



Vehicle Object Detection in Aerial Surveillance

Maheswari¹, Ramaswamy reddy²

Student, Dept. of IT, LBRCE, Vijayawada, India¹

Associate Professor, Dept. of IT, LBRCE, Vijayawada, India²

ABSTRACT: This paper presents vehicle object detection system in Aerial Surveillance. This system is helpful in the field of military recognize and traffic controlling system to automatic target recognition and tracking the moving Objects from the videos. We propose a vehicle detection frame work in two phases like training and Detection Phase. In this video is divided in to frames. We performing a pixel wise classification method among the neighbouring pixels in the regions of objects to obtaining the quantitative observations. Colour transforms are applied to separate vehicle and non-vehicle colours effectively. The Canny edge detection method applied to extract the local features of objects. Thus feature extraction process give exact results to classification by considering the vehicle colours and local features. Further Construct the Dynamic Bayesian network (DBN) for vehicle classification. The experiments conducted on a few varieties of aerial videos. The results demonstrate flexibility and good generalization abilities of the proposed system on a challenging data set with aerial surveillance images taken at different heights and under different camera angles.

Keywords: vehicle detection, Aerial Surveillance, Dynamic Bayesian networks (DBNs).

I. INTRODUCTION

The recent growth in the number of vehicles on the roadway network has forced the transport management agencies to rely on advanced technologies to take better decisions. In this perspective aerial surveillance has better place nowadays. Aerial surveillance provides increased monitoring results in case of fast-moving targets because spatial area coverage is greater. One of the main topics in intelligent aerial surveillance is vehicle detection and tracking. The difficulties involved in the aerial Surveillance include the camera motions such as panning, tilting and rotation. Also the different camera heights largely affect the detection results.

Aerial surveillance has a long history in the military for observing enemy activities and in the commercial world for monitoring resources. Such techniques are used in news gathering and search and rescue aerial surveillance has been performed primarily using film. The highly captured still images of an area under surveillance that could later be examined by human or machine analysts. Video capturing dynamic events cannot be understood when compared with aerial images. Feedback and triggering of actions are enabled based on dynamic events and provides crucial and timely intelligence and understanding that is not available. Video observations can be used to find and geo-locate moving objects in real time. Video also provides new technical challenges. Video cameras have lower resolution when compared to the framing cameras. To get the required resolution and to identify objects on the ground, it is necessary to use the telephoto lens, with narrow field of view. This leads to the shortcoming of video in surveillance. It provides a “soda straw” view of scene. The camera should be scanned to cover the extended regions of interest. Observer who is watching this video must pay constant attention, to the objects of interest rapidly moving in and out of the camera field of view. This video lacks a larger visual context—the observer has difficulty perceiving the relative locations of objects seen at one point of time where the object moments seen before. In addition to that geodetic coordinates for objects of interest seen in the video are not available.

II. RELATED WORK

The system, proposed by Hsu-Yung Cheng [1] escaped from the stereotype and existing frameworks of vehicle detection in aerial surveillance, which are region based or sliding window based. Pixel wise classification method is designed for vehicle detection. Hsu-Yung Cheng proposed Hierarchical model proposed by Hinz and Baumgartner [2] which describes different levels of details of vehicle features and detection method based on cascade classifiers has the disadvantage of lots of miss detections. Vehicle detection algorithm based on symmetric property [3] of car shapes is prone to false detections. The high computational complexity of mean-shift segmentation algorithm is a major concern in the existing methods. One method utilizes colour transformation in case of still images and an approach tends to utilize wide area motion imagery. Linet al.[4] proposed a method by subtracting background colours of each

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 10, October 2013

frame and then refined vehicle candidate regions by enforcing size constraints of vehicles. In [4], the authors proposed a moving-vehicle detection method based on cascade classifiers. Multi scale sliding windows are generated in the detection stage. Disadvantage of this method is that there are lots of miss detections on rotating vehicles. Choi and Yang [5] proposed a vehicle detection algorithm using the symmetric property of car shapes. Therefore, they applied a log polar histogram shape descriptor to verify the shape of the candidates. Unfortunately, the shape descriptor is obtained from the fixed vehicle model, such that the algorithm inflexible. The algorithm in [6] relied on mean-shift clustering algorithm for image colour segmentation. The high computational complexity of mean-shift segmentation algorithm is another concern.

III. METHODOLOGIES

A. Hierarchical Model

This work introduces a new approach on automatic vehicle detection in monocular large scale aerial images. Extraction is based on a hierarchical model that describes the prominent vehicle features on different levels in detail. Besides this object properties, the model comprises contextual knowledge, i.e., relations between the vehicle and objects e.g., pavement beside a vehicle and the sun causing a vehicle's shadow projection. This approach neither relies on external information like digital maps or site models, nor is it limited to vial specific vehicle models.

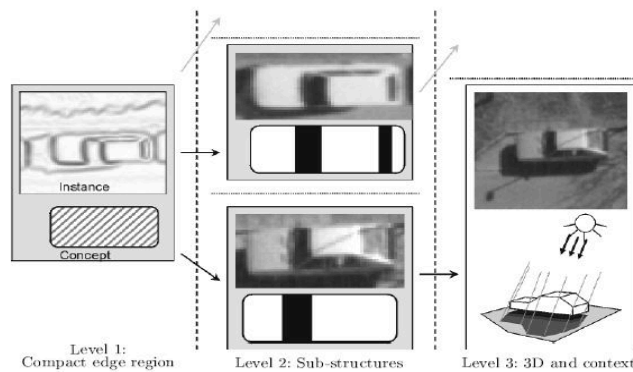
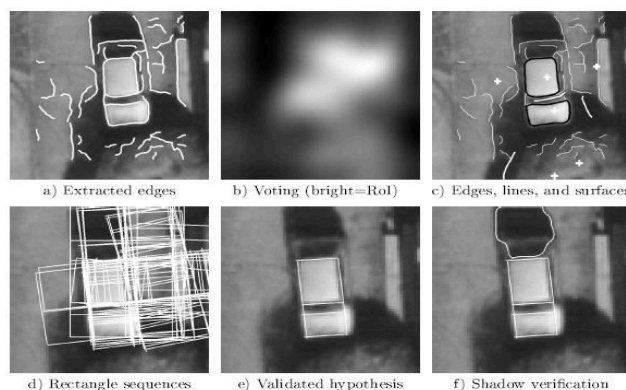


Fig.1.Hierarchical model

B. ROIS (REGIONS OF INTREST)

Which is mainly based on edge voting for convex and compact regions. Hypotheses formation for deriving rectangle sequences front extracted lines, edges, and surfaces. validation and selection which includes the radiometric and geometric analysis of rectangle sequences, and. Verification using vehicle models and their local context. For an illustration of the individual steps. In order to avoid time, consuming grouping algorithms in the early stages of extraction, we first focus on generic image features as edges, lines, and surfaces.



International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 10, October 2013

Fig2: Individual Steps of Vehicle Detection

C. 3D Model Generation and Verification

In order to approximate the 3D shape of a vehicle hypothesis, a particular height profile is selected from a set of predefined profiles. Please note that the hypothesis' underlying rectangles remain unchanged, i.e., the height values of the profile refer to the image edges perpendicular to the vehicle direction. The selection of the profiles depends on the extracted sub structures, i.e., the shape of the validated rectangle sequence. We distinguish rectangle sequences corresponding to 3 types of vehicles: *hatch-back cars*, *saloon cars*, and other vehicles such as vans, small Trucks, etc. In contrast to hatch -back and saloon cars, the derivation of an accurate height profile for the last category would require a deeper analysis of the hypotheses (e.g., for an unambiguous determination of the vehicle orientation). Hence, in this case, we approximate the height profile only roughly by an elliptic arc having a constant height. Offset, above the ground. After creating a 3D model from the 2D hypothesis and the respective height profile we are able to predict the boundary of a vehicle's shadow projection on the underlying road surface. A vehicle hypothesis is judged as verified if a dark and homogeneous region is extracted besides the shadowed part of the vehicle and a bright and homogeneous region besides the illuminated part, respectively.



Fig.3.3D Models of cars

D. 3Planer Detection Technique

If we try to detect the vehicle based on three dimensions at that time detection is very crucial.

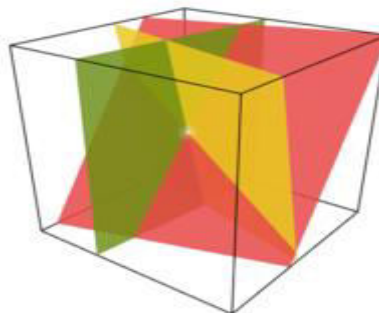


Fig.4.plane shows the Intersect point to the share fields

Secret sharing (also called secret splitting) refers to method for distributing a secret among a group of participants; each of them is allocated to share the secret. The secret is reconstructed only when a sufficient number, of possible different types, of shares are combined together individual shares are of no use on their own.

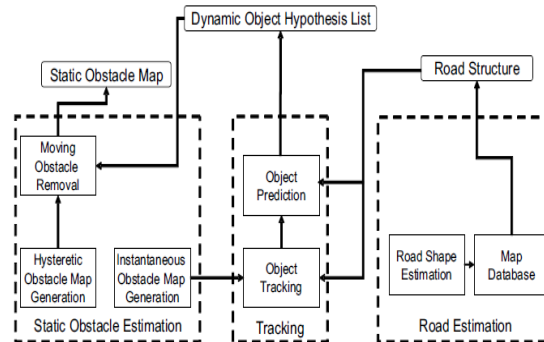
E. Perception system

During the competition the vehicles encountered various typical scenarios of urban driving. They had to interact with other traffic which was human or machine driven. It describes the perception approach given by Team Tartan Racing, winner of the competition. Focus is set on the detection and tracking of other vehicles surrounding the robot. The present approach allows a situation specific interpretation of perception data through situation assessment algorithms keeping the perception algorithms situation independent.

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 10, October 2013



Architecture of Perception System. Encapsulated in dashed lines the subsystems. Arrows show the data flow.

Fig.5. Architecture of Perception System

IV. VEHICLE DETECTION FRAMEWORK

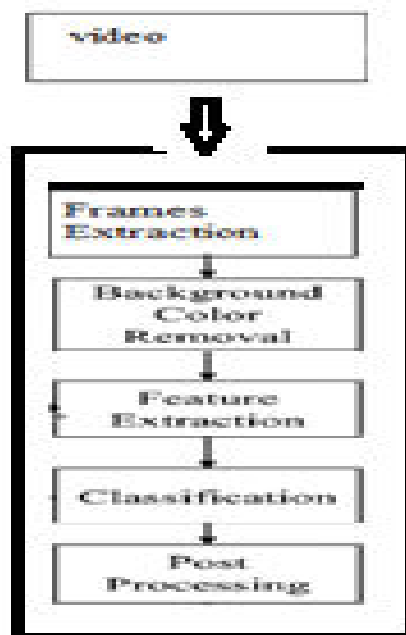


Fig.6. System Frame Work

In this paper, we design a new vehicle detection framework that preserves the advantages of the existing works and avoids their drawbacks.

We read the input video and extract the number of frames from that video. Performing background colour removal cannot only reduce false alarms but also speed up the detection process. We extract the feature from the image frame. We do the following Edge Detecting, Corner Detecting, colour Transform and colour classify. The Frame edge Image is able to transfer by performing Detect edge, corners and places for Transform colour. we perform pixel wise classification for vehicle detection using DBNs. (Dynamic Bayesian Network).we use morphological operations to enhance the detection mask and perform connected component labelling to get the vehicle objects .In this paper, we do not perform region based classification, highly depend on results of colour segmentation algorithm of mean shift. Generating multiscale sliding windows is not necessary.

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 10, October 2013

A. Frame Extraction

In module we read the input video and extract the number of frames from that video. The frames are formed dynamically with pixel calculation, Edge detection and error correction.

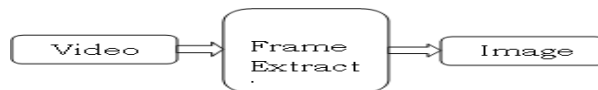


Fig.7.frame Extraction

B. Background Colour Removal

In this module we construct the colour histogram of each frame and remove the colours that appear frequently in the picture. These removed pixels need not to be considered in subsequent detection processes. Performing background colour removal not only reduces false alarms but also speed up the detection process.

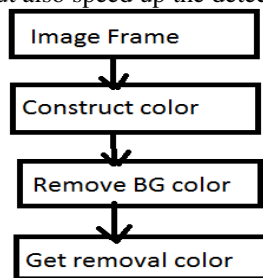


Fig.8.Background Colour Removal

C. Feature Extraction

In this module we extract the features from image frame. We do the following Edge Detecting, Corner Detecting, colour Transform and colour classify as shown in fig 4. Feature extraction is performed in both the training phase and the detection phase we consider local features and colour features in this paper.

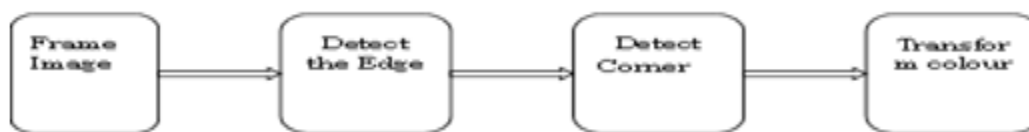


Fig.9.Feature Extraction

D. DBNS for Classification

Dynamic Bayesian Networks (DBNs) are used for vehicle classification in video. A Bayesian network is a directed acyclic graph $G=(V;E)$ where the nodes (vertices) represent random variables from the domain of interest and the arcs (edges) symbolize the direct dependencies between the random variables but a node in a Bayesian network is conditional on its parent's values so

$$p(x_1, x_2, \dots, x_n) = \prod_{i=1}^n p(x_i | \text{parents}(X_i))$$

Where $p(x_1; x_2; \dots; x_n)$ is an abbreviation for $p(X_1=x_1 \wedge X_2=x_2 \wedge \dots \wedge X_n=x_n)$. In other words, Bayesian network models probability distribution if each variable is conditionally independent of all its non-descendants in the graph given the value of its parents.

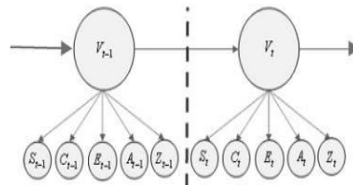
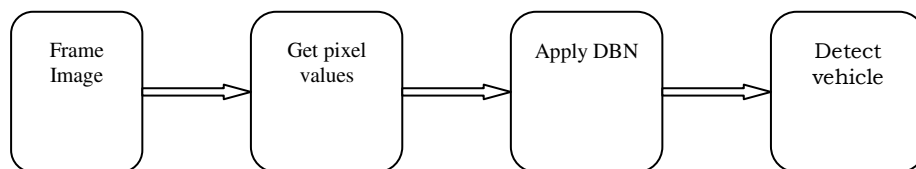


Fig.10. DBN model for pixel wise classification

We perform pixel wise classification for vehicle detection using DBNs [13]. The design of the DBN model is illustrated in Fig. 10. Node V_t indicates if a pixel belongs to a vehicle at time slice t . The state of V_t is dependent on the state of V_{t-1} . Moreover, at each time slice t , the state v_t has influences on the observation nodes S_t, C_t, F_t, A_t and Z_t . The observations are assumed to be independent of one another. The definitions of these observations are explained in the previous subsection. we obtain the conditional probability tables of the DBN model via Expectation-maximization algorithm by providing the ground truth labelling of each pixel and its corresponding observed features from several training videos. In the detection phase, the Bayesian rule is used to obtain the probability of the pixel belongs to the Vehicle,

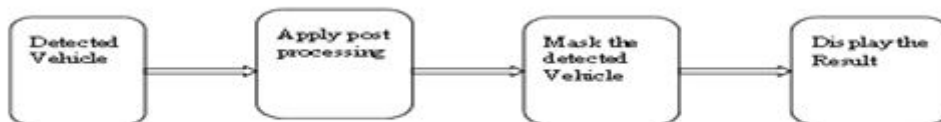
$$P(V_t | S_t, C_t, E_t, A_t, Z_t, V_{t-1}) = P(V_t | S_t)P(V_t | C_t)P(V_t | E_t)P(V_t | A_t)P(V_t | Z_t)P(V_t | V_{t-1})$$

By considering the joint probability $P(V_t | S_t, C_t, F_t, A_t, Z_t, V_{t-1})$ is the probability that a pixel belongs to a vehicle pixel at time slice t given all the observations and the state of the previous time instance. Term $P(V_t | S_t)$ is defined as the probability that a pixel belongs to a vehicle pixel at time slice t , given observation S_t at time instance [is defined in (5)]. Terms $P(V_t | C_t), P(V_t | E_t), P(V_t | A_t), P(V_t | Z_t)$, and $P(V_t | V_{t-1})$ are similarly defined. The proposed framework can also utilize a Bayesian network (BN) to classify a pixel as a vehicle or non- vehicle pixel.



E. Post Processing

Need to identify a vehicle from that we need to perform masking operation of each vehicle. We are removing the frequent pixel because it involves the commonality pixels. We are able to recognise each vehicle by applying post Processing..



V. EXPERIMENTAL RESULTS

Experimental results are demonstrated here. To analyse the performance of the proposed system, various video sequences with different scenes and different filming altitudes are used.

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

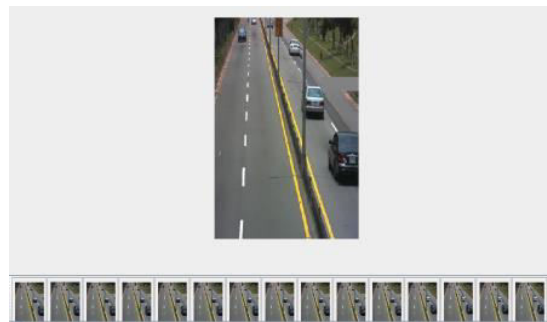
Vol. 2, Issue 10, October 2013

A. Frame Extraction

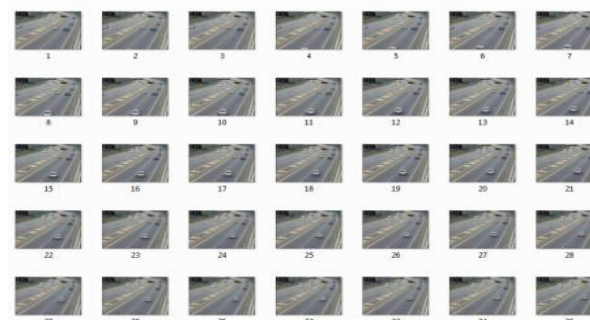
Input video is taken and extract the number of frames from that video .The frames are formed dynamically with pixel calculation.



(a) Input video



(b) Multiple no of Frames generated



c) Dynamic Frame generation

B. Background Colour Removal

These removed pixels need not to be considered in subsequent detection processes. Performing background colour removal not only reduces false alarms but also speed up the detection process.

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

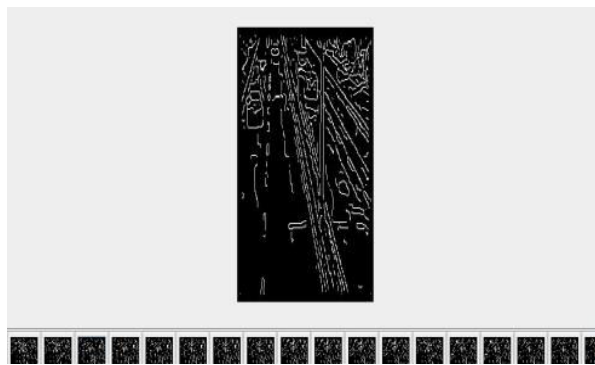
Vol. 2, Issue 10, October 2013



d) Background colour removal results

C. Detect Edge

The Frame edge Image is able to transfer by performing Detect edge.



e) Edge Detection

D. Colour Classification

When employing SVM, we need to select the block size to form a sample and perform vehicle colour classification.



f) Colour classification

E. Post Processing

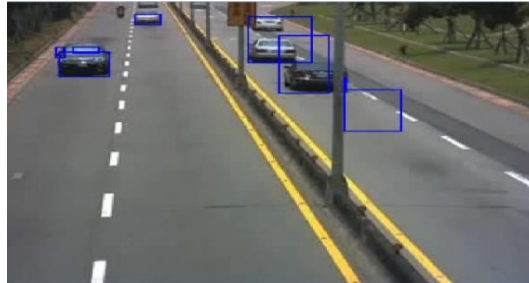
We use morphological operations to enhance the detection mask and perform connected component labelling to get the vehicle. In the post processing stage we eliminate objects that are impossible to be vehicles.



International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 10, October 2013



g) Detecting Each and Every Vehicle

V. CONCLUSION

An automatic vehicle detection system for aerial surveillance does not assume any prior information of camera heights, vehicle sizes, and aspect ratios. This system performs region-based classification, which would highly depend on computational intensive colour segmentation algorithms such as mean shift. We have not generated multi scale sliding windows that are not suitable for detecting rotated vehicles either. Instead, we have proposed a pixel wise classification method for the vehicle detection using DBNs. In spite of performing pixel wise classification, relations among neighbouring pixels in a region are preserved in the feature extraction process. Therefore, the extracted features comprise not only pixel level information but also region-level information. Since the colours of the vehicles would not dramatically change due to the influence of the camera angles and heights, we use only a small number of positive and negative samples to train the SVM for classifying the vehicle colour. The number of frames required to train the DBN is very small. Overall, the entire framework does not require a large amount of training samples. We have also applied moment preserving to enhance the canny edge detector, which increases the adaptability and the accuracy for detection in various aerial images.

REFERENCES

- [1] Hsu-Yung Cheng, Member, IEEE, Chih-Chia Weng, and Yi-Ying Chen, "Vehicle Detection in Aerial Surveillance Using Dynamic Bayesian Networks" March 21, 2012. Lecture Notes in Computer Science Project 99-2628-E-008-098.
- [2] S. Hinz and A. Baumgartner, "Vehicle detection in aerial "images using generic features, grouping, and context," in Proc. DAGM Symp., Sep. 2001, vol. 2191, Lecture Notes in Computer Science, pp. 45–52.
- [3] J. Y. Choi and Y. K. Yang, "Vehicle detection from aerial images using local shape information," Adv. Image Video Technol., vol. 5414, Lecture Notes in Computer Science, pp. 227–236, Jan. 2009.
- [4] R. Lin, X. Cao, Y. Xu, C. Wu, and H. Qiao, "Airborne moving vehicle detection for urban traffic surveillance," in Proc. 11th Int. IEEE Conf. Intell. Transp. Syst., Oct. 2008, pp. 163–167.
- [5] R. Lin, X. Cao, Y. Xu, C. Wu, and H. Qiao, "Airborne moving vehicle detection for video surveillance of urban traffic," in Proc. IEEE Intell. Veh. Symp., 2009, pp. 203–208.
- [6] I. Emst, S. Sujew, K. U. Thiessenhusen, M. Hetscher, S. Rabmann, and M. Ruhe, "LUMOS—Airbone traffic monitoring system," in Proc. IEEE Intell. Transp. Syst., Oct. 2003, vol. 1, pp. 753–759.
- [7] H. Cheng and D. Butler, "Segmentation of aerial surveillance video using a mixture of experts," in Proc. IEEE Digit. Imaging Compute. —tech. Appl., 2005, p. 66.
- [8] S. Srinivasan, H. Latchman, J. Shea, T. Wong, and J. McNair, "Airborne traffic surveillance systems: Video surveillance of highway traffic," in Proc. ACM 2nd Int. Workshop Video Surveillance Sens. Netw., 2004, pp. 131–135.
- [9] C. G. Harris and M. J. Stephens, "A combined corner and edge detector," in Proc. 4th Alvey Vis. Conf., 1988, pp. 147–151.
- [10] J. F. Canny, "A computational approach to edge detection," IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.
- [11] W. H. Tsai, "Moment-preserving thresholding: A new approach," Compute. Vis. Graph, Image Process, vol. 29, no. 3, pp. 377–393, 1985.
- [12] L. W. Tsai, J. W. Hsieh, and K. C. Fan, "Vehicle detection using normalized colour and edge map," IEEE Trans. Image Process., vol. 16, no. 3, pp. 850–864, Mar. 2007.
- [13] S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 2nd ed. Englewood Cliffs, NJ: Prentice