A 2D/3D Vision Based Approach Applied to Road Detection in Urban Environments

Giovani Bernardes Vitor, Danilo Alves de Lima, Alessandro Corrêa Victorino, Janito Vaqueiro Ferreira

HAL Id: hal-00936028
https://hal.science/hal-00936028v1

Submitted on 24 Jan 2014

A 2D/3D Vision Based Approach Applied to Road Detection in Urban Environments

Giovani B. Vitor[1,2], Danilo A. Lima[1], Alessandro C. Victorino[1] and Janito V. Ferreira[2]

Abstract— This paper presents an approach for road detection based on image segmentation. This segmentation is resulted from merging 2D and 3D image processing data from a stereo vision system. The 2D layer returns a matrix containing pixel's clusters based on the Watershed transform. Whereas the 3D layer return labels, that are classified by the V-Disparity technique, to free spaces, obstacles and non-classified area. Thus, a feature's descriptor for each cluster is composed with features from both layers. The road pattern recognition was performed by an artificial neural network, trained to obtain a final result from this feature's descriptor. The proposed work reports real experiments carried out in a challenging urban environment to illustrate the validity and application of this approach.

Index Terms— Road Detection, Computer Vision, Image Segmentation, Watershed Transform, V-Disparity Map.

I. Introduction

A vehicle navigating in an autonomous or semi-autonomous way must have some onboarded capabilities, such as the perception of the neighbor environment, to allow the accomplishment of many tasks. Some of this perception capabilities were presented in the DARPA Grand Challenges, competitions promoted by the American's Defence Advanced Research Projects Agency (DARPA) between 2004 and 2007, where unmanned cars should perform autonomous tasks in a desert rally or in an urban environment. Today, with the diffusion of autonomous and semi-autonomous vehicles, a great number of new applications for environment perception have emerged [1]. Although, for real applications, the number of sensors used and its cost must be considered for viability. Based on the principle of viability, the stereo vision sensors have some advantages, because they provide large amount of data, depending of the camera field of view (FOV) and resolution, with a low cost.

Focusing in urban environments perception, several are the vision applications (mono or stereo) successfully applied for this end. Usually, the mono vision is applied in image segmentation (to detect roads, vehicles and pedestrians) and primitives extraction (like traffic signs and land marks). In the domain of image segmentation, the regions are defined according with some similarity. For road detection the most common segmentations are based on colour ([2], [3], [4]), texture ([5]) and intensity ([6], [7]). In the case of stereo vision applications, the 3D information of the environment is typically used to estimate free spaces and obstacles, with
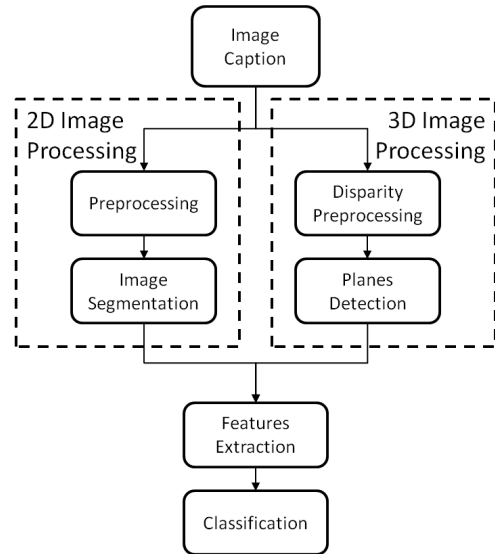
Fig. 1. Solution block diagram.

specific techniques as the V-Disparity Map [8], [9] for example.

However, the stereo vision in urban environments, manly for road detection, must deal with different noise sources, as shadow, road texture, light variations, etc., that difficult the elaboration of the disparity map. Furthermore, due to the high complexity of the urban environments, because of the presence of many different elements like cars, pedestrians, trees, etc., classify a road profile from a 2D image segmented data is also a hard task. As a solution for these problems, some works enhance the 3D information with the 2D, improving the classification and detection of road segments [10], [11]. It is also important to mention that there are many other approaches to accomplish the road detection [12], [13], [14], but they diverge from this work on the sensors used.

This work will address the problem of merging a 2D image segmentation and a 3D image processing data to realize the road detection. Differently from the works [10], [11], the 2D segmentation technique used will be based on the Watershed transformation, which features will be combined with the v-disparity classification elements to compose the feature's descriptor to an artificial neural network (ANN). The usage of an ANN is to better adapt the classification results to different urban environment conditions. The Figure 1 presents a block diagram resuming the steps of this solution.

All the blocks in Figure 1 will be detailed in the next sections of the paper that is divided as followed: Section II

Fig. 2. Example of a stereo pair of an urban view.

presents the 2D and 3D image processing; Section **III** and **IV** presents respectively the features extraction, merging 2D and 3D data, and the artificial neural network used; details of the set-up used and some results obtained are in Section **V**; and, finally, Section 6 presents conclusions and perspectives for future work.

II. IMAGE PROCESSING

As described in the Figure 1, the solution proposed in this article starts with the image caption step to serve the 2D and 3D image processing layers. This was accomplished using a calibrated and synchronized stereo vision system, given two images (an stereo pair), showed in Figure 2. The image data were rectified resulting a left and right images. For the 2D image processing layer were used only the left images. However, to allow the 3D information extraction, the both rectified images are provided to the 3D image processing layer. These layers are presented as follow.

A. 2D Image Processing

All modules in the 2D image processing layer take advantage in the domain of morphological image processing [15], which received considerable attention in the past few decades especially after their theoretical foundations have been demonstrated. These theoretical foundations are known as connected operators [16] and geodesic reconstruction process [17]. The interest above all is due to its fundamental property of simplifying an image without corrupting contour information [18]. The 2D image processing layer performs two modules calculation, the Preprocessing and the Segmentation, that will be explained below.

1) Preprocessing Module:
This preprocessing module is responsible to prepare the rectified image to be segmented and is the key to determine how will be this result. In this work were applied tree different filters: the morphological gradient to obtain the high frequency image; the morphological reconstruction with attribute *Area Closing* to filter out areas smaller than a given threshold; and the morphological reconstruction with attribute *Hmin* which eliminates the local minima.

a) Morphological Gradient: In general, gradient operators are used for image segmentation because they enhance intensity variations, also called as *edge detectors* [19]. Between many gradient operators as Sobel, Prewitt and Roberts, the chosen one was the Morphological gradient due its good

approximation of edges in urban environments, focus of this application .

This filter detects the intensity variations of pixel values in a given neighbourhood. It is obtained by the arithmetic difference between an extensive operator and an anti-extensive operator. The classical one is defined as (1):

$$gradMorph(f) = (f \oplus g_e) - (f \ominus g_i) \qquad (1)$$

where f is the image function, g_e and g_i are structuring elements centered about the origin, and the operators $\oplus$ and $\ominus$ are respectively dilation and erosion.

b) Morphological Reconstruction Area Closing: Since area closing is seen as the complement of area opening, by simplicity, here is showed the definitions to area opening [20].The conception of this filter is to remove from a binary image its connected components with area smaller than a parameter λ. Before demonstrating the grayscale area opening, some definitions should be done in binary area opening. First, the connected opening $C_x(X)$ of a set $X \subseteq M$ at point $x \in X$ is the connected component of X containing x if $x \in X$ and 0 otherwise, and M being the binary image, $M \subset \mathbb{R}^2$. In this mode, the binary area opening is then defined on subsets of M [21].

$$\gamma_\lambda^a(X) = \{x \in X \mid Area(C_x(X)) \geq \lambda\} \qquad (2)$$

The $\gamma_\lambda^a(X)$ denote the morphological area opening with respect to the structure element a and the parameter λ. The $Area(.)$ is the number of elements in a connected component of $C_x(X)$. Its dual binary area closing is obtained as:

$$\phi_\lambda^a(X) = [\gamma_\lambda^a(X^c)]^c \qquad (3)$$

where X_c denotes the complement of X in M. Extending the filter for a mapping $f : M \to \overline{\mathbb{R}}$, in case to grayscale image, then the area opening $\gamma_\lambda^a(f)$ is given by:

$$(\gamma_\lambda^a(f))(x) = sup\{h \leq f(x) \mid x \in \gamma_\lambda^a(T_h(f))\} \qquad (4)$$

In equation 4, $T_h(f)$ represent the threshold of f at value h:

$$T_h(f) = \{x \in M \mid f(x) \geq h\} \qquad (5)$$

As mentioned before, the complement of equation 2 can be similarly extended to the conception of area closing to mappings from $M \to \overline{\mathbb{R}}$.

c) Morphological Reconstruction Hmin: In another point of view, the grayscale morphological reconstruction can be obtained by successive geodesics dilations. This principle employs two subsets of $\mathbb{R}^2$, called *mask image* and *marker image*. Both subsets must have the same size. Moreover, the *mask image* must have intensity values higher than or equal to those from *marker image* [17]. Properly performing the reconstruction *H-maxima* or *Hmax*, is possible to take the *H-minima* or *Hmin* from its complement. Mathematically, defining *mask image* as I and *marker image* as $I - h$, being h the height, the equation is given by [22]:

$$Hmax_h^a(I) = I\Delta_a(I - h) \qquad (6)$$

In this definition, Δ stands for morphological reconstruction with the structure element a. By duality, the *Hmin* is defined as:

$$Hmin_h^a(I) = [I^c \Delta_a (I^c - h)]^c \qquad (7)$$

2) Segmentation Module:
The image features detection needs, in most cases, a segmentation process, some of them with algorithms based on segmentation by discontinuity or similarity. Segmentation itself is not a trivial task, being among the hardest ones for image processing. This module apply a *Watershed transform* based on local elements, named as Local Condition Watershed Transform (LC-WT) [23], with the purpose of mimics the behaviour of a drop of water on a surface. This definition seems to be the steepest descent paths, where the neighbours information is used to create a path to the corresponding minimum, through an arrowing technique.

The arrowing is the algorithmic representation of the drop of water, where, for every pixel in the image(taking an image as a graph), an arrow is drawn from the current to the next one, which creates a path that ultimately leads to a regional minimum. The arrow points to the direction that a drop of water would flow, considering the image as a surface. For the LC-WT definition, every pixel that does not belong to a regional minima will have one and only one arrow. The union-find technique for Watershed transform is based on the algorithm for disjoint sets. Given that the regions of the output image form a partition and these are disjoint by definition, the union-find algorithm process paths to identify the roots - or representatives - for every pixel. At the end, all pixels that falls into the same minimum are labelled with its representative.

The result of 2D image processing layer can be seen in Figure 3. It demonstrates all processing done by this layer. It is important to notice that the λ and h parameters of the preprocessing module give an excellent flexibility to determine the segmentation result of the Watershed transform which is responsible to generate the representatives samples that will be classified. As mentioned early, this is the key of 2D image processing layer.

B. 3D Image Processing

1) Disparity Preprocessing:
The rectified images received from the Image Caption block allow the use of the Epipolar geometry [24]. With this information, the disparity map ($\triangle$) of the stereo pair was built using the Sum of Absolute Differences (SAD) correlation algorithm. A $\triangle$ image from the stereo pair of Figure 2 can be seen in the Figure 4. In this figure, the closest region to the camera are in light grey and far way region are in dark grey. These grey values (1-255) are the disparity for each pixel and they are related to the distance between the camera and the point in the world defined as:
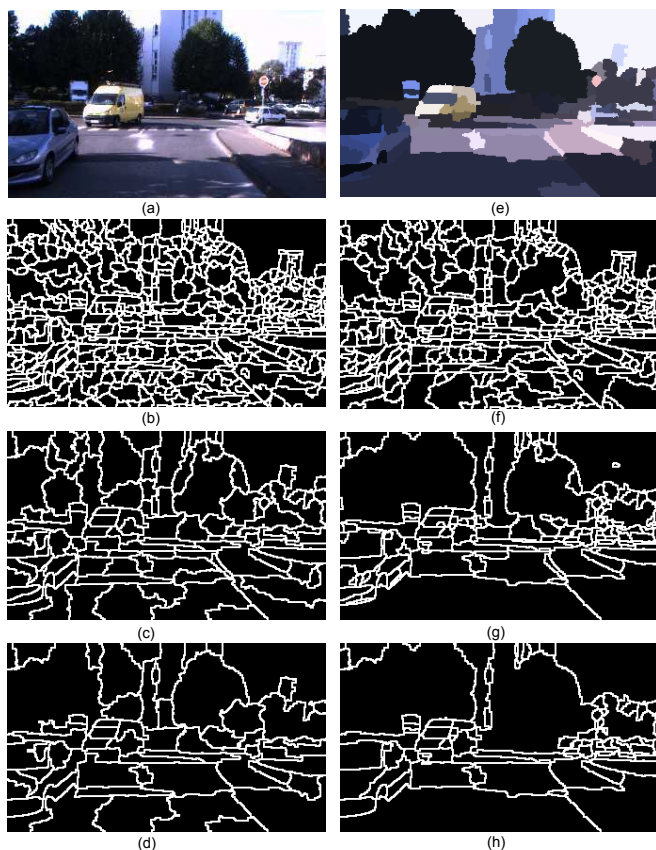
$$Z = \frac{fB}{d} \qquad (8)$$

Fig. 3. The 2D image processing layer result with the demonstration of λ and h parameters. Original image in (a), Showing the RGB mean feature in (e), The segmentation result fixing $h = 1$ and using λ as 5(b) 35(c) 65(d), and the segmentation result fixing $\lambda = 5$ and using h as 2(f) 10(g) 20(h)

Fig. 4. The disparity map from the stereo pair of Figure 2 (left), and the v-disparity map (right).

where f is the focal length of the camera, B is the camera baseline, and d is the disparity value. Others elements observed in the $\triangle$ image is the noise cause by urban environments information. They are mainly caused by shadows, light reflection, and low texture variance, which difficult the correlation algorithm to find a right pixel correspondence in the stereo pair. Some of these noises are detected and eliminated as black points (0 value) in the $\triangle$, the remaining must be worked out to minimize their effect in the planes detection step.

Each disparity map ($\triangle$), constructed with the stereo images, are now converted in a v-disparity map ($I_{v\triangle}$) [10]. Basically, the $I_{v\triangle}$ image parameters are: the *Rows*, which are the same of the $\triangle$ image; the *Columns*, that represent
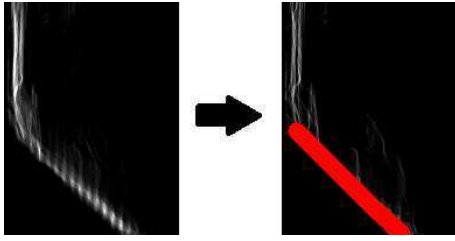
Fig. 5. The v-disparity map from the moving average technique (left), and detected planes (red lines) for the free space (right).

Fig. 6. The original v-disparity map less the detected free space (left), and the obstacles detected (red lines, right).

the disparity value (grey scale) of the △; and the *Pixels Values*, storing the number of pixels with the same disparity in the row analysed.

A side-by-side view between △ and its respective $I_{v\triangle}$ can be seen in Figure 4.

2) Planes Detection:
The v-disparity technique was chosen to allow an easy classification for the image data into drivable and non-drivable areas (obstacles). As described in [8], [10], the world environment can be approximated by horizontal and vertical planes related to the camera. In the disparity map (△), this mean that the road has a continuous variation on its disparity values along the lines, while the obstacles have approximately the same disparity values on the lines. These both effects are represented as small lines with a slope bigger than 90°, in a non-flat world representation, and the obstacles have a slope close to 90° in the v-disparity map ($I_{v\triangle}$), as viewed in the Figure 4.

However, the noisy data, derived from a wrong correlation in △, causes discontinuities in the line segments of the $I_{v\triangle}$, which prevent any straight line extracting procedure, such as the Hough transform, to work as well without detecting the noises. Analysing a $I_{v\triangle}$ image sequence of the environment, is possible to note that the lines slope, that represent the drivable planes, changes smoothly during the time. So, to overcome these wrong data variation, was applied the moving average technique in the $I_{v\triangle}$, as follow:

$$I_{v\triangle A}(i+1) = a * I_{v\triangle} + (1-a) * I_{v\triangle A}(i) \qquad (9)$$

where $I_{v\triangle A}$ is the average image from the $I_{v\triangle}$, and a is the weight. This resulted image are now used to detect the free space, as showed in the Figure 5. The obstacles are detected in the original $I_{v\triangle}$ image less the drivable planes found. This image and the resulted obstacles segments are in Figure 6. The final mapping of these detected elements are in the Figure 7, where the original colours are the free space, the red pixels are the obstacles, and the yellow ones are elements without any classification.

III. FEATURES EXTRACTION

The 2D and 3D image processing, presented in the Section II, returned two data sets: one with segmented regions (clusters) and other with free spaces, obstacles and non-classified data respectively. As seen on the image segmentation (Figure 3), the clusters fit closely the objects contours,
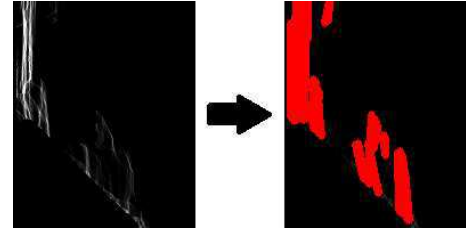

Fig. 7. Final mapping of the free space (original colours), obstacles (red pixels), and elements without any classification (yellow pixels).

but without any label telling what they are. In addition, the free space and obstacle data have imprecise limits (Figure 7), which do not allow the distinction between them. The non-classified areas and wrong classifications have also to be minimized to guarantee the usability of this solution on safe navigation tasks, for example. In this work, the cluster's labels are defined by an artificial neural network classifier (ANN), that will be presented in the Section IV, basing on its features. This section describes some of these possibles features, extracted to each cluster.

The first set of features is based on the intersection of the 3D information with the 2D segmentation. For each cluster, the features are the percentage of free space, obstacles, and non-classified pixels from the 3D image processing result. The second set of features is based on statistical measures, as those presented in [5] like mean, probability, entropy and variance, where they formulations can be found. These values were calculated by the RGB[1] and HSV[2] colour space values of every image cluster.

IV. ARTIFICIAL NEURAL NETWORKS CLASSIFIER

Artificial Neural Network (ANN) has been utilized in several applications as a good tool for data classification. Its wide utilization is based in three fundamentals properties: adaptability, ability to learn by examples and ability of generalization [5]. For this work, was used a Multilayer Perceptron (MLP) to realize a non-linear input-output mapping [25].

The network training is based on the Backpropagation technique and its structure was projected with three layers, being the input and output layers and the hidden one. Normally, the size of the input layer corresponds to the

---

[1] Abbreviation for red, green and blue colour space.
[2] Abbreviation for hue, saturation and value colour space.

number of features extracted, as presented in Section III and which had five neurons in its final version, that will be better explained in Section V. The hidden layer was defined with fifteen neurons, where all neurons use the sigmoid activation function.

The output layer was developed to classify the feature's descriptor as road surface or non-road surface (obstacles, sky, etc.). For this situation is supposed to have only two neurons, but, considering the complexity of the environment to be classified and the different conditions of an urban scenario, was adopted the subclass strategy. It takes three subclass such land marks, normal area and shadow area to represent the road surface. Thus, the output layer has four neurons to classify the feature's descriptor as road surface or non-road surface, which provide responses in decimal values between 0 to 1. The classification result is chosen by the higher output neuron value, where its difference to the other ones is higher than a given threshold, otherwise its classified as unknown.

V. EXPERIMENTAL RESULTS

In this section is presented some qualitative results whose experiments demonstrate the efficiency and robustness of the presented approach. The methodology was tested with a data set acquired in the experimental car CARMEN at the Heudiasyc laboratory (UMR CNRS 7253), in context of project VERVE[3], equipped with a Bumblebee X3 camera installed on its top. The camera's range is up to 70 meters, has a field of view of 66° and the final resolution used was 300x400 pixels. The vehicle was conducted in an urban environment under different conditions, in order to expose the approach presented here to many urban conditions.

The training mode for the ANN classifier concerned about keeping the proportion of samples to each class to perform the cross validation method. The data were divided in train, test and validation set, with respectively 50%, 30%, and 20% of the total, in this case 6924 samples. Within the sets, the amount of data representing the road surface is the same for non-road surface, even if there is subclasses associated to one group. The feature's descriptor was composed for each image cluster by: the 3D image processing data, representing the percentage of free space and obstacles, and the 2D image processing data, based on the best results of [5], that are the mean for the green and hue channels, and the entropy for the hue and value channels.

Some particular results for the image processing layer were already presented in the Subsections II-A and II-B. The complete classification using the ANN for different environment conditions can be seen in the Figures 8 and 9. In these images, the original colour represent the road detected, the red pixels are the non-road surfaces (obstacles, sky, etc.) and the yellow ones represents the unknown areas, where the ANN classification is not in accordance with the output threshold, for these cases equal to 0.2.

(a) The detection result of the highway with obstacle.

(b) The detection result of the road with shadow and obstacle.

Fig. 8. The original image on top and the classification results on bottom for several conditions, where original colour represents the road detected, the red pixels are the non-road surfaces, and the yellow ones representing the unknown areas.

Fig. 9. The original image on top and the classification results on bottom for extreme conditions, where original colour represents the road detected, the red pixels are the non-road surfaces, and the yellow ones representing the unknown areas.

The Figure 8 shows the detection results in a highway and urban streets with obstacles, land marks, and shadow areas. From the results, is possible to see that the method proposed can solve the problem of many different obstacles such as vehicles, pedestrian, trees, road barriers and also the shadow areas of them. However, some segments of road barriers, shadow areas, high-light areas are hard to be classified as road surface or non-road surface, manly due to the feature's descriptor that is not so good to discriminate its class/subclass. Moreover, in some cases image segmentation merges segments with low intensity values, as shadows, mixing a road area with an obstacle area.

The Figure 9 demonstrates the road detection approach exposed to extreme conditions. Even in hard situations of sun and shadows, it is possible to acquire good approximations for the road area. For most of the road conditions our method works well, but improvements must be done in sense of become the system more robust under extreme conditions as shown. The complete result under an image sequence can be seen in [26].

VI. Conclusions and Future Works

This paper presented a 2D/3D vision solution applied to the road detection problem. The 2D vision solution performs a image segmentation, used to extract some statistical data, and the 3D information (stereo vision data) calculates free spaces and obstacles. The image was segmented using a Watershed transformation and the stereo vision data was processed using the V-Disparity technique. The result were merged to compound a feature's descriptor for each image cluster. The road was detected applying these features in an artificial neural network classifier, which reached a accuracy of 93.30%.

As described in the Section V, the solution behaved robustly, presenting good results in different urban situations. Elements which normally has problems to be classified, like low texture variance, shadows and lighting effects, were deal with the neural network classifier. In general, when the V-Disparity technique detects the free space and obstacles with no problem or, in doubt cases, defines as non-classified, the other features extracted from the segmented image were enough to guarantee right classifications. The wrong cases will be treated in future works, improving, for example, the obstacle detection with the U-Disparity technique [13].

In this first moment, the efficiency has not been considered to realize the experiments showed here. The focus was to check the validity of these approach to be used in urban environments. The next steps will be to apply this solution on real navigation tasks, using the segmented data to help in the detection of different urban elements, as vehicles, pedestrians, traffic signs, land marks, etc. For reaching this, the processing time of each step of Figure 1 will be improved, using applications, such the one presented in [27], for GPU cards.

Acknowledgement

References

[1] J. R. McBride, J. C. Ivan, D. S. Rhode, J. D. Rupp, M. Y. Rupp, J. D. Higgins, D. D. Turner, and R. M. Eustice, "A perspective on emerging automotive safety applications, derived from lessons learned through participation in the darpa grand challenges," *Journal of Field Robotics*, vol. 25, pp. 808–840, October 2008.

[2] G. Zhang, N. Zheng, C. Cui, Y. Yan, and Z. Yuan, "An efficient road detection method in noisy urban environment," in *Intelligent Vehicles Symposium, 2009 IEEE*, june 2009, pp. 556 –561.

[3] J. Alvarez and A. Lopez, "Road detection based on illuminant invariance," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 12, no. 1, pp. 184 –193, march 2011.

[4] A. Rahman, B. Verma, and D. Stockwell, "An hierarchical approach towards road image segmentation," in *Neural Networks (IJCNN), The 2012 International Joint Conference on*, june 2012, pp. 1 –8.

[5] P. Y. Shinzato and D. F. Wolf, "A road following approach using artificial neural networks combinations," *J. Intell. Robotics Syst.*, vol. 62, no. 3-4, pp. 527–546, June 2011. [Online]. Available: http://dx.doi.org/10.1007/s10846-010-9463-2

[6] M. Bilodeau, X. Yu, and S. Beucher, *Road segmentation by watersheds algorithms*, ser. CMMMM. Fontainebleau: ENSMP, 1990.

[7] X. Yu, S. Beucher, and M. Bilodeau, "Road tracking, lane segmentation and obstacle recognition by mathematical morphology," *IEEE Transactions on Intelligent Vehicles Symposium.*, vol. 1, pp. 166–172, 1992.

[8] R. Labayrade, D. Aubert, and J. P. Tarel, "Real time obstacle detection in stereovision on non flat road geometry through "v-disparity" representation," in *Proceedings of the IEEE Symposium on Intelligent Vehicles*, vol. 2, 2002, pp. 646–651.

[9] A. Broggi, C. Caraffi, R. I. Fedriga, and P. Grisleri, "Obstacle detection with stereo vision for off-road vehicle navigation," in *Proceedings of the International IEEE Workshop on Machine Vision for Intelligent Vehicles*, 2005, pp. 1–8.

[10] N. Soquet, D. Aubert, and N. Hautiere, "Road segmentation supervised by an extended v-disparity algorithm for autonomous navigation," in *Proceedings of the IEEE Symposium on Intelligent Vehicles*, 2007, pp. 160–165.

[11] F. Dornaika, J. Alvarez, A. Sappa, and A. Lopez, "A new framework for stereo sensor pose through road segmentation and registration," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 12, no. 4, pp. 954 –966, dec. 2011.

[12] S. Nedevschi, R. Danescu, T. Marita, F. Oniga, C. Pocol, S. Sobol, C. Tomiuc, C. Vancea, M. Meinecke, T. Graf, T. B. To, and M. Obojski, "A sensor for urban driving assistance systems based on dense stereovision," in *Intelligent Vehicles Symposium, 2007 IEEE*, june 2007, pp. 276 –283.

[13] M. Teutsch, T. Heger, T. Schamm, and J. Zollner, "3d-segmentation of traffic environments with u/v-disparity supported by radar-given masterpoints," in *Intelligent Vehicles Symposium (IV), 2010 IEEE*, june 2010, pp. 787 –792.

[14] J. Moras, F. Rodriguez, V. Drevelle, G. Dherbomez, V. Cherfaoui, and P. Bonnifait, "Drivable space characterization using automotive lidar and georeferenced map information," in *Intelligent Vehicles Symposium (IV), 2012 IEEE*, june 2012, pp. 778 –783.

[15] J. Serra, *Image Analysis and Mathematical Morphology*, 1st ed. Academic Press, 1982.

[16] J. C. Serra and P. Salembier, "Connected operators and pyramids," in *Image Algebra and Morphological Image Processing IV*, 1993, pp. 65–76.

[17] L. Vincent, "Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms," *IEEE Transactions on Image Processing*, vol. 2, pp. 176–201, 1993.

[18] J. Darbon and C. B. Akgul, "An efficient algorithm for attribute openings and closings," in *In Proceedings of the 13th European Signal Processing Conference (EUSIPCO), Electronic proceedings*, 2005.

[19] J.-F. Rivest, P. Soille, and S. Beucher, "Morphological gradients," *Journal of Electronic Imaging*, vol. 2, pp. 326–336, 1993.

[20] A. Meijster and M. H. F. Wilkinson, "A comparison of algorithms for connected set openings and closings," *IEEE TRANS. PATT. ANAL. MACH. INTELL*, vol. 24, no. 4, pp. 484–494, 2002.

[21] L. Vincent, "Morphological area openings and closings for grey-scale images," in *Proc. of the Workshop on Mathematical Morphology and its Applications to Signal Processing*, Barcelona, Spain, 1993, pp. 22–27.

[22] E. R. Dougherty and R. A. Lotufo, *Hands-on Morphological Image Processing (SPIE Tutorial Texts in Optical Engineering Vol. TT59)*. SPIE Publications, July 2003.

[23] R. Audigier and R. de Alencar Lotufo, "Relationships between some watershed definitions and their tie-zone transforms," *Image Vision Comput.*, vol. 28, no. 10, pp. 1472–1482, Oct. 2010. [Online]. Available: http://dx.doi.org/10.1016/j.imavis.2009.11.002

[24] O. Faugeras, *Three-dimensional computer vision: A geometric view point*. Cambridge: MIT Press, 1993.

[25] S. Haykin, *Neural Networks: A Comprehensive Foundation (2nd Edition)*, 2nd ed. Prentice Hall, July 1998. [Online]. Available: http://www.worldcat.org/isbn/0132733501

[26] "The complete result of this paper to an image sequence." [Online]. Available: http://www.youtube.com/watch?v=rhx8poyLGmY

[27] A. Korbes, G. Vitor, R. de Alencar Lotufo, and J. Ferreira, "Advances on watershed processing on gpu architecture," in *Mathematical Morphology and Its Applications to Image and Signal Processing*, ser. Lecture Notes in Computer Science, P. Soille, M. Pesaresi, and G. Ouzounis, Eds. Springer Berlin / Heidelberg, 2011, vol. 6671, pp. 260–271.