AUTOMATIC TOP-VIEW TRANSFORMATION OF IN-VEHICLE SMARTPHONE CAMERA FOR ROAD CONDITION MONITORING

The University of Tokyo Regular Member ○Xue Kai
The University of Tokyo Student Member GEDA Jose Maria Guyamin
The University of Tokyo Regular Member Tomonori Nagayama

1. INTRODUCTION

A smartphone-based road condition measurement framework that utilizes cameras and GPS is a promising approach for efficient road maintenance. The camera images from their own view angles are, however, difficult to accurately evaluate the crack ratio, pothole areas and other distress indices. Because the interior and exterior parameters of cameras are not easy to obtain, an automatic process for parameter calibration and image transformation is required. In this paper, a vision based top-view transformation that does not need any prior information on camera parameters is developed by combining the parallelism condition of road lane lines and the circular characteristics of road manholes.

2. STRAIGHT LANE LINES DETECTION BY OBJECT DETECTION

In order to calibrate the camera parameters for top-view transformation, a pair of parallel straight lane lines are required. An object detection method is developed for straight lines detection. Four types of road lane lines are defined as straight left, straight right, curve left and curve right. Each class and its corresponding annotation are illustrated in the examples in Fig. 1. Approximately 3000 images of 640×640 pixels are collected as the dataset and the bounding boxes are carefully annotated. The state-of-art YOLOv5 (Glenn et al. 2021) object detector is utilized for this road straight lane lines detection.

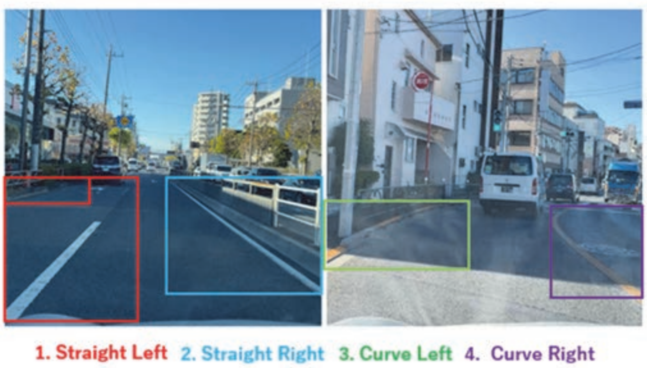

Fig. 1 Four Classes for road lane lines detection

Fig. 2 Road lane lines detected after post-processing

The result is the four coordinates of the bounding box for each detected object. Then, only the straight left and straight right classes are chosen as candidates. The straight lane lines are processed by connecting left-bottom with right-top coordinates for straight left class and right-bottom with left-top coordinates for straight right class. If multiple left (right) straight lines are detected, the end point near the left (right) bottom of the image will be selected, as shown in Fig. 2.

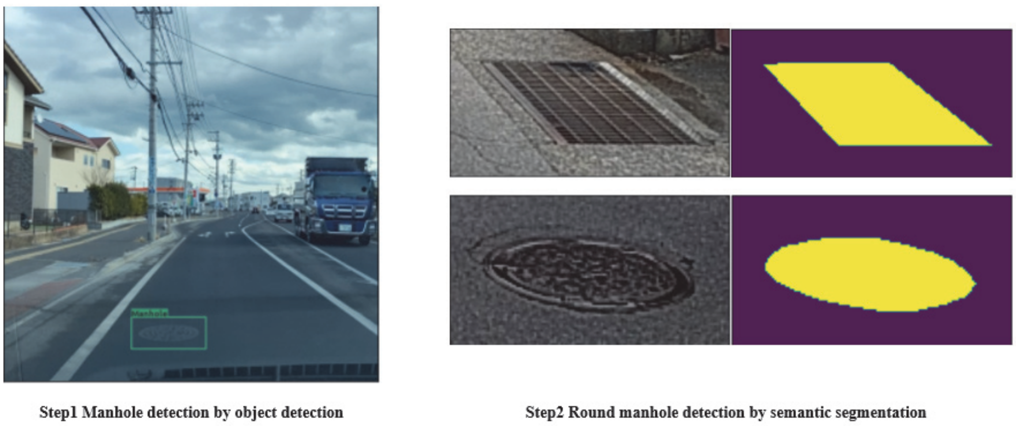

Fig. 3 Manhole Detection

Keywords: Top-view transformation, Lane detection, Crack ratio, Object detection, Semantic segmentation
Contact address: Faculty of Engineering Bldg. 17-3-1 Hongo, Bunkyo-ku, Tokyo. 〒113-8656 TEL: 03-5841-6098

© Japan Society of Civil Engineers

3. MANHOLE DETECTION BY OBJECT DETECTION AND SEMANTIC SEGMENTATION

The camera parameters should also satisfy the condition that round manhole is circular after top-view transformation. The round manhole detection is divided into two steps. In the first step, object detection method is applied to images in order to search for image with manhole. A dataset with approximately 50000 images which are collected from different countries is generated. There are 10 classes with high quality annotations in the dataset and the size of images is 640×640 pixels. A total of 16024 round and rectangular manholes are annotated. YOLOv5 is utilized to achieve the manhole detector (Fig. 3). In the second step, the manholes of bounding box area larger than 3500 $pixel^2$ are cropped from the dataset. The left number of manholes are 1702 and they are pixel-labelled into two classes of round manhole and rectangular manhole for semantic segmentation, as shown in Fig. 3. The FPN architecture with VGG19 encoder is applied as the segmentation model. During the camera parameter calibration, ellipse estimation is utilized to identify the shape of the pixel-level round manhole after semantic segmentation. The length of the major axis and the minor axis will be estimated. With the assumption that the round manhole is circular in shape, the ratio of the minor and major axis should be equal to 1.

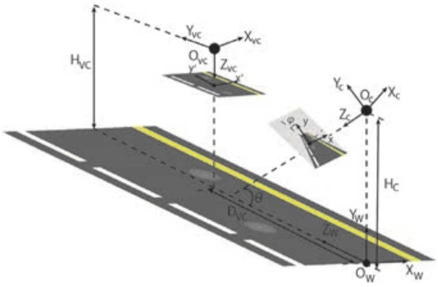

Fig. 4 Coordinate System

Fig. 5 Calibration of f and θ

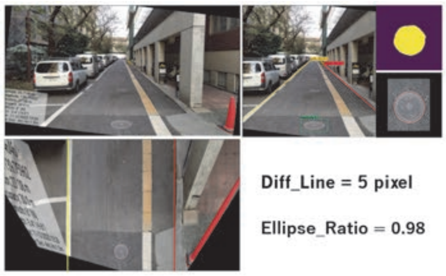

Fig. 6 Calibration of f, θ and φ

4. TOP-VIEW TRANSFORMATION

The coordinate system is shown in Fig. 4 (Tanaka et al. 2011). $O_W - X_W Y_W Z_W$ is defined as the vehicle coordinate, $O_C - X_C Y_C Z_C$ and $O_{VC} - X_{VC} Y_{VC} Z_{VC}$ are defined as the camera-view coordinate and top-view coordinate, respectively. The viewpoint O_C rotates by θ about the X_W axis and by φ about the Z_W axis. H_C and H_{vc} are the heights of in-vehicle camera and the virtual camera, respectively. D_{vc} is the horizontal distance of virtual camera to the vehicle. The top-view image is obtained by transforming the in-vehicle view image with the following equations:

$$x' = \frac{f'}{H_{VC}} \cdot \frac{H_C x}{f sin\theta - y cos\theta} \quad (1)$$

$$y' = \frac{f'}{H_{VC}} \cdot \left\{ \frac{H_C (f sin\theta - y cos\theta)}{f sin\theta - y cos\theta} - D_{VC} \right\} \quad (2)$$

where f and f' are the focal lengths of the in-vehicle camera and virtual camera, respectively. The f'/H_{VC} is the invariance used for determining the range of X_W and Z_W projected onto the top-view image. D_{vc} is determining such that Z_W is started from the bottom position of camera-view image. Therefore, the unknown camera parameters to solve the top-view transformation in Eq. (1) and Eq. (2) are f and θ. In addition, the image should rotate about Z_C axis with $-\varphi$ before the top-view transformation.

5. PARAMETER CALIBRATION

There are 3 parameters in total needing calibration, $[f, \theta, \varphi]$. These parameters are automatically calibrated using the captured image(s) when the vehicle is driving on a normal road with at least a pair of straight lane lines and a round manhole. The same parameters will be applied for all the images if they are successfully calibrated. The parameters are optimized to satisfy the parallelism condition of straight lane lines and the circular characteristics of cropped round manhole after rotate transformation and top-view transformation, as shown in Fig. 5 (only top-view transformation) and Fig. 6 (with rotate transformation).

6. CONCLUSIONS

An automatic top-view transformation method is proposed for road condition monitoring. Straight lane lines detection and manhole detection through object detection and semantic segmentation are developed for the camera parameters calibration. The parameters are optimized such that the detected lane lines are parallel and the cropped round manhole is circular after the top-view transformation. The effectiveness and robustness of the proposed methods are verified by a large number of practical road images. In the future, crack ratio and pothole area are expected to be accurately evaluated after the top-view transformation.

REFERENCES Tanaka, S., Yamada, K., Ito, T. and Ohkawa, T.: Vehicle Detection Based on Perspective Transformation using Rear-view Camera. International journal of vehicular technology, 2011. Glenn, J., Alex, S., ultralytics/yolov5: v4.0 - nn.SiLU() activations, Weights & Biases logging, PyTorch Hub integration (Version v4.0). Zenodo. 2021.

ACKNOWELDGMENT: A part of this research is supported by a research fund by Kanto Regional Development Bureau, MLIT.

© Japan Society of Civil Engineers