

# 基於粒子濾波器技術的多個移動物體 之即時視覺偵測與追蹤

范欽雄

國立台灣科技大學資訊工程系  
csfahn@csie.ntust.edu.tw

宋炫慶

國立台灣科技大學資訊工程系  
loppol12@yahoo.com.tw

## 摘要

本論文提出一個基於粒子濾波器技術的多個移動物體之即時視覺偵測與追蹤系統；在偵測的階段，共分成背景建立、前景偵測、陰影去除和背景更新四部份，其中背景建立，係採用中位數的方法來產生，而前景偵測，則使用一個擷取函數以間接相減的方式分辨目標物體與背景，另在去除陰影的部份，是以決定性無模型為主的方法來去除物體的陰影，至於背景更新，則使用歷史圖來記錄背景像素變動的次數，進而達到更新背景的目的。又在系統的追蹤階段，本論文是採用粒子濾波器的方法來追蹤移動物體，其中係選取顏色分佈來當做目標物的特徵，而目標物的顏色記錄，則使用統計直方圖來表示，並利用背景資訊來調整候選目標物的權重，以獲得更好的物體追蹤效果。實驗結果顯示：在針對多個移動物體的一般情況下，我們的系統可以達到即時視覺偵測與追蹤的速度，並且擁有良好的強健性。

**關鍵詞：**粒子濾波、背景建立、陰影去除、前景更新、視覺偵測與追蹤。

## 一、前言

近十年來，隨著視訊產品的普及化和電腦視覺技術的快速進步，使得動態影像偵測與追蹤的方法被應用在各種領域，例如監視系統、智慧型交通系統、停車場管理系統等，它們可用來取代許多無聊且費時的工作，也可避免因為人類的疲倦所帶來的疏失，而在時效上，它們具有即時回報突發狀況的能力，所以能大幅降低整體系統的時間成本。

這幾年應用在各種領域的視覺偵測與追蹤技術，簡單介紹如下。Horprasert 等人提出利用彩色資訊差異與明亮度資訊差異的方式來偵測前景、背景影像與目標物體影子的演算法，在此方法中，由於門檻值的敏感度較高，因而易受雜訊的干擾，且當前景與背景影像的顏色相似時，很難將兩者分辨出來[1]。陳改良了上述的方法，雖然可以處理背景與前景物相似時所

引起分辨效果不佳的問題，但會犧牲消除影子的功能[2]。Haritaoglu 等人實做出可在戶外場景偵測與追蹤人物的  $W^4$  系統，其效果雖然不錯，可是影子所造成的影響仍然沒有解決[3]。此外，Talukder 和 Matthies 使用光流(optical flow)的技術來即時偵測行進中的車輛[4]。Cucchiara 等人所實做的 Sakbot 系統，已經被應用在車輛的休息站[5]。莊所提出的自動化交通監控系統的目的在于蒐集重要路口及路段的交通資料，諸如車流量、車距、平均車速與加速度、車行方向和軌跡，以及路面佔有率等；將這些數據傳回交通管理中心，可以協助交通號誌的控制[6]。

上述的系統皆運作在一個具有偵測及追蹤移動物體能力的平台上，縱使它們已有實際廣泛的應用，然而這樣的平台仍有許多問題存在，例如：光源變動的影響、低對比環境下的前景物與背景的分辨、影子偵測、如何快速追蹤到所期望的移動物體等等，都有很大的改良空間；另外，監視設備的精密度、品質與成本也是一大考量。基於以上各點，我們提出一個新的背景更新方法以及一個加入背景資訊讓追蹤移動物體更為準確的方法，進而發展出一套可利用電腦視覺技術即時偵測與追蹤多個移動物體的視訊監控系統。

## 二、系統介紹

本論文所提的多個移動物體的即時視覺偵測與追蹤系統，包括背景建立、前景偵測、陰影去除、背景更新與物體追蹤等功能。此系統是把第一次偵測到的物體顏色當做追蹤用的特徵，如果偵測效果不好，後續的追蹤自然就會發生錯誤；對於含有雜訊的影像，我們利用數理形態學的相關技術予以去除。以下我們評述相關論文所做的研究，並簡介我們實作的方法。圖1是本系統偵測與追蹤移動物體的流程圖。

### 1) 背景建立

本系統係以固定式的攝影機監視現場，理論上，於所攝取的影像裡，其背景是不變的。在背景建立時，若在戶外的環境下，則必須考慮天候的變化，因為一成不變的背景，會造成很大的錯誤。在此情況下，我們不期望背景為淨空，因為後續還有更新背景的程序來處理。

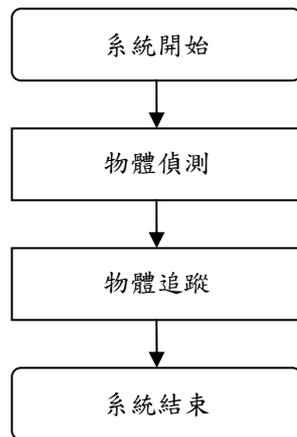


圖1 物體偵測與追蹤的流程圖。

根據Li等人的分類[7]，背景影像可以被描述成靜態(static)和動態(dynamic)兩種像素。靜態像素一般指不動的像素，例如：建築物，而動態像素指會動的像素，例如：電腦螢幕、搖擺的樹枝。動態像素在經過一段時間後，有可能轉成靜態像素，譬如：螢幕關機。又背景影像特徵的擷取，一般被分成光譜(spectral)、空間(spatial)和時間(temporal)三種類型；光譜特徵通常和灰階(gray-scale)有關，空間特徵大致和梯度(gradient)有關，而時間特徵則是和連續兩張圖片的改變有關。近年來，有些研究者係使用光譜的特徵來建造背景模型[3, 7, 8]。為了對抗照度(illumination)的改變，有些研究者採用空間特徵來增加強健性[3, 9, 10]；光譜和空間的特徵足夠用來描述靜態像素，但是對於動態像素必須要配合時間特徵來描述[3, 11]。

本系統所監視的畫面中，背景物基本上是不變的，在統計上，具有中央趨勢(central location)的測量數，可分為三種：平均數(mean)、中位數(median)和眾數(mode)，其中平均數容易受極值的影響，而使用眾數時，如果發生兩個同樣出現次數的數，則會有取捨上的困難。本論文係針對過去 $N$ 張連續影像，求算每一個像素的中位數來建立背景影像。

## 2) 前景偵測

當建立好背景影像之後，每一張輸入影像和背景影像相減，即可得出前景影像。因為在低對比的环境之下，分辨前景與背景的門檻值不易選取，所以我們改用一个擷取函數來取得前景偵測物[12]。

沒有一個偵測演算法是完美的，常見的缺點是：會產生欺騙像素(spurious pixel)和緣於移動物體所產生的空洞(hole)。有鑑於此，我們採用數理形態學(mathematic morphology)的方法，例如：膨脹(dilation)和侵蝕(erosion)運算子來移除欺騙像素和填補空洞；接著，利用相連元件(connected component)偵測演算法將修補後的像素連成數個區塊(block)。當形成的區塊大於畫面的百分之一大小時，就開始進行移動物體的追蹤程序。

## 3) 陰影去除

與背景影像相減得出的餘留影像，並非就是真正的前景物體，因為在戶外的環境下，跟隨物體的陰影，很容易被誤認為前景；於目前電腦視覺的領域裡，在不特定的場景下，還沒有一個可以有效去除物體陰影的理想演算法。陰影分成靜態和動態兩種，前者指的是環境中一些固定物體所產生的陰影，它們會隨著背景更新而被學習當成背景，所以，我們的焦點在於動態陰影的去除，而此種陰影一般指的是源自移動中的物體。

針對陰影去除的研究已有幾種方法被提出，譬如：利用彩色資訊差異與明亮度資訊差異的方式來偵測前景、背景影像與影子的演算法[1]，可是此方法所用的門檻值，其敏感度較高，易受雜訊的影響，而且當前景與背景顏色相近時，有分辨不佳等問題。另外，有利用HSV的色彩資訊來去除陰影的方法[5]，因為色調(hue)值並不會因為陰影而改變其值，可是此方法在每一種天候環境下，所設定的參數都不一樣，對於全天候的條件下，該方法必須先有一組配合每一種天候環境下的參數。還有，利用時空的反照率和雙色的反射模型，以及考慮太陽和天空的散射來偵測陰影的方法[13]，可是此方法必須先將訓練資料放進資料庫，所以儲存空間必須非常大，因而處理速度也會較慢。根據實作評估的結果，本論文決定採用上述第二種方法[5]。

## 4) 背景更新

因為天候或者有背景物體的加入，所以背景更新乃是必要的程序。有關背景更新的情況可分成基於像素(pixel-based)和基於物體(object-based)兩大類，其中前者係根據週期性的亮度變化來改變，而後者則是在有物體進入場景且逗留一段時間後，把物體當成背景學習起來。針對此問題，有研究者採用線性的方法來更新背景，亦即使用一個學習速率 $\alpha$ 來控制學習背景的速度[14]。又有研究者採用建構式的方法記錄影像中每一位置都有自己的特徵個數和每個特徵出現的次數[15]；當某特徵出現一次，則該特徵的次數就增加，反之，則減少，最後把出現次數最多的特徵當做該位置的背景值，然而此方法相當的耗時。

本論文針對上述的方法予以修改，使得整體的運算量大幅減少，亦即影像中每一個位置有自己的歷史圖(history map)，它係用來記錄某位置被當做前景的次數；當次數到達某一程度時，即拿過去 $N$ 張畫面同一位置的像素，取中位數來更新背景。

## 5) 物體追蹤

大體而言，於動態影像中，追蹤物體的做法可分成三種：比對追蹤、預測追蹤與能量函數[16]。比對追蹤必須要整張影像來做偵測，雖然準確率高，但是太花時間，速度沒有辦法有效地提升。能量函數係使用蛇行(snake)模式，它對於輪廓追蹤有很大的幫助，可是本系統針對的前景物體是以人為本，由於人是非剛體，走路時會產生形變，所以在進行物體追蹤的時候，使用輪廓當做特徵會有相當的難度。

本論文係採用預測追蹤的方法，並且以物體的顏色作為特徵的依據。一般預測追蹤的方法有：卡爾曼濾波器(Kalman filter)和粒子濾波器(particle filter)兩種。卡爾曼濾波器對物體運動方程式屬於線性的情況下，有良好的追蹤效果，但是它不適合對非線性、非高斯模式運動的物體做追蹤，因為有許多不確定的因素且物體不一定是單純的線性運動，而粒子濾波器對於處理非線性的問題，有較好的效果，所以，在物體追蹤的階段裡，我們採用粒子濾波器的技術。

### 三、物體偵測

為了使本系統具備視覺監控的功能，首先必須將移動物體順利地偵測並且擷取出來，以供後續物體的追蹤及識別之用，其中物體偵測的階段包括背景建立、前景偵測、背景更新與陰影去除等程序。

#### 3.1 背景建立

建立背景的基本原則是同樣的像素會在影像序列中同一個位置重複出現，也就是每一個位置的像素值會有中央趨勢的特性。如果背景是固定不動的，長期下來一定會向某像素值集中，而具有中央趨勢的測量數主要有三種，其優缺點如下所述：

##### 1)平均數

優點：使用所有數據。

缺點：易受極值影響。

##### 2)中位數

優點：不受極值影響。

缺點：無法進行統計推論。

##### 3)眾數

優點：不受極值影響。

缺點：對數值變化不靈敏。

基於上述的優缺點以及記憶體空間使用上的限制，我們採用中位數來建立背景影像，即利用前 $N$ 張畫面取出每個影像點的中位數當做背景值。因為後續還有更新背景的步驟，所以一開始建立背景時，可以使用較少的張數來取中位數。

#### 3.2 前景偵測

一般情況下，明亮度的改變是由背景影像和目前影像的每一個像素相減而得，可是用來分辨某像素是否為前景或背景的門檻值之設定相當困難，特別是明亮度在低對比的環境之下。為了解決這問題，我們採用一個擷取函數來獲得前景偵測物[12]，如式(1)所示。

$$g(a,b) = 1 - \frac{2\sqrt{(a+1)(b+1)}}{(a+1)+(b+1)} \cdot \frac{2\sqrt{(256-a)(256-b)}}{(256-a)+(256-b)} \quad (1)$$

其中 $a$ 是目前影像的亮度值，而 $b$ 是背景影像的亮度值。此公式是由算術平均和幾何平均的關係： $(a+b)/2 \geq \sqrt{ab}$ ，當 $a$ 與 $b$ 為不相等的正實數推導而來。

圖2顯示擷取函數 $g(a,b)$ 的兩個例子，其中一為 $b=5$ ，另一為 $b=100$ 。當 $a$ 等於 $b$ 時，則擷取函數 $g(a,b)$ 的值為最小，亦即其值為0。當 $a$ 與 $b$ 的差值增加，則擷取函數的值也會增加，特別是在 $b=5$ 的情況下， $g(a,b)$ 的值快速增加，此即表示：在較低程度的明亮度中，很小的變化也可以被偵測出來且忽略明亮度層次是在中間的程度；換言之，我們可以根據背景影像的明亮程度來安排 $g(a,b)$ 的敏感程度，本系統係將門檻值設為0.01。

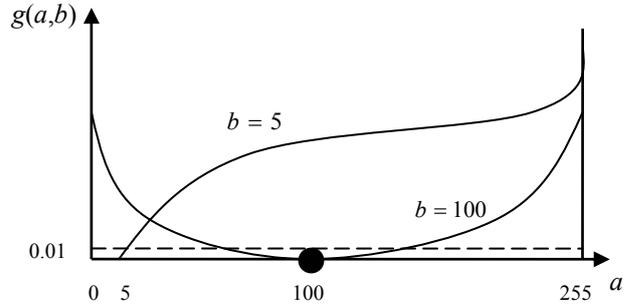


圖2 擷取函數 $g(a,b)$ 的圖例。

#### 3.3 背景更新

在背景更新的步驟裡，我們使用歷史圖來記錄影像中某位置被當做前景的次數；當累積次數達到一定程度，就根據前 $N$ 張畫面來取中位數，並且利用類似王與陳所提的方法來調整學習速率[15]。

為了因應各種不同的場景，必須要有一組設定值來改變背景的學習速率，一般而言，對於物體移動速度較慢的場景，其設定的學習速率較慢。因為上述原來的的方法，計算量太大無法達到我們對於速度的要求，而且記憶體空間的使用也相當龐大。為了解決此問題，我們修改他們的方法，使運算速度可以更快，圖3為新方法的背景學習速率之示意圖，又歷史圖的累加方法如式(2)所示：

$$h_t(x,y) = \begin{cases} h_t(x,y) + A & \text{若}(x,y)\text{為前景像素的位置} \\ h_t(x,y) - D & \text{若}(x,y)\text{為背景像素的位置} \end{cases} \quad (2)$$

其中 $A$ 與 $D$ 分別為累加與遞減值。

像素變動程度

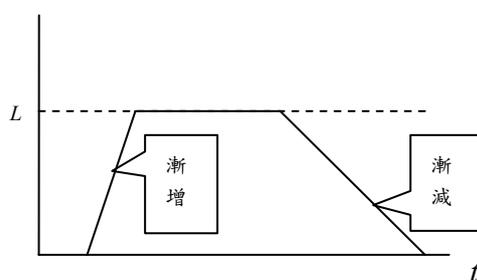


圖3 更新背景的學習速率曲線。

圖4舉一簡單例子說明歷史圖的使用方法；在此例中，圖4(a)與圖4(b)直接相減，如果等於0，代表並無前景物的出現，所以相對於歷史圖的位置所儲存的值必須遞減(-D)；如果不等於0，即代表有前景物的出現，則相對於歷史圖的位置所儲存的值必須被累加(+A)，本例子中假設  $A=3$ ,  $D=1$  且  $L=10$ ，又歷史圖的最大值設為  $L$ ，而最小值設為0。圖4(c)與圖4(d)分別為歷史圖的初始與目前的狀態。

5	5	6
6	7	7
8	8	9

(a)

5	6	6
7	8	9
8	8	10

(b)

0	0	0
0	0	0
0	0	0

(c)

0	3	0
3	3	3
0	0	3

(d)

圖4 歷史圖的簡例：(a)上一張影像；(b)目前的影像；(c)歷史圖的初始狀態；(d)歷史圖的目前狀態。

依照以上所提的方法，圖5為一個序列影像的前景偵測實例，所得的結果以二值化的影像顯示。

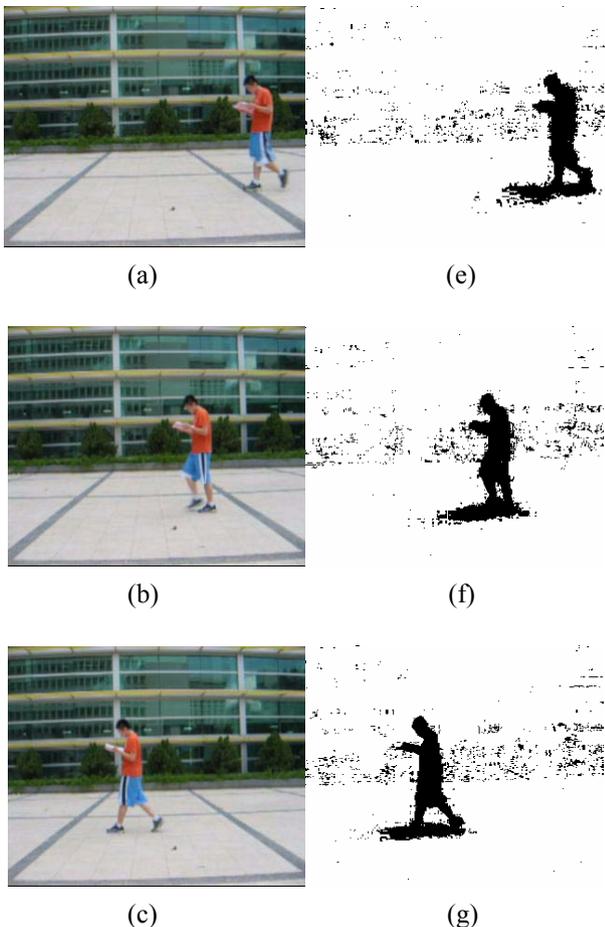


圖5 前景偵測的實例：(a)~(d)為第118、141、164、195個畫面的影像；(e)~(h)為前述畫面所偵測到的前景影像。

### 3.4 陰影去除

在戶外的環境下，陰影在背景相減的時候，很容易被誤認為前景，此時對於前景物的寬與高常會有很大的錯誤，如圖6(a)所示。針對此問題，我們參考Cucchiara等人提出去除陰影的方法[5]，以下為它的原理與本論文的作法。

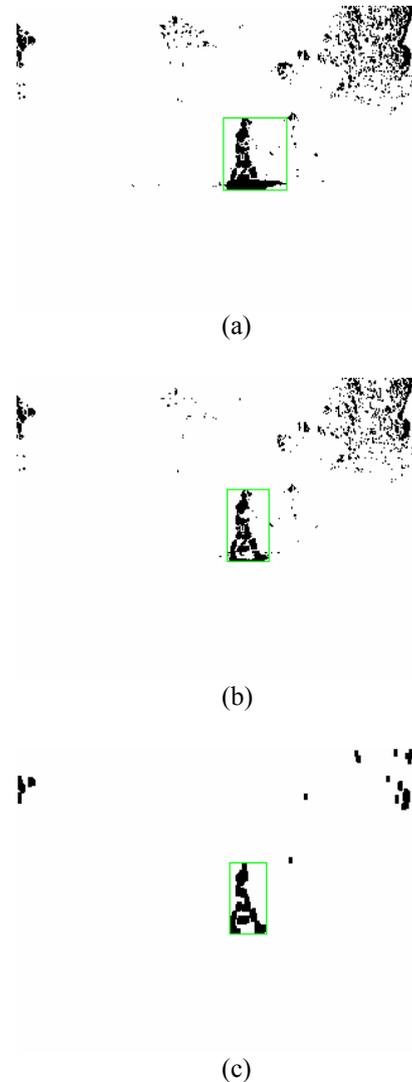


圖6 (a)尚未排除陰影的前景物體；(b)去除陰影後的結果；(c)再經過數理形態學運算後的前景物體。

首先，將影像從RGB的色彩空間轉換成HSV的色彩空間。使用HSV的主要理由在於它和人類的感知比較類似，且用來區別陰影較為正確。事實上，陰影與背景的像素相比對，色調值並沒有明顯的改變，也就是不容易受到光源的影響；又陰影通常有較低的飽和度(saturation)。綜合以上所述，可以根據式(3)決定某像素是否為陰影。

$$SP_k(x,y) = \begin{cases} 1 & \text{若 } \alpha \leq \frac{I_k^V(x,y)}{B_k^V(x,y)} \leq \beta, \\ & (I_k^S(x,y) - B_k^S(x,y)) \leq \tau_S \\ & \text{且 } |I_k^H(x,y) - B_k^H(x,y)| \leq \tau_H \\ 0 & \text{其它} \end{cases} \quad (3)$$

其中  $I_k(x,y)$  和  $B_k(x,y)$  分別代表在第  $k$  個畫面時，目前像素和背景像素在  $(x,y)$  位置上的值(上標  $V, S$  與  $H$  各自表示像素的灰階值、飽和度與色調值)； $\beta$  的使用是為了防止背景遭受細微的雜訊所干擾，而  $\alpha$  是為了考慮陰影的「能量」，亦即光源的強弱會改變  $\alpha$  的值，因此在戶外的場景下且有較亮的陽光時，應該選取較低的  $\alpha$ 。

比較好的方法是：先對各種不同的狀況計算其參數，而在使用的當時，則根據其狀況來選擇參數。本系統是採用直接給定參數的方式，藉由多次實驗的統計，我們選取  $\alpha=0.6, \beta=1.0, \tau_H=0.2$  且  $\tau_S=0.4$ 。圖6(b)為圖6(a)去除陰影後的前景物體影像；由此圖可以看出，伴隨前景物體的陰影已顯著地被去除。

每一個偵測出的前景物體仍殘留一些雜訊和小洞，我們再透過數理形態學當中的侵蝕和擴張兩種運算子來解決這個問題。圖6(c)為圖6(b)經過運算所獲得的結果，可以很明顯看出它們有效率地將雜訊做進一步的消除。圖7為侵蝕與擴張運算子所使用的結構元素(structuring element)。

1	1	1
1	1	1
1	1	1

圖7 侵蝕與擴張運算用的結構元素。

最後，使用相連元件演算法將每一個前景區域連接成一個大區域，並給予單一的標籤(label)，其中我們採用八連通的檢查方式，對整張影像依照由上往下、由左往右的順序檢查。

#### 四、物體追蹤

物體追蹤的主要目的在於：當系統偵測出有前景物的存在時，接下來要持續瞭解物體的行進動作或物

體的移動軌跡。對於物體追蹤的演算法，有兩個主要的考量，即

- 1) 如何適切地描述一個物體？
- 2) 如何有效率地發現物體？

以下闡述本系統對於上述兩個問題的作法。

##### 4.1 物體描述

追蹤移動物體的第一要務是如何適切地表示及描述經由物體偵測程序所獲得的前景影像。基本上，表示一個前景影像有兩種選擇：利用它的外部特性(如：邊界)，或者利用它的內部特性(如：紋理)；接著是根據選擇的表示方法來描述該影像，例如：用前景影像的邊界來表示，而邊界是用它的長度、端點連接的直線，以及邊界凹凸的性質去描述。當前景物體的主要特徵偏重形狀時，則選擇外部的特性表示；反之，則選擇內部的特性表示。在任何一種情況下，做為描述前景影像所選擇的性質應該不會隨它的大小、平移，以及旋轉等幾何轉換而改變。

常見的物體特徵有：顏色、輪廓和光流等。移動物體一般可分為剛體和非剛體，如果使用輪廓為特徵去追蹤諸如人這類非剛體時，由於他們的肢體動作很容易產生輪廓的形變，因而導致物體追蹤發生錯誤。相對地，以顏色當做特徵較為適合，因為它有以下幾種好處：比起輪廓追蹤的方法，使用顏色追蹤較為簡單，容易計算和比對；比起利用樣板匹配的方法，更具有強健性。有鑑於此，本系統選擇以顏色做為描述物體的特徵。

##### 4.2 物體搜尋

為了節省追蹤物體的計算量，有很多搜尋演算法(search algorithm)可以用來發現物體，譬如卡爾曼濾波器、光流的三步法與粒子濾波器等。由於一般化的物體運動模式往往是屬於非線性、非高斯的系統，為能強健地追蹤運動中的物體，我們採用可以針對此系統做最佳化估測的粒子濾波器，它是一種運用於非線性、無高斯約束模型的濾波器，即利用一定數量的隨機樣本(粒子)來表示動態模型的事後隨機分佈[17-19]，以得到一個近似的機率密度函數來描述物體的狀態。粒子濾波器的理論架構主要分為四部份：機率分佈(probability distribution)、動態模型(dynamic model)、測量(measurement)和因子取樣(factored sampling)，分別敘述如下。

###### 1) 機率分佈：

一個物體的特徵狀態是以  $x \in X$  的狀態向量來表示，假設我們不知道精確的狀態，則將此物體的知識用機率密度函數  $p(x)$  來描述，如圖8所示。

###### 2) 動態模型：

當一個物體隨著時間改變，則機率密度函數也必須跟著改變，用以適應改變後的物體狀態。為了計算的可行性，預測是以離散時間的方式執行，又物體的動態模型，係以隨機的微分方程式描述，其中決定性的部份是屬於系統既有的知識，而隨機的部份允許我們自由選定。又微分方程式中，變數的次方可以是任

意階數，且機率密度函數  $p(x_t)$  僅與先前時間的機率密度函數  $p(x_{t-1})$  有關，因此動態模型是以條件式的機率密度函數  $p(x_t | x_{t-1})$  所決定。

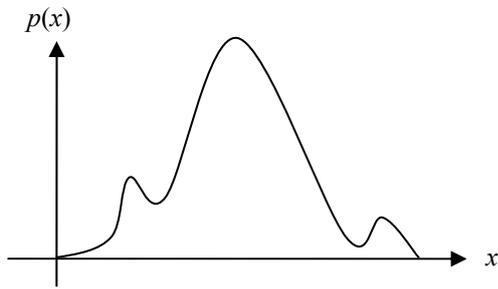


圖8 使用機率密度函數  $p(x)$  表示動態物體的知識。

### 3) 測量：

令  $Z_t = \{z_0, z_1, \dots, z_t\}$  是一測量序列，其中  $z_t$  是在某時間步驟  $t$  的測量。為了追蹤移動物體，我們必須在每一個時間步驟就更新一次，以便包含新的測量  $z_t$ ，此時係以事後機率密度函數  $p(x_t | Z_t) = p(x_t | z_t, Z_{t-1})$  代替機率密度函數  $p(x_t)$ ，而前者與事前機率密度函數  $p(x_t | Z_{t-1})$  的關係可以根據貝氏定理推導獲得，如式(4)所示。

$$\begin{aligned} p(x_t | Z_t) &= \frac{p(z_t | x_t, Z_{t-1})p(x_t | Z_{t-1})}{p(z_t | Z_{t-1})} \\ &= K' p(z_t | x_t, Z_{t-1})p(x_t | Z_{t-1}) \\ &= K p(z_t | x_t)p(x_t | Z_{t-1}) \end{aligned} \quad (4)$$

其中  $K'$  與  $K$  是正規化因子且觀察 (observation) 機率密度函數  $p(z_t | x_t)$  為一個概似 (likelihood) 函數，它係用來表示某一個狀態  $x_t$  測量到  $z_t$  的機率。又事前機率密度函數  $p(x_t | Z_{t-1})$  係將動態模式應用到該機率密度函數在前一個時間步驟的結果，即  $p(x_{t-1} | Z_{t-1})$ ，如式(5)所示。

$$p(x_t | Z_{t-1}) = \int p(x_t | x_{t-1})p(x_{t-1} | Z_{t-1})dx_{t-1} \quad (5)$$

完整的移動物體追蹤步驟，首先必須使用動態模型去計算事前機率密度函數  $p(x_t | Z_{t-1})$ ，然後給予一個新的測量，以評估事後機率密度函數  $p(x_t | Z_t)$ ，如式(6)所示。

$$p(x_{t-1} | Z_{t-1}) \xrightarrow{\text{dynamic model}} p(x_t | Z_{t-1}) \xrightarrow{\text{measurement}} p(x_t | Z_t) \quad (6)$$

### 4) 因子取樣：

一般來說，事後機率密度函數  $p(x_t | Z_t)$  非常複雜，以致無法用一個簡單的式子來表示，也因為  $x_t$  是一個隨時間步驟  $t$  變化的多維度隨機變數，所以我們改用疊代取樣的方法，尋找出一個最相似的機率密度

函數來取代  $p(x_t | Z_t)$ ，以下藉由式(7)來說明此方法：

$$f_1(x_t) = f_2(x_t) \circ f_1(x_{t-1}), \quad x_{t-1}, x_t \in X_t \quad (7)$$

此處  $f_1(\cdot)$  為描述物體的動態模型之隨機取樣函數， $f_2(\cdot)$  為決定所取樣本的隨機權重函數，而「 $\circ$ 」為因子取樣運算子，如下所述。

設一個含有  $N$  個樣本的集合  $S = \{s_1^{(1)}, s_1^{(2)}, \dots, s_1^{(N)}\}$ ，其中  $s^{(j)} \in X$ ，它是利用機率密度函數  $f_1(\cdot)$  隨機抽樣取出，此時，令每個樣本的權重皆相同，如式(8)所示。

$$\pi_1(s^{(j)}) = \frac{\sum_{j=1}^N f_1(s^{(j)})}{N}, \quad j = 1, 2, \dots, N \quad (8)$$

接著，每一個樣本  $s^{(j)}$  的權重，透過機率密度函數  $f_2(\cdot)$  測量所得的值加以調整，如式(9)所示。

$$\pi_2(s^{(j)}) = \frac{f_2(s^{(j)})}{\sum_{j=1}^N f_2(s^{(j)})}, \quad j = 1, 2, \dots, N \quad (9)$$

最後，根據每個樣本的新權重，經過計算選取出新的樣本集合  $S'$ ，而這些樣本所對應的機率密度函數傾向於  $f_1(\cdot)$ ，即為描述物體的動態模型之新隨機取樣函數，圖9為一次疊代中的因子取樣示意圖。

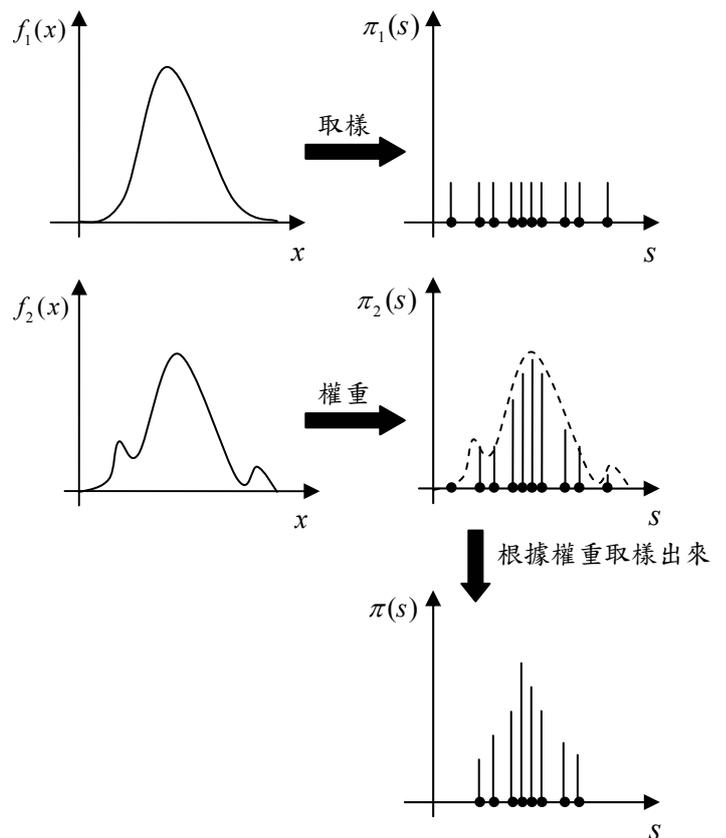


圖9 因子取樣的示意圖。

以下介紹如何實作一個粒子濾波器，它可分成初始化、預測、觀察和選擇四大部分。

1) 初始化：

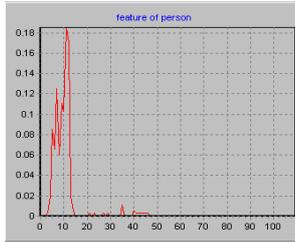
首先，我們選用以顏色為特徵來當做追蹤移動物體的依據，在戶外場景的情況下，RGB色彩空間很容易被陰影所干擾，因此，我們改用HSV的色彩空間，而顏色特徵係以統計直方圖來表示，其中HSV色彩空間映射至一維的區段(bin)編號，如式(10)所示。

$$n_{HSV} = n_H n_S + n_V \quad (10)$$

此處  $n_H \in \{1, 2, \dots, N_H\}$ ， $n_S \in \{1, 2, \dots, N_S\}$  與  $n_V \in \{1, 2, \dots, N_V\}$  分別為色調、飽和度與灰階值的區段編號，因而得知： $n_{HSV} \in \{2, 3, \dots, N_H N_S + N_V\}$ 。圖10(a)為一個物體追蹤的實例，追蹤的目標位置是以胸膛為主，因為此部位比較不容易出現遮蔽等現象，其中取  $N_H = 10$ ， $N_S = 10$  且  $N_V = 10$ ，而圖10(b)為該目標位置的統計直方圖。



(a)



(b)



(c)

圖10 (a)原始的追蹤目標影像；(b)目標位置的統計直方圖；(c)預測出來的候選區域。

接著，我們選擇一個物體的狀態向量來表示該目標影像，如式(11)所示。

$$X_t = [x_t \ y_t \ H_x \ H_y]^T \quad (11)$$

其中  $(x_t, y_t)$  為目標影像在時間  $t$  的區域中心位置，而  $H_x$  和  $H_y$  分別代表它的寬與高。

2) 預測：

樣本集合  $S'$  必須根據物體的動態模型預測出來，於此，我們選擇隨機步行(random walk)的模式來模擬，如式(12)所示。

$$\begin{aligned} x_t &= x_{t-1} + \Delta x_t \\ y_t &= y_{t-1} + \Delta y_t \end{aligned}$$

$$\begin{aligned} H_{x_t} &= H_{x_{t-1}} + \Delta H_{x_t} \\ H_{y_t} &= H_{y_{t-1}} + \Delta H_{y_t} \end{aligned} \quad (12)$$

其中  $\Delta x_t$ ， $\Delta y_t$ ， $\Delta H_{x_t}$  和  $\Delta H_{y_t}$  皆經由高斯分佈(Gaussian distribution)取樣得出。圖10(c)中的黑色框框為利用此方法所預測出來的候選區域。

3) 觀察：

再來，我們令  $b_t(u) \in \{2, 3, \dots, N\}$  為某像素的顏色向量在時間  $t$  於影像位置  $u$  的區段索引， $R(X_t)$  為候選區域的範圍，而  $Q(X_t) = \{q(n; X_t)\}_{n=2,3,\dots,N}$  代表候選區域在時間  $t$  的顏色分佈，計算方法如式(13)所示，其中  $N = N_H N_S + N_V$ 。

$$q(n; X_t) = K'' \sum_{u \in R(X_t)} \omega(|u-n|) \delta[b_t(u)-n] \quad (13)$$

此處  $\delta(\cdot)$  為Kronecker delta函數， $K''$  是正規化常數，用來確保  $\sum_{n=2}^N q(n; X_t) = 1$ ，而  $\omega(\cdot)$  為權重函數，為了簡單起見，本系統將它設為1。

於時間  $t$  的步驟裡，顏色模型  $Q(X_t)$  和參考顏色模型  $Q^* = \{q^*(n)\}_{n=2,3,\dots,N}$  做比對，其中  $\sum_{n=2}^N q^*(n) = 1$ 。在我們的實驗中，初始的參考模型  $Q_{t_0}^*$  是由背景相減之後的前景物體  $X_{t_0}$  所得出，如式(14)所示。

$$Q_{t_0}^* = Q(X_{t_0}) \quad (14)$$

隨後，我們使用Bhattacharyya相似係數[20]當做一個距離函數，來表示候選區域與參考區域的統計直方圖之差異程度，如式(15)所定義。

$$D[Q^*, Q(X_t)] = \left[ 1 - \sum_{n=2}^N \sqrt{q^*(n)q(n; X_t)} \right]^{\frac{1}{2}} \quad (15)$$

對於概似函數的選擇，我們基於高斯機率密度函數去測量統計直方圖，如式(16)所示。

$$p(Y_t | X_t) \propto \exp\{-\lambda D^2[Q^*, Q(X_t)]\} \quad (16)$$

其中  $Y_t$  為候選目標影像的測量狀態向量；此式經過實驗測試後，選取  $\lambda = 20$  有比較好的效果。由於本系統係採用背景相減法，將前景物分離出來，因此可利用背景資訊來增加追蹤移動物體的強健性，方法如下所述。

當背景建立好後，我們把候選區域的部份和相對應的背景相減。接著，計算該候選區域中的前景物佔據整體區域的百分比，然後，把候選區域的概似函數之加權，調整為所對應的百分比；例如：原本概似函數的加權為0.2，前景物佔據30%的範圍，則概似函數的加權變成  $0.2 + 0.2 \times 30\% = 0.206$ 。圖11為加入背景資訊求取新的概似函數之流程圖。

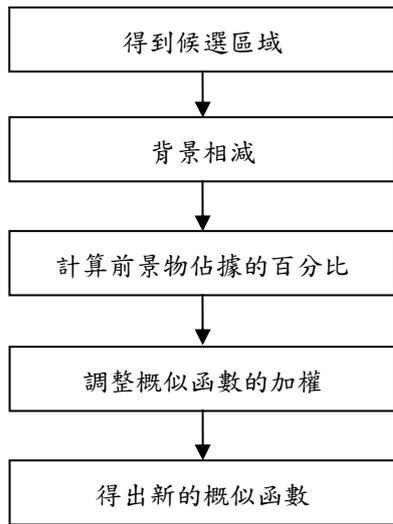


圖11 加入背景資訊求取新的概似函數之流程圖。

圖12為有否加入背景資訊的移動物體追蹤結果之比較；由此圖可知：含有背景資訊所預測的移動物體之位置，其正確性較高。

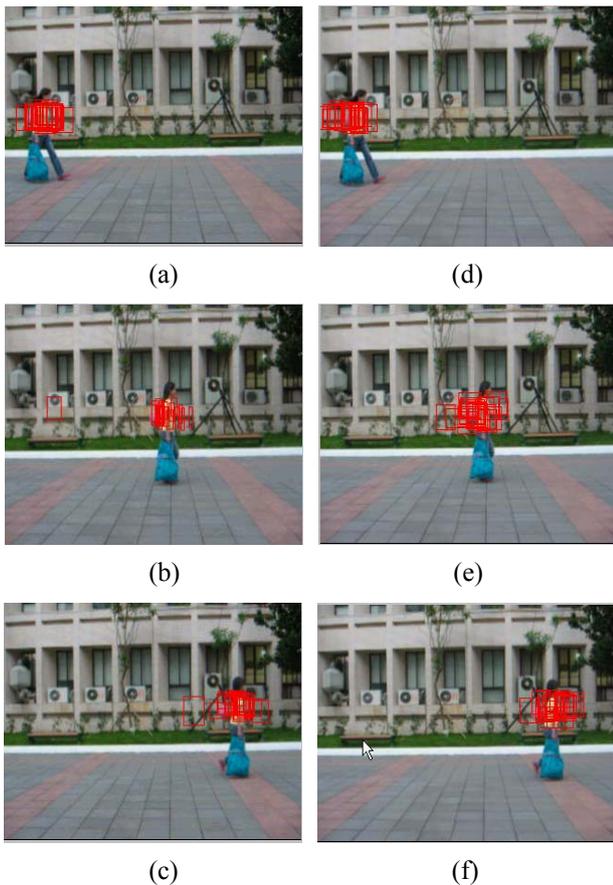


圖12 (a)~(c)為第61、80、90個畫面不含背景資訊的物體追蹤結果；(d)~(f)為前述畫面含有背景資訊的物體追蹤結果。

#### 4) 選擇：

本系統係參考Carpenter等人所提出的方法來取樣[21]，以下是其進行的步驟。首先，以均勻分佈(uniform distribution)的亂數取 $M$ 個值，用 $\{r_i\}_{i=1,2,\dots,M}$ 來表示，再將它們轉換成指數型態的分佈(exponential distribution)，以 $\{t_i\}_{i=1,2,\dots,M}$ 來表示，其中 $t_i = -\log(r_i)$ 。

最後，令 $T_j = \sum_{i=1}^j t_i$  且  $W_j = \sum_{i=1}^j w_i$ ，此處 $w_i$ 為概似函數正規化後的比率，再根據圖13所示的程序來取樣，其中 $s_j$ 表示經過重新取樣後的樣本，此即對應移動物體的候選區域。

```

i = 1 ; j = 1
do while j ≤ M
  if  $W_j T_M > T_i$  then
    i = i + 1 ; output  $s_j$ 
  else
    j = j + 1
  end if
end do

```

圖13 粒子濾波器的取樣程序。

圖14為粒子濾波器的執行示意圖；首先，經初始化後，依次透過預測、觀察、選擇三個程序以疊代的方式進行。

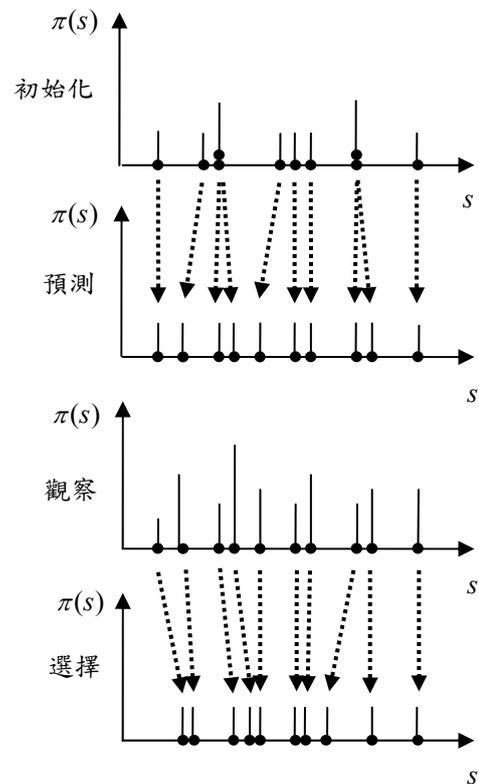


圖14 粒子濾波器的執行示意圖。

## 五、實驗結果與討論

本系統的各项實驗是在下列的電腦軟、硬體環境完成：使用Pentium 4 2.4GHz 中央處理器(CPU)配備512MB記憶體，作業系統為Microsoft Windows XP Professional，而程式開發工具為Borland C++ Builder 6.0。

### 5.1 物體偵測的實驗結果

以下針對戶外與室內幾種場景做移動物體的偵測實驗，用來驗證視訊監控系統的強健性，同時我們也舉出效果較不理想的例子。

#### 1) 戶外場景一：

圖15顯示的場景有一半的光源是被建築物所遮蔽，一半沒有。由於影片的品質比較差，因此產生的雜訊較多，但我們透過相連元件偵測的處理，過濾掉小於百分之一的區塊，所以可以成功地偵測我們所期望的移動物體。

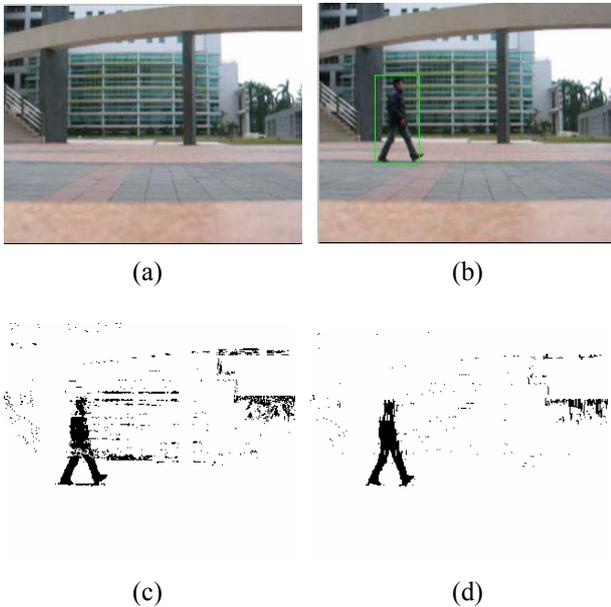


圖15 戶外場景的物體偵測實驗結果一：(a)背景影像；(b)目前影像；(c)前景影像；(d)去除陰影與雜訊之後的影像。

#### 2) 戶外場景二：

在圖16的場景中，背景的樹枝隨風搖動，因此會反映在前景影像，但我們藉由限制大小必須超過整張畫面百分之一的相連元件來消除可能產生的錯誤。

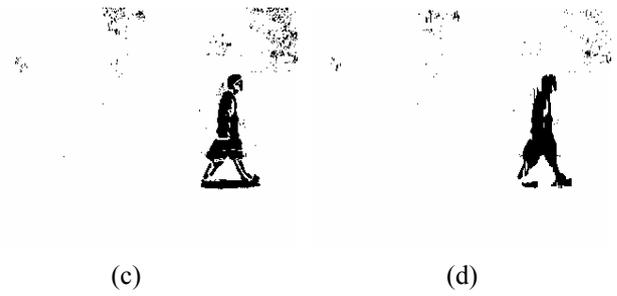
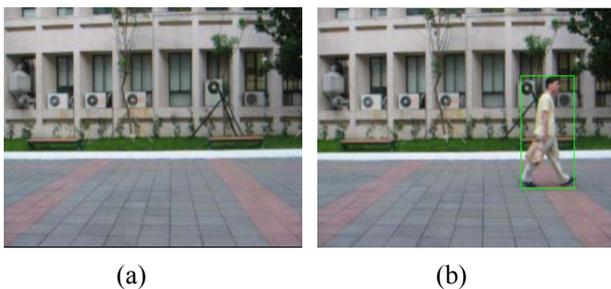


圖16 戶外場景的物體偵測實驗結果二：(a)背景影像；(b)目前影像；(c)前景影像；(d)去除陰影與雜訊之後的影像。

#### 3) 室內場景一：

對於圖17的監視場景，本系統可以有效地把移動物體的陰影所造成的影響降低，不致於發生太離譜的偵測錯誤。

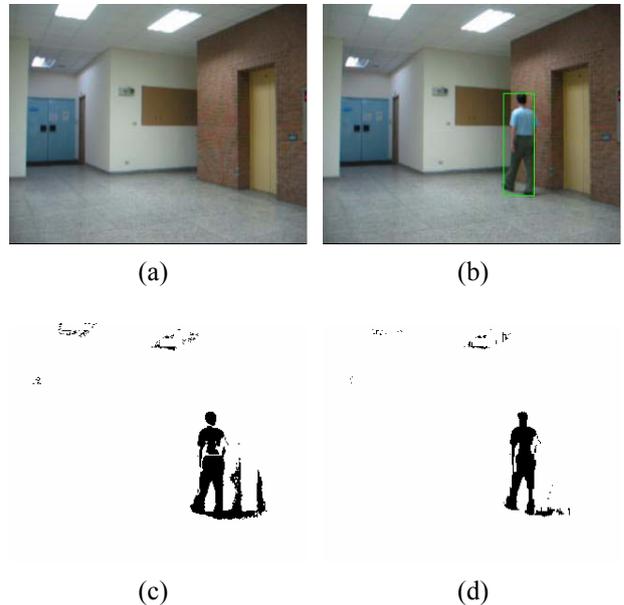


圖17 室內場景的物體偵測實驗結果一：(a)背景影像；(b)目前影像；(c)前景影像；(d)去除陰影與雜訊之後的影像。

#### 4) 室內場景二：

於圖18的場景裡，雖然我們把陰影所造成的影響給去除，但也犧牲了場景中人物腳的部份，以致於沒辦法很完整地框出移動物體的形狀。

### 5.2 物體追蹤的實驗結果

在移動物體追蹤的實驗當中，我們並不是對每一個畫面都做偵測的動作，而是每隔一段時間才需要做一次偵測，如此可以節省計算量並且還能判斷是否有新的移動物體進入監視中的場景。當我們偵測到移動的物體後，則根據所得到的特徵來做追蹤的動作。圖19~圖22係延續5.1節的物體偵測實驗，展示移動物體的追蹤結果，其中每個實驗皆可達到視訊監控系統所要求的即時處理的效能。



圖18 室內場景的物體偵測實驗結果二：(a)背景影像；(b)目前影像；(c)前景影像；(d)去除陰影與雜訊之後的影像。

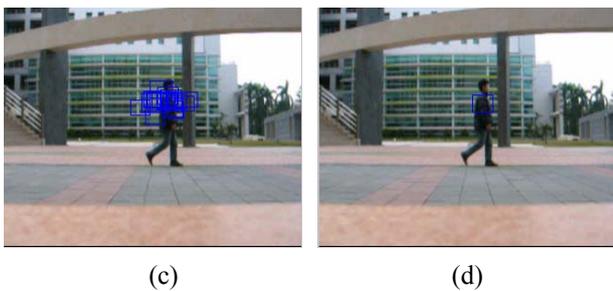
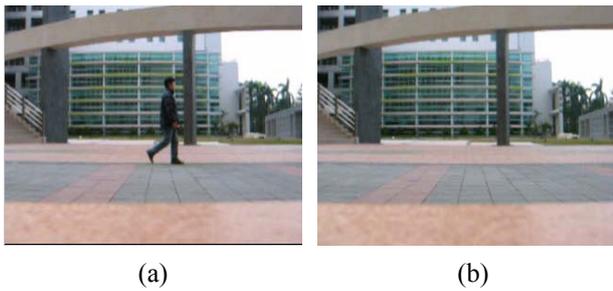


圖19 戶外場景的物體追蹤實驗結果一：(a)目前影像；(b)背景影像；(c)框出移動物體的候選區域；(d)框出最有可能是移動物體的區域。



圖20 戶外場景的物體追蹤實驗結果二：(a)目前影像；(b)背景影像；(c)框出移動物體的候選區域；(d)框出最有可能是移動物體的區域。

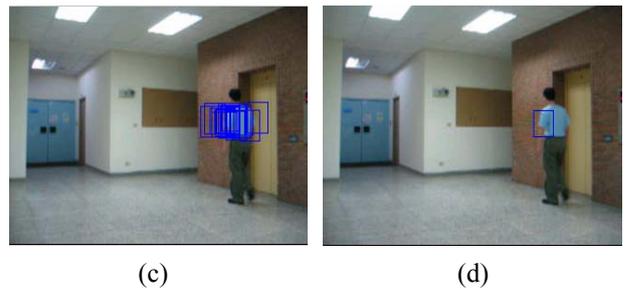


圖21 室內場景的物體追蹤實驗結果一：(a)目前影像；(b)背景影像；(c)框出移動物體的候選區域；(d)框出最有可能是移動物體的區域。

### 5.3 實驗結果討論

以下為多個移動物體的偵測與追蹤之實驗結果的綜合討論。

- 1) 本系統的執行速度為每秒30至35張畫面，可以達到即時視覺偵測與追蹤多個移動物體的效果。
- 2) 改良過的背景建立方法，比起建構式所需的運算速度快得多。
- 3) 當環境的光源過暗時，本系統無法保證將移動的物體偵測出來。
- 4) 若移動物體的顏色和背景的顏色太接近，則無法很完美地將它偵測出來。
- 5) 於物體偵測的階段裡，雖然無法完全去除伴隨前景物的陰影，但是已經能夠有效地改善其結果。
- 6) 本系統加入了背景資訊來提升追蹤移動物體的正確性，比起單純只用顏色資訊來追蹤，所獲得的效果好很多。

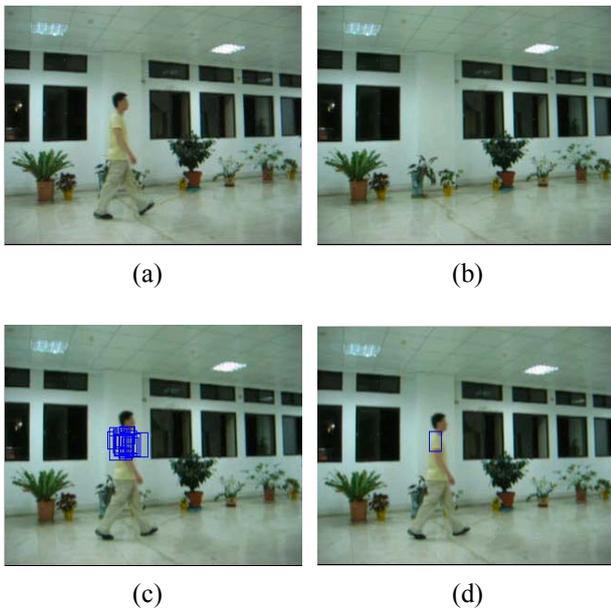


圖22 室內場景的物體追蹤實驗結果二:(a)目前影像;(b)背景影像;(c)框出移動物體的候選區域;(d)框出最有可能是移動物體的區域。

## 六、結論與未來研究方向

近年來，由於電腦視覺技術的快速發展，使得它在視訊監控系統各領域的應用已日益增加，例如：入侵物偵測系統、人體運動與姿勢分析系統，以及智慧型運輸系統中的車流分析系統、道路上危險車輛偵測系統，與以視覺為基礎的駕駛安全輔助系統及停車場安全管理系統等，在這些系統中，移動物體偵測與追蹤的技術扮演了相當重要的角色。

### 6.1 結論

本論文提出一個快速且有效的多個移動物體偵測與追蹤的系統。在物體偵測的階段，我們分成四個程序來實作：

- 1) 背景建立：本系統使用過去 $N$ 張畫面的每個像素之中位數來當做背景的像素。
- 2) 前景偵測：針對低對比度的門檻值問題，本系統採用[12]所提出的擷取函數來解決。
- 3) 陰影去除：本系統使用[5]所提出的去除陰影的方法來改善這一類的問題。
- 4) 背景更新：我們參考漸進式背景影像建構的方法[15]，進而提出自己的背景更新方法。

又在物體追蹤的階段，我們使用粒子濾波器來作為追蹤移動物體的方法。對於欲追蹤的物體，我們用顏色來當做特徵。此外，我們利用背景資訊來調整每一個候選區域的權重，讓移動物體的追蹤更為準確。

### 6.2 未來研究方向

本論文在背景更新的程序和物體追蹤的階段裡所提出的新方法，雖然對於一般環境下有良好的效果，

但是仍有一些缺失需要進一步改善。以下是本系統需要再強化的地方以及解決的方法。

#### 1) 背景稍微變動的影響：

在背景影像中，不全都是靜止不動的像素，例如：搖晃的樹枝、閃爍的螢幕、架在戶外的攝影機受到震動等等，這些情況下都會導致前景偵測的錯誤發生。我們可以試著讓更動頻繁的背景像素記錄下來，以後出現這一類的像素就當做背景處理。

#### 2) 門檻值的自動建立：

目前系統中所使用的門檻值是採取經驗值，比較沒有適應性，應該改用自動建立門檻值的方法，例如：Otsu[22]或Kapur[23]等人所提出的自動建立門檻值的方法。

#### 3) 加入新特徵以達到更準確的追蹤：

因為我們使用顏色特徵作為追蹤移動物體的依據，所以在背景和前景顏色相似時，有可能會發生追蹤錯誤的情況。因此，我們可以在追蹤移動物體時，加入新的特徵，例如：輪廓，以提高物體追蹤的準確度。

#### 4) 學習移動物體的運動方式：

在本系統中，假設所有的物體都是以隨機的方式移動，但實際狀況可能是每個物體有其獨特的運動方式，如果能夠使用機器學習的技巧，例如：隱藏式馬可夫模型(hidden Markov model)，來學習物體的運動模式，可以讓移動物體的追蹤更為準確。

## 七、參考文獻

- [1] T. Horprasert, D. Harwood, and L. S. Davis, "A statistical approach for real-time robust background subtraction and shadow detection," in *Proc. of the IEEE Int. Conf. on Computer Vision*, pp. 918-923, 1999.
- [2] 陳奕光, 「即時多個運動物體之偵測與追蹤於場景監視系統的應用」, 碩士論文, 國立台灣科技大學資訊工程系, 台北市, 民國93年。
- [3] I. Haritaoglu, D. Harwood, and L. Davis, "W<sup>4</sup>: Real-time surveillance of people and their activity," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 809-830, 2000.
- [4] A. Talukder and L. Matthies, "Real-time detection of moving objects from moving vehicles using dense stereo and optical flow," in *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, vol. 4, pp. 3718-3725, 2004.
- [5] R. Cucchiara, C. Grana, G. Neri, M. Piccardi, and A. Prati, "The Sakbot system for moving object detection and tracking," in *Proc. of the 2nd European Workshop on Advanced Video-based Surveillance*, pp. 145-157, 2001.
- [6] 莊建宏, 「自動化交通監控系統」, 碩士論文, 國立台灣師範大學資訊教育學系, 台北市, 民國90年。

- [7] L. Li, W. M. Huang, I. Y. H. Gu, and Q. Tian, "Statistical modeling of complex backgrounds for foreground object detection," *IEEE Trans. on Image Processing*, vol. 13, no. 11, pp. 1459-1472, 2004.
- [8] C. Stauffer and W. Grimson, "Learning patterns of activity using real-time tracking," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 747-757, 2000.
- [9] L. Li and M. Leung, "Integrating intensity and texture differences for robust change detection," *IEEE Trans. on Image Processing*, vol. 11, no. 2, pp. 105-112, 2002.
- [10] O. Javed, K. Shafique, and M. Shah, "A hierarchical approach to robust background subtraction using color and gradient information," in *Proc. of the IEEE Workshop on Motion Video Computing*, pp. 22-27, 2002.
- [11] L. Li, W. M. Huang, I. Y. H. Gu, and Q. Tian, "Foreground object detection in changing background based on color co-occurrence statistics," in *Proc. of the IEEE Workshop on the Applications of Computer Vision*, pp. 269-274, 2002.
- [12] Y. Kuno, T. Watanabe, Y. Shimosakoda, and S. Nakagawa, "Automated detection of human for visual surveillance system," in *Proc. of the Int. Conf. on Pattern Recognition*, pp. 865-869, 1996.
- [13] S. Nadimi and B. Bhanu, "Physical models for moving shadow and object detection in video," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 26, no. 8, pp. 1079-1087, 2004.
- [14] J. Heikkila and O. Silven, "A real-time system for monitoring of cyclists and pedestrians," in *Proc. of the 2nd IEEE Workshop on Visual Surveillance*, pp. 74-81, 1999.
- [15] 王俊明, 陳世旺, 「漸進式背景影像建構」, 師大學報: 數理與科技類, 第47卷, 第2期, 第43-54頁, 民國91年。
- [16] 范國清, 王元凱, 陳炳富, 「追蹤演算法簡介」, 影像與識別, 第8卷, 第4期, 第17-30頁, 民國91年。
- [17] K. Okuma, A. Taleghani, N. de Freitas, J. Little, and D. Lowe, "A boosted particle filter: multitarget detection and tracking," in *Proc. of the European Conf. on Computer Vision*, pp. 28-39, 2004.
- [18] P. Pérez, C. Hue, J. Vermaak, and M. Gangnet, "Color-based probabilistic tracking," in *Proc. of the European Conf. on Computer Vision*, pp. 661-675, 2002.
- [19] Y. Ivanov, C. Stauffer, A. Bobick, and W. E. L. Grimson, "Video surveillance of interactions," in *Proc. of the 2nd IEEE Workshop on Visual Surveillance*, pp. 82-90, 1999.
- [20] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 2, pp. 142-149, 2000.
- [21] J. Carpenter, P. Clifford, and P. Fearnhead, "Improved particle filter for nonlinear problems," *IEE Proc.- Radar, Sonar and Navigation*, vol. 146, no. 1, pp. 2-7, 1999.
- [22] N. Otsu, "A threshold selection method from gray level histogram," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 8, no. 1, pp. 62-66, 1978.
- [23] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram," *Computer Vision, Graphics, and Image Processing*, vol. 29, no. 3, pp. 273-285, 1985.