

# Automatic Detection of Human in Video and Human Tracking

B N Abhinava<sup>3</sup>, Dr. Jharna Majumdar<sup>1, 2</sup>

<sup>1</sup>DEAN R&D Prof. and Head Dept. of MTech Computer Science & Engg.,

<sup>2</sup>Head, Center for Robotics Research

<sup>3</sup>BE Student, Dept. of Computer Science & Engg.,

Nitte Meenakshi Institute of Technology, Yelahanka, Bangalore – 560 064

**Abstract:** - Automatic Human detection and tracking is a vital part of video surveillance. Many human detection and human tracking algorithms have been discussed in literature survey. Authors in this paper have attempted to identify the human in cluttered environment, identify human body (head, body and leg, track the human in the video based on RGB colour model and also detect collision between multiple human.

**Keywords:** - Particle filter tracking, pyramid computation, block based integral, human body identification, frame difference, collision detection.

## I. INTRODUCTION

Trespassing is a major concern; may it be near country border or in residential area. It's a major threat to life, property and security of individual. This criminal activity can be reduced with the use of latest technologies such as security cameras, mobile phones and sensors. These technologies were restricted to government sectors as initial cost of setting up was very high. Due to advances in computer technology, faster computing powers and cheaper initial costs of modern computers, this technology is made available to common people.

Authors in this paper have tried to come up with an efficient algorithm to carry out the task. Section II briefs about some of the research papers. Section III is about methodology used in this paper to accomplish the task. Section IV discuss the proposed method for human detection and tracking, in Section V the various experimental results using the proposed method are given. Finally, we conclude our research in Section VI.

## II. LITERATURE SURVEY

Aras Akbari, U [1], in their work on *Multiple Human Tracking in Dynamic Environment*, they have used both laser-based system and vision-based system to detect object and identify its human or not. Bayes filter, Particle filter, Adaboost algorithms were used.

Alok K. Singh Kushwaha, Chandra Mani Sharma [2], they have used haar-like features for object detection, trained detector is used to identify human and tracking is done using particle filter.

San-lung zhao, Hsi-jian lee[3], in their work they have used Gaussian background model to detect human, color histogram as feature and Particle filter algorithm for tracking. Also, a method for failure adjustment.

Shih-Min Chen, Chen-Kuo Chiang [4], Using methods such as Block based integral image, Spatial Pyramid Ring and Histogram equalization they have presented a feature representation that achieves rotation, translation and scale invariant simultaneously.

E. H. Adelson, C. H. Anderson, J. R. Bergen, P. J. Burt, J. M. Ogden [5], in the journal they have brought out various pyramid methods such as Gaussian pyramid, Laplacian pyramid etc.

Mototsugu Muroi, Heitoh Zen [6], they proposed object tracking method using multi-part color histogram and single-part color histogram for adaptive target color representation. Further, particle filter tracking algorithm is used to track object.

## III. OBJECTIVE OF THE WORK

The main objective of this paper is to propose another efficient method for human detection and tracking.

A schematic representation about the flow of work is shown in figure 1. Initially, Videos of security cameras are taken as data sets and each frame are processed to grey-scale sequentially. Pre-processing by smoothing the input frame using various pyramids is done (section 4.1). Any motion in the scene is detected by frame difference (section 4.2). Block based integral method (section 4.3) is done to identify single or multiple human. Human body parts (head, body and leg) is identified in (section 4.4). Then Particle filter algorithm is

used to track human (section 4.5) and we conclude with detecting collision between multiple human (section 4.6).

**Input:** Video frames

**Output:** Tracked human

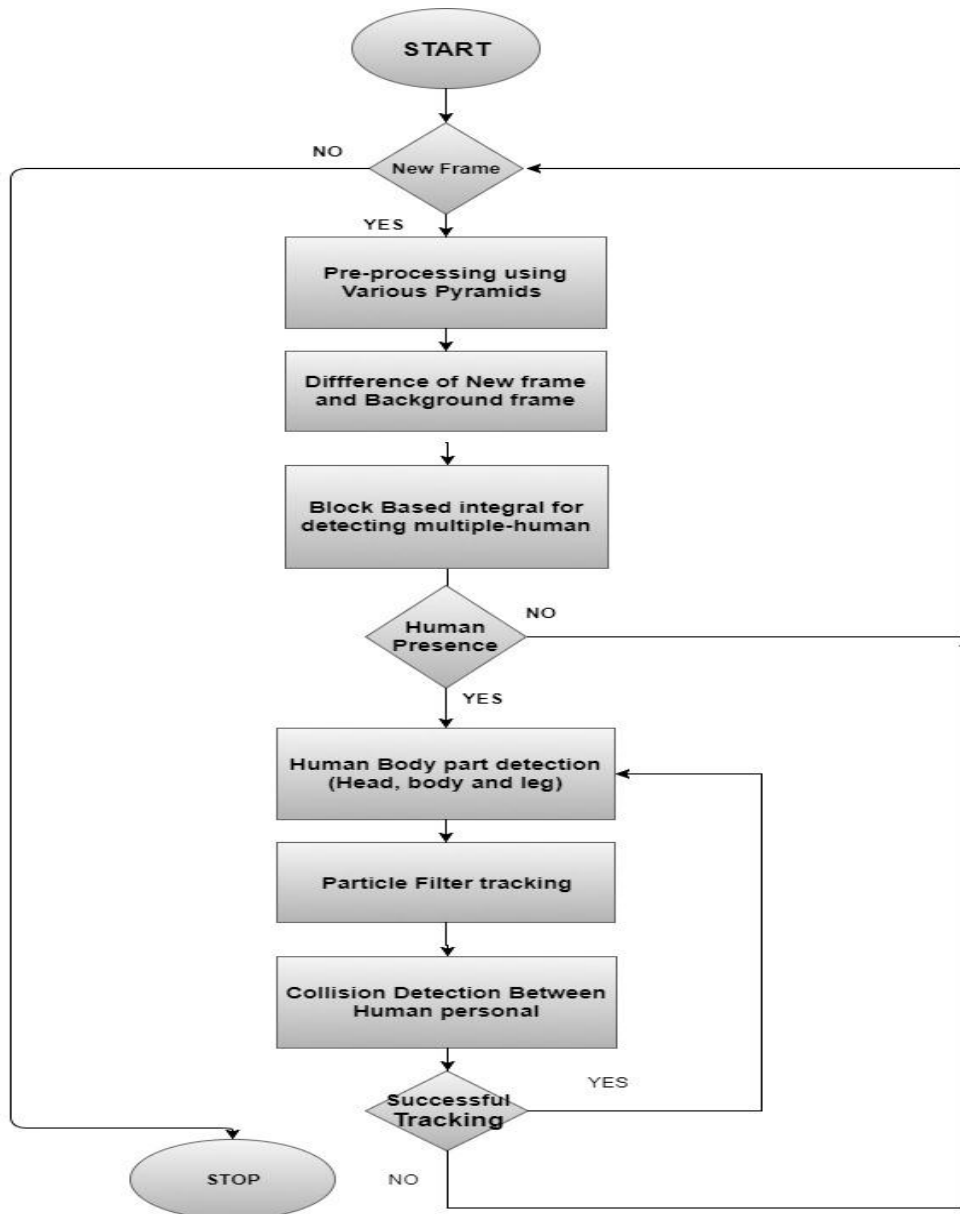


Fig 1: Flow Chart

#### IV. THE PROPOSED METHODS

##### 1) Pre-processing

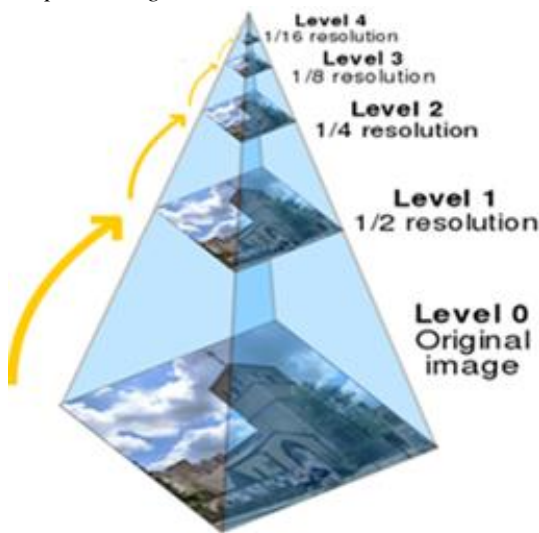


Fig 2: Resolution of image in each level

Various pyramids like Gaussian, Laplacian, contrast, mean and Filter-Subtract-Decimate Pyramids are used to smooth the image. As frame difference method is robust to slight changes in the image. Smoothing helps eliminate this slight variation.

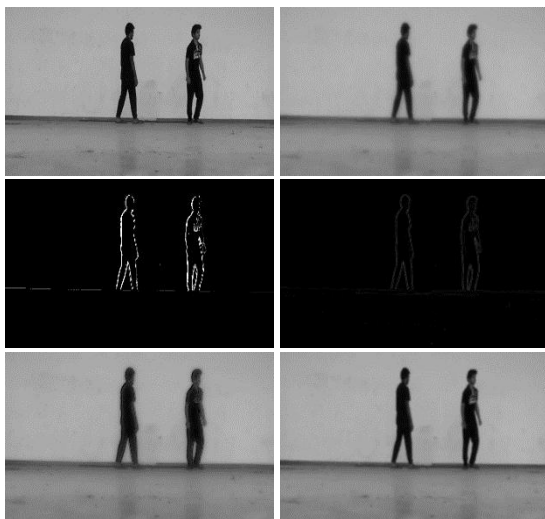


Fig 3: (a) Original image (b) Gaussian (c) Laplacian (d) Filter-Subtract-Decimate (e) Contrast (d) Mean pyramids

Quality Matrix such as Mean square error, mean average error, Entropy and PSNR were used to determine which pyramid was best suitable for this application and results shows that Gaussian Pyramid is best suited for our application.

##### 2) Frame difference

The Video frames are converted to greyscale. Pixelwise difference is performed between Newly Arrived Frame Fig 4(b) and Absolute Background Fig 4(a) of the scene at 2D special coordinate  $[x, y]$  to detect any motion changes in the scene. If difference value is greater than a defined threshold then that pixel is highlighted in the output frame. Other values less than threshold are ignored. Appendix 1.

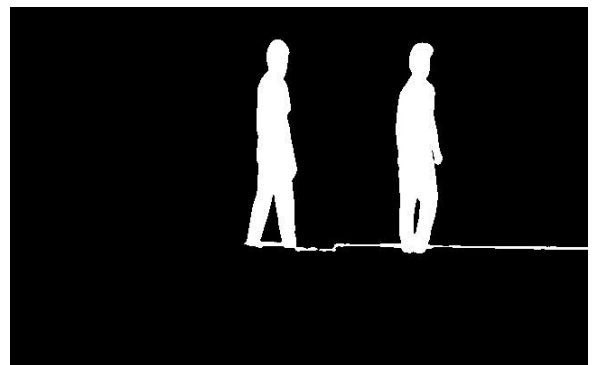
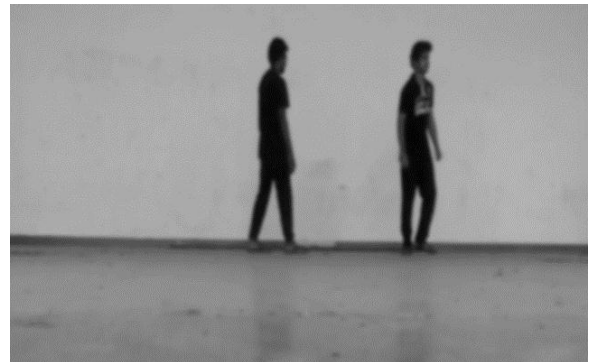


Fig 4: (a) Absolute Background of the scene (b) New frame with foreground objects (c) Pixelwise difference between (a) & (b)

The moving human is separated out of the scene Fig 4(c) (We are taking hypothetical situation where a human walk into the scene) and we get a binary image capturing motion in the scene.

### 3) Block based integral

Efficient method to identify single or multiple human in the scene. First, the frame is divided into blocks Fig 5(a). Density of pixels in the region is counted and mean of all blocks is calculated. Second, Blocks with density more than mean is considered as potential blocks (Fig 5(b)). Third, these potential blocks are clustered based on neighborhood. Each set of Connected blocks are clustered to a group, that means one cluster determines one human Fig 5(c). Therefore, the moving object or the human personal single or multiple are identified through this method.

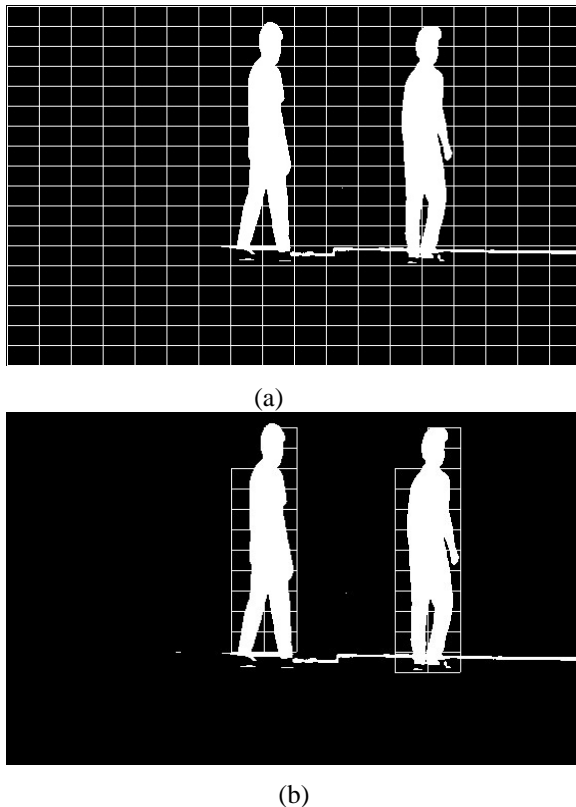


Fig 5: (a) Image divided into blocks.  
(b) Identification of potential blocks.

### 4) Human body parts detection (head, body and leg)



Fig 6: Human Body Proportionality ratio

Detection of human body parts is crucial to identify his face, recognise some of the individuals characteristics like walking pattern, movement of head, etc.

This is possible if we are able to detect head, body and leg. Based on human body proportionately (Fig 7), head, body and leg can be identified.

Algorithm: Human body part detection

1. Initialize: After detection of human from block based integral we get area of human body. Based on the human body proportion we can identify human body part.
2. Detection:
  - Head:  $1/8^{\text{th}}$  of human is head, after identification of human,  $1^{\text{st}}$   $2/8^{\text{th}}$  part is defined as head.
  - Body:  $4/8^{\text{th}}$  of human is body, after identification of human,  $4/8^{\text{th}}$  part after head is designated as body.
  - Leg: bottom:  $3/8^{\text{th}}$  of human body is leg.

This process is repeated for every frame provided the human is detected or present in the scene (Fig 7). This information is further used for tracking human.



Fig 7: Body part identification (head, body and leg)

### 5) Particle filter tracking

Particle filters are state estimation methods for systems with non-linear process and measurement models corrupted with noise which may be non-Gaussian and multimodal. These are recursive Monte Carlo (MC) statistical computing methods. Particle filters are an important alternative to Kalman filters which are optimal to linear systems corrupted with To start with the tracking process we need to determine initial position.

In our tracking algorithm we use two trackers. Based on detection of head and body in Section IV(4) we get the template windows for head and body. We do not consider leg as its movement changes much more frequently in every frame compared to head and body.

Algorithm is explained below:

Initialize: Generate two sets of Particles with a position vector  $[x, y]$  as well as a weight  $w$  inside the template window for head as well as body.



Repeat:

- Predict: The prediction stage moves every particle to a new position (Fig 8). Each particle needs to be moved to where the objects is expected to be next, on the assumption that the particle's value is correct. So, each set of particles simultaneously predicts head and leg based on where it moves.
- Update: The update stage assigns weight to each particle based on the current Observation. Compute the colour histogram for the region and compare it to the desired colour histogram with Histogram correlation. If we normalize the particle weights so their sum is 1.0, then the particles collectively approximate the probability distribution of the object position and take a weighted mean in a region around the best particle (Fig 8).
- Resample: Very low weighted particle is unlikely to represent an accurate estimate of the tracked value. Therefore, Sort the particles by weight, and replace the lowest-weighted half of the particles with duplicates of the highest-weighted half.

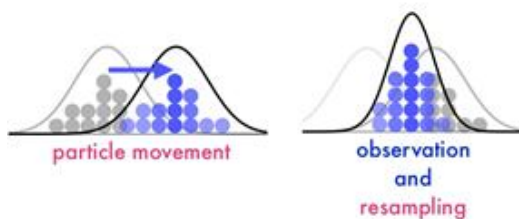


Fig 8: Prediction, observation and resampling

- Highest-weighted particle among head and leg considered, which every has lowest weight among the two, tracker position is update and more particles are added to get better prediction in future (Particle addition is maintained not to reach more than maximum possible particles) and number of particle are reduce for the other.

Updating position of the tracker is curtail as error will not propagate easily to the following frames.



Fig 9: Head and Body tracking with error correction

#### 6) Collision detection

Collision causes most of damage to property and injuries in day to day activities. Detection of collision is one of the major research field in Transportation where collision between vehicles are to be detected.

We restrict our self for human personal collision in surveillance in this paper. Detecting collision between human personal and alerting the personal about the same. From block based integral section IV (3) we get centroid of human. By calculating Euclidian distance, we can detect the collision between personals based on their speed of approach and distance between them to notify and alert about the collision through speakers.

Algorithm:

Euclid's distance algorithm is applied on two individuals at a time (Centroid of each individual). Appendices 2.

Total number of distances between multiple human:

$$n_{C_1}$$

Where n is number of human personal.

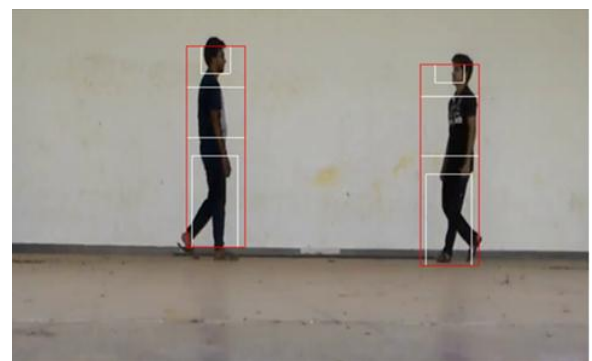


Fig 10: Detecting collision between human.

## V. RESULTS

### a) Pyramid Results

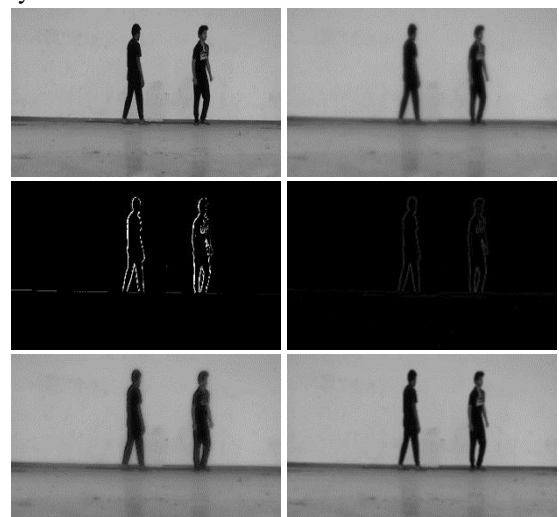


Fig 11: (a) Original image (b) Gaussian (c) Laplacian (d) Filter-Subtract-Decimate (e) Contrast (d) Mean pyramids

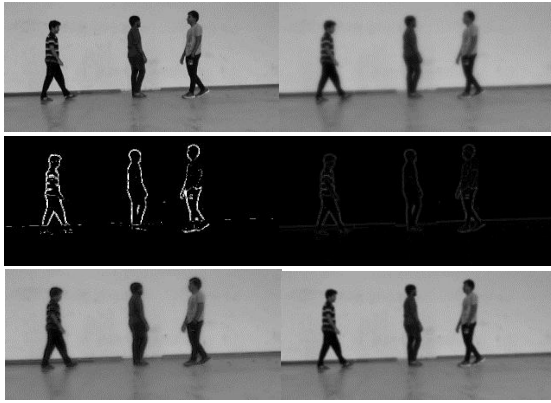


Fig 12: (a) Original image (b) Gaussian (c) Laplacian (d) Filter-Subtract-Decimate (e) Contrast (d) Mean pyramids

Type of pyramid	Kernel & Sigma	MSE	AE	Entropy	PSNR
Gaussian	3, 0.6	43.4318	2.2282	5.94373	92.7901
	3, 1.6	45.8043	2.31743	5.94703	93.2521
	3, 2.6	58.0805	4.53074	5.95953	95.3146
	5, 0.6	62.8479	2.93857	5.96906	95.9998
	5, 1.6	66.9776	4.75114	5.97469	96.5526
	5, 2.6	87.5591	3.69105	5.98953	98.88
Laplacian	3, 0.6	22747.3	145.658	1.90526	147.173
	5, 0.6	22670.2	145.349	2.13632	147.143
Contrast	3, 0.6	72.8445	2.53229	5.99898	97.2819
	5, 0.6	103.093	3.22397	6.03117	100.299
Mean	3, 0.6	44.06198	0.71812	5.86448	93.1628
	3, 2.6	58.9108	2.54233	5.87546	95.504
	5, 0.6	63.9376	1.12406	5.88521	96.2111
	5, 2.6	89.2817	1.5734	5.89784	99.1659
FSD	3, 0.6	22859.2	146.135	1.40953	147.215
	5, 0.6	22762.2	145.744	1.79401	147.178

Fig 13: Quality matrix for Fig 11 determining Gaussian pyramid is most suitable pyramid

Type of pyramid	Kernel & Sigma	MSE	MAE	Entropy	PSNR
Gaussian	3, 0.6	71.0872	3.11776	6.36848	94.5711
	3, 1.6	74.989	3.24985	6.37277	95.0352
	3, 2.6	88.4604	5.26793	6.40992	96.4702
	5, 0.6	101.595	4.10661	6.37937	97.6726
	5, 1.6	102.513	5.59982	6.41569	97.7508
	5, 2.6	139.138	5.12192	6.38453	100.404
Laplacian	3, 0.6	22165.1	142.022	2.12346	144.449
	5, 0.6	22058.2	141.587	2.40736	144.407
Contrast	3, 0.6	120.229	3.45628	6.407	99.1354
	5, 0.6	166.596	4.40005	6.43313	101.968
Mean	3, 0.6	74.3086	1.08506	6.32959	94.8836
	3, 2.6	90.2616	2.88955	6.35339	97.5844
	5, 0.6	102.5876	1.64172	5.86448	97.9955
	5, 2.6	140.7138	2.25248	6.35792	100.8074
FSD	3, 0.6	22284.4	142.562	6.31411	144.495
	5, 0.6	22161.2	142.039	2.05086	144.447

Fig 14: Quality matrix for Fig 12 determining Gaussian pyramid is most suitable pyramid

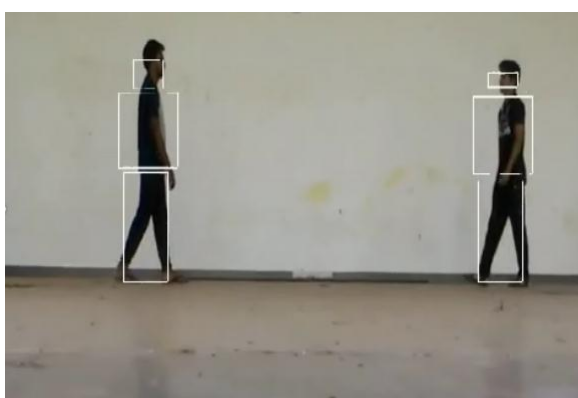
b) Human body part (head, body and leg) identification



Frame 0



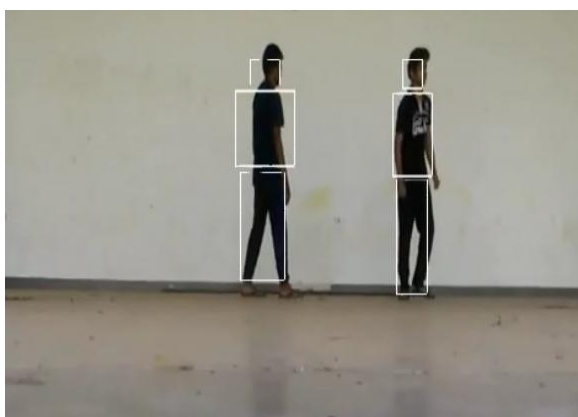
Frame 20



Frame 40



Frame 60



Frame 80



Frame 100

Fig 15: Video sequence where human body parts (head, body and leg) of multiple human is identified.

c) Particle filter tracking

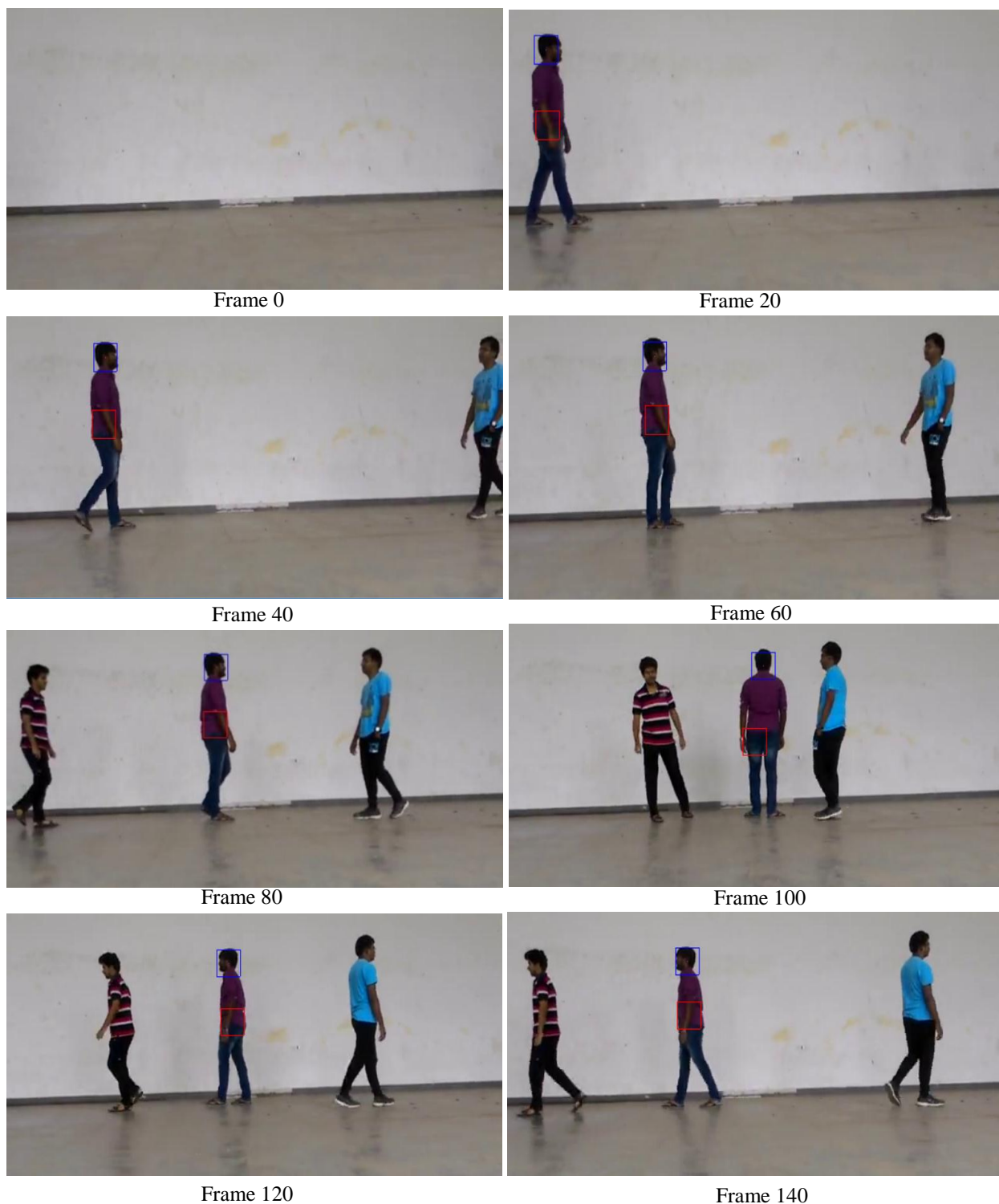


Fig 16: Particle filter tracking using head and body tracker



## VI. CONCLUSION

We were able to successfully implement the proposed work. Our Simulation results has shown that the algorithm can successfully and efficiently detect multiple human personal. It can handle non-rigid deformation of targets, partial occlusion and cluttered background. The key approach is that the background is of sufficiently different colour structure than the object to be tracked. The tracking algorithm with target update and multiple tracker can track human personal more efficiently. The proposed algorithm can be improved by using better regeneration algorithm after pyramid computation, Edge detection algorithms can be incorporated for better body part identification and shadow removal methods can be used to get better accurate area of human personal.

## VII. ACKNOWLEDGEMENT

The authors express their sincere gratitude to Prof N.R Shetty, Advisor and Dr. H.C Nagaraj, Principal, Nitte Meenakshi Institute of Technology for giving constant encouragement and support to carry out research at NMIT.

The authors extend their thanks and gratitude to the Vision Group on Science and Technology (VGST), Government of Karnataka to acknowledge their research and providing financial support to setup the infrastructure required to carry out the research.

## REFERENCES

- [1] Multiple Human Tracking in Dynamic Environment, Aras Akbari, U. of Southern California.
- [2] Alok K. Singh Kushwaha, Chandra Mani Sharma, "Automatic Multiple Human Detection and Tracking for Visual Surveillance System", IEEE/OSAIAPR International Conference on Informatics, 2012
- [3] Mototsugu Muroi and Heitoh Zen, "Human Tracking Based on Particle Filter in Outdoor Scene", MVA2007 IAPR Conference on Machine Vision Applications, May 16-18, Tokyo, JAPAN, 2007
- [4] Rotation, Translation, and Scale Invariant Bag of Feature based on Feature Density, Shih-Min Chen, Chen-Kuo Chiang, 7th International Conference on Intelligent Systems, Modelling and Simulation, 2016.
- [5] E. H. Adelson, C. H. Anderson, J. R. Bergen, P. J. Burt, J. M. Ogden, "Pyramid methods in image processing", RCA Engineer, 29-6, Nov/Dec, pages 33-41, 1984
- [6] San-lung zhao and hsi-jian lee, "particle filter-based multi-part human tracking with failure adjustment in video sequences", journal of information science and engineering 26, 2267-2281 (2010)
- [7] Mohammad Hossein Ghaemini<sup>1</sup>, Amir Hossein Shabani<sup>2</sup>, and Shahryar Baradaran Shokouhi, "Adaptive Motion Model for Human Tracking Using Particle Filter", International Conference on Pattern Recognition, 2010
- [8] [http://docs.opencv.org/2.4/doc/tutorials/imgproc/histograms/histogram\\_comparison/histogram\\_comparison.html](http://docs.opencv.org/2.4/doc/tutorials/imgproc/histograms/histogram_comparison/histogram_comparison.html)
- [9] F. Aherne, N. Thacker, and P. Rockett: "The Bhattacharyya metric as an absolute similarity measure for frequency coded data," *Kybernetika*, vol.32, no.4, pp.1-7, 1997.
- [10] Yuan Chen, Shengsheng Yu<sup>1</sup>, Jun Fan<sup>1</sup>, Wenxin Chen, Hongxing Li, "An Improved Color-Based Particle Filter for Object Tracking", Second International Conference on Genetic and Evolutionary Computing, 2008

## VIII. APPENDIX

1.

$$Io[x, y] = \sum_{i=0}^n |In[x, y] - Ib[x, y]|$$

Where  $Io$  = Binary output frame

$In$  = new frame

$Ib$  = background of frame.

$[x, y]$  = 2D coordinate

$n = l * b$  of the image

2. Euclidean distance

$$\begin{aligned} d(\mathbf{p}, \mathbf{q}) &= d(\mathbf{q}, \mathbf{p}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2} \\ &= \sqrt{\sum_{i=1}^n (q_i - p_i)^2}. \end{aligned}$$

Where  $\mathbf{p} = (p_1, p_2, \dots, p_n)$  and  $\mathbf{q} = (q_1, q_2, \dots, q_n)$  are two points in Euclidean  $n$ -space.