

行政院國家科學委員會專題研究計畫 成果報告

適用於監視系統之人類異常行為自動偵測 研究成果報告(精簡版)

計畫類別：個別型
計畫編號：NSC 95-2221-E-343-006-
執行期間：95年08月01日至96年09月30日
執行單位：南華大學資訊工程學系

計畫主持人：廖怡欽
共同主持人：賴榮滄
計畫參與人員：碩士班研究生-兼任助理：邱舶軒、陳易顯、胡光南
大學生-兼任助理：唐巧薇、劉苓、廖峰熙、張瑋斌、吳誌軒、黃耀震

報告附件：出席國際會議研究心得報告及發表論文

處理方式：本計畫可公開查詢

中華民國 96 年 10 月 13 日

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

適用於監視系統之人類異常行為自動偵測

計畫類別： 個別型計畫 整合型計畫

計畫編號：NSC 95-2221-E-343-006-

執行期間：95年8月1日至96年9月30日

計畫主持人：廖怡欽

共同主持人：賴榮滄

計畫參與人員：邱船舶、陳易顯、唐巧薇、劉苓、廖峰熙、張瑋斌、吳誌軒、黃耀震

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：南華大學 資訊工程學系

中 華 民 國 96 年 10 月 5 日

摘要

由於視訊監視系統的普及，分析監視系統取得之視訊內容的應用也越來越普遍。本計劃主要目的在開發一套人類異常行為分析系統，可針對固定式監視系統取得之視訊內容，自動偵測是否有人類異常行為發生。此系統主要包括：監視系統修正、物件分割、物件追蹤、動作追蹤、與異常行為識別等過程。監視系統修正過程用於去除鏡頭所造成的陰影現象以及修正影像色調曲線的非線性特性；物件分割過程可由輸入畫面中分割出有興趣的物件；物件追蹤過程負責在連續畫面中追蹤物件；動作追蹤過程依物件移動軌跡，物件外形變化，及物件內部變化產生適當的動作集合；最後，所產生的動作集合再與人類異常行為動作集合做比對，以確定該物件的行為是否符合人類異常行為。本計劃成果適合居家照護，保全監測，或犯罪行為偵測等應用，可自動偵測人類異常行為，適時提出警訊。本計劃第一年已完成一套快速物件分割方法的開發及實作，相關研究成果也已在多個研討會[1-3]發表。

關鍵字：背景估測，物件分割

Abstract

Due to the popularity of the surveillance systems, using these systems to analyze captured video stream are becoming more and more popular. This project is planned to develop a system for detecting abnormal human behaviors for fixed surveillance systems. Five processes are included in this system: the correction process of surveillance systems, object segmentation process, object tracking process, action tracking process, and abnormal behavior identification process. The correction process is used to remove the shading effect caused by lens and linearize the tone curves of captured images. In the object segmentation process, the interested objects are separated from an image. The isolated objects are then tracked and their moving paths are recorded in the object tracking process. To identify the behavior of an object, its action set is generated in the action tracking process according to the moving path, variations of contours and contents. Finally, the behavior of an object is determined by comparing its action set with abnormal action sets. The proposed method can detect abnormal human behavior from the video streams of surveillance systems and is especially useful in the in-home caring, security guarding, and criminal detection applications. In the first year of this project, we have developed and implemented an effective fast object segmentation approach. Some results of this project have published in some conferences [1-3].

Keywords: Background Estimation, Object Segmentation

1.前言

由於監視系統的應用愈來愈廣泛，分析監視系統視訊內容的工作也變得愈來愈重要，諸如交通監測、居家照護、遠端監測、犯罪偵查、工業生產、機器人視覺等都有判斷影像或視訊內容的需要。目前視訊內容的判斷工作大多仍需大量人力介入，無法有效快速的完成分析工作。自動視訊分析系統可快速分析視訊內容，找出視訊內容中感興趣的物件，完成視訊檢索或行為辨識工作。本計劃主要目的在開發一套可自動分析固定式監視器內容判斷出人類特定異常行為的人類異常行為分析系統，此系統適用居家照護，保全監測，或犯罪偵測等應用，可自動偵測人類異常行為，適時提出警訊。本計劃第一年工作主要著重在適用於固定式監視系統之物件分割方法開發。物件分割主要目的是將視訊內容中感興趣的區域或者是移動中的物件分割出來，供各種視訊處理應用使用，例如：在保全監測系統[4]中，可將可疑的人或物件分割出來後，再加以追蹤。在 MPEG-4[5]中，使用物件分割技術將視訊內容分割成前景與背景區域，然後再應用不同的壓縮方式處理不同區域。在交通監測[6]上，可將區域內違規的車輛分割出來並自動拍照。目前主要的物件分割方法有光學流程[7]-[8]和背景抑制物件分割方法[9]-[14]兩種。光學流程法主要作法是將畫面中的像素依與相鄰像素間的關係分成不同群體，再計算每個群體由前一張畫面至目前畫面的移動向量，最後再把具相似移動向量的群體合併成最後的物件。此類方法利用物件在時間軸上的移動資訊分割物件，可有效分割物件，但計算複雜度高，處理時間久，不適合具即時要求的應用。背景抑制法利用目前畫面與參考畫面（可使用背景畫面或前張畫面）的差異值來分割物件，此類方法計算複雜度低，處理時間短，但僅適合固定式監視環境。

2.研究目的

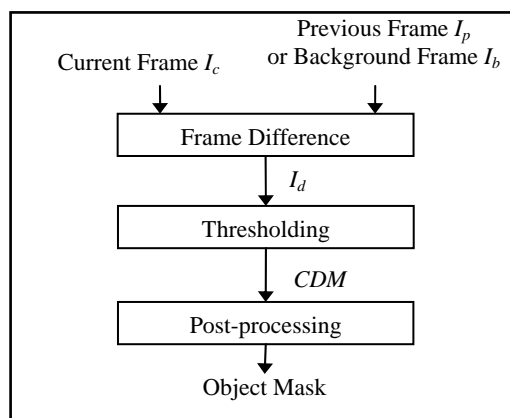
本計劃第一年主要目的在開發適用固定式監視系統之物件分割方法。其中，背景抑制法較適合本計劃需求。在現有背景抑制方法中，背景畫面估測與陰影區域去除是兩個主要的問題。現有背景畫面估測方法在當背景內容有變動時，通常需要一段回復時間，回復時間內背景資訊不全，會導致較高的物件分割錯誤率。另一方面，由於背景抑制方法易受監視環境光源變化與物體陰影的影響而提高物件分割錯誤率，為降低錯誤率，必須有效去除陰影區域，現有陰影區域去除方法，不是陰影去除效果不佳就是需要耗費許多時間。為解決以上問題，本計劃分別提出一個新的多層式背景估測方法[1,15]以及一個具陰影去除能力之快速物件分割方法[2-3,16]。

3.文獻探討

本節將介紹傳統的背景抑制方法、背景估測方法、以及陰影去除方法。

3.1 背景抑制方法

背景抑制方法利用目前畫面與參考畫面的差異值來分割物件，主要流程如圖一所示。首先計算目前畫面(current frame) I_c 與前張畫面(previous frame) I_p 或背景畫面(background frame) I_b 的差異值，產生差異值畫面 I_d ；然後，使用門檻值(threshold)區分每個像素是物件像素或背景像素產生改變偵測遮罩(Change Detection Mask, CDM)；最後，使用後置處理(post-processing)來填補 CDM 的空洞或去除 CDM 的雜訊，產生物件遮罩(object mask)。



圖一·背景抑制方法處理流程

3.2 背景估測方法

目前有許多學者[9]-[11]提出不同的背景畫面估測或更新的方式。Chien 等人[9]使用背景註冊方法估測背景。Spagnolo 等人[10]計算目前影像與背景影像的像素平均亮度增益值，然後再利用像素平均亮度增益值更新背景影像，使背景影像能反映光源變化。Colombari 等人[11]則記錄像素於一連串畫面中的值，再據以找出最適合作為背景的像素值。現有背景估測方法主要問題在於當物件進入畫面並停留一段長時間後，物件會取代原先背景成為新背景，而當物件離開後，需要一段時間才能回復原有的背景；很明顯的，在背景回復期間，由於可用的背景資訊不足，會提高物件分割的錯誤率。

3.3 陰影去除方法

為了去除陰影區域對物件分割的不良影響，有許多改良式背景抑制物件分割方法[9,12-14]被提出。Chien 等人[9]除了提出背景註冊方法外，為了避免光源變化與陰影的影響，改用目前畫面與背景畫面的邊緣強度資訊做為背景抑制方法的輸入找出物件邊緣像素，然後再利用後置處理方法填滿物件內部空洞。Stauder 等人[12]分析陰影的行為，將陰影分成均勻的陰影與半陰影兩種，然後利用像素於目前畫面與前張畫面的差異值、比值、與邊緣強度等資訊識別及去除陰影區域。Xu 等人[13]針對室內環境，使用目前畫面與背景畫面的差異值、邊緣強度差異值、以及陰影成長方法找出及去除陰影區域。Jacques 等人[14]則使用正規化相關值(Normalized Cross-Correlation, NCC)找出目前畫面與背景畫面中符合陰影特性的像素(稱為陰影候選像素)。陰影候選像素及其周圍像素於目前畫面與背景畫面之比值的變異數小於給定門檻值且陰影候選像素於目前畫面的值小於其於背景畫面的值時，可被視為陰影像素去除。以上所提陰影去除方法中，Stauder 與 Xu 的方法需要花費較多運算時間，不適合即時應用。Chien 與 Jacques 等人提出的方法則具有較快的執行速度，但物件分割錯誤率也較高。

4.研究方法

為了降低物件分割錯誤率以及縮短物件分割執行時間，本計劃分別提出一個多層式背景估測方法[1,15]以及一個具陰影去除能力之快速物件分割方法[2-3,16]。這兩個方法分別說明如下：

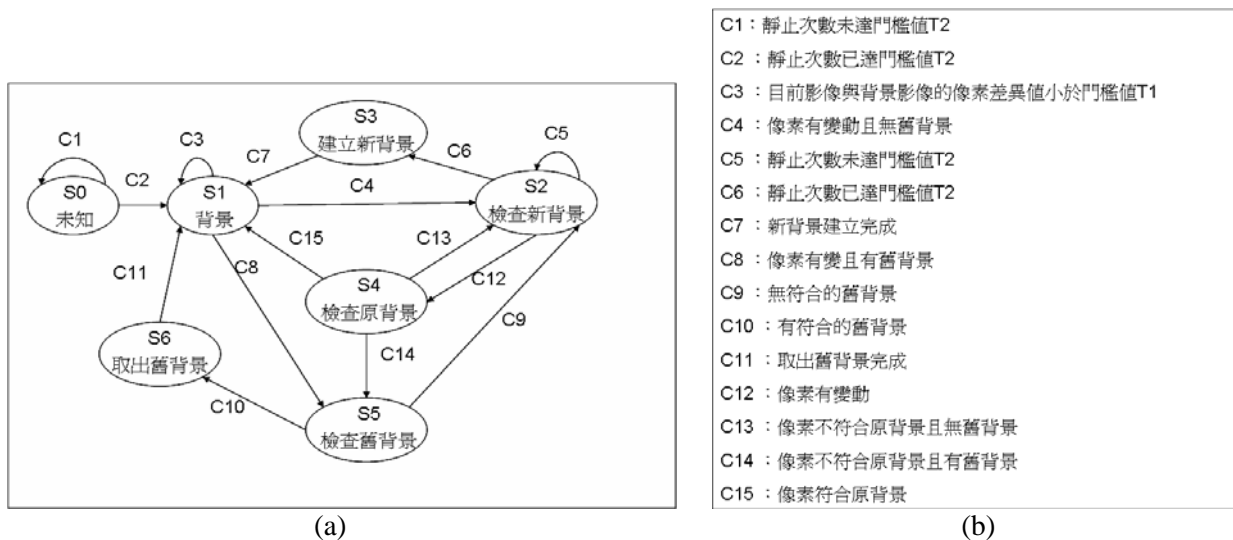
4.1 多層式背景估測方法

為了降低背景回復期間所造成的背景錯誤情況，本計劃提出一個多層式背景估測方

法，所提方法可儲存多層背景，當物件在視訊畫面中停留超過一段時間將被用來更新背景時，我們的方法會將原來的背景記錄下來，並新增一層新的背景；當新背景離開後，原來被遮住的背景內容則可由多層背景中取出，快速回復背景，減少背景回復所需等待時間。為了維護多層背景，針對監視畫面中的每一個位置（點），我們使用如圖二(a)所示的狀態圖來更新每個位置的多層背景資訊。畫面中每個位置的狀態可能位於圖二(a)中的任一狀態(S0-S6)，且當圖二(b)中的轉移條件(C1-C15)成立時可轉移至轉移條件箭頭所指到的目的狀態。圖二(a)所示各狀態分別說明如下：

■ 未知狀態(S0)

背景估測過程一開始，畫面中的每個位置都沒有背景可用，此時每個位置皆處於未知狀態(S0)，且設定靜止累積次數為0。當某一位置處於未知狀態(S0)時，針對每一個畫面，計算此位置於目前畫面與前張畫面的像素差異值，若差異值小於門檻值 T1，則靜止累積次數加 1，否則靜止累積次數歸零。若靜止累積次數未達門檻值 T2(條件 C1)，此位置的狀態仍為未知狀態(S0)；否則(條件 C2)設定此位置的狀態為背景狀態(S1)，設定此位置可用背景層數為 1，並設定此位置第一層(最上層)的背景像素為此位置於目前畫面的像素值。



圖二·多層式背景估測方法 (a)處理狀態圖, (b)狀態轉移條件

■ 背景狀態(S1)

當畫面中某一位置處於背景狀態(S1)時，表示該位置有背景像素可用。此時該位置作用中的背景像素為其最上層的背景像素。針對每個位於此狀態的位置處理過程為，計算此位置於目前畫面與背景畫面的像素差異值。若差異值小於 T1(條件 C3)，表示此位置於目前畫面中的像素不是物件，此時使用目前畫面中的像素來更新背景像素；若差異值大於 T1 且此位置的背景層數大於 1(條件 C8)，表示此位置的背景改變，由於有其它層背景可用，所以將此位置的狀態改為檢查舊背景狀態(S5)，檢查是否為舊背景出現。若差異值大於 T1 且沒有其它層背景可用(條件 C4)，則須檢查是否有新背景出現或只是物件經過，因此須將此位置的狀態改為檢查新背景狀態(S2)並設定靜止累積次數為 0。此時此位置的背景像素仍可使用。

■ 檢查新背景狀態(S2)

針對每個位於此狀態的位置處理過程為，計算此位置於目前畫面與前張畫面的像素差異值，若差異值大於 T1(條件 C12)，表示此位置沒有新背景，此時必須檢查此位置目前畫面的像素是否為原來的背景，而需將此位置的狀態改為檢查原背景狀態(S4)；若差異值

小於 T1，則靜止累積次數加 1，若靜止累積次數小於 T2(條件 C5)，此位置仍維持在檢查新背景狀態(S2)；若靜止累積次數等於 T2(條件 C6)，則設定此位置的狀態為建立新背景狀態(S3)。

■ 建立新背景狀態(S3)

此狀態處理過程為，若此位置的目前背景總使用層數小於最大可用層數，則設定背景總使用層數加 1，然後將每一層背景往下一層搬，所空出的最上層背景(作用中背景)設定為新背景(此位置於目前畫面的像素值)，新背景建立完成(條件 C7)，設定此位置的狀態為背景狀態(S1)。

■ 檢查原背景狀態(S4)

檢查此位置於原背景(最上層背景)的像素是否與目前畫面中的像素相同(差異值小於 T1)，如果是(條件 C15)，則將最上層的背景像素設定為目前影像中的像素，並且設定此像素的狀態為背景狀態(S1)；否則(條件 C14)，設定此位置的狀態為檢查舊背景狀態(S5)。

■ 檢查舊背景狀態(S5)

檢查除了最上層之外的其它層背景是否與目前畫面中的像素相同(差異值小於 T1)，檢查順序為由上至下檢查。如有找到符合的背景，即差異值小於 T1(條件 C10)，將狀態設為取出舊背景狀態(S6)；否則轉移狀態至檢查新背景狀態(C9)並設定靜止累積次數為 0。

■ 取出舊背景狀態(S6)

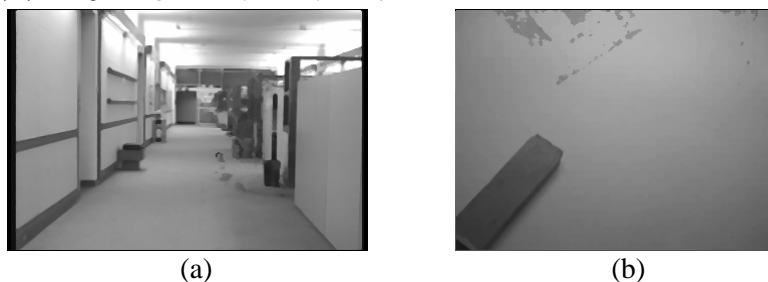
將符合的背景層上方各層往下搬，然後將最上層背景設定為此位置於目前畫面中的像素值，再設定目前像素狀態為背景狀態(S1)，完成取出舊背景工作(條件 C11)。

使用上述方法所產生的背景，會受雜訊與光源變化的影響，導致當背景改變時，舊背景無法完整回復，而產生如圖三所示的結果。



圖三·背景影像：(a)Hall, (b)Block.

為了改善上述問題，產生背景畫面後，再對背景畫面進行中位數濾波(median filter)處理(使用 3×3 遮罩)，處理後的結果如圖四所示。

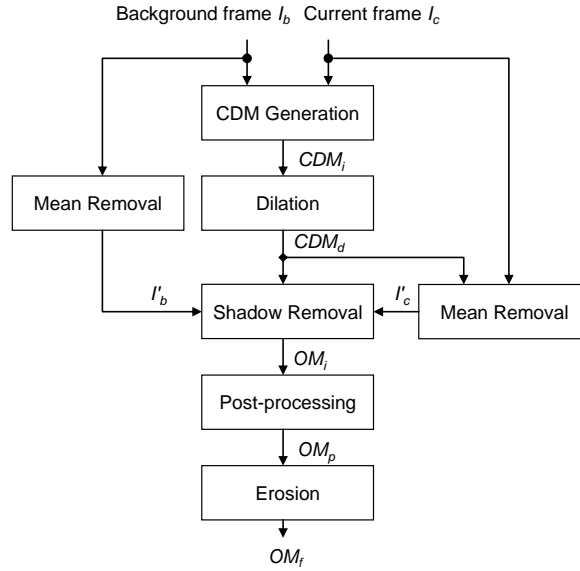


圖四·中位數濾波處理後的背景影像：(a) Hall, (b) Block.

4.2 具陰影去除能力的快速物件分割方法

為了有效且快速的去除陰影區域分割物件，本計劃利用去掉平均值之目前畫面與去掉

平均值之背景畫面的差異值來分割物件。由於去掉平均值的畫面可除去光源對背景的影響並保留畫面的紋理資訊，因此可有效避免背景受光源變化的影響。所提方法比現有方法需要更少的計算量，且由於紋理資訊中也包含邊緣資訊，因此使用紋理資訊分割物件比單純使用邊緣資訊[10]分割物件效果更好。所提方法的主要流程如圖五所示，首先使用背景畫面 (background frame) I_b 與目前畫面 (current frame) I_c 求得初始 CDM ，然後使用去掉均值的背景畫面 I'_b 與去掉均值的目前畫面 I'_c 去除陰影區域，最後再透過後置處理去除雜訊及填補物件內部空洞。



圖五·使用去均值影像之物件分割方法處理流程

圖五所示各功能方塊詳細說明如下：

■ Initial CDM(Change Detection Mask)

此步驟負責計算目前畫面 I_c 和背景畫面 I_b 的差異值，並判斷每個像素的差異值是否小於給定的門檻值 THI 。像素差異值小於 THI ，設定此像素為背景像素(設定此像素的 CDM 值為 0)，否則設定此像素為物件像素(設定此像素的 CDM 為 1)。此步驟輸出初始改變偵測遮罩 CDM_i 。其中， THI 的值需透過實驗決定。

■ Mean Removal

此功能主要目的在去除目前畫面 I_c 與背景畫面 I_b 中每個像素的平均值。針對輸入畫面 I 的每個像素 $I(x,y)$ ，像素 $I(x,y)$ 的平均值 $M(x,y)$ 計算公式如下：

$$M(x, y) = \frac{\sum_{j=-d}^d \sum_{i=-d}^d I(x+i, y+j)}{(2d+1)(2d+1)} \quad (1a)$$

其中 d 的值通常為 1 或 2，實際設定可由實驗決定。

為縮短執行時間，若 $M(x, y-1)$ 的值可用，則 $M(x,y)$ 使用如下的計算公式：

$$M(x, y) = M(x-1, y) + \frac{\sum_{j=-d}^d [I(x+d, y+j) - I(x-d-1, y+j)]}{(2d+1)(2d+1)} \quad (1b)$$

令 $I'(x,y)$ 為像素 $I(x,y)$ 去掉平均值後的結果， $I'(x,y)$ 計算公式如下：

$$I'(x, y) = I(x, y) - M(x, y) \quad (2)$$

為縮短執行時間，在去除目前畫面的平均值時，只計算 CDM_d 中值為 1 的區域。

■ Shadow Removal

此功能主要目的在去除 CDM 中物件區域的陰影像素，首先計算 CDM 中每個物件像素與其周圍像素所組成區塊於 I_c' 與於 I_b' 中的區塊差異值(計算公式如公式(3)所示)，區塊差異值小於給定門檻值 $TH2$ 者即為陰影像素。其中，區塊大小為 $(2d+1) \times (2d+1)$ ， $TH2$ 與 d 的值需透過實驗決定。

$$D(x, y) = \sum_{j=-d}^d \sum_{i=-d}^d [I_c'(x+i, y+j) - I_b'(x+i, y+j)]^2 \quad (3)$$

■ Dilation 與 Erosion

本物件分割方法利用像素與其周圍像素於目前畫面所組成之區塊與於背景畫面所組成之區塊的區塊差異值來判斷像素是否為陰影。此作法在陰影與物件交接處由於有部份物件像素被包含在陰影像素的周圍像素區塊內，以致這類像素通常會被當成物件而無法有效去除。為解決此問題，我們在 Shadow Removal 之前先做 Dilation，將物件邊緣膨脹，陰影區域被膨脹出去的像素位於背景區域，可被當成背景除去，物件區域被膨脹出去的像素雖位於背景區域，因周圍像素區塊包含物件區域，通常無法被當成背景像素去除。因此，在 Shadow Removal 之後，整個物件周圍將會多出一圈背景區域，此時可再使用 Erosion 運算將物件邊緣縮回。其中 Dilation 與 Erosion 運算的球大小固定為 3×3 。

■ Post-processing

此功能主要目的在去除物件遮罩中的小面積區域、雜點、小洞、及平滑邊緣。處理順序為，先利用型態學閉合(close)運算將斷掉的邊緣連結起來，接著使用傳統元件連接方法[17]計算並去除小面積白色與黑色獨立區域，最後利用型態學斷開(open)運算去除雜訊，其中，型態學運算的球大小為 3×3 。

5. 結果與討論

為了比較所提方法與現有方法的好壞，我們分別將兩個方法與現有方法做比較，主要比較項目為物件分割錯誤率，公式定義如下：

$$\text{錯誤率} = \frac{\text{錯誤像素個數}}{\text{一張畫面像素個數}} \quad (4)$$

為求得錯誤像素個數，本計劃以手工分割測試影像，產生一組參考物件遮罩，然後計算待測試之物件分割方法所產生的物件遮罩與參考物件遮罩中不同的像素個數。實驗中，各物件分割方法的參數設定方式以得到最佳錯誤率為主，也就是所設定的參數可讓該物件分割方法產生最佳錯誤率的結果。

5.1 多層式背景估測方法

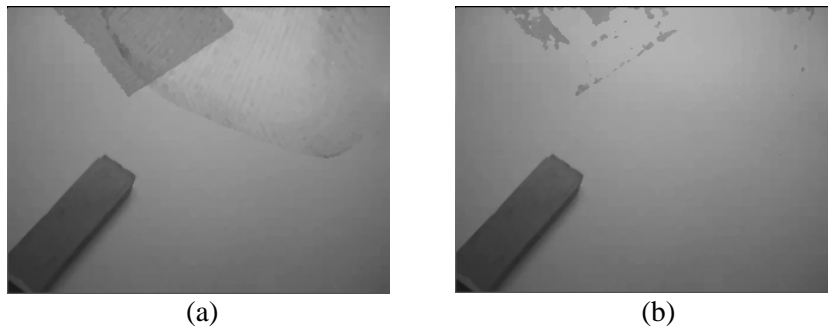
為了解所提背景估測效果的好壞，我們比較了多層式背景估測方法(Proposed)與 Chien 等人[9]所提的背景估測方法(Chien)。實驗使用 Hall 和 Block 兩組測試影片，Hall 與 Block 畫面大小分別為 360×240 與 320×240 ，Hall 與 Block 的測試影像總張數分別為 110 張與 597 張。在我們的方法中，針對 Hall 與 Block 兩組測試影片，T1 與 T2 的值分別設定為(10, 20) 與(20, 50)。實驗開發環境為 MATLAB 7.1。圖六與圖七分別列出使用 Chien 與我們的方法

所產生的兩組背景畫面中差異較大的畫面。



圖六・Hall(第 100 張)：(a) Chien 的方法產生的背景畫面，(b)我們的方法產生的背景畫面。

圖六中，Chien 的方法所產生的背景畫面明顯可見因停留時間太久而被當成背景的身體部份，在身體離開畫面後背景仍需一段重新學習時間才能回復；我們的方法所產生的背景畫面，可在身體離去後快速回復被身體遮住的背景。



圖七・Block (第 387 張)：(a) Chien 的方法產生的背景畫面，(b)我們的方法產生的背景畫面。

圖七所示的背景畫面則是在當畫面左上方的紙張移除後，使用兩種方法所取得的背景畫面。Chien 的方法所產生的背景畫面明顯可見到因停留時間太久而被當成背景的紙張畫面，紙張移除後仍無法馬上回復。我們的方法，在紙張移除後則可快速回復背景。由圖六與圖七可見，我們的方法在畫面中的背景物件移動時，確實可快速還原背景。

為了解背景層數對背景估測執行時間及物件分割錯誤率的影響，我們使用 Hall 影片測試不同背景層數的影響，結果如表一所示。其中，執行時間為使用 Matlab7.1 在 Pentium4 3.0 GHz CPU 的 PC 上估測 110 張背景所需的平均執行時間。

表一・層數對時間與錯誤率之影響(測試影片 Hall)

層數	執行時間(單張)	平均錯誤率
1	2.96 秒	1.93%
2	2.99 秒	1.71%
3	3.00 秒	1.681%
4	3.01 秒	1.6798%
5	3.02 秒	1.6798%

由表一可見層數多寡對執行時間影響不大，而平均錯誤率則會隨層數增加而降低。以平均錯誤率來看，層數由一層增為兩層時改善最多，改善幅度大約 11%；層數由兩層增為三層時，改善幅度則不到 2%，使用四層背景，錯誤率已達最小值。綜合考量執行時間、記憶體需求與平均錯誤率，以 Hall 影片而言，使用兩層背景是較佳的選擇。如果需要更低的錯誤率，則可使用三層背景。表二列出使用兩種方法所產生之背景畫面進行背景抑制物件分割方法後的錯誤率，其中使用的背景層數為兩層。由表二的結果可見，使用我們的方法所產生的背景畫面分割物件，由於我們的方法產生的背景畫面比 Chien 的方法所產生的背景畫面可提早回復被遮住的背景區域，因此可以得到較低的物件分割錯誤率。由表一與

表二可見，當我們的方法只有一層背景時效果也比 Chien 的方法好(錯誤率 2.84%)，主要原因是我們的方法有對背景做中位數濾波而 Chien 的方法沒有。

表二·使用不同背景估測方法所產生之背景畫面進行物件分割的錯誤率比較表

	Chien	Proposed
Hall	2.84%	1.71%
Block	7.91%	6.35%

5.2 具陰影去除能力的快速物件分割方法

為了比較所提具陰影去除能力之快速物件分割方法與現有方法在分割效果與執行時間上的差異，本實驗使用 Hall 與 Frank[9] (如圖八所示) 前 90 張畫面做為測試影片，Hall 與 Frank 畫面大小分別為 360x240 與 320x240。為了降低背景畫面對物件分割方法的影響，兩段影片的第一張畫面被用作背景畫面。此實驗比較我們的方法(Proposed)、Chien [9]、與 Jacques [13]等三種物件分割方法。在我們的方法中，針對 Hall 與 Frank 兩段測試影片(*d*, *TH1*, *TH2*)參數分別設定為(2, 25, 625)與(2, 55, 400)。所有實驗均在同一台個人電腦下執行，電腦配備為 Pentium D 2.8 GHz CPU、1GB 記憶體、作業系統為 Microsoft Windows XP Professional、程式開發環境為 Microsoft Visual C++ 6.0。



(a) (b)

圖八·測試影片：(a) Hall (b) Frank

表三列出各物件分割方法處理 Hall 與 Frank 兩組測試影像的執行時間與錯誤率，表三內所有時間皆為執行 10 次的平均值，時間單位為毫秒(ms)。從執行時間來看，我們所提出的方法速度最快。與 Chien 的方法相比，我們的方法執行時間提昇了大約 47%~58%。與 Jacques 的方法相比，我們的方法執行時間提昇了大約 7%~39%。從物件分割錯誤率來看，我們的方法也比其他方法來的好。

表三、各物件分割方法平均每張畫面執行時間與錯誤率

	Chien		Jacques		Proposed	
	執行時間	錯誤率	執行時間	錯誤率	執行時間	錯誤率
Hall	21.51 ms	0.719%	9.84 ms	0.610%	9.13 ms	0.544%
Frank	17.63 ms	0.426%	15.13 ms	0.459%	9.27 ms	0.412%

5.3 結論

本計劃本年度已開發完成一套適用於固定式監視系統的物件分割方法，所開發的物件分割方法包括多層式背景估測與具陰影去除能力之快速物件分割兩部份。多層式背景估測部份可存放多層背景，當畫面背景物件移動時，所顯露出來的背景區域可使用多層背景畫面中的內容來更新，加快背景畫面回復速度，提昇估測得之背景畫面的可用性。實驗結果顯示，當背景改變時，所提方法比現有方法可更快速的回復背景。使用所產生的背景畫面進行視訊物件分割，也可明顯降低物件分割的錯誤率。在具陰影去除能力之快速物件分割部份，使用去掉平均值之目前畫面與去掉平均值之背景畫面的差異值來分割物件，由於去

掉平均值的畫面可除去光源對背景的影響並保留畫面的紋理資訊，因此可避免背景受光源變化的影響並可有效去除陰影區域。所提方法只需簡單的去除平均值與區塊差異值運算，計算量較現有物件分割方法少，可更快速完成物件分割工作。實驗結果顯示，我們的方法錯誤率比現有具陰影區域去除能力之快速物件分割方法低，執行速度則可有效提昇大約7%~58%。結合本計劃所提兩種方法，確實可有效降低物件分割錯誤率及縮短執行時間。

參考文獻

- [1] 廖怡欽、陳易顯、黃崇仁、與賴榮滄 "針對視訊串流之多層式背景估測方法," The Second conference on Digital Content Management and Applications, Tainan, Taiwan, June 2007.
- [2] 廖怡欽、邱船軒、黃崇仁、與賴榮滄 "去均值影像之視訊內容物件分割方法," The 5th conference on Communication Applications, Taipei, Taiwan, March 2007.
- [3] Yi-Ching Liaw, Bo-Shuan Chiu, Jim Z. C. Lai, and Patrick Huang, "A fast approach of object segmentation for video sequence," The Ninth IASTED International Conference on Signal and Image Processing, Honolulu, USA, August 2007. (EI)
- [4] Po-Nan Cheng, The application of fuzzy inference to the automatic detection and identification of intruders in security system, Master Thesis, Tatung University, Taiwan, June 2003.
- [5] MPEG-4 Video Group: Technical Description of the Video Encoder (Proposal for MPEG-4 Tests), Doc. ISO/IEC JTC1/SC29/WG11 MPEG95/504, December 1995.
- [6] V. Kastinaki, M. Zervakis, and K. Kalaitzakis. "A survey of video processing techniques for traffic applications," Image and Vision Computing, vol. 21, no. 4, pp 359-381, April 2003.
- [7] A. Giachetti, M. Campani, and V. Torre, "The use of optical flow for road navigation," IEEE Transactions on Robotics and Automation, Vol. 14, No. 1, pp. 34-48, February 1998.
- [8] Z. Kalafatic, S. Ribaric, and V. Stanisavljevic, "Real-Time Object Tracking Based on Optical Flow and Active Rays," in Proceedings of MELECON 2000, Cyprus, May 2000.
- [9] Shao-Yi Chien, Shyh-Yih Ma, and Liang-Gee Chen, "Efficient moving object segmentation algorithm using background registration technique," IEEE Transactions on Circuits and Systems for Video Technology, Vol. 12, No. 7, pp. 577-586, July 2002.
- [10] P. Spagnolo, T. D'Orazio, M. Leo, and A. Distanto, "Moving object segmentation by background subtraction and temporal analysis" Image and Vision Computing, Vol. 24, No. 5, pp. 411-423, May 2006.
- [11] A. Colombari, A. Fusiello, and V. Murino, "Segmentation and tracking of multiple video object," Pattern Recognition, Vol. 40, No. 4, pp. 1307-1317, April 2007.
- [12] J. Stauder, R. Mech, and J. Ostermann, "Detection of moving cast shadows for object segmentation," IEEE Transactions on Multimedia, Vol. 1, No. 1, pp. 65-76, March 1999.
- [13] Julio Cezar Silveira Jacques Jr, Cláudio Rosito Jung, and Soraia Raupp Musse, "Background subtraction and shadow detection in grayscale video sequences," in Proceedings of SIBGRAPI 2005, Brazil, October 2005.
- [14] Dong Xu, Jianzhuang Liu, Xuelong Li, Zhengkai Liu, Xiaou Tang, "Insignificant shadow detection for video segmentation," IEEE Transactions on Circuits and Systems for Video Technology," vol. 15, no. 8, pp. 1058-1064, August 2005.
- [15] 陳易顯, 針對視訊串流之多層式背景估測方法, 南華大學資訊管理系, 碩士論文, 2007年6月.
- [16] 邱船軒, 使用去均值影像之物件分割方法, 南華大學資訊管理系, 碩士論文, 2007年6月.
- [17] R. M. Haralick and L. G. Shapiro, Computer and robot vision: volume I, (Addison-Wesely, 1992).

計畫成果自評

■ 原計畫相符程度與達成預期目標情況

本計畫本年度預計完成工作項目有監視器校正與視訊內容物件分割兩部份，在監視器校正部份，我們已實作一套監視器校正方法可校正監視器色調曲線與鏡頭所造成的畫面不均勻現象[1][2]。在視訊內容物件分割部份，我們也開發實作了一個多層式背景估測方法[3,4]以及一個具陰影區域去除能力之快速物件分割方法[5,6,7]。所從事的研究內容與計劃內容完全相符，計劃所訂定的目標也全數達成。

■ 研究成果之學術或應用價值

本計劃研究成果已在多個研討會[3,5,6]發表，其中多層式背景估測方法[3]發表於台南大學舉辦的數位內容管理與應用研討會，具陰影區域去除能力之快速物件分割方法則發表至北台科技學院所舉辦的通訊應用研討會[5]以及 IASTED 所舉辦的訊號與影像處理研討會[6]。研究成果除了投稿至學術研討會之外，為了提昇研究成果的應用價值，我們也跟本實驗室的另一補助單位工研院資通所合作申請專利[8]，希望能讓研究成果發揮最大的效益。

■ 影像處理人才培育

本計劃執行人員，主要包括三位碩士班學生以及多位大學部學生，碩士班學生主要負責實作本計劃所提的多層式背景估測方法[4]與具陰影區域去除能力之快速物件分割方法[7]以及多個相關方法，對其實作能力的提昇有很大的幫助。大學部學生則幫忙處理計劃管理相關事務以及準備實驗所需視訊資料與手動分割物件遮罩資料，參與人員對計劃與實驗的進行方式，都有更深入的了解。

■ 參考資料

- [1] Yi-Ching Liaw, "Implementation of digital still camera ICC profile generator," PICS 2001, Quebec, Canada, April 2001.
- [2] Yi-Ching Liaw, Chao-Hua Wen, Chun-Yen Chen, Jim Z. C. Lai, "Lightness equalization algorithm for tone curves unknown input device," CVGIP2001, KenDing, Taiwan, Aug. 2001.
- [3] 廖怡欽、陳易顯、黃崇仁、與賴榮滄 "針對視訊串流之多層式背景估測方法," The Second conference on Digital Content Management and Applications, Tainan, Taiwan, June 2007.
- [4] 陳易顯, 針對視訊串流之多層式背景估測方法, 南華大學資訊管理系, 碩士論文, 2007年6月.
- [5] 廖怡欽、邱船軒、黃崇仁、與賴榮滄 "去均值影像之視訊內容物件分割方法," The 5th conference on Communication Applications, Taipei, Taiwan, March 2007.
- [6] Yi-Ching Liaw, Bo-Shuan Chiu, Jim Z. C. Lai, and Patrick Huang, "A fast approach of object segmentation for video sequence," The Ninth IASTED International Conference on Signal and Image Processing, Honolulu, USA, August 2007. (EI)
- [7] 邱船軒, 使用去均值影像之物件分割方法, 南華大學資訊管理系, 碩士論文, 2007年6月.
- [8] 黃崇仁, 李政旻, 廖怡欽, 與賴榮滄, "視訊物件分割方法," 2006/12 申請.

可供推廣之研發成果資料表

 可申請專利

 可技術移轉

日期：96年9月7日

國科會補助計畫	計畫名稱：適用於監視系統之人類異常行為自動偵測 計畫主持人：廖怡欽 計畫編號：NSC 95-2221-E-343-006-學門領域：資訊學門二
技術/創作名稱	視訊物件分割方法
發明人/創作人	廖怡欽，賴榮滄，邱船軒，陳易顯，黃崇仁，李政旻
技術說明	中文： 我們開發了一個多層式背景估測方法以及一個具陰影去除能力之快速物件分割方法。多層式背景估測部份儲存多層背景，當背景物件移開時，所顯露出來的背景區域可使用多層背景畫面中的內容快速更新。具陰影去除能力之快速物件分割方法使用去掉平均值之目前畫面與去掉平均值之背景畫面的差異值來分割物件，所提方法只需簡單的運算，可快速完成物件分割工作。實驗結果顯示，跟現有方法比較，我們的方法可減少 7%~58% 的執行時間，且具有更好的分割結果。
	英文： We have developed a multi-layer background estimation method and a fast object segmentation method with shadow removal capability. The multi-layer background estimation method stores multiple layers of background. When the background object removed, the discovered region may be restored from the layered background quickly. The fast object segmentation method with shadow removal capability extracts objects from a frame in a video stream using the difference information between the mean-removed versions of the current and referenced frames. Due to the simplicity of the proposed method, our approach can extract objects rapidly. Experimental results show that compared with the available approaches, our approach reduces the computation time by 7% to 58% with better segmentation accuracy.
可利用之產業及可開發之產品	可利用之產業：保全系統，交通監控，居家照護，網路視訊 可開發之產品：物件分割函式，背景估測函式
技術特點	低物件分割錯誤率 有效去除陰影區域 快速分割物件
推廣及運用的價值	可應用至監視系統自動判別系統，異常行為自動警示系統，網路視訊監控系統

※ 1. 每項研發成果請填寫一式二份，一份隨成果報告送繳本會，一份送 貴單位研發成果推廣單位（如技術移轉中心）。

※ 2. 本項研發成果若尚未申請專利，請勿揭露可申請專利之主要內容。

※ 3. 本表若不敷使用，請自行影印使用。

出席國際學術會議心得報告

計畫編號	NSC 95-2221-E-343-006
計畫名稱	適用於監視系統之人類異常行為自動偵測
出國人員姓名	廖怡欽
服務機關及職稱	南華大學助理教授
會議時間地點	民國 96 年 8 月 20 日至民國 96 年 8 月 22 日
會議名稱	第九屆 IASTED 訊號與影像處理會議(Signal and Image Processing 2007)
發表論文題目	A fast approach of object segmentation for video sequence

一、參加會議經過

此次出國主要目的是至美國檀香山 Signal and Image Processing 2007 會議 (附件一) 發表論文(附件二), 本人報告時間為民國 96 年 8 月 22 日早上 Session 10: Video processing and coding II 的最後一場。此會議由國際科學與科技開發協會(IASTED)舉辦, 會議文章接受率大約 38%, 會議文章也被 EI 所收錄, 是一個很好的會議; 會議地點在檀香山喜來登威基基飯店, 開會地點氣候風景皆宜人, 可有效調節開會氣氛及舒緩報告人的緊張心情。

會議 8 月 20 日(一)開始, 本人搭乘華航班機 8/19 (日) 早上大約 6 點半到, 對開會地點週遭環境及會議室的分佈先做個了解, 會議地點旁邊就是沙灘, 風景宜人, 雖然太陽很大, 但很涼爽。

會議第一天(8 月 20 日)早上 8:00 報到, 人很多, 排了大約半小時才完成報到, 拿到一個背包, 內含一份會議議程及論文光碟。報告完成後, 進 Niihou 廳聽大會主席介紹信號及影像處理的現況及未來展望, 隨後聽取 Illinois 大學香檳分校的 Dr. Nahrstedt 介紹他們開發的一套浸入式互動系統, 該系統可在不同地點佈置立體取像裝置(由四個攝影機組成, 一個彩色三個黑白), 然後將所擷取到的人像於電腦中合成, 不過仍有許多問題有待克服。接著, 進 Honolulu 廳聽 Filters and wavelets I, 由於第三天我也是在這間房間報告, 可以先熟悉一下報告環境。由於附近有工程進行, Honolulu 廳的報告受到很大的影響, 坐在底下聽不到報告人的聲音, 有點可惜。下午則是一場關於如何建立一個自動程式產生器的演講, 演講者介紹現有多種自動程式產生器的建構方法, 以及介紹一個簡單的自動程式產生器開發過程。今天與多個相關專家交談才知道 IASTED 在同一會場同時有五場會議進行, 所以遇到了不同領域的專家, 大部份是網路專家, 也有生醫與水資源及工業工程的專家。

會議第二天(8 月 21 日) 早上安排加州理工學院 Prof. Emmanuel Candes 介紹他們做的研究, 該研究有關於如何在取樣過程中去除無意義的訊號, 降低資訊量, 演講中展示許多證據證明許多取樣訊號可在取樣時期就被丟棄, 不會造成訊號太大的失真; 不過, 這個作法似乎與常用的訊號壓縮方法類似, 因此受到許多質疑。隨後, 又到 Honolulu 廳聽 Session 4: Video processing and coding I。

會議第三天(8月22日)早上在 Honolulu 廳參加 Session 10: Video processing and coding II，這場會議的主持人是台灣中央大學的林教授，我被安排這場最後一個報告，報告過程順利，報告完成，獲得一張證書，證明我確實有出席會議及上台報告。會議結束後也趁機看看夏威夷的風景及放鬆一下心情。期間遇到也是台灣去的中華大學呂教授、雲科大沈教授、海洋大學的曾教授以及醒吾技術學院的教授。

二、與會心得

在這個會議中，除了聽到許多不錯的演講與論文外，也有機會跟來自不同國家的學者專家交換意見，了解同儕工作現況與想法，增加對整體影像處理領域發展現況的了解以及對未來發展趨勢的掌握度。在這個會議中，也發現國內訊號與影像處理的研究在國際上佔有相當大的地位，以論文數量來看，這次訊號與影像處理會議總共發表 94 篇論文，其中，美國 18 篇、中國 10 篇、韓國 8 篇、日本 7 篇、台灣 6 篇，...，以人口比例來看，台灣的論文產量算是最高的。另外一個發現則是許多大型會議在夏威夷舉辦，會議論文作者被要求一定要參加會議，參加會議人數不少，為夏威夷觀光產業帶來許多生意。

附件一：會議議程

■ MONDAY, AUGUST 20, 2007

07:00 – 08:30 REGISTRATION

08:30 – 09:00 WELCOME ADDRESS

09:00– KEYNOTE ADDRESS – “INTERNET AND MULTIMEDIA SYSTEMS CHALLENGES IN NEXT GENERATION TELE-IMMERSIVE ENVIRONMENTS”

10:00 – 10:30 COFFEE BREAK

10:30 – SESSION 1 – DIGITAL IMAGE AND PROCESSING

10:30 – SESSION 2 – FILTERS AND WAVELETS I

14:00 TUTORIAL PRESENTATION – “BUILDING A MULTI-PURPOSE GENERATOR ENGINE AUDITION”

15:30 – 16:00 COFFEE BREAK

16.00 – TUTORIAL CONTINUED

■ TUESDAY, AUGUST 21, 2007

09:00 – SIP KEYNOTE PRESENTATION – “COMPRESSIVE SENSING”

10:00 – 10:30 COFFEE BREAK

10:30 – SESSION 3 –SPEECH PROCESSING AND ACOUSTICS I

10:30 – SESSION 4 – FILTERS AND WAVELETS I

10:30 – SESSION 5 – COMPUTER VISION AND IMAGING I

14:00 – SESSION 6 – VIDEO PROCESSING AND CODING I

14:00 – SESSION 7 – SIGNAL PROCESSING AND APPLICATIONS

14:00 – SESSION 8 – IDENTIFICATION, DETECTION, AND ESTIMATION

15:30 – 16:00 COFFEE BREAK

16:00 – SESSION 6, 7 AND 8 CONTINUED

■ WEDNESDAY, AUGUST 22, 2007

09:00 – SESSION 9 – SPEECH PROCESSING AND ACOUSTICS II

09:00 – SESSION 10 – VIDEO PROCESSING AND CODING II

09:00 – SESSION 11 – MEDICAL IMAGE AND PROCESSING I

10:00 – 10:30 COFFEE BREAK

10:30 – SESSIONS 9, 10 AND 11 CONTINUED

13:30 – SESSION 12 – BIOMETRICS AND HUMAN IMAGING

13:30 – SESSION 13 – COMPUTER VISION AND IMAGING II

13:30 – SESSION 14 – MEDICAL IMAGE AND PROCESSING II

14:30 – 15:00 COFFEE BREAK

16:00 - SESSIONS 12, 13 AND 14 CONTINUED

附件二：發表論文

A FAST APPROACH OF OBJECT SEGMENTATION FOR VIDEO SEQUENCE

Yi-Ching Liaw¹, Bo-Shuan Chiu¹, Jim Z. C. Lai², Tsung-Jen Huang³

¹Nanhua University, Chiayi, Taiwan 622, R. O. C.

²National Taiwan Ocean University, Keelung, Taiwan 202, R. O. C.

³Industrial Technology Research Institute, Hsinchu, Taiwan 310, R. O. C.

ycliaw@mail.nhu.edu.tw, g4141001@mail2.nhu.edu.tw, zclai@mail.ntou.edu.tw, PatrickHuang@itri.org.tw

ABSTRACT

In this paper, we present a novel video object segmentation approach. The proposed approach extracts objects from a frame in a video stream using the difference information between the mean-removed versions of the current and referenced frames. Due to the mean-removed version of a frame reduces the influence of light variation on the frame and reserves the texture information of the frame, the proposed approach can effectively segment objects for video sequences and remove shadow pixels. Experimental results show that the proposed approach has the least computation time among object segmentation approaches with shadow removal capability. Compared with the available approaches, our approach reduces the computation time by 7% to 58% with better segmentation accuracy.

KEY WORDS

Object segmentation, mean removal, and fast algorithm

1. Introduction

Object segmentation is a key technology in the field of video processing. The major purpose of video object segmentation is to extract the region of interest (ROI) from video sequences. The extracted ROI can then be used in various video processing applications, such as video coding [1] and video understanding [2-3]. There are two major methods for object segmentation: optical flow segmentation [4-5] and background suppression segmentation (BSS) [6-9]. For real-time applications [2-3], the BSS approach is usually used due to it requires less computing time. The BSS method uses the gray level difference between a pixel in the current frame and the corresponding pixel in the referenced frame to categorize the pixel into an object or a background pixel. In the case that the difference is greater than a given threshold, the corresponding pixel is classified as an object pixel.

Since the BSS method uses the differences of pixels as the criterion for segmentation, the segmentation efficiency is easily influenced by the variation of lightness on background region. To solve this problem, many modified BSS methods [6-9] are proposed. Stauder et al. [6]

analyzed the behavior of shadow and classified shadow into two types: uniform shadow and penumbra. Each object pixel is checked to determine whether the pixel is a shadow pixel. For the case that an object pixel is determined as a shadow pixel, this pixel is removed from the object mask. Chien et al. [7] proposed a simple background registration method to generate a background frame using a number of consecutive frames. The background frame instead of the previous frame is used as a referenced frame to achieve a better segmentation. The current and referenced frames are also replaced by their edge intensities, respectively, to remove the shadow pixels. In reference [8], the normalized cross-correlation (NCC) values of object pixels are used to find out the candidates of shadow pixels. Through further analysis on these candidates, the shadow pixels can be detected and removed. In reference [9], an effective shadow region growing algorithm is proposed to remove the possible shadow areas.

In most cases, the existing object segmentation algorithms [6-9] can effectively segment objects and remove shadow pixels. However, these algorithms usually take a lot of computing time. To speed up the process of object segmentation and removing shadow pixels, a novel object segmentation algorithm is proposed in this paper.

The rest of the paper is organized as follows. The background suppression segmentation algorithm is introduced briefly in Section II. The proposed object segmentation algorithm is described in Section III. The experimental results are presented in Section IV. Finally, the concluding remarks are given in Section V.

2. Background Suppression Segmentation

Figure 1 shows the procedure for segmenting objects from video sequences using the background suppression segmentation (BSS) method. There are three major processes in the BSS method. The frame difference calculation process evaluates the difference between the current and referenced frames. A threshold is then used to determine whether a pixel is the object or background pixel in the thresholding process. If the gray level

difference of a pixel is greater than the threshold value, the corresponding pixel in the current frame is determined as an object pixel. Otherwise, it is a background pixel. In the thresholding process, the change detection mask (CDM), which records 1s for object pixels and 0s for background pixels, of the current frame will be generated. Finally, the post-processing process is applied to remove unwanted small regions and noisy pixels and generates the object mask for the current frame.

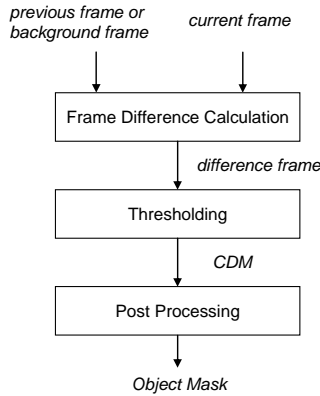


Figure 1. The procedure of BSS method

Figure 2 uses the video stream “Hall” as an example to illustrate the current frame, referenced frame, CDM, and object mask, where the referenced frame is the background frame of the video sequence.

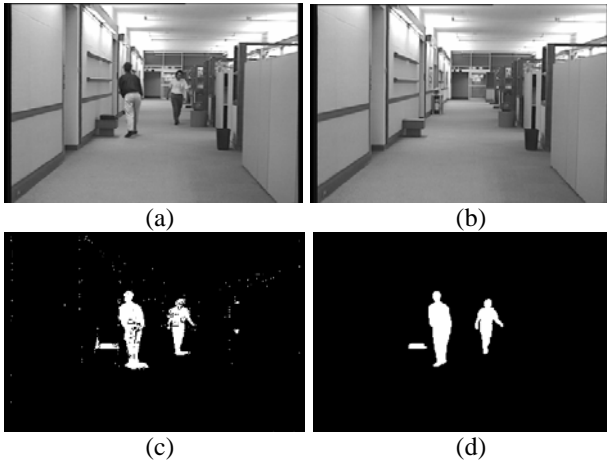


Figure 2. An example of the (a) current frame, (b) referenced frame, (c) CDM, and (d) object mask.

The BSS method as described above is easily influenced by the variation of lighting condition on the background region of current frame. That is, the light variation on the background region may result in poor segmentation using BSS. Once the light condition is changed, the gray level of background region will be changed accordingly. This will misclassify a background pixel as an object pixel. To solve this problem, many modified BSS methods are proposed [6-9]. In this paper, a new fast approach of object segmentation is presented in Section III.

3. Proposed Object Segmentation Method

In this section, we will present our proposed object segmentation approach. The proposed method extracts objects for video streams using the difference information between the mean-removed current frame and background frame. Since the mean-removed version of a frame reduces the influence of light variation on background region and reserves the texture information of the frame, we can extract objects more accurately from video streams and remove shadow area from object regions using mean-removed frames instead of using the original frames. The major procedure of the proposed algorithm is depicted in Fig. 3. A more detail explanation is given in the following subsections.

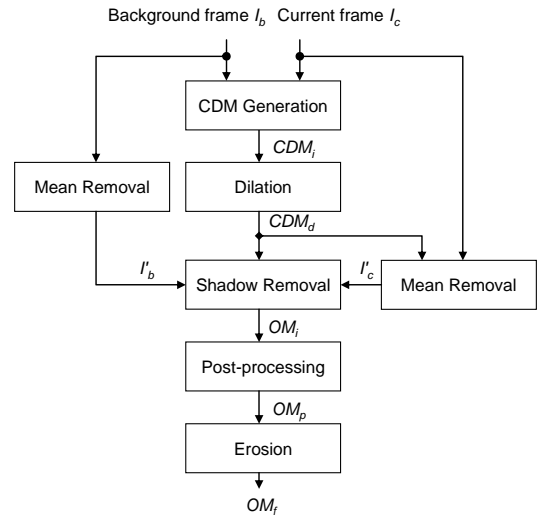


Figure 3. Procedure of object segmentation using mean-removed frames

3.1 CDM Generation

In the process of generating CDM (change detection mask), the gray level difference between a pixel in the current frame I_c and its corresponding pixel in the background frame I_b is calculated and is used to determine whether the pixel is an object or a background pixel. If the difference is less than a given threshold $TH1$, the pixel is classified as a background pixel and the CDM of this pixel is set to 0. Otherwise, it is of an object pixel and the corresponding CDM is set to 1. Finally, the initial change detection mask CDM_i for the current frame is generated.

3.2 Mean Removal

This process is to remove the mean value of each pixel in the input frame. Let $I(x,y)$ be the gray level of the pixel at position (x,y) in frame I , the mean value of gray level for the pixel at position (x,y) is defined as

$$M(x, y) = \frac{\sum_{j=-d}^d \sum_{i=-d}^d I(x+i, y+j)}{(2d+1)(2d+1)} \quad (1a)$$

where d is determined experimentally. In practice, the value of d should be set to 1 or 2. To reduce the computational complexity, the mean value of $I(x, y)$ is calculated using the following equation.

$$M(x, y) = M(x-1, y) + \frac{\sum_{j=-d}^d [I(x+d, y+j) - I(x-d-1, y+j)]}{(2d+1)(2d+1)} \quad (1b)$$

Let $I'(x, y)$ be the mean-removed version of $I(x, y)$. That is

$$I'(x, y) = I(x, y) - M(x, y) \quad (2)$$

In order to reduce the computing time, the above equation is calculated only for a pixel with $CDM = 1$.

3.3 Shadow Removal

The major purpose of this process is to remove the shadow pixels from the object region. To determine if a pixel is a shadow pixel or not, the block difference $D(x, y)$ between a block, in I'_c , of $(2d+1) \times (2d+1)$ pixels with center at position (x, y) and the corresponding block in I'_b is evaluated using equation (3), where I'_c and I'_b are the mean-removed versions of the current frame I_c and background frame I_b , respectively.

$$D(x, y) = \sum_{j=-d}^d \sum_{i=-d}^d [I'_c(x+i, y+j) - I'_b(x+i, y+j)]^2 \quad (3)$$

If the block difference of a pixel is less than a given threshold $TH2$, the pixel is a shadow pixel and should be removed from the object region.

3.4 Dilation and Erosion

In shadow removal process, the block difference of a pixel is used to determine whether the pixel is a shadow pixel. In the intersection area between object and shadow areas, a shadow pixel in object region may be misclassified as an object pixel due to some object pixels are included in the block of shadow pixel. To solve this problem, a morphological dilation operation is performed to extend the boundary area of object region for d pixels before entering the shadow removal process. In the shadow removal process, the extended pixels neighboring the shadow pixels can be removed due to the blocks of extended pixels include only background and shadow pixels. The extended pixels neighboring to object pixels may not be removed in the shadow removal process, since the object pixels are included in the blocks of extended pixels. After the shadow removal process is performed,

some background or shadow pixels around the object pixels may not be removed. Such pixels can be removed easily using a morphological erosion operation in the erosion process. In the dilation and erosion processes, the ball size for dilation and erosion operations is $(2d+1) \times (2d+1)$.

3.5 Post-processing

After the shadow removal process is performed, some unwanted small regions, noise, and holes may be existed. To remove these unwanted regions and pixels, a morphological close operation is performed first to fill the small holes and to connect open regions. Then, the connected component algorithm [10] is applied to remove the small regions. Finally, noises are removed using the morphological open operation. In this process, the ball size for open and close operations is 3×3 .

4. Experimental Results

To evaluate the performance of the proposed approach, two sets of video streams Hall and Frank (as shown in Fig. 4) are used and two segmentation approaches (Chien [7] and Jacques [8]) are implemented. Both video streams with the frame size of 360×240 and 320×240 , respectively, consist of 90 frames. The first frame of each video stream is used as the referenced (background) frame.



Figure 4. Test video streams: (a) Hall, (b) Frank.

In our proposed approach, parameters $(d, TH1, TH2)$ are set to $(2, 25, 625)$ and $(2, 55, 400)$ for Hall and Frank video streams, respectively. The error rate and execution time are used to evaluate the performances of object segmentation approaches. In this paper, the error rate is defined as follows.

$$\text{error rate} = \frac{\text{Error Pixel Count}}{\text{FrameSize}} \quad (4)$$

where the "Error Pixel Count" is the number of misclassified of pixels. That is, the "Error Pixel Count" is the number of misclassified object pixels and misclassified background pixels.

All experiments are performed on a Pentium D 3.2 GHz PC with 512 MB of memory. All programs are implemented as console applications of Microsoft Visual

C++ 6.0 and are executed under Microsoft Windows XP Professional.

Table I shows the execution time of several object segmentation approaches for segmenting Hall and Frank video sequences. The execution time in the unit of millisecond (ms) is the average computing time to segment a video sequence for 10 times.

Table I. Average Execution Time of Several Object Segmentation Approaches

	Hall	Frank
Chien's Approach	21.51 ms	17.63 ms
Jacques's Approach	9.84 ms	15.13 ms
Proposed Approach	9.13 ms	9.27 ms

As shown in Table I, the proposed approach has the least execution time. Compared to Chien's approach, the proposed approach reduces the execution time by 47% to 58%. Comparing with Jacques's approach, the proposed approach reduces the computing time by 7% to 39%.

The error rates of segmenting Hall and Frank video streams using Chien's, Jacques's and the proposed approaches are listed in Table II. From Table II, we may find that compared to Chien's and Jacques's approaches, our approach can extract objects more effectively.

Table II. Error Rates of Several Object Segmentation Approaches

	Hall	Frank
Chien's Approach	0.719%	0.426%
Jacques's Approach	0.610%	0.459%
Proposed Approach	0.544%	0.412%

From these two tables, we may conclude that the proposed approach can effectively reduce the execution time and obtains a lower error rate of segmentation.

5. Conclusions

Object segmentation is a key technology in the field of video processing. To speed up the process of object segmentation for video streams, a novel object segmentation approach is proposed in this paper. The proposed algorithm segments objects from a frame in a video stream using the difference between the mean-removed versions of the current and referenced frames. Due to the mean-removed version of a frame reduces the influence of light variation on the frame and reserves the texture information of the frame, the proposed approach can effectively segment objects and remove shadow pixels. Experimental results show that the proposed approach has the much less computational complexity than available object segmentation approaches with

shadow removal capability. Comparing with Jacques's approach, the proposed approach reduces the execution time by 7% to 39%. Compared to Chien's approach, the proposed approach reduces the computing time by 47% to 58%. Compared to both approaches, our method can also reduce the error rate of segmentation.

Acknowledgements

This work is supported by the Industrial Technology Research Institute and National Science Council (NSC-95-2221-E-343-006). The method presented in this paper is patent pending.

References

- [1] T. Sikora, The MPEG-4 video standard verification model, *IEEE Transactions on Circuits and Systems for Video Technology*, 7(1), 1997, 19-31.
- [2] D. Koller, K. Daniilidis, and H.H. Nagel, Model-based object tracking in monocular image sequences of road traffic scenes, *International Journal of Computer Vision*, 10(3), 1993, 257-281.
- [3] V. Kastinaki, M. Zervakis, and K. Kalaitzakis. A survey of video processing techniques for traffic applications, *Image and Vision Computing*, 21(4), 2003, 359-381.
- [4] A. Giachetti, M. Campani, and V. Torre, The use of optical flow for road navigation, *IEEE Transactions on Robotics and Automation*, 14(1), 1998, 34-48.
- [5] Z. Kalafatic, S. Ribaric and V. Stanislavljevic, Real-Time Object Tracking Based on Optical Flow and Active Rays, in *Proceedings of MELECON 2000*, Cyprus, May 2000.
- [6] J. Stauder, R. Mech, and J. Ostermann, Detection of moving cast shadows for object segmentation, *IEEE Transactions on Multimedia*, 1(1), 1999, 65-76.
- [7] Shao-Yi Chien, Shyh-Yih Ma, and Liang-Gee Chen, Efficient moving object segmentation algorithm using background registration technique, *IEEE Transactions on Circuits and Systems for Video Technology*, 12(7), 2002, 577-586.
- [8] J.C.S. Jacques, C.R. Jung, and S.R. Musse, Background subtraction and shadow detection in grayscale video sequences, in *Proceedings of SIBGRAPI 2005*, Brazil, October 2005.
- [9] D. Xu, J. Liu, X. Li, Z. Liu, and X. Tang, Insignificant shadow detection for video segmentation, *IEEE Transactions on Circuits and Systems for Video Technology*, 15(8), 2005, 1058-1064.
- [10] R. M. Haralick and L. G. Shapiro, *Computer and robot vision: volume I*, (Addison-Wesely, 1992).