Automatic Vehicle Detection and Tracking in Aerial Surveillances using SVM

Divya Michael M G University FISAT, Kerala Paul P Mathai Assistant Professor,M G University,FISAT, Kerala AbhidhaT E M G University FISAT, Kerala

ABSTRACT

Target object detection in aerial surveillance using image processing techniques is growing more and more important. Aerial surveillance is more suitable for monitoring fast moving targets and covers a much larger spatial area. These technologies have a variety of applications, such as traffic management, police and military. Aerial view has the advantage of providing a better perspective of the area being covered and this make use of the aerial videos taken from aerial vehicles. In an automatic vehicle detection system for aerial surveillance background colors are eliminated and then features are extracted. This system extracts features including color feature and local feature points. For vehicle color extraction, system utilizes color transform to separate vehicle colors and non-vehicle colors effectively. For edges detection, system applies moment-preserving method to adjust the thresholds for canny edge detector automatically, which improves the adaptability and accuracy of the system. A support Vector Machine is constructed for classification purpose.

Keywords

Aerial Surveillances, Training, Detection, Classification, SVM, GMM.

1. INTRODUCTION

In recent years, the analysis of aerial videos taken from aerial vehicle has become an important issue. The increase in the number of vehicles on roadway networks based transport management agencies to allow use of technology advances resulting in better decisions. Considering all these perspective aerial surveillance has better place now a days. Aerial surveillance is more suitable for providing increased monitoring results in case of fast-moving targets and covers a much larger spatial area. So these aerial surveillance systems become excellent supplements of ground-plane surveillance systems. 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 tilting, panning and rotation. Also the different camera heights largely affect the detection results. Vision based techniques is one of the most common approach to analyze vehicles from images or videos.

The Gaussian mixture algorithm used for background subtraction developed by Stauffer and Grimson [2] aims to segment moving foreground objects from relatively stationary objects. Recently pixel-based probabilistic model methods gained lots of interests and have shown good detection results. The values of a particular pixel are modeled as a mixture of adaptive Gaussians. At each iteration, Gaussians are evaluated using a simple heuristic to determine which ones are mostly likely to corresponding to the background pixels. Those pixels that do not match with the background Gaussians are classified as foreground. During the training phase the system extracts features including local feature points and color feature points. In the detection phase, first perform background color removal. Afterward, the feature extraction procedure is performed as in the training phase. These features act as the evidence to infer the unknown state of the trained SVM, which indicates whether a pixel belongs to a vehicle or not. There is no need to generate multi-scale sliding windows either. The detection task is based on pixel wise classification and the features are extracted in a neighborhood region of each pixel. Therefore, the extracted features consist of pixel-level information and also relationship among neighboring pixels in a region. Such design is more efficient and effective than existing techniques like region-based or multi-scale sliding window detection methods. A Support Vector machine (SVM) is constructed for classification purpose and based on the features extracted, a well-trained SVM can estimate the probability of a pixel belonging to a vehicle or not.

1.1 Problem Statement

The advancement of computer technology and increasing needs of social security, studies of object detection in aerial surveillance using image processing techniques are gaining more interest. So design a new vehicle detection framework that preserves the advantages of the existing works and avoids their drawbacks has more importance. The tremendous growth in the number of vehicles on the roadway network has forced the transport management agencies to focus on advanced technologies to take better decisions and this make aerial surveillance has better place now a days. Aerial surveillance provides increased monitoring results in case of fast-moving targets because spatial area coverage is greater. Thus aerial surveillance is added for ground-plane surveillance systems. One of the main topics in intelligent aerial surveillance is vehicle detection. The problems involved in the aerial surveillance include the camera motions such as tilting, panning and rotation. Also the different camera heights largely affect the detection results. Vision based techniques is one of the most common approach to analyze vehicles from images or videos.

1.2 Existing System

The existing system uses a design which preserves the merits of the existing vehicle detection framework and avoids their drawbacks. The vehicle detection framework can be classified into the training phase and the detection phase.

During training phase, the system extracts local features as well as vehicle colors to train a dynamic Bayesian network (DBN). In the detection phase, the system first performs background color removal. Then the same feature extraction procedure is performed as in the training phase. These extracted features serve as the evidence to infer the unknown state of the trained DBN. The vehicle detection task is based on pixel wise classification. The features are extracted in a neighborhood region of each pixel and extracted features comprise not only pixel-level information but also relationship among neighboring pixels in a region. So the proposed design is efficient and effective than existing regionbased or multi-scale sliding window detection methods. The existing system uses a color histogram for background removal. The system constructs the color histogram of each frame and removes the colors that appear most frequently in the scene. These are considered as background, which need not to be considered in subsequent detection processes. The background subtraction process speed up the detection process. Since non vehicle regions cover most parts of the entire scene in aerial images, construct the color histogram of each frame and remove the colors that appear most frequently in the scene. The colors are quantized into 48 histogram bins from these histogram bins, the 12th, 21st, and 6th bins are the highest and are thus regarded as background colors and removed. These eliminated pixels do not need to be considered in subsequent detection processes.

1.3 Objective

Objective of the thesis are as under:

•To design a vehicle detection frame work.

•Features are extracted in training phase.

•To design a detection phase and perform background removal.

• Pixel wise classification for detection.

• Support Vector machine to classify vehicle pixel and non-vehicle pixel

•Finally track the detected vehicle.

1.4 Organization of Work

The remaining of the report is organized as follows. Section 2 briefly describes the Literature Survey. Section 3 introduces the methodology adopted used in proposed system. In Section 4, describes the Implementation. Results and Discussions are described in Section 5. Conclusion of the project is discussed in Section 6.

2. LITERATURE SURVEY

S. Srinivasan, H. Latchman, J. Shea, T. Wong, and J. McNair [4] have introduced a method for the airborne video registration and traffic-flow parameter estimation which focus on airborne helicopter video for estimating traffic parameters. The airborne video is taken from a digital video camera attached to the skid of the helicopter. Roll, pitch, and yaw of the helicopter make the video difficult to view, unstable and the derived parameters less accurate. To avoid this, a frameby-frame video-registration technique using a feature tracker to automatically determine control-point correspondences is present in the system. This feature tracker is used to track fixed features through the sequence of images in the video. These feature-location correspondences are used as control points to compute a polynomial transformation function to warp every frame in the sequence successively to the reference. The reference frame can be any one of the frames in the video segment. By changing the tracking-window orientation, the same feature tracker can be used to track moving vehicles within the registered video segment. The advantages of the system include excessive reliance on sensor data, inadequate stabilization and availability of elevation maps.

S. Hinz and A. Baumgartner [5] utilized a hierarchical model that describes different levels of details of vehicle features. In this system, vehicles are represented in different levels and details of vehicle are increasing from left to right. Hierarchical model describes the appearance of vehicle at different levels. At lowest level vehicle substructures are described like type of vehicle and last level represents 3D information and local context of vehicle. The advantage of this system is that it neither relies on external information like digital map or site

models, nor it is limited to any specific vehicle models. As disadvantage, the system would miss vehicles when the contrast is weak or when the influences of neighboring objects are present.

H. Cheng and J.Wus [6] have introduced an adaptive ROI estimation algorithm by analyzing the way the camera is operated. The system proposes an ego-motion of the camera computed from adjacent video frames, and deriving the ROI defined by the operator while capturing the video using an adaptive operator attention model. When monitoring a building beside a highway, only those moving objects associated with the building, such as objects entering or leaving the building, are of interest. Therefore, accurate ROI estimation for aerial surveillance applications must be adaptive and cannot always use the same type of objects, such as moving objects as ROI.

Lin et al [7] proposed a method by subtracting background colors of each frame and then refined vehicle candidate regions by enforcing size constraints of vehicles. The disadvantage of the system is that, it assumed too many parameters such as the smallest and largest sizes of vehicles, height and the focus of the airborne camera etc. Assuming these parameters as known priors might not be realistic in real applications. Also the vehicle detection method is based on cascade classifiers.

Cheng and Butler [8] considered multiple clues and used a mixture of experts to merge the clues for vehicle detection in aerial images. The system consists of color segmentation via mean-shift algorithm and motion analysis via change detection. Also they presented a trainable sequential maximum a posterior method for multi-scale analysis and enforcement of contextual information. The system contains a method by subtracting background colors of each frame and then refined vehicle candidate regions by enforcing size constraints of vehicles. A large number of positive and negative training samples need to be collected for the training purpose. Also considers multi-scale sliding windows which are generated at the detection stage. The main demerit of this method is that there are a lot of miss detections on rotated vehicles. The faces with poses are easily missed if only frontal faces are trained. Also if faces with poses are added as positive samples, then the number of false alarms would surge.

J. Y. Choi and Y. K. Yang [9] have proposed a vehicle detection algorithm using the symmetric property of car shapes. But this is prone to false detections such as symmetrical details of buildings or road markings. In order to avoid this, a log-polar histogram shape descriptor is used to verify the shape of the candidates. The shape descriptor is obtained from a fixed vehicle model, making the algorithm inflexible. Also the algorithm relied on mean-shift clustering algorithm for image color segmentation. Moreover, nearby vehicles might be clustered as one region if they have similar colors. The advantage is that, using symmetric property of car shapes, object can easily detect. Disadvantages include a vehicle tends to be separated as many regions since car roofs and windshields usually have different colors. Also the high computational complexity of mean-shift segmentation algorithm is another concern.

Luo-weitsai, Jun-weihsieh, Kuo-chin fan [10] proposed a novel vehicle detection method using color transform model. The detection procedure consists of different stages. In the first stage a color transformation model is used to separate vehicle colors from non-vehicle colors effectively. This color model transforms (R,G,B) color components into the color domain (u,v). The technique which is adopted for this is dimensionality reduction. Beginning of the technique follows the collection of several thousands of training images and these training images are collected by projecting all colors of input pixels on the color space. The classification is carried out by means of a Bayesian classifier. By using this method the authors insist that vehicles can be very robustly and accurately verified and detected from static images.

B. Morris and M. Trivedi [11] proposed a tracking system with ability to classify vehicles into three classes such as sedan, semi, truck+SUV+van. This system was developed after comparing classification schemes using vehicle images and measurements. The most accurate among this learned classifier was integrated into tracking software which greatly improved the accuracy on low resolution traffic video.

J. W. Hsieh, S. H. Yu, Y. S. Chen, and W. F. Hu [12] proposed a vehicle tracking and classification technique to estimate important traffic parameters from video sequences using one camera. The system has the ability to categorize vehicles into more specific classes by introducing a new linearity feature in vehicle representation. Also this system can easily tackle the problem of vehicle occlusion caused by shadows. This problem is solved by shadow algorithm that uses a set of lines to eliminate all unwanted shadows.

Background subtraction is a process which aims to segment moving foreground objects from a relatively stationary background. Recently pixel-based probabilistic model methods gained lots of interests and have shown good detection results. Background modeling by Gaussian mixtures is a pixel based process. Let x be a random process representing the value of a given pixel in time. A convenient framework to model the probability density function of \mathbf{x} is the parametric Gaussian mixture model where the density is composed of a sum of Gaussians. Let $p(\mathbf{x})$ denotes the probability density function of a Gaussian mixture comprising K component densities. The mixture of Gaussians algorithm, proposed by Stauffer and Grimson [1], estimates these parameters over time to obtain a robust representation of the background. Here every pixel is checked against K Gaussian distributions until a match is found and if no match is found for a pixel, the least probable distribution is replaced with a distribution with current value as its mean value. The advantage includes its good accuracy and disadvantages include high computational cost and high execution time.

Zivkovic [13] proposed an approach which reduces the computation time and memory bandwidth. Here, a Bayesian approach is formulated to select the required number of Gaussian modes for each pixel in the scene. In scenes with static background, this approach assigns a single mode Gaussian to model most of the pixels which helps to reduce average processing time by 32%. However in the outdoor video (trees sequence), results show only a 2% improvement since a significantly large portion of the scene requires a multimodal model. In the GMM algorithm the weights of the Gaussian mixture represent the fraction of the data samples (x(t)) that belongs to the particular mode in the model. This system used a Dirichlet prior with negative coefficients. This is done with an intention of accepting a class only if there is enough evidence from the data samples for the existence of the class. Advantages include fast compared to GMM, efficient and accurate. Disadvantage includes reduced computation time.

D.S. Lee [14] proposed an effective gaussian mixture learning for video background subtraction, each pixel in a frame is modeled as a stochastic process and the modes of the GMM are arranged in decreasing order of their weights. A predefined fraction of the weights proposed to use modes with low weights to model the foreground and higher weight to be selected as the background. A match of an incoming pixel to any of the modes defines the pixel as background or foreground. Advantages are fast compared to GMM and makes fast learning. Disadvantages include high memory requirement.

3. PROPOSED METHODOLOGY

The proposed system presents a new vehicle detection framework that preserves the advantages of the existing works and avoids their drawbacks. The vehicle detection frame work can be classified into training phase and the detection phase. In the training phase, extract multiple features including edge points and corner features. In the detection phase, perform background color removal using Gaussian Mixture Model.

The gaussian mixture model algorithm for background subtraction combines fast learning of effective GMM with reduced selection of number of modes in improved GMM to obtain an efficient and accurate scheme. Then features are extracted in the training phase. The proposed system consist of a new color model to separate vehicle colors from nonvehicle colors effectively and this color model transforms (R,G,B) color components into the color domain (u,v). The support vector machine (SVM) is used to classify vehicle colors and non-vehicle colors effectively. Finally a support vector machine is used for classification purpose. After classification the system can find out whether a pixel belongs to vehicle or not. Finally these classified vehicles are tracked.

3.1 The System Architecture

- The system is split in to three parts
- •Training
- •Detection
- Tracking

Training: - In the training method, extract multiple features including local edge and corner features, as well as vehicle colors to train a Support Vector Machine (SVM).

Detection: - In the detection method, first perform background color removal using Gaussian Mixture Model. Afterward, the same feature extraction procedure is performed as in the training phase. These extracted features are given as input to SVM which represents whether a pixel belongs to a vehicle or not.

Tracking: - Finally these detected vehicles are tracked.

- The system is divided into five modules
- 1. Frame Extraction and Background color removal
- 2. Feature Extraction
- 3. Classification
- 4. Post Processing
- 5. Tracking

4. IMPLEMENTATION

4.1Frame Extraction and Background Removal

In frame extraction, read the input video and extract the number of frames from that video. Video is segmented to frames using the VideoReader class of Matlab.These extracted frames are used as input to background removal. The Table 4.1. Shows the parameters of video.

Table4.1. Video Properties

Input	Size	Leng th	Frame rate	Width* Height	Frame s
Video1	1.2 MB	4	30	320*240	132
Video2	204 KB	8	15	160*120	120
Video3	2.6 MB	3	25	360*168	120
Video4	813 KB	8	15	160*120	80

Background subtraction is a technique in the field of image processing and computer vision where in an image's foreground is extracted for further processing. Generally an image's regions of interest are objects in its foreground. In surveillance applications, background subtraction has become a standard method for video vehicle segmentation. Background subtraction makes object detection much faster. The problem tackled by background subtraction techniques involves the comparison of an observed image with an estimated image that does not contain any object of interest; this is referred to as the background model .This comparison process, called foreground detection, which divides the observed image into two complementary sets of pixels. The foreground that contains the objects of interest and the background is its complementary set.

The proposed system uses adaptive Gaussian mixture model (GMM) [2] for background subtraction. Background modeling by Gaussian mixtures is a pixel based process. Each pixel is temporally modeled as a mixture of two or more The recent history of each pixel, X1, X2 ...,Xt, is modeled by a mi The recent history of each pixel, X1, X2 ...,Xt, is modeled by a mixture of K Gaussian distributions. The probability of observing the current pixel value is [2]

$$P(X_{t}) = \sum_{i=1}^{\kappa} \omega_{i,t} * n(X_{t}, \mu_{i,t,\sum i,t})$$
(1)

4.2 Feature Extraction

Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately. Feature extraction is performed in both the training phase and the detection phase. The system considers local features and color features. Vehicle pixels and nonvehicle pixels are extracted during training phase. These features can be used by the SVM for classification of vehicle. In this module first extract the features from the image frame and perform corner detection, edge detection, color transformation and Color classification.

4.2.1 Local Feature Analysis

Corners and edges are usually located in pixels with more information. The system uses the Harris corner detector [16] to detect corners. Canny edge detector is used to detect edges, which apply moment-preserving thresholding method on the classical Canny edge detector [17] to select thresholds adaptively according to different scenes. The Canny edge detector find out the lower threshold Tlow and the higher threshold Thigh which helps to set background and foreground.

4.2.2Color Transform and Color Classification

The paper [16] uses a new color model to separate vehicle colors from nonvehiclecolors effectively. This color model transforms (R,G,B) color components into the color domain (u,v). [1]

$$Up = \frac{2Zp - Gp - Bp}{Zp} \tag{2}$$

$$Vp = Max \frac{Bp - Gp}{Zp}, \frac{Rp - Bp}{Zp}$$
(3)

where (Rp,Gp,Bp) is the R,G and B color components of pixel p and Zp =(Rp + Gp + Bp)=3. All the vehicle colors are concentrated in a much smaller area on the plane than in other color spaces and are therefore easier to be separated from non-vehicle colors. Although the color transform proposed in [16] did not aim for aerial images, it is found that the separability property still presents in aerial images. Therefore, apply the

color transform to obtain components first and then use a support vector machine (SVM) to classify vehicle colors and non-vehicle colors. When performing SVM training and classification, take n*m block of pixels as a sample. More specifically, each feature vector is defined as $[U_1, V_1, ...,$ Un*m, Vn*m] It is not possible to perform vehicle color classification via SVM for blocks that do not contain any local features. Considering an N*N neighborhood of pixel p. Extract different types of features and save as input to SVM structure. These features include the percentage of pixels that are classified as vehicle color by SVM, the percentage of pixels that are detected as corner points by Harris corner detector, the percentage of pixels that are detected as edge points by enhanced canny edge detector, aspect ratio, pixel count of vehicle color region.

4.3 Classification

The objective of a classification system is to assign a pixel in a frame to a class using feature extraction. There are variety of classifiers available like Neural Network(NN), Support Vector Machine (SVM), K-Nearest Neighbor Classifier(KNN), Decision Trees etc. Classifications of vehicles are implemented using the SVM classifier.

4.4 Support Vector Machine

SVM is considered as a training method for polynomial, radial basis function and multi-layer perceptron classifiers [17] in which the weights of the network are calculated by solving a quadratic problem with linear constraints. SVM works on the basis of Kernel equation and view input data as two sets of vectors in an n-dimensional space. SVM will construct a separating hyper plane in n-dimensional space which maximizes the margin between the two available data sets. After successful training of the SVM classifier, testing can be performed which actually classify the object as either vehicle or non-vehicle.



Length

The features are extracted in a neighborhood region of each pixel in the framework. Considering an N*N neighborhood. Extract five types of features, i.e. S, C, E, A and Z for the pixel. These features are taken as the observations to infer the unknown state of a SVM. The feature S represents the percentage of pixels that are classified as vehicle colors by SVM. Note that Nvehiclecolor denotes to the number of pixels in neighborhood region that are classified as vehicle colors by SVM. Definitions are given below

$$S = \frac{Nvehiclecolor}{N^2}$$
(3)
$$C = \frac{N_{corner}}{N^2}$$
(4)

$$E = \frac{N_{edge}}{N^2}$$
(5)

Similarly, Ncorner denotes to the number of pixels that are detected as corners by the Harris corner detector, and Nedge denotes the number of pixels that are detected as edges by the enhanced Canny edge detector. The connected vehicle-color region contains pixels that are classified as vehicle colors. The last two features and are defined as the aspect ratio and the size of the connected vehicle-color region where the pixel p resides. More specifically, A =Length/Width and feature Z is he pixel count of vehicle region1 in fig.4.1.

Fig.4.1.Neighborhood region for Feature extraction

4.5 The postprocessing

In order to enhance the detection mask and perform connected component labeling to get the vehicle objects use some morphological operations. The size and the aspect ratio constraints are applied again after morphological operations in the postprocessing stage to eliminate objects that are impossible to be vehicles.

4.7 Tracking

When traffic density grows higher, more false positives and false negatives appear in detectors results. All these problems lead to difficulties for tracking cars. Traditional tracking algorithm divided the whole frame into two parts: detecting area and tracking area. A traditional tracker only tracks a detected object in tracking area. Actually, lots of uncertainties make it possible to detect an object for the first time everywhere in a frame. Based on the edge and corner points, it is possible to find out the area of vehicle appear in each frame. Also, centroid of each detected object can be found out from this edge and corner points appear in each frame. In order to track the vehicle, check whether the vehicle with same area appears in the consecutive frames also. When detect a new car or detect an old car in its small predicted area, the input of this state is 1; otherwise, the input value is 0. When confirm the object in a cars predicted area is the car itself, the output of this state is also changed to 1, otherwise is 0. A counter is set for the detected vehicle, and based on corner and edge points, area of vehicle is calculated. If same area is appears in consecutive frames the counter can be incremented. Edge points, corner points and centroid of each vehicle appear in each frame is known and based on these points bounding box can be plot on these area points of consecutive frames.

5. RESULTS AND DISCUSSIONS

Experimental results are demonstrated here. In-order to analysis the performance of the proposed system, various video sequences with different scenes and different altitudes are used. The experimental videos are taken from standard data set. When performing background color removal, the existing system used a histogram based method which quantizes the color histogram bins as 16 * 16 * 16 and colors corresponding to the first eight highest bins are regarded as background colors and removed from the scene. The proposed system used a Gaussian mixture model for background removal which separate background pixels and foreground pixels more accurately. The GMM first learn the background pixel with the help of parameters like mean, variance and Based on learning mechanism, GMM separate weight. foreground and background pixels .In vehicle detection GMM takes number of mode as two and this mode represents background and foreground. If the number modes are less, GMM take less time for background removal.

From paper [1] it is observed that, the training process of DBN does not require a large amount of training samples. DBN demonstrate flexibility and good generalization ability on a wide variety of aerial surveillance scenes under different heights and camera angles. Inside the bounding box, detected vehicles are marked .When observing the detection result the number of misclassification in DBN is more compared to SVM. When employing SVM, it is important to give the block size to form a sample. According to the observation from paper [1] take each 3*4 block to form a feature vector. The color of each pixel would be transformed to u and v color components using (8) and (9). Consider a video with 120 frames. The first 20 frames are taken for training and the remaining frames are considered for testing. The SVM classified the video, with almost all vehicles in a video are detected and tracked. Notice that the blocks that do not contain any local features are taken as non-vehicle areas without the need of performing classification via SVM.

From the paper [1], it is observed that the neighborhood area with the size of 7*7 yields the best detection accuracy. In the experiment the size of the neighborhood area for extracting observations is set as 7*7The neighborhood area remains same for both SVM and DBN classifier. The impacts of the enhanced Canny edge detector on vehicle detection results can be observed [1] and shows the results obtained using the traditional Canny edge detector with non-adaptive thresholds. Non-adaptive thresholds cannot adjust to different scenes and would therefore result in more misses or false alarms. It is observed that the background removal process is important for reducing false positives and the enhanced edge detector is essential for increasing hit rates. In a recent study, experimental results were reported using several feature extraction method and classifiers (i.e., Neural Networks (NN) and Support Vector Machines (SVMs). Testing was performed using a data set obtained from traffic videos. The best approach in terms of accuracy was found to be features with SVMs.

5.1 Evaluation

In vehicle detection and tracking, the number of vehicles identified correctly can be measured using the terms precision, recall. Precision (also called positive predictive value) is the fraction of retrieved instances that are relevant, while recall (also known as sensitivity) is the fraction of relevant instances that are retrieved.

Presicion = A / (A+C)(6)

Recall=A / (A+B)(7)

A = the number of relevant records retrieved,

B = the number of relevant records not retrieved, and

C = the number of irrelevant records retrieved.

Analysis of SVM with their Confusion Table and the results obtained are depicted in Table 5.1.

1 able 5.1	Precision	and Kecall	with SVN

Input	Precision	Recall
Video 1	81.8%	100%
Video 2	90%	75%
Video 3	66%	66%

Precision and Recall parameters for the system using SVM classifier is satisfactory.

6. CONCLUSION

The system proposed automatic vehicle detection and tracking system for aerial surveillance that does not assume any prior information of camera heights, vehicle sizes, and aspect ratios. The introduced system accounts for detecting objects in a video and classifying them either as vehicle or non-vehicle. Detected vehicles are then tracked. It is found in almost all cases that, the system delivers satisfactory results for classification, if the SVM Classifier is trained adequately. The proposed system introduces a pixelwise classification method for the vehicle detection using SVM. Since the colors of the vehicles would not dramatically change due to the influence of the camera angles and heights, so the system uses only a small number of positive and negative samples to train the SVM for vehicle color classification. For background subtraction the system uses a Gaussian mixture model which is more accurate and fast learning compared to histogram based background subtraction. Overall, the entire framework

does not require a large amount of training samples. The system has also applied moment preserving threshold to enhance the Canny edge detector, which increases the accuracy for detection in various aerial images. Automatic vehicle detection and tracking could serve as the foundation for event analysis in intelligent aerial surveillance systems.

7. REFERENCES

- Hsu-Yung Cheng, Chih-Chia Weng, and Yi-Ying Chen" Vehicle Detection in Aerial urveillance Using Dynamic Bayesian Networks."IEEE Trans. On Image Processing, Vol. 21, No. 4, April 2012.
- [2] C Stauffer, W Grimson, "Adaptive background mixture models for real - time tracking". Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1999, 2(6): 248 – 252.
- [3] S. Srinivasan, H. Latchman, J. Shea, T. Wong, and J.Mc Nair, —Airborne traffic surveillance systems: Video surveillance of highway traffic, in Pro.ACm 2nd Int. Workshop Video Surveillance Sens.Netw., 2004,pp. 131-135.
- [4] S. 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.
- [5] 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.
- [6] 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
- [7] 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.
- [8] 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.
- [9] L. W. Tsai, J. W. Hsieh, and K. C. Fan, "Vehicle detection using normalized color and edge map," IEEE Trans. Image Process., vol. 16, no.3, pp. 850–864, Mar. 2007.
- [10] B. Morris and M. Trivedi, —Robust classification and tracking of vehicles in traffic video streams, Proc. IEEE ITSC, 2006, pp. 1078–1083.
- [11] J. W. Hsieh, S. H. Yu, Y. S. Chen, and W. F. Hu, Automatic traffic surveillance system for vehicle tracking and classification, IEEE Trans.Intell. Transp. Syst., vol. 7, no. 2, pp. 175–187, Jun. 2006.
- [12] Z. Zivkovic Improved adaptive gaussian mixture model for background subtraction.Int'l Conf. on Pattern Recognition, 2:28–31, 2004.
- [13] D.-S. Lee. Effective gaussian mixture learning for video background subtraction. IEEE Trans. on Pattern Analysis and Ma-chine Intelligence, 27:827–832, 2005.
- [14] C. G. Harris and M. J. Stephens,"A combined corner and edge detector," linProc. 4th Alvey Vis. Conf., 1988, pp. 147–151.

International Journal of Computer Applications (0975 – 8887) Volume 85 – No 9, January 2014

- [15] 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.
- [16] W. H. Tsai, —Moment-preservingthresholding: A new approach, Comput. Vis.Graph., Image Process., vol. 29, no. 3, pp. 377–393, 1985.
- [17] N. Cristianini and J. Shawe-Taylor, An Introduction to Support VectorMachines and Other Kernel-Based Learning Methods. Cambridge, U.K.: Cambridge Univ. Press, 2000.
- [18] S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 2003.
- [19] Alan.J.Lipton,H.Fujiyoshi, Raju.S.Patil, "M oving Target Classification and Tracking from Real-time Video", WACV '98 Proceedings of the 4th IEEE Workshop on Applications of Computer Vision (WACV'98) Page 8.
- [20] Wenhao Lu, Shengjin Wang, XioaqingDing,Vehicle Detection and Tracking in Relatively Crowded Conditions,Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics San Antonio, TX, USA.