

Pothole Detection Based on Disparity Transformation and Road Surface Modeling

Rui Fan, *Member, IEEE*, Umar Ozgunalp, *Member, IEEE*, Brett Hosking, *Member, IEEE*,
Ming Liu, *Senior Member, IEEE*, Ioannis Pitas, *Fellow, IEEE*

Abstract—Pothole detection is one of the most important tasks for road condition assessment. Computer vision approaches are generally based on either 2D road image analysis or 3D road surface modeling. However, these two categories are always used independently. Furthermore, the pothole detection accuracy is still far from satisfactory. Therefore, in this paper, we present a robust pothole detection algorithm that is both accurate and computationally efficient. A dense disparity map is first transformed to better distinguish between the damaged and undamaged road areas. To achieve greater disparity transformation processing efficiency, golden section search and dynamic programming are utilized to estimate the transformation parameters. Otsu's thresholding method is then used to extract the potential undamaged road areas from the transformed disparity map. The disparities in the extracted areas are modeled by a quadratic surface using least squares fitting. To improve disparity map modeling robustness, the surface normal is also integrated into the modeling process. Furthermore, random sample consensus is utilized to reduce the effects caused by outliers. By comparing the difference between the actual and the modeled disparity maps, we consider that the potholes can be accurately detected. Finally, the point clouds of the detected potholes are extracted from the reconstructed 3D road surface. The experimental results show that the successful detection accuracy of the proposed system is around 98.7% and the overall pixel-level accuracy is approximately 99.6%.

Index Terms—pothole detection, road condition assessment, golden section search, dynamic programming, disparity map modeling.

I. INTRODUCTION

ROAD potholes are considerably large structural failures on the road surface, which are caused by contraction and expansion of the road surface and by rainwater, which permeates into the ground [1]. To ensure traffic safety, it is crucial and necessary to frequently inspect and repair road damage, e.g., cracks and potholes [2]. Currently, potholes are regularly detected and reported by certified inspectors and structural engineers [3]. This task is, however, tedious and time-consuming [4]. Furthermore, the detection results are always subjective, because they entirely depend on personnel experience [5]. Therefore, automated pothole detection sys-

tems have been developed to detect, i.e., to recognize and localize potholes both objectively and efficiently [2].

Over the past decade, various technologies, such as active or passive sensing and the global positioning system (GPS), have been utilized to acquire road data and aid personnel in localizing road damages [5]. For example, Buza et al. [6] mounted two laser scanners on a digital inspection vehicle (DIV) to collect 3D road surface data. Such data were then processed either semi or fully automatically for damaged road area detection. Such systems ensure personnel safety, but also reduce manual interventions to produce pothole detection results [6]. Furthermore, Bristol City Council [7] set up an online system for citizens to report the GPS locations of road damages, which alleviates a significant amount of workload for the qualified personnel. Moreover, by comparing the road data collected over different periods, the traffic flow can be evaluated and the future road condition can be predicted [2]. The remainder of this section presents the state of the art, motivation, contributions and outline of this paper.

A. State of the Art in Road Pothole Detection

1) *2D Image Analysis-Based Pothole Detection Algorithms*: There are typically three main steps used in 2D image analysis-based pothole detection algorithms: image segmentation, shape extraction and object recognition [3]. In the first step, a color or gray-scale road image is segmented using histogram-based thresholding methods, such as Otsu [6] or triangle [8] methods. Otsu's method minimizes the intra-class variance and performs better, in terms of separating the damaged and undamaged road areas [6]. The road image is first processed, e.g., using morphological filters [9], to reduce noise and enhance the target region outline [5], [10], before image segmentation. The extracted region is then modeled by an ellipse [8]. Finally, the image texture within the ellipse is compared with the undamaged road area texture. If the former is grainier and coarser than the latter, the elliptical region is considered to be a pothole [5].

However, color or gray-scale image segmentation is always severely affected by various factors, notably illumination conditions. Therefore, some authors proposed to perform segmentation on depth maps, which has shown to achieve better performance when separating the damaged and undamaged areas [11], [12]. Furthermore, the shapes of actual potholes are always irregular, making the geometric and texture assumptions occasionally unreliable. Moreover, the 3D spatial pothole structure cannot always be explicitly illustrated in 2D road

R. Fan is with the Robotics Institute, the Hong Kong University of Science and Technology, Hong Kong (e-mail: rui.fan@ieee.org).

U. Ozgunalp is with the Electrical and Electronic Engineering, Near East University, Nicosia, Cyprus (e-mail: umar.ozgunalp@neu.edu.tr).

B. Hosking is with the National Oceanography Center, Southampton, U.K. (e-mail: wilski@noc.ac.uk).

M. Liu is with the Robotics Institute, the Hong Kong University of Science and Technology, Hong Kong (e-mail: eelium@ust.hk).

I. Pitas is with the School of Informatics, Aristotle University of Thessaloniki, Thessaloniki, Greece (e-mail: pitas@aiia.csd.auth.gr).

images [4]. Therefore, 3D road surface information is required to measure pothole volumes. In general, 3D road surface modeling-based pothole detection algorithms are more than capable of overcoming the disadvantages mentioned above.

2) *3D Road Surface Modeling-Based Pothole Detection Algorithms*: The 3D surface information used for pothole detection is generally provided by laser scanners [11], Microsoft Kinect sensors [12], or passive sensors [13]–[16]. Laser scanners mounted on DIVs are typically used for road condition assessment and accurate road surface reconstruction. However, the laser scanning equipment purchase and long-term maintenance are still very expensive [3]. The Microsoft Kinect sensors were initially designed for indoor use. Therefore, they greatly suffer from infra-red saturation in direct sunlight [17]. For this reason, passive sensors, such as single movable camera or multiple synchronized cameras, are more suitable for acquiring 3D road data and for pothole detection [4], [18]. For example, Zhang and Elaksher [13] mounted a single camera on an unmanned aerial vehicle (UAV) to reconstruct the road surface via Structure from Motion (SfM) [18]. A variety of stereo vision-based pothole detection methods have been developed as well [14]–[16]. The 3D point cloud generated from a disparity map was interpolated into a quadratic surface using least squares fitting (LSF) [15]. The potholes were then recognized, by comparing the difference between the 3D point cloud and the fitted quadratic surface. In [16], the surface modeling was performed on disparity maps instead of the point clouds, and random sample consensus (RANSAC) was used to improve the pothole detection. An efficient stereo vision-based road surface 3D reconstruction algorithm was proposed in [4] to provide highly accurate disparity information. Therefore, the road surface 3D reconstruction results obtained using [4] were used as the input of our system presented in this paper.

B. Motivation

Currently, laser scanning is still the main technology that is used to provide 3D road information for pothole detection. Other technologies, such as passive sensing, are under-utilized [2]. However, the DIVs are expensive and not widely used. Furthermore, their routine operation and long-term maintenance are still very expensive [8]. Therefore, the trend is to equip DIVs with inexpensive, portable and durable sensors, such as digital cameras, for road data acquisition. Stereo road image pairs can be used to calculate the disparity maps [4], which essentially represent the 3D road surface geometry. Recently, due to some major advances in stereo computer vision, road surface geometry can be reconstructed with a three-millimeter accuracy [4], [19]. Additionally, stereo cameras used for road condition assessment are inexpensive, portable and adaptable for different DIV types. Therefore, there is a strong motivation to explore the use of stereo vision for pothole detection.

So far, comprehensive studies have been made in both 2D image analysis-based and 3D road surface modeling-based pothole detection. Unfortunately, these algorithms are always used independently. Furthermore, pothole detection accuracy is still far from satisfactory [5]. Therefore, there is a strong

motivation to explore efficient approaches for disparity pre-processing by applying 2D image processing algorithms. Only the disparities in the potential undamaged road areas are then used for disparity map modeling.

Moreover, the surface normal vector is a very important descriptor, which is rarely utilized in existing 3D road surface modeling-based pothole detection algorithms. In this paper, we improve disparity map modeling by eliminating the disparities, in regions where surface normals differ significantly from the expected ones.

C. Novel Contributions

In this paper, a robust stereo vision-based pothole detection system is introduced. The main contributions are: a) a novel disparity transformation algorithm; b) a robust disparity map modeling algorithm. Finally, in order to assess pothole detection accuracy, we created three stereo pothole datasets using a ZED stereo camera and made them publicly available for research purposes.

Since the disparities in damaged road areas can severely affect the accuracy of disparity modeling, we first transform the disparity maps to better distinguish between the damaged and undamaged road areas. To achieve greater processing efficiency, we use golden section search (GSS) [20] and dynamic programming (DP) [21] to estimate the transformation parameters. Otsu's thresholding method is then performed on the transformed disparity map to extract the undamaged road areas, where the disparities can be modeled by a quadratic surface using LSF [22]. To improve the robustness of disparity map modeling, the surface normal information is also integrated into the modeling process. Furthermore, RANSAC is utilized to reduce the effects caused by any potential outliers. By comparing the difference between the actual and modeled disparity maps, the potholes can be detected effectively. Finally, different potholes are labeled using connected component labeling (CCL) and their point clouds are extracted from the reconstructed 3D road surface.

D. Paper Outline

The remainder of this paper is structured as follows: Section II details the proposed pothole detection system. The experimental results for performance evaluation are illustrated in Section III. Finally, Section IV summarizes the paper and provides some recommendations for future work.

II. POTHOLE DETECTION ALGORITHM

The block diagram of the proposed pothole detection algorithm is illustrated in Fig. 1, where the algorithm consists of three main components: a) disparity transformation; b) undamaged road area extraction; c) disparity map modeling and pothole detection.

A. Disparity Transformation

The input of this procedure is a dense disparity map having subpixel accuracy. Since the performance of disparity map modeling relies entirely on the disparity estimation accuracy,

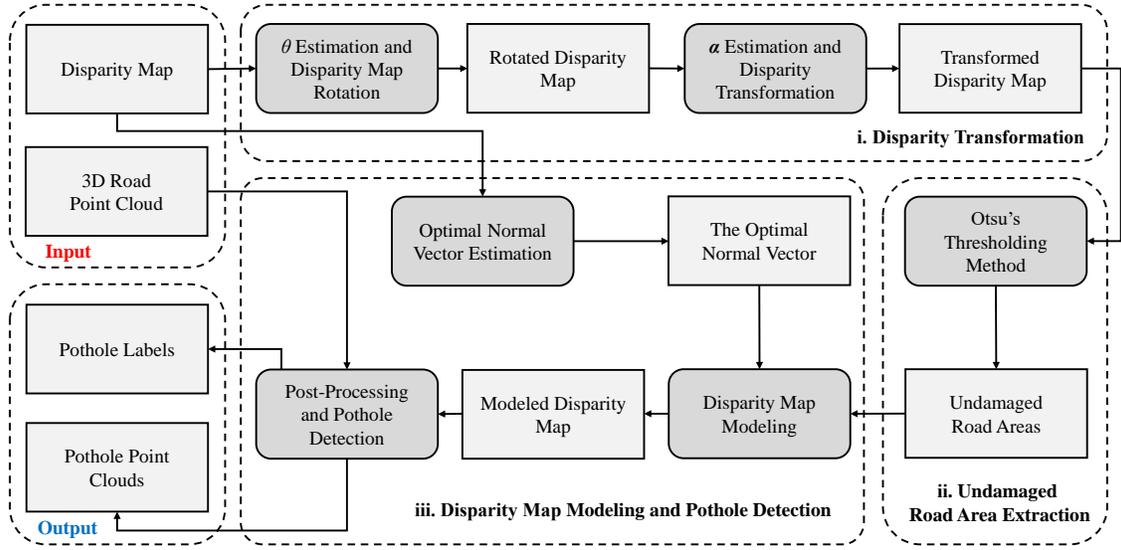


Fig. 1. The block diagram of the proposed pothole detection algorithm.

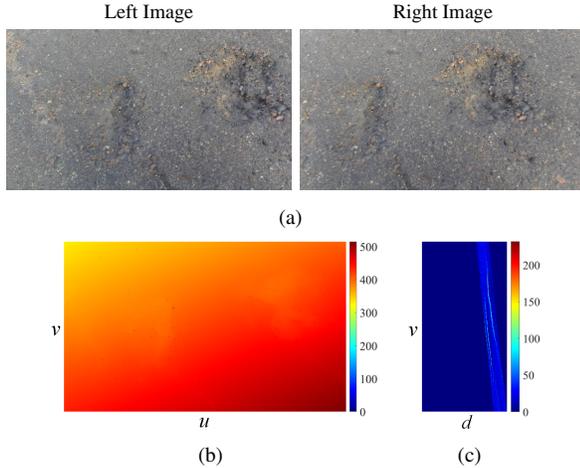


Fig. 2. Disparity map when the roll angle does not equal to zero; (a) stereo road images; (b) disparity map; (c) v-disparity map.

the dense disparity map was obtained from a stereo road image pair (see Fig. 2(a)) through our disparity estimation algorithm [4], where the stereo matching search range propagates iteratively from the bottom of the image to its top, and the subpixel disparity map is refined by iteratively minimizing an energy with respect to the interpolated correlation parabolas. The 3D reconstruction accuracy of [4] is around 3 mm. The disparity map is shown in Fig. 2(b), and its corresponding v-disparity map is shown in Fig. 2(c). A v-disparity map can be created by computing the histograms of each horizontal row of the disparity map. The proposed pothole detection algorithm is based on the work presented in [16], where the disparities of undamaged road surface are modeled by a quadratic surface model as follows:

$$g(u, v) = c_0 + c_1u + c_2v + c_3u^2 + c_4v^2 + c_5uv, \quad (1)$$

where u and v are the horizontal and vertical disparity map coordinates, respectively. The origin of the coordinate system

in (1) is at the center of the disparity map. Since in our experiments the stereo rig is mounted to a relatively low height, the curvature of the reconstructed road surface is not very high. This makes the values of c_1 , c_3 and c_5 in (1) very close to zero when the stereo rig is perfectly parallel to the horizontal road surface. In this case, the projection of the road disparities on the v-disparity map can be assumed to be a parabola of the form [21]:

$$f(v) = \alpha_0 + \alpha_1v + \alpha_2v^2. \quad (2)$$

However, in practice, the stereo rig baseline is not always perfectly parallel to the horizontal road surface. This fact can introduce a non-zero roll angle θ , illustrated in Fig. 3, into the imaging process, where T_c and h represent the baseline and the height of the stereo rig, respectively. o_l^c and o_r^c are the origins of the left and right camera coordinate systems, respectively. O^w is the origin of the world coordinate system. An example of the resulting disparity map is shown in Fig. 2, where readers can clearly see that the disparity values change gradually in the horizontal direction, which makes the approach of representing the disparity projection using (2) somewhat problematic. In this regard, we first estimate the value of the roll angle. The effects caused by the non-zero roll angle are then eliminated by rotating the disparity map by θ . Finally, the coefficients of the disparity projection model in (2) are estimated, and the disparity map is transformed to better distinguish between the damaged and undamaged road areas.

1) *θ Estimation and Disparity Map Rotation*: Over the past decade, great effort has been made to improve the estimation of the roll angle defined in Fig. 3. The most commonly used device for this task is the inertial measurement unit (IMU). An IMU can measure the angular rate of a vehicle by analyzing the data acquired using different sensors, such as accelerometers, gyroscopes and magnetometers [23], [24]. In these approaches, the road bank angle is always assumed to be zero, and only the roll angle is considered in the estimation process. However, the

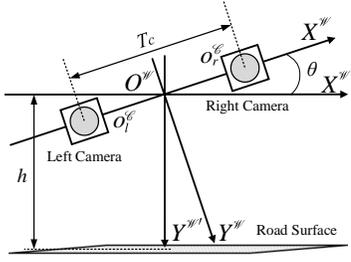


Fig. 3. Roll angle definition in a stereo camera rig.

estimation of both these two angles is always required in many real-world applications. Unfortunately, this cannot be realized with the use of only IMUs [25], [26].

In recent years, many authors have turned their focus towards estimating the roll angle from disparity maps [21], [27]–[29]. For example, the road surface is assumed to be a horizontal ground plane, and an effective roll angle estimation algorithm was proposed based on v -disparity map analysis [27], [28]. In [21], the disparities in a selected small area were modeled by a plane $g(u, v) = c_0 + c_1u + c_2v$. The roll angle was then calculated as $\arctan(-c_1/c_2)$. However, finding a proper disparity map area for plane fitting is always challenging, because the area may contain an obstacle or a pothole, which can severely affect the fitting accuracy [29]. Furthermore, the above-mentioned algorithms are only suitable for planar road surfaces. Hence, in this subsection, we introduce a roll angle estimation algorithm, which can work effectively for both planar and non-planar road surfaces.

When the roll angle is equal to zero, the vector $\alpha = [\alpha_0, \alpha_1, \alpha_2]^T$ containing the disparity projection model coefficients can be estimated by solving a least squares problem as follows:

$$\alpha = \arg \min_{\alpha} E_0, \quad (3)$$

where

$$E_0 = \mathbf{e}^T \mathbf{e}, \quad (4)$$

and

$$\mathbf{e} = \mathbf{d} - \mathbf{V}\alpha. \quad (5)$$

The column vector $\mathbf{d} = [d_0, d_1, \dots, d_n]^T$ stores the disparity values. \mathbf{V} is an $(n+1) \times 3$ matrix shown as follows:

$$\mathbf{V} = \begin{bmatrix} 1 & v_0 & v_0^2 \\ 1 & v_1 & v_1^2 \\ \vdots & \vdots & \vdots \\ 1 & v_n & v_n^2 \end{bmatrix}. \quad (6)$$

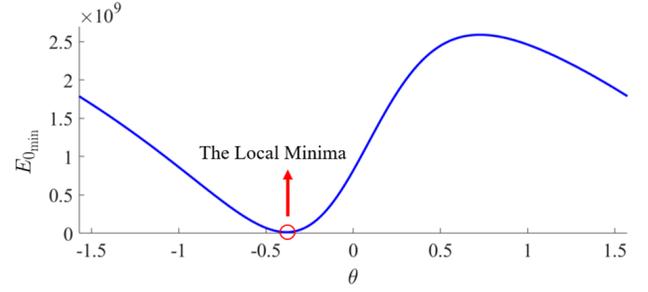
This optimization problem has a closed form solution [30]:

$$\alpha = (\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \mathbf{d}. \quad (7)$$

The minimum energy $E_{0_{\min}}$ can also be obtained by combining (4), (5) and (7):

$$E_{0_{\min}} = \mathbf{d}^T \mathbf{d} - \mathbf{d}^T \mathbf{V} (\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \mathbf{d}. \quad (8)$$

However, when the roll angle does not equal to zero, the disparity distribution on each row becomes less compact (see

Fig. 4. $E_{0_{\min}}$ function versus θ .

Algorithm 1: θ estimation using GSS.

Data: disparity map

Result: θ

- 1 set θ_1 and θ_2 to $-\pi/2$ and $\pi/2$, respectively;
 - 2 compute $E_{0_{\min}}(\theta_1)$ using Eq. 13;
 - 3 compute $E_{0_{\min}}(\theta_2)$ using Eq. 13;
 - 4 **while** $\theta_2 - \theta_1 > \varepsilon_{\theta}$ **do**
 - 5 set θ_3 and θ_4 to $\kappa\theta_1 + (1 - \kappa)\theta_2$ and $\kappa\theta_2 + (1 - \kappa)\theta_1$, respectively;
 - 6 compute $E_{0_{\min}}(\theta_3)$ using Eq. 13;
 - 7 compute $E_{0_{\min}}(\theta_4)$ using Eq. 13;
 - 8 **if** $E_{0_{\min}}(\theta_3) > E_{0_{\min}}(\theta_4)$ **then**
 - 9 θ_1 is replaced by θ_3 ;
 - 10 **else**
 - 11 θ_2 is replaced by θ_4 ;
 - 12 **end**
 - 13 **end**
-

Fig. 2(c)). This greatly affects the accuracy of least squares fitting and produces a much higher $E_{0_{\min}}$.

To rotate the disparity map around a given angle θ , each set of original coordinates $[u, v]^T$ is transformed to a set of new coordinates $[x, y]^T$ as follows [31]:

$$\begin{aligned} x &= u \cos \theta + v \sin \theta, \\ y &= v \cos \theta - u \sin \theta. \end{aligned} \quad (9)$$

(5) can now be rewritten as follows:

$$\mathbf{e}(\theta) = \mathbf{d} - \mathbf{Y}(\theta)\alpha(\theta), \quad (10)$$

where $\mathbf{Y}(\theta)$ is the $(n+1) \times 3$ matrix:

$$\mathbf{Y}(\theta) = \begin{bmatrix} 1 & y_0(\theta) & y_0(\theta)^2 \\ 1 & y_1(\theta) & y_1(\theta)^2 \\ \vdots & \vdots & \vdots \\ 1 & y_n(\theta) & y_n(\theta)^2 \end{bmatrix}. \quad (11)$$

(7) is rewritten as follows:

$$\alpha(\theta) = (\mathbf{Y}(\theta)^T \mathbf{Y}(\theta))^{-1} \mathbf{Y}(\theta)^T \mathbf{d}. \quad (12)$$

(4), (10) and (12) result in the following expression [32]:

$$E_{0_{\min}}(\theta) = \mathbf{d}^T \mathbf{d} - \mathbf{d}^T \mathbf{Y}(\theta) (\mathbf{Y}(\theta)^T \mathbf{Y}(\theta))^{-1} \mathbf{Y}(\theta)^T \mathbf{d}. \quad (13)$$

Therefore, the main consideration of the proposed roll angle estimation algorithm is to rotate the disparity map at different

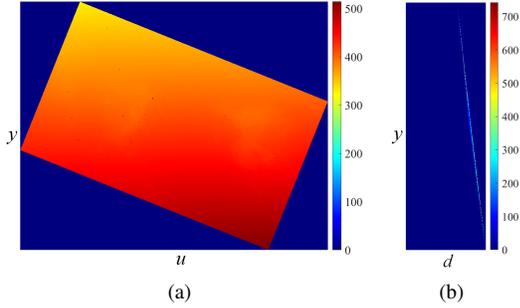


Fig. 5. Elimination of the effects caused by the non-zero roll angle; (a) disparity map; (b) the y-disparity map of Fig. 5(a).

angles, and find the angle proving the minimum $E_{0_{\min}}$. $E_{0_{\min}}$ with respect to different θ is illustrated in Fig. 4. Giving a set of coordinates $[u, v]^T$, the new coordinate y can be calculated using (9). The corresponding $E_{0_{\min}}$ can be computed from (13). Due to the fact that $\cos(\theta+\pi) = -\cos \theta$ and $\sin(\theta+\pi) = -\sin \theta$, the disparity maps rotated around θ and $\theta + \pi$ are symmetric with respect to the origin of the coordinate system. Namely, (13) outputs the same $E_{0_{\min}}$ no matter how the disparity map is rotated around θ or $\theta + \pi$. Therefore, we set the interval of θ to $(-\pi/2, \pi/2]$. The estimation of θ is achieved by finding the position of the local minima between $-\pi/2$ and $\pi/2$.

However, finding the local minima is a computationally intensive task, because it involves performing the necessary calculations through the whole interval of θ . Furthermore, the step size ε_θ has to be set to a very small and practical value, in order to obtain an accurate value of θ . Hence in this paper, GSS is utilized to reduce the searching times. The procedure of the proposed θ estimation algorithm is given in Algorithm 1, where $\kappa = (\sqrt{5} + 1)/2$ is the golden section ratio [20].

The disparity map is then rotated around θ , as illustrated in Fig. 5(a). A y-disparity map (see Fig. 5(b)) can be created by computing the histograms of each horizontal row of the rotated disparity map. We can observe that the disparity values on each row become more uniform. The evaluation of the proposed roll angle estimation algorithm will be discussed in Section III-B.

2) α Estimation and Disparity Transformation: In this subsection, we utilize DP [21] to extract the road disparity projections from the y-disparity map. For notational convenience, the path of the projections is also referred to as target path. The energy of every possible solution is first computed as follows:

$$E_1(d, y) = -m(d, y) + \min_{\tau} [E_1(d+1, y-\tau) - \lambda\tau] \text{ s.t. } \tau \in [\tau_{\min}, \tau_{\max}], \quad (14)$$

where $m(d, y)$ represents the y-disparity value at the position $[d, y]^T$ and λ is a smoothness term [21]. E_1 represents the energy of a possible target path in the y-disparity map. τ_{\min} is typically set to 0. τ_{\max} depends entirely on α , and it is set to 8 in this paper. The target path $M = \{(d_i, y_i), i = 0, 1, \dots, n\}$ can be found by minimizing the energy function in (14), where (d_i, y_i) stores the horizontal and vertical coordinates of the target path, respectively.

By substituting the horizontal and vertical coordinates of the target path into (3), (4), (10) and (11), we can obtain the

Algorithm 2: α estimation using DP.

Input : The target path

Output: α

```

1 create a  $t \times (s + 3)$  matrix  $T$ ;
2 for  $i \leftarrow 1$  to  $t$  do
3   randomly select  $p$  pairs of coordinates from the
   target path;
4   estimate  $\alpha$  using Eq. 3;
5    $T(i, s + 1 : s + 3) \leftarrow \alpha^T$ ;
6   for  $j \leftarrow 1$  to  $s$  do
7     set the tolerance to  $\varepsilon_\alpha/2^{j-1}$ ;
8     compute the number of inliers  $n_{\text{inlier}}$ ;
9     compute the number of outliers  $n_{\text{outlier}}$ ;
10    compute the ratio  $\eta = n_{\text{inlier}}/n_{\text{outlier}}$ ;
11     $T(i, j) \leftarrow \eta$ ;
12  end
13 end
14  $j = 0$ ;
15 do
16    $j \leftarrow j + 1$ ;
17   find the highest  $\eta$  in the  $j$ th column of  $T$ ;
18 while the highest  $\eta$  in the  $j$ th column of  $T$ 
   corresponds to more than one  $\alpha$ ;
19 find  $\alpha$  which corresponds to the highest  $\eta$  in the  $j$ th
   column of  $T$ ;
```

coefficients of the disparity projection model. The disparity map can therefore be transformed using θ and α . However, the outliers in the target path may greatly affect the accuracy of α estimation. We, therefore, use RANSAC to update the values in α . The full list of procedures involved for α estimation are detailed in Algorithm 2.

RANSAC is iterated t times. The selection of a higher t raises the possibility of finding the best α , but on the other hand, increases the processing time. In order to minimize the trade-off between speed and robustness, t is set to 50 in this paper. In each iteration, we select p pairs of coordinates $[d, y]^T$ from the target path to estimate α . For a smaller p , there is less chance that any outliers will influence optimization. In this paper p is set to 3, which is the bare minimal value to determine the values of α . The ratio η of inliers versus outliers can then be computed with respect to a given tolerance ε_α , where ε_α determines the inliers and outliers. The best α corresponds to the highest ratio η . However, the selection of an appropriate ε_α can be challenging, as it is possible that there could be more than one satisfying value for α . Hence, in this paper, the value of ε_α also changes s times. In each iteration of Algorithm 2, the value of ε_α reduces by half and η is computed. The best α can be determined by finding the highest ratio η . In this paper, the values of ε_α and s are both set to 4.

Finally, each disparity is transformed using the following equation:

$$\tilde{d} = [1 \ y(\theta) \ y(\theta)^2]\alpha(\theta) - d + \delta, \quad (15)$$

where δ is a constant value set to guarantee that all the transformed disparity values are non-negative. In this paper,

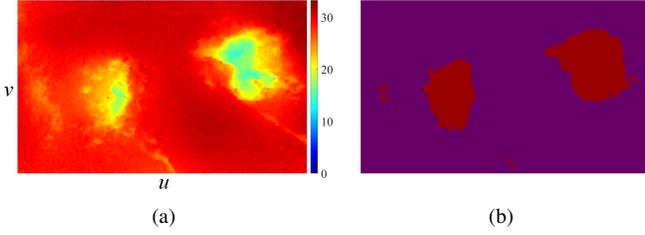


Fig. 6. Disparity transformation and undamaged road area extraction; (a) transformed disparity map; (b) extracted undamaged road area.

we set δ to 30. The transformed disparity map is shown in Fig. 6(a). It can be clearly seen that the disparity values in the undamaged road areas become more uniform, while they differ significantly from those in the damaged areas (potholes). This makes the extraction of undamaged road regions much simpler.

B. Undamaged Road Area Extraction

Next, we utilize Otsu's thresholding method to segment the transformed disparity map. The segmentation threshold T_o can be obtained by maximizing the inter-class variance σ_o^2 as follows [33]:

$$\sigma_o^2(T_o) = \omega_0(T_o)\omega_1(T_o)[\mu_0(T_o) - \mu_1(T_o)]^2, \quad (16)$$

where:

$$P_0(T_o) = \sum_{i=\tilde{d}_{\min}}^{T_o-1} p(i), \quad P_1(T_o) = \sum_{i=T_o}^{\tilde{d}_{\max}} p(i) \quad (17)$$

represent the probabilities of damaged and undamaged road areas, respectively. $p(i)$ is the probability of $\tilde{d} = i$. The average disparity values of the damaged and undamaged road areas are given by:

$$\begin{aligned} \mu_0(T_o) &= \frac{1}{P_0(T_o)} \sum_{i=\tilde{d}_{\min}}^{T_o-1} ip(i), \\ \mu_1(T_o) &= \frac{1}{P_1(T_o)} \sum_{i=T_o}^{\tilde{d}_{\max}} ip(i) \end{aligned} \quad (18)$$

The segmentation result is shown in Fig. 6(b). We can see that the undamaged road area is successfully extracted. However, Otsu's thresholding method will always classify the disparities into two categories, even if the transformed disparity map does not contain any road damage. We therefore carry out disparity map modeling to ensure that the potholes are correctly detected.

C. Disparity Map modeling and Pothole Detection

A common practice in 3D modeling-based pothole detection algorithms [15], [16] is to fit a quadratic surface to either a 3D point cloud or a 2D disparity map. In parallel axis stereo vision, the point cloud is generated from the disparity map as follows [34]:

$$X^{\mathcal{W}} = \frac{uT_c}{d}, \quad Y^{\mathcal{W}} = \frac{vT_c}{d}, \quad Z^{\mathcal{W}} = \frac{fT_c}{d}, \quad (19)$$

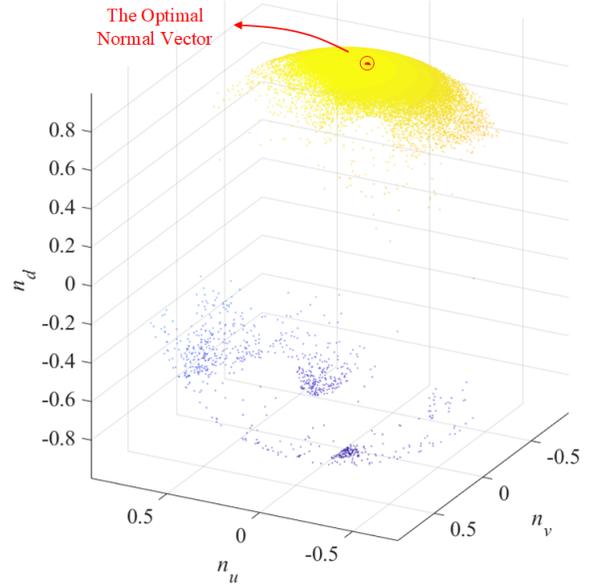


Fig. 7. Normal surface vector mapping on a sphere.

where f and T_c are the focal length in pixels and the stereo rig baseline shown in Fig. 3, respectively. A disparity error larger than one pixel may result in a non-negligible difference in the point cloud [35]. Therefore, disparity map modeling can avoid such errors generated from (19), producing greater accuracy compared to point cloud modeling.

To model the disparity map, $\mathbf{c} = [c_0, c_1, c_2, c_3, c_4, c_5]^T$, the quadratic surface model coefficients can be estimated as follows:

$$\mathbf{c} = (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{d}, \quad (20)$$

where

$$\mathbf{W} = \begin{bmatrix} 1 & v_0 & u_0 & v_0^2 & u_0^2 & v_0 u_0 \\ 1 & v_1 & u_1 & v_1^2 & u_1^2 & v_1 u_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & v_n & u_n & v_n^2 & u_n^2 & v_n u_n \end{bmatrix}. \quad (21)$$

However, potential outliers can severely affect the accuracy of disparity map modeling and therefore need to be discarded beforehand. In this subsection, a disparity map point is determined as an outlier if it fulfills one of the following conditions:

- it is located in one of the damaged road areas.
- its surface normal vector differs greatly from the optimal one.
- its disparity value differs greatly from the one computed using Eq. 1.

In Section II-B, the undamaged road areas are successfully extracted and we only use the disparities in this area to model the disparity map. The rest of this subsection presents the approaches for determining the outliers which satisfy the last two conditions and the process of modeling the disparity map without using these outliers.

1) *Optimal Normal Vector Estimation*: For each point $\mathbf{p}_i = [u_i, v_i, d_i]^T$ in the undamaged road area, we would like to estimate a normal vector $\mathbf{n}_i = [n_{iu}, n_{iv}, n_{id}]^T$ from a set of k

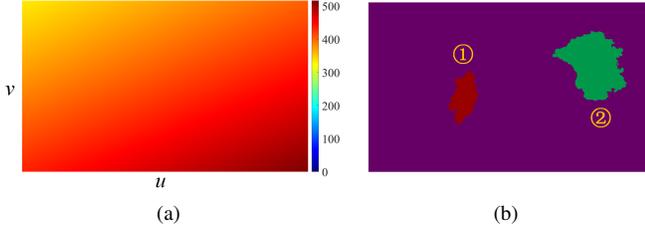


Fig. 8. Disparity map modeling and pothole detection: (a) modeled disparity map, (b) detected potholes.

points in its neighborhood $\mathbf{Q}_i = [q_{i1}, q_{i2}, \dots, q_{ik}]^T$, where $q_{ij} \neq p_i$. Here, we define the augmented neighbour matrix \mathbf{Q}_i^+ which contains all neighbors and the point p_i itself as follows:

$$\mathbf{Q}_i^+ = [p_i, \mathbf{Q}_i^T]^T. \quad (22)$$

Existing normal vector estimation methods are generally classified into one of two categories: optimization-based and averaging-based [36]. Although the performance of normal vector estimation depends primarily on the application itself, *PlanePCA* [37], an optimization-based normal vector estimation method, has superior performance in terms of both speed and accuracy. Hence in this subsection, we utilize *PlanePCA* to estimate the normal vectors of the disparities. \mathbf{n}_i can be estimated as follows:

$$\mathbf{n}_i = \arg \min_{\mathbf{n}_i} \left\| \left[\mathbf{Q}_i^+ - \bar{\mathbf{Q}}_i^+ \right] \mathbf{n}_i \right\|_2, \quad (23)$$

where

$$\bar{\mathbf{Q}}_i^+ = \mathbf{1}_{k+1} \left(\frac{1}{k+1} \mathbf{Q}_i^{+\top} \mathbf{1}_{k+1} \right)^T. \quad (24)$$

$\mathbf{1}_m$ represents an $m \times 1$ vector of ones. Due to the fact that the normal vectors are normalized, they can be projected on a sphere as shown in Fig. 7, where we can clearly see that the projections are distributed in a small area. Therefore, the optimal normal vector $\hat{\mathbf{n}}$ can be determined by finding the position at which the projections distribute most intensively.

Since the projection of $\hat{\mathbf{n}}$ is also on the sphere, it can be written in spherical coordinates as follows:

$$\hat{\mathbf{n}} = [\sin \varphi_1 \cos \varphi_2, \sin \varphi_1 \sin \varphi_2, \cos \varphi_1]^T \quad (25)$$

s.t. $\varphi_1 \in [0, \pi], \varphi_2 \in [0, 2\pi)$.

It can be estimated by minimizing E_2 :

$$E_2 = \mathbf{1}_{n+1}^T \mathbf{m}, \quad (26)$$

where

$$\mathbf{m} = [-\mathbf{n}_0 \cdot \hat{\mathbf{n}}, -\mathbf{n}_1 \cdot \hat{\mathbf{n}}, \dots, -\mathbf{n}_n \cdot \hat{\mathbf{n}}]^T. \quad (27)$$

By applying (25) and (27) to (26), the following expressions are derived:

$$\tan \varphi_1 = \frac{\mathbf{1}_{n+1}^T \mathbf{n}_u \cos \varphi_2 + \mathbf{1}_{n+1}^T \mathbf{n}_v \sin \varphi_2}{\mathbf{1}_{n+1}^T \mathbf{n}_d}, \quad (28)$$

$$\tan \varphi_2 = \frac{\mathbf{1}_{n+1}^T \mathbf{n}_v}{\mathbf{1}_{n+1}^T \mathbf{n}_u}. \quad (29)$$

Due to the fact that φ_2 is between 0 and 2π , (29) will have two solutions:

$$\varphi_2 = \arctan \frac{\mathbf{1}_{n+1}^T \mathbf{n}_v}{\mathbf{1}_{n+1}^T \mathbf{n}_u} + k\pi, \quad k \in \{0, 1\}. \quad (30)$$

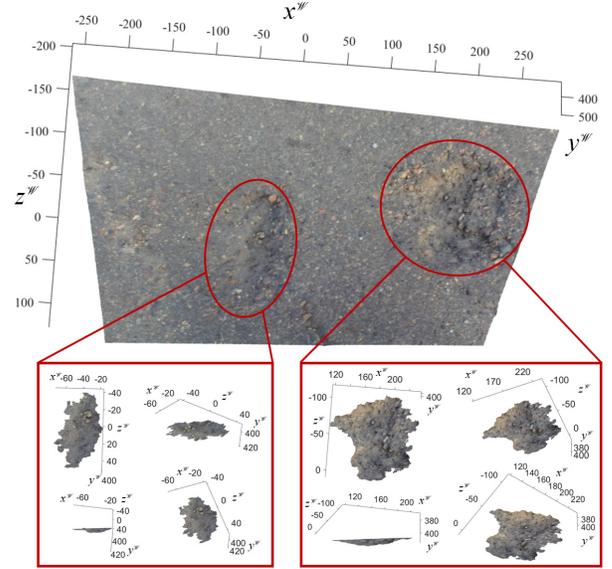


Fig. 9. The point clouds of the detected potholes.

Substituting each φ_2 into (28) produces a value for φ_1 . The two pairs of $[\varphi_1, \varphi_2]^T$ correspond to the maxima and minima of E_2 , respectively. By substituting each pair of $[\varphi_1, \varphi_2]^T$ into Eq. 26 and comparing the two obtained values, we can find the optimal normal vector. If the angle between $\hat{\mathbf{n}}$ and \mathbf{n}_i exceeds a pre-set threshold ε_n , the corresponding disparity will be considered as an outlier and will not be used for disparity map modeling. In our experiments, we assumed that the second category of outliers account for 10% of the undamaged road areas, and therefore, ε_n is set to $\pi/36$. The outliers satisfying the first two conditions can then be successfully removed. The third category of outliers are removed along with the disparity map modeling.

2) *Disparity Map Modeling*: To model the disparity map with more robustness, we use RANSAC to reduce the effects caused by the third category of outliers described in Section II-C. Here, RANSAC iterates t times. In each iteration, a subset of disparities are selected randomly to estimate the values in \mathbf{c} . To ensure uniform distribution of the selected disparities, we equally divide the disparity map into a group of square blocks and select one disparity from each block. The disparity block size is $r \times r$. As r becomes smaller, more disparities will be used for surface fitting, which makes computational complexity more intensive. In contrast, the selection of a higher value for r results in less computational complexity, but potentially increases noise sensitivity. In this paper, the value of r is set to 125, which produces approximately 100 square blocks for our disparity maps. In each iteration, the differences between the actual and fitted disparities are computed and the ratio η between the inliers and outliers are obtained. The values \mathbf{c} which correspond to the highest η are then selected as the desirable surface coefficients. Algorithm 2 presents more details on the least squares fitting using RANSAC. The modeled disparity map is shown in Fig. 8(a).

3) *Pothole Detection*: The potholes can then be detected by finding the regions where the differences between the actual



Fig. 10. Experimental set-up for acquiring stereo road images.

and modeled disparities are larger than a pre-set threshold ε_d . Before labeling different potholes using CCL, the connected components containing fewer than w pixels are removed, because they severely affect the accuracy of the pothole labeling process. Furthermore, the small connected components are filled, as they are considered to be noise. The selection of ε_d and w is discussed in Section III-D. The detected potholes are shown in Fig. 8(b). Finally, the point clouds of the detected potholes are extracted from the 3D point cloud. The corresponding results are shown in Fig. 9.

III. EXPERIMENTAL RESULTS

In this section, the performance of the proposed pothole detection algorithm is evaluated both qualitatively and quantitatively. The proposed algorithm was implemented in MATLAB on an Intel Core i7-8700K CPU using a single thread. The following subsections provide details on the experimental set-up and of the evaluation of the proposed algorithm.

A. Experimental Set-Up

In this work, we utilized a ZED stereo camera¹ to capture visual road data. An example of the experimental set-up is shown in Fig. 10. The stereo camera is calibrated manually using the stereo calibration toolbox in MATLAB R2018b. Using the above-mentioned experimental set-up, we created three datasets containing 67 pairs of stereo images. The image resolutions of dataset 1, 2 and 3 are 1028×1730 , 1030×1720 , 1028×1710 pixels, respectively. The disparity maps are estimated using our previously published algorithm [4]. All datasets are publicly available and can be found at: <http://www.ruirangerfan.com>.

The following subsections analyze the accuracy of roll angle estimation, disparity transformation and pothole detection.

B. Evaluation of Roll Angle Estimation

In this subsection, we first analyze the computational complexity of the proposed roll angle estimation algorithm. When estimating the roll angle without using GSS, we have to search through the whole interval of $(-\frac{\pi}{2}, \frac{\pi}{2}]$ to find the local minima. Therefore, the computational complexity is $\mathcal{O}(\frac{\pi}{\varepsilon_\theta})$. In our method, GSS reduces the interval size exponentially. As a result, the interval size then becomes $\kappa^n \pi$ after the n -th iteration. Therefore, the proposed roll angle estimation

¹<https://www.stereolabs.com/>

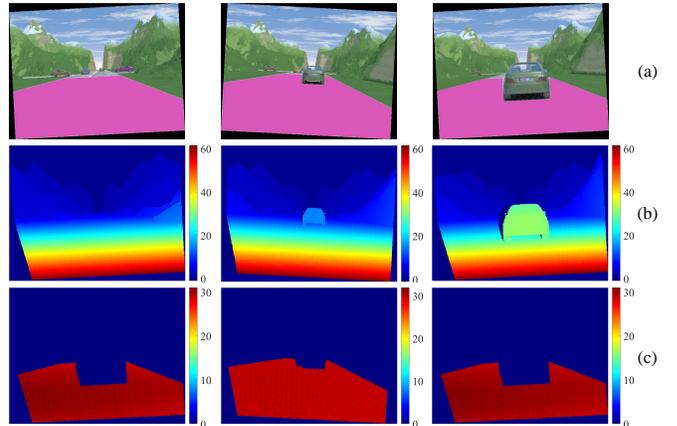


Fig. 11. Experimental results of the EISATS synthesized stereo dataset: (a) left stereo images (the areas in magenta are our manually selected road areas); (b) ground truth disparity maps; (c) transformed disparity maps.

algorithm reduces the computational complexity to $\mathcal{O}(\log_{\kappa} \frac{\varepsilon_\theta}{\pi})$. The proposed roll angle estimation algorithm needs 21 iterations to produce a roll angle, whose accuracy is higher than $\frac{\pi}{18000}$.

To evaluate the accuracy of the proposed roll angle estimation algorithm, we utilize a synthesized stereo dataset from EISATS [38], [39] where the roll angle is perfectly zero. Some experimental results are shown in Fig. 11. The road areas (see the magenta regions in the first column of Fig. 11) are manually selected and the disparities in these areas are utilized to estimate the roll angle $\hat{\theta}$. The absolute difference between the actual and estimated roll angles, i.e., $\Delta\theta = |\theta - \hat{\theta}|$, is computed for each frame. The average values of $\Delta\theta$ is approximately 1.129×10^{-4} rad which is lower than $\frac{\pi}{18000}$ rad. Therefore, the proposed algorithm is capable of estimating the roll angle with high accuracy.

C. Evaluation of Disparity Transformation

Since the datasets we created only contain the ground truth of potholes, KITTI stereo datasets [40], [41] are utilized to quantify the performance of our proposed disparity transformation algorithm (the number of disparity maps in the KITTI stereo 2012 and 2015 datasets are 194 and 200, respectively). Some experimental results are shown in Fig. 12. Due to the fact that the proposed algorithm focuses entirely on the road surface, we manually selected a region of interest (see the purple areas in the first row) in each image to evaluate the performance of our algorithm. Our manually labeled road surface areas are also available at <http://www.ruirangerfan.com>. The corresponding transformed disparity maps are shown in the third row of Fig. 12, where readers can clearly see that the disparities in the road areas tend to have similar values. To quantify the accuracy of the transformed disparity maps, we compute the standard deviation σ_d of the transformed disparity values as follows:

$$\sigma_d = \sqrt{\frac{1}{m+1} \left\| \tilde{\mathbf{d}} - \frac{\tilde{\mathbf{d}}^\top \mathbf{1}_{m+1}}{m+1} \right\|_2^2}. \quad (31)$$

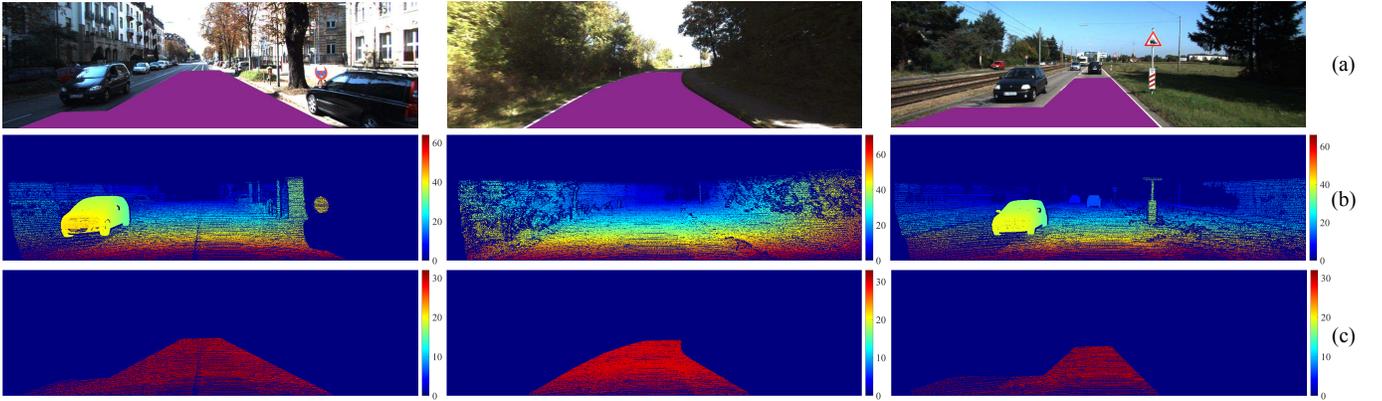


Fig. 12. Experimental results of the KITTI datasets: (a) left stereo images (the areas in purple are the labeled road areas); (b) ground truth disparity maps; (c) transformed disparity maps.

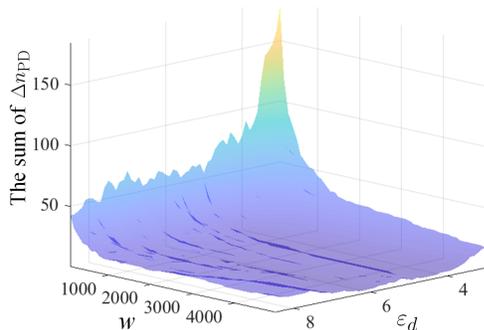


Fig. 13. The sum of Δn_{PD} with respect to different ε_d and w .

where $\tilde{\mathbf{d}} = [\tilde{d}_0, \tilde{d}_1, \dots, \tilde{d}_m]^T$ is a column vector storing the transformed disparity values. The average σ_d values of the KITTI stereo datasets are provided in Table I.

TABLE I
COMPARISON OF THE AVERAGE σ_d VALUES BETWEEN THE TWO CASES.

Dataset	Case 1	Case 2
KITTI stereo 2012 training dataset	0.4862	0.7800
KITTI stereo 2015 training dataset	0.5506	0.9428

In case 1, the effects caused by the non-zero roll angle are eliminated, while in case 2, the roll angle is assumed to be zero. From Table I, we can observe that the average σ_d values of these two datasets reduces by approximately half when the effects caused by the non-zero roll angle are eliminated. The average σ_d value of these two datasets is only 0.5188 pixels. Therefore, the proposed disparity transformation algorithm performs accurately and the transformed disparity values become more uniform. The runtime of the disparity transformation algorithm is about 0.142 seconds. In the next subsection, we will analyze the accuracy of pothole detection.

D. Evaluation of Pothole Detection

In Section II-C, a set of randomly selected disparities are modeled as a quadratic surface. The potholes are detected by

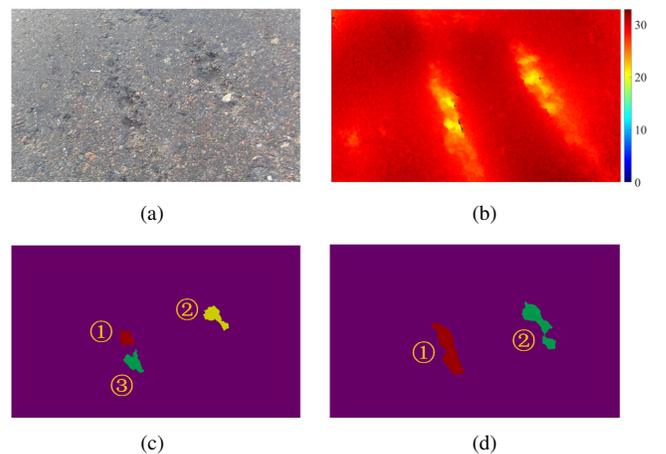


Fig. 14. Experimental result of the incorrect pothole detection: (a) left stereo image; (b) transformed disparity map; (c) detection result; (d) ground truth.

comparing the difference between the actual disparity map and the modeled quadratic surface. If a connected component contains more than w pixels and the disparity difference of each pixel exceeds ε_d , it will be identified as a pothole. In our experiments, we utilize the brute-force search method to find the best values of ε_d and w . The search range for ε_d and w are set to $[3.0, 8.5]$ and $[100, 5000]$, respectively. The step sizes for searching ε_d and w are set to 0.1 and 100, respectively.

For the first step, we go through the whole search range and record the number of detected potholes \hat{n}_{PD} in each frame. The absolute difference Δn_{PD} between each \hat{n}_{PD} and the expected pothole number n_{PD} is then computed. The sum of Δn_{PD} with respect to a pair of given ε_d and w can therefore be obtained, as illustrated in Fig. 13. In our experiments, the least sum of Δn_{PD} is equal to one and it is achieved only when $\varepsilon_d = 6.2$ and $w = 3100$. The corresponding incorrect pothole detection result is shown in Fig. 14. Incorrect detection occurs when the middle of the first pothole subsides and the selected parameters causes the system to detect two potholes instead of one. Some examples of successful detection results are shown in the fifth row of Fig. 15, and the corresponding ground truth is shown

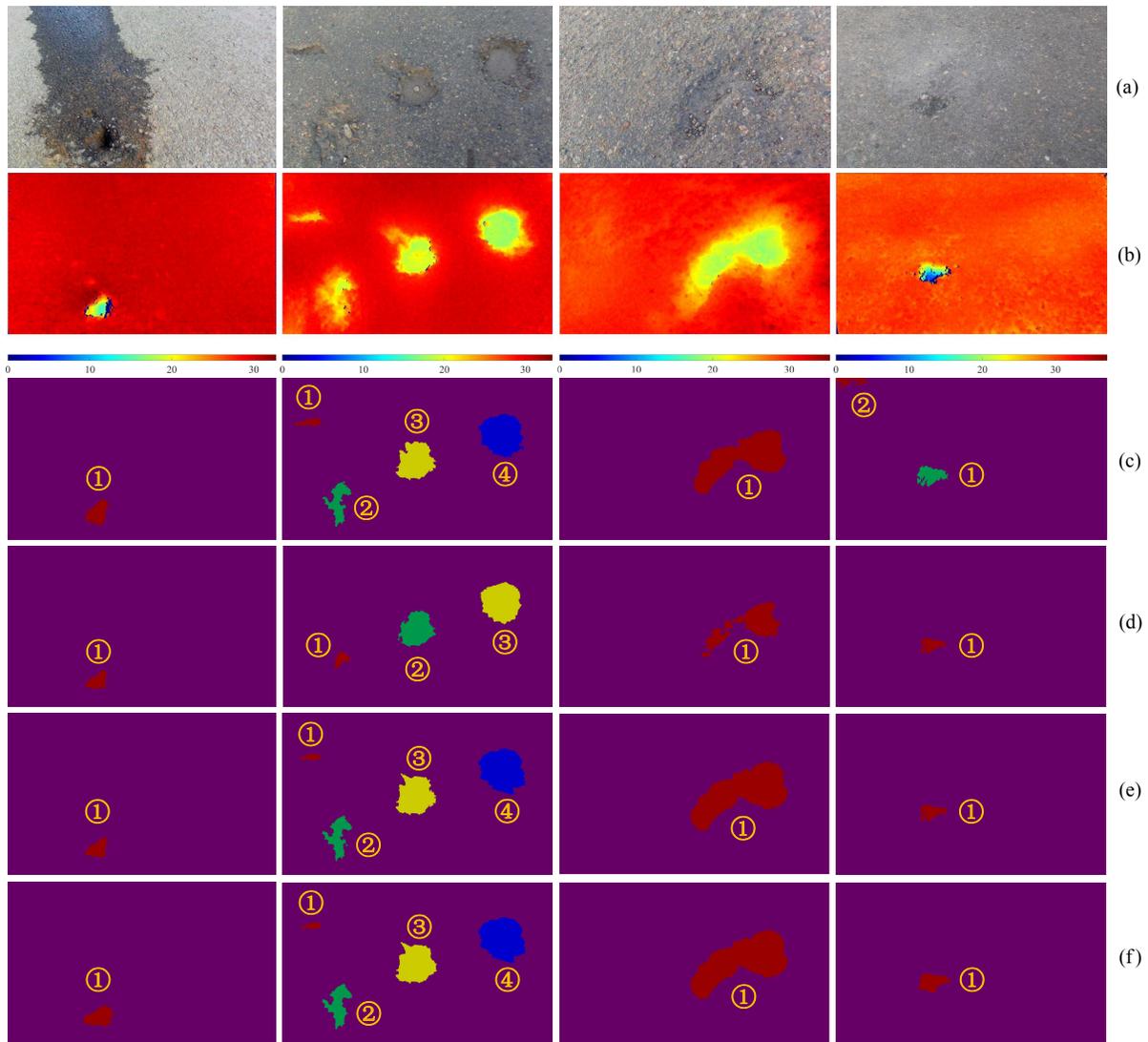


Fig. 15. Experimental results of pothole detection: (a) left stereo images. (b) transformed disparity maps. (c) pothole detection results obtained using the algorithm proposed in [15]. (d) pothole detection results obtained using the algorithm presented in [16]. (e) pothole detection results obtained using the proposed algorithm. (f) pothole ground truth.

TABLE II
COMPARISON OF POTHOLE DETECTION ACCURACY.

Dataset	Method	Total Potholes	Correct Detection	Incorrect Detection	Misdetection
Dataset 1	algorithm in [15]	22	11	11	0
	algorithm in [16]		22	0	0
	our algorithm		22	0	0
Dataset 2	algorithm in [15]	52	42	10	0
	algorithm in [16]		40	8	4
	our algorithm		51	1	0
Dataset 3	algorithm in [15]	5	5	0	0
	algorithm in [16]		5	0	0
	our algorithm		5	0	0
Total	algorithm in [15]	79	58	21	0
	algorithm in [16]		67	8	4
	our algorithm		78	1	0

in the sixth row.

We also compare our proposed algorithm with those produced in [15] and [16]. The pothole detection results obtained

using the algorithms presented in [15] and [16] are shown in the third and fourth rows of Fig. 15, respectively. The comparative pothole detection results are provided in Table

II, where we can see that the successful detection accuracy achieved using [15] and [16] are 73.4% and 84.8%, respectively. Compared to these, our proposed algorithm can detect potholes more accurately with a successful detection accuracy of 98.7%.

We also compare the proposed algorithm with [15] and [16] with respect to the pixel-level precision, recall, F-score and accuracy:

$$\text{precision} = \frac{n_{TP}}{n_{TP} + n_{FP}}, \quad (32)$$

$$\text{recall} = \frac{n_{TP}}{n_{TP} + n_{FN}}, \quad (33)$$

$$\text{F-score} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}, \quad (34)$$

$$\text{accuracy} = \frac{n_{TP} + n_{TN}}{n_{TP} + n_{TN} + n_{FP} + n_{FN}}, \quad (35)$$

where n_{TP} , n_{FP} , n_{FN} and n_{TN} are true positive, false positive, false negative and true negative pixel numbers, respectively.

The comparisons with respect to these four indicators are illustrated in Table III. It can be seen that our proposed algorithm outperforms [15] and [16], in terms of the pixel-level accuracy and it achieves an intermediate performance in terms of precision and recall. In addition, the precision and recall achieved using our proposed algorithm are very close to the highest values between [16] and [15]. Therefore, the proposed pothole detection algorithm performs both robustly and accurately.

TABLE III
COMPARISON OF PIXEL-LEVEL PRECISION, RECALL, ACCURACY AND F-SCORE.

Dataset	Method	precision	recall	F-score	accuracy
Dataset 1	[15]	0.5199	0.5427	0.5311	0.9892
	[16]	0.4622	0.9976	0.6317	0.9936
	proposed	0.4990	0.9871	0.6629	0.9940
Dataset 2	[15]	0.9754	0.9712	0.9733	0.9987
	[16]	0.8736	0.9907	0.9285	0.9968
	proposed	0.9804	0.9797	0.9800	0.9991
Dataset 3	[15]	0.6119	0.7714	0.6825	0.9948
	[16]	0.5339	0.9920	0.6942	0.9957
	proposed	0.5819	0.9829	0.7310	0.9961
Total	[15]	0.7799	0.8220	0.8004	0.9942
	[16]	0.6948	0.9921	0.8173	0.9954
	proposed	0.7709	0.9815	0.8635	0.9964

IV. CONCLUSION AND FUTURE WORK

The main contributions of this paper are novel disparity transformation and disparity map modeling algorithms. Using our method, undamaged road areas are better distinguishable in the transformed disparity map and can be easily extracted using Otsu's thresholding method. This greatly improves the robustness of disparity map modeling. To achieve greater processing efficiency, GSS and DP were utilized to estimate the transformation parameters. Furthermore, the disparities, whose normal vectors differ greatly from the optimal ones, were also discarded in the process of disparity map modeling, which further improves the accuracy of the modeled disparity map. Finally, the potholes were detected by comparing the

difference between the actual and modeled disparity maps. The point clouds of the detected potholes were then extracted from the reconstructed 3D road surface. In addition, we also created three datasets to contribute to stereo vision-based pothole detection research. The experimental results show that the overall successful detection rate of our proposed algorithm is around 98.7% and the pixel-level accuracy is approximately 99.6%.

However, the parameters set for pothole detection cannot be applied to all cases. Therefore, we plan to train a deep neural network to detect the potholes from the transformed disparity map. Furthermore, a road surface cannot always be considered as a quadratic one. Thus, we aim to design an algorithm to segment the reconstructed road surfaces and perform the proposed pothole detection algorithm on each segmented quadratic surface.

REFERENCES

- [1] J. S. Miller and W. Y. Bellinger, "Distress identification manual for the long-term pavement performance program," United States. Federal Highway Administration. Office of Infrastructure Research and Development, Tech. Rep., 2014.
- [2] S. Mathavan, K. Kamal, and M. Rahman, "A review of three-dimensional imaging technologies for pavement distress detection and measurements," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5, pp. 2353–2362, 2015.
- [3] T. Kim and S.-K. Ryu, "Review and analysis of pothole detection methods," *Journal of Emerging Trends in Computing and Information Sciences*, vol. 5, no. 8, pp. 603–608, 2014.
- [4] R. Fan, X. Ai, and N. Dahnoun, "Road surface 3D reconstruction based on dense subpixel disparity map estimation," *IEEE Transactions on Image Processing*, vol. PP, no. 99, p. 1, 2018.
- [5] C. Koch, K. Georgieva, V. Kasireddy, B. Akinci, and P. Fieguth, "A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure," *Advanced Engineering Informatics*, vol. 29, no. 2, pp. 196–210, 2015.
- [6] E. Buza, S. Omanovic, and A. Huseinovic, "Pothole detection with image processing and spectral clustering," in *Proceedings of the 2nd International Conference on Information Technology and Computer Networks*, 2013, pp. 48–53.
- [7] *Report a problem on a street*, Bristol City Council, <https://fixmystreet.bristol.gov.uk/>.
- [8] C. Koch and I. Brilakis, "Pothole detection in asphalt pavement images," *Advanced Engineering Informatics*, vol. 25, no. 3, pp. 507–515, 2011.
- [9] I. Pitas, *Digital image processing algorithms and applications*. John Wiley & Sons, 2000.
- [10] S. Li, C. Yuan, D. Liu, and H. Cai, "Integrated processing of image and gpr data for automated pothole detection," *Journal of computing in civil engineering*, vol. 30, no. 6, p. 04016015, 2016.
- [11] Y.-C. Tsai and A. Chatterjee, "Pothole detection and classification using 3d technology and watershed method," *Journal of Computing in Civil Engineering*, vol. 32, no. 2, p. 04017078, 2017.
- [12] M. R. Jahanshahi, F. Jazizadeh, S. F. Masri, and B. Becerik-Gerber, "Unsupervised approach for autonomous pavement-defect detection and quantification using an inexpensive depth sensor," *Journal of Computing in Civil Engineering*, vol. 27, no. 6, pp. 743–754, 2012.
- [13] C. Zhang and A. Elaksher, "An unmanned aerial vehicle-based imaging system for 3d measurement of unpaved road surface distresses," *Computer-Aided Civil and Infrastructure Engineering*, vol. 27, no. 2, pp. 118–129, 2012.
- [14] A. Barsi, I. Fi, T. Lovas, G. Melykuti, B. Takacs, C. Toth, and Z. Toth, "Mobile pavement measurement system: A concept study," in *Proc. ASPRS Annual Conference, Baltimore*, 2005, p. 8.
- [15] Z. Zhang, X. Ai, C. Chan, and N. Dahnoun, "An efficient algorithm for pothole detection using stereo vision," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 564–568.
- [16] A. Mikhailiuk and N. Dahnoun, "Real-time pothole detection on tms320c6678 dsp," in *Imaging Systems and Techniques (IST), 2016 IEEE International Conference on*. IEEE, 2016, pp. 123–128.

- [17] S. M. Abbas and A. Muhammad, "Outdoor rgb-d slam performance in slow mine detection," in *Robotics; Proceedings of ROBOTIK 2012; 7th German Conference on VDE*, 2012, pp. 1–6.
- [18] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge university press, 2003.
- [19] R. Fan, Y. Liu, X. Yang, M. J. Bocus, N. Dahnoun, and S. Tancock, "Real-time stereo vision for road surface 3-d reconstruction."
- [20] P. Pedregal, *Introduction to optimization*. Springer Science & Business Media, 2006, vol. 46.
- [21] U. Ozgunalp, R. Fan, X. Ai, and N. Dahnoun, "Multiple lane detection algorithm based on novel dense vanishing point estimation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 3, pp. 621–632, 2017.
- [22] A. M. Legendre, *Nouvelles méthodes pour la détermination des orbites des comètes*. F. Didot, 1805.
- [23] H. Eric Tseng, L. Xu, and D. Hrovat, "Estimation of land vehicle roll and pitch angles," *Vehicle System Dynamics*, vol. 45, no. 5, pp. 433–443, 2007.
- [24] J. Oh and S. B. Choi, "Vehicle roll and pitch angle estimation using a cost-effective six-dimensional inertial measurement unit," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 227, no. 4, pp. 577–590, 2013.
- [25] J. Ryu, E. J. Rossetter, and J. C. Gerdes, "Vehicle sideslip and roll parameter estimation using gps," in *Proceedings of the AVEC International Symposium on Advanced Vehicle Control*, 2002, pp. 373–380.
- [26] J. Ryu and J. C. Gerdes, "Estimation of vehicle roll and road bank angle," in *American Control Conference, 2004. Proceedings of the 2004*, vol. 3. IEEE, 2004, pp. 2110–2115.
- [27] R. Labayrade and D. Aubert, "A single framework for vehicle roll, pitch, yaw estimation and obstacles detection by stereovision," in *Intelligent Vehicles Symposium, 2003. Proceedings. IEEE*. IEEE, 2003, pp. 31–36.
- [28] P. Skulimowski, M. Owczarek, and P. Strumillo, "Ground plane detection in 3d scenes for an arbitrary camera roll rotation through 'v-disparity' representation."
- [29] M. Evans, R. Fan, and N. Dahnoun, "Iterative roll angle estimation from dense disparity map," in *Proc. 7th Mediterranean Conf. Embedded Computing (MECO)*, Jun. 2018, pp. 1–4.
- [30] E. W. Weisstein, "Least squares fitting," 2002.
- [31] D. York, "Least squares fitting of a straight line with correlated errors," *Earth and planetary science letters*, vol. 5, pp. 320–324, 1968.
- [32] R. Fan, M. J. Bocus, and N. Dahnoun, "A novel disparity transformation algorithm for road segmentation," *Information Processing Letters*, vol. 140, pp. 18–24, 2018.
- [33] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE transactions on systems, man, and cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [34] U. Ozgunalp, X. Ai, and N. Dahnoun, "Stereo vision-based road estimation assisted by efficient planar patch calculation," *Signal, Image and Video Processing*, vol. 10, no. 6, pp. 1127–1134, 2016.
- [35] I. Haller and S. Nedevschi, "Design of interpolation functions for subpixel-accuracy stereo-vision systems," *IEEE Transactions on image processing*, vol. 21, no. 2, pp. 889–898, 2012.
- [36] K. Klasing, D. Althoff, D. Wollherr, and M. Buss, "Comparison of surface normal estimation methods for range sensing applications," in *Proc. IEEE Int. Conf. Robotics and Automation*, May 2009, pp. 3206–3211.
- [37] C. Wang, H. Tanahashi, H. Hirayu, Y. Niwa, and K. Yamamoto, "Comparison of local plane fitting methods for range data."
- [38] T. Vaudrey, C. Rabe, R. Klette, and J. Milburn, "Differences between stereo and motion behaviour on synthetic and real-world stereo sequences," in *Image and Vision Computing New Zealand, 2008. IVCNZ 2008. 23rd International Conference*. IEEE, 2008, pp. 1–6.
- [39] A. Wedel, C. Rabe, T. Vaudrey, T. Brox, U. Franke, and D. Cremers, "Efficient dense scene flow from sparse or dense stereo data," in *European conference on computer vision*. Springer, 2008, pp. 739–751.
- [40] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2012, pp. 3354–3361.
- [41] M. Menze, C. Heipke, and A. Geiger, "Joint 3d estimation of vehicles and scene flow," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 2, p. 427, 2015.



He received the B.Sc. degree in control science and engineering from the Harbin Institute of Technology in 2015 and the Ph.D. degree in electrical and electronic engineering from the University of Bristol in 2018. Since the July of 2018, he has been working as a research associate with Robotics and Multi-Perception Laboratory (RAM-LAB) in the Robotics Institute at the Hong Kong University of Science and Technology. His research interests include multi-view geometry, real-time depth measurement, high performance computing, and vision-based autonomous applications, e.g., lane detection, obstacle detection, pothole detection, visual tracking, etc.



He received his B.Sc. degree in electrical and electronic engineering from the Eastern Mediterranean University with high honors in 2007. Then, he achieved his M.Sc. degree in electronic communications and computer engineering from the University of Nottingham, UK, in 2009. Later, he pursued his Ph.D. degree from University of Bristol, UK, in 2016. He is currently an assistant professor at the department of electrical and electronic Engineering of Near East University. His research interests include computer vision, pattern recognition, and intelligent vehicles.



He received the M.Sc degree in electronic and electrical engineering from Loughborough University, UK, in 2011, and the Ph.D in electronic and electrical engineering from the University of Bristol, UK, in 2016, supervised by Dr. Dimitris Agrafiotis and Prof. David Bull. In 2017 he joined the National Oceanography Centre, UK, and has primarily been involved in the research and development of object detection and classification tools in underwater imagery using autonomous underwater vehicles. His research interests include signal processing, video coding and compression, computer vision, and machine learning.

He received the B.Sc. degree in control science and engineering from the Harbin Institute of Technology in 2015 and the Ph.D. degree in electrical and electronic engineering from the University of Bristol in 2018. Since the July of 2018, he has been working as a research associate with Robotics and Multi-Perception Laboratory (RAM-LAB) in the Robotics Institute at the Hong Kong University of Science and Technology. His research interests include multi-view geometry, real-time depth measurement, high performance computing, and vision-based autonomous applications, e.g., lane detection, obstacle detection, pothole detection, visual tracking, etc.

He received his B.Sc. degree in electrical and electronic engineering from the Eastern Mediterranean University with high honors in 2007. Then, he achieved his M.Sc. degree in electronic communications and computer engineering from the University of Nottingham, UK, in 2009. Later, he pursued his Ph.D. degree from University of Bristol, UK, in 2016. He is currently an assistant professor at the department of electrical and electronic Engineering of Near East University. His research interests include computer vision, pattern recognition, and intelligent vehicles.

He received the M.Sc degree in electronic and electrical engineering from Loughborough University, UK, in 2011, and the Ph.D in electronic and electrical engineering from the University of Bristol, UK, in 2016, supervised by Dr. Dimitris Agrafiotis and Prof. David Bull. In 2017 he joined the National Oceanography Centre, UK, and has primarily been involved in the research and development of object detection and classification tools in underwater imagery using autonomous underwater vehicles. His research interests include signal processing, video coding and compression, computer vision, and machine learning.



Ming Liu received the B.A. degree in Automation at Tongji University in 2005. He graduated as a PhD student from the Department of Mechanical and Process Engineering of ETH Zurich in 2013, supervised by Prof. Roland Siegwart. He is involved in several NSF projects, and National 863-Hi-Tech-Plan projects in China. He is PI of over 20 projects including projects funded by RGC, NSFC, ITC, SZSTI, etc. He published over 90 papers in major international journals and conferences. He won the second place of EMAV'09 (European Micro Aerial Vehicle Competition). He got two awards from IARC 14 (International Aerial Robotics Contest). He won the Best Student Paper Award as the first author for MFI 2012 (IEEE International Conference on Multisensor Fusion and Information Integration), the Best Paper Award in Information for ICIA 2013 (IEEE International Conference on Information and Automation) as first author and Best Paper Award Finalists as co-author, the Best RoboCup Paper Award for IROS 2013 (IEEE/RSJ International Conference on Intelligent Robots and Systems). He won the Best Student Paper Award of IEEE ICAR 2017, the Best Paper Award in Automation for ICIA 2017. He won twice the innovation contest Chunhui Cup Winning Award in 2012 and 2013. He won the Wu Wenjun AI Innovation Award in 2016. He was the general chair of International Conference on Computer Vision Systems (ICVS) 2017; the program chair of IEEE International Conference on Real-time Computing and Robotics (IEEE-RCAR) 2016; the program chair of International Robotic Alliance Conference 2017. Ming Liu's research interests include dynamic environment modeling, 3D mapping, machine learning and visual control. He is the awardee of IEEE IROS Toshio Fukuda Young Professional Award 2018.



Ioannis Pitas (IEEE fellow, IEEE Distinguished Lecturer, EURASIP fellow) received the Diploma and PhD degree in Electrical Engineering, both from the Aristotle University of Thessaloniki, Greece. Since 1994, he has been a professor at the Department of Informatics of the same university. He served as a visiting professor at several universities. He was the chair of the school of graduate studies in computer science at the University of Thessaloniki. His current interests are in the areas of autonomous systems, image/video processing, machine learning, computer vision, intelligent digital media, human centered interfaces, affective computing, 3D imaging and biomedical imaging. He has published over 1090 papers, contributed in 50 books in his areas of interest and edited or (co-)authored another 11 books. He has also been member of the program committee of many scientific conferences and workshops. In the past he served as Associate Editor or co-Editor of 9 international journals and General or Technical Chair of 4 international conferences. He participated in 69 R&D projects, primarily funded by the European Union and is/was principal investigator/researcher in 41 such projects. He has 28700+ citations to his work and h-index 80+ (Google Scholar). Prof. Pitas leads the big European H2020 R&D project MULTIDRONE: <https://multidrone.eu/>. He is chair of the IEEE Autonomous Systems Initiative (ASI) <https://ieeetasi.signalprocessingsociety.org/>