High-Accuracy Mapping Design Based on Multi-view Images and 3D LiDAR Point Clouds

Jian-Hong Chen  
Department of Electronic Engineering, National Taipei University of Technology, Taipei, Taiwan  
Email address: chenjianhong663@gmail.com

Chitra Meghala Yelamandal  
Department of Electronic Engineering, National Taipei University of Technology, Taipei, Taiwan  
Email address: ychitrameghala@gmail.com

Guo-Han Lin  
Department of Electronic Engineering, National Taipei University of Technology, Taipei, Taiwan  
Email address: w13076334@gmail.com

Yu-Cheng Fan  
Department of Electronic Engineering, National Taipei University of Technology, Taipei, Taiwan  
Email address: skystar@ntut.edu.tw

Abstract
In recent years, the Internet of things (IOT), big data and artificial intelligence are evolving rapidly. The Self-driving vehicle is an important issue. One of the biggest advantages of self-driving vehicles is to reduce traffic accidents. The goal of our efforts is to provide more convenient and safe transportation. Therefore, we present a “high-accuracy mapping design based on multi-view images and three-dimensional (3D) LiDAR point clouds” in this paper. The proposed system for 3D LiDAR combines color camera and three dimensional space coordinate corresponding to the two-dimensional space coordinates technology.

Key Word-- Multi-view; Cameras; Images; LiDAR, Point Clouds; 3D Environment.

1. Introduction
The self-driving technology is one of the important traffic trends. Because in our daily life, we hope to make transportation more convenient and safe. Self-driving technology is one of the most important things to avoid the collision. Therefore, LiDAR can completely detect surrounding environment to accomplish the self-driving technology but the disadvantage is that it misses some information in the generated point cloud map.

To solve the problems for insufficient resolution of the light point cloud map, we combine LiDAR [1] and camera to detect the 3D environment. Velodyne LiDAR PUCK (VLP)-16 is used to scan the surroundings of the environment and generates the point cloud data. Multi-view images can match with point cloud data.

The transmission method is based on the User Datagram Protocol (UDP) technology. Then we can get the 3D LiDAR packet data by User Datagram Protocol (UDP), and convert it into data such as the distance, timing information, and object of angle and so on. Then, the data packet analysis and decoding is performed. In the color camera section, the parameters of the color camera is extracted. Afterwards, we combine the color camera and LiDAR to make the design platform. Finally, we will match the color camera and LiDAR algorithm to complete a highly accurate image matching.

2. Methodology
This paper is focused on the use of images from LiDAR and Logitech C930e. The 2D RGB color image is matched with the 3D point cloud data and the corresponding point can match the correct position of the image.

The matching method requires camera parameters, including extrinsic parameters [2] and intrinsic parameters [3]. In addition, the extrinsic parameters includes rotation matrix and translation matrix, while the intrinsic parameters include the \((f_x, f_y)\) focal length in pixels, and the center coordinates of camera \((c_x, c_y)\).

Under the camera calibration, matching the world coordinates to the plane image, we can use the following Eq.1.

\[
\begin{bmatrix}
\bar{u} \\
\bar{v} \\
1
\end{bmatrix} = M \begin{bmatrix}
x \\
y \\
1
\end{bmatrix}
\]  (1)

\[
\begin{bmatrix}
\bar{u} \\
\bar{v} \\
1
\end{bmatrix} = \begin{bmatrix}
f_u & 0 & c_u \\
0 & f_v & c_v \\
0 & 0 & 1
\end{bmatrix} \begin{bmatrix}
r_1 & r_{12} & r_{13} & R_{11} & R_{12} & R_{13} & T_1 \\
r_{21} & r_{22} & r_{23} & R_{21} & R_{22} & R_{23} & T_2 \\
r_{31} & r_{32} & r_{33} & R_{31} & R_{32} & R_{33} & T_3
\end{bmatrix} \begin{bmatrix}
x \\
y \\
z \\
1
\end{bmatrix}
\]  (2)

Where the \(\bar{u}\) and \(\bar{v}\) are image coordinates, the \(x, y, z\) are world coordinates, and \(M\) is the ratio between the three-dimensional space and the two-dimensional space.

There are many detailed parameters for the camera inside, with external camera parameters. In the matching formula, as shown in Eq.2, the \(\bar{u}\) and \(\bar{v}\) are image coordinates, the \(x, y, z\) are world coordinates, \(f_u\) is the x focal length, \(f_v\) is the y focal length, \(c_u\) and \(c_v\) are the camera center coordinates, and \(r\) is the camera rotation angle, \(R\) and \(T\) are the corresponding relationships between the camera and LiDAR. \(R\) reflects the camera axis relative position with LiDAR. Because LiDAR is differs from the coordinate system of an RGB camera; \(T\) represents the positional relationship between LiDAR and the camera.

Finally, to find the cameras \(\bar{u}\) and \(\bar{v}\) coordinates, the Eq.2 will be simplified, and the Eq.3 can be obtained. The Eq.4 is the coordinate of \(\bar{u}\). The Eq.5 is the coordinate of \(\bar{v}\).

\[
\begin{bmatrix}
\bar{u} \\
\bar{v} \\
1
\end{bmatrix} = \begin{bmatrix}
m_{11} & m_{12} & m_{13} & m_{14} \\
m_{21} & m_{22} & m_{23} & m_{24} \\
m_{31} & m_{32} & m_{33} & m_{34}
\end{bmatrix} \begin{bmatrix}
x \\
y \\
z \\
1
\end{bmatrix}
\]  (3)

\[
\bar{u} = \frac{m_{13}x + m_{12}y + m_{14}}{m_{31}x + m_{32}y + m_{34}}
\]  (4)
\[ v = \frac{m_{21} \times x + m_{22} \times y + m_{23} \times z + m_{24}}{m_{31} \times x + m_{32} \times y + m_{33} \times z + m_{34}} \]  

(5)

3. Experiment

In this experiment, we used Velodyne LIDAR VLP-16 which has a scanning range of up to 100 meters. The vertical scanning angle is from +15.0 to -15.0 degrees and horizontal scanning range is from 0 to 360 degrees. The vertical resolution is about 2 degrees and horizontal resolution is around 0.1 to 0.4 degrees.

Eight cameras and LiDAR are combined into one platform. Then, the platform can capture the multi-view images. The matching picture can provide point cloud depth data for mapping.

The color data recorded by the digital camera is matched to LiDAR coordinates. We use the overlap between camera and the camera, the overlap must be between one-third and two-thirds to facilitate image stitching, so eight cameras establish a 360-degree environment.

In the LiDAR scanning segment, since the LiDAR scanning range is overlapped between the camera and the camera, the point cloud image is sliced by 8 areas, split by 360 degrees into 8 equal sections, and each block angle is sliced at an of 45 degrees.

The black and white squares on the surface of the box are mainly used for the measurement, and the error calculation is performed manually.

![Figure 1: LiDAR point cloud data](image)

Table 1: Matching average error

<table>
<thead>
<tr>
<th>Camera</th>
<th>Scene 1 (pixel)</th>
<th>Scene 2 (pixel)</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camera 1</td>
<td>1.07</td>
<td>1.16</td>
<td>1.11</td>
</tr>
<tr>
<td>Camera 2</td>
<td>0.50</td>
<td>0.60</td>
<td>0.55</td>
</tr>
<tr>
<td>Camera 3</td>
<td>0.62</td>
<td>1.37</td>
<td>0.99</td>
</tr>
<tr>
<td>Camera 4</td>
<td>1.25</td>
<td>0.50</td>
<td>0.87</td>
</tr>
<tr>
<td>Camera 5</td>
<td>0.37</td>
<td>0.75</td>
<td>0.56</td>
</tr>
<tr>
<td>Camera 6</td>
<td>1.20</td>
<td>0.70</td>
<td>0.95</td>
</tr>
<tr>
<td>Camera 7</td>
<td>1.10</td>
<td>0.80</td>
<td>0.95</td>
</tr>
<tr>
<td>Camera 8</td>
<td>0.50</td>
<td>0.70</td>
<td>0.60</td>
</tr>
<tr>
<td>Total Average</td>
<td>0.82</td>
<td>0.82</td>
<td>0.82</td>
</tr>
</tbody>
</table>

We will overlay the 3D point cloud image on the 2D image to analyze the match integrity. The green region represents the distance scanned by the point cloud image matches with the actual matching distance. It represents correct region. The blue area shows that the distance scanned by the point cloud image does not match the actual matching distance. It represents incorrect area. The red region does not part of regional solution.

Table 1 demonstrates two distinct measurements of the scene for eight cameras. The first scene has an average error of 0.82 (pixel). The second scene has an average error of 0.82 (pixel). The total average of the two scenes is 0.82 (pixel).

As a measurement comparison, we compared the products to match average error and integrity. They are respectively 2.8mm and 0.99.

4. Conclusion

In this paper, we combine eight cameras and LiDAR into one platform. Then we scan the environment through LiDAR and match it to the camera. Because the resolution of the LiDAR is insufficient, this paper can solve the problem of the insufficient resolution of the LiDAR by the characteristics of the camera color images. Finally, we will achieve highly accurate 3D environment detection and identification. In addition, the acquisition parameter is an important section. The error of the parameter will cause an error in the match. In the future, we hope to enhance the algorithm to achieve faster and more accurate parameters.

Acknowledgments

The authors gratefully acknowledge the Taiwan Semiconductor Research Institute (TSRI), for supplying the technology models used in IC design.

5. References


