Road Pothole Detection System Based on Stereo Vision

Yaqi Li Christos Papachristou Daniel Weyer
Department of Electrical Engineering and Computer Science
Case Western Reserve University
Cleveland, Ohio 44106
{yxl1544, cap2} @case.edu

***Abstract*—In this paper, we propose a stereo vision system which detects potholes during driving. The objective is to benefit drivers to react to potholes in advance. This system contains two USB cameras taking photo simultaneously. We use parameters obtained from camera calibration with checkerboard to calculate the disparity map. 2-dimensional image points can be projected to 3-dimensional world points using the disparity map. With all the 3-dimensional points, we use the bi-square weighted robust least-squares approximation for road surface fitting. All points below the road surface model can be detected as pothole region. The size and depth of each pothole can be obtained as well. The experiments we conducted show robust detection of potholes in different road and light conditions.**

***Index Terms*—stereo vision, road potholes, surface fitting, disparity**

I. Introduction

Potholes are bowl-shaped openings on the road that can be up to 10 inches in depth and are caused by the wear-and-tear and weathering of the road [1]. They emerge when the top layer of the road, the asphalt, has worn away by lorry traffic and exposing the concrete base. Once a pothole is formed, its depth can grow to several inches, with rain water accelerating the process, making one of the top causes of car accidents. Potholes are not only main cause of car accidents, but also can be fatal to motorcycles. Potholes on roads are especially dangerous for drivers when cruising in high speed. Because, the driver can hardly see potholes on road surface. Moreover, if the car passes potholes at high speed, the impact may rupture car tires. Even though drivers may see the pothole before they pass it, it is usually too late to react. Any sharp turn or suddenly brake, may cause car rollover or rear-end.

Motivated from the above reasons, we decided to investigate a system to detect potholes on roads while driving. The proposed system will produce the 3-dimensional information of potholes and determine the distance from pothole to car for informing the driver in advance.

Currently, the main methods for detecting potholes still rely on public reporting through hotlines or websites, for example, the potholes reporting website in Ohio [2]. However, this reporting usually lacks accurate information of the dimensional and location of potholes. Moreover, this information is usually out of date as well.

A method to detect potholes on road has been reported in a real-time 3D scanning system for pavement distortion inspection [3] which uses high-speed 3D transverse scanning techniques. However, the high-speed 3D transverse scanning equipment is too expensive. Rajeshwari Madli et al. have proposed a cost-effective solution [4] to identify the potholes on roads, and also to measure the depth and height of each pothole using ultrasonic sensors. All the pothole information is stored in database (cloud). Then alerts are provided in the form of a flash messages with an audio beep through android application. To detect the depth of pothole correctly, the ultrasonic sensor should be fixed under the car, which means the car should pass the pothole first.

2D vision-based solutions can detect potholes as well [5]. Regions corresponding to potholes are represented in a matrix of square tiles and the estimated shape of the pothole is determined. However, the 2D vision-based solution can work only under uniform lighting conditions and cannot obtain the exact depth of potholes.

To remove the limitations of the above approaches, we propose a detection method based on computer stereo vision, which provides 3-dimensional measurements. Therefore, the geometric features of potholes can be determined easily based on computer vision techniques. The proposed method requires two cameras to take photos simultaneously. Compared with the expensive high-speed 3D transverse scanning equipment, USB cameras are affordable and flexible. The geometric information of road potholes can be obtained by the stereo camera.

In the remainder of this paper, section 2 briefly discusses the background of this work. Section 3 introduces the technical approach and principles that we applied in this pothole detection system. Section 4 illustrates the experimental setup of the proposed system and results. Finally, section 5 concludes the proposed method and discusses the future work which can be done to improve the pothole detection system.

II. Background

Stereo camera parameters, including intrinsic parameters and extrinsic parameters, are obtained with a checkerboard using Zhang's camera calibration method [6]. Before converting the image coordinates to the world coordinates, some preparation work needs to be done.

Image pairs of road surface should be undistorted and rectified. In undistorted images lens distortion has been removed. Rectification refers to projecting image pairs onto a common

978-1-5386-6557-2/18/$31.00 ©2018 IEEE

image plane, respectively. The rectified and undistorted image pairs are used to calculate the disparity map using the stereo camera parameters obtained before with the semi-global matching algorithm [7] provided in OpenCV (Open Source Computer Vision Library) [8]. This algorithm uses a pixelwise, Mutual Information-based matching cost for compensating radiometric differences of the stereo image pairs.

The disparity map illustrates the corresponding pixels'difference in a pair of stereo images. Thus, with the disparity map, image points can be transferred to world points. The road surface is fitted using the bi-square weighted robust least-squares algorithm [9] with all the points in world coordinate of the road surface image . Subsequently, all the points below the road surface correspond to the pothole region. In case there are more than one pothole in the region of interest, pothole points are labelled into different potholes according to their connections using the connected component labelling algorithm [10].

III. Approach to Pothole Detection System

A. Overview of Pothole Detection System

In this section, we introduce the proposed road pothole detection system. The proposed system consists of 2 modules: Off-line processing and on-line processing. The off-line flowchart of the proposed system is illustrated in Figure 1. Stereo camera parameters, including intrinsic parameters and extrinsic parameters, are obtained using a checkerboard based on Zhang's camera calibration method [6]. The on-line flowchart of the proposed system is shown in Figure 2. There are 3 main modules: Image processing, disparity calculation and pothole detection. Before transferring the image coordinates to the world coordinates, some preparation work needs to be done. The 3 modules will be discussed in following section.

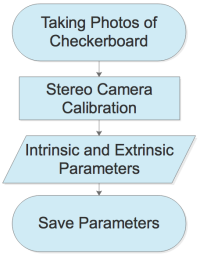

Fig. 1. Off-line Flowchart of the Pothole Detection System.

B. Stereo Camera Calibration

Camera parameters are necessary for disparity calculation. We can obtain camera parameters, both intrinsic and extrinsic parameters, by stereo camera calibration. To calibrate stereo cameras, a flexible camera calibration approach proposed by Zhang [6] is used in this system. Compared with other classic camera calibration methods which use expensive equipment, the method proposed is economical and flexible. Only two

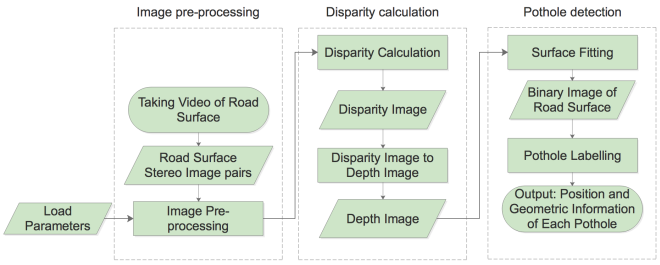

Fig. 2. On-line Flowchart of the Pothole Detection System.

USB cameras are used to observe a checkerboard in different orientations. The proposed system uses an 8×6 checkerboard with 24.5 mm squares. Either two cameras or the checkerboard can be freely moved. Theoretically, a minimum of 2 orientations are needed for camera calibration. However, 20 orientations are used in this system for better quality. Figures 3 and 4 illustrate the relative movement between the checkerboard and the stereo camera pair regarding different frame of reference. All checkerboards'orientations regarding the fixed stereo camera pair are shown in Figure 3. Relatively, all the stereo camera's orientations regarding fixed checkerboards are shown in Figure 4.

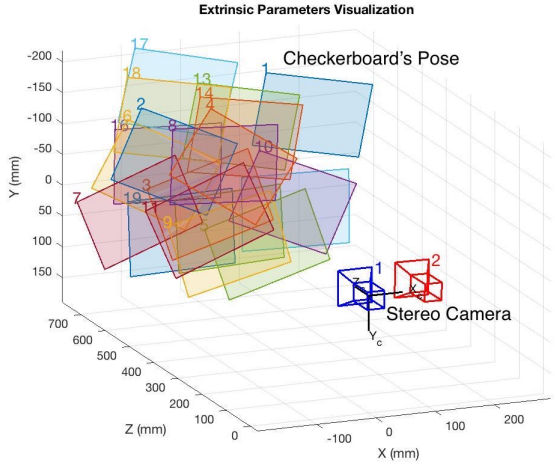

Fig. 3. Checkerboards'Orientations with Fixed Stereo Camera.

Both intrinsic and extrinsic parameters of the stereo camera can be obtained from the camera calibration process. The intrinsic parameters, named camera matrix as well, are independent. Therefore, once the intrinsic parameters are determined, they can be used as long as the focal length of the camera remain unchanged. The output 3×3 camera matrix is shown as follows:

$$A = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}. \quad (1)$$

f_x, f_y are the focal lengths expressed in pixel units. (c_x, c_y)

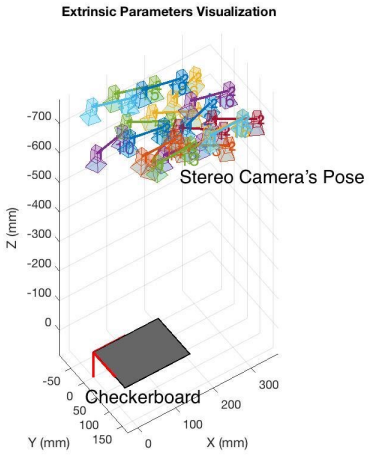

Fig. 4. Stereo Camera's Orientations with Fixed Checkerboards.

is the principal point that is usually the center of the image. The joint rotation-translation matrix is called extrinsic matrix. Extrinsic parameters can translate a point (X, Y, Z) to the coordinate system with fixed camera. The translation is illustrated by the following:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R * \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + t. \tag{2}$$

R is the rotation matrix. t is the translation vector. The joint rotation-translation matrix Rt is the extrinsic matrix. It can transfer image-coordinate points X, Y, Z to world-coordinate pointsx, y, z. With both intrinsic and extrinsic parameters, some pre-processing work including removal of lens distortions and rectification of the stereo image pairs should be done. At this time, light correction can be added to remove the influences caused by lighting condition as well. Therefore, the stereo image pairs can be turned into standard form whose corresponding points located at same horizontal line.

C. Disparity Calibration

In this paper, we use computer stereo vision based methods to detect potholes on road surfaces. Stereo vision is an attempt to imitate the eyes of human beings.

For the single camera, as shown in Figure 5, two different real points P and Q project to a same point in the image plane when they are located in the same line with the optical center. However, when it comes to stereo vision, as illustrated in Figure 6, we are able to obtain the depth by means of triangulation, if we can find the corresponding pixels in the stereo image pairs.

As shown in Figure 7, the B is the distance between the 2 cameras'optical centers. f is the focal length obtained from stereo camera calibration. XLR and X-OL-OR are equivalent triangles. We can write their equivalent equation as following:

$$\frac{B}{z} = \frac{B + XR - XL}{z - f}. \tag{3}$$

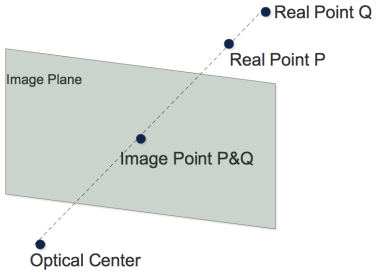

Fig. 5. Single Camera.

Therefore, we can calculate using following equation:

$$z = \frac{B * f}{XL - XR} = \frac{B * f}{d}. \tag{4}$$

z is the depth of point X, and it is inversely proportional to the disparity. So once we find the corresponding points in the stereo image pairs, we can calculate the disparity and the depth of a real point on roads correctly.

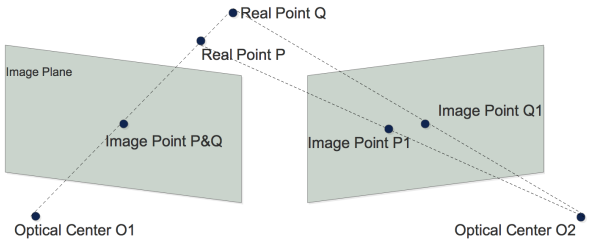

Fig. 6. Stereo Camera.

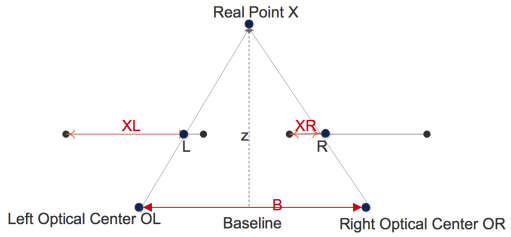

Fig. 7. depthcalculation.png.

The main difficulty is how to find the best corresponding points in the stereo image pairs. In this paper, we use the semi-global Matching algorithm proposed by Heiko [7] which is provided in OpenCV [8].

D. Image Reprojection

Having got the disparity map, we can use a triangulation method to reproject the disparity image to 3D space. For the stereo cameras, given the disparity image and the camera parameters like camera's focal length, we can calculate the 3 dimensional coordinates in real world.

As shown in Figure 8, the optical center of camera L is the origin. f is the focal length of the camera. Triangle OAE and triangle OMP are similar triangles, therefore we can write:

$$\frac{z}{f} = \frac{x}{XL}. \tag{5}$$

For the right camera, triangle OCD and triangle ONP are similar triangles:

$$\frac{z}{f} = \frac{x - B}{XR}. \tag{6}$$

Similarly, along the Y- axis, we can obtain:

$$\frac{z}{f} = \frac{y}{YL} = \frac{y}{YR}. \tag{7}$$

The 3 dimensional points as a function of the disparity can be derived from the above equations and equation (4):

$$x = \frac{B * XL}{d}, \tag{8}$$

$$y = \frac{B * YL}{d}. \tag{9}$$

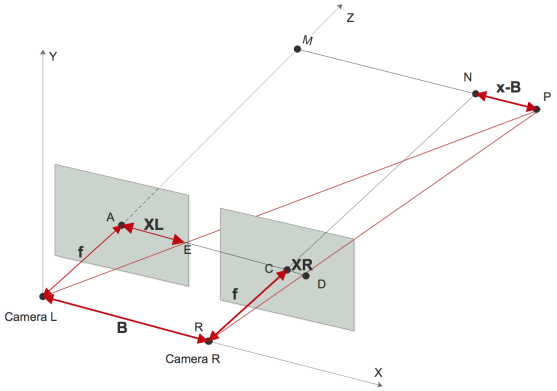

Fig. 8. Triangulation Method to Reproject Disparity Image.

E. Road Surface Fitting

With all the 3 dimensional world coordinates points, we can fit the road surface using the bi-square weighted robust least-squares method [11] [12]. All the points are usually regarded as equal quality when fit to a road surface using the least square method [13]. This includes those pothole points that are below the road surface or those noise points that might influence the accuracy of the fitted road surface. The bi-square weighted robust least-squares method [11] used in this road pothole detection system minimize the outliers'influences during the fitting processes by adding an additional scale factor (the weight). All the outliers below the road surface can be detected as pothole regions.

F. Road Pothole Labelling

We need to label the detected pothole region using the connected component labelling algorithm [10]. Or in case there are more than one road pothole in the region of interest, we number these road potholes with the connected component labelling algorithm [10]. The labelling process consists of 2 passes.

The flowchart of the first pass of judgement is illustrated in Figure 9. We obtain a pixel from the binary image and check whether this pixel is a background pixel. We only label non-background pixel. For non-background pixel, we check whether there is a existing label around this pixel. If there is not a existing label around this pixel, create a new label for it. Otherwise copy the existing label if there is only 1 label around this pixel. Copy the smaller label and mark the bigger label as a child label of the smaller label.

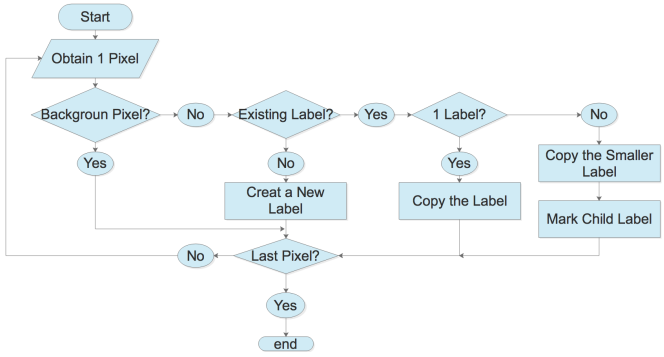

Fig. 9. The First Pass Flowchart.

The flowchart of the second pass of judgement is illustrated in Figure 10. We obtain a pixel from the result labelling image of the first pass and check whether it is a child label. We replace it with the parent label if it is a child label.

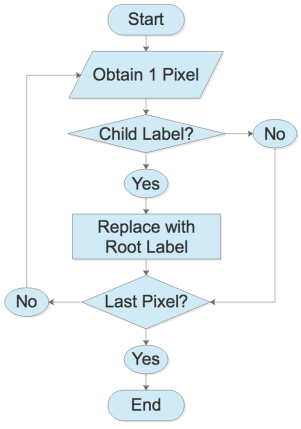

Fig. 10. The Second Pass Flowchart.

IV. Experimental Setup and Results

A. Experimental Setup

In the prosed system, we need to perform two experiments. Firstly, we calibrate the camera which should be done only once unless we change with another pair of stereo cameras. The camera calibration experiment is done in Glennan 519C laboratory. We take 20 image pairs of the checkerboard's different poses including position and posture. The posture of the checkerboard in every image should be as different as possible.

Then, we can employ the experimental setup in real road environment to detect road potholes. We detect a pothole for our experiments in the Case Western Reserve University parking lot 1. We detect the pothole by moving two USB cameras mounted on the roller cart to imitate the cameras moving with the car, but at a low speed. The experimental setup for pothole detection is shown in Figure 11. The experiment is done serval times under different weather including snowy day and sunny day. Both of these two experiments use the same experimental setup including 2 USB cameras, optical rial, tripod and raspberry Pi 2 model B. For the software setup, opencv-python 3.3 and python 3.6 are required.

We use two dependent cameras to imitate the stereo camera, and we should make sure the two cameras are mounted at the same height and same horizontal direction using the optical rail and tripod. In the proposed system, we use the Raspberry Pi 2 Model B [14], which is a single-board computer.

Fig. 11. The Experimental Setup for Pothole Detection.

B. Results

The final road potholes detection result shown in Figure 12 with their geometric information.

Taking a 640×480 image as an example, Figure 13 is the time consumption report for the road pothole detection system.

The single camera calibration using 20 checkerboard images for 2 cameras together takes around $3.52s$. However, we do not need to calibrate every time when we start to detect road pothole. The camera calibration can be done in advance, and all camera parameters saved locally for later use. Of course, the time consumption for road pothole labelling depends on how many potholes detected. If 0 pothole detected, the time consumption for this part is 0 as well. Therefore, the total time of a road pothole detection is around $5s$, which means if the car drives at $30miles/hour$, we should detect the road pothole at least $14meters$ in advance.

The above timing consumption depends on image size as well. The larger the image size is; the more pixels need to be processed; the longer it takes. Therefore, we only select the road surface in front of car as region of interest (ROI) to reduce the timing consumption for the proposed road pothole detection system.

Fig. 12. Pothole with Geometric Information in CWRU Parking Lot 1.

Category	Calibration	Prepocessing	Disparity Calculation	Surface Fitting	Pothole Detection	Pothole Labelling	Total
Time(seconds)	3.52	0.18	0.17	2.75	1.57	0.26	4.94

Fig. 13. Timing Consumption Report of Proposed System.

REFERENCES

[1] "Potholes are a Top Cause of Auto Accidents," Law Offices of Michael Pines, APC. [Online]. Available: https://seriousaccidents.com/legal-advice/top-causes-of-car-accidents/potholes/. [Accessed: 02-May-2018].

[2] Pages - Welcome to The Ohio Department of Transportation Home Page. [Online]. Available: http://www.dot.state.oh.us/pages/home.aspx. [Accessed: 02-May-2018].

[3] Q. Li, M. Yao, X. Yao, and B. Xu, "A real-time 3D scanning system for pavement distortion inspection," Measurement Science and Technology, vol. 21, no. 1, p. 015702, 2009.

[4] R. Madli, S. Hebbar, P. Pattar, and V. Golla, "Automatic Detection and Notification of Potholes and Humps on Roads to Aid Drivers," IEEE Sensors Journal, vol. 15, no. 8, pp. 4313–4318, 2015.

[5] S. B. S. Murthy and G. Varaprasad, "Detection of potholes in autonomous vehicle," IET Intelligent Transport Systems, vol. 8, no. 6, pp. 543–549, Jan. 2014.

[6] Z. Zhang, "A flexible new technique for camera calibration," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 11, pp. 1330–1334, 2000.

[7] H. Hirschmuller, "Stereo Processing by Semiglobal Matching and Mutual Information," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 2, pp. 328–341, 2008.

[8] "OpenCV library," OpenCV library. [Online]. Available: https://opencv.org/. [Accessed: 03-May-2018].

[9] G. H. Golub and U. V. Matt, "Quadratically constrained least squares and quadratic problems," Numerische Mathematik, vol. 59, no. 1, pp. 561–580, 1991.

[10] U. Sinha, "Labelling connected components - Example," AI Shack. [Online]. Available: http://aishack.in/tutorials/labelling-connected-components-example/. [Accessed: 03-May-2018].

[11] X. Ai, Y. Gao, J. Rarity, and N. Dahnoun, "Obstacle detection using U-disparity on quadratic road surfaces," 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013), 2013.

[12] G. H. Golub and U. V. Matt, "Quadratically constrained least squares and quadratic problems," Numerische Mathematik, vol. 59, no. 1, pp. 561–580, 1991.

[13] "Documentation," MATLAB & Simulink. [Online]. Available: https://www.mathworks.com/help/curvefit/least-squares-fitting.html. [Accessed: 03-May-2018].

[14] "Raspberry Pi 2 Model B," Raspberry Pi. [Online]. Available: https://www.raspberrypi.org/products/raspberry-pi-2-model-b/. [Accessed: 03-May-2018].