Real Time Road Line Extraction with Simple Statistical Descriptors

M. Oliveira, V. Santos, *Member, IEEE*
{mriem, vitor}@ua.pt
Department of Mechanical Engineering, DEM/TEMA, University of Aveiro, 3810-193 Aveiro, Portugal

***Abstract* — Most of the autonomous and assisted navigation systems for road-like environments, relying simply on road perception, require lane or track extraction in order to allow safe motion. The problem may be tackled by extracting the road ground plane, using texture continuity, or by extracting the lane delimiting lines, or a combination of both. This paper devises basic tools to develop a robust technique to correctly extract road lines in real time when navigating in variable conditions. The proposed technique uses image perspective correction for geometric reliability and simple statistical descriptors to validate line candidates. The simple descriptors are used for fast computation. The results include the successful application to robots in indoor road-like competitions and also to real roads using images from an ordinary camcorder.**

I. Introduction

Automatic road perception and navigation in real conditions has been a concern for scientists and researchers, both for the utopia of designing autonomous robots and systems, but also for the sake of safety and human assistance in car driving in variable situations. The extraction of road and lane limits poses variable challenges ranging from very well delimited tracks, with high contrast lines, up to poorly defined contours of the pathway in rural and similar environments.

Being complex, the problem must be decomposed, and robust tools and algorithms must be devised first for the simplest cases and then evolve to more demanding tasks. Also, vision based tools appear currently as the best passive sensing systems and are assumed in this paper as the sole perception mechanisms aboard the moving vehicles. Even the apparently simple high-way road with clear delimiting lines turns out challenging when spurious lines appear in the image or lighting conditions vary abruptly.

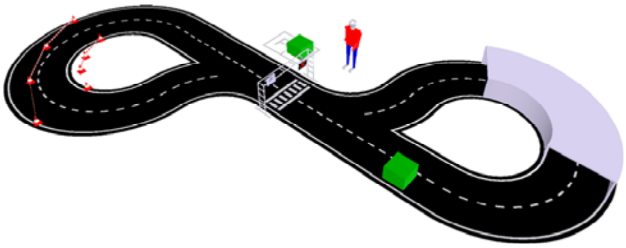

Figure 1 - Autonomous Driving Competition road-like challenge

The concern of the authors for this problem arose when solving a smaller scale problem in the robot competition named *Festival Nacional de Robótica*, the Portuguese Robotics Open [1][2], in the autonomous driving challenge (Figure 1). Within the ATLAS project [3][4], carried out at the Department of Mechanical Engineering at the University of Aveiro, Portugal, some robots for road navigation have been developed since 2003 (Figure 2), and several first and other honorable places in the national competition have been obtained ever since. The methods for lane extraction have relied on line and track detection and are now tested on real roads.

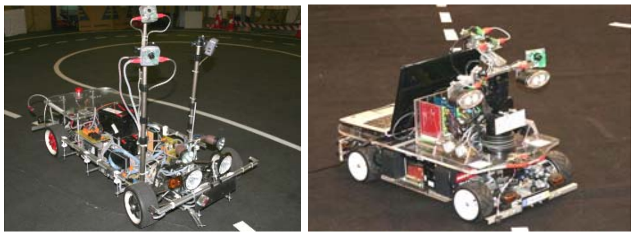

Figure 2 - Examples of ATLAS robots for roal-like navigation

The remainder of the paper presents a brief state of the art in road detection, and then details are given on the method proposed, first in the image preprocessing and, further on, line candidates extraction followed by their validation using simple statistical descriptors. Results on the indoors ATLAS robots and outdoor real road images are given at the end before conclusions and future perspectives.

II. State of the Art

To develop an intelligent transportation system (ITS) several background decisions have to be considered. The first is where to focus the efforts: should the roads infrastructures be purposely designed for robots or should the robots be able to cope with them as they are nowadays. The improvement of roads, in the sense of their adaptation to robotics navigation may pose itself as an interesting alternative. However, building and maintenance costs of an entire road network of such characteristics may become astronomical. Therefore, only small portions of the network can be modified, especially where public transportations with predefined routes make their way, "An ad hoc structuring of the environment can only be considered for a reduced subset of the road network, e.g. for building a fully automated highway in which only automatic vehicles, public or private, can drive", [5]. As a result, the short term implementation of Intelligent Transportation Systems will probably reside in the form of hybrid control (both manual and automatic/semi-automatic). It is common to mount on nowadays ITS several Radar, Laser, GPS and Vision sensors dedicated to the task of perceiving the environment. This trend, despite being more developed nowadays, often involves the usage of active sensors, which may pose a problem regarding the widespread use of such technologies due to sensor interference. Also, the current cost of these sensors and the usage of several per vehicle may prohibit its use on consumer products.

Hence, passive sensors have natural advantages, considering cost and widespread use. Passive sensors, in this context, usually correspond to vision-based sensors. Nowadays, vision-based sensors acquire data typically at

30 to 60 Hz, which is sufficient for ITS applications. Also, considering that it is estimated that humans acquire 90% of the information required for driving visually [5], and that current roads and signalization configuration will probably linger, vision-based sensing assumes a vital role in the development of ITS.

Several approaches have been attempted in order to solve the problem of extracting the road using visual information only. There are two separate threads in what road detection is concerned. Road may be detected based on road color uniformity, or delimiting lines can be looked up.

Some authors try to explore morphological operators in order to extract the road or lines. Liu et. al. make use of Canny and Sobel operators to remove shadows form the image [6], while Zhang et. al. employ Hough Transforms to perform line detection [7].

Some others authors utilize neural networks. Pomerleau uses neural networks that process a low resolution image of a particular area of the road and classify the image according to its similarity to several hypothetical models of road curvature [8]. The low resolution image is extracted from a trapezium, which is vertically positioned according to the current vehicle's speed. The horizontal size of the trapezium is decided on top of perspective transformation considerations, so that particular region of the image can be remapped to a top view of the road.

Foedisch et. al. [9][10][11] also use neural networks to label the image's pixels as "road" or "not road" based on their color similarity to some example pixels taken from sub-windows of the image where road is most likely to be. The positioning of such windows is, obviously, critical to the system's performance. For each analyzed pixel, the neural network's inputs are the pixels RGB color (down-sampled) and the pixels position in the image.

In Italy, Bertozzi and Broggi also use perspective transformation to obtain a bird's eye view of the road. It is performed because "The perspective effect associates different meanings to different image pixels, depending on their position in the image (…). Conversely, after the removal of the perspective effect, each pixel represents the same portion of the road, allowing a homogeneous distribution of the information among all image pixels" [12]. Road lines are afterwards detected on the transformed image, searching for quasi-vertical bright lines of constant width whose pixels are brighter than their neighbors.

In more recent years, authors have made use of state of the art techniques like directional filters, namely Gabor filters. The convolution with these filters extracts several trends of macroscopic features which are then accumulated for different directions [13].

Intel's cooperation with Stanford University has also resulted in some road navigation state of the art techniques. A laser scanner, registered with the camera, is employed to define a polygon that is sure to comprise the nearby road. The color of the pixels inside that polygon is then used as a seed for learning and modeling clusters of road-colored pixels [14] [15].

In the scope of the DIPLODOC project, Lombardi et. al. [16] have developed a model switching architecture that uses pixel color to find the most similar model from a finite database.

The Department of Systems Engineering of the Australian National University has been studying the whole concept of Drivers Assistance Systems (DAS) [17][18]. Regarding road segmentation, they have several different algorithms, i.e., lane tracking, lane keeping, obstacle detection and tracking, gaze monitoring, and others. These algorithms obviously work better in particular conditions, while others fail to do so. Because of this, Zelinski et. al. have developed a visual cue processing framework they called Distillation, whose purpose is to distribute processing time to various visual cues based on the likelihood of each making a more feasible decision. Distillation also integrates the several visual cues and "(…) is able to distill the contributions of various visual cues and test hypotheses into an overall best estimate of the state space distribution" [17]. These authors have noted that the integration of several algorithms clearly enhances the robustness of the DAS, since "the system benefited greatly from the cue fusion and particle filtering technologies used and was shown to perform robustly in a number of situations that are often difficult for lane trackers".

III. Image Pre-Processing

In this work, an alternative line extraction methodology is proposed in order to provide a fast and robust road line detector. In this particular analysis, the input normal view road image (NVI) undergoes a perspective transformation. This transformation generates a new image in a different axis system, placed purposely so that the new transformed image is a bird's eye view of the road, i.e., the bird view image (BVI). A perspective transformation engine has been previously developed by the authors [19]. It makes use of a Denavit-Hartenberg kinematics chain description along with a classic chessboard calibration of the camera's intrinsic parameters to obtain accurate camera registration. Figure 3 displays an NVI and its corresponding BVI.

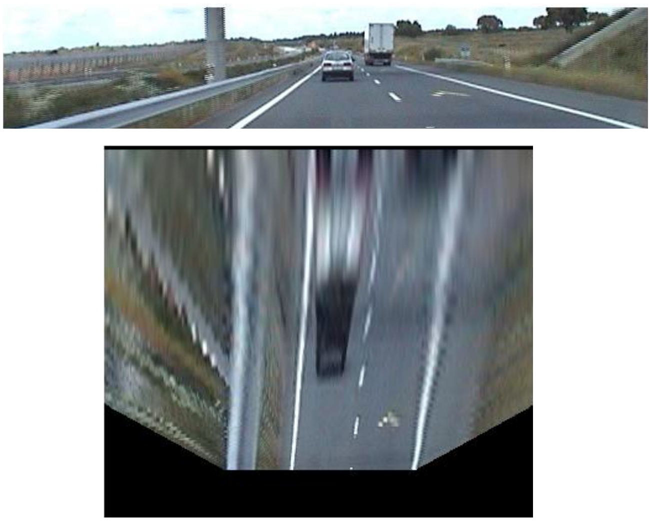

Figure 3 – Normal view image (NVI) and corresponding perspective corrected image (BVI).

After the perspective transformation takes place, the resultant image is a sparse image. Furthermore, the information is not uniformly distributed. Because this is a problem for common image processing algorithms, the sparse image is then interpolated. More details are discussed on [19]. Based on several optimizations, the perspective transformation engine is able to project a 320x240 RBG image and interpolate it in less than 15 milliseconds on a 1.8GHz Dual Core Machine. This enables real time processing of the BVI.

Though quite specific, the problem of road recognition for autonomous driving/driver assistance purposes is not trivial. Most well known researchers refer several issues that hamper visual recognition of the road: changes in illumination due to shadows, entrance or exit from tunnels [5][8][9][10], camera calibration when perspective transformations are employed [12], and also solar glare [8][17]. All these problems are frequent when working in the field. In the current proposal, to extract pixels brighter than their neighbors, a gray level morphological top hat operation is performed, as described in (1), where the '○' operator stands for the morphological opening (erosion followed by dilation)

$$TopHat(A,B) = A - (A \circ B) = A - \max_B\left(\min_B(A)\right) \quad (1)$$

The top hat operation enables the extraction of the lines in demanding illumination conditions, such as shadows, for example. The size of the structuring element B was defined considering that the BVI represents accurate geometrical information: B was made circular with a radius a bit larger than the road line typical width; a 12-pixel diameter disk was used.

IV. BRIGHTNESS PROFILE GENERATION

In order to start looking for road line candidates in the top hat brightness image, and admitting that the road lies ahead and intersects the bottom of the image, a brightness profile is defined along a line as illustrated in Figure 4 with a thicker colored line. This line was named the scan line (SL) and its position is obtained by searching for the first information available for every BVI's column, using a bottom up search in the gray level original BVI.

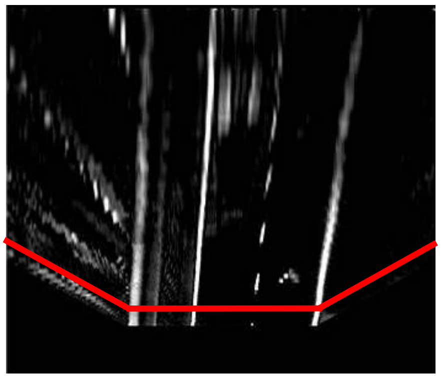

Figure 4 – Result of a top hat operation on BVI. The red full line represents the scan line positioning.

Still reporting to the scan line, a further vertical displacement is imposed so that the scanning occurs in an area slightly above (10 pixels) the limit of image valid information. This small offset was established to avoid possible abrupt gradient transients near the edge of the non-void area, which is typical in perspective corrected images.

The scan line (SL) has been used to extract an intensity profile in the top hat image. This signal is then clustered into several groups after the definition of a maximum gradient value. It is worth mentioning that the top hat image was not immediately thresholded into binary in order to preserve gray level information as long as possible. Therefore, the top hat extracted intensity profile histogram has been stretched to the [0, 255] interval, and a given limit was considered for group clustering (Figure 5).

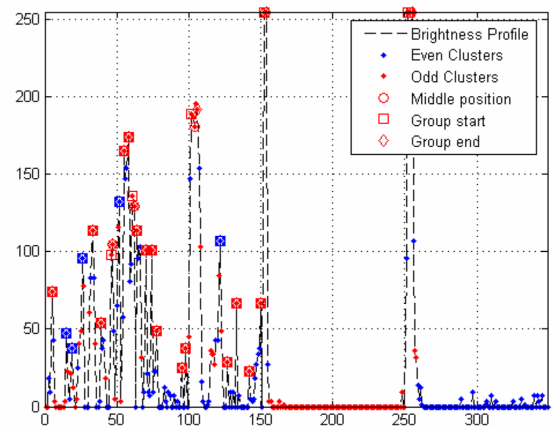

Figure 5 – Brightness profile of the scan line represented in Figure 4.

At the end of this stage, the several clustered groups are filtered to check whether their average brightness is higher than their left and right neighbors. If not, the groups are discarded. Also, a filter for the maximum number of pixels per group is applied, excluding groups whose size is considerably bigger than the lines expected width. Again, using the BVI made the definition of this limit easier.

V. CANDIDATE LINE RECONSTRUCTION AND ANALYSIS

Section IV showed how, departing from the SL, several groups of distinctive pixels can be extracted. Additional filtering based on size and brightness peak criteria output a set of several Candidate Groups. Each group represents a possible intersection of a road line with the SL. The middle point of each group is then employed as a seed point to a flood fill operation, reconstructing the candidate line from that point. By doing so, one assumes that pixels belonging to road lines are similar in brightness. The fill operation may be performed either on the grayscale or on the top hat versions of the BVI.

Each candidate line corresponds to a Boolean image of the same size as BVI, although additional performance optimization has been implemented making use of regions of interest. The images of the candidate lines go trough a search routine that finds several relevant points on the line candidate. The search routine consists of finding, for a set of equally vertically spaced attempts, the coordinate of the first white pixel that is found on a right to left scan (Figure 6). The output of this procedure is a vector of points that are on the right border of the line. Then the segment defined by each two consecutive border points is calculated and the orientation angles are calculated. The vectors normal to those directions (angles) define the line orientation trend throughout the image. Finally, for each normal vector, a correspondent one-pixel wide line segment is defined up to the adequate image limits and its intersection with the Boolean mask is computed, providing an indication of the line width at each point.

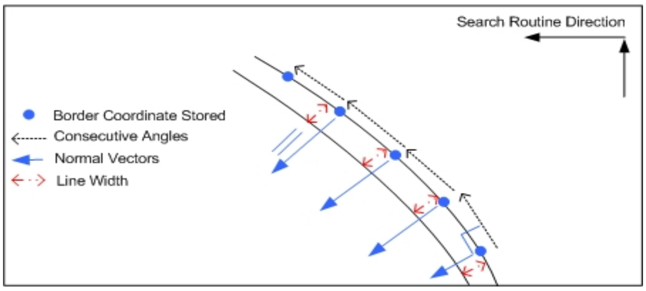

Figure 6 – Line candidate analysis.

The following definitions formalize several related concepts which will be used further for the statistical descriptors: the angles of the M normals along a line candidate (NA) and the variation (gradient) of those angles (GNA) are given in (2) and (3),

$$NA = \{\alpha_1, \alpha_2, \cdots, \alpha_M\} \tag{2}$$

$$GNA = \{\alpha_2 - \alpha_1,\ \ \alpha_3 - \alpha_2,\ \cdots,\ \ \alpha_{M-1} - \alpha_M\} \tag{3}$$

where M is the number of normals to be used and the angles of the normals are computed simply by a quadrant sensitive expression of the type given by (4):

$$\alpha_i = \arctan\left(y_{i+1} - y_i, x_{i+1} - x_i\right),\ i{=}1,\ldots,M \tag{4}$$

The two other indicators to be used are the set of line widths (LW) near the normal origins, and along their direction (as defined earlier), and also the variation (gradient) of the line widths (GLW), given respectively by (5) and (6).

$$LW = \{w_1, w_2, \cdots, w_M\} \tag{5}$$

$$GLW = \{w_2 - w_1,\ \ w_3 - w_2,\ \cdots,\ \ w_{M-1} - w_M\} \tag{6}$$

VI. Statistical Descriptors Generation

In summary, the process consists of generating several groups based on the scan line brightness profile. These groups, after being validated by some criteria related to a minimum size and distinction from their neighbors (as mentioned earlier), provide a seed point for the candidate line reconstruction. Finally, the candidate lines are analyzed by their width properties and set of normal vectors duly calculated. Stated this, the following descriptors were defined:

- Mean value of NA (NA_{avg}): provides and indication of the average line orientation in the image;
- Variance of NA (NA_{var}): is a measurement of the dispersion of the line orientation (how reliable is the mean value of the NA);
- Mean value of GNA (GNA_{avg}): measures how smooth the line inclination evolves (indicates line curvature);
- Variance of GNA (GNA_{var}): translates how constant the variation of line inclination is (smoothness of curvature);
- Mean value of LW (LW_{avg}): provides and indication of the average line width;
- Variance of LW (LW_{var}): is a measurement of the dispersion of the line width;
- Mean value of GLW (GLW_{avg}): translates the average width evolution;
- Variance of GLW (GLW_{var}): indicates how constant the line with variation is;

Figure 7 shows one of the candidate lines of the image in Figure 3 along with the statistical descriptors for that candidate.

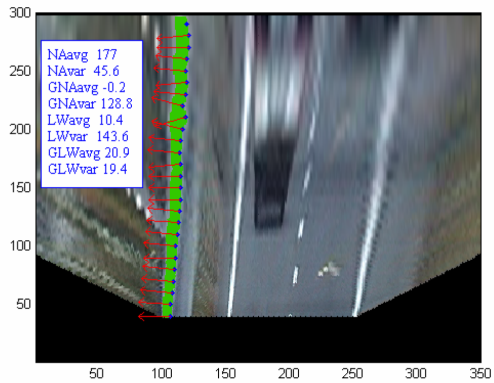

Figure 7 – Analysis of a line candidate.

After all this information is generated, the statistical descriptors of each candidate line are tested to see whether the line is compliant with the predefined intervals for each descriptor. These intervals are, up to now, defined manually. Their tuning is intuitive because the BVI involves real world related parameters. For example, the setting of the maximum and minimum average width is directly related to the road lines width.

VII. Results in the ATLAS Robots

This line detection technique has been originally developed for the ATLAS robots which compete annually in the Portuguese Robotics Open (ROBOTICA), in the Autonomous Driving Competition.

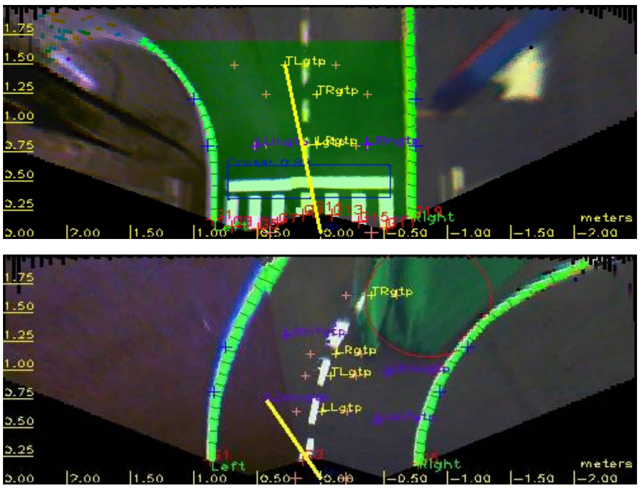

Figure 8 – Line detection using simple statistical descriptors. When detected, lines are superimposed in green.

Line detection, combined with the perspective transformation engine, allowed the ATLAS robot to detect the road lines more than 95% of the times at over 15 Hz.

This enabled a smooth and fast navigation throughout the competition, giving the robot an excellent performance among its competitors. Figure 8 shows two examples of one of the ATLAS robots view of the road.

VIII. Results in Real Roads

Because the results applied in the indoor robot competition were so satisfactory, the authors have decided to try out the algorithm on some images of real highways. The line detector works very well in extracting the lines from the perspective corrected image. Lines detected as such are painted in green. Red blobs are masks that have not been filtered through the scan line brightness profile. However, the information derived from the statistical descriptors has discarded these line candidates. Figure 9 shows the output of the lane detector for Figure 3's image. Notice that the top of the road separator, projected onto the BVI, is actually a bright elongated object very similar to a road line. Nonetheless, it is discarded in this case due to a non compliant line average width LW_{avg}.

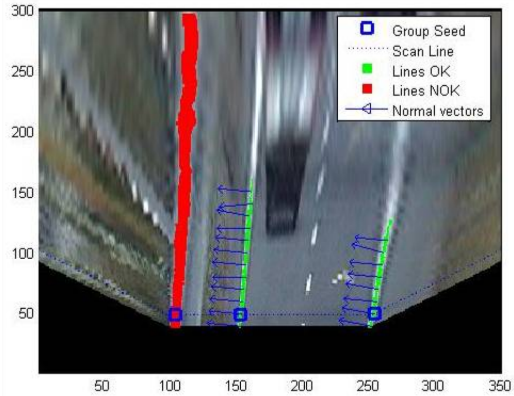

Figure 9 –.Line detection output for Figure 3's image. The leftmost candidate was discarded due to exceeding line average width.

In Figure 10, the "V" shaped line formation was also discarded. In fact, this formation is constituted by two road lines that merge at the bottom of the image.

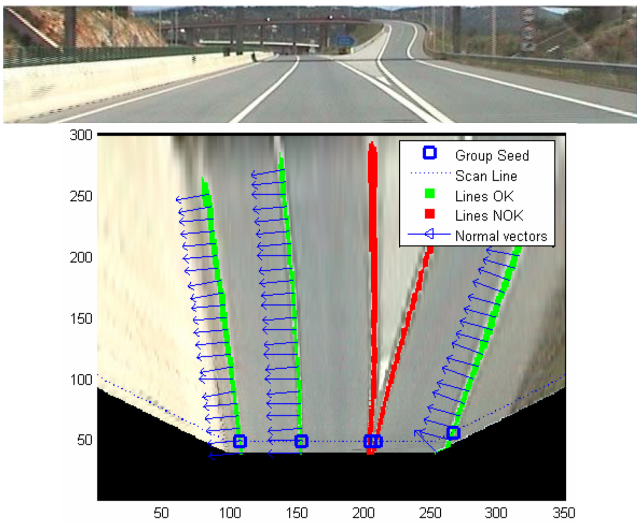

Figure 10 –.Line detection output discards two lines that merge at the bottom of the image.

The two lines of the "V" were discarded since the right from left search has a sudden leap when going from one line to the other causing an abnormal GNA_{var} but also an irregular LW_{avg}. One or two frames later the "V" shape splits into two distinct lines and would be correctly identified as two separate lines on the road.

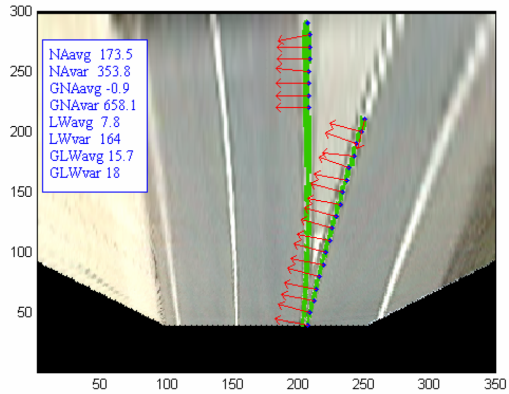

Figure 11 –.Particular analysis of v shaped artifact of Figure 10.

In the case of Figure 12, the left road line is occluded by a vehicle. Despite that, the algorithm reconstructs the line up until where the occlusion begins.

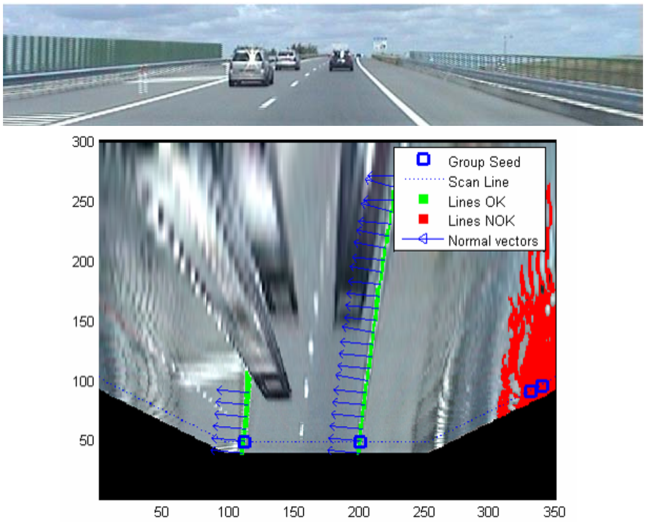

Figure 12 –.Line detection on a busy highway. Left line is detected until it is occluded by the vehicle.

It is also important to note that all these lines were detected using the same parameters, i.e., top hat operation radius, flood fill propagation tolerance and statistical descriptors intervals. No particular tuning for each image was executed.

IX. Conclusions

There are several research teams doing intensive research in ITS and autonomous driving systems which interests automotive industry, especially due to safety concerns. Nowadays, trends in autonomous navigation, shaped particularly by the competitors of the DARPA challenge, still employ an extensive collection of sensors, namely vision, GPS and range finders; therefore, using vision alone still is part of ongoing research. Many different techniques have already been employed due to the diversity of problems and some assumptions are made concerning road properties. Computational power is still an issue when developing ITS, not only because of the demanded processing rates, but also because some of the described algorithms are quite demanding. Most prototypes developed by other researchers have at least 3 dedicated computers, and some even have special hardware for particular tasks. No project reported image processing rates above 20Hz, and the images processed are of medium to low resolution.

The line detector introduced in this paper makes use of several assumptions. The first, required for perspective transformation, assumes the road as flat. This assumption is correct most of the times. The highway images displayed were, at the time, taken without any measurement of the cameras position and orientation. Hence, perspective transformation is not entirely accurate and so the input images are of low quality and also of small resolution. Still, the algorithm shows a high detection rate and a very low false positive rate. One of the advantages of this algorithm is the straightforwardness of setting the statistical descriptors. They are directly related to road lines characteristics in the real world. Also, the algorithm has shown real time performance in an ordinary 1.8GHz Dual Core Laptop, taking the ATLAS robot to a leading position in track detection in the indoor autonomous driving competition.

It must be reminded that the algorithm for line extraction is making decisions based on snapshots only, i.e., neither history nor prediction are, for the moment, employed. The authors are quite confident that after introducing some kind of predictive filtering or other similar techniques to cover for occlusions or other destructive interferences in line extraction, the presented technique is a promising contribution to the framework of ITS, namely in autonomy, assisted navigation and safety at reduced costs.

X. References

[1] Almeida, L., Azevedo, J., Cardeira, C., Costa, P., Fonseca, P. Lima, P., Ribeiro, F., Santos, V., 2000 - Mobile Robot Competitions: Fostering Advances in Research, Development and Education in Robotics. *Proc. of the 4th Portuguese Conference on Automatic Control, CONTROLO2000*, 4-6 October 2000, Guimarães, Portugal, pp. 592-597

[2] P. Afonso, J. Azevedo, C. Cardeira, B. Cunha, P. Lima, V. Santos, 2006 - Challenges and Solutions in an Autonomous Driving Mobile Robot Competition, *Proc. of the 7th Portuguese Conference on Automatic Control, CONTROLO2006*, Lisboa.

[3] M. Oliveira, V. Santos, 2007 - A Vision-based Solution for the Navigation of a Mobile Robot in a Road-like Environment, *Robótica, nº 69, 2007 (ISSN: 0874-9019).*

[4] R. Cancela, M. Neta, M. Oliveira, V. Santos, 2005 - ATLAS III: Um Robô com Visão Orientado para Provas de Condução Autónoma, *Robótica, nº 62, pp. 4-11, (ISSN: 0874-9019).*

[5] Bertozzi M., Broggi A., Fascioli A., 2000 - Vision-based intelligent vehicles: State of the art and perspectives. *Robotics and Autonomous Systems 32 (2000) 1-16.*

[6] Liu T., Zheng N., Cheng H., Xing Z., 2003 - A Novel Approach of Road Recognition Based on Deformable Template and Genetic Algorithm, *Proceedings of the IEEE Intelligent Transportation Systems, 12-15 October 2003, vol. 2, pp. 1251-1256. (ISBN: 0-7803-8125-4).*

[7] Zhang H., Yuan K., Mei S., Zhou Q., 2004 - Visual Navigation of an Automated Guided Vehicle Based on Path Recognition, *Proceedings of the Third International Conference on Machine Learning and Cybernetics, Shanghai, 26-29 August 2004.*

[8] Pomerleau D., 1995 - RALPH: Rapidly Adapting Lateral Position Handler, Proceedings of the Intelligent Vehicles '95 Symposium, Detroit, MI, USA, 25-26 Sep 1995 (ISBN: 0-7803-2983-X).

[9] Foedisch M., Takeuchi A., 2004 - Adaptive Real-Time Road Detection Using Neural Networks, Proceedings of the 7th International Conference on Intelligent Transportation Systems, Washington, DC, October 3-6, 2004.

[10] Foedisch M., Takeuchi A., 2004 - Adaptive Road Detection through Continuous Environment Learning, *Proceedings of the 33rd Applied Imagery Pattern Recognition Workshop (AIPR'04), 1550-5219/04.*

[11] Foedisch M., Schlenoff C., Shneier M., 2005 - Towards an Approach for Knowledge-based Road Detection, Proceedings of the 2005 CIKM Conference: Workshop on Research in Knowledge Representation for Autonomous Systems, October 31 – November 5, Bremen, Germany.

[12] Bertozzi M., Broggi A., 1998 - GOLD: A Parallel Real-Time Stereo Vision System for Generic Obstacle and Lane Detection, *IEEE Transactions on Image Processing, Vol. 7, nº 1, January, 1998.*

[13] Antolovic D., Leykin A., Johnson S., 2005 - Vanishing Point: a Visual Road-Detection Program for a DARPA Grand Challenge Vehicle, *Technical report, Bloomington, Indiana, December 2005.*

[14] Bradski G., Kaehler A., Pisarevsky V., 2005 - Learning-Based Computer Vision with Intel's Open Source Computer Vision Library, *Intel Technology Journal, Volume 9, Issue 2, 2005.*

[15] Dahlkamp H., Kaehler A., Stavens D., Thrun S., Bradski G., 2006 - Self-supervised Monocular Road Detection in Desert Terrain, *Proceedings of Robotics Science and Systems, Philadelphia, USA, August 2006.*

[16] Lombardi P., Zanin M., Messelodi S., 2005 - Switching Models for Vision-based On-Board Road Detection, Proceedings of the IEEE Intelligent Transportation Systems, 13-15 September, 2005, pp. 67-72. (ISBN: 0-7803-9215-9).

[17] Fletcher L., Petersson L., Zelinsky A. 2003 - Driver Assistance Systems based on Vision In and Out of Vehicles, Proceedings of the IEEE Intelligent Vehicles Symposium, 2003., 0-11 June 2003, pp. 322-327. (ISBN: 0-7803-7848-2)

[18] Apostoloff N., Zelinsky A., 2003 - Robust Vision based Lane Tracking using Multiple Cues and Particle Filtering, Proceedings of the IEEE Intelligent Vehicles Symposium, 2003, 9-11 June 2003, pp. 558-563. (ISBN: 0-7803-7848-2).

[19] Oliveira M., Santos V., 2008 - Multi-Camera Active Perception System with Variable Image Perspective for Mobile Robot Navigation, *8th Conference on Mobile Robots and Competitions, Festival Nacional de Robótica 2008, Aveiro, April 2008.*