Real-Time Vision-Based Vehicle Detection for Rear-End Collision Mitigation Systems

D. Balcones, D.F. Llorca, M.A. Sotelo, M. Gavilán,
S. Álvarez, I. Parra, and M. Ocaña

Department of Electronics, University of Alcalá, Alcalá de Henares, Madrid, Spain
{llorca,sotelo,parra,miguel.gavilan}@depeca.uah.es

Abstract. This paper describes a real-time vision-based system that detects vehicles approaching from the rear in order to anticipate possible rear-end collisions. A camera mounted on the rear of the vehicle provides images which are analysed by means of computer vision techniques. The detection of candidates is carried out using the top-hat transform in combination with intensity and edge-based symmetries. The candidates are classified by using a Support Vector Machine-based classifier (SVM) with Histograms of Oriented Gradients (HOG features). Finally, the position of each vehicle is tracked using a Kalman filter and template matching techniques. The proposed system is tested using image data collected in real traffic conditions.

1 Introduction

The rear-end collisions are one of the most common types of automobile accidents. A rear facing camera mounted on the rear of the vehicle can provide an important number of driving assistance functions such as collision warning systems that will alert the driver of an impending collision or pre-crash systems (seat belt pretensioning, intelligent headrest, etc.). Accordingly, the work presented in this paper is directly related with the automotive industry.

The system is divided in four main blocks: rear-lane detection, candidates selection, single-frame classification and multi-frame validation and tracking. The global overview of the system is depicted in Figure 1.

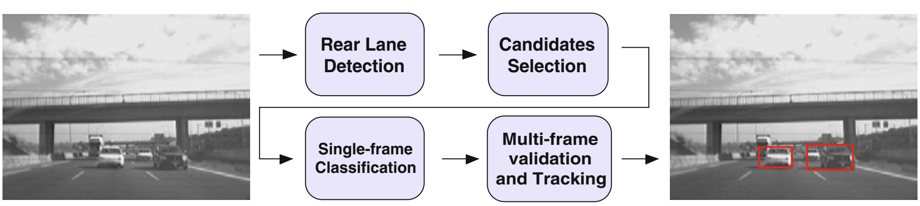

Fig. 1. Global overview of the rear-end vehicle detection system

R. Moreno-Díaz et al. (Eds.): EUROCAST 2009, LNCS 5717, pp. 320–325, 2009.

2 System Description

2.1 Rear-Lane Detection

This stage is carried out by using a Lane Departure Warning (LDW) system previously developed by the authors [1]. The LDW system has been adapted in order to deal with rear conditions. We use this system in combination with flat-world assumption, fixed camera pitch and camera height, so that, the search space is drastically restricted.

2.2 Candidates Selection

Candidates are selected in a three-stage process. Firstly, the vehicle contact point is searched by means of the white top-hat transformation. This operator allows the detection of contrasted objects on non-uniform backgrounds [2]. In our case it enhances the boundary between the vehicles and the road. Horizontal contact points are pre-selected if the number of white top-hat features is greater than a threshold. This process is applied from bottom to top for each detected lane. Then, candidates are pre-selected if the entropy of Canny points is high enough for a region defined by means of perspective constraints and prior knowledge of target objects (see Figure 2).

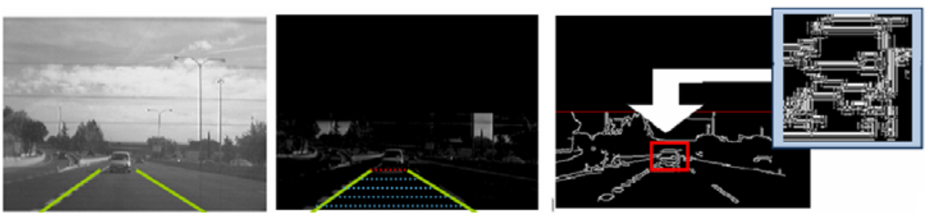

Fig. 2. From left to right: original image; contact point detection on white top-hat image; candidate pre-selected with high entropy of canny points

In a second step, gray level, vertical edges and horizontal edges symmetries are obtained, so that, candidates will only pass to the next stage if their symmetries values are greater than a threshold (see Figure 3). Symmetry axis are linearly combined to obtain the final position of the candidate. Finally, a weighted variable is defined as a function of the entropy of Canny points, the three symmetry values and the distance to the host vehicle. We use this variable to apply a non-maximum suppression process per lane which removes overlapped candidates. An example of this process is depicted in Figure 4.

2.3 Single-Frame Classification

The selected candidates are classified by means of a SVM classifier with RBF kernel, in combination with HOG features [3]. All candidates are resized to a

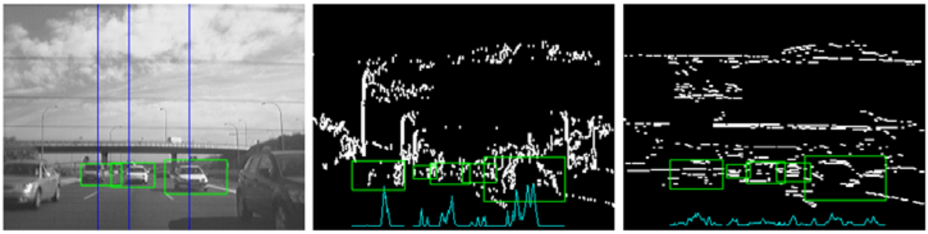

Fig. 3. From left to right: gray-level symmetries, vertical edges symmetries and horizontal edges symmetries

Fig. 4. Left: overlapped candidates. Right: non-maximum suppression results.

fixed size of 64x64 pixels to facilitate the features extraction process. The SVM classifier is trained with 2000 samples and tested with 1000 samples (1/1 positive/negative ratio). Figure 5 depicts some positives and negatives examples of the training and test data sets. The distance to the hyperplane is defined for a detection rate (DR) of 92% and a false positve rate (FPR) of 32%. We have to note that these numbers are defined in a single-frame fashion, so that, they will be improved in subsequent stages.

Fig. 5. Upper row: positive samples (vehicles). Lower row: negative samples.

2.4 Multi-frame Validation and Tracking

The position of the vehicles in the image is tracked by using a Kalman Filter. Data association problem is solved by means of a linear combination of the

Fig. 6. Left: non-rigid grid template. Right: tracking results.

Euclidean distance and the Zero mean Normalized Cross Correlation (ZNCC). Once a vehicle is detected and tracked during a consecutive number of frames, SVM classification is stopped and a non-rigid grid-based matching technique is used until the vehicle disappears from the scene (see Figure 6). Global matching usually fails with close vehicles in lateral lanes. This approach allows to achieve a considerable reduction in the computational cost.

3 Results

The algorithm was implemented on a PC on-board a real automobile. Three different test sequences have been recorded in real traffic conditions with a total duration of 240sec and a traffic density of 1.5 vehicles/frame on average. As we can see in Table 1 the system achieves a detection rate of 92.2% with 1 false positive per minute on average.

Table 1. Overall results

	Duration (sec)	**Number of vehicles**	**Detention Rate (%)**	**False Positives per minute**	**Vehicle density (per frame)**
Sequence 1	115	16	87.5%	1.56	1.58
Sequence 2	65	15	93.33%	0.92	1.778
Sequence 3	54	2	100.0%	0.0	0.645

Regarding the computational cost aspects, in Table 2 we can see that candidates selection and single-frame classification stages are the most time consuming parts of the system. Once the system has a certain knowledge of a candidate of being a vehicle, the proposed approach starts to run really fast since the non-rigid grid-based tracking stage has a low computational cost (less than 3ms). The overall system runs in real-time (23 fps).

The output of the system in a real experiment is depicted on the left side of Figure 7. The distance of each rear-vehicle with regard to the camera is showed on the upper-right corner of the bounding boxes. On the right side of Figure 7

Table 2. Computational cost of the different parts of the system using a Pentium IV 2.8 GHz with 512MB of RAM. The system runs at 23 fps.

	Rear LDW	**Candidates selection**	**Single-frame classification**	**Multi-frame tracking**	**Non-rigid grid-based tracking**
Comp. cost (ms)	5	17.74	16.2	1.9	2.7
Percentage	11.4%	40.7%	37.2%	4.4%	6.2%

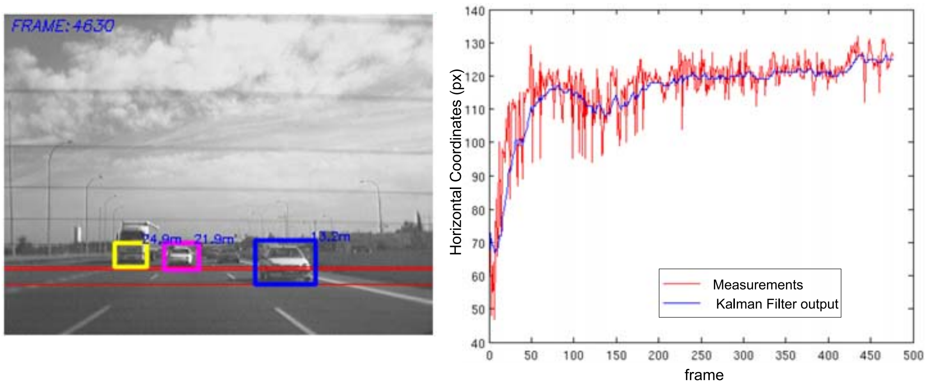

Fig. 7. Left: detected vehicles on a test sequence. Right: filtered y-position of the car located in the middle lane.

the measurement of the y-position in pixel coordinates and its corresponding filtered value for the car located in the middle lane are provided.

At present, the output of the rear-vehicle detection system is being used in combination with Blind Spot Detection (BSD) system developed by the authors [4], resulting in the the so-called Panoramic BSD (see Figure 8).

Fig. 8. Panoramic Blind Spot Detection system (Panoramic BSD)

4 Conclusions and Future Works

This paper presented a real-time vision-based system that detects vehicles approaching from the rear in order to anticipate possible rear-end collisions. The search space is drastically reduced thanks to the rear LDW system which automatically detects the lanes of the road. Candidates are robustly selected using top-hat features and entropy of Canny points in combination with gray-level, horizontal edges and vertical edges symmetries. A single-frame SVM classifier is trained and used with HOG features. This step rejects a considerable amount of false detections. Vehicles are then tracked by using a Kalman filter. Once a vehicle has been classified and tracked during a considerable number of frames, the system performs a non-rigid grid-based matching which decreases the overall computational cost.

The algorithm was implemented on a PC on-board a real automobile. In experiments on datasets captured from a moving vehicle in real traffic conditions, the system achieves a detection rate of 92.2% with 1 false positive per minute on average. Future work involves the detection of vehicles in nighttime conditions by incorporating an infrared camera.

Acknowledgments

This work has been funded by Research Project CCG08-UAH/TIC-3572 (University of Alcalá and CAM).

References

1. Sotelo, M.A., Nuevo, J., Bergasa, L.M., Ocaña, M., Parra, I., Fernández, D.: Road Vehicle Recognition in Monocular Images. In: Proceedings of IEEE ISIE, pp. 1471–1476 (2005)
2. Derong, Y., Yuanyuan, Z., Dongguo, L.: Fast Computation of Multiscale Morphological Operations for Local Contrast Enhancement. In: Proceedings of Engineering in Medicine and Biology 27th Annual Conference (2005)
3. Dalal, N., Triggs, B.: Histograms of Oriented Gradients for Human Detection. In: Proceedings of IEEE CVPR, pp. 886–893 (2005)
4. Sotelo, M.A., Barriga, J., Fernández, D., Parra, I., Naranjo, J.E., Marrón, M., Álvarez, S., Gavilán, M.: Vision-Based Blind Spot Detection Using Optical Flow. In: Moreno Díaz, R., Pichler, F., Quesada Arencibia, A. (eds.) EUROCAST 2007. LNCS, vol. 4739, pp. 1113–1118. Springer, Heidelberg (2007)