Object and ground classification for a mobile robot in urban environment

Jeong Hyo, Ha¹, Sijong Kim², and Myung Jin Chung³*

¹² Department of Electrical Engineering, KAIST, Daejeon, 305-701, Korea
(Tel : +82-42-350-5429; E-mail: pristine08@hanmail.net, terra@cheonji.kaist.ac.kr)
³ Department of Electrical Engineering, KAIST, Daejeon, 305-701, Korea
(Tel : +82-42-350-3429; E-mail: mjchung@ee.kaist.ac.kr) * Corresponding author

Abstract: Recently, an interest about the unmanned vehicles is increasing, and a related research has been actively researched. Application systems using the partial element of technologies are commercialized. The information about surrounding environment should be able to use effectively in order to perform a given task such as robot navigation, path planning, and obstacle avoidance. The essential function for a mobile robot is object perception. This paper proposes an algorithm of object detection using stereo camera. The 3D spatial information is obtained by stereo matching algorithm. The reliability of 3D data is defined according to the distance between the object and the camera, and is used in the filtering process. The geometrical features were analyzed by the continuous characteristic of pixels in image. We achieve classification of ground and object.

Keywords: Object Detection, Ground Rejection, Stereo Matching.

1. INTRODUCTION

For mobile robot navigation, a robot requires information about surrounding environment. Environment information should be classified as object and ground, and used to find a road space. 3D spatial information of surrounding environments is collected by 3D range sensors such as laser range finder, 3D laser scanner and camera sensor. Stereo camera is basic and well-established sensor. Unlike other sensors, stereo camera can provide not only 3D spatial information but 2D image [1]. Generally 3D spatial data is made in form of point clouds. Direct process of point clouds consumes a lot of memory and processing time. To solve this problem, we were processed point clouds using a world modeling techniques such as voxel [2] and elevation map [3]. In this paper, we propose an object perception algorithm using stereo camera. The reliability of 3D spatial data in form of point clouds is defined. As a result of the reliability, reliable spatial data in form of voxel is generated. We extract the geometrical feature of point clouds based on continuity of pixels. Object classification is performed with a geometrical feature. The remainder of this paper is organized as follows. Section 2 presents a reliable data extraction. Section 3 describes ground and object classification. Conclusion is discussed in Section 4.

2. RELIABLE DATA EXTRACTION

The 3D spatial data using a stereo matching algorithm contains many errors that are caused by limitation of stereo matching algorithm. Especially, different objects are linked by incorrect point clouds. Fig.1 (a) is input image, Fig.1 (b) is a result of stereo matching in form of point clouds. In input image (Fig.1 (a)), tree and building are separated. However, connection problem occur in a point clouds map (Fig.1 (b)).

A general solution to solve this problem is the use of the N-nearest neighbor algorithm [4] or the use of the points that exist within a certain distance. However, These algorithms have a disadvantage. First, every point requires an iterative calculation. Second, a search for neighbor point is time consuming processing. So above mentioned algorithms are unsuitable for real time application.

For an efficient calculation of reliability, we use voxel. A point clouds map is divided by 3D grid. The reliability is defined by a number of points that exist in a voxel. Meanwhile, the average number of points that exist in a voxel is inversely proportional to the square of the distance between the camera to the point.

Fig. 1 (a) Input image, (b) a point clouds map, (c) a reliable voxel map
The reliability function is defined with compensation of these characteristics. The use of voxel instead of point clouds has the following advantages. First, the amount of data to be processed was reduced to about 1/10. Second, processing time is faster. Reliable function is the following equation.

The result of reliable data extraction is shown in Fig.1 (c). Fig.1 (c) is reliable data in form of voxel. A point clouds that link different objects is removed. And objects were isolated. Fig. 2 is result of reliable data extraction using a pair of images. Table 1 is the numerical result.

<table>
<thead>
<tr>
<th>Experimental result</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Image size</td>
<td>1344 x 391 pixel</td>
</tr>
<tr>
<td>Map size</td>
<td>33.0 x 13.2 m</td>
</tr>
<tr>
<td># of point clouds</td>
<td>328,888 point</td>
</tr>
<tr>
<td># of voxel</td>
<td>8,953 cell</td>
</tr>
<tr>
<td># of reliable voxel</td>
<td>3,789 cell</td>
</tr>
</tbody>
</table>

3. GROUND AND OBJECT CLASSIFICATION

The core processing for an object perception is ground rejection. The general algorithm for ground rejection is v-disparity [5]. The v-disparity algorithm is designed for road detection using a dominant disparity in row of image. However, v-disparity does not work well in a bumpy road or an environment that appear occlusion. To solve this problem, we analyze a geometrical feature based on continuity on image, and classify as a ground and an object. A red column of Fig. 3 (a) disparity map can be mapped into the 3D space like Fig. 3 (b). A geometrical feature in 3D space can be analyzed using continuity of pixel in image. Point clouds are classified as “horizontal” and “vertical or scatter”.

Fig. 3 3D spatial data about column of image
(a) disparity map(red line), (b) 3D spatial information

A point clouds that is classified according to the geometrical feature can be found in Fig. 4 (a). Red color implies vertical and scatter feature (a candidate of object). Blue color implies horizontal feature (a candidate of ground).

A feature of voxel is classified by a feature of point clouds. Ground and object can be classified by the following two steps classification. In first step, we find the dominant height using a histogram. A voxel is classified as temporary with applying the predefined extra height. In second step, we observe the voxel located over the ground voxel. If voxel was object class, ground voxel replace to object voxel. Result of object classification is shown in Fig. 4 (b).

4. CONCLUSION

This paper proposed object perception algorithm for stereo camera sensor. The proposed reliable data generation step provides error less information, efficient information in form of voxel. Ground and object classification algorithm uses geometrical feature and histogram. The experimental result shows that the object detection is well performed by our algorithm.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the support from UTRC(Unmanned Technology Research Center) at KAIST(Korea Advanced Institute of Science and Technology), originally funded by DAPA, ADD.
Fig. 4 (a) Analysis of geometrical feature (Red : vertical and scatter type, Blue : horizontal type.)
(b) Result of object classification. (Red : object, Blue : ground.)

REFERENCES