

# 建立以動態影像追蹤人臉之二維視覺伺服控制系統

## Build a Two-Dimensional Visual Servo Control System to Track the Human Face by Dynamic Image

參與學生：林高遠

### 摘要

本研究結合微電腦單晶片、串列通訊、電腦視覺函式庫等技術，製作一套二維旋轉的影像追蹤平台應用於動態影像追蹤人臉。不同於國內過去其他文獻採用昂貴的微步進馬達或伺服馬達，本研究採用無內建位置編碼器的減速直流馬達，僅以 Webcam 回傳的影像作為唯一回授訊號，驅動垂直與水平兩顆直流馬達進行人臉追蹤控制，達到二自由度運動追蹤人臉的目的。經由實驗結果顯示，在最差的研究限制下，即當人臉距離畫面中心最遠的四個角落時，本研究所建立的系統平均都可在一秒內完成追蹤定位任務。

**關鍵詞：**影像追蹤、視覺伺服、人臉追蹤、運動偵測、俯仰攝影機

### Abstract

This study combines MCU serial communication, and the Computer Vision library to produce a two-dimensional rotating image tracking platform used in dynamic image to track human face. Unlike other literature expensive micro-stepping motor or servo motor. In the research, deceleration DC motor without position encode, this research only the images Webcam returned as the only feedback signal, driving vertical and horizontal two DC motor for face tracking control, to achieve the purpose of human face Webcam two DOF motion tracking. The results of the experiment show that the worst limitations of the study, that is, when the four corners farthest from the center of the picture

of the face, the system established in this study the average can be completed in the second track positioning.

**Keywords:** Visual Tracking, Visual Servo, Face tracking, Motion Detection, PT camera

### 一、前言

近年來崛起的體感互動遊戲，突破了以往電視遊樂器，只使用手指輸入單調控制訊號的模式，改以人體全身關節的位置與相對角度作為控制訊號，不但能讓玩家有更逼真的遊戲體驗，更希望玩家在打電動的同時，能活動多數較大的骨骼肌，同時增進娛樂與身體健康。人體擁有十數個較大的可動關節，每個肢段間又互相形成複雜的連桿，能夠一次擷取如此豐富訊息的感測器(Sensor)就非攝影機莫屬。諷刺的是，正因為攝影機每次所取得的訊息都太過豐富，從中過濾出有用的資訊相當不易，使得「電腦視覺」或「機器視覺」成為工程技術中的新領域。

另外，目前普遍使用的影像監控系統，多半依賴多組攝影機同時擷取畫面，彌補單一鏡頭視角不夠廣的缺陷，這種監控系統有諸多缺點，如：目標在各鏡頭視野間遊走，不但難以避免死角，且事後追蹤困難，因為觀察的人必須將注意力同時放在多個畫面。為了克服以上缺點，後來又發明全視角攝影鏡頭，但是畫面嚴重變形，不論做電腦視覺自動追蹤或者人工追蹤目標物，都相當困難。

因此，自電腦視覺技術成熟以來，不斷有文獻嘗試將攝影鏡頭安裝在馬達上，

達到如同眼球般，可上下左右旋轉，將畫面中心對準目標物，只需要一個畫面，就可以對數倍於鏡頭視角的廣大範圍監視，並且只保留目標物周圍的重要資訊，不必像現行多通道監控系統一樣，占用  $n$  倍( $n$  為大於一的正整數)頻寬和  $n$  倍的容量貯存影像，卻只有其中幾個畫面帶有重要訊息，相當不經濟。另外，在體感遊戲應用方面，攝影鏡頭的運動過程，亦可作為電腦給人類玩家的回饋與刺激。

過去先有陳佩穎[6]採用雙軸步進馬達搭配 Maxtor Image Library 做頭部追蹤，後有徐煒智[7]以交流伺服馬達達成精密定位影像追蹤。但不論步進馬達或伺服馬達，皆有線路繁多，驅動器龐大且耗電的問題。且步進馬達更有轉動不平滑的現象，即便採用只有一半步進角的一二相激磁都難以消除，必須以微步進驅動器達成細膩的定位控制，但上述的架構又比本研究建立的系統複雜許多，體積更大。

## 二、研究法方與實驗步驟

本研究使用 Visual Studio 2010 C++ 與 OpenCV 2.4.2 撰寫擷取影像與偵測頭部座標的程式，並透過視窗作業系統應用程式介面(Windows API) 將誤差訊號傳送到 USB 所模擬出來的 COM Port，再透過 PL-2303HXA 晶片作為 USB 轉電晶體電晶體邏輯(TTL)通訊界面，連接到 89S51 的通用非同步收發傳輸埠(UART Port)，經 89S51 透過自行撰寫的程式計算，將誤差訊號轉化為 2 組脈衝寬度調變(PWM)送給 TA7279P，對 2 顆減速直流馬達比例控制。這些介面電路與控制電路。

實驗初期，先暫時將所有電子電路插在麵包板上，經過反覆測試、修改，確認可行之後，再全部自行佈線、焊接於一塊空的電路板上(如圖 2)。

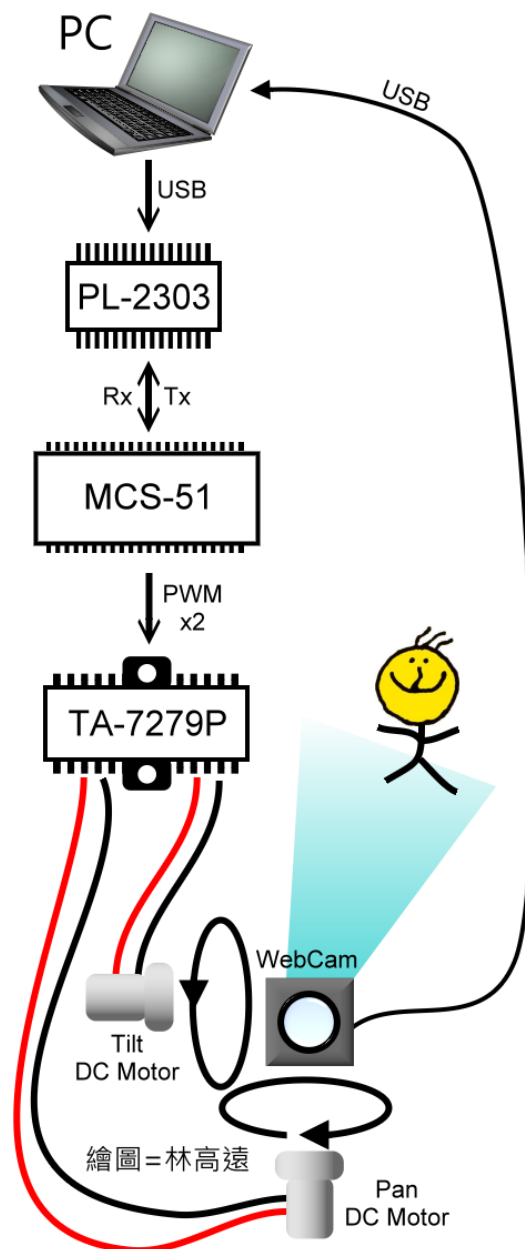


圖 1 整體系統示意圖

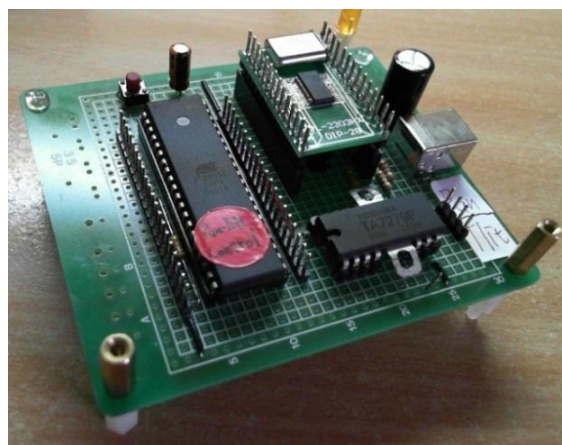


圖 2 焊接完成的介面與控制電路板

另外，機構部分，以兩顆額定電壓 12V、轉速 60rpm 的減速直流馬達，分別以垂直和水平交疊，另外在水平馬達的轉子上，固定一顆 Logitech Webcam C120 作為擷取影像用（如圖 3）。為了減少水平軸馬達負荷的轉動慣量，將較重的垂直軸馬達，置於水平旋轉平面的中心，較輕的 Webcam 置於外側。



圖 3 攝影機與二自由度旋轉馬達平台

## 2.1 影像處理

本研究影像處理程序如下(詳見圖 5)

1. 擷取彩色影像
2. Down Sample 1/2
3. 水平旋轉
4. 灰階化
5. 臉部偵測
6. 計算臉部中心座標

臉部偵測採用 OpenCV 的分類器函式，載入內建的人臉特徵資料庫，對每一個捕捉到的幀(Frame)偵測，取得人臉中心在 Frame 中的座標。為使偵測速度加快，將原始視訊圖形陣列(VGA)畫面的長寬各降取樣(Down-sample)1/2，畫素量為原始畫面的 1/4。

## 2.2 誤差計算

取得臉部中心座標後，計算與畫面中心的 X 軸與 Y 軸像素差距，換算為 PWM 值，送到 COM Port，換算公式如下：

$$H\_PWM = (100/2.5) \frac{H\_ErrCur}{H\_ErrMax} \quad (1)$$

H\_PWM: 水平 PWM 值

H\_ErrCur: 目前水平像素誤差

H\_ErrMax: 最大水平像素誤差

$$V\_PWM = (100/3.5) \frac{V\_ErrCur}{V\_ErrMax} \quad (2)$$

V\_PWM: 垂直 PWM 值

V\_ErrCur: 目前垂直像素誤差

V\_ErrMax: 最大垂直像素誤差

以上兩個 PWM 值皆為 0~±99，100 為將誤差比例標準化至 0~100，1/2.5 和 1/3.5 為 Kp。

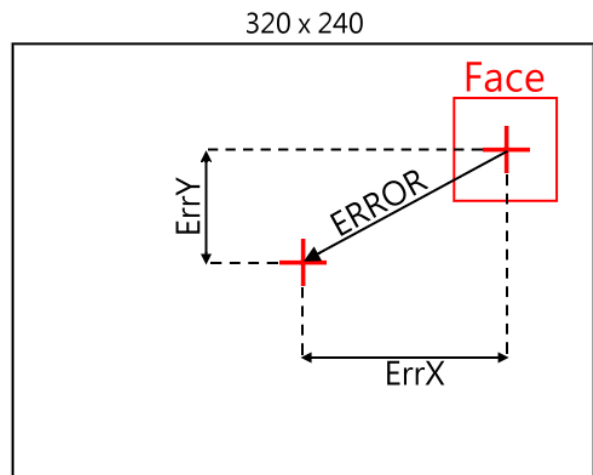


圖 4 像素誤差計算示意圖

## 2.2 微電腦通訊介面

上述的 PWM 值計算完成後，將水平與垂直 PWM 值透過 Visual Studio 2010 C++ 內建的 Windows API，傳送到實體上是 USB 的模擬 COM Port，經過 PL-2303HXA 這顆晶片將 USB 轉為 TTL 電位的 UART 訊號，傳送給 89S51 的 Rx 與 Tx 接腳[1]。

USB 模擬 COM Port 的優點，是直接使用在工業界歷史悠久且廣泛使用的 UART 通訊協定，不必重新學習複雜的 USB 通訊協定；缺點是，每一次使用都要先進

Windows 作業系統中的「裝置管理員」得知 COM Port 號碼，無法自動偵測裝置位址，在操作上稍微麻煩，但是不會對效能造成任何影響。



圖 5 影像處理流程示意圖

### 2.3 雙軸減速直流馬達控制

TOSHIBA TA7279P 是一顆內建兩組 H 橋的直流馬達正反轉控制 IC，每組 H 橋的控制邏輯如下

表 1 TOSHIBA TA7279P 真值表

IN1	IN2	OUT1	OUT2	直流馬達動作
1	1	短路		再生制動
0	1	L	H	正轉/逆轉
1	0	H	L	逆轉/正轉
0	0	高阻抗		摩擦減速

本專題研究輸入給 TA7279P 的 PWM 訊號，在工作週期(Duty Cycle) 時為「正轉/逆轉」(上或左/下或右)，Duty Cycle 以外的時間為「摩擦減速」，誤差=0 時為「再生制動」。

而再生制動，則是利用轉子剩餘的動能發電，產生電樞電流，此電流的電磁轉矩，又剛好與剩餘動能反方向，形成負回授，達成煞車效果，不須加入更多能量。

另外，PWM 並非 TA7279P 原廠所建議的操作方式，但是在本研究的反覆實驗下，證實將 PWM 輸入 TA7279P 可以達到良好的變速效果。

### 2.4 解決雙通道平行 PWM 困境

兩顆直流馬達的變速控制需要兩組 PWM 訊號，且本研究的馬達需追蹤隨機運動的目標，必須上下及左右地來回轉動，因此每個控制訊號還需要多給兩個方向資訊，才能滿足水平垂直方向同時運動的即時控制需求。

在控制 PWM 訊號的方式中，微控制器(MCU)以 Atmel 公司出品的 89S5X 系列最為普及，由於燒錄設備便宜易取得，教材與範例豐富，因此成為專題研究最常用的 MCU。

通常以 MCU 達成 PWM 有三種方式

#### 1. 迴圈[3]

在無限迴圈中，每一次都須進行鋸齒波和輸出值的比較，並將比較結果輸出 pin 腳。因為占用 CPU 週期達 100%，一但為了修改 PWM 值，而被串列通訊函式中斷時，Duty cycle 一定會暫時失準，若在變化慢的裝置上尚可行，但完全無法滿足本專題研究以比例控制達到即時追蹤的需求。

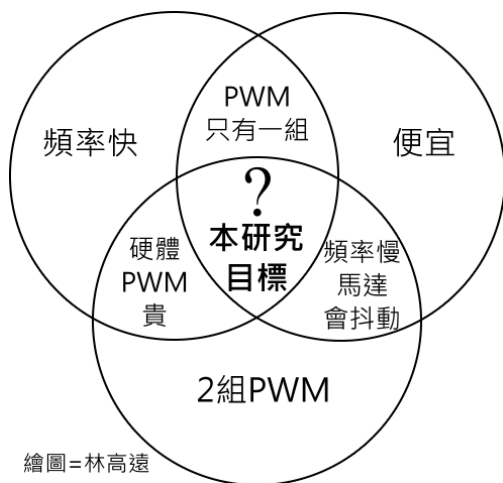
#### 2. Timer 中斷

以 MCU 內建的計數器，定時使其他函式被中斷，去執行 Timer 中斷函式。此作法可保留 Duty cycle 以外的時間給 main 函式，但是當 Duty cycle 一大時(例如：馬達高速運轉)，剩下的 CPU 週期還是不多，若又有串列通訊函式發出中斷時，Duty cycle 還是會暫時失準，且只能夠有一組 PWM 輸出。雖

然已經比上一個方法更高明，但依然無法滿足本研究需同時且快速地控制 2 組 PWM 訊號輸出至馬達的要求。

### 3. 硬體 PWM

市面上已經有許多增強型 8051 內建硬體 PWM 產生器，可以沿用 89S5X 的核心與程式寫法，且多了一到數個 PWM 暫存器。只要將欲輸出的 PWM 值存入 PWM 暫存器，便可產生對應的 PWM 時序(特定時間比例的 Hi 與 Lo 輸出)。此方法之優點是完全不需要顧慮中斷、計時(數)器、CPU 指令週期等時序問題，MCU 便會自動產生穩定且高頻的 PWM 訊號輸出。然而缺點有以下兩點：(1).此類 MCU 價格較昂貴，且零售通路少有販售，取得成本高昂，(2).無法相容於最普及的 89S5X 系列燒錄器，需另購萬用燒錄器或該廠牌 MCU 專用的燒錄器。



繪圖=林高遠

圖 6 PWM 規格衝突示意圖

針對以上所述之 PWM 的困境，除了經費昂貴以外，最主要的原因就是 CPU 的時間資源匱乏。另外，本研究發現在 89S51 中容量 4kB 的程式記憶體使用率極低，所以我嘗試以空間換取時間的做法，先以 Timer 中斷法 PWM 為出發點，採用類似平行處理的方式，將兩個輸出 PWM 的 pin 一起計算

完畢後，再對 pin 腳的電位改變。

經過實驗證實，此做法僅需要一個 Timer 且每 100  $\mu$ S 中斷一次、每個周期可中斷 32 次，可以達到約 310Hz 的兩組 PWM 輸出，且不受 UART Timer 的干擾。在成本、頻率、組數三重限制下，本專題研究以更好的演算法來善用程式記憶體資源，成功實現又快又便宜的兩組平行 PWM 輸出，滿足本裝置水平垂直方向同時運動的即時控制需求，如式(3)。

$$f_{PWM} = \frac{1}{100 \mu S * 32} = 312.5 \text{ Hz} \quad (3)$$

## 三、實驗結果

本實驗完成實體機器後，在光線適中，使人臉輪廓清晰，且只有一人出現在畫面中的複雜背景環境下，對監視畫面的四個角落進行追蹤測試。計算追蹤時間  $t_{Track}$  的公式如式(4):

$$t_{Track} = \frac{\text{Frame}[\text{Stop}] - \text{Frame}[\text{Start}]}{\text{FPS}(\text{每秒處理的 Frame 數})} \quad (4)$$

Frame[Start]:偵測到的第一張

Frame[Stop]:馬達追蹤至停止的第一張

### 3.1 不反應區域

在一般的控制系統中，為了避免系統反應過於靈敏，均會設定一個不反應的公差範圍(Tolerance)。本研究觀察到捨去誤差與馬達死區(Dead zone)兩個現象，可在畫面中心自然提供不反應的公差，因此在程式碼中並未寫入這類判斷式。以下就兩個現象分別說明

#### 3.1.a 捨去誤差

計算出像素誤差後，還須經過標準化與乘上比例控制兩個步驟，才會等於真實輸出的 PWM 值，若計算出來的真實輸出 PWM 值之絕對值未滿 1，會直接捨去，使真實輸出的 PWM 值為 0。以本研究的式(1)為例，若水平像素誤差未滿 8 個像素，真實輸出的 PWM 值都會為 0，像素誤差須達

8 個像素以上，真實輸出的 PWM 值才會  $\geq 1$ 。此現象影響較小。



平均最大追蹤時間 6/7 sec

圖 6 四個角落追蹤測試影片截圖

### 3.1.b 馬達死區

由於轉子轉動，需克服最大靜摩擦力，在有電磁轉矩但轉子轉速為 0 的區間，稱為馬達死區。此現象影響較大。

## 四、結論

針對二自由度視覺伺服控制系統，本專題研究提出一種整合系統的方式，結合影像處理、電腦視覺函式庫、單晶片微電腦控制等技術，使雙軸直流馬達，可以迅速朝目標物的位置轉動，以達到隨機移動的目標物可以保持在畫面中心，不需要用到複雜的影像處理理論或額外的編碼器測量轉子角度，用影像回授取代複雜硬體線路，達到即時追蹤的效能。

## 五、參考文獻與書目

- [1]詹東功 (2012)。微電腦控制-USB TO COM Port 監控實務。臺北市：全華，頁 4-1~7-58。
- [2]范逸之、江文賢、陳立元 (2002)。C++ Builder 與 RS-232 串列通訊控制。臺北市：文魁資訊。
- [3]張義和、王敏男、許宏昌、余春長 (2007)。例說 89S51-C 語言。臺北縣：新文京。
- [4]劉漢平 (2006)。以單晶微電腦製作之多通道波寬調變信號產生器。中華技術學院學報，Vol34，頁 185~196。
- [5]林伸茂 (1993)。8051 單晶片徹底研究-基礎篇，臺北縣：旗標，頁 326-543。
- [6]陳佩穎，“以新型動態輪廓技術完成可追蹤任意形狀物體之強健影像伺服系統”，碩士論文，國立臺灣大學電機工程學系，2003 年。
- [7]徐煒智，“以頻域設計伺服馬達之共振抑制與精密控制實現”，碩士論文，國立交通大學電機與控制工程系，2008 年。
- [8]Robert Laganière(2011)。 *OpenCV 2 Computer Vision Application Programming Cookbook*.USA:Packt
- [9]Gary Bradski, Adrian Kaehler (2008)。 *Learning OpenCV : computer vision with the OpenCV library*.UK:O'Reilly Media
- [10]D. Murray, A. Basu, "Motion Tracking with an Active Camera," IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 449-459, May, 1994
- [11]Ziyan Wu, Richard Radke, "Keeping a Pan-Tilt-Zoom Camera Calibrated," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 99, no. PrePrints, p. 1, , 2012