AN EFFICIENT ALGORITHM FOR POTHOLE DETECTION USING STEREO VISION

Zhen Zhang, Xiao Ai, C. K. Chan and Naim Dahnoun

Department of Electrical and Electronic Engineering
University of Bristol
Merchant Venturers Building
Woodland Road, Bristol
BS8 1UB, United Kingdom

ABSTRACT

In this paper, a stereo vision based pothole detection system is proposed. Using the disparity map generated from an efficient disparity calculation algorithm, potholes can be detected by their distance from the fitted quadratic road surface. The system produces the size, volume and position of the potholes which allows the pothole repair to be prioritised according to its severity. The quadratic road surface model allows for camera orientation variation, road drainage and up/down hill gradients. Experimental results show robust detection in various scenarios.

Index Terms— Stereo vision, pothole detection, disparity calculation

1. INTRODUCTION

Potholes in British roads have become an increasingly serious problem, with a third of vehicles experiencing pothole related damage in the past 2 years [1]. The costs to maintain the road are billions of pounds and local councils pay millions of pounds in compensation [2, 3]. Current methods for detecting potholes in urban roads rely mainly on public reporting, such as through hotlines or social networking websites, which are resource intensive and ineffective. For example, in the most well-known *fixmystreet.com* site [4], only a small percentage of problems have been reported. It is inconvenient to report the size and severity of potholes with the current system. Consequently, a lack of information about the size and severity affects the priority of patching the potholes.

More advanced road surveying techniques exist for highway maintenance, where specialized vehicles equipped with laser scanners are in use [5]. Due to their high costs, typical survey cycles are once every 4 years. For urban roads, Mednis et al. proposed a system that utilises a mobile phone accelerometer to detect and locate potholes [6]. It records the vibration of vehicles when hitting a pothole. By this means, the severity can be approximated by the strength of the vibration. However, the vehicle has to be driven into the pothole, which could cause damage. Machine vision based methods for automatic pothole detection have also been proposed, which only require a camera as input. However, existing approaches rely on the texture of the road surface, resulting in low accuracy. Mis-detections are mainly due to variation of lighting conditions [7].

Stereo vision based techniques, on the other hand, provide 3D measurements, so that the geometric features of a pothole can be determined easily. Stereo vision can provide information on the size of the pothole, without the need for using high cost specialized laser scanners. By putting the system on a vehicle that patrols the roads (whilst undertaking other jobs), it will be possible to continuously detect and evaluate potholes. It is important that the system can perform pothole detection in real-time to reduce data storage. The major difficulty of implementing a real-time and robust stereo vision system resides in the intensive calculations involved in generating the disparity map with high accuracy. Generally, disparity calculation methods [8], such as graph cut (GC) [9], belief propagation (BP) [10] and dynamic programming (DP) [11], would not meet the real-time requirement without specialized hardware accelerators [12].

We have proposed a real-time disparity calculation algorithm [13], in which the disparity search range for a pixel at the image line (v) is supplied by the disparity values calculated at three neighbouring pixels at a lower image line (v-1). This method yields a 90 percent computation saving in comparison to the most basic block matching algorithm. At the same time, the reduction in the search range leads to lower ambiguities which then produces a higher accuracy. Due to the high surface continuity of the road and pothole areas, this algorithm performs exceptionally well in pothole detection applications.

The flowchart of the proposed pothole detection system is illustrated in Figure 1. This paper deals with the highlighted part of the system, in which the disparity map is firstly calculated from left/right image pairs captured by stereo cameras

Thanks to University of Bristol Enterprise Development Fund to provide seed fund for our preliminary studies.

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

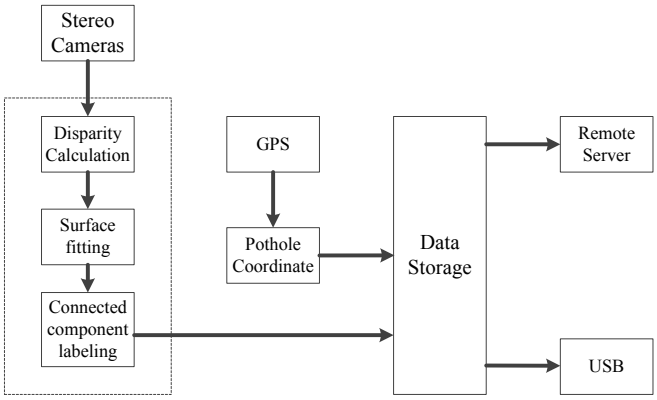

Fig. 1: Pothole detection system.

using a computationally efficient algorithm. Subsequently, the obtained disparity map is converted into a world coordinate system (WCS) point cloud. By fitting these 3D points into a quadratic road surface, outliers that are lower than the fitted surface will be determined as pothole points. The quadratic road surface has its first order terms related to the camera orientation and second order terms indicating the drainage and up/downhill gradients. With the 3D points corresponding to the potholes determined after surface fitting, the relevant pothole areas in the disparity map can be determined. Finally, with the connected component labelling (CCL) algorithm applied, potholes can be labelled as pothole candidates. When a pothole candidate is determined in a particular frame, its GPS location, original image pairs, size and volume of the potholes and 3D point cloud can be saved and accessed later via the Internet or USB port.

In the reminder of this paper, Section 2 introduces details of the proposed algorithm, Section 3 demonstrates the experimental results of the proposed algorithm and Section 4 concludes the paper.

2. POTHOLE DETECTION SYSTEM

2.1. Disparity calculation

In the disparity calculation procedure, it is assumed that the input images are rectified and co-planar, so that the epipolar lines are aligned with the corresponding scanlines. In this case, the correspondence can only exist on the same scanline. If $p(u, v)$ and $q(u_0, v)$ are corresponding pixels in the left and right images respectively, then the disparity d between $p(u, v)$ and $q(u_0, v)$ is defined as $d = u - u_0$. The authors have previously presented a computationally efficient algorithm for disparity calculation in street scenes [13]. Its attained efficiency relies on a reduction in the search range. In a typical street scene, the road is a supporting surface which extends from the near to the far field, such that the near field road pixels should show a large disparity value whilst the far field ones are likely to have a small disparity value. Obstacles protrude from the road surface but potholes will be lower than the road. The base of the obstacle should have a disparity level that is similar to its neighbouring road disparity level; similarly, the pothole disparity is not very different from that of the road. . These means that the disparity search range of a given image line (v) can be supplied by the disparity values of its neighbouring pixels in the lower image line (v-1).

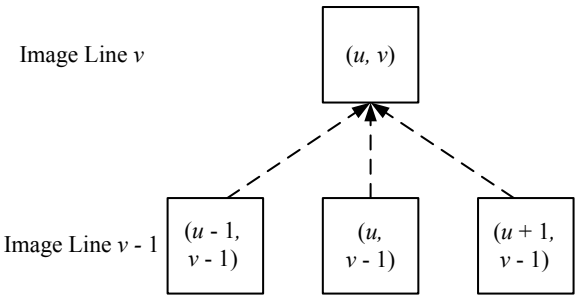

Fig. 2: Controlled search range generation.

The disparity calculation algorithm consists of three steps: matching cost computation, search range recalculation and, finally, disparity enhancement. First of all, it is assumed that every obstacle is on the road surface and that the bottom part of the obstacle has the same disparity as the road surface of the same image line. These steps are discussed below.

2.1.1. Matching cost computation

One of the most important parts in a disparity calculation algorithm is the cost function. The Sum of Absolute Difference (SAD), the Sum of Squared Difference (SSD), the Normalised Cross-Correlation (NCC) and the Sum of Hamming Distances (SHD) are the most used cost functions [12]. By calculating the cost, the cost volume $C(u, v, d)$ for each pixel is obtained. For each pixel's matching cost function, a distinct peak may exist which indicates the correct correspondence. However, this process consumes a large amount of computational power during the exhaustive correspondence searching and, when complex situations are encountered, errors are likely to be introduced in homogeneous regions. To solve this problem, it is proposed that the algorithm should generate a controlled search range.

2.1.2. Controlled search range

Given the road obstacle assumption, the disparity of the current pixel $d(u, v)$ would have connections with the disparities of the pixels of the lower line $d(u - 1, v - 1)$, $d(u, v - 1)$ and $d(u + 1, v - 1)$, as shown in Figure 2. By exploring these connections, the neighbourhood disparities of the pixels in the image line $v - 1$ can be used to generate a much smaller search range for the pixels in the image line v. By controlling the search range $SR(u, v)$ according to the neighbourhood support points, the potential matching ambiguities

can be reduced. Consequently, a much faster and more accurate calculation compared to the exhaustive search algorithms can be achieved. The proposed algorithm generates a controlled search range according to Equation 1.

$$SR(u,v) = SR(u-1,v-1) \cup SR(u,v-1) \cup SR(u+1,v-1) \quad , \tag{1}$$

where

$$SR(u-1,v-1) \in \{(d(u-1,v-1)-\tau)\dots(d(u-1,v-1)+\tau)\} \quad , \tag{2}$$

$$SR(u,v-1) \in \{(d(u,v-1)-\tau)\dots(d(u,v-1)+\tau)\} \quad , \tag{3}$$

$$SR(u+1,v-1) \in \{(d(u+1,v-1)-\tau)\dots(d(u+1,v-1)+\tau)\} \quad , \tag{4}$$

where $d(u,v)$ represents the disparity value of the pixel in position (u,v) and τ denotes the bound of the search range. By minimising τ, the $SR(u,v)$ can be minimised. In this work, we have applied $\tau = 1$, such that only 1 pixel variation is allowed, from one of its three neighbours in the lower image line $(v-1)$, to contribute to the search range of the current line (v).

2.2. Pothole detection

After disparity calculations, the next step is to detect the potholes. In the proposed algorithm, a surface fitting algorithm is used to estimate the road surface. Hence, points that are lower than the road surface can be defined as potholes. The detection and segmentation is achieved by CCL.

2.2.1. Conversion between the disparity and Euclidean

Points in the disparity domain ($[u,v,d]$ space) can be projected back to the world coordinate system ($[x,y,z]$ space) using the relationship as described in Equation 5.

$$x_i = \frac{z_i \cdot u_i}{f}, y_i = \frac{z_i \cdot v_i}{f}, z_i = \frac{b_s \cdot f}{d_i}, \tag{5}$$

where i is the pixel index, $[u,v]$ is referred to as the calibrated coordinates rather than the physical pixel locations, with b_s and f denoting the base-line and focal length distances. For the i_{th} pixel in the disparity image and WCS, one-to-one point associations are maintained as $D_i \sim [u,v,D]_i \quad \sim [x,y,z]_i$. For example, when the o_{th} pixel is classified as the pothole point in the disparity image (denoted by D_p), its equivalent point in the Euclidean domain will be found at $[x,y,z]_o$.

2.2.2. Surface fitting

With the disparity image obtained and the WCS point cloud calculated, potholes can be obtained by surface fitting, subsequently eliminating the road surface areas. Generally, roads can be modelled efficiently by a quadratic surface in the Euclidean domain. The proposed fitting algorithm uses a low computational bi-square weighted robust least-squares method [14, 15]. This method can reduce the effect of outliers. To the fitting procedure, the outliers would be obstacles or potholes. We define the quadratic road model as:

$$z = a_1 + a_2x + a_3y + a_4x^2 + a_5xy + a_6y^2. \tag{6}$$

The equation takes into account the twisting and bending nature of the road surface. A similar model has been used in [16] without the xy term. With the xy term being added in our model, twisting of the road surface can be represented. An estimated surface that is close to the road can be fitted by finding the coefficients $a_1, \dots, a_6$. The pothole can then be detected by taking the actual surface, z, and subtracting the estimated surface, $\widehat{z}$, from it. The difference between the data points on the surface will contain the potholes as its data points should be lower than the fitted surface.

The steps to obtain the coefficients can be broken down to as follows:

1. Fit the surface using normal least-squares:

$$S = \sum_{i=1}^{n} (z_i - \widehat{z_i})^2. \tag{7}$$

2. Minimise the residual, $r = (z_i - \widehat{z_i})^2$ by differentiating the sum with respect to the coefficients. The resulting equations in matrix form are Equation 8, where, for example, $S_{xy} = \sum_{i=1}^{n} w_i x_i y_i$. The initial weights, w, are 1.

$$\begin{pmatrix} n & S_x & S_y & S_{x^2} & S_{xy} & S_{y^2} \\ S_x & S_{x^2} & S_{xy} & S_{x^3} & S_{x^2y} & S_{xy^2} \\ S_y & S_{xy} & S_{y^2} & S_{x^2y} & S_{xy^2} & S_{y^3} \\ S_{x^2} & S_{x^3} & S_{x^2y} & S_{x^4} & S_{x^3y^2} & S_{x^3y} \\ S_{xy} & S_{x^2y} & S_{xy^2} & S_{x^3y} & S_{x^2y^2} & S_{xy^3} \\ S_{y^2} & S_{xy^2} & S_{y^3} & S_{x^2y^2} & S_{xy^3} & S_{y^4} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{pmatrix} = \begin{pmatrix} S_z \\ S_{xz} \\ S_{yz} \\ S_{x^2z} \\ S_{xyz} \\ S_{y^2z} \end{pmatrix} \tag{8}$$

3. Normalise the residuals, where K is the tuning constant 4.685, s represents the robust variance (the median absolute deviation of the residual)/0.6745.

$$r_{adj} = \frac{r_i}{\sqrt{1-h_i}} \qquad u = \frac{r_{adj}}{Ks}$$

4. Compute the bi-square weights for the i_{th} point as follows:

$$w_i = \begin{cases} (1-(u_i)^2)^2 & |u_i| < 1 \\ 0 & |u_i| \geq 1 \end{cases} ,$$

5. Check if the fit converges. If not, recompute the weight with the new residual and iterate until it converges. Typically, convergence can be achieved between 2 to 4 iterations.

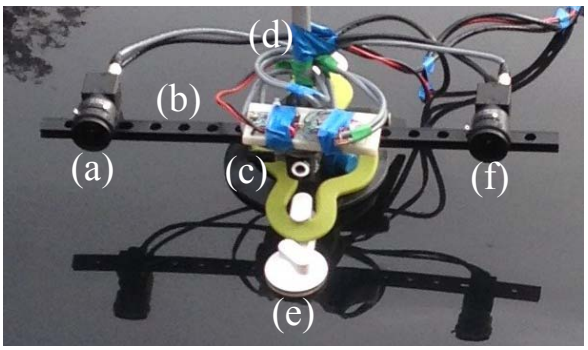

Fig. 3: Experimental setup includes (a,f) PointGrey Flea 3 Cameras; (b) Thorlab optical rail; (c) Synchronisation board; (d,e) Suction pad and vibration damper.

2.2.3. Pothole detection

After surface fitting, the point cloud that belongs to the potholes can be selected by applying a threshold (κ) on the Euclidean distance between each point and the surface ($z_i - \widehat{z}_i > \kappa$). With the pothole 3D points determined, the relevant pixels in the disparity images can be segmented. With the connected component labelling algorithm applied to these segments, pothole areas can be labelled. Finally, the volume of the pothole can be calculated using the pothole point cloud.

3. EXPERIMENTAL RESULTS AND SETUP

The experiments are carried out using two PointGrey Flea 3 Cameras, as shown in Figure 3. Two cameras are mounted in parallel onto the optical rail, with a synchronisation board in the middle (Arduino board). This produces a 20Hz clock signal to synchronise the cameras. The lenses on the cameras are wide angle lenses with $2.8mm$ focal length and $F/1.2$ aperture. They have a $93.2° \times 70.7°$ field of view, with the forward point setup, and we selected a $3m \times 5m$ area in front of the vehicle as the region of interest (ROI). A threshold $\kappa = 0.04m$ is used for detecting the pothole point cloud.

The experimental results are shown in Figure 4. It is observed that the potholes can be detected correctly in most scenarios. In the figure, green areas indicate the ROI, and other highlighted colour areas represent potholes. In Figure 4(e), the blue area should not be counted as a pothole but it is still a road surface defect. As shown in Figure 4(g), a false detection is indicated in blue, which was caused by the errors in the disparity calculation. The high detection accuracy is achieved by defining the ROI in front of the vehicle, so that other obstacles and vehicles would not deteriorate the surface fitting results. The proposed system is not restricted by a fixed camera installation, and road curvatures can be well represented by the quadratic surface. A video result is also available from [17].

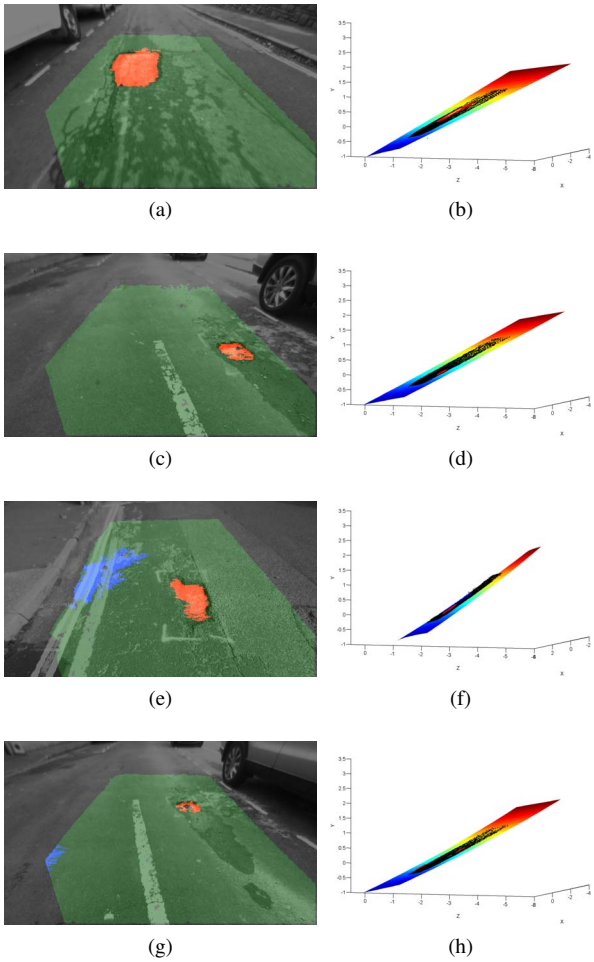

Fig. 4: Pothole detection results. The left column is the detection results, green areas highlight the ROI and the red areas indicate potholes. The right column is the corresponding surface fitting, black areas indicating the ROI and red areas representing detected potholes.

4. CONCLUSION

In this paper, a pothole detection system has been described. With a previously developed computationally efficient and robust disparity calculation algorithm and a quadratic fitting in the WCS point cloud, the road surface could be modelled in real-time. With the road surface fitted, a simple height threshold ($0.04m$ lower) can outline the potholes, utilizing only the geometrical deviation to that of the road. By setting the ROI in front of the vehicle, experimental results show high road fitting and pothole detection accuracies.

5. REFERENCES

[1] BBC News, "'millions of cars' damaged by potholes," http://www.bbc.co.uk/news/uk-21770969/, 2013, Accessed: 2013-11-01.

[2] SKY News, "Pothole damage costs councils 22m in 2011," http://news.sky.com/story/1040136/pothole-damage-costs-councils-22m-in-2011, 2013, Accessed: 2013-11-01.

[3] BBC News, "Potholes which damaged cars prompt council payouts," http://www.bbc.co.uk/news/uk-england-19367206/, 2013, Accessed: 2013-11-01.

[4] "Fixmystreet," http://www.fixmystreet.com/, Accessed: 2013-11-01.

[5] Qingguang Li, Ming Yao, Xun Yao, and Bugao Xu, "A real-time 3d scanning system for pavement distortion inspection," *Measurement Science and Technology*, vol. 21, no. 1, pp. 015702, 2010.

[6] A. Mednis, G. Strazdins, R. Zviedris, G. Kanonirs, and L. Selavo, "Real time pothole detection using android smartphones with accelerometers," in *Distributed Computing in Sensor Systems and Workshops (DCOSS), 2011 International Conference on*, 2011, pp. 1–6.

[7] Ezzatollah Salari, Eddie Chou, and James J Lynch, "Pavement distress evaluation using 3d depth information from stereo vision," Tech. Rep., 2012.

[8] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International journal of computer vision*, vol. 47, no. 1, pp. 7–42, 2002.

[9] Li Hong and G. Chen, "Segment-based stereo matching using graph cuts," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, June-2 July, vol. 1, pp. I–74–I–81 Vol.1.

[10] P.F. Felzenszwalb and D.P. Huttenlocher, "Efficient belief propagation for early vision," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, June-2 July, vol. 1, pp. I–261–I–268 Vol.1.

[11] Jae Chul Kim, Kyoung-Mu Lee, Byoung-Tae Choi, and Sang-Uk Lee, "A dense stereo matching using two-pass dynamic programming with generalized ground control points," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, June, vol. 2, pp. 1075–1082 vol. 2.

[12] N. Lazaros, G.C. Sirakoulis, and A. Gasteratos, "Review of Stereo Vision Algorithms: From Software to Hardware," *International Journal of Optomechatronics*, vol. 2, no. 4, pp. 435–462, 2008.

[13] Zhen Zhang, Xiao Ai, and Naim Dahnoun, "Efficient disparity calculation based on stereo vision with ground obstacle assumption," in *21st European Signal Processing Conference 2013 (EUSIPCO 2013)*, Marrakech, Morocco, Sept. 2013.

[14] GeneH. Golub and Urs Matt, "Quadratically constrained least squares and quadratic problems," *Numerische Mathematik*, vol. 59, no. 1, pp. 561–580, 1991.

[15] Xiao Ai, Yuan Gao, John Rarity, and Naim Dahnoun, "Obstacle detection using u-disparity on quadratic road surfaces.," in *Intelligent Transportation Systems, 2013. ITSC13. 16th International IEEE Conference on.*, 2013.

[16] F. Oniga, S. Nedevschi, M. M Meinecke, and Thanh-Binh To, "Road surface and obstacle detection based on elevation maps from dense stereo," in *Intelligent Transportation Systems Conference, 2007. ITSC 2007. IEEE*, 2007, pp. 859–865.

[17] Zhen Zhang, Xiao Ai, C.K. Chan, and Naim Dahnoun, "Pothole detection video result," http://rwnlabs.co.uk/zhen/, Accessed: 2013-11-01.