Road Damage Detection Based on Unsupervised Disparity Map Segmentation

Rui Fan [ID], *Member, IEEE*, and Ming Liu [ID], *Senior Member, IEEE*

Abstract—This article presents a novel road damage detection algorithm based on unsupervised disparity map segmentation. Firstly, a disparity map is transformed by minimizing an energy function with respect to stereo rig roll angle and road disparity projection model. Instead of solving this energy minimization problem using non-linear optimization techniques, we directly find its numerical solution. The transformed disparity map is then segmented using Otus's thresholding method, and the damaged road areas can be extracted. The proposed algorithm requires no parameters when detecting road damage. The experimental results illustrate that our proposed algorithm performs both accurately and efficiently. The pixel-level road damage detection accuracy is approximately 97.56%. The source code is publicly available at: https://github.com/ruirangerfan/unsupervised_disparity_map_segmentation.git.

Index Terms—Road damage detection, disparity map segmentation, stereo rig roll angle, road disparity projection model, numerical solution.

I. INTRODUCTION

ROAD damage, notably pothole or crack, is not just an inconvenience, but also a safety hazard [1]. Road damage is regularly detected by certified inspectors [2]. This process is, however, cumbersome, costly and time-consuming [3]. Furthermore, the road damage detection results are always subjective, as they depend entirely on the inspectors' experience [4]. Therefore, there is an ever-increasing need for automated road condition assessment systems that can recognize and localize road damage both efficiently and objectively [5]. The rest of this section presents the state of the art in road damage detection and highlights the motivation, contributions and structure of this paper.

A. State-of-the-Art Road Damage Detection Methods

Over the past decade, passive and active sensing technologies have been extensively used to acquire 2D/3D road data [4]. 2D color/gray-scale road images are typically captured by digital cameras [6], while 3D road data, e.g., road point cloud or road depth/disparity map, are supplied by laser scanners [7], Microsoft Kinect sensors [8], or passive sensors (i.e., a single movable camera [9] or an array of synchronized cameras [10]). The state-of-the-art road damage detection methods can be classified as either 3D road surface modeling-based [11] or 2D image analysis-based. The former commonly fits a quadratic surface to the raw 3D road data and detect the damaged road areas by comparing the difference between the raw data and the modeled road surface [11].

On the other hand, 2D image analysis-based road damage detection methods can be grouped into two categories: computer vision-based [6], [7], [12]–[15] and machine learning-based [16]–[19]. The former typically pre-processes a 2D image, i.e.,

an RGB/gray-scale image or a depth/disparity map, using some image processing techniques, e.g., various image filters, to reduce image noise and enhance road damage outline [12], [13]. The pre-processed image is then segmented using some thresholding methods, such as Otsu [14], triangle [6] or watershed [7], to extract damaged road areas. In [13], we proposed a disparity transformation algorithm which can better distinguish between damaged and undamaged road areas. The transformation parameters were estimated by minimizing an energy function using golden section search (GSS) and dynamic programming (DP). Recently, we proposed to minimize the aforementioned energy function using gradient descent (GD), which has shown a better efficiency [15].

With recent advances in supervised learning, deep convolutional neural networks (CNNs) have been used for road image classification and semantic road image segmentation. For example, Cha *et al.* [17] cropped the RGB images into a group of squared image patches and labeled them as either positive or negative ones. The labeled training data were then used to train a CNN for road image patch classification [17]. In [18], the authors utilized thermal images to train a residual network (ResNet) [20] for road image classification. Furthermore, Wu *et al.* [19] developed a robust road image segmentation system based on DeepLabv3+ [21], which employs atrous convolution along with upsampled filters to extract dense feature maps and to capture long-range context.

B. Motivation

Currently, laser scanning is still the main technology used for 3D road data acquisition, while other technologies, such as passive sensing, are under-utilized [5]. However, the long-term maintenance of such laser scanners is still very expensive [6]. Furthermore, Microsoft Kinect sensors were initially designed for indoor use, and they suffer greatly from infra-red saturation in direct sunlight [22]. Therefore, the trend of 3D road data acquisition is to utilize digital cameras, notably stereo cameras.

For 3D road surface modeling-based methods, finding the best parameters is very challenging, as the parameters they select cannot be applied to all cases [11]. On the other hand, computer vision-based methods can recognize road damage with low computational complexity, but the achieved detection accuracy is still far from satisfactory [4]. For machine learning-based methods, training a road image classification/segmentation neural network using supervised learning requires a large amount of labeled training data, and producing such data can be a long and labor-intensive task [4]. Moreover, color/gray-scale image segmentation is always severely affected by various environment factors, notably illumination conditions, but disparity/depth map segmentation is not subject to such environment factors [23]. Therefore, there is a strong motivation to explore unsupervised disparity map segmentation method for road damage detection.

C. Novel Contributions

In this paper, we present a real-time road damage detection algorithm based on unsupervised disparity map segmentation. The proposed algorithm is developed from our previous work [11],

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

2   IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS
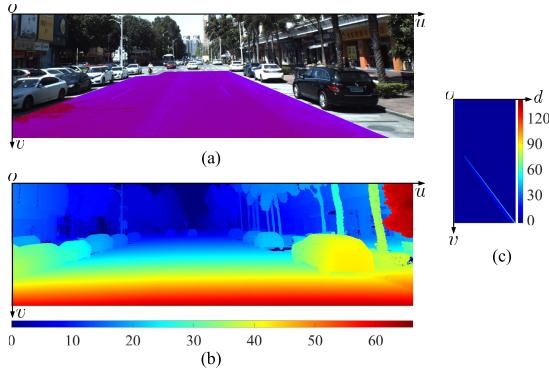

Fig. 1. (a) Left stereo image, where the area in purple is our manually labeled road region; (b) disparity map; (c) v-disparity image.

where road disparity maps were transformed to better distinguish between damaged and undamaged road areas. Instead of estimating the transformation parameters using non-linear optimization methods, such as GSS-DP and GD, we directly find the numerical solution for the energy minimization problem stated in [11]. The proposed algorithm is capable of segmenting dense disparity maps for road damage detection without setting any parameters. Furthermore, the stereo rig roll angle can be accurately estimated from disparity maps, which enables our method to be utilized for vehicle state estimation. We also believe this algorithm can be utilized to automatically label training data for road damage detection.

D. Paper Structure

The remainder of this paper is organized as follows: Section II introduces v-disparity image and road disparity projection model. Section III presents the proposed unsupervised road damage detection algorithm. The experimental results are illustrated and the algorithm performance is discussed in Section IV. Finally, Section V summaries the paper.

II. PRELIMINARIES

Since Labayrade and Aubert [24] introduced the concept of "v-disparity image" in 2003, disparity map has been widely used for road region extraction [25]. An example left stereo image and its corresponding dense disparity map are shown in Fig. 1(a) and 1(b), respectively, where the purple area in Fig. 1(a) is our manually labeled road region. By computing the disparity histogram with respect to each image row, a v-disparity image can be created [26], as shown in Fig. 1(c).

Since the road surface is generally considered as a ground plane, for a stereo rig whose baseline is perfectly parallel to the road surface, its roll angle θ equals 0, and the disparities on each row have similar values [13], as shown in Fig. 1(b). Therefore, the projections of road disparities on the v-disparity image can be represented by a linear model [27]:

$$f(\mathbf{p}) = a_0 + a_1 v, \tag{1}$$

where $\mathbf{a} = [a_0, a_1]^\top$ stores the coefficients of the linear model, and $\mathbf{p} = [u, v]^\top$ is a pixel in the disparity map. $\mathbf{a}$ can be estimated by minimizing the following energy:

$$E = \|\mathbf{d} - \mathbf{Va}\|_2^2, \tag{2}$$

where $\mathbf{d} = [d_1, d_2, \cdots, d_n]^\top$ stores the disparity values. $\mathbf{V} = [\mathbf{1}_n, \mathbf{v}]$, where $\mathbf{1}_k$ represents a $k \times 1$ vector of ones and
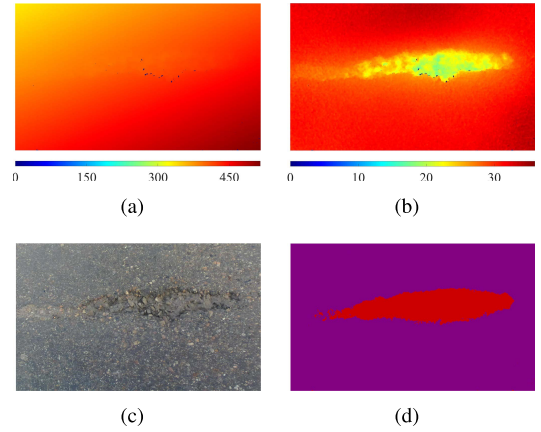

Fig. 2. Unsupervised disparity map segmentation: (a) original disparity map; (b) transformed disparity map; (c); left stereo image; (d) road damage detection result, where the regions in purple and red are undamaged and damaged road areas, respectively.

$\mathbf{v} = [v_1, v_2, \cdots, v_n]^\top$. The above energy minimization problem has a closed form solution:

$$\mathbf{a} = (\mathbf{V}^\top \mathbf{V})^{-1} \mathbf{V}^\top \mathbf{d}. \tag{3}$$

Plugging (3) into (2) obtains the minimum energy:

$$E_{\min} = \mathbf{d}^\top \mathbf{d} - \mathbf{d}^\top \mathbf{V}(\mathbf{V}^\top \mathbf{V})^{-1} \mathbf{V}^\top \mathbf{d}. \tag{4}$$

III. ALGORITHM DESCRIPTION

However, when the stereo rig baseline is not parallel to the road surface, a non-zero roll angle θ will be introduced into the imaging process. This fact leads to gradual disparity change in the horizontal direction (see Fig. 2(a)), making the way of representing road disparity projections using (1) somewhat problematic [13]. Furthermore, compared to the case that the roll angle is zero, the disparity distribution of each row becomes less compact and $E_{\min}$ becomes much higher. Therefore, the roll angle has to be considered when minimizing (2).

To rotate the disparity map at a given angle θ around the image center, each original point $\mathbf{p} = [u, v]^\top$ is transformed to a new point $\mathbf{q} = [s, t]^\top$ using:

$$\mathbf{q}(\theta, \mathbf{p}) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \mathbf{p}. \tag{5}$$

(2) can, therefore, have a more general expression:

$$E(\theta) = \|\mathbf{d} - \mathbf{T}(\theta)\mathbf{a}(\theta)\|_2^2, \tag{6}$$

where $\mathbf{T}(\theta) = [\mathbf{1}_n, \ \mathbf{t}(\theta)]$ and $\mathbf{t} = [t_1(\theta), t_2(\theta), \cdots, t_n(\theta)]^\top$. The closed form solution for (6) is as follows:

$$\mathbf{a}(\theta) = \left(\mathbf{T}(\theta)^\top \mathbf{T}(\theta)\right)^{-1} \mathbf{T}(\theta)^\top \mathbf{d}. \tag{7}$$

(4) can, therefore, be rewritten as follows:

$$E_{\min}(\theta) = \mathbf{d}^\top \mathbf{d} - \mathbf{d}^\top \mathbf{T}(\theta)\left(\mathbf{T}(\theta)^\top \mathbf{T}(\theta)\right)^{-1} \mathbf{T}(\theta)^\top \mathbf{d}. \tag{8}$$

Minimizing (8) is equivalent to maximizing:

$$g(\theta) = \mathbf{d}^\top \mathbf{T}(\theta)\left(\mathbf{T}(\theta)^\top \mathbf{T}(\theta)\right)^{-1} \mathbf{T}(\theta)^\top \mathbf{d} \ \text{ s.t. } \theta \in (-\frac{\pi}{2}, \frac{\pi}{2}]. \tag{9}$$

According to (5), we can obtain:

$$\mathbf{F} = \mathbf{T}(\theta)^\top \mathbf{T}(\theta) = \begin{bmatrix} n & r_0(\theta) \\ r_0(\theta) & r_1(\theta) \end{bmatrix}, \tag{10}$$
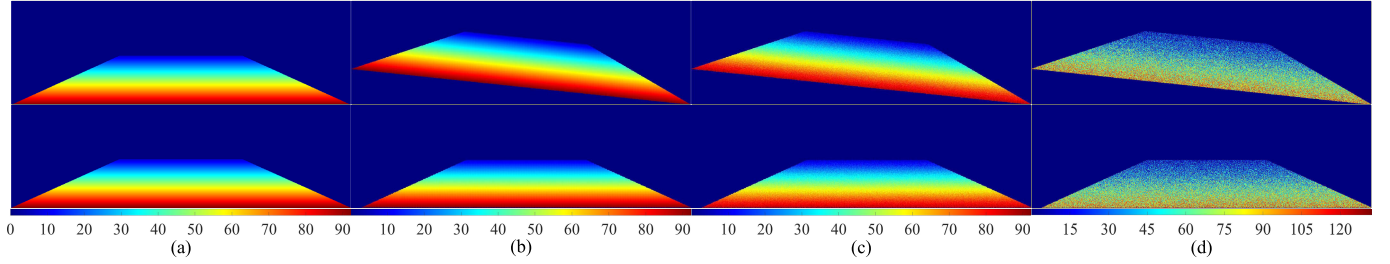
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

FAN AND LIU: ROAD DAMAGE DETECTION BASED ON UNSUPERVISED DISPARITY MAP SEGMENTATION

3

Fig. 3. Experimental results of roll angle estimation: (a) $\theta = 0°$ and $\kappa = 0$; (b) $\theta = 10°$ and $\kappa = 0$; (c) $\theta = 10°$ and $\kappa = 5$; (d) $\theta = 10°$ and $\kappa = 40$.

where

$$r_0(\theta) = \mathbf{v}^\top \mathbf{1}_n \cos \theta - \mathbf{u}^\top \mathbf{1}_n \sin \theta, \tag{11}$$

$$r_1(\theta) = \frac{\mathbf{v}^\top \mathbf{v} + \mathbf{u}^\top \mathbf{u}}{2} + \frac{\mathbf{v}^\top \mathbf{v} - \mathbf{u}^\top \mathbf{u}}{2} \cos 2\theta - \mathbf{u}^\top \mathbf{v} \sin 2\theta, \tag{12}$$

$\mathbf{u} = [u_1, u_2, \cdots, u_n]^\top$ and $\mathbf{v} = [v_1, v_2, \cdots, v_n]^\top$ are two column vectors storing the horizontal and vertical coordinates, respectively. (10), (11) and (12) result in the following expression:

$$\mathbf{F}^{-1} = \frac{1}{nr_1(\theta) - r_0(\theta)^2} \begin{bmatrix} r_1(\theta) & -r_0(\theta) \\ -r_0(\theta) & n \end{bmatrix}, \tag{13}$$

Plugging (11)-(13) into (9) results in the following expression:

$$g(\theta) = \frac{w_3 + w_4 \cos 2\theta + w_5 \sin 2\theta}{w_0 + w_1 \cos 2\theta + w_2 \sin 2\theta} \quad \text{s.t. } \theta \in (-\frac{\pi}{2}, \frac{\pi}{2}], \tag{14}$$

where

$$w_0 = \frac{1}{2}\left[n(\mathbf{v}^\top \mathbf{v} + \mathbf{u}^\top \mathbf{u}) - (\mathbf{v}^\top \mathbf{1}_n)^2 - (\mathbf{u}^\top \mathbf{1}_n)^2\right], \tag{15}$$

$$w_1 = \frac{1}{2}\left[n(\mathbf{v}^\top \mathbf{v} - \mathbf{u}^\top \mathbf{u}) - (\mathbf{v}^\top \mathbf{1}_n)^2 + (\mathbf{u}^\top \mathbf{1}_n)^2\right], \tag{16}$$

$$w_2 = \mathbf{v}^\top \mathbf{1}_n \mathbf{u}^\top \mathbf{1}_n - n\mathbf{v}^\top \mathbf{u}, \tag{17}$$

$$w_3 = \frac{1}{2}\left[(\mathbf{d}^\top \mathbf{1}_n)^2(\mathbf{v}^\top \mathbf{v} + \mathbf{u}^\top \mathbf{u}) + n((\mathbf{d}^\top \mathbf{v})^2 + (\mathbf{d}^\top \mathbf{u})^2)\right]$$
$$- \mathbf{d}^\top \mathbf{1}_n(\mathbf{v}^\top \mathbf{1}_n \mathbf{d}^\top \mathbf{v} + \mathbf{u}^\top \mathbf{1}_n \mathbf{d}^\top \mathbf{u}), \tag{18}$$

$$w_4 = \frac{1}{2}\left[(\mathbf{d}^\top \mathbf{1}_n)^2(\mathbf{v}^\top \mathbf{v} - \mathbf{u}^\top \mathbf{u}) + n((\mathbf{d}^\top \mathbf{v})^2 - (\mathbf{d}^\top \mathbf{u})^2)\right]$$
$$- \mathbf{d}^\top \mathbf{1}_n(\mathbf{v}^\top \mathbf{1}_n \mathbf{d}^\top \mathbf{v} - \mathbf{u}^\top \mathbf{1}_n \mathbf{d}^\top \mathbf{u}), \tag{19}$$

$$w_5 = \mathbf{d}^\top \mathbf{1}_n(\mathbf{v}^\top \mathbf{1}_n \mathbf{d}^\top \mathbf{u} + \mathbf{u}^\top \mathbf{1}_n \mathbf{d}^\top \mathbf{v}) - (\mathbf{d}^\top \mathbf{1}_n)^2 \mathbf{v}^\top \mathbf{u}$$
$$- n\mathbf{d}^\top \mathbf{v} \mathbf{d}^\top \mathbf{u}. \tag{20}$$

The angle θ which maximizes $g(\theta)$ can be obtained by differentiating $g(\theta)$ with respect to θ:

$$\frac{\delta g(\theta)}{\delta \theta} = \frac{-2}{(w_0 + w_1 \cos 2\theta + w_2 \sin 2\theta)^2}\Big((w_4 w_2 - w_5 w_1)$$
$$+ (w_3 w_2 - w_5 w_0) \cos 2\theta + (w_4 w_0 - w_3 w_1) \sin 2\theta\Big). \tag{21}$$

If the denominator of (21) does not equal zero, we can get two angles θ_1 and θ_2 at which $g(\theta)$ achieves the extrema:

$$\theta_1 = \arctan\left(\frac{w_4 w_0 - w_3 w_1 + \sqrt{\Delta}}{w_3 w_2 + w_5 w_1 - w_5 w_0 - w_4 w_2}\right), \tag{22}$$

$$\theta_2 = \arctan\left(\frac{w_4 w_0 - w_3 w_1 - \sqrt{\Delta}}{w_3 w_2 + w_5 w_1 - w_5 w_0 - w_4 w_2}\right), \tag{23}$$

where

$$\Delta = (w_4 w_0 - w_3 w_1)^2 + (w_3 w_2 - w_5 w_0)^2$$
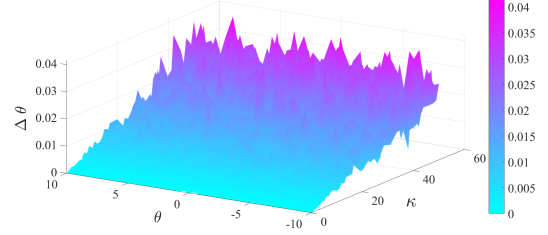$$- (w_4 w_2 - w_5 w_1)^2. \tag{24}$$

Fig. 4. $\Delta \theta$ with respect to θ and κ.

The desirable roll angle θ can, therefore, be determined by finding the highest value between $g(\theta_1)$ and $g(\theta_2)$. $\mathbf{a}$ can then be obtained by substituting θ into (7). Each road disparity can now be represented using:

$$f(\mathbf{p}, \theta) = a_0 + a_1(-u \sin \theta + v \cos \theta). \tag{25}$$

Damaged and undamaged road areas can now be better distinguished by transforming the original disparity map D (see Fig. 2(a)) to a new disparity map $\tilde{D}$ (see Fig. 2(b)) using:

$$\tilde{D}(\mathbf{p}) = D(\mathbf{p}) - f(\mathbf{p}, \theta) + \delta, \tag{26}$$

where δ can be any constant enabling the transformed disparity values to be non-negative. The transformed disparity map is shown in Fig. 2(b). Finally, the damaged road areas can be extracted by applying Otsu's thresholding method on the transformed disparity map. The corresponding result is shown in Fig. 2(d).

IV. EXPERIMENTAL RESULTS

The proposed road damage detection algorithm is programmed in both C++ and Matlab C on an Intel Core i7-8700K CPU (3.7 GHz) using a single thread. The remainder of this section discusses the performance of roll angle estimation, disparity transformation, and road damage detection, respectively.

A. Roll Angle Estimation Evaluation

To quantify the accuracy the proposed roll angle estimation algorithm, we created a synthetic dataset (including 51 dense disparity maps with respect to different roll angles $\theta \in [-10°, +10°]$). The KITTI stereo rig configurations[1] are utilized to create these synthetic disparity maps. To further evaluate the robustness of the proposed roll angle estimation algorithm, we added Gaussian white noise $\kappa \omega$ to the synthetic disparity maps, where $\omega \in [-1, +1]$ is a random decimal value and κ is a scale parameter set to control the intensity of the noise. Some examples of the experimental results are illustrated in Fig. 3, where the first row illustrates the original disparity maps and the second row shows the disparity maps rotated around the estimated roll angles. To quantify the accuracy of our proposed roll

[1] http://www.cvlibs.net/datasets/kitti/setup.php

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

4

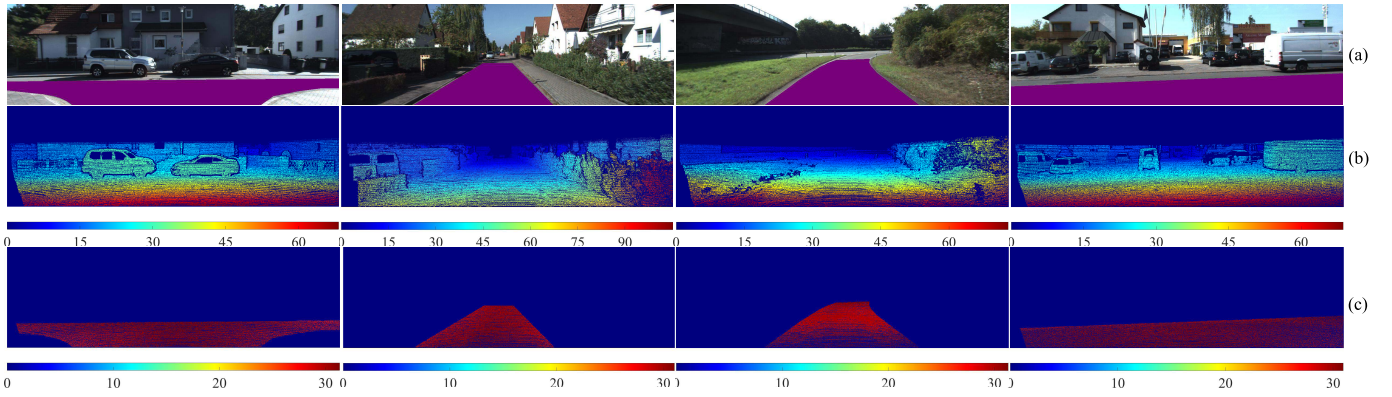IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS

Fig. 5. Experimental results of the KITTI stereo dataset: (a) left stereo images; (b) original disparity maps; (c) transformed disparity maps.
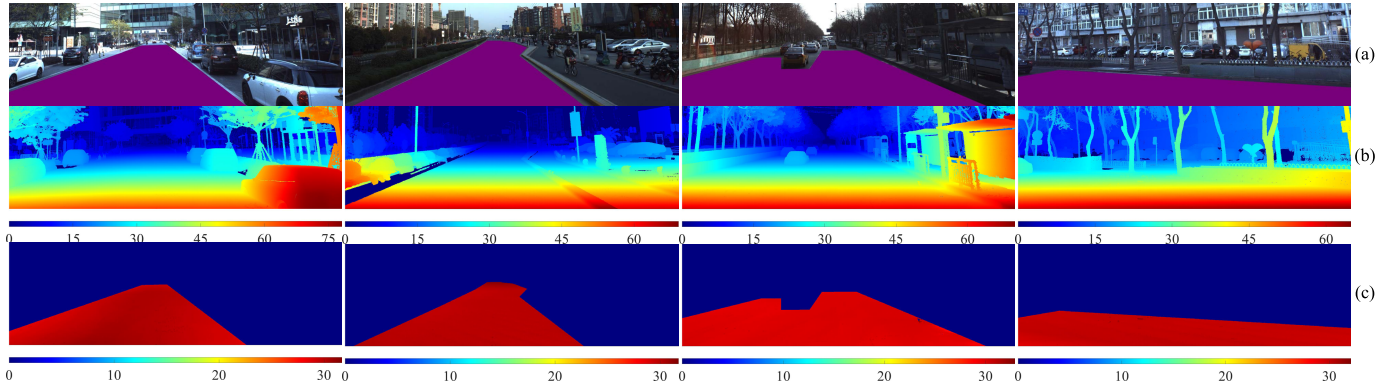

Fig. 6. Experimental results of the ApolloScape stereo dataset: (a) left stereo images; (b) original disparity maps; (c) transformed disparity maps.
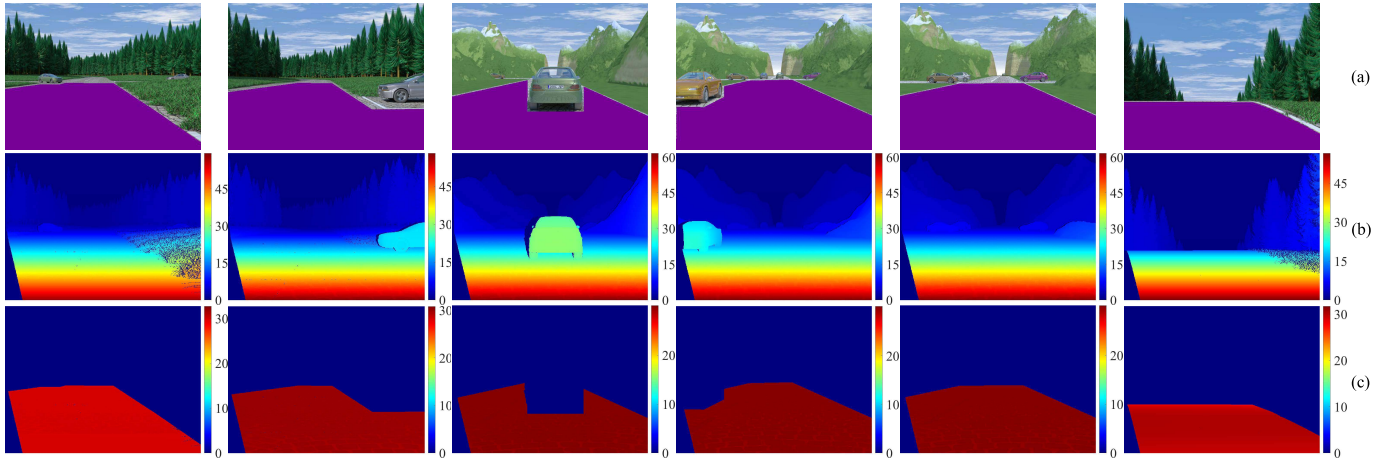

Fig. 7. Experimental results of the EISATS stereo dataset: (a) left stereo images; (b) original disparity maps; (c) transformed disparity maps.

angle estimation algorithm, we compute the absolute difference $\Delta\theta$ between the estimated and actual roll angles, i.e., $\tilde{\theta}$ and θ. $\Delta\theta$ with respect to different θ and κ is shown in Fig. 4, where we can observe that the accuracy of our proposed roll angle estimation algorithm decreases with the increase of κ, but the highest $\Delta\theta$ is only about $0.04°$ ($\kappa = 50$). Therefore, our proposed roll angle estimation algorithm is highly accurate and very robust to noise.

B. Disparity Transformation Evaluation

As discussed in Section III, the road damage becomes highly distinguishable after unsupervised disparity transformation. The transformed disparities in the undamaged road areas tend to have similar values, while they differ greatly from those in the damaged road areas, as shown in Fig. 2(b).

In our experiments, we utilized the KITTI stereo [28], [29], the ApolloScape[2] stereo, and the EISATS stereo [30], [31] datasets to evaluate the performance of our proposed disparity transformation algorithm. The former two datasets are used for the evaluation of sparse and dense real-world disparity map transformation, respectively. The EISATS stereo dataset is utilized to evaluate the performance of our proposed algorithm on synthetic disparity maps. The corresponding experimental results are illustrated

[2]http://apolloscape.auto/stereo.html

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

FAN AND LIU: ROAD DAMAGE DETECTION BASED ON UNSUPERVISED DISPARITY MAP SEGMENTATION 5
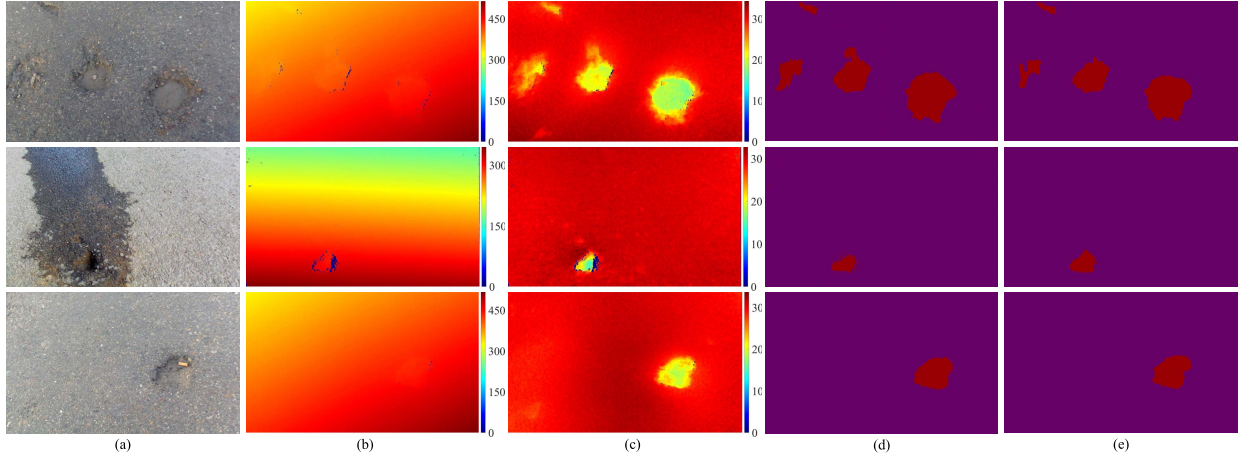

Fig. 8. Experimental results of road damage detection: (a) left stereo images; (b) original disparity maps; (c) transformed disparity maps; (d) detection results; (e) ground truth; the regions in purple and red are undamaged and damaged road areas, respectively.

TABLE I
COMPARISONS OF σ AND RUNTIME

Dataset	Method	σ	runtime (ms)
KITTI	GSS-DP [13]	0.4305	32.3182
	GD [15]	0.4299	7.5107
	Proposed	**0.4289**	**1.1279**
ApolloScape	GSS-DP [13]	0.2614	523.5302
	GD [15]	0.2613	91.8474
	Proposed	**0.2610**	**37.2680**
EISATS	GSS-DP [13]	0.2757	45.6317
	GD [15]	0.2752	8.9177
	Proposed	**0.2703**	**2.3574**

in Fig. 5, 6 and 7, respectively, where the areas in purple are our manually labeled road regions.

To quantify the disparity transformation accuracy, we introduced a measure named transformed disparity standard deviation σ:

$$\sigma = \sqrt{\frac{1}{m} \left\| \tilde{\mathbf{d}} - \frac{\tilde{\mathbf{d}}^\top \mathbf{1}_m}{m} \right\|_2^2},$$ (27)

where $\tilde{\mathbf{d}} = [\tilde{D}(\mathbf{p}_1), \tilde{D}(\mathbf{p}_2), \cdots, \tilde{D}(\mathbf{p}_m)]^\top$ stores the transformed disparity values. We compare our proposed method with GSS-DP [13] and GD [15]. The comparisons of σ and runtime are illustrated in Table I. It can be clearly seen that our proposed algorithm achieves the minimum σ on all the stereo datasets. Furthermore, as our proposed algorithm can directly obtain the numerical solution for (6), it performs much faster than both [13] and [15].

C. Road Damage Detection Evaluation

In this subsection, we utilize our recently published pothole detection dataset[3] [11] to evaluate the performance of road damage detection. Some examples of the detected damaged road areas are shown on the fourth column in Fig. 8. To quantify the accuracy of our proposed road damage detection algorithm, we compute the pixel-level precision, recall, F1-score, IoU, and accuracy, as shown in Fig. 9. It can be seen that our proposed road damage detection algorithm performs accurately. The pixel-level accuracy of the detected road damage areas is approximately 97.56%.
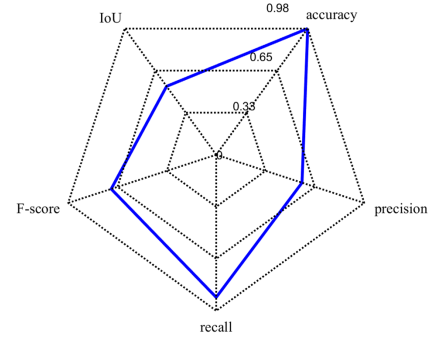
[3]ruirangerfan.com

Fig. 9. The pixel-level accuracy, precision, recall, F1-score, and IoU achieved using the proposed algorithm.

V. CONCLUSION

This paper presented a novel road damage detection algorithm based on unsupervised disparity map segmentation. This was achieved by minimizing an energy function with respect to the stereo rig roll angle and the road disparity projection model. Instead of minimizing this energy function using non-linear optimization methods, such as GSS-DP and GD, we directly found its numerical solution, which enables our proposed algorithm to perform more accurately and efficiently than GSS-DP and GD. A dense disparity map can, therefore, be transformed to better distinguish between damaged and undamaged road areas. By applying Otsu's thresholding method on the transformed disparity map, the road damage can then be effectively detected. The proposed algorithm does not require any parameters when transforming and segmenting road disparity maps. The experimental results also demonstrated that our algorithm can perform in real time. The pixel-level accuracy of the detected road damage areas is approximately 97.56%.

REFERENCES

[1] R. Fan, "Real-time computer stereo vision for automotive applications," Ph.D. dissertation, Univ. Bristol, Bristol, U.K., 2018.

[2] T. Kim and S. K. Ry, "Review and analysis of pothole detection methods," *J. Emerg. Trends Comput. Inf. Sci.*, vol. 5, no. 8, pp. 603–608, Aug. 2014.

[3] R. Fan, X. Ai, and N. Dahnoun, "Road surface 3D reconstruction based on dense subpixel disparity map estimation," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 3025–3035, Jun. 2018.

[4] 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," *Adv. Eng. Inform.*, vol. 29, no. 2, pp. 196–210, 2015.

[5] S. Mathavan, K. Kamal, and M. Rahman, "A review of three-dimensional imaging technologies for pavement distress detection and measurements," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 5, pp. 2353–2362, Oct. 2015.

[6] C. Koch and I. Brilakis, "Pothole detection in asphalt pavement images," *Adv. Eng. Inform.*, vol. 25, no. 3, pp. 507–515, 2011.

[7] Y.-C. Tsai and A. Chatterjee, "Pothole detection and classification using 3D technology and watershed method," *J. Comput. Civil Eng.*, vol. 32, no. 2, 2017, Art. no. 04017078.

[8] 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," *J. Comput. Civil Eng.*, vol. 27, no. 6, pp. 743–754, 2013.

[9] G. Jog, C. Koch, M. Golparvar-Fard, and I. Brilakis, "Pothole properties measurement through visual 2D recognition and 3D reconstruction," in *Proc. Comput. Civil Eng.*, 2012, pp. 553–560.

[10] R. Fan, Y. Liu, X. Yang, M. J. Bocus, N. Dahnoun, and S. Tancock, "Real-time stereo vision for road surface 3-D reconstruction," in *Proc. IEEE Int. Conf. Imag. Syst. Techn. (IST)*, Oct. 2018, pp. 1–6.

[11] R. Fan, U. Ozgunalp, B. Hosking, M. Liu, and I. Pitas, "Pothole detection based on disparity transformation and road surface modeling," *IEEE Trans. Image Process.*, vol. 29, no. 1, pp. 897–908, Dec. 2020.

[12] S. Li, C. Yuan, D. Liu, and H. Cai, "Integrated processing of image and GPR data for automated pothole detection," *J. Comput. Civil Eng.*, vol. 30, no. 6, 2016, Art. no. 04016015.

[13] R. Fan, M. J. Bocus, and N. Dahnoun, "A novel disparity transformation algorithm for road segmentation," *Inf. Process. Lett.*, vol. 140, pp. 18–24, Dec. 2018.

[14] E. Buza, A. Akagic, S. Omanovic, and H. Hasic, "Unsupervised method for detection of high severity distresses on asphalt pavements," in *Proc. IEEE 14th Int. Sci. Conf. Inform.*, Nov. 2017, pp. 45–50.

[15] R. Fan, J. Jiao, J. Pan, H. Huang, S. Shen, and M. Liu, "Real-time dense stereo embedded in a UAV for road inspection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019. [Online]. Available: http://openaccess.thecvf.com/content_CVPRW_2019/html/UAVision/Fan_Real-Time_Dense_Stereo_Embedded_in_a_UAV_for_Road_Inspection_CVPRW_2019_paper.html

[16] A. Cord and S. Chambon, "Automatic road defect detection by textural pattern recognition based on AdaBoost," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 27, no. 4, pp. 244–259, 2012.

[17] Y.-J. Cha, W. Choi, and O. Büyüköztürk, "Deep learning-based crack damage detection using convolutional neural networks," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 32, no. 5, pp. 361–378, 2017.

[18] Y. Bhatia *et al.*, "Convolutional neural networks based potholes detection using thermal imaging," *J. King Saud Univ.–Comput. Inf. Sci.*, to be published.

[19] H. Wu *et al.*, "Road pothole extraction and safety evaluation by integration of point cloud and images derived from mobile mapping sensors," *Adv. Eng. Inform.*, vol. 42, Oct. 2019, Art. no. 100936.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.

[21] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," 2017, *arXiv:1706.05587*. [Online]. Available: https://arxiv.org/abs/1706.05587

[22] S. M. Abbas and A. Muhammad, "Outdoor RGB-D SLAM performance in slow mine detection," in *Proc. 7th German Conf. Robot.*, May 2012, pp. 1–6.

[23] E. Salari and G. Bao, "Automated pavement distress inspection based on 2D and 3D information," in *Proc. IEEE Int. Conf. Electro/Inf. Technol.*, May 2011, pp. 1–4.

[24] R. Labayrade and D. Aubert, "A single framework for vehicle roll, pitch, yaw estimation and obstacles detection by stereovision," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2003, pp. 31–36.

[25] R. Fan and N. Dahnoun, "Real-time stereo vision-based lane detection system," *Meas. Sci. Technol.*, vol. 29, no. 7, 2018, Art. no. 074005.

[26] U. Ozgunalp, R. Fan, X. Ai, and N. Dahnoun, "Multiple lane detection algorithm based on novel dense vanishing point estimation," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 3, pp. 621–632, Mar. 2017.

[27] Z. Hu, F. Lamosa, and K. Uchimura, "A complete U-V-disparity study for stereovision based 3D driving environment analysis," in *Proc. 5th Int. Conf. 3-D Digit. Imag. Modeling*, Jun. 2005, pp. 204–211.

[28] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 3354–3361.

[29] M. Menze, C. Heipke, and A. Geiger, "Joint 3D estimation of vehicles and scene flow," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 2, p. 427, Aug. 2015.

[30] T. Vaudrey, C. Rabe, R. Klette, and J. Milburn, "Differences between stereo and motion behaviour on synthetic and real-world stereo sequences," in *Proc. 23rd Int. Conf. Image Vis. Comput.*, Christchurch, New Zealand, Nov. 2008, pp. 1–6.

[31] A. Wedel, C. Rabe, T. Vaudrey, T. Brox, U. Franke, and D. Cremers, "Efficient dense scene flow from sparse or dense stereo data," in *Proc. Eur. Conf. Comput. Vis.* Marseille, France: Springer, Oct. 2008, pp. 739–751.