

# 以光流為輔助的動態背景視訊物件偵測

紀伯靜 林孝杰\* 戴世峯 L. Lan

國立雲林科技大學 電子與光電工程研究所

\*Email: m10013237@yuntech.edu.tw

## 摘要

在電腦視覺的研究中，移動物件的偵測是相當重要的一環。一般視訊物件偵測方法，大多處於靜態環境的背景下進行，在動態背景的辨識上往往須另行處理。本論文採用移動向量做為特徵，於動態背景干擾下依舊能有效辨別出前景和去除動態背景。所提方法適用時機為攝影機固定並且背景有自然界的動態變化，如：隨風飄動的樹枝、噴泉、水流等場合之視訊監控。我們的方法主要步驟為(1)利用前處理單元篩選出可能的特徵點，減少運算及儲存量；(2)利用光流法預估移動向量，並採用特徵點方向性搭配鄰域投票機制來分離出動態背景和移動物件；(3)採取後處理動作來進行移動物件的位置確認。最後我們進行視訊的實際測試，發現所提出的方法在動態背景擾動下、運動物件移入、移出畫面時，大致都能有效地偵測出移動物件。

關鍵字：移動物件偵測、移動向量、光流、去除動態背景

## 1. 前言

隨著科技發展，在人們生活上也產生重大影響，近年來，在家家戶戶都有車輛的情形下，一到假期就會造成交通上的擁塞，如何管制也讓相關交通監控和流量警報的研究受到注意；另外鄰里間偷竊案件、人與人相處的各種糾紛，甚至國際犯罪(如：恐怖攻擊)等被新聞媒體報導的案件，皆有賴影像監控來還原真相或者解決問題，這都讓影像監控系統的研究備受重視。近年來監控系統的需求增加，數位影像處理、通訊處理、信號處理技術，加上硬體設備的進步，即時監控系統的發展相當快速，數位影像處理將有更廣泛的應用，如：手勢或人臉辨識、影像修正、影像壓縮、娛樂遊戲、人機介面(Human-Computer Interfaces)等等，若能再與其他系統設備搭配將可以做到更多的自動化應用，這都使數位影像處理各應用領域的研究變得相當熱絡並且受到相當的關注。

視訊物件偵測一直是影像處理中重要的一環，同時也是專家學者們感興趣的研究領域，並且嘗試將偵測結果與其他追蹤方式做相關結合。移動物件偵測目的在於正確偵測出影像中可能組成的物件，讓電腦擁有判斷和分析影像特徵的能力，藉由相關特徵分析來分離出我們所要得到或找尋的移動物件及非移動的背景或其他靜態物件。為了達到這樣的效果，需要讓電腦具備理解並分析影像中的資訊，同時因應即時顯示的需求，在實行偵測的演算法複雜度也需要特別注意執行速度上的考量。

近年來移動物件偵測方法也有許多研究與應用在發展，如：背景相減法[1,2]、光流法[3,4]、單高斯模型[5]、混合高斯模型[5,6]等，這些偵測方法皆有以連續時間內像素值的變化來建立一參考背景的方式達到偵測移動物件的目標，但一般來說物件偵測常採用色彩特徵，很容易受到不同視訊的背景雜訊干擾。2010年由W. Zhang與Q.M. Jonathan Wu以及H. Yin提出利用移動向量做為偵測的方式[4]，其方法主要以整張畫面中所有像素點做移動方向的直方圖(histogram)統計，以此偵測出移動物件。避免採用色彩資訊當作特徵，改採用較直觀的移動向量(Motion Vector, MV)當作特徵來辨識移動物件的方向和位置，避免特徵

本身與背景及其他雜訊的相互干擾，達到有效減少錯誤判斷的產生，同時能有效偵測出移動物件。本論文則根據統計移動方向的方式提出以光流法估測移動向量配合鄰域投票機制來有效去除動態背景雜訊，除了避免採用色彩等特徵易產生的背景及雜訊干擾，並使用鄰域投票機制取代統計整張畫面方向的方式，藉由篩選出特徵點的方式進行小範圍的統計，除了可以減少運算量之外，也能專注在可能為移動物件的位置周遭的方向關聯性判斷，避免受到其他統計數據的影響。

本論文提出方法先藉由前處理單元篩選可能為移動物件的特徵點後，再進行光流法估測該特徵點集合的移動向量，再利用去除動態背景的鄰域統計方法分離出動態背景與前景，最後進行後處理的動作，有效估計出移動物件的位置。本文提出之系統，我們稱之為移動向量方向投票法(Motion-Vectors Direction Voting, MVDV)，在本文的實驗結果中，利用此方法測試不同視訊，皆可以有效去除動態景干擾。

本文結構如下，第二節介紹介紹移動向量估測方法，包含光流法預估及應用於影像壓縮中的移動估計演算法，第三節為探討以移動向量作為特徵的視訊物件偵測流程及系統架構分析，第四節為實驗結果與比較，第五節為結論，最後一節為參考文獻。

## 2. 移動向量估測

在使用移動向量估測移動物件的方法中，一般以光流法[4,7]和應用在影像壓縮中的移動估計演算[3,4]為主。光流法可以藉由亮度變化以及在限制條件下進行光流向量的估計，利用此光流向量就可將該像素點位置的方向和速率計算出一預測值；而在影像壓縮方面，移動估計演算法的移動量預估是相當重要的環節，也是視訊編碼器達成高壓縮率的主因。

### 2.1 LK光流法

光流法是從連續影像中找到畫面中像素點的亮度變化，從影像平面上推導出移動物件的相對運動速度，即為光流向量，再以此向量來預估出移動物件的方向等資訊。

本節將介紹以LK光流法[4,8,9]為本論文的移動向量預估方法原理。LK光流法是由Lucas與Kanade在1981年提出的光流演算法[10]，主要原理係利用影像中灰階梯度並用牛頓-拉福生疊代法(Newton-Raphson iterations)來進行影像匹配(image matching)，經常應用在二維物件的偵測和特徵追蹤上。

假設一樣版影像  $T$  (template)，要在輸入影像  $I$  當中找到與樣板影像同樣大的區域並且灰階值相同，由樣板  $T$  到影像  $I$  的轉換可以由一轉換函數  $F$  表示，如式(1)：

$$I(F(x_r, p)) = I(x_r), \quad x_r = F(x_r, p) \quad (1)$$

其中  $P$  為參數， $F$  稱為仿射轉換(affine transform)，求得最佳的參數  $P$  值，使得樣版  $T$  當中每個在  $x_r$  位置像素點的像素值經由轉換函數  $F$  轉換後會與影像  $I$  在  $x_r$  位置像素點的像素值相同。由於影像的亮度分佈是非線性，在經過時間推移後，將前一時刻  $t-1$  的參數  $P$  及目前時刻  $t$  所待估測的參數  $P$  之間的關係寫成式(2)，並藉估測  $P$  求得時間  $t$  的參數值，改寫成式(3)：

$$p_{t-1} + \Delta p = p_t \quad (2)$$

$$I(F(x_t, p_{t-1} + \Delta p), t) = T(x_t) \quad (3)$$

在此影像匹配的計算過程中，找到一組參數  $\Delta p$  使得最後的目標函數  $O(\Delta p)$  最小，此目標函數利用關係式將參數  $p$  改寫，利用疊代的方式趨近最佳值，如式(4)：

$$O(\Delta p) = \sum_{x_t} (I(F(x_t, p_{t-1} + \Delta p), t) - T(x_t))^2 \quad (4)$$

將  $I(F(x_t, p_{t-1} + \Delta p))$  利用  $\Delta p \rightarrow 0$  且泰勒展開來逼近，得到式(5)：

$$I(F(x_t, p_{t-1} + \Delta p), t) \approx I(F(x_t, p_{t-1}), t) + \left( \nabla I \frac{\partial F}{\partial p} \right)^T \Delta p \quad (5)$$

再將式(5)代回式(4)，可得到式(6)：

$$O(\Delta p) = \sum_{x_t} \left( I(F(x_t, p_{t-1}), t) + \left( \nabla I \frac{\partial F}{\partial p} \right)^T \Delta p - T(x_t) \right)^2 \quad (6)$$

其中  $\nabla I$  為  $I(F(x_t, p_{t-1}), t)$  的一階微分，也就是亮度梯度值，即移動向量的垂直與水平分量； $\frac{\partial F}{\partial p}$  為各個轉換參數的微分。後續也有許多基於此LK法的改良[11]。LK光流法相較於其他光流方法可以運用在較稀疏的背景上，也就是利用少數點集合的影像上。

### 2.1.1 金字塔型LK光流法[4,12]

LK光流主要有三項假設條件：(1)連續影像畫面中像素點亮度不變，(2)連續影像當中的物件移動速度慢，(3)像素點周圍有相似的移動向量，但其中第三個假設，影像物件及周圍的移動向量一致性會被相關干擾而破壞，使得第三個假設條件不成立，因此在影像上轉成最小尺度後做光流運算，再到影像下一層的空間尺度運算，以此不斷往下一層運算，直到原來影像的空間尺度，尺寸較小的上層進行分析所得到的資訊可以提供給尺寸較大的下層再進一步分析，以此簡化分析及計算，此法稱為金字塔型LK光流法，示意圖如圖1。

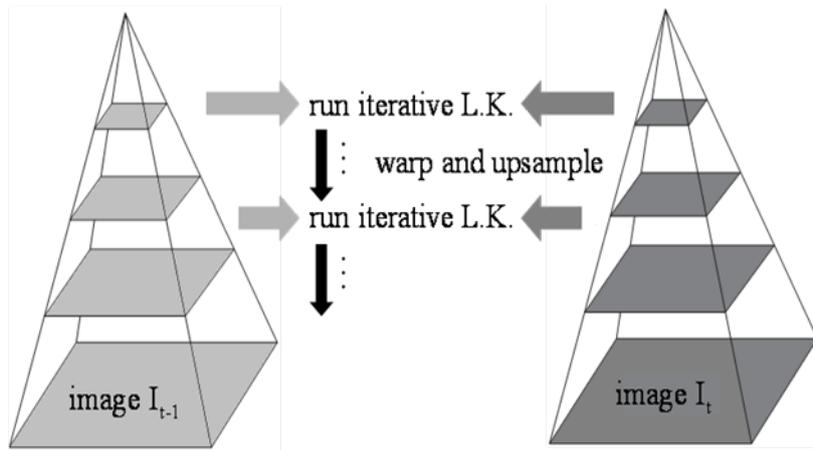


圖1 金字塔光流示意圖

## 2.2 移動估計演算法(Motion Estimation)[7,13,14]

在影像壓縮應用上，經常包含移動向量預估單元，而移動估計演算法的好壞是達成高壓縮率的主要因素。在實際應用上，常以區塊為單位，利用移動向量來描述區塊的移動，這種含有運動資訊的向量稱之為移動向量，之後解碼器只需要利用移動向量的資訊和區塊之間的相似度比對，即可以重建該區塊，而此物體運動過程即是移動估計(Motion Estimation)。

在連續影像畫面中，在  $t$  時刻的影像畫面中找出欲偵測物件的某一位置  $P_t(x, y)$ ，在  $t-1$  時刻的畫面中找出最相似區塊且對應  $P_t(x, y)$  位置的  $P_{t-1}(x, y)$ ，而這個位置的位移量即為我們所要找的移動向量。其數學式如式(7)、(8)，其中  $u$ 、 $v$  分別代表移動向量的垂直和水平的分量，示意圖如下圖2所示：

$$\begin{aligned} u &= P_{t-1}(x) - P_t(x) \\ v &= P_{t-1}(y) - P_t(y) \end{aligned} \tag{7}$$

$$MV = (u, v) \tag{8}$$

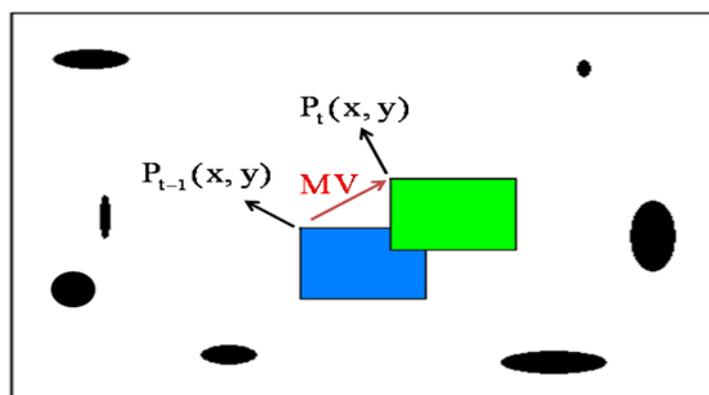


圖2 移動向量MV示意圖

實際應用上以區塊為單位，也就是包含區塊內所有像素點，其中如何找出在  $t-1$  時刻的最相似方塊來求取MV值，其中含有幾種搜索方法[2,3,10,11]，如：全域搜索法(Full search)[2]、三步驟搜索法(Three-step search)[2]、鑽石搜索法(Diamond search)[10,11]，以及比對相似度的計算方式[2,12,13]，一般常見的方法有絕對誤差合(Sum of absolute difference, SAD)、平均絕對誤差合(Mean absolute error, MAE)、均方誤差合(Mean squared error, MSE)等三種。

### 3. 視訊物件偵測流程及系統架構分析

本節描述所提出方法的整體偵測系統架構。主要架構可以分成四個部分，一開始將視訊影像做前處理，留下可能是移動物件的像素點，再進入第二步驟，利用這些可能為物件的像素點集合，搭配LK光流法估計出物件可能的移動向量，接著在步驟三利用移動向量來判斷物件及動態背景的差別，藉此去除動態背景的干擾，最後使用後處理的方式利用剩下的物件像素點集合來有效估計出物件的位置和可能範圍。偵測架構流程圖如圖3所示：

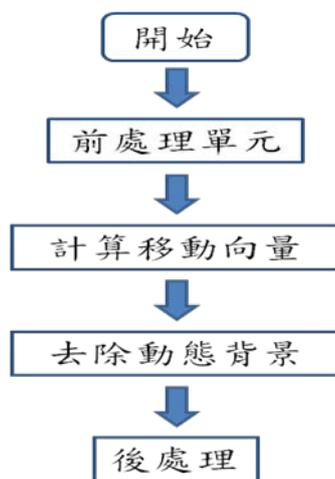


圖3 偵測架構流程圖

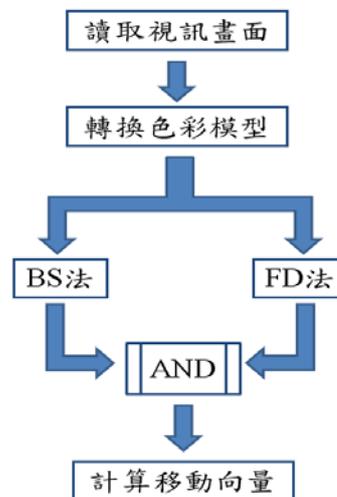


圖4 前處理單元流程圖

#### 3.1 前處理單元

前處理單元目的在於篩選出可能移動物件像素點的位置，保留可能為移動物件的像素點來計算，達到減少計算量的作用。本論文提出的前處理單元是利用HSV色彩模型當中的V值來做單通道估計，在完成色彩轉換的階段後，利用該色彩模型，搭配背景相減法(Background Subtraction, BS)和連續影像相減法(Frame Difference, FD)，就可以達成前處理單元的目標。BS法與FD法做邏輯AND的結果會介於單純個別採用BS法和FD法之間，沒有鬼影和陰影的影響，並且動態雜訊去除較好，但是會產生物件空洞。經過此方法找出可能移動物件像素點，原因取決於減少待觀測的可能移動物件點外，同時保有一定的物件資訊，並且減少其他的雜訊干擾的產生後。最後的二值化結果即為前處理所留下可能為物件的像素點集合。前處理單元的流程圖如圖4所示。

#### 3.2 光流法與移動向量

本論文所提出MVDV法採用OpenCV1.0版本的函式庫當中的金字塔疊代LK光流法的語法來估測出移動向量，利用在前處理單元找到可能物件的點集合去預測可能移動點的移動向量，達到在預估移動向量上盡可能減少運算時間的目標。表1為實際測試不同光流方法所需要計算的特徵點數目，包含HS法[7,8,9]、LK法、BM法[13,14,15,16]，以及本論文所採用的金字塔疊代LK法，平均每張畫面需要計算多少像素點，單位為points/frame，測試視訊皆為  $320 \times 240$  的標準大小。

表1 光流法平均估計像素點數統計表(每張畫面平均計算點數，points/frame)

光流法 視訊名稱	BM法 (block size = 10)	HS法	LK法	金字塔疊代 LK法
Tree	49611.897	76800	76800	653.492
Water	53648.061	76800	76800	727.335
San	47612.583	76800	76800	1512.608

### 3.3 去除動態背景

一般利用移動向量去除動態背景的方法，如文獻[19]提出利用整張影像畫面所統計的方向判斷。我們利用類似概念提出使用八個方向性判斷加上鄰域投票機制的方式來改進，此方式能有效減少計算量並且合理判讀出該像素點為前景或背景。本論文去除動態背景的方法架構可分為八方向法、前景數量濾波、型態學關閉運算，三個方法搭配鄰域投票機制組合而成，採用方法及步驟將在下列各小節詳細介紹。

#### 3.3.1 鄰域投票機制

鄰域投票機制是我們所提出去除動態背景的方法中，經常用到的判斷方式。一般認為某一特徵點附近的像素點與特徵點本身應該會有一定連動的關係，以此假設來做小範圍的判斷，不採用整張畫面的整體結構來觀察，以此做到減少分析整張畫面的計算量，只需以特徵點為中心點，觀察與附近像素點的關係即可。鄰域投票機制示意圖如圖5所示。

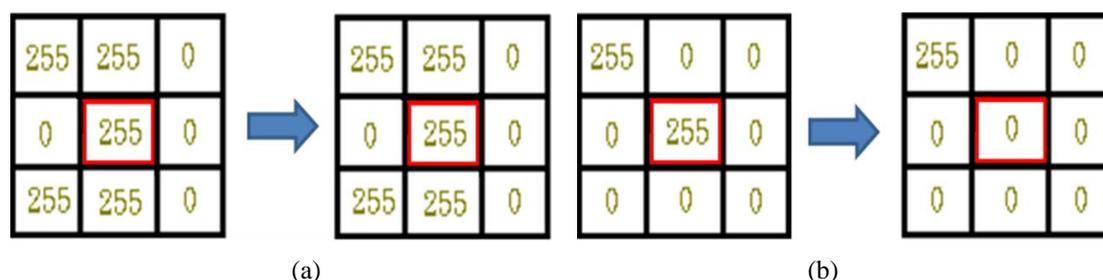


圖5 鄰域投票機制示意圖：(a)統計過半數示意圖。(b)統計未過半數示意圖。

#### 3.3.2 八方向法

八方向法主要是利用移動向量中的 $u$ 分量和 $v$ 分量來代表直角座標中的 $x$ 、 $y$ 分量，把移動方向分成八類，如圖6、表2所示，利用直角座標舉例說明，使用移動向量的正、負號以及兩分量的絕對值大小做為判斷依據，以此分類出八個方向性，其中無法分類出當 $u$ 分量和 $v$ 分量相等時的情形，本論文提出將其歸類於 $u$ 分量大於 $v$ 分量的方向之中，原因取決於在影像偵測或追蹤的研究，常使用的視訊標準大小為 $320 \times 240$ ，換句話說視訊畫面是寬大於高的情形，若兩分量相等時，較可能為 $u$ 分量大於 $v$ 分量的方向，因此將其歸類在此方向內。圖7為以單一像素點為特徵點實際運算八方向法範例圖，圖8為實際進行八方向法前後結果比較圖。

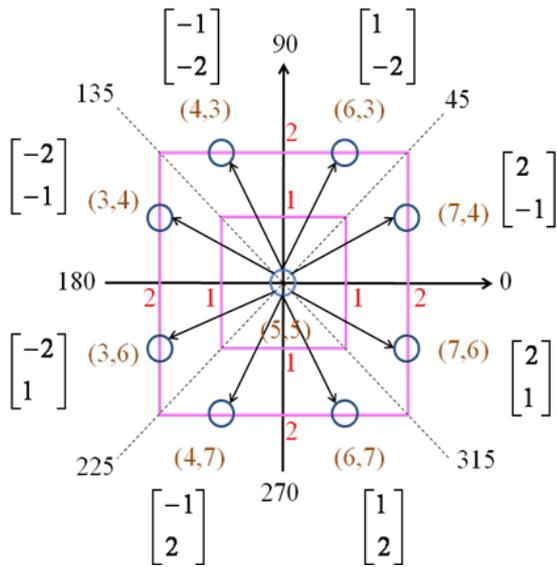
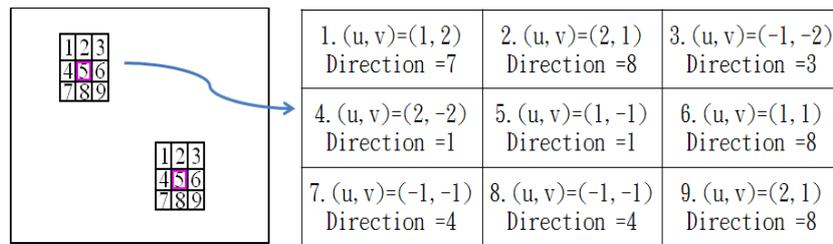


圖6 八方向法分類說明圖

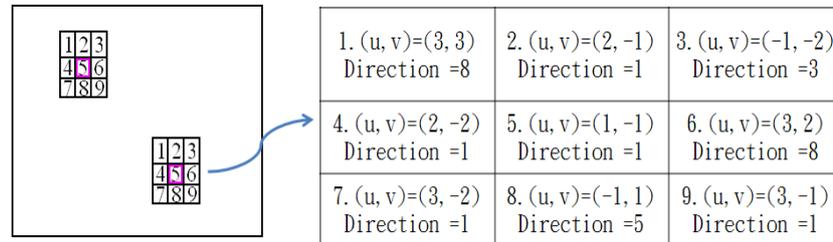
表2 八方向分類統計表

$x(u)$	$y(v)$	$ x(u)  >  y(v) $	$ x(u)  <  y(v) $
正	負	Direction = 1	Direction = 2
負	正	Direction = 4	Direction = 3
負	正	Direction = 5	Direction = 6
正	負	Direction = 8	Direction = 7



Image

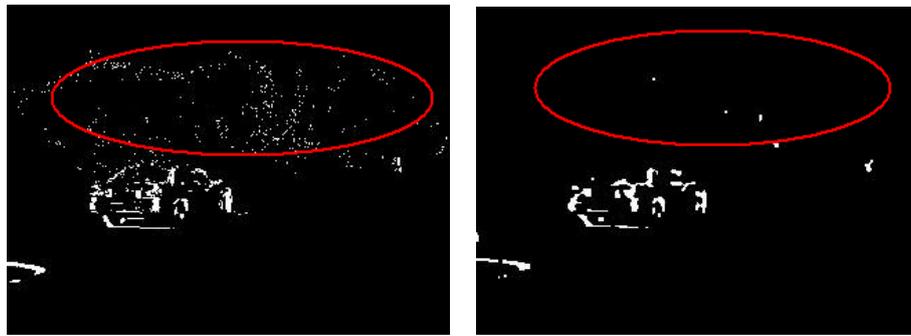
(a)



Image

(b)

圖7 特徵點的八方向法操作示意圖：(a)未過半數同方向示意圖。(b)過半數同方向示意圖。



(a)

(b)

圖8 八方向法前後比較圖：(a)未經過八方向法。(b)經過八方向法。

### 3.3.3 前景數量濾波與關閉運算

一般來說物件特徵點附近也會是物件點，屬於同一移動物件，但同時動態背景的像素點剛好也為同一方向時，利用移動向量的判斷可能會出現誤判而將此點保留。因此利用鄰域投票機制在特徵點附近判斷有多少像素點屬於前景，若超過設定數目則該像素點繼續保留判斷為前景，反之則將該前景點判斷為獨立的雜訊點，將其去除，設定為背景點。此方式即為前景數量濾波。

關閉運算是數學形態學處理的一種常見方法，此方法是由兩種基礎的型態學運算：膨脹(dilation)以及侵蝕(erosion)組合而成，關閉運算主要功能在於將分離的同一物件連結在一起，重新合為同一物件，另外也可以將關閉運算應用在小範圍的區域填滿。

其中膨脹的定義為式(11)所式，主要能產生放大原始物件的效果；而侵蝕定義成式(12)，主要功用是微小雜訊的去除，同時物件受到侵蝕會比原始物件來的更小。

$$A \oplus B \square \{w : B_w \cap A \neq \emptyset\} \quad (11)$$

$$A \ominus B \square \{w : B_w \subset A\} \quad (12)$$

其中 $\oplus$ 代表為膨脹運算符號， $\ominus$ 為侵蝕運算符號， $w$ 為影像 $A$ 上的座標位置， $B$ 則為運算的結構元素而關閉運算是利用先做完侵蝕再做膨脹組合而成的運算，其數學式定義如式(13)：

$$A \bullet B \square (A \oplus B) \ominus B \quad (13)$$

### 3.3.4 後處理

在經過一連串的去掉動態背景的动作後，所保留下來的可能物件點，我們認為在經過所提出的方法淘汰後還能留下來的物件點，為真正物件像素點的可靠性相當高，但是本論文所提出的方法也帶來負面的效應，也就是物件破碎的問題，因此我們需要一個後處理的方式來重新找回真實物件所在的位置，盡可能標示出完整的真實物件所在區域，我們利用Bounding Box的方式來表示移動物件的區域範圍。我們使用的後處理判斷方式有三種，分別為顏色、距離及像素點數量，如下所示：

1. 顏色：我們利用HSV色彩模型當中色調的成分，也就是影像畫面中像素點的H位元，我們假設同一物件的像素點擁有相同色調，利用逐步掃描的前後兩點做色調判斷，若是色調差距在一定範圍內，則此兩點可能為同一物件點，進行其他兩種方式的判斷；反之若色調差距太大，則判斷此兩點為不同物件像素點或雜訊點。

2. 距離：在同一物件的前提下，假設物件像素點彼此間距離會相當接近，因此利用像素點的位置，設定一個限制，例如：距離超過10點的距離即判斷該像素點不屬於同一物件；反之若距離夠接近則判定可能為同一物件。之後進一步做其他判斷。

3. 像素點數量：在數量這個辨別方式中，採用計數可能為同一移動物件的像素點數夠不夠多來取代，若夠多則判定為移動物件，但如果計數的數量太少時，則判斷計數像素點集合所形成的區域可能是受到背景干擾的雜訊點或點集合。

#### 4. 實驗結果與比較

為了測試本論文提出的方法成效，本章採用標準的 MPGE-4 視訊及自行利用攝影機拍攝之視訊做為測試影片。當中我們使用：Tree、Water、San等三個視訊呈現實驗結果，為攝影機固定並採用同一視訊大小 320×240 的規格，利用不同的環境測試視訊與其他常用偵測方法做比較，具體分為比較效果及速度兩類，效果比較是由呈現出來的實驗結果與原始畫面相對照，直接由人眼判斷移動物件位置做評斷；另外我們定義兩種偵測機率，分別為式(15)表示的偵測失敗機率(Detection Failure, DF)與式(16)表示的前景誤判機率(False Alarm, FA)，利用此兩項數據資料做為人眼判斷之外的效果依據。

$$\text{Detection Failure} = \frac{\text{foreground pixels not detected by algorithm}}{\text{total foreground pixels in ground truth}} \quad (15)$$

$$\text{False Alarm} = \frac{\text{pixels misclassified as foreground by algorithm}}{\text{total foreground pixels detected by algorithm}} \quad (16)$$

其中，ground truth是根據人眼判斷所認定的真實物件大小。偵測率DF是指有多少ground truth像素點沒被偵測演算法偵測出來的比例；偵測率FA則是指經過演算法偵測出來的前景像素點，當中有多少比例不符合真實物件的像素點位置。其中MVDV法是以bounding box位置當作偵測結果進行偵測率實驗，而非與一般建立背景方法以每一點比較評估的方式進行。速度比較方面，藉由執行不同偵測演算法所計算出來的平均每張畫面消耗的時間，做為個別偵測演算法執行速度的結果進行評比。

實驗結果中，與本文提出MVDV法進行比較的偵測方法具體如下：W4法[5,20]、未經過改良的單高斯模型及混合高斯模型，原始SGM法[5]、原始GMM法[5,6]、及經文獻[6]改良的GMM改良法。而我們編輯及測試程式方面，使用 Microsoft Visual Studio 2008 C++ 搭配 Open CV 函式庫，電腦系統為 Intel(R) CPU 2.01 GHz，記憶體為 1G Bytes。

##### 4.1 偵測方法效果比較

我們將本論文所提出的MVDV法，與其他偵測方法個別採用不同測試視訊實驗，比較各方法在不同環境下產生的不同情況進行偵測效果比較。

視訊Tree為視訊畫面大小 320×240 的標準影像，以每秒25張圖切割完成，主要測試去除後方動態背景樹叢的影響以及物件由遠至近的移動情況做移動物件偵測的實驗。圖9為Tree視訊第52張畫面(frame52)，往攝影機方向移動的遠處物件是否會受到後方動態背景的干擾。圖10、圖11為視訊Tree的DF及FA偵測率比較圖，以每五張畫面取樣計算。

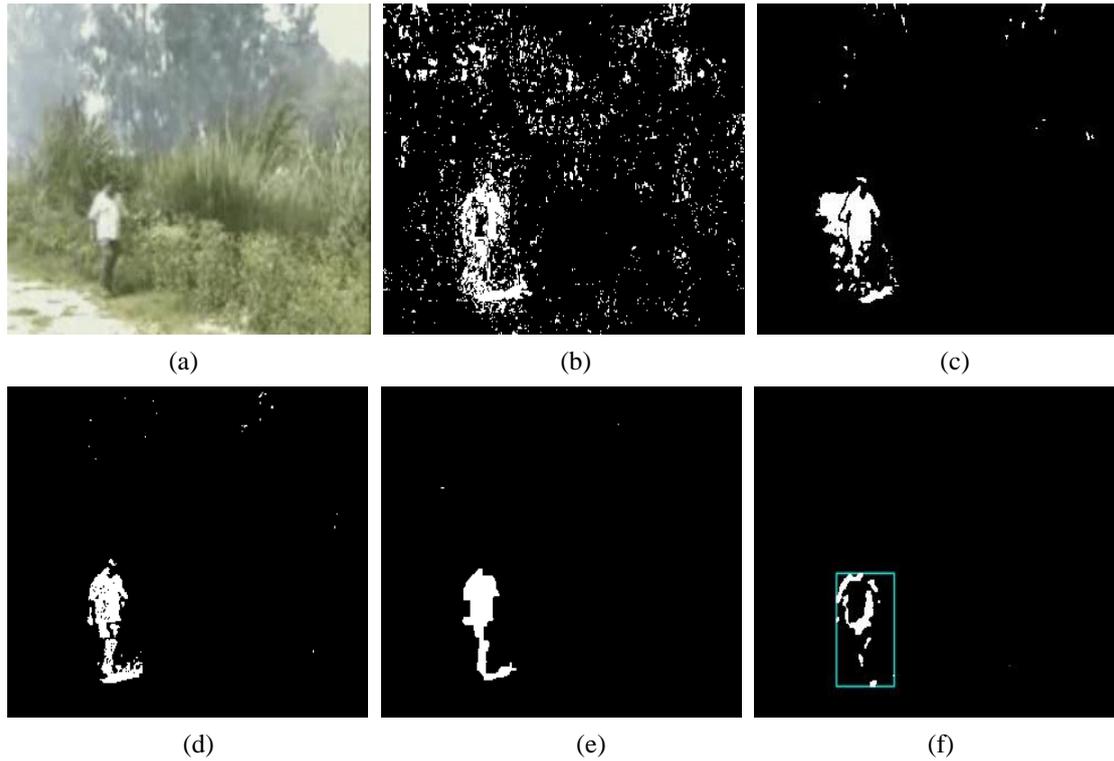


圖9 視訊Tree的第52張畫面：(a)原始視訊。(b)W4法。(c)原始SGM法。(d)原始GMM法。(e)GMM改良法。(f)MVDV法。

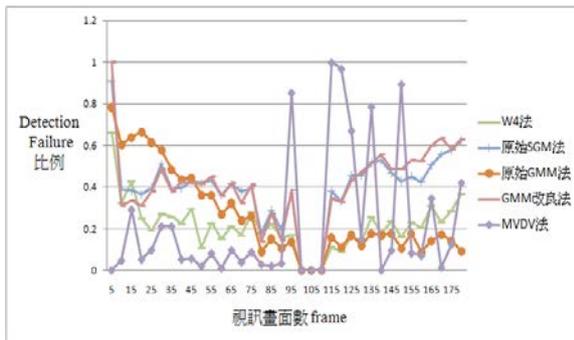


圖10 視訊Tree的DF偵測率比較圖

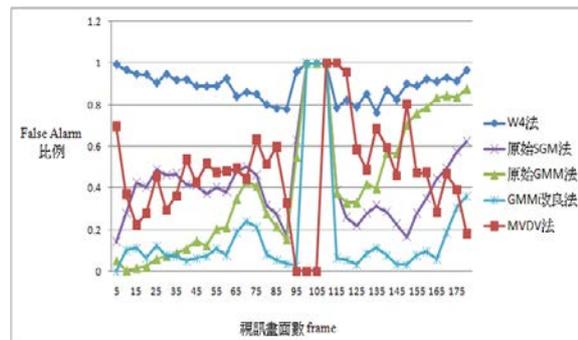


圖11 視訊Tree的FA偵測率比較圖

視訊Water為視訊畫面大小  $320 \times 240$  的標準影像，以每秒25張圖切割，畫面內容會有三個人各別獨自走過後方有一座噴水池的情形，主要測試這三個人經過噴水池的移動過程中是否會受到噴水池所噴出的水花影響到移動物件偵測。圖12為Water視訊frame113，第二個移動物件即將移出畫面時。圖13、圖14為視訊Water的DF及FA偵測率比較圖，以每五張畫面取樣計算。

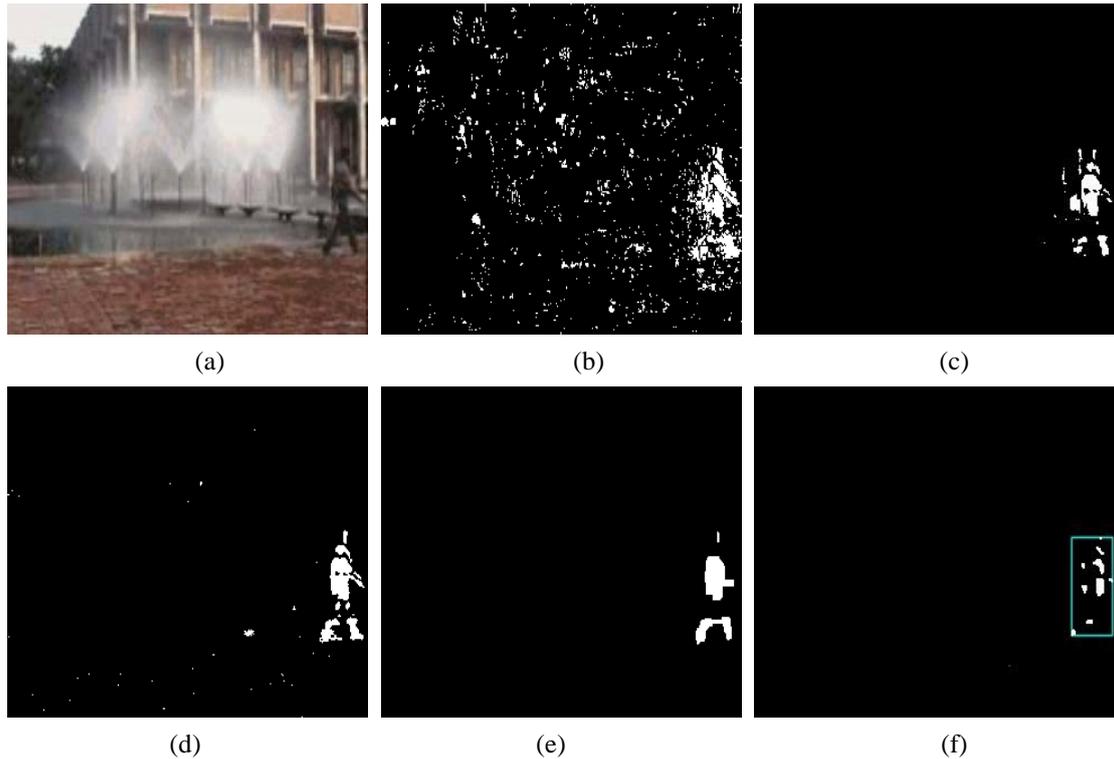


圖12 視訊Water的第113張畫面：(a)原始視訊。(b)W4法。(c)原始SGM法。(d)原始GMM法。(e)GMM改良法。(f)MVDV法。

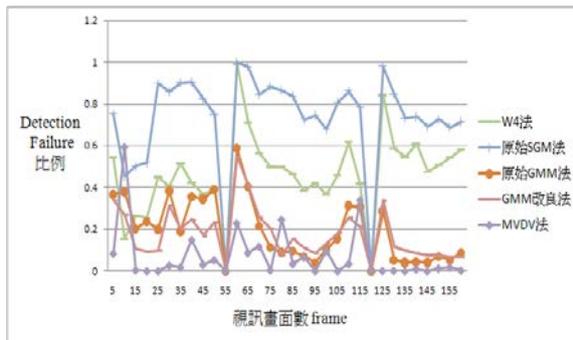


圖13 視訊Water的DF偵測率比較圖

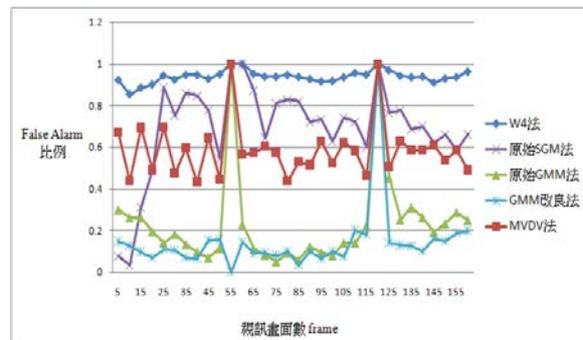


圖14 視訊Water的FA偵測率比較圖

視訊San為視訊畫面大小  $320 \times 240$  的標準影像，以每秒25張圖切割，畫面內容為道路路口固定攝影角度拍攝的擷取畫面，主要變化在於後方種植的樹木受風影響而擺動，加上戶外環境的光線變化，另外在移動物件上從單物件到同時呈現多物件，並且移動物件會產生遮蔽、重疊的情況。圖15為San視訊frame122，畫面中除非光線外，主要有兩個相似的移動車輛重疊造成判斷上的困難。圖16、圖17為視訊San的DF及FA偵測率比較圖，以每五張畫面取樣計算。

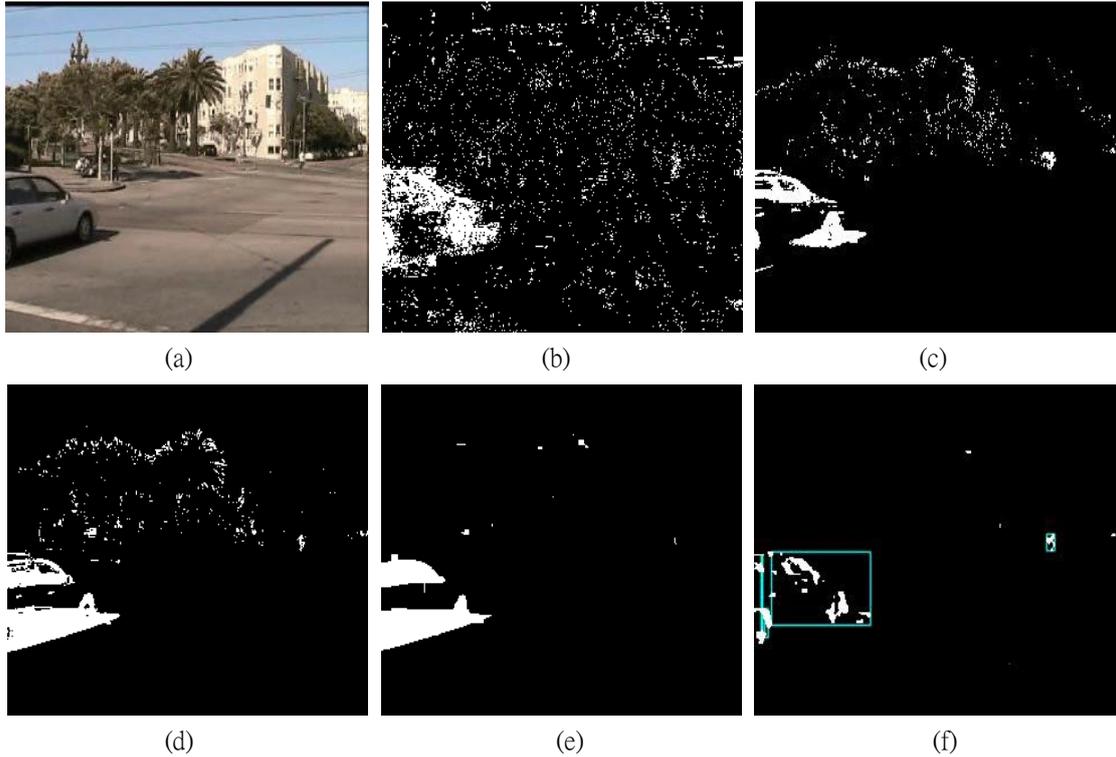


圖15 視訊San的第122張畫面：(a)原始視訊。(b)W4法。(c)原始SGM法。(d)原始GMM法。(e)GMM改良法。(f)MVDV法。

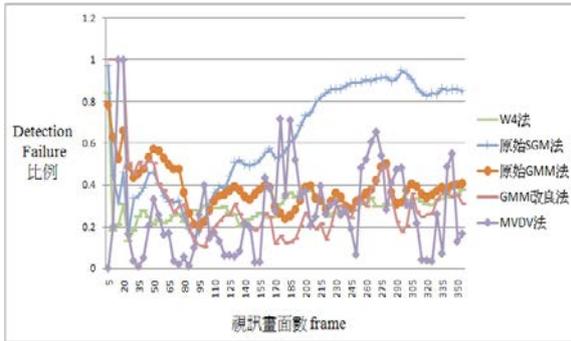


圖16 視訊San的DF偵測率比較圖

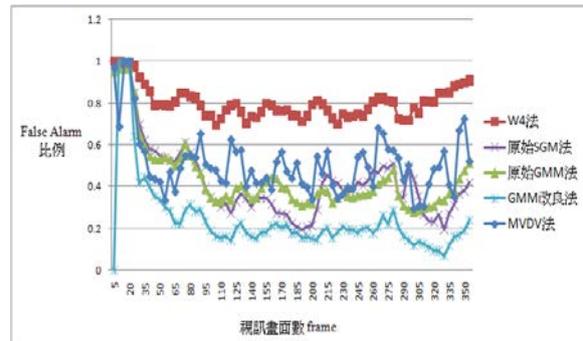


圖17 視訊San的FA偵測率比較圖

#### 4.2 偵測方法速度比較

我們進行各種偵測法的消耗時間測試，以每張畫面平均消耗秒數為單位(second per frame, s/f)來具體評估演算法的執行速度。表3是利用MATLAB工具撰寫測試，根據文獻[15,16]的原始及改良方法所做的相關偵測演算法消耗時間測試。

表3 每張畫面平均消耗時間表(1)

視訊名稱 \ 方法	W4(s/f)	原始 SGM 法(s/f)	原始 GMM 法(s/f)	改良 GMM 法(s/f)
Tree	0.088	0.35	5.7	1.09
Water	0.082	0.33	5.8	1.075
San	0.099	0.3	5.7	1.38

表4為常見的建立背景方式與本論文所提出減少特徵點方式的偵測法做比較，其中以原始SGM法代表建立背景的偵測方法和我們所提出的MVDV法進行消耗時間比較，比較形式採用執行速度較快的程式語言C++撰寫執行。從實驗統計數據來看，表4也證實原始SGM法的平均時間較為固定，而MVDV法則會依據不同視訊的特徵點多寡影響平均時間。總結來說建立背景的偵測方法消耗時間會大於MVDV法只計算特徵點的時間，因此推論MVDV法相較於需要建立整張背景的方式，運行速度較快。

表4 每張畫面平均消耗時間表(2)

視訊名稱 \ 方法	原始 SGM 法	MVDV 法
Tree	0.132231 s/f	0.072806 s/f
Water	0.131783 s/f	0.075283 s/f
San	0.132460 s/f	0.075838 s/f

## 5. 結論

從實驗結果可以發現本論文提出的去除動態背景方法，採用移動向量當作特徵的可行性是相當高的。在移動物件偵測上，傳統使用色彩等特徵會受光線等因素影響，易造成動態背景的誤判，而使用移動向量為特徵的結果，很成功去除了這項憂慮。在演算過程中捨棄了相關移動物件的資訊，因此造成移動物件呈現不完整的形狀，但經過本論文所提出的演算法計算，最後留下的像素點為移動物件點是可以確認的。其後，利用我們提出的後處理方式能有效框出移動物件的所在位置，找到物件的移動區域。在偵測法執行速度比較上，實驗數據顯示本論文提出的方法相較其他非採用移動向量為特徵的偵測方法，每張畫面的消耗時間來得低，並且根據不同的視訊的特徵點數，節省計算時間，執行速度較快而有效率。整體而言，我們的系統在有動態背景的干擾下，進行移動物件偵測，是相當有效的。另外在靜態環境下（由於篇幅所限，本文未展示結果），本論文所提出的系統也有不錯的偵測效果。

## 6. 參考文獻

1. C.E. Smith, C.A. Richards, S.A. Brandt, and N.P. papanikolopoulos, "Visual tracking for intelligent vehicle-highway systems", *IEEE Trans. Vehicular Tech.*, vol. 45, issue.4, pp. 744-759, Nov. 1996.
2. C. Kim and J. N. Hwang, "Fast and Automatic Video Object Segmentation and Tracking for Content-Based Applications", *IEEE Transactions Circuits And Systems for Video Technology*, vol. 12, pp. 122-129, Feb, 2002.

3. Xiaopeng Ji, Zhiqiang Wei, and Yewei Feng, “Effective vehicle detection technique for traffic surveillance systems”, *Journal of Visual Communication and Image Representation*, Volume 17, Issue 3, pp. 647-658, June 2006.
4. 謝宗良,2009, “整合多重線索的視訊追蹤法(Visual Tracking Using Multiple Cues),” 碩士論文,國立雲林科技大學電子工程系.
5. Jacinto C. Nascimento and Jorge S. Marques. “Performance Evaluation of Object Detection Algorithms for Video Surveillance”, *IEEE Trans. On multimedia*, vol. 8, no. 4, pp. 761–774, 2006.
6. 林榮澤,2012, “視訊物件偵測的一些改善做法(Some Improvements to Video Object Detection),” 碩士論文,國立雲林科技大學電子與光電工程研究所碩士班.
7. 景雅新,2003, “光流技術在移動物體影像追尋上之應用(Visual Tracking for a Moving Object Using Optical Flow Technique),” 碩士論文,國立中山大學機械與機電工程學系.
8. David R. Gilland and Bernard A. Mair, “Motion Estimation in Gated Cardiac Emission Tomography by Optical Flow Techniques”, *IEEE Nuclear Science Symposium Conference Recor*, pp. 2699-2702, 2006.
9. B.K.P. Horn and B.G. Schunck, “Determining Optical Flow”, *Artificial Intelligence*, 17, pp. 185-204, 1981.
10. B.D. Lucas and T. Kanade, “An Iterative Image Registration Technique with an Application to Stereo Vision”, *DARPA Image Understanding Workshop*, pp. 121-130, 1981.(see also *IJCAI’81*, pp. 674-679)
11. Simon Baker and Iain Matthews, *Lucas-Kanade 20Years On : A Unifying Framework*, *International Journal of Computer Vision*, Vol 56, pp. 221-255, 2004.
12. 陳柏儒,2011, “人臉朝向偵測應用於平衡復健與人機介面(Detection of Face Orientation and its Applications in Vestibular Rehabilitation and Human-computer Interface),” 碩士論文,國立中央大學資訊工程學系研究所.
13. 侯凱鐘,2005, “適用於H.264可變區塊大小之模式決策與移動估計演算法(Variable Block-Size Mode Decision and Motion Estimation Algorithms in H.264),” 碩士論文,國立東華大學電機工程研究所.
14. 吳崇賓,2004, “低複雜度視訊後處理器與移動估計演算法(Low Complexity Postprocessing and Motion Estimation Algorithms in Video Coders),” 碩士論文,國立成功大學電機工程學系碩博士班.
15. 林炯俊,2007, “適用於H.264之整數點移動估計法則之研究(Fast integer pixel motion estimation in H.264),” 碩士論文,國立中央大學通訊工程研究所.
16. 賴彥任,2005, “應用於快速移動估計之高效能演算法及其硬體實現(Enhanced Algorithm and Hardware Design for Fast Motion Estimation in Video Coding),” 碩士論文,國立臺灣海洋大學電機工程學系.
17. P. Anandan, “A Computational Framework and an Algorithm for Visual Motion”, *International Journal of Computer Vision*, Vol. 2, pp. 283-310, 1989.
18. A. Singh, *Optical Flow Computation : A Unified Perspective*, IEEE computer Society Press, 1992.
19. Wei Zhang, Q.M. Jonathan Wu, and Hai bing Yin “Moving vehicles detection based on adaptive motion histogram”, *ELSEVIER, Digital Signal Processing*, Volume 20, Issue 3, pp. 793-805, May 2010.
20. Ismail Haritaoglu, David Harwood and Larry S. Davis, “W4: Real-Time Surveillance of People and Their Activities”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, No.8, pp. 809-830, Aug 2000.

# **Optical Flow Assisted Video Object Detection in Dynamic Background**

Bo-Jing Ji   Hsiao-Jeh Lin\*   Shih-Fong Dai   L. Lan

Graduate School of Electronic and Optoelectronic Engineering  
National Yunlin University of Science & Technology

\*m10013237@yuntech.edu.tw

## **Abstract**

Moving object detection is currently a very active research area in computer vision. Most detecting methods operate in static environments. Some further processing tasks are usually needed when dynamic background is encountered. This paper presents an optical flow assisted video object detection scheme that can work effectively in dynamic background. The proposed approach is composed of three key steps: (1) It adopts a preprocessing screening step to screen out possible foreground pixels; (2) It utilizes motion vectors generated by the optical flow technique as well as a voting mechanism to identify object pixels; (3) It employs another post-processing step to locate moving objects' positions and draw their bounding boxes. A variety of experiments were conducted to demonstrate the effectiveness of the proposed method. The results indicate the particular usefulness of the presented approach under dynamic background.

**Keywords** : moving object detection; motion vector; removal of dynamic background