

利用視訊壓縮資料進行人體步伐分析

林永欽、鄭智明、韓欽銓

中華大學資訊工程研究所

E-mail : cchan@chu.edu.tw

Abstract

In the paper, human gait classification in compressed domain has been proposed by using the compressed video data in MPEG-1 format. Detection and tracking of moving objects are the primary steps in video surveillance. They are frequently implemented in spatial domain. It needs lots of redundant time to uncompress the video data when the compressed video data are received. The detection and tracking algorithms in compressed domain must be developed to solve this problem. First, moving objects are detected by subtracting the DC values in I frames from those in background. In addition, the DC values of background are also adapted to avoid the noise and illumination change. Second, the tracking process is performed among the consecutive frames. Motion vector information extracted from P frames is used to predict the next position of moving objects. The overlapping table is constructed to determine the relationship among moving objects.

In analyzing the behavior of human, the motion vectors and the velocities of human extracted from P and B frames are utilized to be the feature vectors for gait classification. These features are inputted to train the hidden Markov models of gait classifiers. Four kinds of gait behaviors, including hopping, limping, running, and walking, are classified from the video sequences. The experimental results are given to demonstrate the effectiveness of our proposed method.

摘要

本篇論文提出一個利用 MPEG 壓縮資料進行人體步伐分析。在偵測方面，針對 I 畫面資料，使用 DC based 改良式適應性背景相減法，可以容忍部分監控畫面的雜訊影響並能適應背景場景光線緩慢的變化。在移動物件追蹤方面，利用 P 畫面的運動向量資訊，對已偵測的移動物件做追蹤處理，並建立前後兩張畫面中移動物件之間的歷史重疊關係圖表，了解移動物件之間的關係，以便在重疊分開後，可繼續追蹤正確的移動物件。在人體步伐分析方面，抽取 P、B 畫面裡人體的運動向量特徵，並利用人體頭部及中心軌跡資訊，求出其頭部及中心速度特徵，這些特徵藉由隱藏馬可夫模型訓練，以辨識出視訊資料中人體步伐。

關鍵字：隱藏馬可夫模型(HMM)、陰影偵測、DC based 改良式背景相減法

一、序論

傳統視訊監控系統大部份以閉路電視系統為主，只能以肉眼進行監控，若有必要錄影存證，更需要大量的人力、時間、空間來進行監視。隨著數位化影音技術的進步，視訊資料可以進行壓縮及處理(例如：JPEG, MPEG 1/2, ...等)，再加上網路科技的快速成長，可以將壓縮處理完的視訊資料透過 LAN/WAN 傳輸，經解碼後傳至個人電腦顯示器即可接收到監控影像。因此本論文希望利用 MPEG 以有的硬體技術，提供一套高效能的視訊監控機制，可以具有即時(real-time)、健全(roust)、精確(accurate)、自動化(automatic)、網路化(networked)的特點，主要含有移動物體的偵測、追蹤及行為分析。

在空間領域上應用之監控系統上，Freer[3]等人研究一套使用單色攝影機來偵測及追蹤侵入者(intruder)的自動監控系統，侵入者的偵測是基於背景相減(background subtraction)法進行處理，利用目標的中心點(barycentre)、面積(pixel area)、外形係數(shape factor)去計算相似係數(similarity factor)，藉此判斷是否屬於同一個侵入者。Haritaoglu[4]等人提出一套應用於室外環境的即時監控系統 W^4 (who? when? where? what?)，還運用 Cardboard 模組來定位出人類身體的頭、手、軀幹、腳的位置，藉此達到追蹤與分析人的行為。Collins[2]等人，研發一套自動化的視訊監控系統，搭配經過校正後的多部攝影機來監控寬廣的環境，利用背景相減法來偵測與追蹤移動物件，此外還運用三層的倒傳遞網路 (Back-Propagation Network, BPN) 將偵測到的人和車輛做分類，利用線性判別分析法(linear discriminant Analysis)分析車輛的種類。Black[1]等人建立一套利用多台攝影機的監控系統，運用多台攝影機取得各個視角畫面，建立 3D 地理環境的優點，直接利用較簡單的線性卡曼濾波器(Kalman filter)來追蹤移動物件，還可解決移動物件交會分離(object occlusion)後的身分辨識問題。Wren[14]等人提出一套名稱叫做“Pfinder”的即時追蹤系統，『人』是主要的追蹤對象。此監控系統運用顏色(color)和外型(shape)的多重統計模組(multi-class statistical model)追蹤人的頭和手的位置，並建立人體模型進行分析人的行為。Oren[8]等人運用小波樣板(wavelet templates)的技術，利用其對於顏色與紋理(texture)的不變量性質，可以在複雜環境中精確地偵測出行人，此系統可以透過已存在的樣本(example)，自動地學習意欲辨識之物件，而不用依賴任何事先人

工製作的模型和移動資訊。

在壓縮領域上應用之監控系統, Yoneyama[15] 等人利用 MPEG 壓縮資料進行偵測與追蹤移動物件。在 P- and B- pictures 的偵測方法是利用其 MBs(Macro Blocks, MB 是個 16 x 16 的區塊)的運動資訊(motion information)和 DCT (Discrete Cosine Transformation)係數進行處理, 而 I-pictures 則是利用之前已解碼過的 P-pictures 之移動物件區域大小來判斷目前移動物件的位置, 而追蹤方法則利用目前移動物件的大小(size)、中心點位置(gravity position)、平均運動向量(average motion vector), 藉此判斷與前一個移動物件是否屬於同一個移動物件。Wang[11] 等人提出一套利用 MPEG 壓縮資料的臉部偵測系統。此系統演算法使用 MB 裡屬於色差(chrominance)之 8 x 8 區塊(Cb 與 Cr)的 DC 值、臉部的形狀(shape)和能量分佈(energy distribution)進行臉部偵測。Wang [12]則繼續延伸他們的系統, 利用臉部的運動向量和卡門濾波器達到臉部追蹤的目標。Ozer[9] 等人提出一套利用壓縮資料之人類偵測(human detection)的階層式系統, 在壓縮領域中有兩種偵測方式, 一個是利用 MPEG 視訊檔裡相鄰 MBs 之間的速度來判定可能的人類區域, 另一個是在 JPEG 圖檔裡抽取 AC 值特徵並使用主軸進行人類偵測, 抽取其運動向量特徵, 進行分析人類行為; 在空間領域中則利用人類輪廓(silhouette)之圖形比對(graph matching)演算法, 進行分析人類身體模組, 藉此判斷人類姿勢。Zen[16] 等人提出一套應用於高速公路的監控系統, 利用 MPEG 壓縮資料偵測與追蹤車輛, 採用 MBs 運動向量的大小(magnitude)、方向(direction)、角度(angle)等特徵, 把符合相似條件之 MBs 所合併的區域作為車輛的大小、位置, 利用目前車輛的平均運動向量(average motion vector), 判斷與前一個車輛是否屬於同一車輛, 使用車輛區域區塊 DCT 係數經過主軸成分分析投影, 處理交會分離後車輛的辨識問題。

本篇直接利用 MPEG 壓縮資料[5]進行處理之監控系統如圖一所示, 其優點是可以避免將壓縮資料完全還原, 處理的資料量也相對減少, 可即時偵測出感興趣的移動物件, 並能夠持續追蹤與分析其行為模式。此監控系統的前置作業為移動物件偵測模組, 運用背景相減法對監控畫面區分為前景(即移動物件)和背景, 然後針對包含有移動物件的前景做後續的處理(例如: 濾除雜訊、前景修補等)及背景模組的更新(background updating)。而追蹤模組為偵測模組的後置作業, 當偵測到 I 畫面裡感興趣的移動物件後, 接著可以利用 MPEG 壓縮資料中 P 畫面的運動向量, 直接推測下一個畫面裡移動物件的位置, 藉此可持續地取得移動物件的資訊, 例如移動軌跡(trajjectory)及運動向量、移動物件之間的合併(merge)、分裂(separate)或交錯(occluded)等關係。可利用此些特徵, 經由向量量化(Vector Quantization, VQ)轉換成隱藏馬可夫

模型(hidden Markov model, HMM)所需的可觀察輸出序列(observations), 最後利用隱藏馬可夫模型進行樣本的訓練與比對, 以達到分析移動物件之行為分析的最終目標。

本論文內容架構介紹如下: 第一節簡述研究動機與相關研究以及整個監控系統模組流程與目的。第二節則描述利用 MPEG 壓縮視訊資料, 建立背景模組及其更新方法, 並探討移動物件的偵測、追蹤等問題。第三節將介紹利用 MPEG 壓縮視訊資料中運動向量及軌跡資訊, 抽取移動物件的行為特徵, 經由向量量化轉換成隱藏馬可夫模型所需的可觀察輸出序列, 最後由隱藏馬可夫模型進行樣本的訓練與比對。第四節內容為各項實驗數據及實驗結果的討論。最後一節則為結論與未來工作。

二、移動物件的偵測與追蹤

設計電腦視覺為基礎之監控系統時, 首要任務即偵測視訊畫面中之移動物件, 偵測移動物件的方法, 大致可分為下列三種做法, 分別如下:

(1) 背景相減法(background subtraction)

此方法是事前拍攝背景畫面資料, 然後將攝影機拍到之現在畫面減去背景畫面, 所得到的差異部分, 極可能是系統感興趣之移動物件。此方法容易受到環境雜訊影響, 且常因環境光線變化的因素, 造成整張畫面都有差異而無法正確地分離出前景。

(2) 時間差異法(temporal differencing)

此方法又稱為 frame differencing, 與背景相減法類似, 但無需參考事前拍攝的背景畫面, 直接利用連續或相隔數目的畫面, 直接作差異分析(difference analysis), 藉此判斷畫面中移動物件的位置, 缺點則是通常無法抽取出移動物件的完整特徵資料。

(3) 光流偵測法(optical flow)

此方法是透過較複雜的運算求出連續畫面之間的光流量變化, 以偵測出移動物件, 當取像設備也跟隨之移動時, 偵測效果也依然不錯, 然而這些統計光流量的方法通常需要龐大的計算量, 因此在沒有特殊硬體的支援下, 不適合即時監控系統。

2.1 移動物件偵測

本論文發展的監控系統是直接利用 MPEG 壓縮資料偵測移動物件, 不需將壓縮資料還原後再處理, 因此運用 DC based 改良式適應性背景相減法, 此方法如圖二所示。此方法是利用 I 畫面 8 x 8 Y、Cb、Cr DCT 區塊中的 DC 值直接相減。流程步驟可分為背景的建立、移動目標的追蹤、背景的更新等三部份。

首先建立背景模型, 在圖二中, 取得 MPEG 視訊資料中前幾張包含背景的 I 畫面, 擷取每一個 8 x 8 Y、Cb、Cr DCT 區塊的 DC 值, 形成一段背景畫面序列, 統計此段序列建立背景模組資料。但

考慮色差結構(chrominance scheme)為4:2:0，Cb、Cr的畫面大小僅有Y畫面的四分之一，因此必須將Cb、Cr的畫面放大(up-sampling)，與Y的畫面一樣大小。其次計算背景模組每一個8x8Y、8Y、8Y、Cb、Cr DCT區塊的DC值之平均值(mean, μ)和DC值之標準差(standard deviation, σ)，並表示成向量形式(維度是3)，如方程式(1)、(2)所示。

$$u(i, j) = [u_Y(i, j), u_{Cb}(i, j), u_{Cr}(i, j)] \quad (1)$$

$$\sigma(i, j) = [\sigma_Y(i, j), \sigma_{Cb}(i, j), \sigma_{Cr}(i, j)] \quad (2)$$

$\chi(i, j) = [Y(i, j), Cb(i, j), Cr(i, j)]$ 表示位於目前I畫面第(i, j)個8x8Y、Cb、Cr DCT區塊的DC值， $u(i, j)$ 、 $\sigma(i, j)$ 分別表示位於背景模組中第(i, j)個8x8Y、Cb、Cr DCT區塊的DC值之平均值與標準差。

其次，移動物件的偵測過程，是將後續每一個I畫面和前述建立的背景模組做相減運算(subtraction operator)，然後選擇一個適當的臨界值(threshold)標示出移動物件，定義如下方程式(3)：

$$\begin{cases} |\chi(i, j) - u(i, j)| > t_\beta \times \sigma(i, j) & \chi(i, j) \text{ 屬於前景部分} \\ |\chi(i, j) - u(i, j)| \leq t_\beta \times \sigma(i, j) & \chi(i, j) \text{ 屬於背景部分} \end{cases} \quad (3)$$

其中 $\chi(i, j)$ 、 $u(i, j)$ 、 $\sigma(i, j)$ 分別表示位於目前I畫面第(i, j)個8x8Y、Cb、Cr DCT區塊的DC值、背景模組中第(i, j)個8x8Y、Cb、Cr DCT區塊的DC值之平均值、背景模組中第(i, j)個8x8Y、Cb、Cr DCT區塊的DC值之標準差，而 t_β 表示一個範圍的比例參數(scaling parameter)，取 t_β 倍的標準差當作臨界值。 t_β 一般約設定為3，取得比較好的效果。然而，我們仍會遭遇某些問題，例如：當前景的部分區塊顏色與背景模組顏色相近時，則此部分區塊可能誤判為背景部分，造成前景破碎的情形發生，利用型態學中的閉合及斷開運算分別修補前景部分、去除背景畫面雜訊，其次運用8-連通分量(8-connected component)標示出每一個前景移動物件的大小、位置及其編號(ID)，最後統計每一前景區域的區塊個數，設定一個臨界值，並刪除區塊個數太少的區域。

隨著時間或其它環境因素影響，使得背景畫面的亮度逐漸變亮或變暗，因此需不斷地更新背景模組的參數值 $\mu(i, j)$ 、 $\sigma(i, j)$ ，藉此克服背景畫面緩慢變化的問題，其更新方式如下方程式(4)與方程式(5)所示。

背景模組DC值之平均值計算更新方式：

$$u_i(i, j) = w(i, j) \times \chi_i(i, j) + [1 - w(i, j)] \times u_{i-1}(i, j) \quad (4)$$

背景模組DC值之標準差計算更新方式：

$$\sigma_i^2(i, j) = w(i, j) \times [\chi_i(i, j) - u_i(i, j)]^2 + [1 - w(i, j)] \times \sigma_{i-1}^2(i, j) \quad (5)$$

其中 $\mu_t(i, j)$ 、 $\mu_{t-1}(i, j)$ 分別表示在時間 t 與 $t-1$ 時位於背景模組中第(i, j)個8x8Y、Cb、Cr DCT區塊的DC值之平均值， $\chi_t(i, j)$ 表示在時間 t 時位於目前I畫面第(i, j)個8x8Y、Cb、Cr DCT區塊的DC值， $\sigma_t(i, j)$ 、 $\sigma_{t-1}(i, j)$ 分別表示在時間 t 與 $t-1$ 時位於背景模組中第(i, j)個8x8Y、Cb、Cr DCT區塊的DC值之標準差， $w(i, j)$ 表示混合權重(blending weight)， $w(i, j)$ 值會根據方程式(6)決定。

$$w(i, j) = \begin{cases} \alpha_{bg} & \chi_t(i, j) \text{ 屬於背景部分} \\ \alpha_{fg} & \chi_t(i, j) \text{ 屬於前景部分} \end{cases} \quad (6)$$

在此設定 $\alpha_{bg} = 0.05$ ， $\alpha_{fg} = 0.0125$ 。

2.2 移動物件追蹤

追蹤模組是偵測模組的後續處理，當由偵測模組得到移動物件之後，必須利用追蹤模組對移動物件做持續的監控觀察，藉此可以瞭解移動物件於監控畫面中的大小、位置、軌跡...等資訊，甚至求出移動物件之間的互動關係(例如：交錯、合併、分裂、進入、離開...等)，以達到在連續監控畫面中可以正確地追蹤各個移動物件的目的。

本論文移動物件追蹤之方法，則利用壓縮領域傳輸資料中已具備的運動向量資訊，直接做物件追蹤，無須利用傳統在空間領域(非壓縮領域)最常使用的卡門濾波器[13]來進行預測。如此一來，避免了卡門濾波器的數學計算，可以加快追蹤的速度以達到即時追蹤的目的。在一般常見的MPEG視訊檔中，I畫面與P畫面之間通常只有兩個B畫面(即IBBP格式)且畫面速率是每秒30個畫面，一旦發現I畫面與P畫面的兩個移動物件是相同時，就不需對B畫面做處理，所以當由偵測模組的I畫面中求得移動物件區域之後，根據下一張P畫面中，經過解碼(decoding)後，計算出移動物件區域中(區域大小為m x n區塊)每一個8x8區塊的移動向量，預估此P畫面裡移動物件可能的位置。

三、人體步伐分析與分類

在文獻[4]，提出一個在壓縮領域與空間領域的階層式人類偵測系統。由於本篇論文使用的資料為MPEG壓縮視訊檔，可以利用MPEG視訊資料裡已有的運動向量，抽取出一段連續畫面中，人體步伐的動態特徵。此外，還可利用人體的頭部及中心軌跡(trajecotory)資訊，求出人體頭部及中心速度(velocity)特徵，配合運動向量特徵，可提高人體步伐分析的辨識率。辨識模組，本論文將人體步伐的動態特徵資料套用於隱藏馬可夫模型上判別人體步伐類型，共包含走路、跛行、跑步、跳躍四種步

伐。

步伐的特徵利用 MPEG 視訊資料，抽取 P、B 畫面裡 MBs 已有的運動向量作為特徵。B 畫面的 Inter-Block MBs 有前向預測(即一個根據前一個參考畫面的運動向量)、後向預測(即一個根據後一個參考畫面的運動向量)、雙向預測(同時包含前向與後向預測兩個運動向量)三種型式，不同於 P 畫面的 Inter-Block MBs 只有一種前向預測的型式，所以當遇見 B 畫面裡屬於雙向預測形式的 MBs 時，本篇是取其前向與後向預測的兩個運動向量之平均值當作此一 MB 的運動向量特徵。當人體區域裡每一個 8 x 8 區塊有其運動向量時，將人體區域均分為三個部分(1 x 3 區塊)、九個部分(3 x 3 區塊)，如圖三所示，而區分人體區域裡每一個 8 x 8 區塊所屬的人體部分之方式是依其所涵蓋最大面積的人體部分為原則，最後分別計算各個人體部分的平均運動向量，定義如下方程式(7)所示：

$$\begin{aligned}\bar{u}_i(\Gamma) &= \frac{\sum u_i(x, y, \Gamma)}{\text{人體區域裡 } 8 \times 8 \text{ 區塊數目}} \\ \bar{v}_i(\Gamma) &= \frac{\sum v_i(x, y, \Gamma)}{\text{人體區域裡 } 8 \times 8 \text{ 區塊數目}}\end{aligned}\quad (7)$$

其中 $u_i(x, y, \Gamma)$ 、 $v_i(x, y, \Gamma)$ 分別表示第 Γ 畫面中第 i 個人體部分位於畫面中第 (x, y) 8 x 8 區塊的運動向量之水平與垂直分量， $(\bar{u}_i(\Gamma), \bar{v}_i(\Gamma))$ 是平均運動向量， $i = 1, \dots, 9$ 。

此外，利用追蹤模組所得到的人體頭部及中心軌跡資訊，求出其人體頭部及中心速度特徵，定義如下方程式(8)：

$$\tau_i(\Gamma) = \left(\frac{x_i(\Gamma) - x_i(\Gamma - 1)}{\Delta t}, \frac{y_i(\Gamma) - y_i(\Gamma - 1)}{\Delta t} \right) \quad (8)$$

其中 $x_i(\Gamma)$ 、 $y_i(\Gamma)$ 分別表示人體頭部或中心於第 Γ 畫面軌跡位置的水平與垂直分量， Δt 代表畫面速率(frame rate)， $i =$ 人體頭部或中心。

由於人體在拍攝中個別相距攝影機的距離也會影響到步伐分析(因距攝影機越近，相對移動速度越快)，此時會造成行為分析上的誤判。為了解決此一問題，將上述取得的平均運動向量與人體頭部及中心速度，利用人體區域大小將這些步伐特徵做量化(normalization)，定義如下方程式(9)：

$$\begin{aligned}\bar{u}_i'(\Gamma) &= \frac{\bar{u}_i(\Gamma)}{Z(\Gamma)}, \quad \bar{v}_i'(\Gamma) = \frac{\bar{v}_i(\Gamma)}{Z(\Gamma)} \\ \tau_i'(\Gamma) &= \frac{\tau_i(\Gamma)}{Z(\Gamma)}\end{aligned}\quad (9)$$

其中 $Z(\Gamma)$ 表示在第 Γ 畫面裡人體所佔的 8 x 8 區塊數目。

最後，將上述收集的人體步伐特徵排列為一序列特徵向量，利用向量量化的方式進行分類，轉換成隱藏馬可夫模型所需的可觀察輸出序列。向量量化的過程是將所有 k 維步伐特徵向量分成 ω (這裡設定 $\omega = 4$) 個類別(cluster)，每一類別的中心點

c^i (centroid) 是一個 k 維向量，並會給每一個類別一個量化後的值，稱之為書碼 o^i (codeword 或 symbol)。向量量化的方法首先給定類別個數 ω (即書碼個數)，其次亂數設定每一個類別的中心點，最後藉由 k -mean 演算法不斷地更新其中心點直到收斂為止。

本篇論文基於隱藏馬可夫模型[7][10]具有辨識可變長度樣本的能力，適合發展在壓縮領域中，人體步伐分類的方法。針對步伐分類所設計的隱藏馬可夫模型之敘述如下：

前置處理，向量量化的方法將所有經由步伐特徵抽取步驟得到的特徵向量分成四類，分別用 $\omega_1, \omega_2, \omega_3, \omega_4$ 四個書碼表示之，即意味著人體步伐的各種姿勢，而不同的姿勢代表模型的各個輸出狀態，姿勢的變化也就代表著輸出狀態的改變。其次，利用 Baum-Welch 演算法建立各個人體步伐(走路、跛行、跑步、跳躍)的隱藏馬可夫模型。例如實驗中搜集走路步伐之頭部及中心速度的部分樣本，如序列所示 $(\omega_1 \omega_2 \omega_1 \omega_4 \omega_3 \omega_2 \omega_3 \omega_1 \omega_4 \omega_1 \omega_1 \omega_3)$ ，可觀察輸出狀態個數 $M = 4$ ，而隱藏狀態個數 $N = 5$ ，用 0, 1, 2, 3, 4 代表之，初始隱藏狀態轉移矩陣 A 和初始可觀察輸出矩陣 B 各項值皆為 0.25。

最後，利用維特比演算法，計算出測試樣本之可觀察輸出序列在各個人體步伐隱藏馬可夫模型的機率值，將它歸類為具有最大機率值的步伐隱藏馬可夫模型。

四、實驗結果

本論文實驗結果分為兩部分：移動目標物的偵測追蹤與人體步伐分析。其中偵測、追蹤實驗所使用的視訊資料，是利用 SONY DCR-TRV 120 型的 DV 攝影機於戶外實地拍攝，影像畫面以 MPEG1 的格式壓縮，其中色差結構為 4:2:0，編碼格式為 IBBP，每張畫面大小為 352 x 240 個像素點，畫面速率每秒 30 張畫面；而人體步伐分析實驗，則採用參考文獻[6]拍攝的 100 個 MPEG1 視訊檔與本論文實驗的 100 個 MPEG1 視訊檔作為訓練及測試樣本。

4.1 移動目標物之偵測與追蹤

在移動目標物的實驗中，選取兩個 MPEG 視訊檔，說明本篇論文發展偵測與追蹤方法成效，圖五是一段連續 MPEG 視訊檔的偵測與追蹤結果(第 127~145 畫面)。在圖四中， I_{127} 畫面為使用 DC based 改良式適應性背景相減法所偵測到的移動物件及其 DC 差值二元化影像，一旦偵測到 I_{127} 畫面裡的移動物件，便交由追蹤模組進行追蹤，利用 P_{130} 的運動向量所預測之移動物件的範圍大小(黑色框)，求此範圍內的每一個 8 x 8 DCT 區塊的 DC 值，再利用 DC based 改良式適應性背景相減法，得到較精確的移動物件之範圍大小及其 DC 差值

二元化影像。P₁₃₃ ~ P₁₄₂與 P₁₃₀作法相同，而畫面 I₁₄₅ 為另一個偵測模組的開始，同樣的使用 DC based 改良式適應性背景相減法，可順利地偵測出移動物件及其 DC 差值二元化影像。

同理，圖五為針對多目標物偵測與追蹤的結果，其中包括了移動目標的進出、交錯。在圖五中，I₁₀₉ 為多移動目標物的偵測，P₁₄₈ 為移動目標物的交錯結果，在 P₂₀₈ 中兩移動目標分開後還能順利辨識出原編號並持續追蹤各目標物，此時又有另一目標物進入畫面，I₂₅₃ 有一移動目標物離開畫面。

4.2 步伐分析與分類

在步伐分析與分類實驗中，本論文採用參考文獻[6]所拍攝的 100 個 MPEG1 視訊檔與自行拍攝的 100 個視訊檔作為訓練及測試樣本。包含十個人的四種步伐，分別有 50 個走路、50 個跛行、50 個跑步、50 個跳躍之 MPEG1 視訊檔。首先分別收集四種步伐的 25 個 MPEG1 視訊檔作為訓練樣本(自行拍攝樣本 12 個，參考文獻[6]之樣本 13 個)，然後抽取特徵再利用向量量化的方法轉換成隱藏馬可夫模型所需的可觀察輸出序列，建立各種人體步伐(走路、跛行、跑步、跳躍)的隱藏馬可夫模型。最後，將剩餘的 100 個 MPEG1 視訊檔作為測試樣本，分別利用五組實驗測試其辨識率。

實驗一：利用軌跡資訊求得人體頭部及中心速度作為人體步伐特徵。

實驗二：利用運動向量作為人體步伐特徵，人體區域均分為三個部分(1x3 區塊)。

實驗三：利用運動向量作為人體步伐特徵，人體區域均分為九個部分(3x3 區塊)。

實驗四：利用運動向量與軌跡資訊求得人體頭部及中心速度作為人體步伐特徵，人體區域分為三個部分(1x3 區塊)。

實驗五：利用運動向量與軌跡資訊求得人體頭部及中心速度作為人體步伐特徵，人體區域均分為九個部分(3x3 區塊)，其各辨識結果如表一所示。

表一 各實驗人體步伐辨識結果

實驗編號	1	2	3	4	5
結果					
辨識率	63%	49%	58%	65%	71%

由以上五個實驗結果得知，當利用人體的頭部及中心軌跡的資訊，求出其頭部及中心速度的特徵，並配合運動向量特徵，以隱藏馬可夫模型分類四種不同人體步伐，其辨識率約可以提高 15% 左右。

五、結論與未來工作

在本論文中，利用 MPEG 壓縮資料進行人體步伐分析，對監控畫面裡的移動物件進行偵測、追蹤並分析其行為。提昇監控的電腦化，節省大量人力資源的浪費，並使其達到即時、健全、精確、網

路化與自動化的特點。

由於實驗取景的場合是控制在光線非常均勻且背景較單調的情況之下，所以在光線變化較大且分佈不均(例如拍攝場合裡有多個不同光源)或背景很複雜(例如樹葉晃動、旗幟飄動)時，則可能需要其它的 AC 係數或建立更多的背景模組資訊進行處理。此外，受限於取像環境是利用一台固定位置的攝影機進行實驗，當攝影機的震動或背景突然更換時，會使得整個監控系統陷於癱瘓，則未來可能會使用多部可移動式的攝影機來解決。

人體步伐分析方面，在實驗中的人體是以平行於攝影機的角度移動，未來可嘗試其它的移動角度進行分析。再者，收集更多的樣本數或建立人體模型以抽取其它更細部的人體關節(例如手肘、膝蓋)之動態特徵，藉此提高人體步伐分析的結果。此外，利用人體步伐分析的資訊，尚可進一步達到智慧型監控系統個人身分辨識功能的最終目標。

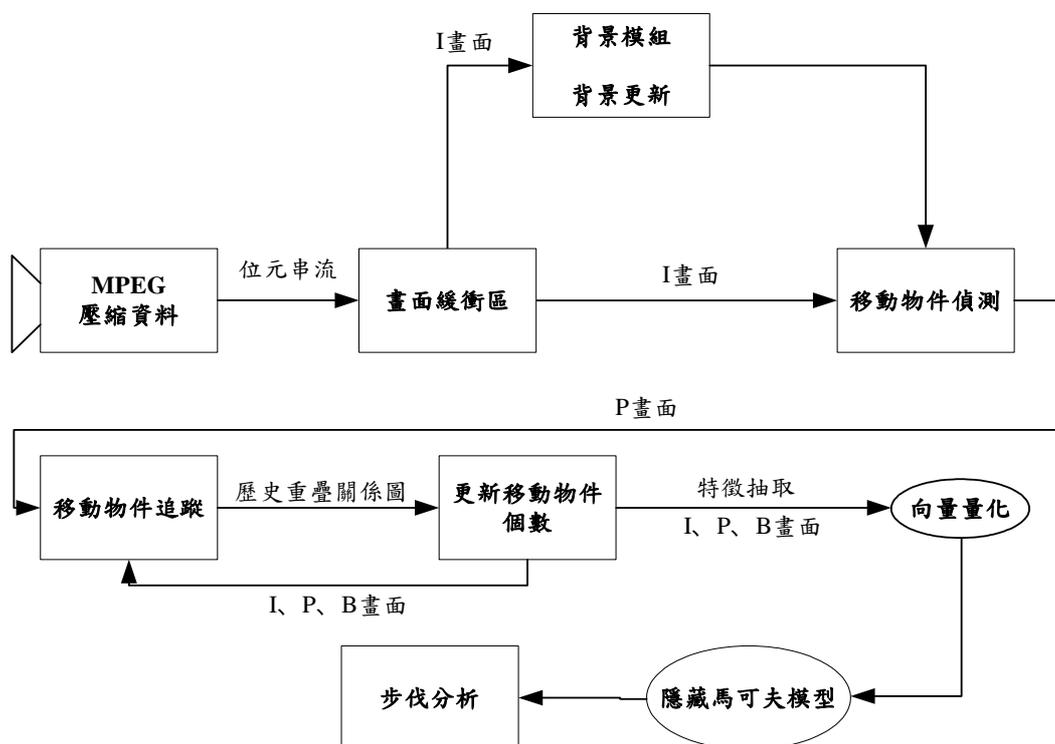
致謝

This work is supported by National Science Council of Taiwan under grant no. NSC91 -2213 -E008 -020.

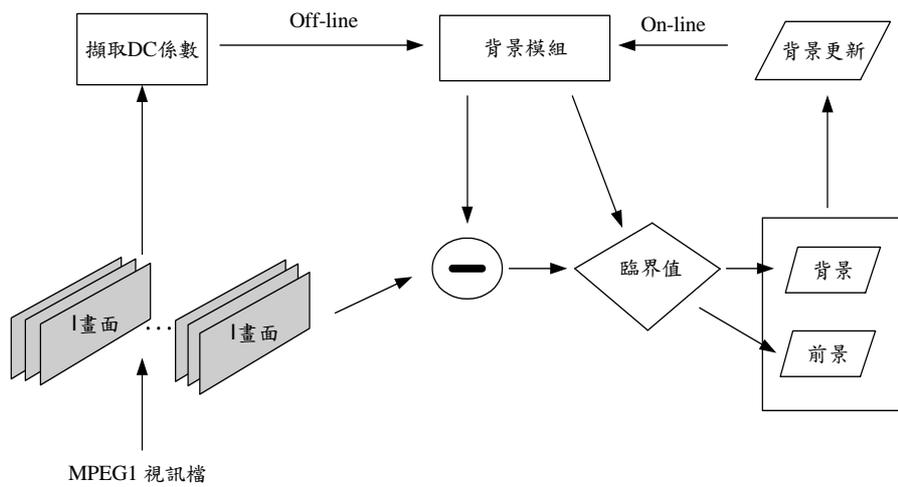
參考文獻

- [1] Black and T. Ellis, "Multi camera image tracking," *Proc. 2nd IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS)*, December 2001.
- [2] R. T. Collins, A. J. Lipton, T. Kanade, H. Fujiyoshi, D. Duggins, Y. Tsin, D. Tolliver, N. Enomoto, O. Hasegawa, P. Burt, and L. Wixson, "A system for video surveillance and monitoring: VSAM Final Report," Technical Report CMU-RI-TR-00-12, Robotics Institute, Carnegie Mellon University, May 2000.
- [3] J. A. Freer, B. J. Beggs, H. L. Fernandez-Canque, F. Chevrier, and A. Goryashko, "Automatic video surveillance with intelligent scene monitoring and intruder detection," *Proc. 30th Annual 1996 International Carnahan Conference*, pp. 89-94, October 1996.
- [4] I. Haritaoglu, D. Harwood, and L. S. Davis, "W4 : Who? When? Where? What? A real-time system for detecting and tracking people," *Proc. of 1998 Automatic Face and Gesture Recognition*, pp. 222-227, April 1998.
- [5] D. Le Gall, "The MPEG video compressed algorithm: A review," *Proc. of 1991 International Society for Optical Engineering (SPIE)*, vol.1452, pp. 444-457, 1991.
- [6] C. H. Liu, C. C. Han, and K. C. Fan, "Human gait classification using video information," *Proc. 15th IPPR Conference on Computer Vision, Graphics and Image Processing*, June 2002.

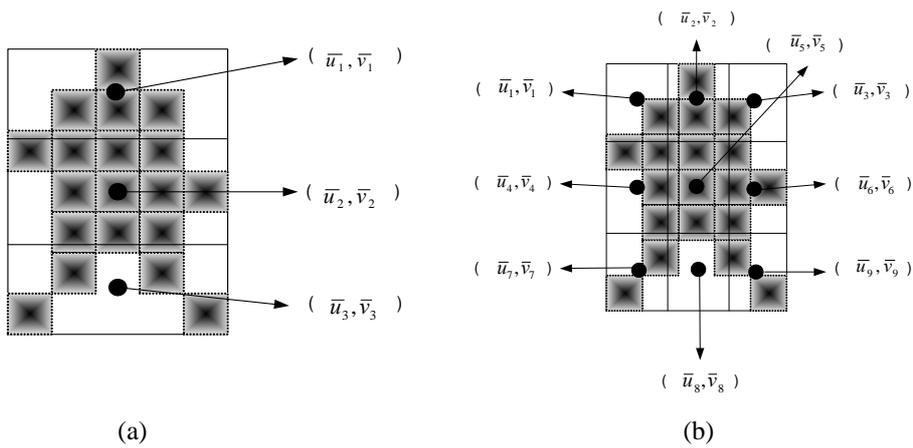
- [7] C. Manning and H. Schutze, "Foundations of statistical natural language processing," *The MIT Press Cambridge*, May 1999.
- [8] M. Oren, C. Papageorgiou, P. Sinha, E. Osuna, and T. Poggio, "Pedestrian detection using wavelet templates," *Proc. of 1997 Computer Vision and Pattern Recognition (CVPR)*, pp. 193-199, June 1997.
- [9] B. Ozer and W. Wolf, "A hierarchical human detection system in compressed and uncompressed domains," *IEEE Transactions on Multimedia*, vol. 4, no. 2, pp. 283-300, June 2002.
- [10] L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. of the IEEE*, vol. 77, no. 22, pp. 257-286, February 1989.
- [11] H. Wang and S. F. Chang, "A highly efficient system for automatic face region detection in MPEG video," *IEEE Transactions on Circuits and Systems for Video Technology (CSVT)*, vol. 7, no. 4, pp. 615-628, August 1997.
- [12] H. Wang, H. S. Stone, and S. F. Chang, "FaceTrack: Tracking and summarizing faces from compressed video," *Proc. of 1998 International Society for Optical Engineering (SPIE), Multimedia Storage and Archiving Systems IV*, pp.19-22, September 1998.
- [13] G. Welch and G. Bishop, "An introduction to the Kalman filter," *University of North Carolina at Chapel Hill, Department of Computer Science, Chapel Hill, NC, USA*. TR95-041.
- [14] C. R. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: Real-time tracking of the human body," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 780-785, July 1997.
- [15] A. Yoneyama, Y. Nakajima, H. Yanagihara, and M. Sugano, "Moving object detection and identification from MPEG coded data," *Proc. of IEEE 1999 International Conference on Image Processing (ICIP)*, vol. 2, pp. 934-938, October 1999.
- [16] H. Zen, T. Hasegawa, and S. Ozawa, "Moving object detection from MPEG coded picture," *Proc. of IEEE 1999 International Conference on Image Processing (ICIP)*, vol. 4, pp. 25-29, October 1999.



圖一 視訊監控系統架構圖。

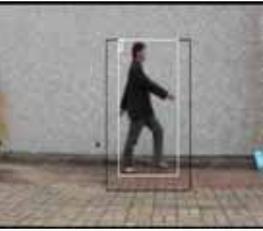
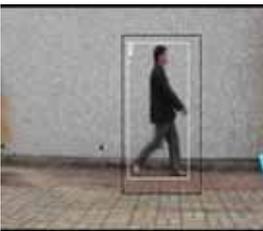
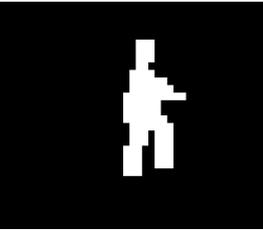
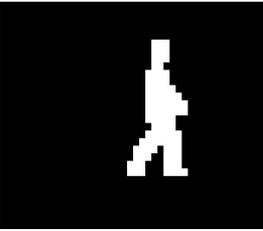
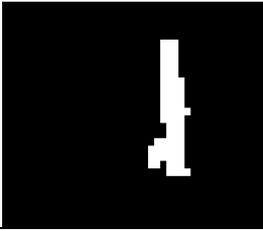


圖二 在壓縮領域裡移動物件偵測之流程圖。

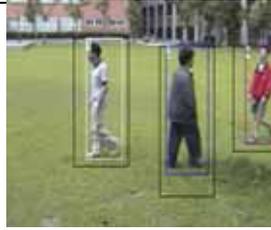
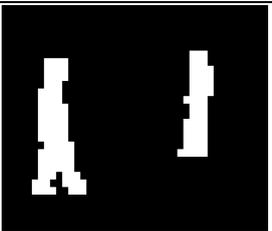
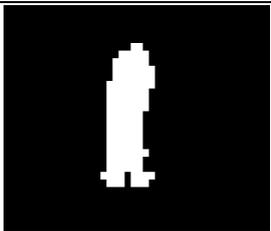
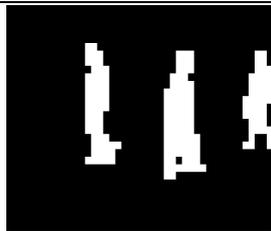
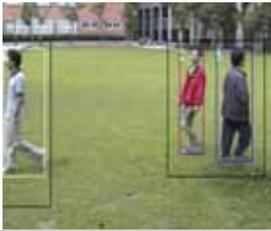
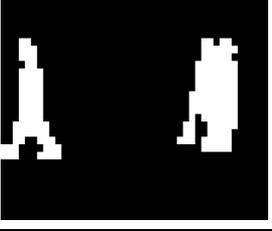
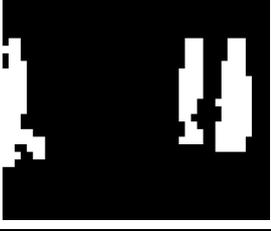
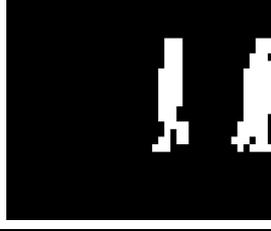


圖三 人體區域均分為 (a) 1 x 3 區塊 (b) 3 x 3 區塊。

畫面	I ₁₂₇	P ₁₃₀	P ₁₃₃
偵測 前景 移動 物件 畫面			
DC 差 值二 元化 影像			

畫面	P ₁₃₆	P ₁₄₂	I ₁₄₅
偵測 前景 移動 物件 畫面			
DC 差 值二 元化 影像			

圖四 單一目標物的偵測與追蹤。

畫面	I ₁₀₉	P ₁₄₈	P ₂₀₈
偵測 前景 移動 物件 畫面			
DC 差 值二 元化 影像			
畫面	I ₂₃₅	P ₂₄₁	I ₂₅₃
偵測 前景 移動 物件 畫面			
DC 差 值二 元化 影像			

圖五 多目標物的 MPEG 視訊檔的偵測與追蹤結果。