

# Vehicle Detection and Tracking from a Video Captured by Moving Host

Thathupara Subramanyan Kavya

School of Electrical Engineering, University of Ulsan  
Ulsan, Republic of Korea  
kavyasubramanyan@gmail.com

Young-Min Jang

School of Electrical Engineering, University of Ulsan  
Ulsan, Republic of Korea

Tao Peng

School of Electrical Engineering, University of Ulsan  
Ulsan, Republic of Korea

Sang-Bock Cho \*

School of Electrical Engineering, University of Ulsan  
Ulsan, Republic of Korea  
sbcho@ulsan.ac.kr \*

**Abstract - An efficient vision-based vehicle detection and tracking system is presented in this paper. One of the most vital stages of the driver assistance system is the detection and tracking. This will help the drivers to fully understand the road conditions. Vehicle detection from a moving camera is difficult compared to a stationary camera. In this method, we used the practical implementation of a vehicle detection based on Viola-Jones algorithm. To track the detected vehicle, we are using a nonlinear filter such as Unscented Kalman Filter. The simulation and experimental results indicate that real-time nonlinear tracking problem can be resolved by using this nonlinear Unscented Kalman Filter.**

**Keywords:** Vehicle Detection, Moving Camera, Vehicle Tracking, Unscented Kalman Filter

## 1. Introduction

Road accidents are increasing on a daily basis. As per statistics, around 1.25 million people are losing their life every year due to fatal road accidents. Non-fatal accidents will make people disabled through injuries and this is another 30 to 50 million. Effective intervention is needed to address road traffic injuries that include proper traffic management system and intelligent transportation system (ITS). Driver visual attention on a road is one of the key points to reduce road accidents. To achieving this, vehicle detection and tracking system performs vital role and this will act as a support system to the driver.

We are proposing an intelligent driver assistance system (IDAS) which continuously track and detect the vehicles travelling at the front. Detecting object under the stationary camera is comparatively easiest method compared to non-stationary cameras, since in non-stationary camera both foreground and background are moving at same time. While there is more literature about vehicle detection and tracking using the static camera, there are only seldom publications about vehicle detection from a moving host. In this paper, we have proposed a method to detect vehicles from a moving host using train cascade object detector. A nonlinear filter such as Unscented Kalman Filter (UKF) is used to track the vehicles. The complexity of the vehicle detection algorithms is due to the variations in illumination, background, visual angle and different weather condition and the implementation is not easy.

Basically vehicle detection can be categories into two. 1) Feature-based 2) Learning-based method. Feature-based method uses the different features of the vehicle images. Learning based method is robust, however the computational complexity and training requirements are intense. Viola-Jones method is one of the most extensively used methods for face detection [1]. We are extending Viola-Jones algorithm to detect vehicles in this paper. The proposed method is superior due to the following reasons.

- Rapid computation of Haar-like features using the integral features.
- Classifier learning with AdaBoost to select the best features.
- The detection process is efficient since the attentional cascade structures reject most of the sub-windows in early layers of the detector.

Due to the simplicity of extracted features, process and selection of the best features, Viola-Jones vehicle detection is fast and robust. Some of the existing vehicle detection methods based on Haar Feature-Based Classifier [2-6], are considered under static conditions.

In linear optimal estimation under Gaussian noise Filter, the Kalman filter is one of the best linear algorithms [7-9]. Harpreet et al. [10] proposed a method for vehicle tracking using Fractional Feedback Kalman filter which modifies the steady-state gain of the Kalman filter. This improves the sudden transitions of the tracked objects. However, we have non-linear systems in the real world. Under the nonlinear conditions, Kalman filter is not ideal. In order to solve this issue, nonlinear filters like Extended Kalman Filter (EKF) [11-12] and Unscented Kalman filters (UKF) are introduced. Unscented Kalman filter performance is superior to EKF.

In recent years UKF has emerged as an efficient algorithm for obtaining approximate solutions to discrete-time nonlinear optimal filtering problems. This has been successfully applied to numerous practical tracking problems [13]. Wang et.al [14] proposed an Unscented Kalman filter for vehicle tracking. However, this paper considered vehicle detection under the static camera. Baofeng et.al [15] proposed another method to detect vehicles based on cascade Adaboost classifier, but for tracking purpose, a linear filter such as adaptive Kalman filter is used.

Last few years has seen a lot of research work in intelligent transportation systems. The task in this paper is to detect a vehicle at the front of a moving host and track the vehicle in real time situations. This will assist the driver for a comfort drive.

The rest of the paper is organized as follows: Section 2 describe the proposed method for detection and tracking, Section 4 shows the experimental results of our method and finally in section 5 briefly explains the conclusion and future work.

## 2. Proposed Vehicle Detection and Tracking

In the last decade or so, there has been a growing interest in computer vision-based driver assistance systems. This system is capable to detect any potentially dangerous situations and warn the driver to avoid any accidents. Vehicle detection and tracking is crucial in assisting the drivers to experience a smooth driving. In a non-stationary camera condition, both the vehicle and the camera are moving and hence it makes very difficult to detect the moving objects from the surroundings. In this research we use the Haar-like feature-based cascade object detector for multiple vehicle detection. Haar-like feature based boosted cascade was initially introduced by Viola and Jones for face detection [1]. We are extending this algorithm to detect vehicles. The general block diagram of the proposed architecture is shown in Fig.1.

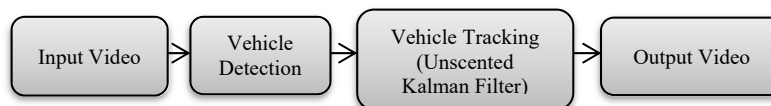


Fig.1. General Block diagram of the proposed method

This paper is mainly divided into two parts. First part explains the detection of a vehicle from a video and the second part explains the tracking of the detected vehicles from the sequence of video frames.

### 2.1. Vehicle detection

Vehicle detection is the initial and foremost task for many applications such as traffic control, Intelligent transportation systems etc. Vehicle detection forms the inevitable part in driver assistance system. Vehicle detection is the process of identifying a vehicle from an image or a sequence of video frames. A set of positive images with regions of interest specified to be used as positive samples must be provided to the cascade classifier. Similarly, a set of negative images must be provided from which the function can generate negative samples automatically. These positive and negative sample images are provided to custom train the cascade classifier. A sample of 10,000 positive and 28,000 negative images are collected from various environments to train the cascade object detector. With the help of a large set of training samples, we could able to produce accurate vehicle detection. The general structure of the cascade classifier is shown in the Fig.2.

Paul Viola and Michael Jones in 2001 proposed the usage of Harr-like feature for object detection. The major advantage of the Haar-like cascade object detection is the ease to remove the negative samples.

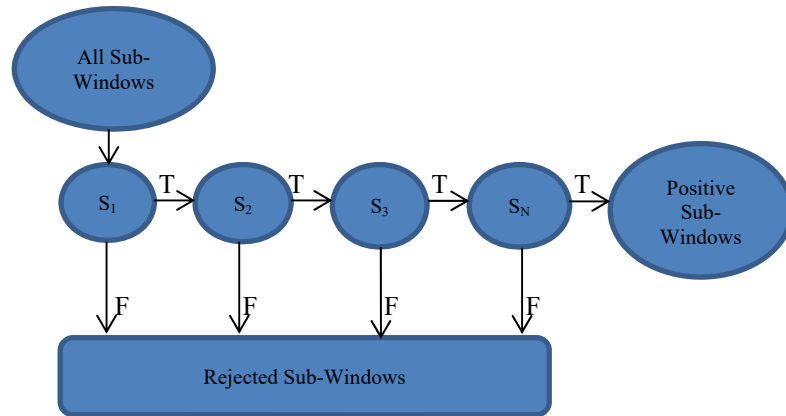


Fig.2. Detection of positive sub-windows using the cascade

This method is relatively robust to noise and lighting changes since they compute the gray level difference between the white and black rectangles. The different kinds of filters based on the Haar-like feature is shown in Fig.3.

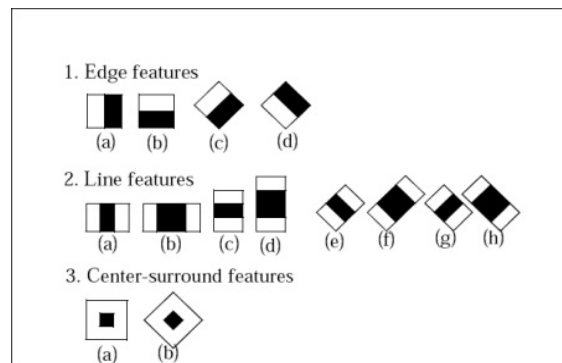


Fig.3. Different Kinds of filters based on Haar-Like feature

Each Haar-like feature consists of two or three connected black and white rectangles. The difference between the total sum of the pixels within the white and the black rectangular region is the value of the feature.

The integral image is a method for a quick and efficient calculation for the sum of intensity values in a rectangular subset of an image. We can express it mathematically by,

$$ii(x, y) = \sum_{x' \leq x} \sum_{y' \leq y} i(x', y') \quad (1)$$

Where  $i(x, y)$ , is the intensity of the gray scale image at pixel  $(x, y)$ . Fig.4 shows the sum of the intensity pixels of any rectangular area ABCD, which can be calculated with only four array references as:

$$\sum_{(x,y) \in ABCD} i(x, y) = ii(D) + ii(A) - ii(B) - ii(C) \quad (2)$$

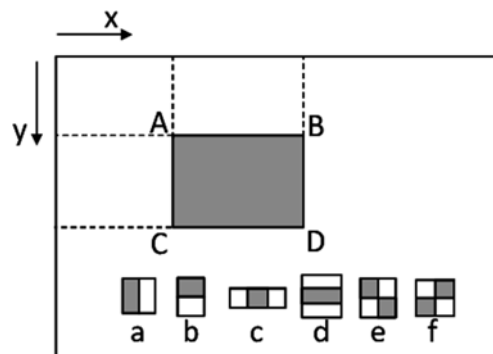


Fig.4. illustration of Eq. (2)

Viola and Jones used this method for fast computation of a huge number of Haar-Like features which are simply defined as the difference between the sum of the intensities in the dark and light shaded regions of simple rectangular patterns as shown in Fig.3.

This computational advantage enabled scaling the features for multi-scale detection at no additional cost because it requires the same number of operations despite of size.

The Adaboost is used both to select features and train the classifier. The weak learner is designed here to select the features which best separates the weighted positive and negative training examples. A weak classifier  $h(x, f, p, \theta)$  is defined as:

$$(x, f, p, \theta) = \begin{cases} 1, & \text{if } pf(x < p\theta) \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

Where  $f$  is a feature from the huge set spanning different sizes of the Haar-Like features,  $p$  is a polarity indicating the direction of the inequality,  $\theta$  is a threshold and  $x$  is a training sub-window of size 24 x24 pixels.

We have tested our algorithm under different weather conditions and illuminations to analyse the performance. Some of the samples which are used for training (both positive and negative samples) is shown in Fig.5 & Fig.6.



Fig.5. Positive samples used for training

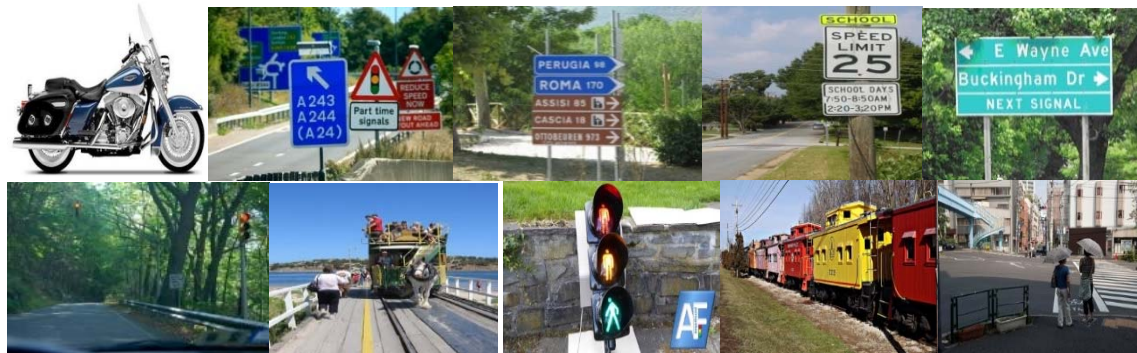


Fig.6. Negative samples used for training

## 2.2. Unscented Kalman Filter for Vehicle Tracking

Object tracking is the most challenging task of any research in the area of computer vision. There are some applications of motion tracking such as:

- Control applications
  - Object Avoidance
  - Automatic Guidance
- Surveillance/Monitoring Applications
  - Security Cameras
  - Traffic Monitoring
  - Vehicle Counting /Tracking

Vehicle tracking is very important in intelligent video surveillance, automatic driving assistance etc. The aim of vehicle tracking is to find the target region in continuous in each frame. In this paper for tracking, we are using a nonlinear version of Kalman filter such as UKF [16] because Kalman Filter is no longer optimal in realistic driving environments. UKF tracking is used for different applications such as person tracking under static camera [17-18] and target tracking [19]. Tracking deals with estimating the trajectory of an object as it moves around a scene in the image plane.

The scope of Kalman filter is extended to non-linear filtering problems using an Extended Kalman Filter (EKF). A Gaussian approximation is formed to the joint distribution of the state  $x$  and  $y$  using a Taylor series-based transformation. However, EKF has some limitations when used in nonlinear conditions:

- The linear and quadratic transformations produce reliable results only when the error propagation can approximate by a linear or a quadratic functions. If this condition fails, the performance of the filter can be poor.
- The Jacobian matrices need to exist so that the transformation can be applied. However, there are cases where this may not be true. For Example, there may be systems where the parameters can change abruptly.
- In plethora cases, the calculation of Jacobians and Hessian matrices may be difficult.

The UKF address these issues by using a deterministic sampling approach. The state distribution, approximated by Gaussian random variables (GRV), is represented with minimal set of finely chosen sample points. These sample points are capable to completely capture the true mean and covariance of the GRV. When this is propagated through the true nonlinear system, it captures the accurate posterior mean and covariance. To understand this concept, we are explaining the unscented transformation with mathematical equations .

### 2.2.1. Unscented Transformation (UT)

The sample points, which is the state distribution, approximated by a Gaussian random variables are called as sigma points. These points have weights associated, and hence these are known as weighted sigma points.

The unscented transformation (UT) is used to calculate the statistics of a random variable which undergoes a nonlinear transformation. Consider a random variable  $x$  (dimension  $n$ ) propagating through a nonlinear function,  $y = g(x)$ . Assume  $x$  has mean  $m$  and covariance  $P$ . To calculate the statistics of  $y$ , we form a matrix  $\sqrt{(n + \lambda)P}$  of  $2n + 1$  sigma points according to the following:

$$x^{(0)} = m \quad (4)$$

$$x^{(i)} = m + (\sqrt{(n + \lambda)P})_i, i = 1, \dots, n \quad (5)$$

$$x^{(i)} = m - (\sqrt{(n + \lambda)P})_i, i = n + 1, \dots, 2n \quad (6)$$

Where  $\lambda = \alpha^2 (n + k) - n$  is a scaling parameter. The constant  $\alpha$  is set to small positive values which determines the spread of the sigma points around  $m$ . The constant  $k$  is a secondary scaling parameter which is usually set to 0 or 3-n, and  $\beta$  is used to incorporate prior knowledge of the distribution of  $x$  ( for Gaussian distributions,  $\beta=2$  is optimal).  $(\sqrt{(n + \lambda)P})_i$  is the  $i^{\text{th}}$  column of the matrix square root. These sigma vectors are propagated through the nonlinear function given in Eq.(7).

$$y^{(i)} = g(x^{(i)}) \quad i = 0, \dots, 2n \quad (7)$$

The mean and covariance of  $y$  are approximated using a weighted sample mean and weighted sample covariance of the posterior sigma points,

$$\mu_U \approx \sum_{i=0}^{2n} \mathcal{W}_m^{(i)} y^{(i)} \quad (8)$$

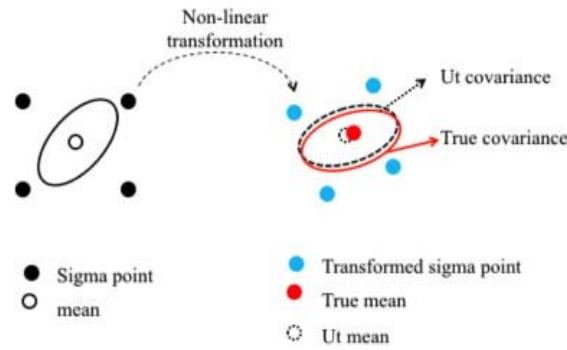
$$S_U \approx \sum_{i=0}^{2n} \mathcal{W}_c^{(i)} \{y^{(i)} - \mu_U\} \{y^{(i)} - \mu_U\}^T \quad (9)$$

With weights  $\mathcal{W}_i$  given by

$$\mathcal{W}_m^{(0)} = \lambda / (n + \lambda) \quad (10)$$

$$\mathcal{W}_c^{(0)} = \frac{\lambda}{n + \lambda} + (1 - \alpha^2 + \beta) \quad (11)$$

$$\mathcal{W}_m^{(i)} = \mathcal{W}_c^{(i)} = 1 / \{2 (n + \lambda)\} \quad i = 1, \dots, 2n \quad (12)$$



Unscented Transformation is a better compared to Taylor expansion. UKF uses the UT during the prediction and correction step.

The UKF make use of the UT described above to give a Gaussian approximation to the filtering solutions of non-linear optimal filtering problems of the form

$$x_k = f(x_{k-1}, k-1) + q_{k-1} \quad (13)$$

$$y_k = h(x_k, k) + r_k, \quad (14)$$

Where  $x_k \in \mathbb{R}^n$  is the state,  $y_k \in \mathbb{R}^n$  is the measurement represent the location information for the  $k^{\text{th}}$  detected object in x-axis and y-axis directions,  $q_{k-1} \sim N(0, Q_{k-1})$  is the Gaussian process noise, and  $r_k \sim N(0, R_k)$  is the Gaussian measurement noise. Using the matrix form of UT, the prediction and update steps of the UKF can computed as follows:

Prediction: Compute the predicted state mean  $m_k^-$  and the predicted covariance  $P_k^-$  as

$$X_{k-1} = [m_{k-1} \dots m_{k-1}] + \sqrt{c} [0 \quad \sqrt{P_{k-1}} \quad -\sqrt{P_{k-1}}] \quad (15)$$

$$\hat{X}_k = f(X_{k-1}, k-1) \quad (16)$$

$$m_k^- = \hat{X}_k W_m \quad (17)$$

$$P_k^- = \hat{X}_k W [\hat{X}_k]^T + Q_{k-1} \quad (18)$$

Update : Updated state is to get measurement data from each frame based on the detection . Compute the predicted mean  $\mu_k$  and covariance of the meaurement  $S_k$ , and the cross-covariance of the state and measurement  $C_k$

$$X_k^- = [m_k^- \dots m_k^-] + \sqrt{c} [0 \quad \sqrt{P_k^-} \quad -\sqrt{P_k^-}] \quad (19)$$

$$Y_k^- = h(X_k^-, k) \quad (20)$$

$$\mu_k = Y_k^- W_m \quad (21)$$

$$S_k = Y_k^- W [Y_k^-]^T + R_k \quad (22)$$

$$C_k = X_k^- W [Y_k^-]^T \quad (23)$$

Then compute the filter gain  $K_k$  and the updated state mean  $m_k$  and covariance  $P_k$

$$K_k = C_k S_k^{-1} \quad (24)$$

$$m_k = m_k^- + K_k [y_k - \mu_k] \quad (25)$$

$$P_k = P_k^- - K_k S_k K_k^T \quad (26)$$

Using the above equations we can easily track the detected vehicle in every frame [20].

Vehicle detection and tracking are one of the supportive systems in ITSs to improve the traffic management systems. In this research, we have taken a large set of positive and negative samples to train the cascade classifier. Most of the literature papers are discussed under static camera conditions however, we are considering moving host. Majority of the paper explains the linear vehicle tracking but we are proposing a non-linear tracking system. This paper is the first to propose the combination of cascade object detector and unscented kalman filter tracking for real-time vehicle detection and tracking.

### 2.3. Experimental results of Vehicle Detection and Tracking on real-time Videos

This method is implemented in the MATLAB platform. For training cascade classifier we used 10,000 positives and 28,000 negative samples. For testing the implemented method we are considering 5 different real-time videos. Table 1 shows the different types of videos used for testing the proposed method.



Table 1.Types of videos used for testing

Input Videos	Number of Frames	Brightness of videos/light intensity	Traffic congestion in roads
Video1	1600	Normal daylight	moderate
Video2	1678	Night time	moderate
Video3	1500	Very brightness	Very high
Video4	300	Normal	less
Video5	300	Less brightness	less

a) Experiment 1

Tested the vehicle detection on video1.Video1 consider as normal illuminance.Fig.7. shows the results after applied vehicle detection on a video1.



Fig.7. vehicle detection algorithm applied to video1



Fig.8. vehicle detection algorithm applied on video2



Fig.9. vehicle detection algorithm applied on video3

b) Experiment 2

Implemented vehicle detection on night videos and the results shows the proposed method is also good in nighttime vehicle detection. Fig.8. shows the results applied after vehicle detection on video2.

c) Experiment 3

Applied vehicle detection method on complicated videos (more brightness video) and the results shows the proposed method is good for detecting all types of real-time applications. Fig.9. show the results after applied vehicle detection on video3.

The cascade classifier performance is then analysed against the parameters such as orientation and color, background contrast, image resolution, and lighting conditions. The experimental results and the analysis shows that the proposed technique is suitable to efficiently detect and track vehicles in real-time applications.

d) Experiment 4

Implemented the training results of vehicle detection and UKF tracking on videos .Fig.10. & Fig.11. shows the results applied after vehicle detection on video and tracked by Unscented Kalman Filter.

To verify the proposed tracking algorithm we used two different real-time videos (video4 & video5). Blue color bounding box represents the detection result and the red color represents the tracking results.



Fig.10. Vehicle detection and tracking in video4





Fig.11.Vehicle detection and tracking in video 5

### 3. Conclusion

In this paper we propose an efficient system to detect and track vehicles from the video frames of a moving host. Vehicle detection and tracking system is an integral part of the driver assistance system. This paper has two major contributions (i) The proposed method has given excellent performance using a moving camera without additional sensors. (ii) The proposed method can detect fast moving and slow-moving objects and hence is suitable to use in real-time applications. Cascade classifiers are used to detect vehicles based on specific features and the classifier is trained using a set of positive and negative images. This technique proved robust against changes in orientation, size and colour. Alongside with Unscented Kalman filter, the limitations of the linear filters can be reduced which makes it suitable for real time vehicle tracking.

As part of future work, we are researching the possibility to add lane detection along with this system. To eliminate the false detections, we are analysing on how to add multiple hypothesis tests in vehicle detection systems.

### Acknowledgments

This work was supported by the 2016 fund of University of Ulsan, Ulsan, Republic of Korea.

### References

- [1] Viola P, Jones M.(2001). "Rapid Object Detection using a Boosted Cascade of Simple Features", Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR 2001; 511-518..
- [2] Choudhury S, Chattopadhyay SP, Hazra TK (2017). "Vehicle Detection and Counting using Haar Feature-Based Classifier". 2017 8th Annual Industrial Automation and Electromechanical Engineering Conference (IEMECON). Bangkok, Thailand.
- [3] Moghimi MM, Nayeri M, Pourahmadi M, Moghimi MK.(2018). "Moving Vehicle Detection Using AdaBoost and Haar-Like Feature in Surveillance Videos". International Journal of Imaging and Robotics ,18 , 94-106.
- [4] Tang Y, Xu YC, Zhang CZ.(2014). "Robust Vehicle Detection Based on Cascade Classifier in Traffic Surveillance System. The Open Automation and Control Systems Journal",6 , 349-354.
- [5] Hasan Y, Asif A, Arif MU, Raza RH.(2016). "Comparative Analysis of Vehicle Detection in Urban Traffic Environment using Haar Cascaded Classifiers and Blob Statistics". FTC 2016 - Future Technologies Conference San Francisco ,USA.
- [6] Zhuang X, Kang W, Wu Q. (2015). "Real-time vehicle detection with foreground-based cascade classifier". IET Image Processing , 10, 289-296.
- [7] Salarpour A, Salarpour A, Fathi M, MirHossein, Dezfoulian.(2011). "Vehicle Tracking Using Kalman Filter And Features". Signal & Image Processing : An International Journal (SIPIJ) .
- [8] Kaattan AH, Saeed TR (2017). "Tracking of Video Objects Based on Kalman Filter". Journal of Babylon University/Engineering Sciences 25 .
- [9] Teoh SS, Bräunl T (2012). "A Reliability Point and Kalman Filter-based Vehicle Tracking Technique".
- [10] Kaur H, Sahambi JS (2016). "Vehicle Tracking in Video Using Fractional Feedback Kalman Filter". IEEE Transactions On Computational Imaging ,2 , 550-561.
- [11] Hostettler R, Birk W, Nordenvaad ML.(2012), "Extended Kalman Filter for Vehicle Tracking using Road Surface Vibration Measurements". 51st IEEE Conference on Decision and Control. Hawaii, USA.
- [12] K MK, R.Velaga N, Ramasankaran R.(2015). "A Two-Stage Extended Kalman Filter Algorithm for vehicle Tracking from GPS Enabled Smart Phones through Crowd-Sourcing". European Transport \ Trasporti Europei.
- [13] Tao J, Klette R.(2012). "Tracking of 2D or 3D Irregular Movement by a Family of Unscented Kalman Filters". JOURNAL OF ICCE , 1-9.
- [14] WANG-lingqun, PAN-shizhu, ZHENG-yingping.(2009). "Moving Vehicle Tracking Based on Unscented Kalman Filter Algorithm". 2009 World Congress on Computer Science and Information Engineering. Los Angeles, California, USA.
- [15] Wang B, Qi Z, Chen S, Liu Z, Ma G.(2017). "Multi-vehicle detection with identity awareness using cascade Adaboost and Adaptive Kalman filter for driver assistant system". PLoS One 2017 March.
- [16] Ponsa D, L'opez A, Serrat J, Lumbrellas F, Graf T.(2005). "Multiple Vehicle 3D Tracking Using an Unscented Kalman Filter". Proceedings of the 8th International IEEE Conference on Intelligent Transportation Systems. Vienna, Austria.
- [17] Chen X, Wang X, Xuan J.(2018). "Tracking Multiple Moving Objects Using Unscented Kalman Filtering Technique". <https://arxiv.org/abs/180201235v1>.

- [18] Gupta M, Behera L, K.Subramanian V, Jamshidi MM,(2015). "A Robust Visual Human Detection Approach With UKF-Based Motion Tracking for a Mobile Robot". IEEE Systems Journal 9.
- [19] Jatoth RK, Kumar TK,(2013). "Real Time Implementation Of Unscented Kalman Filter For Target Tracking". International Journal Of Electronics And Communication Engineering & Technology (IJCET) , 4(208-215).
- [20] Zhang H, Sun J, Dai G, Zhao Y,(2013). "Unscented Kalman filter and its nonlinear application for tracking a moving target". International Journal for Light and Electron Optics (OPTIK).