REAL TIME MOVING OBJECTS TRACKING BASED ON MEAN SHIFT ALGORITHM USING MATLAB

Mrs.T.L. Spandana[1], N.Sree Laya[2], P.Puja Pranathi[3],G.Dileep[4], P.Sai Chand[5]

[1]Assistant professor of ECE, Raghu Institute of Technology, Visakhapatnam, India
[2]UG Students of ECE, Raghu Institute of Technology, Visakhapatnam, India

---***---

Abstract -The paper introduces a modern approach in developing an automated visual surveillance system, which is very effective in detecting and monitoring moving objects in both current video and live streaming videos. Object tracking is one of the most common domains of video processing. The primary objective of object tracking is to estimate the object's location in images continuously and reliably against complex scenes. Here, the proposed system's graphical user interface is implemented using MATLAB 2018b, which operates video images from a stationary camera on both color and gray scale. For a moving object in a video, a rectangular window is described in this approach. In this method, we proposed real-time moving object detection i.e. detection system using static webcam that can process video sequences for catching live scene and also saved segmented video database. Using MEAN SHIFT ALGORITHM. Detecting moving objects serves as a guideline for identification, classification and behaviour analysis. This method can be used for video surveillance, human activity analysis, road condition monitoring, airport safety, marine border security monitoring, and so on.

***Key Words*:**Real-time object tracking, Live scene, Recorded video, MATLAB, Moving objects, MEAN SHIFT ALGORITHM.

1.INTRODUCTION

Tracking moving, non-rigid objects in videos is a significant and challenging task in the field of Computer vision and artificial intelligence with various uses, including video monitoring (of humans and Vehicles), traffic control, sport video , video analysis, encoding, and multimedia mining.And also moving object detection and tracking surveillance system provides crucial information about the object behavior, interaction and the relationship between objects with high level processing since automated video surveillance systems has modules from low-level such as object detection, object tracking, classification, event analysis, efficiency and robustness in each module are particularly important. To design smart computer vision based video surveillance systems which can contribute to the safety of people in the home and in public places such as, airports, railway stations, shopping malls and other public related places has now made it possible with the advent of smart multi mega pixel consumer cameras having higher processing capabilities. The critical threat of public safety due to terrorist attacks especially, explosive attacks and vehicles are repeatedly concentrated on such public places. A key function in such a computer vision-based video surveillance system is the understanding of human behavior in relation with objects left unattended in public places. In this context, real time smart visual surveillance object detection and tracking systems for human behavior understanding have drawn much attention of researchers and investigated worldwide as an active research topic [2]. Thus, a primary goal of video surveillance is to obtain a live description of what is happening in a monitored area and take (or trigger) appropriate action against pertained object and human misbehavior at public place, shopping malls, banks, railway station, airports…. etc.

In video processing, a video can be represented with some hierarchical structure of units, such as video scene, shot and frame. Also, video frame is the lowest level in the hierarchical structure. The content-based video browsing and retrieval uses these structure units for video content analysis. In video retrieval, generally, video applications must first partition a given video sequence into video shots. A video shot is defined as an image or video frame sequence that presents continuous action. The frames in a video shot are captured from a single operation of one camera. The complete video sequence is generally formed by joining two or more video shots consecutively. There are two basic types of video shot transitions, namely abrupt and gradual. Abrupt transitions i.e., cuts are the simplest form that can be occur in a single frame when restarting and stopping the camera. Although many kinds of cinematic

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 05 | May -2020 ISSN: 2582-3930

effects could be applied to artificially combine two or more video shots.

Despite of many research works tracking multiple moving objects through complex scenarios remains a challenging task [1]. An independent robust tracker is used to detect and track each individual object is one category of existing method while using multiple independent trackers are used for multiple object tracking. In our robust visual tracking method that jointly employs a mean shift algorithm, our main objective includes comparison of existing work with mean-shift such as: 1) The mean shift tracking scheme introduces the full combination of functionalities and adjustable bounding box parameters; 2) Partitioning a rectangular bounding box and deriving multi-mode mean shift; 3) Introducing live learning method for the reference object distribution; and 4) Computing bounding box parameters through mean shift algorithm.

2. LITERATURE SURVEY

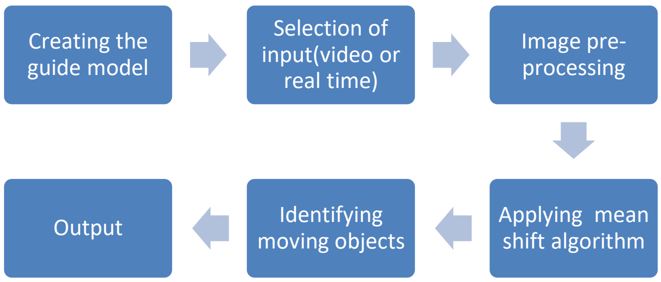
There have been a number of surveys about object detection, classification, tracking and activity analysis in the literature [4]. There are several applications that benefit from smart video processing has divergent needs, thus requires different treatments. However, they have something in common: moving objects. Thus, detecting regions that correspond to moving objects such as people and vehicles at railway station, airports in video is the first step of almost every vision system since it provides a focus of attention and simplified the processing onsubsequent analysis steps. Due to dynamic changes in natural video scenes such as sudden illumination and weather changes, repetitive motions and complex background that cause clutter motion detection is a difficult problem to process reliably. Commonly used techniques for moving object detection and tracking are background subtraction, temporal differencing, statistical methods, and optical flow method.

In object tracking first video is converted into number of consecutive frames then processes for locating moving objects from the scene. Real time object tracking is a challenging problem in the field of computer vision such as motion-based recognition, automated surveillance, traffic monitoring, and object-based video compression etc. Mean shift has drawn much interest in real-time object tracking by maximizing the Bhattacharyya coefficient between the reference and the target i.e. objects region of interest. Early work was proposed by Comaniciu [6], Collins extended his work for mean shift by introducing kernel bandwidth normalization. It performs an extensive search of object within a range of bounding box scales and it is computationally intensive. Using mean shift method, the center, size, shape and orientation of the bounding box are simultaneously estimated during the tracking was proposed by Sumin in [5]. Further improvement is made in visual object tracking by using mean shift.To track the objects motion [5] by using the mean shift where particles with higher weights from the mean shift are combined in the observation model, and it reduces the degeneracy and requires fewer particles than the conventional particle filter [3]. By applying the mean shift on particles with large weights is also as called elite particles to weight particles by using the observation model which is proposed by Zhong [8].

3. PROPOSED METHOD

METHODLGY :

Our method of approach to find moving objects using MATLAB follows six steps. They are:



Approach to moving object detection :

A robust detection system should be able to recognize when objects are moving and these functions usually not possible with traditional motion detection algorithms. So, we use mean shift algorithm which detects the objects based on the pixel'svariation. So here we use pixel analysis.

Pixel analysis :

Pixel analysis determines whether a pixel is stationary or transient by observing its intensity value over time. Moving objects passing through a pixel cause an intensity profile step change, followed by a period of instability; then the profile stabilizes, in a manner dependent on the kind of event. To capture the nature of changes in pixel intensity profiles, a gradient based approach is applied.

Let $It(x)$ be the intensity of pixel x at a time t occurring k frames in the past. The *motion trigger T* prior to the frame of interest t is the maximum absolute difference between the pixel intensity $It(x)$ and its value in the previous l frames:

$$T = max_{j=1,......,l}\{|It(x) - It - j(x)|\} \quad (1)$$

where suggested value for l is 5 [2]. Let us also introduce the stability measure as the variance of the pixel intensity profile from time t to the present:

$$s = \frac{(k+1)\sum_{j=0}^{k} I_{t+j}^{2}(x) - \left(\sum_{j=0}^{k} I_{t+j}(x)\right)^2}{k(k+1)}$$

where k is set to correspond to one second of video [2]. Once$T$ and S have been computed, a *transience map M* can be defined for each pixel, taking three possible values: background (BG), transient (TR), or stationary (ST). For each pixel, the corresponding map value is updated to TR if it was ST or BG and if motion trigger is greater than a given threshold (there has been a step change in intensity). Moreover, if it was TR and stability measure is lower than a given threshold (intensity has been stabilized), it is updated to BG if its stabilized intensity value is equal (within a threshold $T\ h$) to the background intensity value, and to ST otherwise. In order to allow for adaptivity of the background model to slow lighting changes, we update the background B by running average with selectivity. Specifically, background model Bt is initially set to the first image ($B0(x) = I0(x)$ for every pixel x), and then updated as:

$$Bt{+}1(x) = \alpha Bt(x) + (1 - \alpha)It(x),\ x \text{ non} - \text{moving}$$

$$Bt(x), \text{ otherwise } (2)$$

where α is a time constant that specifies how fast new in formation supplants old observations, usually chosen in [0.9, 1].

4. MOVING OBJECTS TRACKINGUSING

MEAN SHIFT

Mean Shift is a powerful and versatile iterative algorithm that can be used for lot of purposes like finding modes, clustering etc. It has been widely used in target tracking field because of some advantages like fewer iteration times and better real-time performance for many years. However, due to only single-color histogram representation of target feature has been used in traditional Mean Shift Algorithm characteristics of mean shift algorithm.

1) Application independent tool
2) Suitable for real data analysis
3) Can handle arbitrary feature space
4) Does not assume any prior shape on data clusters

Mean Shift Algorithm:

The mean shift algorithm is a non-parametric clustering technique that require prior knowledge of cluster number, and does not limit cluster form.

According to n data points xi,I=1,...,non a d-dimensional Rd,multivariate kernel density estimation obtained with K(x) kernel and window radius h

$$f(X) = \frac{1}{nh^d}\sum_{i=1}^{n} k\left(\frac{X-X_i}{h}\right)(1)$$

For radially symmetric kernels, it is necessary to define the kernel profile which satisfies k(x)

$$K(X) = C_{k,d}K(\|X^2\|) \quad (2)$$

Where ck, d is a constant of normalizationassuring that k(x) integrates to

1. The density function modes are located at

The x(x)=0gradient function zeros

The Density Estimator gradient (1) is

$$\nabla f(X) = \frac{2c_{k,d}}{nh^{d+2}}\sum_{i=1}^{n}(X_i - X)g\left(\left\|\frac{X-X_i}{h}\right\|^2\right)$$

$$= \frac{2c_{k,d}}{nh^{d+2}}\left[\sum_{i=1}^{n}(X_i - X)g\left(\left\|\frac{X-X_i}{h}\right\|^2\right)\right]\left[\frac{\sum_{i=1}^{n} X_i g\left(\left\|\frac{X-X_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{X-X_i}{h}\right\|^2\right)} - X\right] \quad (3)$$

where $g(s) = k0(s)$. The first term is proportional to the density estimate at **x** computed with kernel $G(\mathbf{x})=cg, dg$ ($k\mathbf{x}k2$) and the second term

$$m_h(x) = \frac{\sum_{i=1}^{n} X_i g\left(\left\|\frac{X-X_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{X-X_i}{h}\right\|^2\right)} - X (4)$$

is the mean shift. The mean shift vector always points toward the direction of the maximum increase in the density. The mean shift procedure, obtained by successive

- computation of the mean shift vector$m_h(X^t)$,
- translation of the window $X^{t+1} = X^t + m_h(X^t)$

is guaranteed to converge to a point where the gradient of density function is zero. Mean shift mode finding process is illustrated in Figure 1.

The mean shift algorithm is a practical application of the mode finding procedure:

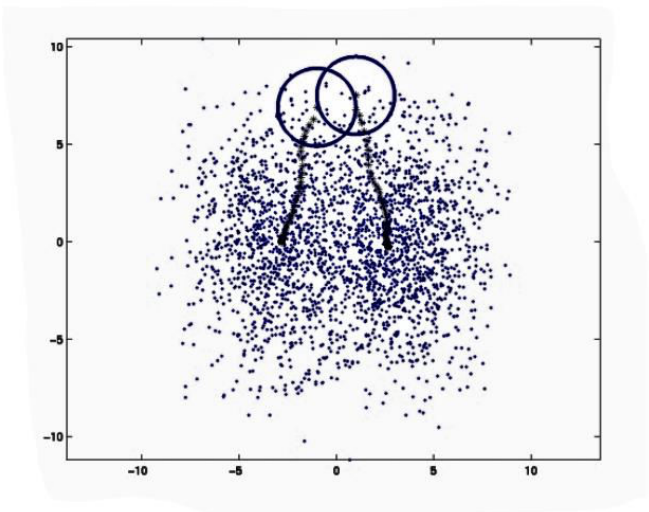
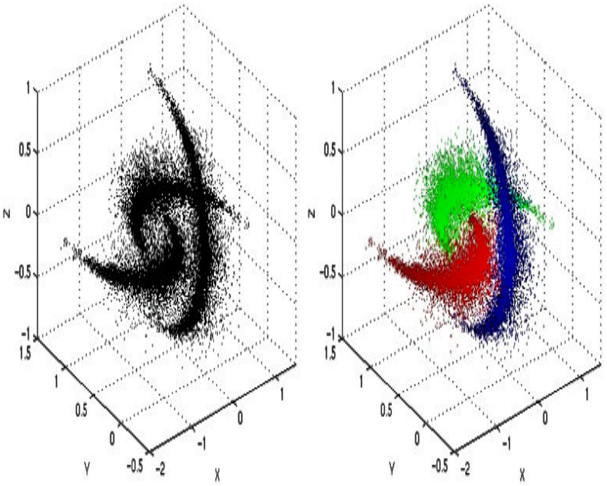
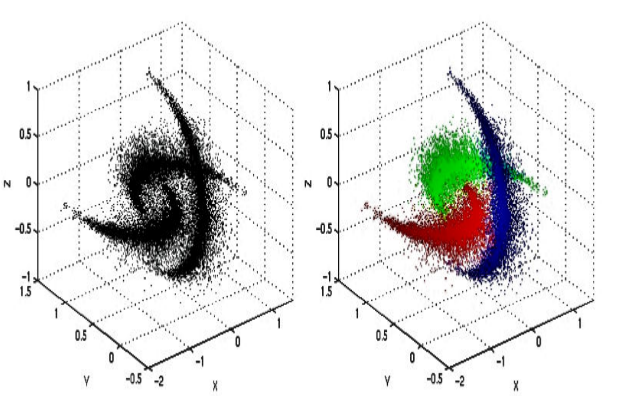


Figure 4.1: Mean Shift Mode Finding

1. starting on the data points, run mean shift procedure to find the stationary points of the density function.
2. prune these points by retaining only the local maxima. The set of all locations that converge to the same mode defines the basin of attractions of that mode. The points which are in the same basin of attraction is associated with the same cluster. Figure 2shows two examples of mean shift on three-dimensional data. More details on mean shift algorithm on Lie Groups can be found in [10].



(a) Synthetic example of three non-linearly separable clusters (32640 points).



(b) Real example of 14826 points in the LUV color space.

Figure 2: Mean Shift Clustering

5. EXPERIMENTAL RESULTS

Moving object detection and tracking is one of the important steps which has attracted a great interest from computer vision and image processing researchers due to its applications in areas, like video surveillance. In case of real time moving object detection from the scene involves identification of an object in consecutive frames. Single object and multiple object movements in a frame with respect to the computed vectors are segmented with the help of specified threshold limit. The extracted movements are tracked using mean shift algorithm. In our proposed MATLAB based moving objects tracking system according to mean shift algorithm outputs are mentioned in the below figures i.e. in Figure: 6.1, Figure: 6.2(a), Figure: 6.2(b), Figure: 6.3, Figure: 6.4. The results

obtained from the live capturing camera datasets is shown in output, from that we can see how the performance parameters are varied accordingly videos taken from different scenarios by using stationary camera as input to the system. Before applying algorithm to input, video is segmented into number of frames and that segmented videoframe sequences are fetched to system. Thresholding is one of the most powerful and important tools in image processing of computer vision for image segmentation. The importance of segmented images obtained from thresholding has the advantages as it requires smaller storage space, performs fast processing and eases in manipulation compared with gray level image.

Object detection and tracking has been used in many applications, with the most popular ones being: security (e.g., recognition, tracking) like Human Activity Analysis, Road Condition Monitoring, Airport Safety, monitoring for protection along marine border. Some existing techniques are now part of many consumer electronics (e.g., face detection for auto-focus in smartphones) or have been integrated in assistant driving technologies, according to this we are still far from achieving human-level performance, in particular in terms of open-world learning. For that we have proposed real time object detection and tracking.

SIMULATION RESULTS :

Figures:6.1 Shows the obtained results, in which an object moving to and fro is tracked in an offline stored video which is given as input, while selecting the background input of the system.

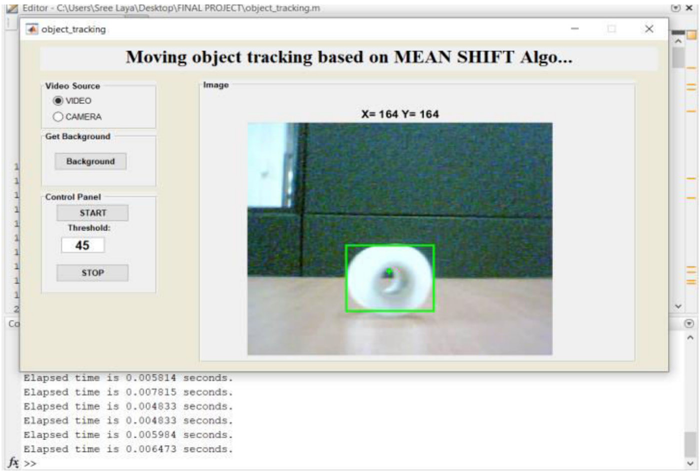


Figure: 6.1- Stimulated output for existing or recorded video

Figure:6.2(a); Figure:6.2(b) Shows the obtained results in which two persons who came across, greets each other and this is tracked (as they are moving) in an offline stored video which is given as input, while selecting the background input of the system.

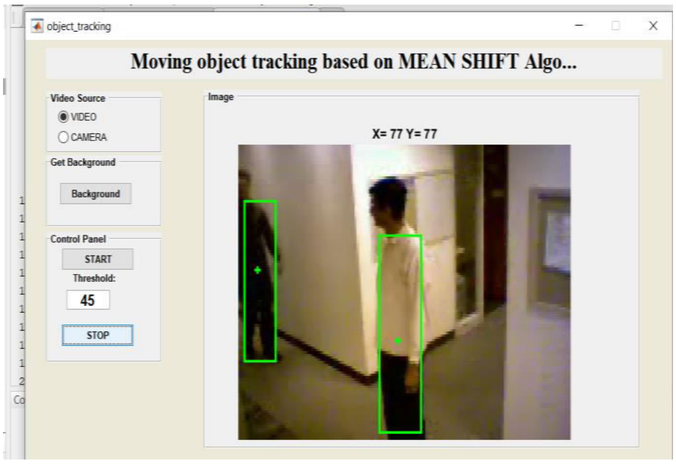


Figure: 6.2(a)- Stimulated output for existing or recorded video

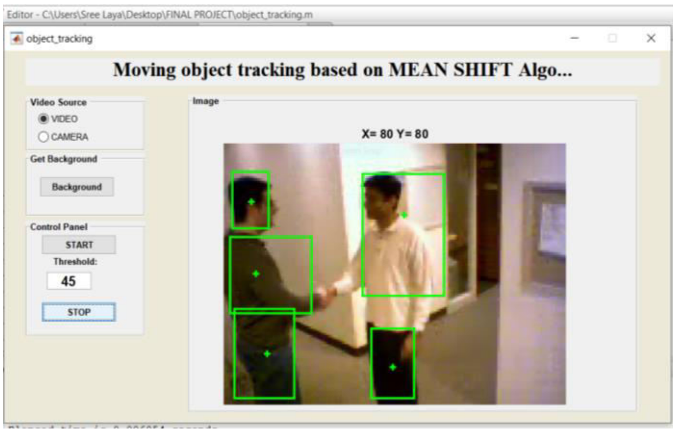


Figure: 6.2(b)- Stimulated output for existing or recorded video

Figure: 6.3 shows the person opening the gate and trying to enter the house is tracked from a live capturing camera which is given as input, while selecting the background input of the system.

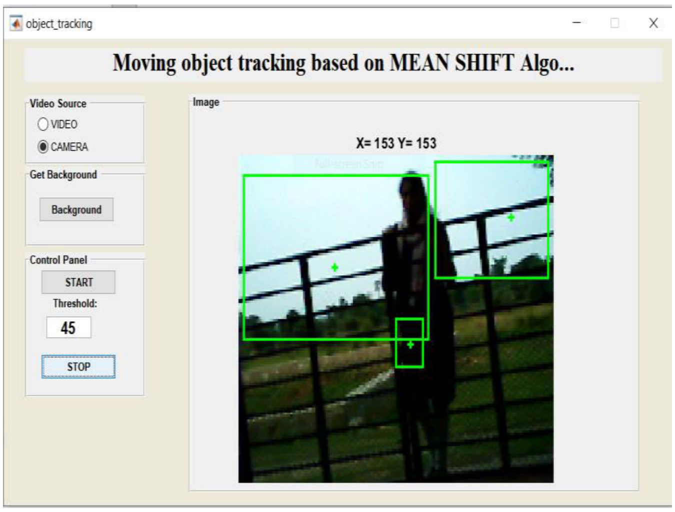


Figure : 6.3- Stimulated output in live

Figure: 6.4 shows the things moving in the balcony is tracked from a live capturing camera which is given as input, while selecting the background input of the system.

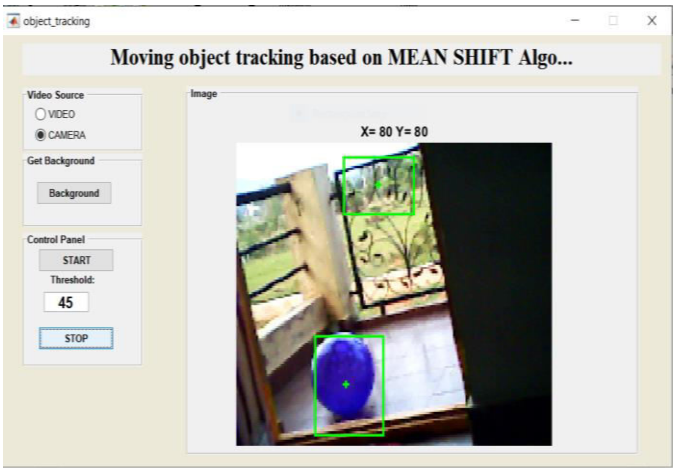


Figure : 6.4- Stimulated output in live

CONCLUSIONS

In our MATLAB based live Object Monitoring, Detection and Tracking system we have tracked multiple objects through multiple complex scenarios from videos captured by a single as well as stationary cameras with stationary or complex moving background. Test result shows that the real

time moving object detection and tracking system is very robust which is resulting in considerable improvement by reducing the object intersections and used in other applications like Video surveillance, Human activity analysis, road condition monitoring, airport safety, monitoring of protection marine board and etc. and long-term partial occlusions in scene, object pose or shape changes, fast motion changes, and cluttered background. The proposed live as well as stored dataset tracking scheme nearly reaches its limit when video scenes contain too many similar objects in terms of their appearance distributions in video scene, frequent intersections of objects at background and occlusions. Other limitation such as the computational speed and tracking drifts of the system requires future improvement.

REFERENCES

1. Z. H. Khan, I. Yu-Hua Gu, and G. Andrew, “Robust visual object tracking using multi-mode anisotropic mean shift and particle filters”, IEEE Trans. Circuits and System for video tech., vol. 21, no.1, January2011.

2. Q. Sumin and H. Xianwu, “Hand tracking and gesture recognition by anisotropic kernel mean shift,” in Proc. IEEE Int. Conf. Neural Netw. Signal Process, vol. 25. pp. 581–585, Jun. 2008.

3. S. Zhong and F. Hao, “Hand tracking by particle filtering with elite particles mean shift,” in Proc. IEEE Workshop Frontier Comput. Sci. Technol., pp. 163–16, Dec.2008.

4. T. B. Moeslund, A. Hilton, and V. Kruger, “A survey of advances in vision-based human motion capture and analysis,” *Computer Vision and Image Understanding*, vol. 104, pp. 90-126, 2006.

5. T. Yang, S. Z. Li, Q. Pan, and J. Li, “Real-time multiple objects tracking with occlusion handling in dynamic scenes,” in Proc. IEEE Conf. Comp. Vision Pattern Recognition, pp. 970–975, Jun. 2005.

6. A. Hampapur, L. Brown, J. Connell, A. Ekin, N. Haas, M. Lu, H. Merkl, S. Pankanti, A. Senior, C. F. Shu, and Y. L. Tian., "Smart video surveillance," IEEE Signal Process. Mag., vol. 22,no. 2, pp. 38–51, Mar. 2005.

7. L. Wang, W. Hu, and T. Tan. Recent developments in human motion analysis. Pattern Recognition, 36(3):585–601, March 2003.

8. D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-based object tracking," IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 5, pp. 564–577, May 2003.

9. R. T. Collins, "Mean-shift blob tracking through scale space", In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2003

10. D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. IEEE Trans pattern anal, machineIntell, 24:603–6192002.

BIOGRAPHIES

T.L. Spanadana pursued M.Tech from Amrita Sai Institute of Science and Technology, Vijayawada. Currently working as Assistant Professor at Raghu Institute of Technology, Vishakhapatnam, Andhra Pradesh, India.

N. Sree Laya Pursuing B. Tech in Electronics and Communication Engineering from Raghu Institute of Technology, Vishakhapatnam, Andhra Pradesh, India.

P. Puja Pranathi Pursuing B. Tech in Electronics and Communication Engineering from Raghu Institute of Technology, Vishakhapatnam, Andhra Pradesh, India.

G. Dileep Pursuing B. Tech in Electronics and Communication Engineering from Raghu Institute of Technology, Vishakhapatnam, Andhra Pradesh, India.

P. Sai Chand Pursuing B. Tech in Electronics and Communication Engineering from Raghu Institute of Technology, Vishakhapatnam, Andhra Pradesh, India.