Detection and Tracking of Objects under Varying Illumination Condition using Wavelet and Discrete Cosine Transform

Naimish Kasundra[1] Krishna K. Warhade[2]

[1]*Department of Electronics & Telecommunication, MIT College of Engineering, Pune, India.*

[2]*Department of Electronics & Telecommunication, MIT College of Engineering, Pune, India.*

Abstract: Most video processing applications require object tracking as it is the base operation for real-time implementations such as surveillance, monitoring and video compression. One of the persistent challenges in computer science has been tracking object under varying lighting condition. There are various difficulties in object tracking like noise in scene, illumination changes, occlusion effect, and pose variation into the scene. Illumination is an important concept in computer science application. It is well known that illumination variations on the observed scene and target are obstacles against robust object tracking causing the tracker lose the target. To address these challenges, we proposed an algorithm for moving object detection and tracking in the presence of illumination change. In this paper, we formulate the Discrete Wavelet Transforms (DWT) and Discrete Cosine Transforms (DCT) based tracking algorithms. Both of the proposed methods provide the illumination invariant features. An experimental result shows that the Discrete Wavelet Transforms (DWT) and Discrete Cosine Transforms (DCT) based tracking technique gives the high success rate. The proposed methods are tested on different video recorded at different location and results are compared with other available techniques.

Keywords: *Object detection, object tracking, illumination, discrete wavelet transforms (DWT) and discrete cosine transforms (DCT)*

I. INTRODUCTION

Visual object detection and tracking in video sequences is a central concern within the field of computer vision. Moving object tracking and detection is an important research area for wide spread application in diverse disciplines like the visual surveillance, human computer interaction, driving assistance system, image stabilization for digital cameras, security surveillance, autonomous navigation, traffic flow analysis and so on. Object tracking is a process of estimating the location of moving object in the current frame. Object detection is defining the process, to find the image region corresponding to that object. For this task, we initially have to detect the objects of interest and then track them across different frames while maintaining the correct identities. Difficulty in tracking objects can arise due to object motion, non-rigid object structure, object to scene and object to object occlusions, illumination and camera motion [1]. These difficulties are unavoidable in real-world environments and hence, we required algorithms which are robust to such condition. A number of techniques have been presented in the literature in the recent past to detect and track objects [1-9, 12-19]. Although partially successful many of these techniques fail on one or more of the following scenarios: the image is too crowded, the ambient lighting changes, they are too slow to compute, static objects become part of the background model. Also another difficulty is finding the exact position of moving object in each frame, Loss of information in projecting the 3-D image to 2-D plane, Complex object shapes, and Need for real-time processing.

Illumination is an important concept in visual arts. Illumination problem is defined as the degree of visibility of the object or change in appearance of the object with different lighting condition. The placement of the light sources can make a difference in the type of any object that is being presented. Outdoor surveillance systems suffer heavily from the change of weather conditions. Rain, sunset, sunrise, as well as floating clouds can have a dramatic impact on the scene illumination. Hence, they will degrade the performance of object detectors and trackers, if these factors are not accommodated properly. Another challenge is the occlusion effect. It is also defined as hidden (occluded) object. In dynamic scenes, the moving objects exhibit many spatial configurations relative to other objects. A relative depth ordering on the objects and the scene background structures is imposed along the lines of sight when observed from a view point. Such a depth ordering leads to the partial or complete viewing obstruction of some of the object of interest by others and the phenomenon is also known as occlusion. Inter-object problem becomes acute when two or more objects enter into the scene while occluding each other,

- Small scene structures: small (thin) objects in the scene such as trees or streetlights break moving objects.

- Large scene structures: because of large scene structures such as towers or buildings, moving objects may disappear completely for a period of time, and then re-appear.

Object tracking algorithm can be categorized into three types. These are point tracking, kernel tracking and silhouette tracking [1]. In point tracking, detected object in consecutive frames are represented by a set of points and kalman filter is widely used in the point based feature tracking [2]. Point tracking is complicated in the presence of occlusion, entries and exists of an object. Kernel tracking is associate with the object shape and appearance. These algorithms differ from the others based on the method used to estimate the object motion and the numbers of the objects tracked [1]. Kernel based object tracking is usually represented with rectangular or elliptical shape of kernel. Silhouette tracking methods provide an accurate shape of the objects [2]. An advantage of the silhouettes is flexibility to handle a large variety of object shapes.

The simplest and effective method for object tracking is proposed by Oksam Chae et al. [2], which is based on the parametric edge of the object. Image information lies on the edge of different objects. Edge information is less sensitive to noise and is more consistent than the pixel values. Also edge-based methods are more robustness as compared to pixel intensity based methods and less sensitive to illumination variation than intensity features. Object boundary shows sharp changes in image intensities. For moving edge detection, Julius et al. [2] have proposed a method to detect the object. Moving edges are extracted from video and segmented using an efficiently designed edge class. In this paper, first the background edges are modeled with an initial reference sequences. Change in background scene and reference edges are updated which will take care of dynamic background. The matching between segment of input edge and reference edge can tolerate fluctuation or camera calibration. Update and detection can be performed by using the list of edge, without accessing input frame. Segment based edge pixels representation is fast compare to all the pixels in the image. Tianzhu Zhang et al. have proposed a robust visual tracking algorithm using multi-task sparse learning [3]. This algorithm handles the particles (target Candidates) independently. In this algorithm, object tracking is to defined or formulated as a particle filter framework (as a multi-task tracking). They have also uses the particle filter to track the target object. Then the particles are randomly sampled according to Gaussian distribution. These particles are represented as a linear combination of updated dictionary template. As particles are densely sampled around the target stage, their representation will be sparse. This is more robust representation for particles. This convex optimization problem can be solve using accelerated proximal gradient method. This algorithm improves the tracking performance and overall computational complexity. Ross et al. [4] proposed an adaptive tracking method that shows robustness to large changes in pose, scale, and illumination via incremental principal component analysis. The online multiple instance learning algorithms successfully tracked an object in real time where lighting conditions change and object occlusion occurs. Object detection and tracking under changing lighting (illumination) conditions studied by Wagas Hassan et al. [5] is based on orientation of the edge. Under lighting conditions, edges are more stable than both edge magnitude and color. Adaptive edge orientation method considers the orientation of the edge rather than the intensity and there is no dependency on color features. This algorithm is also applied to the highly variable lighting video sequence and provides the better results.

The mean shift algorithm is a nonparametric task that shifts each sample to the model to which it statistically belongs [6]. It can be generated like clustering algorithms [7] and was originally developed for data clustering algorithm. An advanced object detection and tracking system was constructed based on a Gaussian Mixture Model (GMM). It requires a fixed number of Gaussian components and stable component [8]. Gaussian Mixture Model with using the Expectation Maximization (EM) algorithms work well in situations where the observation data is incomplete [8]. The optical flow method uses characteristics of flow vectors of moving objects over time to detect moving regions. However, most optical flow methods are with higher complex computation. Rama Chellappa et al. have proposed a method for object detection and tracking using multiple smart cameras [9]. They use the method of background modeling (background subtraction) for moving object detection and tracking. Gaussian distribution functions are used to remove the global changes into the scene such as illumination or camera jitter. This will reduce the effects of illumination into the Scene. Ridder et al. [10,11] modeled each pixel with a Kalman Filter which made their system more robust to lighting changes in the scene. While this method does have a pixel-wise automatic threshold, it still recovers slowly and does not handle bimodal backgrounds well. Koller et al. [10, 12] have successfully integrated this method in an automatic traffic monitoring application. Francois Bardet et al., [13] have proposed a method for illumination invariance where multi-objects are jointly tracked through a Markov chain Monte-carlo Particle Filter (MCMC PF). To allow the object to enter or leave the scene khan et al. [14] have extended their Markov chain Monte-Carlo particle Filter (MCMC PF) method to track the variable number of objects. This extended method is reversible Jump Markov Chain Monte-carlo (RJ MCMC) sampler. Reversible Jump Markov chain Monte-Carlo Particle Filter has become a popular algorithm for real-time tracking. RJMCMCPF samplers allow the object classification as well as detect the object shapes. An RJ MCMC PF sampler algorithm overcomes the problem of Isard et al.,[15] They have proposed a method using SIRPF

(Sequential Importance Resampling Particle Filter) .Online tracking methods under the various outdoor lighting variations with moving cameras are studied by Yanli Liu and Xavier Granier [16]. To design the algorithm, they have assumed a strong correlation in lighting over large spatial and temporal extents. This approach does not require any priori knowledge of 3D scene and works with moving view point. Also algorithm provides a user with an augmented reality experience with its general purpose camera. Without knowledge of lighting direction, algorithm cannot deal with indoor scenes.

Histogram and statistic based methods are sensitive to lighting changes but invariant to change the object motion. The pixel difference comparisons are more robust to lighting changings. The edge-based method is invariant to lighting changes and is generally used with histogram based method, but this approach is computationally expensive. Hence we proposed a new approach for detection and tracking under different lighting condition. We used the two-dimensional Discrete Wavelet Transform (DWT) and Discrete Cosine Transform (DCT) instead of using the traditional methods. Both of these approaches provide illumination invariant features. 2D- Discrete Wavelet Transform (DWT) divides the image sequences into the high- and low frequency components.

Most of the information lies into the high frequency whereas noise lies into the low frequency. In this paper, we use 2D discrete wavelet transform which converts images into “sub-bands. 2D-Discrete Cosine Transform (2D-DCT) is used in various fields. It also removes the illumination into the scene. A DCT is a Fourier-related transform, but using only real numbers. The DCT transforms signals from a spatial domain representation into a frequency domain representation. 2D-Discrete Cosine Transform (2D-DCT) operates left-to-right and top-to-bottom manner on the block. It is used in signal and image processing application, especially for lossy data compression.

The paper is structured as follows. In Section II mean shift algorithm based object tracking and 2D-Cepstrum based tracking methods are described. Section III presents the proposed discrete wavelet transform and discrete cosine transform based object tracking methods. Results of the proposed methods are in Section IV. Paper is concluded in Section V.

II. MEAN SHIFT AND 2D-CEPSTRUM ALGORITHM

A. Mean shift algorithm:

A mean-shift algorithm is a simple interactive method which shifts the data points in its neighborhoods location and it is a basic algorithm that may be used data clustering, analysis and computer vision application. Non-rigid object tracking using mean shift algorithm proposed by Dorin Comaniciu et al. [17] implemented on our video datasets. In this algorithm, tracking of non-rigid objects characterized by the color and/or texture and it is used as the object of interest. Tracking algorithm divided into two modules. First module is used to detect and localize the initial frame of the objects and second module is used for periodic analysis of an object across the different frames. The mean shift iterations are employed to find the target candidate that is the most similar to the target, with the similarity being expressed by a metric based on the Bhattacharyya coefficient. The Chernoff and Bhattacharyya bound have been employed in [18] to determine the effectiveness of edge detectors. Also the kullback divergence [19] has been used to find the pose of an object in an image. First stage of this algorithm is to initialize the location of the target in the current frame and then derive the weight function according to the location of the target. According to the mean-shift vector in eqation-1, find the new location of the target $Yj+1$ and update the weight according to the new location of the target.

$$Yj+1=\frac{\sum_{i=1}^{n} \mathrm{Xi}\, g\left(\left\|\frac{\mathrm{yj-xi}}{\mathrm{h}}\right\|^{2}\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{\mathrm{yj-xi}}{\mathrm{h}}\right\|^{2}\right)}\ ,\ \mathrm{j=1,\ 2\ldots,\ n} \qquad (1)$$

Where, g is the location of kernel, y1 is the center of initial kernel. Various test sequence showed the good tracking performance obtained with high computational complexity.

B. 2D-Cepstrum algorithm:

Fuat corun and A. Enis Cetin have proposed the matrix based method for object tracking under illumination variations using 2D-Cepstrum characteristics of the target. They also describe the co-difference and covariance matrix based object tracking algorithm in [20]. The covariance of feature vectors describing the target is called covariance matrix. Co-difference tracking method describes the co-difference matrix to model the moving objects or target. In both of these methods, the aim is to find the region in a given image frame having the minimum distance from the target matrix, and assign this region as the moving target at that frame. The first stage of these algorithms is to find the feature images and vectors and then co-variance matrix and co-difference matrix is found out by using the feature vectors. The next stage is to estimate the distance metric and target location of the object. This operation is repeated for each frame. Then this algorithm is analyzed by using 2D-Cepstrum [20-21] analysis. 2D-Cepstrum is an amplitude invariant feature extraction method. So, cepstral domain co-efficient of a region remains unchanged under the light intensity variations. This property of cepstrum provides robustness to illumination variation at the target region.

III. PROPOSED ALGORITHM

A. *Background modeling for moving object detection:*

To detect any object background modeling technique is used. It can be used for moving as well as static object detection. One point is that the algorithm based on background subtraction for the video detection and tracking of the moving objects need to convert data format also. In video surveillance systems the first step in detecting objects is to separate the foreground image from the background image to detect motion. So, we should able to detect the regions (apparent shape) of independently moving objects regardless of their speed, direction or texture. An object has two categorize: rigid object and non-rigid object [1, 22]. The major problem comes with non-rigid object detection like human, because human has deformable object. In this paper, we have used background subtraction method by frame differencing. As we want to detect moving object in the video, we can subtract each consecutive frame to get the moving regions in the video. The important thing is that the background is static. If the background is changing, then we have to use EM or GMM based filtering approach. Background subtraction can be classified as: the pixel level, the region level, and the frame level. According to the complexity of the foreground object, we can use frame differencing methodologies. Detecting a vehicle is somewhat easy because they are moving with continuous speed as in human we cannot judge the movement and speed of it.

B. *Extracting the image using 2-D Discrete Wavelet Transform:*

Wavelets are defined by a mathematical expression and are drawn as continuous and infinite. However to use them with our digital signal, they must first be converted to *wavelet filters* having a finite number of discrete points. Discrete Wavelet Transform used in various fields such as signal processing, computer vision, image processing and compression, etc. For image processing, it provides multi-resolution image and can decompose an original image into different sub band images including low- and high-frequencies. This is illustrated in the Figure 1.

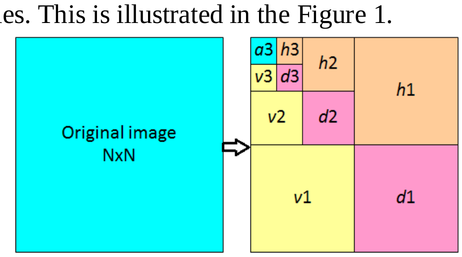

Fig.1: original image of size N*N and its DWT image into the different sub-bends

Whereas: *d* = diagonal detail (HIGH/HIGH)
h = horizontal detail (HIGH/LOW)
v = vertical detail (LOW/HIGH)
a = approximation (LOW/LOW).

The possibility of using two–dimensional discrete wavelet transforms to find out the object feature which is invariant to disturbances caused due to illumination and motion has been explored in our proposed method. Discrete wavelet transforms are used to compute accurate and efficient estimate of the object structure in each video frames, mainly the edges of the object. This feature of the object obtained using discrete wavelet transform which is invariant to illumination condition.

Filtering is required in video surveillance application. Filtering is nothing but the smoothing of the images. The purpose of smoothing is to reduce noise and improve the quality of the image. Here we used the Gaussian filter to remove the noise from the scene. After applying the Gaussian filter, we use the threshold to smooth the image. If the threshold was too low or high, tracking results are not much better. Thus, all the thresholds were kept fixed throughout the testing of the video sequence under the different lighting condition, viewing angles, quality and video resolution and scene density.

C. *Extracting the image using 2-D Discrete cosine Transform:*

There are many transforms available, most of which are very slow. This is important to consider because of video analysis requires real-time encoding and decoding. Discrete cosine Transform is used in number of applications of science and engineering fields. A DCT is a Fourier-related transform (similar to the discrete Fourier transform (DFT), but using only real numbers). The DCT transforms signals from a spatial domain representation into a frequency domain representation. In an image, most of the energy will be concentrated in the lower frequencies, so if we transform an image into its frequency components and throw away the higher frequency coefficients, we can reduce the amount of data needed to describe the image without sacrificing too much image quality. Yn represent the DCT,

$$\mathrm{Yn} = 1/N \sum_{k=0}^{N-1} \mathrm{Xk} \cos\left[\frac{\pi}{\mathrm{N}}\left(\mathrm{k}+\frac{1}{2}\right) \mathrm{n}\right] ;\ \mathrm{k}=0, 1,.., \mathrm{N}-1 \tag{2}$$

To rebuild an image in the spatial domain from the frequencies obtained above, we use the IDCT: the DCT is perfectly reversible and we do not lose any image definition until we start quantizing coefficients. The possibility of using two–dimensional discrete cosine transforms to find out the object feature which is invariant to illumination changes. After applying the 2D-DCT, we make the first row and column elements of the frame to zero that means DC – coefficient makes the zero {J (1:1) = 0}; the output of DCT coefficients contains integers and its range from -1024 to 1023. Mean values are taken to remaining matrices of the image and this values are put into the 4*4 matrices of the image to remove the illumination effect and blurring. Then apply the 2D-IDCT, to find the image.

D. Algorithm Processing Flow:

Initially, we take the video as an input with different resolutions and then pre-processing is done. In pre-processing step frame differencing technique is used to remove the background from the scene. Apply the two-dimensional discrete wavelet transform or two- dimensional discrete cosine transform to find out the object feature which is invariant to disturbances caused due to illumination effects. After applying the two-dimensional discrete wavelet transforms, first level wavelet decomposition is performed using daubechies second order filter. Then HH, HL, and LH components are combined to form the feature of the target. Apply the smoothing and thresholding function and then two-dimensional inverse discrete wavelet transform is applied to rebuild the object features. Using the two- dimensional discrete cosine transform algorithm, DC –Coefficient of the frame makes the zero and then takes the mean value of remaining frame. These average values put into the metrics and track the objects into the multiple frames.

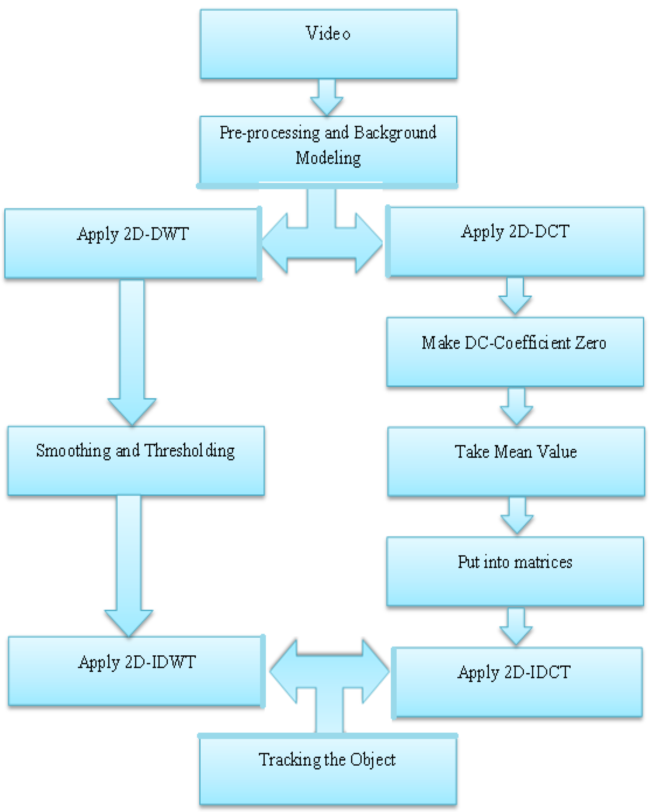

Fig.2: DWT and DCT algorithms

IV. EXPERIMENTAL RESULTS

In this section, we present experimental results that validate the effectiveness and efficiency of our proposed method. We also conduct the comparisons between our proposed method with the Mean-shift algorithm and 2D-Cepstrum method.

A. Test video sequences:

The proposed algorithm has been tested on seven video sequences, taken at different times of the day. We used four video sequences on normal day condition. Video-1 was taken from PETS datasets. We consider the 81 frames into this dataset for the tracking. In video-2 object moves from bright light to building shadow area. Video-3 was recorded into the normal dark light condition and Video-4 was taken into the sun light condition. Remaining three video sequences was taken in heavy bright light condition. These video sequences recorded in indoor environment. Video-5 was recorded in bright light condition with the fast motion. Video-6 was taken with the two flash light and slow motion of object, and video-7 was recorded with slow motion of object and static condition of an object when flash light is on. Threshold was kept fixed throughout the testing of the different video sequences (different quality, lighting condition, video resolution, scene density and viewing angles) for our proposed methods. The proposed algorithm compared with the two different techniques and it gives the high tracking results. All our experiments are done using MATLAB R2013a on Intel P4, CPU 2.40 GHz with 512 MB RAM.

B. Quantitative Comparisons:

The results obtained by our proposed methods are compared with the two methods which are presented in the section-II. Table 1 shows the results obtained using the Mean-shift algorithm [17] on different video sequences with a large number of frames. Average detection rate of mean shift algorithm is 63.71%. Table 2 shows the results obtained using the 2D-Cesptrum method [20] on different video sequences. Average detection rate of this algorithm is 69.65%. Table 3 shows the results obtained using the Discrete Cosine Transform (DCT) on different video sequences with a large number of frames. Average detection rate of the proposed algorithm is 83.36%. Table 4 shows the results obtained using the Discrete Wavelet Transform on different video sequences. Average detection rate of this algorithm is 83.22%. Video-4 which was recorded into the sunlight area and has a total numbers of frames 194. Using the mean-shift algorithm detected frames are 168 and missed frames are 26. So the detection rate using mean shift algorithm is 86.59 %. Tracking results at frames {33, 55, 85, 100, 110, and 130} for video-4 are presented in Figure-3. Figure-3(a) shows that, tracking is failed during the frames 100 and 133. In the 2D-Cesptrum method detected frames are 172 and missed frames are22. Detection rate.

TABLE 1: RESULTS OBTAINED BY MEAN-SHIFT ALGORITHM [17]

VIDEO	TOTAL FRAMES	DETECTED FRAMES	MISSED FRAMES	DETECTION RATE
Video-1	81	75	6	92.59
Video-2	89	77	12	86.51
Video-3	70	56	14	80.00
Video-4	194	168	26	86.59
Video-5	134	41	93	30.59
Video-6	290	92	198	31.72
Video-7	305	116	189	38.03

TABLE 2: RESULTS OBTAINED BY 2D-CESPTRUM ALGORITHM[20]

VIDEO	TOTAL FRAMES	DETECTED FRAMES	MISSED FRAMES	DETECTION RATE
Video-1	81	73	8	90.12
Video-2	89	73	16	82.02
Video-3	70	65	5	92.85
Video-4	194	172	22	88.65
Video-5	134	70	64	52.23
Video-6	290	122	168	42.06
Video-7	305	121	184	39.67

TABLE 3: RESULTS OBTAINED BY DCT ALGORITHM

VIDEO	TOTAL FRAMES	DETECTED FRAMES	MISSED FRAMES	DETECTION RATE
Video-1	81	61	20	75.30
Video-2	89	82	7	92.13
Video-3	70	60	10	85.71
Video-4	194	184	10	94.84
Video-5	134	99	35	73.88
Video-6	290	235	55	81.03
Video-7	305	246	59	80.66

TABLE 4: RESULTS OBTAINED BY DWT ALGORITHM

VIDEO	TOTAL FRAMES	DETECTED FRAMES	MISSED FRAMES	DETECTION RATE
Video-1	81	59	22	72.83
Video-2	89	84	5	94.38
Video-3	70	62	8	88.57
Video-4	194	182	12	93.81
Video-5	134	95	39	70.89
Video-6	290	238	52	82.06
Video-7	305	244	61	80.00

using the 2D-Cesptrum algorithm is 88.65 %. Tracking results of the same frames for video-4 are presented in Figure-3(b). Frame numbers 55 and 110 is missed during the tracking. Our proposed algorithm gives the better results compared to mean shift algorithm and 2D-Cesptrum method. Using the discrete cosine transform, we get the 184 detected frames and 10 missed frames. Detection rate of discrete cosine transform for video-4 is 94.84%. Tracking results at frames {33, 55, 85, 100, 110, and 130} for video-4 are presented in Figure-3(c). Using the discrete wavelet transform, detected frames are 182. Detection rate for video-4 is 93.81using Using the discrete wavelet transforms. Results of our proposed method using video-5 are also shown in Figure-4. Video-5 was recorded with the heavy lighting condition and also motion of an object is fast. Using the mean-shift algorithm, detected frames are 41 out of 134 frames and missed frames are 93. Figure-4(a) shows the result of frame number {33, 55, 85, 100, 110, and 130}. Mean-shift algorithm failed under lighting condition and it is unable to track the object. So the detection rate is low, which are shown in Table 1. Figure-4(a) shows the frame numbers {85, 100, 110, and 130} are missed using mean-shift algorithm. 2D-Cesptrum algorithm is also tested on same video datasets. Using this algorithm detected frames are 70 out of 134 frames. Figure-4b) shows the result of frame number {33, 55, 85, 100, 110, and 130}. Our proposed methods are also tested on same video dataset. Discrete cosine transform algorithm gives the detection rate 73.88%. In this algorithm detected frames are 99 (Video-5) which are high compared to mean-shift algorithm and 2D-Cesptrum method. Result of frame number {33, 55, 85, 100, 110, and 130} is shown in Figure-4(c). Figure-4(c) shows that object is tracked by DCT based algorithm where as mean-shift algorithm and 2D-Cesptrum methods are failed. Using the discrete wavelet transform, detected frames are 95 for Video-5. Figure-4(d) shows that object is tracked using the discrete wavelet transform. Result of frame number {33, 55, 85, 100, 110, and 130} is shown in Figure-4(d) where as mean shift algorithm and 2D-Cepstrum methods are failed.

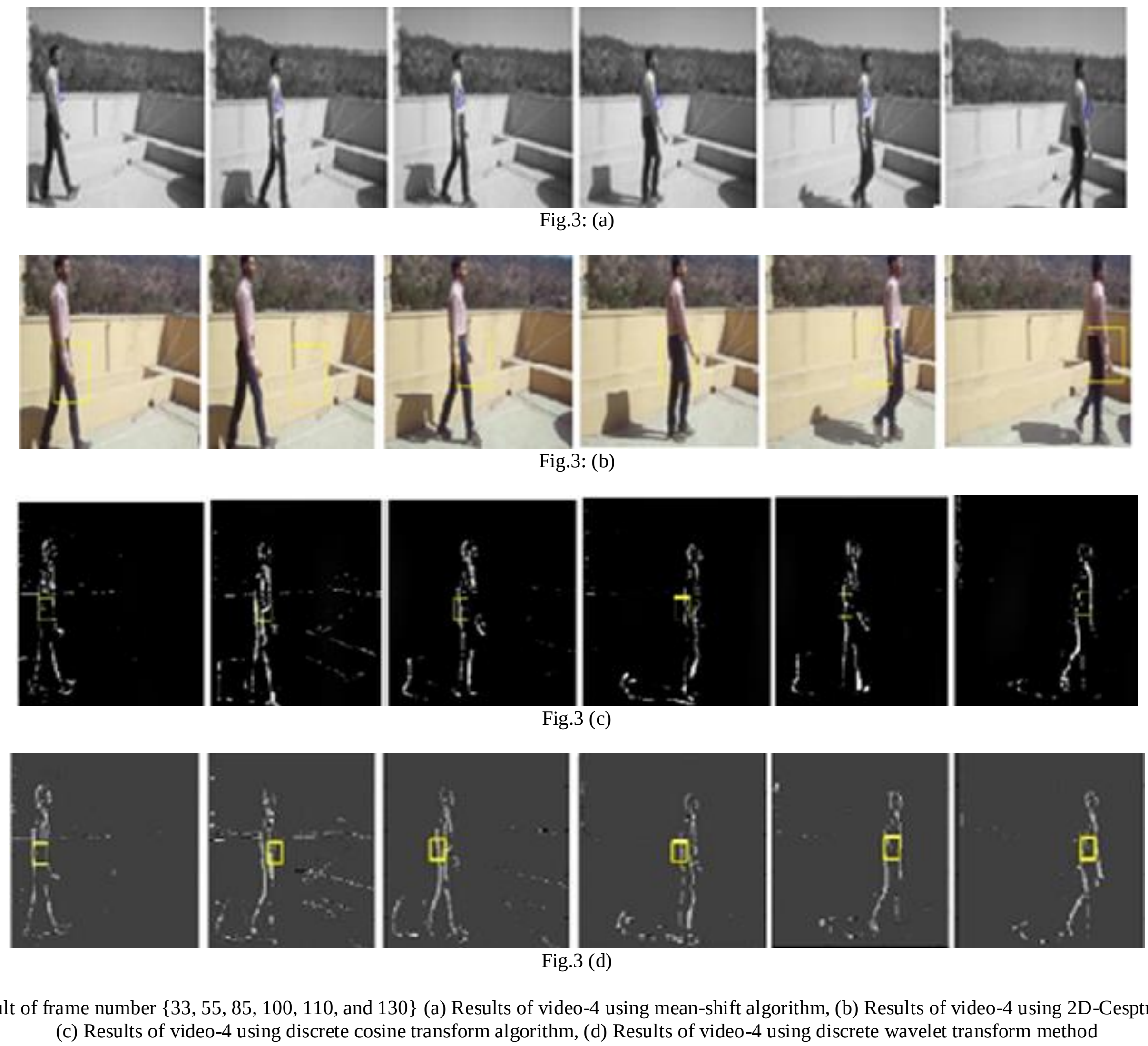

Fig.3: (a)

Fig.3: (b)

Fig.3 (c)

Fig.3 (d)

Fig.3: Result of frame number {33, 55, 85, 100, 110, and 130} (a) Results of video-4 using mean-shift algorithm, (b) Results of video-4 using 2D-Cesptrum method, (c) Results of video-4 using discrete cosine transform algorithm, (d) Results of video-4 using discrete wavelet transform method

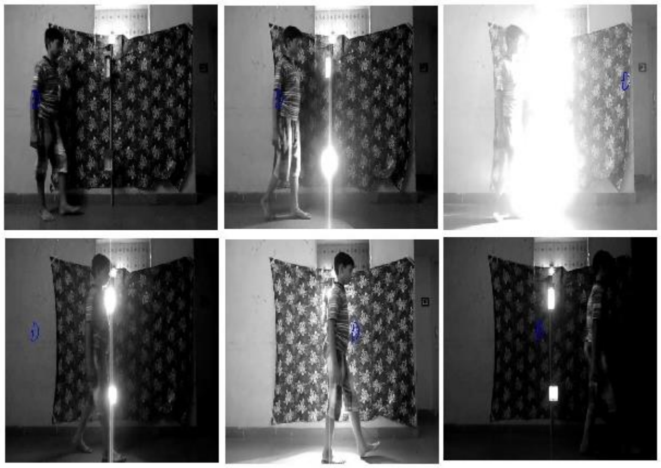

Fig.4: (a) Result of frame number {33, 55, 85, 100, 110, and 130} using mean-shift algorithm

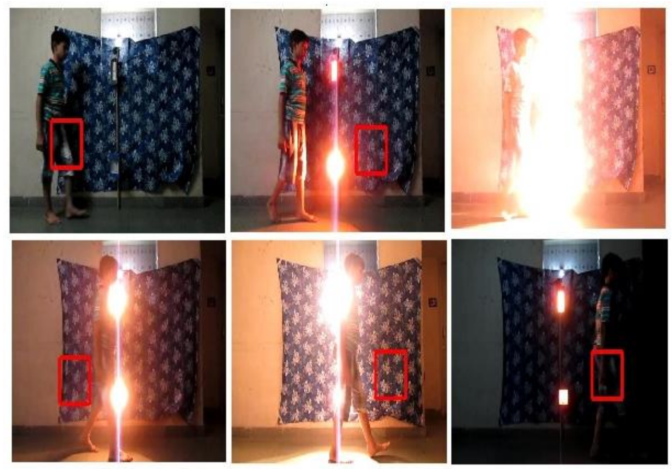

Fig.4: (b) Result of frame number {33, 55, 85, 100, 110, and 130} using 2D-Cesptrum method

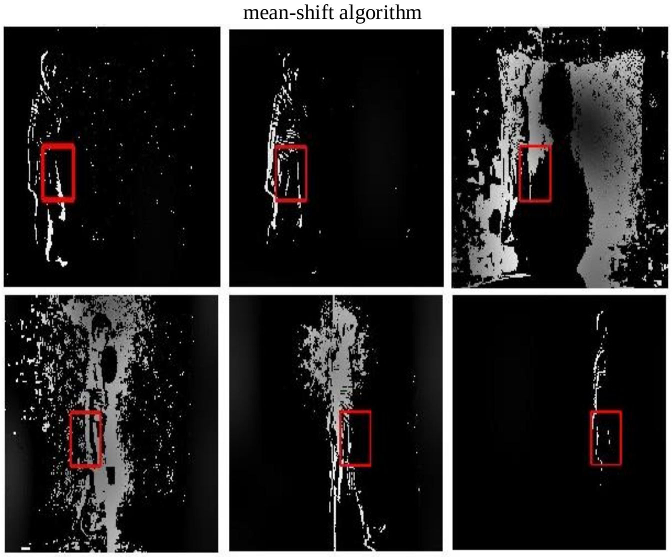

Fig.4: (c) Result of frame number {33, 55, 85, 100, 110, and 130} using discrete cosine transform algorithm

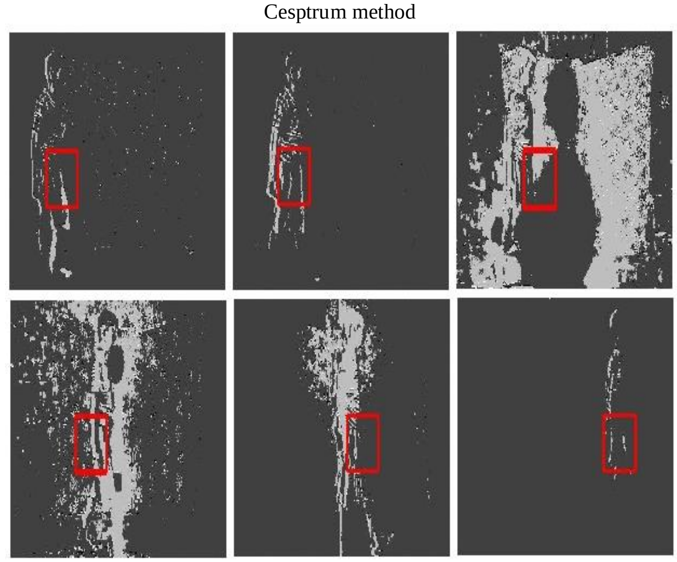

Fig.4: (d) Result of frame number {33, 55, 85, 100, 110, and 130} using discrete wavelet transform algorithm

V. CONCLUSION

A robust object detection and tracking technique has been presented in this paper. Illumination is the major problem for object detection and its elimination is the challenging task. In this paper, we address this problem by developing the Discrete Wavelet Transforms (DWT) and Discrete Cosine Transforms (DCT) based tracking algorithms. The Discrete Wavelet Transforms (DWT) gives the four sub-bend components of the images: HH, HL, LH and LL. We add these components and apply the threshold to remove the blur into the frames. In this paper threshold was kept fixed for all the video sequences. Here, we describe the results obtained using the Discrete Cosine Transforms (DCT) based tracking algorithms and it removes the illumination into scenes by tacking the mean values of the frame matrices. Our proposed algorithm has been tested on different video sequences with a large numbers of the frames. Discrete Wavelet Transforms (DWT) and Discrete Cosine Transforms (DCT) methods gives better performance under the illumination effect compare to Mean shift algorithm and 2D-Cepstrum method. The study has demonstrated that, the proposed tracking method has a high accuracy as compared to other available methods.

VI. REFERENCE

[1] A. Yilmaz, O. Javed and M. Shah. “Object tracking: a survey,” ACM Computing Surveys Vol. 38, 2006.

[2] M. Murshed, M. Ali Akber Dewan and O. Chae. “Moving object tracking- A parametric edge tracking approach,” In IEEE conference on computer and information technology.,pp. 471-476, 2009.

[3] T. Zhang, B. Ghanem and S. Liu.“Robust visual tracking via multi-task sparse learning”.In IEEE con on CVPR, PP.2042-2049, 16-20 June 2012.

[4] D. Ross, J. Lim, R.-S. Lin, and M.-H, Yang, Incremental learning for robust visual tracking, IJCV, Vol. 77, 2008, pp.125 141.

[5] W. Hassan, P. Birch, Chris and Rupert Young. “Illumination invariant stationary object detection”, In IEEE computer vision, IET, vol.7, February, 2013.

[6] Fukunaga K, Hostetler LD (1975) The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition. IEEE Trans. Inf. Theory vol, 21(1): pp. 32–40.

[7] Cheng Y. Mean Shift, Mode Seeking, and Clustering. IEEE Trans. Pattern Anal. Mach. Intell. 17(8): 790–799, 1995.

[8] I. Whoang, Jeong Heon Kim and Kwang Nam Choi. “Object Tracking Using Maximum Colour Distance Under Illumination Change”, Int J Adv Robotic Sy, Vol. 9, 212:2012

[9] Aswin C. S., Ashok V and Rama Chellappa. “Object tracking, detection and recognition for multiple smart cameras,” Proceedings of the IEEE, vol. 96, no.10, pp.1606-1624, 2008.

[10] Christof Ridder, Olaf Munkelt, and Harald Kirchner “Adaptive Background Estimation and Foreground Detection using Kalman-Filtering,” Proceedings of International Conference on recent Advances in Mechatronics, ICRAM’95, UNESCO Chair on Mechatronics, 193-199, 1995.

[11] W.E.L. Grimson, Chris Stauffer, Raquel Romano, and Lily Lee. “Using adaptive tracking to classify and monitor activities in a site,” In Computer Vision and Pattern Recognition 1998(CVPR98), Santa Barbara, CA. June 1998.

[12] D. Koller, J. Weber, T. Huang, J. Malik, G. Ogasawara, B. Rao, and S. Russel. “Towards robust automatic traffic scene analysis in real-time.” In Proc. of the International Conference on Pattern Recognition, Israel, November 1994.

[13] F. Bardet, T. Chateau and D. Ramadasan. “Illumination aware MCMC particle filter for long-term outdoor multi-object simultaneous tracking and classification”, In International IEEE Conference on Computer Vision, pp. 1623-1630, 2009.

[14] Z. Khan, T. Balch, and F. Dellaert. MCMC-based particle filtering for tracking a variable number of interacting targets. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27:1805 – 1918, 2005.

[15] M. Isard and A. Blake. Condensation – conditional density propagation for visual tracking. IJCV : International Journal of Computer Vision, 29(1):5–28, 1998.

[16] Y. Liu and X. Granier. “Online tracking of outdoor lighting variations for augmented reality with moving cameras”. In IEEE Transactions on visualization and computer graphics , vol. 18, no. 4, pp. 573-580, 2012.

[17] D. Comaniciu, V. Ramesh and Peter Meer.“Real time tracking of non-rigid objects using mean shift”, IEEE CVPR, 2000.

[18] S. Konishi, A.L. Yuille, J. Coughlan, S.C. Zhu, Fundamental Bounds on Edge Detection: An Information Theoretic Evaluation of Dierent Edge Cues," IEEE Conf. on Comp. Vis. and Pat. Rec., Fort Collins, 573-579, 1999.

[19] P. Viol, W. M. Wells III, “Alignment by maximization of mutual information,”IEEE Int’lconf. comp. vis., Cambridge, Massachusetts, pp.16-23, 1995.

[20] F. Cogun and A. Enis Cetin.“Object tracking under illumination variation using 2D-Cepstrum characteristics of the target”. IEEE Con. On MMSP’10, pp.521-526, Oct 4-6, 2010.

[21] A. V. Oppenheim and R. W. Schafer. “Frequency to quefrency: A history of cepstrum,” In IEEE signal processing mag., pp.95-106, 2004.

[22] M. J. Hossain, M. A. A. Dewan and O. Chae “Moving object detection for real time video surveillance: An edge based approach,” IEICE Tran Commun., vol. E90-B, no 12, pp. 36543664, 2007.

Naimish Kasundra is pursuing his Master’s degree in “VLSI and Embedded Systems” from Maharashtra Institute of Technology College of Engineering, Pune, India. He received his bachelor’s Degree (Electronics and Communication Engineering) from AITS, Rajkot, India.

Krishna K. Warhade received the Bachelor of Engineering in Electronics in 1995 and Master of Engineering in Instrumentation in 1999 both from Shri Guru Govind Singhaji Institute of Engineering and Technology, Nanded, and Ph.D. in November 2010 from the Department of Electrical Engineering, Indian Institute of Technology Bombay, India. He has 20 years of experience in teaching and research. He is currently working as a Professor in the Department of Electronics and Telecommunication Engineering, Maharashtra Institute of Technology College of Engineering, Pune, India. His research interests are in the area of signal processing, image processing, video segmentation, video retrieval, wavelets, precision agriculture and biomedical.