



Similar Object Detection and Tracking in H.264 Compressed Video Using Modified Local Self Similarity Descriptor and Particle Filtering

Srinivasan Kuppuswamy^{1*}

Balamurugan Panchanathan²

¹*Sri Krishna College of Technology, Coimbatore, Tamilnadu, India*

²*Mount Zion College of Engineering and Technology, Pudukottai, Tamilnadu, India*

* Corresponding author's Email: srinivasank0784@gmail.com

Abstract: Object tracking is a dynamic optimization process based on the temporal information related to the previous frames. Proposing a method with higher precision in complex environments is a challenge for researchers in the field of study. In this paper, we have proposed a novel framework for similar object tracking. In our proposed technique we are considering both PETS and Cricket video for input video sequence. The primary steps of suggested technique are preprocessing, background subtraction and segmentation, similar object detection and object tracking. In the preprocessing stage, the adaptive median filter is used to remove the noise from each frame. Next, the foreground and background images are separated and then segmentation of object is carried out by using morphological operation. For similar object detection, the recommended technique uses the modified local self-similarity descriptor and similar object tracking is done by a particle filter. The performance of the suggested technique is evaluated by means of precision, recall, F-measure, FPR, FNR, PWC, FAR, similarity, specificity, and accuracy. An experimental result shows that the proposed technique attains the maximum tracking efficiency for both videos when compared to the existing techniques.

Keywords: Adaptive median filter, Morphological operation, Local self-similarity descriptor, Particle filter.

1. Introduction

Object tracking assumes a basic part in PC vision with wide usages in video observation, human PC cooperation, vehicle route, and so on. Given the target object at the main edge, the objective of tracking is to find this object for the ensuing casings [1]. Ongoing object tracking is a standout amongst essential prerequisites for independent portable robots. Vision-oriented object tracking needs chose focuses to be tracked and comparing focuses to be looked in every casing [2]. At first video, occasion detection needs detecting and tracking objects initially, and after that perceiving what is going on around those tracked objects [3]. Target objects can be characterized by their appearances, for example, shading, surface, edges, and shape data, which give trademark data about the object. Tracking algorithms then decide the correspondence of the object area in back to back pictures by advancing the pre-decided

comparability utilitarian [4]. There are different techniques utilized for object detection and tracking. Objects are regularly detected by utilizing point detector systems, background subtraction, division, or learning components [5].

Such methods included the constant use of a detection algorithm in individual edges and the relationship of the detections crosswise over edges [6]. A post-processing task refines the detection by evaluating and fitting an arrangement of ovals that characterize the motion of the object to the past arrangement of moving locales [7]. After object detection, its tracking is done. A successful tracker ought to handle the varieties both of the objective and the foundation well [8]. The level set technique is one such capable instrument for object tracking in picture successions because of its adaptability with reference to topological variations of the shapes [9]. And additionally, the particle filter oriented tracker is an inspecting oriented tracking technique, which can

adapt well to the non-direct and non-Gaussian tracking issues [10]. In these days division oriented tracking have pulled in incredible consideration in the field of object tracking. It could give a more exact frontal area/foundation partition, contrasted with established tracking techniques which frequently utilize a bouncing box to characterize the objective question [11].

Despite the fact that similarly object detection and tracking is done utilizing different apparatus [12]. And also tracking the deformable object in reasonable situations is still challenging in light of the fact that the objective appearance may vary continually amid moving. for example, mutilation, turn and scaling [13]. These may bring about confuse or lose object amid tracking and lessen the precision of tracker [14]. Visual tracking is planned as an online twofold characterization issue and the objective appearance designs are overhauled adaptively utilizing the pictures tracked from the past casings [15]. These strategies could be favored for object detection and tracking. The overall objective of the suggested technique is similar object detection and tracking in H.264 video. To detect the similar objects the proposed method use the modified local self similarity descriptor, here we have modified the local self similarity descriptor with the help of correlation value. This is the new features of the recommended technique. The main advantage of the proposed technique is finding the similar objects and tracking in an effective manner when compared with existing technique.

The remaining of this paper is organized as follows. Section 2 gives some brief background of researches related to the proposed technique. Section 3 describes the proposed similar object detection and tracking. The experimental results and discussions of the proposed approach are presented in Section 4. Finally, conclusions are summarized in Section 5.

2. Literature survey

I. Elafi, et al. [16] acquainted another technique to conquer that issue. Without a doubt, another ongoing methodology was built up in view of the molecule channel and foundation subtraction.

D. Riahi and G. A. Bilodeau [17] exhibited a powerful online multiple objects tracking (MOT) method in view of different elements. Their approach could deal with MOT issues, such as long haul and substantial impediments and close comparability between aimed appearance designs.

H. Zhao, et al. [18] broadened sparse representation based classification (SRC) and multi-feature hashing (MFH) into various object tracking

assignment and proposed a joint appearance model of SRC and MFH, which went for separating distinctive objects viably. That pairwise appearance demonstrates concentrated on discernable components from two focuses without focusing on different targets or foundations.

K. Ahmadi and E. Salari [19] showed a novel algorithm for detecting and tracking little diminish focused in Infrared (IR) picture arrangements with low Signal to Noise Ratio (SNR) in light of the frequency and spatial area data. Z. Wu and M. Betke [20] displayed a structure for tracking multiple objects imaged from at least one static camera, where the issues of object detection and data association were communicated by a solitary target work.

F. Sardari and M. E. Moghaddam [21] proposed an object tracking strategy in light of a meta-heuristic method. Despite the fact that there were few meta-heuristic methodologies in the concept, they had adjusted GbSA (galaxy based search algorithm) which was more exact than related works. From the literature survey, they mainly focused the object tracking in video. Similar object detection is not done in the existing research. So that, the suggested technique is mainly concentrate on similar object detection with the help of modified local self similarity descriptor. And also the performance of the existing technique is minimum value when compared to the existing techniques.

3. Proposed methodology

Video tracking is the way towards finding a moving article (or numerous items) over time utilizing a camera. The main drawbacks of the existing object tracking method is a time-consuming approach if the video contains a high volume of information. There arise certain issues in choosing the optimum tracking technique for this huge volume of data. Further, the situation becomes worse when the tracked object varies orientation over time and also it is difficult to predict multiple objects at the same time. In order to overcome these issues here, we have intended to propose an effective method for object detection and movement tracking. The main goal of this paper is to create a system able to detect and track automatically all moving objects in a video surveillance sequence without any prior information about these objects. The overall structure of the suggested technique is shown in Fig. 1,

At first, the video sequence is divided into N number of frames. Then the N number of frames is fed to the preprocessing stage. In preprocessing stage, the adaptive median filter is used to remove the noise from each frame.

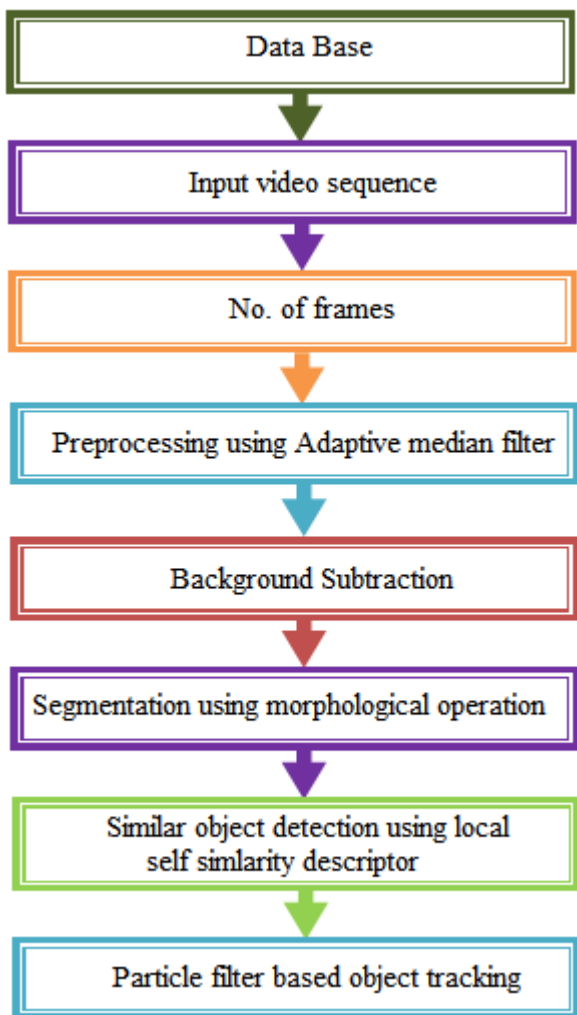


Figure.1 The Semantic Structure of Recommended Technique

Next, the background subtraction and segmentation using morphological operation are done in our recommended technique. And then the similar object detection in each frame is carried out by means of modified local self-similarity descriptor method. It is based on the measurement of appearance change between two consecutive frames. Once the similar object areas are found out in each frame, the object tracking is done from frame to frame by means of the particle filter.

The four stages of proposed method is shown in below,

1. Preprocessing
2. Background subtraction and segmentation
3. Similar object detection
4. Particle filter based object tracking

3.1 Preprocessing

At first, the input image or frame is changed to gray scale configuration. Next, the gray scale image

is preprocessed by means of an adaptive median filter to evacuate the noise. The principle objective of the preprocessing is to enhance the image quality to make it prepared to further handling by evacuating or lessening the random and surplus parts out of sight of the information outlines. In our proposed work, we are applying an adaptive median filter to expel noise. This adaptive median filter works based on the local statistical characters. It detects the impulse by calculating the difference between the standard deviation of the pixels within the filter window and the concerned current pixel. The initial steps of adaptive media filter are initializing the window size and calculate the maximum, minimum and median value from the pixel values. From that value noise candidates or pixels are identified and replaced by the median value and the remaining pixels are unaltered. Based on that, the noises from the input image or frames are removed and it will prepare the input frame for the next process.

3.2 Background subtraction and segmentation

3.2.1. Background subtraction

The process of extracting moving foreground objects (input image) from stored background image (static image) or generated background frame from image series (video) is called background subtraction. In most videos, the objects of interest reside in the foreground of a scene where things are happening and movements are taking place. Therefore, in order to track the motion of these objects, it is necessary to extract and distinguish them from the static background before any further processing. The intensity of the pixels corresponding to the static background remains largely unchanged between two consecutive frames. By computing the difference, these pixels get canceled out and only those of moving foreground objects are retained. After the background subtraction, the segmentation is carried out by means of morphological operation. It is clearly explained in further section,

3.2.2. Morphological operation

Morphological image processing is an assemblage of nonlinear operations connected with the shape or morphology of elements in a picture. A morphological operation on a binary image generates another binary image in which the pixel has non-zero esteem. Morphological operations modify the picture. Basic morphological operations are contracting the closer view ("disintegration"), expanding the frontal area ("widening"), Rejecting gaps in the forefront ("shutting") and eliminating stray closer view pixels

in the background ("opening"). The clear explanation of all the steps is as follows,

- **Erosion**

The erosion operation generates either contracting or diminishing of the item. The level of this operation is made available by the organizing component. Erosion joins two sets (u and v) by method of vector subtraction of set components.

- **Dilation**

The image was diluted by the dilation operation. The level of the amount it ought to be diluted depends on the organizing component. The organizing component is a small amount of the picture. The morphological operation dilation joins two sets by method for vector expansion. The dilation operation can be made via vector expansion of both of the components for both the sets u and v .

- **Opening and Closing**

Erosion sought after by dilation shapes a critical morphological modification called opening. The opening of a picture u by the organizing component v is shown by $(u \circ v)$ and is portrayed as,

$$u \circ v = (u(-)v)(+)v \quad (1)$$

Closing is the Dilation sought after by erosion. The closing of a picture u by the organizing component v is demonstrated by $(u \bullet v)$ and is depicted as,

$$u \bullet v = (u(+)v)(-)v \quad (2)$$

From that morphological operation, we efficiently segment the object and then the resultant output is fed to detect the similar objects from the input video frames.

3.3 Similar Object Detection

The resultant output from the preprocessing stage is used to find the similar object detection. For finding the similar objects, the proposed technique uses the Modified local self-similarity descriptor. The detailed explanation of modified local self-similarity descriptor is as follows,

3.3.1. Modified local self-similarity descriptor

Local self-similarity descriptor captures internal geometric layouts of local self-similarities within images/videos while accounting for small local affine deformations. It captures self-similarity of color,

edges, repetitive patterns and complex textures in a single unified way. A textured region in one image can be matched with a uniformly colored region in the other image as long as they have a similar spatial layout. These self-similarity descriptors are estimated on a dense grid of points in image/video data, at multiple scales. A good match between a pair of images corresponds to finding a matching ensemble of such descriptors with similar descriptor values at similar relative geometric positions, up to small non-rigid deformations. Here the traditional local self-similarity descriptor is modified with the help of correlation value. The step by step explanation of local self-similarity descriptor is as follows,

Step 1: Initially from the input image similarity within the image is detected using the sum of squared differences $SSD_q(a,b)$.

Step 2: The resulting distance surface $SSD_q(a,b)$ is normalized and transformed into a "correlation surface" $S_q(a,b)$:

$$s_q(a,b) = \exp\left(-\frac{SSD_q(a,b)}{\max(\text{var}_{noise}, \text{var}_{noise}(q))}\right) \quad (3)$$

Where,

var_{noise} is a constant that corresponds to acceptable photometric variations;

$\text{var}_{auto}(q)$ takes into account the patch contrast and its pattern structure. In our implementation $\text{var}_{auto}(q)$ is the maximal variance of the difference of all patches within a very small neighborhood of q .

Step 3: The correlation surface $S_q(a,b)$ is then transformed into log-polar coordinates centered at q , and partitioned into 80 bins (20 angles, 4 radial intervals).

Step 4: We select the maximal correlation value in each bin. The maximal values in those bins form the 80 entries of our "modified local self-similarity" descriptor vector d_q associated with the pixel q .

Step 5: Finally, this descriptor vector is normalized by linearly stretching its values to the range [0, 1] in order to be invariant to the differences in pattern and color distribution of different patches and their surrounding image regions.

Based on the above procedure we are finding the similar objects from the input video sequence. Then the main objective of the proposed technique is tracking the similar objects from the input video sequence. For similar object tracking the suggested technique uses the particle filter.

3.4 Particle filter based object tracking

After the detection of similar objects, the objects are tracked. Object tracking is to track the similar object in the video scene. Object tracking is the process of locating a moving object (or multiple objects) over time using a camera. It has a variety of uses, some of which are: human-computer interaction, security and surveillance, video communication and compression, augmented reality, traffic control, medical imaging and video editing. In our proposed method particle filter is used for object tracking.

3.4.1. Tracking object using particle filter

Particle Filtering is a technique for implementing a recursive Bayesian filter by Monte Carlo simulations. The idea is to represent the probability density function with a set of random samples with associated weights and to compute estimates based on these samples and weights.

Particle filter is needed to define the particle and its properties,

$$X_k = \{a, b, \hat{a}, \hat{b}\} \quad (4)$$

- Create particle step consist of N particles creation which has random locations (a, b) and random velocities (\hat{a}, \hat{b}). The step of prediction contains the modification of randomly generated particles using system model, which is in case of object tracking in video sequence equals,

$$S_c = AS_{c-1} + W_{c-1} \quad (5)$$

Where, A defines deterministic and W_{c-1} stochastic part;

This step occurs the change of particle position and its velocity based on mentioned system model.

- The next step is to look at the color of posts and target specific particles on the basis of equality, the value is the actualization. The degree of similarity to the target using the actual color and color difference is calculated as,

$$Z = C - C_{target} \quad (6)$$

Where C is actual color (one-dimensional vector) on position and C_{target} is the color of the target. Scalar value of likelihood is obtained by,

$$L_k = Z' * Z \quad (7)$$

Where, Z' is transposed matrix Z .

This step is important to assign to particles, which position is out of video sequence frame boundaries,

the lowest possible value.

- In step of resampling, cumulative distribution of weights and generation of N random numbers particles are resampled or rearranged, where particles with low weights are relocated to particle positions with higher weights.

The Simple Algorithm of Particle Filtering

Prediction: Predict the present state of each particle using previous information. Present state is represented in below Eq. (8),

$$X_{m|m-1}^k = f(X_{m-1|m-1}^k) + A_m \quad (8)$$

Where, A_m shows the random noise, $X_{m|m-1}^k$ is the predicted state of the particle.

Filtering: Reselection of particle accordingly to their likelihood method which is represented as,

$$L = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\left(\frac{c^2}{2\sigma^2}\right)} \quad (9)$$

Where, L is the likelihood and c are the color value of the pixel.

Hence by using particle filter objects are tracked in input video sequence. Based on the above procedure, we effectively track the similar objects from the input video sequence. The presentation of the projected technique is estimated and the outcomes are elucidated in beneath.

4. Results and discussion

Our proposed similar object detection using modified local self-similarity descriptor and object tracking using particle filter is implemented in the working platform MATLAB. The experimentation is carried out with the aid of various datasets of moving objects and the performance of our work is analyzed with various evaluation metrics.

4.1 Dataset descriptions

Our proposed work is worked out with the two datasets of moving objects. The descriptions are given below. The dataset used here are PETS Dataset and Cricket videos.

The datasets for PETS 2009 mainly consider as crowd image investigation and contain crowd count and density assessment, tracking of individual(s) surrounded by a crowd, and detection of separate flows and definite crowd occasions. The cricket video shows batsman offering a shot.

4.2 Experimentation results

The Proposed tracking system of moving an object is implemented over the Performance Evaluation of Tracking and Surveillance (PETS), Cricket Datasets.

Primarily, the datasets are given as the input video for detecting and tracking the moving objects. The data sets are the input video datasets which are converted into a set of frames. The resultant frames of the videos are shown in the above Fig. 2, These set of frames are then processed for the tracking of objects based on our proposed system. Pre-processing is the first phase, in which, the adaptive median filter is employed on the set of converted video frames. Blurs and noises from the frames of these databases are removed efficiently by this pre-processing filter, which results are illustrated in Fig. 3.

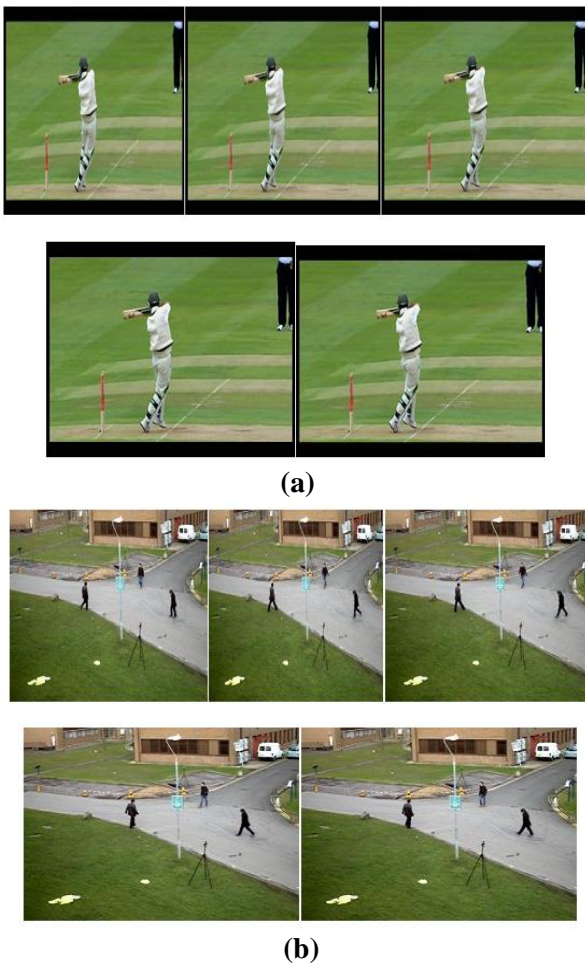


Figure.2 Input video frames (a) Cricket Video (b) PETS

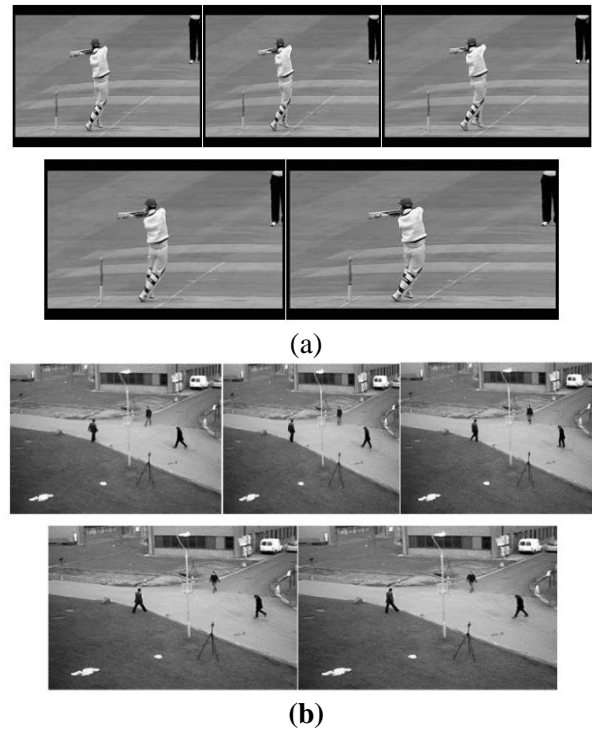


Figure.3 Pre-processing results (a) Cricket video (b) PETS

Followed by the preprocessing phase, moving objects are segmented. The morphological operation is utilized for the process of segmentation. Erosion, dilation, closing and Open in morphological operations are used, which gives a good accuracy of segmentation results with high speed. The output is illustrated in the below Fig. 4,

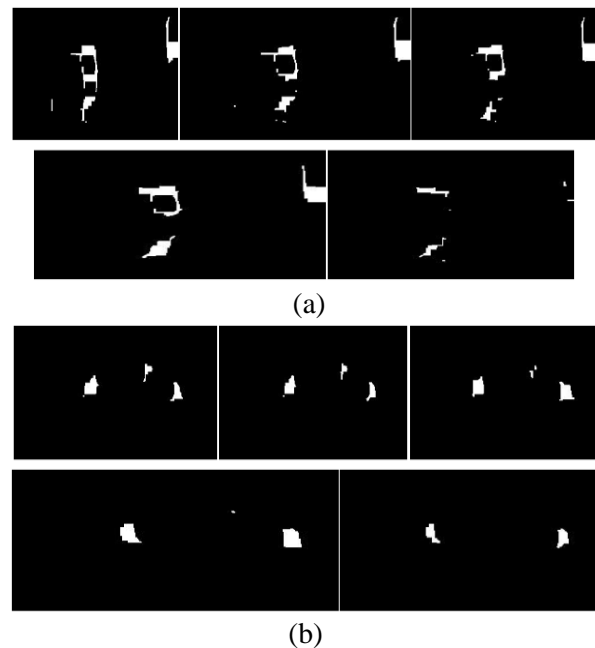


Figure.4 Segmentation results (a) Cricket video (b) PETS

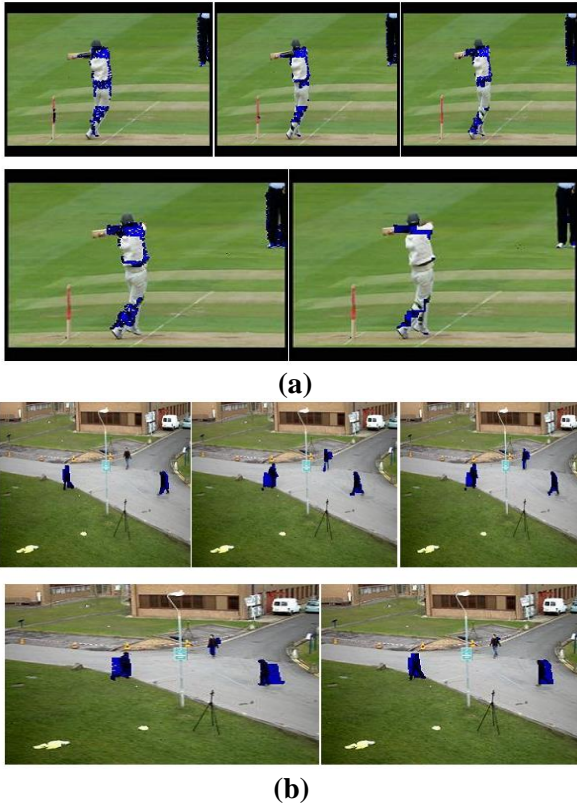


Figure.5 Output from moving object Detection phase (a) Cricket Video (b) PETS

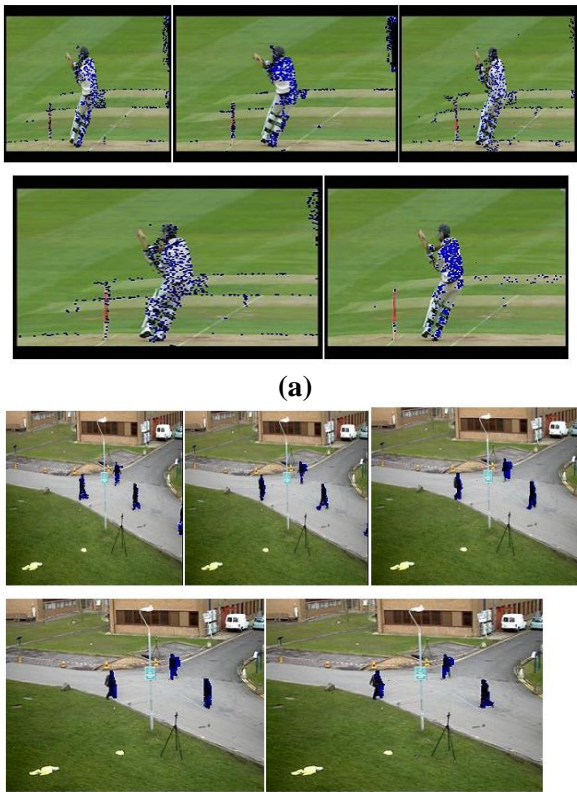


Figure.6 Result of tracking phase (a) Cricket Video (b) PETS

Once the segmentation process is completed, the next process is the detection of the segmented output. The output obtained from the moving object detection phase is illustrated in Fig. 5.

Tracking of the objects are experimented over these segmented frames and the moving objects are tracked with the blue color particles on the object. The task of tracking is handled by the Particle filter. The results of the final phase tracking of the proposed detection and tracking system are given in the Fig. 6.

4.3 Evaluation metrics

We need various evaluation metric values to be calculated in order to analyze the proposed tracking technique of moving objects. The metric values are found based on True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) with the option of pixel differences. The performance of this method is analyzed by ten metrics namely Precision, Recall, F-Measure, False Positive Rate (FPR), False Negative Rate (FNR), Percentage of Wrong Classifications (PWC), False Alarm Rate (FAR), Similarity, Specificity, and Accuracy.

Precision (P)

The precision estimates how many of the pixels in the segmented images segmented to be Positive are actually Positive by means of the Eq. (10),

$$P = \frac{TP}{FP+TP} \quad (10)$$

Recall or Detection Rate (DR)

The percentage of Positives correctly segmented is represented using recall.

$$DR = \frac{TP}{FN+TP} \quad (11)$$

F-Measure ($f-m$)

The percentage of harmonic mean for the combination of precision metric and recall metric gives the metric value of F-Measure, which is given as,

$$f - m = \frac{2(P \times DR)}{P + DR} \quad (12)$$

False Positive Rate (FPR)

The percentage of cases where the results show the video frame is correctly classified, but in fact, it was not successful.

$$FPR = \frac{FP}{FP+TN} \quad (13)$$

False Negative Rate (FNR)

The percentage of cases where the result shows the video frame is not correctly classified, but it was actually successful.

$$FNR = \frac{F_N}{F_N + T_P} \quad (14)$$

Percentage of Wrong Classifications (PWC)

PWC is the measure of wrongly segmented video frame and is expressed in percentage which is given by,

$$PWC = \frac{F_P + F_N}{T_P + F_P + T_N + F_N} \times 100 \quad (15)$$

False Alarm Rate (FAR)

FAR is the ratio between the false detected pixels and the total numbers of positive pixels.

$$FAR = \frac{F_P}{F_P + T_P} \quad (16)$$

Similarity (S)

Similarity measure is the evaluation of pixel similarity between both the segmented frame and ground truth frame is computed by this, which is,

$$S = \frac{T_P}{T_P + F_P + F_N} \quad (17)$$

Specificity (Sp)

The measure of the specificity is the proportion of frames which are exactly segmented. i.e. the measure of how exactly segmentation is done for negative results.

$$Sp = \frac{T_N}{T_N + F_P} \times 100 \quad (18)$$

Accuracy (A)

The weighted percentage of frames that are correctly segmented is measured by accuracy.

$$A = \frac{T_P + T_N}{T_P + F_P + T_N + F_N} \times 100 \quad (19)$$

4.4 Performance analysis

The performance of the proposed detection and tracking methods of moving objects are evaluated by the different metrics. The different metrics are Precision, Recall, F-Measure, False Positive Rate, False Negative Rate, Percentage of Wrong Classifications, False Alarm Rate, Similarity, Specificity, and Accuracy.

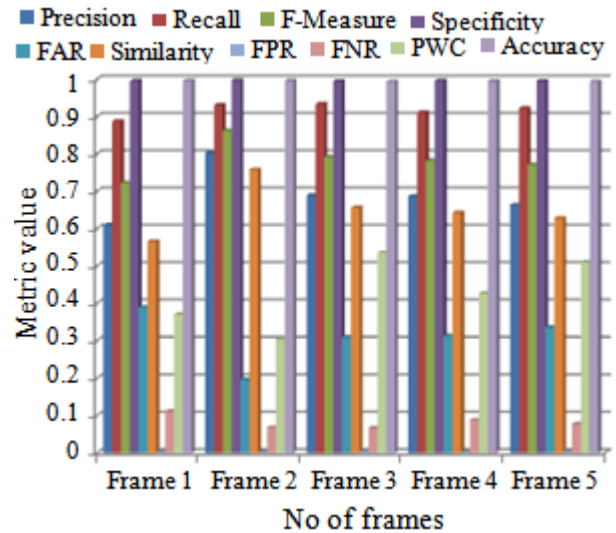


Figure.7 Graphical representation of various evaluation metrics for Different frames of PETS dataset

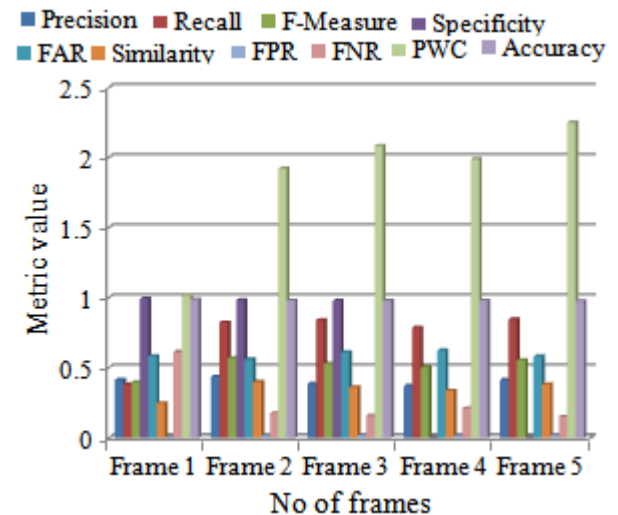


Figure.8 Graphical representation of various evaluation metrics for Different frames of Cricket dataset

The results of proposed work help to analyze the efficiency of the tracking process. The corresponding Graphical representation of these various measures for different frame in PETS dataset is shown in the above Fig. 7,

The corresponding Graphical representation of these various measures for different frame in Cricket dataset is shown in the above Fig. 8.

4.5 Comparative analysis

The existing work will be compared in this section with the proposed work to show that our proposed work is better than the state-of-art work. Here the recommended technique is compared with existing technique [22].

Table 1. Comparison table with Evaluation metrics

| Methods | Recall | Specificity | FPR | FNR | PWC | F-Measure | Precision |
|-------------------------------------|----------|-------------|----------|----------|----------|-----------|-----------|
| Existing Method -PETS video [22] | 0.8122 | 0.9907 | 0.0089 | 0.0268 | 1.13947 | 0.86052 | 0.8372 |
| Proposed Method- PETS video | 0.91749 | 0.996382 | 0.003618 | 0.012624 | 0.430366 | 0.918083 | 0.881306 |
| Existing Method- Cricket Video [22] | 0.8594 | 0.98704 | 0.00295 | 0.00947 | 0.18059 | 0.90297 | 0.8219 |
| Proposed Method- Cricket Video | 0.972042 | 0.995197 | 0.015373 | 0.006393 | 1.854666 | 0.927405 | 0.837335 |

The existing technique uses the fuzzy based algorithm for object tracking. The existing algorithm is dealing with similar object detection; occlusion, Handling Background modeling, and abrupt change in the environmental conditions are some of the challenging work.

In Table 1, the comparison of our proposed method with existing technique is described. The performance can be compared based on precision, recall, specificity, f-measure, and PWC, FPR, and FNR rates. The proposed work outperforms all these four existing works by providing better results of detection and tracking of persons. The lower error rates in FNR, FPR leads to make good tracking results of the moving persons with accurate detection of persons for the proposed system.

5. Conclusion

Similar object detection and tracking in H.264 video is proposed in this paper. For similar object detection, the suggested technique uses the modified local self-similarity descriptor and similar object tracking is done by a particle filter. In our proposed technique we are considering both PETS and Cricket video for input video sequence. The performance of the suggested technique is evaluated by means of precision, recall, F-measure, FPR, FNR, PWC, FAR, similarity, specificity, and accuracy. From the experimental results, the proposed work outperforms than the existing works by providing better results of detection and tracking of persons. The suggested technique attains the FPR and FNR is 0.00361 and 0.0126 but the existing technique [22] attains 0.0089 and 0.0268 which is maximum value when compared to the existing. The lower error rates in FNR, FPR leads to make good tracking results of the moving persons with accurate detection of persons for the proposed system. In future, the researchers may utilize improved some other object tracking algorithm to achieve maximum performance.

Reference

[1] S. Chen, S. Li, R. Ji, Y. Yan and S. Zhu, "Discriminative local collaborative

- representation for online object tracking", *Elsevier on Knowledge-Based Systems*, Vol. 100, pp. 13-24, 2016.
- [2] S. Yasukawa, H. Okuno, K. Ishii and T. Yagi, "Real-time object tracking based on scale-invariant features employing bio-inspired hardware", *Elsevier on Neural Networks*, Vol. 81, pp. 29-38, 2016.
- [3] W. C. Hu, C. H. Chen, T. Y. Chen, D. Y. Huang, and Z. C. Wu, "Moving object detection and tracking from video captured by moving camera", *Elsevier on Journal of Visual Communication and Image Representation*, Vol. 30, pp. 164-180, 2015.
- [4] I. Majeed and O. Arif, "Non-linear eigenspace visual object tracking", *Elsevier on Engineering Applications of Artificial Intelligence*, Vol. 55, pp. 363-374, 2016.
- [5] D. Gregorio, M. Giordano, S. Rossi and M. Staffa, "Experimenting WNN support in object tracking systems", *Neuro computing*, Vol. 183, pp. 79-89, 2016.
- [6] M. M. Azab, H. A. Shedeed and A. S. Hussein, "New technique for online object tracking-by-detection in video", *Elsevier on IET Image Processing*, Vol. 8, No. 12, pp. 794-803, 2014.
- [7] C. R. del-Blanco, F. Jaureguizar and N. García, "An efficient multiple object detection and tracking framework for automatic counting and video surveillance applications", *IEEE Transactions on Consumer Electronics*, Vol. 58, No. 3, pp. 857-862, 2012.
- [8] X. Li, Q. Liu, Z. He, H. Wang, C. Zhang and W. S. Chen, "A multi-view model for visual tracking via correlation filters", *Elsevier on Knowledge-Based Systems*, Vol. 113, pp. 88-99, 2016.
- [9] Y. Li, Y. Su and Y. Liu, "Fast two-cycle curve evolution with narrow perception of background for object tracking and contour refinement", *Elsevier on Signal Processing: Image Communication*, Vol. 44, pp. 29-43, 2016.
- [10] W. Li, P. Wang and H. Qiao, "Top-down visual attention integrated particle filter for robust object tracking", *Elsevier on Signal Processing: Image Communication*, Vol. 43, pp. 28-41, 2016.

- [11] X. Cheng, M. Zeng and X. Liu, "Spatially constrained level-set tracking and segmentation of non-rigid objects", *Elsevier on Journal of Visual Communication and Image Representation*, Vol. 38, pp. 745-752, 2016.
- [12] H. B. Resmi, V. A. Deepambika and M. A. Rahman, "Symmetric Mask Wavelet Based Detection and Tracking of Moving Objects Using Variance Method", *Elsevier on Procedia Computer Science*, Vol. 58, pp. 58-65, 2015.
- [13] P. Zhang, T. Zhuo, L. Xie and Y. Zhang, "Deformable object tracking with spatiotemporal segmentation in big vision surveillance", *Neurocomputing*, Vol. 204, pp. 87-96, 2016.
- [14] G. Han, X. Wang, J. Liu, N. Sun and C. Wang, "Robust object tracking based on local region sparse appearance model", *Elsevier on Neurocomputing*, Vol. 184, pp. 145-167, 2016.
- [15] Y. Chen, Y. Shen, X. Liu and B. Zhong, "3D Object Tracking Via Image Sets And Depth-Based Occlusion Detection", *Elsevier on Signal Processing*, Vol. 112, pp. 146-153, 2015.
- [16] I. Elafi, M. Jedra and N. Zahid, "Unsupervised detection and tracking of moving objects for video surveillance applications", *Elsevier on Pattern Recognition Letters*, Vol. 84, pp.70-77, 2016.
- [17] D. Riahi, and G. A. Bilodeau, "Online multi-object tracking by detection based on generative appearance models", *Computer Vision and Image Understanding*, Vol. 152, pp. 88-102, 2016.
- [18] H. Zhao, S. Qi, and X. Wang, "A joint appearance model of SRC and MFH for multi-objects tracking", *Elsevier on Neurocomputing*, Vol. 213, pp. 75-83, 2016.
- [19] K. Ahmadi and E. Salari, "Small dim object tracking using frequency and spatial domain information", *Elsevier on Pattern Recognition*, Vol. 58, pp. 227-234, 2016.
- [20] Z. Wu and M. Betke, "Global optimization for coupled detection and data association in multiple object tracking", *Elsevier on Computer Vision and Image Understanding*, Vol. 143, pp. 25-37, 2016.
- [21] F. Sardari and M. E. Moghaddam, "An object tracking method using modified galaxy-based search algorithm", *Swarm and Evolutionary Computation*, Vol. 30, pp. 27-38, 2016.
- [22] K. Srinivasan and P. Balamurugan, "Object Detection and Movement Tracking in H.264 Compressed Video Based on Optimal Particle Filtering Technique", *Journal of Computational and Theoretical Nanoscience*, Vol. 13, No. 11, pp. 8171-8181, 2016.