

行政院國家科學委員會補助專題研究計畫 成果報告
 期中進度報告

應用於視訊監控之運動物體追蹤技術研究

計畫類別： 個別型計畫 整合型計畫

計畫編號：NSC 95-2221-E-009-106-MY2

執行期間： 95 年 8 月 1 日至 96 年 7 月 31 日

計畫主持人：王聖智

共同主持人：

計畫參與人員： 陳宜賢、王郁晴、蕭晴駿

成果報告類型(依經費核定清單規定繳交)： 精簡報告 完整報告

本成果報告包括以下應繳交之附件：

- 赴國外出差或研習心得報告一份
- 赴大陸地區出差或研習心得報告一份
- 出席國際學術會議心得報告及發表之論文各一份
- 國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、
列管計畫及下列情形者外，得立即公開查詢

涉及專利或其他智慧財產權， 一年 二年後可公開查詢

執行單位：交通大學電子工程系

中 華 民 國 96 年 5 月 31 日

應用於視訊監控之運動物體追蹤技術研究

計畫編號：NSC 95-2221-E-009-106-MY2

執行期限：95年8月1日至96年7月31日

主持人：王聖智 (交通大學電子工程系教授)

計畫參與人員：陳宜賢 王郁晴 蕭晴駿 (交通大學電子所研究生)

一. 中文摘要

我們提出能夠自動偵測畫面中移動物體並持續追蹤的演算法，並嘗試解決在物體追蹤問題中常遭遇的遮蔽、場景變化以及光源變化等等問題。我們定義一個多團塊模型用以描述移動物體，並基於多團塊模型定義適合的相似度測量測，進而發展出以平均移動法為基礎的追蹤系統。另外，我們也運用等位函式法來做運動物體的偵測與追蹤，此方法亦適用於背景隨時間改變的狀況。實驗結果顯示在室內、室外等不同場景都能夠正確地追蹤移動物體的運動行為。

關鍵詞：區塊追蹤、平均移動法、等位函數法。

In this project, we proposed an object tracking system, which can automatically detect a single moving object in an image sequence and keep tracking of this object. In the proposed system, we deal with the problems of occlusion, scene change and luminance change. A multi-blob model is defined in our approach to represent the moving object. With this multi-blob model, we proposed a new similarity measure and developed a new object tracking algorithm based on the mean-shift method. Additionally, a level set function is used to detect and track the moving objects. A moving object can be well tracked by this method even when the background is under changing. Simulation results demonstrate that the proposed object tracking algorithm can faithfully track the

moving object in different scenes.

Keywords: Region tracking, Mean-shift method, level set function.

二. 進度報告

(1) 背景

簡單來說，物體追蹤是找出連續影像的同一目標物，而在本計畫所使用的平均移動法，是目前其中一個常被使用的追蹤方法，其具有簡單且強韌的特性，追蹤過程中，不需要知道整張影像的色彩分佈特性，即可找到附近的目標物。論文[1]中，Comaniciu和Meer依據移動目標物的模型和影像相似度的比較，應用平均移動法解決物體追蹤的問題，但他們的方法仍存在一些問題，例如：影像中位置的資訊並未適當地保留，以致於可能造成錯誤的追蹤結果。再者，在相似度的運算方面，其所用的 Bhattacharyya coefficient 和 Kullback-Leibler distance，有時候所呈現的目標分辨力並不是那麼理想。而[2]所提的spatiogram方法，同時包含了色彩和位置資訊，但還是可能會失去影像的細節資訊。我們將針對這些問題改善技術。

本計畫另一個偵測和追蹤方法採用等位函數法，主要是由大地測量主動式輪廓模型 (geodesic active contours model) 發展而來 [3]。使用等位面來描述輪廓，而物體邊緣輪廓和此等位面的關係可用等位理論加以建立。等位理論在處理輪廓和記錄多個目標物上是個很好的工具。另外，[4]和[5]所提出的主動式團塊模型則將統計特性列入考慮。[6]的

“active contours without edges”模型在利用統計概念，將影像區分為兩部分方面，其結果更是強韌，可用於背景相減結果中，區分移動區域與靜止區域。對於移動的攝影機，而[7]所提的區域追蹤模型可以在連續兩張影像間找出最相像的區域。而[8]中，Yonggang Shi *et al* 則應用了快速的等位方法到追蹤問題上。我們將嘗試將等位理論應用於區域追蹤的問題，亦即先實現[7]的方法，分析其結果的表現。

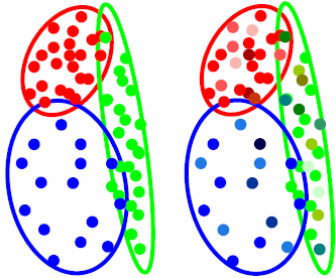
(2) 技術說明

2.1 多團塊模型

在論文[2]中，為了同時保留色彩和位置的資訊，其目標物模型定義如下：

$$h_p(b) = \langle n_b, \mu_b, \Sigma_b \rangle, b = 1, \dots, M. \quad (1)$$

n_b 代表第b個色彩量化區域的像素點數目， μ_b 和 Σ_b 則代表平均向量和協方差矩陣。但是，如此一來，色彩區塊的細節可能被移除，為了呈現[2]的問題，我們產生像圖一這樣具有相同



圖一. 相同spatiogram的樣本。

spatiogram的相異樣本，即使這兩個樣本大致相似，但每個色彩區域的變動，對增加追蹤的可靠性仍有相當大的幫助。為了保留這些特徵資訊，我們在每個色彩區域上收集其統計特性，因此，我們定義了一個新的多團塊模型如下：

$$\text{Model}(d) = \langle n_d, \bar{\mu}_{sd}, V_{sd}, \bar{\mu}_{cd}, V_{cd}, T_d \rangle, \quad (2)$$

$$d = 1, \dots, B.$$

公式(2)中， n_d 代表第d個色彩區域的像素點數目， $\bar{\mu}_{sd}$ 和 V_{sd} 代表影像位置上的平均向量和

協方差矩陣； $\bar{\mu}_{cd}$ 和 V_{cd} 則代表 RGB 色彩空間上的平均向量和協方差矩陣。另外， B 是色彩區域的數目，根據此模型，物體像素點可分成 B 個團塊。除此之外，這些團塊應該會擁有不同的可靠度，所以我們增加參數 T_d 來表示每個團塊的可靠度。此處， T_d 值先設為 1，之後我們將根據追蹤的情況來更新各團塊的可靠度。



圖二. 紅色橢圓目標物中，五個最多像素點的色彩團塊。每個團塊的圈選顏色，以其色彩平均值畫出。

舉例來說，在RGB色彩空間上，每個色彩我們分成四個區域來建立團塊模型，各團塊以所屬像素點的多寡來排列，圖二中，我們以團塊的色彩平均值的顏色畫出前五個擁有最多像素點的團塊，而紅色橢圓代表所圈選的目標物。根據這個團塊模型，大致可以估計出目標物的色彩和位置空間的分佈情況。

2.2 相似度量測

我們以橢圓形圈選移動目標，令 \bar{y} 為目標物的橢圓中心， $\bar{\mu}$ 為模型的橢圓中心，則將影像區域和此多團塊模型的相似度 (similarity) J 定義為

$$J(\bar{y}) = \frac{1}{N} \sum_{i \in I} \sum_{d=1}^B w((\bar{y}_i - \bar{y}) - (\bar{\mu}_{sd} - \bar{\mu}), V_{sd}) \times w(\bar{c}_i - \bar{\mu}_{cd}, V_{cd}) T_d \delta(b(\bar{y}_i) - d). \quad (3)$$

其中 $w(\bar{x}, V) = \exp\left(-\frac{\bar{x}V^{-1}\bar{x}^T}{2}\right)$ 是一個多變異高斯核心 (multi-variant Gaussian kernel)， $I = \{\bar{y}_i, \bar{c}_i\}$ 是橢圓目標物的樣本，而 N 則是樣本數目。每個像素點在位置和色彩空間上都被高斯核心加權，同時再乘上可靠性係數 T_d 。

公式(3)的相似度定義可視為重要性的平均(the mean of the weightings)。圖三描繪出我們所提的相似度量測和 Bhattacharyya coefficient的比較，圖三(a)的左圖是一段影片的第15張影像鏡頭，紅色橢圓為手動圈選的目標物模型，而圖三(a)的右圖則是此影片的第250張影像，綠色和藍色橢圓分別是以 \bar{y}_1 和 \bar{y} 為中心的橢圓目標物，我們特別算出 \bar{y}_1 和 \bar{y} 兩個相異位置附近的相似度，將 Bhattacharyya coefficient 的值畫在圖三(b)，而我們所定義的相似度值則畫在圖三(c)。這裡可以看出，雖然 \bar{y} 是較合理的追蹤結果，但 Bhattacharyya coefficient 在 \bar{y}_1 的相似度表現卻高過 \bar{y} 。相反地，我們的相似度量測對於圖三(a)左右兩張圖，反應了較合於實際情況的相似度大小。再者，圖三(c)中，局部性最大值附近斜率較為陡峭，也顯示了我們所提的相似度量測對於相配區域的偵測擁有較佳的分辨能力。

接下來將推導如何利用平均移動法來達成我們的追蹤模型理論。令 $\nabla J(\bar{y})=0$ ，即

$$\sum_{i \in I} \sum_{d=1}^B [(\bar{\mu}_{sd} - \bar{\mu}) - (\bar{y}_i - \bar{y})] \times g_{sd}(\Delta\bar{y} - \Delta\bar{\mu}) w_{cd}(\Delta\bar{c}) T_d \delta_{id} = 0 \quad (4)$$

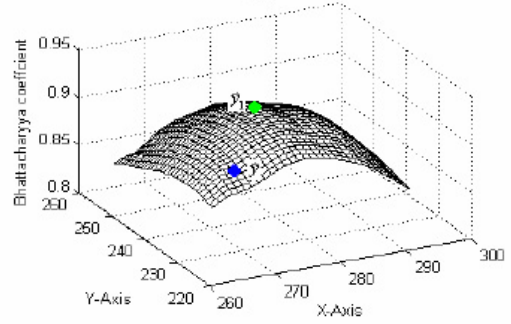
其中 $g_{sd}(\bar{x}) = -w_{sd}(\bar{x}) = -w(\bar{x}, V_{sd})$ 是 shadow kernel。在我們的模型中，shadow kernel 亦是一個多變異高斯核心，且 $\Delta\bar{y} = (\bar{y}_i - \bar{y})$ ， $\Delta\bar{\mu} = (\bar{\mu}_{sd} - \bar{\mu})$ ， $\Delta\bar{c} = (\bar{c}_i - \bar{\mu}_{cd})$ 。另外，如果 \bar{y}_i 的特徵值歸屬於模型中第d個特徵區域，則 δ_{id} 等於1，反之則為0。(4)經過再整理，我們可得

$$\bar{y}_1 = \frac{\sum_{i \in I} \sum_{d=1}^B [\bar{y}_i - \bar{\mu}_{sd}] \tilde{w}_{id} \delta_{id}}{\sum_{i \in I} \sum_{d=1}^B \tilde{w}_{id} \delta_{id}} + \bar{\mu} \quad (5)$$

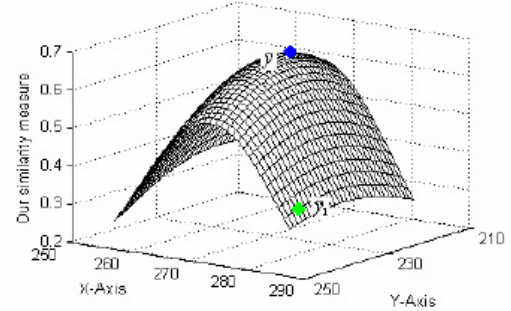
其中 $\tilde{w}_{id} = g_{sd}(\Delta\bar{y}_0 - \Delta\bar{\mu}) w_{cd}(\Delta\bar{c}) T_d$ ， $\Delta\bar{y}_0 = (\bar{y}_i - \bar{y}_0)$ ，而 \bar{y}_0 是初始目標物位置，最後 \bar{y}_1 為新的目標物中心位置。推導出的公式(5)即為我們平均移動法追蹤理論的核心。



(a)



(b)



(c)

圖三。(a)紅色橢圓圈選出目標物模型，藍色和綠色橢圓則代表分別以 \bar{y}_1 和 \bar{y} 為中心的橢圓目標物。(b)依據 Bhattacharyya coefficient 的相似度值。(c) 依據公式(9)的相似度值。

2.3 等位理論

主動式輪廓模型(active contour model)首先在1988年被Kass[9]所提出。先建立一個可以反應此輪廓曲線特性的能量函式，此輪廓曲線將隨著此能量最小化的過程傳遞演變，一般此能量函式可被定義成

$$E(\bar{C}) = \alpha \int_0^1 |\bar{C}'(q)|^2 dq + \int_0^1 g(|\nabla I(\bar{C}(q))|)^2 dq \quad (6)$$

輪廓曲線 \bar{C} 被 $q \in [0,1]$ 所參數化，而 I 是影像

資料， ∇I 則是 I 的梯度 (gradient)。 (6) 式中的第一項表內部能量，如果輪廓曲線越長，內部能量就會越大。第二項為影像能量，越靠近物體的邊緣，則影像能量要越小，因此影像能量採用遞減函式來表示。 α 則決定內部能量和影像能量之間的比例權重。

在 [10] 中，將主動式輪廓模型轉換成大地測量主動式輪廓模型 (geodesic active contour model)，此新的能量函式可表示成

$$L_R = \int_0^1 g(|\nabla I(\bar{C}(q))|) |\bar{C}'(q)| dq \quad (7)$$

這個新模型可視為被 $g(|\nabla I(\bar{C}(q))|)$ 加權的線積分。為了快速地將能量最小化，輪廓曲線的傳遞方向會是

$$\bar{C}_t = \left(g(\bar{C}) \kappa - \langle \nabla g(\bar{C}), \bar{N} \rangle \right) \bar{N} \quad (8)$$

(8) 是由 Euler-Lagrange approach 推導出，這裡的 κ 和 \bar{N} 分別表示輪廓曲線 \bar{C} 的曲率和法線向量，此輪廓曲線傳遞方程式 (8) 包含了一個量速度 (scalar velocity) 和一個法線向量 (normal vector)，此方程式可改寫如 (9)，成為較一般化的形式。

$$\bar{C}_t = F(\kappa) \bar{N} \quad (9)$$

論文 [10] 使用等位理論，進一步地將此傳遞方程式 (9) 加以實現。等位理論是以畫在圖四中的等位面來實行，在物體外側，等位面呈現正值；在物體內側，等位面則呈現負值，而輪廓曲線會恰好落在等位面等於零的位置，如果此等位面向上變動或向下變動，則輪廓曲線的傳遞可以在此等位面和零平面相交處呈現出來。輪廓曲線和等位面兩者的關係過去已經被完善地推導出來，如果等位面的更新如下，

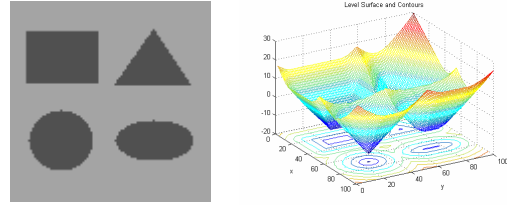
$$\frac{\partial \varphi}{\partial t} = F(\kappa) |\nabla \varphi| \quad (10)$$

則像 (9) 式的輪廓曲線傳遞便會在零平面出現。在速度函數中的曲率 κ 可以直接從等位面

算出，其公式為

$$\kappa = \frac{\varphi_{xx}\varphi_y^2 - 2\varphi_{xy}\varphi_x\varphi_y + \varphi_{yy}\varphi_x^2}{(\varphi_x^2 + \varphi_y^2)^{3/2}} = \text{div} \left(\frac{\nabla \varphi}{|\nabla \varphi|} \right) \quad (11)$$

式中 $\text{div}(\bar{A})$ 代表向量 \bar{A} 的發散性。



圖四. 物體及其相對應之等位面。

2.4 區域追蹤模型

區域追蹤模型是用來找一個特定區域，此區域和前一張影像的某預定區域非常相似。我們的追蹤模型採自於 Abdol-Reza Mansouri [7] 所提出的區域追蹤模型，此模型中，現在時間點影像上的每個像素點 $I^n(\bar{x})$ ，會從前一個時間點的影像中，在一個尋找區域範圍內，從預定區域的內側和外側，找到最相配的像素點。如圖五所示， R_0 是我們的預定區域，假設在 R_0 內側找到的最小誤差值為 V_{in} ，而在外側找到的最小誤差值為 V_{out} ，我們可將 V_{in} 和 V_{out} 公式化成

$$V_{in} = \inf_{\|\bar{z}\| \leq \delta, \bar{x} + \bar{z} \in R_0} \left(I^n(\bar{x}) - I^{n-1}(\bar{x} + \bar{z}) \right)^2 \quad (12)$$

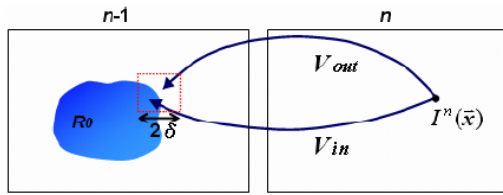
$$\text{和 } V_{out} = \inf_{\|\bar{z}\| \leq \delta, \bar{x} + \bar{z} \in R_0^c} \left(I^n(\bar{x}) - I^{n-1}(\bar{x} + \bar{z}) \right)^2 \quad (13)$$

在 Abdol-Reza Mansouri 的區域追蹤模型中，即將等位面的更新函式定義為

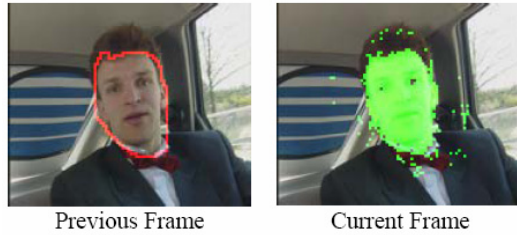
$$\frac{\partial \varphi}{\partial t} = [(V_{in} - V_{out}) + \lambda \kappa] |\nabla \varphi| \quad (14)$$

則在 R_0 內側找到的最小誤差值 V_{in} 會使等位面向上移動，因此像素點最後可能會歸屬於外側區域。從另一方面來看，在 R_0 外側找到的最小誤差值 V_{out} ，可能會使相對應的像素點變成內側區域的一部份，靠著公式 (14)， V_{in} 和 V_{out} 的效應互相競爭拉扯，直到此等位面平衡

穩定為止。我們以圖六為例，再加以說明。



圖五. 區域追蹤模型之圖示說明。



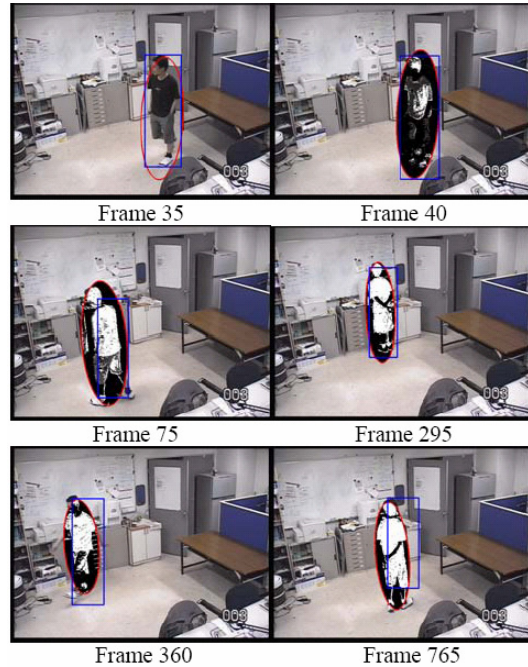
圖六. “Car-phone”影片中的兩張連續影像。左圖前一時間點的影像中，以紅色圈選出追蹤的預定區域，右圖是現在時間點的影像，綠色點是 $V_{in} - V_{out} < 0$ 的像素點。

圖六是“Car-phone”影片中的兩張連續影像，圖中的臉部區域在前一時間點的影像中先被定義圈選出來，而在現在時間點的影像中，綠色點代表 $V_{in} - V_{out} < 0$ 的像素點。在這個例子當中，等位面在這些綠色點的位置向下移動，而在其他位置向上移動，此等位面會反覆更新直到最後收斂為止，而這些綠色區域，可以適當地預測現在時間點的脸部區域位置。

(3) 實驗結果

首先我們呈現使用多團塊平均移動法產生的物體追蹤之結果。影像解析度是 640×480 ，實驗是透過MATLAB 6.5平台實現，且在配有512MB DRAM的3GHz Pentium4電腦上執行，而在目標物模型方面，我們將RGB三個色彩空間各分成8個量化區域，來建立多團塊模型，亦即我們的目標物模型含有512個團塊。另外，我們也實行傳統平均移動法的追蹤過程，目標物模型則採用 $16 \times 16 \times 16$ RGB色彩量化直方圖，測試影片則是用我們自己所拍攝的Hans 影片。圖七為移動物體追蹤結

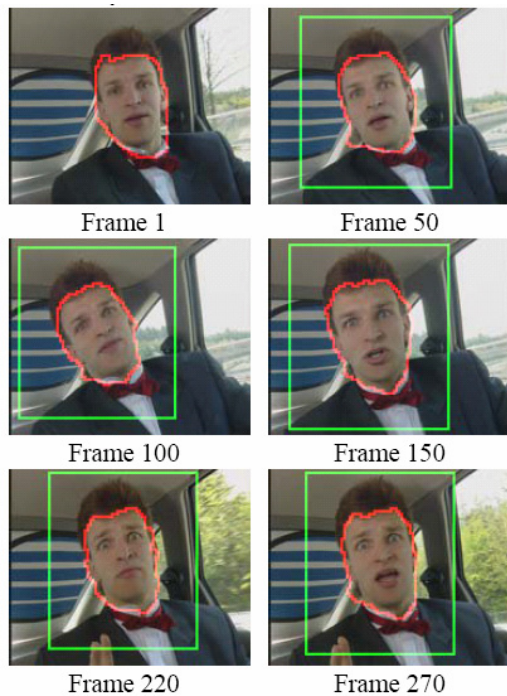
果，多團塊模型在影片第35張時被建立，其中紅色橢圓圈選出我們所提方法的追蹤結果；而藍色矩形則圈選出傳統平均移動法的追蹤結果，由圖七可看出，當追蹤的影像越多、時間越久，我們的多團塊模型較準確。



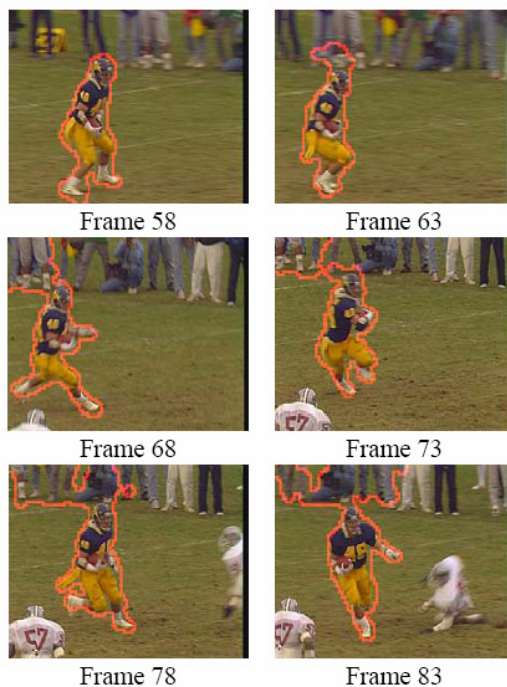
圖七. “Hans”影片的追蹤結果。紅色橢圓圈選出我們所提方法的追蹤結果；而藍色矩形則圈選出傳統平均移動法的追蹤結果。

接下來我們將等位面理論應用在區域追蹤上，以公式(14)當作區域追蹤的模型，我們測試了三種不同的影片，其背景皆是變動的狀態。圖八是car-phone影片的追蹤結果，因為那些距離預定區域 R_0 很遠的像素點，總會在前一時間點影像的 R_0 外側找到相配的像素點，因此，為了節省運算時間，等位面在綠色矩形內的範圍，才做更新動作。從圖八可看出，在這個影片中，應用等位面理論在區域追蹤上的表現良好。圖九是football影片的追蹤結果，這段影片含有一些來自雜亂背景的干擾，其追蹤結果為圖九中橘色曲線圈選區域，我們可看出這樣的追蹤結果受到背景的影響，並不是非常理想。圖十則為coastguard測試影片，同理，等位面只在綠色矩形範圍內做更新動作，以降

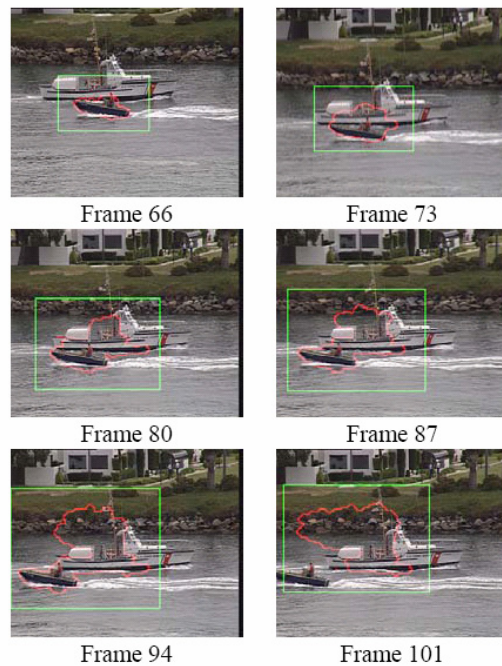
低運算量。在圖十中，因為受到大船的干擾，小船的輪廓曲線傳遞演變的結果並不是很正確，甚至最後完全追丟小船。



圖八. Apply the original region tracking model to the car-phone sequence.



圖九. Apply the original region tracking model to the football sequence.



圖十. Apply the original region tracking model to the coastguard sequence.

(4) 結論

在本計畫中，我們實現了兩種不同種類的偵測和追蹤方法—平均移動法和等位面理論。其中，相較於傳統的平均移動法，我們提出了一個新的目標物模型，稱為多團塊模型，不僅同時保留了目標物的位置和色彩空間的資訊，更進一步地保留了物體的細節資訊，以此多團塊模型為基礎，我們重新定義較符合實際情況的相似度量測，並重新推導其對應的平均移動追蹤程序，未來我們將繼續解決目標物大小和方向的調整，並研究更穩健的方法。另外，我們將等位面理論應用於區域追蹤[7]實作出來，由實驗中我們發現，當移動物體的色彩和亮度很接近背景環境時，此追蹤模型會出現追蹤不當的問題，這也是一般追蹤常會遭遇到的問題，尤其是當背景會隨時間改變下，問題越發難以解決，接下來我們將會著手進行改善此問題。

(5) 参考文献

- [1] Comaniciu D., Ramesh V., Meer P., "Kernel-based object tracking", in Proc. Real-Time Vision & Modeling Dept., Siemens Corporate Res., Princeton, NJ, USA.
- [2] Birchfield, S.T., Sriram Rangarajan, "Spatiograms versus histograms for region-based tracking", in Proc. Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on vol. 2, pp.1158-1163 20-25 June 2005.
- [3] Nikos Paragios, Rachid Deriche, "Geodesic active Contours and Level Sets for the Detection and Tracking of Moving Objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 3, pp.266-280, March 2000.
- [4] Nikos Paragios and Rachid Deriche, "Coupled Geodesic Active Regions for Image Segmentation," *European Conference in Computer Vision*, Nov. 1999.
- [5] Nikos K. Paragios, "Geodesic Active Regions and Level Set Methods: Contributions and Applications in Artificial Vision," PhD Thesis in University of Nice Sophia Antipolis School of Computer Engineering, January 2000.
- [6] Tony F. Chan, Luminita A. Vese, "Active Contours Without Edges," *IEEE Transactions on Image Processing*, Vol. 10, No. 2, pp.266-277, February 2001.
- [7] Abdol-Reza Mansouri, "Region Tracking via Level Set PDEs without Motion Computation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No.7, 947-961, July 2002.
- [8] Yonggang Shi, William Clem Karl, "Real-time Tracking Using Level Sets," *International Conference on Computer Vision and Pattern Recognition (CVPR)* June 2005.
- [9] Kass, M., Witkin, A., and Terzopoulos, D. "Snakes: Active contour models." *International Journal of Computer Vision*, 1:321-331. 1988.
- [10] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic Active Contours," *International Journal of Computer Vision* 22(1), 61-79, 1997.