Illuminant-Invariant Model-Based Road Segmentation

J.M. Álvarez, A. López and R. Baldrich

Abstract— **Road segmentation is an essential functionality for supporting advanced driver assistance systems (ADAS) such as road following or vehicle detection and tracking. Significant efforts have been made in order to solve this task using vision-based techniques. One of the major challenges of these techniques is dealing with lighting variations, especially shadows. Many of the approaches presented within this field use ad-hoc mechanisms applied after an initial segmentation to recover shadowed road patches. In this paper, we present an innovative method to obtain a road segmentation algorithm robust to extreme shadow conditions. The novelty of the proposal is the use of a shadowless feature space in combination with a model-based region growing algorithm. The former projects the color images such that the shadow effect is greatly attenuated. The latter uses histogram models to label the pixels as belonging to the road or to the background. These models are constructed on a frame by frame basis independently of the road shape to avoid limitations when addressing unstructured roads. The results presented show the validity of our approach.**

I. INTRODUCTION

Advanced driver assistance systems (ADAS) have arisen as a contribution to traffic safety, a major social issue in developed countries. A relevant functionality is road segmentation which supports ADAS applications like road departure warning. Road segmentation has been approached using a wide variety of techniques . Even the relevance of some previous works concerning autonomous driving [1], [2] the most important ones referring to this task arise from the first DARPA Grand Challenge[1]. In this context, the vehicles prepared for the challenge fuse the data provided from multiple sensors (LIDAR, cameras and GPS) to achieve the autonomous driving task.

Our interest is the use of computer vision methods to solve the real–time road segmentation. In particular we use a single forward facing color camera placed on the windshield of a vehicle. Color cameras have many advantages over using other sensors: higher resolution, richness of features (color, texture), low cost, easy aesthetic integration, non–intrusive nature and low power consumption. Within this field, road segmentation is also used as an invaluable background segmentation stage for other functionalities such as vehicle and pedestrian detection. Knowing the road surface considerably reduces the image region to search for such objects, thus, reducing false detections and allowing real–time processes. However, road segmentation is a very challenging task. The road is in an outdoor scenario imaged from a mobile platform. Hence, we deal with a continuously changing background, the presence of different vehicles with unknown movement, different road shapes with worn–away asphalt (or not asphalted at all), and different illumination conditions. The underlying idea of vision–based approaches is to consider that the road ahead have some constant features like color or texture that can be used to group road pixels. However, the perception of the road surface not only depends on its own features, which in fact are not constant, but also on unknown lighting conditions which makes the problem very challenging (Fig. 1). A common approach consists in keeping the idea of homogeneous road surface as a simple model, while shadows and reflections are corrected using ad hoc mechanisms applied after an initial segmentation step [3], [4], [5]. Other vision–based approaches use more complex models to overcome road variability and lighting conditions, so that they neither require such hard constraints nor post–processing to deal with shadows and light changes. Some examples are: the SCARF (Supervised Classification Applied to Road Following) [6] which uses a mixture of Gaussians, and the UNSCARF which is the unsupervised version [7]. More recently Tan *et al.* have made use of color histograms to model the road [8] and Sha *et al.* apply a boosting scheme using feature combination [9].

Fig. 1. One of the greatest challenges in road segmentation using computer vision methods is the treatment of shadows.

In this paper, following our preliminary work in [10], rather than the usual segmentation in a color space, we use a physics–based illumination invariant space [11] and a statistical region growing algorithm for reliable road segmentation despite illumination variations. Using this feature space we attenuate the shadow influence from the very beginning even using a simple road model. The invariant space consists of a grey-scale image that results from projecting the log-log pixel values onto the direction orthogonal to lighting change, within and outside the umbra. This projection greatly attenuates the shadows and it is computable in real–time using a single–sensor color camera. Based on such invariant images we also propose a novel approach for the real–time grouping of road surface pixels. Commonly used region growing algorithms characterize intra–region similarity using

J.M. Álvarez, A. López and R. Baldrich are with The Computer Vision Center and with the Dpt. of Computer Science, University Autonoma de Barcelona, 08193 Cerdanyola (Barcelona), Spain. jalvarez@cvc.uab.es

[1]http://www.darpa.mil/grandchallenge04/

first order statistics such as the mean value [12], [13]. Instead, we introduce a histogram-model based approach to describe the homogeneity of regions and decide whether a pixel belongs or not to the road surface. The main difference between our approach and the one presented by Tan *et al.* is the simplicity of the road model. They use four different distributions to build the road model and one more for the background. By contrast, and thanks to the invariant image, we can use a single distribution to model the road and one threshold to label the pixels. In addition, we do not need to estimate the background model nor assume the presence of all the road distributions for algorithm initialization.

The rest of this paper is organized as follows: in Sect. II we review the basics of the illuminant–invariant image space. Special interest is given to the practical problem of the camera calibration to generate the invariant image. The algorithm used to label the pixels as road or background is described in Sect. III. In Sect. IV we present qualitative results to validate the proposal. Finally, in Sect. V we present our conclusions.

II. ILLUMINANT–INVARIANT IMAGE SPACE

A. Theory Overview

Image formation models are defined in terms of the interaction between the spectral power distribution of illumination, surface reflectance and spectral sensitivity of the imaging sensors. Finlayson *et al.* have shown that if the lighting is approximately Planckian and having *Lambertian surfaces* imaged by three delta-function sensors it is possible to generate an illuminant-invariant image ($\mathfrak{I}$) [11], [14], [15]. Under these assumptions, a log-log plot of two dimensional $\{\log(R/G), \log((B/G)\}$ values for any surface forms a straight line provided camera sensors are fairly narrow-band. Thus, lighting change is reduced to a linear transformation along an almost straight line (Fig. 2).

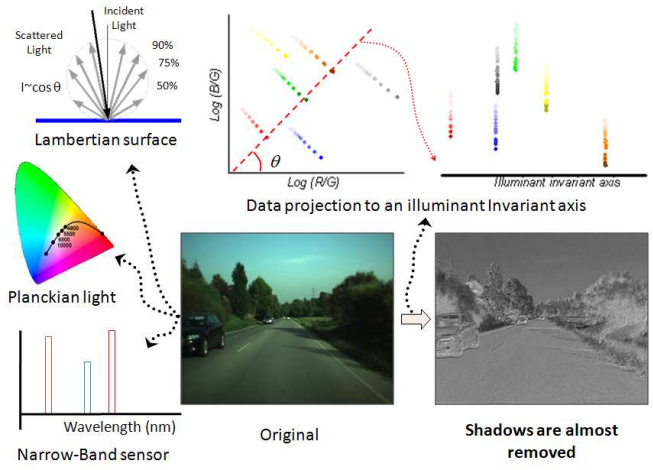

Fig. 2. Under the assumptions of Plankian light, Lambertian surface and narrow-band sensors, we can obtain an illuminant–invariant image which is almost shadow free.

This theory holds even for real data with only approximately Planckian lights, *non-Lambertian surfaces* and real cameras having only approximately non-narrow-band sensors. For instance, our acquisition system includes a Sony CCD camera at 15fps in 640×480 Bayer 8bit mode mounted on a vehicle driving on the road (Fig. 3). A detail to point out is that our acquisition system was operating in automatic shutter mode: i.e., within predefined ranges, the shutter changes to avoid both global overexposure and underexposure. However, provided we are using sensors with linear response and the same shutter for the three channels, we can model the shutter action as a multiplicative constant s, i.e., we have $sI_{RGB} = (sR, sG, sB)$ and, therefore, the channel normalization removes the constant (e.g., $sR/sG = R/G$). In short, $\mathfrak{I}$ is a gray-scale image that is obtained from projecting the log-rgb pixel values of the incoming data onto the direction orthogonal to the lighting change line. This direction θ depends on each camera color characteristics. Hence, before using the algorithm the camera must be calibrated, thus, θ is an additional intrinsic parameter.

Fig. 3. Examples of illuminant invariant images. Upper row contains real images of different scenarios and the lower their corresponding invariant image. These results suggest the invariant theory holds although the theoretical assumptions are not perfectly achieved, i.e., the combination of our camera, the daylight illuminant and the surface we are interested in (the road).

B. Camera Calibration

The calibration of the camera can be done either using calibration patterns [14] or using an entropy minimization technique based on the information in each image [15].

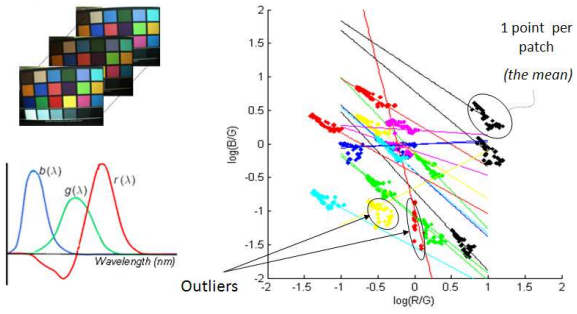

Fig. 4. Log–log space representation for all the patches of every acquired image. Every patch of each color in one image is represented using a single point, the mean. Cameras which do not use narrow–band sensors do not produce parallel lines for each patch.

In order to use the first method the calibration pattern has to be imaged under different lighting conditions. Ideally, RGB triplets of each of the pattern patches will form a straight line in the *log–log* space introduced before. The difference between all the patches is ideally reduced to just an offset. The slope of those lines is the invariant direction and the angle we are looking for, θ, corresponds to the perpendicular one. In practice, those patches are not always distributed as parallel. Moreover, there is a large dispersion within the patch and the corresponding RGB cluster does not clearly form a straight line (Fig. 4). The calibration angle can be obtained using the algorithm proposed by Drew *et al.* in [16]. However, even using the overall mean and a robust estimator, the presence of outliers may bias the final line slope affecting the estimated angle. That is, the solution is unstable. Furthermore, this calibration technique is tedious.

The second method is proposed by Finlayson *et al.* [15]. The characteristic angle is obtained using the entropy histogram of $\mathfrak{I}$ which is formed by projecting the incoming data onto each possible angle θ. Given this histogram, the invariant direction that generates an invariant image with minimum entropy is the correct angle. That is, the method considers the entropy as a measure of the degree of disorder in the data. When the data is projected to the correct angle, under the assumption the camera produces straight lines as the light source varies, the entropy will be minimal since each line will contribute to the same bin of the image histogram used to calculate the entropy value. Projecting to a different direction will scatter points across the invariant axis producing a higher disorder and thus a higher entropy. The algorithm proposed by Finlayson considers the nature of the data, for real images, to estimate the bin width and the data range to be used in the histogram of $\mathfrak{I}$. However, despite all these improvements the algorithm is not repetitive at all as shown in Fig. 5. The angle which produces the minimum entropy varies depending on the image contents.

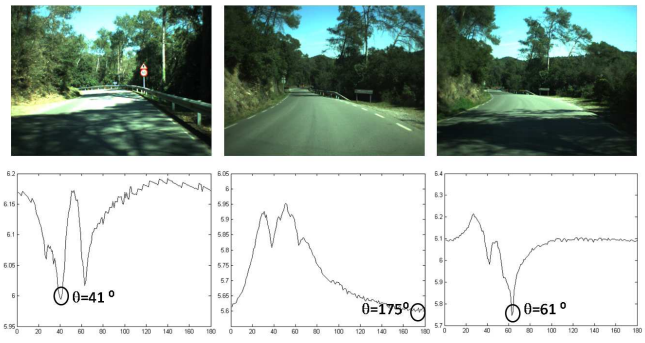

Fig. 5. Upper row shows some images used to calibrate the system by the entropy based method. Lower row shows their respective entropy distributions. These distributions vary considerably affecting the estimated invariant angle. These results suggest the method is not repetitive when used with real images.

We propose a new scheme to obtain θ (Fig. 6). The result is a robust and repetitive calibration method even using real data. Instead of using a single image as in Finlayson's algorithm, our proposal assumes a collection of real images acquired with a monocular color sensor. These images are corrected to linearize the sensor response and lens distortion. Furthermore, the scheme includes an outlier rejection process and a robust entropy density estimator.

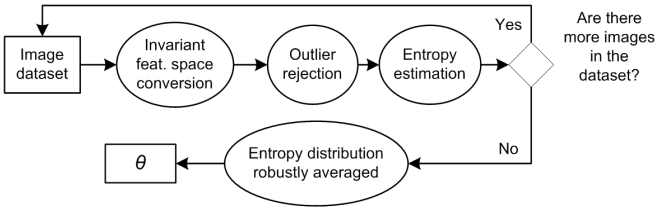

Fig. 6. Algorithm to calibrate the camera using road images. Given a group of images, the entropy distribution is calculated as the robust mean over different entropy distributions. The invariant angle (θ) is the minimum of the final robustly–averaged entropy distribution.

As in the original method proposed by Finlayson, the log-chromaticity image is projected onto a direction $\theta \in \{0,1,..,180\}$. After this, a process to remove outliers is done for every projected data. Rejecting outliers is necessary since subsequent processes use the range of the data to calculate histogram bins. Having an outlier cause all the data get compressed forming an erroneous histogram. Since $\mathfrak{I}$ has an unknown data distribution the *Chebyshev's theorem* is applied to calculate the lower and upper data boundaries [17]. Only 90% of this data range is used to form the histogram. Scott's Rule is used to form the appropriate bin width:

$$bin\_width = 3.5std(\mathfrak{I})N^{-\frac{1}{3}}, \quad (1)$$

where N is the number of non–excluded pixels of the invariant image.

Once the image has been rotated for every θ an entropy distribution is formed for each image in the dataset. Originally, the desired angle was the one having the minimum entropy value. However, in order to improve robustness, we search for the angle which minimizes the averaged entropy distribution of a group of images. This entropy distribution is calculated averaging the entropy obtained for every image using a robust mean approach: for every angle, only the 90% middle range is used to estimate the mean. The highest and lowest values of entropy for each angle are excluded. Although this method has a bigger computational cost, it is less susceptible to extreme scores than the arithmetic mean. Besides, it is intended to be done off–line and only takes some seconds. Using this method a very strong minimum is obtained which represent the desired invariant angle as shown in results section.

III. THE HISTOGRAM-BASED ROAD CLASSIFIER

The road segmentation is performed using a seeded region growing algorithm (SRG) [12], [13]. SRG is a common technique in image segmentation which has been successfully used in other road segmentation schemes [3]. Given a group of N classes $C_i, i \in 1,2..N$ the key step of SGR is

defining the dissimilarity criterion δ used in each iteration to decide whether a pixel $p=(x,y)$ belongs or not to one of these classes. The evaluated pixel is assigned to the class having the lowest δ. This criterion is usually based on image properties such as the mean value of the pixels $p'=(x',y')$ which already belong to the i-th class. Initially, these classes are formed by the seed pixels. Considering $\mathfrak{I}$ as input data, it would be:

$$\delta(p,C_i) = |\mathfrak{I}(p) - \frac{1}{\aleph C_i} \sum_{p' \in C_i} \mathfrak{I}(p')|. \tag{2}$$

However, in order to reach real–time constraints some systems use a fixed class representant (the original seed) rather than the mean [3]. Concerning the road segmentation, there are only two classes, the road and the background. Nevertheless, if only some knowledge about the road is assumed, a maximum dissimilarity value has to be defined to consider p as a road pixel.

This is an elementary approach since it only uses first order statistics to characterize the intra–region homogeneity. However, decisions can not be always made with the help of these statistics alone. For instance, a second-order statistic, such as the covariance between sample vectors of different classes is needed to characterize the feature distribution within a region.

The novelty of our classifier is the use of a density function as classification criterion. We use the road probability density $P(\mathfrak{I}(p)|Road)$ and a fixed threshold λ on this density to decide whether a pixel belongs or not to the road class:

$$\begin{cases} p \text{ is Road}, & \text{if } P(\mathfrak{I}(p)|Road) \geq \lambda, \\ p \text{ is Background}, & \text{otherwise.} \end{cases} \tag{3}$$

Notice that our approach does not need the background model, $P(\mathfrak{I}(p)|Background)$. This is an advantage since estimating it requires assumptions regarding the background position in the image.

It remains to estimate $P(\mathfrak{I}(p)|Road)$, i.e., the road model. Although this can be solved using parametric methods such a mixture of Gaussians (MoG) as in [18], we have considered a non-parametric method to estimate it. In particular, we have used the normalized histogram as an empirical form of probability distribution for a random variable [19]. There are two clear advantages of non-parametric methods: they are fast in training and usage and they are independent to the shape of data distribution. In addition, there are several works which suggest the superiority of histograms compared to MoG [8], [20], [21]. In our case, the road model has been built using the surrounding region of several seeds placed at the bottom part of $\mathfrak{I}$ (Fig. 7). That is, our system uses the only assumption that the bottom region of the image is road. In fact, the lowest row of the image corresponds to a distance of about 4 meters away from the vehicle, thus, it is a reasonable assumption most of the time (other proposals require to see the full road free [3] or all the possible road appearances [8] at the start up of the system, which is more unrealistic).

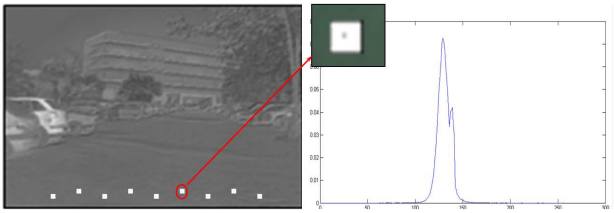

Fig. 7. Road model example: The model is built using the histogram formed with the surrounding region (white blocks) of several seeds. We have used nine seeds placed at the bottom part of $\mathfrak{I}$.

(a) (b) (c)

Fig. 8. Comparison of two different segmentations results when algorithms are applied on $\mathfrak{I}$ which is calculated from the original image as described in Sect. II. (a) Original image, (b) segmentation result using standard seeded region growing and the mean as dissimilarity criteria, (c) segmentation result using the histogram–based region growing algorithm.

Fig. 8 shows the improvement obtained using the statistical approach in comparison with common SRG. Our approach clearly outperforms the other when seeds are placed on points of $\mathfrak{I}$ which do not properly represent the road surface.

IV. RESULTS

In this section we present qualitative results to validate the proposal. We have started by calculating the characteristic angle of the camera using a dataset which includes different scenarios under different lighting and weather conditions (Fig. 9). Results shown in Fig. 10 confirms the proposed algorithm is robust and repetitive. The angle obtained for our camera is $\theta = 44°$.

Fig. 9. Different images from the dataset used to calibrate the camera. The images have been taken under different daytime and weather conditions such as sunny or rainy days, and include different scenarios.

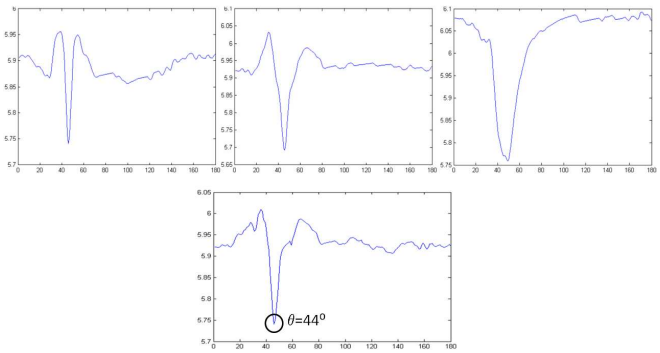

Fig. 10. Upper row shows the entropy obtained using the proposed algorithm for each image in the upper row of Fig. 5. The algorithm is more repetitive than the previous one. The lower row shows the averaged entropy distribution used to obtain a more robust estimation of θ.

Given θ, the algorithm depicted in Fig. 11 has been used to perform a frame by frame segmentation. Images are acquired using the system described in Sect. II, and cover approximately the nearest $80m$ ahead of the vehicle. Each image is converted onto the invariant feature space and the proposed region growing classifier is applied. The result presents some small holes that are filled by standard mathematical morphology.

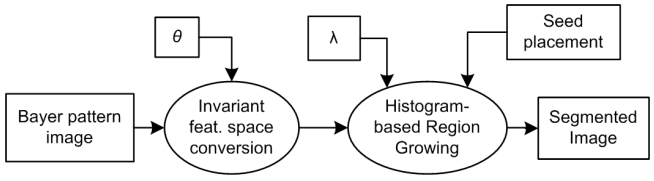

Fig. 11. Algorithm used to validate the road segmentation proposal.

The algorithm has been tested using two different image sequences. The former sequence was taken on a sunny day. Its images include nonhomogeneous roads due to extreme shadows and the presence of other vehicles. The latter was taken on a rainy day and also includes nonhomogeneous roads due to shadows, humidity and the presence of other vehicles. Example results for both sequences are shown in Fig. 12 and 13, respectively. Other results for images with complex road shapes are shown in Fig. 14. Since the algorithm does not use shape constraints it can deal with this kind of situations. All these results suggest that a reliable road segmentation algorithm is obtained by combining the illuminant–invariant image space and a modified region growing classifier. Road surface is well recovered most of the time, with the segmentation stopping at road limits and vehicles, and can deal with complex road shapes.

Currently, the average time required to process each frame in non–optimized MatLab code is approximately $735ms$ at full resolution on a Pentium-4 CPU at 2.8GHz. However, we do not expect any trouble to reach real–time when written in proper C++ code.

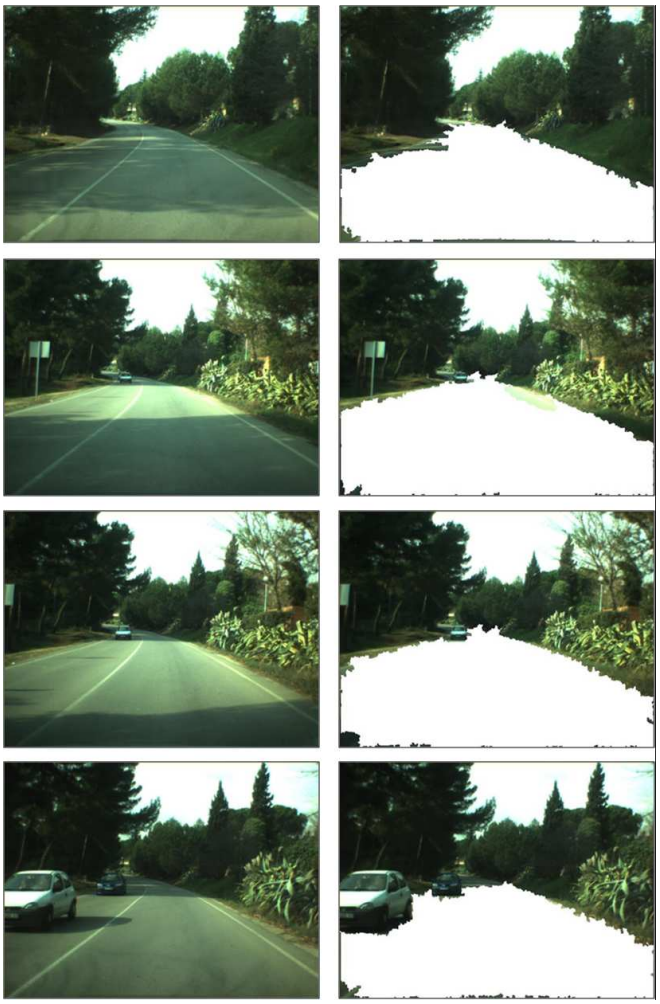

Fig. 12. Results from the sunny day sequence. This sequence includes extreme shadows and other vehicles in the scene. Left column is the original image which covers the nearest $80m$ ahead of the vehicle. Right column shows in white the corresponding segmented road.

V. CONCLUSIONS AND FUTURE WORK

In this paper we have introduced a novel approach for road segmentation. The approach enables robust segmentation despite the presence of strong shadows under different weather conditions. Furthermore, since the approach does not use shape constraints it can deal with complex road shapes and with the presence of other vehicles in the scene. The novelty of the approach resides in using a shadow–invariant image based on a preliminary, simple, camera calibration combined with a statistical region growing algorithm. Our method attenuates the influence of shadows from the beginning rather than using a specific color space either in a complex algorithm or with shadow post–processing. Moreover, it uses a region growing algorithm which includes histogram–models to decide whether a pixel belongs or not to the road surface.

Finally, referring to the required camera calibration using real–data, we have improved the existing algorithm. A more robust and repetitive calibration method has been obtained. The modified algorithm has been successfully tested to

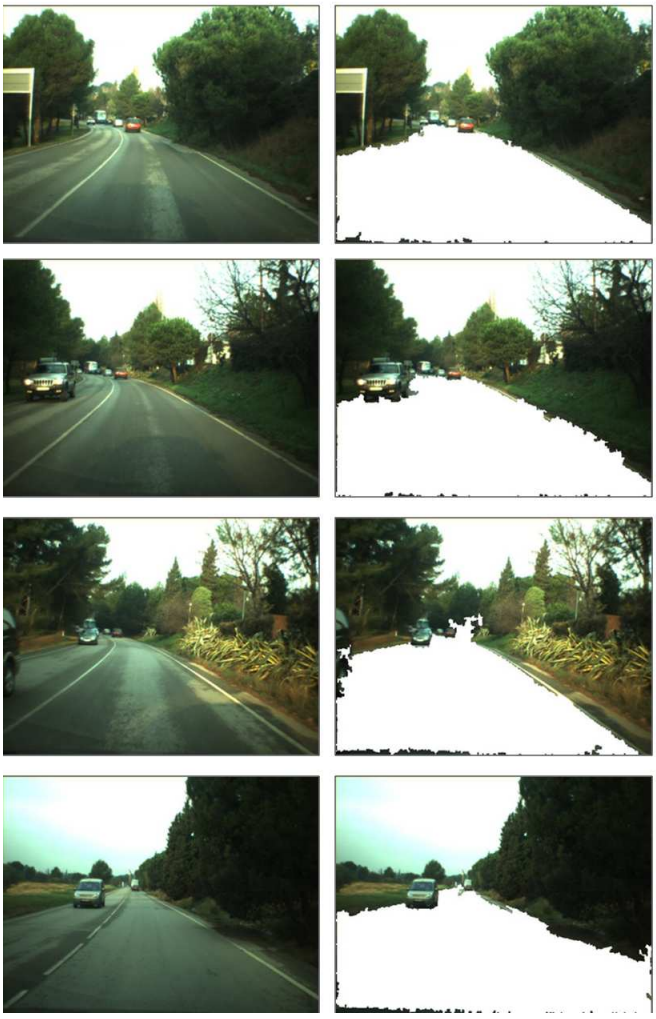

Fig. 13. Results from the rainy day sequence. Road surface is nonhomogeneous because it is non-uniformly soaked.

Fig. 14. Results with complex road shapes. Avoiding shape constraints allows the algorithm to segment the road surface even when the road is unstructured or with no clear shape.

generate the results discussed above.

In the future, we aim to address the challenging problem of evaluating the segmentation results quantitatively [22].

VI. ACKNOWLEDGMENTS

This work was supported by the Spanish Ministry of Education and Science under project TRA2007-62526/AUT and research programme Consolider Ingenio 2010: MIPRCV (CSD2007-00018).

REFERENCES

[1] C. Thorpe, M. Hebert, T. Kanade, and S. Shafer, "Vision and navigation for the carnegie-mellon navlab," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 3, pp. 362 – 373, May 1988.

[2] H. Schneiderman and M. Nashman, "Visual processing for autonomous driving," in *Proceedings of IEEE Workshop on Applications of Computer Vision*, December 1992, pp. 164 – 171.

[3] M. Sotelo, F. Rodriguez, L. Magdalena, L. Bergasa, and L. Boquete, "A color vision-based lane tracking system for autonomous driving in unmarked roads," *Autonomous Robots*, vol. 16, no. 1, 2004.

[4] O. Ramstrom and H. Christensen, "A method for following unmarked roads," in *IEEE Intelligent Vehicles Symposium*, 2005.

[5] C. Rotaru, T. Graf, and J. Zhang, "Extracting road features from color images using a cognitive approach," in *IEEE Intelligent Vehicles Symposium*, 2004.

[6] J. Crisman and C. Thorpe, "Scarf: A color vision system that tracks roads and intersections," *IEEE Trans. on Robotics and Automation*, vol. 9, no. 1, pp. 49 – 58, February 1993.

[7] J. Crisman and C. Thorpe, "Unscarf, a color vision system for the detection of unstructured roads," *Proceedings of IEEE International Conference on Robotics and Automation*, vol. 3, pp. 2496 – 2501, April 1991.

[8] C. Tan, T. Hong, T. Chang, and M. Shneier, "Color model-based real-time learning for road following," *Intelligent Transportation Systems Conference*, pp. 939–944, 2006.

[9] Z. Y. Y. Yun, Sha; Guo-ying, "A road detection algorithm by boosting using feature combination," *Intelligent Vehicles Symposium, 2007 IEEE*, pp. 364–368, 13-15 June 2007.

[10] Omitted for blind Revision.

[11] G. Finlayson, S. Hordley, C. Lu, and M. Drew, "On the removal of shadows from images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, 2006.

[12] R. Adams and L. Bischof, "Seeded region growing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 6, pp. 641–647, 1994.

[13] J. Fan, G. Zeng, M. Body, and M.-S. Hacid, "Seeded region growing: an extensive and comparative study," *Pattern Recogn. Lett.*, vol. 26, no. 8, pp. 1139–1156, 2005.

[14] G. Finlayson, S. Hordley, and M. Drew, "Removing shadows from images," in *European Conference on Computer Vision*, 2002.

[15] G. D. Finlayson, M. S. Drew, and C. Lu, "Intrinsic images by entropy minimization." in *ECCV (3)*, 2004, pp. 582–595.

[16] H. Jiang and M. S. Drew, "Shadow resistant tracking using inertia constraints," *Pattern Recogn.*, vol. 40, no. 7, pp. 1929–1945, 2007.

[17] B. Amidan, T. Ferryman, and S. Cooley, "Data outlier detection using the chebyshev theorem," *IEEE Aerospace Conference*, vol. 5–12.

[18] Z. Pan and J. Lu, "A bayes-based region-growing algorithm for medical image segmentation," *Computing in Science and Engg.*, vol. 9, no. 4, pp. 32–38, 2007.

[19] X. R. Li, *Probability, Random Signals, and Statistics*. CRC Press, 1999, ch. Section 3.11.

[20] M. J. Jones and J. M. Rehg, "Statistical color models with application to skin detection." *International Journal of Computer Vision*, vol. 46, no. 1, pp. 81–96, 2002. [Online]. Available: citeseer.ist.psu.edu/jones99statistical.html

[21] L. Sigal, S. Sclaroff, and V. Athitsos, "Skin color-based video segmentation under time-varying illumination," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 7, pp. 862–877, 2004.

[22] R. Unnikrishnan, C. Pantofaru, and M. Hebert, "Toward objective evaluation of image segmentation algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 6, pp. 929–944, June 2007.