

國立交通大學

電機學院 IC 設計產業研發碩士班

碩 士 論 文

應用於動態背景中的移動物體影像之
偵測與即時追蹤系統

A Real-Time Motion Detection and Tracking System
in the Dynamic Background

研 究 生：鍾宜岑

指導教授：林進燈 教授

中 華 民 國 九 十 六 年 十 月

應用於動態背景中的移動物體影像之偵測與即時追蹤系統
A Real-Time Motion Detection and Tracking System
in the Dynamic Background

研 究 生：鍾宜岑

Student：Yi-Cen Chung

指導教授：林進燈

Advisor：Dr. Chin-Teng Lin

國立交通大學
電機學院 IC 設計產業研發碩士班
碩 士 論 文



Submitted to College of Electrical and Computer Engineering
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master
in

Industrial Technology R & D Master Program on
IC Design

October 2007

Hsinchu, Taiwan, Republic of China

中華民國九十六年十月

應用於動態背景中的移動物體影像之偵測與即時追蹤系統

學生：鍾宜岑

指導教授：林進燈 教授

國立交通大學電機學院 IC 設計產業研發碩士班

摘 要

隨著時代的遷移，人身財物的安全與保全需求逐漸受到大眾的重視，且基於人力資源越來越珍貴的趨勢，自動化即時監視系統(real-time surveillance system)的研究，成為重要的項目之一，然而目前監視系統所能提供的功能有限，原因在於使用的攝影機是固定式的攝影機，會使得監視範圍產生死角；若攝影機在移動的情形之下，則會使得背景與移動物體同處於移動狀態，增加移動物體偵測的困難度；我們希望除了應用在室內的安全監視外，更希望能在複雜的動態背景之下，可以分辨出移動物體，於是提出在室外場景上偵測移動物體，並且進一步地自動追蹤之系統。

在此論文中，我們建立一個基於電腦視覺的移動物體偵測及自動追蹤系統，在移動物體偵測部分，主要是利用連續影像相減法(temporal frame difference)執行快速的優勢來找出移動的畫素部分，因為背景為動態，我們所得到的移動畫素也會有背景的存在，於是在這之前需加入了全域補償，降低背景的影響，再以區塊為基礎(block-based)，計算每個區塊的移動標準差與移動最大值，統計此 2 種數值，我們得到移動物體所佔據的區塊，便能藉此擷取出移動物體的樣版資訊，進而追蹤之。

自動追蹤的系統部分，則是利用平均位移演算法(mean shift)，來計算出影像中與移動物體的樣版相似度最高的影像質心位置；其原理是利用樣版中的色彩分佈與候選影像區塊的色彩分佈，找出相似系數(Bhattacharyya coefficient)最高的影像質心位置，尋找的方法則是利用平均位移疊代(mean shift iteration)，它可以準確且快速的計算出與樣版相似度最高的影像質心位置，達到追蹤移動物體的目的。

藉由以上提出的系統，我們可以在複雜的動態背景之下，做到移動物體的偵測與自動追蹤，讓監視系統可執行環境多了更大的範圍與彈性。

A Real-Time Motion Detection and Tracking System In The Dynamic Background

student : Yi-Cen Chung

Advisors : Dr. Chin-Teng Lin

Industrial Technology R & D Master Program of
Electrical and Computer Engineering College
National Chiao Tung University

ABSTRACT

Safety of our life and asset is becoming more and more important as the times go by. Besides, human resource is on the trend of high cost, thus study of real-time surveillance system becomes one of the important issues.

Due to fixed camera, modern surveillance system can only provide limited function. Under fixed camera model, monitor view will be limited and dead spot can not be observed. But with active camera, background will not stay in static situation, moving object and dynamic background will increase difficulty for moving object detection. Except for indoor surveillance system, we still hope we can detect object under complicated dynamic background, thus we proposed detection of moving object on outdoor plaza and advanced automatic tracking system.

In this thesis, we build up a computer vision based moving object detection and automatic tracking system. We mainly used temporal frame difference technique finding out moving pixel for moving object detection. Temporal frame difference technique can quickly locate all moving pixel. For high computation required moving object detection system, it can improve computation efficiency. But under our dynamic background, our moving pixel will include background; therefore we need to apply global motion compensation first in order to reduce moving background effect. Then we calculate the standard deviation and the maximum of each block. After statistic, we can get moving object area. According to this moving object area, we can

extract moving object template information and track down.

We use mean shift algorithm for automatic tracking system to calculate highest similar image mass center. Our principle is comparing color distribution inside the template and candidate object to find out best Bhattacharyya coefficient location to determine image mass center. We use mean shift iteration which can accuracy and efficiently calculate out the most similar image mass center location and accomplish our purpose of tracking moving object.

We can track moving object under complicate dynamic background under the system we proposed, which will increase the scope of surveillance system and provide more flexibility.



誌謝

首先，我想感謝我的指導老師林進燈教授，在我的研究學習過程中，給了我很多寶貴的意見與悉心地指導，非常感謝老師付出的心力與精神，另外也要感謝張志永教授與陳永平教授能在百忙之中前來擔任我的口試委員，並給予許多珍貴的意見，讓本論文更為完善，在此深深致謝。

其次，感謝學長劉得正博士對我的研究給予指導與幫助，每當遇到瓶頸時，學長總是給我很大的支持，使我能堅持到底。還有實驗室的學長姊、Linda、訓緯、育弘、立倬、亞書和學弟妹們，在研究的過程中給我鼓勵和支持，讓我在研究所的日子中獲益良多。

最後我要感謝我的家人，他們在這段時間內不曾間斷的鼓勵和關懷，使我能夠心無旁騖的進行研究。在此僅以本論文獻給我的家人與其它關心、幫助過我的師長及朋友，非常感謝你們。

目 錄

中文提要		i
英文提要		ii
誌謝		iv
目錄		v
表目錄		vi
圖目錄		vii
第一章	緒論.....	1
1.1	研究動機.....	1
1.2	相關研究.....	2
1.3	論文架構.....	2
第二章	移動物體偵測與自動追蹤系統基礎理論摘要.....	3
2.1	移動物體偵測.....	4
2.2	自動追蹤系統.....	6
2.3	系統程式架構.....	9
第三章	動態背景中移動物體之偵測系統.....	12
3.1	影像前處理.....	13
3.1.1	彩色對灰階的轉換.....	13
3.1.2	平滑化處理.....	14
3.2	移動向量之搜尋方式.....	16
3.3	動態背景之全景補償.....	18
3.4	移動物體之移動統計計算量.....	19
3.5	移動物體之標記.....	19
第四章	移動物體之自動追蹤系統.....	29
4.1	移動物體之色彩資訊分析.....	22
4.1.1	色彩分佈密度函數之定義.....	22
4.1.2	核心函數.....	23
4.2	移動物體之候選位置相似係數.....	26
4.3	應用平均位移法於移動物體之追蹤.....	28
4.3.1	平均位移法理論背景介紹.....	28
4.3.2	平均位移法求取最高的相關係數.....	30
4.4	平均位移法的實作流程.....	33
第五章	實驗結果.....	35
第六章	結論與未來工作.....	41
參考文獻		43

表 目 錄

表 5.1	移動物體偵測與追蹤的執行時間.....	38
-------	---------------------	----



圖 目 錄

圖 2.1	雙分圖.....	7
圖 2.2	階層式人形模型.....	9
圖 2.3	系統流程圖.....	11
圖 3.1	移動物體偵測系統流程圖.....	12
圖 3.2	由(a)彩色模式轉換成(b)灰階模式.....	14
圖 3.3	經均值平滑濾波器的影像.....	16
圖 3.4	區塊匹配法.....	17
圖 3.5	區域移動向量.....	19
圖 3.6	移動邊緣圖.....	19
圖 3.7	移動區塊圖.....	20
圖 4.1	自動追蹤系統流程圖.....	21
圖 4.2	影像像素資訊的重要性示意.....	23
圖 4.3	Uniform Kernel Function.....	24
圖 4.4	Normal Kernel Function.....	25
圖 4.5	Epanechnikov Kernel Function	26
圖 4.6	平均位移法示意圖.....	29
圖 5.1	系統程式輸出畫面.....	35
圖 5.2	系統程式輸出結果.....	36
圖 5.3	與樣版影像相似度太低時的結果.....	36
圖 5.4	部分遮蔽物時的結果.....	38
圖 5.5	連續影像實驗輸出結果.....	40

第一章

緒論

隨著時代的遷移，人身財物的安全與保全需求逐漸受到大眾的重視，且基於人力資源越來越珍貴的趨勢，自動化即時監視系統(real-time surveillance system)的研究，成為重要的項目之一。過去電腦視覺的設備需要仰賴高速的計算量，使得移動物體的偵測與追蹤受到很大的局限，近年來，由於半導體製程的生產成本降低與技術上的進步，電腦設備的運算能力大大地提升，使得電腦能快速處理影像資訊且所用到的硬體設備成本不高，漸漸讓電腦視覺的領域有長足的進展，讓即時性移動物體的偵測與追蹤系統可以在電腦上實現出來。

1.1 研究動機



目前的監視系統的功能與使用環境是有限的，所使用的攝影機是傳統固定式的定點監視模式，會使可觀測的範圍受到局限，而產生死角的區域；架設在室外場景的情形之下，複雜的室外場景增加雜訊的可能性，會使得移動物體的偵測困難度提升，文獻[32]使用建立全景背景作為參考背景，之後的視訊影像與參考背景比較，即可得知移動物體的區域，但此方法需要事先有全景背景的資訊，且背景若改變需再一次建立全景背景，否則會造成誤判的情形，對於非固定背景的監視環境下並不適用。在這篇論文中，便要針對此種情形提出動態背景之下室外場景的移動物體偵測與自動追蹤，以增加監視系統的可應用環境的彈性度。

在此篇論文中，可在複雜的室外背景中，且不完全得到背景資訊的情形之下，利用每個影像區塊(block-based)的移動資訊，偵測到移動物體的區域，再根據偵測到的區域為樣版，以平均位移(mean shift)演算法來快速且準確地執行追蹤的動作，這個方法的好處是所偵測的移動物體的形狀可以是變動的(deformable)，且可以在複雜的背景中分辨出移動物體進而追蹤之。

藉由此移動物體偵測及自動追蹤系統，可以改進傳統固定式的定點監視的缺點，使得可監視的範圍大大的提升，且不需先建立起背景環境的資料，可用非固定的背景環境之下，大大增加監視系統的可應用環境的。

1.2 相關研究

移動物體的偵測與追蹤系統一直是電腦視覺重要的研究課題，相關的應用也不斷地推出，主要分為兩大部分：(一)移動物體的偵測部分及(二)動態追蹤部分。在移動物體的偵測部分的方法主要有以下幾種，光流法(Optical Flow)、連續影像相減法(Temporal Differencing)、背景相減法(Background Subtraction)；而在追蹤部分的方法，主要有以下幾種，區域為基礎的追蹤(region-based tracking)、主動輪廓為基礎的追蹤(active-contour-based tracking)、特徵點的追蹤(feature-based tracking)、以模型為基礎的追蹤(model-based tracking)，各個方法的詳細部分將在第二章再做仔細的說明。



1.3 論文架構

此論文主要分為六章，第一章為緒論，主要是說明研究動機與介紹相關研究。第二章則是摘要移動物體偵測與自動追蹤系統基礎理論，介紹相關方法，並討論其優劣，其後介紹此論文的系統架構及完整流程。第三章則是詳述此系統中移動物體偵測的方法，藉由移動物體偵測所引發的問題及研究現況，推演出此篇論文的移動物體偵測的演算法，繼而細述每個流程所運用到的方法，以得到完整的移動物體區域，當成之後追蹤的樣版。第四章是說明追蹤的系統，在此之前，先介紹動態追蹤所引發的問題及研究現況，再來說明主要的理論基礎，與實作後的成效。第五章為實驗結果，在此章節將對整體移動物體偵測及追蹤系統做一個完整的說明，在實際運作後對效能進行評估的動作，藉由顯示的結果，說明此系統的能力。第六章是包含此論文中的研究心得，在此檢討系統中不夠完善的地方，以供未來可能研究的方向。

第二章

移動物體偵測與自動追蹤系統基礎理論摘要

在監視系統中，主要分為兩大部分：一是移動物體偵測系統部分，二為自動追蹤系統部分。移動物體偵測是為了偵測出畫面中是否有移動的物體，針對執行的環境分為兩大種類，一為固定式攝影鏡頭之下的靜態背景，一為非固定式攝影鏡頭之下的動態背景。在固定式攝影鏡頭之下是較簡易的移動物體偵測環境，靜態背景可透過時間域上連續的影像相減而移除，建立一個影像差異圖(difference image)，再定義一個臨界值(threshold value)過濾掉因擷取影像畫面所產生的雜訊，由此方法可有效率地得到移動物體；但對於非固定式攝影鏡頭之下的動態背景卻無法直接使用上述方法，因為非固定式攝影鏡頭是可移動的，其背景是動態的，除了移動物體本身的移動量(object motion)外，尚有因鏡頭的移動而產生的全域移動量(global motion)，若直接使用連續影像相減所得到的影像差異圖之中，除了得到移動的物體之外，還會包含移動的背景，換句話說，若我們可以將影像中的全域移動量移除，便可將固定式攝影鏡頭之下的移動物體偵測方法，應用於動態背景的移動物體偵測之中，因此非固定式攝影鏡頭之下的移動物體偵測系統中，首要之務便是需從視訊畫面中擷取出全域移動量，得到全域移動量後便可對影像補償，使背景類似於靜態背景，便可利用靜態背景的移動物體偵測方法將移動的物體與背景分離出來。

自動追蹤系統的目的是在對特定的移動目標物，在視訊影像進行的同時，持續地跟隨此目標物。追蹤移動物體的方法基礎的想法是在多個候選影像中，選出與被追蹤的目標影像相似度最高的候選影像，相似度最高的候選影像則可判斷與目標物是吻合的結果，因此我們需要被追蹤的移動物體所含有的相關資訊，以及相似度的定義與計算的方式，用以選出相似度最高的候選影像，並在視訊影像中持續地追蹤移動物體。在追蹤系統中廣泛用到的移動物體相關資訊諸如色彩、紋理、移動速度、邊緣等，在目標物的形狀或相關資訊未改變的情況之下，此時追蹤的條件是較容易處理的，若目標物產生形變

(deformed)的情形之下，其原有的資訊可能有所改變，進而導致追蹤錯誤，因此在追蹤系統中，需對被追蹤之目標物的資訊作動態的更新，才能在物體形狀改變後仍然持續有效地追蹤之。

2.1. 移動物體偵測

移動物體偵測的部分一直是電腦視覺中十分重要的一環。我們可以觀察到，在場景中可察覺的移動量，會對視訊影像造成變化量，根據以上的觀察結果，我們可以說，在視訊影像中有變動的部分，通常是移動物體所造成的，而如何找出這些變動的部分，便是移動物體偵測主要的用途。以下介紹幾種主要的移動物體偵測方法，並討論之。

I. 光流法(Optical Flow)：當物體在移動時，在影像中對應物體的亮度也會產生變化，光流則是指影像亮度的表面變化(apparent motion)分佈情形[1]-[3]。因需給影像中每一個像素速度向量(motion velocity)，其運算量很高，且對雜訊非常敏感，若是用於視訊影像無法執行即時性的應用，需要特殊的硬體設備。

II. 連續影像相減法(Temporal Differencing)：在連續的影像中，讓相隔兩、三張的影像相減，若相差為零，則表示此像素不屬於移動物體，若不為零，則為移動物體的像素，利用相減後所得的結果來辨別移動的區域[4]，與光流法相較之下，連續影像相減法在運算上有大幅降低的情形，在實作在即時性的應用上，較為適用。但在動態背景之下，對於背景移動與實際移動物體無分辨的能力，且對剛性物體造成內部破碎，對移動物體偵測無法提供完整的資訊。

III. 背景相減法(Background Subtraction)：背景相減法是種易於實現且廣受應用的方法，特別是在靜態背景的场景中；首先是建立好參考背景，再用現在的影像與參考背景影像相減，以去除背景相同的部分，留下來的即為移動物體的區域[5][6]。此種方法可以完整且清楚地辨別出移動物體的部分，但它受光影響很大，光線的變化會導致相減後的結果不精確，因此良好的參考背景影像是很重要的，另一方面它也不適用於動態背景，因為背景的資訊隨時在更動的情形之下，無法找到特定的背

景作為參考影像。

C. Kim[5]中運用了連續影像相減法和背景相減法，利用邊緣的資訊建立兩張邊緣圖 (Edge Map)，一個為參考背景邊緣圖，另一個則為由連續兩張的視訊影像相減而得的移動邊緣圖，再將移動邊緣圖與參考背景邊緣圖相減，便可求得移動物體的區域。這是運算速度快的求得移動物體的方法，但參考背景若不在光源固定的地方，參考背景便不再是固定了，需隨著時間對參考背景做更新，在 A. H. S. Lai[7]中為了得到穩定的參考背景，紀錄現在的影像與前個參考背景中每個像素的灰度值的變化值，根據不同的變化值來估計現在的參考背景。

A. J. Lipton[4]使用連續影像相減法得到移動物體的區域；假設 I_n 為第 n 個影像的灰度值，連續影像相減法的移動量估計 Δ_n ：

$$\Delta_n = \|I_n - I_{n-1}\| \quad (2.1)$$

連續影像相減法可以偵測到移動物體的微量變化，當移動量太小，可以累積相加數個連續影像的移動量，以得到移動物體的區域；相對地，若累積相加的影像數量過多，會造成同個目標物判別為不同的移動區域，因此應用上需對累積相加的影像數量多加考量。

B. K. P. Horn[1]假設在連續的影像中某一像素的投影亮度是為固定的，即亮度守恆 (conservation of image intensity)[8]如下：

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t) \quad (2.2)$$

I 代表影像像素的灰度值， (x, y) 為像素的位置， t 代表時間，對式(2.1)做泰勒展開到一階，可得：

$$I(x + \delta x, y + \delta y, t + \delta t) = I(x, y, t) + \delta x \frac{\partial I}{\partial x} + \delta y \frac{\partial I}{\partial y} + \delta t \frac{\partial I}{\partial t} \quad (2.3)$$

將式(2.2)代入式(2.3)，，推導出影像中的像素在水平與垂直方向位移的一階線性方程式，即光流方程式如下：

$$\delta x \frac{\partial I}{\partial x} + \delta y \frac{\partial I}{\partial y} + \delta t \frac{\partial I}{\partial t} = 0 \quad (2.4)$$

式(2.3)的光流方程式，再加上光流變化的平滑限制，透過疊代法可求出像素位移量；但

光流法需對整張影像做微分，除了計算量高外，也容易造成誤差過大。C. J. Li [16]中以區塊匹配法(block matching)偵測區塊周圍的鄰近區塊中是否具有相似區塊，即區域亮度分佈守恆(conservation of local intensity distribution)[8]，給定第一張影像的像素，在第二張影像中相對應的像素的區塊，在給定的搜尋視窗(search window)中找出相關度最高的區塊，其移動量即為此區塊的位移向量，用此方式可以大大降低光流法計算量高的缺點。

本論文是的研究重點是在動態背景之下偵測移動物體，背景是隨時間而改變的狀態下，是無法建立屬於靜態的參考背景，以背景相減法得到移動物體；由於在動態背景之下整個影像畫面皆是處於移動的情形，需考慮如何將背景與前景(移動物體的區域)分離；我們可以觀察到背景與前景的不同之處，在於背景所含有的位移向量是相同的，而前景的位移向量則為不固定，且在一個影像中背景佔大多數的面積，我們可以說，數量最多的移動向量便是屬於背景，利用此特性，我們使用區塊匹配法找出每個區塊的位移向量，找出數量最多的移動向量為全域移動向量(global motion vector)，用以補償移動後的背景，便可得到幾乎靜態的背景，再搭配連續影像相減法，以求出可能的移動物體的區域，其方法的細節會在第三章中再詳細的介紹。



2.2. 自動追蹤系統

在監視系統中另一個重要部分，即為自動追蹤系統，追蹤的目的是利用移動物體偵測到的目標物體，在連續的視訊影像中持續地找到目標物移動後的位置，因此我們需要尋找與目標物吻合的候選影像，藉由目標物所含有的相關資訊量，計算觀察的候選影像與參考物體之間的相似度是否夠高，相似度最高即可判別與目標物吻合，以下介紹幾種追蹤的方法。

- I. 區域為基礎的追蹤(region-based tracking)：區域追蹤主要是用影像中有變動的區域為基礎，得到此區域的方法，例如用上一節所提到的背景相減法、連續影像相減法得到移動物體的區域，再根據此區域內部參數來進行追蹤[9]。
- II. 主動輪廓為基礎的追蹤(active-contour-based tracking)：輪廓追蹤是將移動物體

用輪廓線(contour)來表示，如影像邊緣、形狀等，再利用之後影像的輪廓線更新資訊來追蹤物體[11]，在[12]將之應用於車輛的追蹤上，可得到良好的追蹤效果。

III. 特徵點的追蹤(feature-based tracking)：特徵點追蹤是先分析影像中物體的成分特徵，將其群集到較高階的特徵，再比對之後的影像的特徵資訊，追蹤目標物。利用到的特徵可能有質量中心、面積、色彩[13]等。

IV. 模型為基礎的追蹤(model-based tracking)：模型追蹤首先會建立起物體的模型結構，再利用此模型進行比對追蹤，它需要較多的模型參數，才能精確地取得移動物體[15]。

S. Gupte[9]藉由背景相減法得到連續影像*i*與影像*i+1*的移動物體區域，利用空間匹配的方法(spatial matching method)追蹤移動物體，因此需建立影像*i*與影像*i+1*的移動物體區域之間的關係圖(association graph)，在此利用雙分圖(bipartite graph)(圖 2.1)，圖中每一個點代表一個區域，*P*分部(partition)內的點代表前一張連續影像內的區域，而*C*分部的點代表下一張連續影像內的區域，分部間的邊 E_{ij} (edge)連結起 P_i 與 C_j 間的關係，給予每一個邊 E_{ij} 一個權重 w ：

$$w(E_{ij}) = A(P_i \cap C_j) \quad (2.5)$$

E_{ij} 的權重即是 P_i 區域與 C_j 區域之間重疊的面積，

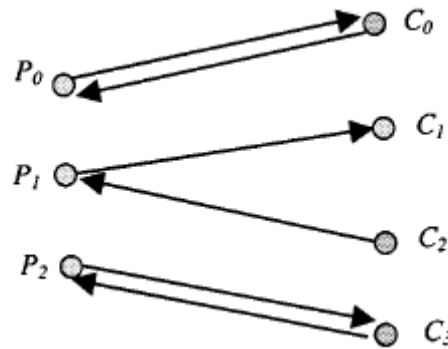


圖 2.1：雙分圖[9]

追蹤的問題即為尋求最大權重圖(maximal weight graph)；此方法需對整張影像做處理，以求得到每個時間點的移動物體區域，運算量取決於整張影像的大小，運算量較高，不易做到即時性追蹤。

M. Kass[10]提出了主動輪廓法(Active contour)，主要目的是在於找出物體區域的輪廓線，一開始以一初始輪廓朝物體區域做輪廓演化以將物體分割出來，輪廓演化是藉由最小化能量函數(energy function)來達成，當能量最小處所形成的輪廓即為所求，以此輪廓的位置為起始位置，尋找下一時間影像的目標物的輪廓線，以達到追蹤的效果；輪廓追蹤的缺點在於其效果很容易受到初始位置的影響，需要一個較準確的初始輪廓，且物體在高速運動時，能量函數無法明確地定義出物體的輪廓線。

A. Chachich[14]利用強度正規色彩長條圖(histograms of intensity normalized color)，使用強度正規色彩公式得到色彩向量(r, g)：

$$r = \frac{R}{R+G+B} \quad g = \frac{G}{R+G+B} \quad (2.6)$$

其中(R, G, B)為 RGB 色彩座標的三個成份，經過正規化後，可定義目標物的二維色彩長條圖 $h_{target}(r, g)$ ，第二個色彩長條圖則是由整個影像的像素而得之， $h_{total}(r, g)$ ，經由 Bayes rule 得到影像中給定一色彩向量，(r, g)，此色彩向量為目標物的機率：

$$p(target | r, g) \approx h_{ratio}(r, g) = \frac{h_{target}(r, g)}{h_{total}(r, g)} \quad (2.7)$$

將目標物在位置(i, j)的色彩像素($r(i, j), g(i, j)$)群集為一機率圖：

$$P_{target}(i, j) = p(target | r(i, j), g(i, j)) \quad (2.8)$$

根據式(2.8)得到的目標物區域來追蹤之。

I.A. Karaulova[15]使用階層式人形模型來追蹤移動的人形。在分析整個人形的架構後，在最高階層中包含整個人形的資訊，較低階層則包含更細部的資訊，將人形分為幾個分部描述可能的姿勢，當進行追蹤時，利用較低階層的分部資訊來增加準確性，利用此種追蹤模式，可以準確地追蹤到目標物，但所使用的模型參數繁多，因而增加了系統的複雜性，也相對提高運算量。

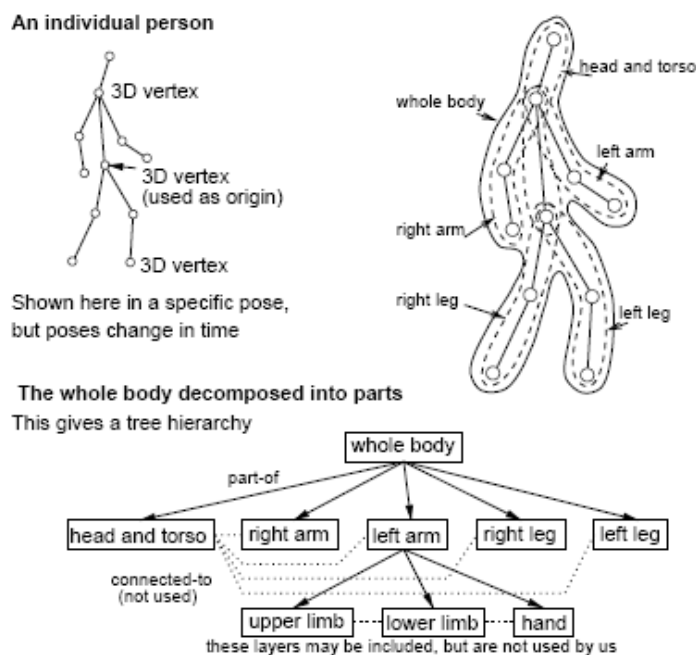


圖 2.2：階層式人形模型[15]

由移動物體偵測系統，我們得到了移動物體的區域，目的為有效持續地追蹤此區域內的物體，考慮到即時性的追蹤應用，我們無法對整張影像做空間匹配，也無法使用建立模型參數來追蹤之，這會使得運算量大增，因此運算量高的追蹤方法我們不與採用，為了能夠應用於可形變的物體，我們觀察到對於可形變的物體，在連續的影像中，其色彩的變化差異不大，因此我們採用以色彩為其特徵，使用移動物體的色彩資訊來追蹤之，在此利用平均位移(mean shift)演算法來進行移動物之追蹤，利用平均位移法運算快速計算出相似度最高的候選影像，進行追蹤的動作，其方法的細節及理論會在第四章中再詳細的介紹。

2.3. 系統程式架構

本論文是的研究重點是在動態背景之下偵測移動物體並追蹤之，論文的系統程式中，主要分為兩大部分：移動物體的偵測部分及自動追蹤部分。

在移動物體的偵測中，為了從動態的背景中取得移動物體，首先經過影像前處理，

再以區塊為基礎，利用連續的 2 張影像尋找每個區塊的移動向量，經過統計得到全域移動向量，用來補償影像背景部分，進行到此的處理程序後，接下來就可以執行連續影像相減法來求得移動的像素，再以求得每個區塊的統計值，用以辨別移動區塊，再進行移動區塊的群聚，即可得到移動物體的區域。

因移動物體的偵測需對整張影像做處理，會使得計算量太高，我們從移動物體偵測中得到移動物體的區域後，以此區域製造移動物體的樣版，使用色彩資訊來判別相似度是否相近，利用運算速度快的平均位移法計算影像中與樣版相似度最高的候選影像，完成移動物體的追蹤，系統流程圖如下：



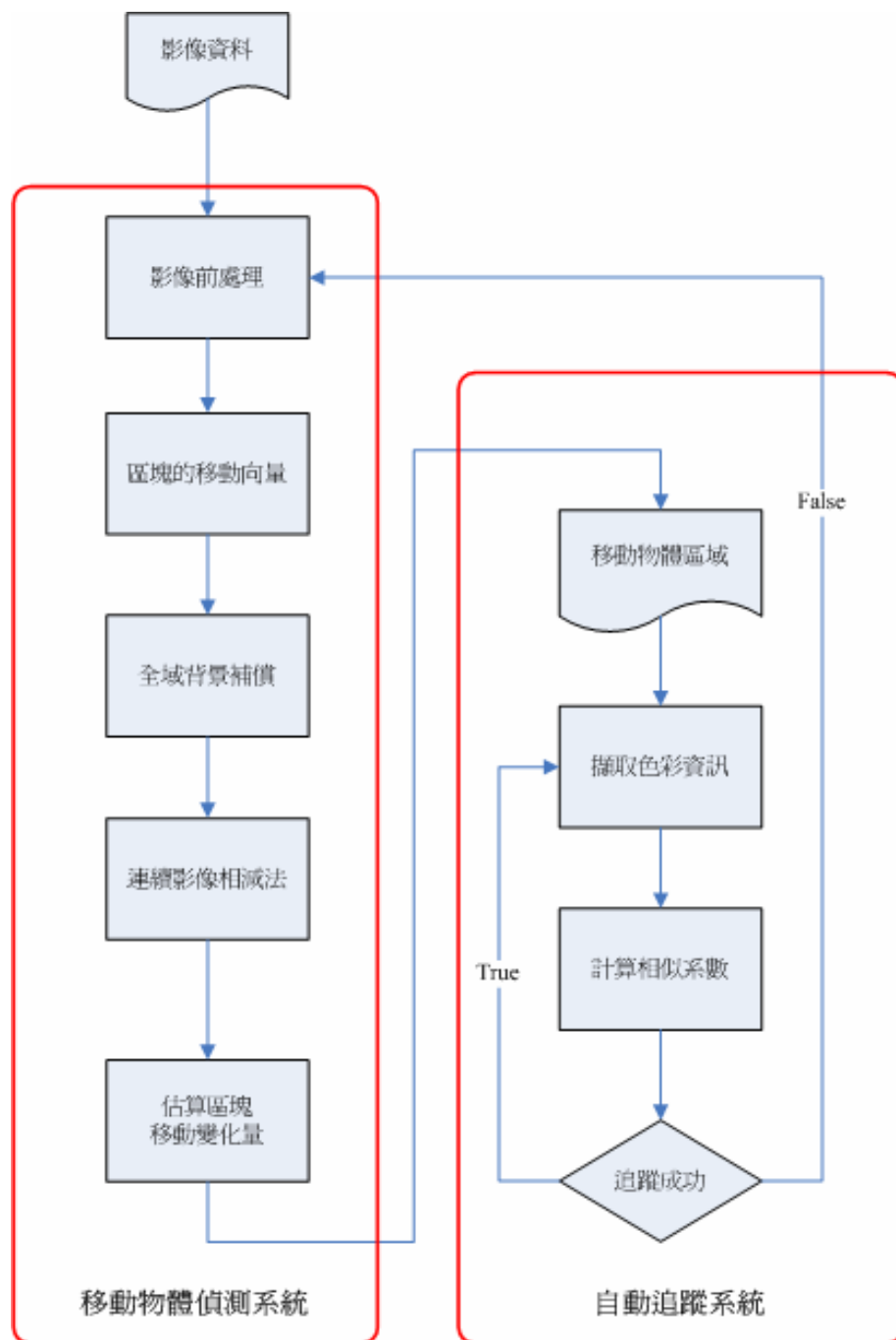


圖 2.3 系統流程圖

第三章

動態背景中移動物體之偵測系統

在此章節介紹移動物體偵測系統，並對其方法做詳細的說明。此系統包含了 5 個模組，如圖 3.1 所示，影像前處理將視訊影像從彩色影像轉換為灰階影像，得到灰階影像後被為數個 $N*N$ 的小區塊，經由區塊匹配求得每個區塊的區域移動向量(Local Motion Vector, LMV)，由區域移動向量找出全域的移動向量(Global Motion Vector, GMV)將移動的背景補償回去，這時便可執行連續影像相減法得到移動的區域，經過每個區塊的統計值[16]，找尋有可能的移動區塊並標記之，分別在以下各小節說明。

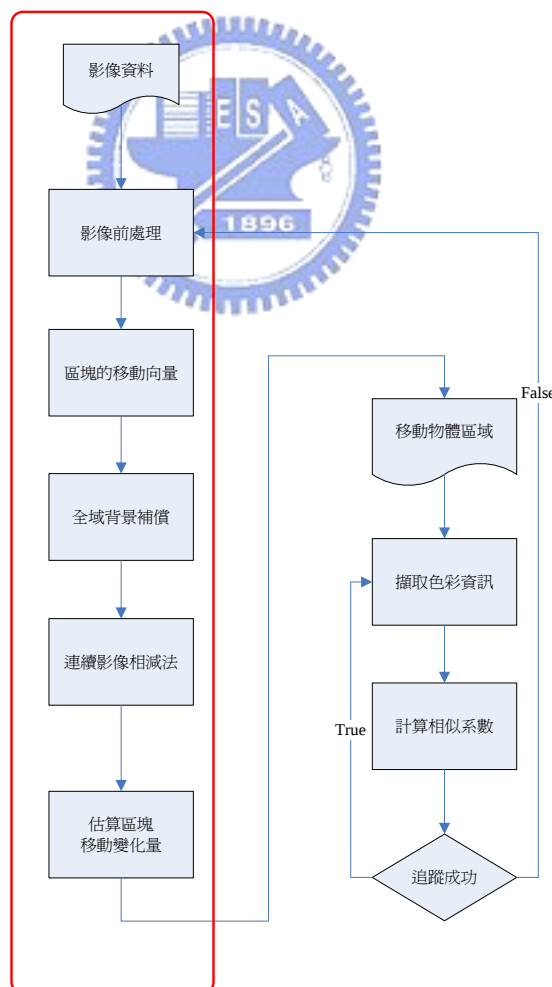


圖 3.1：移動物體偵測系統流程圖(方框部分)

3.1. 影像前處理

影像在進行主要的處理及判斷之前，通常會經過一道前處理的手續，主要的原因是降低一些干擾因素(例如：光源的影響，雜訊的干擾等)，藉以提升影像的品質，這可使之後誤判的情形降低，而使影像處理的正確性加以提升，這即是影像前處理主要的功用。以下即針對此論文中所用到的影像前處理[17]-[20]，做一簡短的說明介紹。

3.1.1 彩色對灰階的轉換：

基本上對人類視覺來說，所謂的“色彩”是自然光照射到物體，反射後被眼睛接收的資訊，經大腦分析後，最後所得到的結果，人眼所能看見的光在光譜上非常的窄，可見光的波長為 400nm 至 700nm，其它頻率低於可見光(紅外線、微波、無線電波)或高於可見光(紫外線、X 射線、伽瑪射線)便無法辨視。若將所有可見光混合，即成為白光，白光又可以紅、藍、綠三種基本色光混合而得之，因此若將三種基本色光以不同的比例混合，可讓我們得到各種不同的顏色，此三種基本色光紅、藍、綠因而被稱為三原色。

我們使用的視訊影像是彩色的，首先要對影像做彩色與灰階的轉換，為什麼要將彩色的影像轉換成灰階呢？其主要的原因是我們希望能降低在移動物體偵測部分的運算量。在電腦視覺中，每個彩色像素包含 3 個資料，分別為之前所提到的三原色，紅、藍、綠，經過轉換成為灰階值，只留下一個資料，運算上可降低為原本的三分之一，對速度的提升有一定的幫助。

影像的彩色模式上，較常被應用的有下列幾種模式：1.RGB 模式，2.HSV 模式，3.YIV 模式，4.YUV 模式。基於應用層面上的不同，各個模式各有其使用上的優點，在此我們需要知道的資訊是影像變動的情形，因此我們選擇 YUV 模式，因為它在數位色彩中是最常被用到的格式，除此之外，Y 值包含了影像明暗的資訊，可以藉此來判斷影像是否變動。從三原色 RGB 模式轉換成 YUV 模式，有以下的轉換關係：

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3.1)$$

由上式(3.1)我們可得到 $Y=0.299*R+0.587*G+0.114*B$ ，在 YUV 模式中，Y 佔了全部影像百分之九十以上的能量，利用上式我們可以保有高灰階影像，又可以大大地降低運算量。

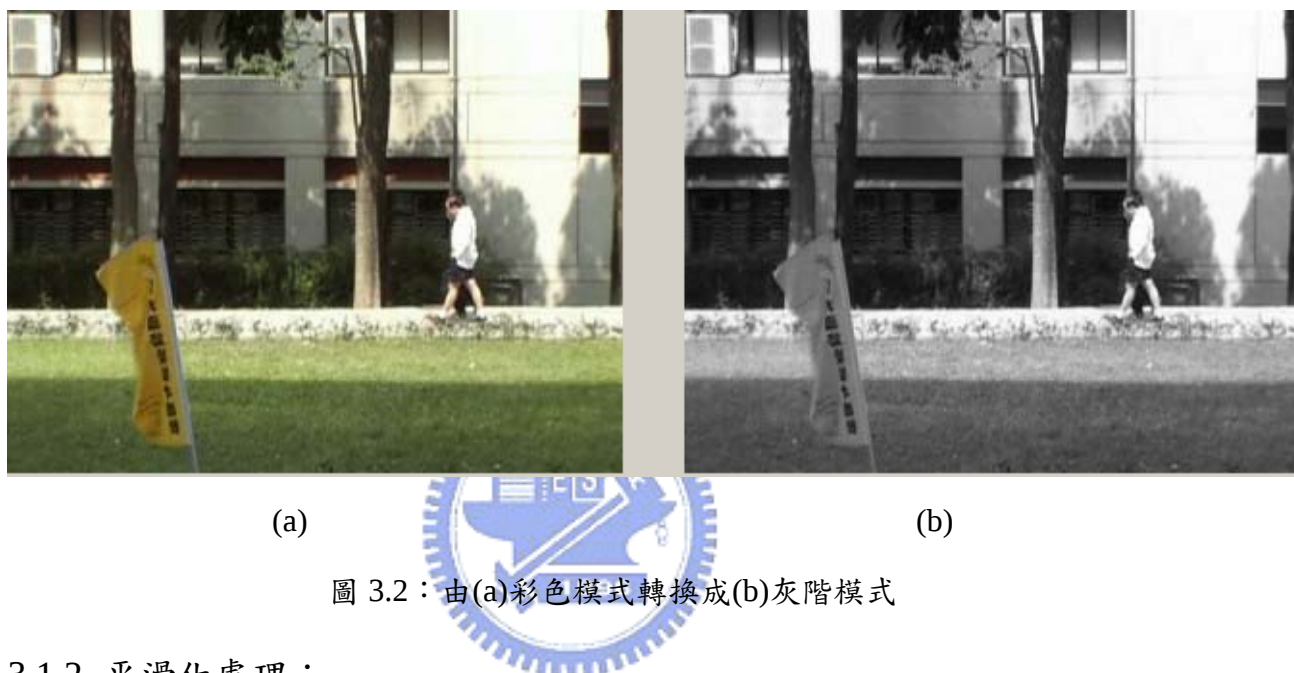


圖 3.2：由(a)彩色模式轉換成(b)灰階模式

3.1.2 平滑化處理：

這個作用是在抑制雜訊的干擾，此論文的場景是在室外，易受到光源的變動、背景細小的變化所影響，而所謂的雜訊是指，在影像中包含了突波或衝擊波，以訊號處理的觀點而言，這些雜訊是屬於高頻的成份，因此我們主要的目的便是除去這些屬於高頻的訊號，我們可以透過低通濾波器與原始影像訊號做迴旋積分(Convolution)，從而減低雜訊的影響，在影像處理中，是用二維的低通濾波器來降低雜訊訊號，我們選用了均值平滑濾波器(Mean Filter)來達到我們的目的。

均值平滑濾波器是利用一個 $N*N$ 的矩陣，對原始影像做迴旋積分，其轉移數函數為：

$$h(x, y) = \begin{cases} \frac{1}{n * n} & \text{if } x - m \leq x \leq x + m \quad \text{and} \quad y - m \leq y \leq y + m \\ 0 & \text{otherwise} \end{cases} \quad (3.2)$$

在此 $m = \frac{n-1}{2}$ ， n 為奇數。

在此論文是用 3*3 的轉移矩陣，其各元素值標示如下：

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

在圖 3.3 中可觀察到，原本因樹葉搖晃而造成的雜訊，經均值平滑濾波器後，被抑制後的結果。圖 3.3(a)是原始視訊影像，在圖 3.3(b)是未經均值平滑濾波之前的影像，有很明顯樹葉搖晃的情形及邊緣，在經過均值平滑濾波器後，在圖 3.3(c)中可看出大大地抑制這些雜訊。在此論文中移動物體偵測部分，是利用影像中的區塊做移動量的統計，這些雜訊會影響到移動量的估計，若在做移動量的統計之前，讓影像經過均值平滑濾波以降低雜訊，對之後的移動物體偵測結果可以降低其誤判的可能性，在 3.4 節會詳細地說明。



(a)



(b)

(c)

圖 3.3：(a)原始影像(b)未經均值平滑濾波器的影像(c)經均值平滑濾波器的影像

3.2. 移動向量之搜尋方式

在移動的背景之下，首先遇到的問題便是：何者為移動物體，何者為背景，建立全域背景[21]是可快速偵測到移動物體的方法，首先將攝影機可能移動的區域建立起各個部分的背景，找出各個背景的相似特徵，建立攝影機可照攝到的全域背景，當有移動物體進入，先做背景的匹配，再與背景相減即可得到完整的移動物體區域[22]，但這方法的缺點在於光源的影響，不同的光源下，可能使得背景匹配錯誤，或是移動物體區域的誤判，此外最大的缺點是在於需要全域的背景，這也表示它不能應用於未知的背景中。

我們想用於可變換的背景中，必須讓背景的移動量是可知的，用以補償背景的變動，於是我們利用區塊匹配法(Block Matching)[23]-[26]，尋找每個區塊的移動向量，從

而求得全景的移動向量。

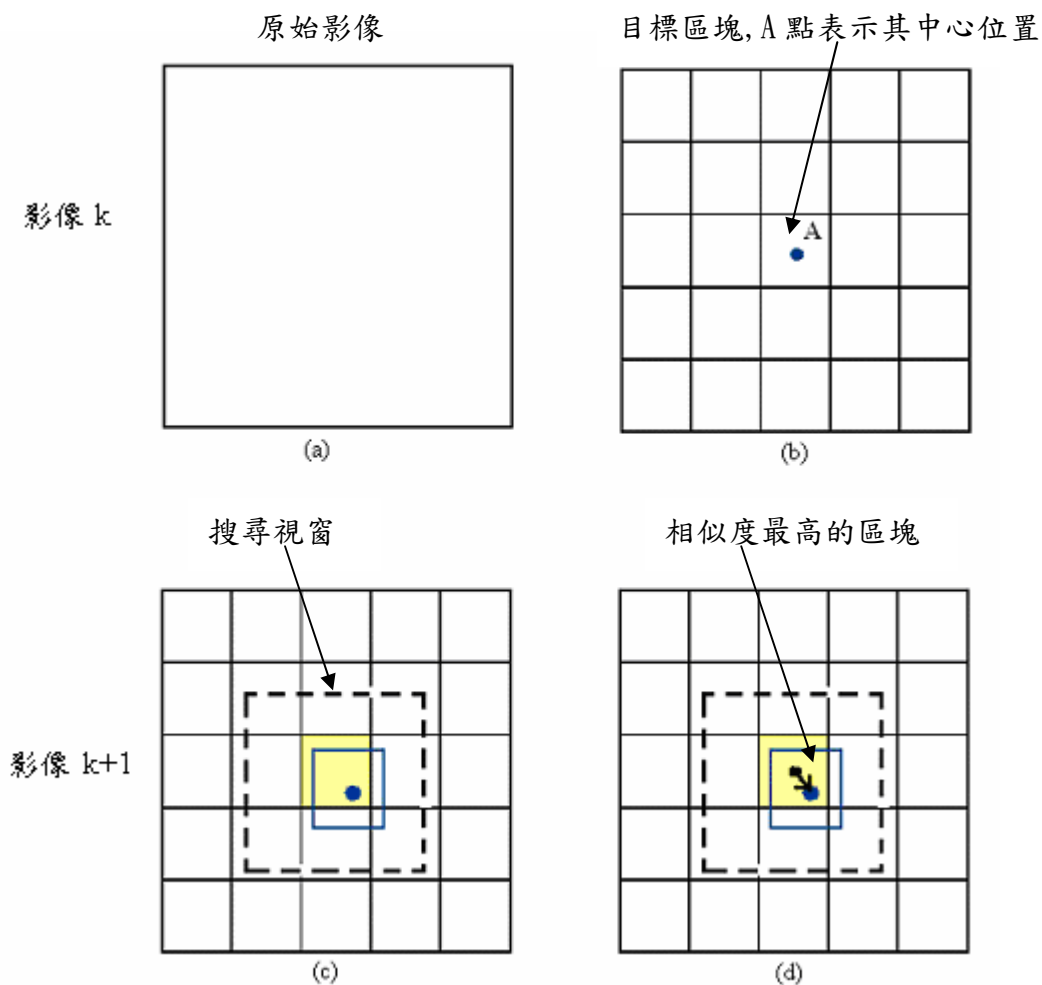


圖 3.4：區塊匹配法

區塊匹配法主要是將影像區分為以 $n*n$ 像素為單位所組成的區塊，尋找最相近的區塊，舉例來說，在圖 3.4(a)和(b)中為時間 k 的影像，圖 3.4(c)和(d)則為時間 $k+1$ 的影像；圖 3.4(a)為時間 k 原始影像，在圖 3.4(b)將影像 k 等分為一個個的區塊，以中央的區塊 A 為例，其中間的點代表此區塊的中心像素，在圖 3.4(c)影像 $k+1$ 中建立搜尋視窗(Search Window)，即虛線方框的部分，在影像 $k+1$ 的搜尋視窗內找出與圖 3.4(b) 的區塊 A 最相似的區塊，如此搜尋過整個搜尋視窗，相似值最近的我們就標示為區塊 A 的移動向量圖 3.4(d)。

在此論文中所採用的相似值量測，是以平均絕對差值(Mean Absolute Difference, MAD)來做判斷，如下式：

$$MAD = \frac{1}{nn} \sum_{i=1}^n \sum_{j=1}^n |A(i, j) - B(i, j)| \quad (3.3)$$

在搜尋視窗內擁有最小的 MAD 值，即為此區塊的區域移動向量(Local Motion Vector)，當計算完所有的區塊，我們得到全部的區域移動向量。

在區塊匹配中，我們為了精準度，且不影響即時性的應用之下，採取了全域搜尋(Full Search)，這時我們可以計算出全域的移動向量，以做背景的全景補償。

3.3. 動態背景之全景補償

從區塊匹配法中，我們求得了每個區塊的區域移動向量，接下來便是如何在這些區域移動向量中，找出全域移動向量。在這裡我們先做個觀察，圖 3.5 藍點是區塊的中心，從藍點延伸出去的紅線，便是移動向量的方向及移動量。我們發現這些區域移動向量有個規則性，同一方向的區域移動向量是佔大多數的。雖然在物體移動的部分，移動向量會散亂，但大多數的區域移動向量是同一個方向的，我們可以得到一個假設，背景的移動佔大多數的區域移動向量，因此找到最常出現的區域移動向量，即是全域移動向量。

在這我們使用的方法是尋找區域移動向量的中位向量(median vector)。

Global_Motion_Vector = Median(Local_Motion_Vectors)

再依得到的全域移動向量，將補償整張影像。

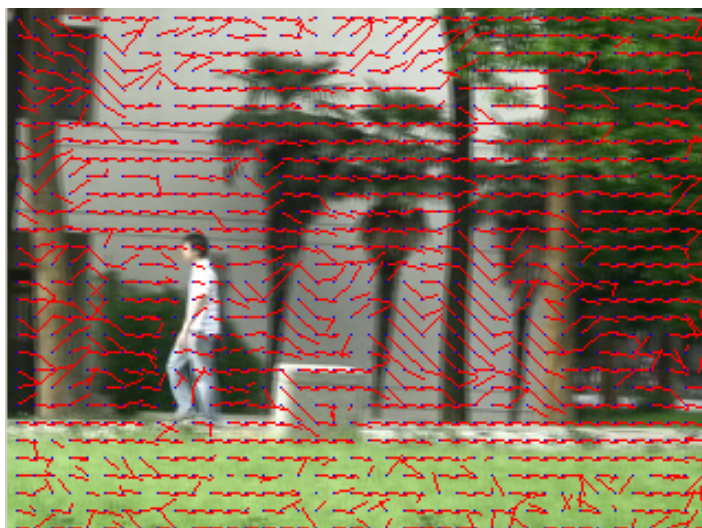


圖 3.5：區域移動向量

再以補償後的影像，與其前張影像相減，得到背景被抑制的結果圖 3.6 移動邊緣圖。



(a)



(b)

圖 3.6：移動邊緣圖(a)未補償的連續影像相減結果(b)補償後的連續影像相減結果

3.4. 移動物體之移動統計計算量

得到的移動邊緣圖中，移動的區域不僅僅移動物體，尚有抑制不完全的背景部分，如何分辨此二者的差異呢？在此我們利用了統計的方式。首先觀察移動物體的區域與被抑制的背景區域的相異點，背景部分的移動量較微小，移動物體的部分則擁有較大的變化量，於是利用標準差(Standard Deviation)來量測變化量，但標準差只有變化情形，無法得知變化量值的資訊，為了彌補此處資訊的不足，再加上每個區塊內的最大值計算，

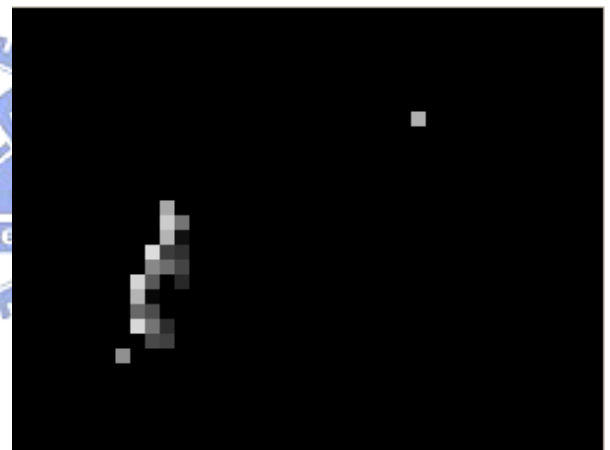
因為我們觀察到，最大的變化量是出現在移動物體部分。由此二值，我們得到影像中每個區塊的資訊，在下節便可就這些資訊得到移動物體的區域。

3.5. 移動物體之標記

將每個區塊的標準差與最大值求出後，兩者皆大於臨界值的，才視之為移動區塊，我們將符合此條件的區塊填入它的標準差，可得到移動區塊圖 3.7，因為最大值出現在移動物體的部分，我們便以擁有最大值的區塊為開始，利用八相鄰(8-Neighborhood)的性質，將相鄰的區塊標記成移動區塊，再以已標記後的區塊繼續尋找八相鄰的區塊，直到全部相鄰的區塊皆標記過，經由此過程，我們便可得到移動物體的區域，完成移動物體偵測部分。



(a)



(b)

圖 3.7：(a)圖 3.6(b)，(b)移動區塊圖

第四章

移動物體之自動追蹤系統

在第三章的移動物體偵測系統，我們以區塊為基礎求出區域移動向量並補償背景，再利用連續影像相減法來求得可能移動的像素。接著，將影像以等分的區塊大小做分割，統計各個區塊移動量的最大值與標準差之後，最後標定移動物體的位置與大小，此移動物體樣版影像的相關資訊便可作為移動物體追蹤時的依據。

在這章節要探討如何在視訊影像中，準確且有效率的執行即時移動物體的追蹤，本論文利用第三章所求得的移動物體的區域為樣版影像，使用以色彩分佈資訊為基礎的追蹤法，採用平均位移法(Mean Shift)來計算候選影像與樣版的相似度，以達成追蹤的目的。

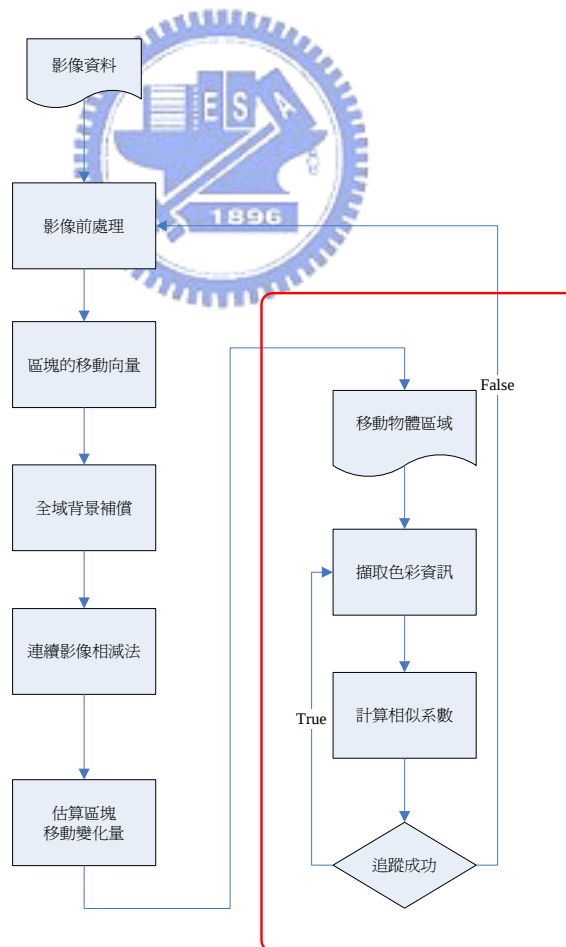


圖 4.1：自動追蹤系統流程圖(方框部分)

移動物體的追蹤可視為在多個候選影像中，找尋與樣版影像相似值最高的移動物體。而相似值的估算可利用諸如色彩分佈、形狀模型、移動向量等資訊的變化，在本論文我們利用樣版影像的色彩分佈密度函數，並使用 Bhattacharya 係數[29]來判斷與候選影像的相似度，判斷的方法則使用平均位移疊代(Mean Shift Iteration)[28]-[32]，最後則說明整體追蹤的系統流程。

4.1. 移動物體之色彩資訊分析

在本節，我們將介紹如何利用移動物體的色彩分佈資，並搭配平均位移法來追蹤移動物體。首先，建立樣版影的色彩分佈密度函數，然後介紹平均位移法如何從候選影像尋找相似的移動物體。

4.1.1. 色彩分佈密度函數之定義

給定所偵測的移動物體位置後，我們要抽取其色彩分佈資訊的特徵。令樣版影像的色彩分佈密度函數為 $q(u)$ ，而候選影像在位置 y 的色彩密度函數為 $p(u, y)$ ，我們的任務則是要找到位置 y ，使得 $p(u, y)$ 與 $q(u)$ 有最高的相似度。

$q(u)$ 的定義為：

$$q(u) = C \sum_{i=1}^n k(\|x_i^*\|^2) \delta[b(x_i^*) - u], \quad (4.1)$$

其中， $\{x_i^*\}_{i=1, \dots, n}$ 表示樣版內所有像素位置的集合，其像素總數為 n ，令其中心點為 0；

函數 $b: R^2 \rightarrow \{1 \dots m\}$ ， $b(x_i^*)$ 回傳 x_i^* 相對應的色彩索引(Color Index)， $u; [1:m]$ 為色彩索引的範圍，在本論文中我們用 24 位元全彩影像，所以色彩索引的範圍為 $[1:2^{24}]$ ； $\delta(x)$ 為

Kronecker delta function； C 為正規化常數， $C = \frac{1}{\sum_{i=1}^n k(\|x_i^*\|^2)}$ ，使得 $\sum_{u=1}^m q(u) = 1$ ； $k(\|x_i^*\|^2)$

為核心函數，關於核心函數將於下一節再進一步說明。色彩分佈密度函數的求取可視為，先獲得樣版影像的色彩直條圖，再將其正規化。

候選影像色彩分佈密度函數的定義 $p(u, y)$ 與 $q(u)$ 類似，差別是前者以 y 為中心，以 $\{x_i\}_{i=1 \dots n}$ 來表示候選影像內所有像素位置的集合，並使用相同的核心函數 k ，其定義為：

$$p(u, y) = C_h \sum_{i=1}^{n_h} k\left(\left\|\frac{x_i - y}{h}\right\|^2\right) \delta[b(x_i) - u], \quad (4.2)$$

其中， C_h 為正規化常數， $C_h = \frac{1}{\sum_{i=1}^{n_h} k\left(\left\|\frac{x_i - y}{h}\right\|^2\right)}$ ，使得 $\sum_{u=1}^m p(u, y) = 1$ ； h 為核心函數之

半徑長度。

我們定義了移動物體的色彩特徵，我們便可以樣版影像為參考影像，使之與候選影像做比較，相似度越高的候選影像有越高的機率是與樣版影像相似的，下一章便介紹關係係數。

4.1.2. 核心函數

本節討論為什麼需要核心函數及其定義。我們可以觀察，樣版影像在中心點的像素 (Pixel) 含有較高的資訊量，愈接近邊緣的影像則愈容易受到背景的干擾，其有效資訊量愈低，核心函數的作用在於讓影像樣版中，接近中心點的像素的權重較高，而離中心點愈遠的像素，則權重逐漸降低，如圖 4.2 所示。

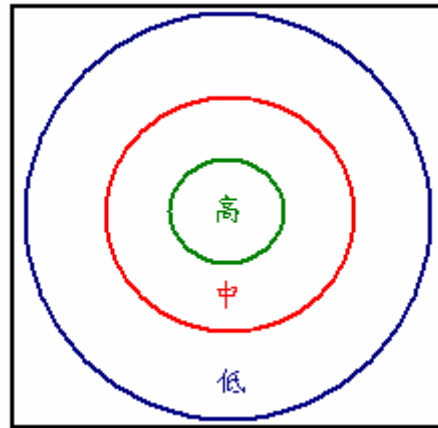


圖 4.2 影像像素資訊的重要性示意

根據核心函數的功能，我們可以知曉核心函數大體上為中間高起，四周低下的函

數，核心函數的定義：

X 代表 d 維的歐氏空間(Euclidean space)， x 為此空間中的一點， x 的歐氏長度(Euclidean Length)為 $\|x\|^2 = x^t x$ ， R 為實數域，設一函數 $K: X \rightarrow R$ 存在一剖面函數(Profile Function) $k: [0, \infty] \rightarrow R$ ，即 $K(x) = k(\|x\|^2)$ ，滿足以下條件：

1. k 為非負的
2. k 是非增的，若 $a < b$ ，則 $k(a) \geq k(b)$
3. k 為分段連續的，且 $\int_0^{\infty} k(t) dt < \infty$

則 $K(x)$ 為核心函數。

以下介紹幾種較常被應用到的核心函數，分別為：

1. Uniform Kernel
2. Normal Kernel Function
3. Epanechnikov Kernel Function

Uniform Kernel 它的定義如下：

$$K(x) = \begin{cases} 1 & \text{if } \|x\| < 1 \\ 0 & \text{if } \|x\| \geq 1 \end{cases} \quad (4.3)$$

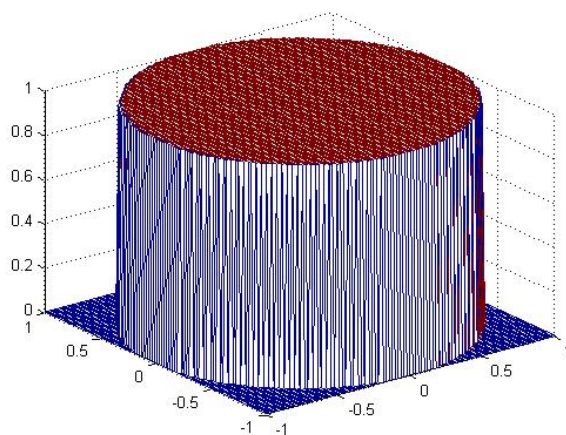


圖 4.3 Uniform Kernel Function

Normal Kernel Function 又稱為 Gauss Kernel Function，它的定義如下：

$$K_N(x) = \begin{cases} (2\pi)^{-\frac{d}{2}} \exp(-\frac{1}{2}\|x\|^2) & \text{if } \|x\| < 1 \\ 0 & \text{otherwise} \end{cases} \quad (4.4)$$

d 為空間的維度，我們要處理的是二維影像，因此 d 值為 2。

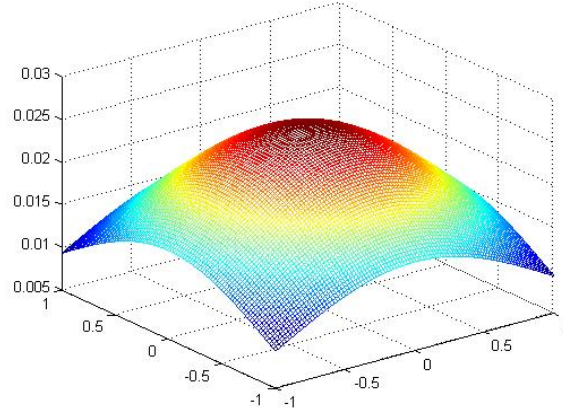


圖 4.4 Normal Kernel Function

Epanechnikov Kernel Function，它的定義如下：

$$K(x) = \begin{cases} \frac{1}{2} c_d^{-1} (d+2) (1 - \|x\|^2) & \text{if } \|x\| < 1 \\ 0 & \text{otherwise} \end{cases} \quad (4.5)$$

d 為空間的維度，我們要處理的是二維影像，因此 d 值為 2。

C_d 為 d 維度空間中，單位圓的面積，因此 C_d 值為 π 。

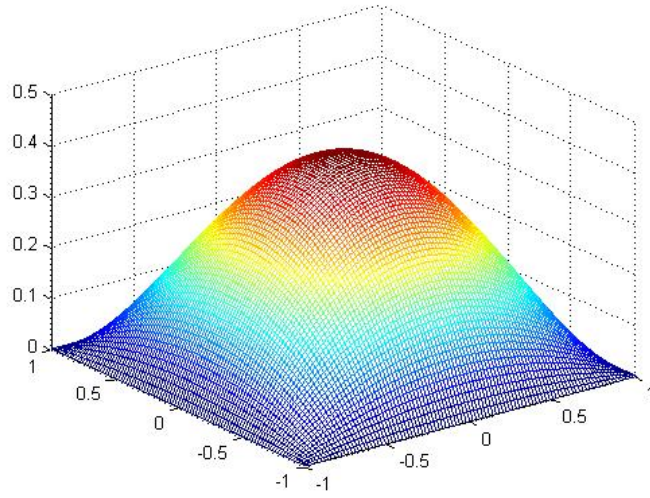


圖 4.5 Epanechnikov Kernel Function

根據文獻[31]的機率密度函數估計，Epanechnikov Kernel Function 所得的平均全體誤差(Average Global Error)會是最小，因此在此論文中選用 Epanechnikov Kernel Function 為核心函數。



4.2. 移動物體之候選位置相似係數

在本論文中，我們採用 Bhattacharyya 係數[33]量測樣版與候選影像的相似度，其定義如下：

$$\rho(y) \equiv \rho[p(u, y), q(u)] = \int \sqrt{p(u, y)q(u)} du \quad (4.6)$$

為了將 Bhattacharyya 係數運用至色彩分佈密度函數上，其型式可調整如下：

$$\rho(y) \equiv \rho[p(u, y), q(u)] = \sum_{u=1}^m \sqrt{p(u, y)q(u)} \quad (4.7)$$

由式(4.7)我們可以定義兩個色彩分佈密度函數的距離(Distance)為：

$$d(y) = \sqrt{1 - \rho[p(u, y), q(u)]} \quad (4.8)$$

根據式(4.7)， Bhattacharyya 係數有下列三個特性：

1. $0 \leq \rho(y) \leq 1$ 。
2. $\rho(y) = 1$ 則 $p(u, y) = q(u)$ ，反之亦然。
3. $\rho(y) = 0$ 則 $p(u, y)$ 與 $q(u)$ 為正交，反之亦然。

根據以上三個特性做推論，當樣版影像與候選影像的色彩分佈密度函數最相似時，便是 Bhattacharyya 係數最大的時候。

接下來，我們將利用泰勒展開式(Taylor Expansion) 對 Bhattacharyya 係數做展開，以推導出如何尋找與樣版影像色彩分佈密度函數最相似的候選影像的方法，使得 Bhattacharyya 係數有最大值。在泰勒展開式中，越高階的導數項，其重要性越低，於是我們只展開至一階導數項，而忽略其高階項，泰勒展開式對函數 $\rho[p(u, y)]$ 在 $p(u, y_0)$ 的一階導數項展開為：

$$\rho[p(u, y)] \approx \rho[p(u, y_0)] + \rho'[p(u, y_0)][p(u, y) - p(u, y_0)] \quad (4.9)$$

式(4.7)代入式(4.9)可得到以下結果：

$$\rho[p(u, y), q(u)] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p(u, y_0)q(u)} + \frac{1}{2} \sum_{u=1}^m p(u, y) \sqrt{\frac{q(u)}{p(u, y_0)}} \quad (4.10)$$

將式(4.2)代入上式：

$$\rho[p(u, y), q(u)] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p(u, y_0)q(u)} + \frac{C_h}{2} \sum_{i=1}^n w_i k\left(\left\|\frac{x_i - y}{h}\right\|^2\right) \quad (4.11)$$

$$\text{其中 } w(x_i) = \sum_{u=1}^m \delta[b(x_i) - u] \sqrt{\frac{q(u)}{p(u, y_0)}} \quad (4.12)$$

為了得到最大的相似度，式(4.11)的 Bhattacharyya 係數必須達到最大值。我們由式(4.11)中可以看出第一項的值由 y_0 決定，為一常數。因此，主要影響 Bhattacharyya 係數是第二項，為了使 Bhattacharyya 係數達到最大，式(4.11)的第二項需達到最大值，接下來我們將推導以平均位移法進行疊代，以求得最大值，並找出相對應的候選影像。

4.3. 應用平均位移法於移動物體之追蹤

平均位移法(Mean Shift)最早是由 Fukunaga 等人[34]在 1975 年提出，最初是用於機率密度函數(Probability Density Function)的梯度(Gradient)估算，在 1995 年由 Yizong Cheng[35]做了兩方面的推廣，首先定義了一個核心函數(Kernel Function)，讓位於樣版中心的值給予較高的權重，越遠離樣版中心則權重越低，接下來則設定權重係數(Weight Coefficient)，不同的樣版的點，給予的重要性不同。之後應用平均位移法在圖像平滑與圖像分割中，進而應用於移動物體追蹤，將移動物體追蹤近似於平均位移法最優化的問題，使得追蹤可在即時性的系統中運作。在這一章介紹平均位移法的理論基礎，與如何運用至移動物體的追蹤應用上。

4.3.1. 平均位移法理論背景介紹

在 d 維空間 R^d 中，給定 n 個點 x_i ， $i=1, \dots, n$ ，則在參考點 x 的平均位移向量(Mean Shift Vector)的定義如下：

$$M_h(x) = \frac{1}{n'} \sum_{x_i \in S_h} (x_i - x), \quad (4.13)$$

其中 S_h 是半徑為 h 的高維球形區域，滿足以下關係式的集合：

$$S_h(z) \equiv \{z : (z - x)^t (z - x) \leq h^2\} \quad (4.14)$$

n' 則為在這 n 個 x_i 中，落在 S_h 內的個數。

由式(4.13)可看出， $x_i - x$ 是 x_i 相對於參考點 x 的偏移量， $M_h(x)$ 即為落在 S_h 內 n' 個 x_i 對參考點 x 偏移量的平均。同時， $M_h(x)$ 也可看成是參考點 x 指向 x_i 分佈密度較高的方向向量，若以機率密度函數來看，因為機率密度函數的梯度指向機率密度增加最大的方向，因此 S_h 區域內有較多的 x_i 落在機率密度梯度的方向，所以平均位移向量 $M_h(x)$ 便會指向機率密度梯度的方向。

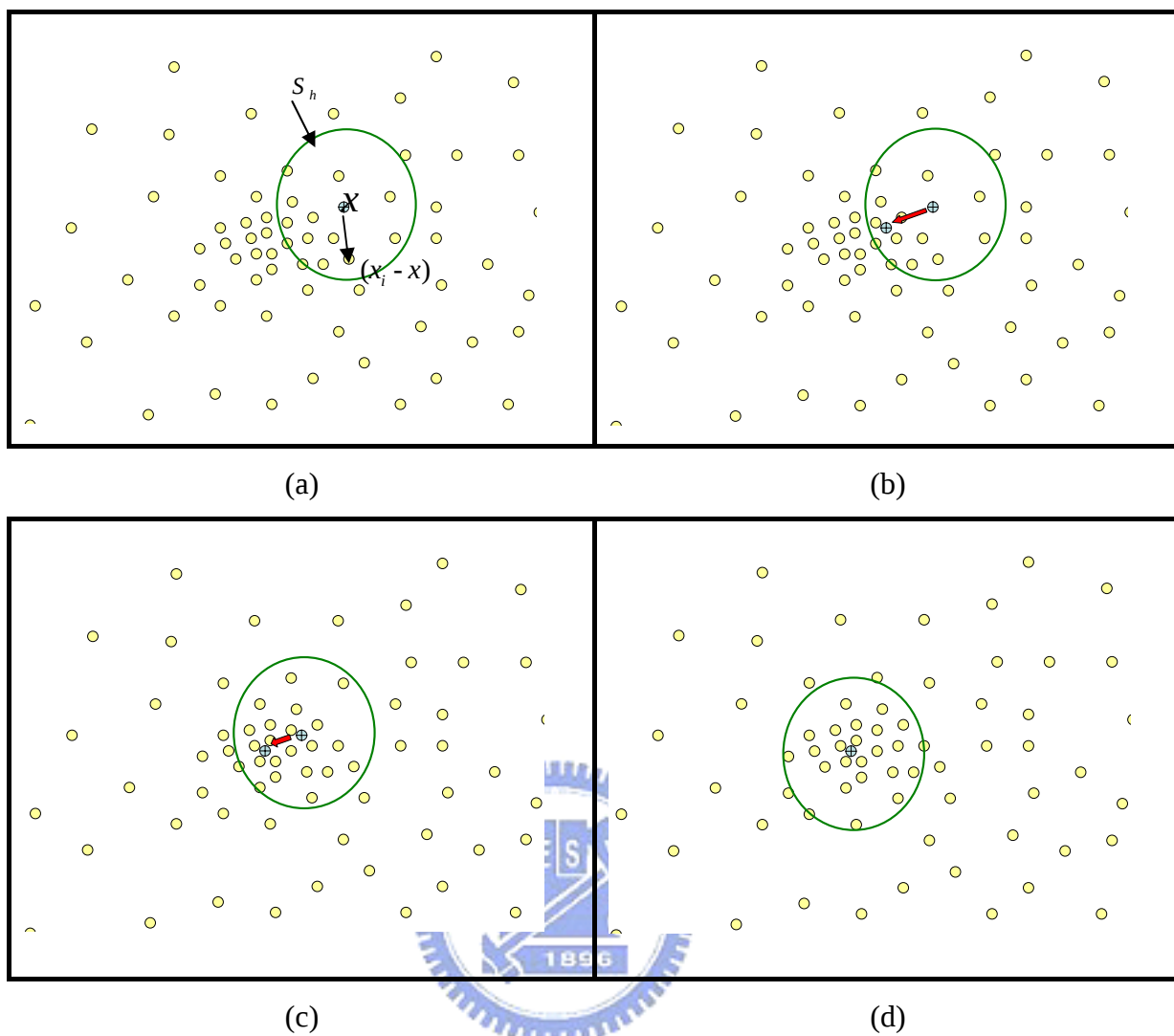


圖 4.6 平均位移法示意圖

圖 4.6 為平均位移法示意圖，大圓圈的範圍為 S_h ，小圓點代表落在 S_h 的 x_i ，數量有 n' 個，中間的黑點為參考點 x ，箭頭表示 $(x_i - x)$ 是 x_i 相對於參考點 x 的偏移向量如圖 4.6(a) 所示；由圖 4.6(b) 可看出平均位移向量 $M_h(x)$ 會指向 x_i 分佈較多的方向，即機率密度函數的梯度方向；若新的參考點 x' 以舊的參考點 x 加上 $M_h(x)$ 取代，則經過幾次疊代後，新的參考點 x' 便會逐漸移至密度分佈最高的地方，如圖 4.6(b)(c)(d) 所示。

在介紹完平均位移向量的標準形式，接下來引進核心函數的概念，我們考慮將距離的影響加入平均位移向量，同時我們也將落在 S_h 內的點 x_i 根據重要性的不同，加入權重係數，於是我們便可將基本的平均位移向量形式，擴充成下式：

$$M(x) = \frac{\sum_{i=1}^n G\left(\frac{x_i - x}{h}\right) w(x_i) (x_i - x)}{\sum_{i=1}^n G\left(\frac{x_i - x}{h}\right) w(x_i)} \quad (4.15)$$

其中， $G(x)$ 是一個核心函數， h 為核心函數的半徑長度。

4.3.2. 平均位移法求取最高的相關係數

我們使用 Bhattacharyya 係數來計算樣版影像與候選影像的色彩分佈密度函數的相似度，Bhattacharyya 係數越高，代表樣版影像與候選影像越相似，如何找尋最高 Bhattacharyya 係數便是本節中所要探討的主題。

前面定義了樣版影像與候選影像的色彩分佈密度函數，分別為式(4.1)及式(4.2)，我們將這兩個函數在此重寫一次如下：

樣版影像色彩分佈密度函數：

$$q(u) = C \sum_{i=1}^n k\left(\left\|x_i^* - u\right\|^2\right) \delta[b(x_i^*) - u] \quad (4.16)$$

候選影像色彩分佈密度函數：

$$p(u, y) = C_h \sum_{i=1}^{n_h} k\left(\left\|\frac{x_i - y}{h}\right\|^2\right) \delta[b(x_i) - u] \quad (4.17)$$

我們的目的是要估算出最高 Bhattacharyya 係數：

$$\rho(y) \equiv \rho[p(u, y), q(u)] = \sum_{u=1}^m \sqrt{p(u, y) q(u)} \quad (4.18)$$

平均位移向量(Mean Shift Vector)是一個函數指向機率密度梯度的方向，它亦是此一函數增加量最大的方向，以下我們便要導證如何用平均位移法求得 Bhattacharyya 係數最大值。

給予一個機率密度函數 $f(x)$ ，在 d 維空間中已知 n 個點 x_i ， $i=1, \dots, n$ ， $f(x)$ 的核密度估計(Kernel Density Estimation)為：

$$\bar{f}(x) = \frac{\sum_{i=1}^n k\left(\frac{x_i - x}{h}\right) w(x_i)}{h^d \sum_{i=1}^n w(x_i)} \quad (4.19)$$

其中 $w(x_i) \geq 0$ 是給點 x_i 的權重， $k(x)$ 是核心函數。

我們對 $k(x)$ 做變數代換，滿足 $k(x) = K(\|x\|^2)$ 。

接下來我們便要求得機率密度函數 $f(x)$ 的梯度 $\nabla \bar{f}(x)$ ：

$$\nabla \bar{f}(x) = \frac{-2 \sum_{i=1}^n (x_i - x) K'\left(\left\|\frac{x_i - x}{h}\right\|^2\right) w(x_i)}{h^{d+2} \sum_{i=1}^n w(x_i)} \quad (4.20)$$

定義 $g(x)$ 為 $K(x)$ 的負導函數 $g(x) = -K'(x)$ 代入式(4.20)，對分子分母同乘上

$$\sum_{i=1}^n g\left(\left\|\frac{x_i - x}{h}\right\|^2\right) w(x_i) :$$

$$\begin{aligned} \nabla f(x) &= \frac{2 \sum_{i=1}^n (x_i - x) g\left(\left\|\frac{x_i - x}{h}\right\|^2\right) w(x_i)}{h^{d+2} \sum_{i=1}^n w(x_i)} \\ &= \frac{2}{h^2} \left[\frac{\sum_{i=1}^n g\left(\left\|\frac{x_i - x}{h}\right\|^2\right) w(x_i)}{h^d \sum_{i=1}^n w(x_i)} \left| \frac{\sum_{i=1}^n (x_i - x) g\left(\left\|\frac{x_i - x}{h}\right\|^2\right) w(x_i)}{\sum_{i=1}^n g\left(\left\|\frac{x_i - x}{h}\right\|^2\right) w(x_i)} \right| \end{aligned} \quad (4.21)$$

式(4.21)的第 1 項為以 $g(x)$ 為核心函數，對 $f(x)$ 作核密度估計，而第 2 項即是式(4.15)所定義的平均移動向量，我們可將之表示如下式：

$$M(x) = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x_i - x}{h}\right\|^2\right) w(x_i)}{\sum_{i=1}^n g\left(\left\|\frac{x_i - x}{h}\right\|^2\right) w(x_i)} - x \quad (4.22)$$

對 $f(x)$ 做梯度運算後，得到 $f(x)$ 的梯度向量式(4.21)，其為 $f(x)$ 增加量最大的方向，可經由式(4.22)的平均移動向量計算出來。

現在我們將式(4.19)與式(4.11)做比較，分別重寫如下：

$$\bar{f}(x) = \frac{\sum_{i=1}^n k\left(\frac{x_i - x}{h}\right) w(x_i)}{h^d \sum_{i=1}^n w(x_i)} \quad (4.23)$$

$$\rho[p(u, y), q(u)] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p(u, y_0) q(u)} + \frac{C_h}{2} \sum_{i=1}^n w_i k\left(\left\|\frac{x_i - y}{h}\right\|^2\right) \quad (4.24)$$

由上式可看出式(4.23)與式(4.24)的第二項有相同的數學形式，所以式(4.22)所推導出來的平均移動向量，可以用來計算 Bhattacharyya 係數最大值，亦即式(4.2)所定義之候選影像色彩密度分佈函數在 $y = x + M(x)$ 時，Bhattacharyya 係數有最大值，其中，

$$M(x) = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x_i - x}{h}\right\|^2\right) w(x_i)}{\sum_{i=1}^n g\left(\left\|\frac{x_i - x}{h}\right\|^2\right) w(x_i)} - x \quad (4.25)$$

現在我們考慮核心函數 $K(x)$ 為式(4.5)的 Epanechnikov Kernel Function 如下所示：

$$K(x) = \begin{cases} \frac{1}{2} c_d^{-1} (d+2) (1 - \|x\|^2) & \text{if } \|x\| < 1 \\ 0 & \text{otherwise} \end{cases} \quad (4.26)$$

經過變數代換，滿足 $K(x) = k(\|x\|^2)$ ，因為我們處理的影像是二維，所以 d 值為 2，

C_d 為單位圓的面積 $C_d = \pi \times 1^2 = \pi$ ，代入式(4.26)得到：

$$k(x) = \begin{cases} \frac{2}{\pi} (1 - x) & \text{if } x < 1 \\ 0 & \text{otherwise} \end{cases} \quad (4.27)$$

令 $g(x)$ 為 $k(x)$ 的負導函數， $g(x) = -k'(x)$ 代入式(4.27)：

$$g(x) = \begin{cases} \frac{2}{\pi} & \text{if } x < 1 \\ 0 & \text{otherwise} \end{cases} \quad (4.28)$$

將式(4.28)代回式(4.25)可得：

$$M(x) = \begin{cases} \frac{\sum_{i=1}^n x_i w(x_i)}{\sum_{i=1}^n w(x_i)} - x & \text{if } \left\|\frac{x_i - x}{h}\right\|^2 < 1 \\ -x & \text{otherwise} \end{cases} \quad (4.29)$$

因為 $\left\| \frac{x_i - x}{h} \right\|^2$ 必定小於 1，式(4.29)可改寫為：

$$M(x) = \frac{\sum_{i=1}^n x_i w(x_i)}{\sum_{i=1}^n w(x_i)} - x \quad (4.30)$$

我們使用 Epanechnikov 核心函數在平均位移向量中，利用平均位移向量求得 Bhattacharyya 係數的最大值，此步驟可簡化以上式表示，於是我們便可用平均位移法計算與樣版影像有最大的相似度的候選影像，藉由此法可快速且有效的找出和移動物體的樣版影像最相似的候選影像，達到追蹤的效果。

4.4. 平均位移法的實作流程

之前章節已介紹平均位移法的基礎知識及應用原理，在此節將介紹在二維的視訊影像中，以平均位移法實現追蹤的系統流程。

以上定義了樣版影像與候選影像的色彩分佈密度函數，以 Bhattacharyya 係數計算兩者間的相似度，再利用平均位移疊代法求得最高的 Bhattacharyya 係數。在平均位移疊代法中，候選影像的起始位置的選擇是很重要，若選擇錯誤會令追蹤的結果完全失敗，然而我們觀察到，在現實世界中移動物體的運動一定是有連續性的，只要在取樣時間的間隔夠短，可以認為下一張影像物體移動後的位置與前張影像的位置相差不遠，因此可解決起始點選擇的問題，於是我們便以前張影像的位置做為起始點，以進行平均位移的疊代。我們在第三章已得知移動物體的區域及座標，此移動物體的中心點為 y_0 ，經由以上觀察結果的基礎下，我們以 y_0 為平均位移法的起始位置，執行平均位移法的疊代，其流程如下：

1. 在第三章偵測到的移動物體的區域及座標做為樣版影像，在目前的視訊影像

中，根據 $q(u) = C \sum_{i=1}^n k(\|x_i^*\|^2) \delta[b(x_i^*) - u]$ 計算出樣版影像相對應的色彩分佈密度函數 $\{q_u\}_{u=1\dots m}$ ，以樣版影像的中心點 y_0 為起始位置。

2. 在下個時間點的影像中，計算出候選影像相對應的色彩分佈密度。 y_0 為起始位

置，根據 $p(u, y) = C_h \sum_{i=1}^{n_h} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \delta[b(x_i) - u]$ 計算出候選影像相對應的

$\{p(u, y_0)\}_{u=1\dots m}$ 。

3. 將樣版影像色彩分佈密度 $\{q_u\}_{u=1\dots m}$ 與候選影像色彩分佈密度 $\{p(u, y_0)\}_{u=1\dots m}$ 代入 Bhattacharyya 係數 $\rho[p(u, y_0), q(u)] = \sum_{u=1}^m \sqrt{p(u, y_0)q(u)}$ ，求得 y_0 位置的 Bhattacharyya 係數。

4. 由 $w(x_i) = \sum_{u=1}^m \delta[b(x_i) - u] \sqrt{\frac{q(u)}{p(u, y_0)}}$ 計算出 $\{w(x_i)\}_{i=1\dots n_k}$ 。

5. 由 $M(y_0) = \frac{\sum_{i=1}^{n_h} x_i g\left(\left\|\frac{x_i - y_0}{h}\right\|^2\right) w(x_i)}{\sum_{i=1}^{n_h} g\left(\left\|\frac{x_i - y_0}{h}\right\|^2\right) w(x_i)} - y_0$ 計算出候選影像的新位置

$$y_1 = \frac{\sum_{i=1}^{n_h} x_i g\left(\left\|\frac{x_i - y_0}{h}\right\|^2\right) w(x_i)}{\sum_{i=1}^{n_h} g\left(\left\|\frac{x_i - y_0}{h}\right\|^2\right) w(x_i)}。$$

6. 以新位置 y_1 ，更新 $\{p(u, y_1)\}_{u=1\dots m}$ 與 $\rho[p(u, y_1), q(u)] = \sum_{u=1}^m \sqrt{p(u, y_1)q(u)}$ 。

7. 當 $\rho[p(u, y_1), q(u)] < \rho[p(u, y_0), q(u)]$ ，修正 $y_1 = \frac{1}{2}(y_0 + y_1)$ 。

8. 若是 $\|(y_1 - y_0)\| < \varepsilon$ 則停止疊代，否則更新 y_0 ，使 $y_0 = y_1$ ，回到步驟 2 繼續疊代的動作。在此 ε 為我們定義的門檻值，若 y_1 與 y_0 的位移量小於此門檻值，便會停止疊代。

在步驟 5 中平均位移量不一定會使得 Bhattacharyya 係數增加，於是加入步驟 7 來驗證 Bhattacharyya 係數是否增加，若 Bhattacharyya 係數沒有增加，則修正新位置 y_1 。

經由以上步驟的疊代，我們可以藉著計算出相似度最高的候選影像，達到追蹤的效果。

第五章

實驗結果

此系統的實驗影片是經由手持式攝影機，拍攝在戶外的環境中目標物體移動的影像，並透過影像擷取卡將影像轉換成為 320*240 像素大小的影像序列，所擷取的影像資料格式為 24 位元彩色的未壓縮 AVI 影像檔，將此 AVI 影像檔輸入至本論文的實驗系統，產生輸出影像。此系統的測試硬體設備為 Pentium 4 2.8GHz 中央處理器，512MB 的記憶體，作業環境是 Microsoft Windows XP，此實驗的開發平台為 Borland C++ 6.0，圖 5.1 為實驗的輸出畫面。

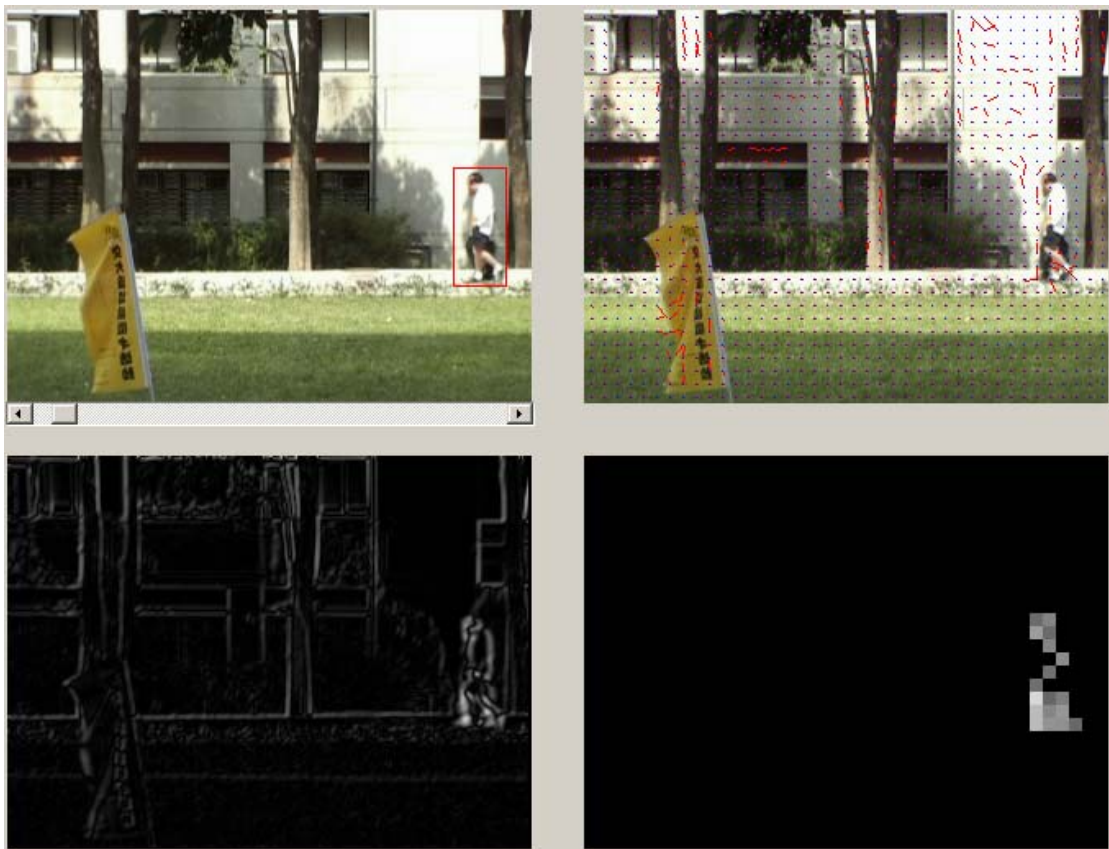


圖 5.1：系統程式輸出畫面

在左上角的部分為前一張視訊影像，紅色方框為移動物體偵測的結果，右上角為下一張視訊影像，同時也顯示出每個區塊的區域移動向量，而在左下角為經由全域區域移動向量補償後所得到的連續影像相減的結果，右下角為左下角的影像經過移動統計後所求得的移動區塊。

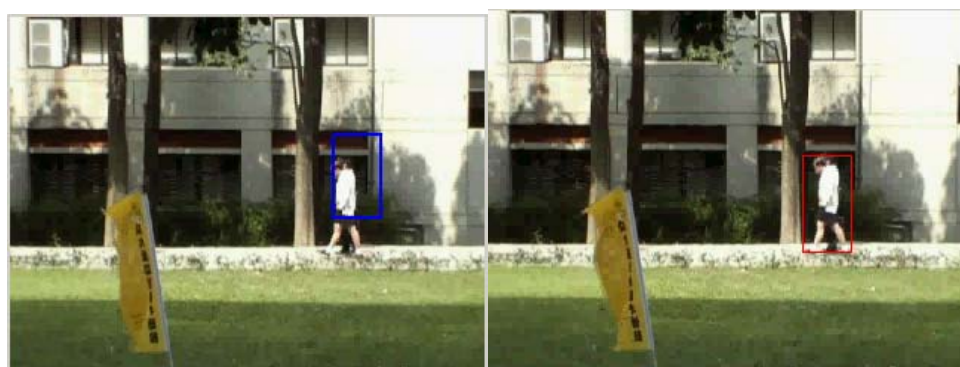
在我們在實驗中，偵測到移動物體後，即以紅色方框來表示，進入追蹤部分，即以藍色方框來表示。



frame 75 frame 76

圖 5.2：系統程式輸出結果

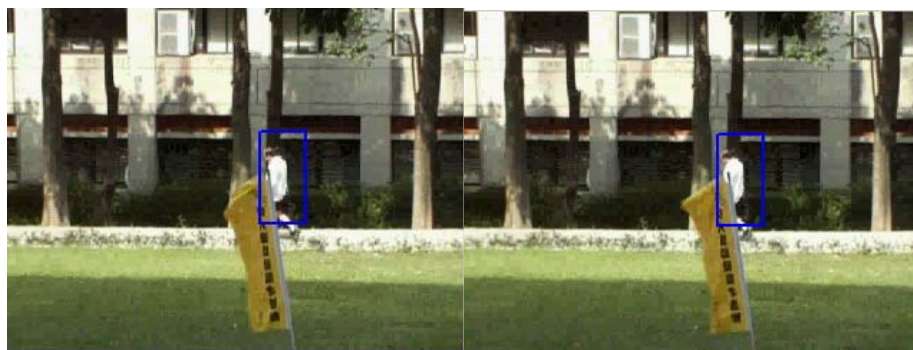
當與樣版影像相似度太低時，令其進入移動物體偵測程序。



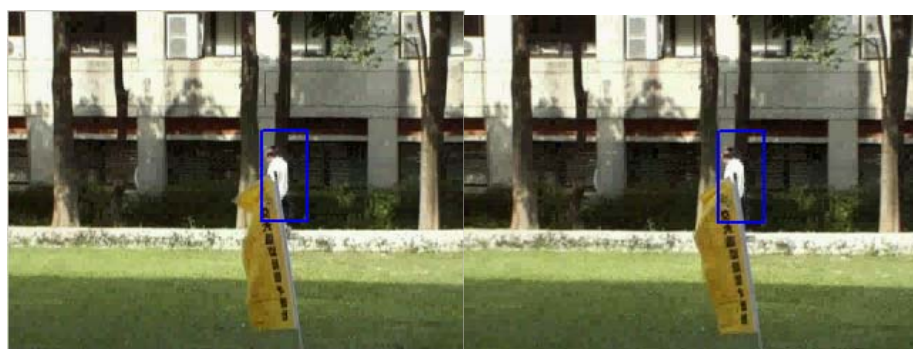
frame 113 frame 114

圖 5.3：與樣版影像相似度太低時的結果

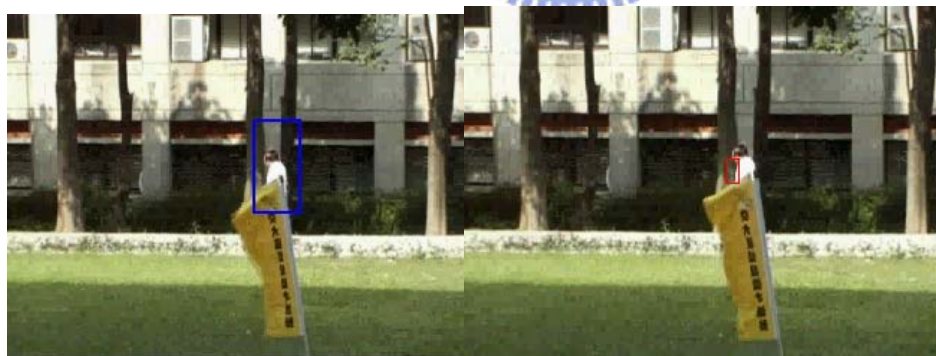
追蹤系統可以容忍部分遮蔽物的阻礙：雖然有黃色的旗子干擾，在干擾部分不大時，仍能執行追蹤功能。



frame 210 frame 212



frame 214 frame 216



frame 219 frame 221



frame 227 frame 229

圖 5.4：部分遮蔽物時的結果

樣版影像中背景部分的影像變動，可使系統重新偵測移動物體，以更新移動物體的資訊。實際執行三個影片，統計移動物體偵測與追蹤的執行時間與執行次數，完成下表：

	影片 1	影片 3	影片 2
總圖幅	1179	2391	2070
整體效能 fps	27.681	24.277	20.135
MD 總秒數	19.857	52.752	70.805
MD 圖幅	112	298	404
MD fps	5.640	5.649	5.706
TR 總秒數	22.735	45.738	32.002
TR 圖幅	1067	2093	1666
TR fps	46.932	45.761	52.059
MD: Motion Detection TR: Tracking			

表 5.1：移動物體偵測與追蹤的執行時間

在表 5.1 中我們先統計影片的總圖幅數與計算其整體效能，在此效能的計算單位為 fps(frames per second)，我們可以得到第一行及第二行的數據，再來針對影片中的移動物體偵測部分(MD)和追蹤部分(TR)分別統計其處理所耗費的時間與圖幅數，即 MD(/TR)總秒數和 MD(/TR)圖幅，計算其處理效能，得到 MD fps 與 TR fps 的數據。

由表 5.1 中可得移動物體偵測處理的速度為 5.6~5.7fps，平均位移追蹤處理的速度為 45~52fps，我們可以得知整體系統主要的載荷是在移動物體偵測系統的部分，因此若使整體系統的效能提升，移動物體偵測系統的執行速度會是個瓶頸，因為它必須得到整個影像的移動資訊，勢必經過很多的計算，因此在系統中，我們針對移動物體偵測系統的執行速度不夠快，利用將影像分為區塊來估計移動資訊，以增加執行速度，實驗結果顯示每秒處理 5.6~5.7 張影像，但尚未達到即時性的應用，為了能適應即時性的應用，我們在其後使用平均位移追蹤系統，實驗結果顯示每秒可處理 45~52 張影像，執行的速度相當快速，使整體系統可提升至每秒處理 20~27 張影像，藉由 Bhattacharyya Coefficient 判斷追蹤物體是否相似，來偵測追蹤錯誤，若追蹤錯誤，即回到移動物體偵測系統：



frame 414 frame 415



frame 418 frame 423



frame 426 frame 430



frame 433 frame 435



frame 436 frame 437

圖 5.5：連續影像實驗輸出結果

第六章

結論與未來工作

在此論文提出了一個系統對室外場景之下動態背景中的移動物體偵測及追蹤，在此系統中分為兩個子系統：移動物體偵測系統及追蹤系統。在移動物體偵測系統中，使用基本的影像前處理來消除雜訊與壓縮資料量，以連續影像相減法在執行速度上的優勢對連續的兩張影像做相減的程序，從而得到影像移動的部分，再將影像以區塊為基礎，計算其移動量得到可能移動的區塊，再將移動量大的區塊群集起來，其結果即為移動物體的區域；而在追蹤系統中，以移動物體偵測系統所得到的移動物體的區域為樣版，因中間畫素的資訊重要度較邊緣高，乘上核心函數賦予不同的權重，在之後的影像中，以平均位移演算法反覆疊代出相似度最高的平均位移向量，此結果便為物體移動後的位置，以達到追蹤的效果。



在整體系統的運作部分，移動物體偵測系統因其運算量大，執行速度較慢，是整個系統的瓶頸之處，而追蹤系統的平均位移演算法計算量少執行速度快，在此運用兩者各自的優勢組成整體系統；在偵測到移動物體後，即將移動物體的區域交由追蹤系統，藉由執行效率快速的平均位移演算法達成追蹤的效果，以 Bhattacharyya coefficient 來判斷是否追蹤錯誤，在 Bhattacharyya coefficient 過低，即相似度不高的情形下，將控制權交還給移動物體偵測系統，再度偵測移動物體的區域，交由追蹤系統追蹤之，如此反覆此運作模式，以提升整體系統的效能，適用於即時性的應用。

在此系統中仍存在些缺點，在有樹葉或草地的背景之下，易受到樹葉的雜訊而造成移動物體偵測的誤判，此項缺點可藉由前處理雜訊的抑制，使其影響下降，但若移動物體與樹葉太接近時，雜訊抑制的效果便會降低而造成誤判；在多個移動物體時，只能偵測到移動量較大的，若兩個移動物體太接近時，會被認為是同一個物體而造成誤判，這

是未來需要克服及改進的地方。移動物體偵測系統需要對整個畫面做運算，一直是偵測時的困難點，若能在此提高運算效能，可使整體系統效能提升，也可以再處理更大的影像畫面。



參考文獻

- [1] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, pp. 185-203, 1981.
- [2] J. Barron, D. Fleet, and S. Beauchemin, "Performance of Optical Flow Techniques," *International Journal of Computer Vision*, vol.12, no.1, pp.42–77, Jan. 1994.
- [3] A. M. Tekalp, "Digital video processing," *Prentice Hall PTR*, 1995.
- [4] A. J. Lipton, H. Fujiyoshi, and R. S. Patil, "Moving target classification and tracking from real-time video," in *Proc. of the IEEE Workshop on Applications of Computer Vision*, pp. 8–14, Oct. 1998.
- [5] C. Kim and J. N. Hwang, "A fast and robust moving object segmentation in video sequences," in *Proc. of the IEEE International Conference on Image Processing (ICIP'99)*, vol.2, Kobe, Japan, pp. 131–134, Oct. 1999.
- [6] 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.
- [7] A. H. S. Lai and N. H. C. Yung, "A fast and accurate scoreboard algorithm for estimating stationary backgrounds in an image sequence," in *Proc. of the IEEE International Symposium on Circuits and Systems*, vol. 4, pp. 241-244, 1998.
- [8] A.G. Nguyen and J. N. Hwang, "Scene context dependent key frame selection in streaming," in *Proc. of the 22nd International Conference on Distributed Computing Systems Workshops*, pp.208–213, Jul. 2002.
- [9] S. Gupte, O. Masoud, R. F. K. Martin, and N. P. Papanikolopoulos, "Detection and classification of vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 3, no. 1, pp. 37-47, Mar. 2002.
- [10] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," *International Journal of Computer Vision*, vol. 1, pp. 321–332, 1988.

- [11] N. Peterfreund, "Robust Tracking of Position and Velocity With Kalman Snakes", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 6, pp. 564-569, Jun. 1999.
- [12] D. Koller, J. Weber, T. Huang, J. Malik, G. Ogasawara, B. Rao, and S. Russell, "Towards robust automatic traffic scene analysis in real-time," in *Proc. of the 12th IAPR International Conference on Pattern Recognition*, vol. 1, pp. 126-131, 1994.
- [13] A. Chachich, A. Pau, A. Barber, K. Kennedy, E. Olejniczak, J. Hackney, Q. Sun, and E. Mireles, "Traffic sensor using a color vision method," in *Proc. of SPIE: Transportation Sensors and Controls: Collision Avoidance, Traffic Management, and ITS*, vol. 2902, pp. 156-165, 1996.
- [14] K. Schwerdt and J. L. Crowley, "Robust face tracking using color," in *Proc. of the 4th IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 90-95, Mar. 2000.
- [15] I. A. Karaulova, P. M. Hall, and A. D. Marshall, "A hierarchical model of dynamics for tracking people with a single video camera," in *Proc. of British Machine Vision Conference*, pp. 262-352, 2000.
- [16] C. J. Li, and S. J. Wang, "Detection and Tracking of a Single Deformable Object on an Active Surveillance Camera," in *Proc. Computer Vision, Graphics, and Image Processing, Kinmen, Taiwan*, Aug. 2003.
- [17] R. C. Gonzalez, and R. E. Woods, "Digital Image Processing," *Prentice Hall*, 2002.
- [18] M. Sonka, V. Hlavac, and R. Boyle, "Image Processing, Analysis, and Machine Vision," *PWS Publishing*, 1999.
- [19] I. Pitas, "Digital Image Processing Algorithms and Applications," *John Wiley & Sons*, 2000.
- [20] J.R. Parker, "Algorithms for Image Processing and Computer Vision," *John Wiley & Sons*, 1997
- [21] R. Szeliski, "Video Mosaics for Virtual Environments," *IEEE Computer Graphics and Applications*, vol. 16, no. 2, pp. 22-30, Mar. 1996.

- [22] J. W. Hsieh, "Moving Object Detection and Mosaic Construction by Image Stitching," *2003 National Computer Symposium*, pp.183-190, Dec. 18-19, 2003.
- [23] A. M. Tekalp, "Digital video processing," *Prentice Hall PTR*, 1995.
- [24] L. G. Shapiro, and G. C. stockman, "Computer vision," *Prentice Hall Inc.*, 2001.
- [25] J.R. Jain, and A.K. Jain, "Displacement measurement and its application in interframe image coding", *IEEE Transactions on Communications*, vol. 29, no. 12, pp. 1799-1808, Dec. 1981.
- [26] J. R. Corbera and D.L. Neuhoff, "On the optimal block size for block-based, motion compensated video coders," *SPIE Proceedings of Visual Communications and Image Processing*, vol. 3024, pp 1132-1143, Feb. 1997.
- [27] D. W. Scott, "Multivariate Density Estimation," *John Wiley & Sons, Inc.*, 1992.
- [28] D. Comaniciu, and P. Meer, "Mean Shift Analysis and Applications," *IEEE Intelligence Conference Computer Vision (ICCV'99)*, Kerkyra, Greece, pp. 1197-1203, 1999.
- [29] D. Comaniciu, and P. Meer, "Mean Shift: A Robust Approach toward Feature Space Analysis," *IEEE Transactions Pattern Analysis Machine Intelligence*, vol. 24, no. 5, pp. 603-619, 2002.
- [30] D. Comaniciu, and V. Ramesh, "Mean Shift and Optimal Prediction for Efficient Object Tracking, " *IEEE Intelligence Conference Image Processing (ICIP'00)*, Vancouver, Canada, vol. 3, pp. 70-73, 2000.
- [31] D. Comaniciu, and V. Ramesh, "Robust Detection and Tracking of Human Faces with an Active Camera," *IEEE Transactions Pattern Analysis Machine Intelligence*, vol. 24, no. 5, pp. 603-619, 2002.
- [32] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-Based Object Tracking, " *IEEE Transactions Pattern Analysis Machine Intelligence*, vol. 25, no. 5, pp. 564-575, 2003.
- [33] T. Kailath, "The Divergence and Bhattacharyya Distance Measures in Signal Selection," *IEEE Transactions Communications Technology*, vol. 15, pp. 52-60, 1967.

- [34] K. Fukunaga, and L.D. Hostetler, "The estimation of the gradient of a density function, with applications in pattern recognition," *IEEE Trans. Information Theory*, vol. 21, pp. 32-40, 1975.
- [35] Y. Cheng, "Mean Shift, Mode Seeking, and Clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 8, pp. 790-799, Aug., 1995.
- [36] 張龍生, "動態場景中的人物之偵測與追蹤," 大同大學電機工程研究所碩士論文, 2004.

