ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

A Real Time Video Object Tracking Using SVM

Asha G.S[1], K.Arun Kumar[2], D. David Neels Pon Kumar[3]
1 PG scholar, Department of ECE, Einstein College of Engineering, TN, India
2Assistant Professor, Department of ECE, Einstein College of Engg., TN, India
3 Professor, Department of ECE, Einstein College of Engineering, TN, India

Abstract- This paper presents the vehicle detection in aerial surveillance using Dynamic Bayesian Networks. A Dynamic Bayesian Network (DBN) is a Bayesian Network which relates variables to each other over adjacent time steps. This is often called a Two-Time slice BN because it says that at any point in time T, the value of a variable can be calculated from the internal regressors and the immediate prior value (time T-1). DBNs are common in video tracking, and have shown potential for a wide range of digital image applications. It is a method to estimate the real value of an observed variable that evolves in time. We present a pixel wise classification is performed, relation among neighbouring pixels preserved as feature extraction process. The SVM classifier is used to train the image samples for the classification. For edge detection canny edge detector used. In recent years, the analysis of aerial videos taken from aerial vehicle has become an important issue. These technologies have a variety of applications, such as military, police, and traffic management. DBN is used for the node classification. 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.

Index Terms—**Aerial Surveillance, Dynamic Bayesian Networks (DBNS), Vehicle Detection.**

I. INTRODUCTION

In recent years, the analysis of aerial videos taken from aerial vehicle has become an important issue. The Moving Object Detection module is made up of three sub modules, as illustrated in Fig. 1. The background sub-traction model adopted by was adopted since thesis one of the most robust real time background subtraction models found in literature. This background subtraction model first applies a pixel wise median filter over time (20-40 seconds) to discriminate between moving and stationary pixels to construct the initial background model. Therefore, aerial surveillance systems be-come an excellent supplements of ground-plane surveillance systems. 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.

Hinz and Baumgartner [7] 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 neighbouring objects are present. Cheng and Butler [8] considered multiple clues and used a mixture of experts to merge the clues for vehicle detection in aerial images. They performed colour 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, the motion analysis algorithm applied in their system cannot deal with aforementioned camera motions and complex background changes. Proposed a method by subtracting background colours of each frame and then refined vehicle candidate regions by enforcing size constraints of vehicles. The moving tracking algorithm adopted in this work employs a correspondence based strategy where every previous moving object is compared to all current moving objects to and its match if it is still there in the scene. The tracking algorithm also ensures that the region occupied by the moving in the current frame must overlap the location of its corresponding object in the previous frame. If the overlapping area is larger than a specified threshold T, the overlapping objects are transformed in the HSV colour space. All the pixels in the object area in the current and previous frame are associated to one of the 21 bins, as specified in [3].The resulting histograms are compared using the insertion distance matrix. 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.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

Choi and Yang proposed a vehicle detection algorithm using the symmetric property of car shapes. However, this cue is prone to false detections such as symmetrical details of buildings or road markings. Therefore, they applied a log-polar histogram shape descriptor to verify the shape of the candidates. Unfortunately, the shape descriptor is obtained from a fixed vehicle model, making the algorithm in flexible. Moreover, the algorithm in relied on mean-shift clustering algorithm for image colour segmentation. The major drawback is that a vehicle tends to be separated as many regions since car roofs and windshields usually have different colours. Moreover, nearby vehicles might be clustered as one region if they have similar colours. The high computational complexity of mean-shift segmentation algorithm is another concern. In this paper, we design a new vehicle detection framework that preserves the advantages of the existing works and avoids their drawbacks. The modules of the proposed system frame-work.. In the detection phase, first perform background colour removal similar to the process proposed. 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, there is no need to generate multiscale sliding windows either. This paper has presented the implementation of a real-time autonomous video surveillance system. 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 neighbourhood region of each pixel. Therefore, the extracted features comprise not only pixel-level information but also relationship among neighbouring pixels in a region. Such design is more effective and efficient than region-based or multiscale sliding window detection methods. 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. Cheng and Butler considered multiple clues and used a mixture of experts to merge the clues for vehicle detection in aerial images. They performed colour 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.Lin et al. proposed a method by subtracting background colours of each frame and then refined vehicle candidate regions by enforcing size constraints of vehicles. 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.

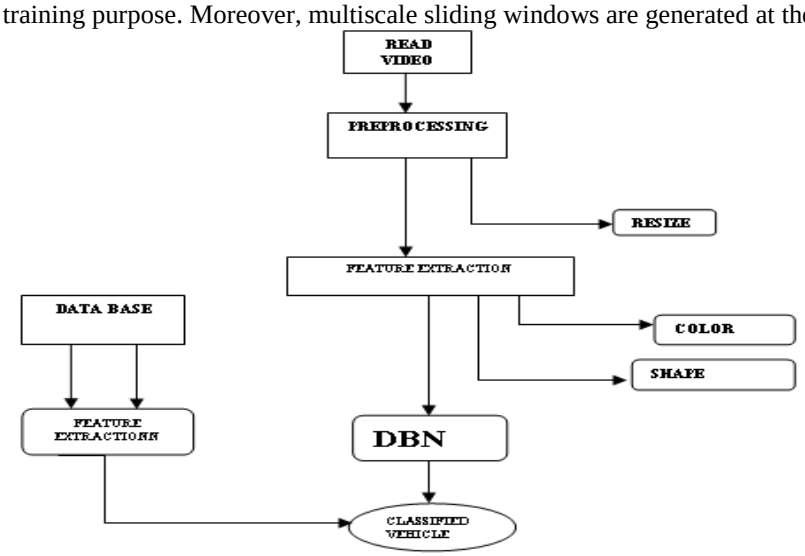


Fig: 1 Proposed System Framework

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. Unfortunately, the shape descriptor is obtained from a fixed vehicle model, making the algorithm in flexible. Moreover, similar relied on mean-shift clustering algorithm for image colour segmentation. The system would miss vehicles when the contrast is weak or when the influences of neighbouring objects are present. Segmentation algorithm is another concern. In this paper, we design a new vehicle detection framework that preserves the advantages of the existing works and avoids their drawbacks. Modules of the proposed system framework are illustrated in Fig. 1. The framework can be divided into the training phase and the detection 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

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

this paper, we do not perform region-based classification, which would highly depend on results of colour 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 pixel wise classification. However, the features are extracted in a neighbourhood region of each pixel. Therefore, the extracted features comprise not only pixel-level information but also relationship among neighbouring pixels in a region. Such design is more. The major drawback is that a vehicle tends to be separated as many regions since car roofs and windshields usually have different colours. Moreover, nearby vehicles might be clustered as one region if they have similar colours. The high computational complexity of mean-shift is that a vehicle tends to be separated as many regions since car roofs and windshields usually have different colours. Moreover, nearby vehicles might be clustered as one region if they have similar colours.

II. PROPOSED VEHICLE DETECTION FRAMEWORK

We present an automatic vehicle detection system for aerial surveillance in this project. 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.

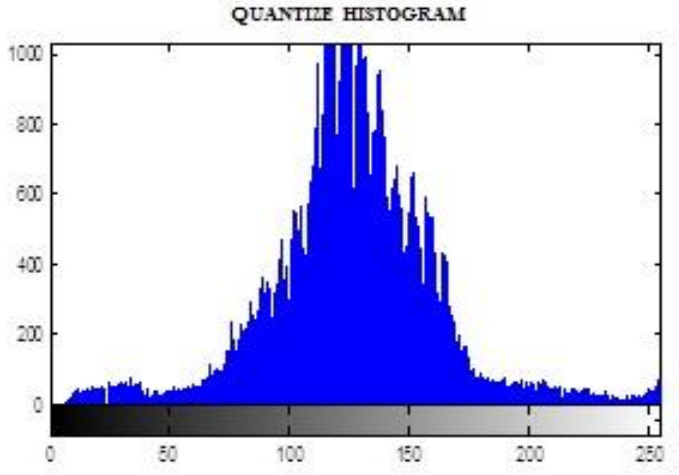


Fig. 2.Colour Histogram of a Frame

A. Background Colour Removal

Since no vehicle regions cover most parts of the entire scene in aerial images, we construct the colour histogram of each frame and remove the colours that appear most frequently in the scene. Take Fig. 2 for example, the colours are quantized into 48 histogram bins. Among all histogram bins, the 12th, 21st, and 6thbins are the highest and are thus regarded as background colours and removed. These removed pixels do not need to be considered in subsequent detection processes. Performing background colour removal cannot only reduce false alarms but also speed up the detection process.

B.Feature Extraction

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

1. Local Feature Analysis

Corners and edges are usually located in pixels with more information. We use the Harris corner detector [13] to detect corners. To detect edges, we apply moment-preserving thresholding method on the classical canny edge detector [14] 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 the higher thresholdT_{high}.As the illumination in every aerial image differs, the desired thresholds vary and adaptive thresholds are required. The computation of Tsai's moment-preserving method [15] is deterministic without iterations for level thresholding with L< 5. Its derivation of thresholds is described as follows. Let be an image with pixels and *(x,y)*denote the Gray value at pixel (x,y) . Thei_{th} moment of is defined asm_i of f is defined as

$$m_i=(\tfrac{1}{n})\sum_j n_j\left(z_j\right)^i=\sum_j p_j\left(z_j\right)^i \text{ ,i=1,2,3,.(1)}$$

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

Where n_i is the total number of pixels in image with Gray valueZ_j and$P_j = n_i/n$. For bit level thresholding, we would like to select threshold *T*such that the first three moments of image Fare preserved in the resulting bit level image. Let all the below-threshold Gray values in *f*be replaced byZ_0 and all the above-threshold grey values be replaced by Z_1; we can solvep_0 andp_1 for and based on the moment preserving principle [15]. After obtaining p_0 andp_1, the desired threshold is computed

Using

po=(1/n)$\sum_1^T nj$(2)

In order to detect edges, we use the gradient magnitude*(x,y)*of each pixel to replace the greyscale value *F(x,y)*inTsai's method. Then, the adaptive threshold found by (2) is used

As the higher thresholdT_{high} in the Canny edge detector. We set the lower threshold as,

$T_{low} = 0.1 \times (G_{max}) - (G_{min})$

Where$G_{max}$and G_{min}represent the maximum and minimum gradient magnitudes in the image. Thresholds automatically and dynamically selected by our method give better performance on the edge detection. We will demonstrate the performance improvement on the edge detection with adaptive thresholds and the corresponding impact on final vehicle detection results in Section 3

2. Colour Transform and Colour Classification:

In [16], the authors proposed a new colour model to separate vehicle colours from nonvehicle colours effectively. This colour model transforms*(R, G, B*) colour components into the colour domain *(u,v),*i.e.,

$u_p = \frac{2B_p - G_p - R_p}{Z_p}$(3)

$u_p = Max\left\{\frac{B_p - G_p}{Z_p}, -\frac{R_p - B_p}{Z_p}\right\}$(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$\times$ $m$pixels as a sample. More specifically, each feature vector is defined as$[u_1, v_1, \ldots u_{n\times m}, v_{n\times m}]$.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. Considering N$\times$ Nan neighbourhoodΛp, of pixel*P*, as shown in Fig. 4, we extract five types of features, i.e., 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 Λpthat are classified as vehicle colours by SVM, as defined in (5). Note that $N_{vehicle\ color}$ denotes to the number of pixels inΛpthat are classified as vehicle colours by SVM, i.e.,

$S = \frac{N_{vehicle color}}{N^2}$ (5)

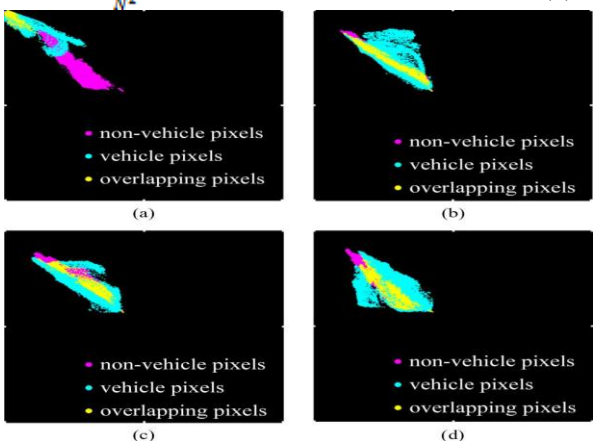


Fig. 3.Vehicle Colours and Nonvehicle Colours in Different Colour Spaces. (A) ,U-V,(B) R-G ,(C) G-B ,(D) B-R Planes

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

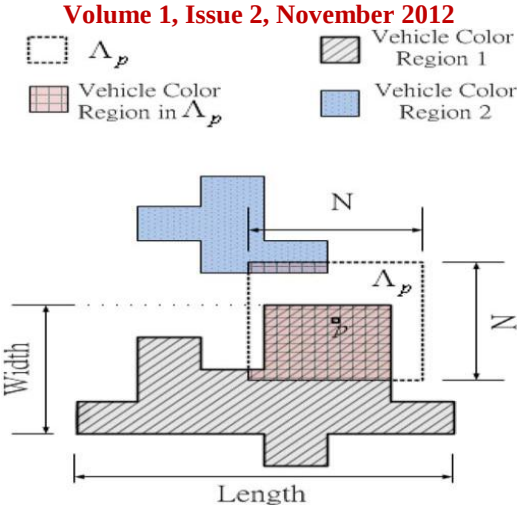


Fig. 4.Neighbourhood Region for Feature Extraction

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

$$C=\frac{N_{corner}}{N^2} \quad (6)$$

$$E = \frac{N_{Edge}}{N^2} \quad (7)$$

Similarly,N_{corner} denotes to the number of pixels inΛp that are detected as corners by the Harris corner detector, andN_{Edge}denotes the number of pixels inΛp 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.

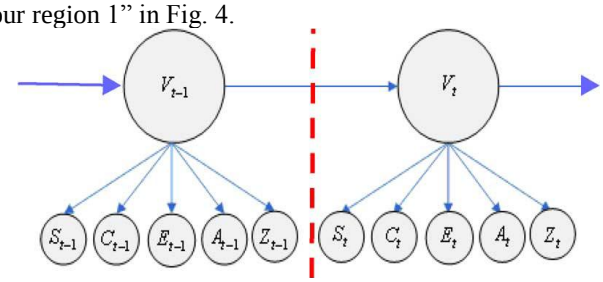


Fig. 5. DBN Model for Pixel Wise Classification

C. DBN

We perform pixel wise classification for vehicle detection using DBNs [18]. The design of the DBN model is illustrated in Fig. 5. Node V_t indicates if a pixel belongs to a vehicle at time slice *t* . The state ofV_2 is dependent on the state ofV_{t-1} .Moreover, at each time slice *t*, state has influences on the observation nodesS_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.,

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

$P(V_t \mid, E_t,) P(V_t \mid A_t) \, P(V_t \mid Z_t) \, P(V_t \mid V_{t-1}) \, P(V_{t-1})$ (8)

The joint probability ($V_t \mid S_t, C_t, E_t, A_t, Z_t, V_{t-1}$)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 ($V_t \mid S_t$) is defined as the probability that a pixel belongs to a vehicle pixel at time

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

slice given observation $S_t$at time instance *t*[*S*is defined in (5)]. Terms(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. We will compare the detection results using BN and DBN in the next section.

III. 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 $16\times 16 \times$ 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

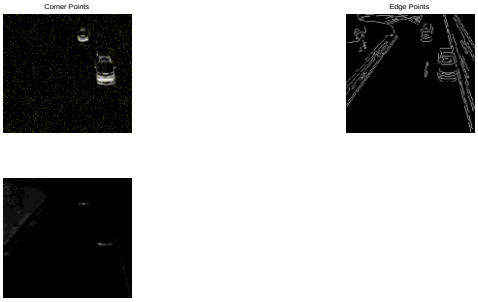
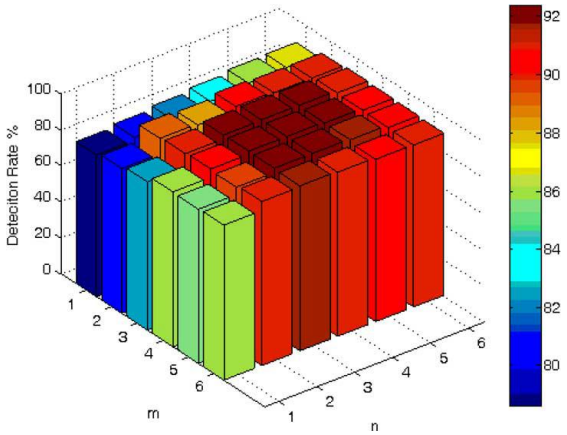


Fig 7: Background Colour Removal Results



.Fig 8. Influence Of Different Block Sizes On The Detection Rate.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

Fig 9: Results of Vehicle Colour Classification

When employing SVM, we need to select the block size to form a sample (see Fig.7)<<AQ1>>. Fig. 8 displays the influence of different block sizes on the detection rate. According to the observation from Fig. 9, we take each 3 4 block to form a feature vector. The colour of each pixel would be transformed to u and v colour components using (3) and (4). More specifically, each feature vector has a dimension of 24. The training images used to train the SVM are displayed in Fig 7. Notice that the blocks that do not contain any local features are taken as nonvehicle areas without the need of performing classification via SVM. Fig 9 shows the results of colour classification by SVM after background colour removal and local feature analysis. To obtain the conditional probability tables of the DBN, we select the training clips from the first six experimental videos displayed in Fig. 5. The remaining four videos are not involved in the training process. Each training clips contains 30 frames, whose ground-truth vehicle positions are manually marked. That is, we use only 180 frames to train the DBN. To select the size of the neighbourhood area for feature extraction,

Table.1 Detection Accuracy Using Different Neighbourhood Sizes

Size of Λp	Hit rate	Number of positives per frame
5×5	70.91	0.523
7×7	92.31	0.278
9×9	87.06	0.281
11×11	82.35	0.401
13×13	75.58	0.415

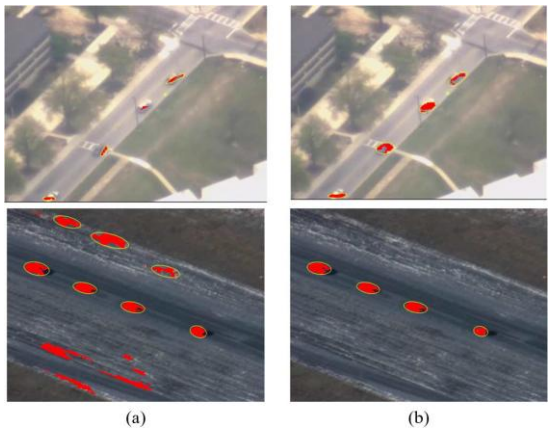


Fig10: Impact of the Enhanced Edge Detector on Vehicle Detection Results

We list the detection accuracy using different neighbourhood sizes in Table I. The detection accuracy is measured by the hit rate and the number of false positives per frame. There are a total of 225 025 frames in the data set. When calculating the detection accuracy, we perform evaluation every 100 frames. We can observe that the neighbourhood area with the size of 7× 7 yields the best detection accuracy. Therefore, for the rest of the

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

experiments, the size of the neighbourhood area Λp for extracting observations is set as 7×7.The impacts of the enhanced Canny edge detector on vehicle detection results can be observed in Fig10. Fig 10(a) shows the results obtained using the traditional canny edge detector with non adaptive thresholds. Fig 10(b) shows the detection results obtained using the enhanced canny edge detector with moment preserving threshold selection. No adaptive thresholds cannot adjust to different scenes and would therefore result in more misses or false alarms, as displayed in Fig 10(a). To examine then necessity of the background removal process and the enhanced edge detector, we list the detection accuracy of four different scenarios in Table II. 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. We compare different vehicle detection methods in Fig. 11.The moving-vehicle detection with road detection method in [9] requires setting a lot of parameters to enforce the size constraints. in order to reduce false alarms. However, for the experimental data set, it is very difficult to select one set of parameters that suits all videos. Tracking algorithm is used to track every human present in the scene throughout the video sequence.

Table.2 Detection Accuracy of Four Different Scenarios

Scenarios	Hit rate	Number of false positives per frame
Without background removal and without enhanced edge detector	75.08	0.399
Without background removal and with enhanced edge detector	92.35	0.459
With background removal and without enhanced edge detector	74.96	0.297
With background removal and with enhanced edge detector	92.31	0.278

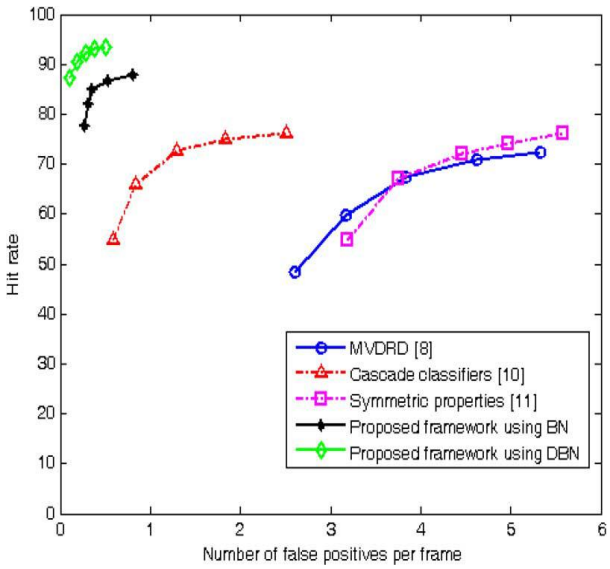


Fig. 11. Comparisons of Different Vehicle Detection Methods

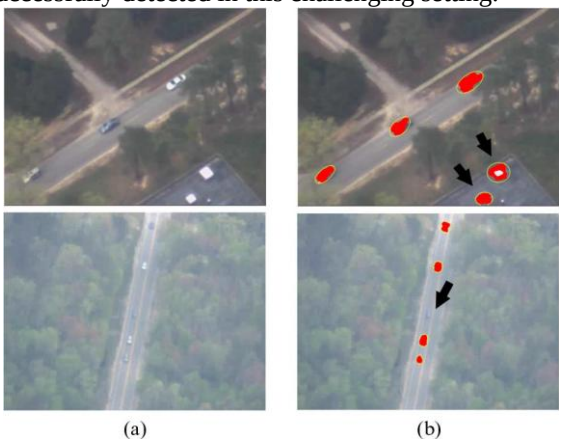
Setting the parameters heuristically for the data set would result in low hit rate and high false positive numbers. The cascade classifiers used in [11] need to be trained by a large number of positive and negative training samples. The number of training samples required in [11] is much larger than the training samples used to train the SVM classifier in Fig 7.The colours of the vehicles would not dramatically change due to the influence of the camera angles and heights. Moreover, if the aspect ratio of the multiscaled detection windows is fixed, large

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

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.

Fig12: Vehicle Detection Results
A)Input Image,B) Classified Image, C) Binary Image D)Post Processing Image

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 DBNis better than that of the BN. In Fig. 12, we display the detection results using BN and DBN [see Fig. 12(a) and (b), respectively].The coloured pixels are the ones that are classified as vehicle pixels by BN or DBN. The ellipses are the final vehicle detection results after performing post processing. DBN outperforms 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 P ($V_t \mid V_{t-1}$) by helps stabilize the detection results in the DBN.Fig:13 shows some detection error cases. Fig. 13(a) display the original image frames, and Fig. 13(b) displays the detection results. The black arrows in Fig 13(b) indicate the misdetection or false positive cases. In the first row of Fig 13, 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 Fig13, 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.



(a) (b)

Fig. 13.Detection Error Cases.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

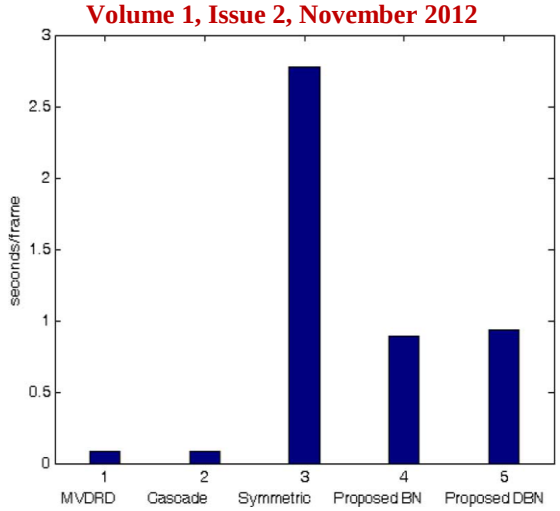


Fig. 14. Processing Speeds of Different Methods

Fig. 14 shows the average processing speeds of different vehicle detection methods. The experiments are conducted on a personal computer with 2.66-GB central processing unit and4-G random access memory. The frame size of the experimental videos is 640 480 pixels. Although the proposed framework using BN and DBN cannot reach the frame rate of the surveillance videos, it is sufficient to perform vehicle detection every50–100 frames. Therefore, high detection rate and low false alarm rate should be the primary considerations of designing detection methods given the condition that the execution time is reasonable.

IV. 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.

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.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

[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.

AUTHOR BIOGRAPHY

Asha. G.S born in 1988 completed B.E in Electronics and Communication from, Einstein College of Engineering, Tirunelveli through Anna University, and Chennai in 2010. Now she is pursuing M.E Applied Electronics from the same college

K. Arun kumar born in 1973 completed his B.E. in ECE from the Govt. College of Engineering, Tirunelveli and M.E. (Optical Commn.) from ACCET Karaikudi. He has more than 10 years of teaching experience and 7 years of Industrial experience. Presently he is working as Associate professor in ECE department, at Einstein College of Engg., India

D. David Neels Pon Kumar was born in India, in 1971. He completed his B.E degree in ECE in 1992 and M.E degree in Digital Communication and Networking through Anna University Chennai in 2004. He is pursuing PhD in Wireless Networks through Anna University Chennai since 2007.He has 10 years of Industrial experience in India and abroad apart from 9 years of teaching experience at various cadres. Presently he is working as Associate professor in ECE department, at Einstein College of Engg., India. He has 8 publications in International journals and presented 10 papers in International and National conferences. He is a member of ISTE, IEEE and IAENG. He is a reviewer in IET Networks