

國立交通大學

資訊科學與工程研究所

碩士論文

動態攝影機下的前景背景分割



Separation of Foreground and Background for Dynamic
Cameras

研究生：邱永昌

指導教授：莊仁輝 教授

中華民國九十八年六月

Separation of Foreground and Background for Dynamic Cameras

研 究 生：邱永昌

Student : Yung-Chang Chiu

指 導 教 授：莊仁輝

Advisor : Jen-Hui Chuang

國 立 交 通 大 學

資 訊 科 學 與 工 程 研 究 所

碩 士 論 文

A Thesis

Submitted to Institute of Computer Science and Engineering

College of Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Computer Science

June 2009

Hsinchu, Taiwan, Republic of China

中華民國九十八年六月

動態攝影機下之前景背景分割

學生：邱永昌

指導教授：莊仁輝 教授

國立交通大學

資訊科學與工程研究所

摘要

在本篇論文中，我們提出一個方法，能在動態式攝影機拍攝影像中偵測前景物體。本方法一開始會需要一張或數張單純只有背景物的影像以建立背景高斯混合模型，之後再利用此高斯混合模型所提供的背景影像建立以顏色為基礎的像素群組、以邊緣 (edge) 所切割出的區塊以及用角點偵測器所獲得之角點。當有新進的影像須要做前景物偵測時，以同樣方式對該新進影像建立對應的群組、區塊以及角點，接著藉由配對上述三種背景影像與新進影像的對應特徵，便可以估算兩張影像間的移動向量。接著根據此移動向量，將背景高斯混合模型移動到正確的位置，意即我們對齊背景高斯混合模型與新進影像，然後才進行前景背景的分割，之後再進一步地消除雜訊以及攝影機些微旋轉的影響，以得到最後的前背景分割結果。從實驗結果中可以看到，在背景景深差異不是太大的前提下，即使兩張影像間的移動量並不是十分的小，我們的方法仍可以成功地切割出前景物。

Separation of Foreground and Background for Dynamic Cameras

Student : Yung-Chang Chiu

Advisor : Jen-Hui Chuang

Institutes of Computer Science and Engineering
National Chiao Tung University

ABSTRACT

In this thesis, we propose an algorithm to detect foreground objects in an image sequence obtained with dynamic cameras. The algorithm needs one or several initial images containing pure background environment to establish its Gaussian mixture model. Then, we establish color-based pixel clusters, edge-secluded image blocks, and corners of image regions from the background image supported by such a Gaussian mixture model. In order to detect moving objects inside an upcoming input image, the above three types of image features are extracted. By matching these features between the background image and the input image, the motion vector between the background image and the input image can be estimated. Such a process corresponds to an alignment between the Gaussian mixture model and the input image, and the moving objects can be detected using aligned Gaussian mixture model. After removing some noise and partly eliminating the effect due to small camera rotation, the final foreground separation result can be obtained. According to the experiment results, even when the displacement between two consecutive images is not very small, the proposed algorithm can still detect moving objects if the background has a small range of depth.

致謝

本篇論文得以完成，中間經歷了許多的困難，從問題的發現到解決、器材的選擇、演算法的測試、場地的設定以及論文內容的撰寫與修正，深深地讓我體會到原來做研究真的要投資許多的時間與勞力。但是，從研究的過程中也讓我獲益良多，另外，我衷心地感謝我的指導教授莊仁輝博士，他對電腦視覺領域精闢的見解與廣博的見識往往是我遇到瓶頸時候最大的幫助，尤其是他對研究的熱忱與認真的態度更是讓我印象深刻，也感謝我的口試委員王聖智教授、顏嗣鈞教授與雷欽隆教授在口試時給我寶貴的意見，讓我的論文能夠更加地完善。

同時我也要感謝實驗室博士班學長林泓宏與羅國華對我熱心的指導，讓我在研究中遇到瓶頸的時候有了新的方向。另外，我也要感謝尚一、維辰、宣良、世旻、聖中以及邦展，讓我在研究所的生涯中得到的不只是學術上的知識，還有彼此之間的鼓勵以及幫助，讓我了解到友情的可貴以及朋友的重要。最後我要感謝我的家人，因為有他們對我的支持，才能讓我沒有後顧之憂地全力攻讀碩士學位，感謝所有曾經幫助過我的人。

目錄

摘要.....	I
ABSTRACT.....	II
致謝.....	III
目錄	IV
圖目錄.....	V
表目錄.....	VI
一、 緒論.....	1
1.1 研究動機.....	1
1.2 相關研究.....	4
1.3 系統流程.....	10
1.4 各章簡介.....	12
二、 建立用於影像對齊之重建的背景影像.....	13
2.1 高斯混合模型.....	13
2.2 建立重建的背景影像.....	16
三、 三種可用於影像對齊之影像特徵擷取.....	18
3.1 以顏色為基礎的群組.....	18
3.2 用邊緣所切割出的區塊.....	22
3.3 Harris Corner Detector與角點配對.....	24
四、 影像對齊與前景背景分割.....	30
4.1 尋找影像間之候選移動量.....	30
4.2 找出最佳移動量.....	33
4.3 影像對齊與前景背景分割.....	36
4.4 新進與移出的背景像素之處理.....	38
五、 實驗結果與討論.....	40
六、 結論與未來展望.....	49
參考文獻.....	51
附錄1. 找出相連的區域 (Connected Component Search)	53

圖目錄

圖1.1 本論文之想法示意圖。.....	3
圖1.2 背景移動超過兩個像素對文獻[1]的影響。.....	6
圖1.3 本論文所提出的動態攝影機下前景背景分割法的流程圖。.....	11
圖2.1 一個高斯模型的例子。.....	14
圖2.2 GMM提供的背景影像。.....	17
圖3.1 群組的結果。.....	21
圖3.2 群組模型顏色值有無更新對群組模型結果的影響。.....	21
圖3.3 sobel edge detector。.....	22
圖3.4 用邊緣所切割出的區塊的結果。.....	23
圖3.5 群組可以成功地切出區域但區塊卻不行的例子。.....	24
圖3.6 不同的 R 值所得到的角點偵測結果。.....	27
圖3.7 消除角點群聚現象的結果。.....	28
圖4.1 差異值與總數值的計算範圍。.....	34
圖4.2 不準確的移動量對差異值和總數的影響。.....	35
圖4.3 有無濾除雜訊的結果。.....	37
圖4.4 有無消除攝影機旋轉影響的結果。.....	37
圖4.5 有無強制設定背景，對新進影像與背景影像的同步的影響。.....	38
圖5.1 模擬影像的測試結果。.....	40
圖5.2 出錯後又回復正常的過程。.....	41
圖5.3 靜態式攝影機下的實驗結果。.....	42
圖5.4 動態式攝影機下的實驗結果。.....	43
圖5.5 動態式攝影機下的第二個實驗結果。.....	44
圖5.6 動態式攝影機下的第三個實驗結果。.....	45

表目錄

表1 本論文所引用的相關文獻與本論文之間的優缺點比較。.....	10
表2 群組、區塊以及角點的效率比較。.....	48
表3 群組、區塊以及角點的穩定性比較。.....	48



一、緒論

1.1 研究動機

在電腦視覺的領域中，為了達到減少計算量的目的，許多的方法希望能先在影像中找出需要做進一步處理的區域，再作接下來的處理。例如在安全監控相關應用下的監視系統，就需要做人體姿態的辨識，以判斷監視場景中存在的人物是否有可疑的行為。但是企圖在整張影像中作人體姿態辨識容易有很高的false alarm，更會大大地增加所需的計算時間。因此如果能事先提供一個精準的方法，過濾掉與前景物（例如人體）不相關的背景區域，僅僅留下前景物所在的像素，讓整個辨識的方法只需要專注在切割出來的前景物，而非連背景環境也一起加入考慮，如此便能為接下來的辨識過程省去極為可觀的時間。

對於靜態的攝影機，現今之高斯混合模型已經能夠有效率地處理前景背景分割的問題，但前提是攝影機必須要是靜態的，如此才能夠正確地切割出前景物，即使攝影機只稍微移動了一段很短的距離，也足以造成嚴重的錯誤。由於這個條件大大地限制了整個前景背景分割的實用性，因此我們希望發展一個能被應用於動態式攝影機的前景背景分割方法。此方法之一種可能的應用是指定物體的追蹤，而且追蹤的範圍非常大，超過一台靜態式攝影機可以拍攝的範圍，因此追蹤的工具由靜態式攝影機轉變成動態式攝影機，甚至是裝設在自動車上的攝影機。在這種應用下，一個常見的問題是被追蹤物體本身的位置在兩張連續的影像中有可能差非常多，有些假設被追蹤物體在連續影像中不可移動過多的方法，例如[1]，自然就無法處理這樣的問題。另有一種可行的方法為在影像中找出與被追蹤物體具有相似特徵的區塊，這些特徵可能是顏色或是形狀等資訊，再進一步地判斷該區塊是被追蹤物的可能性有多高，或是特徵有多相似，最後再找出一個最可能或是最相似的區塊，將其當成被追蹤的物體。這樣的追蹤法，例如[2]，必須要在

整張影像中尋找特徵點，才有辦法做進一步的判斷，但是尋找特徵點這一步必然會花費非常多的時間，且接下來的特徵點配對一樣是非常花費時間的步驟，因為其特徵點來自於整張影像。所以針對這種問題，最好能夠事先提供一個有效的動態式攝影機下的前景背景分割法，先過濾出前景物出現的區域，讓追蹤的方法需要檢視的可能特徵點存在區域減到最少，以增進整體追蹤方法的效率。

有鑑於以上的需求，本論文希望藉由維護一張純粹只有背景環境的背景影像（在本論文中稱此背景影像為重建的背景影像，且此重建的背景影像由背景高斯混合模型提供），之後當有新進影像要做前景背景分割時，先找出新進影像與重建的背景影像之間的移動向量，之後利用此移動向量將背景高斯混合模型對齊新進影像，再進一步地找出哪些像素是原來就存在且沒有移動過的背景像素，以及哪些像素是有移動過的前景像素。如此，在能夠對齊高斯混合模型與新進影像的情況下，就能夠達到動態式攝影機下前景背景分割的目的。至於找出移動向量的方法，可以藉由配對新進影像與重建的背景影像之間重複出現的區域或是特徵點（例如角點），來計算移動向量。圖1.1表示藉由配對區域與角點以對齊影像，再切割出前景物的想法。圖1.1(a)為一重建的背景影像，圖1.1(b)為一新進影像，我們在重建的背景影像中找出區域和角點，如圖1.1(c)中紅色的圈圈和矩形分別表示一個重建的背景影像中的角點和區域，而該角點和區域的質心座標分別為 $(144, 28)$ 、 $(169, 76)$ ，同時我們也在新進影像中找出區域和角點，如圖1.1(d)表示一個在新進影像中找到的角點和區域，該角點和區域的質心座標為 $(66, 106)$ 、 $(91, 154)$ ，之後將這兩個角點的座標相減或是兩個區域的質心座標相減就可得到重建的背景影像相對於新進影像的移動向量 $(-78, 78)$ 。接著將重建的背景影像根據找到的移動向量移動到正確的位置以對齊新進影像，意即我們將重建的背景影像（即圖1.1(a)）往左移動78個像素後再往下移動78個像素，移動後的結果如圖1.1(e)所示，其中黑色的區域表示原來不存在重建的背景影像中的區域，因此新進影像位在黑色區域的部分將無從判斷是否有前景物，本論文將該部分的新進影像通通視為背景。在對齊新進影像和重建的背景影像後，就可以利用任何一

個靜態式攝影機下的前景背景分割法切割出前景物（本論文利用GMM來達到對齊重建的背景影像與新進影像後的前背景分割），而整個前景背景分割的結果如圖1.1(f)所示。最後再將重建的背景影像中黑色的部分用新進影像中相同位置的像素填滿，讓重建的背景影像能夠跟最新的背景環境同步，成為一張新的重建的背景影像，如圖1.1(g)所示，這樣就能夠繼續做下一張新進影像的前背景分割。另外因為錯誤地配對區域或角點的影響，會在新進影像與重建的背景影像間找出數個可能的候選移動向量，因此我們需要過濾這些候選移動向量，以找出新進影像與重建的背景影像間最佳的移動向量。

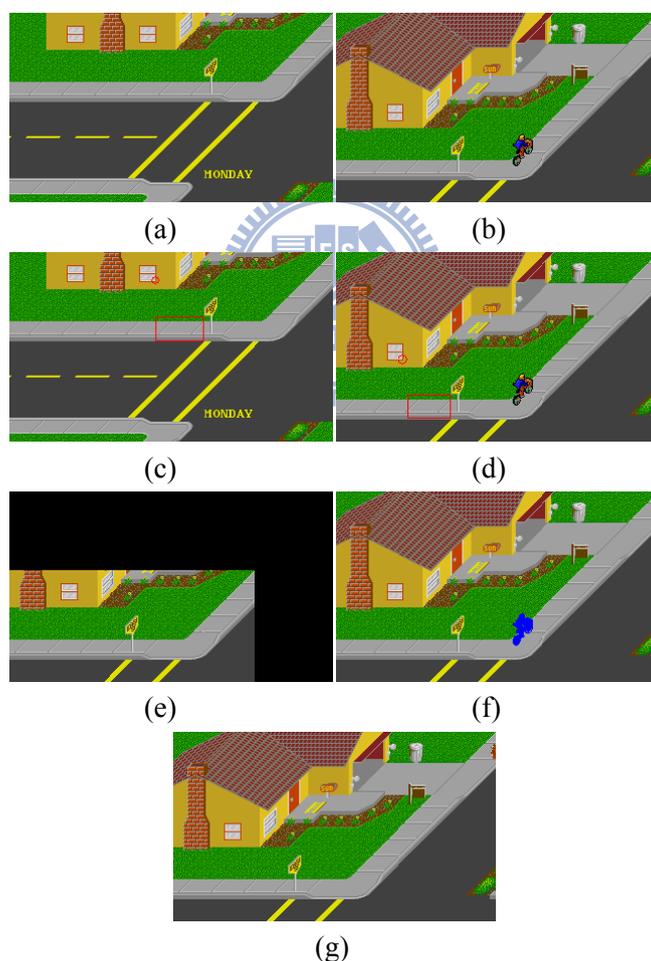


圖1.1 本論文之想法示意圖。(a) 重建的背景影像，(b) 新進影像，(c) 重建的背景影像中的角點和區域，(d) 新進影像中的角點和區域，(e) 將重建的背景影像根據找到的移動向量移動到正確的位置，(f) 前景背景分割的結果，(g) 新的重建的背景影像。

但是因為將影像間的運動僅用一個移動向量來表示，所以本論文的方法不適合應用在攝影機有前進後退、透式投影的影響強烈以及像素間景深差異太大的情況，因為以上三種情況會造成影像間的運動無法僅用一個移動向量來表示，因此本論文所提出的方法就不能夠有效的解決這三種情況。

本論文提出的方法除了解決攝影機移動的問題外，同時也加入了Gaussian Mixture Models的觀念（即高斯混合模型，以下簡稱GMM），以處理場景中有規律運動現象的背景物，例如吊鐘的鐘擺、較規律的樹影晃動或甚至是其他更複雜但有規律的運動。且GMM亦可解決光影變化的問題，對真實環境下場景因光影變化造成影像不穩定時做前景背景分割的幫助相當大。但GMM的缺點是無法直接應用在動態式攝影機的環境下，因此我們希望藉由估算出重建的背景影像相對於新進影像（由於加入了GMM，所以重建的背景影像就來自於GMM，詳細的方法請參考2.2節）的移動向量，再利用此移動向量對齊GMM與新進影像，達到前景背景分割的目的。



1.2 相關研究

前景背景分割根據其應用環境的不同，可以區分為靜態式以及動態式攝影機下的前景背景分割兩個領域，其中靜態式攝影機下的領域已經有了相當不錯的研究結果，例如文獻[3]。文獻[3]中提及，我們將每個像素都用數個高斯模型來模擬該像素過去曾經出現過的顏色，即完成背景模型之建構。對於後續的新進影像，我們把每個新進影像中像素的顏色值，與位在該像素位置的數個高斯模型做比較，看看是否有可以配對成功的模型存在。然後再判斷這些配對成功的模型是否屬於背景模型。如果以上的兩個條件成立，將該像素視為背景，反之則視為前景。我們可以根據使用者提供的參數來更新每個模型的平均值、變異數以及權重。文獻[3]提出的方法，在靜態的環境下已經能夠很穩定地達成前景背景分割的目

的，但是這個方法的缺點是沒有辦法應用在動態式攝影機的環境下，因為若是攝影機有稍稍的偏移，我們將無從判斷像素與高斯模型之間的空間對應關係。因此要是沒有提供一個對齊的機制，則新進影像中的像素將會對應到不正確的高斯模型，整體的結果就會因為沒有對齊而切割出許多的前景物。儘管這個問題在一段時間後，會因為該方法本身隨著時間新增模型以及更新模型參數的特性而消失，但是這段時間內因為沒有對齊高斯模型而被誤認為前景物的背景像素，與真正進入場景中的前景物，將無從區別。

對於動態式攝影機，因為背景環境是會移動的，因此需要一套機制來區分移動中的像素究竟是屬於前景物還是背景環境。在這種需求下，有兩種領域可以解決移動中的背景像素所產生的問題，一個是不需要對齊影像的追蹤法，例如文獻[2][4]，另一個則是需要對齊影像的前景背景分割法。至於需要對齊的影像，通常是新進影像跟上一個影格的影像，例如文獻[5][6][7]，或者是維護一個純粹只有背景的影像，例如文獻[1]與本論文所提出的方法，再對齊新進影像與該純背景影像。對於不需要對齊影像的追蹤法，例如文獻[4]，雖然其功能在於追蹤一個指定的前景物，但最終目的則與本論文一樣，都是希望能夠準確地定位出有興趣的區域，只不過追蹤法有興趣的區域是一開始就指定好的區域，例如一個人，而本論文則是對會移動的區域有興趣。文獻[4]整體的方法大致是：記住前景物（或說被追蹤物）過去曾經出現過的位置，利用一個固定的移動向量計算模型以及一些亂數，以估測前景物在下一個影格（frame）中可能的位置。但是因為需要提供一個固定的移動向量計算模型來模擬前景物的移動行為，若是前景物出現在場景中的位置較無規律，造成其移動行為無法以提供的移動向量計算模型來模擬時，文獻[4]就無法成功地找出該前景物，因此不適合被應用在需要準確區分前景背景的情況。

至於需要對齊影像的文獻[1][5][6][7]，就需要提出一個對齊的機制，例如文獻[1]是一個在動態式攝影機下切割出前景物的方法，且與本論文所提出的方法一樣地，都需要一張純背景影像作為前景背景分割的依據。當要對新進影像做前

背景分割時，同樣也跟本論文所提出的方法一樣地，將背景影像中顏色相近且空間上相連的像素組成一個群組，但是接下來的過程就與本論文不同。文獻[1]假設攝影機的移動不會超過上下左右各兩個像素，所以對接下來新進影像中的每個像素，以 5×5 的範圍內作新進影像像素與背景影像群組的配對，如果可以配對成功，則將該新進影像像素判斷為背景，反之則判斷成前景。文獻[1]的缺點是限制了攝影機的移動向量只有上下左右最多各兩個像素，一旦攝影機的移動向量超過了這個限制，這個方法就會發生嚴重的錯誤，這個限制大大地降低了文獻[1]的實用性，圖1.2表示文獻[1]處理攝影機移動向量超過上下左右各兩個像素的結果，在這個例子中背景影像相對於第一張新進影像的移動向量是 $(-78, 78)$ ，我們希望將新進影像中的腳踏車騎士切割出來，但是因為新進影像相對於背景影像的移動向量超過了上下左右各最多兩個像素的限制，因此文獻[1]將很多的背景像素誤判為前景物。

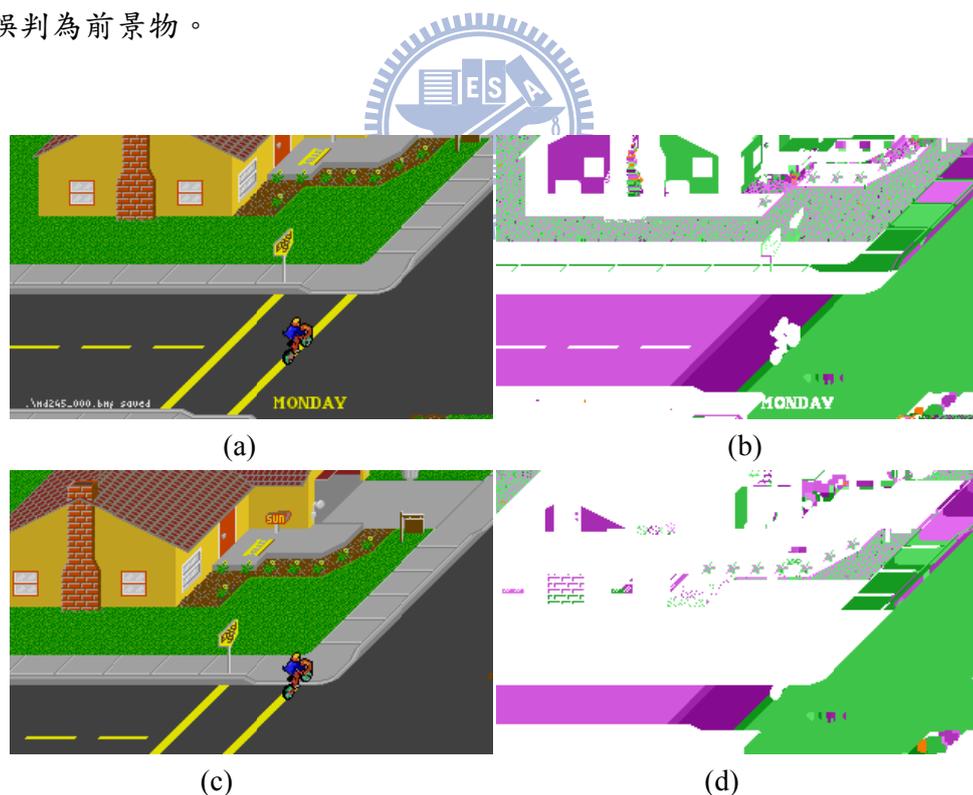


圖1.2 背景移動超過兩個像素對文獻[1]的影響。白色區域表示切割出來的前景，由此可見文獻[1]無法處理影像移動向量超過上下左右各兩個像素的問題。(a) 第1個影格，(b) (a)的前景背景分割結果，(c) 第2個影格，(d) (c)的前景背景分割結果。

文獻[5][6][7]雖然與文獻[1]和本論文一樣，都需要對齊影像才能進行前景背景的分割，但是文獻[5][6][7]是對齊新進影像與上一個影格的影像，而非維護一個純背景影像。例如文獻[5]介紹了一個在動態式攝影機下找出移動中行人的方法，然而該方法需要有特殊的儀器提供攝影機的旋轉跟移動向量，才能夠找出前後兩張影像之間的關係，再進一步地切割出正在移動的前景物件。但是就算有了攝影機的旋轉跟移動向量，該方法仍然遇到一個問題，那就是沒有影像中像素的深度資訊，因為計算一個像素在旋轉移動後的位置需要有深度的資訊。對此，該方法建議事先測量出攝影機所有可能拍攝的深度範圍，接著對這個範圍做量化，再把這些量化後的深度資訊當作一個像素可能的深度，這意味著一個像素會有很多個可能的深度值。因此往後當該方法欲計算上一張影像的像素在下一張影像的位置時，便需要考慮所有量化後的深度，這會造成一個像素在下一張影像中會有很多個可能的位置。該方法建議只要在這些可能的位置中有一個能夠跟下一張影像中相同位置的像素相似，則該下一張影像中的同位置像素便可以歸類成背景。因此這個方法的問題為深度量化的不精準性，以及容易將前景物誤判為背景物。

文獻[6]為一動態式攝影機下切割出移動前景物的方法，該方法跟文獻[5]一樣，需要由機器輸入特定的參數，如果是移動式的攝影機則輸入攝影機的移動向量，但如果是旋轉式攝影機則輸入攝影機的旋轉量，之後當要從輸入影像切割出前景物時，本方法將上一個影格的輸入影像根據機器提供的移動或旋轉量來做移動或是旋轉，意即與本方法一樣地，先對齊兩張影像，接著再將這兩張對齊後的影像相減，得到移動中的區域。但是相減後得到的移動中區域有兩種可能：第一個是在上一張影像中為背景但在下一張影像中有前景物出現的區域，第二個是在上一張影像中為前景但在下一張影像中因為前景物離開而再次露出來的背景區域。為了消除因為第二種現象而出現的移動中區域，文獻[6]建議對輸入影像尋找影像中所有的邊線，之後對每一個像素判斷其是否為同時為邊線和移動中區域，如果是則表示該像素屬於移動中前景物的輪廓，如此便可以找出影像中所有移動前景物的輪廓。但是也因為限制一個前景物像素必須要同時屬於邊線和移動中區

域，導致文獻[6]只能切割出前景物的輪廓而已，而非完整的前景物像素區域。文獻[6]與本論文提出的方法相較下有兩個缺點：第一個是需要由機器輸入特定的參數，第二個是對於前景物，文獻[6]只能切出該前景物的輪廓，無法跟本論文提出的方法一樣地找出完整的移動中前景物。

文獻[7]為一個在動態式攝影機下判斷是否有行人的方法，該方法一開始先找出移動中的前景物，接著將找出的前景物交與一個訓練好的TDNN(Time Delay Neuron Network)分類器來判斷該前景物是否為移動中的行人。與本論文所提出的方法一樣，該方法要求第一張影格須為純背景的影子，接著指定一個數字 N 表示影像中有多少個顏色相似且空間上相鄰的區塊，再將影像中的像素根據其顏色和位置分類到這 N 個不同的區塊。當之後有新進影像要做前景背景分割時，試著將新進影像中所有的像素根據其顏色分類進這 N 個區塊中，如果像素可以成功地被分類進這 N 個區塊中，表示該像素為背景物，否則將該像素判斷為前景物。此方法的缺點是 N 必須手動指定，無法自動化判斷，且藉由將像素分類進 N 個區塊以判斷該像素是否為前景物的步驟僅僅考慮像素與區塊顏色上的關係，並沒有考慮到像素與區塊空間上的關係，這是因為攝影機是會移動的，如此將會造成新進影像中的像素與區塊間的空間關係無法決定，因此如果新進前景物與 N 個區塊中的其中一個區塊顏色相似，則該新進前景物將全部被誤判成背景。而本論文只有在前景物經過顏色相似的背景區域時，才會將前景物誤判成背景。

其他還有一些可以發展成為動態式攝影機下前景背景分割法的方法，例如文獻[8][9]，文獻[8]原本是一個計算兩張影像間每個區塊的移動向量的方法，因此我們只需要將移動向量跟其他區塊特別不一樣的區塊找出，便可以將該區塊判斷為移動中的前景物。文獻[8]一開始先將影像切割成數個區塊(block)，接著計算這些區塊各自的移動向量(motion vector)，要是發現有一個區塊的移動向量跟其他區塊的移動向量差異非常多，該區塊就有可能為前景物。套用此方法，前景和背景可以很輕易的被分割出來，甚至可以結合文獻[3]，成為一個更穩定且更能容忍雜訊的方法。但其缺點在於移動向量的計算是先將影像分割成數個區

塊，再對每一個區塊，先從小的移動向量開始判斷是否為該區塊合適的移動向量，若否，則慢慢加大移動向量，直到找到該區塊合適的移動向量為止。一旦一個區塊找到了合適的移動向量，該區塊傳遞自己的移動向量給附近的區塊，讓附近的區塊確認該移動向量是否也是適合自己的移動向量。若是，則繼續傳遞該移動向量給其他鄰近的區塊，讓鄰近的區塊免於費時地尋找移動向量，以達到加速的目的。這個方法適合小範圍的影像移動，若遇到較劇烈的影像移動，其花費的時間將會非常可觀。

文獻[9]為一個對齊兩張影像間角點的方法，如果我們跟文獻[1]一樣維護一張純背景的影像，且利用文獻[9]來對齊新進影像與純背景影像間的角點，則所有在新進影像中無法在背景影像中找到對應點的角點便可歸類為前景物角點。文獻[9]一開始先從兩張影像中找出角點，之後粗略地配對這些角點再利用八點演算法以求出影像之間可能的基本矩陣（fundamental matrix）。由於考慮到不準確的角點配對，在測試基本矩陣是否夠準確的過程中利用極線幾何（epipolar geometry）來測試在一張影像中的角點在另外一張影像上的極線是否亦有角點存在，如果有則認為該角點能夠藉由該基本矩陣成功地找到在另外一張影像中的對應點。如果大部分的角點都能利用該基本矩陣成功地找出對應的角點，則認定該基本矩陣為兩張影像間最準確的基本矩陣，反之則需另外選擇其他八個粗略配對的角點，再求出新的基本矩陣，重複上述的步驟，直到大部分的角點都能利用該基本矩陣成功地找出對應的角點為止。文獻[9]與本論文所提出的方法相比下有一個缺點：利用文獻[9]所找出的前景物僅僅是角點而以，並無法像本論文所提出的方法一樣完整地找出所有前景物所在的像素，因此若要使用文獻[9]來切割前景物，必須要解決這個問題才行。

表1整理出以上所列出的相關文獻與本論文之間的優缺點比較，以方便了解各個方法相對於本論文的優缺點。

表1 本論文所引用的相關文獻與本論文之間的優缺點比較。

	相對於本論文的優缺點
文獻[1]	攝影機的最大移動向量有限，僅有上下左右各兩個像素
文獻[2]	需要提供前景物的移動模型、尋找特徵點費時
文獻[3]	只適用於靜態式攝影機
文獻[4]	需要提供前景物的移動模型
文獻[5]	需要輸入攝影機移動的參數，容易將前景物誤判成背景物
文獻[6]	需要輸入攝影機移動的參數，僅能找出前景物的輪廓
文獻[7]	需要手動指定影像中區塊的數目，容易將前景物誤判為背景
文獻[8]	計算非常費時
文獻[9]	只能找出前景物的角點
本論文	無法處理太複雜的背景運動

1.3 系統流程

圖1.3為本論文所提出的動態攝影機下前景背景分割法的流程圖，在流程圖的頂端為新進影像以及重建的背景影像，其中新進影像為我們想要找出移動前景物的目標影像，而重建的背景影像則是由高斯混合模型則所提供的純背景影像，因此至少第一個影格需要是單純背景的背景影像，用以建立純背景高斯混合模型。有了重建的背景影像後，才能建立以顏色為基礎之純背景群組、用邊緣所切割出之純背景區塊以及背景角點，其中一個群組或區塊代表影像中顏色相似且空間上相連的區域，而角點則是在影像中 x 與 y 軸顏色差異度都很大的像素。在取得了新進影像後，同樣對該新進影像建立群組、區塊及角點。接著藉由配對重建的背景影像與新進影像的群組、區塊以及角點以計算出重建的背景影像與新進影像間所有的候選移動向量，之後再找出能提供重建的背景影像與新進影像最小差異度的候選移動向量為最佳的移動向量。在取得了最佳移動向量後便可以將高斯混合模型移動到正確的位置，意即我們對齊高斯混合模型以及新進影像，接著執行前景背景分割。在得到切割完的結果後，把面積過小的前景物視為雜訊刪除，只留下面積夠大的前景區域，再消除攝影機些微旋轉對影像造成的影響，就是本方法最終的結果。最後更新高斯混合模型的位置，使其能與新進影像中的背景同步。

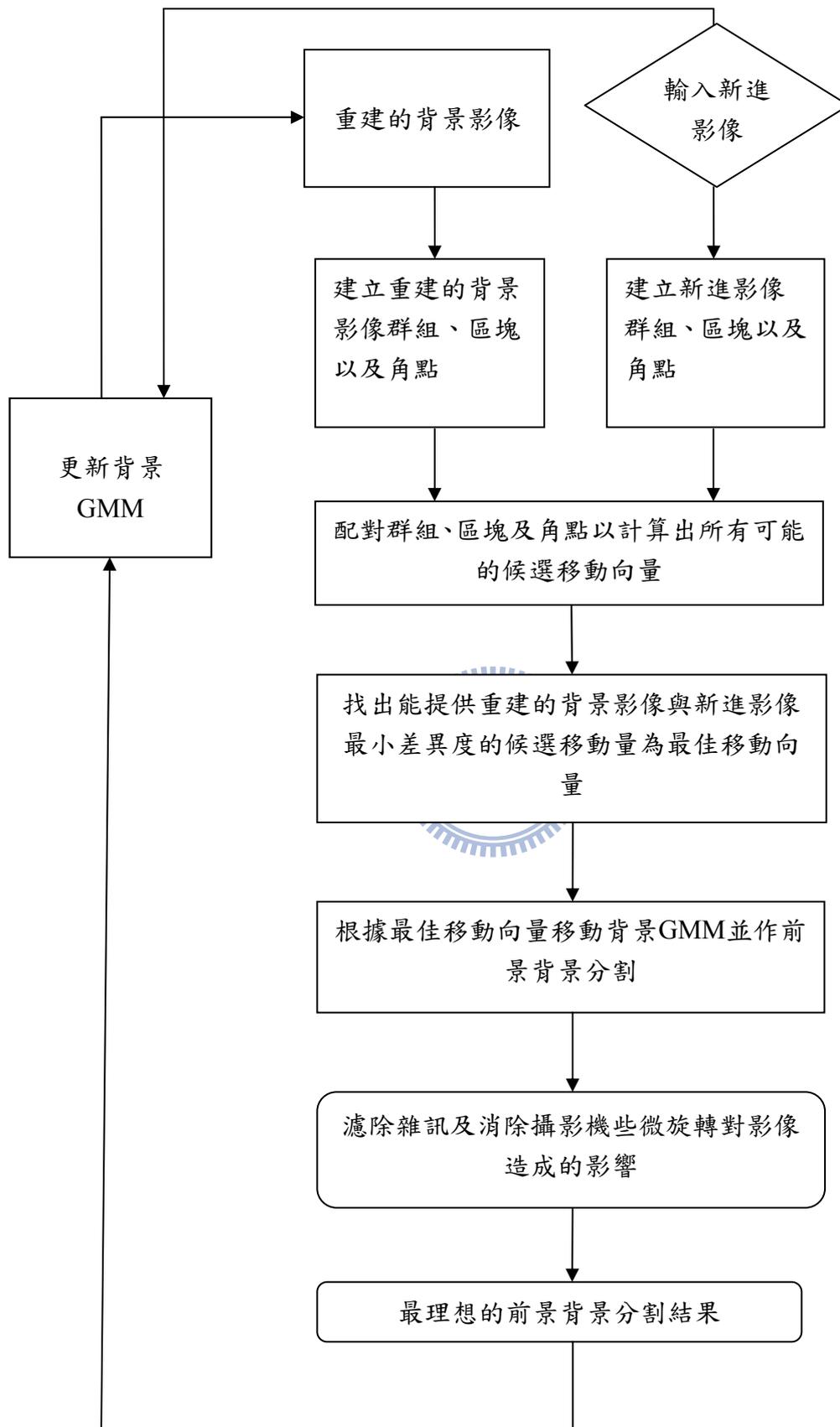


圖1.3 本論文所提出的動態攝影機下前景背景分割法的流程圖。

1.4 各章簡介

本論文共分成六章，除了本章緒論之外，我們將在第二章介紹高斯混合模型的原理，以達到在高斯混合模型以及新進影像已經對齊的前提下，完成前景背景分割的目的，同時利用高斯混合模型建立重建的背景影像，以提供對齊高斯混合模型與下一個新進影像的依據。第三章將介紹在影像中擷取特徵的三個方法，分別為以顏色為基礎建立群組的方法、在影像中尋找邊線以建立用邊線所切割出的區塊的方法、在影像中尋找角點以及配對角點的方法。第四章說明了如何藉由配對群組、區塊以估算候選移動向量，再根據每個候選移動向量找出能提供重建的背景影像與新進影像最小差異度的候選移動向量為最佳移動向量，之後利用該最佳移動向量將背景高斯混合模型移動到正確的位置以對齊新進影像，再作前景背景的分割，接著再消除面積過小的雜訊，並試著將剩下仍被判斷成前景的像素與背景高斯混合模型中相對應位置像素的鄰近像素的模型做配對，以消除攝影機些微旋轉對影像造成的影響。第五章顯示本方法的實驗結果，第六章則說明未來可能的研究方向。



二、 建立用於影像對齊之重建的背景影像

在本章中，我們在2.1節介紹高斯混合模型如何在新進影像已經對齊高斯混合模型的前提下，達成前景背景分割的目的。並在2.2節中介紹如何利用高斯混合模型來建立對齊高斯混合模型與新進影像所需的重建的背景影像，在有了重建的背景影像之後，才能建立對齊所需要的以顏色為基礎的群組、用邊線所切割出的區塊以及用角點偵測器所獲得之角點。

2.1 高斯混合模型

Grimson[1]在1999年所提出的Gaussian Mixture Model (GMM)，由於其對光影變化以及背景規律性運動現象的處理能力優秀，已經成為靜態式攝影機下前景背景分割的重要方法。因此本論文亦想借重其能力，讓前景背景分割的結果更加地穩定。

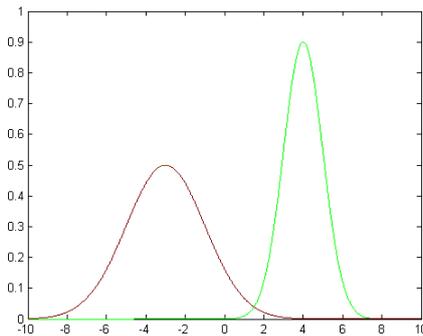
GMM的精神為對每一個像素建立數個高斯模型，代表該像素曾經出現過的顏色，例如在樹影搖曳的環境下，一個正好處在樹葉跟樹枝擺盪之間的像素就會有代表綠色樹葉以及褐色樹枝的兩個高斯模型，如圖2.1所示（值得一提的是，Grimson提出來的方法主要是利用強度 (intensity) 來建立高斯模型，因此在這裡說綠色跟褐色的高斯模型其實並不精準，但我們可以對GMM做一點修改，將其維度從一維的強度增加至三維的RGB，使其亦能處理彩色的狀況）。而這數個高斯模型又有各自的權重，一個權重可以被理解為其對應的高斯模型的出現頻率，越常出現的高斯模型其權重會越高，表示其出現頻率高。

假設一個像素已經有了數個高斯模型，在新的影像進來後，我們想判斷在新進影像中該位置的像素是否為新出現的前景，對每一個該像素所擁有的高斯模



(a)

(b)



(c)

圖2.1 一個高斯模型的例子。(a)與(b)為樹影搖曳的圖片，其中位於紅點的像素的高斯混合模型如(c)所示。

型，以下式判斷該像素是否能夠配對 (match) 到該高斯模型

$$(i - \mu_t)^2 \leq 2.5 \times \sigma \quad (2.1)$$

其中 i 為像素的強度， μ_t 為該高斯模型的平均強度， σ 為該高斯模型的標準差，如果可以配對則以式子 (2.2)、式子 (2.3) 和式子 (2.4) 更新該模型的權重、平均強度以及變異量 (variance)

$$\omega_{k,t} = (1 - \alpha) \omega_{k,t-1} + \alpha \quad (2.2)$$

$$\mu_t = (1 - \rho) \mu_{t-1} + \rho X_t \quad (2.3)$$

$$\sigma_t = (1-\rho) \sigma_{t-1} + \rho (X_t - \mu_t)^2 \quad (2.4)$$

其中 α 以及 ρ 分別為指定的更新速度， X_t 為該新進影像像素的強度，但如果無法配對到該高斯模型，則以下式更新該模型的權重即可

$$\omega_{k,t} = (1-\alpha) \omega_{k,t-1} \quad (2.5)$$

無需更新該模型的平均強度以及變異量。在每個高斯模型都測試完畢後，正規化每個高斯模型的權重，使其總合為固定數值。接著我們選出 B 個權重除以變異量最大的高斯模型，使其權重合小於等於指定的閾值 T ，而這 B 個高斯模型就代表該像素最有可能的背景模型。選擇權重除以變異量最大的模型的理由是因為這樣的模型必然有著很大的權重或是很小的變異量，很大的權重表示該模型常出現，因此會是背景的機率較大；而很小的變異量表示該模型的顏色很穩定，穩定的模型會是背景的機率也較大。最後我們判斷該新進像素的強度是否至少能夠配對到這 B 個模型的其中一個，如果是，則表示該像素為背景，否則為前景。

考慮到在真實的應用下我們不可能給每一個像素無限制數量的高斯模型，因此通常我們會指定一個數目 K ，表示一個像素至多只能有 K 個高斯模型。因此若發現一個新進像素無法配對到任何一個已經存在的高斯模型（如此則該新進像素必然為前景，因為無法配對到任何一個已經存在的模型，自然也不會配對到上一個段落中最後提到的 B 個最有可能的背景模型），我們需要為該新進像素建立一個高斯模型。如果已經存在的高斯模型數量少於 K ，則我們直接新增一個新的高斯模型，其強度為該新進像素的強度，但是權重和變異量則需要手動指定一個初始值，因為我們無法僅靠一個樣本算出變異量。但若已經存在了 K 個高斯模型，則我們刪除權重除以變異量最小的高斯模型。選擇權重除以變異量最小的模型的理由是因為這樣的模型必然有著很小的權重或是很大的變異量，很小的權重表示該模型不常出現，通常是過去曾經出現過的前景，但是出現的次數很少，或是很

久沒再出現過了，因此可以推論其不會是背景的機率很大，所以在必須要刪除一個模型的要求下，我們選擇刪除這樣的模型，而很大的變異量表示該模型的顏色很不穩定，不穩定的模型通常也不會是背景，因此也可以刪除之。在刪除了一個最不可能代表背景的高斯模型後，我們新增一個強度為該新進像素的強度，但是權重和變異量為手動指定的高斯模型。

在真正讓GMM處理前景背景分割的工作前，最好讓GMM先在純背景的環境下執行一段時間，這樣才可以讓每一個像素建立出所有可能代表該像素的背景高斯模型，以及計算出每個高斯模型各自的變異量。例如在上述樹枝擺盪的例子中，我們需要足夠數目的影格，讓這些影格足以包含到樹葉和樹枝晃動的情況，使位在其中的像素可以有機會建立代表綠色樹葉與褐色樹枝的高斯模型。

2.2 建立重建的背景影像



由於我們的方法需要利用GMM來做前景背景的分割，所以為了將GMM整合進我們的方法，需要對GMM做一點修改。第一個需要修改的地方是之前提過的，將一維的強度推廣到三維的RGB，讓GMM能夠處理彩色的影像。修改的方法很簡單，對於每一個高斯模型，我們僅需將代表強度的數值重複三份，並將每一份數值的意義分別改成RGB，而當測試一個像素能否配對到該高斯模型時，將式子(2.1)分別對RGB三個色頻做測試，唯有當三個色頻都能配對成功的時候，該像素才能被視為配對成功，否則為配對失敗。至於平均顏色與變異量的更新則對RGB三個色頻分別更新即可。另外在第三章中需要每一個背景像素都提供自己的背景顏色，以建立如圖1.1(a)與圖1.1(g)的重建的背景影像，如此才能利用該重建的背景影像來建立背景群組、區塊以及角點。如果我們真正拿每個像素的B個最能代表背景的高斯模型的顏色來做排列組合，以求出所有可能的群組和區塊，這樣的計算量會十分龐大，因此我們限制一個像素在提供背景顏色的時候，只提

供最可能代表背景的高斯模型的顏色，即權重除以變異量最大的高斯模型所擁有的顏色。

以圖2.2為例子，當圖2.2(a)的新進影像在經由本論文所提出的方法做前景分割後，得到的結果如圖2.2(b)所示。之後除了GMM本身的更新機制外，再加上本論文建議的強制設定背景模型法（詳細的過程與說明請參考4.4節：新進與移出的背景像素之處理），可得到如圖2.2(c)所示的重建的背景影像，而該重建的背景影像可以在下一張新進影像要做前景背景分割時，提供建立以顏色為基礎的群組、用邊緣所切割出的區塊以及用角點偵測器所獲得之角點的依據。

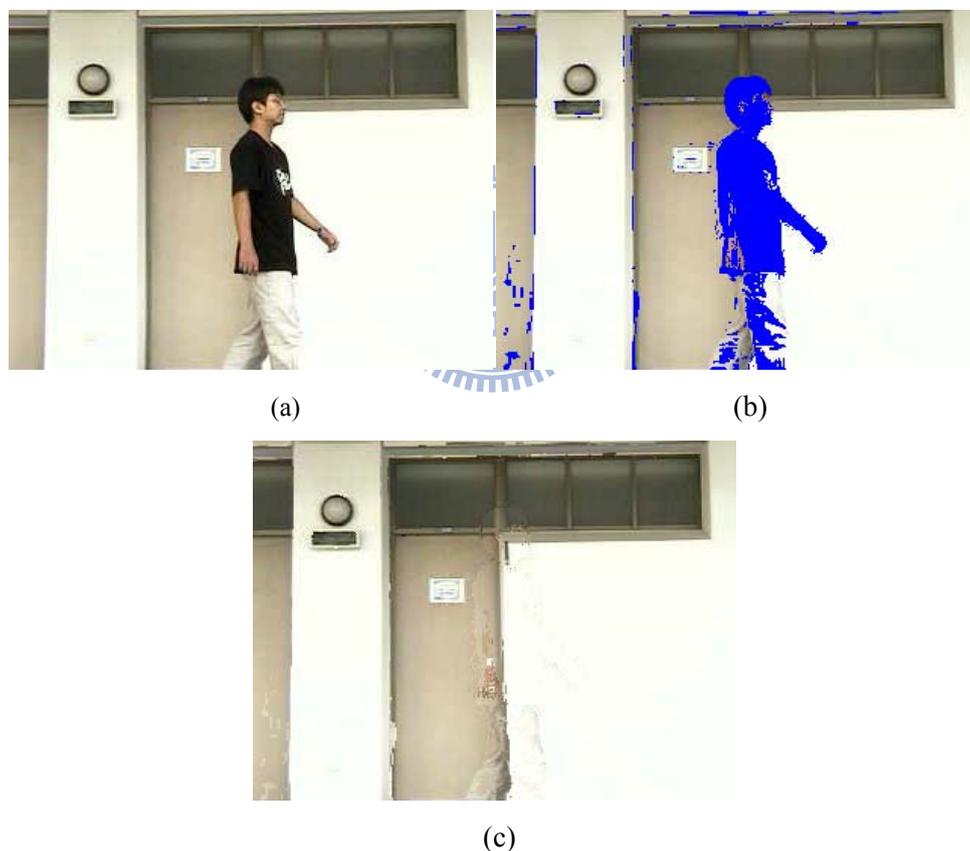


圖2.2 GMM提供的重建的背景影像。(a) 新進影像，(b) (a)的前背景分割結果，(c) 結合了GMM本身的更新機制以及本論文建議的強制設定背景模型法，所得到的重建的背景影像。

三、三種可用於影像對齊之影像特徵擷取

由於本論文的主要精神為先找出重建的背景影像（由GMM提供）與新進影像間的最佳移動向量，之後再將背景GMM根據此最佳移動向量移動到新的位置，意即我們對齊GMM與新進影像，最後利用GMM達到前景背景分割的目的，因此我們需要提供可行的方法來尋找GMM與新進影像間的最佳移動向量，如此才能對齊GMM與新進影像。第一個可行的方法為找出重建的背景影像與新進影像中顏色相似且空間上相連的區域，如圖1.1(c)與(d)中紅色矩形所包住的人行道上的磁磚，接著計算這兩個區域各自的質心座標，最後將這兩個區域的質心座標相減，就是重建的背景影像與新進影像之間一個可能的移動向量。對此本論文提出兩個不同的方法來尋找影像中的區域：以顏色為基礎的群組以及用邊線所切割出的區塊，3.1節將介紹以顏色為基礎的群組方法，而3.2節將介紹如何找出用邊線所切割出的區塊。第二個可行的方法為找出重建的背景影像與新進影像之間的角點，如圖1.1(c)與(d)中紅色圓圈所包住的窗戶角落，之後將這兩個角點的座標相減，同樣也是一個重建的背景影像與新進影像之間可能的移動向量，3.3節將介紹尋找角點以及配對角點的方法。

3.1 以顏色為基礎的群組

第一個可行的移動向量尋找方法為對重建的背景影像（由背景GMM提供）與新進影像建立各自的顏色相近群組模型，再利用附錄1的連通物件搜尋法將空間上不相鄰的群組模型分開成獨立的群組，如此群組所代表的意義就是影像中顏色相近且空間上相連的區域，接著配對新進影像與重建的背景影像的群組，最後將所有配對成功的群組的質心座標相減，以估算出所有可能的候選移動向量。因

此在本節我們介紹對重建的背景影像以及前景影像中的每個物件(即相連在一起的相似顏色的區域)建立以顏色為基礎的群組模型,之後再利用附錄1的連通物件搜尋法建立以顏色為基礎的群組。

我們利用區域之間顏色不同的特性,設計了一個群組影像區域的方法。建立的步驟如下:首先將影像中位置為(0,0)的像素建立成第一個群組模型,這個群組模型中有該像素的RGB三個顏色值,但在往後的過程中,這三個顏色值會一直被更新,其代表的意義便成為所有屬於該群組模型的像素之平均顏色值。接下來對影像中所有的像素,逐一判斷該像素能不能配對到任何一個已經存在的群組模型。判斷的方法直接採用下式

$$| (I_R - V_R) | < T_l \text{ 且 } | (I_G - V_G) | < T_l \text{ 且 } | (I_B - V_B) | < T_l \quad (3.1)$$

其中 I_R 表示該像素的紅色值, V_R 表示該群組模型的紅色值, I_G 表示該像素的綠色值, V_G 表示該群組模型的綠色值, I_B 表示該像素的藍色值, V_B 表示該群組模型的藍色值, T_l 為一個閾值,經由實驗發現,若 T_l 設為10能得到較好的群組模型結果。

如果RGB三個色頻都可以滿足式子(3.1),表示該像素與該群組模型可以成功配對。若該像素只配對到一個群組模型,則我們將該像素分類到該群組模型,並且更新該群組模型的顏色值,更新的方法為對RGB三個色頻各自採用下式

$$V_{updated} = (1 - \sigma) \times V_{old} + \sigma \times I \quad (3.2)$$

其中 V 為群組模型的顏色值, I 為新進像素的顏色值, σ 為更新速度,經由實驗發現,將 σ 設為0.1能得到較好的群組模型結果。但若該像素無法配對到任何一個已經存在的群組模型,那麼我們就為這個像素建立一個新的群組模型,而該新建立的群組模型的顏色值則指定成該像素的顏色。若該像素同時配對到數個已經存在

的群組模型，我們選出最適合該像素的群組模型當作該像素所屬的群組模型，選擇的依據採用下式

$$R = (I_R - V_R)^2 + (I_G - V_G)^2 + (I_B - V_B)^2 \quad (3.3)$$

因為最小的 R 值表示該群組模型和該像素最為接近，所以我們選擇能提供最小 R 值之群組模型為最適合該像素的群組模型，並將該像素歸類給該群組模型，最後同樣以式子(3.2)更新該群組模型的RGB值，如此一來影像中只要是顏色相近的像素都會被歸類到同一個群組模型。圖3.1(a)為一原始影像，影像中有多個不同顏色的物體，也有顏色相同的物體。圖3.1(b)則為圖3.1(a)經由以上介紹的方法處理過後的結果，我們對每一個群組模型給定隨機的顏色，因此不同顏色的像素代表被分在不同的群組模型裡，方便視覺驗證。

注意到在圖3.1(a)中的兩顆黑球在圖3.1(b)中被歸類到同一個群組模型，但是我們希望能夠將這兩顆黑球歸類到不同的群組，因為若將空間上不相連的像素分類到一個群組，很多資訊都將變得毫無意義，例如群組的質心以及 x 與 y 軸方向的變異量。且有可能在原來的影像中存在兩個或多個不相連但顏色相似的區域，我們把它們歸類成一個群組模型，並且計算其質心、面積以及 x 與 y 軸的變異量，但在接下來的影格，因為攝影機移動的關係，導致其中一個區域消失，因此利用剩下來的區域所計算出來的質心、面積以及 x 與 y 軸的變異量將變得與之前算出來的完全不同，而無法利用。由此可知，將空間上不相連的區域分成不同的群組是必須的，這樣當我們希望利用群組的質心、面積以及 x 與 y 軸的變異量來配對兩個群組時，才能將出錯的機率降到最低。對此我們利用附錄1的連通物件搜尋法將空間上不相鄰但屬於同一個群組模型的區域分開成兩個群組，而分開後的結果如圖3.1(c)所示。

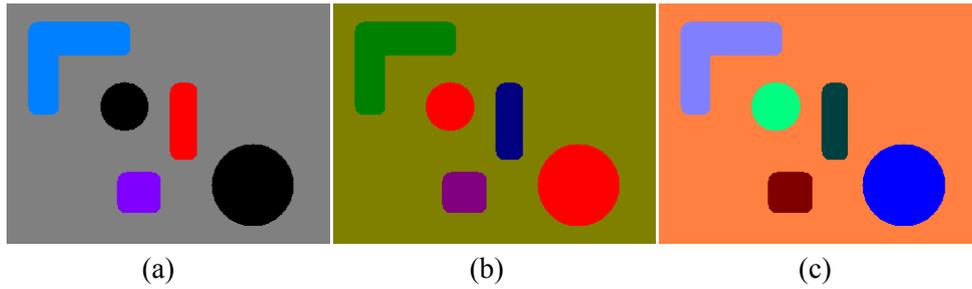


圖3.1 群組的結果。(a) 原始影像，(b) (a)用群組模型來模擬空間的顏色分佈，(c) (b)做連通物件搜尋法後的結果。

群組模型顏色值更新與否對整體群組模型的結果影響很大，在有更新的情況下，建立出的群組模型較能符合人眼對區域的認知，若沒有更新，則建立出來的群組模型會比有更新的结果更細碎，而且因為群組模型一開始的顏色值被設定為第一個建立該群組模型的像素顏色值，這樣的情況將使得群組模型在稍為有光影影響下會變化地非常劇烈，而且也很容易受到雜訊的干擾。圖3.2顯示群組模型顏色值在有更新及無更新的情況下，建立出的群組模型差異。

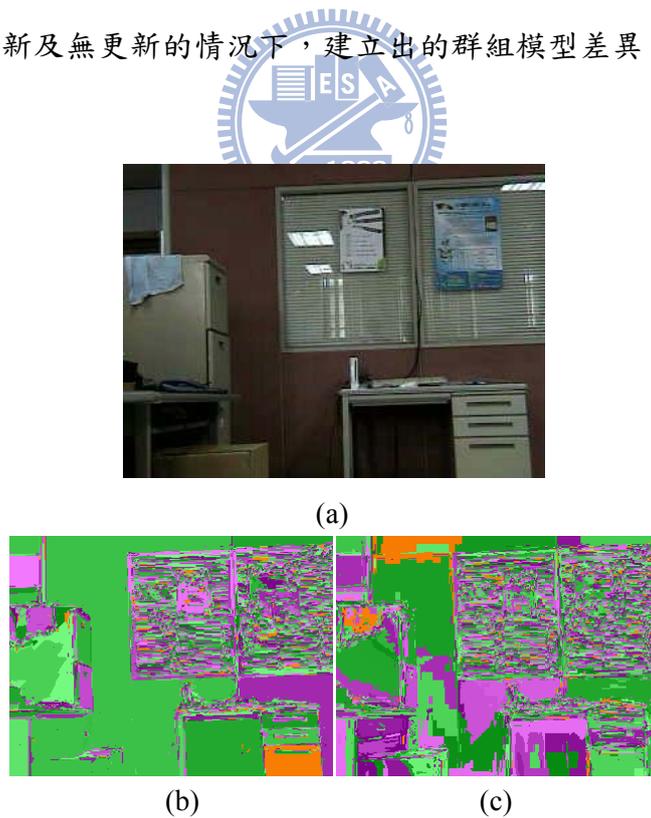


圖3.2 群組模型顏色值有無更新對群組模型結果的影響。我們透過給予每個群組模型隨機的顏色來達到視覺化驗證的目的。(a) 重建的背景影像，(b) $\sigma=0.1$ 的背景群組模型結果，(c) 不更新群組模型顏色值的群組模型結果。

3.2 用邊線所切割出的區塊

有鑑於3.1節所介紹的以顏色為基礎的群組在真實環境下受到光影變化的影響太大，我們提出第二個尋找影像中區域的方法。此方法先利用邊線偵測器（edge detector）找出影像中所有的邊線，接著再利用附錄1所介紹的連通物件搜尋法找出用邊線切割出的區塊（或說被邊線所包圍的區塊）。在本論文中，我們選擇使用文獻[10]中第三章第七節所介紹的Sobel edge detector來尋找邊線。Sobel edge detector的作法為：對於影像中的每一個像素，如圖3.3中所示的 Z_5 ，以下式判斷該像素是否為邊線上的像素

$$\sqrt{(v_1)^2 + (v_2)^2} \geq T \quad (3.4)$$

其中 v_1 為圖3.3(a)與圖3.3(b)做點對點相乘後相加的結果， v_2 為圖3.3(a)與圖3.3(c)做點對點相乘後相加的結果， T 為一個閾值。經由實驗發現，若 T 設為40能得到較好的邊線偵測結果，如果該式成立，則代表 Z_5 為一個邊線像素。如此找完了一張影像上所有的邊線像素後，再利用附錄1所介紹的連通物件搜尋法，找出所有用邊線切割出來的區塊。這些用邊線切割出來的區塊，與3.1節所介紹的以顏色為基礎的群組一樣，都可以提供新進影像與背景GMM間可能的移動向量。

Z_1	Z_2	Z_3
Z_4	Z_5	Z_6
Z_7	Z_8	Z_9

-1	-2	-1
0	0	0
1	2	1

-1	0	1
-2	0	2
-1	0	1

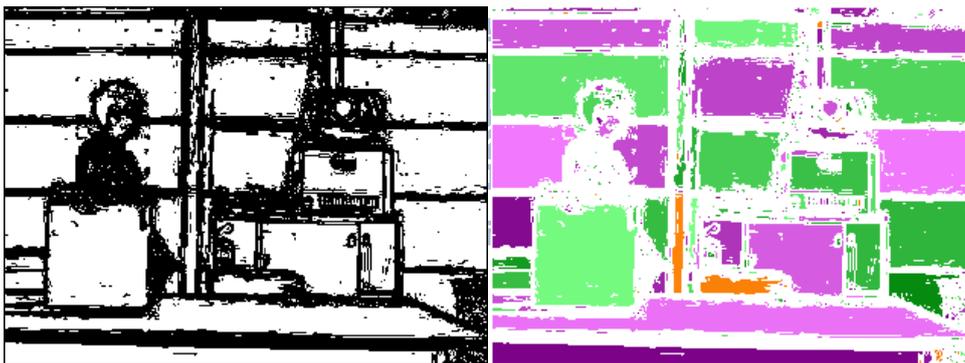
(a)
(b)
(c)

圖3.3 Sobel edge detector。(a) 像素的索引(index)，其中 Z_5 為要測試是否為邊線像素的像素，(b) 測試是否為水平邊線像素的mask，(c) 測試是否為垂直邊線像素的mask。

圖3.4為用邊線所切割出的區塊的結果，圖3.4(a)為一輸入影像，圖3.4(b)為利用Sobel edge detector所找出來的邊線，圖3.4(c)為圖3.4(b)利用附錄1的連通物件搜尋法所找出的區塊，我們透過給予每個區塊隨機的顏色來達到視覺化驗證的目的。



(a)



(b)

(c)

圖3.4 用邊線所切割出的區塊的結果。(a) 輸入影像，(b) 利用Sobel edge detector所找出來的邊線，(c) 利用附錄1的連通物件搜尋法所找出的區塊。

在大部分的情況下，用邊線所切割出來的區塊都有不錯的穩定性，因為相較於群組模型，區塊對光影變化的容忍度相當高，但是偶爾還是會遇到失敗的情況。例如我們曾經發現攝影機拍出來的影像會突然變模糊，這樣邊線就會不明顯，切割出來的區塊就有可能會跟從重建的背景影像切割出來的區塊差很多，這時第三章以顏色為基礎的群組模型就可以發揮其作用。原因是：假設現在因為影像變模糊的關係，存在兩個顏色差異頗大的區域，中間是以模糊的漸層相連，這個時

候Sobel edge detector會因為該模糊漸層內相鄰的像素顏色太過於相似，而無法在這個漸層內找出一條邊線將這兩個區域切開。但是以顏色為基礎的群組模型就可以避免將這兩個區域分在同一個群組模型，因為群組模型的顏色是漸漸地更新，當由其中一個區域所建立出來的群組模型想繼續往另外一個區域擴張時，會因為該群組模型的顏色更新速度趕不上漸層顏色的變化速度，而停在該漸層上，不會將另一個區域也一起併入該群組模型。圖3.5顯示這個現象。

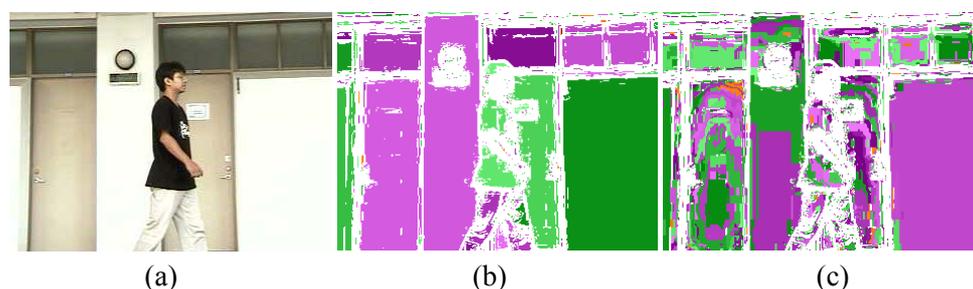


圖3.5 群組可以成功地切出區域但區塊卻不行的例子。(a) 原始影像，(b) 利用邊線切出來的區塊，其中白色的像素表示邊線像素，注意到右上角的兩個窗戶同屬於一個區塊，(c) 使用群組後的結果，右上角的窗戶成功地被分開。

至此我們已經能夠對重建的背景影像跟新進影像建立以顏色為基礎的群組以及用邊線所切割出的區塊，在第四章中我們將說明如何配對重建的背景影像及新進影像間的群組和區塊來找出所有可能的候選移動向量。

3.3 Harris Corner Detector與角點配對

第三個可行的對齊方法為配對影像中的角點以估測影像間的移動向量，在本論文中我們利用文獻[11]所提出的Harris corner detector來偵測角點，之後再對每個偵測出來的角點建立各自的描述子，以利判斷在兩張不同影像中的角點是否可以視為同一個角點。因此當我們想要估測重建的背景影像相對於前景影像的移動向量時，將重建的背景影像中所有的角點與新進影像中所有的角點做配對，如果有任兩個角點能配對成功，則將該新進影像中角點的座標減掉配對成功的重建

的背景影像中的角點座標，即為一個可能的候選移動向量。

由於本論文應用了Harris corner detector來尋找影像中的角點，因此我們將在本段中介紹Harris corner detector的運作方式，並且將其作一點修改，使其能夠處理彩色的影像。當我們想判斷影像中一個像素是否為角點時，對該像素的RGB計算各自的 x 與 y 方向的梯度， x 方向的梯度可用下式計算

$$\partial I / \partial x = I \otimes (-1, 0, 1) \quad (3.5)$$

其中 I 表示像素的顏色值， \otimes 表示旋積運算子 (convolution)，而 y 方向的梯度則用下式計算

$$\partial I / \partial y = I \otimes (-1, 0, 1)^T \quad (3.6)$$

之後再進一步對該像素的RGB色頻各自建立一個 2×2 的對稱矩陣 M 如下

$$M = \begin{pmatrix} A & C \\ C & B \end{pmatrix}$$

其中 A 、 B 、 C 定義如下

$$A = \sum_W \left(\frac{\partial I}{\partial x} \right)^2 \quad B = \sum_W \left(\frac{\partial I}{\partial y} \right)^2 \quad C = \sum_W \left(\frac{\partial I}{\partial x} \right) \left(\frac{\partial I}{\partial y} \right)$$

其中 W 表示以該像素為中心的一個矩形範圍，在本論文中將 W 設定為以該像素為中心的九宮格。接下來以下式對該像素的RGB色頻各自計算一個 R 值

$$R = \text{Det}(M) - k \cdot [\text{Tr}(M)]^2 \quad (3.7)$$

其中

$$\text{Det}(M) = AB - C^2$$

$$\text{Tr}(M) = A + A$$

當任一個色頻的 R 為一個夠大的正數時，表示該像素為一個角點。在本論文中，將常數 k 設定為0.04，因此式子(3.7)代入 k 並且展開後可以改寫如下式

$$R = 0.92 AB - C^2 - 0.04 (A^2 + B^2) \quad (3.8)$$

根據實驗得到的結果，將 R 大於 10^9 的像素判斷為角點能夠得到比較合理的角點偵測結果，我們將不同的 R 值所得到的角點偵測結果整理於圖3.6，從圖中可以看出過小的 R 值會找出過多的角點，而太大的 R 值則會造成找出的角點數量太少。

由於Harris corner detector對一個影像中的角落所偵測出來的角點往往不只一個，而是一群角點圍繞著該角落，這會造成影像中偵測出的角點太多，使得在做角點配對的時候會需要過多的時間，而且圍繞在一個角落周圍的角點彼此間的鑑別度不高，這會造成許多錯誤的配對，而使得找出的候選移動向量的數目大增，關於角點配對的詳細方法，我們將在本節的最後兩段說明。為了消除角點群聚的現象以及維持角點必須在角落上的要求，我們利用附錄1的連通物件搜尋法將所有相連在一起的角點找出，之後對每一群相連在一起的角點選出 R 值最大的角點為真正的角點，如此便可大大地減少角點的數量，進而增加角點配的效率以及降低錯誤配對的機率。圖3.7(a)為一個將 R 值設為 10^9 所得到的角點偵測結果，圖3.7(b)為消除角點群聚現象後的結果。

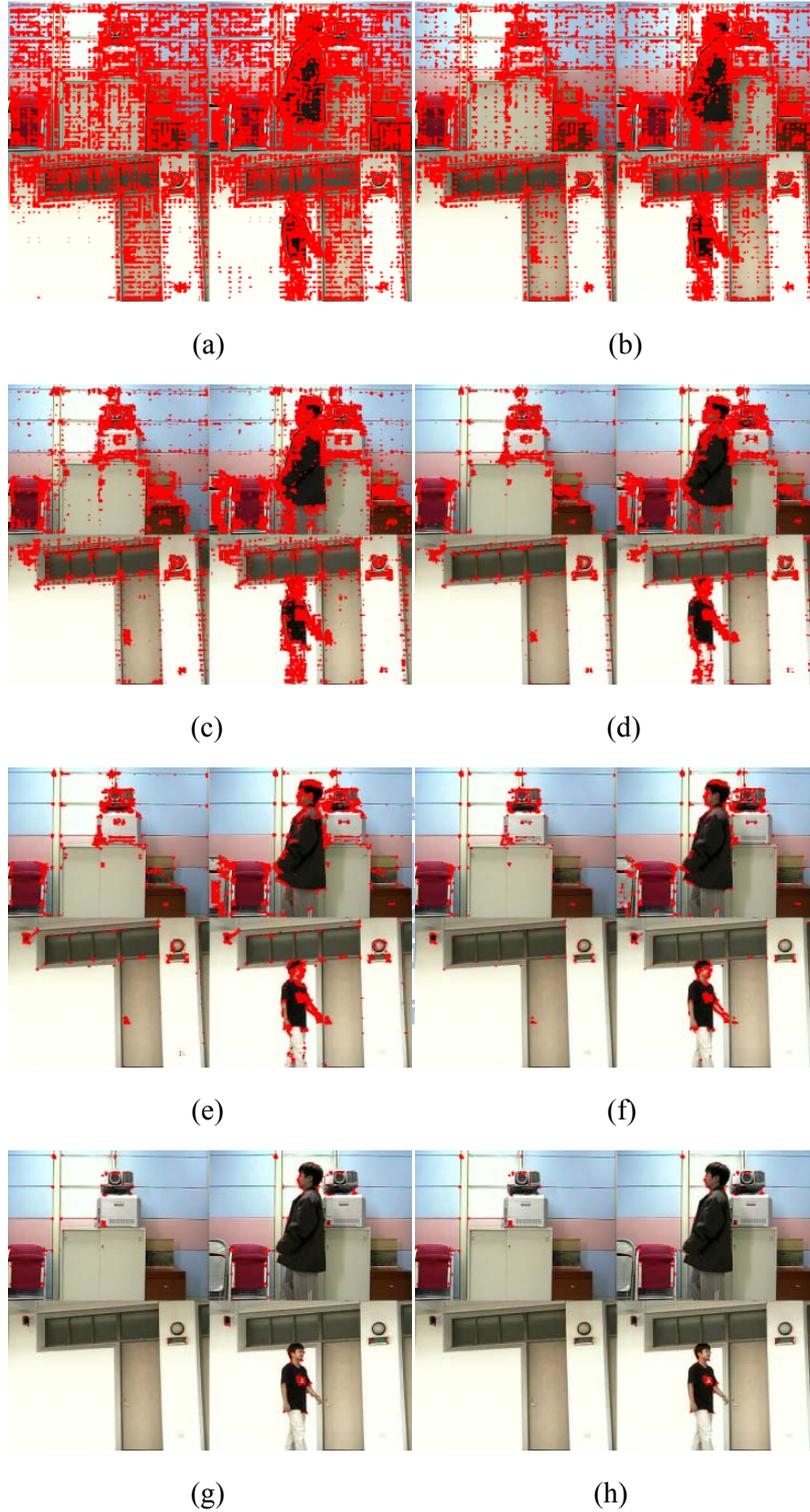


圖3.6 不同的 R 值所得到的角點偵測結果。(a) $R=10^4$, (b) $R=10^5$, (c) $R=10^6$, (d) $R=10^7$, (e) $R=10^8$, (f) $R=10^9$, (g) $R=2 \times 10^9$, (h) $R=3 \times 10^9$ 。



(a)



(b)

圖3.7 消除角點群聚現象的結果。(a) R 值為 10° 所得到的角點偵測結果，(b) 消除角點群聚現象後的結果。

在找完了一張影像所有的角點後，我們對每個角點建立各自的描述子，當做之後配對角點的依據。本論文建立描述子的方法很簡單且直覺，我們使用以該角點為中心且面積為 7×7 的正方形內所有像素正規化後的顏色值為該角點的描述子，由於一個像素有RGB三個色頻，因此每個角點的描述子都會有 $7 \times 7 \times 3$ 個數值。其中正規化的方法為先對RGB三個色頻各自算出在 7×7 正方形內所有像素的平均顏色值，再將該 7×7 正方形內所有的顏色值減掉這個平均值，即為正規化後的顏色值。正規化的目的是希望能夠降低光影變化對描述子造成的影響，因為突然的光影變化會造成一個描述子內所有的顏色值都加上或減掉一個數值，但是這些數值之間的差異並沒有改變，因此加上正規化的步驟則可以抵銷因為光影變化所造成的加上或減掉一個數值的影響。

之後當要判斷兩個來自於不同影像的角點是否為同一個角點時，我們判斷兩個角點描述子內相同顏色值的像素個數是否有超過一個指定的數目，如果有即可判斷兩個角點配對成功，意即可視為同一個角點，反之不為同一個角點。經由實驗發現，將判斷兩個顏色值是否相同的條件設為RGB三個正規化後的色頻差值皆須小於30，且兩個角點描述子內相同顏色值的像素數目須大於描述子內像素數目的三分之二才可將兩個角點視為同一個角點（由於本論文中一個描述子有 $7 \times 7 = 49$ 個像素，因此兩個角點的描述子至少需要有33或34個相同顏色值的像素，才會被視為同一個角點），能夠得到較穩定的配對結果。最後再將每一組配對成功的兩個角點的座標值相減，就是重建的背景影像與新進影像間可能的候選移動向量。

四、影像對齊與前景背景分割

在3.1節與3.2節中說明了如何建立新進影像與重建的背景影像的以顏色為基礎的群組與用邊線所切割出的區塊，下一個步驟是藉由配對這些群組與區塊來尋找新進影像與重建的背景影像間的候選移動向量，這個部分我們將在4.1節中介紹。在找出了利用4.1節的方法與3.3節的方法所得到的候選移動向量後，我們需要從這些候選移動向量中找出最佳移動向量，當作對齊新進影像與背景GMM的依據，4.2節將介紹如何找出最佳移動向量。4.3節介紹在有了最佳移動向量後，我們對齊新進影像與背景GMM，以達到前景背景分割的目的，並試著消除雜訊與攝影機些微旋轉對前景背景分割造成的影響。最後，我們在4.4節中說明如何處理新進與移出的背景像素，以建立新的重建的背景影像，讓下一張新進影像能有對齊的依據。



4.1 尋找影像間之候選移動量

當要對一張攝影機擷取到的影像作前景背景分割時，依照3.1節、3.2節及3.3節所介紹的方法，對新進影像建立群組、區塊以及角點，同時對GMM亦要求其提供最能代表每個背景像素的顏色，以建立重建的背景影像，再對此重建的背景影像建立群組、區塊以及角點。現在，我們有兩組群組、兩組區塊以及兩組角點，分別來自於新進影像與重建的背景影像。

利用角點來尋找候選移動向量的方法已於3.3節中介紹過，因此在這裡就不再重複敘述，下面我們將介紹如何藉由各自配對群組及區塊，以求出所有可能的候選移動向量。依照人眼對於移動向量的計算方式，要判斷下一個影像相對於上一個影像移動了多少，人眼會找出影像中幾乎一模一樣的部分，進而判斷出相對

的移動向量，本章節的步驟也是遵循這樣的想法。而且在真實的應用上，給定的影像往往是時間上相鄰的，這意味著兩張影像之間至少有部分場景是重複出現的，而非截然不同的場景，所以可以假設這兩個影像所建立出來的群組、區塊或角點至少有一個是重複出現的，且在面積、形狀及顏色上變化很小。因此我們對新進影像跟重建的背景影像所建立出的群組以及區塊，先刪除面積過小的群組和區塊，再以顏色差異、面積差異以及x跟y軸的變異量差異判斷新進影像與重建的背景影像中的兩個群組是否可以視為同一個群組，或兩個區塊是否可以視為同一個區塊，判斷的方法為式子（4.1）、式子（4.2）以及式子（4.3）都需要成立

$$(I_R - V_R) < T_2 \text{ 且 } (I_G - V_G) < T_2 \text{ 且 } (I_B - V_B) < T_2 \quad (4.1)$$

$$|(S_b - S_f)| < T_3 \quad (4.2)$$

$$Var_b(x) = Var_f(x) \text{ 且 } Var_b(y) = Var_f(y) \quad (4.3)$$

其中 I_R 表示重建的背景影像的群組或區塊的平均紅色值， V_R 表示新進影像的群組或區塊的平均紅色值， I_G 表示重建的背景影像的群組或區塊的平均綠色值， V_G 表示新進影像的群組或區塊的平均綠色值， I_B 表示重建的背景影像的群組或區塊的平均藍色值， V_B 表示新進影像的群組或區塊的平均藍色值， T_2 為一個閾值， S_b 為重建的背景影像的群組或區塊的面積， S_f 為新進影像的群組或區塊的面積， T_3 為一個閾值， $Var_b(x)$ 為重建的背景影像的群組或區塊的x變異量， $Var_f(x)$ 為新進影像的群組或區塊的x變異量， $Var_b(y)$ 為重建的背景影像的群組或區塊的y變異量， $Var_f(y)$ 為新進影像的群組或區塊的y變異量。式子（4.1）表示兩個群組或區塊之間的顏色差異不能太大，式子（4.2）則表示兩個群組或區塊之間的面積差異也不能過大，式子（4.3）則規定兩個群組或區塊在形狀上必須有一定程度的相似性。

若兩個群組被判斷為同一個群組，或是兩個區塊被判斷為同一個區塊，則計算面積差異值 v_1 ，計算的方法為直接將兩個群組或區塊的面積相減，再取絕對值。對於算出來的兩個群組的面積差異 v_1 ，將此差異值除以該兩個群組中面積較大的群組的面積，再乘上100，得到一個值 v_2 ；而對於兩個區塊的面積差異 v_1 ，同樣將此差異值除以該兩個區塊中面積較大的區塊的面積，再乘上100，得到一個值 v_2 。 v_2 代表的意義可以視為該群組或區塊在新進影像以及重建的背景影像間的面積變化比例，因此 v_2 必然小於等於100。若 v_2 等於0表示該群組或區塊在新進影像與重建的背景影像中幾乎沒有變化，因此很有可能可以提供準確的移動向量，而把兩個群組或區塊的質心座標相減就是一個可能的候選移動向量。

在對所有的群組以及區塊做完以上的計算後，可以得到許多 v_2 ，但我們只注意數值小的 v_2 ，最理想的情況為0，這樣表示該 v_2 是從非常穩定的群組或區塊所得到的結果。在本論文的實驗中，我們將所有 v_2 小於等於5的群組或區塊所提供的移動向量當成可能的移動向量，因此會得到數個候選移動向量。

在真實環境不穩定的考量下，藉由群組或區塊的質心相減所得到的候選移動向量可能無法十分精準，為了得到最準確的移動向量，建議可以對每個找出來的候選移動向量上下左右各擴大兩個像素，亦即一個候選移動向量會衍生出25個候選移動向量。例如目前有一個候選移動向量其值為(100, 50)，表示重建的背景影像相對於新進影像在 x 方向移動了100，且在 y 方向移動了50，我們對其上下左右各擴大兩個像素，因此除了(100, 50)這個候選移動向量，會另外新增(98, 48)、(99, 48)、(100, 48)、(101, 48)、(102, 48)、(98, 49)、(99, 49)、(100, 49)、(101, 49)、(102, 49)、(98, 50)、(99, 50)、(101, 50)、(102, 50)、(98, 51)、(99, 51)、(100, 51)、(101, 51)、(102, 51)、(98, 52)、(99, 52)、(100, 52)、(101, 52)、(102, 52)這24個候選移動向量。

4.2 找出最佳移動向量

接著我們對4.1節中所找出來的候選移動向量各自作統計，找出能提供新進影像與重建的背景影像間最小差異度的候選移動向量為最佳移動向量。首先對每一個候選移動向量，將原來的重建的背景影像根據該候選移動向量移動到新的位置，意即我們根據該候選移動向量對齊重建的背景影像與新進影像，並且計算一個差異值跟一個總數值，差異值表示新進影像與重建的背景影像的差異度，總數值表示同時存在於新進影像與重建的背景影像的像素數目，也就是有重疊或重複出現的像素數目。差異值與總數值一開始為0，接下來對新進影像中每個像素，先以該候選移動向量判斷該像素是否有在重建的背景影像的範圍內，若不在重建的背景影像的範圍內則跳過該像素，繼續處理下一個像素；但若是重建的背景影像的範圍內則將總數值加一，再依據下式計算差異值

$$D = \sum_x \sum_y | R_b(x,y) - R_f(x,y) | + | G_b(x,y) - G_f(x,y) | + | B_b(x,y) - B_f(x,y) | \quad (4.4)$$

其中 D 表示差異值， $R_b(x,y)$ 表示重建的背景影像中座標為 (x,y) 的像素的紅色值， $R_f(x,y)$ 表示新進影像中座標為 (x,y) 的像素的紅色值， $G_b(x,y)$ 表示重建的背景影像中座標為 (x,y) 的像素的綠色值， $G_f(x,y)$ 表示新進影像中座標為 (x,y) 的像素的綠色值， $B_b(x,y)$ 表示重建的背景影像中座標為 (x,y) 的像素的藍色值， $B_f(x,y)$ 表示新進影像中座標為 (x,y) 的像素的藍色值。至於式子(4.4)所包含的範圍則如圖4.1中黑色矩形所包含的所有像素，為將圖1.1(a)與圖1.1(b)對齊後的示意圖，其中黑色的矩形表示同時存在於圖1.1(a)與圖1.1(b)的像素，因此差異值所代表的意義即為黑色矩形中圖1.1(a)與圖1.1(b)的差異度，而總合值則代表該矩形的面積。

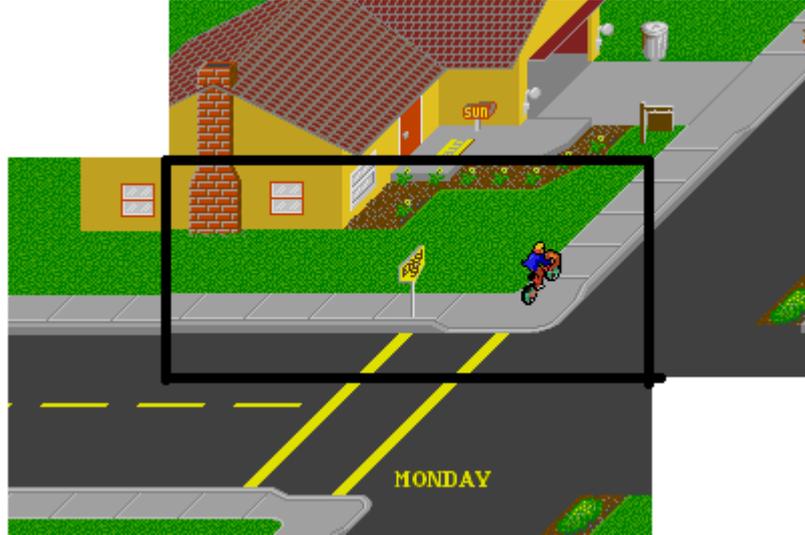


圖4.1 差異值與總數值的計算範圍。

在計算完所有候選移動向量各自的差異值以及總數值後，將差異值除以總數值，即代表在一個候選移動向量下，每個像素在新進影像與重建的背景影像的平均差異度。我們選擇能提供最小平均差異度的候選移動向量為真正的移動向量。這樣的理由是對於錯誤的候選移動向量，因為無法對齊重建的背景影像到新進影像，將會導致原來該對應到新進影像中某個位置的背景像素被移動到其他的位置，使得新進影像中該位置的像素會被對應到錯誤的背景像素，這意味著新進影像中的像素與對應到的背景像素的顏色可能會差非常多，因為它們可能分別屬於場景中不同的物件，進而造成式子(4.4)中絕對值內的差值變得很大，使得差異值也會變大很多。因此對於錯誤的候選移動向量，其差異值一定會大於最佳移動向量所計算出來的差異值。

圖4.2表示不準確的移動向量對差異值和總數值的影響，圖4.2(a)為重建的背景影像，圖4.2(b)為新進影像，且重建的背景影像相對於新進影像的移動向量為 $(-78, 78)$ ，圖4.2(c)~圖4.2(f)中左邊的圖示表示重建的背景影像根據不同的候選移動向量移到相對應位置的結果，而右邊的圖示表示不同的候選移動向量造成移動後的重建的背景影像和新進影像的對齊結果，其中藍色的像素表示式子(4.4)中 $|R_b(x, y) - R_f(x, y)| + |G_b(x, y) - G_f(x, y)| + |B_b(x, y) - B_f(x, y)|$

$(x, y) \mid$ 超過30的像素，由圖4.2(e)可知最準確的移動向量有最少的藍色像素，表示重建的背景影像與新進影像之間的對齊越精準。

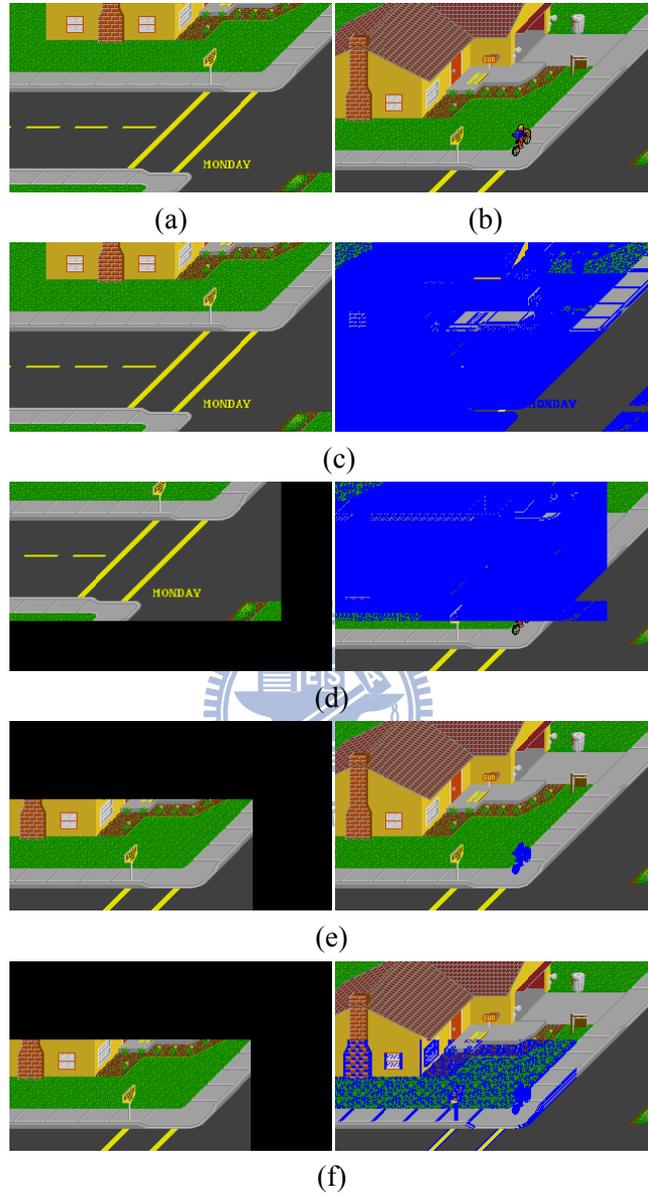


圖4.2 不準確的移動向量對差異值和總數的影響。(a) 重建的背景影像，(b) 新進影像，(c) 移動向量 $(0, 0)$ 造成的結果，(d) 移動向量 $(-50, -50)$ 造成的結果，(e) 移動向量 $(-78, 78)$ 造成的結果，(f) 移動向量 $(-80, 78)$ 造成的結果。

至於將差異值除上總數值，是為了避免大的候選移動向量可能會造成重建的背景影像與新進影像之間的重疊區域太小，因此式子 (4.4) 所需要計算的矩形面積也會很小，結果最後的差異值反而會小於小的候選移動向量所算出來的差

異，因為小的候選移動向量表示重建的背景影像與新進影像之間重疊的區域很大，因此式子(4.4)所需要計算的矩形面積就會很大，進而造成差異值變得很大。所以我們需要正規化差異值，將選擇最佳移動向量的依據定義為對重建的背景影像與新進影像，能夠提供每一個像素最小平均差異度的候選移動向量。

例如圖4.2(c)為移動向量(0, 0)造成的結果，其差異值為10333264，總數值為60480，平均差異度為170，圖4.2(d)為移動向量(-50, -50)造成的結果，其差異值為8057424，總數值為37530，平均差異度為214，圖4.2(e)為移動向量(-78, 78)造成的結果，其差異值為109456，總數值為26862，平均差異度為4，圖4.2(f)為移動向量(-80, 78)造成的結果，其差異值為1046480，總數值為26640，平均差異度為39，由此可知最準確的移動向量能夠提供最小的平均差異度。

4.3 影像對齊與前景背景分割



在有了最佳移動向量後，將背景GMM根據此最佳移動向量移動到新的位置，再處理新進影像中每個像素，進而判斷出每個新進像素為前景或背景。在處理完所有的像素後，得到一張二值化影像，其值只有0或1，0表示為背景，而1表示前景，我們再利用附錄1介紹的連通物件搜尋法將此二值化影像中所有相連的前景區域（即值為1的像素）找出，並且計算其面積大小，若面積過小，則表示該前景區域可能為雜訊，因此須將這塊區域視為背景。刪除完所有面積過小的前景區域後，最後留下來的面積夠大的前景物件。圖4.3顯示在有應用此步驟的情況下，能得到較不受雜訊影響的結果。

為了處理攝影機些微旋轉對影像造成的影響，我們將在上一個消除雜訊步驟後所剩下的前景像素，與在重建的背景影像中相對應位置一定範圍內的GMM做配對，若有一個GMM能夠成功的配對到該像素，便將該像素視為背景。這個步驟的代價是若前景物像素與相對應位置附近的GMM有顏色相似的情況，則該

前景物像素也會被誤判為背景，因此若配對的範圍越大，處理攝影機旋轉的效果也越好，但是也增加了前景物像素被誤判成背景像素的機率。在本論文中，將本步驟配對的範圍設定成上下左右各三個像素，圖4.4顯示在有應用擴大配對的情況下，在邊界上能得到比較好的前景背景分割結果。

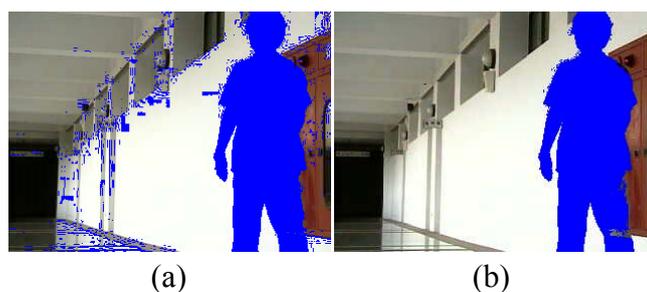
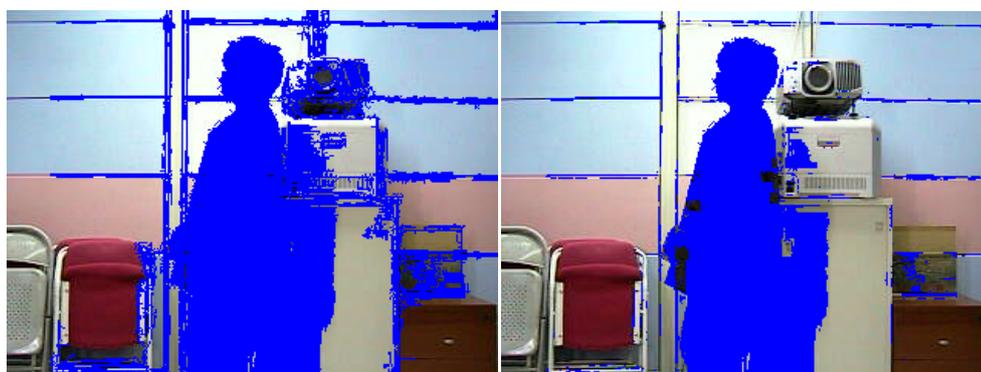


圖4.3 有無濾除雜訊的結果。其中純藍色區域為切割出的前景。
(a) 沒有濾除雜訊的結果，(b) 有濾除雜訊的結果。



(a)



(b)

(c)

圖4.4 有無消除攝影機旋轉影響的結果。(a) 原始影像，(b) 沒有擴大配對的結果，邊界上前景背景分割的結果並不好，(c) 有擴大配對的結果，邊界的地方改善很多，但有部分前景物像素被判斷成背景。

4.4 新進與移出的背景像素之處理

在完成一個影格的前景背景分割後，我們需要更新背景GMM，使其能夠在下一個影格繼續提供最新的重建的背景影像來建立以顏色為基礎的群組、用邊緣所切割出的區塊以及用角點偵測器所獲得之角點，以達到對齊下一張新進影像以及GMM的目的。由於GMM本身已經有了更新的機制，因此對於原來就存在於場景中的GMM，我們不必特別設計一個新的更新機制，但如果影像有移動過，則勢必有新加入以及移出的背景像素以及其對應的GMM，對於移出的背景像素，刪除其對應的高斯混合模型，而對於新加入的背景像素，在該背景像素的位置新增一個新的高斯模型，其RGB色頻為該像素的RGB值，但是權重和變異量則需要手動指定一個初始值。另外對於經由消除雜訊與攝影機旋轉對影像造成的影響而被判斷成背景的像素，我們需要將它們所對應到的高斯模型強制設定為該像素的背景模型，這對於新進影像與重建的背景影像的同步有相當大的幫助，圖4.5顯示在有強制設定的幫助下，重建的背景影像與新進影像的同步有顯著的改進。





(a)



(b)



(c)



(d)

圖4.5 有無強制設定背景，對新進影像與重建的背景影像的同步的影響。(a) 上一個影格的重建的背景影像，(b) 新進影像，(c) 沒有強制設定的結果，重建的背景影像出現些許的斷層，(d) 有強制設定的結果，新進影像與重建的背景影像的同步改善很多，幾乎沒有斷層。

五、實驗結果與討論

在本章中，我們提供了一個非真實的模擬影像、一個靜態的實際測試影像以及兩個動態式攝影機所拍攝的測試影像的前景背景分割結果，並且分析以顏色為基礎的群組、用邊緣所切割出的區塊以及用角點偵測器所獲得之角點在尋找最佳移動向量時的準確率以及平均在每個影格提供候選移動向量的數目。

圖5.1為一個人造圖片下的實驗結果，圖5.1(a)表示有一個腳踏車騎士進入，圖5.1(c)為整個攝影機向右上角移動了一段距離所的結果，圖5.1(b)與圖5.1(d)為圖5.1(a)與圖5.1(c)的前景背景分割結果，其中純白色的像素為切割出的前景物像素。由此可見，我們的方法不受限於攝影機的移動大小，都能準確的計算出移動向量，並且成功地切割出前景物件。

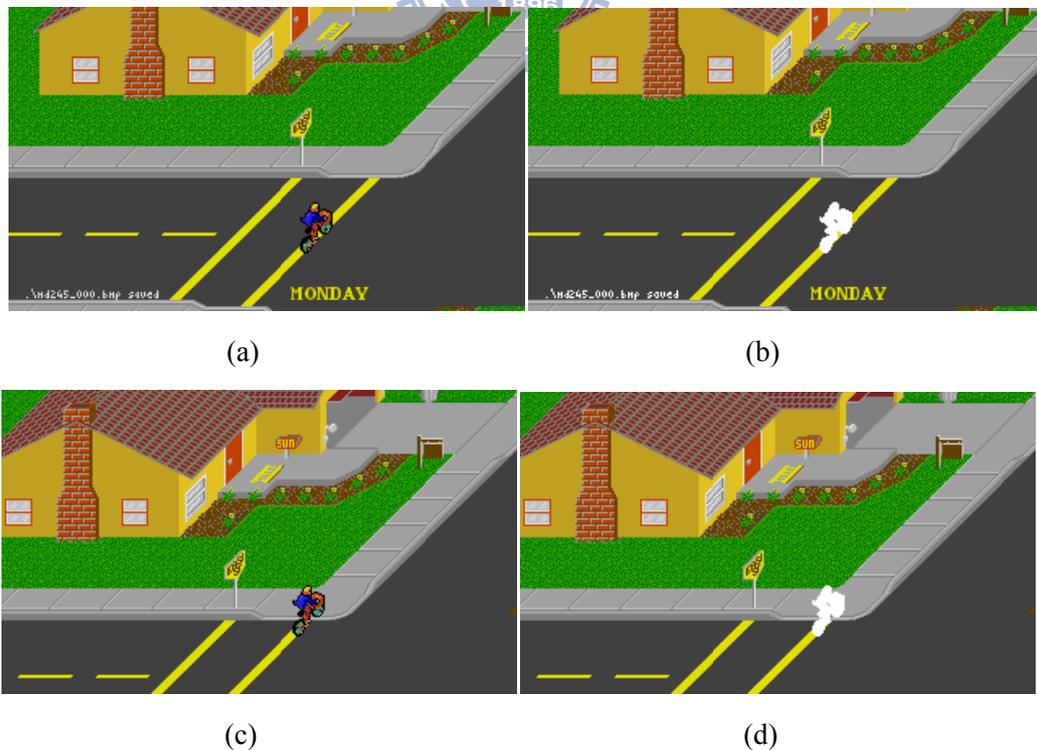


圖5.1 模擬影像的測試結果。(a) 腳踏車騎士進入場景，(b) (a)的分割結果，(c) 攝影機往右上角移動，(d) (c)的分割結果。

另外，我們的方法就算在一個影格出現了問題，例如不準確的移動向量，甚至是完全計算錯誤的移動向量，仍然有機會在下一個影格補救回來。主要的原因是即使在一個影格算錯了移動向量，導致整個背景GMM被移動到錯誤的位置，但若下一個影格進來後，GMM所提供的重建的背景影像中僅僅是因為被移動到錯誤的位置，但結構都還很完整的區域在該影格中還持續存在，則我們的方法會因為判斷該影格中與重建的背景影像中有非常相似的區塊，而重新找到正確的移動向量。圖5.2表示這個過程，其中純藍色的像素表示切割出的前景物。



圖5.2 出錯後又回復正常的過程。(a) 出錯的影像，(b) 回復回來的影像，本圖僅為GMM前景背景分割後的結果，並未經過消除雜訊與攝影機旋轉的處理。

圖5.3為靜態式攝影機下的實驗結果，在這個例子中，我們同時顯示了群組的結果以及切割出的前景物像素。圖5.3(a)表示重建的背景影像，圖5.3(c)表示前景物出現的影像，圖5.3(b)表示程式輸出的重建的背景影像的群組分佈，圖5.3(d)表示前景影像的切割結果。從這個實驗結果可知，我們的方法不但能夠應用在動態式攝影機下，同時也能夠應用在靜態式攝影機下。

圖5.4為動態式攝影機下純粹前景背景分割的實驗結果，我們並未套用消除雜訊以及攝影機旋轉對影像造成的影響的步驟，在這個例子中藍色的像素表示切割出來的前景物像素。圖5.4(a)為重建的背景影像，圖5.4(b)表示前景物出現的影像，圖5.4(c)為攝影機往右移動一段距離。圖5.4(d)表示圖5.4(b)的切割結果，圖

5.4(e)表示圖5.4(c)的前景背景切割結果。圖5.5與圖5.6為第二個和第三個動態式攝影機下的實驗結果，我們套用了消除雜訊以及攝影機旋轉對影像造成的影響這兩個步驟，其中藍色的像素表示切割出的前景物像素，我們可以看到結果跟圖5.4相較下有顯著的改進。

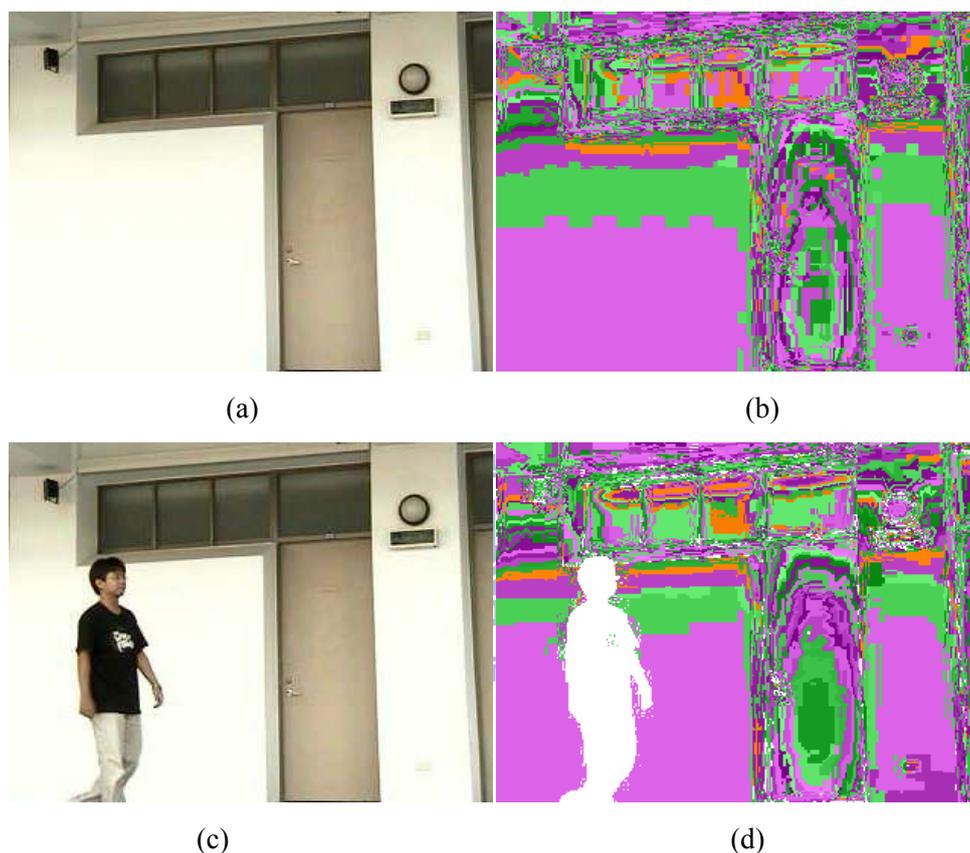


圖5.3 靜態式攝影機下的實驗結果。(a) 重建的背景影像，(b) (a)的群組分佈，(c) 前景物出現，(d) (c)的前景背景分割結果。

我們將群組、區塊及角點在五個測試影像中找出正確移動向量的機率以及提供不同的候選移動向量的數目整理於表2，其中找出正確移動向量的機率越高越好，而提供不同的候選移動向量的數目越少越好，因為過多不同的候選移動向量意味著需要花費更多的時間去找出最佳移動向量。注意在表2中的平均候選移動向量的數目並未經過利用一個候選移動向量衍生出25個候選移動向量的步驟，因此表2的平均候選移動向量的數目不會是25的倍數，另外本論文中所用的測試影像之影格間格都在一秒左右。例如圖5.5的測試影像原本共有10個影格，群組

能夠找到正確移動向量的機率為100%，平均一個影格提供51個可能的候選移動向量；區塊能夠找到正確移動向量的機率為90%，平均一個影格提供30個可能的候選移動向量；角點能夠找到正確移動向量的機率為100%，平均一個影格提供34個可能的候選移動向量。圖5.6的測試影像原本共有21個影格，群組能夠找到正確移動向量的機率為76%，平均一個影格提供25個可能的候選移動向量；區塊能夠找到正確移動向量的機率為90%，平均一個影格提供20個可能的候選移動向量；角點能夠找到正確移動向量的機率為81%，平均一個影格提供10個可能的候選移動向量。



圖5.4 動態式攝影機下的實驗結果。(a) 重建的背景影像，(b) 前景物出現，(c) (b)的切割結果，(d) 攝影機往右移動，(e) (d)的切割結果。

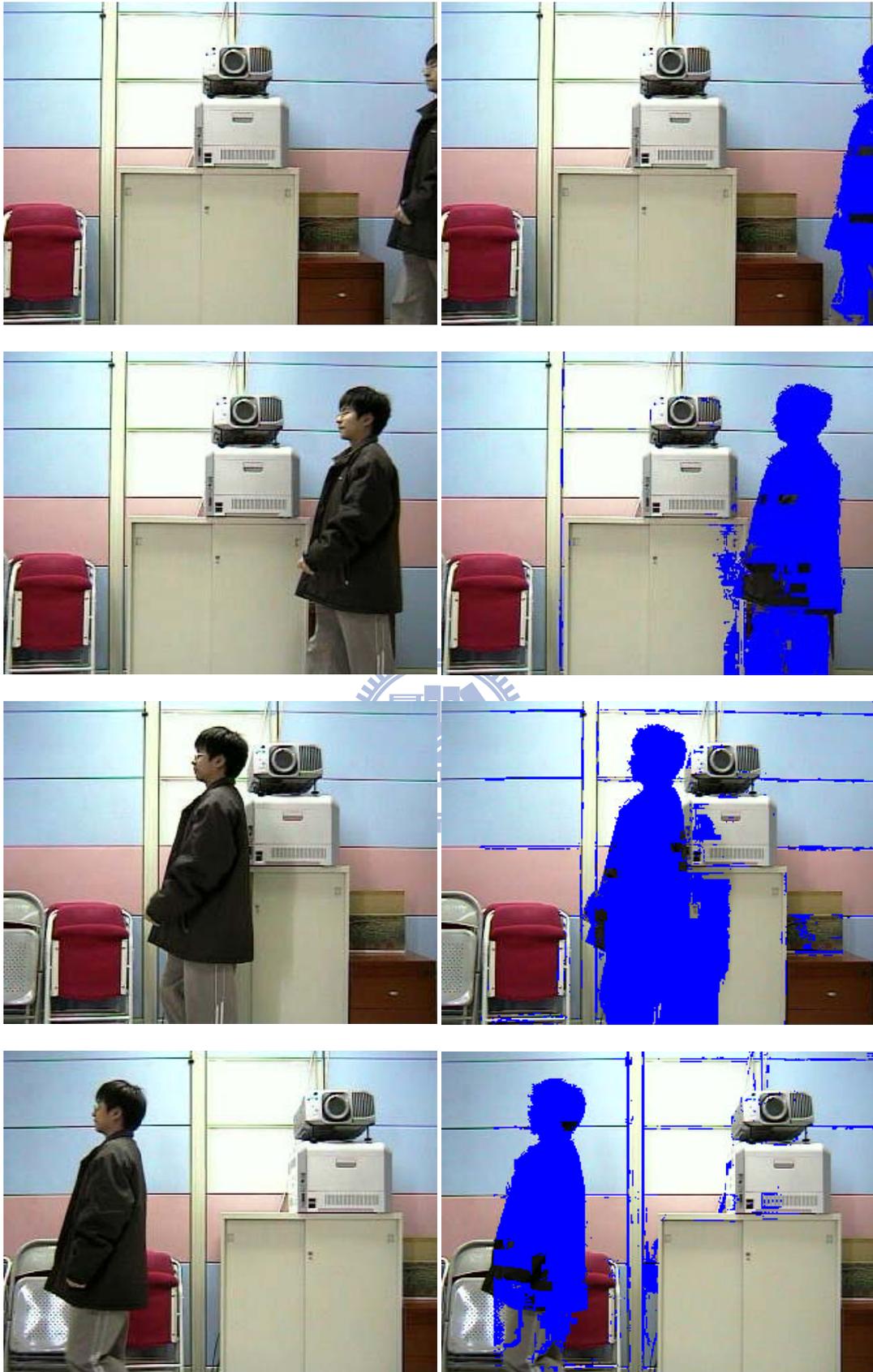


圖5.5 動態式攝影機下的第二個實驗結果。

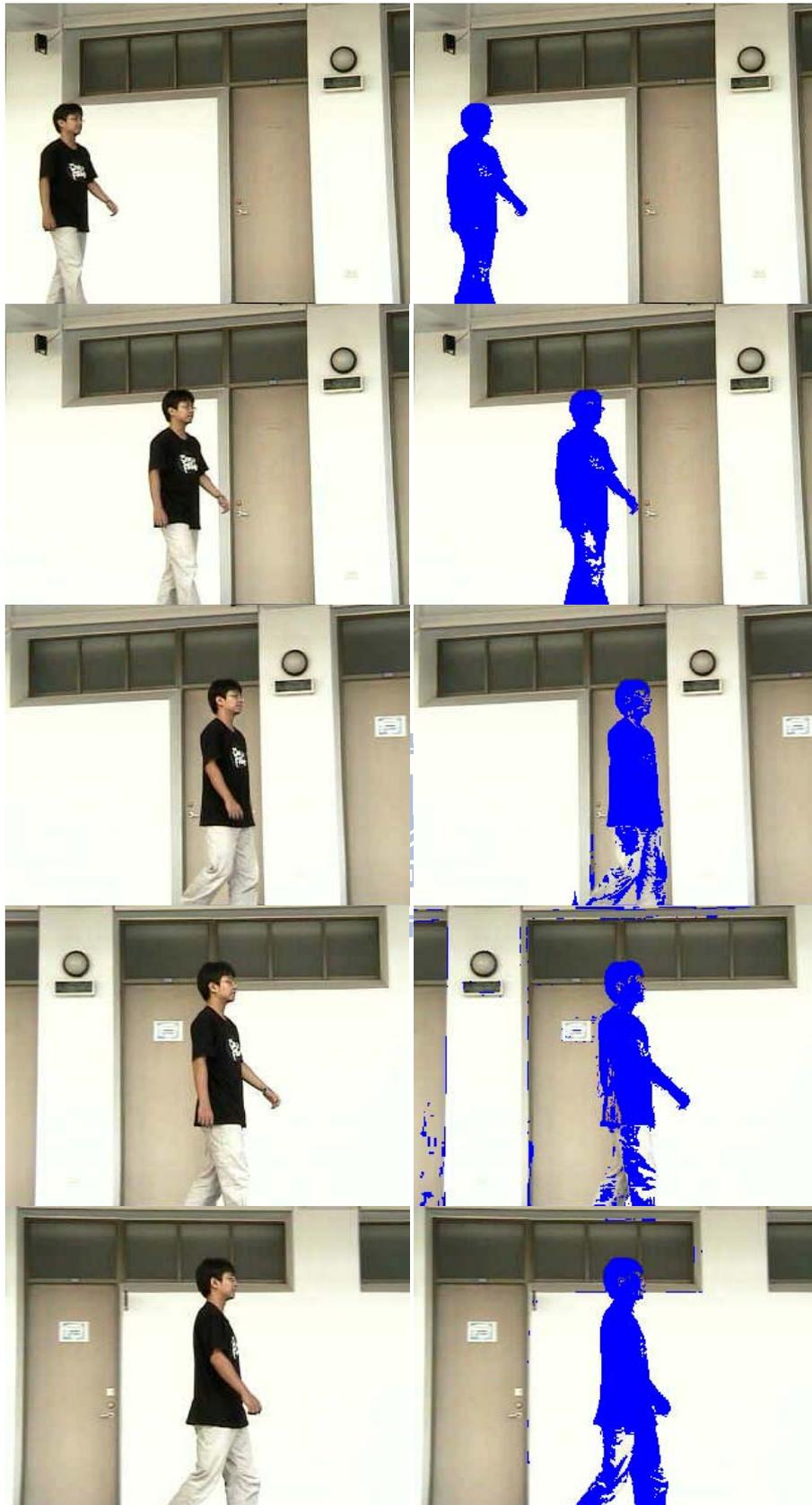


圖5.6 動態式攝影機下的第三個實驗結果。

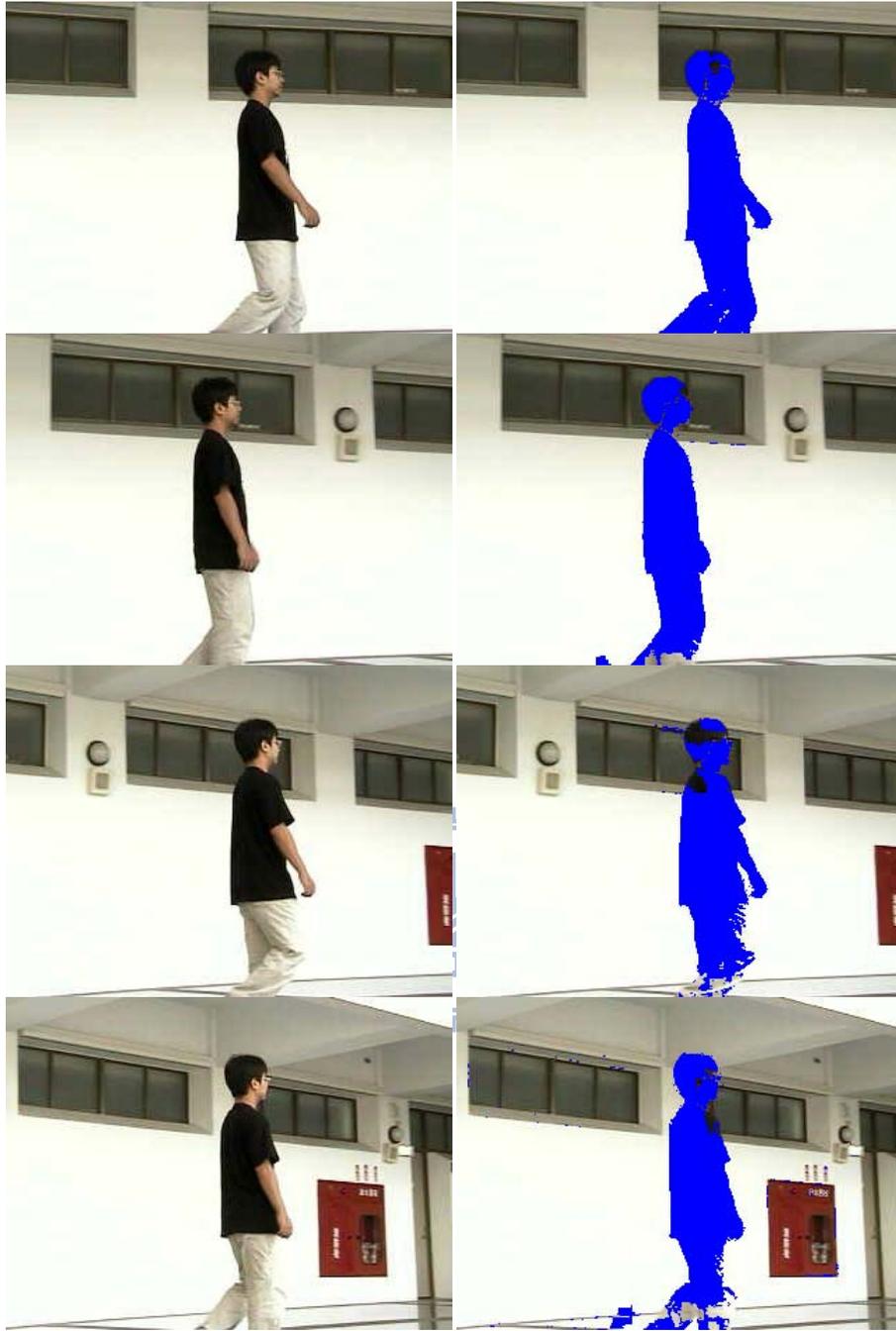


圖5.6 (續)

表3分別針對5個測試影片，顯示了個別使用群組、區塊或角點來尋找最佳移動向量時，最長可以連續執行前背景分割的工作多少個影格，若可以完整地對一個測試影像執行前背景分割的工作，則以”完整”來表示，否則以一個數字來表示從測試影像的開始，該方法可以連續且成功地執行前背景分割的工作多少個影格而不發生錯誤。

觀察表2與表3，我們可以發現角點在圖4.4的環境下表現得特別差，這是因為圖4.4跟其他的測試影片比起來相對模糊，所以偵測出的角點數目少而且不穩定，平均在一張新進影像中只能找到一個角點。儘管我們利用Harris corner detector可以在圖3.7(即圖5.5與圖5.6中的兩個影格)中找到令人滿意的角點數量，同時也兼顧了角點需要在角落上的需求，但是在圖4.4的環境下就無法得到相同的表現。減少Harris corner detector的 R 值雖然有助於改善在圖4.4中角點偵測不良的問題，但卻會造成圖3.7中找到的角點數目大增，如圖3.6所示。至於造成新進影像變模糊的原因，除了攝影機本身的品質外，也來自於攝影機的快速移動，因為本論文的目的是在動態式攝影機下分割前景背景，所以攝影機的移動是免不了的，但若是攝影機移動過於強烈造成影像過度模糊，則角點就可能無法提供精準的移動向量。

從表2與表3的統計中我們可以知道無論是群組、區域或角點，沒有一個方法可以完美地找出影像間的移動向量，因此我們必須同時利用配對群組、區域及角點這三種尋找移動向量的方法，讓其中一個或是兩個方法在某個影格無法正確地找到移動向量時，可以靠著第三種方法來找到正確的移動向量。或者是提出其他的方法，來動態地判斷目前的影像是否適用群組、區塊或角點來估算出移動向量，例如3.2節以及上一個段落中曾經提到，區塊和角點並不適用於模糊的影像，因此如果能夠找出一個方法來衡量影像是否夠清晰，讓影像在夠清晰的時候能利用區塊和角點來尋找最佳移動向量，但在影像變模糊的時候則自動停用區塊和角點，便能夠為整體的方法加速，因為我們不需要在影像變模糊的時候建立區塊和角點，也不用費時地測試從不穩定的區塊和角點所找出來的候選移動向量。另一個可能的加速方法是在影像光影變化太強烈時停用群組，因為群組是以顏色為基礎建立的，因此對於光影的變化相當敏感，如果我們能夠提出一個方法來衡量影像間的光影變化是否過於強烈，並在影像間的光影變化過於強烈時停用群組，也有助於加速本論文所提出的前景背景分割方法。

表2 群組、區塊以及角點的效率比較。

	群組	區塊	角點
在圖4.2的測試影像中找出正確移動向量的機率	59%	100%	100%
在圖4.2的測試影像中平均每個影格提供的候選 移動向量數目	19	34	17
在圖4.4的測試影像中找出正確移動向量的機率	81%	56%	33%
在圖4.4的測試影像中平均每個影格提供的候選 移動向量數目	33	7	1
在圖5.1的測試影像中找出正確移動向量的機率	100%	100%	100%
在圖5.1的測試影像中平均每個影格提供的候選 移動向量數目	23	46	112
在圖5.5的測試影像中找出正確移動向量的機率	100%	90%	100%
在圖5.5的測試影像中平均每個影格提供的候選 移動向量數目	51	30	34
在圖5.6的測試影像中找出正確移動向量的機率	76%	90%	81%
在圖5.6的測試影像中平均每個影格提供的候選 移動向量數目	25	20	10

表3 群組、區塊以及角點的穩定性比較。

	總長度(影格數)	群組	區塊	角點
圖4.2	27	4	完整	完整
圖4.4	27	完整	18	9
圖5.1	92	完整	完整	完整
圖5.5	10	完整	完整	完整
圖5.6	21	14	12	13

六、 結論與未來展望

由於前景背景分割在安全監控以及其他應用的計算效率上一直扮演著重要的角色，因此本論文設計了一個能在動態攝影機條件下偵測移動物體的方法，讓視訊偵測的範圍跳脫出靜態攝影機的限制。本方法會持續維護一張只有單純背景環境的背景影像，之後當有新進影像要做前景背景分割時，同時對背景影像及新進影像建立以顏色為基礎的群組、用邊線所切割出的區塊以及用角點偵測器所獲得之角點，接著藉由配對上述三種背景影像與新進影像的對應特徵，得到兩張影像間所有可能的候選移動向量，之後再選出能提供新進影像與背景影像之間最小差異度的候選移動向量，做為最佳的移動向量。接著根據此最佳移動向量對齊背景高斯混合模型與新進影像，然後才進行前景背景的分割，之後在進一步地消除雜訊以及攝影機旋轉的影響，以得到最後的前背景分割結果。

在本章中我們將探討本論文所提出的動態攝影機下前景背景分割法的缺點以及還可以發展的方向，一共有四個，分別為：（1）群組及區塊配對的方法應該要將群組或區塊的細部形狀納入考量（2）混雜在新進背景中的前景物無法被正確地切割出來（3）攝影機的運動包括了移動和旋轉，這會造成比較複雜的背景變化（4）景深造成新進影像與背景GMM間的相對運動無法只以一個移動向量表示，我們將在下面的段落中依續討論這四個問題。

因為一個成功的配對代表一個可能的候選移動向量被發現，但是測試候選移動向量非常浪費時間。因此若能找到更好的群組及區塊配對方法，降低錯誤配對的機率，便可以有效地降低候選移動向量的數目，以加快整體方法的速度。目前我們只用到顏色、面積差異以及 x 跟 y 軸的變異量來判斷兩個群組或區塊可否配對成功，但是用形狀來配對會是一個比較理想且準確的方法。

因為本論文提出的方法主要是希望能夠應用在使用動態式攝影機的環境，

所以我們還需要想辦法克服混雜在新進背景中的前景物的問題，因為現在對於新進的區域要判斷是否有前景物的存在有困難。根據目前的進度，是通通歸類成背景，唯有在前景物移動到新的位置且與上一張影像中原來的的位置沒有重疊的前提下，才能夠完整地找出該前景物，但是即便該前景物被正確地找出，因為該位置的背景GMM是儲存著前景物的顏色，而造成背景物像素被判斷成前景物像素，這顯然不合理。雖然GMM允許該露出的背景物像素在出現一段時間後被正確的分類成背景像素，但是出現多久的像素應該要被分類成背景，才不會將前景像素也誤判為背景像素，也是一個大問題。對此，我們想到可能的解法是試著把這套方法做小幅度的修改，可能只有更改各個閾值的數值而已，再套用到其他性質的攝影機下，例如中紅外線攝影機，然後試著在這兩套異質平台影像上面找一個能夠互補的方法，讓結果能夠更合理。但是我們目前還沒有達到這個目標，因此我們還需要繼續設計一套方法，讓我們的方法能夠在動態攝影機下面運作。

另外我們的方法假設所有的影像之間都是平移的關係，沒有旋轉的存在，但是這對於一般情況都是不成立的。儘管我們在4.3節中提出了比對付近背景GMM的方法來降低攝影機旋轉的影響，但卻要冒著前景物像素被誤判為背景物像素的風險。對此，我們想到可能的解法是，將新進影像與重建的背景影像之間的關係以真實的平移旋轉模型來模擬，接著估算出可能的平移向量和旋轉角度，再找出最佳的平移向量與旋轉角度，最後將重建的背景影像像素根據找出的旋轉平移模型移動到正確的位置，再作前景背景的分割。

最後一個問題是景深相關的議題，因為本論文假設所有背景像素的移動都是同一個移動向量，但這個假設只有在影像中的景深差異不大的時候才成立，例如室內的場景，對於室外的場景就不能成立。因為對於像素間景深差異很大的場景，即使靠近攝影機的物件因為攝影機的移動而在影像中移動了一個很大的距離，影像中離攝影機很遠的物件卻可能只有移動一點而已。因此對於室外的場景，影像中像素的移動可能就必須以多個移動向量來代表，而非只有一個移動向量而已。

參考文獻

- [1] C.-C. Chen, "A Novel Method of Foreground Object Detection in Infrared Images," Master's Thesis, Institutes of Multimedia Engineering, National Chiao Tung University, June 2008.
- [2] S.-W. Gao, L. Guo, L. Chen, and Y. Yu, "Adaptive Visual Tracking Using Particle Filter," *International Conference on Information Technology: New Generations*, pp. 1117-1122, April 2008.
- [3] C. Stauffer and W. E. L. Grimson, "Adaptive Background Mixture Models for Real-time Tracking," *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 246-252, 1999.
- [4] M. Isard and A. Blake, "Condensation-Conditional Density Propagation for Visual Tracking," *International Journal of Computer Vision*, vol. 29, pp. 5-28, November 1998.
- [5] T. Hashiyama, D. Mochizuki, Y. Yano, and S. Okuma, "Active Frame Subtraction for Pedestrian Detection from Images of Moving Camera," *IEEE International Conference on Systems, Man and Cybernetics*, vol. 1, pp. 480-485, 2003.
- [6] D. Murray and A. Basu, "Motion Tracking with an Active Camera," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, pp. 449-459, 1994.
- [7] B. Heisele and C. Wohler, "Motion-Based Recognition of Pedestrians," *International Conference on Pattern Recognition*, vol. 2, pp. 1325-1330, 1998.
- [8] S.-C. Tai, Y.-R. Chen, Z.-B. Huang, and C.-C. Wang, "A Multi-Pass True Motion Estimation Scheme with Motion Vector Propagation for Frame Rate Up-Conversion Applications," *Journal of Display Technology*, vol. 4, pp. 188-197, June 2008.
- [9] Z. Zhang, R. Deriche, O. Faugeras, and Q.T. Luong, "A Robust Technique for Matching Two Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry," Research Report no. 2273, INRIA Sophia-Antipolis.
- [10] R. C. Gonzales and R. E. Woods, *Digital Image Processing 2nd ed.*, Prentice Hall New Jersey, 2002.

- [11] C. Harris and M. Stephens, "A Combined Corner and Edge Detector,"
Fourth Alvey Vision Conference, Manchester, pp. 147-151, 1988.



附錄1. 連通物件搜尋法 (Connected Component Search)

在傳統的連通物件搜尋法中，先將影像中所有的像素標記為未處理過，表示該像素尚未被歸屬到任何一個群組，並且維護一個後進先出的堆疊 (stack)，以及代表一個最新被加入堆疊的像素 P 。一開始這個堆疊會是空的，我們在影像中尋找一個未處理過的像素，將該像素加入這個堆疊且將其設定成 P ，並將 P 記錄為已處理過。接著判斷 P 周圍的像素是否與 P 屬於同一個群組模型且尚未被處理過，加入每個符合條件的像素到堆疊中，並將它們記錄為已處理過。當 P 周圍的八個像素都判斷完畢後，我們將 P 從堆疊中移除。接著取出堆疊中最上層的元素 (即最後一個被加入堆疊的像素)，將其設定為 P ，並且回到判斷 P 周圍的像素的步驟。當堆疊中沒有元素後，就代表一個連通物件被找到，因此回到在影像中尋找一個未處理過的像素這個步驟，繼續尋找下一個連通物件。整個連通物件搜尋法的演算法如下：

- (a) 將影像中所有的像素標記為未處理過。
- (b) 設定一個像素 P 表示最新一個被加入堆疊的像素。
- (c) 在影像中尋找一個未處理過的像素，將該像素加入這個堆疊並且將其設定為 P ，同時將該像素記錄為已處理過。如果找不到一個未處理過的像素，則整個連通物件搜尋法結束。
- (d) 試著將跟 P 相鄰且屬於同一個群組模型且尚未被處理過的八個像素加入堆疊，並將它們記錄為已處理過。
- (e) 將 P 從堆疊中移除。
- (f) 如果堆疊變成空的，回到(c)。否則將堆疊中最上層元素設定成 P ，並回到(d)。

以上介紹的方法已經能夠用來尋找3.1節所需的群組，而同樣的方法只需要做一點小修改，亦可用來搜尋在3.2節中用邊線所切割出來的區塊。在上一個段落中我們提到，當我們判斷堆疊中最上層元素(意即最後一個被加入堆疊的像素)周圍的像素能否被加入堆疊時，純粹取決於周圍像素與堆疊中最上層元素是否屬於同一個群組模型且尚未被處理過，如果是，則加入周圍像素到堆疊；否則不對該像素做任何動作。現在對於邊線所切割出來的區塊，如果我們想判斷一個像素是否能夠被加入堆疊，我們僅僅只需判斷該像素是否不為邊線像素且尚未被處理過，若該像素不為邊線像素且尚未被處理過，則加入堆疊，否則不做任何動作。因為如果一個區塊被邊線所包圍，則所有不屬於該區塊的像素一定不會被加入堆疊，因為邊線像素不會被加入到堆疊中，所以所有在該邊線外面的像素都不會有機會被加入到堆疊中，這是由於搜尋的過程是在一個像素的周圍尋找，因此如果邊線像素沒有被加入堆疊，則所有在邊線像素外面且與邊線像素相鄰的像素自然也就沒有機會被加入到堆疊中。如此便可確保所有被邊線像素包圍的區塊都能被正確的找出，而不會超出邊線。

