



Deposited via The University of Leeds.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/157459/>

Version: Accepted Version

---

**Proceedings Paper:**

Saleem, NH, Rezaei, M and Klette, R (2017) Extending the stixel world using polynomial ground manifold approximation. In: 2017 24th International Conference on Mechatronics and Machine Vision in Practice (M2VIP). 2017 24th International Conference on Mechatronics and Machine Vision in Practice (M2VIP), 21-23 Nov 2017, Auckland, New Zealand. IEEE. ISBN: 978-1-5090-6546-2.

<https://doi.org/10.1109/m2vip.2017.8211440>

---

©2017 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

**Reuse**

See Attached

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

# Extending the Stixel World Using Polynomial Ground Manifold Approximation

Noor Haitham Saleem

*School of Engineering, Computer and  
Mathematical Sciences, EEE Dep.  
Auckland University of Technology  
Auckland, New Zealand*

Mahdi Rezaei

*Faculty of Computer,  
Information Technology Engineering  
Qazvin Islamic Azad University  
Qazvin, Iran*

Reinhard Klette

*School of Engineering, Computer and  
Mathematical Sciences, EEE Dep.  
Auckland University of Technology  
Auckland, New Zealand*

**Abstract**—Stixel-based segmentation is specifically designed towards obstacle detection which combines road surface estimation in traffic scenes, stixel calculations, and stixel clustering. Stixels are defined by observed height above road surface. Road surfaces (ground manifolds) are represented by using an occupancy grid map. Stixel-based segmentation may improve the accuracy of real-time obstacle detection, especially if adaptive to changes in ground manifolds (e.g. with respect to non-planar road geometry). In this paper, we propose the use of a polynomial curve fitting algorithm based on the  $v$ -disparity space for ground manifold estimation. This is beneficial for two reasons. First, the coordinate space has inherently finite boundaries, which is useful when working with probability densities. Second, it leads to reduced computation time. We combine height segmentation and improved ground manifold algorithms together for stixel extraction. Our experimental results show a significant improvement in the accuracy of the ground manifold detection (an 8% improvement) compared to occupancy-grid mapping methods.

## 1. Introduction

Applications of computer-vision based driver-assistance systems are increasing every year thanks to their great potentials to improve traffic efficiency, to reduce accident rates, and to enhance driver’s comfort, even in heavy traffic. According to the 2016 “International Transport Forum Report”, the number of pedestrians dying in traffic accidents has greatly reduced over the past few years [1]. Studies have also identified different approaches and technologies that have their individual advantages and drawbacks. For instance, radar and LiDAR, are probably two of the most accurate sensor technologies available that support driver-assistance systems for obstacle detections [2]. Due to higher production costs of such sensors, besides the working principles, the priority is given to metal reflectors, and therefore camera-based systems are more preferable [4]. In fact, *vision-based driver assistance* (VBDA) systems are very likely to become the most widely used systems, as cameras are a relatively inexpensive and reliable sensor, and they can be installed or embedded in any car [15], [16]. Using cameras for stereoscopic vision is an advanced field of research. In the

last few years it has made a remarkable leap. Subjects related to stereoscopic vision, such as scene analysis [5], feature descriptors [6], learning [7], and reducing processing time [8], are still under consideration.

Blending *stixels* (i.e. defining a thin column of defined height on a base rectangle, with fixed pixel width and in vertical pose) with semantic segmentation may result in efficiency for global optimized segmentation for scene analysis. The *stixel world* [3] is constructed by cascading multiple independent techniques: mapping disparities to occupancy grids, ground plane computation, height segmentation, and a final stixel extraction step. This can be used to effectively model the scene content of arbitrary 3D traffic environments [9]. For exploring stixels, one way is to calculate the distances between obstacles and the camera based on an occupancy grid [2], [3]. The calculation of the *base-point* of an obstacle, i.e. where it touches the ground, is not a trivial projection equation due to existence of non-planar road surfaces. Research efforts in this field need to seek alternatives which can lead to reliable ground plane estimation and base-points of stixels.

The *free space* [10] is defined as a road area without any obstacle, i.e. regions ahead of the ego-vehicle where the vehicle may potentially drive in safely (e.g. within the next, say, 3-5 seconds). The *ego-vehicle* is the vehicle in which the system is operating in. A proper and accurate detection of free space is important for intelligent transportation control [11], to prevent a collision of the ego-vehicle with any obstacle, or to assist a blind pedestrian [12]. Free-space analysis in a traffic scene can be solved by vision-based control mechanisms, designed towards autonomous cars, either using monocular or multi-ocular vision. Applications for fine-grained, responsive vision-based autonomous car operations require better accuracy to provide location estimates in real time to avoid traffic collisions, for traffic safety, and for driver’s comfort [13].

*Advanced driver assistance systems* (ADAS) aim to provide a better understanding of the environment in order to improve traffic safety and efficiency [14], and may use other types of sensors, not just cameras. In this context, VBDA have emerged as a significant contribution to ADAS [15], [16]. Different computer vision technologies can support a VBDA to estimate the free space either by using monocular

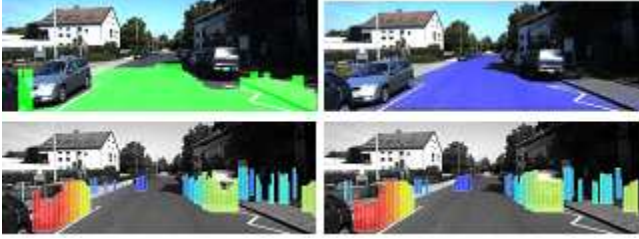


Fig. 1. Stixel world. *Top-left*: Ground manifold estimated via occupancy grid (i.e. a plane). *Top-right*: Polynomial fitting of a ground manifold as estimated in this paper. *Bottom-left*: Stixels via occupancy grid which do not fit to obstacles in terms of distances. *Bottom-right*: Stixels via polynomial fitting as proposed in this paper. Green (see *top-left*) illustrates a ground plane, and purple (see *top-right*) a possibly curved ground manifold (by polynomial fitting).

or multi-ocular vision. Due to variations in traffic scenes (e.g. weather, road conditions, road geometry, or traffic density) this remains an active field of research [14]. See Fig. 1.

The road surface is not a perfect plane in general. We may even expect more irregularities in slope changes than represented by a continuous curve. Thus, the common stixel calculation using the assumption of a ground plane can lead to many false obstacle detections. Figure 1, bottom-left, shows an example.

To gain more accuracy, we consider (as a possible alternative) polynomial line fitting for extracting a piecewise best line fit to  $v$ -disparities, defining a piecewise linear curve as our envelop function. This extends straight line fitting, using (e.g.) a Hough line and an occupancy grid mapping method. We are interested in using a robust and straightforward method with a minimised number of parameters for detecting the lower envelop. The rest of the paper is structured as follows: Section 2 briefly reviews related literature. Section 3 discusses the proposed method towards stixel extraction. Section 4 provides an experimental analysis. Section 5 concludes.

## 2. Ground-Obstacle Segmentation

Road-boundary segmentations are applied for modelling the space in which an ego-vehicle can drive in. This also supports the concept of ground- versus obstacle-segmentation.

Ground manifold detection, i.e. the detection of surfaces ahead of the ego-vehicle which can potentially be used for driving in (e.g. in the case of a suddenly arising emergency), can be done using both, either monocular or stereo vision.

For the monocular case, various monocular features have been exploited as cues for ground manifold estimation such as colour [17], [18], intensity [19], shape [20], boundary [21], or vanishing points [22], [23]. For the stereo-vision case, the ground plane is typically estimated by using normal vectors in disparity space [11]. Urban environments feature complex surroundings in which the ground plane is limited by large or just relatively flat obstacles, such as cars or curbs [24]. Ground plane methods rely on single frame

measurements, suffer from sensor noise and depth artifacts, which lead to increased deviations from correct estimations. A temporally filtered ground plane estimation method, using dense disparity images (from stereo vision), is proposed in [24]. Stereo vision supports ground manifold estimation techniques by providing  $v$ -disparity analysis [25], object-related disparity analysis [26], or occupancy-grid generation [3], [27], [36].

Automated vehicle control may actually induce safety risks when relevant information is missing [29]. In a given scenario, ground-plane detection based on a stereo-vision system may contribute to improved high-level driver assistance functionality [29]. Major challenges for this functionality may appear due to arising deviations which may require confidence estimation in parallel to the supported decision process [24].

[30] introduces a new approach for detecting free-spaces and obstacles using omnidirectional images. Such images support the discussed application purely for vision-based robot navigations in indoor environments; the system employs naïve Bayes classifiers for fusing multiple visual cues and features generated from heterogeneous segmentation schemes. The considered schemes maintain separate appearance models and initiate seeds for ground manifold and obstacle regions.

The robustness of ground manifold detection may be increased by the exploration of using convolutional networks in supervised learning [28]. In another research, a disparity-based obstacle detection method was implemented [31] (based on a SORT-SGM stereo matcher) in order to gain appropriate density information for stereo reconstruction that can support subsequent free-space calculation.

Our method for stereo reconstruction, reported below, continues the work as initiated by [32]. We also take the benefits of reduced computation time, offered by the rSGM stereo matcher [32], to enable a fast and robust ground manifold detection.

## 3. The Stixels World

In order to solve the segmentation task efficiently, this research employs stixel representations for a given situation, which is obtained in four steps. First, generate a disparity image for the given stereo image pair. Second, determine the base points by computing the ground manifold using  $v$ -disparity. Third, perform a height segmentation by using membership votes, a benefit image, and dynamic programming to find the top-points of the objects. Fourth, extract the stixel's depth by using a histogram-based disparity registration scheme.

### 3.1. Base-points in Ground Manifold

The  $v$ -disparities are calculated based on disparity map values estimated using a stereo matcher. As test data we employ stereo-vision traffic images with multiple marked, or unmarked lanes, as available on the KITTI website [33], with a resolution of  $1242 \times 375$  pixels. These images are

used for testing a variant of semi-global matching which is a commonly followed stereo matching paradigm.

For stereo matching we decided for *rapid semi-global matching* (rSGM) on a (simple) CPU [32]. A dense disparity map is computed by matching all the pixels in one image of a stereo pair with their corresponding pixels in the second image; see Fig. 2.

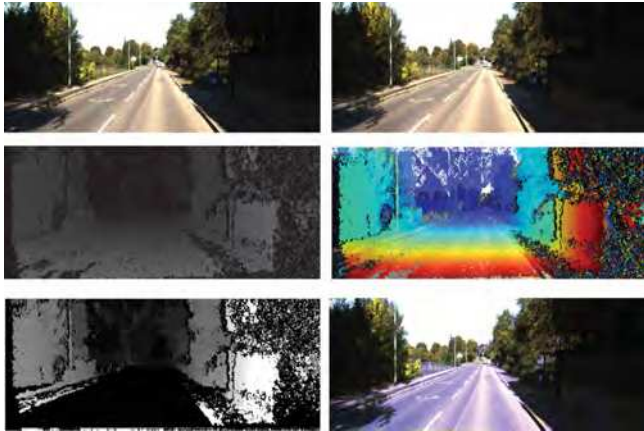


Fig. 2. Free-space detection using stereo vision. *Top row*: Stereo pair of the KITTI urban dataset (multiple-lanes, marked). *Middle left*: Disparity map using rSGM, visualized with a gray-level key. *Middle right*: Disparity map visualized with a color key. *Bottom left*: Free-space (subset of the black ground manifold) approximated based on disparity. *Bottom right*: Free-space visualised in the original scene.

After the disparity map is computed, we populate the  $v$ -disparity space by accumulating pixels with the same disparity value on a horizontal row of the disparity map. The  $v$ -disparity space is a matrix which stores frequencies of disparity values, row by row for the disparity map. Row  $v$  of a  $v$ -disparity map corresponds to row  $v$  of the disparity map; the number of columns of the disparity map correspond to the maximum disparity value in the disparity map. The value of an element in the  $v$ -disparity map represents the number of pixels which has the same disparity value in the corresponding horizontal row  $v$  of the disparity map [35]:

$$V_v(d) = \text{card}\{u : 1 \leq u \leq N_{\text{cols}} \wedge D(u, v) = d\} \quad (1)$$

where  $0 \leq d \leq d_{\text{max}}$ , and  $V_v(d)$  represents the value from the  $v$ -disparity space which accumulates the number of pixels with disparity  $d$  from row  $v$  in the disparity map. The next step is to extract the ground manifold from the generated  $v$ -disparity map. The ground manifold is identified based on a calculated lower envelop in the  $v$ -disparity space.

The Hough transform was used in previous work [25], [31] to detect a lower envelop function in form of a straight line in the  $v$ -disparity map, starting with calculating a lower and an upper envelop. Envelop estimation is based on calculating the intensity sum of all pixels along a considered curve in the  $v$ -disparity map, and then selecting the envelop for which this sum is the minimum.

Considering that a road surface is not a perfect plane, and possibly also more irregular in slope changes than a continuous curve, we consider here polynomial line fitting

for extracting a piecewise best linear fit to the  $v$ -disparities. In each row  $y_i$ , let  $x_i$  be the identified point to be used for lower envelop fitting:

- 1) Find the coefficients  $a_0, a_1, \dots, a_n$  of a polynomial

$$y = P(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_0 \quad (2)$$

where  $n$  is the degree.

- 2) Apply least-square error minimisation for obtaining  $a_0, a_1, \dots, a_n$ , which minimize

$$E(a_1, a_2, \dots, a_n) = \sum_{i=1}^{N_{\text{row}}} [y_i - P(x_i)]^2 \quad (3)$$

for rows  $y_i = i$ ,  $1 \leq i \leq N_{\text{row}}$ .

The generated polynomial  $P(x)$  fits given data in general better than just a fitted straight line; see Fig. 3.

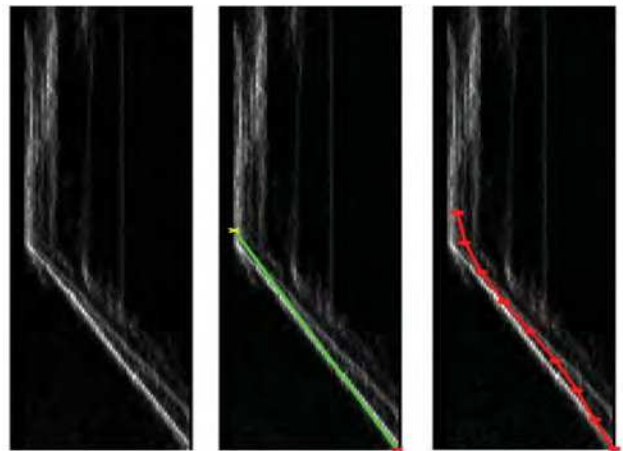


Fig. 3. Example for lower envelop calculation in  $v$ -disparity space. *Left*:  $v$ -disparity map. *Middle*: Detected straight-line fit using a Hough transform. *Right*: Detected polynomial fit.

The final step, calculating base-points  $b_1, b_2, \dots, b_{N_{\text{col}}}$  of stixels for row  $v$ , is now straightforward. Consider a pixel with a given disparity, thus also a defined depth, in row  $v$ . The projection ray of this pixel intersects the ground manifold at a particular point in 3-dimensional (3D) camera space.

### 3.2. Top-points of Stixels

The height of the obstacles, which “sit” on the ground manifold, is obtained by seeking an ideal segmentation between *foreground* and *background* disparities, see [2]. The estimation of top-points  $t_1, t_2, \dots, t_{N_{\text{col}}}$  of stixels for row  $v$  begins with selecting membership votes for pixels (pixel in background or foreground?). Afterwards, a benefit image is produced and used for approximate calculation of those top-points [2].

The membership values rely on the selection of a disparity in a column with respect to the other disparities in this column for defining, or not defining a foreground obstacle. The membership value is positive if it does not exceed the

maximum distance of the expected obstacle disparity; otherwise negative. [2] suggests this exponential membership function:

$$M_{u,v} = 2 \left( 1 - \left( \frac{D_{u,v} - d_u}{\Delta D_u} \right)^2 \right) - 1 \quad (4)$$

where  $\Delta D_u$  is a computed parameter which determines the difference between the disparity obtained from the ground manifold vector and the disparity corresponding to depth.

See Fig. 4, top row, for a visualization of membership votes where green represents a true-positive region (belonging to an object), pale shows a neutral region (e.g. free space), and blue shows true negatives (background).

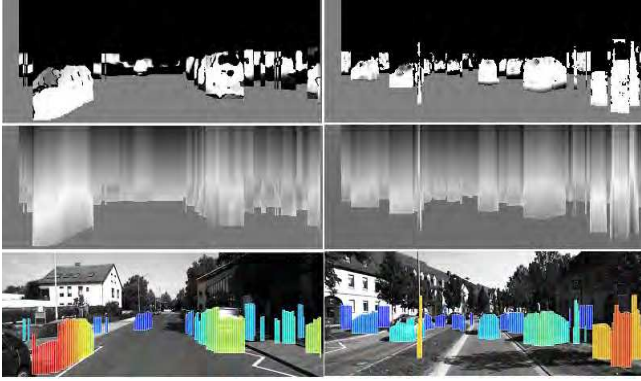


Fig. 4. Stixels world. *First row*: Membership votes. *Second row*: Benefit image. *Third column*: Extracted stixels (of substantial height above ground manifold).

From the membership values, the *benefit image* is computed:

$$C_{u,v} = \sum_{i=v}^{b_u} M(u, i) - \sum_{i=1}^{v-1} M(u, i) \quad (5)$$

(Note that row 1 is the topmost row in the image.) The benefit image is used for calculating top-pixels  $v = t_u$  which maximize the benefit. See Fig. 4, middle. Bright pixels indicate most likely positions of top pixels.

### 3.3. Stixel Extraction

By clustering line segments, defined by one base-point and one related top-point in 3D camera coordinates, into an occupancy grid defined by  $w \times w$  cells in a plane above the ground manifold, we create the input data for individual stixels.

One stixel is defined by all the covered disparities and image points within one  $w \times w$  cell. The covered disparity values are used to accurately estimate the distance related to a stixel. For this purpose, we use a histogram-based disparity registration scheme to extract a stixel’s depth. The stixel generation is a basic step towards object segmentation using geometric data with the aim of improving segmentation accuracy. A stixel represents the closest significant object towards the ego-vehicle. Two examples of resulting stixel worlds are shown in Fig. 4, lower row. The colours of

the stixels encode the distance to the ego-vehicle; a red-range colour represents closeness while a blue-range colour represents far away.

## 4. Experimental Results

In this section, we evaluate the accuracy of stixel detection using either the common ground-plane model, or our proposed way of using a polynomially approximated ground manifold instead. In both cases we use the same stereo matcher. Our local processing platform is a standard PC with a CPU Intel I7 of 3.4 GHz clock and 32 GB of RAM. Both methods were tested using a dataset of 184 stereo pairs from KITTI [34]. For selecting our dataset, we aimed at having a wide diversity of challenging traffic situations including different lighting conditions, different road views, shades, and colourings.

We adopt the pixel-based measurements as employed for the KITTI dataset [34] with the proposed and implemented four error measures: precision, recall, accuracy, and the F-measure (the harmonic mean of recall and precision) calculated for the *birds-eye view* (BEV) projection of a given image; see Fig. 5.

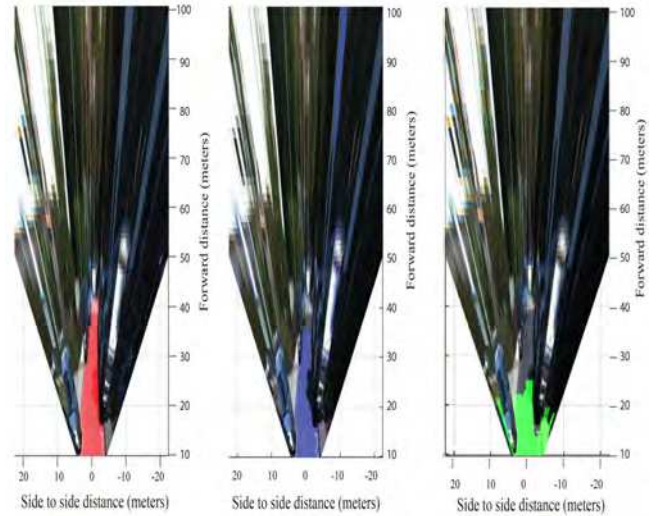


Fig. 5. Classification results for challenging UU road area image using BEV. The horizontal axis represents side-to-side distance (meters), while the vertical axis represents forward-to-side distance (meters). *Left*: Ground-truth image. *Middle*:  $v$ -disparity using the proposed polynomial ground manifold approximation. *Right*:  $v$ -disparity using the common ground plane model.

For clarity, we recall the used measures:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (6)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (7)$$

$$\text{F-measure} = \frac{2(\text{Precision} \cdot \text{Recall})}{\text{Precision} + \text{Recall}} \quad (8)$$

TABLE 1. APPLICATION OF COMMON CLASSIFICATION MEASURES IN CASE OF FREE-SPACE DETECTION

		Ground truth	
		Occupied space	Ground plane
Result	Occupied space	$TN$	$FN$
	Ground plane	$FP$	$TP$

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (9)$$

An application of these measures for free-space detection is given as an example in Table 1.

A weakness of ground-plane (or occupancy-grid) mapping is that many false alarms and misses of true-positive can be expected. Compared to that, having the ground manifold approximated by a polynomial curve, improves the situation in general. Although the process is slightly more complex compared to the use of the ground plane model, the polynomial-based stixel-extraction method clearly outperforms the plane-based method. See Tables 2 and 3.

TABLE 2. RESULTS FOR URBAN UNMARKED LANES

	Accuracy	Recall	F-measure	Precision
Plane-based	0.68	0.73	0.31	0.43
Polynomial-based	0.76	0.85	0.47	0.61

TABLE 3. RESULT FOR MULTIPLE MARKED LANES

	Accuracy	Recall	F-measure	Precision
Plane-based	0.66	0.52	0.41	0.34
Polynomial-based	0.69	0.70	0.52	0.41

Comparative qualitative results are presented in Fig. 6, also illustrating our way of quantitative evaluation. Defining the border of ground plane or polynomial ground manifold by detected stixels, we can compare the KITTI-provided ground truth for free space with the remaining free space of one of the two methods.

The use of a ground plane results in a lower detection rate and a higher rate of false positives in urban unmarked areas due to many obstacles, while the polynomial fitting method outperforms with a smaller rate of false alarms. The polynomial fitting method still suffers from some noise, especially in case of shades on roads which lead to false detections; this is also visible in the example shown in Fig. 6. Under such more difficult circumstances, containing strong noise, the method can consequently also lead to false alarms. For urban marked multiple-lane situations, the polynomial fitting method shows relatively close behaviour to the plane-based detection. See Table 3.

We noticed that the border line of the ground plane may occasionally be far away from the object, and this leads to an “early” detection of stixels although they are not apparent as obstacles in the occupancy grid. False rendering also occurs because of digitization errors when mapping measurements into grid or voxel spaces. Furthermore, as experiments show, the occupancy-grid mapping required more computation time for triangulation or projection, estimated in average

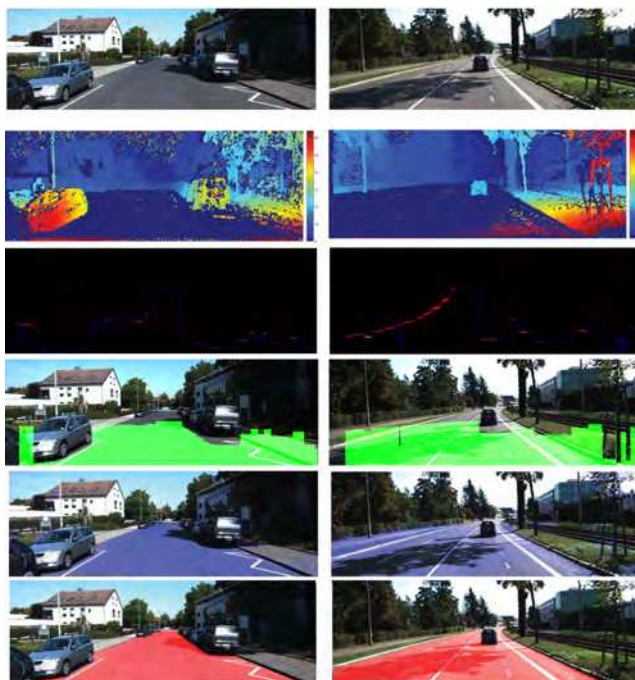


Fig. 6. Qualitative results on KITTI datasets. *First row:* An original image. *Second row:* Visualization of polynomial ground manifold. *Third row:* Visualization of occupancy grid after background subtraction. *Fourth row:* Ground plane for challenging UU, and UMM road. *Fifth row:* Polynomial ground manifold for challenging UU, and UMM road. *Sixth row:* KITTI-provided ground truth for challenging UU, and UMM road.

around 2.58 s per frame. The use of  $v$ -disparity requires less computation time which was on average around 1.6 s per frame. Having in mind that the ground plane is the navigable area for vehicles, accuracy is the most important factor for safe driving. As shown, polynomial ground-manifold fitting can guarantee safety better than just considering a plane (i.e. occupancy grid mapping) for the considered methods.

## 5. Conclusions

We proposed polynomial fitting for ground-manifold estimation using  $v$ -disparity estimation, which is based on polynomial curve fitting for lower envelop detection. The main advantage of our work is that the polynomial curve fitting matches ground manifold to the actual road geometry, thus reducing sensitivity of stixel extractions to road-surface slope changes. This also reduces additional quantization artifacts which is an arising problem when mapping measurements into grid or voxel spaces.

In order to demonstrate our idea in this paper, we reported on tests on two-hundred stereo frames including road labels, supplied in the KITTI datasets, to generate ground truth images for the ground manifold. Based on these ground truth labels, experiments were performed for various road scenes. We showed that the proposed framework robustly estimates stixels, and, in conclusion, free space. We also compared the proposed method with the common ground-plane-based (i.e. occupancy grid mapping-based) method to

show the potential of optimization towards more accurate stixel extractions.

## References

- [1] International Transport Forum: *Road Safety Annual Report for 2016*, OECD Publishing, Paris. DOI= <http://dx.doi.org/10.1787/irtad-2016-en> (2016)
- [2] D. Pfeiffer: *The Stixel World*. Doctoral Thesis, Humboldt University Berlin (2011)
- [3] H. Badino, U. Franke, and D. Pfeiffer: The stixel world - A compact medium level representation of the 3d-world. In *LNCS*, 51–60. [https://doi.org/10.1007/978-3-642-03798-6\\_6](https://doi.org/10.1007/978-3-642-03798-6_6) (2009)
- [4] J. M. Alvarez, T. Gevers, Y. LeCun, and A. M. Lopez: Road scene segmentation from a single image. *LNCS 7578*, 376–389. [https://doi.org/10.1007/978-3-642-33786-4\\_28](https://doi.org/10.1007/978-3-642-33786-4_28) (2012)
- [5] K. Kaaniche, C. Demonceaux, and P. Vasseur: Analysis of low-altitude aerial sequences for road traffic diagnosis using graph partitioning and Markov hierarchical models. In *Int. Multi-Conf. Systems Signals Devices*, 656–661. <https://doi.org/10.1109/SSD.2016.7473665> (2016)
- [6] J. Wu, Z. Cui, V. S. Sheng, P. Zhao, D. Su, and S. Gong: A comparative study of SIFT and its variants. *Measurement Science Review* 13, 3: 122–131. <https://doi.org/10.2478/msr-2013-0021> (2013)
- [7] C. Farabet, C. Couprie, L. Najman, and Y. Lecun: Learning hierarchical features for scene labeling. *IEEE Trans. Pattern Analysis Machine Intelligence* 35, 8: 1915–1929. <https://doi.org/10.1109/TPAMI.2012.231> (2013)
- [8] J. Anders, M. Mefenza, C. Bobda, F. Yonga, Z. Aklah, and K. Gunn: A hardware/software prototyping system for driving assistance investigations. *J. Real-Time Image Processing* 11, 3: 559–569. <https://doi.org/10.1007/s11554-013-0351-4> (2016)
- [9] T. Scharwächter, M. Enzweiler, U. Franke, and S. Roth: Stixmantics: A medium-level model for real-time semantic scene understanding. *LNCS 8693*, 533–548. [https://doi.org/10.1007/978-3-319-10602-1\\_35](https://doi.org/10.1007/978-3-319-10602-1_35) (2014)
- [10] B. S. Shin, Z. Xu, and R. Klette: Visual lane analysis and higher-order tasks: A concise review. *Machine Vision Applications* 25, 6: 1519–1547. <https://doi.org/10.1007/s00138-014-0611-8> (2014)
- [11] K. Onoguchi, N. Takeda, and M. Watanabe: Obstacle location estimation using planar projection stereopsis method. *Systems Computers Japan* 32, 14: 67–76. <https://doi.org/10.1002/scj.1093> (2001)
- [12] T. Shioyama, H. Y. Wu, M. Takebe, and N. Shimaoka: Segmentation and free space detection using Gabor filters. In *Proc. Scandinavian Conf. Image Analysis LNCS 2749*, 1:311–319 (2003)
- [13] K. Schauwecker, S. Morales, S. Hermann, and R. Klette: A comparative study of stereo-matching algorithms for road-modeling in the presence of windscreen wipers. In *Proc. IEEE Intelligent Vehicles Symp.* 7–12. <https://doi.org/10.1109/IVS.2011.5940392> (2011)
- [14] J. Seo, C. Oh, and K. Sohn: Segment-based free space estimation using plane normal vector in disparity space. *Int. Conf. Connected Vehicles Expo: 144–149*. <https://doi.org/10.1109/ICCVE.2015.6> (2015)
- [15] R. Klette: Vision-based driver assistance. In *Wiley Encyclopedia of Electrical and Electronics Engineering*. John Wiley & Sons, Hoboken, NJ, 1–15. <https://doi.org/10.1002/047134608X.W8272> (2015)
- [16] M. Rezaei and R. Klette: *Computer Vision for Driver Assistance*. Springer, Amsterdam (2017)
- [17] K. Lu, J. Li, X. An, and H. He: A hierarchical approach for road detection. *IEEE Int. Conf. Robotics Automation: 517–522*. <https://doi.org/10.1109/ICRA.2014.6906904> (2014)
- [18] A. Rahman, B. Verma, and D. Stockwell: An hierarchical approach towards road image segmentation. In *Proc. Int. Joint Conf. Neural Networks*, 1–8. <https://doi.org/10.1109/IJCNN.2012.6252403> (2012)
- [19] C. Rasmussen: Combining laser range, color, and texture cues for autonomous road following. In *IEEE Int. Conf. Robotics Automation (Cat. No.02CH37292)*, <https://doi.org/10.1109/ROBOT.2002.1014439> (2002)
- [20] Z. He, T. Wu, Z. Xiao, and H. He: Robust road detection from a single image using road shape prior. In *IEEE Int. Conf. Image Processing, 2757–2761*. <https://doi.org/10.1109/ICIP.2013.6738568> (2013)
- [21] Y. He, H. Wang, and B. Zhang: Color-based road detection in urban traffic scenes. In *IEEE Trans. Intelligent Transportation Systems*, 309–318. <https://doi.org/10.1109/TITS.2004.838221> (2004)
- [22] H. Kong, J. Y. Audibert, and J. Ponce: General road detection from a single image. *IEEE Trans. Image Processing* 19, 8: 2211–2220. <https://doi.org/10.1109/TIP.2010.2045715> (2010)
- [23] O. Miksik: Rapid vanishing point estimation for general road detection. In *IEEE Int. Conf. Robotics Automation*, 4844–4849. <https://doi.org/10.1109/ICRA.2012.6225206> (2012)
- [24] H. Harms, E. Rehder, M. Lauer: Grid map based free-space estimation using stereovision. In *Workshop Environment Perception Automated On-road Vehicles, IEEE Intelligent Vehicles Symp.* (2015)
- [25] R. Labayrade, D. Aubert, and J.-P. Tarel: Real time obstacle detection in stereovision on non flat road geometry through v-disparity representation. *IEEE Intelligent Vehicles Symp.* 646–651, <https://doi.org/10.1109/IVS.2002.1188024>
- [26] J. Suhr and H. Jung: Noise-resilient road surface and free space estimation using dense stereo. *IEEE Intelligent Vehicles Symp.*, 0–5. <https://doi.org/10.1109/IVS.2013.6629511> (2013)
- [27] M. Brandao, R. Ferreira, K. Hashimoto, A. Takanishi, J. Santos-Victor: On stereo confidence measures for global methods, evaluation, new Model and integration into occupancy grids. In *Proc. Pattern Analysis Machine Intelligence*, 116–128 (2016)
- [28] W. P. Sanberg, G. Dubbelman, P. H. N. de With: Free-space detection with self-supervised and online trained fully convolution networks. In *Computer Vision Pattern Recognition*, 8 pages, arXiv:1604.02316 (2016)
- [29] L. Neumann, B. Vanholme, M. Gressmann, A. Bachmann, L. Kahlke, and F. Schule: Free Space Detection: A Corner Stone of Automated Driving. In *IEEE Conf. Intelligent Transportation Systems*, 1280–1285. <https://doi.org/10.1109/ITSC.2015.210> (2015)
- [30] L. Posada, K. Narayanan, F. Hoffmann, T. Bertram: Detecting free-space and obstacles in omnidirectional images. In *Intelligent Robotics Applications*, 610–619 (2011)
- [31] A. Ioie, I. Giosan, and S. Nedeveschi: UV disparity based obstacle detection and pedestrian classification in urban traffic scenarios. In *IEEE Int. Conf. Intelligent Computer Communication Processing*, 119–125. <https://doi.org/10.1109/ICCP.2014.6936963> (2014)
- [32] R. Spangenberg, T. Langner, S. Adfeldt, and R. Rojas: Large scale Semi-Global Matching on the CPU. In *IEEE Intelligent Vehicles Symp.*, 195–201. <https://doi.org/10.1109/IVS.2014.6856419> (2014)
- [33] A. Geiger, P. Lenz, and R. Urtasun: Are we ready for autonomous driving? the KITTI vision benchmark suite. In *Computer Vision Pattern Recognition*, 3354–3361. <https://doi.org/10.1109/CVPR.2012.6248074> (2012)
- [34] J. Fritsch, T. Kuhn, and A. Geiger: A new performance measure and evaluation benchmark for road detection algorithms. *IEEE Conf. Intelligent Transportation Systems*, 1693–1700. <https://doi.org/10.1109/ITSC.2013.6728473> (2013)
- [35] N. H. Saleem and R. Klette: Accuracy of Free-space Detection for Stereo versus Monocular Vision. In *Image Vision Computing New Zealand*, 48–53. <https://doi.org/10.1109/IVCNZ.2016.7804422> (2016)
- [36] M. Perrollaz, J.D. Yoder, A. Negre, A. Spalanzani, C. Laugier: A visibility-based approach for occupancy grid computation in disparity space. *IEEE Trans. Intelligent Transportation Systems*, 13:1383–1393 (2012)