講演番号	122

文献番号	20226122

Direction estimation of pedestrian based on 2D poses detected on a monocular image

Akisue Kuramoto [1,2] Keisuke Yoneda [2]
Ryo Yanase [2] Naoki Suganuma [2]

1) Tokyo Institute of Technology, School of Engineering, Department of Systems and Control Engineering,
2-12-1 O-okayama, Meguro, Tokyo, 152-8552, Japan (E-mail: akisuekura@sc.e.titech.ac.jp)
2) Kanazawa University.
Kakuma-cho, Kanazawa, Ishikawa, 920-1192, Japan

KEY WORDS: Pedestrian direction, Monocular camera, Multilayer perceptron, Image processing, Deep neural network (C1)

Recognizing the direction of movement of pedestrians on footpaths and pedestrian crossings is an important factor in determining the safe response of vehicles in urban areas. This study aims to enable recognition of the direction of movement of pedestrians around the vehicle using cameras, which are relatively inexpensive and easily available among the various sensors used to realize preventive safety functions in vehicles. This paper proposes a method for recognizing the orientation of pedestrians in images from only a single frame by training using images captured by an in-vehicle monocular camera while driving in urban areas and test facilities. The accuracy of the proposed method was tested using the detection rate of the 2D posture of the human body using OpenPose and the classification accuracy of the pedestrian's direction.

Figure 1 shows an overview of the procedure of the proposed method. First, an image of a pedestrian is extracted from an image captured by an in-vehicle camera. In this paper, we used YOLOv4, a popular method for object detection from images. Next, the 2D pose data of the pedestrian is obtained from the extracted images. We used OpenPose as the 2D pose detector. The 2D pose data is then passed to a classifier composed of MLP (Multilayer Perceptron), which classifies the orientation of the pedestrian into the eight categories composed of Right, Diagonally forward right, Forward, Diagonally forward left, Left, Facing diagonally left, Facing and Facing diagonally right. This pedestrian orientation classification follows the TUD Multiview dataset.

Fig. 1 Overview of the proposed method.

A dataset of pedestrian images and orientations was prepared from videos taken by three monocular cameras with different viewing angles boarded on an experimental vehicle driving through urban areas and experimental facilities. The pedestrian images were cropped from human body regions detected by YOLOv4. For each cropped image, the pedestrian direction was annotated based on visual inspection. Among the prepared datasets, images taken before November 2021 were used as training data for the MLP. The remaining images and the TUD Multiview dataset were used to validate the trained MLP.

Validation datasets were classified by image width to validate the recognition accuracy of the MLP in a small pedestrian. The categorical accuracy results are shown in Figure 4. The MLP correctly recognized the direction of pedestrians with 90.9% accuracy for pedestrian images with a width less than 40 px in the dataset we prepared. The accuracy was slightly lower (66.7%) in the range of 40 px to 60 px widths but achieved an accuracy of about 80% in the larger images. Moreover, in the TUD Multiview dataset, the MLP achieved an accuracy rate of about 90% for the images of 40 px to 60 px width and around 95% for the larger sizes. Pedestrian images with a width of less than 40 px were not found and could not be evaluated. These results show that the proposed method can recognize the direction of a pedestrian not only near the vehicle but also in the distance where her/his size in the image becomes almost the minimum detectable size.

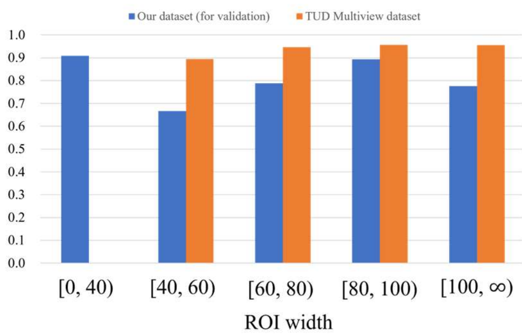

Fig. 2 Recognition accuracy (categorical accuracy) by size of ROI width