CETRA2020* 20-21 May 2021, Zagreb, Croatia
6th International Conferenceon Road and Rail Infrastructure

DEVELOPMENT OF AUTOMATIC ROAD WIDTH AND POTHOLE SIZE ESTIMATION METHOD FROM DASHCAM VIDEO FOR UNDER DEVELOPING COUNTRIES

Hidekazu Fukai[1], Fernão A.L.N. Mouzinho[1,2], Ryo Nagae[1], Masayuki Uchida[1]
[1] *National University of Timor Lorosae, Faculty of Engineering, Science & Technology, Timor-Leste*
[2] *Gifu University, Faculty of Engineering, Japan*

Abstract

Road condition monitoring usually requires extremely expensive special vehicles, equipment, or many human resources. On the other hand, with the development of ICT and data science technologies in recent years, there are several research trials in which the heavy technical tasks of road asset condition monitoring are replaced by automatic inspection systems consisting of common devices such as smartphones and dashcam videos. As the system consists of low-price devices, it also suitable for developing countries. However, there are many differences in the situation and the inspection items on road condition monitoring between advanced countries and developing countries. There are few trials to develop such a road condition monitoring system in developing countries. Our project is developing an integrated road condition monitoring system focusing on developing countries like Timor-Leste. In developing countries, many parts of the road are still unpaved, and the “road width” is an important item to be inspected. In this paper, we discuss the road width and pothole size estimation as a part of the integrated system we are developing. We survey the road width of both paved and unpaved roads. We use a common dashcam to take video along the road. The estimated values are integrated into a database with GPS information and visualized in Google Map, QGIS, or the original visualization system which we developed. To estimate the real width of the road and pothole size, we need to transform the captured forward view image of dashcam video into bird’s-eye-view. For the transformation, we need to estimate the vanishing point in a captured image. However, unlike the advanced countries, it is difficult to detect the vanishing point in developing countries because there are usually no straight lines in the images in the unpaved road of the province. In this study, we propose to use the optical flow method to detect the vanishing point in the rural road. To identify the area of road and the existence of potholes in images, we apply state-of-the-art semantic segmentation using deep learning.

Keywords: developing countries, road condition monitoring, road width, pothole, dashcam, semantic segmentation, deep learning

1 Introduction

A huge budget is allocated for regular road asset condition monitoring to maintain its essential infrastructures in each country and local government. In advanced countries, there are several sophisticated ways of road surveys. However, they are too expensive to introduce in developing countries. In developing countries, road survey is usually achieved manually by humans, but the manual way has problems such as time-consuming, requires many human resources, individual variation of the results, a troublesome task of data organization, and so on.

By the way, the penetration rate of the smartphone is remarkably increasing even in under developing countries. The smartphone has many high-performance sensors, such as accelerometers, gyro sensors, GPS sensors, communication functions, touch panels, etc., in the small body at a low price. The utilization of the sensors is receiving increased attention in recent years. This is the central part of a stream of ICT framework for Developing Countries (ICT4D).

In our project, we are developing an automatic road surface condition survey system using smartphone sensors, a dashcam, and a server for calculation and database [1], [2], [3], [4], [5]. This approach is advantageous to under developing countries because the system consists of a common cheap smartphone and dashcam, and no expert knowledge or trained skills are required for inspectors on the field site. These types of trials have been conducted so far in advanced countries [6], [7], [8], [9], [10]. However, not many in developing countries as far as we know.

This study is conducted by the collaboration of faculty of engineering of National University of Timor-Leste and Gifu University in Japan, with the ministry of public works of Timor-Leste under financial support of Japan International Cooperation Agency (JICA). The survey items of the road surface are different between advanced and developing countries like Timor-Leste. For instance, more than 50 % of national road is still unpaved in Timor-Leste [11], [12], and we need to specify the unpaved sections and also road width of both paved and unpaved road for the national plan of road construction. The survey items of our whole project include paved and unpaved classification, road width estimation for both paved and unpaved, roughness estimation for both paved and unpaved, pothole detection and its size estimation, crack detection and its type and size estimation, markers on road surface detection, rutting estimation, roughness (IRI) estimation, etc. All of the inspection items are estimated by analysis of time series data acquired by smartphone sensors and dashcam video, using signal processing, image processing, and state-of-the-art deep learning techniques.

This paper focus on the road width estimation for both paved and unpaved especially in developing countries, and pothole detection and its size estimation from dashcam video.

2 Methods

2.1 Video recording of road using in-vehicle dashcam

We can use any common dashcams that can easily get in any country to take the video. However, there are several requirements for use in this study.

First, the dashcam must have functions to get GPS signal and embed the GPS information in the video file. Second, the angle of view of the lens must be specified. In this study, we used Transcend® DrivePro™ 230. The dashcam is attached to the front window as usual. We used full HD 1080P, 30 fps, and MOV (H.264) format as the specification of the video file. The viewing angle for the lens of DrivePro™ 230 is 130° (diagonal).

2.2 Extraction of GPS value from video and clipping of frames every 10 m

The information of date, time, and GPS is embedded almost every 1 second in the video file. In this study, we extract the information every 1 second by binary analysis and estimate the vehicle velocity from the GPS values. The estimated vehicle velocity is used to clip the frames of the video every about 10 m.

2.3 Estimation of the vanishing point in the frame by optical flow tracking

It is necessary to convert the captured image to the real scale for the estimation of the real road width and pothole size, that is, bird's-eye-view transformation (Fig. (1)).

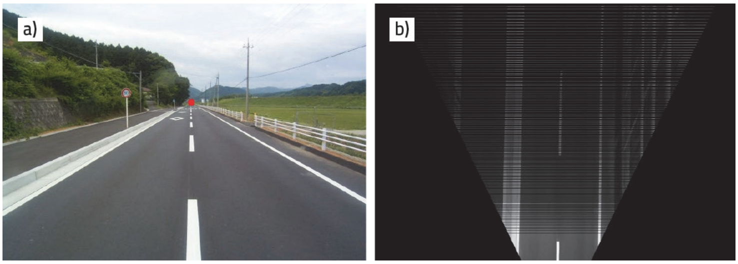

Figure 1 (a) Original dashcam image taken in Japan, (b) Transformed bird's-eye-view image

For the conversion, several parameter values are required. We can know the angles of view θ_H, θ_W from the specification of the dashcam (Fig. 2). The pixels of width and height of the video frame W, H can be set manually on the recording. We measure the height of the lens from ground level on the recording. However, we need to estimate somehow the depression angle of the optical axis of the lens of the dashcam from the videos or images taken by the dashcam. To estimate the depression angle, firstly, we need to find the vanishing point in the image.
We estimated the depression angle of dashcam using pixel sizes of the image. Let *D* be pixels from the center horizontal line in the image to the vanishing point v (Fig.2). Then, the depression angle θ_C is calculated as the following formula:

$$\theta_C \sim \sin\theta_C = \frac{2D}{H}\tan\frac{\theta_H}{2}$$

Here, we need to identify the vanishing point v. The vanishing point can be estimated as an intersection of multiple straight lines in the image (Fig. 1(a)). Usually, the straight lines in an image are detected by Hough transform. In advanced countries, there are many straight lines in the road scene, such as white center and sidelines on the road, the edges of buildings, and so on. On the other hand, it is usually challenging to find straight lines in the scene of a road in under developing countries, and we cannot use Hough transform.

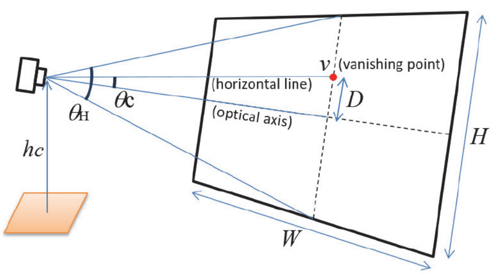

Figure 2 Schematic diagram of dashcam position and posture and its view. The vanishing point is described as a red point

In this study, we propose to apply optical flow tracking on video flames to detect the vanishing point. To reduce the influence of vibration of the vehicle on driving, we used a total of 10 seconds with 30 fps. The directions of flow of grid points are estimated using OpenCV, and we took the average of the intersections of extended lines of flow vectors as the vanishing point. The accuracy of the position of the vanishing point is evaluated by comparing it with the vanishing point estimated by Hough transform approach.

2.4 Identification of road area and pothole using semantic segmentation

To identify the road surface area and detect the pothole, we used state-of-the-art semantic segmentation by deep learning. We compare the performance of FCN [13], SegNet [14], and U-net [15] and found that SegNet and U-Net have an advantage on calculation time. On the other hand, we got a maximum IoU (Intersection over Union) value with FCN.

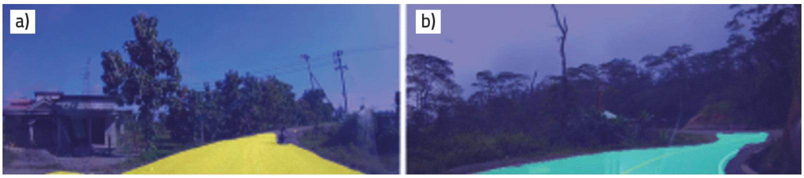

Figure 3 Results of semantic segmentation of road area using SegNet. (a) unpaved road, (b) paved road

2.5 Measurement of road width and pothole size by bird's-eye-view transformation

The estimation of road width is easier than that of pothole size. By taking extension lines with the vanishing point (red point) and each edge of road area (yellow lines), we can estimate the road width on the lower edge of flame as *L* pixels and l meter (Fig. 4).

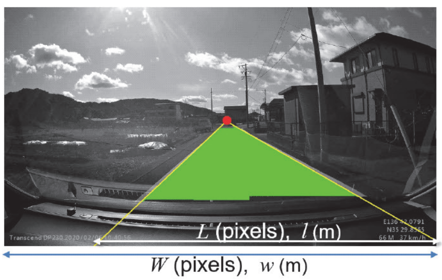

Figure 4 Estimation of the road width on the lower edge of flame

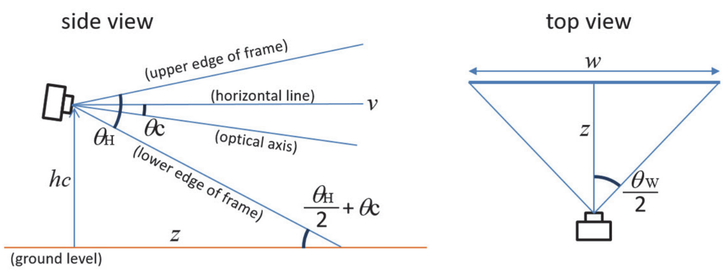

Figure 5 Calculation of the real size of lower edge of flame as meter unit

The formula to estimate the road width L (m) is as follows (Fig. 5):

$$l = w\frac{L}{W} = \frac{L}{W}2z\tan\frac{\theta_W}{2} = \frac{L}{W}2\tan\frac{\theta_W}{2} \times \frac{h_c}{\tan\left(\frac{\theta_H}{2} + \theta_c\right)}$$

To estimate the pothole size, we need to transform the front view image to the bird's-eye-view image (Fig. 1) and count pixels of the pothole area.

2.6 Storing the GPS and road width information into database

The road width information is stocked every 10m along the road with longitude and latitude values. When the pothole is detected, its size and longitude, and latitude values are also stocked. Our analyzing results are visualized on the map using Google Map, Open Street Map, or QGIS (a free and open-source cross-platform desktop geographic information system [16]).

3 Results

We took the dashcam video of more than 1,000 kilometers in Timor-Leste and a certain degree distance in Japanese for system development and comparison.
To estimate the accuracy of the optical flow method on detecting the vanishing point, we first applied the method on paved road video in urban areas in Japan (Fig. 6 (a)). Besides applying the optical flow to identify the vanishing point, we also identified the vanishing point by Hough transform using white straight lines on the road as a correct reference. As a result, the difference in the degree of the vanishing points between the optical flow method and the Hough transform method was about 0.2°. The results show that we can replace the Hough transform method with the optical flow method.

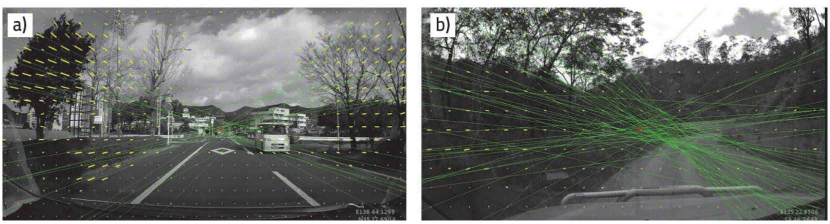

Figure 6 Optical flows of mesh points and detection of the vanishing point, a) Paved road in urban area in Japan, b) Unpaved road in rural area in Timor-Leste. The identified vanishing point is described as red point, the flows of grid points are described as yellow lines, and the extension lines of the flows are described as green lines

We applied the optical flow method on the video taken in Timor-Leste (Fig. 3 (b)). Even if there are no straight lines, we could estimate the vanishing point. The position of the estimated vanishing point does not stable in the short term because of the vibration of the vehicle on unpaved roads. We used a total of 300 flames of the video, which correspond 10 seconds with 30 fps, to remove the vibration noise.
We applied the proposed method to the dashcam video taken on a paved rural road in Japan to evaluate the road width estimation accuracy. As a result, the error of the road width estimation compare to the actual measured value was 0.25 %.

4 Conclusion

This paper proposed a method to estimate road width and pothole size from a dashcam video taken on an unpaved rural road in under developing countries. Since there are neither straight lines nor edges in the scenes of rural unpaved road, we cannot use Hough transform. We found that the optical flow method can be a good alternative. However, because of the vehicle's vibration on driving, we need to take statistics about 10 seconds driving.
Before the road width and pothole size estimation, we need to identify the area of the road surface and detect potholes. To identify the road area and detect the pothole in the image, we used state-of-the-art semantic segmentation by deep learning. Although we have evaluated the accuracy of road width estimation, the evaluation of pothole size estimation is the next task. We already have more than 1,000 km recording in Timor-Leste. We are in the process of analyzing all the data and making a database of road width and pothole size in Timor-Leste.

References

[1] Cabral, F.S., Pinto, M., Mouzinho, F.A., Fukai, H., Tamura, S.: An Automatic Survey System for Paved and Unpaved Road Classification and Road Anomaly Detection using Smartphone Sensor, IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI), pp.65-70, Singapore, 31 July–2 August 2018.

[2] Pereira, V., Tamura, S., Hayamizu, S., Fukai, H.: A deep learning-based approach for road pothole detection in Timor-Leste, IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI), pp. 1-6, Singapore, 31 July – 2 August 2018.

[3] Pereira, V., Tamura, S., Hayami, S., Fukai, H.: Classification of paved and unpaved road image using convolutional neural network for road condition inspection system, IEEE International Conference on Advanced Informatics, Concepts, Theory, and Application (ICAICTA), pp. 1–6, 2019.

[4] Cabral, F.S., Fukai, H, Tamura, S.: Feature Extraction Methods Proposed for Speech Recognition Are Effective on Road Condition Monitoring Using Smartphone Inertial Sensors, Sensors 19 (2019) 16, doi: https://doi.org/10.3390/s19163481

[5] Cabral, F.S., Mouzinho, F.A.L.N., Pereira, V., Fukai, H., Tamura, S.: The development of integrated road condition monitoring system for under developing countries using smartphone sensors and dashcam in vehicles, 6th International Conference on Road and Rail Infrastructure –CETRA 2020 (in press)

[6] Eriksson, J., Girod, L., Hull, B., Newton, R., Madden, S., Balakrishnan, H.: The pothole patrol: Using a mobile sensor network for road surface monitoring, 6th International Conference on Mobile Systems, Applications, and Services, pp. 29–39, Breckenridge, Colorado, 17–20 June 2008.

[7] Tai, Y., Chan, C., Hsu, J.Y.: Automatic road anomaly detection using smart mobile device. In Proceedings of the Conference on Technologies and Applications of Artificial Intelligence, Hsinchu, Taiwan, 18–20 November 2010.

[8] Seraj, F., Zwaag, B.J., Dilo, A., Luarasi, T., Havinga, P.: RoADS: A road pavement monitoring system for anomaly detection using smart phones, 1st International Workshop on Machine Learning for Urban Sensor Data Switzerland, pp. 128–146, 2014.

[9] Bhoraskar, R., Vankadhara, N., Raman, B., Kulkarni, P.: Traffic and road condition estimation using smartphone sensors, 4th International Conference on Communication Systems and Networks (COMSNETS 2012), pp.1-6, Bangalore, India, 3–7 January 2012.

[10] Tomiyama, K., Kawamura, A., Nakajima, S., Ishida, T., Jomoto, M.: A mobile profilometer for road surface monitoring by use of accelerometers, 7th Symposiumon Pavement Surface Characteristics: SURF 2012, Norfolk, VA, USA, 19–22 September 2012.

[11] Timor-Leste Strategic Plan 2011–2030, http://timor-leste.gov.tl/wp-content/uploads/2011/07/Timor-Leste-Strategic-Plan-2011-20301.pdf

[12] Building a Sustainable Rural Road Network in Timor-Leste, https://www.dfat.gov.au/sites/default/files/roads-for-development-ilo-ausaid-concept-note.pdf

[13] Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation, IEEE conference on computer vision and pattern recognition, pp. 3431-3440, 2015.

[14] Badrinarayanan, V., Kendall, Cipolla, R.: SegNet. A deep convolutional encoder-decoder architecture for image segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence, 39 (2017) 12, pp. 2481-2495, doi: 10.1109/TPAMI.2016.2644615

[15] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation, International Conference on Medical image computing and computer-assisted intervention, pp. 234-241, Munich, Germany, 5-9.10.2015.