Pillars Detection for Side Viewed Vehicles

Raluca Brehar
Computer Science Department
Technical University Of Cluj-Napoca
Email: Raluca.Brehar@cs.utcluj.ro

Sergiu Nedevschi
Computer Science Department
Technical University of Cluj-Napoca
Email: Sergiu.Nedevschi@cs.utcluj.ro

Lorena Dăian
Computer Science Department
Technical University of Cluj-Napoca
Email: LorenaIulia@yahoo.com

***Abstract*—Detecting the parts of a vehicle represents a topic of major interest for computer vision applications, especially for precrash systems. This paper proposes an artificial vision based technique that identifies the pillars of the lateral viewed cars. The novelty of the approach resides in the multi-layer classification scheme applied within the context of a stereo-based object detection system. From all the objects deetected by stereovision the side viewed cars are recognized, and for them the pillars are identified. This process of pillar identification is the result of a multi-layer classification that comprises: a rough object hypothesis refinement that selects only those objects that are likely to have one or two wheels, followed by an adaptive boosting classifier build using histograms of oriented gradient features. The boosted classifier realizes a fine selection of the wheel-based hypotheses and discriminates between side viewed vehicles and other objects in a traffic scene. The last step consists in the construction of a geometrical model of the pillars' region of interest for the identified side vehicles.**

I. Introduction

The part of the car in which the driver and passengers sit is bordered by the pillars of the vehicle. Pillars represent the vertical supports of the greenhouse of a closed automobile body. A car can have from two to four pillars, depending on the type of car. The four pillars are named A, B, C, and D-pillars. All the pillars form the greenhouse of the car, which is the point of interest of the current paper. Figure 1 presents the A, B and C pillars from the greenhouse of a car. This paper

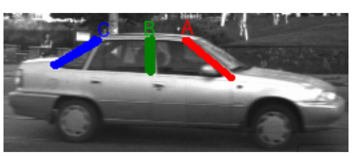

Fig. 1. The pillars of a car

presents an approach for identifying the A and C/D pillars of side viewed sedans or coupes cars in a traffic environment. We have used a stereo system [1] that provides 3D and 2D information for constructing the object assumption. The 3D object hypothesis is projected onto the image plane generating a 2D object hypothesis represented by a 2D image window. A first analysis is done on the motion direction, speed and aspect ratio of the obstacle. We have identified some initial constrains for an obstacle to possible represent a side viewed car: motion direction relatively perpendicular to the ego motion direction, a given aspect ratio (height usually smaller than the width). Also, to reinforce these assumptions we apply a circular symmetry algorithm to the lower half of the 2D image window in order to identify the possible presence of wheels for the given obstacle. This step performs a rough classification that has the role of eliminating most objects that are not likely to be side vehicles.

To further confirm the side vehicle assumption we apply a more powerful classification to the objects selected by the wheel based analysis. The 2D image window is scaned by a boosted classifier build upon histogram of oriented gradient features. The role of this classifier is to perform a fine selection of the detected objects and recognize the ones which have a high probability of being side cars. Next, only on those objects we build a geometric model for the greenhouse, considering the position of the wheels and the diagonal edges in the upper half of the 2D image window.

II. Related Work

Our work relates mainly to recognizing the vehicles and their parts based on the information provided by stereo sensors and is applicable to precrash applications.

The existing approaches in vehicle recognition are based on shape detectors [2] and pattern matching [3] or component based recognition. The system proposed by [3] is based on recognising rigid structure samples, such as the number plate, obtained using specific feature extraction techniques from an image of the vehicle. Feature vectors are classified using simple nearest neighbour classification. A novel set of image strip features is proposed by [4] to describe the appearances of vehicles. The features represent various types of lines and arcs with edge-like and ridge-like strip patterns, which significantly enrich the simple features such as haar-like features and edgelet features. A multivehicle detection system based on stereo vision has been developed by [5]. This system utilizes morphological filters, feature detection, template matching, and epipolar constraint techniques in order to detect vehicles.

III. Method Description

The proposed approach is formed of a multi-layer classification scheme that performs a coarse to fine classification of the side car objects, and for the identified side cars it applies a geometric-based recognition model that identifies the pillars. The classification is done on several layers and each has the role of refining the side car object assumption, such that, the goemetric based recognition of the pillars will be applied only to those objects that have been recognized as being side viewed cars by the previous layers of the classification scheme. Figure 2 depicts the whole process.

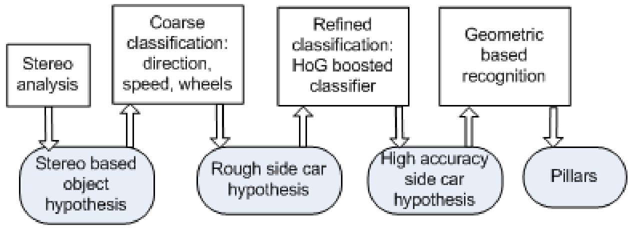

Fig. 2. Overview of method for pillars detection of side cars

A. *Coarse classification module*

The first level of the classification scheme applies fast and simple criteria to all object hypotheses that come from the stereo module. These criteria are:

- Movement direction: if the objects are in motion then check if the movement direction is perpedicular to the direction of the ego-vehicle.
- Presence of wheels: apply a circular symmetry detection algorithm in the lower half of the objects. If circular shapes are encountered, than they are likely to represent the wheels of the side viewed car.

The speed and the direction of the objects is provided by the stereovision system that generates the 3D objects' points and the corresponding 2D window.

A circular symmetry algorithm is applied to the 2D window. The algorithm extracts points of high radial symmetry, such as wheels, and has been successfully used to detect circles of different radii in images where noise was present or the circles were not so clearly seen. We give a brief overview of the algorithm and further details can be found in [6]. Being given a radius $n \in N$, the value of the fast radial symmetry transform indicates the contribution to radial symmetry of the gradients at a distance n away from each point. For each radius n an orientation projection image O_n and a magnitude projection image M_n are formed. These images are generated by examining the gradient g at each point p from which a corresponding positively-affected pixel $p_{+ve}(p)$ and negatively affected pixel $p_{+ve}(p)$ are determined. Their coordinates are:

$$p_{+ve}(p) = p + round(\frac{g(p)}{\|g(p)\|}n) \quad p_{-ve}(p) = p - round(\frac{g(p)}{\|g(p)\|}n)$$

The orientation and projection images are initially zero and they are updated using the following rules:

$$O_n(p_{+ve}) = O_n(p_{+ve}) + 1 \qquad O_n(p_{-ve}) = O_n(p_{-ve}) - 1$$
$$M_n(p_{+ve}) = M_n(p_{+ve}) + \|g(p)\| \quad M_n(p_{-ve}) = M_n(p_{-ve}) - \|g(p)\|$$

The radial symmetry contribution at a radius n is defined as the convolution: $S_n = F_n * A_n$. Where: $F_n(p) = \|\tilde{O}_n(p)\|^\alpha * \tilde{M}_n(p)$ $\tilde{O}_n(p)$ and $\tilde{M}_n(p)$ are obtained from O_n and M_n by dividing each of the element in the two matrices with the corresponding maximum of the absolute value and α is the radial strictness parameter and A_n is a two dimensional Gaussian. The full transform is defined as the sum of the symmetry contributions over all the considered radii: $S = \sum_{n \in N} S_n$.

B. *Side car classification module*

An adaptive boosted classifier [7] trained on histogram of gradient orientations features has been used in the next layer of the classification scheme.

1) Dataset and methodology : We have considered images of side cars having all the same orientation. Samples from the database are shown in Figure 3. The positive dataset was

Fig. 3. Samples from the positive and negative sets.

formed of images that we have manually croped from traffic scenes and images from the UIUC dataset [8], [9]. All the images were scaled to the dimension 128×48. The set of negatives was formed of patches cropped from images of traffic scenes that did not contain side cars. For the positive training set we have used sedans and coupes. The method we propose is appropriate for these models of cars. The positive and negative sets have been divided into train and test. In the train set we have used 4000 positive images of side cars and 100000 negative images. The test set has been formed of 1000 positive image and 200000 negatives.

2) Histogram of Oriented Gradients: Histogram of Oriented Gradients turned out to have good results for object classification [10]. For each point of an image I in the dataset we have computed the gradient magnitude, M and orientation θ. Next, the image is divided into non-overlapping rectangular cells of equal dimension. For each cell a weighted histogram of gradient orientations is computed. In each pixel location the orientation gives to the histogram a vote weighted by the gradient magnitude at the position of the respective pixel. The orientation bins are evenly spaced over 0°- 360°. The last step for HOG descriptors extraction is represented by normalization. The cells are grouped into overlapping blocks of various dimensions. We use L2-Hys normalization algorithm applied on image blocks. The final vector of descriptors will contain all the components of the normalized cell responses from all of the blocks in the detection window.

3) Adaboost classification: Using the images from the dataset and the extracted HoG features, we have trained an adaptive boosting classifier. Adaptive Boosting is a learning technique that has been used for object recognition providing good results [7]. The classifier is a linear combination of weak learners arranged in cascade. Each classifier in the cascade is a thresholded sum of weighted features and each feature is a thresholded sum of HoG features.

C. *Geometric based pillar recognition*

If the HoG based boosted classifier returns a high probability that the object is a side-car then a geometric model is applied to that object in order to find the pillars. Three factors influence the appearence of the model: (1) the position of the wheels; (2) the diagonal edges that define the A and C or D pillars; (3) the 3D points in the range image.

The greenhouse is defined by the A and C or D pillars, depending on the type of car. Our purpose is to find the lateral pillars, the ones that define the extremities of the greenhouse for sedans and coupes.

Figure 4 shows the steps used for defining the geometric model and the search areas for the lateral pillars.

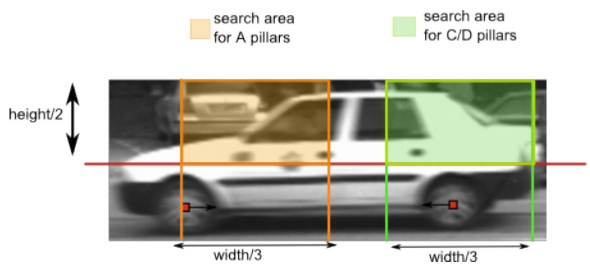

Fig. 4. Geometric model - defines the search area for diagonal edges

The steps of the algorithm are:

1) Consider the radii and the coordinates of the centers of the front and back wheels, $r_f, (x_f, y_f)$ and $r_b, (x_b, y_b)$.
2) Define the region of interest of the A pillar being a rectangle having top left coordinates: $(x_f, 0)$, $width = \frac{object\_width}{3}$, $height = \frac{object\_height}{2}$
3) Define the region of interest of the C/D pillars being a rectangle having top left coordinates equal to: $(x_f - \frac{object\_width}{3} \times \frac{1}{2}, 0)$, $width = \frac{object\_width}{3}$ and $height = \frac{object\_height}{2}$

In all the equations we have considered as $(0, 0)$ the top left corner of the 2D image window associated to the object. After the two regions of interest, for the A pillar and the C/D pillars are defined we extract all diagonal edges in those regions, and we use the range image as a mask for diagonal edge points.

IV. Experimental Setup and Results

Several parameters should be tuned during our experiments, and we will present the values for which the best results have been obtained.

A. Orientation of the side-car

For the direction analysis we have considered all the objects having a direction angle in the range $[80°, \ldots 100°]$ or $[-80°, \ldots -100°]$. The angle is defined as the angle formed by the direction of the ego car and the motion direction of the object.

B. Wheels detection parameters

For wheels detection we have used $\alpha = 2$. In order to compute the posible values of the radii we have used the distance of the side car with respect to the ego-car (provided by stereovision) and knowning an average value in centimeters for the diameter of a wheel, we have computed the diameter in pixels. Further processing was applied to the result of the free radial symmetry transform by applying a non-maxima supression and some morphologic operations that elliminated small groups of pixels, and kept only the ones having a high symmetry. Figure 5 presents the results. A segmentation

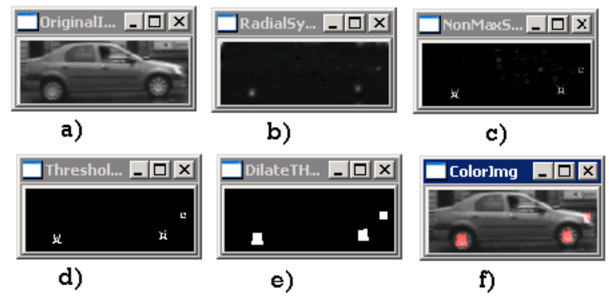

Fig. 5. Results of free radial symmetry algorithm: a) the original image b) the result of free radial symmetry c) nonmaxima suprassion d) thresholded non-maxima supression e) dilation and erosion f) final result, possible position of the wheels

operation is applied to the binary image that results after the morphologic operations. For each region in the segmented image we compute the center of gravity and the area. If more than two regions are present we consider the ones whose centers have aproximately the same coordinates on the y-axis and also have the largest areas.

C. Histogram of oriented gradients parameters

The parameters we have used for training the HoG based classifier are the following: cell size: 8×8 pixels ; block size: 16×16 pixels; block stride for block overlapping displacement: 8×8 pixels; number of bins in the histogram: 9. For the images in the database of side cars that have a dimension of 128×48 pixels the total number of features is 2870.

D. Boosted classifcation

The classifier was trained using decision trees as weak learners. The cascade is formed of ten weak learners. The accuracy of the classifier on the test set that contains 1042 positives and 200000 negatives is provided in Table I:

TABLE I
HoG Boosted Classifier Results on Test Dataset

Number of positives	Number of negatives	Correct predictions positives	Correct predictions negatives
1042	200000	936 = 90%	198922 = 99%

E. Geometric Model for the pillars

After an object has been identified as being a side-car the search areas (or regions of interest) for the A pillar and C/D pillars are defined. Within those areas the diagonal edge points are searched for.

- For the A pillar (corresponding to the front of the greenhouse) we have considered as diagonal edge point all the points that have an orientation in the range: $[22.5°, 67.5°]$, or $[202.5°, 247.5°]$.

- For the C/D pillar (corresponding to the rear of the greenhouse) we have considered as diagonal edge point all the points that have an orientation in the range: $[112.5°, 157.5°]$, or $[292.5°, 337.5°]$.

A non-maxima supression is applied to the resulting diagonal edges, and it is followed by a processing with a morphological operator aimed at enhancing groups of pixels while discarding single pixels. The results contain more diagonal edge points, not just the ones belonging to the A and C/D pillars (for example they can include points that are in the background). In order to elliminate those points and find the exact location of the pillars we use the 3D points associated to the 2D image window for an object. All the 3D points are used as a mask, and only the diagonal edge points that are within this mask are kept. A final edge linking and grouping step is applied to the result. We keep the longest segments, situated at the extremities of the range window. The result is show in Figure 6. The measurements taken on different video sequences have proved the method works in real time. We have tested the pillars extraction algorithm on traffic scenarios that contained overall about 2000 side viewed cars. The accuracy of the method is about 85%. Further adjustments of the parameters used can improve the precision of the greenhouse recognition. The method fails when the position of the wheels is not detected correctly and when the reconstruction of the 3D points is not so accurate, i.e. we have too few points.

(a) Range image

(b) Candidate points for A pillars

(c) Candidate points for C/D pillars

(d) Intersection with range image

(e) Edge linking

Fig. 6.

V. Conclusions and Future Work

A method for extracting the lateral pillars of a side-viewed car in a traffic environment has been described. The proposed method has direct applications in colision avoidance systems. For example for imminent side crashes the ego car can be guided to hit the other car outside the passenger area, which is delimited by the lateral pillars. Hence, getting the location of the pillars in real time would be extremely important. The originality of the method resides in the multi-layer recognition model that implies a stratified classification formed of 1)wheels detector 2) boosted classifier and 3) geometric model for the pillars. As future work we propose to adapt the algorithm for partially side viewed cars (vehicles whose motion direction is not perpendicular to the direction of the ego-car but may have an absolute value of the orientation in the interval $60° - 90°$) and also experiment with other classificatin algorithms that may provide a higher accuracy than the boosted classifier.

References

[1] S. Nedevschi, R. Danescu, T. Marita, F. Oniga, C. Pocol, S. Sobol, C. Tomiuc, C. Vancea, M. M. Meinecke, T. Graf, T. B. To, and M. A. Obojski, "A sensor for urban driving assistance systems based on dense stereovision," in *Proceedings of Intelligent Vehicles*, June 2007, pp. 278–286.

[2] J. Thureson and S.Carlsson, "Finding object categories in cluttered images using minimal shape prototypes," *Proceedings of 13th Scandinavian Conference on Image Analysis SCIA, Goteborg, Sweden*, 2003.

[3] V. Petrovic and T. Cootes, "Analysis of features for rigid structure vehicle type recognition," *Proceedings of BritishMachine Vision Conf.*, pp. 587–596, 2004.

[4] W. Zheng and L. Liang, "Fast car detection using image strip features," *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, pp. 2703–2710, 2009.

[5] J. Hwang, K. Huh, and D. Lee, "Vision-based vehicle detection and tracking algorithm design," *Optical Engineering*, pp. 127–201, 2009.

[6] G. Loy and A. Zelinsky, "A fast radial symmetry transform for detecting points of interest," in *ECCV '02: Proceedings of the 7th European Conference on Computer Vision-Part I*. London, UK: Springer-Verlag, 2002, pp. 358–368.

[7] D. Snow, P.Viola, and M. Jones, "Detecting pedestrians using patterns of motion and appearance," in *International Journal of Comupter Vision*, vol. 63, 2005, pp. 153–161.

[8] S. Agarwal and D. Roth, "Learning a sparse representation for object detection," in *Proceedings of the European Conference on Computer Vision*, vol. 4. Copenhagen, Denmark: Springer-Verlag, May 2002, pp. 113–130.

[9] S. Agarwal, A. Awan, and D. Roth, "Learning to detect objects in images via a sparse, part-based representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 11, pp. 1475–1490, November 2004.

[10] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005.*, June 2005, pp. 886–893.