

國立交通大學

資訊學院資訊科技（IT）產業研發碩士班

碩 士 論 文

建構軌跡式之監控影像資料庫索引系統

Constructing a trajectory-based indexing system
for a database of surveillance images

研 究 生：吳忠衛

指導教授：莊仁輝 教授

中 華 民 國 九 十 七 年 六 月

建構軌跡式之監控影像資料庫索引系統

Constructing a trajectory-based indexing system
for a database of surveillance images

研 究 生：吳忠衛

Student：Chung-Wei Wu

指導教授：莊仁輝

Advisor：Jen-Hui Chuang



Industrial Technology R & D Master Program on
Computer Science and Engineering

June 2008

Hsinchu, Taiwan, Republic of China

中華民國九十七年六月

建構軌跡式之監控影像資料庫索引系統

學生：吳忠衛

指導教授：莊仁輝 教授

國立交通大學資訊學院產業研發碩士班

摘 要

本論文研製之目的是對監控影像資料庫建構以物體移動軌跡為基礎的索引系統。首先利用蒙恬公司所提供的監控影像偵測技術，擷取出資料庫影片中各前景物體的外接矩形(Bounding Box)。接著考慮各物體間交會狀況提出一套追蹤方法，對物體在連續兩張影像畫面的外接矩形進行追蹤，依時間順序建構出物體的移動軌跡，並考量多個物體移動軌跡會有分離與合併的狀態。因此我們索引系統引用上述連續外接矩形所組成的軌跡與其附加屬性來作為搜尋的依據。最後我們採用友善的圖形化搜尋介面，來幫助使用者對資料庫中物體移動軌跡的索引進行不同查詢方式。而實驗的結果說明本論文所提出的索引系統在不同的監控場景下，仍可有效進行概略的篩選，濾除大部份不相關的資料，讓使用者可快速找到所需的影片片段。

Constructing a trajectory-based indexing system for a database of surveillance images

Student : Chung-Wei Wu

Advisor : Dr. Jen-Hui Chuang

Industrial Technology R & D Master Program of
Computer Science College
National Chiao Tung University

ABSTRACT

The goal of this thesis is to construct a trajectory-based indexing system for a database of surveillance images. For each image sequence in the database, moving objects are first identified with a PenPower software, by finding a bounding box for each object. The object trajectories are then established with proposed tracking scheme by considering intersections of the bounding boxes between two consecutive image frames, which also takes into account merging and splitting of trajectories. Accordingly, the proposed indexing system uses various attributes of the sequence of bounding boxes associated with a trajectory as searching clues. Finally, a friendly graphic/text user interface is implemented to facilitate various forms of queries of an object trajectory. Experimental results show that the proposed indexing system works effectively in terms of a rough screening of object trajectories in different surveillance environments.

致謝

本論文得以順利完成，首先必須要感謝我的指導教授莊仁輝老師。在這段日子裡，老師不僅教導我在學術上的知識，更教導我一個研究生應有的研究態度及解決問題的方法。再者必須要感謝蒙恬科技公司與徐繼聖處長，無論在監控場景資料庫計劃合作或論文上，都予以最大的支持與幫助，使我可以順利完成這些工作。也非常謝謝未來的公司新眾電腦與林洽民董事長，在學期間不斷鼓勵我努力向學，並提供獎助金來予以栽培與支持。

在兩年裡的日子，感謝博班學長泓宏、國華的細心指導，讓我在研究上可以更上一層樓。也謝謝實驗室的學弟妹、學長姐們及隔壁實驗室所有伙伴，由於你們帶給大家的歡樂，使得實驗室的氣氛一直很和樂，使得在研究的生活並不會孤單及煩悶。更謝謝同屆的同學證中、寧玄、諾白與俊宏，由於大家的相互扶持及鼓勵，使得大家得以一起順利畢業。

最後，感謝一直以來支持我、為我操心的家人，以及所有幫助過我的朋友們，謹以此篇表達我誠摯的謝意。

目錄

第一章 緒論	1
1.1. 研究動機與目的.....	1
1.2. 相關研究.....	2
1.3. 系統流程.....	3
1.4. 本論文各章節介紹.....	4
第二章 監控影像資料庫系統介紹	5
2.1. 硬體設備.....	5
2.2. 軟體架構與操作介面.....	6
2.3. 資料庫內容.....	9
2.3.1. 影片名稱編碼規則.....	9
2.3.2. 視訊特徵.....	10
2.3.3. 各場景視訊檔內容.....	11
第三章 以蒙恬 TRACKIN 前景物偵測功能為基礎建構軌跡式索引	13
3.1. 蒙恬“TRACKIN 智慧型全域追蹤監控系統”簡介.....	13
3.2. 使用代表前景物體的外接矩形進行軌跡的追蹤.....	15
3.2.1. 分析區塊重疊來追蹤.....	15
3.2.2. 利用顏色直方圖來輔助追蹤.....	18
3.2.3. 單一物體移動.....	19
3.2.4. 多個物體移動.....	21
3.3. 雜訊偵測與附加軌跡屬性.....	24
3.3.1. 雜訊偵測（Detecting Noise）	25
3.3.2. 對實際物體的移動軌跡附加屬性.....	26
3.4. 圖形化使用者搜尋介面.....	28
第四章 實驗結果與討論	31
4.1. 物體各移動型態之追蹤結果.....	31
4.1.1. 單一物體移動.....	31
4.1.2. 多個物體移動.....	32
4.2. 追蹤失誤原因.....	38
4.3. 雜訊偵測問題.....	39
第五章 結論與未來工作	41
參考文獻	42

圖目錄

圖 1-1：系統流程圖。	3
圖 2-1：硬體設備。	5
圖 2-2：彩色攝影機。	5
圖 2-3：紅外線攝影機。	6
圖 2-4：監控影像資料庫系統架構。	6
圖 2-5：使用者之操作介面。	7
圖 2-6：依視訊內容的關鍵特徵篩選影片的搜尋功能。	7
圖 2-7：搜尋“室外-大樓門口”條件所篩選出的影片一覽表。	7
圖 2-8：管理者之操作介面（一）- 對影片的索引進行新增、修改、刪除功能。	8
圖 2-9：管理者之操作介面（二）- 對視訊的特徵進行新增、修改、刪除功能。	8
圖 2-10：視訊檔名稱編碼	9
圖 2-11：一般彩色各場景的視訊內容。	11
圖 2-12：夜間紅外線各場景的視訊內容。	12
圖 3-1：TrackIN 系統架構。	14
圖 3-2：TrackIN 偵測過程與結果。	14
圖 3-3：外接矩形(Bounding Box)的資料格式。	15
圖 3-4：對外接矩形進行追蹤找出物體的移動軌跡。	15
圖 3-5：區塊重疊示意圖。	16
圖 3-6：單一物體移動狀況。(a)~(c)分別是該影片第 6、42 和 64 張畫面。	16
圖 3-7：使用區塊重疊方法追蹤圖 3-6 結果。	17
圖 3-8：多個物體移動-人與狗發生交會狀況。(a)~(c)分別是該影片中第 115、139 和 210 張畫面。	17
圖 3-9：使用區塊重疊方法追蹤圖 3-8 結果，人與狗的移動軌跡會混在一起。	17
圖 3-10：擷取代表物體的外接矩形內所有像素建立顏色直方圖。	18
圖 3-11：追蹤 (a)、(b)、(c)、(d) 單一物體移動，產生(e)移動軌跡。	20
圖 3-12：多個物體移動下的兩種狀態。	21
圖 3-13：多個物體在分離狀態產生的移動軌跡(續)。	22
圖 3-14：多個物體移動交會成為合併狀態產生的移動軌跡。	22
圖 3-15：多個物體交會後分開狀態產生的移動軌跡。	23
圖 3-16：偵測軌跡是否為背景雜訊並附加屬性。	25
圖 3-17：因光線變化被 TrackIN 所偵測的狀況。	25
圖 3-18：極座標角度圖。	27
圖 3-19：圖形化使用者搜尋介面。	28
圖 3-20：手繪軌跡輸入的搜尋操作。	29
圖 3-21： T_{user} 與 T_{db} 軌跡相似度比較。	30
圖 4-1：物體單一移動的情況。(a), (b), ..., (f)分別為影片中第 17、29、47、67、	

85 和 103 張畫面。.....	31
圖 4-2：單一物體移動之追蹤結果。.....	32
圖 4-3：偵測為背景雜訊之部分結果。.....	32
圖 4-4：多個物體移動（Case 1）。(a), (b), (c)分別為影片中第 34、61 和 75 張畫面。.....	32
圖 4-5：多個物體移動（Case 1）追蹤所得結果。.....	33
圖 4-6：多個物體移動（Case 1）偵測為背景雜訊之部分結果。.....	33
圖 4-7：多個物體移動（Case 2）。(a), (b), (c)分別為影片中第 38、43 和 46 張畫面。.....	33
圖 4-8：多個物體移動（Case 2）追蹤所得結果。.....	34
圖 4-9：多個物體移動（Case 2）偵測為背景雜訊之部分結果。.....	34
圖 4-10：多個物體移動（Case 2）。(a), (b), (c)分別為影片中第 115、139 和 210 張畫面。.....	34
圖 4-11：多個物體移動（Case 2）追蹤所得結果。.....	35
圖 4-12：多個物體移動（Case 2）偵測為背景雜訊之部分結果。.....	35
圖 4-13：多個物體移動（Case 3）。(a), (b), (c)分別為影片中第 30、38 和 53 張畫面。.....	35
圖 4-14：多個物體移動（Case 3）追蹤所得結果。.....	36
圖 4-15：多個物體移動（Case 3）偵測為背景雜訊之部分結果。.....	36
圖 4-16：多個物體移動（Case 4）。(a), (b), (c)分別為影片中第 7、22 和 67 張畫面。.....	37
圖 4-17：多個物體移動（Case 4）追蹤所得結果。.....	37
圖 4-18：多個物體移動（Case 4）偵測為背景雜訊之部分結果。.....	37
圖 4-19：TrackIN 偵測前景物體發生問題的情形。(a), (b), ..., (i)分別為影片中第 18、22、25、26、27、32、33、34 與 36 張畫面。.....	38
圖 4-20：追蹤失誤結果：單一物體移動卻有三個移動軌跡。.....	39
圖 4-21：樹枝搖晃之背景物擾動。.....	39
圖 4-22：玻璃造成的反光現象。.....	40
圖 4-23：雜訊偵測之問題。.....	40

第一章 緒論

1.1. 研究動機與目的

近年來由於科技進步與硬體技術突破，促使監控攝影機成本降低且體積更小，而可廣泛裝設於都市大樓、街道、角落與重要公共設施中。又因電腦和網路的快速發展，促進了監控影像的數位化，使得所拍攝的資料日益龐大，“該如何”藉由軟體更有智慧的自動管理這些龐大資料已成為趨勢。

而本實驗室與蒙恬科技公司合作，共同建構了一個監控影像資料庫，提供各場景日間彩色及夜間紅外線的監控影像資料。為了讓使用者有效率的使用此資料庫，在本篇論文中，提出了一套以前景物的移動軌跡為主的搜尋方法，藉由蒙恬所提供的監控影像偵測技術為基礎，假設已得知前景物體的外接矩形(Bounding Box)條件下，對物體在不同情況下的移動狀態進行分析，並依其移動狀態將這些外接矩形按照時間順序建構出移動軌跡，作為監控影像資料庫的索引。

而在搜尋介面上，由於傳統文字介面為主的查詢對於使用者而言，比較無法直觀找出所需要的資訊。因此我們採用手繪輸入圖形的搜尋介面，讓使用者在畫面上繪出欲搜尋的移動軌跡，再與資料庫建構好的索引比對後，即可找到最相似的內容，如此一來使用者能更便於操作此系統。

此外由於我們資料庫僅使用外接矩形、座標、移動軌跡等基本資訊來描述物體，可以廣泛包含不同場景的影片，來供給其他數位監控系統來評估該系統在物體偵測（object detection）、物體追蹤（object tracking）、物體分類（object classification）與事件辨識（event recognition）等各功能的優缺[13]。此外還可與其他系統評測結果作一個比較對照，藉此我們希望能建立出開放的共同檢測平台，進而促進數位監控產業的升級與發展。

1.2. 相關研究

對於影像資料庫而言，一般可分為以單一張畫面得到的圖片格式，或由多張連續畫面組成的影片格式來做為儲存單位。而在影像的內容檢索上，常見一種是以空間的視覺特徵為主，如顏色（color）、紋理（texture）和外形（shape）等，通常用於單張畫面的圖片搜尋上，其資料庫系統目前有QBIC、VisualSEEK、PhotoBook和Virage等[1]。而另一種則是物體在時間軸上的變化特徵，通常以物體的移動軌跡（trajectory）來表示，而VideoQ[2]就是一個提供依此特徵對連續畫面的影片進行搜尋之資料庫系統。

在數位監控技術中如行人行為辨識[3][4]、交通安全監控[5]等都是利用物體的移動軌跡來做擷取、比對與應用，而監控影像資料庫也常以軌跡作為索引，如行人移動行為資料庫，以行人移動路徑為基礎[6]，藉由訓練資料建立行人的常用路徑，依此來查詢有多少行人經過或知道此行人在哪一路徑上，並判斷是否為異常行為。而在交通監控資料庫上，會先以小區塊為單位計算出行人、車輛的移動向量，濾除掉雜訊的移動向量後，再將這些移動向量整合成一總向量代表物體的移動方向與位置[7]，最後透過軌跡比對尋找出最相似的前景物體。

在建立索引值之後，接著是如何對物體的移動軌跡進行比對，達到資料庫查詢的功能。在軌跡比對上時間長度（duration）、取樣頻率（frame rates）與時間間距（time difference）各因素都會影響比較結果，當在相同時間長度、取樣頻率與時間間距下，通常會取兩軌跡在各時間的位置之總差距[10]來做比對，但此方法在兩軌跡具相同時間長度卻有不同取樣頻率（different frame rates）或是不同規則的時間間距（uneven time difference）就無法使用。上述狀況可使用面積差距公式[11]來加以克服，但假若兩軌跡對應之時間長度不相同（different duration），就得使用隱藏馬可夫模型（Hidden Markov Model）的方法[12]來進行軌跡比對。

1.3. 系統流程

圖 1-1 為系統流程圖。首先取出監控影像資料庫內的連續視訊資料，經過蒙恬公司的追蹤監控系統 TrackIN 偵測，產生代表前景物體的外接矩形，接著對這些外接矩形進行追蹤，建立出前景物的移動軌跡後，再對這些軌跡做雜訊偵測。濾除雜訊後剩下的實際物體，我們將對該物體的移動軌跡附加上移動狀態、移動方向、移動速度等屬性來簡單描述此一物體行為，並作為監控影像資料庫的索引。最後透過圖形化的使用者搜尋介面，藉由手繪軌跡方式輸入查詢條件，來提供使用者快速、方便的對前述所建立的索引進行搜尋並取得監控影像資料庫的所需視訊片段。

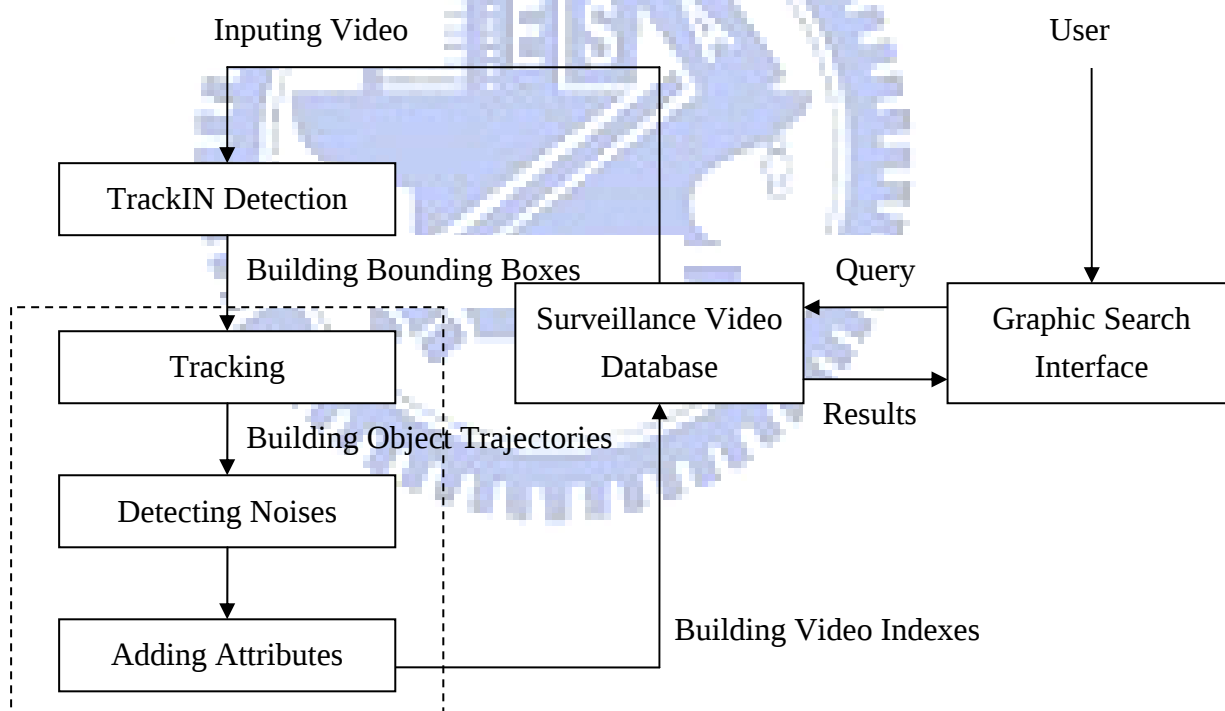


圖 1-1：系統流程圖。

1.4. 本論文各章節介紹

本篇論文全文共分為五章，除了本章為緒論外，第二章介紹監控影像資料庫系統的架構與目前所收集到日間彩色及夜間紅外線各場景的影片資料。第三章則是說明對代表前景物的外接矩形進行追蹤的方法，以及對追蹤所建立物體的移動軌跡如何做雜訊偵測，並在濾除雜訊後對實際物體的移動軌跡附加屬性，最後透過圖形化軌跡搜索介面來實現資料庫的查詢。而第四章是對單一物體移動、多個物體移動等各種狀態實驗的結果，以及探討追蹤失誤的原因與雜訊偵測的問題。最後第五章為結論並說明本研究未來的一些研究方向。



第二章 監控影像資料庫系統介紹

本章將介紹本實驗室與蒙恬公司所合作開發的監控影像資料庫系統，我們將在2.1節先簡介開發使用的硬體設備，在2.2節介紹系統架構與軟體操作，最後於2.3節說明資料庫制定規則與目前所儲存收集的資料內容。

2.1. 硬體設備

在本系統中，我們所使用的硬體設備由攝影機、3 公尺高腳架、類比轉數位視訊轉接器、PC 等四部份所組成（圖 2-1）。系統運作是攝影機所拍攝到的類比訊號，經由視訊轉接器轉換成數位視訊資料，再傳至 PC 上觀看或儲存。其中攝影機包含了拍攝一般行人的彩色攝影機（圖 2-2(a)）、針對快速車輛具電子快門攝影機（圖 2-2(b)）與用於拍攝夜間場景的近紅外線攝影機（圖 2-3(a)、(b)）。而本系統工作環境是以校園室內、室外場景為主，並使用高度腳架模擬真實監控攝影機所拍攝的角度。

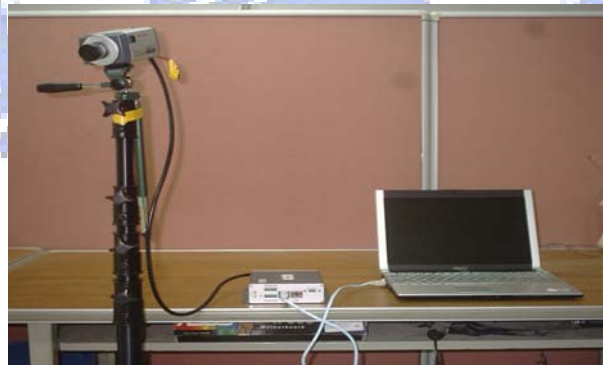


圖 2-1：硬體設備。



(a) CCD Camera (不具電子快門)



(b) CCTV Camera (具電子快門可拍快速移動物體)

圖 2-2：彩色攝影機。



(a) IR Camera KMS-63A3-61T (彩色顯示)



(b) IR Camera (灰階顯示)

圖 2-3：紅外線攝影機。

2.2. 軟體架構與操作介面

本系統是以 Office Acces2003 軟體作為開發環境，以方便資料庫系統擴充。圖 2-4 為監控影像資料庫系統架構，其功能包含了僅能進行資料查詢的使用者權限，與可增刪影片索引與視訊特徵的管理者權限兩部分。

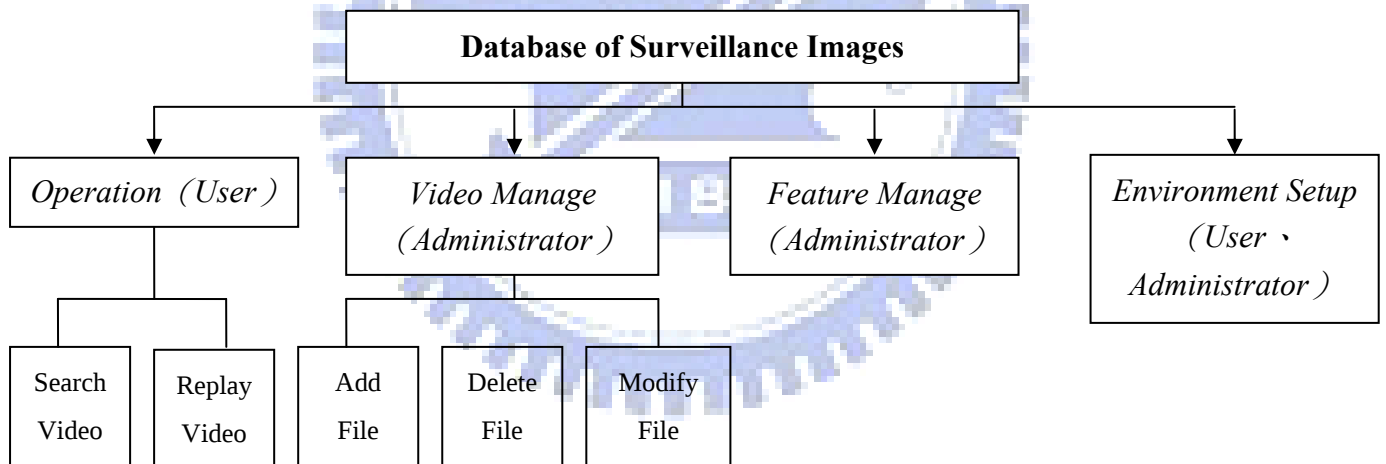


圖 2-4：監控影像資料庫系統架構。

圖 2-5 為使用者權限的資料庫系統操作介面，使用者可先從“監控影像資料一覽”表來瀏覽所有的影片，並從其中點選所需的影片。而操作介面會列出所點選影片其視訊內容的關鍵特徵，並顯示該影片的摘要畫面，讓使用者能夠快速知道此影片的概略內容，而不用逐一開啟播放。其次使用者還可進一步透過搜尋功能（圖 2-6），輸入所需要的視訊特徵當做條件，即可篩選出如圖 2-7 具有該特徵的影片。

監控影像資料明細

建構編號	1	開始目標物存在	無
影片檔名	CEle0B001	光源變化	NP
攝影型態	彩色 (720x480)	搖晃干擾	NP
目標物	行人	其他干擾	反光輕微
場景	室內-電梯門口	影像摘要畫面	
視角距離	中視角		
影片長度	短 (< 1Min)		
Max 目標物數量	單獨 (1)		
Min 目標物數量	單獨 (1)		

記錄: 1 之 103

監控影像資料一覽

Index	VName	Camera	Object	Scene	View	Time
1	CEle0B001	彩色 (720x480)	行人	室內-電梯門口	中視角	短 (< 1Min)
2	CEle0B002	彩色 (720x480)	行人	室內-電梯門口	中視角	短 (< 1Min)
3	CEle0B003	彩色 (720x480)	行人	室內-電梯門口	中視角	短 (< 1Min)
4	CEle0B004	彩色 (720x480)	行人	室內-電梯門口	中視角	短 (< 1Min)
5	CEle0B005	彩色 (720x480)	行人	室內-電梯門口	中視角	短 (< 1Min)
6	CEle0B006	彩色 (720x480)	行人	室內-電梯門口	中視角	短 (< 1Min)
7	CEle0B007	彩色 (720x480)	行人	室內-電梯門口	中視角	短 (< 1Min)

圖 2-5：使用者之操作介面。

監控影像資料搜尋

影片檔名		開始存在目標物	
攝影型態		光源變化	
目標物		搖晃干擾	
場景		其他干擾	
視角距離	室內-房間門口 室內-電梯門口 室內-地下室門口 室外-大樓門口 室外-餐廳門口 室外-宿舍門口 室外-進出大門口 室外-停車場門口		
影片長度			
Max 目標			
Min 目標			

✓

圖 2-6：依視訊內容的關鍵特徵篩選影片的搜尋功能。

監控影像資料明細

建構編號	101	開始目標物存在	無
影片檔名	CBui0C007	光源變化	漸進變化
攝影型態	彩色 (720x480)	搖晃干擾	NP
目標物	行人	其他干擾	反光輕微
場景	室外-大樓門口	影像摘要畫面	
視角距離	中視角		
影片長度	長(10Min~30Min)		
Max 目標物數量	大量 (> 6)		
Min 目標物數量	單獨 (1)		

記錄: 7 之 8 (篩選)

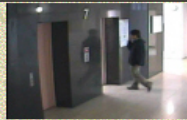
監控影像資料一覽

Index	VName	Camera	Object	Scene	View	Time
97	CBui0C003	彩色 (720x480)	行人	室外-大樓門口	中視角	長(10Min~30Mi
98	CBui0C004	彩色 (720x480)	行人	室外-大樓門口	中視角	長(10Min~30Mi
99	CBui0C005	彩色 (720x480)	行人	室外-大樓門口	中視角	長(10Min~30Mi
100	CBui0C006	彩色 (720x480)	行人	室外-大樓門口	中視角	長(10Min~30Mi
101	CBui0C007	彩色 (720x480)	行人	室外-大樓門口	中視角	長(10Min~30Mi
102	CBui0C008	彩色 (720x480)	行人	室外-大樓門口	中視角	長(10Min~30Mi

圖 2-7：搜尋“室外-大樓門口”條件所篩選出的影片一覽表。

而具管理者權限的操作介面，還會比使用者多出對影片的索引與視訊的特徵進行新增、修改、刪除功能（圖 2-8、圖 2-9），以便對資料庫系統來進行管理與維護。

監控影像資料明細

建檔編號	1	開始目標物存在	無
影片檔名	CEle0B001	光源變化	N/A
攝影型態	彩色 (720x480)	搖晃干擾	N/A
目標物	行人	其他干擾	地板牆壁反光
場景	室內-電梯門口	行為模式	正常行進
視角距離	中視角	影像摘要畫面	
影片長度	短 (< 1Min)		
Max 目標物數量	單獨 (1)		
Min 目標物數量	單獨 (1)		

監控影像資料一覽

	VName	Camera	Object	Scene	View	Time	Max Amount
▶	CEle0B001	彩色 (720x480)	行人	室內-電梯門口	中視角	短 (< 1Min)	單獨 (1)
	CEle0B002	彩色 (720x480)	行人	室內-電梯門口	中視角	短 (< 1Min)	單獨 (1)
	CEle0B003	彩色 (720x480)	行人	室內-電梯門口	中視角	短 (< 1Min)	單獨 (1)
	CEle0B004	彩色 (720x480)	行人	室內-電梯門口	中視角	短 (< 1Min)	少量 (2~3)
	CEle0B005	彩色 (720x480)	行人	室內-電梯門口	中視角	短 (< 1Min)	少量 (2~3)
	CEle0B006	彩色 (720x480)	行人	室內-電梯門口	中視角	短 (< 1Min)	中量 (4~6)
	CEle0B007	彩色 (720x480)	行人	室內-電梯門口	中視角	短 (< 1Min)	少量 (2~3)

記錄: 1 之 205

圖 2-8：管理者之操作介面（一）- 對影片的索引進行新增、修改、刪除功能。

特徵內容增刪

特徵種類: Camera

特徵內容

Feature
▶ 彩色 (720x480)
紅外線 (720x480)
高畫質 (1024x480)

記錄: 1 之 3

圖 2-9：管理者之操作介面（二）- 對視訊的特徵進行新增、修改、刪除功能。

2.3. 資料庫內容

本資料庫系統是以監控為目的來建構各場景影像資料，為了達到此一要求，目前是以“門口”為目標來進行拍攝與蒐集資料，接下來將介紹資料庫系統如何分類管理所建構的影片。其中 2.3.1 是說明如何對影片名稱進行編碼，而 2.3.2 是說明各影片的視訊特徵內容，以利快速搜尋所需的影像，最後 2.3.3 是介紹目前資料庫系統所儲存的各類場景影片。

2.3.1. 影片名稱編碼規則

為了有效管理檔案讓資料庫更容易搜尋，我們對影片名稱提出了一套規則來管理。圖 2-10 為影片名稱編碼，其編碼共分為五類，以下對這些分類來詳細說明。分類 1 代表攝影機類型，“C”表示彩色影像，“I”表示紅外線影像。分類 2 代表場景，取場景英文前三字母縮寫，如“Elv”表示電梯門口場景，而“Bld”則表示大樓門口場景。分類 3 代表相同場景但不同地點的區別，如“0”表示在 A 大樓門口所拍攝影片，而“1”表示在 B 大樓門口所拍攝影片，其餘依此類推。分類 4 代表相同場景且同一地點但不同視角的區別，如“A”表示近視角，而“B”表示中視角。分類 5 代表影片數量編號，相同場景且同一地點、視角的影片數量最多可到 999。

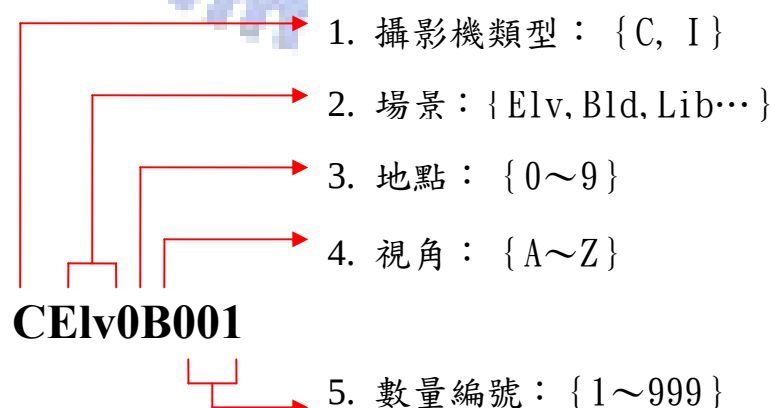


圖 2-10：視訊檔名稱編碼

2.3.2. 視訊特徵

視訊特徵是在影像連續播放過程中，將重要表徵、容易識別景物或特殊狀況下用人工方式擷取出來，以作為資料庫系統分類、搜尋的依據。本資料庫系統對視訊特徵也具有擴充功能，未來針對不同的目標物、攝影機或新場景，都可以繼續擴增特徵來描述。以下是目前資料庫系統已建立的視訊特徵項目：

- (1) 攝影機型態 (Camera)：彩色 (720×480)、紅外線 (320×240)。
- (2) 目標物 (Object)：行人、車輛、人車混合。
- (3) 場景 (Scene)：室內—房間門口、室內—電梯門口、室外—大樓門口、室外—餐廳門口、室外—圖書館門口、室外—進出大門口。
- (4) 視角 (View)：近視角、中視角、遠視角。其視角是以“門口”高度佔整畫面百分比來做區分，如“近視角”是門口高度佔滿或超過整畫面，“中視角”是門口高度佔整畫面 3/4~1/2，而“遠視角”是門口高度佔整畫面 1/2 以下。由於近視角拍攝會有個人隱私權問題，無法取得大量資料，因此在本資料庫系統中都以中、遠視角影片做為資料來源。
- (5) 影片長度 (Time)：短 (< 1Min)、中 (1Min~5Min)、長 (10Min~30Min)。
- (6) Max 目標物數量 (Max Amount)：單獨 (1)、少量 (2~3)、中量 (4~6)、大量 (>6)。
- (7) Min 目標物數量 (Min Amount)：單獨 (1)、少量 (2~3)、中量 (4~6)、大量 (>6)。
- (8) 起始影片畫面是否有目標物存在 (Initialization)：無、有 (畫面開始有目標物出現)。
- (9) 光源變化 (Light)：無、快速變化、漸進變化、昏暗。其中“快速變化”是指畫面中有出現瞬間閃光的狀況，“漸進變化”是由白天到傍晚或傍晚到晚上，光源緩慢變化狀況，而“昏暗”是指在室內未開燈或夜晚時採光不足狀況。
- (10) 搖晃干擾 (Shake)：無、有 (背景搖晃、鏡頭搖晃)。

(11) 其他干擾 (Noisy)：無、有 (地板牆壁反光、玻璃反光、水面反光)。

(12) 行為模式 (Behavior)：正常行進、異常行進、撿放遺留物。

2.3.3. 各場景視訊檔內容

首先我們介紹現階段本資料庫系統所建構一般彩色的視訊內容，其中各場景所拍數量如圖 2-11(a)~(f)所示：室內—電梯門口場景大約 40 部、室內—房間門口場景大約 55 部、室外—大樓門口場景大約 35 部、室外—圖書館門口場景大約 40 部、室外—餐廳門口場景大約 50 部、室外—進出大門口場景大約 20 部。



圖 2-11：一般彩色各場景的視訊內容。

另外圖 2-12 (a)~(c)所顯示的，是夜間紅外線的視訊內容，其中各場景所拍攝的數量如下：室外一大樓門口場景（灰階顯示）大約 25 部、室外一圖書館門口場景（灰階顯示）大約 35 部、室內-大廳場景（彩色顯示）大約 5 部。目前除了仍持續拍攝以“行人”為主的內容來擴充資料量，之後重點是針對“車輛”作為拍攝目標，使的本資料庫的資料能更加多元化，用途更加廣泛。



(a) 室外一大樓門口場景（灰階顯示）



(b) 室外一圖書館門口場景（灰階顯示）



(c) 室內-大廳場景（彩色顯示）

圖 2-12：夜間紅外線各場景的視訊內容。

第三章 以蒙恬 TrackIN 前景物偵測功能 為基礎建構軌跡式索引

在本章中我們會應用蒙恬 TrackIN 程式偵測出資料庫影片中前景物，並對此前景物進行追蹤，以建立物體的移動軌跡作為該影片索引。在 3.1 節會先對蒙恬公司所提供“TrackIN 智慧型全域追蹤監控系統”作簡介，並以此系統的偵測技術為基礎，偵測出影像資料庫系統中影片裡所有前景物體，並用個別的外接矩形(Bounding Box)來表示。接著 3.2 節將說明如何以這些外接矩形為基礎，在連續的視訊中進行追蹤，以建立出物體的移動軌跡，做為資料庫的索引。而後，在 3.3 節會進一步做雜訊偵測處理，來區分背景雜訊與實際物體。若真的是實際物體，就再附加移動狀態、移動方向與移動速度等基礎屬性，讓資料庫影片中的物體描述除了外接矩形外，還有行為說明與實際數值來加強完整性。最後我們將在 3.4 節介紹如何透過圖形化使用者搜尋介面，提供使用者快速、便利找到所需的視訊片段內容。

3.1. 蒙恬 “TrackIN 智慧型全域追蹤監控系統” 簡介

蒙恬 “TrackIN 智慧型全域追蹤監控系統” 是結合高速球型攝影機 (Speed Dome Camera)、與數位視訊錄影機 (Digital Video Recorder) 而成的全自動影像追蹤系統，並具有人臉辨識、行人追蹤等功能。系統結構是以分析各影像資料的伺服器主機 (Analytic Server) 為核心，加上一個高速球型攝影機與三個監控攝影機 (Monitor Cameras) 所組成 (如圖 3-1)。此系統可主動分析多支監控攝影機的監視，一旦發現可疑的人車出現時，系統會立即驅動高速球型攝影機對可疑人車進行放大畫面與錄影，以擷取遠端物體的近照或臉部特寫並持續追蹤記錄動線與軌跡。而對於區域內可疑份子的臉部特寫，更運用了人臉分析引擎 (Face Metrixa)，提供快速錄影檔案及搜尋功能，以最高效率爭取用戶需要的影像片段。

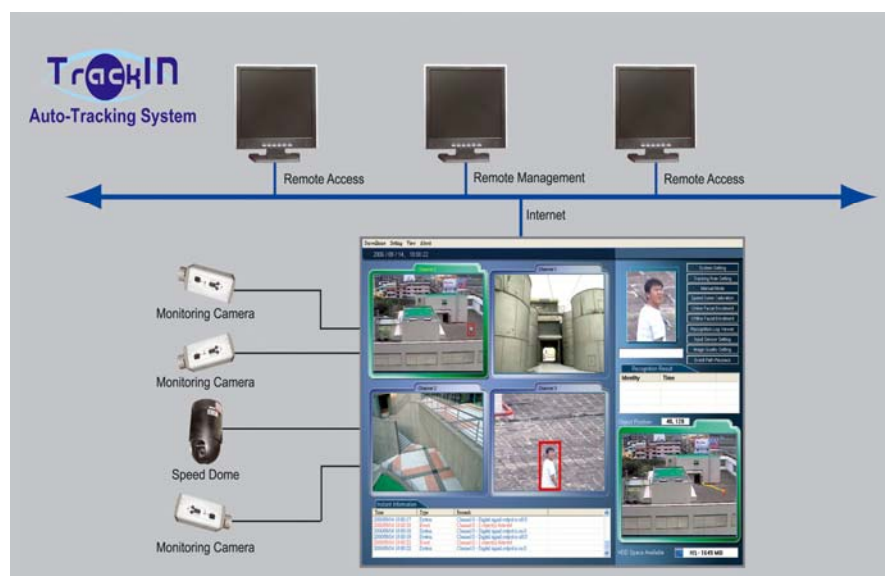


圖 3-1：TrackIN 系統架構。

在 TrackIN 中前景物偵測的技術是對影像資料庫中影片作前置處理，如圖 3-2(a) 是利用綠色的外接矩形(Bounding Box)，來代表所偵測到的前景物體。當分析完一段影片後，將每一時間得到的外接矩形全部顯示如圖 3-2(b)。此處我們所使用外接矩形資料格式為（圖 3-3）：

$$B = (x, y, w, h) \quad (3-1)$$

x 、 y 為外接矩形左上角的起始座標，而 w 、 h 為矩形的寬度、高度，這些資料均以像素為單位。



(a) 偵測到的前景物體用綠色的外接矩形框住 (b) 每一時間得到的外接矩形依時間順序全部顯示

圖 3-2：TrackIN 偵測過程與結果。

下一節中，藉由追蹤找出具時間順序的外接矩形集合，我們將使用這些外接矩形來代表物體的移動軌跡。

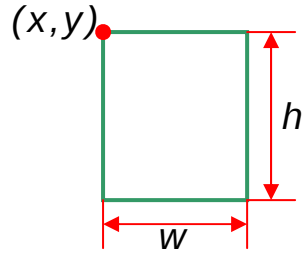


圖 3-3：外接矩形(Bounding Box)的資料格式。

3.2. 使用代表前景物體的外接矩形進行軌跡的追蹤

接著我們希望找出物體的移動軌跡，來描述一段影片中各物體的具體行為，以作為資料庫的索引。因此，我們使用代表前景物體的外接矩形進行追蹤，建立具時間順序的外接矩形集合—軌跡 T (如圖 3-4 所示)。

$$T = \{B_{t-1}, B_t, B_{t+1}, \dots\} \quad (3-2)$$

其中 B_{t-1} 為時間 $t-1$ 的外接矩形, B_t 為時間 t 的外接矩形以此類推。而我們將分析(1)區塊之重疊與(2)顏色直方圖，來對外接矩形“進行追蹤”，以建立出各物體的移動軌跡，以下是(1)、(2)的內容說明。

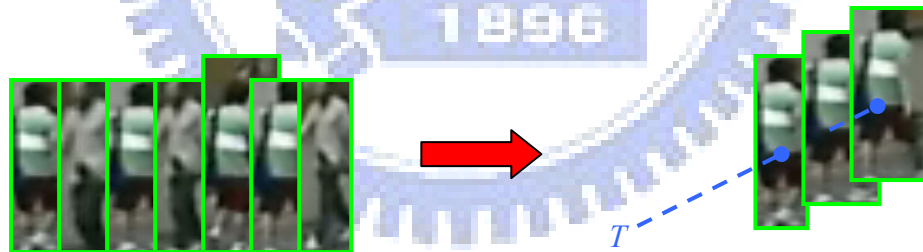


圖 3-4：對外接矩形進行追蹤找出物體的移動軌跡。

3.2.1. 分析區塊重疊來追蹤

在一連續視訊中，假設物體的移動幅度不大，則在前後兩張影像某些前景區塊必有些許的重疊 (圖 3-5)，藉由此特性就可偵測物體在下一張影像中可能的位置。而判斷兩張影像的外接矩形是否同一物體的方式是以兩個矩形交集 (Intersection) 除以聯集 (Union) 所佔的重疊比率來決定：

$$Overlap\ ratio = \frac{B_{t-1} \cap B_t}{B_{t-1} \cup B_t} > Th_{overlap} \quad (3-3)$$

在本論文中是對重疊比率（Overlap ratio）取一個門檻值 $Th_{overlap} = 10\%$ ，超過此值才代表同一物體。

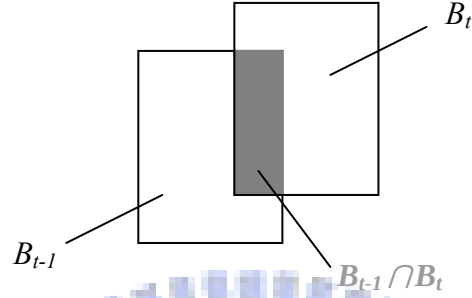


圖 3-5：區塊重疊示意圖。

利用區塊重疊方法就可對物體進行追蹤，進而建立該物體的移動軌跡，這對於處理如圖 3-6 單一物體移動的簡單狀況下，其追蹤結果顯然不錯，如圖 3-7 所示。圖中藍色方塊代表外接矩形的中心點，青色連線代表物體移動的路徑，而外接矩形顏色由淺至深變化，代表物體的移動方向。然而當兩個以上物體在移動過程中常會發生交會遮蔽(Occulsion)的複雜狀況，由於多個物體已重疊在一起（圖 3-8），使區塊重疊方法分不出差異性，而誤認都是同一物體，導致追蹤結果出現錯誤，如圖 3-9 所示。圖中人與狗的移動軌跡原本應各自獨立，但因區塊重疊追蹤誤判，無法分出人與狗是不同物體，造成人與狗移動軌跡混在一起。



(a)

(b)

(c)

圖 3-6：單一物體移動狀況。(a)~(c)分別是該影片第 6、42 和 64 張畫面。



圖 3-7：使用區塊重疊方法追蹤圖 3-6 結果。

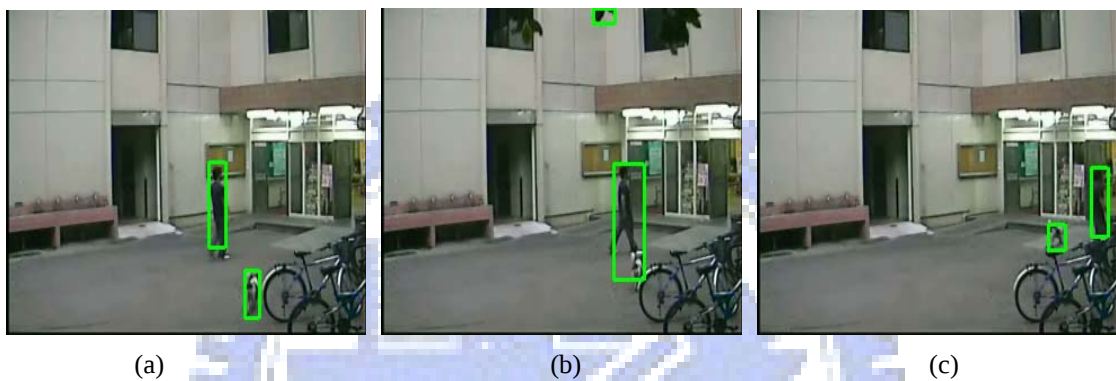


圖 3-8：多個物體移動-人與狗發生交會狀況。(a)~(c)分別是該影片中第 115、139 和 210 張畫面。



圖 3-9：使用區塊重疊方法追蹤圖 3-8 結果，人與狗的移動軌跡會混在一起。

有鑒於一般監控影像的畫面通常都會有多個物體出現，假若我們想要更廣泛的將我們追蹤演算法應用在更多不同的監控場景上，就必須能處理物體發生交會的狀況。因此在下一小節中，我們將進一步取得物體的顏色特徵，並統計成直方圖的形式(Histogram)，以期能識別各物體的差異。

3.2.2. 利用顏色直方圖來輔助追蹤

為了能處理上述交會狀況，我們使用直方圖(Histogram) [8]來記錄各前景物顏色的分佈，並對顏色直方圖持續更新藉以識別不同物體的差異。首先我們對物體的顏色特徵來建立直方圖 H ，其作法是將外接矩形範圍內各像素的顏色數值，對應到 RGB 三維空間顏色分佈的位置來累計數量。而由於一般監控攝影機所拍攝影像畫面其 RGB 每一維空間大約只用到 4~5 bits 來表示，利用此一特性我們可以將原本 RGB 顏色表示從 8 bits 變成 4 bits，因此只用了 $2^{12} = 4096$ bins，可以達到節省儲存空間、加快運算的優點。圖 3-10 是擷取代表物體的外接矩形內所有像素來建立顏色值方圖，其中橫軸是 0~4095 個 bins，而縱軸是各像素在同一 bin 的累加數。

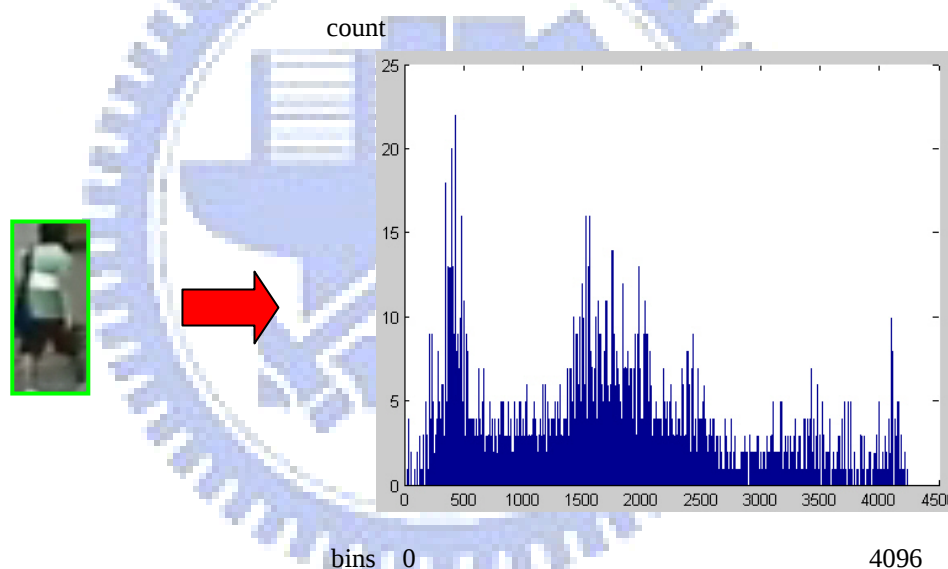


圖 3-10：擷取代表物體的外接矩形內所有像素建立顏色直方圖。

由於我們是對代表物體的外接矩形範圍來計算顏色直方圖，導致直方圖不單只有前景物顏色特徵，也會包含到部份的背景，而混入了背景的顏色特徵。不過由於 TrackIN 大部分所偵測到的前景物體都能被外接矩形足夠緊密的框住，使的背景所佔比例相對於前景物體少許多，因此在實驗過程中其顏色直方圖仍可有效區別不同前景物體的差異性。

接下來我們將說明如何使用(1)區塊重疊與(2)顏色直方圖對單一物體移動、多個物體移動等各種型態來進行處理。

3.2.3. 單一物體移動

對於連續視訊中只有單一物體移動的情況（圖 3-11(a)、(b)、(c)、(d)），我們首先利用區塊重疊找出在下一時間的外接矩形，接著統計目前與下一時間外接矩形的顏色直方圖，並使用 Swain and Ballard 的直方圖交集公式(Histogram Intersection)[9]計算兩直方圖差異來辨識其相似度，其公式的表示如下：

$$HDist(H_{t-1}, H_t) = \sum_{k=1}^n \frac{\min(H_{t-1}(k), H_t(k))}{H_{t-1}(k)} \quad (3-4)$$

其中 H_{t-1} 是前一時間外接矩形的顏色直方圖， H_t 是目前時間外接矩形的顏色直方圖， n 則代表在 RGB 空間的顏色分佈範圍 4096 bins，而差異度 $HDist$ 之數值是介於 0 到 1 之間。

若前一時間外接矩形的顏色直方圖與目前外接矩形的顏色直方圖差異不大時，就代表追蹤到該物體的移動位置，並將找到的外接矩形 B_t 加入此物體的移動軌跡 T 中。同時為了避免物體因移動造成位置改變而被誤認不是同一個，因此對顏色直方圖持續進行更新動作，以維持同一物體在不同位置的顏色特徵。其更新公式為：

$$H'_t = \beta H_{t-1} + (1 - \beta) H_t, \quad 0 < \beta < 1 \quad (3-5)$$

其中 H_{t-1} 是前一時間外接矩形的顏色直方圖，而 H_t 是目前時間找到外接矩形的顏色直方圖。此外 β 值影響更新率，在這裡使用 $\beta = 0.8$ 。圖 3-11(e) 顯示最後產生之該物體的移動軌跡。



(a)



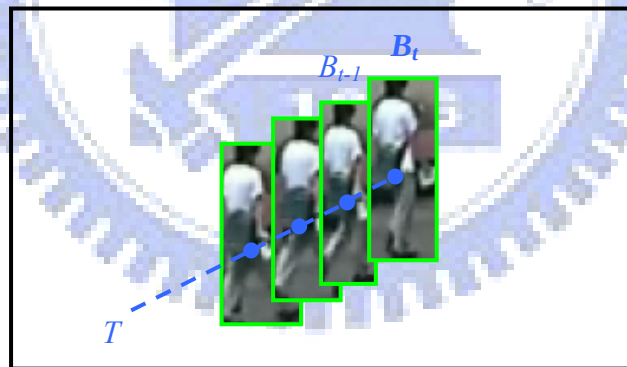
(b)



(c)



(d)



(e)

圖 3-11：追蹤 (a)、(b)、(c)、(d) 單一物體移動，產生(e)移動軌跡。

3.2.4. 多個物體移動

若考慮比較複雜的場景，假設有兩個以上物體出現在畫面中，且有發生交會狀況，此時我們會將這個過程分成兩種狀態（圖 3-12）：多個物體各自獨立移動之分離狀態（N1：No Occlusion），多個物體發生交會成為合併狀態（N2：Occlusion），以下分別對各種狀態及其變化來詳細說明。

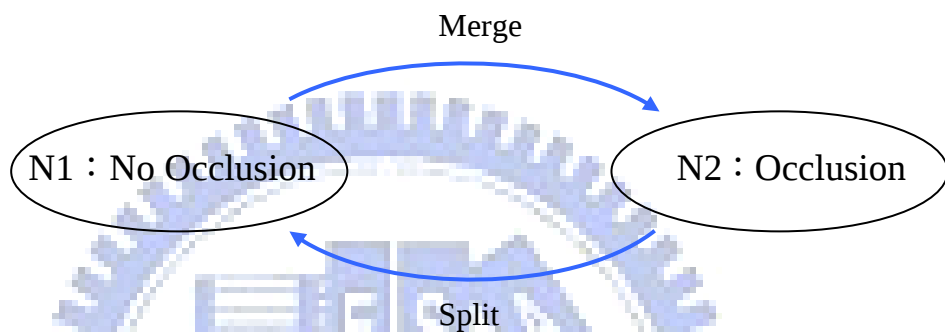


圖 3-12：多個物體移動下的兩種狀態。

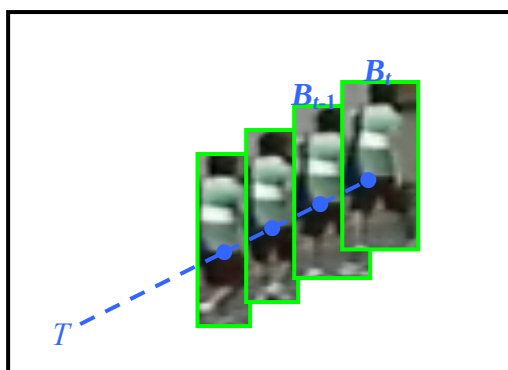
N1：分離狀態

當多個物體一開始出現在畫面中各自移動，且沒有發生交會狀況下（圖 3-13 (a)），此時將數個物體視為個別單一物體的情況來處理，而產生圖 3-13 (b)、(c) 各物體的移動軌跡。

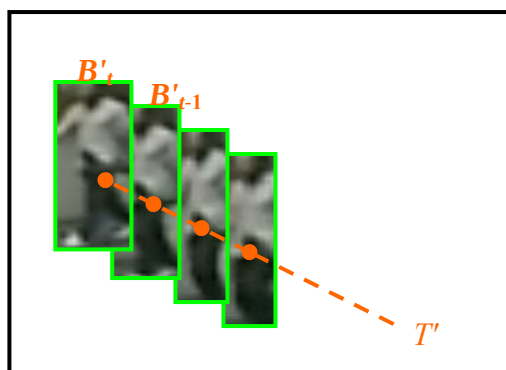


(a) 多個物體在分離狀態

圖 3-13：多個物體在分離狀態產生的移動軌跡。



(b) 物體 A 的移動軌跡

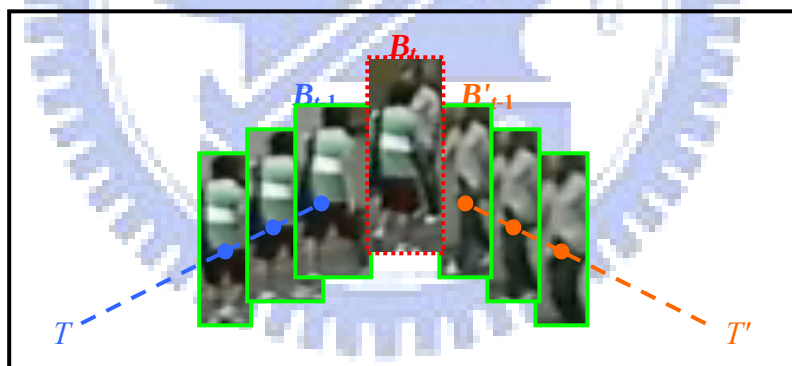


(c) 物體 B 的移動軌跡

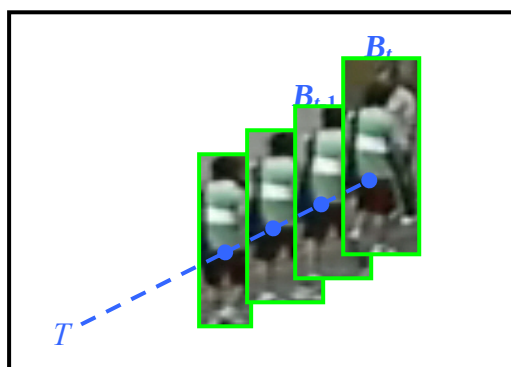
圖 3-13：多個物體在分離狀態產生的移動軌跡(續)。

N1→N2：交會成為合併狀態

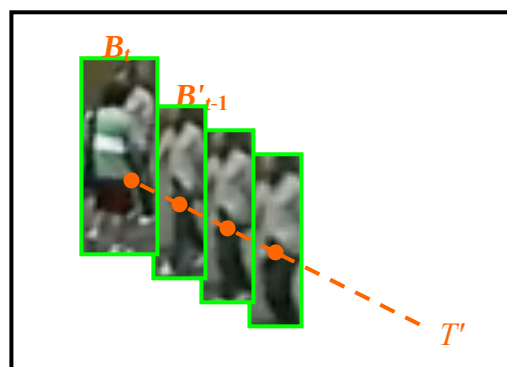
若多個物體發生交會狀況，如圖 3-14(a) 所示，當 $B_t \cap B'_t = \emptyset$ 、 $B_{t-1} \cap B_t \neq \emptyset$ 、 $B'_{t-1} \cap B'_t \neq \emptyset$ ，表示前一時間外接矩形 B_{t-1} 、 B'_{t-1} 所各代表的物體發生交會，因此將目前時間的外接矩形 B_t 加入以上有發生交會的物體移動軌跡 T 、 T' 中。但由於 B_t 混雜了多個物體，為了維持該物體原本的顏色特徵，因此我們對各物體的顏色直方圖不進行更新的動作，產生圖 3-14 (b)、(c) 各物體的移動軌跡。



(a) 多個物體交會成為合併狀態



(b) 物體 A 的移動軌跡

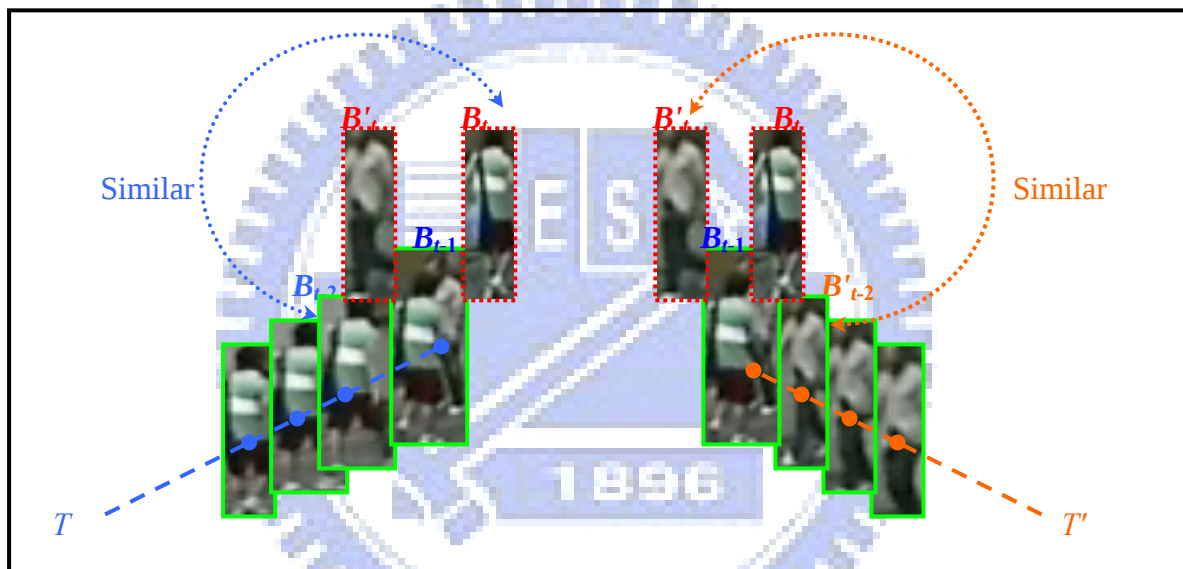


(c) 物體 B 的移動軌跡

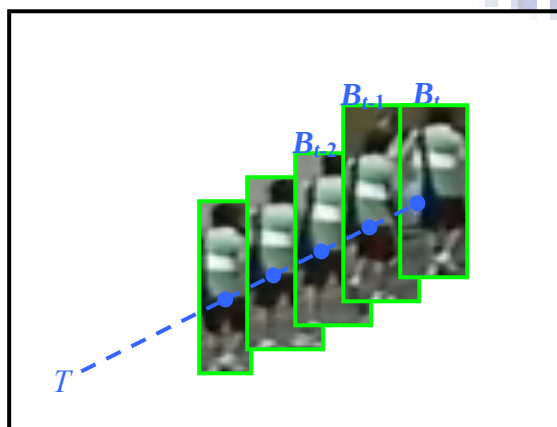
圖 3-14：多個物體移動交會成為合併狀態產生的移動軌跡。

N2→N1：交會後分開狀態

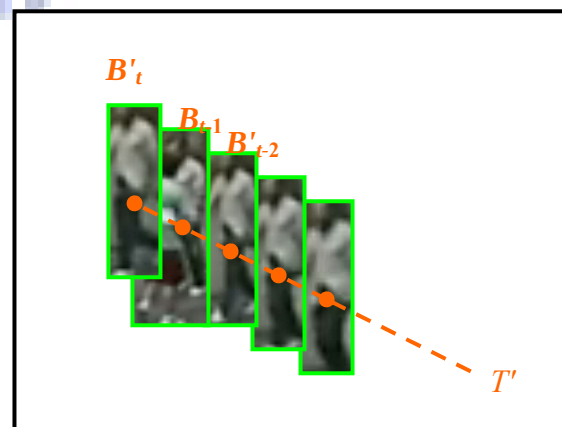
而在發生交會後各物體又分開情況下，以圖 3-15(a) 為例物體 A 的移動軌跡 T 其目前時間的外接矩形會有兩個以上 (B_t, B'_t)，此時我們可藉由顏色直方圖交集公式比較 B_{t-2} 與 B_t 、 B_{t-2} 與 B'_t ，並找出哪組顏色直方圖差異最小，則該組中目前時間的外接矩形即是我們所要追蹤的物體。當知道物體 A、B 的移動軌跡在目前時間是要追蹤哪一個物體後，接下來就同單一物體移動的處理方式，分別將追蹤到物體加入各物體的移動軌跡 T, T' 中，並進行該物體之顏色直方圖更新的動作，最後產生圖 3-15 (b)、(c) 各物體的移動軌跡。



(a) 多個物體交會後分開狀態



(b) 物體 A 的移動軌跡



(c) 物體 B 的移動軌跡

圖 3-15：多個物體交會後分開狀態產生的移動軌跡。

而多個物體移動會產生幾種由上述不同狀態及其變化情形，以下分別說明之。

Case 1 : N1

多個物體從頭至尾移動中都未發生交會狀況，可視同多個單一物體移動情況。

Case 2 : N1→N2→N1

如圖 3-15 順序包含所有狀態，多個物體在開始時是各自獨立移動，後來發生交會狀況，並在交會後各物體會分開各自移動。

Case 3 : N1→N2

多個物體一開始雖各自移動，但發生交會後多個物體就重疊合併在一塊，一起移動直到離開畫面都不會分開。

Case 4 : N2→N1

多個物體一開始就處於合併狀態，移動一段時間後再分開各自移動。對於此種情況，由於一開始多個物體就重疊在一塊而被視為單一物體，等到這些物體分開時，因為多個物體合併時與各自單獨的顏色直方圖有所差異而可能不會繼續追蹤，造成多個物體只有合併時的移動軌跡，而無分開後各物體的移動軌跡。因此我們會對交會後分開卻未被追蹤到的物體，重新追蹤來產生一個新的軌跡，避免上述移動軌跡遺失的問題。

而下一節會對物體所追蹤到的移動軌跡做雜訊偵測，區分出哪些為背景雜訊，哪些是實際物體的移動，並對實際物體的移動軌跡附加移動狀態、移動方向、移動速度等屬性，來描述各物體的移動行為。

3.3. 雜訊偵測與附加軌跡屬性

由上節對各物體追蹤得到由外接矩形集合而成的移動軌跡後，我們會先對此軌跡進行雜訊偵測，若不是背景雜訊就視為實際物體的移動軌跡，並附加移動狀態、移動方向與速度等屬性（圖 3-16）。

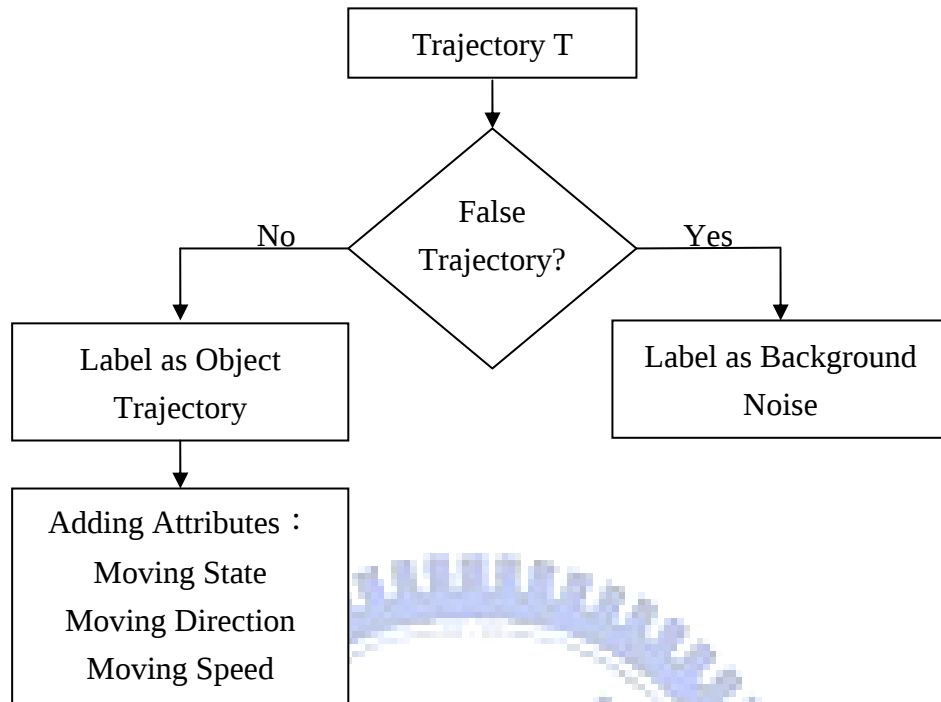


圖 3-16：偵測軌跡是否為背景雜訊並附加屬性。

3.3.1. 雜訊偵測 (Detecting Noise)

在使用 TrackIN 偵測時，由於有些光線變化或干擾（圖 3-17），都會被視為前景物體而產生外接矩形，為了避免這些背景雜訊被誤認為是實際物體的移動軌跡，我們將進行雜訊偵測，將這些背景雜訊濾除。

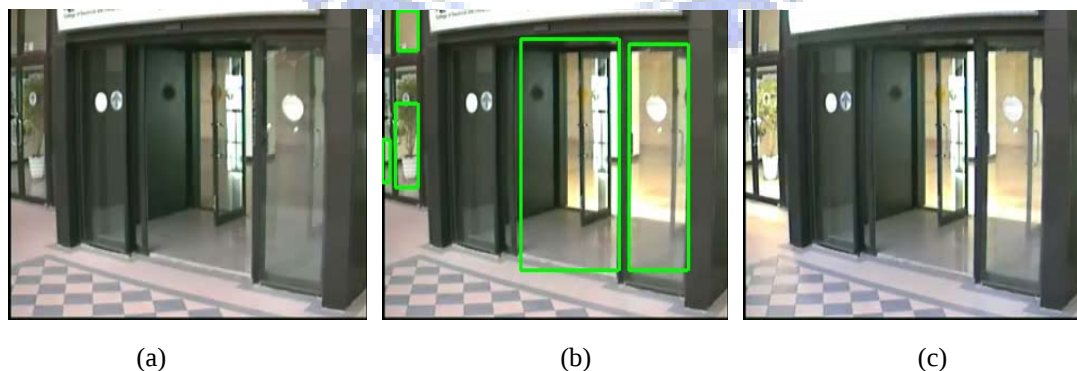


圖 3-17：因光線變化被 TrackIN 所偵測的狀況。

在實驗中，對這些光線變化產生的雜訊經觀察後發現有以下特性：

- (1) 通常位置變動不大。
- (2) 常出現一下又馬上消失，不具長時間連續性。

因此我們將以這些特性做為偵測軌跡是否為背景雜訊的依據。首先定義物體的外接矩形質心為：

$$P_i = (C_{ix}, C_{iy}) = (x_i + \frac{w_i - 1}{2}, y_i + \frac{h_i - 1}{2}), 1 \leq i \leq N \quad (3-6)$$

其中 N 為此軌跡 T 內所包含外接矩形的總數，而 x_i 、 y_i 、 w_i 、 h_i 見圖 3-3。再來計算軌跡 T 內所有外接矩形的質心位置平均值與變化量。

$$\mu_x = \frac{\sum_{i=1}^N C_{ix}}{N}, \quad \mu_y = \frac{\sum_{i=1}^N C_{iy}}{N} \quad (3-7)$$

$$\sigma_x = \frac{1}{N} \sqrt{\sum_{i=1}^{N-1} C_{ix} - \mu_x}, \quad \sigma_y = \frac{1}{N} \sqrt{\sum_{i=1}^{N-1} C_{iy} - \mu_y} \quad (3-8)$$

然後依上述雜訊的特性(1)、(2)點針對 $\sigma_{\text{positionX}}$ 、 $\sigma_{\text{positionY}}$ 與 N 來進行辨別，假若 $\sigma_{\text{positionX}}$ 、 $\sigma_{\text{positionY}}$ 變動不大以及 N 太少不具長時間連續性，就將此軌跡視為背景雜訊

3.3.2. 對實際物體的移動軌跡附加屬性

在濾除背景雜訊後，其他剩下的物體之移動軌跡就被視為實際物體的移動軌跡，我們將附加移動狀態、移動方向與移動速度等屬性來描述物體的移動行為。其中移動狀態的屬性我們是依照 3.2 節中分類的方法，區分成單一物體移動、多個物體移動（Cases 1 – 4）來描述。

其次有關移動方向之屬性是以八個方向來表示。首先算出物體移動的軌跡中從起點到終點的總向量，如下所示：

$$Direction_x = \sum_{i=1}^{N-1} (C_{i+1,x} - C_{ix}) = C_{Nx} - C_{1x} \quad (3-9)$$

$$Direction_y = \sum_{i=1}^{N-1} (C_{i+1,y} - C_{iy}) = C_{Ny} - C_{1y} \quad (3-10)$$

之後再依圖 3-18 所示各角度範圍分成八個方向，來描述該物體的移動方向。

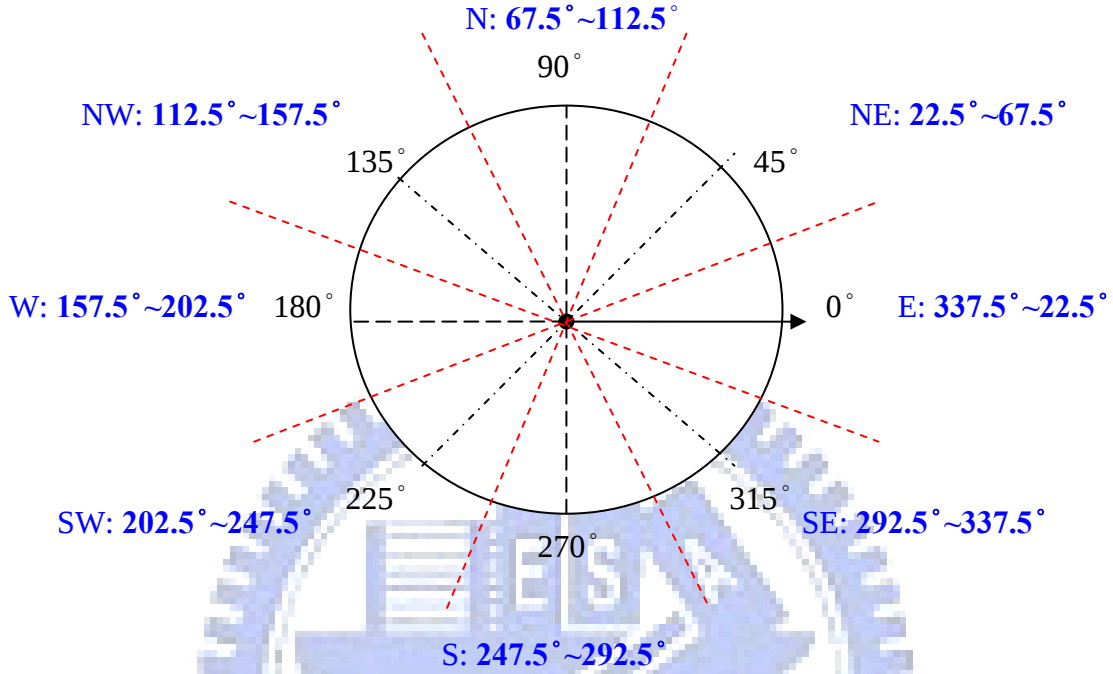


圖 3-18：極座標角度圖。

而移動速度屬性其計算方式如下所示：

$$Speed = \left(\frac{\sum_{i=1}^{N-1} \sqrt{(C_{i+1,x} - C_{ix})^2 + (C_{i+1,y} - C_{iy})^2}}{t} \right) \quad (3-11)$$

其中 t 為物體移動的總時間。在實驗中我們得知一般行人移動速度大約在 8 pixel/sec （以行人大小約佔影像畫面 1/2~3/4 為基準的狀況下），因此當算出的平均速度超過 8 pixel/sec，我們就判定此物體移動速度為快，而小於 8 pixel/sec 就描述為慢。

有了以上附加的屬性，我們即可進一步知道監控影像資料庫中各影片有哪些物體以及該物體的相關行為，並透過下一節所介紹的圖形化使用者搜尋介面，來讓使用者更快速、方便取得資料庫內所需要的視訊片段。

3.4. 圖形化使用者搜尋介面

經 3.2 與 3.3 節對監控影像資料庫內影片建立物體的移動軌跡作為索引並附加屬性後，我們即可使用圖 3-19 更直觀的圖形化使用者介面，來提供給使用者快速、便利依索引來搜尋所需的影像內容。而搜尋的操作是以手繪軌跡來作為輸入，其步驟是先按下介面中“Search Trajectory Mapping”按鍵，接著在右上方影片的摘要視窗上逐次按下滑鼠左鍵點選座標點（圖 3-20(a)），慢慢形成使用者指定的軌跡 T_{user} 。當確定輸入完畢後，只要連按滑鼠左鍵兩下系統即會開始執行搜尋，並會列出與手繪輸入軌跡在門檻值 $Th_{distance}$ 範圍內所有相似的結果（圖 3-20(b)）。若想觀看個別的結果，可點選右下方表格內的各列的軌跡，並顯示於右上方的視窗中（圖 3-20(c)、(d)）。



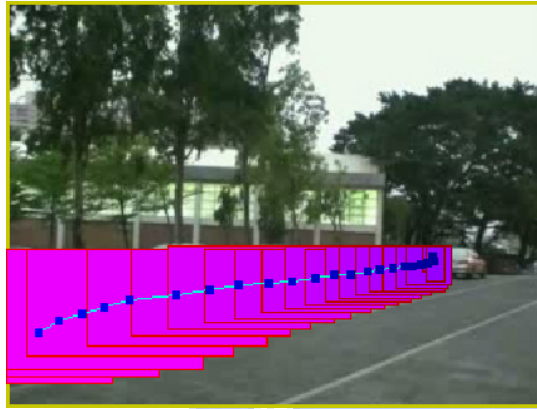
圖 3-19：圖形化使用者搜尋介面。右下方表格是代表一段影片內所包含物體的移動軌跡，而右上方視窗是顯示出目前所選擇到的物體其移動軌跡，其移動方向顏色是由淺至深表示。當想瀏覽該物體移動的視訊片段內容，點選“Show Trajectory Video”按鍵就會顯示在左上方視窗提供給使用者觀看。



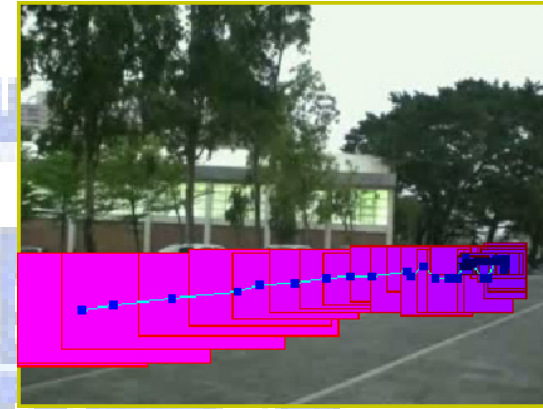
(a) 手繪輸入欲搜尋軌跡



(b) 相似於手繪輸入軌跡的所有結果



(c) 物體 A 的運動軌跡



(d) 物體 B 的運動軌跡

圖 3-20：手繪軌跡輸入的搜尋操作。

介紹完整個介面的操作後，接下來詳細說明使用者輸入手繪軌跡後，系統是如何與資料庫索引軌跡比對，搜尋到最相似的軌跡。其搜尋的方法是以資料庫中索引軌跡各座標點與手繪軌跡最近距離總差距來進行比對。

剛開始由使用者逐次點選的座標點其兩點之間會自動產生連線，當連續輸入三個座標點 P_1 、 P_2 、 P_3 時，如圖 3-21 所示，就會產生線段 $\overline{P_1P_2}$ 、 $\overline{P_2P_3}$ 來連結各點，最後形成使用者手繪軌跡 T_{user} 。而 T_{db} 是資料庫已建立的索引軌跡，其中 P'_1 、 P'_2 、 P'_3 、 P'_4 為外接矩形質心座標點。當進行搜索比對時，則會先從 T_{db} 各質心座標點分別找出距離 T_{user} 線段 $\overline{P_1P_2}$ 、 $\overline{P_2P_3}$ 最近的各點 u_1 、 u_2 、 u_3 、 u_4 ，再計算出 Euclidean 距離 d_1 、 d_2 、 d_3 、 d_4 ，最後將這些距離差加總並取平均，可得到 T_{user} 、 T_{db} 軌跡比對的相似度，其計算公式為 (3-12)。

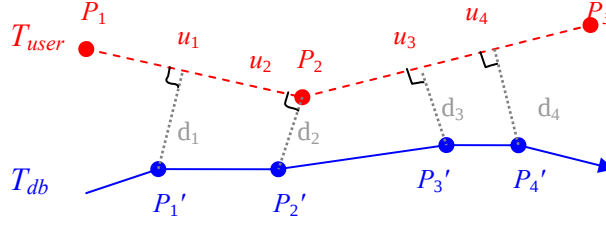


圖 3-21： T_{user} 與 T_{db} 軌跡相似度比較。

$$TDist\ Ratio = \frac{\sum_{i=1}^N d_i}{N} < Th_{distance} \quad (3-12)$$

其中 N 為 T_{db} 座標點總數，而因為我們的目的是想把一段影片內，所有相似於使用者所輸入的軌跡的影片片段通通搜尋出來，而不是只找最相似的，所以我們依 (3-12) 設一門檻值 $Th_{distance}$ 來協助使用者濾除掉不需要的視訊內容，找出有相關的部份。

對於所搜尋出來的影片片段，我們會對這些片段進一步分析與篩選，針對移動軌跡 T_{user} 與 T_{db} 間其移動方向與移動速度屬性來判斷相似度，計算公式如 (3-13) 和 (3-14) 所示：

$$|Direction\ Angle_{T_{db}} - Direction\ Angle_{T_{user}}| < 45^\circ \quad (3-13)$$

$$|Speed_{T_{db}} - Speed_{T_{user}}| < \varepsilon \quad (3-14)$$

其中當資料庫的索引軌跡與使用者手繪軌跡之移動方向相差 45° 以內，就判定兩軌跡方向相似。而在速度相似度的判定上，由於用手繪方式所點選的軌跡速度計算不易掌握，因此我們在 (3-14) 中設 ε 為一範圍變數，可由使用者自行調整。初始值 $\varepsilon = 5\text{ pixel/sec}$ ，當資料庫的索引軌跡與使用者手繪軌跡之移動速度相差此值以內，就判定兩軌跡速度相似。

第四章 實驗結果與討論

本章將對各種場景下不同物體的移動情況來進行實驗，並討論結果不理想的問題與原因。在 4.1 節中，我們對影像資料庫中不同物體的移動情況利用 3.2 節追蹤的方法，依單一物體移動、多個物體移動各型態進行實驗，並顯示其追蹤結果。在 4.2 節探討各結果之優缺與錯誤原因，而 4.3 節說明雜訊偵測時會影響準確度的問題為何。

4.1. 物體各移動型態之追蹤結果

4.1.1. 單一物體移動

我們對一段影片裡只有單一物體移動的情況（圖 4-1）來進行追蹤，可得到圖 4-2(a)由右往左移動的行人 A 與圖 4-2(b)由左往右移動的行人 B 各移動軌跡之結果，而偵測出為背景雜訊的部份結果在圖 4-3 中。

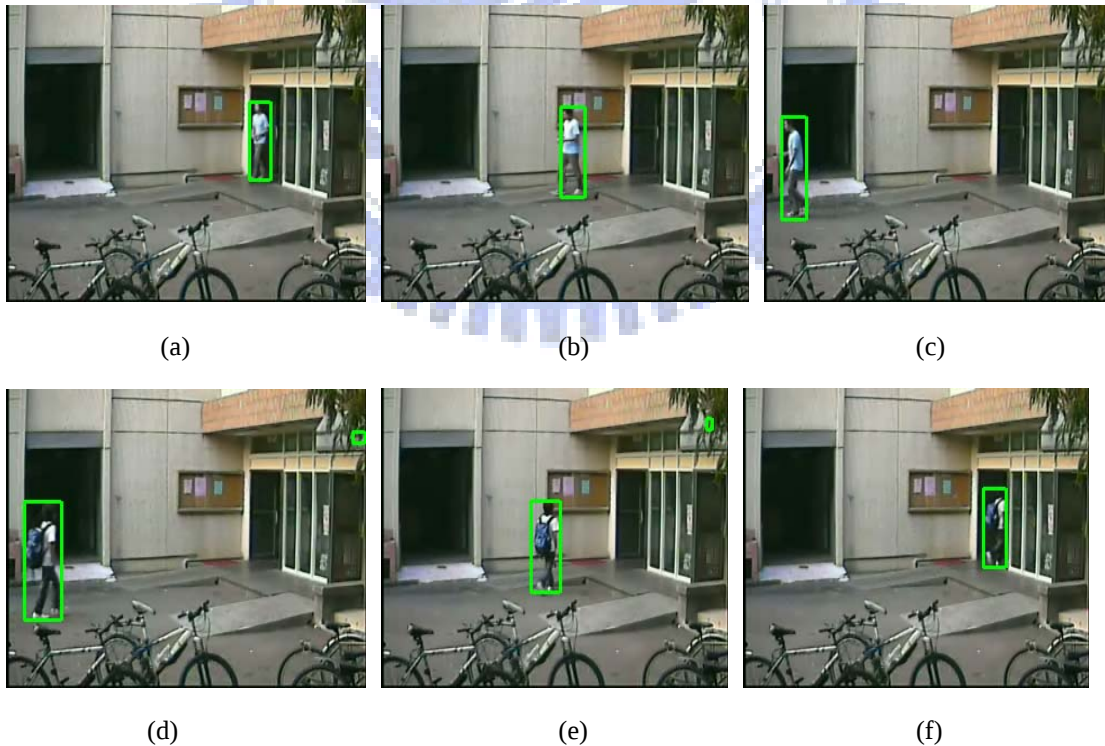
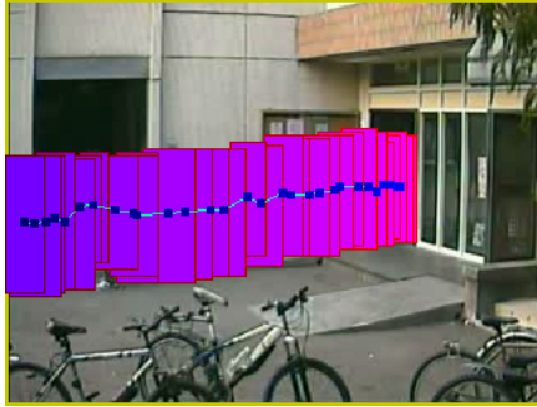
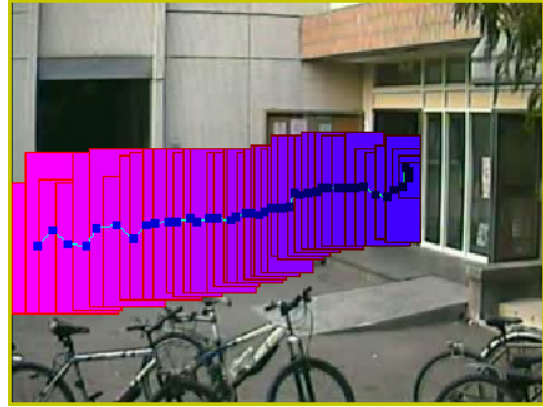


圖 4-1：物體單一移動的情況。(a), (b), ..., (f)分別為影片中第 17、29、47、67、85 和 103 張畫面。



(a) 由左往右移動的行人 A 之移動軌跡：移動方向 10.26° ，移動速度 6.53 pixel/frame



(b) 由右往左移動的行人 B 之移動軌跡：移動方向 185.36° ，移動速度 6.78 pixel/frame

圖 4-2：單一物體移動之追蹤結果。



(a) 畫面 2-5



(b) 畫面 88-90



(c) 畫面 96-97

圖 4-3：偵測為背景雜訊之部分結果。

4.1.2. 多個物體移動

Case 1 (N1)

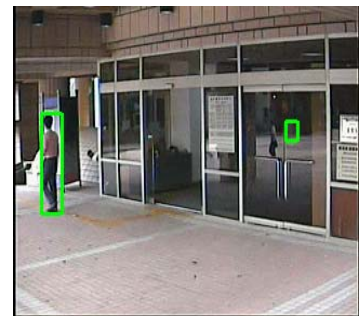
當一段影片裡有多個物體移動但都未發生交會（圖 4-4），則追蹤結果可得到圖 4-5 (a)由上往左下移動的女生與圖 4-5(b)由左往右移動的男生之移動軌跡，而偵測出為背景雜訊的部份結果在圖 4-6 中。



(a) 先出現單一物體移動

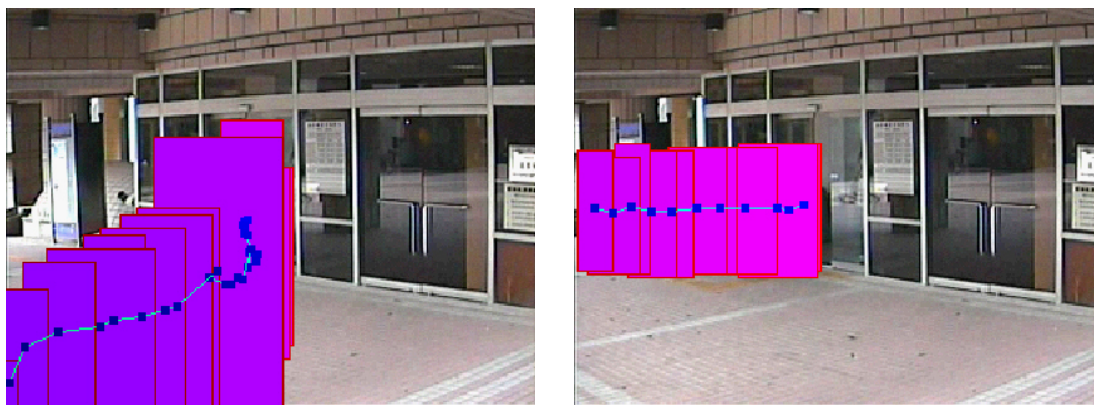


(b) 同時出現多個物體但未交會



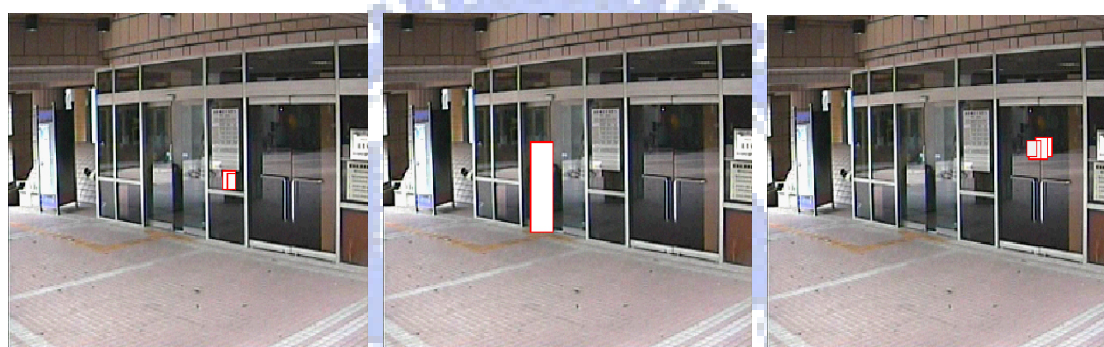
(c) 只剩另一個物體移動

圖 4-4：多個物體移動（Case 1）。(a), (b), (c)分別為影片中第 34、61 和 75 張畫面。



(a) 由上往左下移動的女生之移動軌跡：移動方向 214.06° ，移動速度 6.29 pixel/frame
(b) 由右往左移動的男生之移動軌跡：移動方向 180.92° ，移動速度 6.48 pixel/frame

圖 4-5：多個物體移動 (Case 1) 追蹤所得結果。



(a) 畫面 22-23

(b) 畫面 59

(c) 畫面 73-76

圖 4-6：多個物體移動 (Case 1) 偵測為背景雜訊之部分結果。

Case 2 (N1→N2→N1)

接著對依照 N1、N2、N1 狀態順序多個物體發生交會之狀況進行實驗。第一個例子場景是在大樓門口，兩個行人從左右兩邊往中間移動，並在中間發生交會的情況 (圖 4-7)，其追蹤結果為圖 4-8 所示，而偵測出為背景雜訊的部份結果在圖 4-9 中。



(a) 兩物體交會前

(b) 兩物體正交會中

(c) 兩物體交會後各自分開

圖 4-7：多個物體移動 (Case 2)。 (a), (b), (c) 分別為影片中第 38、43 和 46 張畫面。



(a) 由左往右移動的行人 A 之移動軌跡：移動方向 305.17° ，移動速度 12.10 pixel/frame
(b) 由右往左移動的行人 B 之移動軌跡：移動方向 175.75° ，移動速度 12.54 pixel/frame

圖 4-8：多個物體移動 (Case 2) 追蹤所得結果。



(a) 畫面 33-34

(b) 畫面 48

(c) 畫面 60

圖 4-9：多個物體移動 (Case 2) 偵測為背景雜訊之部分結果。

而第二個例子場景是在餐廳門口，為行人與小狗發生交會狀況 (圖 4-10)，其結果說明同前例 (圖 4-11、圖 4-12)。

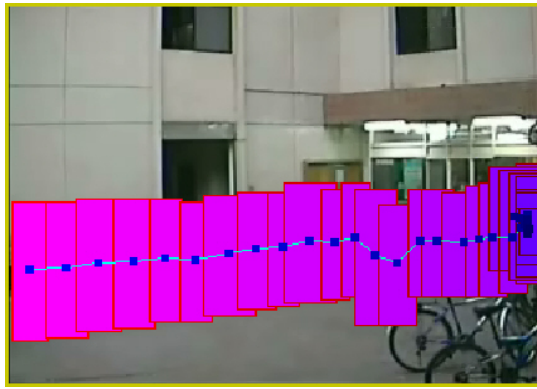


(a) 兩物體交會前

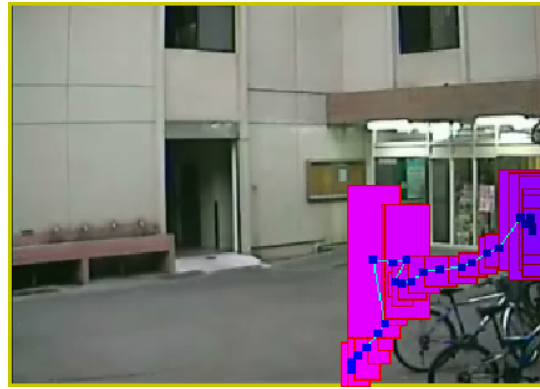
(b) 兩物體正交會中

(c) 兩物體交會後各自分開

圖 4-10：多個物體移動 (Case 2)。 (a), (b), (c) 分別為影片中第 115、139 和 210 張畫面。



(a) 行人之移動軌跡：移動方向 4.8° ，移動速度 4.58 pixel/frame



(b) 小狗之移動軌跡：移動方向 39.17° ，移動速度 4.36 pixel/frame

圖 4-11：多個物體移動（Case 2）追蹤所得結果。



(a) 畫面 137-141



(b) 畫面 160



(c) 畫面 198

圖 4-12：多個物體移動（Case 2）偵測為背景雜訊之部分結果。

Case 3 (N1→N2)

若一開始多個物體各自移動，後來發生交會狀況就一直處於合併狀態不再分開，即依序發生 N1、N2 狀態後就一直維持在 N2 的情況（圖 4-13），其追蹤結果如圖 4-14 所示分別可得到兩個行人的移動軌跡，而偵測為背景雜訊的部份結果顯示在圖 4-15。



(a) 一開始多個物體各自移動

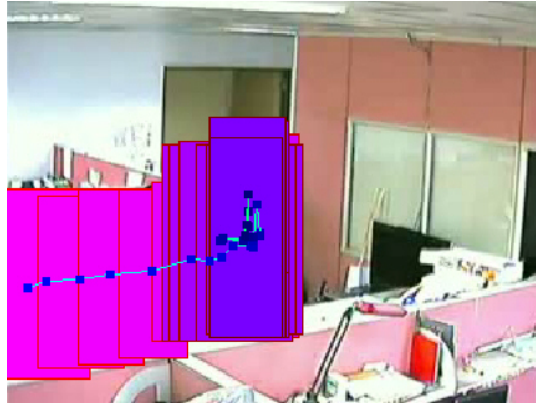


(b) 後來多個物體發生交會

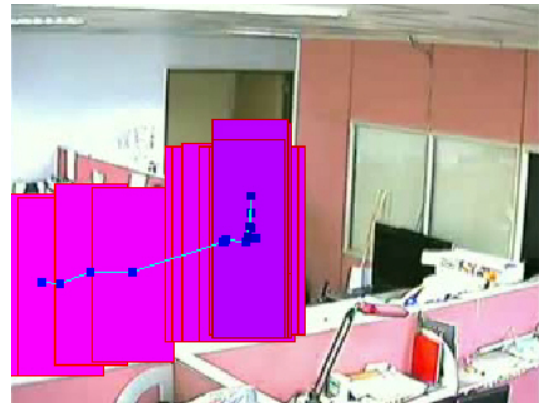


(c) 交會後在合併狀態未分開

圖 4-13：多個物體移動（Case3）。(a), (b), (c)分別為影片中第 30、38 和 53 張畫面。



(a) 穿著藍色衣服之行人的移動軌跡：移動方向 13.51° ，移動速度 6.93 pixel/frame



(b) 穿著紅色衣服之行人的移動軌跡：移動方向 12.38° ，移動速度 7.11 pixel/frame

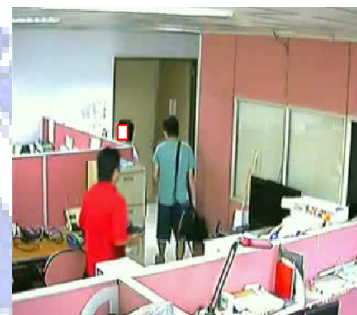
圖 4-14：多個物體移動（Case 3）追蹤所得結果。



(a) 畫面 15



(b) 畫面 25-26



(c) 畫面 40

圖 4-15：多個物體移動（Case 3）偵測為背景雜訊之部分結果。

Case 4 (N2→N1)

若畫面一開始多個物體並非各自獨立移動，而是處於交會影像重疊的情況，經過一段時間後才各自分開之情況，即從 N2 變成 N1 狀態後就一直維持在 N1 狀態（圖 4-16），此情況下追蹤結果如圖 4-17 所示。因為兩人一開始移動時影像就重疊，但是由於女生較靠近鏡頭而男生離鏡頭較遠，造成女生的影像大部份都覆蓋在男生的影像上，使得兩人重疊時後顏色特徵是以女生為主，所以兩人分開的時候會繼續追蹤女生的影像（圖 4-17(a)），而男生的影像則再產生一個新的移動軌跡來代表他（圖 4-17(b)），而另外偵測出為背景雜訊的部份結果顯示在圖 4-18 中。



(a) 兩人一開始即處於交會情況 (b) 兩人移動時仍交會重疊 (c) 之後兩人各自分開

圖 4-16：多個物體移動（Case4）。(a), (b), (c)分別為影片中第 7、22 和 67 張畫面。



(a) 女生的移動軌跡：移動方向 11.03° ，
移動速度 3.43 pixel/frame

(b) 男生的移動軌跡：移動方向 7.03° ，
移動速度 3.93 pixel/frame

圖 4-17：多個物體移動（Case 4）追蹤所得結果。



(a) 畫面 10-13

(b) 畫面 31-32

(c) 畫面 77

圖 4-18：多個物體移動（Case 4）偵測為背景雜訊之部分結果。

4.2. 追蹤失誤原因

圖 4-19 為 TrackIN 偵測前景物體發生問題的情形。(a) - (i)分別為影片中第 18、22、25、26、27、32、33、34 與 36 張畫面。由於追蹤方法是針對 TrackIN 所偵測代表前景物體的外接矩形來處理，假若 TrackIN 偵測過程出現如圖 4-19(c)因背景因素造成偵測錯誤，產生外接矩形與物體大小不合狀況，這會使圖 4-19(c)與圖 4-19(d)顏色直方圖差異太大導致使追蹤中斷。以及如圖 4-19(g)原本應偵測到物體卻未偵測出的問題，也會讓原本已被追蹤的物體誤認已經消失而中斷追蹤。而上述追蹤失誤的原因會造成如圖 4-20 所示，對單一物體進行追蹤卻會產生三個移動軌跡的錯誤結果。

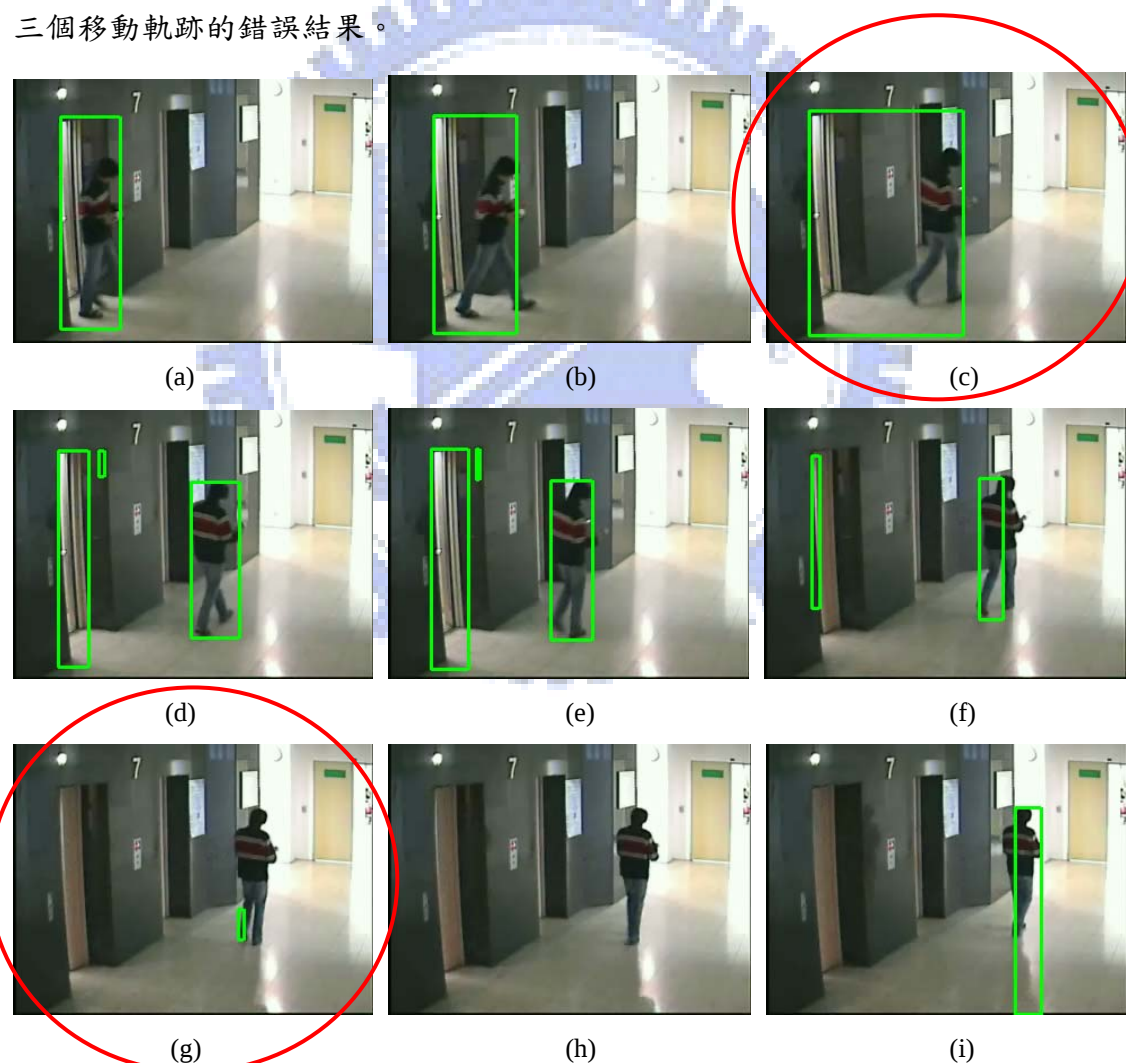


圖 4-19：TrackIN 偵測前景物體發生問題的情形。(a), (b), ..., (i)分別為影片中第 18、22、25、26、27、32、33、34 與 36 張畫面。其中(c) 因背景因素造成偵測錯誤，產生外接矩形與物體大小不合狀況，而(g) 為原本應偵測到物體卻未偵測出的問題。



圖 4-20：追蹤失誤結果：單一物體移動卻有三個移動軌跡。

4.3. 雜訊偵測問題

我們使用雜訊偵測的方法主要是以光線變化作為考量，對於因光變化而誤認為前景物的情況，其雜訊偵測結果非常不錯。然而對於背景物擾動如樹枝搖晃（圖 4-21）或玻璃、地板造成的反光現象（圖 4-22）等狀況，在我們的雜訊偵測上卻仍有問題。因為背景如樹枝原本就是實際存在物體，當風很大造成搖晃非常劇烈時，就會視同物體正在移動的情況（圖 4-23(a)），而另外一種因玻璃或地板造成的反光現象，若反光面積過大使得反射成像趨近完整，也會造成近似物體在鏡面上移動的現象（圖 4-23(b)）。因此對於上述的特殊狀況，若想要有效的將這些狀況區別出來，則需針對它們的特性再做更深入的研究。

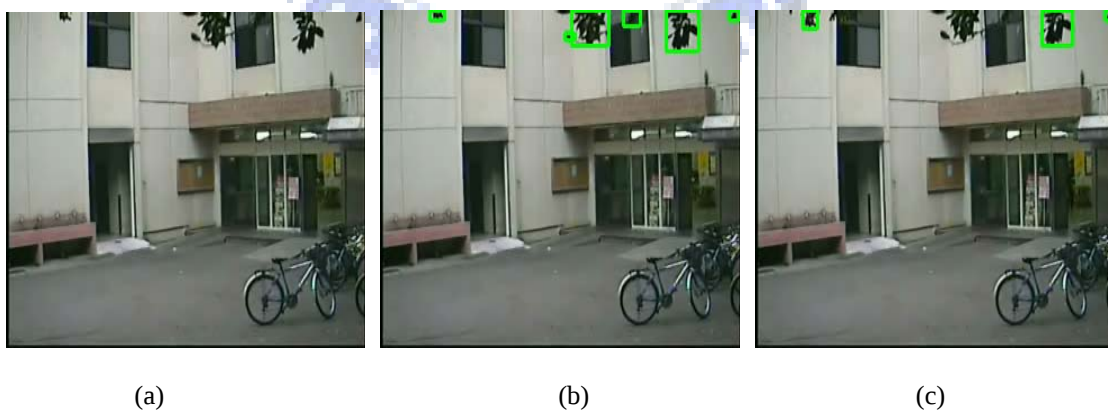


圖 4-21：樹枝搖晃之背景物擾動。



(a)

(b)

(c)

圖 4-22：玻璃造成的反光現象。



(a) 樹枝搖動之追蹤結果

(b) 玻璃反光之追蹤結果

圖 4-23：雜訊偵測之問題。

第五章 結論與未來工作

在本篇論文中，我們對監控影像資料庫提出了一套以物體的移動軌跡來搜尋的方法，對蒙恬的監控影像追蹤技術先偵測出連續視訊中各個前景物體，並以外接矩形表示。接著分析各物體的移動狀態，依照單一物體移動、多個物體移動等不同型態進行處理，建構出由外接矩形所組成表示物體移動的軌跡做為索引。同時進一步做雜訊偵測區別出背景雜訊，而對實際物體移動附加狀態、方向、速度等描述屬性。最後提供給使用者圖形化的軌跡搜尋介面，來快速並便利的得到所需資訊。

由於我們事前不知道使用者所要的資料為何，而提供的資料要多詳細才是對使用者有意義的也不清楚，因此我們以外接矩形組成代表物體移動的軌跡作為基礎資料，可先提供比原始影片還多一點的資訊給使用者參考。若使用者還想得到更多詳細的資料，可以更進一步萃取這些基礎資料來得到，譬如說想要知道某行人穿著何種顏色的衣服，我們就可以使用之前提供的基礎資料，知道此行人所佔的像素範圍來進行人體軀幹識別，進而取得上半身的顏色的資料。此外由於外接矩形所組成的軌跡，其結構簡單且清楚描述物體移動變化，因此本資料庫可適合提供給其他監控系統使用，並藉由互相比較與評估各自系統效能，來推廣出一個共同測試的平台，以幫助監控影像產業的進步與發展。

然而我們的追蹤方法若畫面中物體移動超過三個以上，如人潮來回等更複雜場景就容易發生失誤，在這方面需要再努力改善與克服。而以監控系統角度而言，夜間所拍攝的視訊資訊比一般日間來的有意義也更重要，且用於夜間拍攝的紅外線影像其特性也不大同於一般彩色影像，因此如何有效對夜間紅外線視訊加以應用，並與原本彩色視訊進行整合，是我們未來所要研究的重要課題。

參考文獻

- [1] S. Deb and Y. Zhang, "An Overview of Content-Based Image Retrieval Techniques," *Proceedings of International Conference on Advanced Information Networking and Applications*, vol. 01, no. 1, p. 59, 2004.
- [2] F. Arman, A. Hsu, and M. Chiu, "Image Processing on Compressed Data for Large Video Databases," *ACM Multimedia*, pp. 267-272, 1993.
- [3] D. Makris and T. Ellis, "Path Detection in Video Surveillance," *Image and Vision Computing Journal*, vol. 20, no. 12, pp. 895-903, 2002.
- [4] I. Junejo, O. Javed, and M. Shah, "Multi Feature Path Modeling for Video Surveillance," *Proceedings of International Conference on Pattern Recognition*, vol. 2, pp. 716-719, 2004.
- [5] Z. Fu, W. Hu, and T. Tan, "Similarity Based Vehicle Trajectory Clustering and Anomaly Detection," *Proceedings of IEEE International Conference in Image Processing*, vol. 2, pp. 602-605, 2005.
- [6] J. Black, T. Ellis, and D. Makris, "A Hierarchical Database for Visual Surveillance Applications," *Proceedings of IEEE International Conference on Multimedia and Expo*, pp. 1571-1574, 2004.
- [7] E. Sahouria and A. Zakhori, "A trajectory Based Video Indexing System for Street Surveillance," *Proceedings of IEEE International Conference on Image Processing*, pp. 24-28, 1999.
- [8] S. McKenna, S. Jabri, Z. Duric, and A. Rosenfeld, "Tracking Groups of People," *Computer Vision and Image Understanding*, no. 80, pp. 42-56, 2000.
- [9] M. Swain and D. Ballard, "Colour Indexing," *International Journal of Computer Vision*, vol. 7, no. 1, pp. 11-32, 1991.

- [10] C. Jaynes, S. Webb, R. Steele, and Q. Xiong, "An Open Development Environment for Evaluation of Video Surveillance Systems," *Proceedings of International Workshop on Performance Evaluation of Tracking and Surveillance*, pp. 32-39, 2002.
- [11] C. Needham and R. Boyle, "Performance Evaluation Metrics and Statistics for Positional Tracker Evaluation," *Proceedings of the Third International Conference on Computer Vision Systems*, vol. 2626, pp. 278-289, 2003.
- [12] F. Porikli, "Trajectory Distance Metric Using Hidden Markov Model Based Representation," *International Workshop on Performance Evaluation of Tracking and Surveillance*, 2004.
- [13] A. Nghiem, F. Bremond, M. Thonnat, and V. Valentin, "ETISEO, Performance Evaluation for Video Surveillance Systems," *IEEE International Conference on Advanced Video and Signal based Surveillance*, 2007.

