

# FPGA based Object Detection using Background Subtraction and Variable Threshold Technique

Sridevi N<sup>1</sup>, M Meenakshi<sup>2</sup>

<sup>1</sup>Associate Professor, Department of EIE, Dr. Ambedkar Institute of Technology, Bangalore, Karnataka, India.

<sup>2</sup>Professor, Department of EIE, Dr. Ambedkar Institute of Technology, Bangalore, Karnataka, India.

## Abstract

Visual surveillance plays a vital role in last few years, because of the increasing need for security Object detection is mainly used in surveillance purpose and computer vision systems and is an integral part of any surveillance system. In this paper an efficient FPGA architecture for detection of moving object based on background subtraction algorithm and Haar wavelet transform is proposed. The wavelet transform is mainly used to reduce the storage space without compromising each of the video frames unique characteristics which occurs due to the motion of the objects. The proposed architecture is designed using VHDL language and the simulation is performed by Xilinx ISE tool. The proposed architecture can be able to operate at high speed with less hardware consumption. The main reason for this is that the total architecture is built by basic logic gate and a simplification is done at the gate level which is very efficient. The architecture is validated by inputting the video frames and implementing it on Spartan-6 (XC6SLX45-2csg324) FPGA kit.

**Keywords:** Xilinx, FPGA, Background subtraction, Haar,

## INTRODUCTION

Security is a serious concern in today's modern society. Surveillance is one among various techniques to address the security issue with the help of video camera installed in many places. For further analysis or processing the video, the captured videos must be stored in secured locations, which require huge storage spaces increasing the cost of the total system. Otherwise a person should continuously monitor the video frame at the base station to detect objects in the video which may cause error intentionally or unintentionally. This necessitates the need for developing an algorithm for automatically detecting moving object present in any video frames. The goal of video surveillance is to find the object that is present in the video frame. Moving object detection plays very important role in computer vision, object tracking, surveillance traffic monitoring, video analysis and human action detection [1].

There are three methods available in the literature for detecting the moving objects. The simplest of all such technique is the background subtraction for detecting moving objects [2]. In background subtraction technique the current video frame is subtracted from the static background frame to obtain the additional component that is present in the actual frame. The significance of this technique is the simplicity of

implementation and the suitability for the application where the background is constant.

Now a day's computer vision is an emerging field in the engineering domain due to increase in automations in every field. Normally computer vision is mainly consists of various types of image processing algorithms. Such algorithms are implemented for different applications on software simulation basics. This is because most of the software available in the market are having inbuilt image processing algorithms which reduces the effort of the designer. The implementation of these algorithms will not allow the designer to make any customizations, which is necessary to make the algorithm more suitable for any related applications. Moreover the requirement for the simulation of the software increases in proportional to its complexity and size of the frame.

In the present scenario the minimum resolution of all videos is around 720p, which is very high and hence, it's processing with the developed algorithm takes few milli seconds to few seconds. Therefore its real time implementation is practically impossible, because of the computational time constraints, which necessitates taking an alternate method. To overcome from this kind of problem, a dedicated processing architecture implemented on real hardware instead of virtual simulation is proposed in this work.

Many researchers have proposed various methods for detecting the moving object. In [3] proposed object detection using background subtraction technique. This technique is mainly based on morphological operation. The morphological operation is mainly used to remove noise efficiently and proposed architecture is implemented on real time embedded systems. To remove noise effectively Gaussian filter with 3x3 mask is used which shows good accuracy of detected image in the case of noisy input. Authors in [4] proposed FPGA architecture for detection of motion present in video frames using background subtraction technique. Here the subtraction of the current frame from the background frame which will give the total position in X-Y coordinate is performed. This position values are used to locate the centre of the object in X-Y coordinate system. The design related to [4] techniques is implemented on Spartan-3 FPGA board. Survey on various existing object detection and tracking system for video surveillances is conducted in [5]. In this paper, the explanation of the various problems present in the existing algorithm with implementation in MATLAB and comparison of all techniques is presented. In [6] it is proposed a new technique for object detection where the object can be detected by

separating the background moving object directly from the video frames. This will reduce the need of background frame. The main disadvantage is the complexity of the algorithm, hence reducing the processing speed. This makes the algorithm unsuitable for real time high speed applications. Moving human detection and tracking from real time video sequences is proposed in [7]. The authors of this paper present new technique in MATLAB to remove the shadow present in the video frames for proper detection. This way they can be able to generate improved background image from the video frames which will help to detect the movement very accurately. A review on various types of techniques available for moving object detection is presented in [8]. In this paper the types with advantages and disadvantages of existing moving object detection techniques were discussed. Also the authors of the paper conclude that the background subtraction algorithm alone will not detect the objects, an additional hardware is required for real time implementation of the algorithm. Authors in [9] proposed moving object detection based on local value a change that occurs in the video frames due to the movement of the object. The proposed algorithm calculate the regions of changes inside one image present in the video frame and then separate adaptive threshold value for each corresponding regions. But this process lacks in detection accuracy, which was eliminated with the help of entropy based processing. This algorithm is tested on MATLAB software and the main disadvantage of this algorithm is high architectural complexity which makes it unsuitable for VLSI implementation. Object detection based on contour present in an image is presented in [10]. Optical flow based on gradient calculation is mainly used to develop the algorithm. To increase the accuracy, the authors perform the optical flow based on different contours. Those contours are calculated by normal clustering method. This technique can detect object at high rate which makes it suitable for real time applications. Authors in [11] proposed a new moving object detection which can operate on heterogeneous conditions. The presented algorithm in this paper is implemented using real time embedded system components. To reduce the hardware requirements the whole image is divided into a large number of small sub-parts and at one time the system process only one sub-part. Object detection based on background subtraction and adaptive threshold technique is presented in [13]. Here the background subtraction is used to detect the moving object from video frames, and then the detected image is compressed by applying stationary discrete wavelet transform onto it. Kalman filter is used to trace back the moving object directly in video frames. Authors in [14] proposed a new algorithm for real time object detection. The total algorithm is developed using learning and decision based techniques. Here some initial frames were fed to the system for learning purpose which increases the detection accuracy. But due to the statistical method which is used in this case, designing hardware architecture is very difficult. Searching of object in any video frames is a difficult task. Normally the whole frame is used for searching some predefined objects present in the frame. This needs a large amount of computational time. Also increase in pixel size increases the delay correspondingly. To overcome from the drawbacks discussed above, in this paper the architecture capable of

operating at high speed with less hardware for detection of object using background subtraction and variable threshold technique is proposed.

The paper is organized as follows. Section II describes the block diagram of the proposed architecture. Results and discussion of the proposed approach for the FPGA implementation and comparison with the existing techniques is given in section III, and finally conclusions are drawn in section IV.

## SYSTEM ARCHITECTURE FOR OBJECT DETECTION

The block diagram representation of system architecture for moving object detection is given in fig. 1. The entire block diagram is divided in to eight blocks. That is image conversion, modified median filter, modified two dimensional Haar wavelet transform, variable threshold calculation, inverse transform, background subtraction, thresholding, and intensity mapping. The functionality of each block is explained below.

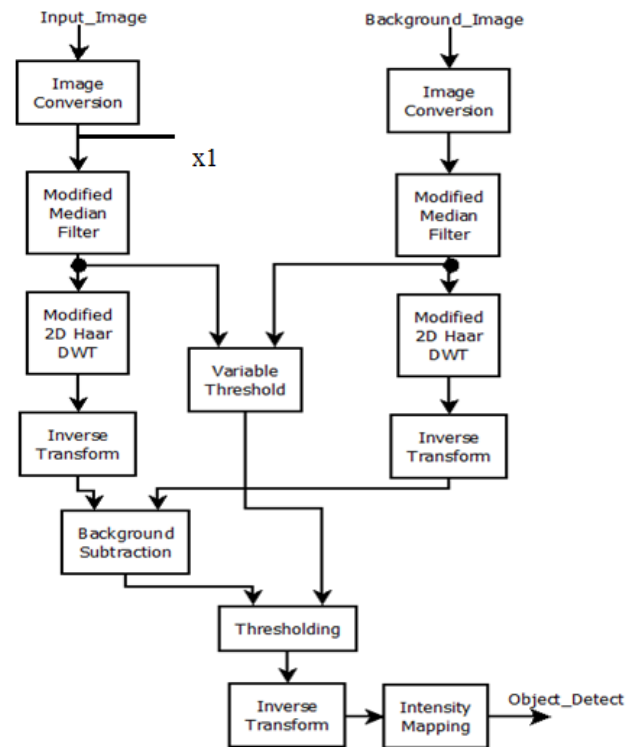


Figure 1: Proposed Object Detection Architecture

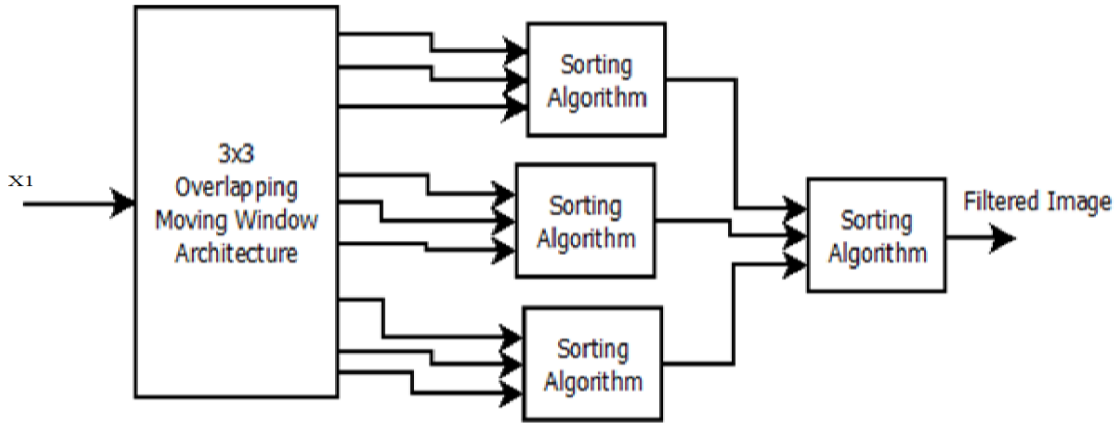
### Image Conversion

The resolution of the images varies from one image to another. The hardware requirement for processing the image also varies. The input and background image are processed by image conversion block independently. In this so as to normalize the different size of the image into 256x256. In addition this block also converts colour image into gray scale image to simplify its hardware implementation.

**Modified Median Filter**

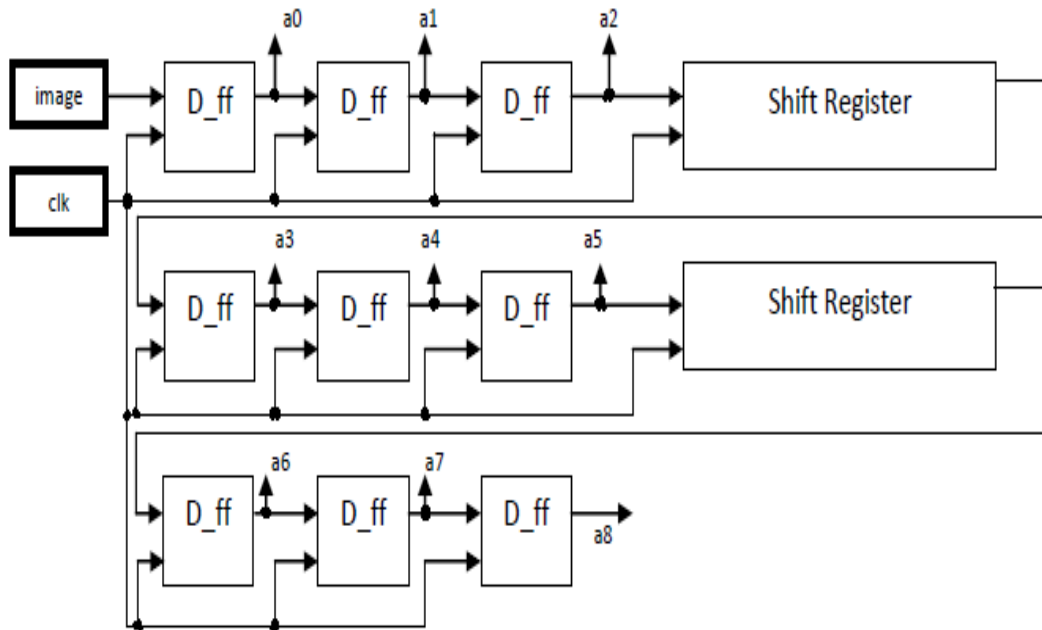
Median filter is mainly used to minimize the salt and pepper noise present in any noisy image. In this method large amount of intensity variations of both actual and background images

introduced similar kind of effect in the detected object. To get optimum filtering 3x3 mask filtering technique is considered. The block diagram of proposed median filter is given in Fig.2 below.



**Figure 2:** Modified Median Filter Architecture

The 3x3 overlapping window selection from an input image of size 256x256 is shown in the Fig. 3 [13].



**Figure 3:** 3x3 Moving Window Architecture [13]

The 3x3 input sub-matrix is generated by the architecture shown in Fig. 3 and are stored into buffer block of separate rows. This will help to find the mid-pixel value at early processing stage by arranging the pixel values in ascending order. Then, finally the accurate mid-pixel values are calculated with the help of D\_FF and Sorting Blocks.

The Sorting block is used to decide the medium values used to smooth the input image to reduce the effect of light variations.

**Modified Two Dimensional Haar Wavelet Transform**

The DWT provides spatial and frequency characteristics of an image. Temporal resolution is an advantage of this method over Fourier transform where both frequency and location information can be obtained. The signal is translated into shifted and scale versions of the mother wavelet to generate

DWT bands. Also image is decomposed into multi-resolution representation. Features are extracted from LL, HH, LH and HL sub bands by applying Haar wavelet transform. The overall information of the original image is obtained from the LL sub band, the diagonal details of the image is given by HH. LH sub band represents horizontal information of the image, and HL gives vertical characteristics of the image. The extraction of LL-Band using Haar DWT is given in fig. 4. For

2x2 input matrices, the wavelet transform can be represented as equation (1) given in [17].

$$y_{LL} = \frac{1}{4} [(a_0 + a_1) + (a_2 + a_3)] \quad (1)$$

Where,  $a_0, a_1, a_2, a_3$ , are 2x2 input image sub-matrix.

The architecture or the implementation of Haar DWT in FPGA is given in fig. 5

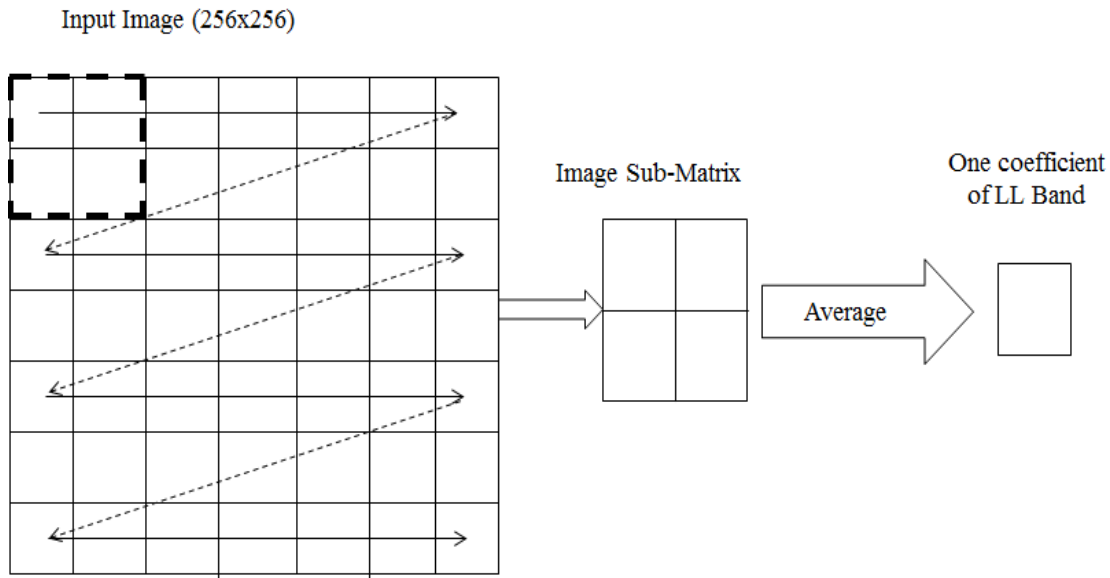


Figure 4: Basics of Haar DWT LL-Band

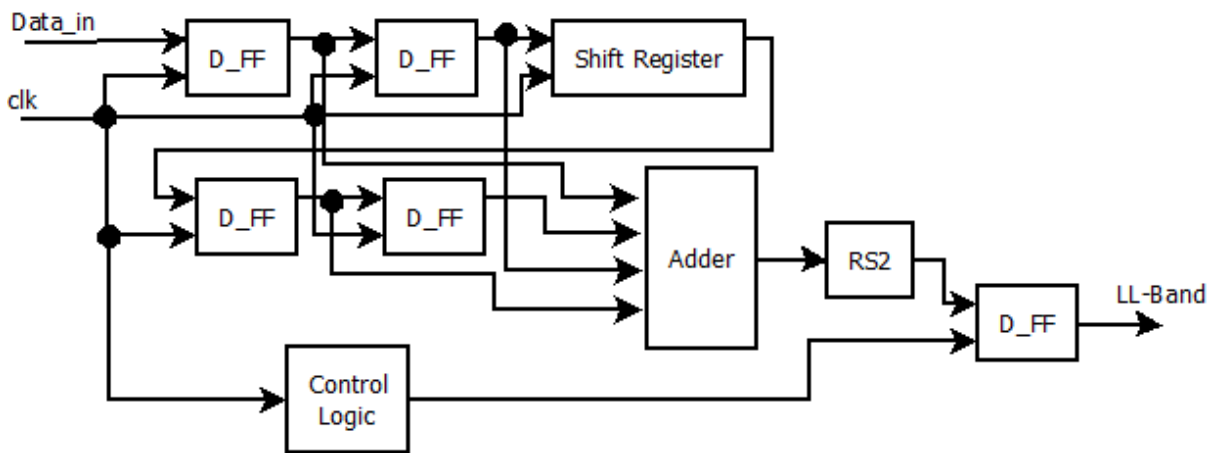


Figure 5: Hardware Structure for 2D-Haar DWT Block

### Variable Threshold Calculation

With the constant threshold value it is difficult to generate accurate threshold value using any mathematical operations which is suitable for all pixels present in image. Therefore variable threshold block is used to calculate variable threshold value which is different for different set of images. The threshold value is considered as a variable of the both input and wavelet transformed image.

### Inverse Transform

The image negative is considered as inverse transform. The linear transformation is negative transformation, which is inversion of identity transformation. In negative transformation, each value of the input image is subtracted from the maximum possible image intensity level and mapped onto the output image. The equation of inverse transform [12] is given in equation (2).

$$y = \{(maximum\ possible\ image\ pixel\ value) - x\} \quad (2)$$

Where, x is input image pixel value.

y is resultant image pixel value.

The maximum possible image pixel value is 255.

Fig. 6 shows the working of inverse transform.

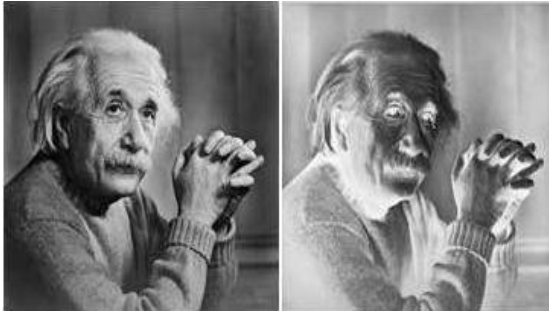


Figure 6: Negative Transform result

### Background Subtraction

The background subtraction block takes the filtered input image and the filtered background image as inputs calculates the distinction of two pictures exploitation constituent by constituent and within the frame continues for all the pixels and the same is repeated for all the frames of the captured video. The changes may be identified when there's transition edge from foreground to background leading to larger distinction. The mathematical expression representing background subtraction [19] is given in equation (3).

$$Subtracted_{Image_{Pixel}} = |Actual\_Image_{Pixel} - Background\_Image_{Pixel}| \quad (3)$$

### Thresholding

The background subtraction cannot alone remove all background pixels due to light intensity variations in the image. To remove extra background pixels thresholding is mainly used and the equation is given in (4)

$$Output_{Pixel} = \begin{cases} Subtracted\_Value; & \text{if } Subtracted\_Value \leq threshold \\ 0; & \text{Otherwise} \end{cases} \quad (4)$$

### Intensity Mapping

In many cases the intensity region of the camera, processing unit differs from the intensity region of many display devices. Many times to get very good clarity the manufacturer of display device increases the intensity scale. As a result if the processed image is directly fed then it may look darker than input image. To solve this problem intensity mapping block is used which convert the processed pixel intensity to the corresponding display device equivalent intensity scale. The equation for conversion is given in (5)

$$Output_{Intensity} = \{f * (Input_{Intensity})\} \quad (5)$$

Where, f is a function which will give the multiplication value for scale to make both equals.

## 1. Results and Discussion

The experimental results obtained by implementing the background subtraction algorithm using FPGA are discussed in detail. The entire system for object detection is presented in this section. Fig 7.a and b. shows the schematic representation of Modified Median Filter and its detailed mapped structure generated by FPGA.

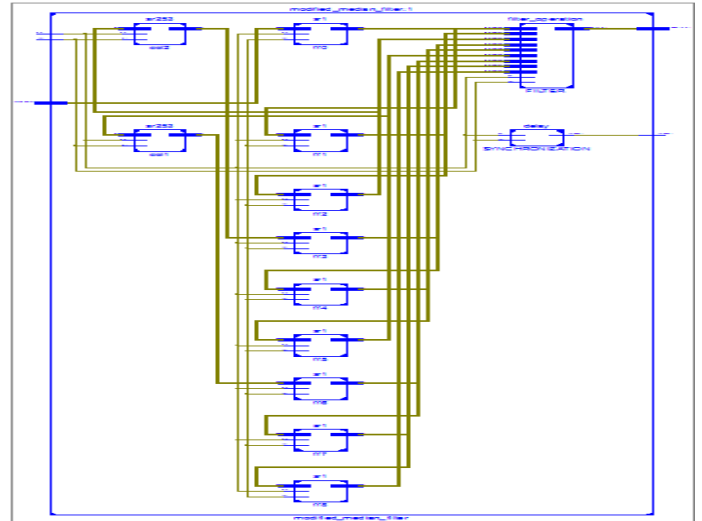


Figure 7.a: RTL Schematic of Modified Median Filter

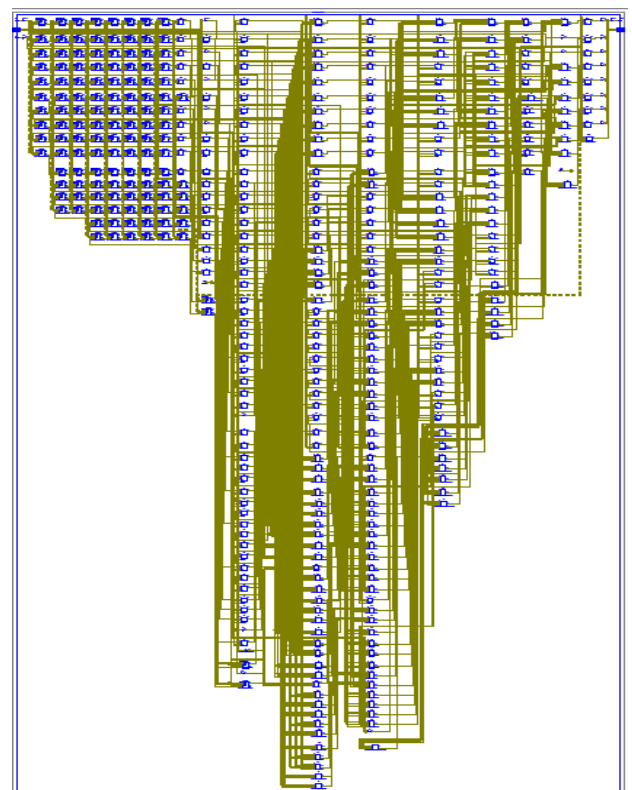
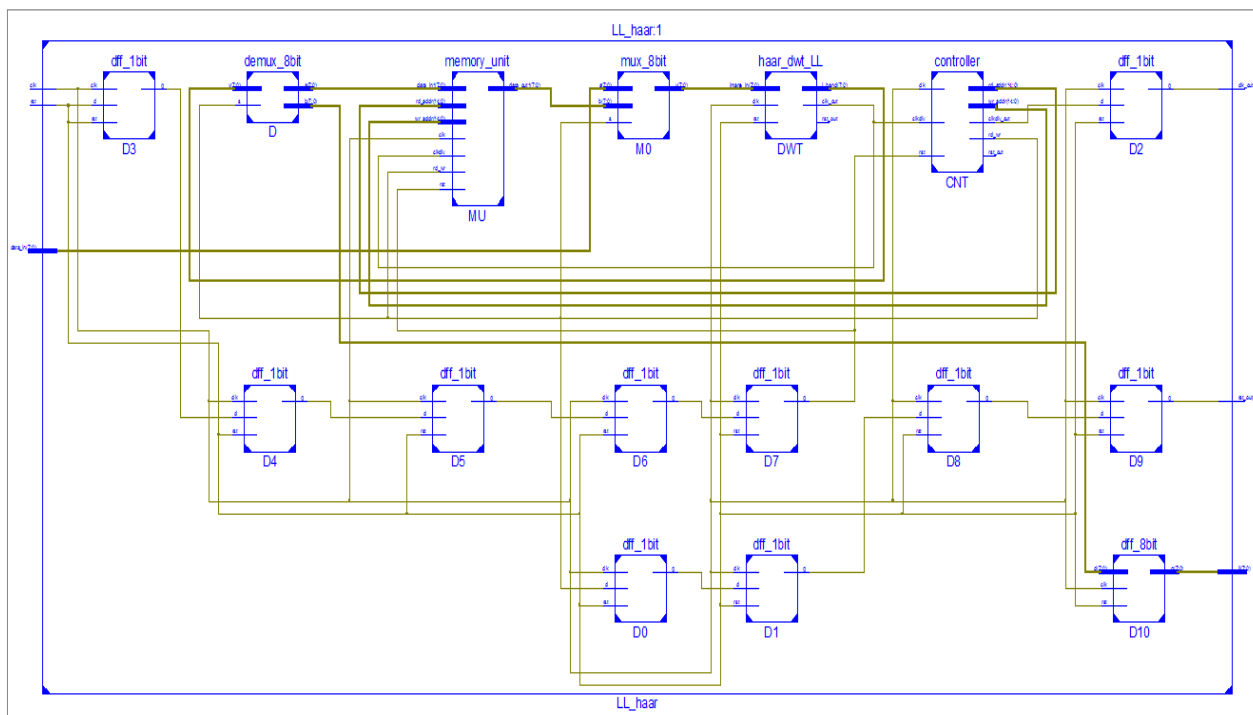


Figure 7.b: Technology Schematic of Modified Median Filter

The schematic representation of Modified Haar DWT and its detailed mapped structure generated by FPGA is given in Fig 8.a and b respectively.



**Figure 8.a:** RTL Schematic of Modified Haar DWT



**Figure 8.b:** Technology Schematic of Modified Haar DWT

Architecture of the object detection and its detailed structure is given in fig9.a and b.

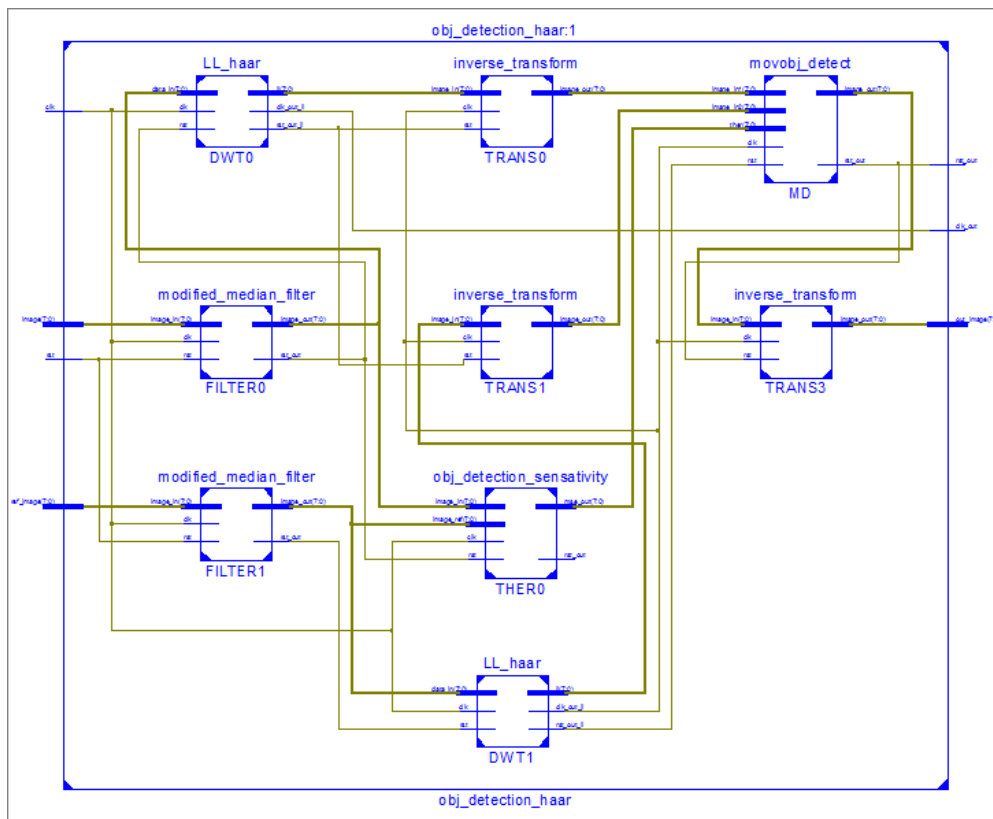


Figure 9.a: RTL Schematic of Total Object Detection Architecture



Figure 9.b: Technology Schematic of Total Object Detection Architecture



**FPGA Implementation**

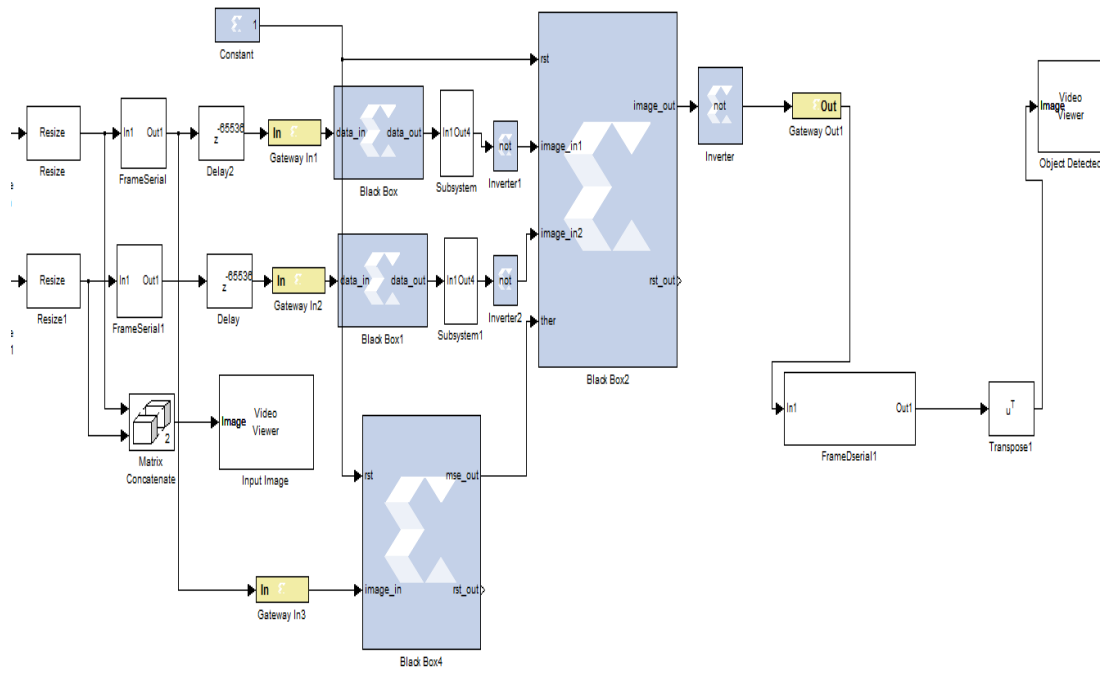
Table 1 gives the hardware deployment of proposed object detection method along with the utilizations of individual

blocks. It is observed that the time requirement for the entire module is more than the individual module and hence the operating frequency of the entire module is reduced due to the synchronization between the modules.

**Table 1:** Utilization of Hardware in Proposed system

Parameters	Modified Median Filter	Background Subtraction	Haar DWT	Inverse Transform	Total Module
<b>FPGA Board</b>	Virtex-5 (XC5VLX110T-3FF1136)				
<b>Number of Slice Registers</b>	146	21	214	0	827
<b>Number of Slice LUTs</b>	273	50	4887	9	10453
<b>Number of LUT-FF Pairs</b>	58	19	199	0	638
<b>Number of BUFG/BUFGCTRLs</b>	2	1	2	1	5
<b>Number of DSP48A1s</b>	0	0	0	0	1
<b>Maximum Operating Frequency (MHz)</b>	349.455	541.448	258.358	----	139.455

The model is implemented on hardware using System Generator Tool. The snapshot of which is given in Fig. 10.



**Figure 10:** Model of Proposed Architecture in System Generator

**Object Detection**

The results of object detection obtained by processing the raw image as mentioned in block diagram given in fig 1 is shown in Fig. 11 (a) to (c). Here, fig 11a,b and c corresponds to three different cases. That is human detection, outdoor surveillance and indoor surveillance.

**Human Detection:** Fig. 11(a) gives the detection of a person by considering a random frame from video sequence. Here background subtraction is used to remove back ground which is captured along with the raw image so as to generate the object accurately.





**Figure 11(a):** Detected Image of a person from a Video Frame

**Outdoor Surveillance:** The detection of a person is given in fig. 11(b) by considering a random frame from outdoor surveillance video sequence. Here the object is generated accurately by implementing background subtraction algorithm

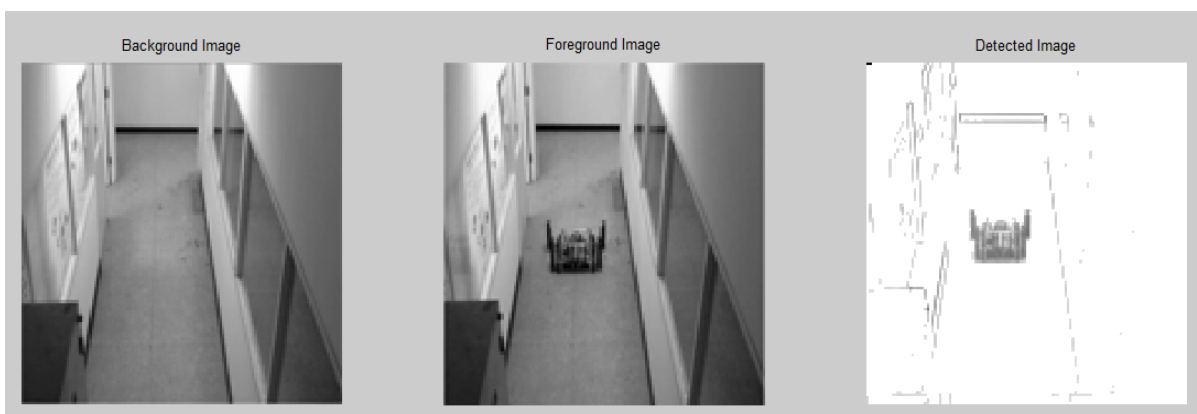
and variable threshold using FPGA architecture. The algorithm removes the background which is captured along with the raw image.



**Figure 11(b):** Detected Image from Outdoor Surveillance Video Frame

**Indoor Surveillance:** Detected image from background by considering a random frame from the indoor video surveillance is given in fig. 11(c). It is observed that the object

is accurately extracted from the background by removing the background which is captured along with the raw image.



**Figure 11(c):** Detected Image from an Indoor Surveillance Video Frame

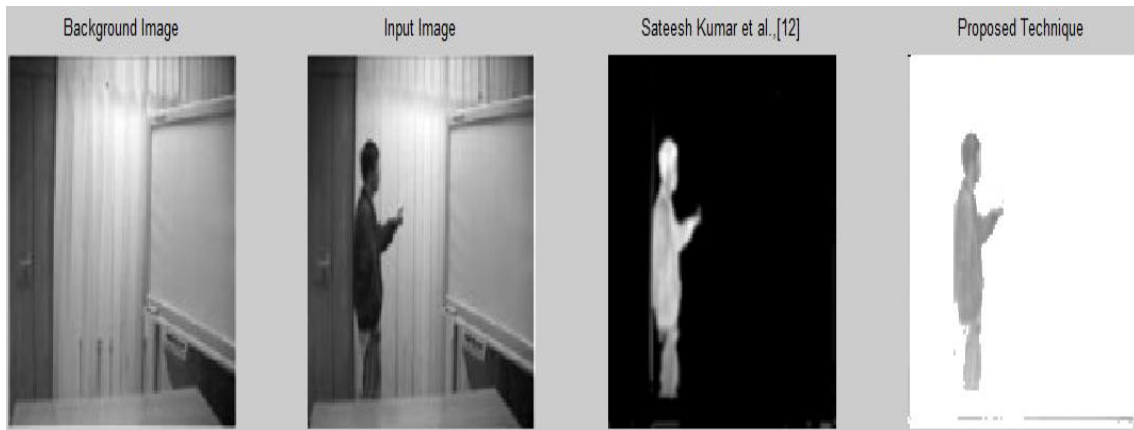
### Comparisons with Existing Techniques

The results obtained by implementing the background subtraction algorithm with variable threshold method is compared with some of the existing techniques from the literature in terms of accuracy of detection, utilization of resources and speed of operation.

### Image Comparisons

The comparison of the detected accuracy between various existing techniques with implementation background subtraction method in terms of detection accuracy is given in fig 12 and 13 respectively.

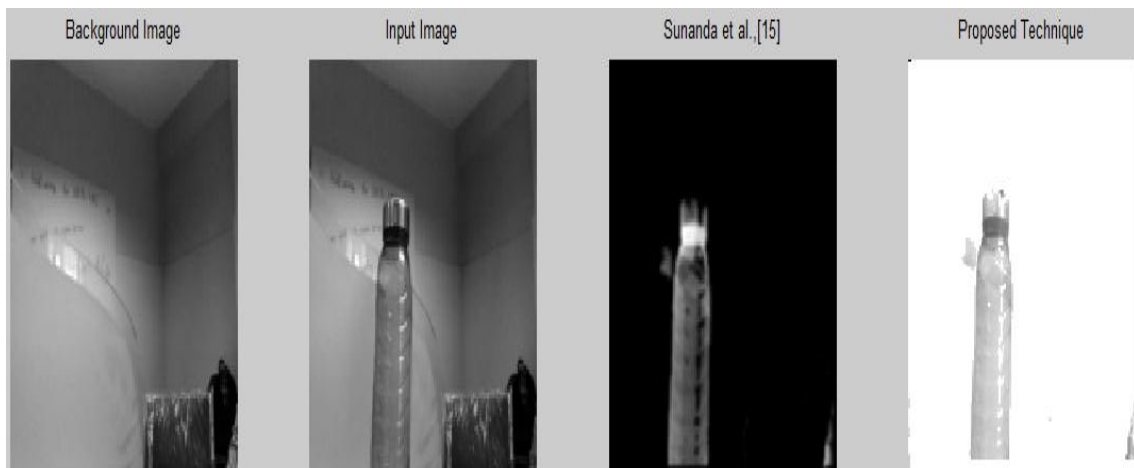
Fig.12 shows the comparison of these results by implementing the background subtraction and variable threshold method with the one obtained from FPGA based modified background subtraction which is explained in paper 12. By observation the proposed method can accurately detect the object from the background.



**Figure 12:** Comparisons with Existing Technique in terms of Detection Accuracy

Similarly, the result obtained from the proposed architecture is compared with the architecture designed using background subtraction method which is presented in paper [15]. By

observation it is clear that the results obtained by implementing background subtraction with variable threshold method accurately extract the object from the background.



**Figure 13:** Comparisons with Existing Technique in terms of Object Recognition

### Hardware Comparisons

To check the effectiveness of the proposed architecture in terms of various hardware requirements such as area and

maximum operating frequency, the background subtraction with variable threshold technique is compared with few of the techniques present in the literature and is given in Table 2.

**Table 2:** Comparisons of Hardware used in Proposed Architecture with Existing Architectures

Sl No	Authors	Board	Number of Slices	Maximum Operating Frequency (MHz)
1	Sateesh Kumar et al.,[16]	Virtex-5	----	118
2	Sunanda Hanchinamani et al., [19]	Spartan-6	1607	--
3	Hung and Yuan [20]	Cyclone-II	1618	113
4	Proposed Method	Virtex-5	827	139

From the table the results clearly shows that the proposed architecture of background subtraction and variable threshold technique is better than the other method in terms of speed and hardware utilization as it operate at 139 MHz and used 827 slices compared with other architecture which is presented in paper [12] and is implemented on Vertex 5 board with 118MHz operating frequency. The architecture proposed in paper [15] based on modified background subtraction method used inbuilt Xilinx block sets hence, they utilized 1607 slices. By implementing background subtraction method with frame buffer scheme presented in paper [20] obtained 113 MHz and 1618 slices. The main reason is the optimization of the architecture which makes it possible to implement the structure using basic logic gates only.

## CONCLUSION

The synthesis results demonstrated the ability of object detection using background subtraction with variable threshold method. This technique is simple and is utilizing lesser area as compared to other methods is presented. The performance comparisons in the existing few techniques indicate the betterment of background subtraction with variable threshold method in terms of LUT utilization, Speed and quality. Thus background subtraction with variable threshold method for object detection can be extended to detect multiple targets presented in any video sequences.

## REFERENCES

- [1] Mubarak Shah, Asaad Hakeem and Arslan Basharat "Detection and Tracking of object from Multiple Airborne Cameras" The International Society for Optical Engineering SPIE
- [2] Juan Shi, Jinling Wang, Yaming Xu " Object Based Change Detection Using Georeferenced UAV Images" International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences , S witzerland 2011
- [3] Mahesh C. Pawaskar, N. S.Narkhede and Saurabh S. Athalye, "Detection of Moving Object Based on Background Subtraction", *International Journal of Emerging Trends and Technology in Computer Science*, Vol. 3, No. 3, pp. 215-218, May 2014.
- [4] Megha Mahesh Chakorkar and M. M. Patil, "Motion Detection by Background Subtraction Algorithm in FPGA", *IOSR Journal of Electronic and Communication Engineering*, Vol. 9, No. 4, pp. 85-88, August 2014.
- [5] Pawan Kumar Mishra and G. P. Saroha, "A Study on Video Surveillance System for Object Detection and Tracking", *3<sup>rd</sup> IEEE International Conference on Computing for Sustainable Global Development*, pp. 221-226, 2016, India.
- [6] Nitya Raviprakash, Meggha Suresh, Asmitha Rathis, Divija Devarala, Aakanksha Yadav and G. S. Nagaraja, "Moving Object Detection Content Based Video Retrieval", *IEEE International Conference on Communication and Signal Processing*, pp. 322-326, 2016, India.
- [7] Ling Li and Wentao Hu, "New Moving Human Detecting and Tracking Algorithm", *3<sup>rd</sup> IEEE international Congress on Image and Signal Processing*, pp. 345-348, 2010.
- [8] Amandeep and Er. Monica Goyal, "Moving Object Detection Techniques", *International Journal of Computer Science and Mobile Computing*, pp. 345-349, Vol. 4, Issue. 9, 2015.
- [9] Anuradha S. G., K. Karibasappa and B. Eswar Reddy, "Video Segmentation for Moving Object Detection using Local Change and Entropy Based Adaptive Window Thresholding", *2<sup>nd</sup> International Conference on Information Technology Convergence and Services*, pp. 155-166, 2013, India.
- [10] Maasayuki Yokoyama and Tomaso Poggio, "A Contour Based Moving Object Detection and Tracking", *2<sup>nd</sup> IEEE International Workshop on Visual Surveillance and Performance Evaluation and Tracking and Surveillance*, pp. 1-6, 2005, China.
- [11] Alessandro Garbo and Stefano Quer, "Moving Object Detection in Heterogeneous Conditions in Embedded Systems", *International Journal of Sensors*, pp. 1-25, 2017.
- [12] Sateesh Kumar H. C., Sayantam Sarkar, Satish S Bhairannawar, Raja K B and Venugopal K R, "FPGA Implementation of Moving Object and Face Detection using Adaptive Threshold" ,*International Journal of VLSI design & Communication Systems*, Vol. 6, No. 5, pp. 15-35, 2015.
- [13] Oussama Boufares, Nouredine Aloui and Adnene Cherif, "Adaptive Threshold for Background Subtraction in moving Object Detection Using

- Stationary Wavelet transform 2D”, *International Journal of Advanced Computer Science and Applications*, Vol. 7, No. 8, pp. 29-36, 2016.
- [14] Paul Viola and Michael Jones, “Real Time Object Detection”, Second International Workshop on Statistical Modeling and Computational Theories, pp. 1-25, 2001, Canada.
- [15] Sunanda R Hanchinamani, Sayantam Sarkar and Satish S Bhairannawar, “Design and Implementation High Speed Background Subtraction Algorithm for Moving Object Detection”, *6<sup>th</sup> International Conference on Advances in Computing and Communication*, pp. 367-374, 2016, India.
- [16] J R R Uijlings, K E A Van de Sande, T Gevers and A W M Smeulders, “Selective Search for Object Recognition”, *IJCV*, pp.1-14.
- [17] Yanwei Pang, Jiale Cao and Xuelong Li, “Learning Sampling Distribution for Efficient Object Detection”, *IEEE Transactions on Cybernetics*, Vol. 47, pp. 117-129, 2017.
- [18] Rohan K Naik, “A Robust Background Subtraction Technique for Object Detection”, *International Journal of Advanced Research in Computer Engineering and Technology*, Vol. 4, pp. 362-366, 2015.
- [19] Yan Lu, Lei Jiang, Ming Dai and Shi Li, “Sort Optimization Algorithm of Median Filtering Based on FPGA”, *IEEE International Conference on Machine Vision and Human-machine Interface*, pp. 250-253, 2010.
- [20] Hung-Yu Chen and Yuan-Kai Wang, “Hardware Design of Moving Object Detection on Reconfigurable System”, *Journal of Computer and Communications*, pp. 1-14, Vol. 4, 2016.
- [21] Satish S Bhairannawar, Anand R, Raja K B and Venugopal K R, “FPGA Implementation of Fingerprint Recognition System using Adaptive Threshold Technique”, *IEEE International Conference on Electrical, Electronics, Signals, Communication and Optimization*, pp. 1-5, January 2015, India.
- [22] Rafael C Gonzalez and Richard E. Woods, “Digital Image Processing”, PHI, 2<sup>nd</sup> Edition, 2005.