A Fast Pavement Location Approach for Autonomous Car Navigation

Thiago Rateke^{1,2}, Karla A. Justen², Vito F. Chiarella²,
Rodrigo T.F. Linhares¹, Antonio Carlos Sobieranski^{2,3},
Eros Comunello^{1,2}, and Aldo von Wangenheim^{1,2,3}

¹ Federal University of Santa Catarina,

Graduate Program in Computer Science, Brazil

² Laboratory of Image Processing and Computer Graphics,

National Brazilian Institute for Digital Convergence, Brazil

³ Federal University of Parana, Department of Informatics, Brazil

Abstract. This paper describes a fast image segmentation approach designed for pavement detection in a moving camera. The method is based on a graph-oriented segmentation approach where gradient information is used temporally as a system of discontinuities to control merging between adjacent regions. The method presumes that the navigable path usually is located at specific positions on the scene, and a predefined set of seed points is used to locate the region of interest. The obtained results shown the proposed approach is able to accurately detect in an inexpensive computation manner the navigable path even in non-optimum scenarios such as miss-painted or unpaved dirt roads. Validation was conducted using a dataset with 701 samples of navigable paths, presenting a very high precision for real-time applications.

1 Introduction

Image segmentation has been extensively studied over the past decades in many application areas. Its goal is to divide an input image I into a set of homogeneous regions W , where each W_i represents a particular object on the scene. This process can be performed by grouping similar patterns based on some similarity measure, such as vector norm [1], statistical metrics [2], specific color metrics (HSV, CIE-lab) for better color discrimination, and others. Once obtained, these regions can be used as primary information in high-level algorithms specifically designed for object recognition.

Specifically for the road segmentation problem, three main approaches can be found in the literature: *feature-based*, *model-based* and *lane-region* detection [3]. The feature-based approaches employ of some kind of markers to identify lane boundaries, typically extracted from the scene with gradient information proceeded by Hough transform [4]. Methods based on this approach can produce excellent lane detection results [5], but they are highly dependent on an optimum scenario to be applied [6]. In a non-optimum scenario where the lanes are not entirely painted, methods from this category tend to produce unsatisfactory

results [6]. On the other hand the model-based detection approaches try to use a robust system to predict the road shape from a straight line or a parabolic curve [3][6]. Since this approach is less dependent on the lane markers it normally produces more satisfactory results than the feature-based algorithms, but at some extra computational cost. And even though they may take into account complex roads, as S curves [7], they do not recognize bifurcations, an important requirement for a completely autonomous navigation. The last category, the lane-region detection approaches, attempt to tag each pixel as road or non-road, generally using color and/or texture segmentation methods [3]. These methods normally use a stationary camera and background subtraction although some work has been performed considering a camera in motion [8][9].

It is important for a good road segmentation method to be reliable even in adverse conditions like: faded or unpainted roads, and shadows being projected over the pavement. It should also keep tracking on curved roads [6], since for an autonomous car navigation system curves are the place where the system will most likely be needed. In this paper a real-time navigable path detection approach for road segmentation in video sequences is proposed. Our method is based on a powerful, yet simple, temporal graph-oriented region growing segmentation method that considers the similarity of adjacent pixels to merge similar patterns into the region-of-interest (ROI). We propose a method for locating a road from a vehicle in moving which does not depend on an optimum scenario, such as a specific type of pavement, lanes entirely painted, roads with a unique shape, or interference of external phenomenon (dust, lightness). The proposed approach, in contrast to solutions presented in the literature works in non-optimum real world scenarios, segmenting different types of pavements and unpaved dirt roads with excellent response time. The obtained results were validated against a dataset of real-world scenarios, presenting very high correlation in terms of specificity and sensibility.

2 Fast Pavement Location (FPL) Approach

The proposed approach can be divided into 3 steps: (1) Preprocessing, (2) Image Segmentation and (3) Automated Pavement Location.

For the first step, the pre-processing, the objective is to reduce noise and at the same time preserve discontinuities related to boundaries of the segmented regions. Most related works employ some kind of blur algorithm to smooth regions and reduce noise, however, as such methods are linearly applied over the whole image, they tend to smooth also discontinuities important to control the segmentation process. For this purpose, a combined approach where gradient information and Anisotropic Diffusion Filtering [10] were used to suppress noise and preserve relevant aspects of boundaries for each segment. Anisotropic Diffusion Filtering, different from the isotropic or linear filtering approaches, applies a selectivity kernel over the image which measures adjacent pixels based on a λ value. In other words, this kernel defines which pixels should be merged to smooth this region or preserved (boundary) when compared to a hotspot pixel.

Fig. 1. General overview of the FPL approach. (a) The input image. (b) Preprocessing step combining gradient and Anisotropic Diffusion Filtering. (c) Segmentation result. (d) Navigable path.

Gradient information can be combined in this process by weighting this diffusivity function, allowing higher lambda values in order to reduce the number of iterations and improve the filtering results.

This filtering approach is well-known to be a very time-expensive approach. A GPGPU-based Anisotropic Filtering Approach was implemented, therefore a real-time performance can be achieved. The result of this combined filtering approach is shown in the Figure 1-(b), where the input frame in (a) was submitted to the combined filtering approach.

The second step is the image segmentation applied on the filtered image. The segmentation step aims to divide the image into distinct homogeneous areas, by creating an indexing map that holds a value corresponding to which area some given pixel belongs. This division can be observed in Figure 1-(c), where representative partitions were obtained from the filtered step. The segmentation method proposed in this paper is based on a measure of the similarity for each adjacent pixels, taking into account a gradient network method responsible to determine restrictions during the merging process. These restrictions are applied to ensure that no region will be merged outside of its boundaries, and then a region growing method is applied. The algorithm starts with a gradient network map, where the initial seed point is also defined the start point for the algorithm. A network of similarities is filled over the gradient map, where pixel intensities are measured according to a similarity function defined by [11], to extract 2-dimensional information of the navigable path. After extracted, temporal information is also analyzed where the gradient network method is expanded to the previous frames to provide a very robust segmentation scheme. This is given by $D^2 = \sum_{k=1}^n \Delta C_n$, where n is the dimensionality of the color space used, and $\Delta C = ||C_p - C_s||$ is the color intensity invariant to a specific color space, computed over the past frames. ΔC is the difference between the analyzed pixel

and the seed point for a given channel k , where $\Delta D < th$ is determined by a thresholding value where pixels are considered to be of the same area. When the aforementioned condition is satisfied, the same procedure is applied to the current pixel.

But unlike most region growing algorithms, this method uses an adaptive seed scheme, which means that the values of the seed will be changed to the values of the new pixel of interest. This allows the region to get easily into or out of shadowed regions in the pavement, and since the border image is delimiting the outside of the pavement this does not allow the pavement to be recognized incorrectly under normal circumstances. Without the adaptation of the seed of interest, the algorithm would require a considerably large threshold value for more accurate segmentation. That could cause two problems: the road not being entirely recognized because of light changes and the segmentation extrapolating the road boundaries. When no more pixels can be added to region-of-interest, the algorithm searches for the next seed point. The area map is temporally scanned to find the first non-area point, and when it is found, the previous procedure is applied to it with the correspondent area. This is repeated until all pixels in the image belong to some area.

The last step of the proposed approach is the pavement location itself. The objective is to find which of the previously obtained areas is the navigable path. This step is performed considering that the navigable path will exist within a given region of interest in the image, shown by a red rectangle in Figure 1-(d) used in the first frames. Based on this initial calibration, the gradient network method previously explained is used to find the navigable path temporally. On the other hand, since in all real world applications the camera should be located on the center of the motor hood, the road will always exist inside of that area and presumably occupy it almost entirely. So the simplest approach is to search the most representative area for that rectangle. For images where the motor hood appears, the rectangle can be simply moved upwards, since it will occupy always the same space on the image. Additionally, this rectangle can be automatically updated based on the gradient method.

3 Results

3.1 Experimental Environment and Video Material

Video data were acquired in form of FullHD color videos from three different sets of locations: (a) urban, asphalt-paved roads; (b) suburban streets paved with asphalt sett of hexagonal concrete pavers; and (c) countryside roads that were either paved with asphalt or concrete pavers, or unpaved dirt roads. Unpaved roads were acquired only in locations where there is a clear road boundary, normally given naturally occurring vegetation. The videos used in our experiments were obtained from a GoPro HD camera fixed on the hood of a moving car that recorded several videos under different circumstances. These videos had a 30 frames per second (fps) rate and a 1080p resolution. They were cropped

into shorter films to change their resolution to different values as well as to provide different scenarios for testing. The resolution was reduced to provide both a faster outcome and a more versatile algorithm that could correctly detect the road also in low resolution videos, demonstrated in the next subsection. The results presented in this paper were processed using a Intel(R) Core(TM) i5 CPU, with 6GB of RAM and a Video Board NVIDIA GeForce GT 520 for the Anisotropic Diffusion Filter. Remaining procedures were executed on CPU.

3.2 Experimental Results

The experimental results of the FPL approach were evaluated using different resolutions of the same video to compare the differences in terms of execution time. Since our method is based on region growing a larger image implies in bigger areas to be scanned and segmented, resulting in a slowdown of the execution time. Besides, the location step requires to analyze a specific portion of the image, which for bigger images consequently means more pixels to be verified. This means that the execution time will be affected by the video resolution, but a real-time point can be determined for a comparison of the size of the video against the total amount of frames played in 10 seconds (17 *fps* were achieved), when a resolution of 800x600 pixels is used.

Some difficult scenarios were investigated in our approach. The hood of the car appeared in the video where Figure 2 was taken, the region of interest was moved as shown by the red rectangle in the figures. Furthermore, because the road was clear, there was no need for a diffusion filter and it could be turned off for a faster execution time (22 *fps*).

The FPL approach worked well in several difficult scenarios. In Figure 2-(a), the road changes its color and intensities, and since the difference is not totally abrupt, the algorithm was able to keep track of it as shown in Figure 3-(a); in Figure 2-(b) the road is mispainted and predictive methods based on lines (Hough transform) have a considerable chance to consider a possible turning left, when it was in fact straight; In Figure 2-(c) there is a lack of lane markers on intersections. Methods using the boundaries of lanes in the literature would perform mistakes for the both of these scenarios. However, this did not affect our results as shown in Figures 3-(b) and 3-(c). In Figure 2-(d), a relatively complex scenario is presented. The road is divided into one entrance path and two possible navigable paths. This would confuse most of the methods that work by assuming a lane format, but it was correctly detected by our approach as shown in Figure 3-(d); Most of the approaches limit themselves to asphalt location, while our method presented good results even in unpaved dirt roads such as presented in Figures 2-(e) and 3-(e).

3.3 Results Validation

In order to validate the results obtained by the proposed segmentation approach, a Ground Truth (GT) of a road segmentation is needed [11]. A review of the literature, however, could not find any specific validation approach for a pavement

Fig. 2. (a) Changes in the road color and intensities. (b) Mispainted road lane. (c) Boundaries not clear. (d) Bifurcation region. (e) Dirt Road.

Fig. 3. Road detection of the figures 2(a to e)

detection. On the other side, a set of images recorded from a moving vehicle with their respective manual segmentation for image categorization was found [12], and by manually excluding all non-road objects we obtained a valuable Ground-Truth to validate our approach, as shown in Figure 4-(left side). Also, this work can be used as a base for future comparisons of other lane or pavement detection methods.

Fig. 4. Ground Truth selected images and their road detection. First row: original image GT image. Second row: GT detection. Third row: our detection.

The GT consisted of four sets of 960x720 images named 001E5 (305 images), 001TP (124 images), 006R0 (101 images), and Seq05VD (171 images). All of them were used as a golden standard to test our approach and a chart summarizing the results is presented in Figure 4-(right side). In those results, sensitivity is the percentage of road pixels detected correctly over the total of pixels marked as road, so this measure is the percentage that a given pixel will be correctly detected as a road by our method. While specificity is the same for non-road pixels, meaning that it measures the possibility that a non-road pixel will be tagged correctly.

Some mistaken detection can be verified with this evaluation strategy. Most of them occurred due to poor image quality and pavement discontinuity. Specifically, for the 001TP set of images the low sensitivity obtained was caused because of the low quality of the original images, which presented a significant amount of noise due the high frequencies concentration. Some of the error on the other sets of images was caused because the Ground Truth accounted for all the pavement, while our approach concentrated only on navigable paths, which excluded all the non-continuous pavement from the scene. One of the most representative examples of this limitation (but expected) can be observed in Figure 4-(i), where the region of the pavement was interrupted by the visual division that can be observed in Figure 4-(c).

4 Conclusion and Discussions

In this paper a real-time pavement location method was proposed by combining filtering procedure and a fast temporal image segmentation method. The preliminary obtained results show this method can be used to automatically locate navigable paths in a moving car. Although some limitations can be observed, our method proved itself well adaptive without any parametrization after the initial calibration for a given camera in a given position. The FPL approach

manages to overcome most of the common difficulties existent in distinction of navigable paths with a fast *fps* rate, and works relatively well for very complex scenarios. There is still a necessity of integration of the FPL approach into an autonomous navigation scenario, since we only propose the method for road recognition without contextualizing it to a given scenario. Further research needs to be performed in the integration of this method into embedded computers. The next step of this research is to integrate our approach to receive the video of a radio-controlled car and send signals to control it, and then evaluate its response to several scenarios and adapt the method until it shows acceptable precision. A more detailed description of our results, as well as video references, both of original and post-FPL can be found in the project folder publicly available¹.

References

1. Xia, Y., Feng, D.D., Wang, T., Zhao, R., Zhang, Y.: Image segmentation by clustering of spatial patterns. *Pattern Recognition Letters* 28(12), 1548–1555 (2007)
2. Duda, R.O., Stork, D.G.: *Pattern Classification*, 2nd edn. John Wiley and Sons (2001)
3. Kastrinaki, V., Zervakis, M., Kalaitzakis, K.: A survey of video processing techniques for traffic applications. *Image and Vision Computing* 21(4), 359–381 (2003)
4. Kang, D.-J., Jung, M.-H.: Road lane segmentation using dynamic programming for active safety vehicles. *Pattern Recognition Letters* 24(16), 3177–3185 (2003)
5. Lee, J.W., Yi, U.K.: A lane-departure identification based on lbpe, hough transform, and linear regression. *Computer Vision and Image Understanding* 99(3), 359–383 (2005)
6. Wang, Y., Teoh, E.K., Shen, D.: Lane detection and tracking using b-snake. *Image and Vision Computing* 22(4), 269–280 (2004)
7. Bai, L., Wang, Y., Fairhurst, M.: Multiple condensation filters for road detection and tracking. *Pattern Analysis and Applications* 13, 251–262 (2010), doi:10.1007/s10044-010-0175-9
8. Zhang, L., Zhou, W.-H., Liu, J.-L.: A robust road segmentation method. In: *International Conference on Wavelet Analysis and Pattern Recognition, ICWAPR 2007*, vol. 2, pp. 912–917 (November 2007)
9. Chen, C.-L., Tai, C.-L.: Adaptive fuzzy color segmentation with neural network for road detections. *Engineering Applications of Artificial Intelligence* 23(3), 400–410 (2010)
10. Weickert, J.: *Anisotropic diffusion in image processing* (1996)
11. von Wangenheim, A., Bertoldi, R.F., Abdala, D.D., Richter, M.M.: Color image segmentation guided by a color gradient network. *Pattern Recognition Letters* 28(13), 1795–1803 (2007)
12. Brostow, G.J., Fauqueur, J., Cipolla, R.: Semantic object classes in video: A high-definition ground truth database. *Pattern Recogn. Lett.* 30, 88–97 (2009)

¹ <https://www.dropbox.com/sh/2xdew61tikqq3yn/AAB7I0NytX0qCxUt3B4JzPNda>