

使用單攝影機之入侵者自動追蹤監控系統

People Auto Tracking Subsystem of Visual Surveillance System

by Single PTZ Camera

(一) 摘要

本計畫擬設計一套以電腦視覺為基礎的監控系統，將拍攝到的連續影像輸入至電腦後由監控系統判斷是否為入侵者，並做出即時的反應與適當的處理。本計畫使用 PTZ 攝影機進行辦公大樓或是研究室等室內建築內部的監控，當樓層遭到入侵時，監控系統發出訊息通知相關人員，讓相關人員快速的得知入侵者的位置及其行為，未來更可進一步對入侵者進行臉部辨識，在事件發生當下進行入侵者的身分比對；或是對入侵者的行為進行即時的分析，讓人員能夠立即對入侵者的行為做出適當的反應，讓監控環境的控管更有效率。

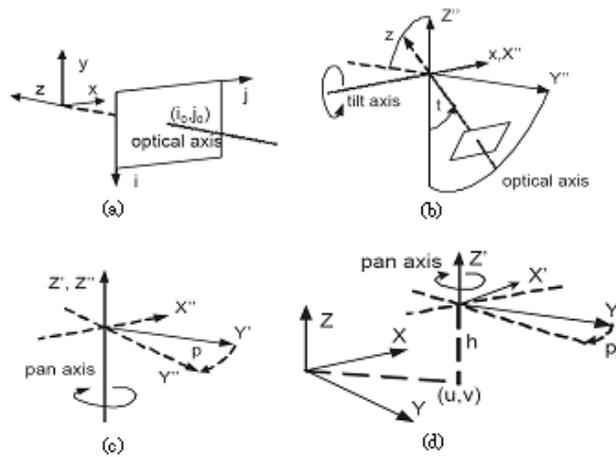
由於本計畫目前為初始階段，將由基礎的入侵者自動追蹤監控系統開始處理。在系統判斷為入侵者後，攝影機便會以入侵者的位置為中心點自動追蹤拍攝，直到入侵者離開攝影機的拍攝範圍為止，接著系統再令攝影機回到原位待命。自動追蹤監控入侵者的行動，可以有效的達到人力資源的運用，不需要依賴人員監控，更能夠避免因為人為疏忽而導致的監控上的死角[1]。

(二) 研究動機與研究問題

為了因應安全意識的逐漸高漲，監控系統也逐漸走向智慧化，不僅限於安全防護的功能，而是多了監控與管理等更廣泛的應用層面。自動追蹤監控、人臉辨識還有安全機制的回覆等，已經成為目前在安全監控的範疇裡相當重要的一環。而上述的自動追蹤監控更是能夠拿來進行其他的應用。

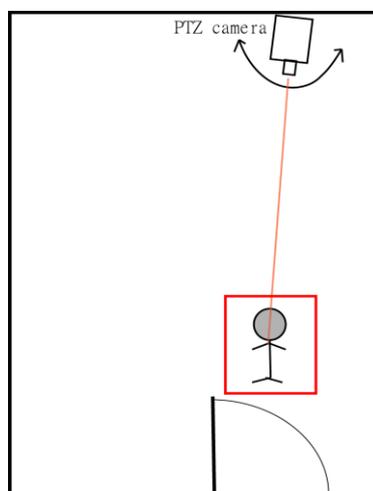
目前監控系統相當多樣化，許多室內影像監控系統多使用環場攝影機，其監控的範圍雖然比傳統固定式攝影機大上許多，但是隨著監控範圍的變大，在影像品質控制上的困難也隨之增加。其一困難來自於環場攝影機的鏡頭無法放大或縮小，對於監控畫面中的移動物體經常無法正確的攫取，導致影像可能過小或是模糊不清。

而本系統使用的攝影機為 PTZ(Pan-Tilt-Zoom)攝影機，所謂的 PTZ 是表示攝影機的鏡頭可以進行左右轉動(Pan)、上下傾斜(Tilt)與改變焦距、放大(Zoom)等不同的功能[2]。透過 PTZ 攝影機，不僅可以隨時改變拍攝的角度與所涵蓋的範圍，更能透過焦距的改變調整影像的大小與清晰程度，相較於環場攝影機可以獲得更好的監控效果與影像品質的維持。圖一中以座標的形式來解釋 PTZ 攝影機在改變角度方面的各種功能[3]，其中(a)為一般攝影機的平面座標和與影像平面的關係圖，(b)(c)分別利用座標表示 PTZ 攝影機進行左右轉動(Pan)、上下傾斜(Tilt)功能時的角度變化，(d)則是進一步顯示出 PTZ 攝影機在現實三維座標中的相對轉動情形。



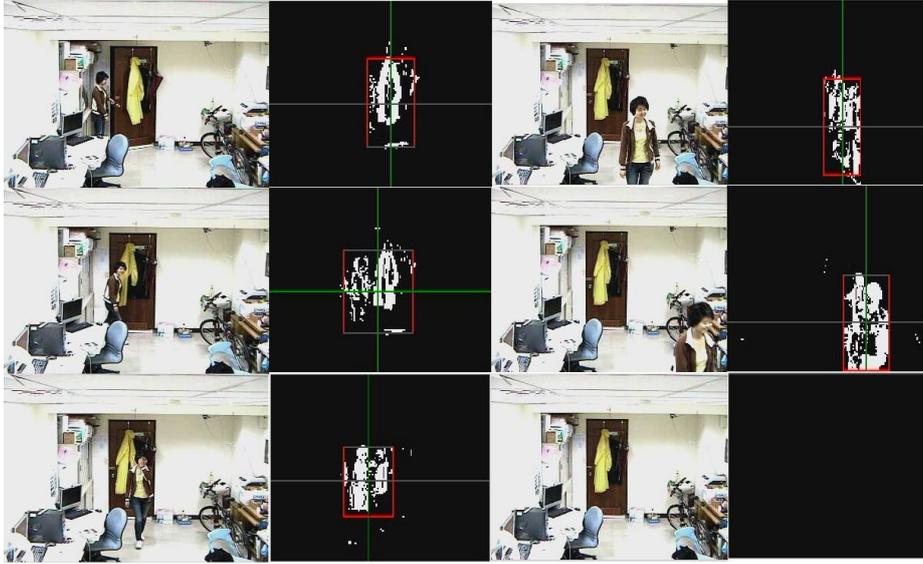
圖一：PTZ 攝影機之座標示意圖。

在本計畫進行階段先於研究室天花板架設 PTZ 攝影機，圖二為 PTZ 攝影機在室內架設之簡易平面圖。若有入侵者從出入口進入，系統即會鎖定入侵者，接著 PTZ 攝影機則會自動追蹤拍攝該入侵者，圖三為預期之鎖定結果。在圖三中，畫面右邊黑色部分為系統鎖定結果。入侵者進入研究室時系統便自動鎖定該入侵



圖二：PTZ 攝影機室內架設平面圖。

者，即使其位置移動仍會持續鎖定該入侵者。直到入侵者離開畫面，系統便不再顯示鎖定之結果。



圖三：PTZ 攝影機鎖定進入畫面入侵者預期之結果。

本系統在攝影機轉動角度時，可能會遇到影像處理和系統程式上的問題，有許多需要考慮到的困難。

- (1) 光線因素：雖然本系統使用於室內，但攝影鏡頭的每一次轉移都會導致光線的改變，會造成入侵者辨識和追蹤上的困難。
- (2) 背景改變：當鏡頭轉移時的背景也會隨之改變，會增加入侵者位置判斷上的難度與複雜度。
- (3) 入侵者之衣著：若是入侵者之衣著顏色與背景太過相似，在人物鎖定時可能會無法正確找出入侵者，會造成影像辨識上的困難，因而降低入侵者的辨識率。
- (4) 障礙物之影響：若是入侵者的影像受到障礙物的遮蔽，像是書櫃、樓層的梁柱等，可能會使攝影機無法正確鎖定入侵者之位置。
- (5) 即時性：為了要跟上入侵者移動之速度，對於影像之分析處理一定要快速。若是攝影鏡頭轉移不夠迅速，可能無法自動追蹤入侵者的行進，則無法達到監控之成效。

(三) 文獻回顧與探討

近年來國外內對於 PTZ 攝影機應用上的相關研究相當多，而在進行物體追蹤

(object tracking)之前，必須要先考慮 PTZ 攝影機與監控環境的空間轉換，以便於攝影機進行角度的改變。

首先必須將影像平面轉換至攝影機空間座標，再進行攝影機的校正以及水平和垂直角度的旋轉。而在這裡的討論為不需進行校正之情形。

PTZ 攝影機的 tilt 值指的是攝影機以 X 軸為旋轉軸做旋轉產生的偏差角。影像中的某一個點 $C_0 = (x, y, z)^T$ 在對應於攝影機做上下旋轉後可以得到一新座標 C_T 。 C_0 和 C_T 的關係如(3.1)式：

$$C_T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{bmatrix} C_0 \quad (3.1)$$

t_0 為對應至攝影機原始水平軸的 tilt 值， t_1 為攝影機原始水平軸對應至監視環境水平軸的 tilt 值， $t = t_0 + t_1$ 。

水平旋轉的處理與上下旋轉相類似。同樣先求得對應於攝影機原始 Y 軸的 pan 值 p_0 ，再求對應於監視環境 Y 軸的 p_1 值。接著算出水平旋轉後對應至監視環境的 pan 值 $p = p_0 + p_1$ 。將先前得到的座標 C_r 再經過旋轉矩陣如式(3.2)轉成新座標 C_p ：

$$C_p = \begin{bmatrix} \cos p & -\sin p & 0 \\ \sin p & \cos p & 0 \\ 0 & 0 & 1 \end{bmatrix} C_r \quad (3.2)$$

最後再將攝影機空間轉換至監視環境座標，定義 UVH 表示攝影機於監視環境中的座標位置， XYZ 代表監視環境的三維座標軸。利用 C_p 和 C_r 的轉換求出座標轉換公式如(3.3)， C_w 為監控環境中的點座標。在完成座標的轉換後即可在求出 t 和 p 值用來求得攝影鏡頭需要移動的角度。

$$C_w = C_p + \begin{bmatrix} U \\ V \\ H \end{bmatrix} = \lambda R \begin{bmatrix} \alpha(j - j_0) \\ -\beta(i - i_0) \\ -1 \end{bmatrix} + \begin{bmatrix} U \\ V \\ H \end{bmatrix} \quad (3.3)$$

$$R = \begin{bmatrix} \cos p & -\sin p \cos t & \sin p \sin t \\ \sin p & \cos p \cos t & -\cos p \sin t \\ 0 & \sin t & \cos t \end{bmatrix}$$

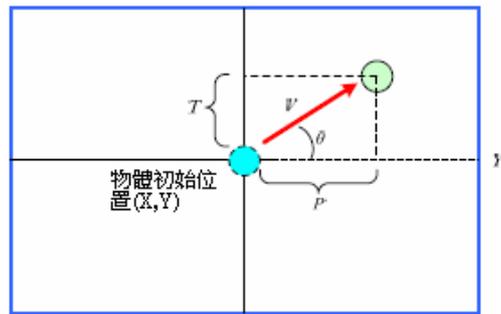
$$\lambda = \frac{H - Z}{\beta(i - i_0) \sin t + \cos t}$$

在 PTZ 攝影機可以運作後，接著進行自動追蹤拍攝部份之探討。為了使該物體能夠一直在拍攝範圍內，其位置的判定是相當重要的。因此 Yang[4] 等人利用一連串向量與角度的運算進行追蹤物體的位置判定。他們先在攝影機未移動前，先求出該物體的移動向量 \vec{V} ，再利用角度的改變，使用向量 \vec{V} 算出 P (pan) 和 T (tilt) 這兩個參數，以便於之後用來計算物體的實際位置，再配合攝影機的移動進行追蹤拍攝。圖四即為模擬之攝影機拍攝物體移動之畫面。

$$\begin{cases} \vec{P} = \vec{V} \times \cos \theta \\ \vec{T} = \vec{V} \times \sin \theta \end{cases} \quad (3.4)$$

因此可由(3.4)式求出：

$$\theta = \tan^{-1} \frac{v_y}{v_x}$$



圖四：模擬攝影機畫面中的物體移動。

Everts[3] 等人的研究中則是使用 Mean Shift 演算法來進行 PTZ 攝影機的追蹤。Mean Shift 演算法可以在大量資料樣本當中找到資料密度最高的地方。此方法並沒有限制使用在某些領域，它是個不需要設定參數的技術，可以使用在任可以大量數據資料表示的問題，運用在影像的處理上能夠幫忙自動擷取出需要的影像。

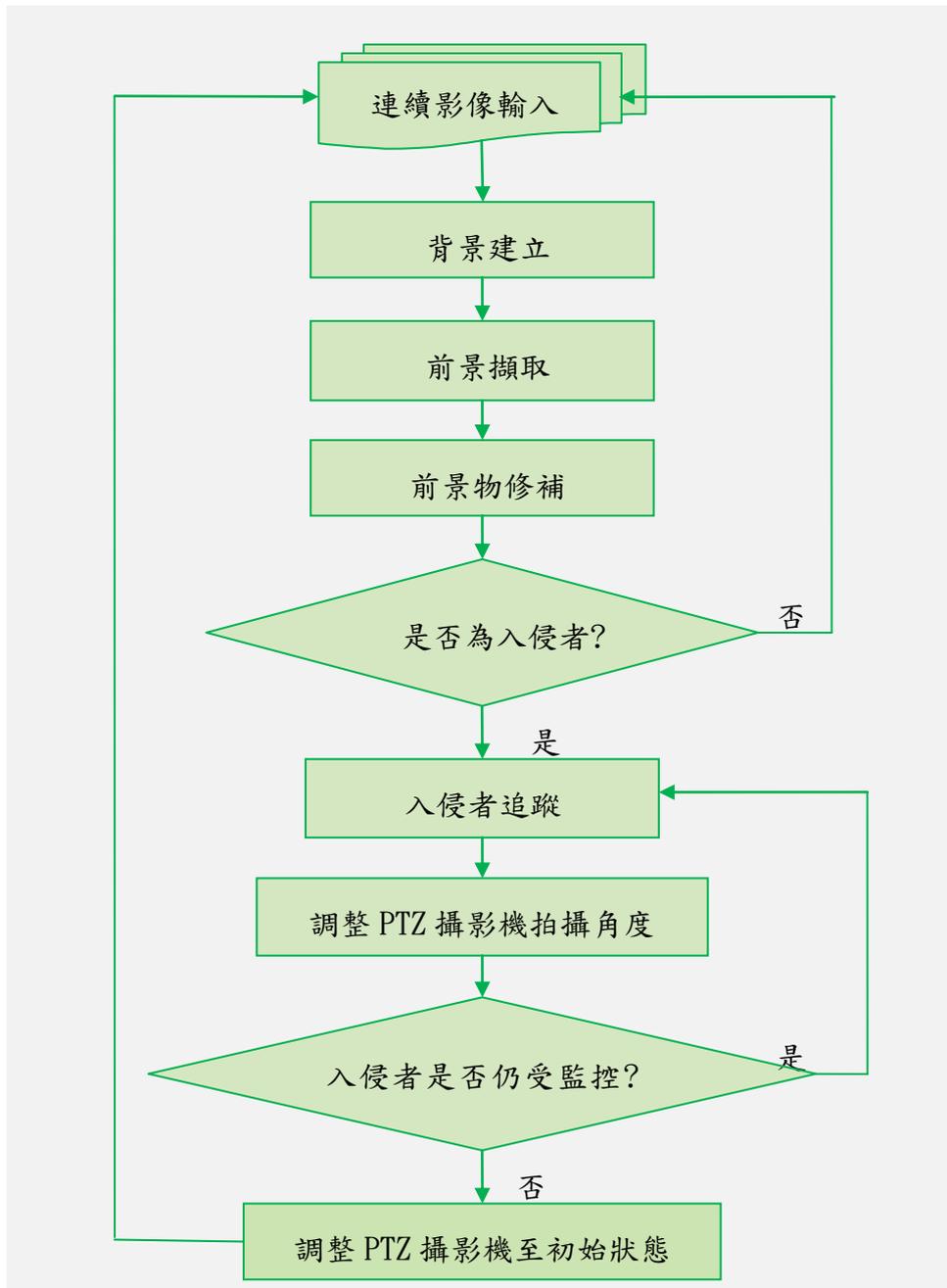
Everts 等人應用該演算法在物體追蹤方面，先假設追蹤的物體為顏色密度較高的區域，再利用該方法自動求出該物體的位置。Mean Shift 演算法的搜尋視窗大小固定，假設目前目標位置為 y_0 ，密度為 \hat{q} ，接著演算法在下一個影像中求出新的密度 \hat{p} 。利用 \hat{q} 和 \hat{p} 由(3.5) (3.6) 求出新的目標位置 y_1 。

$$y_1 = \frac{\sum_i^n w_i x_i}{\sum_i^n w_i} \quad (3.5)$$

$$w_i = \sum_{b \in B} \sqrt{\frac{\hat{q}_b}{\hat{p}_b(y_0)}} \delta(b(x_i) - b) \quad (3.6)$$

x_i 是 candidate region 內的第 n 個 pixel， B 為 bin indices 的集合， $b(x)$ 則是 x 的 bin index 函數。

(四) 研究方法與步驟



圖五：系統流程圖。

以下為流程圖的說明步驟：

1. 連續影像輸入

由架設在研究室內的 PTZ 攝影機拍攝連續彩色影像，本系統將會對其做進一步的分析與處理。

2. 背景建立

由於攝影機在尚未有入侵者時為靜止狀態，所以本計畫採用一般靜態環境常用的背景相減法來進行背景的建立。此方法會事先建立一個背景影像，接著再和連續輸入影像進行比較，在對應位置變化較大的區域即為環境變化區域，可藉此得知前景物的可能所在位置。

3. 前景擷取

本系統利用膚色偵測來進行前景物的擷取(如圖六)，由於在 $Y C_r C_b$ color space 中擷取膚色比較不容易受到光線的干擾，但是系統的輸入是 RGB 影像(R、G、B 分別表示紅色、綠色和藍色)。因此可使用下列數學公式將 RGB color space 成 $Y C_r C_b$ color space：

$$\begin{bmatrix} Y \\ C_r \\ C_b \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.5 & -0.4187 & -0.0813 \\ -0.1687 & -0.3313 & 0.500 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 1 \\ 128 \\ 128 \end{bmatrix}$$

其中 Y 是亮度值， C_r 和 C_b 都是表示色度值。



圖六：影像進行膚色擷取後經過二值化的處理結果。

4. 前景物修補

在進行膚色擷取時，有可能也會擷取出顏色與膚色較為接近的物體，必須將這些會影響物體判斷的雜訊排除，以利於後續之處理。圖七為進一步進行修補後的結果。



圖七：進行前景物修補後之影像。

5. 入侵者追蹤

由於畫面中的入侵者極有可能會改變位置或轉移角度，導致畫面上所見之入侵者的影像大小改變。所以本系統採用 CAMSHIFT(Continuously Adaptive Mean Shift)追蹤法[5]，此方法為 Mean Shift 演算法的改良，但此法能夠在進行影像處理時不斷調整該入侵者在畫面上的色彩分佈機率，改變儲存的影像大小，以達到更準確的位置鎖定與追蹤，相較於 Mean Shift 演算法，CAMSHIFT 更適合動態影像之追蹤。以下為使用此種追蹤法進行追蹤的五個步驟[6]：

步驟一：先令目前全部的影像為機率分佈的計算區塊。

步驟二：選擇出 Mean Shift 搜尋視窗的初始位置。

步驟三：計算該視窗中央位置的色彩機率分佈。

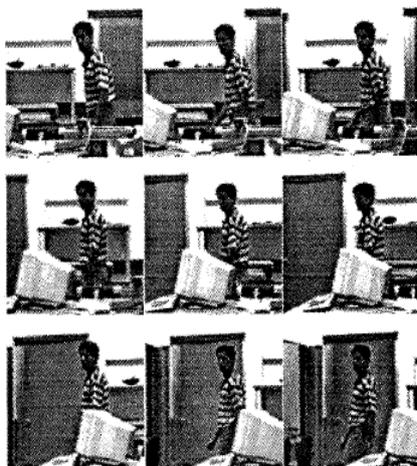
步驟四：可選擇直接傳送目前位置，或是進入步驟五重新設定。並儲存最初求出的物體顯示框的影像大小與位置。

步驟五：在下一個輸入影像中，使用步驟四存下的資料，初始化搜尋視窗的大小和位置，再回到步驟三。

在上述步驟當中，步驟一到三為一般 Mean Shift 演算的前置步驟。CAMSHIFT 演算法多了步驟四可以選擇是否重設顯示框的大小，可使目前儲存的顯示框相關資料更新為物體移動後的最新資訊，以增加動態追蹤的準確度。

6. 調整 PTZ 攝影機拍攝角度

使用先前提過的攝影機定位轉換公式，當入侵者位置移動時，系統計算出 PTZ 攝影機需要移動的角度 t (tilt axis) 和 p (pan axis)，改變攝影方向使入侵者依然在畫面的中間位置(如圖七所示)。



圖七：攝影機追蹤拍攝入侵者[7]。

7. 調整 PTZ 攝影機至初始狀態

在入侵者離開攝影機的拍攝範圍後，系統通知攝影機回到初始位置待命。直到有其他入侵者始重新進行自動追蹤。

(五) 預期結果

本計畫將會設計出一套以電腦視覺為基礎的自動追蹤監控系統。希望使用在單一出入口的室內空間，將 PTZ 攝影機架設在可拍攝到出入口處，當有入侵者進入時，系統便會鎖定該入侵者，並且自動追蹤拍攝。可以使環境監控完全自動化，達到人力資源的節省，也能使監控做的更完善。

(六) 參考文獻

- [1] 王春清，陳坤霖，“RFID 全路徑影像追蹤監控系統”，*Proceedings of the Conference on Ubiquitous Home*，2008。

- [2] 陳志銘，“利用環場及PTZ攝影機建構室內環境監控系統作臉部辨識”，國立中央大學資訊工程研究所碩士論文，2004。
- [3] I. Everts, N. Sebe, and G. A. Jones, “Cooperative Object Tracking with Multiple PTZ Cameras,” *Proceedings of the 14th International Conference on Image Analysis and Processing (ICIAP2007)*, 2007.
- [4] C. S. Yang, R. H. Chen, C. Y. Lee, and S. J. Lin, “PTZ Camera Based Position Tracking in IP-Surveillance System,” *Proceedings of the 3rd International Conference on Sensing Technology*, Tainan, Taiwan, pp.1 - 5, 2008.
- [5] S. Huttunen and J. Heikkila, “An Active Head Tracking System for Distance Education and Videoconferencing Applications,” *Proceedings of the IEEE International Conference on Video and Signal Based Surveillance (AVSS 2006)*, pp.30, 2006.
- [6] 蔡建戊，王翊航，鄭俊傑，張雅淇，陳佑杰，吳雋湧，“具人臉追蹤功能之網路監控系統”，銘傳大學電腦與通訊工程學系，2004。
- [7] J. Batista, P. Peixoto, and H. Aratijo, “Real-Time Visual Behaviors with a Binocular Active Vision System,” *Proceedings of the 1996 IEEE/SICE/RSJ International Conference on Multisensor Fusion and Integration for Intelligent Systems*, Vol.4, pp.3391 - 3396, 1997.

(七) 需要指導教授指導內容

在研究剛開始時和教授訂定工作進度時程表，需要和教授討論實作方法和實驗步驟。每個星期安排固定時間討論，過程中向教授請教論文的相關問題與研究，並在技術上做更深入的探討，最後再配合教授的指引撰寫出最終研究報告書。