Jurnal Teknologi, 54 (Sains & Kej.) Keluaran Khas, Jan. 2011: 94–108

AUTO PAN TILT MOTION SURVEILLANCE SYSTEM

PEI SONG CHEE [1*], PEI LING LEOW [2*] & MOHOMAD SHUKRI ABDUL MANAF[3]

Abstract. This research develops webcam base pan and tilt camera with multiple tracking ability. This pan tilt surveillance system is small in size, portable and easy for installation. Convention surveillance system is limited to single function operation and have bulky camera system. The key component of this surveillance system is the attachment of low cost webcam onto pan and tilt servo motor. The movement of the webcam results from pulls and push coupling unit which attach to the motor. The smart surveillance system able to perform motion detection task, color blob tracking and laser light tracking. Automatic system enhanced its ability into real-time auto motion video record, photo snap shot and trigger alarm. This multi function system is developed with improve algorithm combination from different type of multi-filter. It is experimented under indoor and outdoor environment. The result shows the system able to compensate with the noise disturbance. The reported maximum speed is 0.125 ms^{-1} with 145° pan movement and 60° tilt movements. Such automated system is cost effective and can be used as robot vision, automated video conference and UAV application.

Keywords: Motion detection; color blob tracking; laser tracking; smart surveillance system; object tracking camera

Abstrak. Projek ini bertujuan untuk membangunkan satu sistem yang dilengkapi dengan pelbagai jenis jejakkan. Sistem tersebut kecil dari segi saiz, mudah alih dan memudahkan pemasangan. Konvensyen sistem terdiri daripada fungsi operasi tunggal dan mempunyai sistem yang besar. Projek ini boleh dicapai dengan pemasangan webcam murah dengan dua motor servo yang memainkan peranan sebagai sendi gerakan. Gerakan tersebut berdasarkan proses menarik dan mendorong dengan sambungan yang terlekat pada motor. Sistem tersebut mampu melakukan tugas pengesanan objek, jejakkan warna dan jejakkan gumpalan warna cahaya laser. Selain itu, video rakaman, gambar tangkapan dan pencetusan penggera boleh dilakukan. Sistem pelbagai fungsi ini dibangunkan dengan algoritma gabungan dari pelbagai jenis penapis. Alat tersebut telah dicuba dalam keadaan bilik dan keadaan luaran. Kajian menunjukkan sistem mampu mengkompensasi dengan gangguan hingar. Sistem tersebut mampu mencapai kelajuan0.125 ms^{-1} dengan 145° dan 60°periuk gerakan miring. Pemasangan system tersebut

[1-3] Control & Instrumentation Engineering Department, Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor Darul Ta'azim, Malaysia
[*] Corresponding author: Telephone: +6075535421, Fax: +607-5566272; Email: [1]pschee2@live.utm.my or [2]leowpl@fke.utm.my

melibatkan kos yang murah dan boleh diaplikasikan dalam robot visi, persidangan video dan aplikasi UAV automatik.

Kata kunci: Pengesanan gerakan; pelacakan gumpalan warna; pelacakan laser; sistem pemantauan pintar; kamera pelacakan objek

1.0 INTRODUCTION

Surveillance system has been increasingly demand with the daily increase of crime rate. Hence, continuous monitoring activity over a wide environment is needed. Close Circuit Television System (CCTV) has provided a network system to extend the eye limitation from one observer location to another observer location.

Traditional CCTV requires real time monitoring from surveillance officer. According to the research carried out by the UK Police Scientific Research Branch,[1] the observer will experience "video blindness" after twenty to forty minutes of observation. Intelligent surveillance system [2-5] provides the technology to analyze information from situated sensor. Maturity of this kind of technology has brought to the deployment of isolation application which provides a certain type of functionality such as facial recognition, motion detection, etc. [6-8]

Huibin *et al.*[9] proposed multi-camera tracking based on information fusion in video surveillance application. This multi-camera technology offers extending surveillance function and can solve many complex situations such as occlusion, brightness and shadows. However, this technology needs installation of many vision sensors which will create problem when involving maintenance work. To overcome this drawback, pan and tilt mechanism [10-12] was implemented to this smart surveillance system which will enlarge the system Field Of View (FOV).

Integrated functionality such as motion detection in pan and tilt camera has raise robustness issue when camera functioning in complex environment. Detection algorithm must equip with shadow and light noise elimination. Frame differencing technique [13-14] is the algorithm that was used to compare the differences between two consecutive frames. Karan *et al.*[15] improved this algorithm with the addition of erosion filter. However, consecutive update of image consumes a lot of memory space and slow object's motion cannot be detected. Another method used is the Background Modeling [16-17] technique by comparing the foreground pixel frame and picture frame to provide the solution for detecting slow object movements.

From previous review, the mechanism of pan and tilt camera is bulky, complex, not portable and only integrated to one functionality. Susan O'Donnell *et al.*[18] proposed the idea of converting municipal surveillance camera into municipal webcams. This idea certainly brings to the benefit of small size mechanism, portable and low cost. In this paper, we proposed and developed low cost auto motion video surveillance system with the implementation of multiple tracking abilities. The system built was able to detect intruders and trigger alarm with Background Modeling (BM) and Otsu threshold algorithm [19]. This motion detection features also equipped with motion video record and photo capture where video or photo will be captured when the object is detected.

In order to enhance the auto surveillance system, special features of interest such as human face, vehicle and eyeball tracking are included in the algorithm. Computer vision is developed to enable the system user to define object's color as well as laser light tracking ability. This can be implemented with color filtering and dilation method. The surveillance system also designed to be able to follow the movement of the laser light and centered it at the FOV.

2.0 SYSTEM DESIGN AND DESCRIPTION

Figure 1 shows the architecture of advance auto motion surveillance system. The system consists of modified servo actuated Pan Tilt Webcam Unit (PTWU), servo controller circuit and GUI tools. The PTWU was designed to be small, compact and portable. This design features enable system to operate in complex environment for example high hill surveillance control.

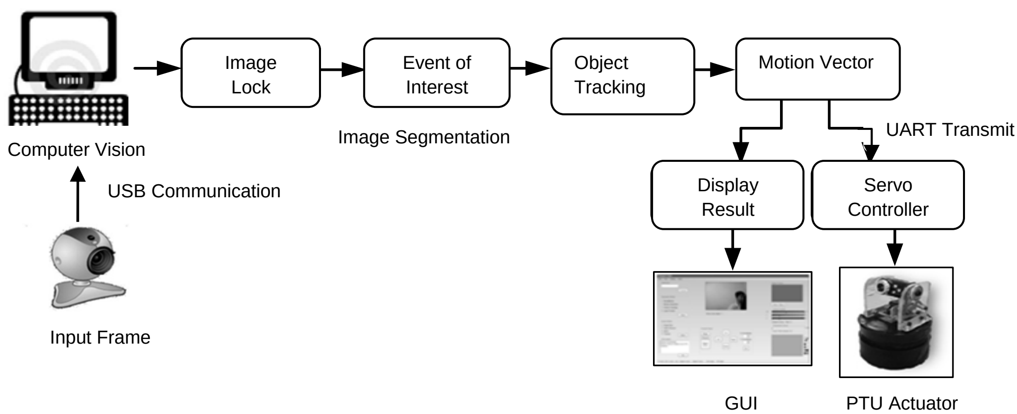


Figure 1 Architecture of auto motion surveillance system

From Figure 1, webcam acts as motion sensor that transfers the input frame to the computer via USB communication with the speed of 30 frames per second. The picture frame was then transferred to the computer in the form of Quarter Video Graphic Array (QVGA).The pixel data is limited to 240×320 resolutions. This is to ensure that the digital video recording required only very little amount of data storage. Hence, image processing can be performed faster. Consequently, the picture frame will be locked where image is segmented based on the event of interest. Object's motion detection can be performed with BM algorithm which computes the difference between pixel intensities of the input pair for each time instant. When the difference is high, it indicates that an object is detected and initiates the video recording or photo capturing based on user requirement. Video recording task was performed by the Video Compressor Manager (VCM). VCM encodes the content in Audio-Video Interleaved (AVI) file. The file will be stored in a formatted time and date when the object is detected. For security enhancement purposes, alarm will be triggered to alert the owner of the premises and to frighten the intruders.

Auto surveillance system supports a variety of user-defined behavior tracking capabilities via GUI tools which is written in Visual Basic.Net language. Drag and drop function was designed to allow the user to choose any object to be tracked based on the object's color properties. This feature can be achieved with color filtering algorithm. This algorithm manipulates Red-Green-Blue (RGB) image with three dimensional $M \times N \times 3$ double matrixes. Laser tracking involved high intensity light which required brightness and color filtering. Thresholding algorithm and color filtering algorithm was implemented to reject or keep pixel below some value.

After identify the tracking features, object can be tracked based on motion vector algorithm. Within the algorithm, coordinate is updated to the Motor Positioning Module (MPM). MPM converts the coordinate to servo pulse and send to servo controller via Universal Asynchronous Transmitter (UART).A PIC 30F4011 microcontroller was programmed as a servo controller using C language. It continuously monitors and sends pulse to the servo motor with timer control and interrupt algorithm. Meanwhile, the result will be displayed onto the screen panel with tracked object be highlighted in red region.

3.0 SYSTEM CONTROL UNIT

In this section, the software algorithm and hardware setup employed in this auto tracking surveillance system are discussed. The algorithm has been optimized with the integrated real-time filtering technique.

3.1 Algorithm

3.1.1 Motion Detection Algorithm

Figure 2 shows the Background Modeling (BM) algorithm. The Grayscale image is represented by matrix of double data type of size M × N where M is tall pixel and N is wide pixel. The pixel intensities denote in [0, 1] with 0= black and 1= white. Absolute difference is calculated between two consecutive frame which represented by $f_{t-1}(x,y)$ and $f_t(x,y)$. If the difference is close to zero, it indicates that no changes occurred. The result, D(x,y) is converted to binaries with Otsu's threshold method to produce black and white image [19]. White image indicates object detection. If background $f_t(x_1, y_1)$ is changed, a new background is modeled again in background update block.

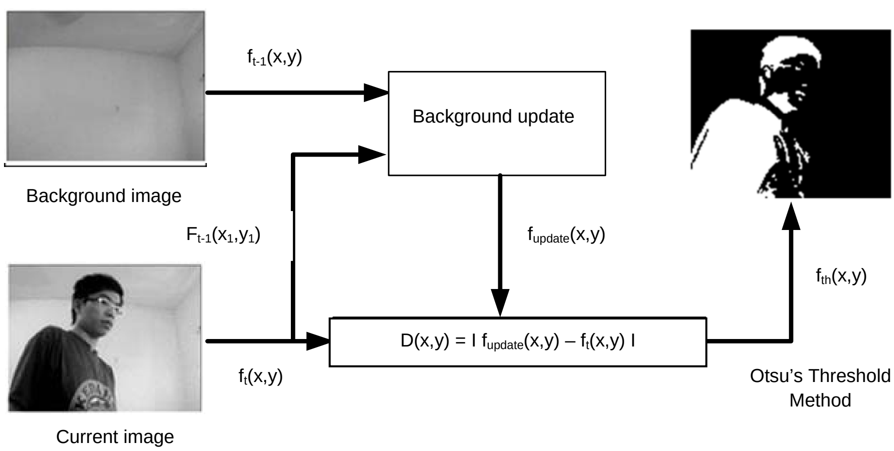


Figure 2 Background modeling algorithm

BM algorithm performance is fast, however it is very sensitive to noise due to illumination changes. This affects the gray level which gives false signal to the system. Otsu method [19] provides an auto threshold selection from image histogram. This can be achieved by minimizing the weight within-class variance. The weight within class variance can be represented by Equation 1:

$$\sigma_w^2(t) = q_1(t)\sigma_1^2(t) + q_2(t)\sigma_2^2(t) \tag{1}$$

Two classes are estimated which include class probability and class mean,
Two class probabilities, q_1 and q_2 are estimated in Equation 2:

$$q_1(t) = \sum_{i=1}^{t} P(i) \quad ; \quad q_2(t) = \sum_{i=t+1}^{I} P(i) \tag{1}$$

Class means estimation:

$$\mu_1(t) = \sum_{i=1}^{t} \frac{iP(i)}{q_1(t)} \quad ; \quad \mu_2(t) = \sum_{i=t+1}^{I} \frac{iP(i)}{q_2(t)} \tag{2}$$

Which finally result in:

$$\begin{aligned} \sigma_1^2(t) &= \sum_{i=1}^{t} [i - \mu_1(t)]^2 \frac{P(i)}{q_1(t)} \\ \sigma_2^2(t) &= \sum_{i=t+1}^{I} [i - \mu_2(t)]^2 \frac{P(i)}{q_2(t)} \end{aligned} \tag{3}$$

This integrated Otsu's algorithm increases the robustness of the Background model technique. It can adapt to different kind of scene and perform more accurate detection.

3.1.2 Color Tracking Algorithm

In RGB, three indexed images exist. First image contains red portion, second portion contains green and third portion contains blue potion of image. Each of the portion images contribute to 8 bit, which result to 24 bit. Figure 3 shows color filtering algorithm flow diagram. The RGB color file format in full saturated faces of 24 bit per pixel RGB can be represented within a plane (0,255) where 0 is minimum value of the color range and 255 is maximum value of color range. The purpose of this filter is to filter out red color object. Hence, the value set for the color range consists of R (110, 255), G (0, 20), B (0, 20). The pixel fall within the range will remain the color and pixel fall out of range will be converted into black color. The image is then undergoing binarization.

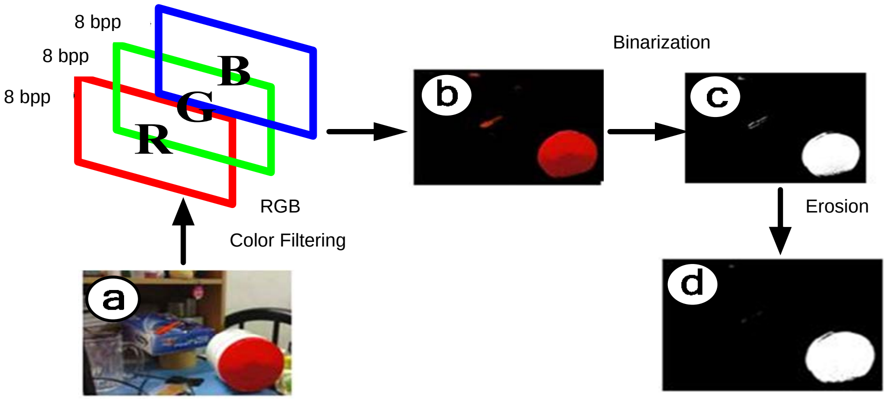


Figure 3 Color filtering algorithm

Figure 3(c) shows the binarization result. From the observation, the picture contains of noise. Hence, erosion method was utilized to eliminate irrelevant noise. Erosion is basis morphology which compute local minimum over the specified area. Figure 3(d) shows that erosion can be used to shrink isolated individual element.

3.1.3 Laser Tracking Algorithm

Figure 4 shows the algorithm which makes use of the characteristic of the laser which is high color pixel intensity and brightness compared to normal ambient light. The initial picture undergoes color filter with the range of R (200, 255), G (0, 150) and B (0, 150). Figure 4(b) shows the color filter's result.

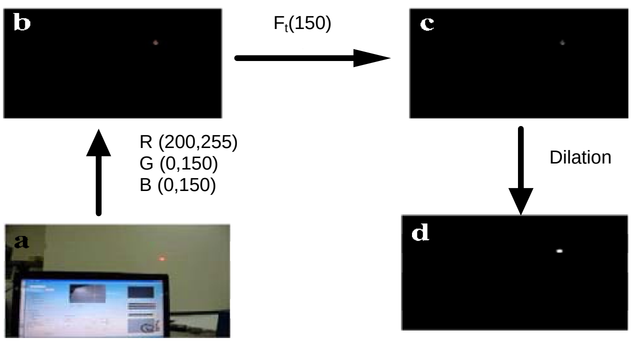


Figure 4 Laser tracking algorithm

In order to ensure that the picture does not influence by the ambient light, brightness threshold filter was implemented. This threshold filter consists of 8 bitsper-pixel (bpp) image, which the maximum value can up to 255. Figure 4(c) corresponds to comparison between the source picture and threshold value. The pixel fall below the threshold value was rejected and the pixel above threshold value were remained. The technique can be applied as follow:

```
Let T=threshold value, f[m,n]=frame
If f [m,n] ≥ T
Then fb[m,n]=object=1
Else fb[m,n]=background=0
```

Figure 5 Algorithm for threshold filter technique

After thresholding filter, only some pixels exist which is hardly seen due to the small size. Dilation is a mathematical morphology to enlarge the boundary of region. It is implemented to the picture frame so that the laser dot can be clearly seen.

3.1.4 Positioning Algorithm

Stereo Visual [20] is 3-D stereoscopic which implemented the concept of both our eyes to estimate the distance between object and camera. Figure 6 shows the 3-D coordinate parameter [x, y, z] where x and y is 2-D axis, z is the distance between object and camera. Stereo Visual implementation involved two cameras to estimate the depth of z axis. Hence, for single camera application, it is difficult to estimate the distance.

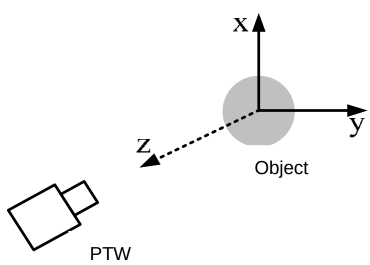


Figure 6 Coordinate parameter [x, y, z]

In order to find the coordinate, the picture frame can be divided into four quadrant which is I, II, III and IV. Figure 7 shows the four quadrants coordinate. Difference between two successive frames of image sequence, f_{t-1} (x,y) and f_t (x,y) is taken into account. For example, calculated result, x appeared to be positive value, and y is negative value, it falls into quadrant IV. The movement of the servo motor is controlled by servo controller. Servo controller is programmed to function as stepper motor where the motor will move in certain steps. For this system, each step will contribute to 4° of movement. The pan and tilt servo will move to the particular quadrant until the object is centered to FOV. For example, the servo motor needs 16° to reach its center; it will turn 4 steps to reach the center. Then, the particular coordinate will be reset to zero. This algorithm enables estimation of motion vector with 2-D modeling. Since the coordinate accuracy is not a major concern, the system still able to compensate with the high speed servo control update.

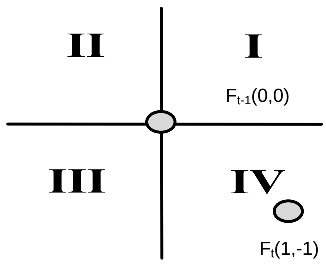
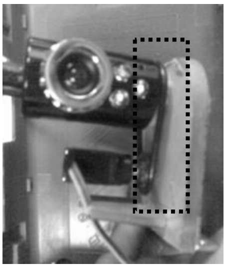


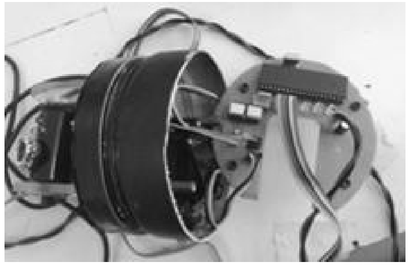
Figure 7 Four quadrant coordinate

3.2 Hardware Setup

This section describes the configuration of hardware setting which consists of two pan tilt servo motors and one webcam. As illustrated in Figure 8(a), the hardware utilized the pull push mechanism where the coupling unit pulls the camera up and down with pan angle. This mechanism enable the rotation angle become wider and minimized the camera size.



(a)



(b)

Figure 8 Hardware configuration

The circuit board used to control the servo motor was attached under the round bottle. The circuit board was cut in round shape to fit into a plastic bottle. Figure 8(b) shows the circuit design where a microcontroller was utilized to control pan and tilt motion. The ease of circuit design will reduce system complexity.

4.0 EXPERIMENTS AND RESULT

Figure 9 shows the experiment setup which was used in the tracking analysis. The experiment was performed with Intel Core Duo laptop, modified PTWU and ATX Power Supply. To simplify the system, the power supply can be generated from battery as the PTWU operate at 5 V. The webcam configured to capture 240 × 320 resolution image at 30 frames per second.

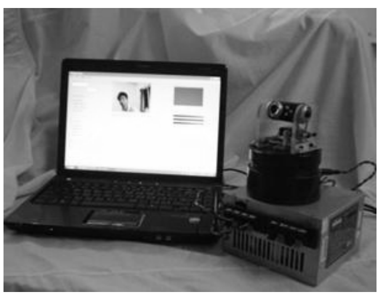


Figure 9 Experiment setup

The experiments were carried out at indoor and outdoor condition to test the robustness of the automated tracking system. Figure 10 shows the highlighted object from entering the screen until the object moved to the center of the camera under room condition.

Figure 10 Motion detection in room condition

The automatic detection experiment also carried out at outdoor environment where it would expose to massive sunlight and showdown influence. Figure 11 shows the car being detect when it passed through the camera view.

Figure 11 Car detection at outdoor condition

Both detection results showed that the automated surveillance system is able to detect moving objects whether in ideal condition (indoor condition) and noisy condition (outdoor condition). The result was then video recorded and photo snapshot.

To investigate tracking of event of interest, the experiment was carried out on object of different color. Figure 12 shows the different colored object tracking result. Figure 12(a) shows the red color bottle being tracked and face tracking was performed in Figure 12(b) whereas Figure 12(c) shows the hair color tracking result. The tracked color could be selected by the user via GUI. When the color object moved, the center of FOV would be defined and the camera would move onto the direction.

(a)

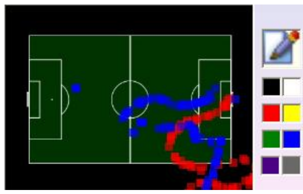
(b)

(c)

Figure 12 Color tracking (a) red color moving object, (b) face tracking and (c) black hair tracking

Laser tracking experiment was done with the projection of red laser light. Figure 13 shows the laser tracking result in Figure 13(a) and laser trajectory in Figure 13(b).

(a)



(b)

Figure 13 Laser tracking experiment

The intersection lines at Figure 13(a) were drawn across the frame to indicate the position of laser light. The color of laser dot trajectory can be freely chosen by user to differentiate various laser light sources.

To operate the automated system, the characteristic of the surveillance unit need to be determined. Figure 14 shows the graph of the maximum speed which could be achieved by the camera. The tracking PTWU takes 0.8 second to reach distance of 10 cm. The calculated maximum speed was 0.125 ms^{-1}. The camera able to achieve maximum 145° pan movement and 60° tilt movements.

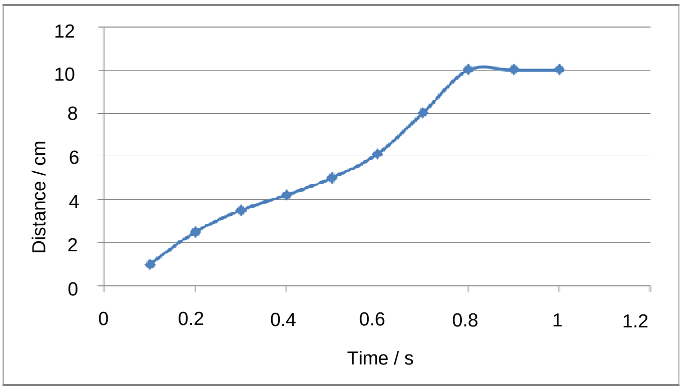


Figure 14 Maximum speed of tracking system

5.0 CONCLUSION

In this paper, the efficiency, fast and robust motion surveillance system was presented. The main improvement of this surveillance system is the integrated of multiple tracking purposes into single low cost webcam. The smart surveillance system is able to perform motion detection function, color blob tracking and laser light tracking. These abilities were developed with improvement of algorithms combination from different types of filter. This algorithm does not require the initial positioning of the tracking camera. The system was experimented indoor and outdoor environment and managed to show ability to track moving objects. The reported maximum speed is 0.125 ms^{-1} with 145° pan movement and 60° tilt movements. This automated surveillance system is suitable for various kind of application such as robot eye, smart home security system and UAV application.

ACKNOWLEDGEMENTS

The work is financed by Zamalah/Institutional Scholarship provided by Universiti Teknologi Malaysia and the Ministry of Higher Education Malaysia. The authors would like to acknowledge the Control and Instrumentation Engineering Department of the Faculty of Electrical Engineering, Universiti Teknologi Malaysia for providing laboratory space and technical support. The authors would like to thanks Jun Kong who have become the study subjects in this research.

REFERENCES

[1] Wallace, E. and C. Diffley, 1998. CCTV Control Room Ergonomics. Police Scientific Development Branch (PSDB), in UK Home Office.

[2] Sedky, M. H., M. Moniri, and C. C. Chibelushi. 2005. Classification of Smart Video Surveillance Systems for Commercial Applications. In Advanced Video and Signal Based Surveillance, AVSS 2005. IEEE Conference on 2005.

[3] Greiffenhagen, M., *et al.* 2001. Design, analysis, and Engineering of Video Monitoring Systems: An Approach and a Case Study. Proceedings of the IEEE. 89(10): 1498-1517.

[4] Cucchiara, R. and G. Gualdi. 2010. Mobile Video Surveillance Systems: An Architectural Overview, in Mobile Multimedia Processing. 89-109.

[5] Chiao-Fe, S., *et al.* 2005. IBM Smart Surveillance System (S3): A Open and Extensible Framework for Event Based Surveillance. In Advanced Video and Signal Based Surveillance. AVSS 2005. IEEE Conference on 2005.

[6] Xie, X. and K.-M. Lam. 2009. Facial Expression Recognition Based on Shape and Texture. *Pattern Recognition.* 42(5): 1003-1011.

[7] Simpson, W. A. and A. Newman. 1998. Motion Detection and Directional Tuning. Vision Research. 38(11): 1593-1604.

[8] van den Berg, A.V. and W.A. van de Grind. 1989. Reaction Times to Motion Onset and Motion Detection Thresholds Reflect the Properties of Bilocal Motion Detectors. *Vision Research.* 29(9): 1261-1266.

[9] Huibin, W., *et al.* 2009. Multi-Camera Tracking Based on Information Fusion in Video Surveillance. In Image and Signal Processing. CISP '09. 2nd International Congress on 2009.

[10] [Mian, A. 2008. Realtime Face Detection and Tracking Using a Single Pan, Tilt, Zoom Camera. In Image And Vision Computing New Zealand. IVCNZ 2008. 23rd International Conference.

[11] Yu-Wen, H., *et al.* 2002. Simple And Effective Algorithm For Automatic Tracking Of A Single Object Using A Pan-Tilt-Zoom Camera. In Multimedia and Expo. ICME '02. Proceedings. 2002 IEEE International Conference on 2002.

[12] Chu-Sing, Y., *et al.* 2008. PTZ Camera Based Position Tracking In IP-Surveillance System. In Sensing Technology. ICST 2008. 3rd International Conference on 2008.

[13] Jain, R. and H. H. Nagel. 1979. On the Analysis of Accumulative Difference Pictures from Image Sequences of Real World Scenes. Pattern Analysis and Machine Intelligence, IEEE Transactions on PAMI-1(2): 206-214.

[14] Haritaoglu, I., D. Harwood, and L. S. Davis. 2000. W⁴: real-time Surveillance of People and Their Activities. Pattern Analysis And Machine Intelligence, IEEE Transactions on. 22(8): 809-830.

[15] [15] Gupta, K. and A.V. Kulkarni, 2008. Implementation of an Automated Single Camera Object Tracking System Using Frame Differencing and Dynamic Template Matching, in Advances in Computer and Information Sciences and Engineering, T. Sobh, Editor. Springer Netherlands. 245-250.

[16] Stauffer, C. and W. E. L. Grimson. 1999. Adaptive background Mixture Models For Real-Time Tracking. In Computer Vision and Pattern Recognition. IEEE Computer Society Conference on 1999.

[17] Wada, T. and T. Matsuyama. 1996. Appearance Sphere: Background Model For Pan-Tilt-Zoom Camera. In Pattern Recognition. Proceedings of the 13th International Conference on 1996.

[18] O'Donnell, S. and M. Richard. 2006. Turning Municipal Video Surveillance Cameras Into Municipal Webcams. in Technology and Society. ISTAS 2006. IEEE International Symposium on 2006.

[19] Otsu, 1979. A Threshold Selection Method from Gray-Level Histograms. Systems, Man and Cybernetics, IEEE Transactions on 9(1): 62-66.

[20] Jiandan, C., S. Khatibi, and W. Kulesza. 2007. Planning of a Multi Stereo Visual Sensor System - Depth Accuracy and Variable Baseline Approach. in 3DTV Conference, 2007.