**Road detection from single images in challenging illumination conditions**

Hadi Ghahremannezhad

Department of Computer Science

New Jersey Institute of Technology

Newark, NJ 07102

**Abstract**

Detecting the road region is one of the most important initial steps in several computer vision applications, such as ADAS (Advanced Driver Assistance Systems) and autonomous driving vehicles. Yet, the process of segmenting the road from non-road regions in a single image is challenging due to various weather and illumination conditions. In this paper, a novel road detection method is proposed in order to overcome these challenges without strong presumptions about the road structure. Specifically, an illumination invariant image is generated to weaken the effects of illumination and shadows. Then the dominant boundaries of the road are estimated in order to limit the leak-segmentation errors. The final road region is estimated by a weighted combination of several probability maps. Experimental results justify the effectiveness of the proposed method.

*Index Terms* – Road detection, Intrinsic image, Illuminant invariance.

# Introduction

Region of interest (RoI) determination is one of the most important and fundamental pre-processing steps in many image and video processing applications. A well-defined region of interest contains all the important location of the image while excluding the unnecessary regions from the processed data in the tasks of image and video analysis. Lowering the amount of required computational resources, increasing the processing speed, and reducing faulty results are some of the benefits of determining an RoI. The focus of this study is automatic RoI determination in traffic videos which is mostly associated with the road region where the vehicles are located. Road region recognition is a crucial step in many computer vision applications such as self-driving vehicles, intelligent driver assistant systems, traffic surveillance, and navigation systems.

There have been many studies addressing the issue of vision-based road recognition in recent years in applications regarding in-vehicle cameras [1]-[3] and traffic video surveillance [4]-[7]. Most recently, convolutional neural networks have attracted a lot of attention in computer vision applications including road segmentation [8]-[13]. However, the need for large amount of training data and computational resources along with lack of sufficient generalization ability, makes it difficult to apply these methods in real-world applications.

In this paper, we propose an adaptive road recognition method that extracts the road location from single traffic frames. First, an illumination invariant gray-scale image is extracted from the RGB image in order to weaken the effects of shadows that decrease the segmentation performance. Second, a number of probability maps are generated based on color differences and gradient distortion. The generated probability maps are combined in a weighted manner and global thresholding is applied on the resulting merged map. Third, the boundaries of the road and the horizontal line are estimated in order to limit the road map from the previous step and avoid possible leak-segmentation errors.

The remainder of this paper is organized as follows. Section II describes the details of generating a gray-scale illumination invariant image in details. Section III demonstrates the approach for generating and combining road probability maps. Section IV describes the details of extracting the road boundaries. The performance of the proposed method is evaluated in section IV, section V contains the concluding remarks and perspectives.

# Generating the illumination invariant image

The shadows on the objects of an image captured by a regular camera have negative effects on most computer vision tasks such as segmentation and object detection, especially in outdoor scenes. Therefore, eliminating or weakening the effects of illumination and shadows as a preprocessing step can improve the performance of vision tasks. One of the main methods for weakening the effects of shadows is to derive a one-dimensional illumination invariant image from the three channel color image based on the relations between the three color values.

Assuming Planckian light source and Lambertian surface for the objects in the natural environment, we can denote the spectral power distribution (SPD) of the light with which is incident on a surface with reflectance . Then the response of the camera sensor is as follows:

Where , is a constant equal to the dot product of the illumination direction and the surface normal at location and is the sensitivity of the kth camera sensor.

If we drop the indices for the locations and assume the camera sensors are based on Dirac delta functions , we would have:

If the illumination is modeled by Wien’s approximation to Planck’s law, the SPD can be demonstrated by its color temperature as follows:

With and being constants being the overall intensity of the light. Therefore, the response of each camera sensor to can be expressed as follows:

If we calculate the ratio chromaticities using the color channels, we would have:

In logarithmic space, we would have:

Which indicates that by varying the illumination (T) the vector moves along a straight line in the log-chromaticity space for each surface. Therefore, by determining the direction of vector , we can specify the changes in illumination which is only camera-dependent and by projecting the vector onto the vector orthogonal to , a one-dimensional grayscale image is generated as follows:

Where the effect of the illumination is weakened.

Here, if we represent the triangular area containing the road samples as , for each road image the RGB values of pixel where are indicated by and the corresponding intrinsic image is calculated as follows:

In order to calculate the angle , we have used the median of four different values based on moment 1, linear regression, moment 3, and principal component analysis (PCA) to estimate a more accurate and general value as follows:

Where is the first principal component. In Figure 1 we can see an example of weakening the shadow effect by projecting the log ratio values onto orthogonal vector to e.

|  |  |  |
| --- | --- | --- |
|  |  |  |
| Original image | Log-chromaticity space | Intrinsic image |

Figure 1. Transformation from RGB to 1D intrinsic image

# Road boundaries extraction

In order to extract the dominant boundaries of the road, first we need to weaken the shadow effects, yet preserving the gradient information corresponding to the material changes. Therefore, we cannot use the intrinsic image from the previous step since it also reduces the amount of gradient information at important edges. Here, we have used another shadow feature which is robust to strong cast shadows in order to reduce the illumination effects while intensify the edges corresponding to material changes. This feature extraction method has less dependence on the camera settings and relies on the fact that road color values in the RGB space are close to each other.

In the most cases, the road surface has relatively similar values in red, green, and blue components whereas the surrounding vegetation has one component considerably higher than the another. This fact can be used as a discriminating feature between shadow edges and the edges corresponding to material changes. By assuming the road to be a homogeneous dielectric surface, the values of pixels of the same material lie on a line passing the origin in the RGB space with a small offset. We can assume a vector for each pixel that belongs to the road surface in the RGB space. The pixels of the same surface that are under the umbra of the shadows fall in a conic area between the lit pixel and the origin (as illustrated in Figure 2).

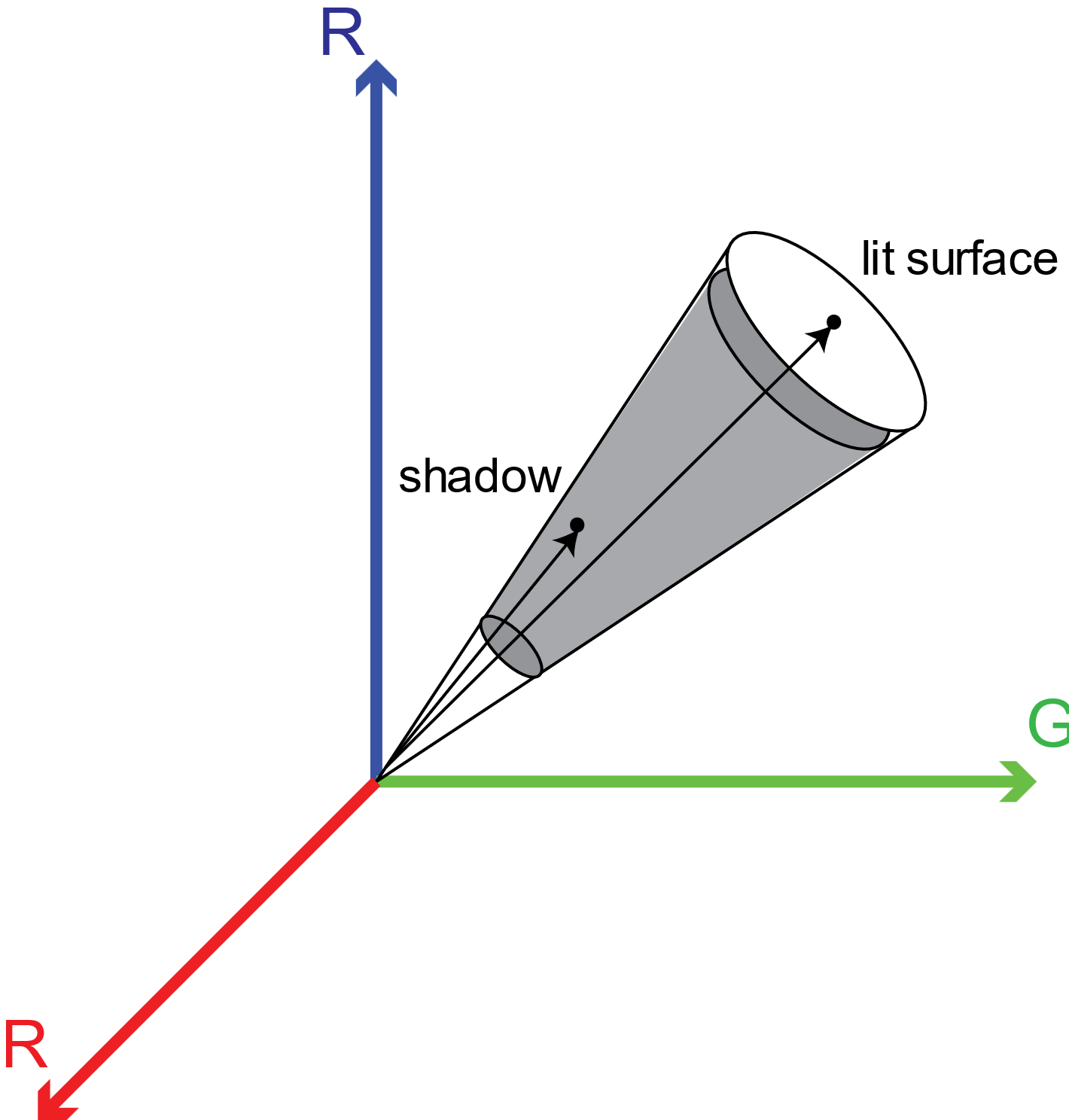


Figure . non-linear attenuation property of shadowed regions

The cause of shadows is the occlusion of sunlight by objects and in the shadow areas the road is illuminated by skylight. On the other hand, we know that in the outdoor scenes the white light emitted from the sun is scattered in all directions by molecules in the air. The Rayleigh scattering effect is higher in the shorter wavelengths such as blue which causes the sky to appear bluish. Whereas the light in the longer wavelengths such as red passes through the atmosphere with less scattering effect. Therefore, we can confidently state that the attenuation is non-proportional due to the ambient illumination which is blue in this case. As the ambient term can have a Spectral Power Distribution (SPD) different from that of incident light, the decrease in luminance when a surface is under shadow is not proportional among the color channels.

Here, we have used the HSV color-space in order to generate a grayscale image where the gradient information is stronger in edges corresponding to material changes in comparison to edges corresponding to illumination changes. The RGB image is converted to HSV color-space where the V component represents the maximum value among the red, green, and blue channels and the S component denotes the saturation and is calculated as follows:

Taking into account that the road surfaces have similar values among the three components, we introduce a feature to weaken the shadow effects while preserving the discriminating properties of material changes as follows:

Where F represents a feature matrix with the same size as the image, b is a bias that is dependent on the camera sensors and can be estimated by polynomial fitting, and is a small positive constant. Some examples of the extracted feature can be seen in Figure 3.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |

Figure . Examples of the proposed shadow feature extractor

After weakening the illumination effects, we can extract the candidate road boundaries by using global thresholding method and only filtering the bottom section of each image followed by a morphological operation and connected component analysis to remove the small blobs which are considered to be noise. Middle-to-side operation is performed in order to extract the pixels corresponding to the road boundaries. A bottom-up scan is applied at each column and the first non-zero pixels are marked as boundary candidates. Then a middle to left and to right scan is performed on the candidate pixels to extract the left and right boundaries, respectively. Finally, Hough transformation is applied to fit a straight line on each boundary and the vanishing point is defined as the intersection of the two lines which is assumed to be located on the horizontal line. These boundaries are later used in order to limit the segmentation process and reduce the amount of leak-segmentation errors. An example of road boundary detection is shown in .



Figure . road boundary detection example

# Road region segmentation

In order to segment the road region, the Chan-Vese segmentation algorithm is applied. This model for active contours is more robust than traditional segmentation methods such as thresholding or gradient based methods. The Chan-Vese model is based on Mumford-Shah function and is mostly used in medical images. In case of road segmentation we have already extracted the road boundaries and a region-growing strategy such as active contour seems to be a choice. The neighboring pixels of a set of initial seed points are examined and decided to be added to the region or not in an iterative manner. In the Chan-Vese model the goal is to minimize the energy defined as the weighted values corresponding to the sum of intensity variations among the pixels inside and outside of the currently segmented region, and a term indicating the arc-length of the region’s boundary.

Here, we first apply the segmentation method on the illumination invariant image extracted in the first step. Then the extracted region is filtered by removing the possible pixels segmented as part of the road region as leak-segmentation error. This error is usually caused by the similarities between the sky and road chromaticity proportions. An example of the iterative region-growing method is illustrated in Figure 4. The initial seed points are chosen from the triangular area assumed to belong to the road region and the region is iteratively growing until the boundaries of the road are covered. After this step the leak-segmentation errors must be removed by using the horizontal line calculated in the previous step.



Figure . iterative region-growing segmentation

The similarities between the sky and road region in terms of chromaticity and color component proportions tends to cause leak-segmentation errors in some images. In order to deal with these type of errors we can apply limiting boundaries from horizontal line estimated in the previous steps. The resulting mask of the segmentation step is intersected with the resulting binary mask of the road boundary detection step in order to remove areas that are outside of the boundaries (such as sky) while preserving the exact boundary points of the road instead of the straight lines.

# Experimental results

The performance of the proposed method is evaluated using several videos. The ROMA dataset is used for testing the performance which contains various illumination and weather conditions and different road types. All the experiments were conducted using a DELL XPS 8900 PC with a 3.4 GHz processor and 16 GB RAM using Matlab R2020a software. The average processing speed for video frames of size pixels was around ~41.28 frames per second, which is inline with real-time requirements of video analysis applications.

The following metrics are used in order to evaluate the quantitative results:

where TP and FP refer to the number of pixels correctly and incorrectly detected as part of the road region, and TN and FN are the number of pixels that are correctly and incorrectly detected as part of the non-road region, respectively. FPR, PRE, REC, ACC, and F1 refer to false positive rate, precision, recall, accuracy, and F-measure respectively. The number of pixels classified as road and non-road are compared with the ground-truth data to calculate each measure. In the images of the dataset, the average FPR was 84%, average accuracy was around 87%, and average F1-score was around 0.82. Figure 5 illustrates some of the results of the road region extraction method. Leak-segmentation errors can be seen in some cases which is due to similarity in texture of the road and surrounding areas.



Figure 6. Sample results of the road region extraction

# Conclusion

Determining the region of interest (RoI) in video analysis applications is an important pre-processing step. An accurate RoI can keep all the essential information while reducing the need for computational resources and eliminating a portion of potential false outputs. In applications of traffic video analysis the RoI usually refers to the road region. In this paper, an adaptive and fully automatic method is proposed for road recognition in real-time. An illumination invariant gray-scale image is generated in order to reduce the illumination effects. Then another gray-scale image is generated where the shadow effect is weakened while the edges corresponding to the boundaries of the road are intensified. Based on this image, the boundaries of the road are extracted in order to limit the segmentation process. In the next step, an iterative segmentation algorithm is applied to fill the road region and finally, the segmentation mask is intersected with the boundary mask in order to define an exact estimation of the road region. This method is not limited to geometrical models since no assumptions have been made about the structure of the road. The extracted road region can further be utilized as the RoI in different traffic video analysis tasks. The experimental results from evaluating the method using real traffic images indicate the feasibility of the proposed method for real-world applications.

# References

[1] C.-K. Chang, J. Zhao, and L. Itti, “Deepvp: Deep learning for vanishing point detection on 1 million street view images,” in 2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2018, pp. 1–8.

[2] T. Kim, Y.-W. Tai, and S.-E. Yoon, “Pca based computation of illumination-invariant space for road detection,” in 2017 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2017, pp. 632–640.

[3] T. Rateke and A. von Wangenheim, “Road surface detection and differentiation considering surface damages,” arXiv preprint

arXiv:2006.13377, 2020.

[4] M. A. Helala, K. Q. Pu, and F. Z. Qureshi, “Road boundary detection in challenging scenarios,” in 2012 IEEE Ninth International Conference on Advanced Video and Signal-Based Surveillance. IEEE, 2012, pp. 428–433.

[5] M. A. Helala, F. Z. Qureshi, and K. Q. Pu, “Automatic parsing of lane and road boundaries in challenging traffic scenes,” Journal of electronic imaging, vol. 24, no. 5, p. 053020, 2015.

[6] M. Santos, M. Linder, L. Schnitman, U. Nunes, and L. Oliveira, “Learning to segment roads for traffic analysis in urban images,” in 2013 IEEE Intelligent Vehicles Symposium (IV). IEEE, 2013, pp. 527–532.

[7] Q.-J. Kong, L. Zhou, G. Xiong, and F. Zhu, “Automatic road detection for highway surveillance using frequency-domain information,” in Proceedings of 2013 IEEE International Conference on Service Operations and Logistics, and Informatics. IEEE, 2013, pp. 24–28.

[8] J.-S. Lee and T.-H. Park, “Fast road detection by cnn-based camera-lidar fusion and spherical coordinate transformation,” IEEE Transactions on Intelligent Transportation Systems, 2020.

[9] L. Caltagirone, L. Svensson, M. Wahde, and M. Sanfridson, “Lidarcamera co-training for semi-supervised road detection,” arXiv preprint arXiv:1911.12597, 2019.

[10] Z. Chen, J. Zhang, and D. Tao, “Progressive lidar adaptation for road detection,” IEEE/CAA Journal of Automatica Sinica, vol. 6, no. 3, pp. 693–702, 2019.

[11] J. W. Perng, Y. W. Hsu, Y. Z. Yang, C. Y. Chen, and T. K. Yin, “Development of an embedded road boundary detection system based on deep learning,” Image and Vision Computing, p. 103935, 2020.

[12] Y. Lyu, L. Bai, and X. Huang, “Road segmentation using cnn and distributed lstm,” in 2019 IEEE International Symposium on Circuits and Systems (ISCAS). IEEE, 2019, pp. 1–5.

[13] N. Y. Q. Abderrahim, S. Abderrahim, and A. Rida, “Road segmentation using u-net architecture,” in 2020 IEEE International conference of Moroccan Geomatics (Morgeo). IEEE, 2020, pp. 1–4.