International Journal of Electrical, Electronic and Communication Sciences
ISSN: 2517-9438
Vol:4, No:6, 2010

Smart Surveillance using PDA

Basem Mustafa Abd. Amer , Syed Abdul Rahman Al-Attas

***Abstract*—**The aim of this research is to develop a fast and reliable surveillance system based on a personal digital assistant (PDA) device. This is to extend the capability of the device to detect moving objects which is already available in personal computers. Secondly, to compare the performance between Background subtraction (BS) and Temporal Frame Differencing (TFD) techniques for PDA platform as to which is more suitable. In order to reduce noise and to prepare frames for the moving object detection part, each frame is first converted to a gray-scale representation and then smoothed using a Gaussian low pass filter. Two moving object detection schemes i.e., BS and TFD have been analyzed. The background frame is updated by using Infinite Impulse Response (IIR) filter so that the background frame is adapted to the varying illuminate conditions and geometry settings. In order to reduce the effect of noise pixels resulting from frame differencing morphological filters erosion and dilation are applied. In this research, it has been found that TFD technique is more suitable for motion detection purpose than the BS in term of speed. On average TFD is approximately 170 ms faster than the BS technique

***Keywords*—**Surveillance, PDA, Motion Detection, Image Processing , Background Subtraction.

I. Introduction

PDA is an abbreviation for Personal Digital Assistant, a handheld device that combines computing, telephone/fax, Internet and networking features. A typical PDA can function as a cellular phone, fax sender, Web browser and personal organizer. Unlike portable computers, most PDAs began as pen-based, using a stylus rather than a keyboard for input. This means that they also integrated handwriting recognition features. Some PDAs can also react to voice input by using voice recognition technologies. PDAs of today are available in either a stylus or keyboard version

PDAs have become devices many of us use on a daily basis. They are relatively inexpensive; offer a many of applications while also being small, portable, and easy to use. Many standard PDA interfaces have been developed for a wide range of applications. In addition to standard applications such as word processing, calendar management, and calculators, some more specialized systems have been developed. For example, character recognition [1, 2, 3], educational tools [4, 5], and face recognition [6].

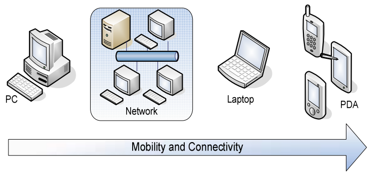

Fig.1 Evolution of Mobility & connectivity

Therefore, it seems relatively obvious that PDAs may represent a good base for supervision. Their small size, processing capability, and portability make them suitable platform for surveillance

II. Related Work

There are few on going research work in this field, and most of them developed intelligent applications using image processing such as character recognitions. Examples are Z. Luo et al[1] developed a Chinese character recognition for PDA and H. Kang et al [2] designed an Interface on PDA for Korean language. Moreover, Face recognitions as in J. Yang, et al. [3] signed and developed a Face Recognition System based on PDA. Besides that, A few of researchers have developed a mobile surveillance. Sheng-Tun et. al [7] developed a real-time surveillance services to handheld devices over the Internet. They developed a surveillance system, which delivers various surveillance services over the Internet for supporting mobile PDAs users. J2ME technology is used to developing this system; the resulting system is composed of a surveillance server and a mobile thin-client monitor, PDA Watcher. The surveillance server is built upon a set of distributed components which cooperates each other through message.

III. The over all system

In this Research. The Overall system architecture block diagram as shown in Fig. 2 gives an over view of the main stages for the smart surveillance system based on PDA In this research Pocket PC (HP iPAQ hx4700) is chose as a platform for this project. The main reasons to choose Pocket PC rather than Palm PDA are : Palm is an Organizer, Pocket PC is a computer. Also Pocket PC is multitasking (several programs can run at once).The Pocket PC uses a windows interface which is already user-friendly to many users. Finally Pocket PCs is more popular and used by many people. There are many kinds of camera for PDA. There are built–in cameras, wireless camera and digital card camera. The most of modern smartphones and PDAs have built-in camera.

Basem Mustafa Abd Muftah Amer is with Data Analysis Department, Faculty of Economy , The 7th April University, Sabratha Almadina 307. Libya Arab Aljamhiriya. (Phone: +218 92 4645356 P. Box: 307 Sabratha –Libya email: basim_Zawaghy@yahoo.com).

Syed Abdul Rahman Al-Attas is with the Computer Vision, Video, and Image Processing Research Group (CVVIP), Department of Microelectronics and Computer Engineering, Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 Skudai, Johor Darul Takzim, Malaysia (e-mail: syed@fke.utm.my).

International Journal of Electrical, Electronic and Communication Sciences
ISSN: 2517-9438
Vol:4, No:6, 2010

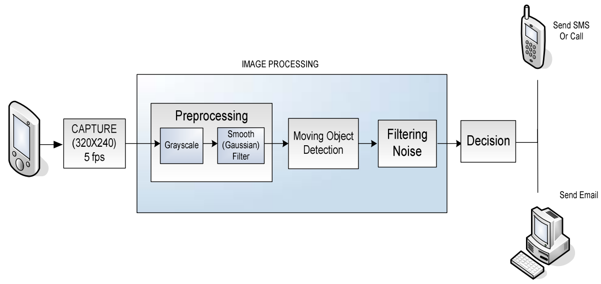

Fig. 2 Overall system architecture block diagram

In this Project, HP photosmart mobile camera has been used with frame rate of 5 FPS and dimensions of 320X240 pixels.

Images obtained from the above mentioned source are mainly 24-bit RGB images. When an image is obtained, some preprocessing is carried out to prepare the image/frame for next processes. The first preprocessing is to convert the color image to grayscale image. Secondly, the image is filtered by using Gaussian (smooth) filter to remove noise from the image. The following sections elaborate on these processes

Grayscaling is an importance process that has been used to remove color values in the image. As it is known that the colored image is composed of three color channels of Red, Green and Blue. Basically grayscaling converts a color image (a 24 bit RGB, 8 bit per channel in this case) to an 8 bit gray level image. There are many formulas to calculate the grayscale. Equation (1) is one of these formulas that produce an acceptable result:

$$Grayscale = \frac{R+G+B}{3} \approx 0.333R + 0.333G + 0.333B \quad (1)$$

The Gaussian smoothing operator is a 2-D convolution operator that is used to blur images and remove detail and noise. In this case it is similar to the mean filter, but it uses a different kernel that represents the shape of a Gaussian (bell-shaped). In 2-D, an isotropic (*i.e.* circularly symmetric) Gaussian can be expressed as shown in the Equation (2):

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\left(\frac{x^2+y^2}{2\sigma^2}\right)} \quad (2)$$

The distribution of Gaussian is shown in Fig. 3.

The filter kernel of the Gaussian blur filter is decomposable in the product of a vertical vector and a horizontal vector. Fig. 4 shows the possible vectors, multiplied by each other they will produce Gaussian blur filters:

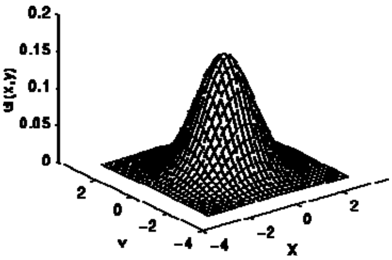

Fig. 3 2D Gaussian distribution with mean (0, 0) and σ =1

$$\begin{pmatrix}1\\2\\1\end{pmatrix} \times \begin{pmatrix}1 & 2 & 1\end{pmatrix} = \begin{pmatrix}1 & 2 & 1\\2 & 4 & 2\\1 & 2 & 1\end{pmatrix}$$

Fig. 4 The mask of Gaussian filter

There are two major types of motion detection: static background motion detection and dynamic background motion detection. The first will be used for surveillance means: the surveillance camera does not move and has a capture (in a bitmap for example) of the static background it is facing. The second is much more complex. Indeed, in the entering streaming video the background moves at the same time as moving objects.

Motion detection uses image subtraction to determine object motion between two images. Typically, two images of the same scene are subtracted, yielding an image that shows only the differences. Constant image information is removed and only the objects that have removed between two images will appear. There are many methods to implement detection of moving object, but the most common three methods for Moving Detection: Temporal differencing, Background subtraction algorithm and optical flow. In this paper, background subtraction (BS) and Temporal Frame Differencing (TFD) methods are tested to decide on their performance and the suitability for a surveillance system based on PDA platform.

TFD attempts to detect moving regions by making use of the pixel-by-pixel difference of consecutive frames (two or three) in a video sequence. This method is highly adaptive to dynamic scene changes. Lipton et al. [9] presented a two-frame differencing scheme where the pixels that satisfy the following Equation are marked as foreground.

$$\left| I_t(x,y) - I_{t-1}(x,y) \right| > T \quad (3)$$

where

I_t is the current frame.

I_{t-1} is the previous frame.

T is threshold value.

Background subtraction is a widely used approach for detecting moving objects in videos from static cameras. The rationale in the approach is that of detecting the moving objects from the difference between the current frame and a reference frame, often called the "background image", or "background model". The background image must be a representation of the scene with no moving objects and must be kept regularly updated to adapt to the varying luminaries conditions and geometry settings. Models that are more complex have extended the concept of "background subtraction" beyond its literal meaning. In general, BS can be simplified as shown in the Equation (4):

$$f(x,y) = abs(Frame(x,y) - Background(x,y)) \quad (4)$$

It is necessary to update the background image frequently in order to guarantee reliable motion detection. The basic idea in background adaptation is to integrate the new coming information into the current background image using the following first order Infinite Impulse Response (IIR) filter as shown in the Equation (5).

$$B_{(t+1)} = \alpha I_t + (1-\alpha)B_t \quad (5)$$

In our approach, α (Adaptation coefficient) is kept small (0.1) and the update process based on Equation. (3.5) is only indented for adapting to slow changes in overall lighting. The activity of each pixel is monitored during several consecutive frames. The intensity values of those pixels that are active most of the time are directly copied from the latest I_t to B_t. In this way we can adapt reasonably fast to new static objects appearing on the scene, like stationary objects, or to sudden changes in the illumination level

Fig. 5 (a) The background image (b) The adapted background when α =0.5 (after 2 seconds) (c) The final adaptation of background (after 5 seconds)

In this project, the threshold value was fixed to the value (T=32) and followed by the Equation (6), assume that $g(x,y)$ is the output image (binary image).

$$g(x,y) = \begin{cases} 0 & if\ f(x,y) > T \\ 1 & Elsewhere \end{cases} \quad (6)$$

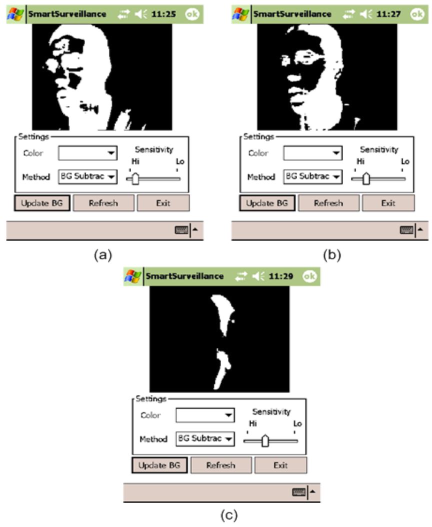

Fig. 6 The result of threshed (a)T=32 (b) T=60 and (c) T-100

In order to reduce the effect of noise pixels resulting from frame differencing, Morphological operators Erosion and Dilation are applied to the detected moving object. By using erosion and dilation operators in turn, some of the noise (grainy noise) can be removed from the mask. Apart from the noise removal, erosion operation might disconnect the links between loosely connected regions, which are not the desired foreground objects most of the time. When the connectedness of a region is lost and the region area is below a threshold, it is not treated as a foreground object any more. On the other hand, strongly connected regions are not affected from this operation (except from their boundaries) and a subsequent dilation operation recovers the shrinkage caused by erosion.

Fig. 7 Illustrates the result of Morphological operators (Erosion and followed by Dilation)

For tracking the detected moving object, the tracking process is carried out recursively for all the input frames captured by the camera. Abounding box will be drawn at the location of the detected moving object in each frame. Projection Histogram techniques have been used for tracking an object. The process of this technique involves the size of the entire image frame (320X240 pixels). Figure 3.10 shows this projection histogram technique .The projection of the horizontal axis is obtained based on the Equation (7):

$$P_h(x) = \sum_y f(x,y) \quad (7)$$

And the projection of the vertical axis is obtained based on the Equation (8):

$$P_v(y) = \sum_x f(x,y) \quad (8)$$

where:

x is column, x =0,1,2,3,4,......,255

y is row, y =0,1,2,3,4,......,255

$f(x,y)$ is the pixel value at column x and row y .

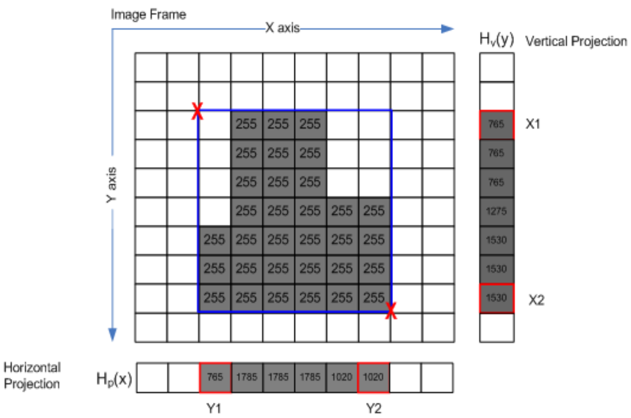

Fig. 8 Summations of rows and columns to obtain the vertical and horizontal

IV. Program Flow Chart

The surveillance system program is written using Microsoft embedded visual C++ which is part of Microsoft Visual Studio 2005 Professional Edition. The flow chart of the program is given in Figure below. In Initialization stage the program runs these steps:

1- Check if the camera is available.
2- Configure the camera to run at normal mode, and set the size of the image to 320 X 240 pixels.
3- Disable screensaver and / or auto off screen.
4- Set the frame rate at 5 FPS.

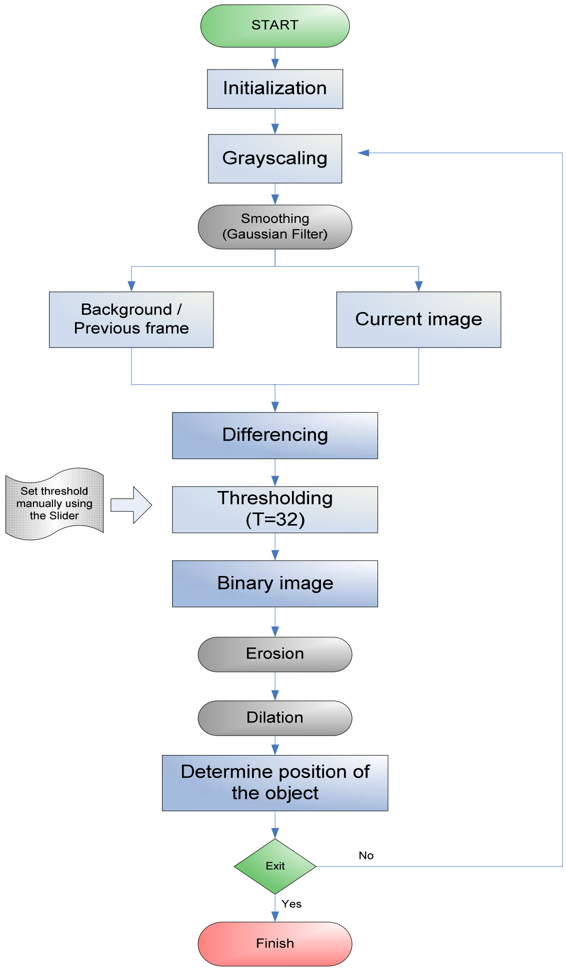

Fig. 9 Program flow chart

The computation time of the processes of the system is shown in Fig. 10. As it is illustrated Gaussian filter has larger computation time (around 25 millisecond) this due to the convolution operation over the a whole frame. Besides that, Frame Differencing takes 20 milliseconds. Grayscaling and thresholding spend 15 and 9 milliseconds respectively. Finally, Morphological operation Dilation and Erosion takes 15 milliseconds for each.

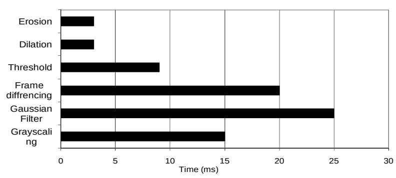

Fig. 10. Computation time for the processes of the system

International Journal of Electrical, Electronic and Communication Sciences
ISSN: 2517-9438
Vol:4, No:6, 2010

A whole surveillance system used adaptive BS takes around 370 milliseconds to achieve processing two frames. TFD Technique takes around 200 ms to complete processing two consecutive frames. Fig. 11 shows the performance comparison between BS and TFD.

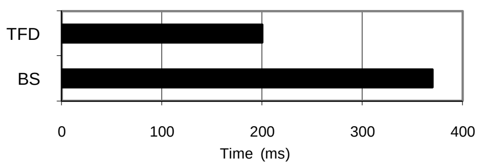

Fig. 11 The computation performance comparison between BS and TFD

V. Conclusion

In general, the main objective of this project is to develop a surveillance system based PDA platform and it is programmed to fit for online (real-time) application. In order to reduce noise and to prepare frames for the moving object detection part, each frame is first converted to a gray-scale representation and then smoothed using a Gaussian low pass filter.

The entire system which has been developed shows perfect result for detection the moving objects. TFD and BS have been used because of their high performance of handling and the speed of the process. A number of preliminary processes have been applied to prepare the image for next stages. These processes have proved high accuracy of detection. The program performed well with rate of 5 FPS on HP iPAQ hx4700 with frame size of 320x240 pixels Finally it has been found that TFD technique is more suitable for motion detection purpose than the BS in term of speed. On average TFD is approximately 170 ms faster than the BS technique.

References

[1] Z. Luo and C.H. Wu, "A Chinese Character Recognition Interface for Mobile Communication Devices Using Fuzzy Logic and Unit Extraction", Proceedings of the IEEE IECON 22nd International Conference on Industrial Electronics, Control, and Instrumentation, August 1996. 572-577

[2] H. Kang and H.J. Kim, "Design of an Interface on PDA for Korean", IEEE Transactions on Consumer Electronics, Vol.46, Issue: 3, pp.834-838, August 2000.

[3] Yu-Fei Ma, Hong-Jiang Zhang. Detecting Motion Object By Spatio-Temporal Entropy Microsoft Research, China 5F, Sigma Center, 49 Zhi Chun Road Beijing, China.

[4] J. Breitbart, D. Balakrishnan, and A. Ganz, "Pocket-IMPACT Software for Delivering Online Courseware on a PDA: Challenges, Design Guidelines and Implementation", *Proceedings of IEEE Frontiers in Education Conference*, Vol.1, pp. T3F-5, November 2002.

[5] X. Vila, A. Riera, E. Sanchez, M. Lama, D. L. Mureno, "A PDA-based Interface for a Computer Supported Educational System", *Proceedings of the 3rd IEEE International Conference on Advanced Learning Technologies*, pp.12-16, July 2003.

[6] J. Yang, X. Chen, and W. Kunz, "A PDA-based Face Recognition System", Proceedings *of the Sixth IEEE Workshop on Applications of Computer Vision*, pp. 19-23, December 2002.

[7] Sheng-Tun Li, Huang-Chih Hsieh, Ly-Yen Shue, Wen-Shen Chen "PDA Watch for Mobile Surveillance Services". Department of Information Management National Kaohsiung First University of Science Technology 2 Juoyue Rd., Nantz District, Kaohsiung 811, Taiwan.

[8] Gregory A. Baxes. 1996. |Digital Image Processing: Principles and Applications. New York: John willey & Sons.

[9] A. J. Lipton, H. Fujiyoshi, and R.S. Patil. Moving target classification and tracking from real-time video. In Proc. of Workshop Applications of Computer Vision, pages 129–136, 1998.

[10] Jeff Prosise. Programming Windows with MFC. Redmond, Washington: Microsoft Press. 1999.

[11] Richardh.Carverkuo Chungtai. Modern Multithreading. Hoboken, New Jersey .John Wiley & Sons. 2006.