

Received March 26, 2022, accepted May 21, 2022, date of publication May 26, 2022, date of current version June 1, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3178154

Deep-Learning-Based Stair Detection Using 3D Point Cloud Data for Preventing Walking Accidents of the Visually Impaired

HARUKA MATSUMURA¹ AND CHINTHAKA PREMACHANDRA², (Senior Member, IEEE)

¹Department of Electronics Engineering, School of Engineering, Shibaura Institute of Technology, Tokyo 135-8548, Japan

²Graduate School of Engineering and Science, Shibaura Institute of Technology, Tokyo 135-8548, Japan

Corresponding author: Chinthaka Premachandra (chintaka@sic.shibaura-it.ac.jp)

This work was supported in part by the Branding Research Fund of the Shibaura Institute of Technology.

ABSTRACT Visually impaired individuals worldwide are at a risk of accidents while walking. In particular, falling from a raised place, such as stairs, can lead to serious injury. Therefore, we attempted to determine the best accident prevention method that can notify visually impaired individuals of the existence, height, and step information when they approach stairs. In this study, we have investigated stair detection through deep learning. First, the three-dimensional point cloud data generated from depth information are learned by deep learning. Stairs were detected using the results of deep learning. To apply the point cloud data for deep learning-based training, we proposed preprocessing stages to reduce the weight of the point cloud data. The accuracy of stair detection was 97.3%, which is the best performance compared to other conventional methods. Therefore, we confirmed the effectiveness of the proposed method.

INDEX TERMS Visually impaired support systems, depth sensor, 3D point cloud data, deep-learning, PointNet.

I. INTRODUCTION

There are visually impaired people in every country in the world. According to World Health Organization (WHO) [1], as of 2019, more than 2.2 billion people in the world have various visual impairments ranging from visual impairment to blindness. In Japan, the number of visually impaired people in 2016 was 312,000 [2]. Furthermore, according to a study by a research team at Anglia Ruskin University in the UK [3], the number of visually impaired people worldwide will increase unless current treatment methods are improved. The number of partially sighted people is expected to increase from 36 million in 2015 to 115 million in 2050, and the number of people with moderate-to-severe visual impairment is expected to increase from 216.6 million in 2015 to 550 million in 2050. They also reported that even moderate visual impairment can have a substantial impact on people's lives. Therefore, it is important to provide support to the visually impaired. The visually impaired are assisted in various ways.

The associate editor coordinating the review of this manuscript and approving it for publication was Wu-Shiung Feng.

In Japan, Article 14 of the Road Traffic Law stipulates that persons with visual impairment (including people equivalent to visually impaired persons) must carry a cane or having a guide dog specified by Cabinet Orders when crossing the road." Thus, carrying a white cane or having a guide dog is common because it is specified by the government ordinance. According to the International Guide Dog Federation [4], there are 20,000 guide dogs in 31 countries, but this is not sufficient for the number of visually impaired people. In contrast, white canes are readily available, but it takes time to get used to them, and walking training is required to move safely with them [5].

However, several visually impaired people are reluctant to carry a white cane owing to various concerns, such as stares from people around them and the inconvenience to their families. According to a survey that analyzed the behavior of visually impaired people [6], 76.6% of visually impaired people walk alone without the assistance of a caregiver or advanced technology. According to a Japanese survey, "Walking Accident National Survey to Maintain Visually Handicapped Persons Walking Environment" [7], 47% of the people who walk alone have experienced a walking accident. Therefore, the risk of accidents while walking is

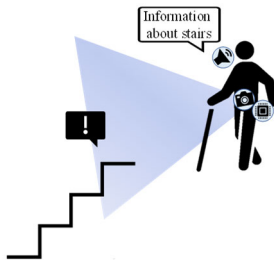


FIGURE 1. Operation image of this system.

high. In particular, a fall from a raised place, such as a station platform or stairs, can lead to serious injuries. Thus, according to the survey, stairs are one of the most dangerous places for visually impaired people and the most likely place for them to fall outdoors. In addition, it is the third most dangerous indoor location. From the aforementioned information, it can be observed that stairs are very dangerous for the visually impaired.

To prevent accidents, we developed a wearable system that notifies visually impaired people of the existence of stairs as they approach them, and also informs them of the height of the stairs and the difference in steps. The operation of this system is illustrated in Fig. 1. In this study, we focused on the detection of stairs.

Conventional research exists in stair detection, for example, the Hough transform using grayscale images [8] and the RGB image [9]. Conventional studies on stair detection include methods based on two-dimensional (2D) images [10]–[13]. In addition, methods have been proposed to detect stairs using three-dimensional (3D) data, such as methods using RGB-D images [14]–[16] and methods using 3D data with stereo images [17]. In several cases, these methods do not distinguish between upward and downward stairs or detect stairs. Therefore, we believe that we can detect stairs with a higher probability than that achieved using conventional methods by acquiring information regarding the steps in stairs from the depth information captured by the depth camera and analyzing the information using deep learning. In addition, because each pixel of the depth camera has distance information instead of color information, it can be applied to stair detection in dark environments such as nighttime or during power outages due to disasters.

In this study, we pursued the detection of stairs using deep learning. Specifically, we pre-processed the depth information of the stairs and non-stairs places from the depth camera attached to the user, and trained it using deep learning. The system then uses the learning results to detect stairs.

We used the RealSense depth camera D435i (RealSense) [18] because it is small, easy to wear, and can acquire depth information in addition to 2D images.

In this study, we confirm the effectiveness of the proposed method through several verification experiments. The results reveal that the proposed method can detect stairs with a very high accuracy rate of 97.3%, exhibiting the best performance compared to other conventional methods.

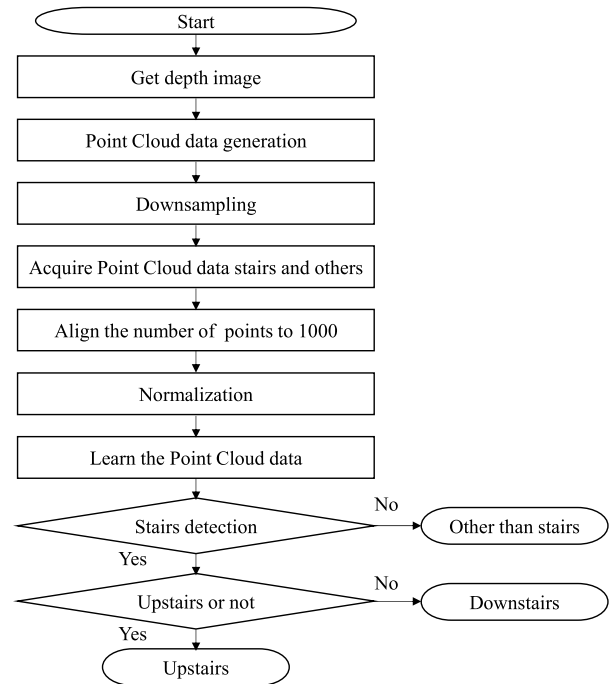


FIGURE 2. Flow of the proposed method.

II. OVERVIEW OF THE PROPOSED SYSTEM AND EQUIPMENT USED

A. DEVELOPMENT ENVIRONMENT

The programming language used in this study was Python, which has extensive libraries for image processing and AI. For deep learning using 3D point cloud data, we used PointNet [19], a deep learning model that can easily handle point cloud data.

B. OVERALL FLOW OF THE PROPOSED STAIRS DETECTION METHOD

The overall flow of the proposed stair-detection method is illustrated in Fig. 2. We generated 3D point cloud data from depth information. Using the point cloud data, we performed stair detection using deep learning. In this study, we used PointNet, a deep learning model that can easily handle point cloud data, to perform class classification and segmentation using point clouds as inputs. Therefore, we used a class classification method to detect and classify stairs into three types, namely, downstairs, upstairs, and other than stairs, and trained and verified the classification.

C. DEPTH CAMERA USED IN THIS STUDY

A depth camera can detect the distance between the camera and an object using a sensor, and each pixel of the image represents that distance. In this study, we used the RealSense depth camera D435i (RealSense), which can be used indoors and outdoors as a depth camera. The equipment used is shown in Fig. 3. This depth camera can output depth information in addition to 2D information and is small and easy to wear. In recent years, RealSense has been used to build



FIGURE 3. RealSense depth camera D435i.

TABLE 1. The specifications of Realsense d435i.

Use environment	Indoor/Outdoor
Image sensor technology	Global Shutter
Depth Field of View (FOV) (Horizontal × Vertical × Diagonal)	87°±3°×58°±1°×95°±3°
Depth output resolution	Up to 1280×720
Depth frame rate	Up to 90 fps
Minimum depth distance (Min-Z)	0.105m
Ideal range	.3m to 3m
Length × Depth × Height	90mm×25mm×25mm

support systems for the physically challenged [20], [21]. The specifications of the depth camera are listed in Table 1.

III. STAIRS DETECTION USING DEEP LEARNING OF 3D POINT CLOUDS

A. PREPARING THE DATA SET

In this study, we first generated 3D point-cloud data based on the depth information captured by a depth camera. Thereafter, considering the reduction in processing time, we performed downsampling [22] on the number of points in each point cloud data sample to prepare a lightweight 3D point cloud dataset.

We prepared 1000 training and 500 validation datasets for each of the above classes and conducted experiments with 3000 training and 1500 validation datasets. Fig. 4 shows the sample 3D point cloud data stairs with their 2D image (Fig. 4 (a)). Fig. 4 (b) shows an example of depth data from the Realsense while Fig. 4 (c) shows extraction of the approximate stair region by Open 3D. Fig. 4 (d) shows the down-sampled results of the depth image in Fig. 4(c). The down-sampling process is explained in the next sub-section.

B. DOWNSAMPLING

In this study, to reduce the processing time, down-sampling was performed on the number of point clouds of each point cloud data sample to make the 3D point cloud data lightweight. Down-sampling refers to thinning of the soil. Without down-sampling, the 3D point cloud data are too detailed and the data size is large, which is very large when

collecting a large amount of 3D point cloud data for deep learning. Down-sampling has an impact on the processing time. For example, it may take a long time to read the point cloud data (pcd) file and process 3D point cloud data. Therefore, down-sampling is performed. This reduces the number of point clouds and accelerates the subsequent processing time. In the down-sampling of 3D point cloud data, a new point is placed at the center of the points within a certain range, and all other points are deleted to thin out the data. Because the points are thinned out at equal intervals, the overall structure can be relatively preserved.

C. NORMALIZATION OF POINT CLOUD DATA

When captured by a depth camera, nearby objects have smaller values than distant objects. These differences may prevent deep learning from working well and may require more training time. Therefore, as pre-processing of the data, we perform normalization [23] based on the vector distance that is the largest from the origin in the sample. Normalization is the process of transforming the range of values of a feature such that they fall within a certain range. The formula for normalizing the original data is expressed by equation (1) as follows:

$$x_{norm} \mathcal{D} \frac{x_i - \sigma}{\max\{|x_i - \sigma|\}} \quad (1)$$

D. DEEP LEARNING FOR 3D POINT CLOUD DATA

3D point cloud data is a method of describing 3D shapes expressed as a set of 3D points (x, y, and z). 3D point cloud data has two important properties that must be considered when handling it in deep learning; order and translation invariances [24].

First, let us discuss order invariance. It is the property that the output is invariant even if the order of the points is changed and input into the model. Because point cloud data does not have a fixed format and the order of points cannot be assigned to each element, the order of input to the model is arbitrary. Therefore, for a point cloud of N points, there will be N! different inputs, but the object represented by the point cloud will be similar, even if the order of the inputs changes. Therefore, a deep learning model is required to output the same value each time for different permutations of point cloud inputs. Next, we discuss translation invariance. Translation invariance is a property in which the output is invariant, even if point cloud data are input to a deep learning model under parallel or rotational translation. First, the invariance to translation is expressed by (2) as follow:

$$f(x_1 + \mathbf{r}, x_2 + \mathbf{r}, \dots, x_M + \mathbf{r}) = f(x_1, x_2, \dots, x_M) \quad (2)$$

Here, x_M denotes the input and \mathbf{r} denotes an arbitrary vector. This equation demonstrates that moving the input x_M by an arbitrary vector \mathbf{r} will not change the output. Next, the invariance to the rotational movement is expressed as follows:

$$f(Rx_1, Rx_2, \dots, Rx_M) = f(x_1, x_2, \dots, x_M) \quad (3)$$

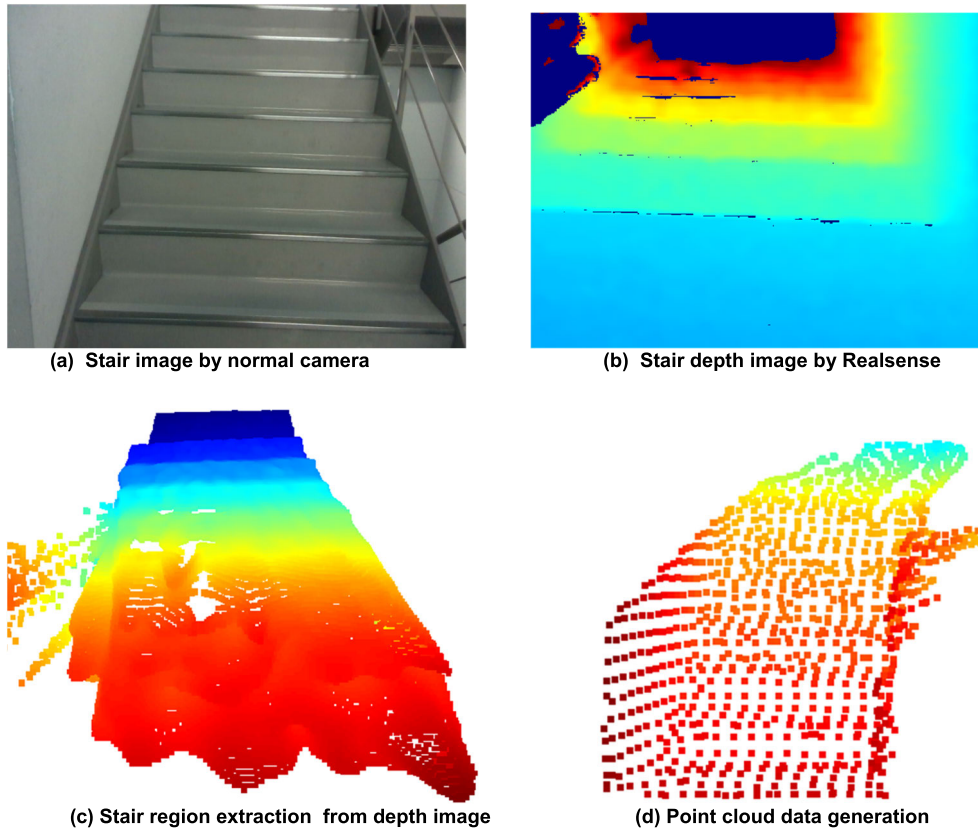


FIGURE 4. 3D point cloud data samples and their appearance as a 2D image.

where R denotes the rotation matrix. This equation demonstrates that the output is invariant even if the input data are rotated and shifted by an arbitrary rotation matrix R . Point clouds are rotating and moving, and not all point cloud data will stay in the same position without rotating. Therefore, the deep learning model must be able to maintain the output, even when the point cloud data are transformed by translation or rotation.

E. POINTNET

PointNet is a deep-learning model that considers the order and movement invariances described above [19]. In conventional 3D convolutional neural networks, point clouds are voxelized and one layer is treated as an image that is used as the input. On the other hand, PointNet accepts point clouds as input, which facilitates the handling of point cloud data and solves the shortcomings of conventional methods.

This section describes how PointNet considers the two points of order and movement invariances, as described above. A symmetric function is a function whose value does not change even if the order of the variables is changed [19]. PointNet obtains order invariance by using a symmetric function called MaxPooling, which outputs the largest element among the input elements. In other words, even if the input elements of MaxPooling are replaced, the

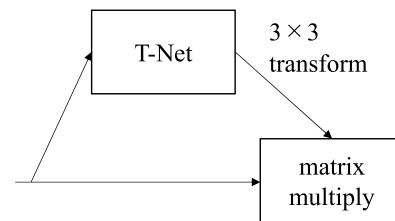


FIGURE 5. Network structures.

output will be the same as the output before the replacement, because the function outputs the largest element.

Next, we describe the movement invariance of PointNet, which estimates the affine transformation matrix of the input point cloud and multiplies it by the transformation matrix to obtain approximate movement invariance. The structure of this network is illustrated in Fig. 5. The affine transformation matrix is a transformation of rotation, translation, and scaling, and can be represented by a single 3×3 matrix. The affine transformation matrix is estimated using T-Net [25], and by multiplying the input point cloud by this estimated matrix, the output does not change even if the point cloud data are transformed by translation or rotation. Here, T-Net is a network consisting of feature extraction, max-pooling, and total joins.

We describe the flow of the PointNet classifications. The structure of this network is illustrated in Fig. 6.

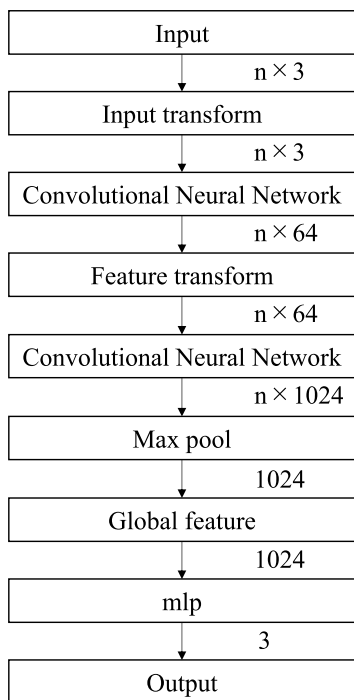


FIGURE 6. Network structures of PointNet.

Here, n denotes the number of points. Because we are dealing with 3D point cloud data, the input data are $n \times 3$. mlp is the multilayer perceptron. First, we input the input data to the transform layer. The structure of this layer is illustrated in Fig. 5. This structure allowed us to approximate the movement invariance of the input data in the transform layer. Next, we used a convolutional neural network. By repeating these steps, we obtained the feature values for the points. MaxPooling was then performed on the resulting values to obtain the order invariance, and the features of the entire 1024-dimensional point cloud were obtained. Finally, by passing the features through mlp, the classification scores of the three classes were obtained.

IV. VERIFICATION EXPERIMENT

A. EXPERIMENTAL ENVIRONMENT

In this study, a depth camera (RealSense) was attached to the waist position of the subject to obtain data. Fig. 7 shows a scene captured on the stairs. In other cases, the camera was placed at waist level in a room with obstacles, and depth data were taken from various angles. We used a computer with the specifications listed in Table 2 to perform the learning process.

B. EXPERIMENTAL RESULTS

The confusion matrix predicted from the validation data when the number of training epochs was 10 is shown in Table 3. In this confusion matrix, the vertical and horizontal axes represent the correct and predicted labels, respectively. Each

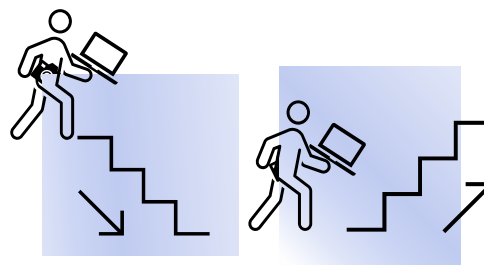


FIGURE 7. Shooting environment.

TABLE 2. Computer specifications.

CPU	Intel Core i5-7500 @3.40GHz
GPU	Intel(R) HD Graphics 520
RAM	16GB

TABLE 3. Confusion matrix (3-class classification).

		Predicted class		
		Downstairs	Others	Upstairs
True class	Downstairs	500	0	0
	Others	0	460	40
	Upstairs	0	0	500

TABLE 4. Confusion matrix (2-class classification).

		Predicted class	
		Stairs	Others
True class	Stairs	1000	0
	Others	40	460

element of the matrix represents the predicted number of samples for correct labels. The confusion matrix illustrates that all downstairs and upstairs are detected correctly, but various non-stair places are detected as upstairs.

The confusion matrix in Table 3 is summarized in Table 4, where the upstairs and downstairs are combined as one element, and two classes of non-stairs are considered. Accuracy, precision, and recall levels are described. Here, each element of the confusion matrix of the two-class classification is a true positive (TP), true negative (TN), false positive (FP), or false negative (FN).

First, accuracy refers to the percentage of correct answers to all predictions and is calculated as follows.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{4}$$

Thereafter, the rate of fit refers the percentage of data that is actually positive among the data predicted to be positive and is calculated as expressed by equation (5).

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

Finally, recall refers the proportion of predicted positive values among the actual positive values and is calculated as

TABLE 5. Performance evaluation.

Accuracy [%]	Precision [%]	Recall [%]
97.3	96.2	100

TABLE 6. Comparison with conventional methods.

Method	Accuracy [%]
Conventional method ①[9]	80.81
Conventional method ②[11]	89.74
Conventional method ③[15]	92.7
Conventional method ④[16]	90.2
Proposed method	97.3



FIGURE 8. A sample 2D image showing the direction of the shot.

expressed by equation (6).

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

Table 5 summarizes the results of accuracy, precision, and recall. A comparison of the accuracies of the conventional methods for stair detection is presented in Table 6. Conventional methods ① [9] and ② [11] are based on 2D images, whereas methods ③ [15] and ④ [16] are based on RGB-D images.

In this study, we acquired the depth information of stairs and non-stairs from various directions. Fig. 8 shows a sample 2D image that shows an example of the orientation of each scene. Fig. 9 shows a sample of the results of randomly selecting and estimating the 3D point cloud data generated from depth information acquired from various directions. In Fig. 9, label indicates the correct label, and Pred indicates the estimated label.

C. DISCUSSIONS

In this study, we used the class classification of PointNet, one of the deep learning model, to detect stairs from 3D point cloud data generated using depth information captured by

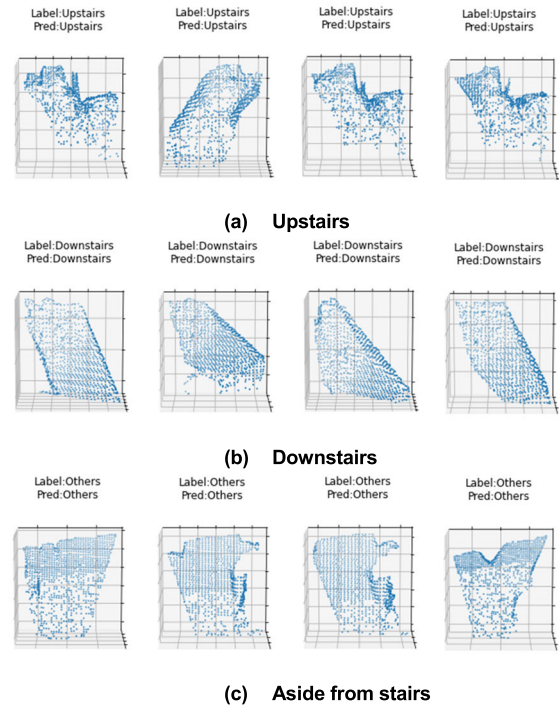


FIGURE 9. Sample estimation results for validation data.

a depth camera with a high probability of 97.3% accuracy, 96.2% precision, and 100% recall.

In addition, Table 6 summarizes a comparison of the correctness rate between our method and conventional methods, which demonstrates that our method can detect stairs with a higher accuracy rate than conventional methods. In addition, the samples of the estimation results for the validation data in Fig. 9 indicate that the estimation results are correct.

Therefore, the effectiveness of the proposed method is considered very high. However, Table 3 summarizes that all the downstairs and upstairs were detected correctly, but some of the non-stairs were detected as upstairs. The reason for this is that some of the data other than stairs include obstacles, and some of them were detected upstairs. Therefore, it is necessary to add preprocessing steps, such as noise processing, to eliminate this false detection to increase the correct answer rate in the future. The average learning time was 842.406s, and the average estimation time was 2.13s on a computer with the configuration mentioned in Table 2.

V. CONCLUSION

In this study, stairs detection was realized using point cloud data captured by a depth camera by applying deep learning. To use the point cloud data for deep learning model training, the point cloud had to be made lighter through processes such as downsampling and max pooling. We confirmed the effectiveness of the proposed method through several verification experiments by preparing the relevant depth data. The results reveal that the proposed method can detect stairs

with a very high accuracy rate of 97.3% and exhibited the best performance compared to other conventional methods.

REFERENCES

- [1] *Fact Sheet Blindness and Vision Impairment*, World Health Org., Geneva, Switzerland, 2019.
- [2] *Results of the 2016 Survey on Difficulties in Daily Life (National Survey on Children and Persons With Disabilities at Home)*, Department Health Welfare Persons Disabilities, Social Welfare War Victims Relief Bureau, Ministry Health, Labour Welfare, 2018.
- [3] R. R. A. Bourne, S. R. Flaxman, T. Braithwaite, M. V. Cicinelli, A. Das, and J. B. Jonas, "Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: A systematic review and meta-analysis," *Lancet Global Health*, vol. 5, no. 9, pp. 888–897, Sep. 2017.
- [4] (2020). International Guide Dog Federation. *Guide Dogs Worldwide*. [Online]. Available: <https://www.guidedogs.org.U.K./about%20us/what%20do%20dog%20guide%20dogs%20worldwide>
- [5] J. Ballemans, G. I. Kempen, and G. R. Zijlstra, "Orientation and mobility training for partially-sighted older adults using an identification cane: A systematic review," *Clin. Rehabil.*, vol. 25, no. 10, pp. 880–891, Oct. 2011.
- [6] A. Nobuyuki and H. Norihisa, "Walking accident national survey to maintain visually handicapped persons walking environment," *Bull. Hachinohe Inst. Technol.*, vol. 24, pp. 81–92, Feb. 2005.
- [7] N. Abe and N. Hashimoto, "Walking accident national survey to maintain visually handicapped persons walking environment," *Bull. Hachinohe Inst. Technol.*, vol. 24, pp. 81–92, Feb. 2004.
- [8] N. Molton, S. Se, M. Brady, D. Lee, and P. Probert, "Robotic sensing for the partially sighted," *Robot. Auto. Syst.*, vol. 26, nos. 2–3, pp. 185–201, Feb. 1999.
- [9] U. Patil, A. Gujarathi, A. Kulkarni, A. Jain, L. Malke, R. Tekade, K. Paigwar, and P. Chaturvedi, "Deep learning based stair detection and statistical image filtering for autonomous stair climbing," in *Proc. 3rd IEEE Int. Conf. Robotic Comput. (IRC)*, Feb. 2019, pp. 159–166.
- [10] S. Carbonara and C. Guaragnella, "Efficient stairs detection algorithm assisted navigation for vision impaired people," in *Proc. IEEE Int. Symp. Innov. Intell. Syst. Appl. (INISTA)*, Jun. 2014, pp. 313–318.
- [11] A. Ramteke, B. Parabattina, and P. K. Das, "A neural network based technique for staircase detection using smart phone images," in *Proc. 6th Int. Conf. Wireless Commun., Signal Process. Netw. (WiSPNET)*, Mar. 2021, pp. 374–379.
- [12] E. Mihankhah, A. Kalantari, E. Aboosaedan, H. D. Taghirad, S. Ali, and A. Moosavian, "Autonomous staircase detection and stair climbing for a tracked mobile robot using fuzzy controller," in *Proc. IEEE Int. Conf. Robot. Biomimetics*, Feb. 2009, pp. 1980–1985.
- [13] C. Zhong, Y. Zhuang, and W. Wang, "Stairway detection using Gabor filter and FPPG," in *Proc. Int. Conf. Soft Comput. Pattern Recognit. (SoCPar)*, Oct. 2011, pp. 578–582.
- [14] S. Murakami, M. Shimakawa, K. Kivota, and T. Kato, "Study on stairs detection using RGB-depth images," in *Proc. Joint 7th Int. Conf. Soft Comput. Intell. Syst. (SCIS) 15th Int. Symp. Adv. Intell. Syst. (ISIS)*, Dec. 2014, pp. 699–702.
- [15] R. Munoz, X. Rong, and Y. Tian, "Depth-aware indoor staircase detection and recognition for the visually impaired," in *Proc. IEEE Int. Conf. Multimedia Expo Workshops (ICMEW)*, Sep. 2016, pp. 1–6.
- [16] S. Wang, H. Pan, C. Zhang, and Y. Tian, "RGB-D image-based detection of stairs, pedestrian crosswalks and traffic signs," *J. Vis. Commun. Image Represent.*, vol. 25, no. 2, pp. 263–272, Feb. 2014.
- [17] M. Hayami and M. Hild, "Detection of stairs using stereo images as a walking aid for visually impaired persons," in *Proc. Conf. Inf. Process. Soc. Japan*, 2010.
- [18] (2021). Intel RealSense Technology. *Intel RealSense Camera D400 Series Product Family Datasheet*. [Online]. Available: <https://www.intelrealsense.com/wp-content/uploads/2020/06/Intel-RealSense-D400-Series-Datasheet-June-2020.pdf>
- [19] R. Q. Charles, H. Su, M. Kaichun, and L. J. Guibas, "PointNet: Deep learning on point sets for 3D classification and segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 652–660.
- [20] Y. Endo and C. Premachandra, "Development of a bathing accident monitoring system using a depth sensor," *IEEE Sensors Lett.*, vol. 6, no. 2, pp. 1–4, Feb. 2022.
- [21] Y. Ito, C. Premachandra, S. Sumathipala, H. W. H. Premachandra, and B. S. Sudantha, "Tactile paving detection by dynamic thresholding based on HSV space analysis for developing a walking support system," *IEEE Access*, vol. 9, pp. 20358–20367, 2021.
- [22] E. Nezhadarya, E. Taghavi, R. Razani, B. Liu, and J. Luo, "Adaptive hierarchical down-sampling for point cloud classification," in *Proc. CVPR*, Jun. 2020, pp. 12956–12964.
- [23] Z. Yang, Y. Sun, S. Liu, X. Qi, and J. Jia, "CN: Channel normalization for point cloud recognition," in *Proc. ECCV*, 2020, pp. 600–616.
- [24] Y. Liu, C. Wang, Z. Song, and M. Wang, "Efficient global point cloud registration by matching rotation invariant features through translation search," in *Proc. ECCV*, 2018, pp. 448–463.
- [25] J. Kossaiji, A. Bulat, G. Tzimiropoulos, and M. Pantic, "T-Net: Parametrizing fully convolutional nets with a single high-order tensor," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 7822–7831.



HARUKA MATSUMURA received the B.S. degree in electronic engineering from the Shibaura Institute of Technology, Tokyo, Japan, in 2022.

Her research interests include depth image processing, visually impaired support systems, and 3D vision.



CHINTHAKA PREMACHANDRA (Senior Member, IEEE) was born in Sri Lanka. He received the B.Sc. and M.Sc. degrees from Mie University, Tsu, Japan, in 2006 and 2008, respectively, and the Ph.D. degree from Nagoya University, Nagoya, Japan, in 2011.

From 2012 to 2015, he was an Assistant Professor with the Department of Electrical Engineering, Faculty of Engineering, Tokyo University of Science, Tokyo, Japan. From 2016 to 2017, he was an Assistant Professor. From 2018 to 2022, he was an Associate Professor with the Department of Electronic Engineering, School of Engineering, Shibaura Institute of Technology, Tokyo. In 2022, he was promoted to a Professor with the Department of Electronic Engineering, Graduate School of Engineering, Shibaura Institute of Technology, where he is currently the Manager of the Image Processing and Robotic Laboratory. His research interests include AI, UAV, image processing, audio processing, intelligent transport systems (ITS), and mobile robotics.

Dr. Premachandra is a member of IEICE, Japan; SICE, Japan; and SOFT, Japan. He received the FIT Best Paper Award and the FIT Young Researchers Award from IEICE and IPSJ, Japan, in 2009 and 2010, respectively. He was a recipient of the IEEE Japan Medal, in 2022. He has served many international conferences and journals as a steering committee member and an editor, respectively. He is the Founding Chair of the International Conference on Image Processing and Robotics (ICIPRoB) which is technically co-sponsored by the IEEE.

• • •