International Journal of Engineering Research and General Science Volume 4, Issue 3, May-June, 2016
ISSN 2091-2730

# An effective algorithm to detect vehicles in real time

Shruti B Shettar[1], Mallikarjun Anandhalli[2], M.M.Raikar[3], Dr.Vishwanath Baligar[4]
Dept of Computer science
BVB College of engineering Hubli,India
E-MailId[1]: shrut.shettar@gmail.com
Contact No[1]: 9036349513

**Abstract**— With the development of technology, automated vehicle detection in aerial surveillance keeps on improving and plays an important role in modern wars, transport system and industries. Due to significant importance in many fields such as terrorist activities monitoring, traffic accidents avoidance, congestion avoidance, toll collection, military, police, security and surveillance system, it has become an important field to study. Different technologies have been proposed and implemented for detection of vehicles and aerial view is considered, which gives better appearance of area being covered. In this paper an efficient algorithm for detection of vehicles in aerial surveillance is proposed. The proposed algorithm first isolates channels from the input video, applies morphological operations on each channel and performs background subtraction in identifying the foreground objects followed by vehicle detection method. Vehicle detection is based on low level feature extraction like edges, shape and detection of grouped objects. Finally system performs post processing for the detection of vehicles. Experiments were conducted on different types of aerial videos. The results illustrate flexibility and accuracy of the proposed algorithm on different aerial videos taken under different camera angles and heights.

**Keywords**— aerial surveillance, congestion, toll collection, morphological operations, background subtraction, foreground objects, grouped objects.

## 1. INTRODUCTION

Many different kinds of techniques have been proposed, but very effective and novel technique has been introduced in this paper. Surveillance of traffic area is leading nowadays, it is because of rapid increase in the number of vehicles but roads area remains the same causing congestion. Determining the density of the vehicles in the road and analysis of the surveillance of the video will help a lot in traffic management systems.

An organization named intelligent transport systems (ITS) has invested a huge amount in traffic surveillance. The technology has replaced the surveillance system of sensors with the video camera and computer vision techniques. Number of vehicles can be detected as well as tracked with the features of vehicles in the video. Also the detected vehicles can be classified into heavy vehicles, light vehicles etc. This kind of video surveillance systems holds good for different traffic and environmental conditions.

Hinz and Baumgartner [1] proposed hierarchical model which describes vehicle features at distinct levels. There is no explicit vehicle models are assumed, which defines the flexibility of the model. With the influence of neighboring objects that is present and by the weak contrast the system has disadvantage of many of false detection. Cheng and Butler [2] considered many clues and used mixture of experts to combine the clues for vehicle detection in aerial surveillance. Mean shift algorithm is considered for performing color segmentation and motion analysis is done by change detection. In inclusion, for enforcement of contextual information and multisacale analysis they consider a trainable sequential maximum and posterior method. However, the system cannot deal with complex background changes and previously mentioned camera motions, when motion analysis algorithm is applied. Further, there is high dependency on the results of color segmentation in information fusion step.

Lin et al. [3] utilized a background subtraction method which performs subtraction of static background color of each frame then refines the foreground vehicle regions by imposing size constraints of vehicles. However, too many parameters are assumed such as the smallest and largest sizes of vehicles, and the focus and the height of the airborne camera. Assuming these parameters the known priors are not realistic in real applications In [4], the authors described a moving-vehicle detection method which is based on cascade classifiers.

The system is trained with large number of positive and negative samples. Furthermore, the detection stage generates multiscale sliding windows. This method has the disadvantage of lot of miss detections on the rotated vehicles. If the system is trained with only frontal faces, then the faces with poses are missed absolutely. However, the positive samples of faces with poses are considered, it results in increase number of false alarms.

Choi and Yang [5] suggested a vehicle detection algorithm based on symmetric property of car shapes. However, this indication is prostrate to miss detections such as road markings or symmetrical details of buildings. So, log-polar histogram shape descriptor is applied to verify the shape of the candidates. But, the fixed vehicle Model defines the shape descriptor which makes the algorithm inflexible. Moreover, identical to [2], the algorithm in [5] depends on mean-shift clustering for image color segmentation. Since vehicle has different colors of roofs and windshields, it results in separating vehicle as many regions which is the major disadvantage of this method. Furthermore, adjacent vehicles may be joined as one region if they have identical colors. The mean-shift segmentation which has high computational complexity is included in this algorithm.

In this paper, we propose an efficient and novel vehicle detection method which preserves the advantages of existing systems and avoid their defects. The proposed system design is illustrated in Fig. 1. The proposed algorithm is purely based on morphological operations, it extracts frames from the given video, convert it into HSV format and the channels of the frames are split into separate channels. Each channel is processed separately with the series of morphological operation such as erosion and dilation succeeded by tophat and bottomhat transformation. Frames are converted from HSV to grayscale color model, so that low level feature can be extracted easily. Moving foreground region is separated from the static background by using background subtraction method, followed by low level feature extraction by canny edge detection. The proposed algorithm considers shape and size as low level feature and detects grouped object which results in blob detection followed by post processing such as drawing bounded rectangle on the detected vehicle and counting the number of vehicles. Section II elaborates the proposed vehicle detection mechanism in detail. Section III comprises of experimental results. Finally, conclusions and future works are made in Section IV.

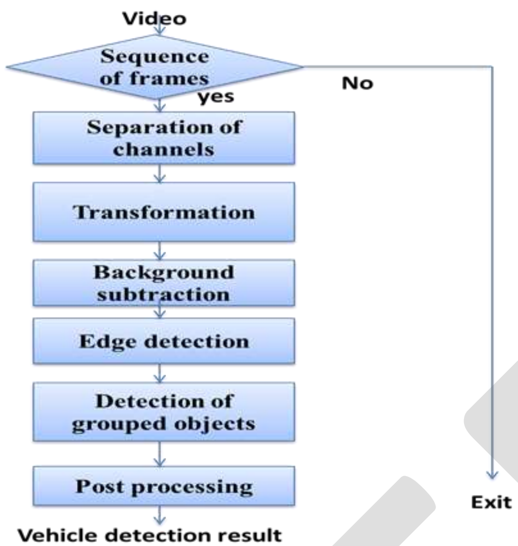



Figure 1 Proposed system Framework

Here, we explain each module of the proposed method in detail.

## 2. PROPOSED METHODOLOGY

### 2.1 Separation of channels

HSV formats are most suitable for detection, since HSV provide absolute color space of each vehicle. Hence we convert we convert from RGB to HSV where Hue (H) and Saturation (S) has only color information and Value(V) has only intensity information. And the intensities of the channels are processed without altering the color information. After processing, intensity information is recombined with the color information. We want to extract color information from the image so we split the colored image of HSV color space to three different channels and process the pixel intensity value in each channel. Splitting an image in its color channel decreases the time complexity of algorithm.

### 2.2 Transformation

Morphological processing is basically collection of non-linear operations related to the shape or features of an image. It depends on relative ordering of pixel values and not on their numerical values. Morphological operations are usually performed to remove noise, isolating the individual elements and joining separate elements in an image and also for finding intensity bumps or holes in an image. Morphological techniques examine an image with a pattern or small shape called structuring element.

The structuring element is located at all possible positions in an image and it is analyzed with the corresponding proximity of the pixels. Morphological operation differs in how we carry out this comparison. Structuring element is also called as kernel and consists of a template specified as the co-ordinates of a number of distinct points. Consider an example of 8X8 image as shown in the fig below where the structuring element is of 2X2. White color box represents zero pixel value and grey color box represents non zero pixel value. The structuring element is slide across the image and examines the image with the structuring element values. The label 'A' indicates the structuring element not fitting in the image, 'B' indicates structuring element intersecting the image and 'c' indicates structuring element fitting in the image.

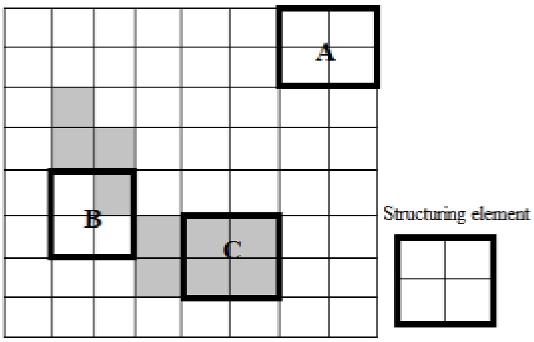



Figure 2 Illustration of structuring element

The fundamental morphological operations are erosion and dilation.

### 2.2.1 Erosion

Erosion operation is to gradually decrease the boundaries of region of foreground pixel, the area of foreground pixels shrink in size and strips away a layer of pixels from an object where the holes within those areas become larger.

It considers two pieces of data, one is the input image and other is structuring element. For each of the foreground pixel, structuring element is superimposed on the top of the input image. So that origin of structuring element coincides with the input pixel co-ordinates and the pixels which are not completely enclosed by other foreground pixels are removed, these pixels lay at the edges of foreground regions. So it results in shrinking of foreground region where holes inside a region grows. It results in removal of small spurious bright spots in images.

### 2.2.2 Dilation

Dilation operation gradually enlarges the boundaries of regions of foreground pixels, thus area of foreground pixels grow in size where holes within these region become smaller.

This operation considers two input data sets, one is the input image and other is structuring element, it considers the background pixel in the input image as input pixel, structuring element is superimposed on the top of the input image. So that structuring element coincides with the input pixel co-ordinates and if there is an intersection of at least one pixel between structuring element and foreground pixel then the input pixel is set to foreground. This operation sets the background pixels to the foreground pixel value if it has a neighbor foreground pixel and these pixels lay at the edges of foreground regions. So it results in in growing of foreground regions where the holes inside a region shrink and erosion is used in filling the small spurious holes in images.

### 2.2.3 Opening and closing

An opening is defined as erosion succeeded by dilation with the same structuring element. It considers two input data sets, one is the input image and other is structuring element.

Let 'f' be the input image and 'õ' be the structuring element. 'Ó' indicates opening operator, 'Ê' indicates erosion and 'Đ' indicates dilation. The opening operation is represented in equation (1).

**Ó (f, õ) = Đ (Ê (f, õ), õ) --------------(1)**

This operation retains foreground regions that have identical shape of structuring element and eliminates other regions of foreground pixels.

Closing operation is performed in revere of opening. It is defined as dilation succeeded by erosion with the same structuring element. It considers two input data sets, one is the input image and other is structuring element.

'Ç' indicates closing operator. The closing operation is represented in equation (2).

**Ç (f, õ) = Ê (Đ (f, - õ),- õ) --------------(2)**

This operation retains background regions that have identical shape of structuring element and eliminates other regions of background pixels.

### 2.2.4 Tophat and Bottomhat Transformation

Tophat transformation is performed by subtracting the original frame from the opening operator. The tophat transformation is represented in equation (3). Let I(x,y) be the original frame. 'Ť' indicates tophat transformation.

**Ť(x,y) = Ó (f, õ) - I(x,y) ---------------(3)**

Bottomhat transformation is performed by subtracting the closing operator frame from the original frame. The bottomhat transformation is represented in equation (3). 'Ḃ' indicates bottomhat transformation.

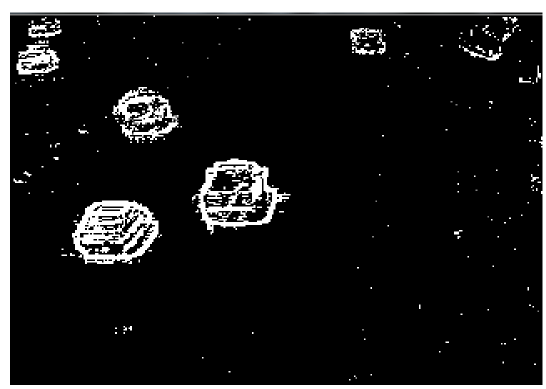
**Ḃ(x,y) = I(x,y) - Ç (f, õ ) ----------------(4)**

After the transformation method, channels are merged and further processed with the combined channels.

### 2.3 Background subtraction

Moving foreground objects are isolated from the static noisy background using background subtraction method [6]. It is the most important technique used to extract foreground objects from the image. It performs frame by frame subtraction which results in the removal of static noise in the video. The objects are segmented using Gaussian mixture model of background subtraction method. The proposed system extracts the foreground region by background subtraction [7] which gives the region where the vehicle exists in the scene. A mask is created for this region, convolved over the input image to obtain the vehicle and subtracts the background. Figure 3 shows the result of foreground object extraction.

(a)



(b)

Figure 3: (a) Input image (b) Object extraction

### 2.4 Edge detection

For extracting low level feature edge, classical canny edge detection is used which is less sensitive to noise and reduce the noise by smoothing. Two thresholds $T_{high}$ and $T_{low}$ are set and avoid streaking problem. It provides good localization by considering gradient orientation. Once the edges are extracted from the object, blurring of the extracted object is performed so that superfluous interior edges are not considered in further processing.

### 2.5 Blob detection

Grouped object which is called as blob can be identified by using eight way connectivity. In this method the pixels are scanned across the image and foreground pixel is checked with the eight neighboring pixels, whenever the input pixel is connected with the foreground pixel it checks for the label of that pixel if the neighboring foreground pixel is already labeled, then it assigns the same label to the input pixel or it assigns unique label to the input pixel. This process is carried out till it reaches the last pixel of the frame.

Consider an image of 8X8 as shown in the figure below, which consist of three objects which are uniquely labeled as L1,L2,L3. Gray color indicates the foreground pixel i.e ‘0’ and white color indicates background pixel i.e ‘1’.

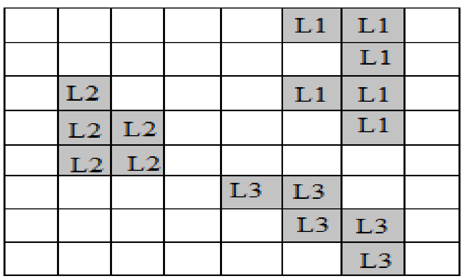



Figure 4 Blob detection

Since the grouped object identified has irregular shape convex hull is used to define a regular shape. It considers the exterior points of the blob and polygon is constructed from these points.

### 2.6 Post processing

Size of the blobs are calculated which is based on the calculating the number of occurrence of the foreground pixel in the blob. Size constraint is enforced on the object so that small objects which can act as noise can be removed. Bounded rectangle is drawn on the objects which confines the detected vehicle and a virtual line is drawn to count the number of vehicles, the centroid of the bounded rectangle is calculated. When the centroid intersects the virtual line, count is incremented.

## 3. EXPERIMENTAL RESULTS

Experimental results of the proposed algorithm is demonstrated in this section which is conducted on different datasets of aerial videos which is taken under different camera angles and heights, the frame rate of the input video considered is 25 frames per second and video is of length 15 minutes. Figures in this section demonstrate the detected vehicles with the bounded rectangle on the vehicles and the red color line in the figure is a virtual line which is drawn to count the number of vehicles. Fig (6) shows the detection of all the vehicles in the complex environment which confines the accuracy of the proposed algorithm in the complex environment. Fig (9) shows that the bounded rectangle on the detected vehicle remains until the vehicle goes out of the frame which confines the stability of the algorithm. Red color circle in the Fig (11) shows the total number of vehicles in the video and the blue color circle indicates the vehicles coordinates intersecting the virtual line. Time complexity of the algorithm for processing each frame is less which is 15 frames frames per second since the numbers of instructions used are less and the computational complexity of the algorithm is less

ISSN 2091-2730

which makes use of basic techniques like morphological operations, etc. Table (1) shows the results

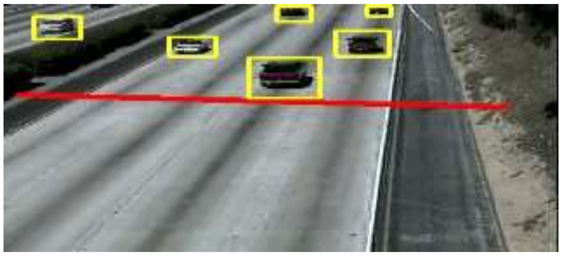


Figure 5

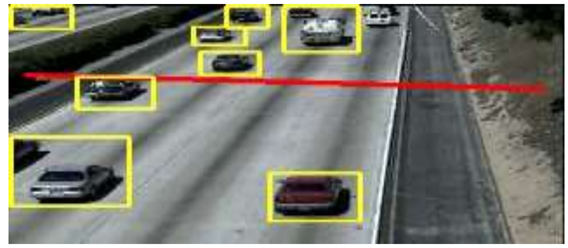


Figure 6

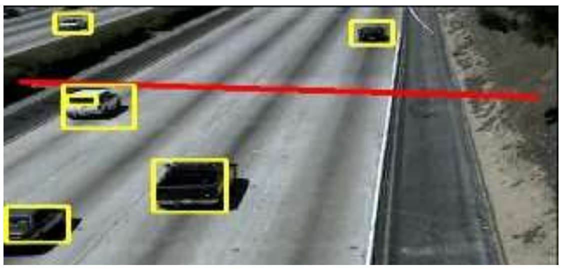


Figure 7

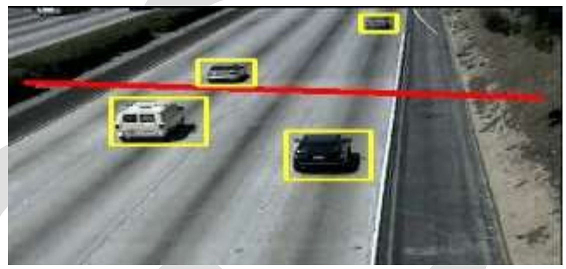


Figure 8

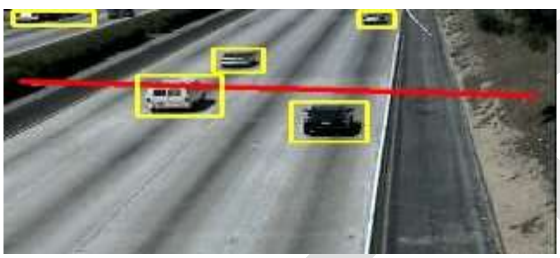


Figure 9

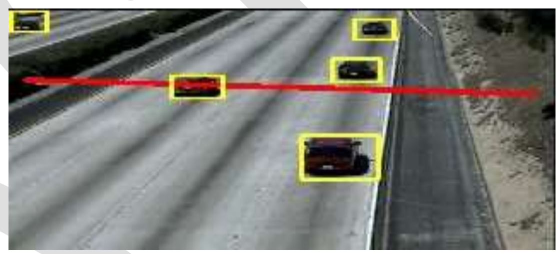


Figure 10

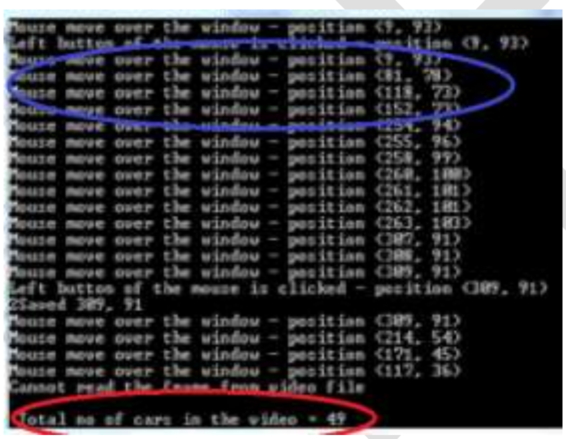



Figure 11: command prompt of vehicle count shown with red color circle.

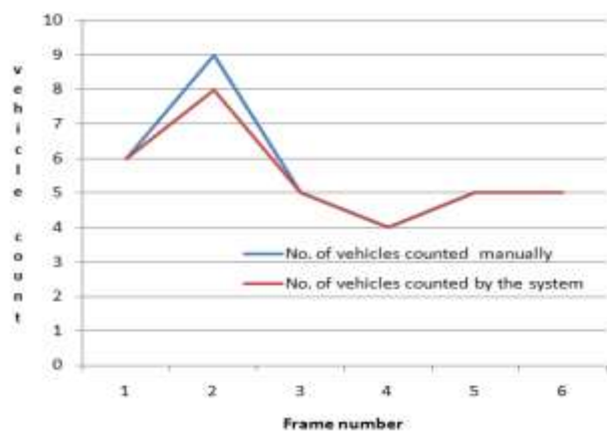



Figure 12: Analysis of vehicle detection

of the algorithm with number of vehicles detected by the system, number of vehicles not detected and the vehicles detected manually, precision is calculated using the equation (5)

$$\mathbf{P} = \frac{\mathfrak{M}}{\mathfrak{w}} \quad \text{----------------------------------} \ \mathbf{(5)}$$

'𝔐' indicates the number of vehicles detected manually and '𝔴' indicates number of vehicles detected by the system.

Precision value close to '1' indicates the accuracy of the algorithm. Graph in the fig (12) with x-axis frame number and y-axis vehicle count demonstrates the algorithm pictorially with line graph, blue line indicates the number of vehicles counted by the system and the red line indicates number of vehicles counted by the system with implementing the proposed algorithm. Overlapping of the two lines defines the feasibility of the algorithm in the complex environment even when the vehicles are close enough.

| Scene | No. of vehicles counted manually | No. of vehicles counted by the system | No. of vehicles not detected | Precision |
|---|---|---|---|---|
| 1 | 6 | 6 | 0 | 1 |
| 2 | 9 | 8 | 1 | 0.89 |
| 3 | 5 | 5 | 0 | 1 |
| 4 | 4 | 4 | 0 | 1 |
| 5 | 5 | 5 | 0 | 1 |
| 6 | 5 | 5 | 0 | 1 |

Table 1: Tabular representation of video frame results

## 4. CONCLUSION AND FUTURE WORK

The proposed algorithm provides efficient and novel methodology for vehicle detection in aerial surveillance without assuming any details of camera heights, views and vehicle sizes. Canny edge detection is used which increase the accuracy and adaptability of detection in aerial images. Proposed algorithm has less computation complexity which increases the efficiency of the algorithm. The experimental result illustrates flexibility and accuracy of the proposed algorithm on different aerial videos taken under different camera angles and heights. The proposed method can be extended to perform tracking on the detected vehicle which stabilizes the results. It can also be improved in solving merging problem and working on shadows.

**REFERENCES:**

[1] 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.
[2] 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.
[3] 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.
[4] 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.
[5] 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.
[6] Jin-Cyuan Lai ;Shih-Shinh Huang ; Chien-Cheng Tseng , "Image-based vehicle tracking and classification on the highway", Green Circuits and Systems (ICGCS), 2010 International Conference, 21-23 June 2010
[7] Zhong Qin, Guangzhou, "Method of vehicle classification based on video", Advanced Intelligent Mechatronics, 2008. IEEE/ASME International Conference, 2-5 July 2008
[8] 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.
[9] Omar Javed, Khurram Shafique, and Mubarak Shah. Appearance modeling for tracking in multiple non-overlapping cameras. In CVPR, 2005.
[10] A.G.A. Perera, C. Srinivas, A. Hoogs, G. Brooksby, and W.S. Hu. Multi-object tracking through simultaneous long occlusions and split-merge conditions. In CVPR, 2006.
[11] Robert Kaucic, Amitha Perera, Glen Brooksby, John Kaufhold, and Anthony Hoogs. A unified framework for tracking through occlusions and across sensor gaps. In CVPR, 2005.

[12] 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.
[13] M. Betke, D. E. Hirsh, A. Bagchi, N. I. Hristov, N. C. Makris, and T. H. Kunz. Tracking large variable numbers of objects in clutter. In CVPR, 2007.