International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-6, April 2020

Efficient FPGA Implementation of Human Detection from Video Sequences

Sharath S, Rangaraju H G

Abstract: ***Detection of Human is a vital and difficult task in computer vision applications like a police investigation, vehicle tracking, and human following. Human detection in video stream is very important in public security management. In such security related cases detecting an object in the video, sequences are very important to understand the behavior of moving objects which normally used in the background subtraction technique. The input data is preprocessed using a modified median filter and Haar transform. The region of interest is extracted using a background subtraction algorithm with remaining spikes removed using threshold technique. The proposed architecture is coded using standard VHDL language and performance is checked in the Spartan-6 FPGA board. The comparison result shows that the proposed architecture is better than the existing method in both hardware and image quality.***

Keywords: Adaptive Threshold, Background Subtraction, FPGA Implementation, Human Detection, Modified Median Filter.

I. INTRODUCTION

In today's trendy life the protection is ought in each field, which will increase the need for video processing algorithms. Video-based investigations normally require human supervision to detect the required evidence from any video frames and also it's very useful in several digital image analyses. Moving object detection [1] in video police investigation systems is an important part of security applications. The goal of human detection is to detect the person that is having some motion in a video with same background. The moving object detection technique helps in distinguishing the kind of the objects within the path for artificial intelligence applications that results in more selections on the trail designing. In increased reality, the video information is used in any of the ways on a live platform, which is utilized and accustomed to tracking the interesting objects. This may even be accustomed to counting the number of persons within the crowd or in the interloper region. There exist several techniques for human detection through motion comparison. Background Subtraction method is one of the very important techniques to observe the moving objects. During this methodology the segmentation of objects is completed by exploitation of subtraction. In police video investigations related to the borders across the countries, quicker transmission and fewer information measures becomes crucial issue for information transmission. Background Subtraction helps in dividing the object/person of interest by removing the background that more reduces the information measure information.

This info is more transferred at an awfully quicker rate through closed-circuit television.

Contribution: In this paper, we proposed a novel hardware employment of object/human detection based on background subtraction technique. The video frame data is processed using modified median filter and Haar transform. The region of interest is extracted by means of background subtraction algorithm using adaptive threshold technique.

Organization: The rest of the paper is organized as follows, literature review of existing methods of face recognition is given in section 2. The proposed architecture is deliberated in section 3. The results and performance analysis is discussed in section 4. The final concluding remarks are given in section 5.

II. LITERATURE SURVEY

Mahesh Pawaskar et al., [1] proposed object detection using background subtraction technique. The main technique used in this paper is morphological operation. To remove noise effectively the morphological operation is used and the architecture is implemented on real time embedded systems. In case of noisy input Gaussian filter with 3x3 masks is used which shows good accuracy of detected image. Megha and Patil [2] proposed FPGA architecture for motion detection present in video frames using background subtraction technique. To obtain the total positions in X-Y coordinate the subtraction of the current frames from the background frame is performed. These positions are used to locate the center of the object in X-Y coordinate system. The design was implemented on Spartan-3 FPGA. Pawan and Saraoha [3] conducted survey on various existing tracking system and object detection for video surveillances. In this paper, implementation and comparison of all techniques in MATLAB software are presented with the explanation of the various problems present in the existing algorithm. Nitya et al., [4] proposed a moving object technique from video sequences. In this paper the authors proposed a technique for object detection, separating the background moving object directly from video frames, where the objects are detected, that reduced the needs of background frame. The disadvantage was the algorithm was complex and unsuitable for real time high speed applications.

Revised Manuscript Received on April 30, 2020.
***** Correspondence Author
Sharath.S*, Department of Electronics and Communication Engineering, Government College of Engineering, Krishnarajapete, Mandya, Karnataka, India. Email: ss.sharath@gmail.com
Rangaraju H G, Department of Electronics and Communication Engineering, Govt. Sri Krishnarajendra Silver Jubilee Technological Institute, Bengaluru, Karnataka, India. . Email: rangraju@gmail.com

Ling and Wentao [5] proposed moving human detection and tracking from real time video sequences. In this algorithm authors presented a technique for proper detection to remove the shadow present in the video frames. From the video frames this way they are able to generate improved background image and it will help to detect the movement very accurately. Amandeep and Monica [6] perform a review on various types of techniques available for moving object detection with advantages and disadvantages of existing moving object detection techniques. The authors conclude that only background subtraction algorithm does not need a large amount of extra hardware and for real time hardware processing it makes the implementation of the algorithm inefficient. Anuradha et al., [7] proposed moving object detection based on local value changes occurs in the video frames due to the movement of the object. In the video frame, the proposed algorithm calculates the regions of changes inside one image and for each corresponding regions, it calculates separate adaptive threshold values. This process lacks in detection accuracy and to correct the detection process the author added entropy based processing. The disadvantage of this algorithm is high complexity in the architecture which makes it unsuitable for VLSI implementation.

Masayuki and Tomaso [8] proposed object detection based on contour present in an image. To develop the algorithm Optical flow based on gradient calculation is mainly used. The authors perform the optical flow based on different contours to increase the accuracy. These contours are calculated by normal clustering method. This technique can detect object at high rate and makes it suitable for real time applications. Alessandro and Stefano [9] proposed a new moving object detection which can operate on heterogeneous conditions. Real time embedded system components are used to implement the algorithm. The whole image is divided into a large number of small sub-parts to reduce the hardware requirements. Oussama Boufares et al., [10] proposed object detection based on background subtraction and adaptive threshold technique. To detect the object from video frames background subtraction is used. The stationary discrete wavelet transform is used to compress the image using Kalman filter algorithm, which able to trace back the moving object directly in video frames. Viola and Jones [11] proposed an algorithm for real-time object detection. The algorithm is based on decision and learning-based. For learning purpose, some initial frames must have fed to the system which will increase the detection accuracy. The hardware architecture design is difficult, due to which the statistical methods were used in this case. The whole frame is used for searching some predefined objects present in the frame.

III. PROPOSED ARCHITECTURE

The proposed hardware architecture is used to detect the human from video frame is shown in Figure 1. The Video to frame conversion block is used to convert the video into a finite number of frames. These frames and foreground image are converted separately into a format which is suitable for hardware processing. The high frequency components present in each frame is reduced by the Modified Median Filter block which is then compressed by the Modified Haar Wavelet Transform block by considering LL-Band only. At the same time, the Adaptive thresholding block is used to calculate the appropriate threshold values using adaptive calculation method. The human is detected using the Modified background subtraction block which detects many unwanted objects present in the frame. These unwanted objects are removed by the thresholding block with the help of Adaptive thresholding block. Despite some of noises present in the processed frame will then be removed by the Modified Median Filter block to get proper output whose background is changed by Negative Transformation.

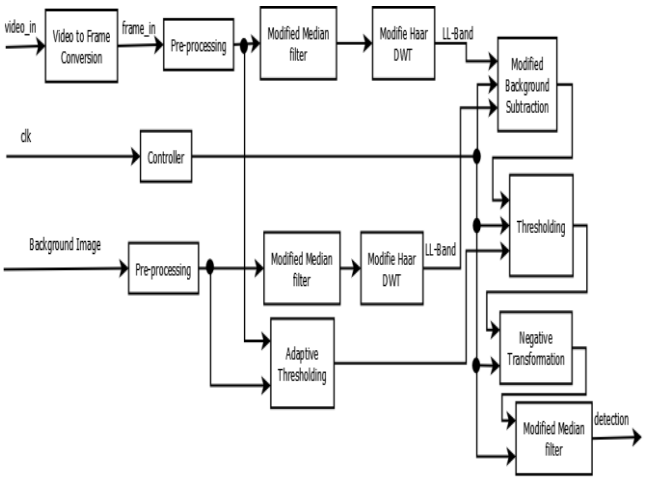

Fig.1. Proposed Architecture for Human Detection

A. Video to frames conversion

Any video file cannot be processed directly and hence it is essential to convert it into a finite number of frames. The video to frame conversion block is used to perform this conversion through the video processing toolboxes present in MATLAB [12] tool. The video to frames conversion is shown in Figure 2 and a sample frames of video is given in Figure3

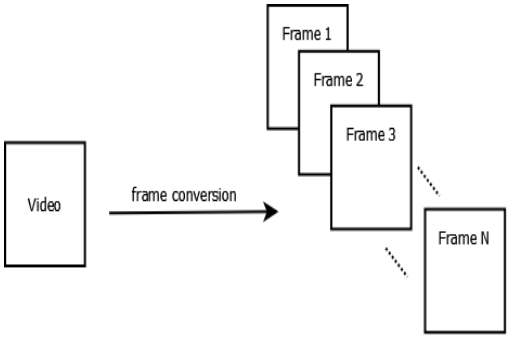

Fig 2: video to frames conversion

Fig 3. Samples of video to frame conversion

B. Pre-Processing

The video frames are not directly suitable for hardware processing. As a result, it is needed to process the frames to make it suitable for hardware processing. In this architecture, the Pre-processing block is used to convert the color frame image into gray format and resize the frame image into a fixed uniform standard size of 256x256.

C. Modified Median Filter

It is mainly used to minimize the salt and pepper noise present in any image. In our cases great amount of light intensity variations of each actual and background images introduced similar effects in the detected object. To reduce this effect, the Median filter is used before feeding the input pictures to any process block. To get optimum filtering, 3x3 mask is used to implement the Modified Median Filter block. The hardware architecture to generate 3x3 overlapping window of input image of size 256x256 is shown within the Figure 4.

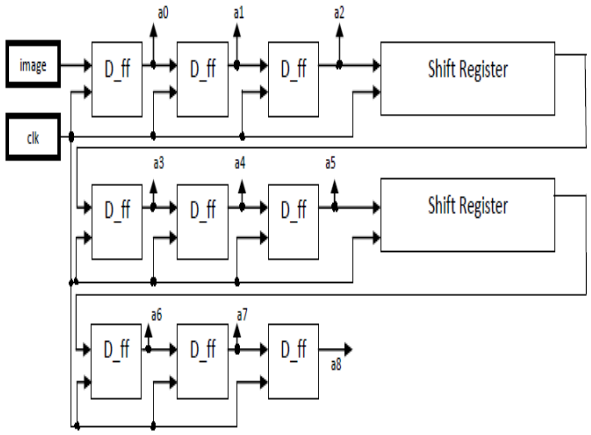

Fig. 4: 3x3 Moving Window Architecture [13]

The pseudo code is used to build the Modified Median filter is given in Table 1 which accepts nine inputs from Moving Window Architecture in overlap manner and generates median value

Modified Haar DWT:

It is a Discrete Wavelet Transform (DWT) and generates four sub-bands namely LL, LH, HL and HH respectively. Among those bands, only LL-Band consists of most of the valuable information of that image and the noises are in the remaining bands. Also the size of each band is half of the original image size in both directions (i.e., horizontal and vertical) respectively. This makes the DWT is suitable for image compression also. In this paper, the DWT is used to compress each frames of any video, which reduces the storage space and processing time. The Haar DWT [12] is the simplest of all wavelet which is used for compression. For this implementation, image matrix is converted into vector and separable Haar DWT is used for which L-Band equation is given in equation as

$$L - Band = \frac{1}{2}(a + b) \quad (1)$$

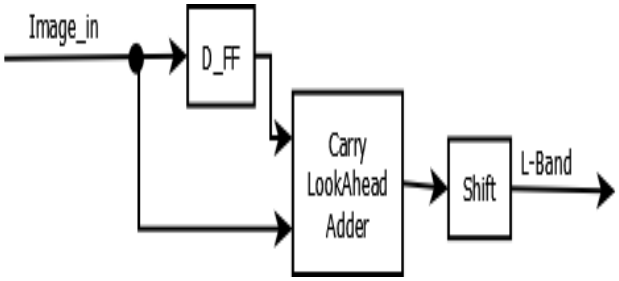

Fig 5: 3x3 Moving Window Architecture [12]

The architecture of 1D-DWT is given in Figure 5, where Carry Lookahead adder and Shifters are used to get optimum hardware architecture. Normally Ripple Carry adder is used in most of the cases for implementing addition operation which wok's fine for smaller number of bits, but for higher bits, the total delay propagation time increases, which reduces overall speed of the architecture.

Table 1: The pseudo code of Modified Median filter

To avoid this problem, Carry LookAhead adder [15] is commonly used, which reduces the propagation time with small area overhead. The equations for Carry LookAhead adder [15] is given in equations 2-5.

$$propogation\ (p_i) = a_i \oplus b_i \quad (2)$$

$$generation\ (g_i) = a_i . b_i \quad (3)$$

The Sum and Carry equations are then

$$Sum_i = p_i \oplus g_i \quad (4)$$

$$c_{i+1} = c_i . p_i + g_i \quad (5)$$

The diagram of 4-bit Carry LookAhead adder is shown in Figure 6.

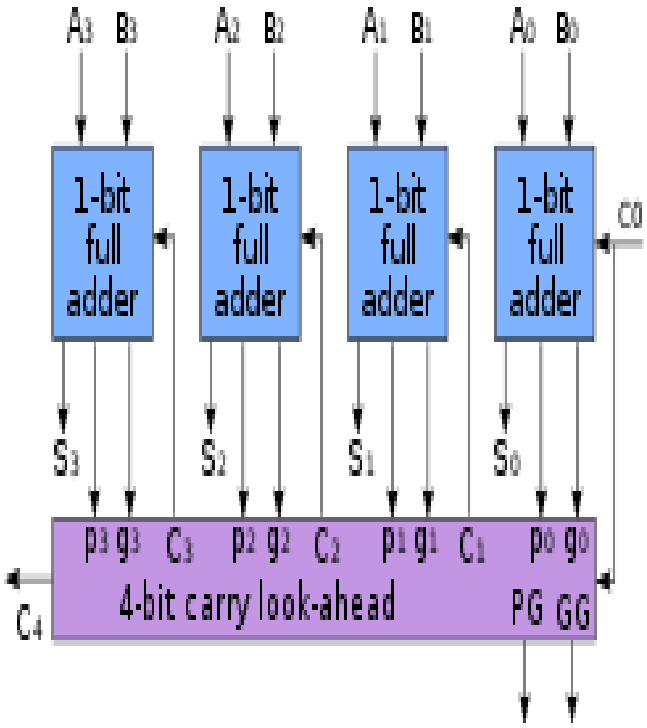

Fig. 6: 3x3 Moving Window Architecture [15]

By using this 1D-DWT, the architecture of 2D-DWT is built [13] which is shown in Figure 7

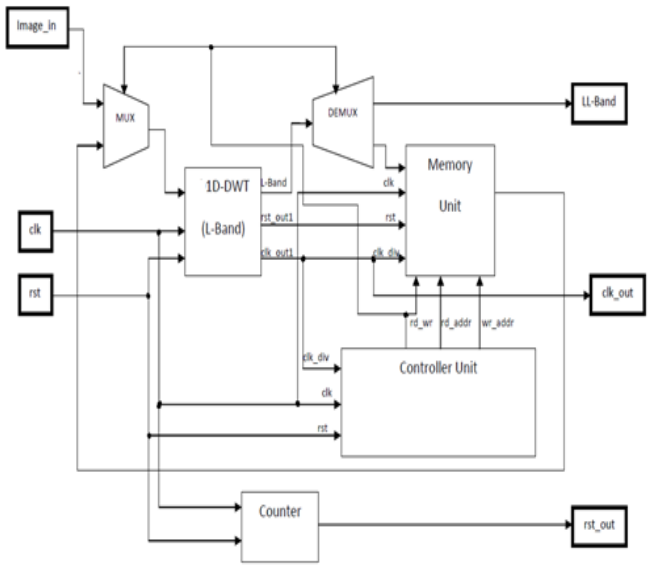

Fig. 7: 2D-DWT Architecture for LL-Band [13]

E. Modified Background Subtraction

The compressed foreground and background images from the output of respective Haar DWT block is fed to the background subtraction block which is used to remove the background objects. This is mainly done by using subtraction operation as given in equation 6

$$Foreground = |Foreground\ Image - Background\ Image| \quad (6)$$

To optimize the hardware architecture, Carry Lookahead adder is used which is given in Figure 8.

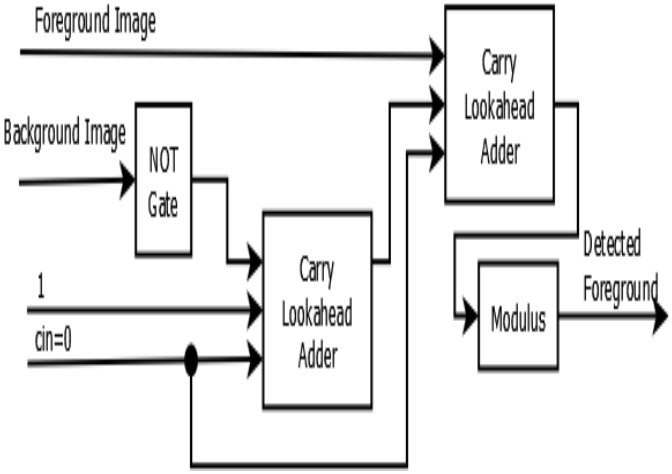

Fig. 8: Modified Background Subtraction Architecture

F. Adaptive Threshold

The Modified Background Subtraction block cannot remove all background components present in the frame. As a result, the thresholding block is used which needs some value to compare and remove the background components. To get better result, it is necessary to calculate the value in dynamic manner instead of static manner. The adaptive threshold block is used to calculate this value depending upon both foregrounds and background frames. The equation of the Adaptive Threshold [16] are given in equations 7 and 8.

Pseudo Code 1: Modified Median Filter

Inputs: a0, a1, a2, a3, a4, a5, a6, a7, a8;
Variables: *value0, value1, value2;*
Output: *mid_value;*

```
{
{
if(a0 > a1 and a0 < a2),   value0 = a0;
else if(a1 > a0 and a1 < a2),   value0 = a1;
else if(a2 > a1 and a2 < a0),   value0 = a2;
else value0 = a2;
end if;
}
{
if(a3 > a4 and a3 < a5),   value1 = a3;
else if(a4 > a3 and a4 < a5),   value2 = a4;
else if(a5 > a4 and a5 < a3),   value1 = a5;
else value1 = a5;
end if;
}
{
if(a6 > a7 and a6 < a8),  value2 = a6;
else if(a7 > a6 and a7 < a8),  value2 = a7;
else if(a8 > a7 and a2 < a8), value2 = a8;
else value2 = a8;
end if;
}
}
{
if(value0 > value1 and value0 < value2)
mid_value = value0;
else if(value1 > value0 and value1 < value2)
mid_value = value1;
else if(value2 > value1 and value2 < value0)
mid_value = value2;
else mid_value = value2;
end if;
}
```

$$WMSE = \frac{(P_{a0}-P_{b0})^2+(P_{a1}-P_{b1})^2+\ldots\ldots+(P_{ax}-P_{bx})^2}{8*N*M} \quad (7)$$

$$WMSE = \sum_{i=0}^{(N*M)-1} \frac{[P_{ai}-P_{bi}]^2}{8*N*M} \quad (8)$$

Where, P_{a0} ,P_{a1}, ……P_{ax} are the background pixels. P_{b0}, P_{b1}, ……P_{bx} are the actual pixels. N and M are the total image dimension in matrix format. The modified architecture [13] is used to generate the adaptive threshold value is given in Figure 9 which consists of Vedic multiplier, carry lookahead adder etc.

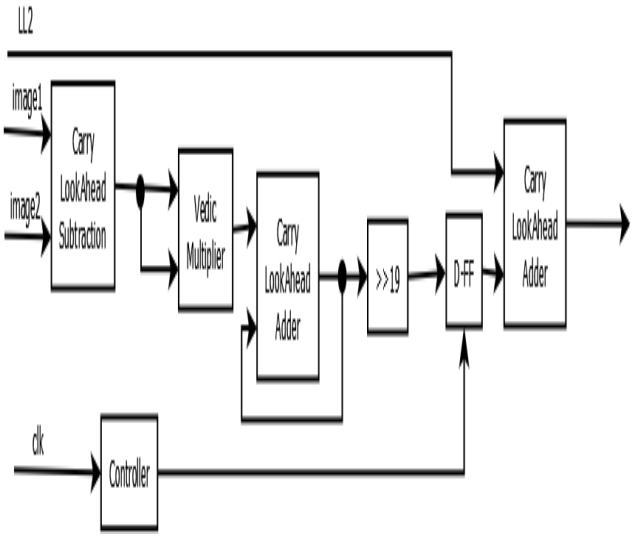

Fig. 9: Modified Adaptive Threshold Calculation Architecture

Normal multipliers are usually requiring large amounts of hardware resources and also slow down the overall design speed. To overcome this problem, one of the multiplier architecture available from ancient Vedic mathematics which is known as Vedic multiplier is used in this paper. Normally Vedic multiplication is divided into *Urdhva Tiryagbhyam Sutra* and *Nikhilam Sutra* multiplication techniques. Among those *Urdhva Tiryagbhyam Sutra* gives optimal hardware utilizations which is widely known as *Urdhva* multiplier [17]. To optimize the hardware utilizations, Carry Lookahead adder architecture is used instead of normal adder which is shown in Figure 10 for 4x4 Modified Vedic Multiplier. The architecture presented in Figure 11 is used to build 8x8 modified Vedic multiplier.

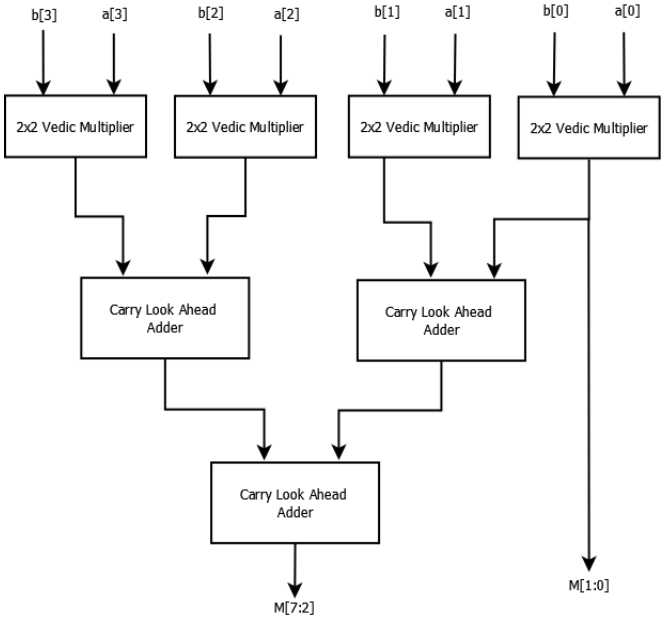

Fig 10. Hardware Architecture for 4x4 Modified Vedic Multiplier

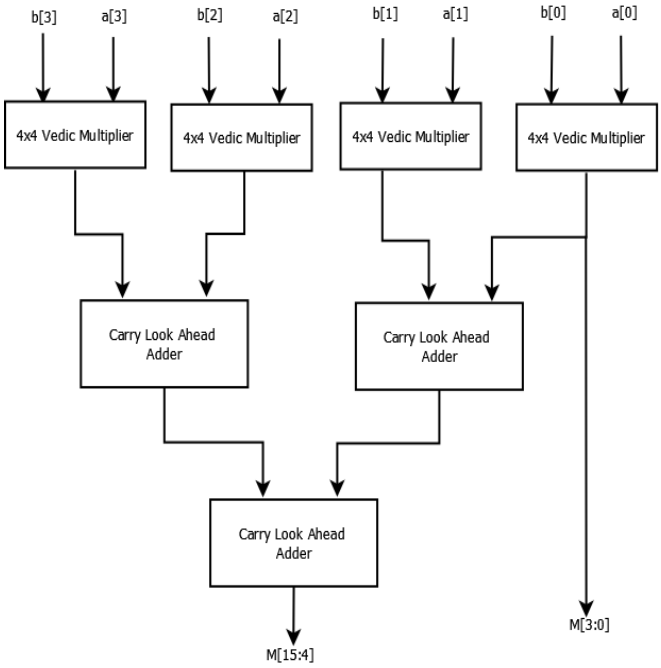

Fig 11. Hardware Architecture for 8x8 Modified Vedic Multiplier

G. Thresholding

The resultant image generated by the Modified Background Subtraction consists of various unwanted distortions which occur due to various reasons such as light intensity variations, camera focusing issues and algorithmic problem etc., which are removed by this block. The mathematical model is given in equation 9 and is implement using normal comparator circuit [14].

$$Modified\ Thresholding = \begin{cases} original\ image\ pixels; \\ 0 \end{cases} \quad (9)$$

H. Negative Transformation

The images are generated by the Modified Thresholding block consists of dark background which makes it difficult for the viewer to recognize the object. As a result, Negative Transformation is used to get the negative of the processed image whose background is white. The equation for this transform [12] is given in 10 and the regular and its corresponding negative image is as shown in Figure 12.

$$Negative\ transform = 255 - Input\ Image\ Pixels \quad (10)$$

Fig. 12: Negative Transform result

IV. RESULTS AND EXPERIMENTAL ANALYSIS

A. FPGA Implementation

The proposed architecture is implemented on Digilent ATLYS FPGA board where standard VHDL coding is used to code the architecture and System Generator tool with the help of compatible XILINX and MATLAB tools. Table-2 shows the hardware utilization of Median filter, Adaptive threshold, Modified Background subtraction and moving object detection with area consumption in terms of hardware parameters such as number of slices, flip-flops, LUTs and DSP48Es.

Table 2. Hardware utilization of proposed method

Parameters	Modified Median filter	Adaptive threshold	Background Subtraction	Proposed Technique
Slices	96	69	25	289
Slice Flip-Flops	17	64	23	65
4 input LUT's	207	19	19	216
DSP48E's	0	0	0	0

The object detection result of proposed method is as shown in the Figure 13. It is clearly seen that; the object/human being is very clearly observed with the proposed background subtraction technique. This can be used to detect any object in an image or video frames effectively.

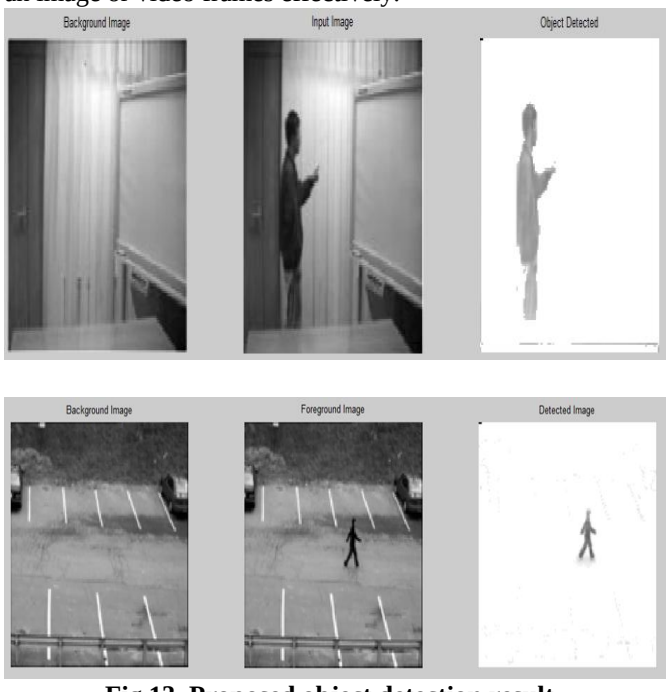

Fig 13. Proposed object detection result

B. Comparison of proposed method results with existing method:

In this section the planned design is compared with existing method to ascertain the performance. The detected image quality of proposed and existing moving object detection technique is shown in Figure 14. It can be seen clearly that the moving object detected by the existing technique [13] is not very clear for proper understanding, whereas in proposed technique, an object visibility and understandability is proper

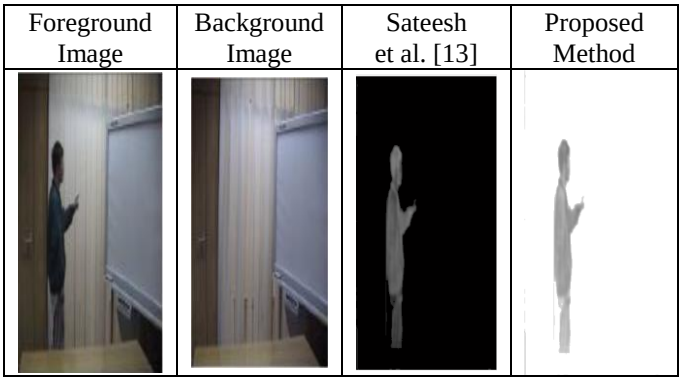

Figure 14. Image Comparison of existing and proposed method

The hardware utilizations of proposed moving object detection architecture is compared with the architectures presented by Sateesh Kumar et al. [13], and Susrutha Babu et al., [19] are shown in Table 3. It is seen that the proposed architecture requires less hardware than the existing technique. It is confirmed that, the proposed technique is superior to the existing technique in terms of hardware utilizations and the quality of a detected object.

Table 3: Hardware Utilization Comparisons of Moving Object Detection

Parameters	Sateesh et al., [13]	Susrutha Babu et al.,[19]	Proposed Method
Slice Registers	365	409	289
Slice Flip-Flops	90	269	65

V. CONCLUSION

In this paper, the novel hardware architecture of human detection using a background subtraction technique is proposed. To detect the human presence in any video frames we use the background subtraction technique, but this technique alone cannot be able to produce good quality of detected images. The filter wavelet and thresholding blocks are used to reduce hardware utilizations. The complex internal architecture is replaced by an equivalent hardware architecture using basic gates, which reduces the total hardware requirement and increase frequency. The comparison of the proposed method results with the existing method validates the superiority. In the future, the proposed method can be implemented using Open CV and appropriate hardware.

REFERENCES

1. 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.
2. 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.
3. Pawan Kumar Mishra and G. P. Saroha, "A Study on Video Surveillance System for Object Detection and Tracking", 3rd IEEE International Conference on Computing for Sustainable Global Development, pp. 221-226, 2016.
4. Nitya Raviprakash, Meggha Suresh, Asmitha Rathis, Divija Devarala, Aakanksha Yadav and G. S. Nagaraja, "Moving Object Detection Content Based Video Retrival",

IEEE International Conference on Communication and Signal Processing, pp. 322-326, 2016.
5. Ling Li and Wentao Hu, "New Moving Human Detecting and Tracking Algorithm", IEEE international Congress on Image and Signal Processing, pp. 345-348, 2010.
6. Amandeep and Er. Monica Goyal, "Moving Object Detection Techniques", International Journal of Computer Science and Mobile Computing, Vol. 4, Issue. 9, pp. 345-349, 2015.
7. Anuradha S. G., K. Karibasappa and B. Eswar Reddy, "Video Segmentation for Moving Object Detection using Local Change and Entropy Based Adaptive Window Thresholding", International Conference on Information Technology Convergence and Services, pp. 155-166, 2013.
8. Maasayuki Yokoyama and Tomaso Poggio, "A Contour Based Moving Object Detection and Tracking", 2nd IEEE International Workshop on Visual Surveillance and Performance Evaluation and Tracking and Surveillance, pp. 1-6, 2005.
9. Alessandro Garbo and Stefano Quer, "Moving Object Detection in Heterogeneous Conditions in Embedded Systems", International Journal of Sensors, pp. 1-25, 2017.
10. Oussama Boufares, Noureddine 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.
11. Paul Viola and Michael Jones, "Real Time Object Detection", Second International Workshop on Statistical Modeling and Computational Theories, pp. 1-25, 2001, Canada.
12. Rafael C. Gonzalez, Richard E. Woods and Steven L. Eddins, "Digital Image Processing Using MATLAB", Gatesmark Publishing, 2nd Edition, 2009.
13. 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 and Communication Systems, AIRCC, Vol. 6, No. 5, pp. 15-35, October 2015.
14. Charls H. Roth (Jr.), "Fundamentals of Logic Design", Jaico Publishing House, First Edition, 1992.
15. Jia Miao and Shuguo Li, "A Novel Implementation of 4-bit Carry Look-Ahead Adder", IEEE International Conference on Electron Devices and Solid-State Circuits, pp. 1-5, 2017.
16. Sharath S and Rangaraju H G, "FPGA based Human Detection using Background Subtraction", International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques ICEECCOT-2018
17. Siba Kumar Panda, Ritisnigdha Das, S K Saifur Raheman and Tapasa Ranjan Sahoo, "VLSI Implementation of Vedic Multiplier using Urdhva–Tiryakbhyam Sutra in VHDL Environment: A Novelty", IOSR Journal of VLSI and Signal Processing, Vol. 5, Issue. 1, pp. 17-24, 2015.
18. Vrushali and Charudutta, "FPGA based Moving Object Detection", IEEE International Conference on Computer Communication and Informatics, 2014.
19. S. Susrytha Babu, S. Suparshya Babu, Habibulla Khan and M. Kalpana Chowdary, "Implementation of Running Average Background Subtraction Algorithm in FPGA for Image Processing Applications", International Journal of Computer Applications, Vol. 73, No. 21, pp. 41-46, 2013.

AUTHORS PROFILE

Sharath S is currently working as Assistant Professor in Electronics and Communication Engineering Department at Government College of Engineering Krishnarajapete, Mandya, Karnataka, India. Received B. E. degree in Electronics and Communication Engineering from APSCE, Bengaluru, M.Tech degree in VLSI Design and Embedded Systems from RVCE, Bengaluru. He has more than 12 years of experience in Teaching and Industry. His major areas of interest include VLSI Design, Image Processing, Communications and Wireless Network

Dr. Rangaraju H G is currently working as Associate Professor and Head of Electronics and Communication Engineering Department at Govt. Sri Krishnarajendra Silver Jubilee Technological Institute, Bengaluru, Karnataka, India.. Received B. E. degree in Electronics and Communication Engineering from SIT, Tumkur, M. E degree in Electronics and Communication Engineering from University Visvesvaraya College of Engineering, Bengaluru university and Ph. D., degree from Visvesvaraya Technological University, Belagavi. He has more than 20 years of experience in Teaching and Industry. He has to his credit more than 20 research publications in National/International Journals and Conferences. Now, He is guiding six research scholars for their PhD degrees under VTU. His major areas of interest include VLSI Design, Signal Processing, Communications and Wireless Networks.