The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

Evaluating Road Surface Condition by using Wheelchair Driving Data and Positional Information based Weakly Supervision

Takumi Watanabe[*1] Hiroki Takahashi[*1] Yusuke Iwasawa[*2] Yutaka Matsuo[*2] Ikuko E. Yairi[*1]

[*1] Graduate School of Science and Engineering, Sophia University
[*2] Graduate School of Technology Management for Innovation, The University of Tokyo

Providing accessibility information on sidewalks for mobility impaired people is an important social issue. Until now, the authors have evaluated the accessibility of sidewalks by estimating the road surface condition by supervised learning on the accelerometer data mounted on wheelchairs. Video recording and data labeling to accelerometer data based on the video for teacher data require enormous costs and become problematic. This paper proposed and evaluated a novel weakly supervised road surface condition evaluation system of using positional information automatically acquired at driving as a label. The evaluation result showed that weakly supervised learning method using locational label captured detailed features of road surfaces, and classified moving on slopes, curb climbing, moving on tactile indicators, and others with a mean F-score of 0.57 and accuracy of 0.71 close to those of supervised learning method.

1. Introduction

Providing accessibility information on sidewalks for mobility impaired people, such as elderly people and wheelchair users, is one of the important social issues. The conventional methods for gathering accessibility information are as follows: a system that experts evaluate images of sidewalks for each case [Ponsard 06], a crowdsourcing method to recruit volunteers to take pictures of sidewalks and evaluate them [Hara 14, Cardonha 13]. All these methods are based on human power and thus gathering large-scale accessibility information is difficult. Because of the recent expansion of intelligent gadgets, such as smartphones and wristwatch-shaped vital sensors, there is a growing movement of sensing human activities [Swan 13, Nagamine 15]. The authors have been proposing a system which evaluates road surface condition by machine learning using accelerometer data. This system focuses on the fact that the observed values of the accelerometer mounted on a wheelchair is influenced by the condition of the road surfaces. In machine learning, however, video recording and video-based labeling to acceleration data for teacher data require a huge cost and become a serious problem. In various machine learning fields, weakly supervised learning [Zhou 18] methods that do not require conventional detailed teacher labels have been proposed [Oquab 15, Gidaris 18, You 18]. In this paper, the authors propose and evaluate the road surface condition evaluation system by weakly supervised learning which uses positional information labels which can be automatically acquired at the time of driving and thus does not require conventional detailed labels. Our contributions are as follows: we propose a novel method of weakly supervised learning of extracting feature representations of the road surface condition from accelerometer data without conventional detailed labels; we verify the effectiveness of our method using actual data.

Contact: Takumi Watanabe, Graduate School of Science and Engineering, Sophia University, Tokyo, 102-8554 Japan, 03-3238-3280, watanabe@yairilab.net

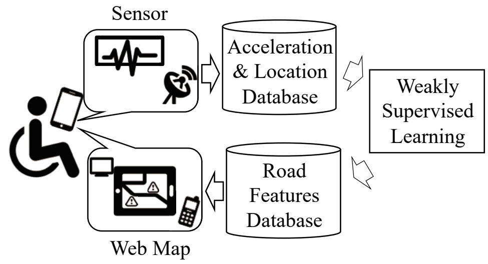

Figure 1 Outline drawing of road surface condition evaluation system by wheelchair sensing and weak supervision.

2. Road surface condition evaluation system

Figure 1 shows an outline of the proposed system. Vibration waveforms of wheelchair movement are collected by an accelerometer mounted on a wheelchair. Extracting road surface information from vibration waveforms using machine learning, the extracted data is accumulated and visualized on a web map. Extracting influence of the road surface condition from the raw accelerometer data is not easy [Lara 12, Liu 17]. Therefore, it is important to convert observed accelerometer data to indexes which represent the condition of the road surface. Some methods for expressing the road surface condition in several discrete classes by creating acceleration data classifiers using machine learning have been proposed [Iwasawa 12, Iwasawa15, Iwasawa 16], and a method for acquiring more detailed road surface features than applied several discrete labels by using feature values extracted from pre-trained DCNN is proposed [Takahashi 18]. However, these methods depend on detailed road surface condition labels and require enormous costs.

3. Road surface condition evaluation by weakly supervised learning

Dataset

The total of nine wheelchair users, including six manual wheelchair users(M1~M6) and three Powered wheelchair users(P1~P3), participated in the experiment. Their actions while driving about 1.4 km of a specified route around Yotsuya station in Tokyo were measured by an accelerometer (iPod touch) mounted on the lower part of the wheelchair seat, and positioning data of Quasi-Zenith Satellite System (QZSS) was measured at the same time. In order to confirm the situation where the acceleration data sample was acquired, the video of the participant's driving state and the driving road surface condition were taken during the experiment. Acceleration values in the x, y, and z axes of the accelerometer were sampled at 50 Hz, and the total of 1,341,142 samples (about 8 hours) was obtained.

Weakly supervised label

For the training of DCNN, positional information was used as a weakly supervised label as a method of weakly supervised learning. For the positional information in this paper, we checked the location where the accelerometer data was measured by visual observation of the recorded video and used the GPS data (latitude, longitude) acquired on Google Map website as positional information. In assigning labels to the acceleration data, all the sidewalks traveled at the time of the experiment were divided into meshes of an uniform width, and a grid belonging at the time of the measurement of the acceleration data was assigned as a weakly supervised label (as shown in Figure 2). Labels were generated under the conditions of a grid width of 3 m, 4 m, and 5 m.

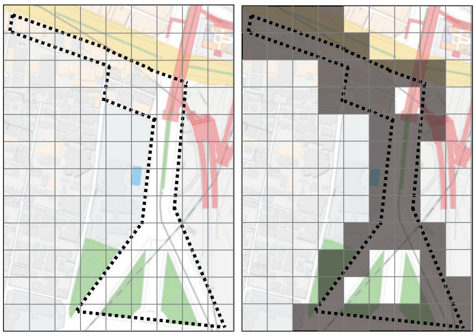

Figure 2 Outline of grids to be weakly supervised label. The left shows how all the sidewalks were divided into meshes, and the right shows how each grid was assigned as a label.

Weakly supervised DCNN

The three axes of acceleration data were segmented into 28502 and 6692 pieces by a sliding window method with a window size of 400 (about 8 seconds) and 100 (about 2 seconds) respectively and overlapping rate of 0.5. As shown in Figure 3, the DCNN used for weakly supervised learning is composed of 7 layers of an input layer, 4 convolutional layers, one fully connected layer, and an output layer. By using the hierarchical structured network and training functions in layers from input to output, feature extractor h and the classifier f those are effective for classification are trained simultaneously.

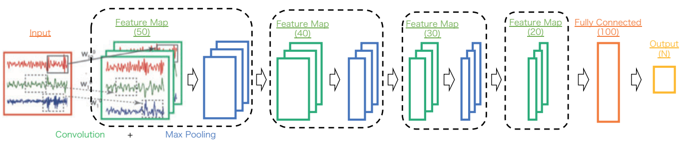

Figure 3 Illustration of the DCNN structure.

Acquisition and Clustering of feature representations

The procedure of acquiring road surface feature representations from weakly trained DCNN and clustering the similar condition road surfaces based on the extracted feature representations is described in order from Step 1 to Step 5.

Step 1: Acquisition of an output pattern of all data

For the DCNN model trained with eight participants data sets as a training data, the remaining one participant data set was input to the DCNN and 100 units output pattern in the fully connected layer was extracted as feature representations of each segmented data.

Step 2: Clustering of feature representations

After compressing the acquired 100-dimensional feature values to a dimension whose cumulative contribution rate exceeds 80% by principal component analysis, clustering was performed on the compressed feature values using the k-means method.

Step 3: Visualization on a map

Clusters generated in Step 2 were color-coded and each point of each cluster was visualized on a map.

Step 4: Analysis of clustering results

Visually comparing the plot result obtained in Step 3 and the recorded video during driving, the road surface condition belonging to each cluster was analyzed.

Step 5: Optimum grid width, window size, and number of clusters

Based on Step 3 and Step 4, the optimum grid width and window size were selected, then a number of clusters that captures the most detailed features of the road surface conditions were selected under best grid width and window size.

4. Qualitative evaluation by clustering

a) Selection of optimum grid width

Plot results with a grid width of 3 m, 4 m, and 5 m with a cluster number of 5 and a window size of 400 were compared.

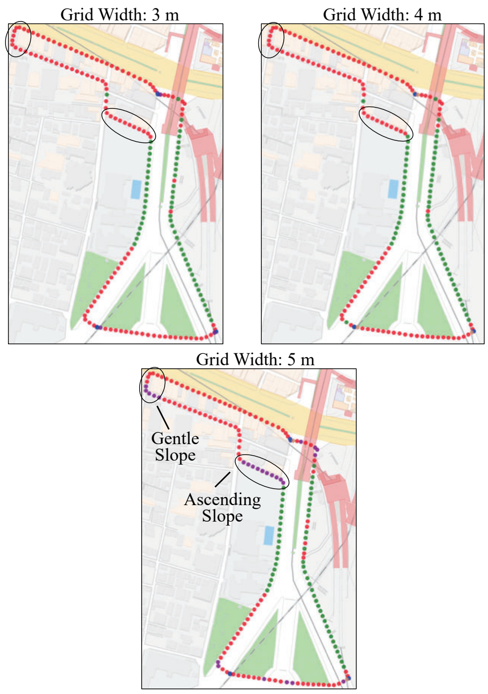

Figure 4 Comparison of clustering results under each condition of a grid width of 3 m, 4 m, and 5 m.

As shown in Figure 4, at the grid width of 5 m, the DCNN captured the features of the ascending slope the most. From this result it is considered that the larger the grid width is, the larger the range of the road surface learned as one label in the DCNN, and DCNN captured an ascending slope where features are easier to read in the larger range.

b) Selection of optimum window size

Plot results with a window size of 100 and 400 with a cluster number of 5 and a grid width of 5 m were compared. As Shown in Figure 5, at the window size of 400, DCNN captured the features of the ascending slope the most. From this result, it is considered that the larger window size is, the larger each segmented training sample in DCNN, and DCNN captured an ascending slope where features are easier to read in the larger range.

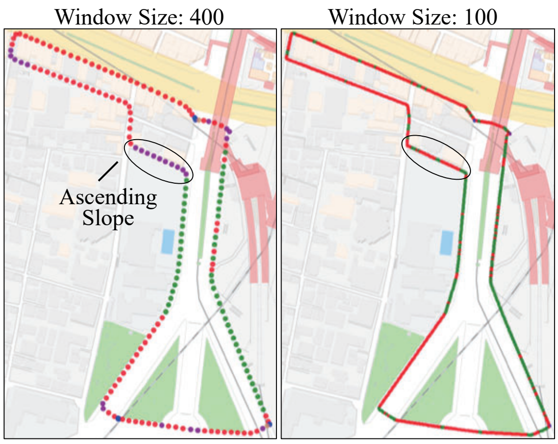

Figure 5 Comparison of clustering results under each condition of a window size of 400 and 100.

c) Selection of the optimum number of clusters

Plot results with the number of clusters 5 to 10 with a grid width of 5 m window size of 400 were compared. As shown in Figure 6, the ascending slope and descending slope were classified into one cluster respectively, and curbs were classified into a specific cluster. Table 1 shows the number of clusters that classified the most detailed road surface condition for each user.

d) Comparison with conventional labeling method

As a result of comparing Figure 5 and clustering result of feature values extracted from detailed labeled trained DCNN, it was shown that weakly supervised method captured more detailed road surface features than the conventional DCNN.

5. Quantitative evaluation of features acquired from weakly supervised DCNN

Evaluation method

Using feature values extracted from weakly trained DCNN as an input, a new classifier was trained as a classification task of four types of labeled road surfaces: slope, curb, braille block, and others. These four types represent typical features of road surfaces, so this method evaluates feature values extracted from weakly supervised DCNN whether they are useful as an indicator of the condition of the road surface.

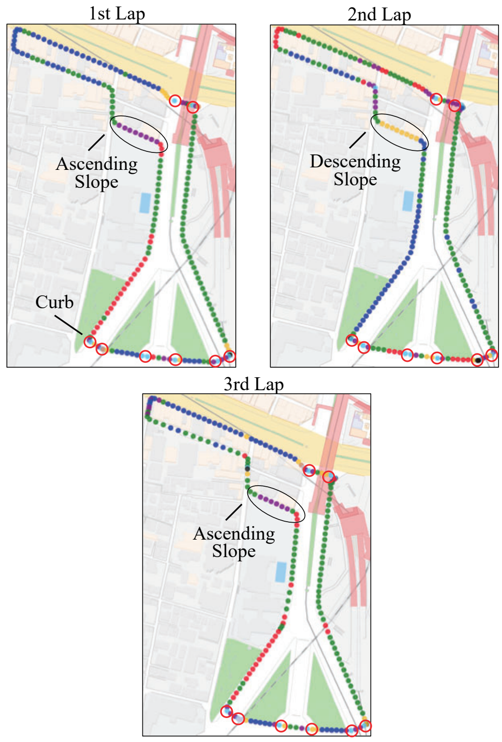

Figure 6 Clustering result with the number of clusters 9 in grid width 5 m and window size 400.
The 1st and 3rd lap are clockwise, so the slope is ascending.
The 2nd lap is counterclockwise, so the slope is descending.

Table 1 Optimum number of clusters for each participant.

Participant	M1	M2	M3	M4	M5	M6	P1	P2	P3
number of clusters	10	8	9	8	9	8	9	9	9

Table 2 Performance comparison between supervised DCNN and weakly supervised DCNN method.

Method	Supervised DCNN	SVM	LR
Mean F-Score	0.58	0.57	0.54
Accuracy	0.81	0.71	0.74

Comparison with Supervised DCNN

Table 2 is a comparison of the classification score. The mean F-score and the accuracy of each class were used as evaluation indexes. Supervised DCNN is the conventional method of training the dataset with DCNN labeled four types of road surfaces. SVM

uses feature values extracted from weakly supervised DCNN as an input and uses Support Vector Machine as a classifier. LR uses feature values extracted from weakly supervised DCNN as an input and uses Logistic Regression as a classifier. As a result, in LR, the mean F-Score was 0.04 points lower than the supervised DCNN, and the accuracy was 0.07 points lower than the supervised DCNN. From this result, it is considered that weakly supervised method misclassified others which occupy more than 70 % of the four labels into a slope, curb, or braille block, and classified the three types of the road surface to the same degree as Supervised DCNN.

6. Conclusion

In this paper, we proposed a novel method to evaluate road surface condition by weakly supervised learning using positional information as a label and accelerometer data. As a result, it was shown that feature representations acquired by Weakly Supervised DCNN can capture more detailed features of road surfaces than feature values by conventional supervised DCNN, and quantitatively estimate road surface condition. As a future work, we will propose a method to acquire higher-precision feature representations based on new weak label generation method and conduct a detailed analysis of what kind of road surface conditions DCNN with the position information extracts.

Acknowledgments.

We are deeply grateful to all participants of the experiment. This study was supported by Tateishi Science and Technology Foundation in FY2011-2012, the research grant by Chiyoda-ku(CHIYODAGAKU) in FY 2014-2016, and JSPS KAKENHI Grant Number 17H01946.

References

[Ponsard 06] Ponsard, C. and Snoeck, V.: Objective accessibility assessment of public infrastructures, in *10th Int. Conf. on Computers Helping People with Special Needs,* pp. 314–321, Linz, Austria (2006)

[Hara 14] Hara, K.: Scalable methods to collect and visualize sidewalk accessibility data for people with mobility impairments, in *Proc. Adjunct Publication of the 27th Annual ACM Symposium on User Interface Software and Technology*, pp. 1–4, Honolulu, HI, USA (2014)

[Cardonha 13] Cardonha, C., Gallo, D., Avegliano, P., Herrmann, R., Koch, F., and Borger, S.: A crowdsourcing platform for the construction of accessibility maps, in *Proc. 10th Int. Cross-Disciplinary Conf. on Web Accessibility*, p. 26 Rio de Janeiro, Brazil (2013)

[Swan 13] Swan, M.: The Quantified Self: Fundamental Disruption in Big Data Science and Biological Discovery, *Big Data*, Vol. 1, No. 2, pp. 85–99 (2013)

[Nagamine 15] Nagamine, K., Iwasawa, Y., Matsuo, Y., and Yairi, I. E.: An Estimation of Wheelchair User's Muscle Fatigue by Accelerometers on Smart Devices, in *Adjunct Proc. 2015 ACM Int. Joint Conf. on Pervasive and Ubiquitous Computing and Proc. 2015 ACM Int. Symposium on Wearable Computers*, UbiComp/ISWC'15 Adjunct, pp. 57–60, New York, NY, USA (2015)

[Zhou 18] Zhou, Z. H.: A brief introduction to weakly supervised learning, in *National Science Review*, Vol. 5, No. 1, pp. 44-53, (2018)

[Oquab 15] Oquab, M., Bottou, L., Laptev, I., and Sivic, J.: Is object localization for free?-weakly-supervised learning with convolutional neural networks, in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 685-694, Boston, MA, USA (2015)

[Gidaris 18] Gidaris, S., Singh, P., and Komodakis, N.: Unsupervised Representation Learning by Predicting Image Rotations, in *6th Int. Conf. on Learning Representations*, Vancouver, BC, Canada, (2018)

[You 18] You, Z., Raich, R., Fern, X. Z., and Kim, J.: Weakly Supervised Dictionary Learning, in *IEEE Trans. Signal Proc.*, Vol. 66, No. 10, pp. 2527-2541, (2018)

[Lara 12] Lara, O. D. and Labrador, M. A.: A survey on human activity recognition using wearable sensors, in *IEEE Commun. Surveys Tutorials*, Vol. 15, No. 3, pp. 1192–1209, (2012)

[Liu 17] Liu, H., Taniguchi, T., Tanaka, K., Takenaka, Y., and T. Bando: Visualization of driving behavior based on hidden feature extraction by using deep learning, in *IEEE Trans. Intelligent Transportation Systems*, Vol. 18, No. 9, pp. 2477–2489, (2017)

[Iwasawa 12] Iwasawa, Y. and Yairi, I. E.: Life-logging of wheelchair driving on web maps for visualizing potential accidents and incidents, in *PRICAI 2012: Trends in Artificial Intelligence*, pp. 157–169, Springer (2012)

[Iwasawa 15] Iwasawa, Y., Nagamine, K., Matsuo, Y., and Eguchi Yairi, I.: Road Sensing: Personal Sensing and Machine Learning for Development of Large Scale Accessibility Map, in *Proc. 17th Int. ACM SIGACCESS Conf. on Computers & Accessibility,* pp. 335–336ACM (2015)

[Iwasawa 16] Iwasawa, Y., Yairi, I. E., and Matsuo, Y.: Combining Human Action Sensing of Wheelchair Users and Machine Learning for Autonomous Accessibility Data Collection, in *IEICE Trans. on Information and Systems*, Vol. E99-D, No. 4, pp. 115–124 (2016)

[Takahashi 18] Takahashi, H., Nagamine, K., Iwasawa, Y., Matsuo, Y., and Yairi, I. E.: Quantification of Road Condition from Wheelchair Sensing Data Using Deep Convolutional Neural Network, in *IEICE Trans. Information and Systems*, Vol. J101-D, No. 6, pp. 1009-1021 (2018) (In Japanese)