Shadow Resistant Road Segmentation from a Mobile Monocular System

José Manuel Álvarez, Antonio M. López, and Ramon Baldrich

Computer Vision Center and Computer Science Dpt.,
Universitat Autònoma de Barcelona
Edifici O, 08193 Bellaterra, Barcelona, Spain
{jalvarez,antonio,ramon}@cvc.uab.es
http://www.cvc.uab.es/adas

Abstract. An essential functionality for advanced driver assistance systems (ADAS) is road segmentation, which directly supports ADAS applications like road departure warning and is an invaluable background segmentation stage for other functionalities as vehicle detection. Unfortunately, road segmentation is far from being trivial since the road is in an outdoor scenario imaged from a mobile platform. For instance, shadows are a relevant problem for segmentation. The usual approaches are ad hoc mechanisms, applied after an initial segmentation step, that try to recover road patches not included as segmented road for being in shadow. In this paper we argue that by using a different feature space to perform the segmentation we can minimize the problem of shadows from the very beginning. Rather than the usual segmentation in a color space we propose segmentation in a shadowless image which is computable in real–time using a color camera. The paper presents comparative results for both asphalted and non–asphalted roads, showing the benefits of the proposal in presence of shadows and vehicles.

1 Introduction

Advanced driver assistance systems (ADAS) arise as a contribution to traffic safety, a major social issue in modern countries. The functionalities required to build such systems can be addressed by computer vision techniques, which have many advantages over using active sensors (e.g. radar, lidar). Some of them are: higher resolution, richness of features (color, texture), low cost, easy aesthetic integration, non–intrusive nature, low power consumption, and besides, some functionalities can only be addressed by interpreting visual information. A relevant functionality is road segmentation which supports ADAS applications like road departure warning. Moreover, it is an invaluable background segmentation stage for other functionalities as vehicle and pedestrian detection, since knowing the road surface considerably reduces the image region to search for such objects, thus, allowing real–time and reducing false detections.

Our interest is real–time segmentation of road surfaces, both non–asphalted and asphalted, using a single forward facing color camera placed at the windshield of a vehicle. However, road segmentation is far from being trivial since the

J. Martí et al. (Eds.): IbPRIA 2007, Part II, LNCS 4478, pp. 9–16, 2007.

Fig. 1. Roads with shadows

road is in an outdoor scenario imaged from a mobile platform. Hence, we deal with a continuously changing background, the presence of different vehicles of unknown movement, different road shapes with worn–out asphalt (or not asphalted at all), and different illumination conditions. For instance, a particularly relevant problem is the presence of shadows (Fig. 1). The usual approaches found in the literature are ad hoc mechanisms applied after an initial segmentation step (e.g. [1,2,3]). These mechanisms try to recover road patches not included as segmented road for being in shadow. In this paper we argue that by using a different feature space to perform the segmentation we can minimize the problem of shadows from the very beginning. Rather than the usual segmentation in a color space, we propose segmentation in a shadowless image, which is computable in real–time using a color camera. In particular, we use the grey–scale illuminant invariant image introduced in [4], $\mathcal{I}$ from now on.

In Sect. 2 we summarize the formulation of $\mathcal{I}$. Moreover, we also show that automatic shutter, needed outdoors to avoid global over/under–exposure, fits well in such formulation. In order to illustrate the usefulness of $\mathcal{I}$, in Sect. 3 we propose a segmentation algorithm based on standard region growing applied to $\mathcal{I}$. We remark that we do not recover a shadow–free color image from the original, which would result in too large processing time for the road segmentation problem. Section 4 presents comparative road segmentation results in presence of shadows and vehicles, both in asphalted and non–asphalted roads, confirming the validity of our hypothesis. Finally, conclusions are drawn in Sect. 5.

2 Illuminant Invariant Image

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.* [4] show that under the assumptions of *Planckian illumination*, *Lambertian surfaces* and having three different *narrow band sensors*, it is possible to obtain a shadow–free color image. We are not interested in such image since it requires very large processing time to be recovered. We focus on an illuminant invariant image ($\mathcal{I}$) that is obtained at the first stage of the shadow–free color image recovering process. We briefly expose here the idea behind $\mathcal{I}$ and refer to [4] for details.

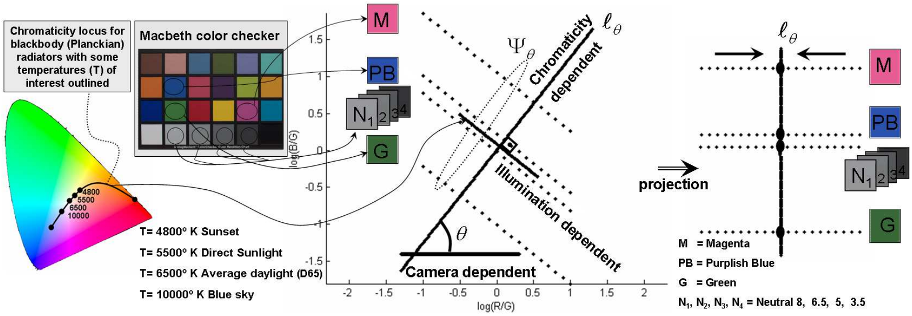


Fig. 2. Ideal log–log chromaticity plot. A Lambertian surface patch of a given chromaticity under a Planckian illumination is represented by a point. By changing the color temperature of the Planckian illuminator we obtain a straight line associated to the patch. Lambertian surface patches of different chromaticity have different associated lines. All these lines form a family of parallel lines, namely Ψ_θ. Let ℓ_θ be a line perpendicular to Ψ_θ and θ the angle between ℓ_θ and the horizontal axis. Then, by projection, we have a one–to–one correspondence between points in ℓ_θ and straight lines of Ψ_θ, so that ℓ_θ preserves the differences regarding chromaticity but removes differences due to illumination changes assuming Planckian radiators.

Let us denote by R, G, B the usual color channels and assume a normalizing channel (or combination of channels), e.g. without losing generality let us choose G as such normalizing channel. Then, under the assumptions regarding the sensors, the surfaces and the illuminators, if we perform a plot of $r = \log(R/G)$ vs. $b = \log(B/G)$ for a set of surfaces of different chromaticity under different illuminants, we would obtain a result similar to the one in Fig. 2. This means that we obtain an axis, ℓ_θ, where a surface under different illuminations is represented by the same point, while moving along ℓ_θ implies to change the surface chromaticity. In other words, ℓ_θ can be seen as a grey–level axis where each grey level corresponds to a surface chromaticity, independently of the surface illumination. Therefore, we obtain an illuminant invariant image, $\mathcal{I}(\mathbf{p})$, by taking each pixel $\mathbf{p} = (x, y)$ of the original color image, $I_{RGB}(\mathbf{p}) = (R(\mathbf{p}), G(\mathbf{p}), B(\mathbf{p}))$, computing $\mathbf{p}' = (r(\mathbf{p}), b(\mathbf{p}))$ and projecting $\mathbf{p}'$ onto ℓ_θ according to θ (a camera dependent constant angle). The reason for $\mathcal{I}$ being shadow–free is, roughly, that non–shadow surface areas are illuminated by both direct sunlight and skylight (a sort of scattered ambient light), while areas in the umbra are only illuminated by skylight. Since both, skylight alone and with sunlight addition, can be considered Planckian illuminations [5], areas of the same chromaticity ideally project onto the same point in ℓ_θ, no matter if the areas are in shadow or not.

Given this result, the first question is whether the working assumptions are realistic or not. In fact, *Finlayson et al.* [4] show examples where, despite the departures from the assumptions that are found in practice, the obtained results are quite good. We will see in Sect. 4 that this holds in our case, i.e., the combination of our camera, the daylight illuminant and the surface we are interested in (the road) fits pretty well the $\mathcal{I}$ theory.

A detail to point out is that our acquisition system was operating in automatic shutter mode: i.e., inside 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 addition, we expect the illumination invariant image to reduce not only differences due to shadow but also differences due to asphalt degradation since, at the resolution we work, they are pretty analogous to just intensity changes. Note that the whole intensity axis is equivalent to a single chromaticity, i.e., all the patches of the last row of the Macbeth color checker in Fig. 2 (N_i) project to the same point of ℓ_θ.

3 Road Segmentation

With the aim of evaluating the suitability of the illuminant invariant image we have devised a relatively simple segmentation method based on region growing [6], sketched in Fig. 3. This is, we do not claim that the proposed segmentation is the best, but one of the most simplest that can be expected to work in our problem. We emphasize that our aim is to show the suitability of $\mathcal{I}$ for road segmentation and we think that providing good results can be a proof of it, even using such simple segmentation approach.

The region growing uses a very simple *aggregation criterium*: if $\mathbf{p} = (x, y)$ is a pixel already classified as of the road, any other pixel $\mathbf{p}_n = (x_n, y_n)$ of its 8–connected neighborhood is classified as road one if

$$diss(\mathbf{p}, \mathbf{p}_n) < t_{agg} \ , \tag{1}$$

where $diss(\mathbf{p}, \mathbf{p}_n)$ is the dissimilarity metric for the aggregation and t_{agg} a threshold that fixes the maximum dissimilarity to consider two connected pixels as of the same region. To prove the usefulness of $\mathcal{I}$ we use the simplest dissimilarity based on grey levels, i.e.,

$$diss_{\mathcal{I}}(\mathbf{p}, \mathbf{p}_n) = |\mathcal{I}(\mathbf{p}) - \mathcal{I}(\mathbf{p}_n)| \ . \tag{2}$$

Of course, region growing needs initialization, i.e., the so–called *seeds*. Currently, such seeds are taken from fixed positions at the bottom region of the image (Fig. 3), i.e., we assume that such region is part of the 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 at the start up of the system, e.g. [1]).

In order to compute the angle θ corresponding to our camera, we have followed two approaches. One is the proposal in [7], based on acquiring images of

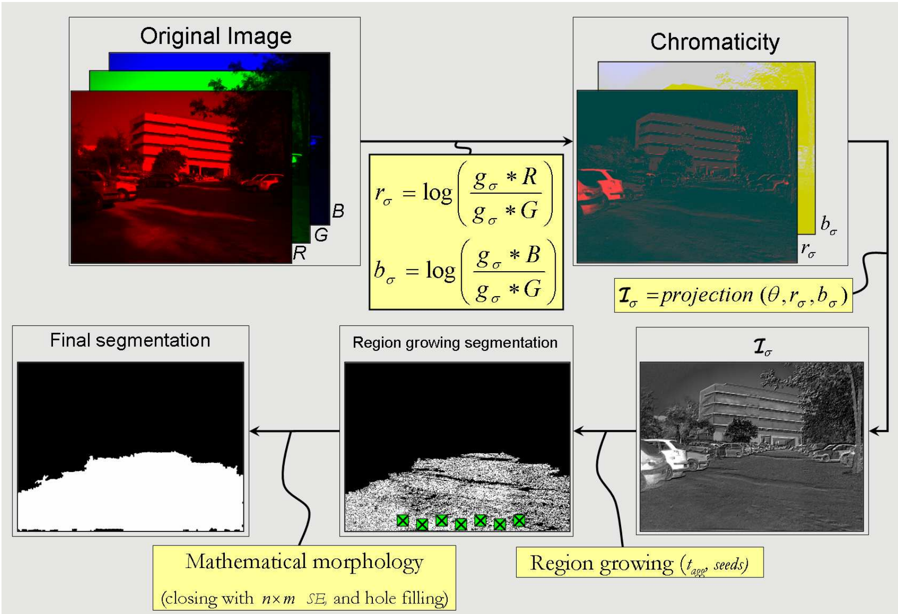


Fig. 3. Proposed algorithm. In all our experiments we have fixed values for the algorithm parameters: $\sigma = 0.5$ for Gaussian smoothing (Gaussian kernel, g_σ, discretized in a 3×3 window for convolution '$*$'); $\theta = 38°$; $t_{agg} = 0,007$ and seven *seeds* placed at the squares pointed out in the region growing result; structuring element (SE) of $n \times m = 5 \times 3$. Notice that we apply some mathematical morphology just to fill in some small gaps and thin grooves.

the Macbeth color checker under different day time illuminations and using the (r,b)–plot to obtain θ. The other approach consists in taking a few road images with shadows and use them as positive examples to find θ providing the best shadow–free images for all the examples. The values of θ obtained from the two calibration methods basically give rise to the same segmentation results. We have taken θ from the example–based calibration because it provides slightly better segmentations. Besides, although not proposed in the original formulation of $\mathcal{I}$, before computing it we regularize the input image I_{RGB} by a small amount of Gaussian smoothing (the same for each color channel).

4 Results

In this section we present comparative results based on the region growing algorithm introduced in Sect. 3 for three different feature spaces: intensity image (I; also called luminance or brightness); hue–saturation–intensity (HSI) color space; and the illuminant invariant image ($\mathcal{I}$).

The intensity image is included in the comparison just to see what can we expect from a monocular monochrome system. Since it is a grey level image, its corresponding dissimilarity measure is defined analogously to Eq. (2), i.e.:

$$diss_I(\mathbf{p}, \mathbf{p}_n) = |I(\mathbf{p}) - I(\mathbf{p}_n)| \ . \quad (3)$$

The HSI space is chosen because it is one of the most accepted color spaces for segmentation purposes [8]. The reason is that by having separated chrominance (H & S) and intensity (I) such space allows reasoning in a closer way to human perception than others. For instance, it is possible to define a psychologically meaningful distance between colors as the cylindrical metric proposed in [8] for multimedia applications, and used in [1] for segmenting non–asphalted roads. Such metric gives rise to the following dissimilarity measure for HSI space:

– Case *achromatic pixels*: use only the definition of $diss_I$ given in Eq. (3).
– Case *chromatic pixels*:

$$diss_{HSI}(\mathbf{p}, \mathbf{p}_n) = \sqrt{diss^2_{HS}(\mathbf{p}, \mathbf{p}_n) + diss^2_I(\mathbf{p}, \mathbf{p}_n)} \ , \quad (4)$$

taking the definition of $diss_I$ from Eq. (3), and given

$$diss_{HS}(\mathbf{p}, \mathbf{p}_n) = \sqrt{S^2(\mathbf{p}) + S^2(\mathbf{p}_n) + S(\mathbf{p})S(\mathbf{p}_n)\cos\varphi} \ ,$$

being
$$\varphi = \begin{cases} diss_H(\mathbf{p}, \mathbf{p}_n) & \text{if } diss_H(\mathbf{p}, \mathbf{p}_n) < 180^\circ \ , \\ 360^\circ - diss_H(\mathbf{p}, \mathbf{p}_n) & \text{otherwise} \ , \end{cases}$$

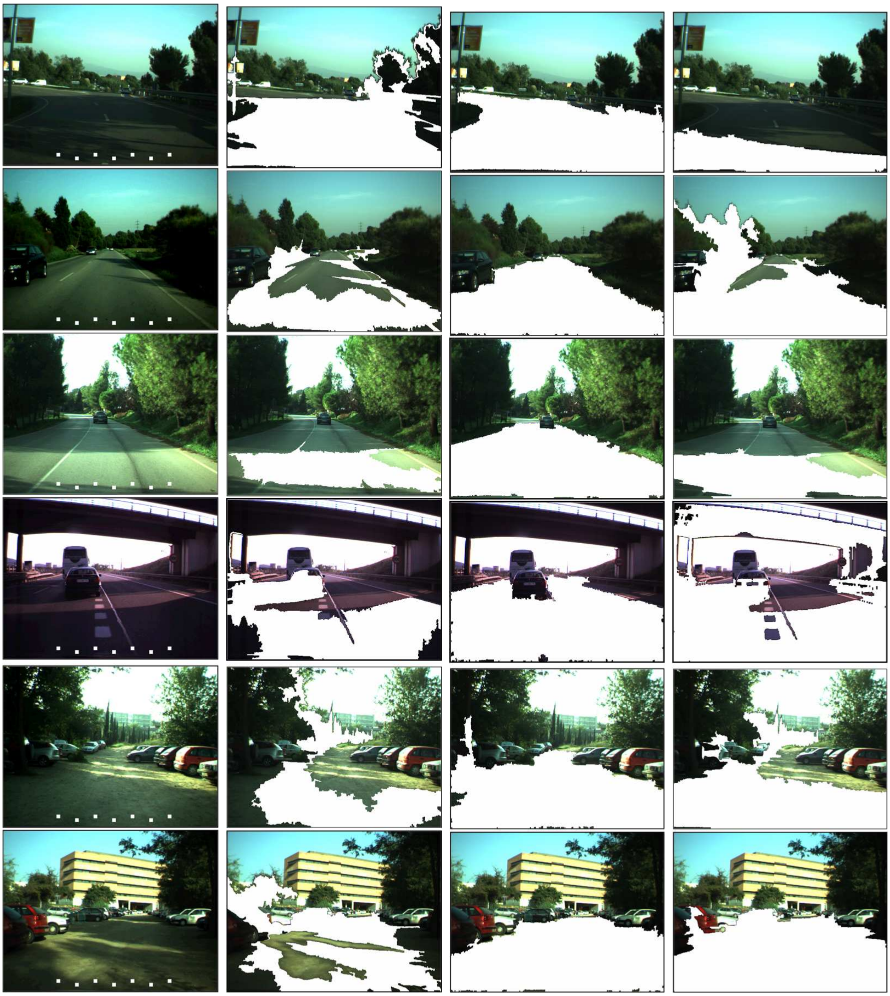
$$\text{for } \ diss_H(\mathbf{p}, \mathbf{p}_n) = |H(\mathbf{p}) - H(\mathbf{p}_n)| \ ,$$

where the different criterion regarding chromaticity is used to take into account the fact that hue value (H) is meaningless when the intensity (I) is very low or very high, or when the saturation (S) is very low. For such cases only intensity is taken into account for aggregation. We use the proposal in [8,1] to define the frontier of meaningful hue, i.e., $\mathbf{p}$ is an achromatic pixel if either $I(\mathbf{p}) > 0.9I_{\max}$ or $I(\mathbf{p}) < 0.1I_{\max}$ or $S(\mathbf{p}) < 0.1S_{\max}$, where $I_{\max}$ and $S_{\max}$ represent the maximum intensity and saturation values, respectively.

In summary, to compute Eq. (1) we use Eq. (2) for $\mathcal{I}$ with threshold $t_{agg,\mathcal{I}}$, Eq. (3) for I with threshold $t_{agg,I}$, and Eq. (4) for HSI with thresholds $t_{agg,ch}$ (chromatic case) and $t_{agg,ach}$ (achromatic case). Figure 4 shows the results obtained for examples of both asphalted and non–asphalted roads. We have manually set the $t_{agg,\mathcal{I}}$, $t_{agg,I}$, and $t_{agg,ch}, t_{agg,ach}$ parameters to obtain the best results for each feature space, but such values are not changed from image to image, i.e., all the frames of our sequences have been processed with them fixed.

These results suggest that $\mathcal{I}$ is a more suitable feature space for road segmentation than the others. Road surface is well recovered most of the times, with the segmentation stopping at road limits and vehicles[1], even with a simple

[1] Other on going experiments, not included here for space restrictions, also show that segmentation is quite stable regarding the chosen aggregation threshold as well as the number and position of seeds, much more stable than both I and HSI.



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

Fig. 4. From left to right columns: (a) original 640 × 480 color image with the seven used seeds marked in white; (b) segmentation using I with $t_{agg,I} = 0,008$; (c) segmentation using $\mathcal{I}$ with $t_{agg,\mathcal{I}} = 0,003$; (d) segmentation using HSI with $t_{agg,ch} = 0,08$, and $t_{agg,ach} = 0,008$. The white pixels over the original image correspond to the segmentation results. The top four rows correspond to asphalted roads and the rest to non–asphalted areas of a parking.

segmentation method. Now, such segmentation can be augmented with road shape models like in [9,10] with the aim of estimating the not seen road in case of many vehicles in the scene. As a result, road limits and road curvature obtained will be useful for applications as road departure warning. The processing time required in non–optimized MatLab code to compute $\mathcal{I}$ is about $125ms$ and $700ms$ for the whole segmentation process. We expect it to reach real–time when written in C++ code.

5 Conclusions

We have addressed road segmentation by using a shadow–free image ($\mathcal{I}$). In order to illustrate the suitability of $\mathcal{I}$ for such task we have devised a very simple segmentation method based on region growing. By using this method we have provided comparative results for asphalted and non–asphalted roads which suggest that $\mathcal{I}$ makes the segmentation process easier in comparison to other popular feature space found in road segmentation algorithms, namely the HSI. In addition, the process can run in real–time. In fact, since the computation of $\mathcal{I}$ only depends on a precalculated parameter, i.e., the camera characteristic angle θ, it is possible that a camera supplier would provide such angle after calibration (analogously to calibration parameters provided with stereo rigs).

Acknowledgments. This work was supported by the Spanish Ministry of Education and Science under project TRA2004-06702/AUT.

References

1. Sotelo, M., Rodriguez, F., Magdalena, L., Bergasa, L., Boquete, L.: A color vision-based lane tracking system for autonomous driving in unmarked roads. Autonomous Robots **16**(1) (2004)
2. Rotaru, C., Graf, T., Zhang, J.: Extracting road features from color images using a cognitive approach. In: IEEE Intelligent Vehicles Symposium. (2004)
3. Ramstrom, O., Christensen, H.: A method for following unmarked roads. In: IEEE Intelligent Vehicles Symposium. (2005)
4. Finlayson, G., Hordley, S., Lu, C., Drew, M.: On the removal of shadows from images. IEEE Trans. on Pattern Analysis and Machine Intelligence **28**(1) (2006)
5. Wyszecki, G., Stiles, W.: Section 1.2. In: Color science: concepts and methods, quantitative data and formulae (2nd Edition). John Wiley & Sons (1982)
6. Gonzalez, R., Woods, R.: Section 10.4. In: Digital Image Processing (2nd Edition). Prentice Hall (2002)
7. Finlayson, G., Hordley, S., Drew, M.: Removing shadows from images. In: European Conference on Computer Vision. (2002)
8. Ikonomakis, N., Plataniotis, K., Venetsanopoulos, A.: Color image segmentation for multimedia applications. Journal of Intelligent Robotics Systems **28**(1-2) (2000)
9. He, Y., Wang, H., Zhang, B.: Color–based road detection in urban traffic scenes. IEEE Transactions on Intelligent Transportation Systems **5**(24) (2004)
10. Lombardi, P., Zanin, M., Messelodi, S.: Switching models for vision-based on–board road detection. In: International IEEE Conference on Intelligent Transportation Systems. (2005)