University of Waikato
Faculty of Computing and Mathematical Sciences

Preliminary Study report of CSMAX570-23A

Mori Cultural Library

Realized by :
Yulin Gu

Song Li

Zhen Chen (John)

University year :
2023A

Contents

List of Figures

List of Tables

General introduction

Over these years, accurate prediction of distances to objects within 2D images has become increasingly important to the industry in computer vision. As the industry grows, more and more methods are being discovered by researchers. The most popular of these should be depth prediction using machine learning.

The hardware approach is also sought after by researchers.

Chapter 1

Data collection

Introduction

Lonrix's business scenario requires the scanning of the road by means of a camera placed in a moving vehicle at about 50km/h to produce measurements of the target in both horizontal and vertical directions.

In this first chapter, we will start with some basic definitions and then focus on how ZED cameras can collect data at a variety of speeds, angles and altitudes.

1.1 ZED camera

1.2 Camera optical concepts

1.2.1 Polariser

1.2.2 Shutter mode

1.2.3 Automatic exposure

1.2.4 Focal length lens

1.3 Basic assumptions for data collection

1.4 Data collection environment

Figure 1.1 – Screenshot of Zed Explorer

You can find settings to turn off auto-exposure and a range of other parameters. They are only valid before each recording, and these values can no longer be changed in the finished SVO file.

The resolution and FPS can be adjusted in the multi selection in the top left corner, but only before the svo recording starts. If this option only has a very low resolution, it is because the usb cable is faulty and cannot support high speed transfers.

1.5 All collected datasets

1.5.1 Feb. 10 at Hillcrest

SN31380347_10-54-05.svo (this file is from Feb. 15 to fix the frame lost issue).

1.5.2 Feb. 15 at Hillcrest

This dataset tried to fix the issue of previous one.

1.5.3 Feb. 10 at MacDonald

This data was collected for very slow speeds (less than 20km/h) with obvious problems on the ground.

1.5.4 Feb. 20 at Lonrix Office

This dataset was collected on the ground in the car park outside the Lonrix office and the distances were marked using a measuring tape.

1.5.5 Feb. 21 near Lonrix Office

This dataset collected vertical poles and traffic signs within different ranges inlcuding 5m and 10m.

1.5.6 Mar. 02 near the University of Waikato

This dataset collected some vertical traffic signs that have been measured in person to obtain ground truth.

1.6 Overview of the collection

Overall, care should be taken when collecting data to make adjustments to optical parameters.

1.6.1 Camera positions and angles

- As the accurate range of the camera is limited, the closer the camera is positioned to the target object the better, but as the camera requires a USB cable connection, be sure to note that a long USB cable is required.
- Whether on the top or on the bonnet of the car, the difference is not that great.
- The angle of the camera has an effect, especially on more distant vertical targets.
- It is important to keep the camera angle as perpendicular to the ground as possible, otherwise this will result in more long-tailed vertical pole point clouds.

1.6.2 USB3.0

ZED cameras use USB 3.0 in order to transfer data in high resolution, but the camera interface is USB-C.

1.6.3 ZED SDK and GPU

The ZED SDK only supports Windows and Ubuntu and does not support macOS.

Conclusion

Collecting data with the ZED is relatively easy, but hardware preparation especially USB cables, GPU machines, or the right model of camera itself can make collection difficult.

Chapter 2

Problem objectives and findings

Introduction

As some parameters are already given in the official documentation, a qualitative analysis based on the data we have collected, only for each particular type of scenario, makes it possible to derive through this report the effective distance and the conditions to be noted for the ZED camera in outdoor conditions.

2.1 A general comparison of horizontal and vertical measurements

2.1.1 Datasets used

In this section the Hillcrest data is selected for an experimental summary.

2.1.2 Objective and method

In the hillcrest dataset there is a traffic sign and some small short lines.

2.1.3 Results

To better display the results, the data presentation has been condensed into a table, as shown in table 2.1.

Position	Resolution	Horizontal	Vertical
Top	720	0.9m-1.1m/1m/<10m/easy	1.92m-1.98/2.1m/<7m/intermediate
Top	1080	0.88m-1m/1m/<10m/easy	2.05m-2.13/2.1m/<9m/intermediate
Top	2k	0.85-1.01m/1m/<9m/intermediate	2.49-2.93m/2.1m/<9m/hard
Bonnet	720	0.9-1.01m/1m/<10m/intermediate	1.92-2.09m/2.1m/<9m/easy
Bonnet	1080	0.8-1.2m/1m/>10m/intermediate	2.38-2.57m/2.1m/>11m/hard
Bonnet	2k	0.8-0.97m/1m/8.6m/intermediate	2.06-2.08m/2.1m/1.8m/easy

Table 2.1 – Results of Horizontal and Vertical objectives

No.	Range	Resolution	Mode	Ground Truth	Range in the view	Result
1	5m	2k	Neural	0.3m*0.5m	1.2m	0.322m*0.429m
2	5m	1080	Neural	0.3m*0.5m	1.2m	0.291m*0.464m
3	5m	720	Neural	0.3m*0.5m	2.2m	0.297m*0.423m
4	10m	2k	Neural	0.3m*0.5m	6.2m	0.293m*0.446m
5	10m	1080	Neural	0.3m*0.5m	5.5m	0.304m*0.522m
6	10m	720	Neural	0.3m*0.5m	7m	0.303m*0.545m
7	5m	2k	Ultra	0.3m*0.5m	1.5m	0.309m*0.457m
8	5m	1080	Ultra	0.3m*0.5m	1.5m	0.309m*0.469m
9	5m	720	Ultra	0.3m*0.5m	2.5m	0.312m*0.462m
10	10m	2k	Ultra	0.3m*0.5m	6.2m	0305m*0.333m
11	10m	1080	Ultra	0.3m*0.5m	5.3m	0.298m*0.857m
12	10m	720	Ultra	0.3m*0.5m	6.1m	0.299m*0.497m

Table 2.2 – Results of Horizonal pavements

Several different values for the data in the result column of the table represent, respectively, the distribution of measured fetch values, ground truth, the estimated distance from the camera, and the perceptual measurement difficulty.

2.1.4 Findings

- ☐ The higher position of the camera is not a significant improvement in accuracy.
- ☐ The higher camera resolution is not a significant improvement in accuracy.
- ☐ Horizontal measurements are significantly easier than vertical ones.

2.2 General comparison of hand-held camera recordings

2.2.1 Datasets used

Because the previous measurements would have revealed difficulties with vertical targets, a series of handheld data collection was done.

2.2.2 Objective and method

In this dataset there is a total of 1 horizontal target and 3 vertical targets at different ranges of 5m and 10m from the camera.

2.2.3 Various scenarios: Horizonal pavements

2.2.3.1 Results

From table 2.2, there are 12 different cases.

2.2.3.2 Findings

It works very well under NEURAL depth mode within the range between 0-10m (1.5m-7m showing in the picture) when it comes pavements.

No.	Range	Resolution	Mode	Ground Truth	Range in the view	Result
13	5m	720	Neural	Pole:2.5m	1.5m	6.19m
14	5m	720	Neural	Sign:4m	1.5m	4.0m
15	5m	720	Ultra	Pole:2.5m	1.5m	2.7m
16	5m	720	Ultra	Sign:4m	1.5m	NaN

Table 2.3 – Results of Vertical signs under backlights

No.	Range	Resolution	Mode	Ground Truth	Range in the view	Result
17	5m	720	Neural	Pole:2.5m	1.7m	2.62m
18	5m	720	Neural	Sign:4m	1.7m	5.1m
19	5m	1080	Neural	Pole:2.5m	0.8m	2.65m
20	5m	1080	Neural	Sign:4m	0.8m	5.37m
21	10m	1080	Neural	Pole:2.5m	6.5m	3.69m
22	10m	1080	Neural	Sign:4m	6.5m	4.13m
23	10m	1080	Ultra	Pole:2.5m	6.5m	2.76m
24	10m	1080	Ultra	Sign:4m	6.5m	NaN

Table 2.4 – Results of Vertical signs under normal sun light

2.2.4 Various scenarios: Vertical signs under backlights

2.2.4.1 Results

It is shown as table 2.3.

2.2.4.2 Findings

Lights make a seriously impact on the Neural depth mode results.

2.2.5 Various scenarios: Vertical signs under normal sun light

2.2.5.1 Results

It is shown as table 2.4

2.2.5.2 Findings

Ultra mode will enhance the pole height accuracy but will not get the sign height.

2.2.6 Various scenarios: Small reflection objects

2.2.6.1 Results

The result is shown as table 2.5.

2.2.6.2 Findings

Neural mode is trying to merge closed pixels by nerual network models so the border area will be not accurate under it.

Conclusion

In this chapter, we have completed a perceptual evaluation of the accuracy of the ZED camera data under a wide range of conditions.

No.	Range	Resolution	Mode	Ground Truth	Range in the view	Result
25	5m	1080	Ultra	Pole:0.97m	1.2m	0.82/0.906
26	5m	720	Ultra	Panel:0.38m*1.01m	1.2m	NaN * 1.08m Top width: NaN
27	10m	1080	Ultra	Pole:0.97m	6.6m	0.807/0.919
28	10m	1080	Ultra	Panel:0.38m*1.01m	6.6m	NaN*3.21m
29	5m	1080	Neural	Pole:0.97m	1.2m	0.861/0.818
30	5m	1080	Neural	Panel:0.38m*1.01m	1.2m	2.66m*0.9m
31	10m	1080	Neural	Pole:0.97m	6.6m	1.13/1.33
32	10m	1080	Neural	Panel:0.38m*1.01m	6.6m	2.94m*1.24m

Table 2.5 – Results of Small reflection objects

In the next chapter, we will introduce algorithms to accomplish the enhancement of the depth accuracy and thus the measurement accuracy.

Chapter 3

In-depth study of 3D Registration

3.1 Introduction

In order to obtain highly accurate depth images over a large range on a low accuracy device, it becomes possible to stitch multiple frames from the camera by aligning the point cloud. The high accuracy over a short range is extended to the accuracy over a long range by algorithms.

3.1.1 3D reconstruction and Depth estimation

It requires a device that supports time-of-flight (TOF) lasers, microwaves or 3D ultrasound.

3.1.2 Depth and Point clouds

The depth information can be used to generate point clouds, or to augment existing point clouds with additional depth information.

3.1.3 Point clouds and Registration

Point cloud registration is often used to combine or merge multiple point clouds into a single, comprehensive model of a scene or object.

3.1.4 ICP

Iterative closest point (ICP) is a widely used algorithm in 3D computer vision and robotics for aligning and registering two or more point clouds or 3D models.

3.1.5 SVO file format

SVO is the recorded video file format from the ZED camera.

3.2 Direct performance of the ZED camera without any processing

For fast svo file processing, a toolbox project was implemented. In this project, the most important data is the depth data that can be dump from SVO file and stored as npz file.

```
distance = math.sqrt((X1 - X2) * (X1 - X2) +
                       (Y1 - Y2) * (Y1- Y2) +
                      (Z1 - Z2) * (Z1 - Z2))
```

3.2.1 Parameters of ZED

Because ZED uses both Neural models and stereo methods to predict the depth value, there are a lot of parameters.

```
depth_mode = NEURAL
depth_minimum_distance = 0.01 m
depth_maximum_distance = 30 m
```

These parameters can be modified after the SVO has been generated.

3.2.2 Best working distance

The camera, which is ZED 2i 2.1mm with polarizer, works well in the outdoor environment when the depth value is in the near range (0m-7.5m).

3.2.3 Resolutions

The higher resolution will provide more details but the FPS will be reduced.

3.2.4 The shortage of ZED

In Lonrix scenario, the most common objective is the traffic sign, which is a white or blue metal panel. (Fig. 3.1.)

Figure 3.1 – Different surfaces and ranges

It is not too bad when close enough in the actual test results.

3.2.5 Results

There are a lot of highly accurate measurements that results in different resolutions and speeds, which is shown in Fig. 3.2.

Figure 3.2 – Results in different ranges

Fig. 3.3 shows an example of the different values for the same white line.
A sufficiently dense and accurate point cloud is the basis for accurate measurements.

Figure 3.3 – 2.86m vs 4.08m

3.3 Optimising point clouds

There are quite a few issues if the range is beyond 7m.

Figure 3.4 – HD2k/50km/h, open3d image, 14m

3.3.1 ICP registration

Fig. 3.5 shows the single frame that only keeps 7m.

Figure 3.5 – HD720 50km/h Frame No. 38

With the camera pose data, a pose matrix can be generated, which makes it easier to convert from local to world coordinates

3.3.2 More on 3d reconstruction

Because of the ICP registration, it is possible to get entire 3d model in an area, so that it is possible to get the measurements at various angles.

3.4 Code

The toolbox project can deal with the SVO files and do the measurements.

3.4.1 Dump files from SVO file

This file is used for dumping the SVO file into there kinds of file: JPG, PCD and NPZ.

```
mkdir -p example_files/UNI-HD720/
python runtools.py -f ~/Downloads/mar02/HD720_SN31380347\_14-31-25.svo -s
    example_files/UNI-HD720/ -g 1
```

The parameters represent the address of the SVO file, the directory to which it is exported, and the number of frames between each export.

3.4.2 Do frame measurement showing as single color image

```
python runtools.py -t 3 -f example_files/UNI-HD720/ -frame 730 -show 1
```

The result will display an image, and the selected distance can be shown by clicking with a mouse.

3.4.3 Do frame measurement showing as left depth plus image mode

```
python runtools.py -t 3 -f example_files/UNI-HD720/ -frame 730 -show 2
```

3.4.4 Do frame measurement under the YOLO result

This command requires YOLO installation in advance.

```
python runtools.py -t 3 -f example_files/UNI-HD720/ -frame 730 -show 3
```

3.4.5 Run colored ICP

This tool will remove the pavement and the long tails of the pole.

```
python runtools.py -t 4 -f example_files/UNI-HD720/ -frame 730 -rp 1 -rl 1
```

Figure 3.6 – Screenshot of colored ICP

3.4.6 Run manual registration then run a micro distance ICP

```
python runtools.py -t 4 -f example_files/UNI-HD720/ -frame 730 -a 2
```

This tool will not remove the pavement and the long tails of the pole.

3.4.7 Run RANSAC registration

```
python runtools.py -t 4 -f example_files/UNI-HD720/ -frame 730 -a 3
```

This tool will not remove the pavement and the long tails of the pole.

Figure 3.7 – Screenshot of manual ICP

Figure 3.8 – Screenshot of RANSAC

3.5 Conclution and future work

The ZED camera provides a reliable and accurate 3D sensing solution with high-quality depth sensing and stereo imaging capabilities in specific range.

3.5.1 Future work

There is some work that needs to be done in the future to ensure this camera is used in online products.

3.5.2 Todo List