

# 行政院國家科學委員會專題研究計畫 成果報告

## 總計畫

計畫類別：整合型計畫

計畫編號：NSC91-2213-E-002-126-

執行期間：91年08月01日至92年07月31日

執行單位：國立臺灣大學資訊工程學系暨研究所

計畫主持人：洪一平

共同主持人：陳祝嵩，劉庭祿

計畫參與人員：王嘉銘、劉憲璋、李子洋

報告類型：完整報告

處理方式：本計畫可公開查詢

中 華 民 國 92 年 10 月 31 日

# 行政院國家科學委員會專題研究計畫成果報告

## 視訊中多物體之偵測、分割與追蹤(2/2)－總計畫

### Detection, Segmentation, and Tracking of Multiple Objects in Videos

計畫編號：NSC 91-2213-E-002-126-

執行期限：91年8月1日至92年7月31日

主持人：洪一平

國立台灣大學資訊工程學系

共同主持人：劉庭祿、陳祝嵩 中央研究院資訊科學研究所

計畫參與人員：王嘉銘、劉憲璋、李子洋

#### 一、中文摘要

本整合型計畫的目的在於研究及整合有關「視訊中多物體之偵測、分割與追蹤」之相關技術，其內容包含：視訊中物體之偵測與辨識(子計畫一)、視訊物件之分割(子計畫二)、多個物體之即時追蹤(子計畫三)以及視訊尋人器之實作(總計畫)。總計畫的主要目標在於整合各子計畫所發展的技術，以建立一個以人為目標物的視訊偵測與追蹤系統，即所謂的「視訊尋人器」。基於第一年的執行成果與經驗，我們設計了一個雙攝影機模組，完成其控制校正，並結合各子計畫的相關研究成果，可正確地偵測並追蹤視訊中移動的人物。

**關鍵詞：**視訊尋人器、攝影機校正、視訊追蹤、運動偵測、背景模型學習。

#### Abstract

The goal of this project is to investigate and integrate the techniques for detecting, segmenting and tracking multiple objects in videos. Our research focus includes: investigating techniques for detecting and recognizing 3D objects (sub-project 1), developing new methods for video object segmentation (sub-project 2), developing new techniques for real-time visual tracking of multiple objects (sub-project 3), and implementing a video people tracker (the overall project). In this project, we have built a system for detecting, segmenting, and

tracking people, which we called the *ViPtracker* (Video People tracker). This system adopts two cameras to accomplish its mission: a wide-angle camera is used to monitor the whole area and is responsible for detecting moving targets; a PTZ (pan-tilt-zoom) camera is used to track the detected targets and provides the higher-resolution image detail for future recognition. These two cameras are calibrated off-line, and can be used to detect and track moving people successfully.

**Keywords:** Video People Tracker, Camera Calibration, Visual Tracking, Motion Detection, Background Modeling.

#### 二、緣由與目的

視訊中多個目標物的偵測、分割與追蹤是電腦視覺領域中十分重要且實用的研究課題。其應用範圍相當廣泛，其中又以視訊安全監控系統在過去幾年來特別廣受學術界與產業界的重視[3][4]。

本整合型計畫的目的在於研究與整合相關的視訊安全監控技術，其工作內容包含三個研究子題，分別對視訊物體之偵測、分割與追蹤作深入探討。其實每個子題均有其重點及特殊的困難需要解決，但彼此之間的關係非常密切，不容易找出清楚的界限，所以希望透過此一整合型計畫來研究及整合這些技術，並藉由總計畫來實作一個可以自動偵測與追蹤人物的視訊

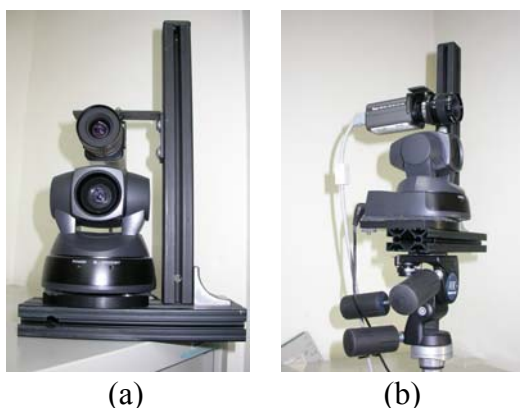
監控系統，驗證相關技術之可行性。在第一年中，總計畫已建立一套共用的視訊處理模組（包含攝影機校正、深度估算和光流計算等），完成系統整合的準備工作。在第二年的工作中，我們進一步建立一套「視訊尋人系統」，可利用廣角攝影機自動偵測進入監控區域內的目標物，然後控制 PTZ (pan-tilt-zoom) 攝影機，結合其視訊回饋訊號與廣角攝影機的偵測結果，穩定持續地追蹤目標物，取得高解析度的監控影像。

### 三、結果與討論

本年度的工作內容及具體成果如下：

#### 總計畫：

為了建立一套穩定的「視訊尋人系統」，我們採用兩部攝影機來完成視訊目標物之偵測與追蹤任務。其中，廣角攝影機的目的在於將整個監控區域同時納入影像的視域中，以確保不會遺漏任何目標物。另外，我們再利用一部 PTZ 攝影機，使用較高解析度的影像來追蹤目標，以彌補廣角攝影機解析度太低的缺點。如圖一(a)(b)所示，我們係將一部廣角攝影機和一部 PTZ 攝影機以垂直相鄰的方式固定在支架上，然後進行兩部攝影機相對關係的校正。當目標物進入監控區域後，首先會被廣角攝影機偵測到，然後驅動 PTZ 攝影機持續地追蹤目標物的移動，直到其離開監控範圍。以下分別對雙攝影機組校正、目標物偵測、以及目標物追蹤等工作做進一步的說明。



圖一：動態雙攝影機組，上為廣角攝影機，下為 PTZ 攝影機。(a) 正面圖，(b)側面圖

#### (1) 廣角攝影機和 PTZ 攝影機的校正

我們首先在廣角攝影機所拍攝到的影像中找出一些控制點（如圖二之上圖所示）。接著轉動 PTZ 攝影機去對準這些控制點，並記錄下 PTZ 攝影機的每個旋轉角度與廣角影像座標間之對應關係。然後再採用內插法決定廣角攝影機的影像座標與 PTZ 攝影機控制角度間的幾何對應關係。一旦找到這個對應關係之後，我們的 PTZ 攝影機即可經由電腦自動控制，依實際需要，對準廣角攝影機之畫面上的每一個影像點（如圖二之下圖所示）。



圖二：上圖為廣角攝影機所看到的畫面以及其所使用的校正控制點。下圖為某一個校正控制點之下 PTZ 攝影機所對準的畫面。

#### (2) 目標物偵測

在目標物偵測部分，經由分析比較後，我們決定採用並修改 Larry Davis 所提出的方法[2]。這個方法係針對影像中的每個像素，利用最近一連串所觀測到的顏色分佈來一個建立背景模型。接著再比較新進像素與背景模型的差異，以決定該像素是否為移動物體或雜訊。背景模型依其更新方式可分為「短期模型」與「長期模型」兩種。短期模型是將所有新進之影像資料一

律加入背景模型中。好處為可以適應背景情況之改變。缺點則是由於不論前景或背景都被當作背景模型的學習資料，故前景亦有可能被納入背景模型中而導致某些前景物體可能無法被正確的偵測出來。長期模型僅僅將被歸類為背景的部分加入背景模型的學習當中。優點為偵測效果較佳，缺點則是較不易反應出背景光線或其他因素的變化，而造成持續性的警報。在此一計畫中，我們結合了長期與短期的背景模型，各取其優點，以提升目標物偵測的正確率。

### (3) 目標物追蹤

由於廣角攝影機的視角涵蓋範圍較廣，因此其所偵測到的目標物之影像解析度通常都不是很高（除非目標物距離攝影機非常近）。為了能作後續的影像分析與目標辨識，我們採用 PTZ 攝影機以 zoom-in 的方式來持續追蹤鎖定的目標物，以取得解析度較高的視訊。首先，我們的系統會利用前述的校正結果，控制 PTZ 攝影機自動鎖定廣角攝影機所偵測到的移動目標物，然後再利用視訊回饋的方式，來持續追蹤移動的目標物。在此，我們係採用子計畫三所發展的 trust-region method [9]，其追蹤效果相當不錯。圖三顯示某個移動目標物的偵測追蹤結果。在 PTZ 攝影機的追蹤過程中，我們的系統會不斷地拿廣角攝影機的偵測結果來對照。如果因為目標物移動速度太快而導致 PTZ 攝影機的追蹤失敗，我們的系統可隨時使用廣角攝影機的偵測結果來恢復 PTZ 攝影機正確的追蹤程序。

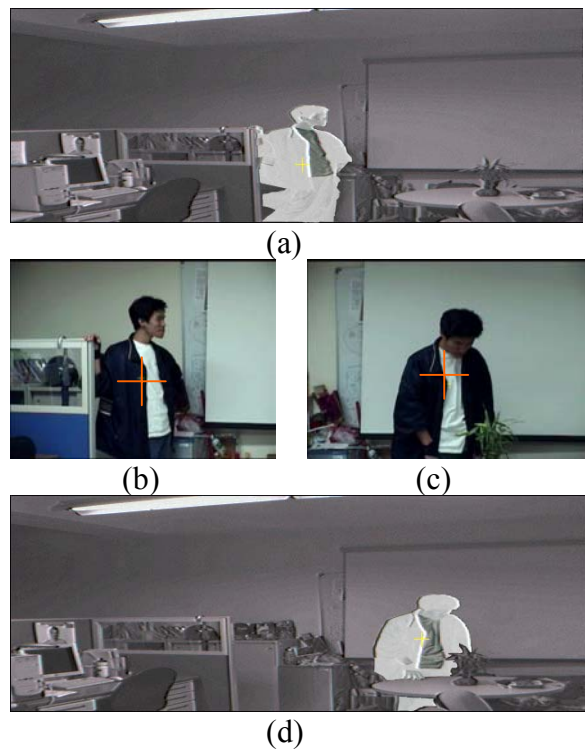
#### 子計畫一：

在第一年的計畫中，我們提出一個以外觀印象為基礎的方法，來追蹤與辨別視訊中的三維物體。在第二年中，我們發展了由連續影像序列辨識人臉的方法，並設計了一套即時的人臉影像樣本取樣系統。

在人臉追蹤部分，我們利用方位估測與位置估測交替執行來求最佳解，改善第

一年計畫中需要進行費時的最佳化運算，以達到即時追蹤的效果。首先我們將第一張影像進行人臉的偵測，接著進行方位估測並依結果產生樣板，再送至位置估測的步驟進行位置比對，找出更適合的人臉位置，再送回方位估測步驟中估測，如此交替執行直到找出最佳解。

在位置估測中，我們採用贏者更新樣本比對方法(winner-update strategy) [10] 快速地找出影像中人臉的位置。在方位估測中，我們共發展了四種方法，分別為直接樣本比對+人臉特徵線、Fuzzy Kernel Perception (FKP)、Support Vector Regression (SVR)、人臉特徵平面+動態規劃法，來估測查詢影像的人臉方位。



圖三：目標物偵測與追蹤的實驗結果。(a)白色區域為偵測到的移動目標物，(b)PTZ 攝影機開始鎖定目標物，(c)PTZ 攝影機的追蹤過程畫面，(d)廣角攝影機所拍攝到的對應畫面及其即時偵測之結果。

我們事先收集不同使用者在預定的 25 個方位之下的平均臉。直接樣本比對+人臉特徵法是將詢問影像對最相近的兩張平均人臉樣本的特徵直線做投影，並利用

線性內插估測出方位。在 FKP 的方法中，除了平均臉之外，更利用到每個方位下所收集到的所有人臉來幫助方位估測，並且把方位估測視為一個有 25 類的方位分類問題，利用 FKP 分類器來進行分類。由於方位估測並不適合被視為一個離散的分類問題，SVR 的方法就是將方位估測視為一個迴歸問題，利用事先收集的人臉影像當訓練資料，訓練出估計 *pan*、*tilt* 的兩個迴歸模型。

在連續影像的人臉辨識部分，我們採取最近特徵線法，這個方法的效果比傳統的最近鄰居法還要好。最近特徵線法一般只考慮查詢樣本跟特徵線的距離資訊，而忽略了相鄰特徵線的相關資訊。在我們的方法中，將一併考慮相鄰特徵線的資訊。我們定義給定查詢影像之下，特徵線集合的事後機率，並且將其轉換成圖形理論的最短路徑問題，再利用現有的方法求其最佳解。詳細內容請參閱子計畫一的報告。

### 子計畫二：

在第一年的計畫中我們將傳統的分水嶺分割法由單張影像之分割推展到視訊物件之分割，並發展一套三維時空容積的分水嶺分割技術。今年我們除了改進去年所發展的方法之外，並進一步發展多層次的視訊物件分割技術。

本年度的新方法可以分為兩個階段。(1)前處理階段：首先我們會為輸入影像建立不同解析度的多階層影像，再利用分水嶺分割法把每一階層的影像切割成分水嶺區域，接著再估算每個分水嶺區域的運動向量，然後為每一階層的分水嶺區域建立一個時空相鄰區域圖。(2)標記階段：此階段的目的是要將時空相鄰區域圖的每一個節點標記成「前景」、「背景」或「不確定」區域。我們的方法是從最粗略的階層依序往最細緻的階層進行切割。在每一階層中我們的系統都會使用最大事後機率法來求得一個較佳的切割。較粗略階層所切割的結果會被拿來當做是較細緻階層的最佳化之初始值。最粗略的初始值則可以由使用者給定或是由某些較粗略的全自動方

法求得。當進行完最精細層的最大事後機率分割後，使用者可以檢查分割結果。如果有些地方不如使用者預期，使用者可以透過系統所提供之界面來修正。當使用者在少量的影像進行修正後，系統會自動地重新執行最大事後機率分割，以把正確的資訊傳遞到其它影像中。詳細內容請參閱子計畫二的報告。

### 子計畫三：

去年的目標物追蹤成果係使用固定式攝影機所取得的視訊。今年則是進行在動態背景下的追蹤研究。對於動態追蹤問題來說，在被追蹤物體的形狀改變時，直接的圖案比對法所獲得的結果會變得比較不準確。因此，我們決定發展新的方法。

對於物體所在的位置，我們用一個橢圓形來表示物體涵蓋的範圍，並且以離橢圓中心的遠近來給予權重。我們計算橢圓形內每個像素所形成的色彩分佈直方圖，並訂定橢圓區域中物體的顏色機率分佈函數。建立物體模型之後，我們將追蹤問題轉換成最佳化問題。在資訊理論中，Kullback-Leibler Divergence 可以用來比較兩個機率分佈的差異。我們要解的最佳化問題就是找到最好的參數(包含位置、大小、方向資訊)，使其所在位置的顏色機率分布和目標物體的顏色機率分佈最相像。除了顏色分佈外，輪廓的資訊也被我們放入最佳化問題中，並且以火山口形的權重函數調整兩者的重要性。

為了快速解這個最佳化問題，我們設計了一個信賴區域法(trust-region method)，利用疊代方式收斂。每一次疊代可分為三個步驟：(1)解決子問題，我們利用泰勒展開式定義一個二次函數來估計原始要最佳化的函數，並假設我們要的解落在某個信賴區域之內，(2)衡量信賴區域的可信度，測試在某個信賴區域下，目標函數的值在每一步確實降低，(3)依照上一步所計算的可信度調整信賴區域。如果可信度很大，就可以放大信賴區域來做下一步的搜尋，反之，就要在更小的範圍內做下

一步的搜尋。在我們的實驗中，信賴區域法比傳統數值方法中的 Steepest descent 可以獲得較佳的結果。詳細內容請參閱子計畫三的報告。

#### 四、計畫成果自評

本整合型計畫之目標在於結合電腦視覺、影像處理與圖形識別等相關技術，偵測、分割與追蹤視訊中所含的移動人物。子計畫一的主要研究工作項目是視訊中物體之偵測與辨識，子計畫二的目標是要發展視訊物件之多層次分割技術，子計畫三則要研發多物體之即時視訊追蹤技術。整體而言，各子計畫均獲得相當可觀的研究成果，分別都有一些相當不錯的會議論文發表，而且也有期刊論文正在審查中。總計畫的主要目標不在新技術的開發，而是在於利用各子計畫所研發的相關技術，實作一個「視訊尋人系統」。我們最終所整合出來的系統具有兩個攝影機：廣角攝影機負責周遭環境的監控偵測與較低解析度的追蹤，而 PTZ 攝影機則負責高解析度的目標物追蹤。此一系統的功能大致都能滿足原先的預期。

基於執行此一計畫之成果，我們正進一步與法國 INRIA 的 Thonnat 研究團隊進行跨國合作，探討有關視訊安全監控方面的研究議題。當然，我們的系統還有不少需要改進的地方。例如，當偵測到的目標物數目大於 1 時，我們的系統目前係選擇讓 PTZ 攝影機追蹤最大目標。事實上，如何選擇適當的追蹤目標是一個相當複雜的問題，牽涉到監控的目的與任務，同時系統還必須能辨識目標物的身份，因此仍有待未來繼續努力，才能逐一解決這些問題，達到利用視訊從事智慧型的自動安全監控之最終目標。

#### 五、參考文獻

[1] D. Comaniciu, V. Ramesh and P. Meer, "Real-Time Tracking of Non-Rigid Objects using Mean Shift," *IEEE CVPR*, 2000.

- [2] A. Elgammal, D. Harwood, and L. Davis, "Non-parametric Model for Background Subtraction," *European Conference on Computer Vision*, 2000, p.p. 751-767.
- [3] A. Mittal and L. Davis, "M2Tracker: A multi-view approach to segmenting and tracking people in a cluttered scene using region-based stereo," *Seventh European Conference on Computer Vision*, 2002.
- [4] I. Haritaoglu, R. Cutler, D. Harwood, and L. Davis, *Backpack: Detection of People Carrying Objects Using Silhouettes*, Proceedings of International Conference on Computer Vision, ICCV99, pp. 102-107.
- [5] B. K. P. Horn and B. G. Schunck, "Determining Optical Flow", *Artificial Intelligence*, vol. 17, pp.185-203, Aug. 1981.
- [6] Y.-P. Hung, D. B. Cooper, B. Cernuschi-Frias, "Asymptotic Bayesian Surface Estimation Using an Image Sequence," *International Journal of Computer Vision*, 6(2): 105-132, June 1991.
- [7] M. Irani, "Multi-frame Optical Flow Estimation Using Subspace Constraints", *International Conference on Computer Vision*, pp.626-633, 1999.
- [8] T. Kanade and M. Okutomi, "A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, No.9, pp.920-932, 1994.
- [9] Tynn-Luh Liu and Hwann-Tzong Chen, "Real-time Tracking Using Trust-Region Methods", to be appeared in PAMI, 2003
- [10] Y. S. Chen, Y. P. Hung and C. S. Fuh, "Fast Block Matching Algorithm Based on the Winner Update Strategy," *IEEE Transactions on Image Processing*, Vol. 10, pp. 1212-1222, 2001.