{B}}{|\bar{A}| |\bar{B}|} \quad (1)$$

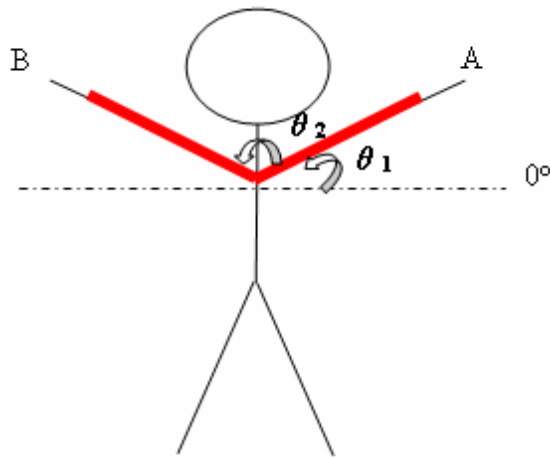


圖 6 計算角度示意圖

$$|\theta_1 - \theta_1^*| < \varepsilon_1, |\theta_2 - \theta_2^*| < \varepsilon_2 \quad (2)$$

4.實驗結果

本實驗測試有五位測試人員，個別測試所有五個手勢指令並在不同環境下分別各測試十次，而表 1 則為所測試後的辨識結果，從表中的數據上可看出向右與向左的指令辨識結果較低，大多是因為這兩個手勢的姿勢不是對稱的，所以測試人員在此兩個手勢上的角度會有較大的差異，導致在辨識率上會有些許的影響，而整體辨識率上為 75%。

表 1 手勢辨識率圖表

	前進	後退	向左	向右	停止	統計
1	$\frac{8}{10}$	$\frac{7}{10}$	$\frac{7}{10}$	$\frac{8}{10}$	$\frac{10}{10}$	$\frac{40}{50}$

2	$\frac{10}{10}$	$\frac{10}{10}$	$\frac{5}{10}$	$\frac{7}{10}$	$\frac{9}{10}$	$\frac{41}{50}$
3	$\frac{7}{10}$	$\frac{10}{10}$	$\frac{3}{10}$	$\frac{1}{10}$	$\frac{4}{10}$	$\frac{25}{50}$
4	$\frac{10}{10}$	$\frac{10}{10}$	$\frac{10}{10}$	$\frac{9}{10}$	$\frac{10}{10}$	$\frac{49}{50}$
5	$\frac{10}{10}$	$\frac{1}{10}$	$\frac{9}{10}$	$\frac{6}{10}$	$\frac{8}{10}$	$\frac{34}{50}$
統計	$\frac{45}{50}$	$\frac{38}{50}$	$\frac{34}{50}$	$\frac{31}{50}$	$\frac{41}{50}$	$\frac{189}{250}$
辨識率	75%					

5.結論

本論文提出以手勢識別為基礎的人機互動技術，主要是藉由以直覺性的手臂肢體動作作為與機器人的溝通橋樑，而本方法能應用在不同的環境下，並且在執行時能夠達到即時的辨識；辨識率則達到 75%，在未來還有改善的空間，以提高辨識率。

參考文獻

- [1] Capurro, R., Hausmanninger, T., Weber, K., Weil, F., "Ethics in Robotics," *International Review of Information Ethics*, Vol. 6, 2006.
- [2] Mahlknecht, S., Oberhammer, R. and Novak, G., "A Real-time Image Recognition System for Tiny Autonomous Mobile Robots," *Springer Science + Business Media*, 2005.
- [3] Hirai, K., Hirose, M., Haikawa, Y., Takenaka, T., "The Development of Honda Humanoid Robot," *International Conference on Robotics & Automation*, 1998.
- [4] Kaneko, K., Kanehiro, F., Kajita, S., Hirukawa, H., Kawasaki, T., Hirata, M., Akachi, K. and Isozumi, T., "Humanoid Robot HRP-2," *International Conference on Robotics & Automation*, 2004.
- [5] Lison, P. and Kruijff, G.J., "Salience-driven contextual priming of speech recognition for human-robot interaction," *In Proceedings of ECAI*, 2008.
- [6] Fong, T., Nourbakhsh, I., Dautenhahn, K., "A survey of socially interactive robots," *Robotics and Autonomous Systems*, PP.

- 143-166, 2003.
- [7] Myers, B. A., "A brief history of human computer interaction technology," **Human Computer Interaction Institute**, School of Computer Science, Carnegie-Mellon University, Pittsburgh, PA, Tech. Rep. CMU-CS-TR-96-163, 1996.
 - [8] Triesch, J. and Malsburg, C. V. D., "Robotic gesture recognition," **In Gesture Workshop**, PP. 233–244, 1997.
 - [9] Cheng, K., Pulo, K., "Direct Interaction with Large-Scale Display Systems using Infrared Laser Tracking Devices," **Proceed-ings of the Australian symposium on Information visualization**, volume 24, PP.67-74, 2003.
 - [10] Lee, T. and H"ollerer, T., "Handy AR: Markerless inspection of augmented reality objects using fingertip tracking," **In International Symposium on Wearable Computers**, October 2007.
 - [11] Park, J. and Yoon, Y. L., " LED-Glove Based Interactions in Multi-Modal Displays for Teleconferencing," **Artificial Reality and Telexistence-Workshops, 2006. ICAT'06. 16th International Conference on**, PP. 395–399, 2006.
 - [12] Maraqa, M., Abu-Zaiter, R., "Recognition of Arabic Sign Language (ArSL) Using Recurrent Neural Networks," **Applications of Digital Information and Web Technologies**, 2008.
 - [13] Chang, C. C., Liu, C. Y., Tai, W. K., "Feature alignment approach for hand posture recognition based on curvature scale space," **Neurocomputing for Vision Research**, PP. 1947-1953, 2008.
 - [14] Nickel, K. and Stiefelhagen, R., "Visual recognition of pointing gestures for human-robot interaction," **Image and Vision Computing**, 2006, in Press, available online since October 2006.