International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 1, Issue 5, May 2014)

FPGA Implementation of Moving Object Detection in Frames by Using Background Subtraction Algorithm for Visual Surveillance

VENU HB, M.tech, VLSI, TOCE, CAHNDRAKALA, Prof, TOCE
Venu.sbr.reddy@gmail.com

***Abstract*—In this paper a new background subtraction algorithm was developed to detect moving objects from a stable system in which visual surveillance plays a major role. Among all existing algorithms it was chosen because of low computational complexity which is the major parameter of time in VLSI. The concept of the background subtraction is to subtract the current image with respect to the reference image and compare it with to the certain threshold values. We have written the algorithm in VHDL, implemented using XILINX ISE 8.1 Design suite and tested in SPARTAN-3 FPGA kit by interfacing a test circuit with the PC using the RS232 cable. The test results are seen to be satisfactory. The area taken and the speed of the algorithm are also evaluated.**

***Index Terms*— Background Subtraction, Video input file, VHDL.**

I. INTRODUCTION

Object detection and tracking is one the most important tasks in computer vision. In video surveillance, it assists understanding the movement patterns of people to uncover suspicious events. It is a key technology in traffic management to estimate flux and congestion statistics. Advanced vehicle control systems depend on the tracking information to keep the vehicle in lane and prevent from collisions. In physical therapy, analyzing the mobility of patients improves the accuracy of their diagnosis. Learning the shopping behavior of customers by tracking assists the architecture design in retail space instrumentation. In robotics, tracking bridges the gap between the raw visual information and environmental awareness [3].

In video summarization, it is applied to generate object-based representations and automatic content annotations. Tracking is also a fundamental technology to extract regions of interest and video object. Even though it is essential to many applications, robust object tracking under uncontrolled conditions still poses a challenge [14]. Real-life systems are required to track objects not only when the background scene is static but also when lighting changes suddenly, camera-object motion becomes large, color contrast becomes low, image noise soars to an unacceptable level, etc. In addition, the computational complexity is required to be kept minimum for real-time performance.

This paper is organized as follows in the section I. introduction to object detection in video surveillance and in the section II. previous algorithms and disadvantages after that in section III.bProposed background modeling IV. Conclusion.

II. REVIEW OF PREVIOUS ALGORITHMS

There are several approaches for moving detection task namely, (a) the optical flow (b) the temporal difference of two (c) background subtraction.

A. Optical Flow Method

In a video frame, the field of motion vector per pixel or sub-pixel is called optical flow. There are many a methods for computing optical flow among which few are partial differential equation based methods, gradient consistency based methods and least squared methods[6]. The objective in optic flow calculation is to find the 2D-motion field in an image sequence. As a pixel at location *(x,y, t)* with intensity *I (x, y, t)* will have moved by $\delta x, \delta y$ and δt between the two frames the following image constraint equation can be given:

$$I(x,y,t) = I(x+\delta x, y+\delta y, t+\delta t)$$

Assuming that the movement is small enough, the image constraint at I(x, y, t) with Taylor series can be derived as

$$(x+\delta x, y+\delta y, t+\delta t) = I(x,y,t) + \frac{\partial I}{\partial x}\delta x + \frac{\partial I}{\partial y}\delta y + \frac{\partial I}{\partial t}\delta t + H.O.T$$

Where H.O.T. means those higher order terms, which are small enough to be ignored.

From these equations follows that

$$\frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y + \frac{\partial I}{\partial t}\Delta t = 0$$

or

$$\frac{\partial I}{\partial x}\frac{\Delta x}{\Delta t} + \frac{\partial I}{\partial x}\frac{\Delta y}{\Delta t} + \frac{\partial I}{\partial x}\frac{\Delta t}{\Delta t} = 0$$

International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 1, Issue 5, May 2014)

This results in:

$$\frac{\partial I}{\partial x} v_x + \frac{\partial I}{\partial y} v_y + \frac{\partial I}{\partial t} = 0$$

The main disadvantage of this Optical flow approach is hard to apply in real-time due to its high computational cost.

B. Temporal Diference Method

The Frame difference is arguably the simplest form of background subtraction. The current frame is simply subtracted from the previous frame, and if the difference in pixel values for a given pixel is greater than a threshold (Th), the pixel is considered part of the foreground. A major flaw of this method is that for objects with uniformly distributed intensity values (such as the side of a car), the interior pixels are interpreted as part of the background. Another problem is that objects must be continuously moving [7]. Let Fi-1 be the previous frame and Fi be the consecutive frame, where i= 1 to n. The psuedocode is given below:

```
For(i=1; i<n; i++)
{
If (Fi-1 - Fi ) > Th
Then accept and process the frame
Else
Reject the frame
}
```

Where this called as Threshold value.

Temporal differencing is very adaptive to dynamic environments, but generally does a poor job of extracting all relevant feature pixels. Due to the disadvantages of this two algorithms we go for background subtraction algorithm.

III. PROPOSED BACKGROUND MODELING

In past years the moving object detection system is designed using software using different algorithms and is system dependent. The past and present moving object detection system like CCTV system which is controlled by software and maintained by server system. If we want to change the setup location, we can't carry the server system everywhere. Cost of the whole system is more. If some technical problems happened in the server system, the captured data will be lost. Our proposed system gives better solution for the above problems.

Background subtraction is a commonly used class of techniques for segmenting out objects of interest in a scene for applications such as surveillance. It compares an observed image with an estimate of the image if it contained no objects of interest. The areas of the image plane where there is a significant difference between the observed and estimated images indicate the location of the objects of interest.

The name "background subtraction" comes from the simple technique of subtracting the observed image from the estimated image and thresholding the result to generate the objects of interest. Here we survey several techniques which are representative of this class, and compare three important attributes of them, how the object areas are distinguished from the background, how the background is maintained over time and how the segmented object areas are post-processed to reject false positives, etc.

In this context, the moving object detection algorithm by background subtraction can be described as shown in Fig 1. Firstly, sequence of image (frames) is generated and registered on black RAM. Then background image is set by user which is subtracted from sequence of frames. At the same time, the current frame can be used for background actualization (using a control bit). Afterward, the resulting image from the subtraction is segmented in order to produce a binary image that highlights the moving regions on the image that belongs to the moving objects.

Once background subtraction is done, comparison the background subtraction with thresholding, then de-noising using morphological operation for objects then perform the tracking to get detected objects. If objects are detected then perform the displacement calculation. If displacement is zero, then shows the static present frame objects and displacement is more than thresholding then detect the abandoned objects or taken objects then displayed on monitoring system.

MORPHOLOGICAL OPERATIONS:

Morphological operations are used to understand the structure and form of an image. This usually means identifying boundaries or objects within an image. Morphological operations play a key role in applications such as automatic and machine vision object detection.

There are three primary morphological functions: *erosion*, *hit-or-miss*, and *dilation*. Morphological operations are usually performed on binary images wuth the pixel values are either 0 or 1. For simplicity, we will show a value of 1 as white and a value of zero as black. While most morphological operations focus on binary images, some also can be applied to grayscale images.
Binary Erosion and Dilation: Erosion is done by doing AND operation between the mask and correspond binary input image values.
Binary erosion uses the following for its mask:

1 1 1
1 1 1
1 1 1

International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 1, Issue 5, May 2014)

We can see that if neighboring pixels have value 1 then the outcome will be 1. Otherwise, the pixel will be 0. Whatever may be value the neighboring pixels have, if the central pixel value is 0 the output pixel value is also 0. Just a single 0 value pixel anywhere in the neighborhood will cause the output pixel to become value 0. Erosion can be used to eliminate unwanted white noise pixels from an otherwise black area. The only condition in which a white pixel will remain white in the output image is if all of its neighbors are white.

The effect on a binary image is to diminish, or erode, the edges of a white area of pixels.

Dilation is the opposite of erosion. Its mask is:

0 0 0
0 0 0
0 0 0

This mask will make white areas grow, or dilate. The same rules that applied to erosion conditions apply to dilation, but the logic is inverted - use the NAND rather than the AND logical operation. Being the opposite of erosion, dilation will allow a black pixel to remain black only if all of its neighbors are black. This operator is useful for removing isolated black pixels from an image.

Outlining:

Other functions can be performed using erosion and dilation as their basic operation. One of these is *outlining*. It is possible to perform a single erosion operation and then subtract the resultant image from the original. The result will be an image that shows a one-pixel outline of all objects. If two erode operators are performed before the subtraction, a two-pixel outline would be created.

Binary Hit-or-Miss Operators:

Two operator masks have been discussed so far, one filled with l's to perform erosion and another filled with 0's to perform dilation. There are other masks that could be useful for other types of conditional processing. For example, the following masks can be used to check to see if a pixel is four-connected to its neighbors:

0 0 0 0 1 0 0 0 0 0 0 0
0 1 1 0 1 0 1 1 0 0 1 0
0 0 0 0 0 0 0 0 0 0 1 0

A similar set of masks can be used to check for eight-connectivity. Bridges, which are defined to be single-pixel connections between groups of similar pixels, can be identified by the following masks:

1 0 1 1 1 1
1 1 1 0 1 0
1 0 1 1 1 1

There also are masks that check for corners or interior pixels or other conditions.

Performing multiple passes on the same image to check for every possible condition of interest can become time consuming. To solve this problem, a concept can be borrowed from the image point operators - look-up tables. Because each pixel in a binary image is either one or zero, it can become a bit that is grouped with other pixels in the neighborhood to form a numerical value.

The neighborhood of 9 binary pixels becomes a 9-bit number that can be used as an index into a look-up table to determine if the output pixel should be a hit or a miss. This table is known as a 9-to-1 LUT since the 9-bit input value results in a 1-bit output value. The table has 512 entries, the number of possible conditions of the 3 x 3 binary pixel neighborhoods.

Obviously, the challenge of using this technique is generating the proper look-up table, because all possible conditions of pixel neighborhoods must be considered. Once this task is completed, however, the resultant processing is much faster.

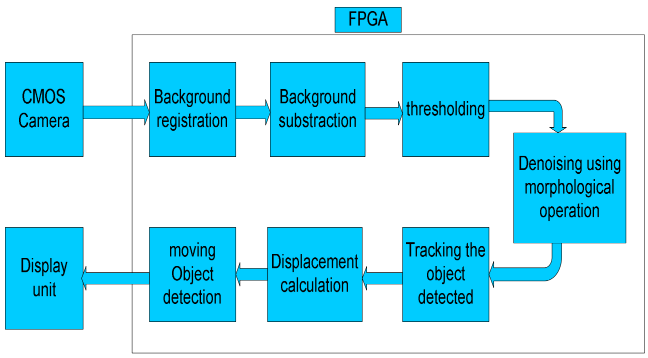


Fig 1: Flow diagram for proposed algorithm

In order to allow high-resolution images of the people in the scene to be acquired it is reasonable to assume that such people move about in the scene. To monitor the scene reliably it is essential that the processing time per frame be as low as possible. Hence it is important that the techniques which are employed are as simple and as efficient as possible.

For that reason the well known technique of background subtraction [15,16] was selected for this application. Background subtraction allows moving objects to be detected by taking the point-by-point absolute difference of the current image and a background image which must be acquired when there are no moving objects in the scene

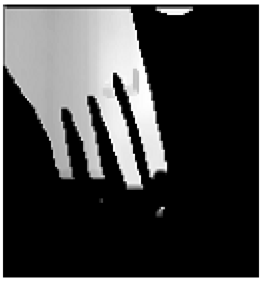
$$Moving_t(i; j) = |Image_t(i; j) - Background(i; j)|$$

International Journal of Ethics in Engineering & Management Education

Website: www.ijeee.in (ISSN: 2348-4748, Volume 1, Issue 5, May 2014)

Background frame.

Upcoming frame.



Background subtracted frame.

Fig: Background subtraction of frames.

When the difference value is greater than the threshold value it is considered as a foreground otherwise it is background object.

In the first step we consider the video and convert it into frames, from that frames we can select any of two images for that two images we create a text file using mat lab tool and then we can write an algorithm for back ground subtraction by using system 'VHDL' language.

IV. CONCLUSION

In this work a moving object detection based on background subtraction algorithm was developed on a reconfigurable hardware. when compared to the other algorithms like temporal difference and optical flow methods our background subtraction gives a better performance in the utilization of memory and logic elements.

REFERENCES

[1]. K. Toyama, J. Krumm, B. Brumitt, and B. Meyers, "Wallflower: principles and practice of background maintenance," in Proc. 7th IEEE Conf. Computer Vision, 1999, vol. 1, pp. 255–261.

[2]. G. Backer, B. Mertsching, and M. Bollmann, "Data- and model-driven gaze control for an active-vision system," IEEE Trans. Pattern Anal

[3]. G.L.Foresti,"A Real Time System for Video Surveillance of Unattended Outdoor Environments".

[4]. J. M. Ferryman, Ed., in Proc. 9th IEEE Int. Workshop on Performance Evaluation of Tracking and Surveillance, 2006.

[5]. 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," Tech. Rep. CMU-RI-TR-00-12, The Robotics Inst., Carnegie Mellon Univ., Pittsburgh, PA, 2000.

[6]. J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Performance of optical flow techniques," Int. J. Comput. Vis., vol. 12, no. 1, pp. 42–77, 1994

[7]. L. Maddalena and A. Petrosino, "A self-organizing approach to detection of moving patterns for real-time applications," in Proc. 2nd Int. Symp. Brain, Vision, and Artificial Intelligence, 2007, pp. 181–190, Lecture Notes Comput. Sci. 4729

[8]. S.-C. Cheung and C. Kamath, "Robust techniques for background subtraction in urban traffic video," in Proc. EIVCIP, 2004, pp. 88 1–892.

[9]. M. Piccardi, "Background subtraction techniques areview," in Proc.IEEE Int. Conf. Systems, Man, Cybernetics, 2004, pp. 3099–3104.

[10]. B. P. L. Lo and S. A. Velastin, "Automatic congestion detection system for underground platforms," in Proc. ISIMP, 2001, pp. 158–161.

[11]. R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: a systematic survey,"IEEE Trans. Image Process., vol. 14, no. 3, pp. 294– 307, Mar. 2005.

[12]. C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: Realtime tracking of the human body," IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 7, pp. 780–785, May 1997.

[13]. A. Elgammal, D. Hanvood, and L. S. Davis, "Nonparametric model for background subtraction," in Proc. ECCV, 2000, pp. 75 1–767

[14]. K. Kim, T. H. Chalidabhongse, D. Harwood, and L. S. Davis, "Real-time foreground-background segmentation using codebook Model," Real-Time Imag., vol. 11, pp. 172–185, 2005.

[15]. G. Backer, B. Mertsching, and M. Bollmann, "Data- and model-driven gaze control for an active-vision system," IEEE Trans. Pattern Anal

[16]. G.L.Foresti,"A Real Time System for Video Surveillance of Unattended Outdoor Environments".

[17]. 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," Tech. Rep. CMU-RI-TR-00-12, The Robotics Inst., Carnegie Mellon Univ., Pittsburgh, PA, 2000.