DETECTION OF DYNAMIC VEHICLES USING DYNAMIC BAYESIAN NETWORKS AND PIXEL WISE CLASSIFICATION

Mustafa khader Syed[1]

[1]*Student, CSE Department, Prakasam Engineering College, A.P., India,* ***syed.soft7@gmail.com***

Abstract

There is a fast growth in computer technology and increasing needs in security and studies of target vehicle detection in aerial surveillance using image processing techniques. This paper presents an automatic vehicle detection system for aerial surveillance. 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. A Dynamic Bayesian Network (DBN) is constructed for classification purpose. A well trained DBN can estimate the probability of a pixel belonging to a vehicle or not. It also relates among neighboring pixels in a region.

Index Terms: *Vehicle Detection, Aerial Surveillance, Dynamic Bayesian Networks (DBNs), Canny Edge Detector*

--- *** ---

1. INTRODUCTION

OVER the past few years vehicle classification has been widely studied as part of the broader vehicle recognition research area. A vehicle classification system is essential for effective transportation systems (e.g., traffic management and toll systems), parking optimization, law enforcement, autonomous navigation, etc. A common approach utilizes vision-based methods and employs external physical features to detect and classify a vehicle in still images and video streams. A human being may be capable of identifying the class of a vehicle with a quick glance at the digital data (image, video) but accomplishing that with a computer is not as straight forward. Several problems such as occlusion, tracking a moving object, shadows, rotation, lack of color invariance, and many more must be carefully considered in order to design an effective and robust automatic vehicle classification system which can work in real-world conditions.

The increase in the number of vehicles on the roadway network has forced the transport management agencies to depend on advanced technologies to take better decisions. In this perspective aerial surveillance has better place nowadays. Aerial surveillance provides 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. 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. 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.

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).

2. 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 meanshift 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.

3. DISADVANTAGES

• Hierarchical model system would miss vehicles when the contrast is weak or when the influences of neighboring objects are present.
• Existing method result highly depends on the color segmentation
• a lot of miss detections on rotated vehicles
• a vehicle tends to be separated as many regions since car roofs and windshields usually have different colors
• high computational complexity

4. 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 shown in figure 1. 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.

Here, we elaborate each module of the proposed system framework in detail.

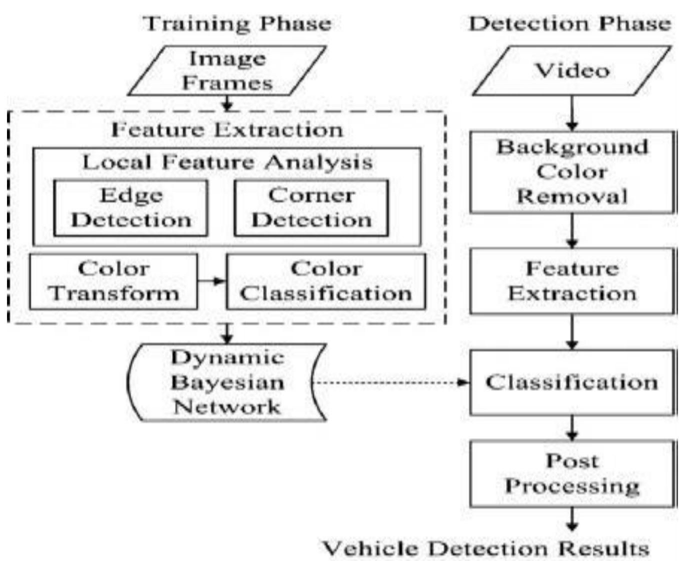

Fig-1: Proposed system framework

3.1 Background Color Removal

Since nonvehicle regions cover most parts of the entire scene in aerial images, we construct the color histogram of each frame and remove the colors that appear most frequently in the scene. Take Fig. 2 for example, the colors are quantized into 48 histogram bins. Among all histogram bins, the 12th , 21st , and 6th bins are the highest and are thus regarded as background colors and removed. Quantize the color histogram bins as 16*16*16. Colors corresponding to the first eight highest bins are regarded as background colors and removed from the scene.

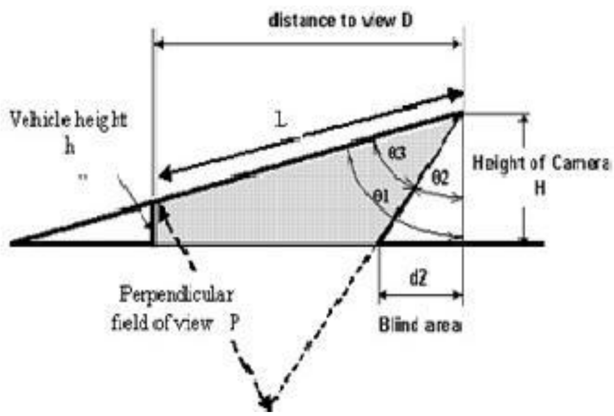

Fig- 2: Camera setting

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.

3.2 Feature Extraction

Feature extraction is performed in both the training phase and the detection phase. We consider local features and color features in this paper.

Local Feature Analys : Corners and edges are usually located in pixels with more information. TheHarries cornerdetector can be used to detect corners. To detect edges, moment-preserving thresholding method on the classical Canny edge detector to select thresholds adaptively according to different scenes. In the Canny edge detector, there are two important thresholds.i.e., the lower threshold T_{low} and higher threshold T_{high} . As the illumination in every aerial image differes, the desired thresholds vary and adaptive thresholds are required. The computation of Tsai's moment-preserving method is deterministic without iterations for L-level thresholding with L < 5. Its derivation of thresholds is described a follows.

Let f be an image with n pixels and f(x,y) denoted the gray value at pixel (x,y). The i^{th} moment m, of f is defined as

$$m_i = (1/n) \sum n_j (z_j)^i \quad \text{for each } j .$$

$$= \sum p_j (z_j)^I \quad \text{for each j with } I = 1,2,3,\ldots. \qquad (1)$$

Where n_j is the total number of pixels in image f with gray value z_j and $p_j = n_j / n$. For bi-level thresholding, the threshold T is selected such that the first three moments of image f are preserved in the resulting bi-level image g. Let all the below-thresholdgray values in f be replaced by z_0 and all the above-threshold gray values be replaced by z_1 , we can solve for p_0 and p_1 based on the moment-preserving principle.

The desired threshold T is computed using

$$P_0 = (1/n) \sum n_j \quad \text{for j from 1 to T} \qquad (2)$$

In order to detect edges, the gradient magnitude G(x,y) of each pixel is used to replace the grayscale value f(x,y) in Tsai's method. Then, the adaptive threshold found by (2) is used as the higher threshold T_{high} in the Canny edge detector.

The lower threshold can be defined as $T_{low} = 0.1 \times (G_{max} - G_{min}) + G_{min}$, where G_{max} and G_{min} represent maximum and minjimum gradient magnitude in the image respectively. Thresholds automatically and dynamically selected by our method give better performance on the edge detection.

Color Transform and Color Classification : New color model to can be separate Vehicle colors from non vehicle colors effectively. This color model transforms (R,G,B) color components into the color domain (u,v), i.e.,

$$u_p = (2Z_p - G_p - B_p) / Z_p. \qquad (3)$$

$$v_p = \text{Max}((B_p - G_p)/Z_p , (R_p - B_p) / Z_p) \qquad (4)$$

Where (R_p, G_p, B_p) is the R,G,B and B colour components of pixel *p* and $Z_p = (R_p + G_p + B_p) / 3$. It has been shown in [16] that all the vehicle colours are concentrated in a much smaller area on the *u-v* plane than in other colour spaces and are therefore easier to be separated from nonvehicle colours. Although the colour transform proposed in [16] did not aim for aerial images, we have found that the separability property still presents in aerial images. As shown in Fig. 3, we can observe that vehicle colours and nonvehicle colours have less overlapping regions under the (u,v) colour model. Therefore, we apply the colour transform to obtain (u,v)components first and then use a support vector machine (SVM) to classify vehicle colours and nonvehicle colours. When performing SVM training and classification, i take a block of n X m pixels as a sample. More specifically, each feature vector is defined as $[u_1,v_1,\ldots\ldots.,u_{nXm},v_{nXm}]$. Notice that we do not perform vehicle colour classification via SVM for blocks that do not contain any local features. Those blocks are taken as non vehicle colour areas. As mentioned in Section I, the features are extracted in a neighbourhood region of each pixel in our framework.

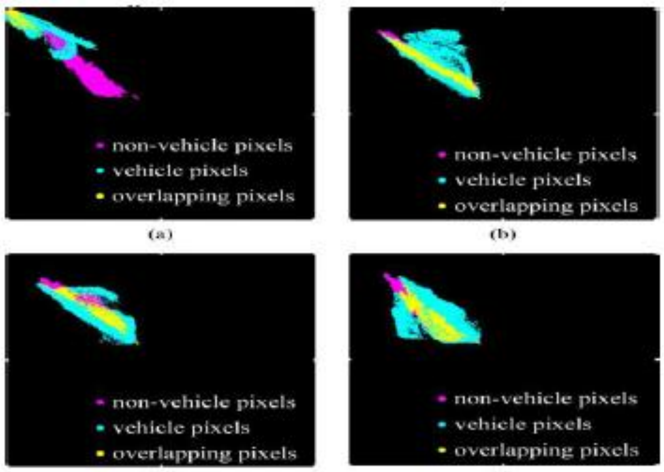

Fig-3 :Vehicle Colours and Nonvehicle Colours in Different Colour Spaces. (A) U-V, (B) R-G , (C) G-B , (D) B-R Planes

Considering N an neighbourhood of pixel *P*, as shown in Fig. 4, we extract five types of features, i.e., S, C ,E ,A , and Z, for the pixel. These features serve as the observations to infer the unknown state of a DBN, which will be elaborated in the next subsection. The first feature S denotes the percentage of pixels in that are classified as vehicle colours by SVM, as defined in (5). Note that denotes to the number of pixels in that are classified as vehicle colours by SVM, i.e.,

$$S = N_{vehiclecolor} / N^2 \quad (5)$$

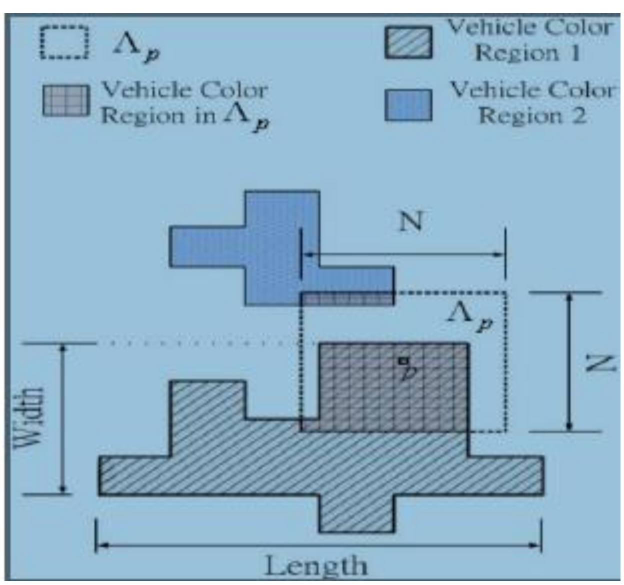

Fig-4:Neighbourhood Region for Feature Extraction

Features *C* and *E* are defined, respectively, as

$$C = N_{Corner} / N^2 \quad (6)$$
$$E = N_{Edge} / N^2 \quad (7)$$

Similarly, N_{Corner} denotes to the number of pixels in A_p that are detected as corners by the Harris corner detector, and N_{Edge} denotes the number of pixels in A_p that are detected as edges by the enhanced Canny edge detector. The pixels that are classified as vehicle colours are labelled as connected vehicle-colour regions. The last two features *A* and *Z* are defined as the aspect ratio and the size of the connected vehicle-colour region where the pixel resides, as illustrated in Fig. 4. More specifically, A = Length / Width, and feature *Z* is the pixel count of "vehicle colour region 1" in Fig. 4.

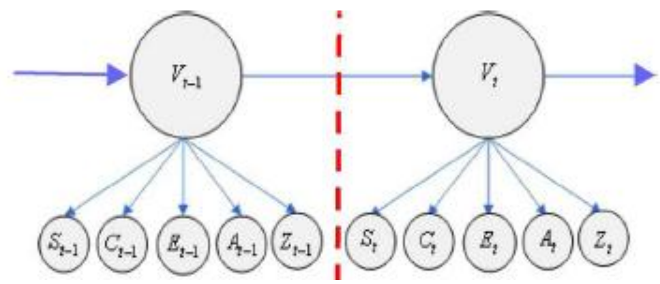

Fig- 5: DBN Model for Pixel Wise Classification

DBN : We perform pixel wise classification for vehicle detection using DBNs [18]. The design of the DBN model is illustrated in Fig. 5. Node indicates if a pixel belongs to a vehicle at time slice *t* . The state of V_i is dependent on the state of V_{t-1}. Moreover, at each time slice *t*, stateV_t has influences on the observation nodes S_t,C_t E_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. Discrete observation symbols are used in our system. We use*K*-means to cluster each observation into three clusters, i.e., we use three discrete symbols for each observation node. In the training stage, we obtain the conditional probability tables of the DBN model via expectation-maximization [18] 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 that a pixel belongs to a vehicle, i.e.,

$$P(V_t|S_t, C_t, A_t, Z_t, V_{t-1}) = P(V_t|S_t)P(V_t|C_t) \times P(V_t|E_t)P(V_t|A_t)P(V_t|Z_t)P(V_t|V_{t-1})p(V_{t-1}). \quad (8)$$

The joint probability $P(V_t|S_{t,}\ C_t\ E_t,\ A_t,\ Z_{t,}\ V_{t-1})$ is the probability that a pixel belongs to a vehicle pixel at time slice given all the observations and the state of the previous time instance. According to the naive Bayesian rule of conditional probability, the desired joint probability can be factorized since all the observations are assumed to be independent. Term P $(V_t|S_t)$ is defined as the probability that a pixel belongs to a vehicle pixel at time slice given observation at time instance *t* [*S* 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 vehicle detection framework can also utilize a Bayesian network (BN) to classify a pixel as a vehicle or non vehicle pixel. When performing vehicle detection using BN, the structure of the BN is set as one time slice of the DBN modelling Fig. 5.

4. 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. The experimental videos are displayed in Fig6. Note that it is infeasible to assume prior information of camera heights and target object sizes for this challenging data set. When performing background colour removal, we quantize the colour histogram bins as 16 16.Colours corresponding to the first eight highest bins are regarded as background colours and removed from the scene. Fig 6(a) displays an original image frame, and Fig 6(b) displays the corresponding image after background removal.

Fig- 6: Snapshots of the Experimental Videos
(a) Original image (b) image after removing background

5. 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. Instead of region based classification, we have proposed a pixel wise classification method for the vehicle detection using DBNs. In pixel-wise classification, relations of 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 change due to the influence of the camera angles and heights, we use only a small number of positive and negative samples. Which are trained by the SVM for vehicle colour 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 canny edge detector, which increases the adaptability and the accuracy for detection in various aerial images.zsThe experimental results demonstrate flexibility and good generalization abilities of the proposed method. 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.

ACKNOWLEDGEMENT

We would like to express our sincere thanks to Sri. Dr. Kancharla Ramaiah Secretary and Correspondent Prkasam Engineering College, Kandukur, A.P. India for his support with providing research environment. We are extremely thankful to our colloquies, friends and family members who are cooperated in this work.

REFERENCES

[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, andM. Ruhe, ―LUMOS—Airbome traffic monitoring system,‖ in Proc.IEEEIntell. 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, ―Airborn 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 vide 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 usin 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.