The Eighth International Conference on Intelligent Technologies (InTech'07)
Sydney, Australia, 12-14 December 2007

Multiple Cue Based Vehicle Detection and Tracking for Road Safety

Silas Wan, Sarath Kodagoda* and Stephan Sehestedt*
*ARC Centre of Excellence for Autonomous Systems (CAS)
Faculty of Engineering
University of Technology, Sydney
Broadway, NSW, Australia
Email: { s.kodagoda, s.sehestedt}@cas.edu.au, Silas.Wan@uts.edu.au

Abstract **– With the rise in accident related fatalities on roads, the researchers around the world are looking for solutions including integrating intelligence to vehicles. One cruicial aspects of it is the robust detection and tracking of other vehicles in the visinity. In this paper, we have proposed a probabilistic way of incorporation of several visual cues in vehicle detection and a particle filter based tracking strategy. Visual cues used are, lane markings, symmetry, shadows and edge information. Combination of visual cues provided us with robust results when compared with their individual counterparts. The definition of a region of interest lowers the computational requirements with improved robustness. Experimental results of the algorithm in Sydney urban areas are presented.**

Index Terms – Vehicle Detection, particle filter, road safety, tracking, Fuzzy c means clustering, computer vision

I. INTRODUCTION

Road safety is a worldwide concern [1]. Developed countries have managed to significantly reduce the impact of road traffic injuries/fatalities by developing infrastructure, enforcing legislation to control speed and alcohol consumption, mandating the use of seatbelts and crash helmets. A comprehensive five year study of road accidents by Treat et al [2] found that the human error was the sole factor in 57% and a contributing factor in 92.6%. They also concluded that over 80% of all automobile accidents are due to human perceptual error. Therefore, further reduction of road related accidents can be achieved by enhancing driver perception by means of intelligent sensing and advanced processing.

Vehicle detection and tracking by means of sensors fitted on a vehicle is important towards enhancing driver perception. In this paper, we focuse our attendtion on using camera based methodology for vehicle detection and tracking. Main advantages of using camera based methods are the low cost when comparing with counterparts such as, laser range finders, infra red sensors and radars. Camers also provide large amount of information, which can effectively be utilized. However, the downside of using cameras are the presence of significant variability depending on the time of the day and viewing direction. Forexample, the visual cues that are suitable for day time operation may be uselees in night. The visual cues best suited in detecting mid to far range vehicles may not be usable for the vehicles that are at a short range. Therefore, in this paper, we focus our attention mainly on detection and tracking mid to far range vehicles in day time images.

Camera based techniques for vehicle detection is a challenging problem due to the large variability of visual cues and environmental conditons. Vehicles may have different shape, colour and texture. The shape is dependent on the viewing direction. The environment conditions such as lighting and occlusions can adversely affect the detection process.

Camera based vehicle detection problem is widely researched in the literature [11]. Kuehnle et al [3] used symmetry of vehicles as a visual cue to detect vehicles. Tzomakas et al [4], use transition between paved road and information about shadows to locate a vehicle. Kalinke et al [5] use texture to detect vehicles. Individually, these algorithms are somewhat successful in locating vehicles but suffers from false-positives. Van Leeuwen et al [6] combine these three visual cues into a single detection algorithm which performs well in enviroments with simple background. However, the algorithm is restricted to frame by frame detection without using temporal information.

This paper brings together camera based vehicle detection algorithms and tracking algorithms. Different to most approaches, combination of the visual cues are achieved in a probabilistic manner. First, the lane detection is carried out to define the region of interest (ROI). Then the features in the ROI are extracted using edge detection. For each pixel, a small area around it is examined for the three visual cues. Based on the strength of each visual cue in each pixel, a weight is assigned. The weights are normalized and considered as probability density functions. When the three distributions are combined, an overall probability distrubution is obtained. The resulting distribution reflects the likelihood of a vehicle in the image. The possible vehicle locations are tracked using a particle filter. The particles are then grouped together with clustering algorithms to differentiate individual vehicles.

Unlike most papers in the literature which only focuses on vehicle detection, this paper acknowledges the advantages of vehicle detection and tracking. Also, most papers merge different visual cues in a sequential manner. We combine different visual cues through a probabilistic manner treating them as independent quantities, which improves detection capabilities. Further, the parallel merging of the visual cues allow the flexibility of adding or removing visual cues without drastically altering the algorithm, which is important for generalization.

This paper is organized as follows. Section II discusses the visual cue extraction procedure. Section III outlines the partical filter based tracking methodology. Clustering algorithm is given in Section IV. Experiment results are given in Section V. Section VI provides a discussion about the limitations of the algorithm. Section VII concludes the paper.

II. EXTRACTION OF VISUAL CUES

Visual cues are features in an image that will assist in locating the object of interest (in our case it is a vehicle). A camera image contains a large amount of information or features, which can be due to vehicles or other roadway structures. For example, most of the man made structures including vehicles are symmetric around some axis. Therefore, symmetric property alonne can not be used to distinguish a car from a building. This ambiguity can be minimized by incorporating other visual cues such as symmetry, shadow and edge information.

A. ROI based on Lane Detection

Firstly, lane markings in the visual images are detected. Here, we have more focused on urban roads rather than highways. Detection of road markings on highways are simpler when comparing with urban roads. Assuming geometric models of roads, such as lanes being straight or curved is only appropriate for highways but is probable to fail in urban areas. Main reasons for the complexity of roads include bad visibility of markings, occlusions or missing markings over extended periods, and complexity of markings themselves. Therefore, strong assumptions about the lane geometry can lead to failure and hence weaker models are preferable.

Low level image processing for lane marking detection is performed on the Inverse Perspective Map (IPM) applying a lane model which stipulates that a road marking is represented by a predominantly bright line (lane marking) of constant width surrounded by a darker region (the road). Thus, the pixels belonging to a road marking have a brightness value higher than their left and right neighbors do at a given horizontal distance. A vertical edge in an image conforms similarly to the same principle; however, the intensity difference between neighboring pixels must be over a threshold to be validated as a lane marking. Therefore, an exhaustive search across each row of the image will produce potential lane marking candidates where matching probability can be measured with edge quality.

Then we use a particle filter based approach, where particles are used to represent a possible position of a piece of lane marking, including a weight that is proportional to the probability that a lane marking is intact present. Particles are initialized at the bottom of the image and move to the top seeking lane markings [7]. Some lane detection results in urban roads are shown in Fig. 1.

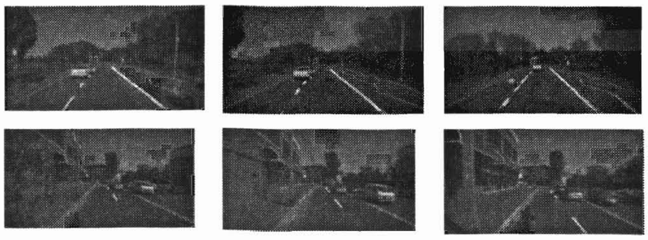


Fig.1. Lane marking detection

Once lane markings are detected, a ROI in the image is defined. Based on the lane markings, the vanishing piont is determined. A line connecting lower middle of the image and vanishing point is used to determine the driving lane of the prototype vehicle and the ROI is defined covering two adjescent lanes on both sides. Fig. 2 shows the ROI detected based on lane markings.

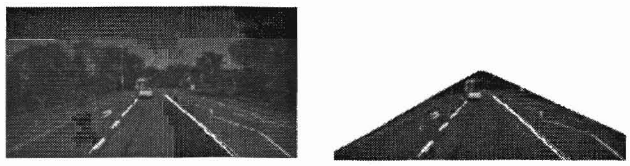
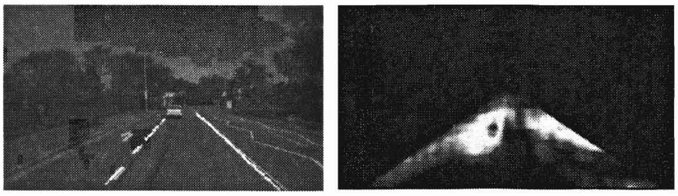


Fig.2. ROIs based on detected lane markings

B. Symmetry

Vehicles in visual images are symmetric when viewed from an appropriate direction, for example front view or rear view. Here, we use grey level image to calculate the symmetry. Firstly, the ROI determined by the lane detector is further processed with Sobel edge detector. For every pixel on the edge image, 40 horizontal fixels (20 on each side) are examined for symmetry and are given a weighting (symetric weighting). Higher symmetric weights are given for the symmetric pixels that are further away from the center pixel. This is to discourage small 'skinny' items causing spurious detections, which are caused by for example, poles and pedestrians. Typical objects that possess high level of symmetry in road environments are vehicles, buildings, windows and trees. However, the ROIs defined earlier eliminate most of the offroad objects improving the reliability. Fig.3 shows the normalized symmetric weights calculated.



(a) input image (b) symmetricity image:

Fig.3. Symetricity

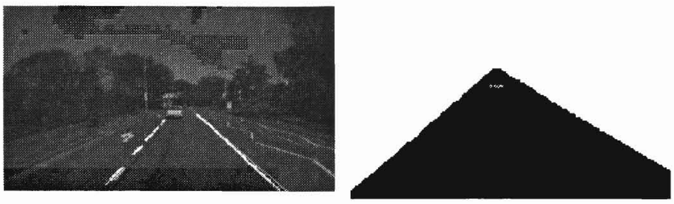
C. Shadow

Vehicle underneath shadows provide distinctly darker areas on asphalt paved roads. However, this is nontrivial to extract due to environmental conditions such as intensity variations and other shadows caused by nearby structures. In our shadow detection algorithm, areas of low pixel intensity values are determined from the grey scale image. Fig. 4 shows detected shadows in a roadway image. As can be seen, the vehicle shadow and a darker part of the road (right, bottom of Fig. 4 (c)) are picked as shadows.

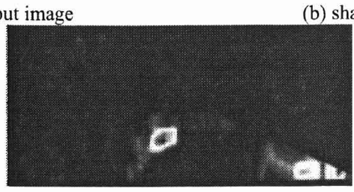
Assuming the vehicle in an image will always be located above its own shadow, a normal distribution with a vertical offset is assigned to every 'dark' pixel whose mean be located above the pixel.

$$w_{shadow}(i, j+t) = \eta e^{\frac{-(t-\frac{R}{2})^2}{2\sigma^2}} \tag{1}$$

where η is a constant that reflects the vehicle size, R and σ are mean and standard deviation of the distribution, both depend on the location of the shadow. This reflects the fact that the vehicles at far have tighter distributions than that are closer. Once all the individual contributions of each 'dark' pixel are summed up, the entire distribution is normalized. It can then be thought of as a probability density function of possible vehicle locations based on shadow information (see Fig. 4).



(a) input image (b) shadows

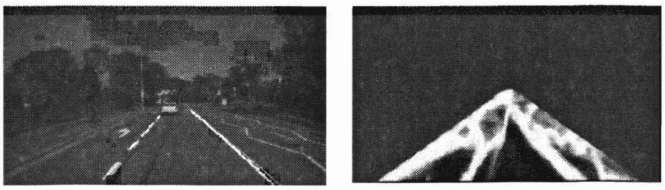


(c) weights based on shadows

Fig.4. Shadows

D. Edge information

Edge information can be used as a measure of information content. First, the image is edge detected (Sobel). Then the number of edge pixels present within an interested area is calculated as a weight. By using the edge information, we could reduce confusion caused by features that lack edge information but still satisfies the symmetry and shadow criteria such as a complete black blob on road.

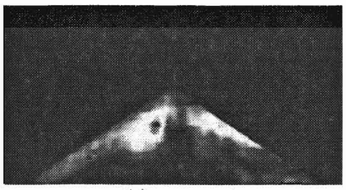


(a) original image (b) weights based on edge information

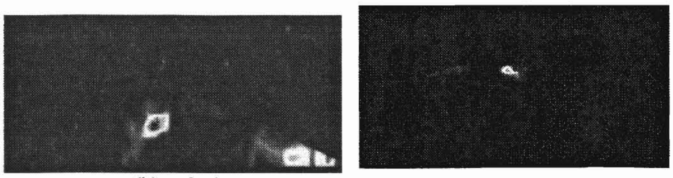
Fig.5. Edge information

D. Integration of Visual Cues

Three cues, namely, symmetry, shadow and edge information, are combined together to represent the most probable vehicle locations in the ROI (see Fig. 6).

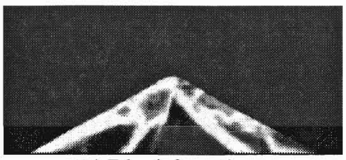


(a) symmetry



(b) shadow

(d) Combined weights



(c) Edge information

Fig.6. Combined cues

Higher 'peaks' show the most possible vehicle locations on the image. This distribution is then used as an input to the tracking algorithm discussed in the next section. This method of integrating visual cues facilitates integration of new visual cues or deleting unwanted visual cues without having to change the overall framework significantly.

III. TRACKING ALGORITHM

Vehicle detection is the most difficult when compared with vehicle tracking, which has additional information, such as temporal continuity and structured nature of the dynamics. Here, we propose to synthsize the vehicle detection and tracking by utilizing the probabilistic weights calculated in the previous section and a particle filter tracker.

A particle filter is a Bayesian tracking algorithm that simulates the prior and posterior state distributions with a large number of particles. It is widely used in the modelling of non-linear and non-Gaussian physical systems. Like the Kalman filter, it has a prediction and an update step. In our approach, a particle filter is used instead of the kalman filter because of the difficulty of identifying a 'crisp' tracking point on the vehicle. Centre of area occupied by the vehicle in the image could be used, however, it has adverse effects due to change in viewpoints and occlusions. Number plate can be another option, however, once the vehicle is far away, image resolution is too small for the purpose.

The particle filter used is a Sampling Importance Resampling particle filter [9]. It is one of the basic particle filter types. It includes a resampling step which allows particles to be redistributed according to the level of attention that is necessary in different areas. The number of particles used for the implementation is 500. For the prediction step, it is assumed that all vehicles are moving forward. On the image plane, this involves moving the particles towards the vanishing point of the image. For the update step, the importance density is chosen to be the prior and is taken as the combined distribution from the vehicle detection algorithm described in the previous section. Therefore, the particle weights are simply taken as the local values of the combined distribution.

$$p(x_k \mid z_{1:k}) \approx \sum_{i}^{N} P_{total}(x_k)\delta(x_k - x_k^i) \quad (2)$$

where, x_k is the state vector which contains both x and y coordinates of the vehicle location in the image. The purpose of the particle filter is to 'discretize' the combined distribution into 500 numbers. These numbers are later grouped together with the use of clustering algorithms, which is discussed in the next section.

Resampling involves the regeneration of a new set of particles based on the current weights. A larger number of particles will be distributed to areas of higher weights than areas of lower weights. This forces the particles to concentrate in areas with high probability of vehicle existance. Fig. 7 shows that the particles converged in 3 frames with additional information provided by the ROI.

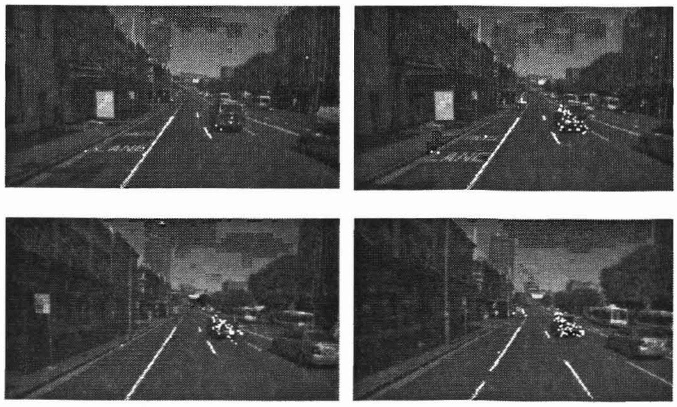


Fig.7. Convergence of the partical filter

IV. CLUSTERING ALGORITHMS

In Fig. 7, particles are distributed all over a vehicle. As there is a single particle filter is used, a clustering algorithm is needed to distinguish two objects. Two clustering algorithms were attempted, namely, fuzzy c means [8] and the Bachelor Wilkins algorithms [10]. The Bachelor-Wilkins algorithm has the advantage of using it without knowing the number of clusters *apriori*. Instead, it can determine the number of clusters from the data. It is done by controlling a threshold value, which governs the minimum distance between clusters. If the minimum intercluster distance is not satisfied, the data is combined. One weakness of this algorithm is that it does not rely on the cluster centers of previous frames, which makes the cluster centers quite unstable in implementation.

The fuzzy c means algorithm offers a more stable cluster centers. The new cluster centers are calculated from a set of existing cluster centers and the new data points, which offers stability over the cluster centers. This arrangement protects the stability offered by the Bayesian nature of the particle filter. Fuzzy c means however does require a predetemined number of clusters to start off, which is its major weakness. The fuzzy c-means update for each cluster j is [8],

$$u_{ij} = \frac{1}{\sum_{k=1}^{c}\left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|}\right)^{\frac{2}{m-1}}} \qquad c_j = \frac{\sum_{i=1}^{N} u_{ij}^m x_i}{\sum_{i=1}^{N} u_{ij}^m} \quad (3)$$

where,

u_{ij}	Cluster weight of i^{th} point relating to j^{th} cluster
c_j	Cluster centre of j^{th} cluster
N	Number of clusters
m	Fuzzification Parameter (usually 2)
x_i	Data points

V. EXPERIMENTAL RESULTS

A. Experimental Set up

Experiment test-bed used is the CAS Research Ute for Intelligence, Safety and Exploration (CRUISE), which is a Ford courier fitted with variety of sensors and computing hardware (see Fig.8). CRUISE was driven along Sydney urban roads logging sensory data. For the purpose of this paper, we have only utilized the images captured by roof rack mounted 3 CCD camera.

Fig. 8 CRUISE test-bed

B. One Vehicle

Here, we show the results of most simplified single vehicle detection and tracking. It can be seen that the algorithm tracks a vehicle as long as the vehicle is not too small on the image or dissapears from the end of the lane markings (see Fig. 9). The ROI has eliminated the other spurios detections due to trees, road signs and nearby buildings.

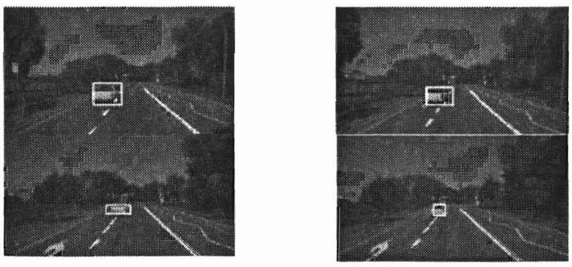


Fig.9. Single vehicle tracking

A. Complex Traffic Scene with multiple vehicles

Fig. 10 shows a more complex scenario with tracking of three vehicles. The complex features in the image provide good reasons for using a combination of visual cues. For example, the window frames on the left hand side of the image are symmetrical and has a shadow beneath it. However, they do not have sufficient edge information nor do they fall within the ROI given by the frames. Similarly, the trees on the right hand side are rejected due to lack of edge information and being out side of the ROI. The tall building in the far background is outside of the ROI and lacks a shadow underneath. Therefore, the vehicles in front are the only features that satisfies all requirements of visual cues.

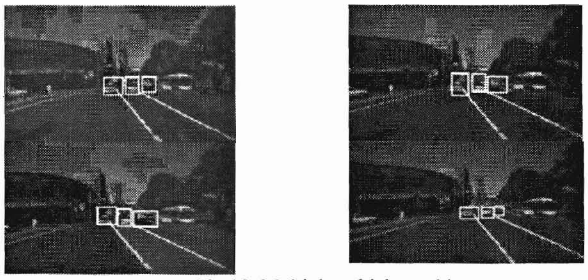


Fig.10. Multiple vehicle tracking

The sequence of the images in the Fig. 11 shows that the vehicles on the right most lane are not being tracked. That is mainly due to the fact that the vehicles are situated outside the ROI. Furthermore, it can be seen that such vehicles are not very symmetrical due to the viewing direction.

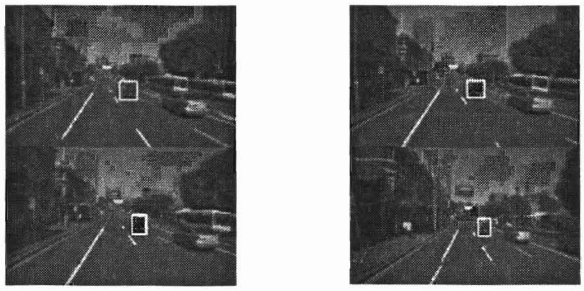


Fig.11. Single vehicle tracking

VI. LIMITATIONJS

Camera images provide large amount of information, however with high variability. Therefore, feature detection is non trivial. In this paper, we have only concentrated on the medium to far range vehicle detection. The detection was further resctricted to adjacent lanes. Detection and tracking of closeby vehicles, overtaking vehicles or the ones at larger lateral displacements need different cues and are like to fail with the current algorithm.

More importantly, relying on visual cues for vehicle detection can only be robust within limits. Combination of visual cues increases robustness, but there are scenarios where non-vehicles satisfy the visual cues. Also, special vehicles may not satisfy the presently defined visual cues, for example a bus with complex advertisements. Therefore, it is essential to investigate on an automatic way of selecting proper visual cues based on the visual image.

In this paper, we tracked the vehicles in the image plane due to the lack of depth information. However, we have future plans of tracking targets in vehicle corodinates by utilizing the camera calibration parameters and flat road assumption. The clustering algorithm will need more attention.

VII. CONCLUSION

This paper illustrates the feasibility of combining visual cues in a probabilistic manner to detect vehicles and applying a particle filter to track them. The visual cues include, lane markings, symmetry, shadows and edge information. Lane markings are used to define a region of interest. Weights of symmetry, shadow and edge information are only calculated with in the ROI improving the computation requirements as well as robustness. Those weights are combined and a particle filter is used to track the high weight (most probable to be a vehicle) portions. Data captured on Sydney urban roads are used to analyse the results. The methodology of combining visual cues in a probabilistic manner allows new visual cues to be added or existing ones to be removed easily, which would provide a framework for future research on the subject of interest.

ACKNOWLEDGMENT

This work is supported by the ARC Centre of Excellence program, funded by the Australian Research Council (ARC) and the New South Wales State Government. The authors gratefully acknowledge the partial funding of this work by the German Academic Exchange Service (DAAD) with a PhD student scholarship.

REFERENCES

[1] World report on Road Traffic Injury Prevention: Summary, World Health Organization, Geneva, 2004.

[2] Treat, J. R., Tumbas, N. S., McDonald, S .T., Shinar, D., Hume, R. D., Mayer, R. E., Stanisfer, R. L., and Castillan, N. J. (1977) Tri-level study of the causes of traffic accidents. Report No. DOTHS-034-3-535-77, Indiana University.

[3] A. Kuehnle, "Symmetry-Based Recognition for Vehicle Rears," Pattern Recognition Letters, vol. 12, pp. 249-258, 1991.

[4] C. Tzomakas and W. Seelen, "Vehicle Detection in Traffic Scenes Using Shadows," Technical Report 98-06, Institut fu¨ r Neuroinformatik, Ruht-Universitat, Bochum, Germany, 1998.

[5] T. Kalinke, C. Tzomakas, and W. von Seelen, "A Texture-Based Object Detection and an Adaptive Model-Based Classification," Proc. IEEE Int'l Conf. Intelligent Vehicles, pp. 143-148, 1998.

[6] van Leeuwen, M.B. Groen, F.C.A. ,Vehicle detection with a mobile camera: spotting midrange, distant, and passing cars, The IEEE Robotics & Automation Magazine, Vol.12, Issue.1, March 2005, pp. 37- 43.

[7] Stephan Sehestedt, Sarath Kodagoda, Alen Alempijevic and Gamini Dissanayake, "Robust Lane Detection in Urban Environments", in press, the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS 2007) , Oct., 2007, pp123-128.

[8] Rui Xu and Donald Wunsh II, "Survey of Clustering Algorithms", IEEE Transactions on Neural Networks, VOL 16, May 3 2005.

[9] M Arulampalam, Simon Maskell, Neil Gordon and Tim Clapp, "A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking", IEEE Transactions on Signal Processing, VOL 50, NO 2, FEB 2002.

[10] Amy Soller, "Computational Analysis of Knowledge Sharing in Collaborative Distance Learning", "http://www.cscl-research.com/Dr/documents/Soller-Dissertation-Proposal.doc".

[11] Zehang Sun and George, "On Road Vechicle Detection – A Review", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 28, No5, 2006