# Drivable Road Detection with 3D Point Clouds Based on the MRF for Intelligent Vehicle

Jaemin Byun, Ki-in Na, Beom-su Seo, and Myungchan Roh

**Abstract.** In this paper, a reliable road/obstacle detection with 3D point cloud for intelligent vehicle on a variety of challenging environments (undulated road and/or uphill/ downhill) is handled. For robust detection of road we propose the followings: 1) correction of 3D point cloud distorted by the motion of vehicle (high speed and heading up and down) incorporating vehicle posture information; 2) guideline for the best selection of the proper features such as gradient value, height average of neighboring node; 3) transformation of the road detection problem into a classification problem of different features; and 4) inference algorithm based on MRF with the loopy belief propagation for the area that the LIDAR does not cover. In experiments, we use a publicly available dataset as well as numerous scans acquired by the HDL-64E sensor mounted on experimental vehicle in inner city traffic scenes. The results show that the proposed method is more robust and reliable than the conventional approach based on the height value on the variety of challenging environment.

# 1 Introduction

The accurate perception of the environment is a very important step to drive autonomously for an intelligent vehicle, namely the detection of road area and obstacle. The road and obstacle detection are being nearly performed by using a various kind of sensors. Many teams participating in the DARPA Urban Challenge have nearly performed based on the data acquired by 2D range sensors for road region and obstacle detection [17, 10, 1, 2]. However, these sensors scan the environment along a plane within a limited viewing angle, thus the objects above or below this plane cannot be detected. A number of approaches focused on the use of vision exclusively have been studied for decades [18],[12]. The fusion of the range and

Jaemin Byun · Ki-in Na · Beom-su Seo · MyungChan Roh

Robot and Cognitive System Research Department(RCSRD) in Electronics and Telecommunications Research Institute (ETRI), Daejeon, South Korea e-mail: {jaemin.byu, kina, bsseo, mcroh}@etri.re.kr

vision data that allows a richer description of the world have been developed [5],[14]. Recently, a three-dimensional range-scanner which provides 3D point cloud instead of 2D slice of the environment have been commercially introduced. Although there has been an overwhelming amount of work on perception in 2D and 2.5D, but the problem of perception in 3D has been addressed by comparatively fewer researchers yet. One of its main reasons is the enormous amount of data provided by 3D sensors. The amount of data in a single scan of 3D sensor is usually several times larger than that of a 2D scan. Therefore, how to build consistent and efficient 2D representations out of 3D range data is important for the sensor data processing as well as road/obstacle detection. Improving on earlier these work, the main contribution of this paper is a method for efficient road detection based on MRF with LBP(Loopy Belief Propagation). We also employ a cylindrical 2D grid map with the different size of cell corresponding to the distance from vehicle. Besides, our objective is to present a method that can detect accurately drivable road and obstacle regions in a variety of challenging environment such as undulated road, uphill/downhill, rolling /pitching of the host vehicle as shown in Fig. 1.

In our work, the 3D range data is acquired by a Velodyne HDL-64E sensor as shown in the Fig.1-(b), which is mounted on the top of the vehicle, it is covering a total vertical range of approximately 25 degrees. To obtain data from the whole environment, the laser scanner rotates at a speed of 10 Hz. A data packet from the LIDAR consists of the rotational angle of the scanner itself, the range and the intensity measurement of each laser. From this data, a complete scan of the environment can be computed. By the way, the motion of vehicle itself would affect these 3D data information as the sensor is mounted moving on the vehicle. To remove the distortion that is caused by the movement of the vehicle during one revolution, the paper presents the strategy that involve in correction process through estimating the posture of vehicle. The paper is organized as follows. In the next section we give an outline of relevant works, followed by the detailed description of our approach. Experimental results are given in section 5. Section 6 concludes this paper and provides a perspective for future research in this area.



Fig. 1 (a)Uphill road.(b) 3D point clouds acquired by Velodyne LIDAR.

### 2 Previous Works

With range scanning devices becoming standard equipment in mobile robotics, the task of 3D scan segmentation and classification is one of increasing practical relevance. Typical algorithms for road and obstacle detection with 3D LIDAR are as follows: One of the most widely used method is the projection of 3D point clouds on the assumed or estimated ground plane and finds similar x-y coordinates whose height exceeds a given threshold value. This is represented by a grid in which each cell contains only one height by selection of the average, max, min height of the sensor data located in each grid cell [16],[15]. One of the advantages is that the several sensors can be fused easily and that mapping is straightforward. Many teams participating in the DARPA Urban Challenge successfully applied this method. However, the difficulty of road detection has still in sloped terrains or the situations with big rolling/pitch angle of the host vehicle. Both Leonard et al.[8] and Himmelsbach et al.[7] describe a method that identifies points in the point clouds that are likely to be on the ground, and then fit a ground model through those ground points. And other points above the ground model are deal with as obstacle point. Douillard et al.[3] proposes a strategy that utilizes ground models of non-constant resolution either providing a continuous probabilistic surface or a terrain mesh built from the structure of a range image. Moosmann et al.[11] proposes graph-based approach to segment ground and objects from 3D LIDAR scans using a novel unified, generic criterion based on local convexity measures. Guo et al.[6] use a graph-based approach for 2D road representation of 3D point clouds with respect to the road topography. The method describes also the gradient cues of the road geometry to construct a MRF and implements a belief propagation (BP) algorithm to classify the road environment into four categories, i.e. the reachable region, the drivable region, the obstacle region and the unknown region. However their method uses only gradient value for labeling so that it cant sometimes be distinguished the ground and the roof of vehicle. Li and Li[9] proposes a method of Four Directions Scan Line Gradient Criterion (4DSG) that is calculated the gradient with neighboring points. These features can not only reflect the flatness of pavement, but also reflect the distinguishing feature of point cloud on curbs in four directions. Bohren et al.[1] addresses a method that road points can also be detected based on the reflectivity of the ground in the Velodyne scans. However, such approach can only work well under good conditions so that their road/obstacle detection has to be supplemented by other sensor. The outline of our work can be show in the Fig.2. The proposed approach differs from previous related work. Main contributions that we propose are as follow

- Unlike most of the previous works, we focus on the correction of distorted 3D point cloud occurred by motion of vehicle (high speed/ move up and down) in practical road situation.
- We employ the approach that 3D point clouds are projected on the grid map in the cylindrical coordination. We have considered the fact that the point cloud by Velodyne sensor will be gradually sparse from near to far and the dramatic change happens between adjacent beams can reveal the vertical change of the environment along the circular direction as shown in Fig.2-(b).



**Fig. 2** Illustration of our method.(a) 3D point clouds aquired by Velodyne LIDAR. (b) cylindrical grid map. (c) extraction of feature values in each cell. (d) multi-labeling and classification with LBP.

- We propose a robust method of road detection with 3D data in undulated road such as down/uphill by the best selection of the proper features such as gradient value, height average of neighboring node as shown in Fig.2-(c).
- We formulate the road detection problem based on MRF with the loopy belief propagation to find the different regions with different classes as shown in Fig.2-(d).

# 3 3D Points Representation and Grid Map Building

# 3.1 Correction of Distorted 3D Point Cloud by Considering Vehicle Motion

As shown in Fig.3, the Velodyne LIDAR that is mounted on the top of the vehicle uses 64 lasers, which cover in different vertical angle, and it can also provide 360 degrees field of view for surrounding environment with more than 1.3 million points per second. The LIDAR returns deliver spherical point coordinates so it needs to transformation that data into Cartesian space. To do the transformation, we have to consider calibration parameters such as distance correction factor  $\Delta r$ , vertical/horizontal correction angle  $\Delta \theta_v$ ,  $\Delta \theta$ , rotation angle  $\varphi$ , measured distance rand vertical/ horizontal offset  $r_v$ ,  $r_h$ , The 3D point cloud computation in the cartesian coordinates is as below

$$p(x, y, z) = \begin{pmatrix} (r + \Delta r) \cdot \cos(\varphi + \Delta \theta) \cdot \cos(\Delta \phi) + r_h \cdot \cos(\varphi) \\ (r + \Delta r) \cdot \sin(\varphi + \Delta \theta) \cdot \cos(\Delta \phi) + r_h \cdot \sin(\varphi) \\ d \cdot \sin(\varphi + \Delta \theta) - r_v \end{pmatrix}^T$$
(1)

Here, we have to consider that the scanner takes a no negligible amount of time to complete one rotation, so the observed 3D point clouds with LIDAR are distorted by the motion of the vehicle. For instance, if the speed of vehicle runs as about 100km/h (27.8 m/s) in the highway, it would be unfortunately given the distorted information past 2.7 meter on the every scan due to the fact that the scanner is rotating with a frequency of 10 Hz(0.1s). Furthermore, we should consider the situation that the vehicle passes over speed bumper with big rolling/pitch angle and then the front road is classified as obstacle due to the downward pitching of the vehicle. We are able to solve these problems by using information about the ego-motion of the car. In other word, the resulting frame is an approximation of how the environment would have looked like if the car had not moved.

To correct the distorted laser measurement by the vehicle's movement, we utilize a GPS/INS unit that provides highly the accurate motion information of the vehicle. Each laser measurement *i* during one revolution is referenced with respect to vehicle position and orientation  $O_{t+i}$  from the start of a sensor revolution, and afterwards transformed such that the coordinates are referenced with respect to  $O_{t+\Delta t}$  at the end of the revolution. For transformation with rotation  $\tilde{R}$  and translation  $\vec{T}$  for each data, the undistorted coordinates  $p_{O_{t+\Delta t}}^i$  of a point  $p_{O_t}^i$  referenced with respect to  $O_{t+\Delta t}$  as shown in Fig.3 can be calculated as follow as

$$p_{O_{t+\Lambda t}}^{i} = \tilde{R}(p_{O_{t}}^{i} - \vec{T})$$
<sup>(2)</sup>

#### 3.2 Grid Map Building

The 3D point clouds which obtained by Velodyne LIDAR need to expensive costs to deal a large amount of data for real time processing, in our work we try to reduce it with using a 2.5D ego-centered cylindrical grid. Some relative works use the rectangular grid map projected by 3D point cloud points. Others approaches use



**Fig. 3** Correction of distorted 3D point cloud by estimation the motion of vehicle

a mesh-grid map which directly is decomposed of a neighborhood graph from a scanner. Here, we focus on considering the manner a LIDAR scan, we know that it can give a different density of point depend on distance from vehicle despite the Velodyne sensor can take 3D scans of environment and provide millions of points per second. The points cloud will be gradually sparse from the near to the far. So we set two direction as direction and circular direction in cylindrical coordination. We can know that the gradient value is dramatically changed at some place for object/structure along the circular scan direction. Therefore, it is can be separated the coverage area with more high resolution along the circular direction than radial direction as shown in the Fig.4

### 4 Feature Extraction and Road Classification

#### 4.1 Features(Gradient Value and Height) Extraction

We assume that the road surface is continuous and there is high correlation between neighborhood data. Therefore, given the world coordinates of the 3D point clouds, the gradient value at each node can be computed by using known neighboring nodes. To get the gradient value at each node as shown in Fig. 4, we need to height of neighboring nodes along the radial direction and circular direction. This gradient value  $G_m(p)$  can reflect the geometrical character of roads. To obtain neighborhood points, it searches for closest nodes that have height values in four directions along the radial axis and circular axis respectively. We denote them as  $z_m^{c1}, z_m^{c2}, z_m^{r1}, z_m^{r2}$ , where *c* means circular direction and *r* is radial direction.

The gradient of radial direction can be computed as

$$G_m^r(p) = \frac{z_m^{r1} - z_m^{r2}}{\|P^{r1} - P^{r2}\|}$$
(3)

Next, the gradient of circular direction can also be computed as

$$G_m^c(p) = \frac{z_m^{c1} - z_m^{c2}}{\|P^{c1} - P^{c2}\|}$$
(4)

where p is referred in the cylindrical coordinate. The gradient value is described by

$$G_m(p) = \sqrt{G_m^r(p)^2 + G_m^c(p)^2}$$
(5)

The height average of neighborhood nodes can be described as follows

$$H(p) = \frac{1}{n} \left(\sum_{n} Z_{n}\right) \tag{6}$$

where  $Z_n$  is an average of height value on the neighborhood nodes surrounding current node *p*. Finally, we describe a feature function as follow

$$g(p) = \alpha G_m^*(p) \cdot H(p)^* \tag{7}$$

The  $\alpha$  is weight constant ane  $G_m^*(p)$  and  $H(p)^*$  are normalized with  $G_m(p)$  and H(p).

#### 4.2 Classification Based on MRF

The goal of this step is to find the different regions with different classes and inference the area that the LIDAR does not cover. We take a graph-based approach for classification. Let G = (V, E) be an undirected graph with nodes  $v_i \in V$ , the set of elements to be segmented, and edges  $(v_i, v_j) \in E$  in corresponds to pairs of neighboring nodes. Each edge has a weight  $w(v_i, v_j)$  which is a non-negative measure of the dissimilarity between neighboring elements  $v_i$  and  $v_j$ . We present the classification problem as LBP approach for performing inference on MRF as formed by the standard 4-connected neighborhood system since it models the spatial interactions present in the scene so that the labels of the points are determined jointly.

Let *P* be a set of node in the cylindrical grid map and *L* be as set of labels. The labels correspond to quantities that we want to estimate at the each node, such as gradient value and height average of neighborhood. A labeling *f* assigns a label  $f(v_i) \in L$  to each node  $v_i \in V$ . We assume that the labels should vary slowly almost everywhere but may change dramatically at some places such grids along object boundaries. The quality of a labeling is given by an energy function,

$$E(f) = \sum_{(v_i, v_j) \in E} V(f(v_i), f(v_j)) + \sum_{(v_i) \in V} D(f(v_i))$$
(8)

Where *E* are the edges in the four-connected grid graph.  $V(f(v_i), f(v_j))$  is the cost of assigning labels  $f(v_i)$  and  $f(v_j)$  to two neighboring nodes, and it is referred to as the discontinuity cost.  $D(f(v_i))$  is the cost of assigning label  $f(v_i)$  to node  $v_i$ , which is referred to as the data cost.

Finding a labeling that minimizes this energy corresponds to the maximum a posteriori (MAP) estimation for MRF in the form of Eq.8. Normally this algorithm is defined in terms of probability distributions, but an equivalent computation can be



**Fig. 4** The computation of gradient with neighboring nodes

performed with negative log probabilities, where the max-product becomes a minsum. We use this formulation because it is less sensitive to numerical artifacts, and it uses the energy function definition more directly.

The max-product BP Algorithm works by passing message around the graph defined by the four-connected grid. Each message is a vector of dimension given by the number of possible levels. Let  $m_{v_iv_j}^t$  be the message that node  $v_i$  sends to a neighboring node  $v_j$  at time t. When using negative log probabilities all entries in  $m_{v_iv_j}^0$  are initialized to zero, and at each iteration new messages are computed in the following way,

$$m_{v_{i}v_{j}}^{t}\left(f(v_{j})\right) = min_{f(v_{i})}\left(V\left(f(v_{i}), f(v_{j})\right) + D\left(f(v_{i})\right) + \sum_{N(v_{i})v_{j}} m_{v_{i}v_{j}}^{t-1}\left(f(v_{j})\right)\right)\right)$$
(9)

Where  $N(v_i) \setminus v_j$  denotes the neighbors of  $v_i$  other than  $v_j$ . After *T* iterations a belief vector is computed for each node,

$$b_{\nu_{j}}\left(f(\nu_{j})\right) = D_{\nu_{j}}\left(f(\nu_{i})\right) + \sum_{p \in N\left(\nu_{j}\right)} m_{\nu_{i}\nu_{j}}^{T}\left(f(\nu_{j})\right)$$
(10)

Finally, the  $f_{v_j}^*$  label that minimizes  $b_{v_j}(f(v_j))$  individually at each node is selected.

In the work, the labels correspond to different gradient value and height average that should be assigned to grids in the map. Thus the data costs can be defined as

$$D(f(v_i)) = \min(||g(v_i)| - f(v_i)||, \tau)$$
(11)

We use a truncated step function for the data cost,  $\tau$  is a truncation value,  $g(v_i)$  is the feature of node  $v_i$ . The truncation makes the data cost robust to abnormally large feature values.

Another class of cost functions is based on the degree of difference between labels. The cost of assigning a pair of labels to neighboring node is generally based on the amount of difference between these quantities. In order to allow for discontinuities, as the values are not smoothly changing everywhere, the cost function should be robust, becoming constant as the difference become large. So it can be used the truncated linear model, where the cost increases linearly based on the distance between the labels  $f(v_i)$  and  $f(v_i)$  up to some level,

$$V\left(f(v_i), f(v_j)\right) = \min\left(s \left\| f(v_i) - f(v_j) \right\|, d\right)$$
(12)

Where s is the rate of increase in the cost, and d controls when the cost stops increasing.





## **5** Experimental Result

We have evaluated the proposed algorithm using both a publicly available dataset[4], [13] as well as numerous scans acquired by the HDL-64E sensor mounted on an experimental vehicle(Hyun-dai Sorrento) in inner city traffic scenes as shown Fig.6. As no ground truth information is available, a qualitative performance evaluation is conducted.

For cylindrical grid map, we use  $\Delta \theta = 0.5$  and  $\Delta r = 0.2m$  (*range* :  $0 \sim 30m$ ) and  $\Delta r = 0.5m$  (*range* :  $30 \sim 60m$ ) and  $\Delta r = 1m$  (*outof* 60m) throughout all experiments. For classification based on MRF, we set 10 as number of labels and the truncation value were respectively fixed to  $\tau = 5$  and d = 3.

Since the classification of the 3D point clouds in normal road environment is well demonstrated. So we focus on a variety of challenging environment. Fig.7 shows example result of a slope road, an uphill road and a downhill road, which substantiated that the proposed method is more robust and reliable than the conventional approach based on the height.

As visible in the first column, though the road is able to show normally flat in Fig7 (a), we can see that laser returns corresponding to the area are irregular as Fig.7 (e). As shown in Fig.7 (i), the conventional approach based on the height can give the wrong result that drivable space is as obstacle region, as indicated by the red circle in the figure. Whereas the proposed approach successfully classify the slope area as the drivable with feature values such as gradient and average height of road space as shown in Fig.7 (m). Furthermore, the spatial interactions based on the smoothness term in the MRF can also ensure the local consistency in such scenarios so that all of the rough region will be classified into to same category, even when some of the gradient are abnormal due to noise. Besides, we can see in the second column, when our vehicle drives on the slope road which is more high right than the left side, we can see that the result of the convention approach misrecognizes partially some road and some vehicles as obstacles indicated by the red circle in the Fig.7 (j). But our proposed method gives the robust result of detection according to this height variance of road because of shown in the Fig.7 (n). Also we can see that some area that the LIDAR does not cover is interpolated by inference algorithm based on MRF with the loopy belief propagation through the comparison of area indicated by the red circle in the Fig.7 (k) and in the Fig.7 (o). As shown in the fourth column, it is caused misunderstanding that there is a big obstacle in front of vehicle on the road by the conventional approach based on the height in the Fig.7



**Fig. 6** Classification result of a variety of environment(flat road, slope road, down hill ,uphill,), the pictures of environment(first row), the 3D point clouds by LIDAR, the result of conventional approach (third row), the result of our proposed work(fourth row)

(l). However the proposed approach successfully classified the road as the drivable space since our feature values are in the drivable space for vehicle in the Fig.7 (p).

## 6 Conclusion

In this paper, we have presented a robust method of road detection with 3D point clouds on the challenging road environments such as down/uphill, sloped road. Our first contribution is correction of 3D point cloud distorted by the motion of vehicle incorporating vehicle posture information. Our second contribution is guideline for the best selection of the proper features such as gradient value, height average of neighboring node. Our third contribution is transformation of the road detection problem into a classification problem of different features. Our fourth contribution is inference algorithm based on MRF with the loopy belief propagation for the area that the LIDAR does not cover. In experiments, we use a publicly available dataset as well as numerous scans acquired by the HDL-64E sensor mounted on experimental vehicle in inner city traffic scenes. The results proved that the proposed method is more robust and reliable than the conventional approach based on the height on the variety of challenging environment. Our future work will focus on the detection of dynamic road environment with the supervised/unsupervised learning approach and the fusion of the LIDAR and vision data.

Acknowledgements. This work was supported by the IT R/&D program of MSIP/KEIT. [KI0041417, Development of Decision Making/ Control Technology of Vehicle/Driver Cooperative Autonomous Driving System (Co- Pilot) Based on ICT].

## References

- [1] Bohren, J., Foote, T., Keller, J., Kushleyev, A., Lee, D., Stewart, A., Vernaza, P., Derenick, J., Spletzer, J., Satterfield, B.: Little ben: The ben franklin racing team's entry in the 2007 darpa urban challenge. Journal of Field Robotics 25(9), 598–614 (2008)
- [2] Buehler, M., Iagnemma, K., Singh, S. (eds.): The DARPA urban challenge: Autonomous vehicles in city traffic. STAR, vol. 56. Springer, Heidelberg (2009)
- [3] Douillard, B., Underwood, J., Kuntz, N., Vlaskine, V., Quadros, A., Morton, P., Frenkel, A.: On the segmentation of 3d lidar point clouds. In: 2011 IEEE International Conference on Robotics and Automation (ICRA), pp. 2798–2805. IEEE (2011)
- [4] Geiger, A., Lenz, P., Stiller, C., Urtasun, R.: Vision meets robotics: The kitti dataset. International Journal of Robotics Research, IJRR (2013)
- [5] Guo, C., Mita, S., McAllester, D.: Hierarchical road understanding for intelligent vehicles based on sensor fusion. In: 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), pp. 1672–1679. IEEE (2011)
- [6] Guo, C., Sato, W., Han, L., Mita, S., McAllester, D.: Graph-based 2d road representation of 3d point clouds for intelligent vehicles. In: 2011 IEEE Intelligent Vehicles Symposium (IV), pp. 715–721. IEEE (2011)
- [7] Himmelsbach, M., Luettel, T., Wuensche, H.-J.: Real-time object classification in 3d point clouds using point feature histograms. In: IEEE/RSJ International Conference on

Intelligent Robots and Systems, IROS 2009, pp. 994–1000. IEEE (2009)

- [8] Leonard, J., How, J., Teller, S., Berger, M., Campbell, S., Fiore, G., Fletcher, L., Frazzoli, E., Huang, A., Karaman, S., et al.: A perception-driven autonomous urban vehicle. Journal of Field Robotics 25(10), 727–774 (2008)
- [9] Li, M., Li, Q.: Real-time road detection in 3d point clouds using four directions scan line gradient criterion. Future, 5 (2009)
- [10] Montemerlo, M., Becker, J., Bhat, S., Dahlkamp, H., Dolgov, D., Ettinger, S., Haehnel, D., Hilden, T., Hoffmann, G., Huhnke, B., et al.: Junior: The stanford entry in the urban challenge. Journal of Field Robotics 25(9), 569–597 (2008)
- [11] Moosmann, F., Pink, O., Stiller, C.: Segmentation of 3d lidar data in non-flat urban environments using a local convexity criterion. In: 2009 IEEE Intelligent Vehicles Symposium, pp. 215–220. IEEE (2009)
- [12] Nguyen, T.-N., Michaelis, B., Al-Hamadi, A., Tornow, M., Meinecke, M.: Stereocamera-based urban environment perception using occupancy grid and object tracking. IEEE Transactions on Intelligent Transportation Systems 13(1), 154–165 (2012)
- [13] Pandey, G., McBride, J.R., Eustice, R.M.: Ford campus vision and lidar data set. International Journal of Robotics Research 30(13), 1543–1552 (2011)
- [14] Strom, J., Richardson, A., Olson, E.: Graph-based segmentation for colored 3d laser point clouds. In: 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 2131–2136. IEEE (2010)
- [15] Tay, M.K., Mekhnacha, K., Chen, C., Yguel, M.: An efficient formulation of the bayesian occupation filter for target tracking in dynamic environments. International Journal of Vehicle Autonomous Systems 6(1), 155–171 (2008)
- [16] Thrun, S.: Learning occupancy grid maps with forward sensor models. Autonomous Robots 15(2), 111–127 (2003)
- [17] Urmson, C., Anhalt, J., Bagnell, D., Baker, C., Bittner, R., Clark, M.N., Dolan, J., Duggins, D., Galatali, T., Geyer, C., et al.: Autonomous driving in urban environments: Boss and the urban challenge. Journal of Field Robotics 25(8), 425–466 (2008)
- [18] Wedel, A., Badino, H., Rabe, C., Loose, H., Franke, U., Cremers, D.: B-spline modeling of road surfaces with an application to free-space estimation. IEEE Transactions on Intelligent Transportation Systems 10(4), 572–583 (2009)