Performance Analysis and Augmentation of K-means Clustering, based approach for Human Detection in Videos

Mr. Pruthviraj Parmar, Mrs. Gayatri Sunilkumar Pandi(Jain)
ME, Information Technology, PG Department and PG Co-ordinate
LJIET, Ahmedabad, India

Abstract **- "Performance Analysis and Augmentation of K-means Clustering, based approach for Human Detection in Videos" takes us to video level processing techniques to identify human from video. Surveillance System have become increasingly popular in the globalization process. Intelligent Video Surveillance system based on image recognition is widely used to effectively help to provide security, safety and prevents many crimes. Due to the high Complexity in techniques such as real-time processing and image contents analysis/understanding. However detecting humans in images and videos still challenging task owing to their variable appearance caused by variety of clothes shadows, articulation and illumination situations, and unpredictable poses that they can adopt. In this thesis there is a brief survey of different object detection techniques, as well as human detection techniques like fuzzy logic, single Gaussian model, mixture of Gaussian model, Background Subtraction Technique, human body Pose Recognition ,appearance based ,motion-based and hybrid methods. Hear Adaptive Background subtraction method and K-means clustering used for detection of the human .It works efficient for the indoor environment. There is near about 67-70% human detection from the video frame for the indoor. For the outdoor there are so many parameters which effect the human detection, so it is not efficient.**

Keywords **- Human Detection, Background Subtraction, Human Detection technique-means clustering, skin detection**

I. INTRODUCTION

Image processing is a technique to convert an image into digital form in order to get an enhanced image or to extract some useful information from images and videos. Image processing is a type of signal dispensation in which input is image, like video frame or photograph and output may be image or characteristics associated with that images. Usually Image Processing system includes treating images or video as two dimensional signals while applying already set signal processing methods to them. Now a days it's become very important technologies, with its applications in various aspects of a business. Image Processing forms core research area within engineering and computer science disciplines too for extracting information from the images or videos. Intelligent Surveillance Systems' requires fast, robust and reliable algorithms for object detection and tracking.

The proposed definition aims to achieve the performance of the smart surveillance system to detect human in a prohibited area and automatically generate an alarm which will enable the human operators to take action quickly. For this work various method are analyzed to improve the performance.

II. LITERATURE SURVEY

1. Human Detection using Background Subtraction Techniques.

Identifying moving objects from a video sequence is a fundamental and critical task in many computer-vision applications. A common method is to perform background subtraction, using Background Subtraction this technique [1] identifies moving objects from the portion of a video frame that differs significantly from a background model. There are following challenges in developing a good background subtraction algorithm.

- It must be robust against changes in lighting.
- It would avoid detecting non-stationary background objects such as moving leaves, rain, snow, and shadows cast by moving objects.
- Its internal background model would react quickly to changes in background such as starting and stopping of vehicles.
- For performing background subtraction, there are different approaches to this basic scheme of background subtraction in terms of foreground post processing, region detection and background maintenance.
  Here some of the main techniques are surveyed for deciding that which method will be best suitable for Human Detection.

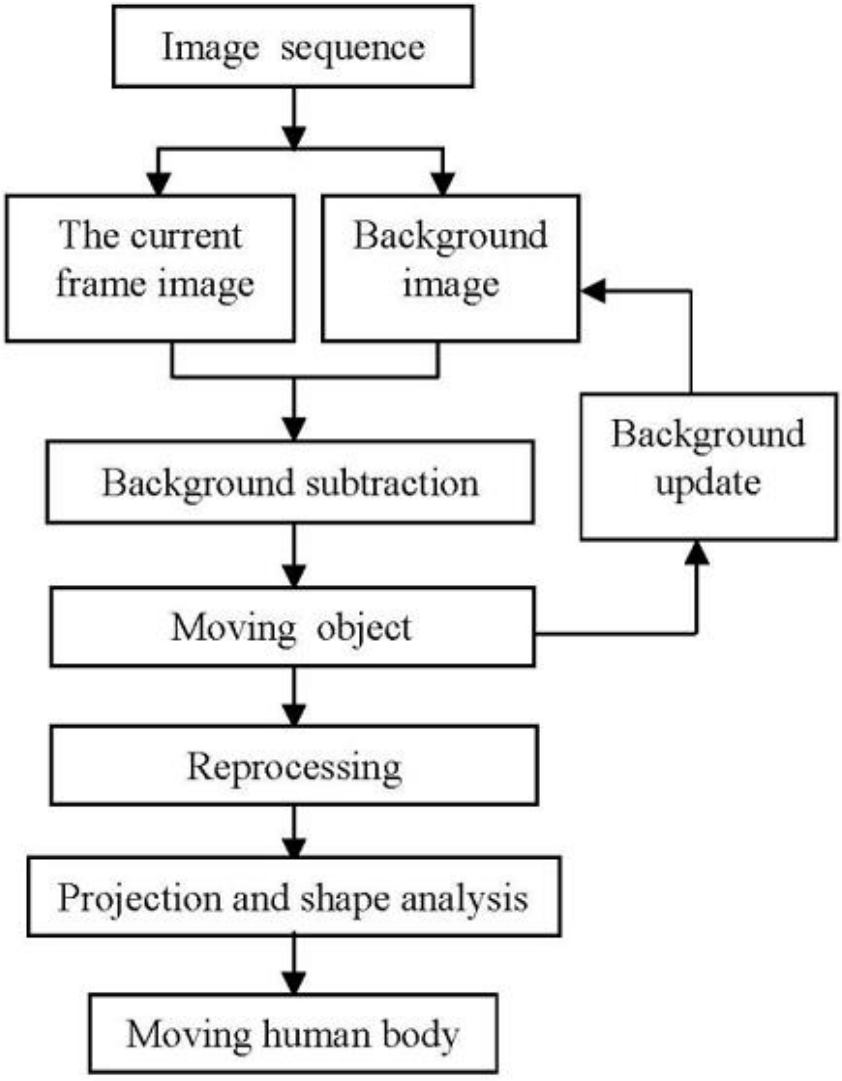

Figure 1: the flow chart of moving human body extraction [1]

2. Human detection Algorithm using Gaussian Mixture model (GMM).

The background of the scene contains many non-static objects such as tree branches whose movement depends on the wind in the scene. Such background motion causes the pixel intensity values to vary significantly with time. Therefore, a single Gaussian assumption for the pdf of the pixel intensity won't hold. Instead, a generalization based on a Gaussian Mixture Model(GMM)[2][3] has been used to model such variations. The pixel intensity was modeled by a mixture of K Gaussian distributions (K is a small number from 3 to 5). A mixture of three Gaussian distributions was used to model the pixel value for traffic surveillance applications, corresponding to road, shadow, and vehicle distribution. Although, in this case, the pixel intensity is modeled with three distributions, still unimodal distribution assumption is used for the scene background, i.e. the road distribution. Mixture of Gaussians is tremendous popular since it was first proposed for background modeling. The generalized Mixture of Gaussians has been used to model complex, non-static backgrounds. However, the Mixture of Gaussians (MoG) has its own drawbacks. First, it parameters require careful tuning and its computationally intensive. Second, it is very sensitive to sudden changes in global illumination. If a scene remains stationary for a long period of time, the variances of the background components may become very small. A sudden change in global illumination can then turn the entire frame into foreground. Backgrounds having fast variations cannot be modeled with just a few Gaussians accurately.

3. Human detection in video using Fuzzy Approach.

Detection of the human is based on the skin of the human, skin detection is the preliminary step to a wide variety application such as personal identification and video surveillance system using fuzzy logic [4] .In this method human can be detected from the surveillance video by skin region from the selected frame.in this paper there are step for detecting the human first by using histogram differences shot detection process is done. Than the process to select the key frame based on average pixel value of select the key frame based on average pixel of every shot .than in next stage fuzzy logic rule is applied for the skin color identification.

- **Shot Detection using Histogram Difference :-**
  HSV color model has a good liner Scalability where the color corresponds to the weight value proportional to the Euclid distance .HSV color space are independent and gives best result than RGB.
  The following figure 3.3.2 shows flowchart of shot detection process:-

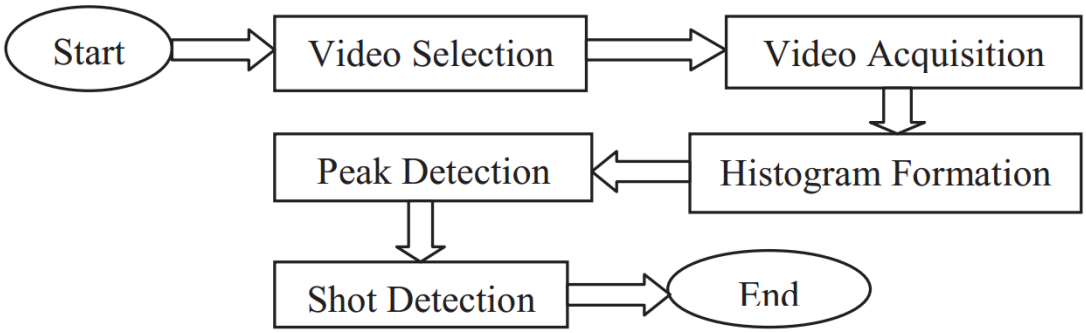

Figure 2:- Flowchart of shot detection [4]

All the frames lying between the two transition frames are considers as single shot. Different gradual transition methods proposed by different researchers are fade in/fade out, wipe detection and dissolve detection.

- **Frame selection:-**

With the reference to figure 3.3.3 frames are selected by the mean value of color pixels are calculated the mean value of gray level pixels can be calculated the mean value of gray level and R,B,G level and then find the frame which has the closed value to the mean.

The following figure 3.3.3 shows flowchart of Frame Selection process:-

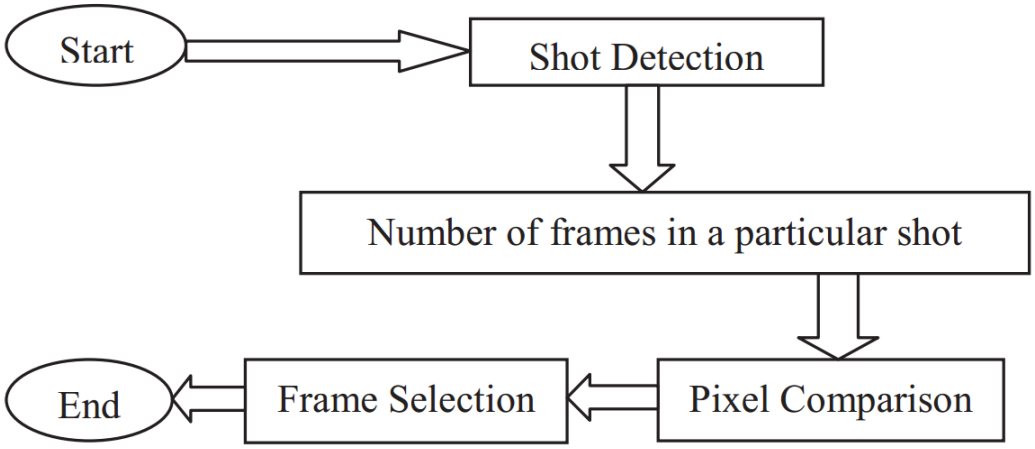

Figure 3:- Flowchart of Frame Selection [4]

- **Human skin detection using fuzzy logic:-**

The color of human detection is totally different from the color of many other objects and for this reason the statistical measurements of this area is very important for face detection gesture recognition and personal identification. The fuzzy rules are characterized by a collection of different fuzzy membership function, various types of logical operations logic is using the image luminance of the image intensity.

The following figure 3.3.4 shows flowchart of Skin Detection process:-

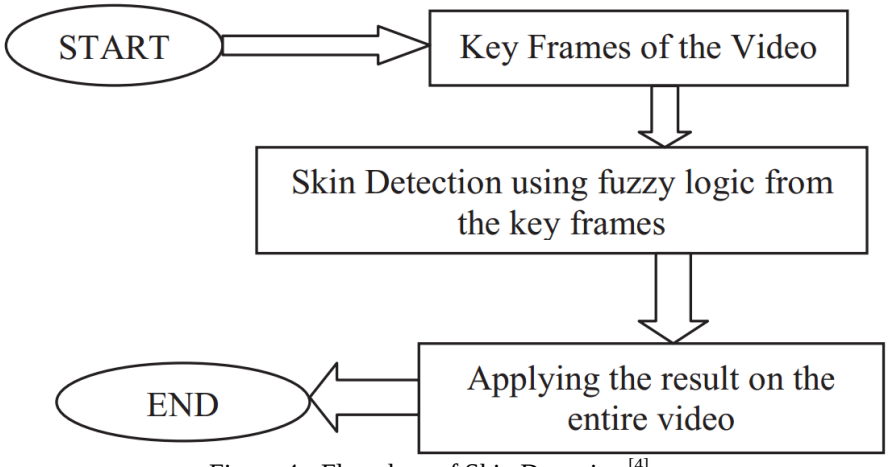

Figure 4:- Flowchart of Skin Detection [4]

In this methods system has used to histogram based algorithm to detect shots, which has been compare to other for better results, after the frame section skin regions are Detected using the fuzzy membership function in HSV plane.

4. A Quantitative & Qualitative Comparison of Real-time Background Subtraction Algorithms for Video Surveillance Applications

This basic Gaussian model can adapt to slow changes in the scene like, gradual illumination changes by recursively updating the model using a simple adaptive filter. The main feature of modeling [5] the probability distribution of the pixel intensity that differentiates it from other ways such as predictive filters is that it ignores the order in which observations are made and focuses on the distribution of the pixel intensities. Then they use a clustering method to fit the data with an approximation of a mixture of Gaussians. At each pixel, one of the clusters (Gaussians) is selected as the background process, the others are considered to be caused by foreground processes. They are working on extensions which will allow dynamic background estimation based on the previous N frames.

5. Human Behavior Understanding

In Human Behavior Understanding [6] there are some detection technique based on appearance based, motion based, and hybrid methods.

Appearance based human detection: - In this method there are large pedestrian databases, this algorithm scan frame and searching the pattern which match with our pedestrian databases. This method can be directly applied to the Non-static cameras. Using Histograms of Oriented Gradients (HOG) the shape of the human can be represented by edge direction and intensity gradients, usually combined with Support Vector Machine (SVM). Appearance based people detection have rich database with videos and discuss evolution measures for perform training comparison.

Motion based Human detection:-In this method, it detects the cyclic motion of the legs and assume a static camera to identify the moving foreground. Discrete Fourier Transform(DFT) used to quantify pixel oscillation and variation of the method analyses the power spectral similarity in the walking patterns or the amount of changes in motion history image another algorithms are ARMA(Auto Regressive Moving Average Models) which used for theoretical performance bounds and MUSIC(Multiple Signal Classification) which used for frequency estimation .People detection can me improve by Considering scene modelling that helps reducing the search space. Knowledge of the scene can be used train specific area. But when the homograph between the ground and the camera is known size features can used for the detection.

Hybrid Methods:-this method is the combination of the Appearance based detection method and motion based detection. These two methods can be individually analyses the data, merging the final result according to a given function, or apply a still image detector to regions potentially containing pedestrians as indicated by object tracker. The various methods for implementing this methods are a Viola and Jones Detector ,2D head contour,3D head contour.by this methods detection of human was 85% in ranges up 8 m achieved .Data sets which used for the human detection are: MIT,CALTECH,INRIA,PETS,ETH used for this technique.

III. Proposed Work

From the literature survey it is clear that, Due to Gaussian Mixture Mode (Background subtraction) human detection was not done if there is small changes in background this techniques is very sensitive to sudden changes in global illumination, false alarm of human being detection is occur and in the fuzzy logic if there is low light or high light it cannot work effectively.

From the literature survey, all the techniques are having some problems to detect the human effectively the system can improve the results by using the combination of two models for effect human detection

1. K-means Clustering.
2. Adaptive Background Method

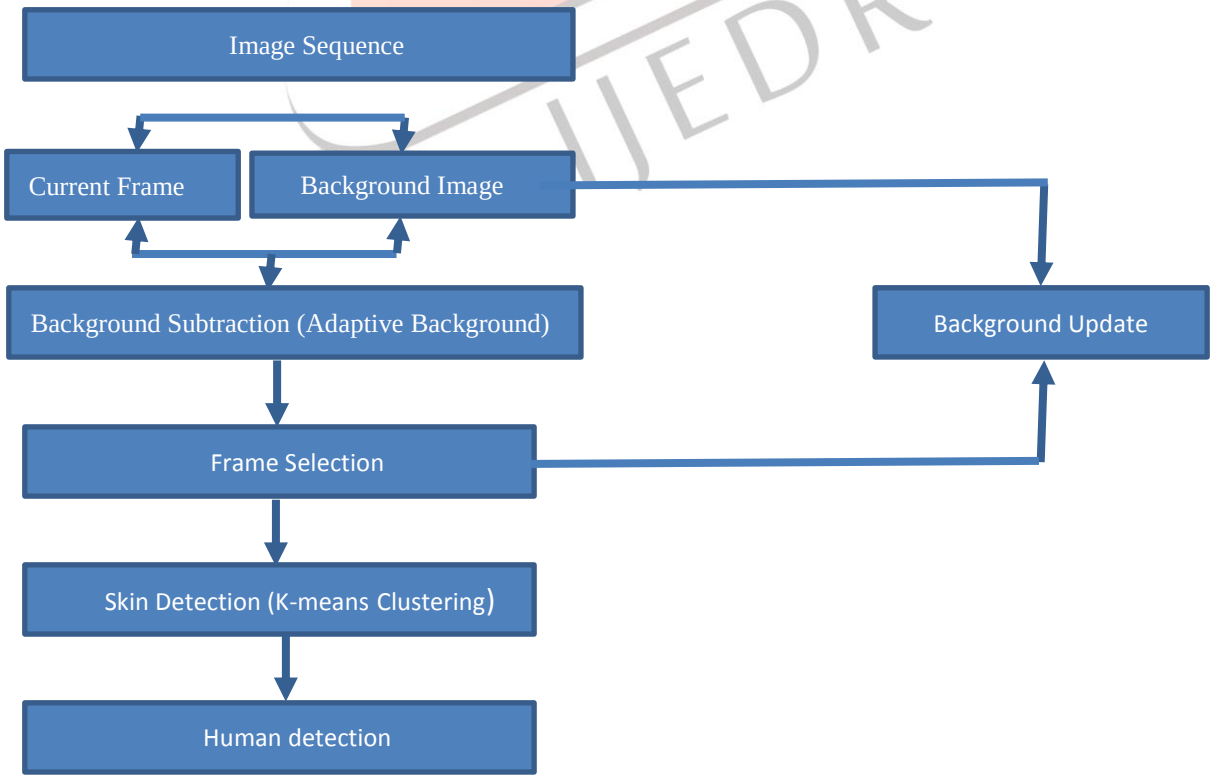

Figure 5:- Block Diagram of proposed system

In the proposed model, system will first capture the current image from the real time video, then by using the background subtraction method Gaussian Mixture; the common background of Images will be subtracted t, consider the output from the GMM as input for the fuzzy logic and by selecting the frame, detect the skin and improve in detection technique for the human. When the human detect by the Camera, the alarm will be rang, notification will be given by the system.
Using the combination of this two model and take advantage of this two model can improve the speed of the detection as well as the efficiency for the detecting the human via Camera and it also reduce the false alarm notification.

IV. IMPLEMENTATION

Images are inserted as the raw material to MATLAB, and accordingly the algorithm and flow it will process on the algorithm.
This is the input image for this algorithm. On this image algorithm try to detect using the skin. For the skin detection color scheme were used. By this color recognition we can detect the human from the images. And then it subtract the background which is non-human things or does not match skin color of the human.

Selection of the frame

For the selection of the frame we choose a video and, at T time if we select the one frame F from the video that is inserted as the the input image for our algorithm. Figure 6 is the input image for our algorithm.

K-mean Clustering[7]

K mean algorithm work only on the double data type for the we convert the binary string of the image in the double Before applying the K-mean algorithm we convert out raw image RGT to LAB.it gives more accurate result than the RBG.As shown in Figure 7 , we apply the k mean clustering, in this matlab make cluster ,it is possible to set the no clustering, normally 3 cluster is the identical for the best result , in this they made the three cluster according the color, in this cluster they made a matrix of the 3 color which is mostly used. This clusters will be the raw material for the skin detection algorithm, if Number of cluster is more the procedure time will be increase. But accuracy of the algorithm also will be increase. Figure 12 shows how the cluster were made and how it's work.

Figure:-6 Input Image

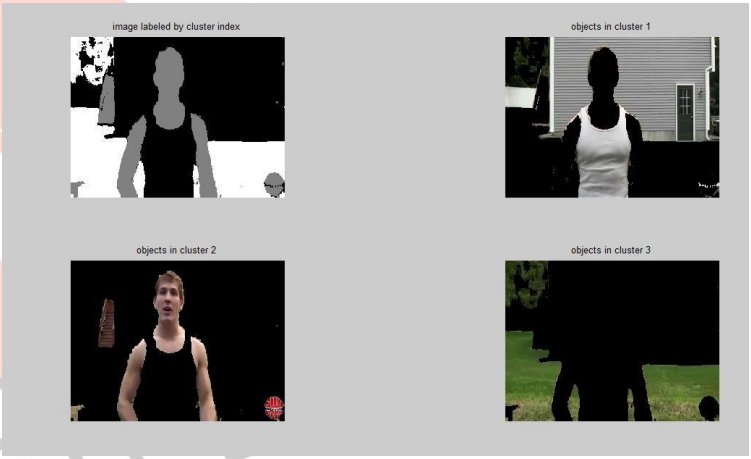

Figure:7 Procedure on the input Image

Skin Detection

After the k means algorithm we match this 3 cluster with HSV value of the of the human skin , the HSV value is depends of the are at where the human is living , for the USA the HSV value is 0.1 to 0.5. if it match with this skin color it consider as the input for the next step , we convert that cluster in to the black &white .After that algorithm give skin color to the black and white image and it will be apply to the detected image. That's known as masking figure 13 .in the masking we also make the edge soft for the better result. After feeling the color we put the square on the detected part to show this is the skin as shown in Figure 14. And as the result we show the only detected part of the human as shown in Figure 15.

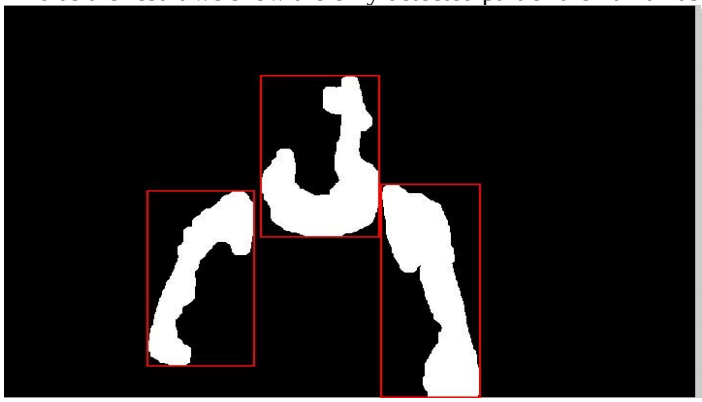

Figure:-8 Morphological Process

Figure:-9 Result of skin Detection

Adaptive Background[8]

This method is used for the background subtraction in this method this take last frame as the background for the new frame so it reference image is changed every time for the better background subtraction.

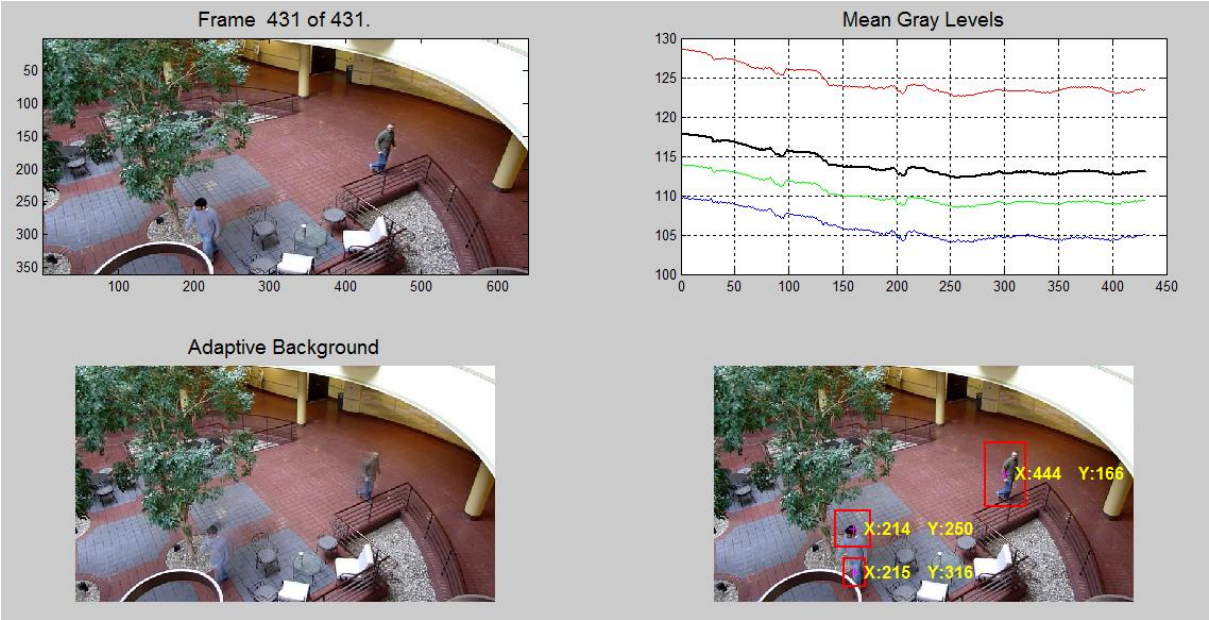

Figure:10 Result of Adaptive Background for the video

In this method there is video that direct converted into the frames and adaptive background method was applied to that video that , detection of the updated background was highlighted with red box . This take the last updated image as reference for finding the new updating in the image as shown in Figure 15 Mean Gary values show the frames Vs. RGB color detected in the current frame.

Results

Table 3: Adaptive Background Results

Video	Total Expected	Perfectly Detected Result	False Positive	Missed	Recall	Precision
Medona_Vid_1.avi	72	65	10	7	0.902778	0.866667
Scott_Her_Vid_1.avi	85	77	1	8	0.905882	0.987179
Jobs_vid_1.avi	112	110	8	12	0.901639	0.932203
atrium.avi	430	422	0	8	0.981395	1
walk1.avi	585	578	5	7	0.988034	0.991424

V. CONCLUSION

This paper includes survey on various system for Human Detection. There are some problems for the Human Detection by the CCTV camera using the various algorithms. Efficiency, speed and false detection are the major issue of the human detection. From literature survey, Conclusion is that if the combination of two algorithm which is used to detect the human, Taking advantage of this two algorithm which is "K-means" and "adaptive Background" system try to improve in the result of human detection via video surveillance system. Result of the "K-means", "Skin detection" and "face detection" on the video frames and it give the accurate result of the skin of the human.

ACKNOWLEDGMENT

The authors would like to thank anonymous reviewers for their valuable comments and suggestions on the part of security analysis, which certainly led to improvements my work.

REFERENCES

[1] Lijing Zhang,Yingli Liang ,"Motion human detection based on background subtraction", Education Technology and Computer Science (ETCS) 6-7 March 2010 Second International Workshop ,IEEE March 2010,page(s): 284 - 287 .

[2] Jae Kyu Suhr, Ho Gi Jung, Gen Li, and Jaihie Kim, "Mixture of Gaussians-based Background Subtraction for Bayer-Pattern Image Sequences", Circuits and Systems for Video Technology, IEEE ,Volume-21,Issue-3, 18th October 2010,pages(s):365 – 370.

[3]LI Li, XU Jining,"Moving Human Detection Algorithm Based On Gaussian Mixture Model", Control Conference (CCC), 2010 29th Chinese ,IEEE, 29-31 July 2010,page(s):-2853-2856

[4] Arnav Chowdhury,Sanjaya Shankar Tripathy,"Detection of Human in a Surveillance video using Fuzzy Approach",IEEE International Conference On Signal Processing and Integrated Networks, IEEE 20-21 Feb. 2014 ,pages(s):-216-219.

[5] M. Hedayati, Wan Mimi Diyana Wan Zaki, Aini Hussain, "A Qualitative and Quantitative Comparison of Real-time Background Subtraction Algorithms for Video Surveillance Applications", Journal of Computational Information Systems 8: 2 (2012).
[6] Paulo Vinicius, Koerich Borges, Nicola Conci and Andrea Cavallaro,"Video-Based Human Behavior Understanding: A survey", IEEE Transaction on Circuits and systems for Video Technology Vol. 23 No.11 November 2013, page(s):1993-2008.
[7] Grigorios Tzortzis, Aristidis Likas, " The MinMax k-Means clustering algorithm", Pattern Recognition , Volume 47, July 2014,ISSN: 0031-3203, page(s):2505–2516.
[8] Ruolin Zhang, Jian Ding, "Object Tracking and Detecting Based on Adaptive Background Subtraction", International Workshop on Information and Electronics Engineering, Volume 29, ISSN: 1877-7058 , page(s):1351–1355.