ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645
Indian Journal of Science and Technology, *Vol 10(19), DOI: 10.17485/ijst/2017/v10i19/95427, May 2017*

Moving Object Detection and Segmentation using Background Subtraction by Kalman Filter

Archita Tah[1]*, Sudipta Roy[2], Prasenjit Das[1] and Anirban Mitra[1]

[1]Department of Computer Science and Engineering, Academy of Technology, Adisaptagram, Hooghly – 712121, West Bengal, India; archita.kp@gmail.com, p.das@aot.edu.in, anirban.mitra@aot.edu.in
[2]Department of Computer Science and Engineering, U. V. Patel College of Engineering, Ganpat University, Kherva – 384012, Gujarat, India; sudiptaroy01@yahoo.com

Abstract

Objectives: Object tracking and detection are significant and demanding tasks in the area of computer vision such as video surveillance, vehicle navigation, and autonomous robot navigation. **Methods/Statistical Analysis**: This paper presents the moving object tracking using Kalman filter and reference of background generation. Kalman filter is based on two types of filters: cell Kalman filter and relation Kalman filters. The process entails separating an object into different sub-regions and discovering the relational information between sub-regions of the moving objects. **Findings**: In this paper, the precise and real-time method for moving object detection and tracking is based on reference background subtraction and use threshold value dynamically to achieve a more inclusive moving target. This method can effectively eliminate the impact of luminescence changes. Due to deployment of Kalman filter this fast algorithm is very straightforward to use to detect moving object in improved way and it has also a broad applicability. This technique is very authentic and typically used in video surveillance applications. **Application/Improvements:** This technique is very legitimate and typically used in video surveillance applications. The Kalman filtering algorithm upgrades the model and enlarges the dimensionality of the moving system state.

Keywords: Background Subtraction, Detection and Segmentation, Moving Object, Kalman Filter, Object Tracking

1. Introduction

Security is needed in every working environment. It can be accomplished in audio, video or by any other means. Multiple moving objects tracking from a video series of surveillance camera is nowadays a demanding application. Some of the existing object tracking method used contour-based, region-based and feature point–based method; but there is some drawback this leads us to generate a motion based recognition a rewarding. The object identification and tracking is the task of movement based identification, computerized examination, video indexing, human-computer interaction and traffic supervise and vehicle direction-finding. It is most difficult task to design a system which is capable of identifying objects where objects and background both are changing. It may be complicated to detect unfamiliar objects with considerable changes like colour, silhouette and surface. Most surveillance classification use stationary camera. It makes the object revealing much easier. In such problems, a background model is accomplished with data obtain from vacant sight and foreground section are acknowledged using the variation between the skilled model and new clarification.

**Author for correspondence*

Survey on this paper is persistent on techniques for tracking objects in general and not to trackers customized for explicit objects. Kalman filter has been widely used for tracking. The random variation and inaccuracies are observed over a time then Kalman filter is used to approximation the unknown variable which is closer to the measurement's true values. This filter is used to estimate the ambiguity of an expect value and it computer a weighted average of the guessed value and considered value. In varied domain of applications like control, navigation, computer vision, and time series econometrics it can be applied. The filer is based on the following three parameters: object's future locality prediction, decrease of noise introduced by mistaken detections and facilitates the process of organization of multiple objects to their tracks. By taking the difference between two frames we can track the object then using domain knowledge along with the displacement of scale and displacement vector the motion of the object can have characterized. In this paper we use Kalman filter for tracking object by assuming variables are normally distributed.

Object tracking is the significant problem in human motion investigation. It is upper level computer vision problem. Tracking engage corresponding detected foreground objects between successive frames using dissimilar feature of object like motion, velocity, color and texture. In the surveillance system, locate the object in each frame then find its location in each frame trajectory of object is generated. Object tracker may also supply the whole section of the object in the frame at every instant of time at the same it also provides correspondences between the object instances between two frames.

We review some papers on background estimation, background subtraction and object tracking. In [1] studied at background estimation of a scene using a Kalman filter approach. In the previous research introduces the concept to track background intensity and adjust without human intervention to modify in the scene. There is two phase in the algorithm in 1st step track update mean intensity and in the 2nd step track standard deviation update. Updating can take place according Kalman filter equation. Using standard deviation, it modeled how the object transits background and foreground. In result, the foreground objects are basically detected from frame to frame with particularly no sequential tracking. But a straightforward improvement to the object detection would be to include a simple tracker for the detected objects.

In [2,3] compare various background subtraction algorithms for distinguish moving vehicles and pedestrians in urban traffic video series. In their experiments, they vary the attributes in every algorithm to take dissimilar precision working points. Then evaluate them base on how they diverge in preprocessing, background representation, foreground finding, and data justification. Five definite algorithms are tested on city traffic video succession: frame differencing, adaptive median filtering, median filtering, mixture of Gaussians, and Kalman filtering. Mixture of Gaussians generates the overriding results, while adaptive median filtering suggests a straightforward substitute with competitive performance.

In [4] proposed a spatio and temporal approaches for Background representation by the region which can be predicted by automatic process.

In [5] proposed an algorithm that is capable to contract with both steady and unexpected total enlightenment changes. They had accessible an innovative algorithm for background updating in support on Kalman filtering method, which is strong to regular and jagged enlightenment transform. The most important originality they have establish, which make the algorithm strong to steady and jagged illumination changes, is the evaluation of the total illumination dissimilarity and its use as an exterior control of the Kalman filter. The testing had shown that the algorithm spread out the diversity of development where it cans effort successfully. Its most important drawback is incapability to run dynamic background pixels.

In [6] proposed an algorithm which was able to differentiate actions of interest in the land and imitation events on the ocean face with enormously low false alarm rate. Again, this technique was capable to contract with the movement of the trees and outperformed obtainable techniques. In their paper they proposed a technique for modeling of dynamic scenes for the purpose of background foreground differentiation and change detection. The method relies on the deployment of visual flow as a characteristic for change detection. In order to correctly utilize the doubts in the features, an original kernel based multivariate concentration assessment procedure is proposed, that adjusts the bandwidth according the suspicions in the examination and trial measurements.

In [7] proposed a general practice for dispensation a video series is as follows: Initialize mixture models for every pixel with a weak preceding. For every frame: (a) modernize the predictable mixture model for every pixel using incremental EM (b) Heuristically label the mixture components, (c) Classify every pixel according its modern mixture model. The heuristic labeling improvement is desirable in cases where the mixture mechanisms are not in the similar order as the previous models designation. Their heuristics are: label the darkest component as silhouette; remain two components; label the one with the biggest variance as vehicle and the other as road. But their heuristic advancement may not effort well in limits of illumination circumstances.

In [8] propose to merge an incorporated method, the Principal Component Analysis - Gaussian Mixture Model (PCA-GMM) method that produce a comparatively enhanced segmentation product as contrast to predictable GMM with Kalman Filtering (KF) in their article. Background subtraction engages manipulating the reference image and labeling the pixels equivalent to foreground objects. The collective new method of PCA-GMM-KF challenge tracking several moving objects; the size and location of the objects alongside the series of their images in dynamic prospect. The tracking algorithm corresponds to performing the following: First they concern the PCA-GMM algorithm to take out the moving objects from background in every video frame, and then calculate the subsequent position candidates in the next frame by guessing it from KF.

In [9] proposed a tracking algorithm based on a multi-attribute combined sparse illustration. The patterns for the sparse demonstration consist of pixel values, consistency, and edges. The surveillance with the minimum renovation error is preferred under the particle filtering framework as the tracking result. Main drawback of the paper is the runtime of solitary object tracking algorithm for every frame is less than one seconds and the runtime of multiple object tracking algorithm is less than five seconds per frame in the case of occlusions. Here as the number of objects enhance, the runtime does not rapidly improve. The computational complexity of their planned process was not satisfactory.

In [10] suggest a new method to identify and track several moving targets. Their approach focuses on challenging urban scenarios. Main steps of the designed approach are demonstrated as: Video stabilization, Foreground extraction, Detect the Moving object and tracking. In result, the videos which incomplete to the corner points above the entire image, so the tracking was also accurate. In urban locality the algorithm was inaccurate in middle traffic and also in the existence of a bridge which occluded the cars. The errors are occurred because of dense traffic, because in this case the algorithm can hardly differentiate between the moving vehicles and the background.

In [11] proposed a multiple object tracking method based on a new weighted Kanade-Lucas-Tomasi (KLT) tracker. Using original KLT tracking algorithm trajectory of the target objects is generated from an image sequence. The algorithm is based on the three basic steps: Object detection, track the attribute within the object mask a biased function is used to identify the object trajectory. Object can be represented by any, such as edges, shape, colors, and primal information than conventional KLT detector is used to choose the primary attributes.

In [12] divide tracking methods into two different approaches: Merge Split (MS) approaches and Straight-Through (ST) approaches. Here, we are focusing to identify the most challenging and promising method for trade with occlusions in common and with soccer request in particular. Here problem will be solved based on MS approaches and ST approaches. The problem for tracking in general and for soccer in particular is solved. The main challenge in MS approaches is to renovate object identities following a split. In regions-based ST approaches, main dispute is project to an explicit object of pixels that could fit in to several objects ("disputed" pixels). In contour-based ST approaches, the dispute is the obligation to a precise object of some incomplete contours.

In [13], a precise and flexible method for strong identification and tracking of several objects in video series was proposed. This paper integrates color instant and wavelet instant together for recognition and tracking. It is proposed to utilize a fuzzy operator i.e. the Choquet fundamental to combined the results acquire in every dimension to keep away from crisp decision. In this paper, the tracking algorithm merges the attribute identical model and Kalman filter framework to resolve the existing problems. A strong and practical multiple objects identification and tracking method is proposed here. The experimental results demonstrate that the anticipated method can perfectly identify and energetically track several objects. Furthermore, the detection of several moving objects with occlusion is effectively finished which is a problem for identification and tracking based on characteristic matching.

In [14] proposed an object tracking method. They divide the method into three categories for object representation point correspondence, primitive geometric models, and contour evolution methods. They find out the correspondence points, detect the object and determine the regions of object in the scene.

In [15] proposed an algorithm using frame difference method and removal of noise using threshold technique. In his research approach he is proposed environment modelling using distance transform technique. In the result, here a video is taken and converted into frames. The object is detected from the video frames by taking difference between two frames then threshold is applied. Distance between one edges of the road and one edge of the vehicle is used to distance transform technique. Particle filter is applied to track multiple moving objects. The system is designed in such a way it can refresh automatically and track multiple objects efficiently by taking relationship among environment states and object characteristics. Tracking error is highly minimized by considering a large number of particles in spite of dynamic changes of environment.

In [16] apply discrete wavelet transform for moving objects detection. Their proposed methods performing better and it is capable to alleviate problems attached with other spatial domains as noises, clutters and ghost etc.

The rest of the paper has formatted in the way as: the details of our proposed methodology has described in section 2. The result on input video has been described on section 3. Finally, we conclude our paper in section 4

2. Proposed Methodology

Once the entity is distinguished from frame then tracking is possible beside its path. A number of standard methods are accessible for the tracking of objects, such as Kalman filter, Particle filter, and Mean-shift based kernel tracking etc. A Kalman filter is a best possible estimator that gather parameters of understanding from indirect, inaccurate and uncertain interpretation. It is recursive, so that new capacity can be processed as they appear. Flow chart diagram of our proposed planned method has been shown in Figure 1.

In our method, we first take a video and split it into some frame. Then estimate the background using background estimation method. After that we take a frame from which we will find our object. Now, convert both images in their binary form. Then subtract the background image from the frame in their binary form. Now perform a function for morphologically open binary image (remove small objects). At last the object is in front of us. Feedback control is used to estimates a process in Kalman filter. In moving object, we need to predict and correct a frame. For prediction, time update equation which is extrapolative onward from the present status and estimate a priori error covariance in subsequent step. For the correction measurement equation where to obtain an improved a posterior estimate.

Now let's find the formula. It is easiest to look at this first in one dimension. A 1D Gaussian bell curve with variance σ2 and mean μ is defined as:

$$N(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (1)$$

Now we multiply two Gaussian curves together:

$$N(x, \mu_0, \sigma_0) N(x, \mu_1, \sigma_1) ? = N(x, \mu', \sigma') \quad (2)$$

Next we substitute equation (1) into equation (2) and then do some algebra (being careful, so that the total probability is 1) to obtain:

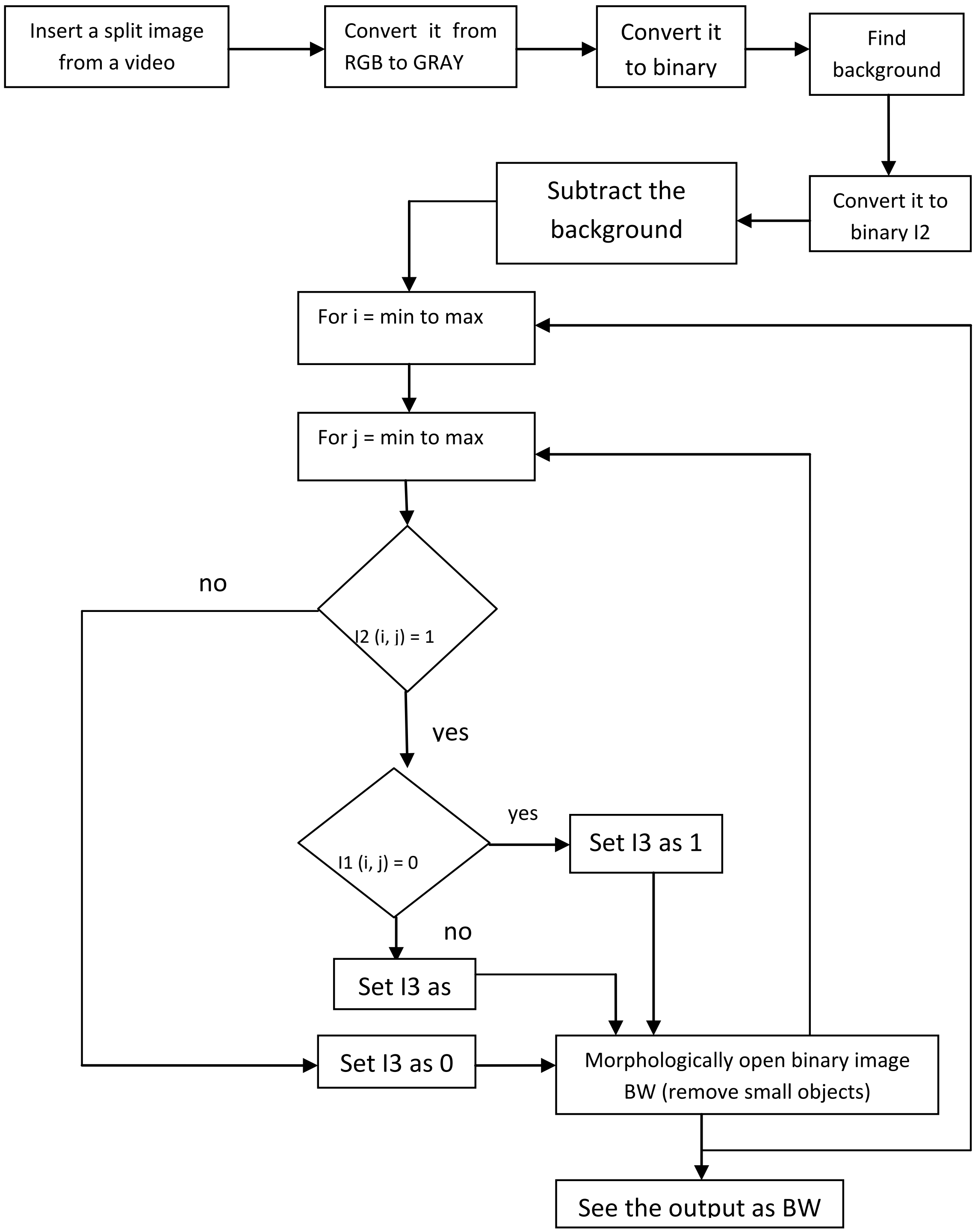

Figure 1. Flowchart of proposed method.

$$\mu' = \mu_0 + \frac{\sigma_0^2(\mu_1 - \mu_0)}{\sigma_0^2 + \sigma_1^2}$$

And

$$\sigma'^2 = \sigma_0^2 - \frac{\sigma_0^4}{\sigma_0^2 + \sigma_1^2} \quad (3)$$

We now simplify by factoring out a little piece k:

$$k = \frac{\sigma_0^2}{\sigma_0^2 + \sigma_1^2} \quad (4)$$

$$\mu' = \mu_0 + k(\mu_1 - \mu_0)$$

$$\sigma'^2 = \sigma_0^2 - k\sigma_0^2 \quad (5)$$

Here, we take our previous estimate and add something to make a new estimate. Now, we check the matrix version. Well, now re-write equations (4) and (5) in matrix form. If we consider Σ as the covariance matrix of a Gaussian blob, and $\vec{\mu}$ its mean along each axis, then:

$$K = \sum_0 \left(\sum_0 + \sum_1\right)^{-1} \quad (6)$$

$$\vec{\mu}' = \vec{\mu_0} + K(\vec{\mu_1} - \vec{\mu_0})$$

$$\sum' = \sum_0 - K\sum_0 \quad (7)$$

K is the matrix called the Kalman gain. We'll use it in just a moment.

Putting it all together, we have two distributions: The predicted measurement and the observed measurement. The predicted measurement is with $\left(\mu_0, \sum_0\right) = \left(H_k\hat{x}_k, H_k P_k H_k^T\right)$, and the observed measurement with $\left(\mu_1, \sum_1\right) = \left(\vec{z_k}, R_k\right)$. We plug these into equation (7) to find their overlap:

$$H_k P_k' H_k^T = H_k P_k H_k^T - K H_k P_k H_k^T$$

$$H_k P_k' H_k^T = H_k P_k H_k^T - K H_k P_k H_k^T \quad (8)$$

From (6), the Kalman gain is:

$$K = H_k P_k H_k^T \left(H_k P_k H_k^T + R_k\right)^{-1} \quad (9)$$

We can knock an H_k off the front of every term in (8) and (9) (note that one is hiding inside K), and an H_k^T off the end of all terms in the equation for P_k'.

$$\hat{x}_k' = \hat{x}_k + K'(\vec{z_k} - H_k\hat{x}_k) \quad (10)$$

$$P_k' = P_k - K' H_k P_k$$

$$K' = P_k H_k^T \left(H_k P_k H_k^T + +R_k\right)^{-1} \quad (11)$$

K' is giving us the complete equations for the update step.

And that's it! $\hat{x}_k'$ is our new best estimate, and we can go on and feed it (along with P_k') back into another round of **predict** or **update** as many times as we like [17].

3. Result and Discussion

The primary objective our work is to detect the motion object and segmentation using background subtraction. In single particular object follow-up can considered in two steps and they are: Average of all pixel can be considered as the background framework. Then the frame number is chosen from the several frame of video that contains the moving object. Subtract the background image from that

(a) (b) (c) (d)

Figure 2. (a)-(d) are some input sequences.

Figure 3. Background image.

frame. We use Kalman filter for finding the object. Then the location and error is calculated for the chosen frame.

The Kalman filter has wide applications in technology like navigation and control specifically in air craft and space craft. It is based on the time series analysis and we can measure recorded data over a long time span which normally contains random variation and inaccuracies are observed over a time then Kalman filter is used to estimate the unknown variable which is closer to the measurement's true values. The main motive to track an object is to finding position by generating the path of the object over long time span. It is also generating the region

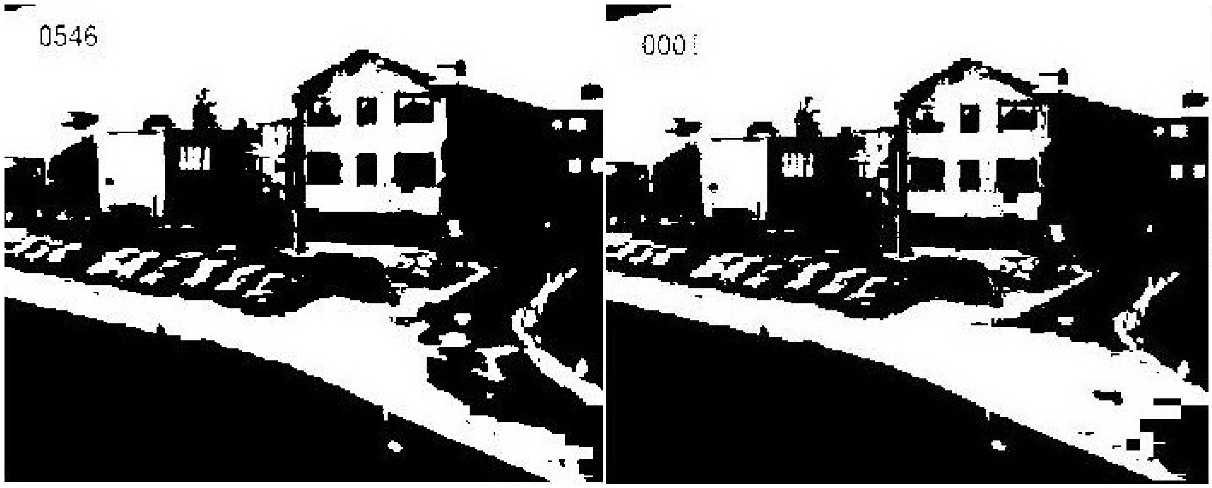

Figure 4. (a) Binary image of Frame No 546 and (b) Background image in Binary Form.

Figure 5. Output object.

covered by the object at every instant of time. It will detect the object by measuring the association between the objects occurrences in the frames that either be achieved individually or mutually.

We first split the frames from the video. In Figure 2(a)-(d) we show some splitting frame, frame num 104,546,608 and 800 Frames are stored in a separate file. Here this video split on 800 frames. Then find out the background image Figure 3. Then a particular frame is chosen from form the large number of frame for object identification. We take the frame num 546. Now it is necessary to convert the two frames into binary image Figure 4. After that we have do background subtraction using STATE SPACE representation. The background reference image is deducted from the individual frames after removing the small noise. At last, the object is in front of us.

Final segmented output of moving object car has been shown in Figure 5. A performance metric is a basic parameter and metric which is used to measure the performance of any proposed approach. At the following section we are comparing our proposed methodology with other techniques by the performance metric. Our proposed method is tested with various videos with adaptable parameters. Performance measurement of the proposed methodology is improved tracking in terms of visual performance and mathematical measures accuracy, correlation coefficient and similarity. This section gives the performance measurement metrics that is used to apply in order to study the performance.

3.1 Mean Square Error (MSE) and Peak Signal to Noise ratio (PSNR)

The mean square error or MSE and peak signal to noise ratio or PSNR are the two common error measurement metrics that are used to evaluate the superiority of background image. The MSE correspond to the increasing squared error between the original and the background image. The smaller value of MSE is good for the system. To compute PSNR, the block first computes mean-squared error using the equation given:

$$\mathbf{MSE} = \frac{\left(\sum_{M,N} [I1(M,N) - I2(M,N)]\right)^2}{M * N} \quad (12)$$

Here, M and N represents the sizes of the input images. Now the PSNR is calculated as the following equation:

$$PSNR = 10\log_{10}\left(\frac{R^2}{MSE}\right) \quad (13)$$

The PSNR is calculated between two images comprising of 8 bits per pixel in terms of decibels (dBs). In general, when PSNR is 40dB or greater than, then the reconstructed image and the original image are almost difficult to differentiate by human eyes. In this equation, R denotes utmost instability for the input images. For example, if the input image possesses double-precision of floating-point data, then R is 1. For an 8-bit unsigned integer data type, R is 255, etc.

3.2 Correlation Coefficient

Correlation coefficient defines statistical association among two or more arbitrary variable and experiential information. This calculates the association coefficient among matrices of the same size.

3.3 Similarity

We measured pixel based match or similarity assessment as

$$\text{Similarity} = \frac{\mathbf{tp}}{\mathbf{tp} + \mathbf{fn} + \mathbf{fp}} \quad (14)$$

Larger value of similarity index demonstrates a precise recognition of moving object. Table 1 shows the MSE metric, PSNR metric, correlation coefficient metric and similarity metric for frame no 546.

A good image sequence gives the value of MSE <10 and value of PSNR, Correlation coefficient and are more than 75%. In our project min value of MSE is 2.90 and maximum value of MSE is 8.40, which are less than 10.

Min value of PSNR is 84.83 and max value is 95.30, which are more than 75%. Min value of Correlation coefficient is 90.6 and max value is 95.0, which are more than 75%. Min value of Similarity is 90.11 and max value is 95.44, which are more than 75%. So after performance evaluation, we can say that our algorithm gives a good result.

Table 1. Performance matrices

Image sequence	Metrics in %			
	MSE	PSNR	Correlation coefficient	Similarity
1	8.60	84.83	90.9	91.77
2	5.60	87.20	90.6	92.44
3	7.90	90.40	93.7	94.22
4	2.90	85.00	94.5	90.11
5	4.40	87.50	92.7	90.88
6	8.18	89.60	95.0	94.99
7	3.97	91.40	91.6	93.77
8	7.12	86.40	93.7	95.44
9	6.80	88.10	92.9	94.55
10	5.25	92.50	94.7	92.88
11	8.40	95.30	91.3	95.00
12	7.90	93.40	94.1	91.66
13	6.44	90.20	94.7	91.44
14	7.20	87.30	93.2	93.55
15	8.10	87.50	92.9	92.66

4. Conclusion

We survey a number of background subtraction and moving object detection algorithms. STATE SPACE representation without user involvement provides sufficient output. Our design method of background subtraction is considerably increasing the competence of computer vision. The proposed Kalman filter based algorithms for moving object detection execute pretty sound for an early simplified example. In addition, we have given a general survey of object tracking methods. We believe that, this article can give important insight into this significant research topic and support new research.

Our algorithm cannot recognize the very small objects. For future work, we are planning to initiate multimodal features to characterize objects so that the identification and tracking performance can be further enhanced.

5. References

1. Scott J, Pusateri MA, Cornish D. Kalman filter based video background estimation; 2009.
2. Cheung S-CS, Kamath C. Robust techniques for background subtraction in urban traffic video. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series; 2004 Jan. p. 881–92.
3. Cheung S-CS, Kamath C. EURASIP Journal on Applied Signal Processing. 2005:14:2330–40.
4. Monnet A, Mittal A, Paragios N, Ramesh V. Background modeling and subtraction of dynamic scenes. ICCV, Nice, France; 2003 Oct. p. 1305–12. CrossRef.
5. Messelodi S, Modena CM, Segata N, Zanin M. A Kalman filter based background updating algorithm robust to sharp illumination changes .
6. Mittal A, Paragios N. Motion-based background subtraction using adaptive kernel density estimation.
7. Chinchkhede DW, Uke NJ. International Journal of Computer Science & Information Technology. 2012 Feb; 4(1).
8. Noureldaim E, Jedra M, Zahid N. Intelligent Information Management. 2013; 5:42–7.
9. Hu W, Li W, Zhang X. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2014.
10. Mattyus G, Benedek C, Sziranyi T. Modeling of optical airborne and spaceborne Sensors. ISPRS Istanbul Workshop 2010, WG I/4, Oct. 11-13. 2010; Vol. XXXVIII-1/W17.
11. Han B, Paulson C, Lu T, Wu D, Li J. Tracking of multiple objects under partial occlusion. 2015. p. 1–10.
12. Gabriel PF, Verly JG, Piater JH, Genon A. The state of the art in multiple object tracking under occlusion in video sequences.
13. Ramya G, Srilatha. Global Journal of Researches in Engineering . 2014; 14(6).
14. Yilmaz A, Javed O, Shah M. Object tracking: A survey. ACM Computing Surveys; 2006. CrossRef.
15. Gogulamudi SL, Kumari BMK. International Journal of Information Technology, Control and Automation. 2012 Oct; 2(4).
16. Gupta P, Singh Y, Gupta MK. International Journal of Advanced Research in Computer Science and Software Engineering. 2015 May; 5(5).
17. Available from: http://www.bzarg.com/p/how-a-kalman-filter-works-in-pictures/.