

嵌入式智慧型視訊分析系統之實作

廖文淵¹、歐陽起鵬²

¹ 德霖技術學院資訊工程系助理教授

² 德霖技術學院資訊工程系講師

中文摘要

近幾年來，數位影像分析技術用於智慧型監視系統的方法，已經發展出來，並具有良好的效能。除了相關技術的突飛猛進，再加上大家對於安全與犯罪預防的問題日益重視，使得智慧型監視系統的應用已廣泛出現在我們的日常生活中。本論文提出以數位訊號處理晶片為基礎，將智慧型視訊監視系統實作於嵌入式系統上，與一般將所有影像視訊回傳到視訊主機再作分析處理不同的是，利用數位訊號處理器的計高速計算特性，在遠端便能即時分析與異常事件偵測，以提高傳輸效率及異常事件判斷之正確性。本系統發展原理是利用嵌入式數位訊號處理器的計高速計算特性與操作性，在遠端即時進行異常事件的偵測，以提高傳輸效率及異常事件判斷之正確性，將視訊影像資料與事件判斷結果，以網路傳送至監控中心，在遠端處並同時進行必要之處置，如發出聲響或燈號，與遠端攝影機之參數調整，使智慧型監控達到更符合人性並提高功能性。

關鍵字：智慧型監視系統、數位訊號處理晶片、嵌入式系統、異常事件偵測。

Implementation of the Embedded Video Analysis System for Intelligent Surveillance

Wen-Yuan Liao¹

C.-P.Ouyang²

¹Assistant Professor, Department of Computer Science and Engineering De-Lin Institute of Technology

²Lecturer, Department of Computer Science and Engineering De-Lin Institute of Technology

Abstract

Intelligent video surveillance system (IVS) is the most popular topics for researches and applications. The development of technology and growth of market for IVS is rapid in recent years. Furthermore, IVS is made use of in many applications, including traffics, crime events detect, and so on. In this paper, we developed an embedded IVS system. The system mainly built up from remote system, which is composed of a dual core DSP system and video camera, and server system established from a PC-based monitoring system. The system is performed by applying principles of the digital image processing, video analysis, events detection and computer network. The properties of fast computation by using DSP and convenient human interface design for embedded system are the advanced improvements in our system. The large amount of analysis and processing on remote area, actually reduce the bandwidth requirement of network and the load of the server system effectively. Moreover, the more robust function and appliance are set in the remote system, including alarm signals and sounds. Finally, some experimental results are presented to show the performance of the proposed system.

Key word: Intelligent video surveillance system (IVS), embedded IVS system, dual core DSP system, events detection.

前言

隨著時代的進步，人們對安全相關的議題益加重視，使得現今攝影監視系統的應用如雨後春筍不斷開發出來，其所用的技術更是蓬勃發展，可說是已經到了隨時隨地皆可發現其蹤影。常見的用途包括：居家安全照護、一般保全、交通流量監控、公共場所安全監控等應用[1,3,5,12]。數位監視系統中的影像已成為各國警政單位的重要破案工具。例如汐止殺警案、台中毒蠻牛案及雲林縣立肉品市場運鈔車搶案。因此，重要物件之影像的清晰與否，成為能否破案的重要關鍵。一般的視訊監控系統存在著管理不易、施工花費過高、錄製的異常事件影像儲存位置不具彈性，此外尚需利用人力來監看畫面，浪費人力且易造成關鍵畫面因為人對畫面的麻痺疲勞而造成畫面遺失的風險等問題。此外，傳統的視訊監控皆將拍到的視訊傳送至主機後，就將視訊資料予以紀錄與儲存，若遇到異常狀況發生，就只能從歷史檔案或資料庫中尋找，如此大大降低了即時監控的意義。

上述一般網路式遠端監視技術[7,9,18]，其架構如圖 1。網路攝影機將視訊傳回到監控伺服器主機後，進行影像處理與事件分析，並將視訊全部儲存與和加上記錄檢索。在此架構下，網路頻寬與伺服器主機計算能力之限制乃成為此技術之一大瓶頸，其中影像分析伺服器主機除需不斷儲存視訊，若要同時對每架攝影機具有影像分析與事件判斷之功能，在實作層面上確有困難之處。且若再加上欲透過手機等行動通訊網進行遠端查詢、監控，礙於行動通訊網路之頻寬有限，需將視訊作即時的格式轉換，以能符合行動通訊網路進行通訊時，對於頻寬的需求。因此，若加上此一功能，以現有伺服器主機實難實現。

現今 TI 市場上智慧型監視系統技術的突飛猛進，再加上大家對於安全與犯罪預防的問題日益重視，使得智慧型監視系統的應用已廣泛出現在我們的日常生活中。從智慧型監視系統的發展過程來看，可以分成三個歷程[22]，分別為(1)類比影像錄影(VCR, Video Camera Recorder)階段;(2)數位影像錄影(DVR, Digital Video Recorder) 階段及(3)網路視訊錄影(NVR 或 IPVR, Network Video Recorder) 階段。最早的類比影像錄影階段，只是透過類比的記錄方式，將影像保留下來，以作為預防與事後追查的用途，其缺點是類比資料只能以循序的方式搜尋，查閱非常費時，且錄影帶的品質較差，保存不易。此外其硬體設備架設較為複雜，傳輸纜線的重複佈置，更是浪費許多空間與時間上的資源。透過數位化的影像攝影機，記錄方式由類比轉為數位，許多視訊上的儲存、傳輸與處理，便可透過電腦的演算法則加以處理與分析，而不像類比方式需要電子元件來修正。這個階段大部份可處理的方式是以離線方式處理視訊上的雜訊、影像增強等，屬於數位影像處理的範疇。由於影像資料經過數位化的處理，利用資料備份與區域網路的技術，便可達到內部存取相當容易，這可以視第三階段的起步雛型。

現階段是網路視訊錄影應用最多的階段，其興起主要針對 DVR 架構僅適合於單一區域，不易進行整合，必須以獨立的系統存在著[15,16,19]。當監視系統不斷地擴張，新舊系統與不同廠牌的整合與相容性便顯得非常重要，如此只要設立一個中控中心，便可監視到所有鏡頭畫面。因此網路視訊錄影技術就是解決此問題的最好方法。因此，網路視訊技術是非常有彈性與延伸性的架構，數位化與網路化，讓監看者或管理者可以迅速的查詢與分析，但管理層面的問題卻逐步浮現。

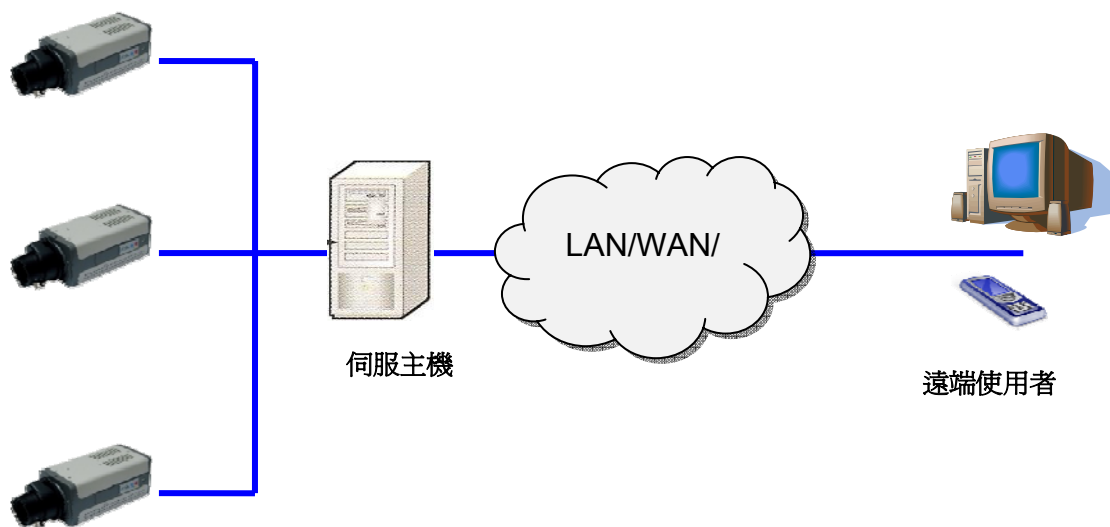


圖 1 傳統網路式遠端監視之架構

首先，隨著技術的進步，影像壓縮與儲存能力越來越強，網路傳輸頻寬或品質越來越好，一個監控中心所連結的攝影機越來越多，在畫面上同時呈現的頻道數也越來越多，使得管理人員經常無法掌握每個子畫面。根據安全監控業者的統計，在警衛緊盯眾多監視螢幕 22 分鐘之後，視覺疲勞將造成警衛對 95% 的活動失去反應[21]。因此如何建立有效的智慧式遠端監視系統，提供完整的異常狀況或事件功能給監控中心，而不是由監控中心人員隨時盯著螢幕，藉以提高安全人員工作效率與事件預警的正確性，這是一個刻不容緩的工作。因此，本論文提出以數位訊號處理晶片為基礎，將智慧型視訊監視系統實作於嵌入式系統上，與一般將所有影像視訊回傳到視訊主機再作分析處理不同的是，利用數位訊號處理器的計高速計算特性，在遠端便能即時分析異常事件的發生，以提高傳輸效率及異常事件判斷之正確性，其架構如圖 2。

壹、 設計原理

本計畫提出的架構流程，大致可將系統分成四大功能方塊，我們使用的嵌入式視訊架構包含下列六個模組，包括(1)背景影像建構模組；(2)移動物體偵測與分割模組；(3)事件追蹤模組；(4)物件定位模組；(5)攝影機參數設定模組；(6)感測與告警 I/O 介面模組。當系統自攝影機接收影像視訊進來，須從背景中找出移動物體，背景是利用背景影像建構模組建立出背景資料庫。移動物體偵測後再傳遞到移動物體偵測與分割模組進行物件的圖像切割。事件追蹤模組為在移動物件偵測出來後決定出事件圖像與定位，最後將結果利用網路傳回至監控伺服主機作儲存與資料庫的處理。感測與告警 I/O 介面模組則進行必要之處置如發出音響或燈號，使用感測元件作為攝影機參數設定模組參數調整之依據。

近幾年來，智慧型視訊監控的發展，已有許多國內外學者提出各種方法與學術論文與技術報告，總結發展的系統流程，大致可將系統分成四大功能方塊：背景影像之建構、移動物體偵測、事件追蹤及物件定位。本系統分析流程圖如圖 3 所示。

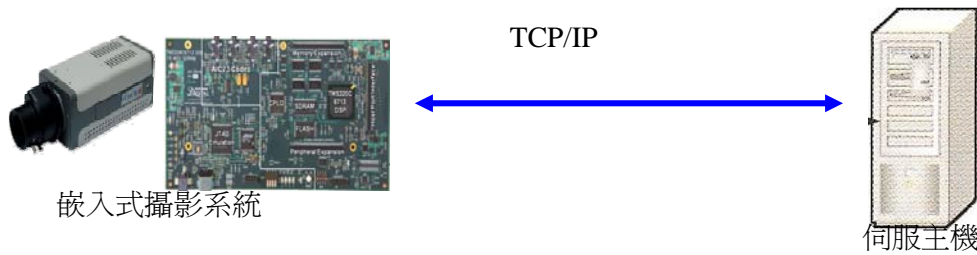


圖 2 系統架構圖

一、背景影像前處理與建構

在建立背景影像前須經過前處理，使獲得之影像適合作進一步處理。影像前處理的方法已有許多研究提出，其中 RGB 色彩模型為電腦視覺應用中最廣泛使用的一種色彩模型，其原理使用紅色、綠色與藍色(RGB)作為三原色，但其在辨識或分開顏色中亮度(luminance)與彩度(chromatic)的成份時，並無法有效分辨而提供明顯特徵供後端作進一步的處理供影像辨識與應用。為了達到有效分辨並擷取亮度、彩度的特徵值，各種不同的色彩模型轉換已發展出[2,4,8]，包括正規 RGB 色彩空間 (Normalized RGB Space)、HSL 色彩空間(Hue、Saturation 及 Luminance)、HIS 色彩空間(Hue、Intensity 及 Saturation)與 HSV 色彩空間(Hue、Saturation 及 Value) 與 YCbCr 色彩空間。

因此，選擇一個適當的色彩空間作為影像前處理就非常重要，對於特定入侵物件而言，要取得一個適切的色彩空間會因入侵物件的不同而有不同的作法，我們首先考慮的物件是以人為主，由於不同膚色的反射光頻譜能量分布對於 RGB 色彩非常敏感，因此我們選擇使用正規 RGB 色彩空間作為影像前處理的根據[10,11,13]。將 RGB 色彩空間轉換到正規 RGB 色彩空間如公式(1)、(2)所示:

$$r = R / (R+G+B), \quad (1)$$

$$g = G / (R+G+B). \quad (2)$$

以上轉換公式可以降低影像中膚色對亮度造成的影響，根據我們實驗所顯示的結果可以獲得其中的 r 值約介於 0.3 到 0.6 之間; g 值則約介於 0.2 到 0.4 之間. 在取得色彩空間中的 r 值與 g 值之後，亮度(luminance) Y 值即由公式(3)計算得出:

$$Y = 0.3R + 0.59G + 0.11B. \quad (3)$$

接下來即可依據 r 、 g 、 Y 值作二值運算，將前景物件自影像中分割出來，以取得背景影像。在建立連續背景影像部份，一般常用的是時間平均法，此方法將一段時間內所有的影像作平均運算，作為建立新的背景影像的依據。所謂背景，是指具有相對於前景或偵測物件較為穩定的像素。背景建置的方式是取出從輸入視訊畫面的 n 個連續畫面(Frame)，作為背景之初始基礎，計算法式如公式(4)：

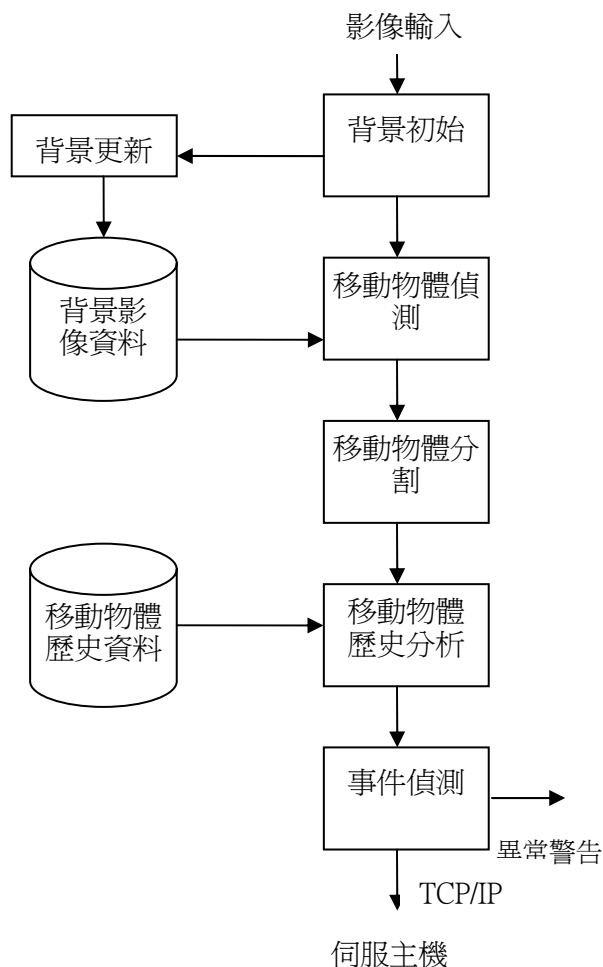


圖 3 系統分析流程圖

$$Bg(x, y) = \frac{1}{n} \sum_{i=1}^n I_i(x, y) \quad (4)$$

其中 $i(x,y)$ 為第 i 張背景影像， n 為背景影像數。因此 $Bg(x,y)$ 即為建構出來的始初背景。

二、 移動物體偵測

物件偵測之目的在於由影像中標示出具有某一特定物件的影像區塊，以取得物件之位置及類別。移動物體偵測之目的則在於給定一張場景視訊影像，計算出畫面中與之前影像具有變異之像素，形成該畫面的移動物件遮罩(Moving object mask)。移動物體偵測亦稱為前景擷取(Foreground Extraction)或背景影去除(Background subtraction)，目的是先用一個時間區段的影片，把初始的背景模型先建好，之後每一張進來的影像，都先跟高斯混合模型(Gaussian mixture model, GMM)做比較看看有沒有符合，若是有的話就更新那個符合的模型。因此，移動物體偵測所關心的是因物體之進入與離開、物體之移動及形變等因素造成之影像變異，而應避免計入其他不重要因素如攝影機之雜訊、光照之改變、背景之週期性運動等因素造成之變異。

接著說明移動物體偵測的方法，國內外學者提出的許多的技術[6]，歸納這些方法，大致使用影像在空間或者空間的資訊，這些方法主要包括 1)背景相減法(Background Subtraction), 2)時間差異法(Temporal differencing)與 3)邊緣偵測。本系統使用的背景相減是先建立背景影像，再將新影像畫面與背景相減，取得完整的移動物件的形狀動態資訊，以利進一步處理複雜背景問題，對雜訊有較好的容忍度。計算法式如公式(5)、(6)所示：

$$D(x, y) = |I(x, y) - I_b(x, y)| \quad (5)$$

及

$$B(x, y) = \begin{cases} 1, & D(x, y) > Threshold \\ 0, & Otherwise \end{cases} \quad (6)$$

其中， $I(x,y)$ 與 $I_b(x,y)$ 分別為目前與背景影像， $B(x,y)$ 為二值影像。

時間差異法是利用時間上連續的連續影像進行像素相減，如果結果為零，表示此為移動物件的像素;反之則表示不屬於移動物件的像素。使用時間差異法，其偵測物件的結果內會有缺口、外型不完整、及對偵測的環境變化相當敏感。因此一般的研究，若先處理環境變化的因素，所得到的移動物件形狀則可能不完整;反之亦然。

三、物件追蹤與定位

本系統中所使用的技術為物件影像切割法，再利用移動物體之移動慣性特性，將前一張影像所取得的物件加上其移動向量，可以計算預估在目前畫面上會出現的位置，進而依據各預估向量尋找出與原質心最小距離的物體，若其實際位置與預估的位置十分接近，即做為物件追蹤的下一位置。

貳、系統架構

本系統使用嵌入式系統來實作網路攝影前端系統，嵌入式系統可以說是「嵌入」到設備的一種電腦系統控制、監視這個設備，因此可定義成「一種作為設備或工廠監視、控制用途的專用電腦系統」，結合微處理機或微控器之系統電路與其專屬之軟體，以達到系統獨立運作。其涵蓋軟硬體綜合體並與機構與週邊裝置連結，主要特性包含：主要用來執行特定功能、以微電腦與週邊構成核心、需要嚴格的時序與穩定度、全自動操作循環。近來，嵌入式多核心系統正應用於各個系統，因此也充滿挑戰，包括晶片體積變小(Chip size getting small)、成本需能更低(Cost down)、內建更多功能(Embedded more functionality)、內建更多記憶體(Embedded more memory)、獨立的即時作業系統(RTOS)。因此，本論文欲發展的網路攝影前端系統特別符合嵌入式系統的特性與要求，並配合雙核心之數位訊號處理器作高速影像處理運算，更能將此一複雜系統整合在單一系統晶片。

在實驗環境建置方面，本系統的監控伺服器平台使用桌上型電腦。遠端系統使用 CCTV 攝影機、TI 多媒體晶片系統與一個包含光感測器與輸出接點之 I/O 板，此 I/O 板與嵌入式 DSP 板連接，以作為遠端監控之用，以每秒 30 張影像所拍攝的路口車輛與行人，每張彩色色階影像解析度為 360×240 像素。在此監控的內容中，將發生移動的物件、駐留物件與取出物件的情形視為異常事件處理。

參、系統實驗

一、實驗方法與步驟

在實驗方法方面，可分為背景建構與更新、事件偵測、物件追蹤等分析步驟，系統流程如圖 3 所示。在背景建構與更新部份，將實驗各影片中移除之背景物件是否該區塊會被更新，而新加入的前景物件則被保留。因此，在發展階段中，係利用自行拍攝一小段測試用影片，在影片中移動其中一個背景物件，觀察其背景與前景的變化是否符合預期。

在事件偵測方面，若系統連續偵測出異常事件時，就將此異常事件連續出現的下一畫面作為轉換此段偵測事件監視時所需事件的第一個畫面。若未偵測到異常事件出現在系統連續的影像中，則將此段畫面視為異常事件的最後一個畫面。

當系統偵測完事件後，即偵測到異常事件的最後一個畫面時，會將事件的第一張影像到最後一張影像的異常事件之所有靜態畫面轉換成傳輸率為 15fps、位元傳輸率為 128Kbps、及影像大小為 176x144 像素的 QCIF 網路多媒體視訊檔案格式，並利用網路回傳至監控伺服器。

本系統中利用移動物體之移動慣性特性，將前一張影像所取得的物件加上其移動向量，可以計算預估在目前畫面上會出現的位置，進而依據各預估向量尋找出與原質心最小距離的物體，若其實際位置與預估的位置十分接近，即做為物件追蹤的下一位置。

二、實驗結果

如上所述，目前智慧型監控關鍵技術與整合應用仍在發展階段，若能利用嵌入式數位訊號處理器的計高速計算特性與操作性，在遠端便能即時分析異常事件的發生，以提高傳輸效率及異常事件判斷之正確性，並進行必要之處置如發出音響或燈號，與參數調整，相信可使智慧型監控達到更符合人性與提高功能性的境界。因此本論文提出以數位訊號處理晶片為基礎，將智慧型視訊監視系統實作於嵌入式系統上，利用數位訊號處理器的計高速計算特性，在遠端便能即時分析異常事件的發生，以提高傳輸效率及異常事件判斷之正確性。

我們實驗使用攝影機所拍攝出來的影像如圖 4 螢幕畫面，影像可以說是非常清楚，攝影機可以隨者偵測物的遠近，用手動來調整焦距。

我們將把攝影機架設好來進行偵測，當然光線也是很重，所得到的畫面如圖 4 所示。圖 5 為透過網路傳回畫面圖。當有人入侵時系統馬上發覺到，且可以獲得相當準確的結果。入侵畫面如圖 6。圖 7 是透過網路傳回畫面放大圖。當有人入侵時，利用自己所寫的 DSP 程式透過喇叭發出類似救護車的聲音來警告入侵者，或者來通知監視人員。



圖 4 嵌入式平台與偵測到的畫面



圖 5 傳回畫面放大圖

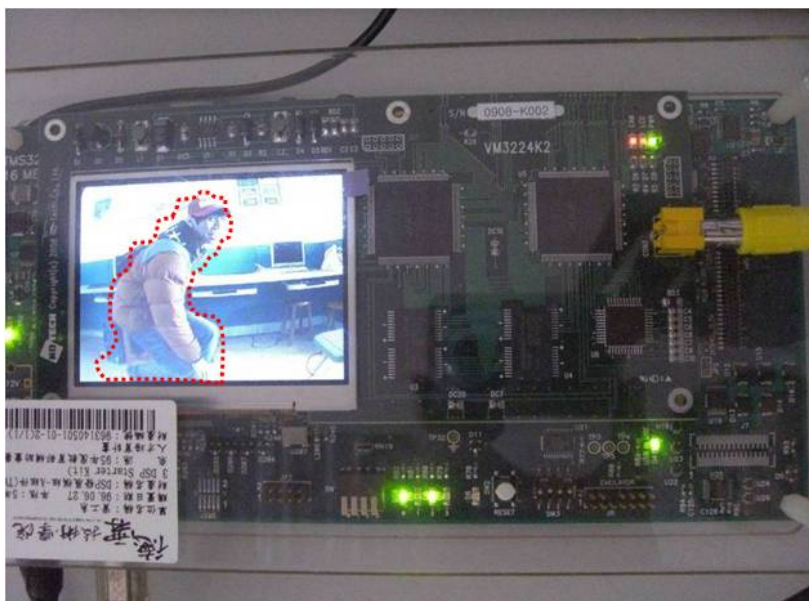


圖 6 嵌入式平台與入侵偵測結果畫面



圖 7 入侵畫面偵測結果傳回圖

肆、 結論

本系統獲得之初步成果：包括(1)利用數位訊號處理器的計高速計算特性，在遠端便能即時分析異常事件的發生，以提高異常事件判斷之正確性。(2)嵌入式攝影機端可即時分析異常事件，以提高傳輸效率及減輕伺服器主機端的工作負載。(3) 利用 I/O 介面，進行必要之處置，如發出警響或燈號。(4) 嵌入式攝影機端可藉由感測裝置，依據現場光線與亮度狀況，作拍攝與分析之參數調整，提高影像視訊品質，以利分析。(5) 除了嵌入式攝影機端可自行調整參數，主機監控人員亦可透過指令，遠端調整拍攝參數。透過一些實驗獲得初步的成果，且顯示出可以正確地偵測出異常狀況並提出適當的告警機制。

伍、 References

1. L. Brown, A. Hampapur, J. Connell, M. Lu, A. Senior, C. F. Shu, and Y. Tian, "IBM Smart surveillance system(S3):an open and extensible architecture for smart video surveillance," *Proc. of IEEE Conf. on Advanced Video and Signal Based Surveillance*, pp. 318-323, Sep. 2005.
2. E. Durucan, F. Ziliani, and O. N. Gerek, "Change detection with automatic reference frame update and key frame detector," *Proc. of IEEE-EURASIP. Workshop Nonlinear Signal and Image Processing (NSIP)*, pp. 57-60, 1999.
3. F. Y. Hu, Y. N. Zhang, and L. Yao, "An effective detection algorithm for moving object with complex background," *Proc. of Machine Learning and Cybernetics*, vol. 8, pp. 5011-5015, Aug. 2005.
4. C. Kim and J. N. Hwang, "Object-based video abstraction for video surveillance systems," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, pp. 1128-1138, Dec. 2002.
5. Q. Liu, D. Kimber, J. Foote, L. Wilcox, J. Boreczky, "FLYSPEC: a multi-user video camera system with hybrid human and automatic control", *ACM Conference on Multimedia*, pp. 484-492, 2002.
6. M. Piccardi, "Background Subtraction Techniques: A Review", *Proceedings of IEEE SMC 2004 International Conference on Systems, Man and Cybernetics*, Vol.4, October 2004, pp.3099-3104.
7. M. Shah, O. Javed, K. Shafique, "Automated Visual Surveillance in Realistic Scenarios," *IEEE Multimedia*, pp. 30-39, 2007.
8. O. Steiger, T. Ebrahimi, and A. Cavallaro, "Surveillance video for mobile devices," *Proc. of IEEE Conf. Advanced Video and Signal Based Surveillance*, pp. 620-625, Sep. 2005.
9. E. Stringa, and C. S. Regazzoni, "Real time video-shot detection for scene surveillance applications," *IEEE Transactions on Image Processing*, vol. 9, pp. 69-79, Jan. 2000.

10. L. Sigal, S. Sclaroff and V. Athitsos, "Skin Color-Based Video Segmentation under Time-Varying Illumination", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, Issue: 7, pp. 862-877, July 2004.
11. M. Soriano, B. Martinkauppi, S. Huovinen and M. Laaksonen, "Skin detection in video under changing illumination conditions", *Proceedings of 15th International Conference on Pattern Recognition*, Vol. 1, pp. 839 – 842, 3-7 Sept, 2000.
12. M. Valera and Velastin, "Intelligent distributed surveillance systems: a review", *Vision, IEE Proceedings- Image and Signal Processing*, vol. 152, Issue 2, 8 April 2005, pp. 192 - 204
13. M. Valsta and M. Pantic, "Fully Automatic Facial Action Unit Detection and Temporal Analysis", *Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06)*, pp. 149 – 149.
14. D. Vukadinovic, and M. Pantic, "Fully automatic facial feature point detection using Gabor feature based boosted classifiers", in *2005 IEEE International Conference on Systems, Man and Cybernetics*, Vol. 2, pp. 1692 – 1698, 10-12 Oct. 2005.
15. J. Wang, W.-Q. Yan, Kankanhalli, M.S., Ramesh Jain, and Reinders, "Adaptive monitoring for video surveillance", *Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia.*, vol. 2, 15-18 Dec., 2003 pp. 1139 – 1143
16. T. Yamazaki, K. Kayama; S. Igi, "Detection of People and Tracking Test Bed Using Multiple Sensors," *Multimedia and Ubiquitous Engineering*, pp. 135–140, 2007.
17. 吳國修, "基於離散小波轉換與形態學硬體實現多個車牌偵測", 國立台灣科技大學電子工程系碩士論文, 2008.
18. 祝珮軒, 遺留物及持有人自動偵測並具關鍵影像提供能力之視訊監控系統, 碩士論文, 國立中央大學資訊工程研究所, 台灣, 2006 年。
19. 陳又慈, 應用於監控系統上之即時人物追蹤、區分與異常行為偵測演算法, 碩士論文, 國立成功大學電腦與通信工程研究所, 台灣, 2005 年。
20. 拓璞產業研究所, 2008 年 5 月。
21. 資策會 MIC, 2007 年 11 月。