International Journal of Automotive Technology, Vol. ?, No. ?, pp. ?−?(year)

Serial#Given by KSAE

ROAD TYPE IDENTIFICATION AHEAD OF THE TIRE USING D-CNN AND REFLECTED ULTRASONIC SIGNALS

Min-Hyun Kim[1)], Jongchan Park[2)] and Seibum Choi[1)*]

[1)] Automotive Control Laboratory, Department of Mechanical Engineering, KAIST, Daejeon 34141, Korea
[2)] Lunit, 6th floor, 175 Yeoksam-ro, Gangnam-gu, Seoul 06247, Korea

(*Received date* ; *Revised date* ; *Accepted date*) * Please leave blank

ABSTRACT− Every land moving object accelerates or decelerates based on the frictional coefficient of the road surface. It has been known that this coefficient on the road is determined by the type of road surface. In this work, we propose a simplistic, machine-learning based solution to estimate the road type using the reflected ultrasonic signals paired with ultrasonic transmitter and receiver. Since the reflected signal contains the material information of the surface due to the difference in the surface roughness and acoustic impedance, different characteristics can be observed for each frequency of the reflected signal. To exploit such characteristics, the signals are transformed into the frequency domain using short-time Fourier transform. In addition, a deep convolutional neural network is applied as the road identifier due to its well-known representational power. In order to verify the aforementioned ideas, the ample database consisting of eight types of road surfaces are obtained with the ultrasonic sensors. And then, the database is used to train the model, as well as to evaluate the accuracy of the trained model. It can be seen that the proposed method makes it easier and more accurate to identify the type of road surface than the conventional methods.

KEY WORDS : Ultrasonic sensor; Road type identification; Friction coefficient; Short-time Fourier transform; Machine-learning; Deep convolutional neural network

NOMENCLATURE

F_l : feature of l-th layer

b : bias of convolutional layer

w : weight of convolutional layer

B : bias of classifier layer

W : weight of classifier layer

C_{in} : number of input channel

1. INTRODUCTION

All land mobile objects, including humans and robots, cannot accelerate or decelerate well on slippery surfaces like ice. That is, the friction coefficient of the road surface where the wheel will step on is a very important parameter for the movement of land moving objects (Li et al., 2016; arnioli et al., 2016; Persson, 2013). The dynamic-based methods, and vision-based methods have been used to obtain this parameter.

The dynamic-based method (Li et al., 2016; Rajamani et al., 2012; Dahiya et al., 2010; Han et al., 2016) is able to measure this value using the dynamic information of the land vehicles. However, additional vehicle dynamics models, GPS and accelerometer sensors are required to check the state of the vehicle. It is possible to estimate the condition of the road surface only on which the tire is stepped by applying the brakes once in a while without driver's intention. That is, there is a limitation in that this process may affect ride comfort and the condition of the road surface cannot be known in advance. In addition, this physical excitation process affects the energy efficiency for the land vehicle.

While on the subject, the vision-based method (Omer and Fu, 2010; Raj et al., 2012; Elunai et al., 2011) has the advantage as it allows for prior knowledge of the type of road surface approaching; a physical excitation process is unnecessary. However, expensive vision sensors and processing devices are essentially required. Also, a calibration process should be done based on the sensor setup, and a costly pre-processing must be performed from the acquired vision data to locate the road and to eliminate outliers. Furthermore, the use of a vision-based

* *Corresponding author*. e-mail: sbchoi@kaist.ac.kr

approach can potentially mislead non-road surfaces to road surfaces due to ambient illumination effects.

To overcome all of aforementioned limitations, we propose a method for road type identification using the fact that the acoustic characteristics of each object are different (Kinsler et al., 1999). When sound waves are emitted to each material, the reflected waves are dependent on material properties such as the acoustic impedance and surface roughness of the material. That is, even though ultrasonic waves are transmitted in the same waveform each time, different reflected waves are generated according to the material of the surface and can be classified using the difference. Therefore, this paper proposes a method of classifying the type of the road surface by transmitting ultrasonic waves perpendicular to the road surface using an ultrasonic transmitter, and by receiving sound waves reflected from the road surface using an ultrasonic receiver. This paper also proposes a modeling method using machine-learning based on acoustic big data from the reflected wave to distinguish the type of road surface, and the effectiveness of the proposed method is verified through experiments.

Figure 1. An ultrasonic look-ahead sensor module under the front of the vehicle for vehicle predictive control.

2. PREVIOUS WORKS

Ultrasonic sensors have been applied to various fields because they are inexpensive and easy to apply. Almost of the previous methods (Kim and Choi, 2016; Carullo and Parvis, 2001) using ultrasonic signals on the road focused on estimating the distance from the road surface to the sensor system as shown in Fig. 1. These studies showed that the road surface can be profiled before the vehicle passes by using the time-of-flight (ToF) preview information. On the other hand, several studies were done that estimated the friction coefficient of the road surface by passively analyzing the acoustic components from the tire made noise when slipping on the road surface (Masino et al., 2017; Gailius and Jačėnas, 2007; Alonso et al., 2014; Kongrattanaprasert et al., 2009; Nakashima et al., 2016). In more detail, (Nakashima et al., 2016) is to analyze the intensity of the reflected signal after transmitting ultrasonic waves to the road surface. The road surface was classified by simply fitting the intensity of the signal reflected from the road surface according to the type and distance of the road surface. In other words, it is impossible to apply to various road surfaces due to the inferiority in the case of other road surfaces with similar intensity reflection. On the other hand, (Gailius and Jačėnas, 2007) conducted a study to detect ice road surfaces by frequency analysis of the friction noise produced by tires. Similarly, (Kongrattanaprasert et al., 2009) classified the dry, wet, or snowy state into microphones and showed 81% accuracy. Recent studies (Masino et al., 2017; Alonso et al., 2014) used a machine-learning technique called support vector machine (SVM) to classify the road surface, and showed its classification performance from 70% to 92%. However, (Masino et al., 2017) used not only acoustic sensor but also additional pressure and temperature sensors are required inside each tire. In addition, since these aforementioned methods measure both tire noise and ambient noise, the SNR is relatively low. Also, the condition of the road surface can be known only after the tire has stepped on the road surface. Another study

The SVM used as a machine-learning model for these studies (Masino et al., 2017; Alonso et al., 2014) is widely used for its robustness and quick convergence. However, the SVM requires a well-designed feature extractor, which may require much expert knowledge and efforts. That is, a lot of human effort is required for parameter tuning. On the other hand, a deep neural network (DNN) (LeCun et al., 2015; Krizhevsky et al., 2012) jointly learns the feature extractor and the classifier, and does not require much heuristic hand-tunings. Also, with the recent advances of DNN (Goodfellow, et al, 2016), 1D convolutional neural network (CNN) is widely used as a universal function approximator to process large collected dataset. Due to its simplicity and representational power, it shows faster data processing speed and accuracy than other models.

In order to use a DNN as the function approximator by the data-driven method, the data type of input must be specified first. According to the study (Kinsler et al., 1999), all materials have different acoustic impedances. That is, when observing the reflected signal in the frequency domain, the differences can be seen depending on the type of road surface. Therefore following previous approaches (Arandjelovic and Zisserman, 2017), a spectrogram extracted from short-time Fourier transform (STFT) is used as the input to analyze the frequency domain over time.

In summary, this paper proposes a method of classifying the types of road surfaces with different acoustic characteristics for each material. These acoustic properties include the material's acoustic impedance and

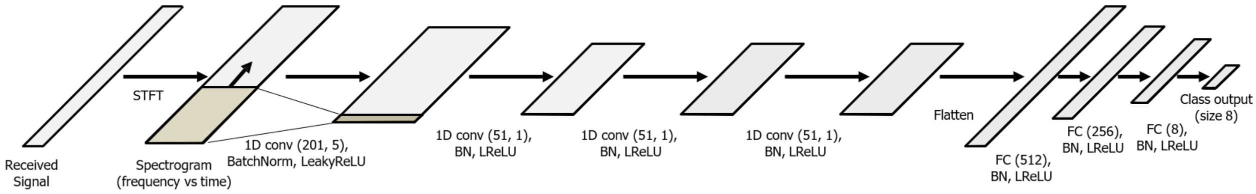

Figure 2. Proposed network architecture. The numbers for 1D convolution layer are (window size, stride), the number for fully connected (FC) layer is the hidden output size, BN denotes the batch normalization layer, and LReLU denotes LeakyReLU. The number of convolutional filters is 64 for all layers.

surface roughness. In order to increase the SNR, an active method of transmitting ultrasonic waves that are invisible to the human ear is actively used on the road surface. In addition, the spectrogram converted to STFT is used as the input of DNN to observe the frequency domain of the received ultrasound. One type of DNN, 1D CNN, is used as a function approximator and consists of MLP as the final classifier.

Here, the vertical velocity due to the heave motion of the ground moving object and the height change of the road surface is very small compared to the speed of the sound wave, so the Doppler effect is ignored. In addition, the tire road friction coefficient is assumed to be determined by the slip ratio of the wheel as suggested in the previous studies (Burckhardt, 1993; Rajamani et al., 2010) if the type of road surface is determined through the output of our trained model.

3. IDENTIFICATION MODEL AND ARCHITECTURE

To identify the type of road surface from the reflected ultrasound waves, the modeling process must be performed first. The reflected ultrasonic signals from the road largely depends on the road material, surface roughness. However, it is almost impossible to mathematically model the relationship between sound waves and road surfaces by considering all circumstances. Therefore, a modeling method based on machine-learning is used to generalize various input signals from the ultrasonic receiver.

Our DNN is configured to consist of several 1D convolutional layers followed by non-linear layers and batch normalization (Ioffe and Szegedy, 2015) for feature extraction, as well as a multi-layered perceptron (MLP) at the end for classification process. It is common to treat spectrograms as images and may seem more sensible to use 2D CNN for feature extractor, but as the time dimension of spectrogram is narrow, we intuitively use 1D CNN instead. Also, 2D CNNs have more parameters on average, which may lead to over-fitting in some cases.

The design principles of the deep convolutional neural network (D-CNN) are simple and lightweight, so we designed them using only a few common layers. The designed CNN consists of a feature extractor of four convolution layers and a classifier of three layers fully connected. The filter size and stride for 4 convolutional layers are (201,5), (51,1), (51,1), (51,1), respectively as shown in Fig. 2. In addition, all convolutional layers have 64 filters each. The padding for each layer is half of the filter size. Each convolutional layer is followed by a batch-norm layer and a LeakyReLU (Xu et al., 2015) layer:

$$F_{l+1} = LReLU(BN(bias + \sum_{k=0}^{C_{in}-1} w_k * F_l^k)) \tag{1}$$

where $F_l \in R^{C_{in} \times L_{in}}$ and $F_{l+1} \in R^{C_{out} \times L_{out}}$ denotes the feature map after layer l and layer $l+1$, $w_k \in R^{C_{out}}$ and $bias \in R^{C_{out}}$ denote conv weight and bias, respectively. In addition, BN and LReLU mean batch-norm layer and LeakyReLU layer. In other words, this equation shows that batch normalization and LeakyReLU are applied to the results of a general CNN.

In the classifier, the hidden sizes of the MLP are 512 and 256, respectively, and the final output dimension is the number of classes. The LeakyReLU layers are also inserted in between all layers for non-linearity:

$$F_{l+1} = LReLU(W * F_l + B) \tag{2}$$

where W and B denote weights and bias for the fully connected layer, respectively.

The convolutional feature map F between convolutional layers are 2D shaped, and the last convolutional feature map is flattened to 1D shape. Thus, all feature maps F in MLP are 1D shaped. The last fully-connected layer has no non-linear layer. Furthermore, a standard soft-max cross entropy loss is used for training. Additionally, we use the Adam optimizer with a constant learning rate of 0.0001 and a dropout with 0.3 rate in the classifier for regularization. To compare 2D CNN and 1D CNN, additional experiments are conducted, and 2D CNN has the same architecture with 1D, but the filter size along time axis is set to 3.

4. EXPERIMENTAL DESCRIPTION

This section describes the experimental environment for estimating the type of road surface in the manner suggested above. In addition, the dataset in which ultrasonic signals reflected on the road surface are collected is described.

4.1. Experimental Equipment and Settings

The ultrasonic sensor system, which is the set of transmitter and receiver, is installed perpendicular to the road surface as shown in Fig. 1. This sensor is used with a main frequency of 40kHz, which is twice the human audible frequency. To clearly sample the analog signal from the sensor, it is sampled at a frequency of 20 times higher than the main frequency that satisfies the condition of the Nyquist theorem (Nyquist, 1928). That is, the analog sampling module is used as an instrument that can obtain 1 million signal samples per second from an ultrasonic receiver. In addition, we apply GPU (NVIDIA GeForce GTX 1080 Ti) for parallel processing to train and test the road type model from the collected database. The details of the instruments used are shown in Table 1.

4.2. Dataset Composition

Well-trained models should be able to distinguish various road types, so the dataset must be collected from diverse road types in various environments. For this experiment, the reflected ultrasounds are collected on a total of eight road surfaces: asphalt, cement, dirt, marble, paint, snow, water, and ice. In addition, the time to receive the period is set to 20ms per sample such that the transmitted ultrasonic signal is sufficient to disappear, as shown in Fig. 3.

More than 6,000 datasets from eight different types of road surface are collected to validate the approach, and the numbers of obtained dataset are shown in Table 2. In addition, the following efforts are made to ensure the generality of the dataset with more intra-class diversity. First of all, the dataset is collected while moving the sensor along the road to ensure diversity even on one class of road surface, and the dataset is obtained by varying the distance between the sensor and the road surface in preparation for the unevenness of the road surface. Furthermore, the database of the same class is collected over several days to accommodate multiple environments.

Among the collected data, we use 70% for training and 30% for testing as in other studies. All the data are pre-processed with STFT (nperseg=5k, 10k, and 15k) without any augmentations. Here nperseg is used as one of STFT tuned parameter, and it represents the ultrasonic signal length of segment for the FFT. That is, it is a variable that determines the size of local section to which the Fourier transform is applied.

Table 1. The equipment used for this experiment.

Equipment Type	Manufacturer	Model Name	Specification
Ultrasonic Transmitter	Hagisonic	HG-M40T	40kHz
Ultrasonic Receiver	Hagisonic	HG-M40R	40kHz
Controller	NI	cRIO-9036	Ethernet based
Module of Controller	NI	NI 9223	1MSamples/second
DC Battery	ROCKET	Lead Battery	12V
GPU for training	NVIDIA	GTX 1080 Ti	3584 CUDA Cores

Table 2. The number of obtained samples for each type of road surface

Road Surface Type	Train Set	Test Set
Asphalt	1661	712
Cement	440	189
Dirt	633	272
Ice	447	191
Marble	211	91
Paint	214	93
Snow	208	89
Water	431	185
Total	4247	1820

5. RESULTS

Through the dataset collected using the sensor module and controller described above, the training process converges within approximately 20 minutes, except for the STFT time, and each test inference takes 1 millisecond per sample on our single GPU system described in Table 1. Here, all dataset for training and testing are randomly selected without overlaps, and the test results of our 1D D-CNN method are shown in Table 4, 5, and 6. In addition, the result of 2D D-CNN for comparison with 1D is shown in Table 3. In these tables, each row represents the ground truth class label, and each column represents the inferenced class.

For the asphalt, marble, and snow surfaces, this method provides complete estimation performance for all three

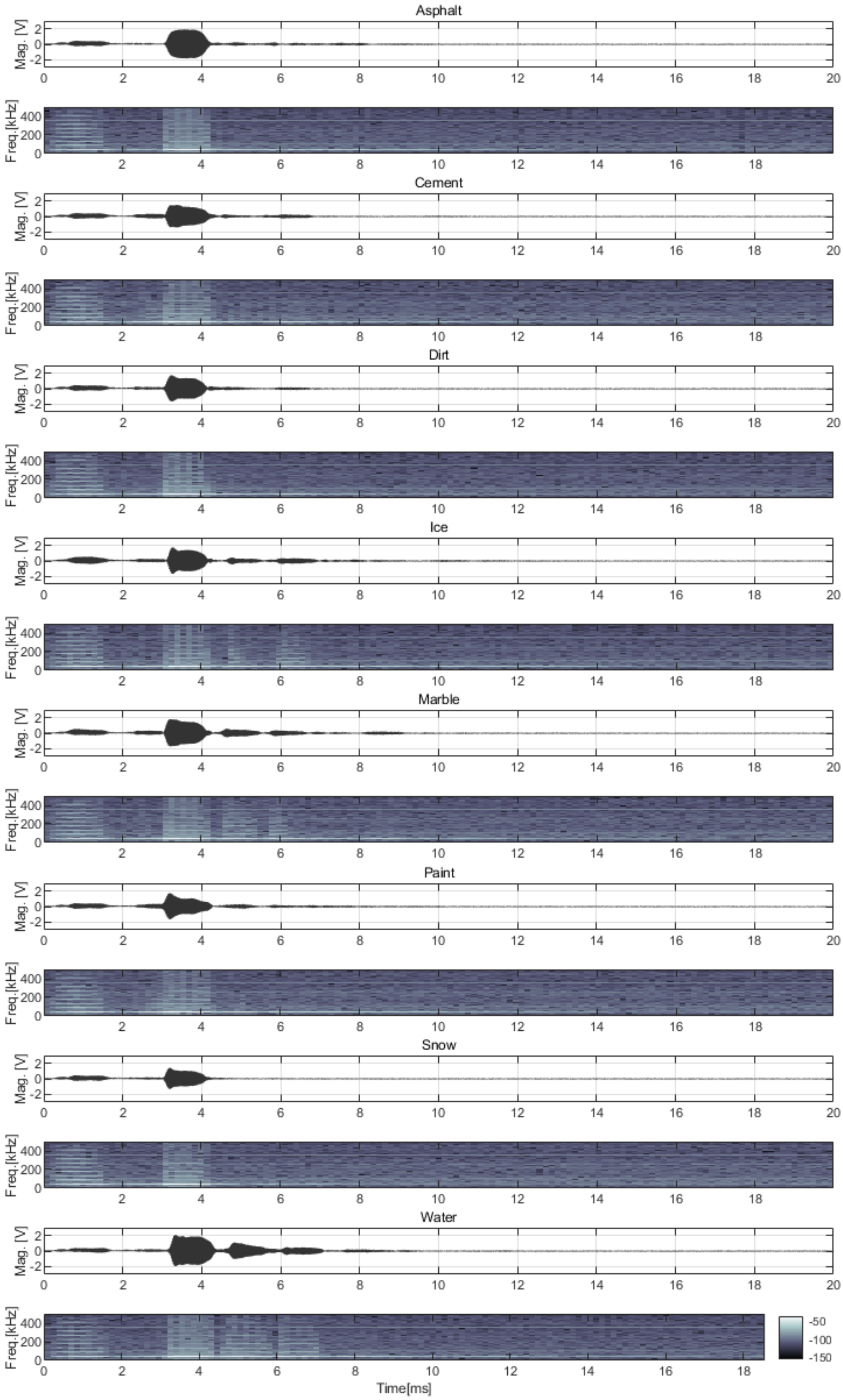

Figure 3. Sample signals of raw and spectrogram for all road types. In the spectrogram, the color represents the power of the corresponding frequency at the time. In other words the bright color means strong, the dark means weak.

parameter settings in Table 4-6. In detail, when nperseg is 5k, it shows more than 99% accuracy on roads that are asphalt, cement, dirt, marble, paint, water and ice. Also, when nperseg is 10k, it shows 100% accuracy for asphalt, cement, marble, ice, and all roads except for snow and water at 15k. Comparing the result between Table 3 and Table 6, the average accuracy of 1D CNN (97.5%) is higher than 2D CNN (95.2%). This is partially due to increased parameter size, and support our assumption that 1D CNN is more intuitive to use.

The proposed method shows totally over 97% accuracy for most of the road surface types. For some road surface types, the validation accuracies achieve a near-perfect level. Compared with the previous study (Alonso et al., 2014), it can be confirmed that the accuracy is improved by 1.8 to 2.5%. Especially in the case of asphalt, the performance is 12% higher. It also shows superiority in response time. In the study (Alonso et al., 2014), it takes a minimum of 0.2 seconds to recognize that the road surface has changed, but our proposed algorithms can fully appreciate the changes in the road surface in tens of milliseconds.

In addition, Table 7 is the comparison result of modeling using the decision tree (Breiman N. 2017) and the SVM (Cristianini et al., 2000) technique, which are frequently used in machine-learning studies. The comparison is conducted using the same dataset we collected. From the table, the superiority of our proposed D-CNN method is validated. The proposed method shows the highest accuracy for all type of road surfaces. However, other methods (Breiman N. 2017; Cristianini et al., 2000) also have low accuracy for snow road surface.

It is notable that the accuracy of ice road. The characteristics of ice are usually intuitively hard to be captured by visual information. But, our approach that incorporates ultrasonic signals allows for successful distinguishing between ice and other road types.

6. DISCUSSION AND CONCLUSION

Identifying the type of road surface is an important factor in the movement of land vehicles including walking robots. In the beginning of this study, it is assumed that the reflected ultrasonic signals can convey the characteristics of the road material types. Therefore, a new method for estimating the type of road surface using ultrasonic transmitters and receivers with D-CNN is proposed. In addition, the assumption is verified and the type of the road surface can be estimated by analyzing the frequency component using the STFT from the road reflected ultrasonic wave. Furthermore, throughout the comprehensive ultrasonic sensor-based dataset collection and experiments, a high-performance road identification model is trained.

Since ultrasonic sensors are inexpensive and easy to operate, they can be effortless to use and apply. Therefore, by attaching the sensor to the front of the robot or the front bumper of vehicle as shown in Fig. 1, it is possible to preview control the motion suitable according to the type of road surface. Moreover, this method can easily pre-estimate the road surface by installing an ultrasonic sensor without any physical excitation. In addition, since the ultrasonic sensor module transmitted ultrasonic waves with a cycle of 20 ms, the type of road surface can be estimated every 20 ms. Therefore, this sensor module can react robustly to the rapid change of the surrounding environment. Furthermore, since the database has been collected in various environments, the type of road surface can be estimated regardless of the flight time of ultrasound, ambient temperature, etc.

Above all things, the main advantage of this proposed method is the ability to detect the ice-covered roads that are not visually recognizable, such as black ice. Thus, a notification can be sent in advance

Table 3. Confusion matrix for test data with nperseg=15k, with 2D convolutions.

	asphalt	**cement**	**dirt**	**marble**	**paint**	**snow**	**water**	**ice**	**Accuracy (%)**
asphalt	712								100
cement		189							100
dirt			272						100
marble				91					100
paint					93				100
snow	45					44			49.4
water						41	143	1	77.3
ice								191	100
Avg.									95.2

Table 4. Confusion matrix for test data with nperseg=5k, with 1D convolutions.

	asphalt	**cement**	**dirt**	**marble**	**paint**	**snow**	**water**	**ice**	**Accuracy (%)**
asphalt	712								100
cement		188	1						99.5
dirt			272						100
marble				91					100
paint					93				100
snow	46					43			48.3
water							185		100
ice								191	100
Avg.									97.4

Table 5. Confusion matrix for test data with nperseg=10k, with 1D convolutions.

	asphalt	**cement**	**dirt**	**marble**	**paint**	**snow**	**water**	**ice**	**Accuracy (%)**
asphalt	712								100
cement		189							100
dirt	2		270						99.3
marble				91					100
paint	2				91				97.8
snow	48					41			46.1
water	2						183		98.9
ice								191	100
Avg.									97.0

Table 6. Confusion matrix for test data with nperseg=15k, with 1D convolutions.

	asphalt	**cement**	**dirt**	**marble**	**paint**	**snow**	**water**	**ice**	**Accuracy (%)**
asphalt	712								100
cement		189							100
dirt			272						100
marble				91					100
paint					93				100
snow	45					44			49.4
water	1						184		99.5
ice								191	100
Avg.									97.5

Table 7. Comparison of decision tree and SVM with proposed method.

	asphalt	**cement**	**dirt**	**marble**	**paint**	**snow**	**water**	**ice**	**Accuracy (%)**
Proposed 15k	100	100	100	100	100	49	99	100	97
Decision Tree	90	67	84	91	87	49	49	90	79
SVM	98	69	98	97	90	37	97	97	86

to all moving objects, and acceleration and deceleration preview control becomes entirely possible before stepping on the ice.

A limitation of our method is that the STFT conversion takes a certain amount of time, so real-time properties may not be guaranteed. In addition, the snow road database is classified to asphalt road. It is believed that both asphalt and snow are porous, and the acoustic characteristics for a particular frequency band are similar to each other, as well as the conversion through STFT results in loss of information in the acoustic signal.

By observing the collected signal in Fig. 3, it can be observed that the ultrasonic signal disappears faster than the collection time of 20ms. Therefore, in the future work, it is necessary to shorten the signal collection time to improve the real-time performance by reducing the time required for signal processing and classification. In addition, the frequency domain observations through signal processing method other than STFT will complement performance on the snow roads.

REFERENCES

Li, X., Izawa, D., Minami, M., Matsuno, T. and Yanou, A., 2016, December. Dynamical model of humanoid considering slipping with nonlinear floor friction and internal force during free-fall motion. In *System Integration (SII), 2016 IEEE/SICE International Symposium on* (pp. 69-74). IEEE.

Farnioli, E., Gabiccini, M. and Bicchi, A., 2016, October. Toward whole-body loco-manipulation: Experimental results on multi-contact interaction with the Walk-Man robot. In *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on* (pp. 1372-1379). IEEE.

Kinsler, L.E., Frey, A.R., Coppens, A.B., Sanders, J.V. Fundamentals of acoustics, 4th Edition, pp.560. ISBN 2510-471-84789-5. Wiley-VCH, December 1999.

Persson, B.N., 2013. *Sliding friction: physical principles and applications*. Springer Science & Business Media.

Rajamani, R., Phanomchoeng, G., Piyabongkarn, D. and Lew, J.Y., 2012. Algorithms for real-time estimation of individual wheel tire-road friction coefficients. *IEEE/ASME Transactions on Mechatronics*, *17*(6), pp.1183-1195.

Rajamani, R., Piyabongkarn, N., Lew, J., Yi, K. and Phanomchoeng, G., 2010. Tire-road friction-coefficient estimation. *IEEE Control Systems Magazine*, *30*(4), pp.54-69.

Dahiya, R.S., Metta, G., Valle, M. and Sandini, G., 2010. Tactile sensing—from humans to humanoids. *IEEE transactions on robotics*, *26*(1), pp.1-20.

Han, K., Hwang, Y., Lee, E. and Choi, S., 2016. Robust estimation of maximum tire-road friction coefficient considering road surface irregularity. *International journal of automotive technology*, *17*(3), pp.415-425.

Omer, R. and Fu, L., 2010, September. An automatic image recognition system for winter road surface condition classification. In *Intelligent transportation systems (itsc), 2010 13th international ieee conference on* (pp. 1375-1379). IEEE.

Raj, A., Krishna, D. and Shantanu, K., 2012, September. Vision based road surface detection for automotive systems. In *Applied Electronics (AE), 2012 International Conference on* (pp. 223-228). IEEE.

Elunai, R., Chandran, V. and Gallagher, E., 2011. Asphalt concrete surfaces macrotexture determination from still images. *IEEE Transactions on Intelligent Transportation Systems*, *12*(3), pp.857-869.

Kim, M.H. and Choi, S.B., 2016, September. Road Surface Height Estimation for Preview System Using Ultrasonic Sensor Array. In *FISITA World Automotive Congress*. FISITA.

Carullo, A. and Parvis, M., 2001. An ultrasonic sensor for distance measurement in automotive applications. *IEEE Sensors journal*, *1*(2), p.143.

Masino, J., Pinay, J., Reischl, M. and Gauterin, F., 2017. Road surface prediction from acoustical measurements in the tire cavity using support vector machine. *Applied Acoustics*, *125*, pp.41-48.

Kongrattanaprasert, W., Nomura, H., Kamakura, T., Ueda, K. Detection of road surface conditions using tire noise from vehicles. IEEJ Transactions on Industry Applications, 2009, 129, 761–767.

Gailius, D. and Jačėnas, S., 2007. Ice detection on a road by analyzing tire to road friction ultrasonic noise. *Ultragarsas" Ultrasound"*, *62*(2), pp.17-20.

Alonso, J., Lopez, J.M., Pavón, I., Recuero, M., Asensio, C., Arcas, G. and Bravo, A., 2014. On-board wet road surface identification using tyre/road noise and support vector machines. *Applied Acoustics*, *76*, pp.407-415.

Burckhardt, M., Fahrwerktechnik: Radschlupf-Regelsysteme. 1993. *Vogel-Verlag, Wurtzburg*.

LeCun, Y., Bengio, Y. and Hinton, G., 2015. Deep learning. *nature*, *521*(7553), p.436.

Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).

Arandjelovic, R. and Zisserman, A., 2017, October. Look, listen and learn. In *2017 IEEE International Conference on Computer Vision (ICCV)* (pp. 609-617). IEEE.

Ohtani, K. and Baba, M., 2006, April. A simple identification method for object shapes and materials using an ultrasonic sensor array. In *Instrumentation and Measurement Technology Conference, 2006. IMTC 2006. Proceedings of the IEEE* (pp. 2138-2143). IEEE.

Moritake, Y. and Hikawa, H., 2005. Category recognition system using two ultrasonic sensors and combinational logic circuit. *Electronics and Communications in Japan (Part III: Fundamental Electronic Science)*, *88*(7), pp.33-42.

Goodfellow, I., Bengio, Y., Courville, A. and Bengio, Y., 2016. *Deep learning* (Vol. 1). Cambridge: MIT press.

Breiman, L. Classification and regression trees, Routledge, 2017.

Cristianini, N., Shawe-Taylor, J. others. An introduction to support vector machines and other kernel-based learning methods, Cambridge university press, 2000.

Nakashima, S., Aramaki, S., Kitazono, Y., Mu, S., Tanaka, K., & Serikawa, S., 2016. Application of ultrasonic sensors in road surface condition distinction methods. Sensors, 16(10), 1678.

Xu, B., Wang, N., Chen, T., & Li, M., 2015. Empirical evaluation of rectified activations in convolutional network. arXiv preprint arXiv:1505.00853.

Ioffe, S. and Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167.

Nyquist, H., 1928. Certain topics in telegraph transmission theory. Transactions of the American Institute of Electrical Engineers, 47(2), pp.617-644.