International Journal of Communication Networks and Security

Volume 2 | Issue 1 Article 13

January 2013

AERIAL SURVEILLANCE FOR VEHICLE DETECTION USING DBN AND CANNY EDGE DETECTOR

SK. MADEENA
QCET, Nellore,India, SK.MADEENA@gmail.com

SD.AFZAL AHMED
QCET, Nellore,India, AFZALAHMED@gmail.com

P. BABU
QCET, Nellore,India, P.BABU@gmail.com

Recommended Citation

MADEENA, SK.; AHMED, SD.AFZAL; and BABU, P. (2013) "AERIAL SURVEILLANCE FOR VEHICLE DETECTION USING DBN AND CANNY EDGE DETECTOR," *International Journal of Communication Networks and Security*: Vol. 2 : Iss. 1 , Article 13.
DOI: 10.47893/IJCNS.2013.1070
Available at: https://www.interscience.in/ijcns/vol2/iss1/13

AERIAL SURVEILLANCE FOR VEHICLE DETECTION USING DBN AND CANNY EDGE DETECTOR

SK. MADEENA[1], SD.AFZAL AHMED[2], P.BABU[3]

[1]PG Student, QCET, [2,3,]Associate Professor, QCET, Nellore

Abstract- We present an automatic vehicle detection system for aerial surveillance in this paper. In this system, we escape from the stereotype and existing frameworks of vehicle detection in aerial surveillance, which are either region based or sliding window based. We design a pixel wise classification method for vehicle detection. The novelty lies in the fact that, in spite of performing pixel wise classification, relations among neighboring pixels in a region are preserved in the feature extraction process. We consider features including vehicle colors and local features. For vehicle color extraction, we utilize a color transform to separate vehicle colors and non-vehicle colors effectively. For edge detection, we apply moment preserving to adjust the thresholds of the Canny edge detector automatically, which increases the adaptability and the accuracy for detection in various aerial images. Afterward, a dynamic Bayesian network (DBN) is constructed for the classification purpose. We convert regional local features into quantitative observations that can be referenced when applying pixel wise classification via DBN. Experiments were conducted on a wide variety of aerial videos. The results demonstrate flexibility and good generalization abilities of the proposed method on a challenging data set with aerial surveillance images taken at different heights and under different camera angles.

I. INTRODUCTION:

Aerial surveillance has a long history in the military for observing enemy activities and in the commercial world for monitoring resources such as forests and crops. Similar imaging techniques are used in aerial news gathering and search and rescue aerial surveillance has been performed primarily using film or electronic framing cameras. The objective has been to gather high-resolution still images of an area under surveillance that could later be examined by human or machine analysts to derive information of interest. Currently, there is growing interest in using video cameras for these tasks. Video captures dynamic events that cannot be understood from aerial still images. It enables feedback and triggering of actions based on dynamic events and provides crucial and timely intelligence and understanding that is not otherwise available. Video observations can be used to detect and geo-locate moving objects in real time and to control the camera, for example, to follow detected vehicles or constantly monitor a site. However, video also brings new technical challenges. Video cameras have lower resolution than framing cameras. In order to get the resolution required to identify objects on the ground, it is generally necessary to use a telephoto lens, with a narrow field of view. This leads to the most serious shortcoming of video in surveillance— it provides only a “soda straw” view of the scene. The camera must then be scanned to cover extended regions of interest. An observer watching this video must pay constant attention, as objects of interest move rapidly in and out of the camera field of view. The video also lacks a larger visual context—the observer has difficulty perceiving the relative locations of objects seen at one point in time to objects seen moments before. In addition, geodetic coordinates for objects of interest seen in the video are not available.

One of the main topics in aerial image analysis is scene registration and alignment. Another very important topic in intelligent aerial surveillance is vehicle detection and tracking. The challenges of vehicle detection in aerial surveillance include camera motions such as panning, tilting, and rotation. In addition, airborne platforms at different heights result in different sizes of target objects.

In this paper, we design a new vehicle detection framework that preserves the advantages of the existing works and avoids their drawbacks. The framework can be divided into the training phase and the detection phase. In the training phase, we extract multiple features including local edge and corner features, as well as vehicle colors to train a dynamic Bayesian network (DBN). In the detection phase, we first perform background color removal. Afterward, the same feature extraction procedure is performed as in the training phase. The extracted features serve as the evidence to infer the unknown state of the trained DBN, which indicates whether a pixel belongs to a vehicle or not. In this paper, we do not perform region-based classification, which would highly depend on results of color segmentation algorithms such as mean shift. There is no need to generate multiscale sliding windows either. The distinguishing feature of the proposed framework is that the detection task is based on pixelwise classification. However, the features are extracted in a neighborhood region of each pixel. Therefore, the extracted features comprise not only pixel-level information but also relationship among neighboring pixels in a region. Such design is more effective and efficient than region-based or multiscale sliding window detection methods.

Existing System

Hinz and Baumgartner utilized a hierarchical model that describes different levels of details of vehicle features. There is no specific vehicle models assumed, making the method flexible. However, their system would miss vehicles when the contrast is weak or when the influences of neighboring objects are present.

Cheng and Butler considered multiple clues and used a mixture of experts to merge the clues for vehicle detection in aerial images. They performed color segmentation via mean-shift algorithm and motion analysis via change detection. In addition, they presented a trainable sequential maximum a posterior method for multiscale analysis and enforcement of contextual information. However, themotion analysis algorithm applied in their system cannot deal with aforementioned camera motions and complex background changes. Moreover, in the information fusion step, their algorithm highly depends on the color segmentation results.

Lin et al. proposed a method by subtracting background colors of each frame and then refined vehicle candidate regions by enforcing size constraints of vehicles. However, they assumed too many parameters such as the largest and smallest sizes of vehicles, and the height and the focus of the airborne camera. Assuming these parameters as known priors might not be realistic in real applications.

The authors proposed a moving-vehicle detection method based on cascade classifiers. A large number of positive and negative training samples need to be collected for the training purpose. Moreover, multiscale sliding windows are generated at the detection stage. The main disadvantage of this method is that there are a lot of miss detections on rotated vehicles. Such results are not surprising from the experiences of face detection using cascade classifiers. If only frontal faces are trained, then faces with poses are easily missed. However, if faces with poses are added as positive samples, the number of false alarms would surge.

Proposed System

In this paper, we design a new vehicle detection framework that preserves the advantages of the existing works and avoids their drawbacks. The framework can be divided into the training phase and the detection phase. In the training phase, we extract multiple features including local edge and corner features, as well as vehicle colors to train a dynamic Bayesian network (DBN). In the detection phase, we first perform background color removal. Afterward, the same feature extraction procedure is performed as in the training phase. The extracted features serve as the evidence to infer the unknown state of the trained DBN, which indicates whether a pixel belongs to a vehicle or not. In this paper, we do not perform region-based classification, which would highly depend on results of color segmentation algorithms such as mean shift. There is no need to generate multi-scale sliding windows either. The distinguishing feature of the proposed framework is that the detection task is based on pixel wise classification. However, the features are extracted in a neighborhood region of each pixel. Therefore, the extracted features comprise not only pixel-level information but also relationship among neighboring pixels in a region. Such design is more effective and efficient than region-based or multi scale sliding window detection methods

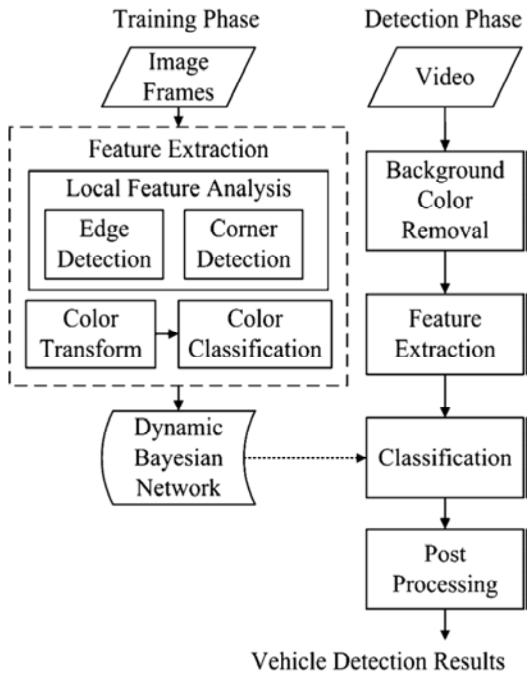


Fig. 1. Proposed system framework

Software Requirement Specification:
Software Requirement
Operating System: Windows Xp
Technology : java 1.6
IDE : Net Beans
Hardware Requirement
Processor : Any Processor above 500 MHz
Ram : 1 GB.
Hard Disk : 10 GB.
Compact Disk : 650 Mb.
Input device : Standard Keyboard and Mouse.

Modules

1.Frame Extraction

In module we read the input video and extract the number of frames from that video.

2. Background color removal
In this module we construct the color histogram of each frame and remove the colors that appear most frequently in the scene.

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

3. Feature Extraction
In this module we extract the feature from the image frame. In this module we do the following Edge Detection, Corner Detection, color Transformation and color classification.

4. Classification
In this module we perform pixel wise classification for vehicle detection using DBNs. (Dynamic Bayesian Network). In the training stage, we obtain the conditional probability tables of the DBN model via expectation-maximization algorithm by providing the ground-truth labeling 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 that a pixel belongs to a vehicle.

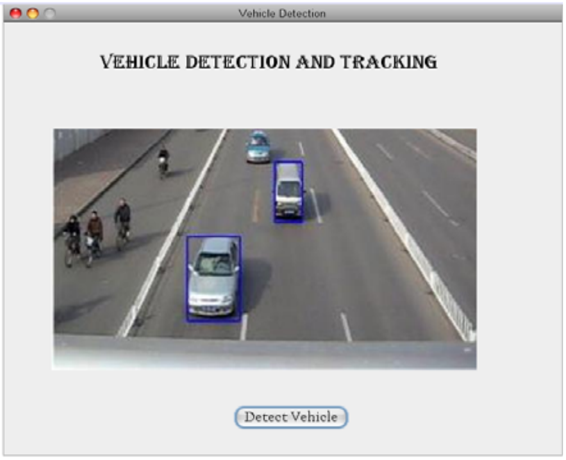
5. Post processing
In this module we use morphological operations to enhance the detection mask and perform connected component labeling to get the vehicle objects. The size and the aspect ratio constraints are applied again after morphological operations in the post processing stage to eliminate objects that are impossible to be vehicles.

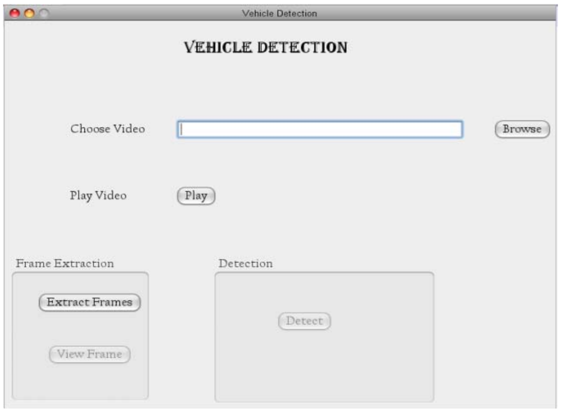
Experimental Results
Experimental results are demonstrated here. To analyze the performance of the proposed system, various video sequences with different scenes and different filming altitudes are used.

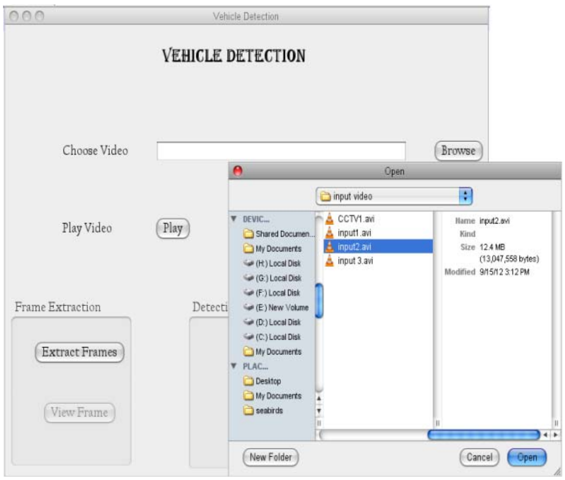
The experimental videos are displayed in. Note that it is infeasible to assume prior information of camera heights and target object sizes for this challenging data set. When performing background color removal, we quantize the color histogram bins as 16x 16 x16. Colors corresponding to the first eight highest bins are regarded as background colors and removed from the scene.

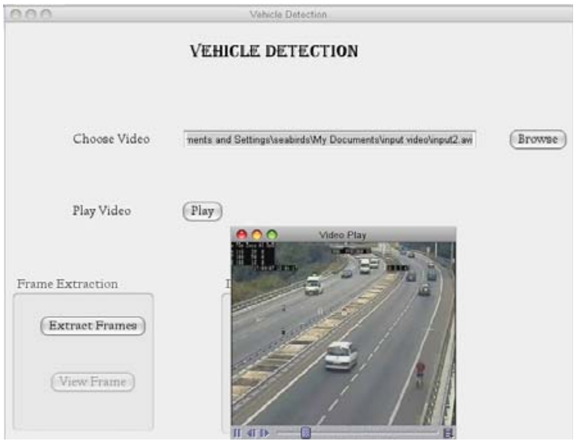
To examine the necessity of the background removal process and the enhanced edge detector, we list the detection accuracy of four different scenarios.

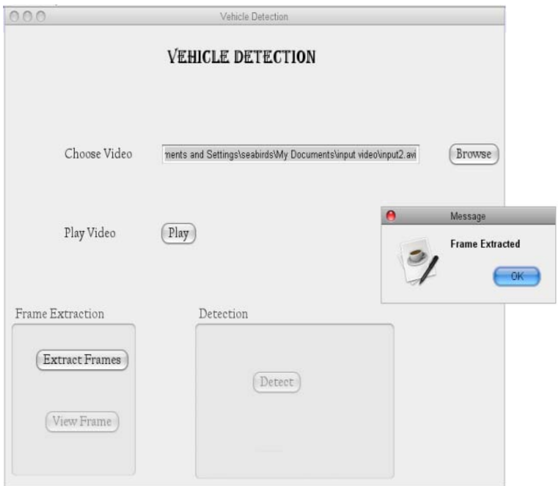
We can observe that the background removal process is important for reducing false positives and the enhanced edge detector is essential for increasing hit rates











However, for the experimental data set, it is very difficult to select one set of parameters that suits all videos. Setting the parameters heuristically for the data set would result in low hit rate and high false positive numbers. The cascade classifiers used in [11] need to be trained by a large number of positive and negative training samples. The number of training samples required in [11] is much larger than the training samples used to train the SVM classifier . The colors of the vehicles would not dramatically change due to the influence of the camera angles and heights. However, the entire appearance of the vehicle templates would vary a lot under different heights and camera angles. When training the cascade classifiers, the large variance in the appearance of the positive templates would decrease the hit rate and increase the number of false positives.Moreover, if the aspect ratio of the multiscale detection windows is fixed, large and rotated vehicles would be often missed. The symmetric property method proposed in [12] is prone to false detections such as symmetrical details of buildings or road markings. Moreover, the shape descriptor used to verify the shape of the candidates is obtained from a fixed vehicle model and is therefore not flexible. Moreover, in some of our experimental data, the vehicles are not completely symmetric due to the angle of the camera. Therefore, the method in [12] is not able to yield satisfactory results. Compared with these methods, the proposed vehicle detection framework does not depend on strict vehicle size or aspect ratio constraints. Instead, these constraints are observations that can be learned by BN or DBN. The training process does not require a large amount of training samples. The results . demonstrate flexibility and good generalization ability on a wide variety of aerial surveillance scenes under different heights and camera angles. It can be expected that the performance of DBN is better than that of the BN., we display the detection results using BN and DBN The colored pixels are the ones that are classified as vehicle pixels by BN or DBN. The ellipses are the final vehicle detection results after performing postprocessing. DBN outperforms BN because it includes information along time. When observing detection results of consecutive frames, we also notice that the detection results via DBN are more stable. The reason is that, in aerial surveillance, the aircraft carrying the camera usually follows the vehicles on the ground, and therefore, the positions of the vehicles would not have dramatic changes in the scene even when the vehicles are moving in high speeds. Therefore, the information along the time contributed by helps stabilize the detection results in the DBN. some detection error cases display the original image frames, and displays the detection results. The black arrows in) indicate the miss detection or false positive cases. In the first row of, the rectangular structures on the buildings are very similar to vehicles. Therefore, sometimes, these rectangular structures would be detected as vehicles incorrectly. In the second row of, the miss detection is caused by the low contrast and the small size of the vehicle. However, other vehicles are successfully detected in this challenging setting. For future work, performing vehicle tracking on the detected vehicles can further stabilize the detection results. Automatic vehicle detection and tracking could serve as the foundation for event analysis in intelligent aerial surveillance systems.

Conclusion

In this paper, we have proposed an automatic vehicle detection system for aerial surveillance that does not assume any prior information of camera heights, vehicle sizes, and aspect ratios. In this system, we have not performed region-based classification, which would highly depend on computational intensive color segmentation algorithms such as mean shift. We have not generated multiscale sliding windows that are not suitable for detecting rotated vehicles either. Instead, we have proposed a pixelwise classification method for the vehicle detection using DBNs. In spite of performing pixelwise classification, relations among neighboring 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 colors 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 SVMfor vehicle color classification. Moreover, 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. The experimental results demonstrate flexibility and good generalization abilities of the proposed method on a challenging data set with aerial surveillance images taken at different heights and under different camera angles. For future work, performing vehicle tracking on the detected vehicles can further stabilize the detection results. Automatic vehicle detection and tracking could serve as the foundation for event analysis in intelligent aerial surveillance systems.

REFERENCE

[1] R. Kumar, H. Sawhney, S. Samarasekera, S. Hsu, T. Hai, G. Yanlin, K. Hanna, A. Pope, R. Wildes, D. Hirvonen,M. Hansen, and P. Burt, "Aerial video surveillance and exploitation," Proc. IEEE, vol. 89, no. 10, pp. 1518–1539, 2001.

[2] I. Emst, S. Sujew, K. U. Thiessenhusen,M. Hetscher, S. Rabmann, and M. Ruhe, "LUMOS—Airbome traffic monitoring system," in Proc. IEEE Intell. Transp. Syst., Oct. 2003, vol. 1, pp. 753–759.

[3] L. D. Chou, J. Y. Yang, Y. C. Hsieh, D. C. Chang, and C. F. Tung, "Intersection- based routing protocol for VANETs,"Wirel. Pers. Commun., vol. 60, no. 1, pp. 105–124, Sep. 2011.

[4] 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.

[5] A. C. Shastry and R. A. Schowengerdt, "Airborne video registration and traffic-flow parameter estimation," IEEE Trans. Intell. Transp. Syst., vol. 6, no. 4, pp. 391–405, Dec. 2005.

[6] H. Cheng and J.Wus, "Adaptive region of interest estimation for aerial surveillance video," in Proc. IEEE Int. Conf. Image Process., 2005, vol. 3, pp. 860–863

[7] 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.

[8] H. Cheng and D. Butler, "Segmentation of aerial surveillance video using a mixture of experts," in Proc. IEEE Digit. Imaging Comput. —Tech. Appl., 2005, p. 66.

[9] 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.

[10] L. Hong, Y. Ruan, W. Li, D. Wicker, and J. Layne, "Energy-based video tracking using joint target density processing with an application to unmanned aerial vehicle surveillance," IET Comput. Vis., vol. 2, no. 1, pp. 1–12, 2008

[11] 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.

[12] 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.

[13] C. G. Harris and M. J. Stephens, "A combined corner and edge detector," in Proc. 4th Alvey Vis. Conf., 1988, pp. 147–151.

[14] 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