ROBUST VEHICLE DETECTION THROUGH MULTIDIMENSIONAL CLASSIFICATION FOR ON BOARD VIDEO BASED SYSTEMS

Daniel Alonso, Luis Salgado, and Marcos Nieto

Grupo de Tratamiento de Imágenes – E.T.S. Ing. Telecomunicación
Universidad Politécnica de Madrid – Madrid – Spain
{dar, lsa, mnd}@gti.ssr.upm.es

ABSTRACT

This paper presents a new in-vehicle real-time vehicle detection strategy which hypothesizes the presence of vehicles in rectangular sub-regions based on the robust classification of features vectors result of a combination of multiple morphological vehicle features. One vector is extracted for each region of the image likely containing vehicles as a multidimensional likelihood measure with respect to a simplified vehicle model. A supervised training phase set the representative vectors of the classes *vehicle* and *non-vehicle*, so that the hypothesis is verified or not according to the Mahalanobis distance between the feature vector and the representative vectors. Excellent results have been obtained in several video sequences accurately detecting vehicles with very different aspect-ratio, color, size, etc, while minimizing the number of missing detections and false alarms.

***Index Terms*—** Vehicle detection, morphological feature extraction, Mahalanobis distance

1. INTRODUCTION

Automatic vehicle detection is one of the most challenging research fields within video-based driver assistance systems. The main objective is to provide reliable and real-time detections of the vehicles ahead based on the analysis of the images captured from in-vehicle video cameras. Major difficulties arise as the mobile environment generates images usually containing cluttered background consisting of buildings, trees and road signs.

Due to the real time processing constraint, only those parts of the images likely showing vehicles are computed. For that purpose, most of works reported in the literature follow two main phases [1]: (i) Generation of hypotheses, in charge to select those image regions likely to hold vehicles; and (ii) Verification of hypotheses, whose objective is to verify the presence of vehicles, among the selected areas. The hypothesis generation has been solved in the literature by using different approaches [1], such as knowledge-based, stereo-based, and motion-based strategies. On the other hand, the verification of these hypotheses is still a widely open field of research, with significant contributions which can be classified into two main categories: (i) template-based methods and (ii) appearance-based methods.

Template-based methods extract morphological characteristics to be compared with a predefined model of a vehicle [2]. These methods main limitation is that they make no decision about whether the hypotheses created correspond to a real vehicle or not, and they do not give any confidence measure of the detection.

Appearance-based methods are trained with a set of images to describe the variability of vehicle appearance in order to achieve an adequate classification, based on spatial features [3] or frequency information [4]. Although these methods allow providing confidence measures for the detections, they show a lack of robustness due to the fact that features used for classification are not directly linked to vehicles morphological structure.

In this paper, a new vehicle detection strategy is presented designed to overcome the abovementioned limitations. The proposed approach performs an efficient combination of selected aspects of template-based and appearance-based methods, within the verification phase, to reach simpler and more robust results. The proposed strategy is based on an efficient computation of heterogeneous model likelihood measures which drive a classification process to decide which hypotheses most probably correspond to vehicles.

The generation phase of the proposed strategy selects regions of interest (ROI) through an adaptive edge-based split-and-merge segmentation, and for each ROI it proposes candidate sub-regions which are classified in the verification phase. For that purpose, a Mahalanobis minimum distance classification [5] is applied in a multidimensional feature space. In this work, the selected features are the likelihood measures with respect to a simplified vehicle model of three morphological characteristics of the candidate sub-region: shadows, symmetry and corners. This way, the classification result identifies each candidate as belonging to the class *vehicle* or *non-vehicle* while providing a confidence measure.

1-4244-1437-7/07/$20.00 ©2007 IEEE

2. SYSTEM OVERVIEW

A block diagram of the proposed vehicle detection approach is presented in Fig. 1. The hypotheses generation phase takes each image, *I(n)*, of the sequence, and compute the edge image, *E(n)*, which is used to automatically define a set of ROIs, R_i, through an split and merge segmentation strategy. Candidate sub-regions, the so-called hypothesis, H_i^j, are generated by analyzing the lateral histograms of each R_i. Each hypothesis is then represented by a model likelihood feature vector, $\mathbf{f} = (f_0, f_1, f_2)$, whose components are the normalized measures derived from the different morphological characteristics considered: shadows, symmetry and corners. Hypotheses are classified according to their likelihood vector using a simple Mahalanobis minimum distance classifier [5]. Finally an image with the detected vehicles, *DV(n),* is generated together with the associated confidence measures.

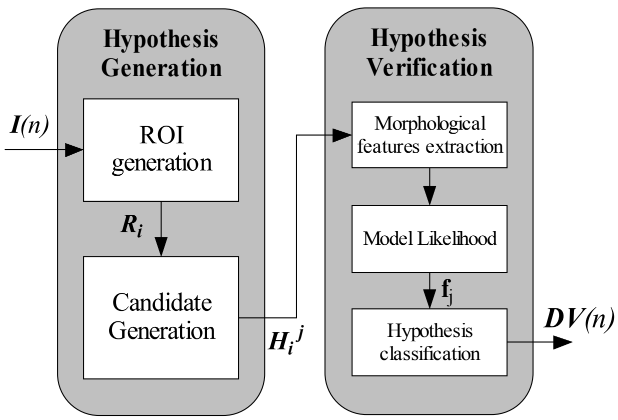


Fig. 1: Vehicle detection strategy block diagram

3. HYPOTHESIS GENERATION STRATEGY

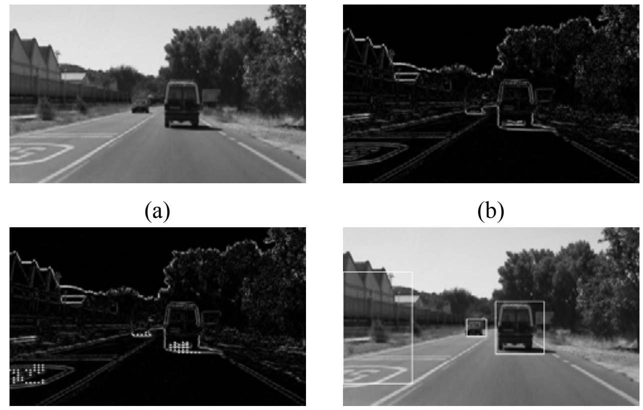
It is composed by two processing phases: (i) selection of regions of interest, and (ii) identification of candidate sub-regions likely containing vehicles to be further analyzed in the hypothesis verification module.

3.1. Regions of interest selection

An edges image, *E(n)*, is computed as in (1) so as to enhance vertical and horizontal edges, which are supposed to appear mostly in regions of the image containing vehicles:

$$E(n) = \left| E_h(n) - E_v(n) \right| \tag{1}$$

where $E_h(n)$ and $E_v(n)$ represent the horizontal and vertical gradient images, computed with the Sobel filter [5]. This way, a horizontal edge obtains high values in $E_h(n)$, and very low ones in $E_v(n)$ so that their absolute difference is high; analogously, vertical edges obtain high values in *E(n)*. On the other hand, edges likely belonging to lane markings, trees, shadows, and other irregular shapes, will obtain lower values in *E(n)*. An example of this edges image is shown in Fig. 2 (b), where Fig. 2 (a) shows the original image.



(a) (b) (c) (d)

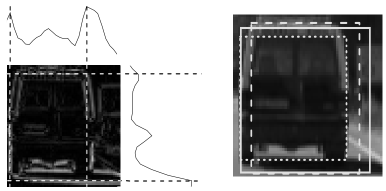
Fig. 2: Selection of regions of interest. (a) original image, (b) edges image, (c) high edge density areas image, (d) ROI image.

The ROIs definition is the result of the application of a modified split and merge segmentation strategy on *E(n)* through the use of an adaptive threshold selection. Splitting is done down to 4x4 pixel blocks. At each iteration *k*, a block is split if the sum of edge intensity values inside it, s_k, exceeds a threshold. This threshold is automatically obtained from the previous split step, as the minimum of the mean and the median of the values s_{k-1}. This automatically adaptive threshold selection strategy ensures that split down to 4x4 pixels blocks is only raised on those regions showing the most significant values of *E(n)*.

Merging is performed grouping those 8-connected 4x4 blocks. As some of these regions may not fully include the vehicle, their width and height are resized up to reach an adequate aspect ratio according to the expected vehicle dimensions. Fig. 2 (c) shows in white the 4x4 blocks in which the higher intensity edges are found; those mainly lying on the vehicle bottom parts. Fig. 2 (d) shows the final R_i obtained after merging and the resizing process. As it can be observed, some R_i contain vehicles, while others do not.

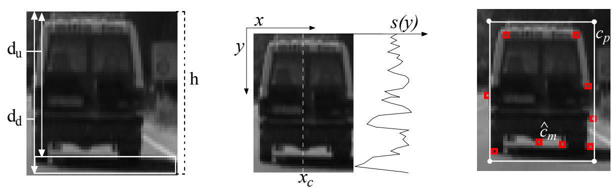
3.2. Candidate sub-regions identification

For a ROI R_i, to select the set of candidate sub-regions H_i^j to be further analyzed by the verification phase, the lateral histograms are computed on *E(n)*. Fig. 3 (a) shows the lateral histograms of an example R_i. As it is shown, the peaks of the histograms determine the position of the strongest horizontal and vertical edges. Pairs of peaks are linked to determine the bounding box of each candidate, setting minimum and maximum distance between peaks based on the assumed width and height of a vehicle given the perspective effect within the image. The resulting candidate sub-regions are presented in Fig. 3 (b).



(a) (b)

Fig. 3: ROI lateral histogram analysis and hypotheses extraction.



(a) (b) (c)

Fig. 4: Model likelihood measures construction.

4. HYPOTHESIS VERIFICATION STRATEGY

A verification algorithm is applied to differentiate those hypotheses that correspond to vehicles from those which not. Three heterogeneous morphological features are extracted for each H_i^j : shadows, symmetry and corners. The obtained features are compared with a simplified vehicle model to obtain likelihood measures. The model used as vehicles theoretical appearance is a rectangular symmetric box with an underneath shadow. This way it is possible to apply a classification system which allows to efficiently segment vehicles from the rest of the image.

4.1. Morphological Characteristics Extraction

4.1.1. Shadows

Shadows are obtained based on vertical profiles analysis [7]. Each image vertical line is scanned bottom-up looking for grey value transitions from road to vehicle shadows (expected darker than road gray values). The shadow model likelihood measure is constructed as in (2):

$$f_0 = \frac{d_u}{h-(d_d-d_u)} \tag{2}$$

where d_u is the vertical distance of the top boundary of the detected shadow to the top boundary of the candidate area, d_d is the vertical distance of the bottom boundary of the detected shadow to the top boundary of the candidate area and h corresponds to the height of the candidate area, and therefore, $f_o \in [0,1]$. Fig. 4 (a) depicts this scheme.

4.1.1. Symmetry

The symmetry algorithm is based on the one proposed in [8]. The likelihood value f_1 is obtained as the sum of the symmetry values, *s(y)*, computed for each pixel belonging to the central column, x_c, of the hypothesis region H_i^j:

$$f_1 = \sum_{y=0}^{h} s(y) \tag{3}$$

where h is the height of the rectangular hypothesis region, and *s(y)* is compute as in (4).

$$s(y) = \frac{\int_0^w H_e^2(x,y)dx - \int_0^w H_o^2(x,y)dx}{\int_0^w H_e^2(x,y)dx + \int_0^w H_o^2(x,y)dx} \tag{4}$$

In these equations H_e and H_o represent, respectively, the normalized even and odd part of the hypothesis H_i^j, obtained as in (5), while w is the width of H_i^j.

$$\begin{aligned} H_e(x,y) &= \frac{H(x,y)+H(-x,y)}{2} \\ H_o(x,y) &= \frac{H(x,y)-H(-x,y)}{2} \end{aligned} \tag{5}$$

Fig. 4 (b) shows the *s(y)* of an example H_i^j.

4.1.1.Corners

Corners within a hypothesis region are detected based on the Harris detector [9]. The most significant ones are expected to correspond to a vehicle, assuming it is present in the region. Although only four corners are strictly needed to describe a vehicle, robustness against outliers is achieved increasing the number of detected corners. Considering $\mathbf{c}_p$ as the coordinates of the four corners of the rectangular hypothesis region H_i^j, its likelihood measure is given by (6):

$$f_2 = 1 - \frac{1}{4}\sum_{p=0}^{4} D(\mathbf{c}_p) \tag{6}$$

where $D(\mathbf{c}_p)$ is the minimum of the normalized distances between each $\mathbf{c}_p$ and the detected corners $\hat{\mathbf{c}}_m$:

$$D(\mathbf{c}_p) = \min\{d(\mathbf{c}_p, \hat{\mathbf{c}}_m)\}, \forall p, m \tag{7}$$

Fig. 4 (c) depicts the corners $\mathbf{c}_p$ of the bounding box of a hypothesis H_i^j, and the detected corners $\hat{\mathbf{c}}_m$.

4.2. Hypothesis Classification

Once the likelihood vector $\mathbf{f} = (f_0, f_1, f_2)$ has been computed for each hypothesis, $H_i{}^j$, the classification system decides whether the estimated hypothesis is or not a vehicle. Two classes are defined: *vehicle* and *non-vehicle*. A supervised training phase has been applied to compute class representatives as the centroid of each class distribution: $\mathbf{v} = (v_0, v_1, v_2)$ for the *vehicles* class, and $\mathbf{n} = (n_0, n_1, n_2)$ for the *non-vehicles* class.

Considering that the selected features have different autocovariance values and are dependant, i.e. the covariance matrix is not the identity, the most suitable classification strategy is the one provided by the minimization of the Mahalanobis distance [5], as it differs from Euclidean distance in that it takes into account the data correlation.

Fig. 5: Examples of vehicles detected.

Therefore, a hypothesis H_i^j with a computed likelihood vector $\mathbf{f}$, is classified as *vehicle* if $d_M(\mathbf{f}, \mathbf{v}) < d_M(\mathbf{f}, \mathbf{n})$, where d_M is the Mahalanobis distance operator between vectors. Otherwise, $d_M(\mathbf{f}, \mathbf{v}) \geq d_M(\mathbf{f}, \mathbf{n})$, H_i^j is set as *non-vehicle*.

It may happen that within a ROI, more than one hypothesis is verified as *vehicle*. In these cases, the best hypothesis is selected as the one with higher values in the likelihood feature vector $\mathbf{f}$. A confidence measure of the decision made is constructed as the relation between the distance of the candidate model likelihood vector to the vehicle class centroid and the sum of distances to the vehicle and non-vehicle class centroids. This way, candidates whose model likelihood distance to vehicle class centroid is much lower than to the non vehicle centroid will get a high confidence measure.

5. RESULTS

The proposed algorithm has been tested over several real situation images, which have been captured under different weather conditions and roads, such as highways and local roads to verify its robustness, obtaining excellent results.

Fig. 4 shows vehicle detection results in different scenarios. As can be observed, very different vehicles, in terms of size, colour or aspect-ratio are successfully detected, highly accurately delimiting their contour. It is remarkable that the system has demonstrated to get excellent results even on the detection of very distant vehicles and approaching vehicles. Weather conditions are another circumstance that may affects the performance of the system; nevertheless it has obtained very good results in extremely sunny and cloudy weather conditions. The percentage of vehicles correctly detected along the tested sequences is 92,63 %, and only generating a 3,68 % of false alarms. Classification results obtained on the same test sequences but using only one model likelihood feature drop down to an 83,04 % of correct detections with a very significant increase the non-detected vehicles, up to 16,96%, and a 16,06% of false alarms generated.

6. CONCLUSIONS

A new vehicle detection strategy is proposed based on the classification of multidimensional likelihood measures. Robustness and efficiency are achieved considering a combination of three morphological vehicle features: shadows, symmetry and corners. Candidate sub-regions of the image likely containing vehicles are analyzed to obtain a feature vector, which drives to a classification into *vehicle* or *non-vehicle* classes. This process is based on the minimization of the Mahalanobis distance to the representative vectors of the vehicle and non-vehicle classes obtained through a supervised training phase.

Excellent results have been obtained, offering more than 90% of correct detections in different situations, including several types of road and illumination conditions, while few missing detections and false alarms.

7. ACKNOWLEDGEMENTS

This work has been partially supported by the European Commission 6th Framework Program under project IST-2004-027195 (I-WAY). This work is also supported by the Comunidad de Madrid under project P-TIC-0223-0505 (PRO-MULTIDIS).

8. REFERENCES

[1] Z. Sun, G. Bebis and R. Miller, “On Road Vehicle Detection Using Optical Sensors: A Review”, IEEE ITSC’04, Washington D.C., USA, October 3-6, 2004 pp. 585-590.

[2] P. Parodi and G. Piccioli, "A feature-based recognition scheme for traffic scenes," IEEE IV’95. Detroit, MI, USA, Sept. 25-26, 1995, pp. 229-234.

[3] J. Wu and X. Zhang, “A PCA Classifier and its Application in Vehicle Detection”, IEEE CVPR’01. Washington D.C., USA, Jul. 15-19, 2001, pp. 600-604.

[4] Z. Sun, G. Bebis and R. Miller, “Quantized wavelet features and support vector machines for on-road vehicle detection”, IEEE ICARCV’02, Dec, 2-5 2002, pp. 1641-1646.

[5] R.C. González and R.E. Woods, Digital Image Processing using MATLAB, Addison-Wesley, 2001.

[6] M. Betke, E. Haritaoglu and L. S. Davis, “Multiple Vehicle Detection and Tracking in Hard Real-Time”, Int. Vehicles Symposium, Tokyo, Japan, Sept. 19-20, 1996, pp. 351-356.

[7] T. K. ten Kate, *et al*, “Mid-range and Distant Vehicle Detection with a Mobile Camera”, IEEE Intelligent Vehicles Symposium, Jun. 14-17, 2004, pp. 72-77.

[8] T. Zielke, M. Brauckmann, and W. von Seelen, "Intensity and Edge-based Symmetry Detection with an Application to Car-following", Computer Vision, Graphics, and Image Processing: Image Understanding, vol. 58, 1993, pp. 177-190.

[9] C. Harris and M. Stephens, “A Combined Corner and Edge Detector”, Proc. of The Fourth Alvey Vision Conference, Manchester, 1988, pp 147-152.