SVM-based Obstacles Recognition for Road Vehicle Applications

M. A. Sotelo, J. Nuevo, D. Fernandez, I. Parra, L. M. Bergasa, M. Ocana, R. Flores
Department of Electronics
University of Alcala
Campus Universitario s/n Alcala de Henares, Madrid, Spain
michael@depeca.uah.es

Abstract

This paper describes an obstacle Recognition System based on SVM and vision. The basic components of the detected objects are first located in the image and then combined with a SVM-based classifier. A distributed learning approach is proposed in order to better deal with objects variability, illumination conditions, partial occlusions and rotations. A large database containing thousands of object examples extracted from real road images has been created for learning purposes. We present and discuss the results achieved up to date.

1 Introduction

This paper describes an SVM-based object recognition system that can recognise both vehicles and pedestrians using vision. In our approach, the basic components of the objects are first located in the image and then combined with a SVM-based classifier. Our object detection technique is characterised by example-based learning algorithms. The salient features of a class are learnt by the system based on a set of examples. Example-based techniques have been previously used in natural, cluttered environments for pedestrian detection [Shashua, 2004]. In general, these techniques are easy to use with objects composed of distinct identifiable parts arranged in a well-defined configuration. This is the case of road vehicles, where a distributed learning approach based on components [Mohan, 2001] is more efficient for object recognition in real cluttered environments than holistic approaches [Papageorgiou, 2000]. Distributed learning techniques can deal with partial occlusions and are less sensitive to object rotations. The use of SVMs is a viable option as long as we intend to discriminate between two classes: car and non-car.

2 System Description

The system is divided in two modular subsystems. The first subsystem is responsible for vehicle detection and tracking. The second subsystem provides pedestrians detection using the information obtained by the vehicle detection module. In this paper, we focus on the vehicle recognition system alone, working with 320x240 monochrome images. The objects searching space is reduced by using the limits established by the estimated lane markings. This helps reduce the rate of false positive detections. In case that no lane markings are detected, a basic area of interest is used instead covering the front part ahead of the host-vehicle. In a first stage, an attention mechanism has been devised with the intention of filtering out inappropriate candidate windows based on the lack of distinctive features, such as horizontal edges and symmetrical structures, which are essential characteristics of road vehicles. This has the positive effect of decreasing both the total computation time and the rate of false positive detections. Each road lane is sequentially scanned, from the bottom to the horizon line of the image, looking for collections of horizontal edges that might represent a potential obstacle. We use a distributed learning approach in which each individual part of the vehicle is independently learnt by a specialized classifier in a first learning stage. The local parts are then integrated by another classifier in a second learning stage. We have considered a total of 3 different sub-regions for each candidate region, covering the most characteristic parts of the vehicle. Two small sub-regions have been located in the area of the region where the wheels are supposed to be. A third sub-region is located in the central part of the region, covering the area where car plates and rear windshield are usually placed. The locations of the three sub-regions have been chosen in an attempt to detect coherent and structural car features. A set of features must be extracted from each sub-region and fed to the classifier. Before doing that, the entire candidate region of interest is pre-processed using a Canny operator in order to enhance the differential information contained in it (edges). The Canny image provides a good representation of the discriminating features of the car class. On the one hand, edges, both horizontal and vertical, are clearly visible and distinguishable. On the other hand, the vertical symmetry of a car remains unchanged. In addition, edges are not affected by colours or intensity. This property makes the use of edges robust enough to different car models of the same type. The pre-processed sub-region is directly applied to the input of the classifier. The dimensions of the entire region of interest are normalized before being fed to the classifier. A size of 70x80 pixels has been chosen, as it is adequate for detecting vehicles at long distances (up to 80 meters).

The global training strategy is carried out in two stages. In a first stage, separate SVM-based classifiers are trained using individual training sets that represent a subset of a sub-region.

This stage provides classification of individual parts of the candidate sub-regions. In a second step, the outputs of all classifiers are merged in a single SVM classifier in order to provide the final classification result.

3 Results

The system was implemented on a Apple PC at 2.0 GHz running the Debian GNU/Linux Operating System. The complete algorithm runs at 25 frames/s. We created a preliminary database containing 2000 samples of road vehicles. The samples were extracted from recorded images acquired in real experiments onboard a road vehicle in real traffic conditions in Madrid. Two different training sets were built for the same sub-region in different conditions in order to decrease the complexity of the training process. This yields a total of 6 training sets (2x3). All training sets were created at day time conditions using the TsetBuilder [Nuevo, 2005] tool, specifically developed in this wok for this purpose. By using the TsetBuilder tool different candidate regions are manually selected in the image on a frame-by-frame basis. This allows to select candidate regions containing vehicles of different size, from different manufacturers, and so on. The number of non-vehicle samples in the training sets was chosen to be similar to the number of vehicle samples. We obtained a detection rate of 85% in a test set containing 1000 images, and a false detection rate of 5%. No image from the training set was used in the test set. As an example, figure 1 shows a sequence of images in which a vehicle is detected and tracked along the lane of the host vehicle. A blue box is overprinted over the detected vehicle indicating the estimated distance measured from the host vehicle. Other vehicles appearing along the adjoining lane are marked with a horizontal red line.

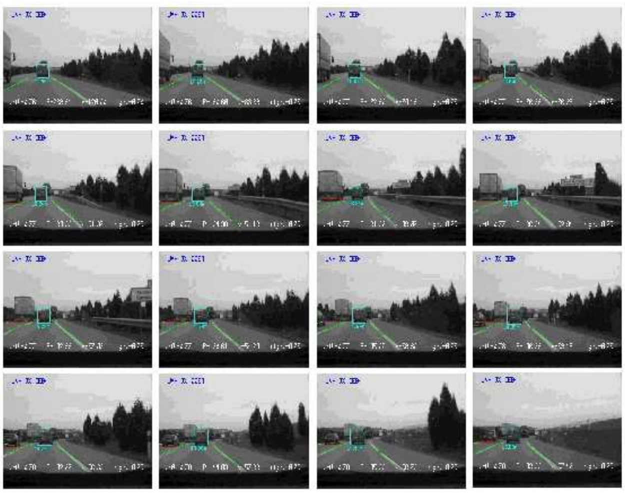

Figure 1: Vehicle detection and tracking in a sequence.

4 Conclusions and Future Work

We have developed a visual multi-frame two-stage object classification system based on Support Vector Machines (SVM) [Boser, 1992]. The learning process has been simplified by decomposing the candidate regions into 3 local sub-regions that are easily learned by individual SVM classifiers. Several training sets have been built for each sub-region in order to cope with different weather and illumination conditions. The results achieved up to date using a set of 2000 samples are encouraging. Nevertheless they still need to be improved before being safely used as an assistance driving system onboard road vehicles in real conditions. For this purpose, the content of the training sets will be largely increased by including new and more complex samples that will boost the classifier performance. In addition, the attention mechanism will be refined in order to provide more candidates around the original candidate region. This will reduce the number of candidate regions that only contain a part of the vehicle, i.e., those cases in which the entire vehicle is not completely visible in the candidate region due to a misdetection of the attention mechanism.

Acknowledgment

This work has been funded by Research Projects CICYT DPI2002-04064-05-04 and FOM2002-002 (Ministerio de Fomento, Spain).

References

[Shashua, 2004] A. Shashua, Y. Gdalyahu, and G. Hayu. Pedestrian detection for driving assistance systems: single-frame classification and system level performance. *Proceedings of the IEEE Intelligent Vehicles Symposium*, pages 1–6, Parma, Italy, June 2004.

[Mohan, 2001] A. Mohan, C. Papageorgiou, and T. Poggio. Example-based object detection in images by components. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23, April 2001.

[Papageorgiou, 2000] C. Papageorgiou and T. Poggio. A trainable system for object detection. *International Journal of Computer Vision*, 38,1:15–33, 2000.

[Boser, 1992] B. Boser, I. Guyon, and V. Vapnik. A training algorithm for optimal margin classifiers. *Proceedings of the Fifth Annual Workshop on Computational Learning*, 1992.

[Nuevo, 2005] J. Nuevo. TsetBuilder User Manual. *Technical Report. University of Alcala. ftp://www.depeca.uah.es/pub/vision/SVM/manual.pdf*, 2005.