Multiple Objects Tracking Using the Kalman Filter Method in Outdoor Video

[1]P.P Halkarnikar, [2]Amol Dhakne, [3]Pratiksha Shevatekar, [4]Prateek A. Meshram, [5]Shivaji R. Vasekar, [6]Jitendra Garud

[1,2,3,4,5,6]Department of Computer Engineering, Dr D Y Patil Institute of Engineering, Management and Research, Akurdi, SPPU, Pune, Maharashtra, India.

Article Info
Page Number: 7711-7719
Publication Issue:
Vol. 71 No. 4 (2022)

Article History
Article Received: 25 March 2022
Revised: 30 April 2022
Accepted: 15 June 2022
Publication: 19 August 2022

Abstract
In video surveillance systems, one of the most critical tasks is object tracking. Due to the constantly shifting circumstances outside, following moving objects in outdoor footage may be challenging. Inside the scope of this article, we investigated a variety of factors that might impact the tracking of objects within outdoor movies. Occlusion of objects may occur while monitoring many objects in an outdoor video system since the lighting conditions in the environment are always changing. When opposed to an interior setting, which has illumination conditions that are constant, tracking things in an outdoor area is much more challenging as a result of this factor. The shadows that fall on top of each other in the movie make it difficult to follow objects and produce occlusion of some of them. Tracking is one of the most common applications for the Kalman filter. When an item disappears, however, the system becomes unstable, and it takes the Kalman filter a significant amount of time to begin tracking again. In order to prevent this from happening, we have suggested using a modified version of the Kalman filter, which will make it easier to track objects in environments that are constantly changing outside.
Keywords: - video surveillance, multiple object tracking, Kalman filter.

I. Introduction

A video surveillance system is currently considered mandatory hardware for the physical protection of many different types of establishments. These days, video surveillance is required in a wide variety of establishments, including banks, hospitals, government organisations, ATMs, temples, airports, train stations, shopping malls, residences, and even private homes. Both the inside and outside of an organization's facility, particularly strategic points, need to be monitored by surveillance equipment. The success of the identification of the item is contingent on the precision with which the object is segmented and on the correctness with which its placement in the video sequences is located. [1] This process of segmentation is influenced by various elements, including environmental factors, hardware factors, and software considerations. When there is more than one item in the frame, the segmentation process becomes more difficult since the different objects are interacting with one another. Shadow, overlap, and camouflage with the backdrop may all have an effect on the segmentation of an item. When trying to track an item in video, having other moving or stationary objects in the frame might cause the tracking to become impossible. The automated tracking system will become unreliable as a result of this.

Throughout the course of this study, we have investigated many issues that arise while attempting to follow an object using a computer vision system. Objects that have been accurately divided in a video recording may be followed. According to the properties that are given to the numerous objects, the segmented item is very straightforward to identify. Increasing the number of characteristics allows for improved object detection, both anticipated and unexpected. In this study, we make use of the Kalman filter, which is often used for purposes of tracking. In order to appropriately segment and track the needed object in video, a modification of the Kalman filter to add shadow removal would be of great assistance.

Video surveillance is quickly becoming the most common method used in public places for the purposes of monitoring, management, and law enforcement. Because there are currently millions of CCTVs used in different areas, it is necessary to engage a big number of human operators in order to monitor each camera. This means that there is a need for more employment opportunities. On the other hand, it has been shown that human operators are not able to maintain a satisfactory level of performance when the number of CCTV cameras that they are responsible for increasing. Because of this, the automation of visual surveillance systems as an application of different object identification and tracking has been a highly significant study area in recent years. In this research, one of the challenges associated with attaching the item was investigated. Since the shadow of an item in an outside setting is constantly shifting, the tracking process is often disrupted. In this section, we attempted to find a solution to the problem of shadows affecting object tracking in videos.

II. Related work:

The most common reasons for losing track of a moving object in video are shifts in the illumination and shadows. Jinhai Xiang and colleagues have come up with a local intensity ratio model (LIRM) that is resilient to variations in the lighting. [4] This model was presented. The study of the lighting and shadow model is used as the basis for the calculation of the distribution of the local intensity ratio. In order to do segmentation without employing shadows, the Gaussian mixture model (GMM) is put to use. The erasure approach is used in order to get the shapes of the moving objects and to get rid of the dispersed shadow patches and sounds. Camilo Aguilar et al. [5] describe the difficulties that arise when attempting to track down tiny objects in low-resolution photos obtained using remote sensing. Because targets are on the smaller side, they do not have distinguishing traits, and this makes it easier for them to be overlooked in busy settings. They have successfully detected and followed tiny moving objects by using a convolutional neural network and a Bayesian tracker. This strategy is known as a track-by-detection approach. The general position of the target was determined in the first step by using motion detection, and in the second stage, this information was fed into a CNN in order to improve the detection findings. The Probability Hypothesis Density (PHD) filter makes use of the method known as "track-by-detection" in order to transform detections into tracks. In order to monitor several objects in surroundings with a lot of clutter, a Bayesian data-association framework is utilised. A system for scalable object tracking was presented by Niranjil Kumar et al. [6], which is capable of eliminating shadows using fuzzy logic and tracking individuals with object boundaries. The Kalman filter technique is used to keep track of the various individuals. Poonam B. Linghate and colleagues [7] have analysed and compared a number of different approaches that may be used to track objects in video surveillance. The purpose of object tracking is to identify and link target items appearing in several consecutive video frames. They provided a list of difficulties in tracking objects over successive frames. Some of the difficulties stem from the complicated motion of the item, the uneven form of the object, and the blockage of one object by another. Having a low computing In order to improve the tracking of objects in a variety of demanding environments, the Kalman filter is often combined with a number of additional tracking approaches. Kalman filter was used for the purpose of multiple object tracking by SanjivaniShantaiya et al. [8]. The authors of the paper developed an enhanced optical flow method in addition to the Kalman filter in order to track many objects. This is because the Kalman filter becomes unstable if the priori centre prediction varies by a considerable number. This tweak to the algorithm will allow for the occlusion issue to be handled appropriately.

A significant number of scholars have focused their attention specifically on Kalman filter, Mean shift, and Particle filter. A number of different efforts have been presented in order to combine the possibilities offered by these trackers in order to produce superior tracking outcomes. In addition, machine learning methods are used so that objects may be recognised. [9],[10],[11]

III. Kalman filters theory

A Kalman Filter is an optimal recursive data processing algorithm. In image sequence processing Kalman filtering is used for adaptive background estimation, in order to separate the foreground from the background. The equations for the Kalman Filter fall into two steps: time update (or prediction) and measurement update (or correction). The time update equations are responsible for projecting forward (in time) the current state and error covariance to obtain the a priori estimates for the next time step. The measurement update equations deal with the feedback, which is new measurements incorporated into the a priori estimates to obtain improved a posteriori values. The Kalman filter estimates the position of the object in each frame of the sequence. The variable parameters of the Kalman filter are the state vector and measurement vector. The state vector is composed of the initial position; width and length of the search window and the centre of mass of the object (x_c, y_c)at time t_k .This vector are presented by following equation.

$$s_k = (x_k, y_k, W_k, L_k, x_c, y_c) \quad \ldots (4.1)$$

The measurement vector of the Kalman filter is composed of the initial position, length and width of the search window of the object at timet_k. This vector is given by following equation.

$$z_k = (x_k, y_k, W_k, L_k) \quad \ldots \quad (4.2)$$

A. Major steps of Kalman filtering

1) Process to estimate: The Kalman filter estimates the statesa discrete process, this state is modelled by the linear equation as below

$S_k = A\, S_{k-1} + w_k$ …….(4.3)

With A is the transition matrix, $w_k$is the noise process.

The measurement model is defined by equation below.

$Z_k = H\, S_k + v_k$……(4.5)

With H as the measurement matrix.

Noise process w_k and v_k are assumed independent of state vectors and measurement and are normal distributions with zero mean and are presented by equations as below

$p(w) \sim \mathcal{N}(0, Q)$………….(4.7)

$p(v) \sim \mathcal{N}(0, R)$…………(4.8)

Finally, the output equations for the two blocks of prediction and correction of Kalman filter are: (4.13) and (4.14):

$\hat{S}_K = A \times \hat{S}_{K-1}$ ….…. (4.13)

$P_K^- = A \times P_{K-1} \times A^T + Q$…….. (4.14)

2) The equations for updating: measurement residual:

$\tilde{Y}_K = Z_K - H_K \times \hat{S}_K^-$……(4.15)

Innovation covariance:

$S_K = H \times P_K^- \times H^T + R$..……(4.16)

Updated state estimate:

$\hat{S}_K = \hat{S}_K^- + K_K \times \tilde{Y}_K$…….(4.17)

Optimal Kalman gain:

$K_k = P_K^- \times H^T \times S_K^{-1}$……(4.18)

Updated estimate covariance:

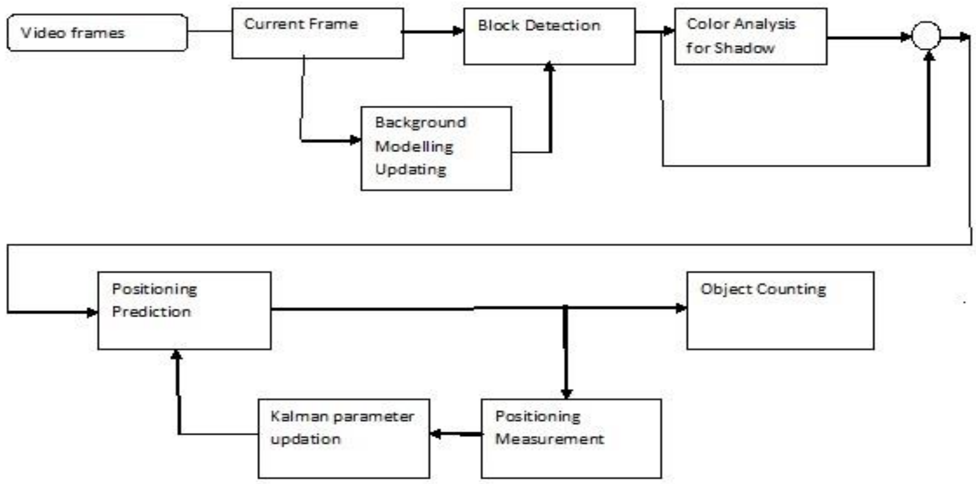
$P_{K-1} = P_K^- - K_k \times H \times P_K^-$…..(4.19)

The iterative predictor-corrector nature of the Kalman filter can be helpful, because at each time instance only one constraint on the state variable need to be considered. The process is repeated, considering a different constraint at every time instance. All the measured data are accumulated over time and help in predicting the state.

IV. Block diagram of modified Kalman filter implementation:

Even when there are numerous objects, the Kalman filter is able to properly follow the object blog. However, it is unable to monitor the blog when there is clutter in the circumstance. Shadow clutter is the focus of this discussion since there are a variety of other circumstances that can generate clutter. A revised version of the process block diagram is shown in figure1. After compiling all of the pixel blog posts, a mask will be created. Pixels that are included inside the mask are analysed for their chromaticity before being deleted from the blog. The Kalman parameters are updated based on the remaining pixels in the image. There is no instability in the Kalman tracking since the updated parameters are not diverted in the same way that blog localization is.

Figure1. Block diagram of modified Kalman filter



V. Experimental results:

Because the focus of our study is on handling tracking in environments with a lot of clutter, we chose test movies that were set outside. These films were collected from a public database and were captured with a single camera in an outdoor setting. We are responsible for recording one of the test videos that was chosen (kop comb1.avi) for the experiment. Occlusion and shadow provide the impression that there are a lot of objects in these films. During these frames, we have seen focused behaviour from the object detection algorithm. In the interest of experimentation and analysis, we took into consideration the possibility of objects melting together owing to shadows. The characteristics of these films are outlined in Table1. below. These films were shot outside and have a frame size of 320x240. They include automobiles and objects that blend together in many frames because of shadows. The highway scene is shown in both of the videos, and it is used to identify and tally the number of automobiles. The objects in films take on a variety of forms, including that of automobiles, trucks, and scooters. The descriptions of the films may be found in Table1.

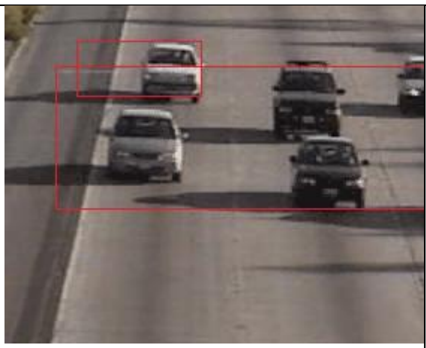
ISSN: 2094-0343
2326-9865

Video in .avi	**Frame rate (fps)**	**duration**	**environment**	**Total No. of Frames**	**Object Class**	**Object Size (relative)**
Highway_I	14	29 sec	Outdoor	440	Cars	Medium
Kop1_comb	25	31 sec	Outdoor	770	Cars, Trucks, Scooter	Medium, Small

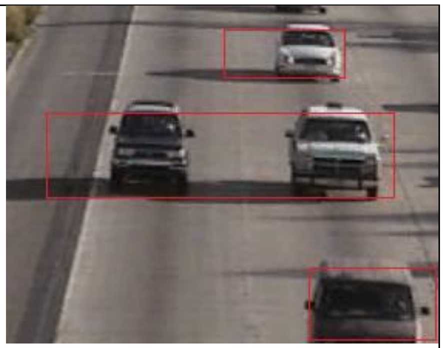
Table1. Profiles of videos used for experiments.

a. Experiment:

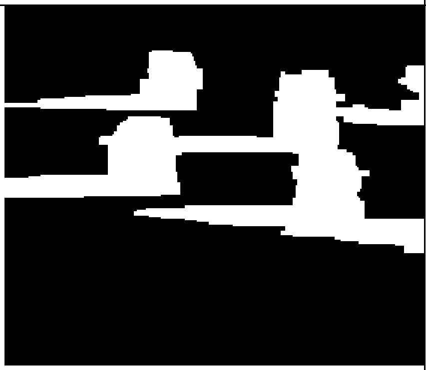
The first experiment is being conducted on the highway-I.avi. Noted below are the frames in which Kalman lost his bearings about the item. Unstable tracking was a result of the presence of clutter on the scene, which was brought about by shadow and occlusion. Figure2 illustrates the current state of affairs. Figure2 depicts the pixel blob that should be used for this detection once frame differencing has been performed. B. The chrominance test reveals the existence of a shadow pixel, which is seen in Figure2.C. Figure 2.D depicts the track that was generated using the suggested modified Kalman algorithm after shadow pixels were eliminated. The modified Kalman technique, produces satisfactory results when applied to transient situations like clutter encountered when tracking.



(A)

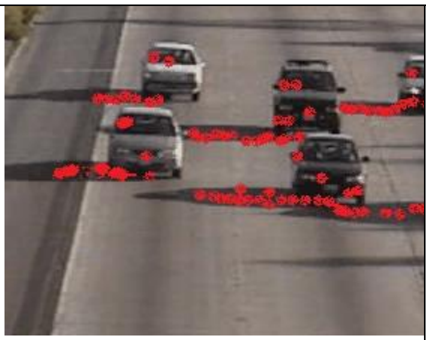


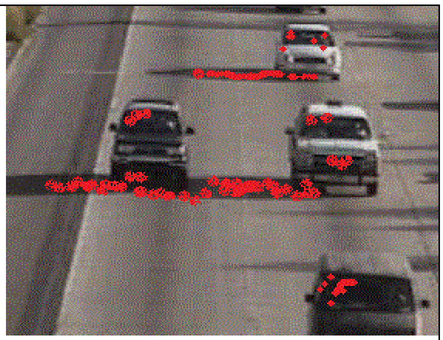
(A)



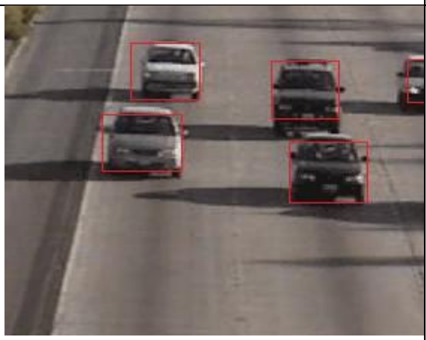
(B)

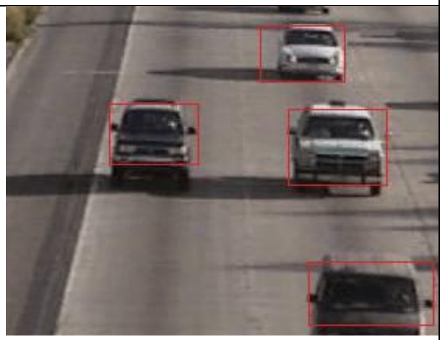
(B)





(C) (C)





(D) frame #142 (D) frame # 183

Figure2. Output for Frame no. #142 and #183 (A) showing clutter effect observed in video. (B) foreground pixels deteced at back ground subtaraction stage. (C) shadow pixels detected (D) clear segmentation of objects.

VI. Performance Analysis:

The object-oriented analytical approach is being used in the work that we conduct. Analysis based on pixels is used for the vast majority of the work that is done in the field of literature. One disadvantage of using such an analysis for object-based apps or surveillance is that the performance of user satisfaction is dependent on the number of objects that are accurately recognised. As a result, pixel analysis does not accurately represent the performance of the system, despite the fact that it is superior for analysing the algorithms that are used. The success of pixel-based approaches is strongly dependent on the ground truth creation of pixel, which may be challenging and does not have perfect consistency. [12]. Because the success of many tracking and classification applications is contingent not on the quantity of pixels identified by the system but rather on the precise location of the objects being tracked, object-based performance offers many benefits to video object detection systems.

Recall and precision values are calculated as follow

$$Recall = \frac{T_p}{T_p+F_n} \text{..........}(4.22)$$

$$Precision = \frac{T_p}{T_p+F_p} \text{.............}(4.23)$$

$$Fscore = \frac{recall*precision}{recall+presion} \quad \ldots\ldots\ldots(4.24)$$

VII. Conclusion:

The overlapping of objects and shadows is a common factor for clutter in outdoor recordings, which disrupts the identification and tracking of many objects. The nature of shadow is arbitrary, and its appearance may change depending on the time of day. In this kind of scenario, the fix geometric form shadow detection algorithm cannot be used. The chromaticity property of shadow pixels is important for detecting shadows as a result. The pixel that remains after the shadow pixels have been removed is the one that is used for object tracking. This results in a decrease in the Kalman error because the noise contributed by shadow pixels is eliminated before the Kalman filter is applied. Even if the improvement is significant, the recommended change does not take into account overlapping, which is another factor that might lead to undetected objects. Sometimes, a split in an item may emerge because of pixel clustering. This is a transient circumstance that cannot be fixed by just adding alteration. When there is a change in the total number of objects in the frame, the Kalman filter requires more time to settle the tracking, which lowers its level of trustworthiness. In statistical analysis, the F-score is used to determine how accurate a test is. Experiments employing the Kalman filter and a modified version of the Kalman filter were performed on test movies taken from our database. The results showed an increase in the accuracy of item identification. Utilizing the traditional Kalman filter, the average F-score for the test film (Highway 1.avi) is 0.534, however using the suggested modified Kalman filter, this score rises to 0.832. Another one of our test videos, kop1 comb.avi, had an average F-score of 0.177 when it was run through the case Kalman filter. Using the suggested improved Kalman filter results in an improvement to 0.433.

References:

[1] DivyaMaheshwari, Dr. S. R. Gengaje, “Abandoned and Removed Object Detection in a Video”, International Journal of Innovative Research in Science, Engineering and Technology, Vol. 4, Issue 9, September 2015. Pp.8394- 8399.

[2] HrishabhP, SLongowal, B Singh, “An Approach for Shadow Removal inMoving Object”, International Journal of Computer Science and Information Security (IJCSIS),Vol. 14, No. 8, August 2016.Pp 781- 790.

[3] Shiping Zhu*, ZhichaoGuo and Li Ma, “Shadow removal with background differencemethod based on shadow position and edgesattributes”, EURASIP Journal on Image and Video Processing 2012. pp. 1-15.

[4]Jinhai Xiang, Heng Fan, Honghong Liao, Jun Xu, Weiping Sun, Shengsheng Yu, “Moving Object Detection and Shadow Removing underChanging Illumination Condition”, Mathematical Problems in Engineering, 2014. Pp 1-11.

[5] Camilo Aguilar, Mathias Ortner,JosianeZerubia, “Small Object Detection and Trackingin Satellite Videos With MotionInformed-CNN and GM-PHD Filter”, Frontiers in Signal Processing, Vol. 2, 2022.pp1-15.

[6] Niranjil Kumar A, Sureshkumar C, “Fuzzy Based Shadow Removal and Integrated Boundary Detectionfor Video Surveillance”, Journal of Electrical Engineering Technology, Vol. 9, No. 6, 2014. pp. 2126- 2133.

[7] Poonam B. Linghate,Prof.K.R.Ingole, “ Detection And Tracking Of Object In Real Time Video”, International Journal of Creative Research Thoughts (IJCRT), Volume 9, Issue 1 January 2021, pp 2443- 2449.

[8] Shantaiya, S. &Verma, Kesari& Mehta, Kamal. “ Multiple object tracking using kalman filter and optical flow” . European Journal of Advances in Engineering and Technology.Vol. 2.2015, pp. 34-39

[9] Fancy Joy, V. Vijayakumar“ A Review on Multiple Object Detection and Tracking in Smart City Video Analytics”, International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-2S2 December, 2018

[10] Mohankumar S, Madigondanahalli T,Gopalakrishna, Chikkaguddaiah N, “Approach for shadow detection and removalusing machine learning techniques”, IET Image Process., 2020, Vol. 14 Issue. 13, pp. 2998-3005.

[11] P. P. Halkarnikar, Dr. S.N. Talbar, Dr. P.N. Vasambekar, “Object Tracking using Association Rule in Clutter Environment “, International Journal of Innovative Research in Science, Engineering and Technology, Vol. 4, Issue 1, January 2015, pp 18645 – 18652.

[12] Komal Gogi1, Dr. M. Uttarakumari, “Detection of Moving Cast Shadow and Removalfor Video Surveillance using HSV Color Space”, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 6, Issue 4, April 2017, pp 733-737.

[13] Thuc Trinh Le, AAlmansa, Yann Gousseau, Simon Masnou, “Object removal from complex videos using a few annotations”, Computational Visual Media, 2019,pp. 267–291, https://doi.org/10.1007/s41095-019-0145-0.