Sensors and Actuators A 291 (2019) 7–12

Contents lists available at ScienceDirect

Sensors and Actuators A: Physical

journal homepage: www.elsevier.com/locate/sna

Using a convolutional neural network to construct a pen-type tactile sensor system for roughness recognition

Satoshi Tsuji*, Teruhiko Kohama

Dept. of Electrical Engineering, Fukuoka University, Japan

ARTICLE INFO

Article history:
Received 28 September 2018
Received in revised form 12 February 2019
Accepted 24 February 2019
Available online 23 March 2019

Keywords:
Tactile texture recognition
Tactile sensor
Convolutional neural networks

ABSTRACT

The tactile texture of a product is one of many important factors that influences an impression of a product. Numerical evaluation methods for recognizing tactile texture measures, such as roughness, are required because there are indeed individual differences in how users experience tactile textures. In this paper, we propose a tactile sensor to recognize roughness using a convolutional neural network (CNN). Our sensor system consists of a pressure sensor and a six-axis acceleration sensor for detecting time-series data. The sensor measures time-series data, which are the pressure, speed, and posture of the sensor when the sensor is touching an object and is moved by a user. The surface roughness is calculated from these time-series data using a CNN. Here, our system configuration is simple and therefore easy and inexpensive to construct. To evaluate our approach, we constructed a prototype sensor and measured the roughness of six objects. The average correct recognition rate proved to be 71% for the experimental data acquired by one user, which are categorized into learning data and evaluation data. Further, the total average recognition rate for evaluation data by our five users for considering each individual using the sensor system was 42%. While the problem of roughness recognition by each individual user remains, we were able to show the possibility of roughness recognition via our approach. We conclude that our proposed sensor system is useful as a functional and useful device.

1. Introduction

Tactile texture is one of many important factors that influences an impression of an object. Human users naturally identify the tactile texture of an object by touch, often placing substantial value on the detected tactile texture in selecting personal belongings, such as clothes. Therefore, design efforts that include tactile texture are crucial in the development of a product. Generally, the quality of the tactile texture is identified by a human. The feeling invoked by the tactile texture is typically expressed in words, e.g., “smooth,” “slippery,” and “flat” can represent slight differences in texture; however, these words are ambiguous, and the feeling invoked by the tactile texture varies from one individual to another based on contact conditions, temperature, and humidity. Given this, a numerical evaluation method for measuring tactile texture is required.

A number of recent studies have focused on recognizing objects. In [1], Asaga et al. proposed a tactile evaluation method based on a human tactile perception mechanism; in their system, a piezoelectric element measures vibration information of a contact with an object set on a sample table rotated at a constant speed. In [2], Song et al. proposed a fabric surface texture sensor using PVDF film. Here, a sensor uses polyvinylidene difluoride (PVDF) to measure the surface texture when the sensor moves along an object at a certain speed. While the measured information has high levels of precision and recall, the sensors require the driving device to measure at a constant speed, a constant motion, and a constant pressure. In [3], Liu et al. proposed a fingertip piezoelectric tactile sensor array to encode the roughness of a surface. Here, the sensor measures roughness without the need to maintain a constant speed for the sensor by automatically correcting based on speed information; however, their sensor setup requires a driving device to move the sensor.

Next, we turn to research involving human users moving a sensor along a given product. In [4], Tanaka et al. developed a tactile sensor using humans' ability of material discrimination based on haptic bidirectionality. Here, the sensor consists of two microphones mounted on a human fingertip. The finger with the sensor moves along an object, and the sensor detects the surface conditions of the given object. Similarly, in [5], Ye et al. proposed a pen-type sensor for sensing surface roughness. Their sensor consists of strain gauges, a force sensor, and a PVDF sensor. Further, in [6], Liu et al. proposed a surface material categorization method

* Corresponding author at: Dept. of Electrical Engineering, Fukuoka University, 8-19-1 Nanakuma, Jonan-ku, Fukuoka 814-0180, Japan.
E-mail address: tsuji@fukuoka-u.ac.jp (S. Tsuji).

https://doi.org/10.1016/j.sna.2019.02.032
0924-4247/

using both sound and acceleration measurements from the contact between a sensor and an object. Although the above approaches are able to measure surface conditions via a simple sensor, measurement data is influenced by contact pressure, speed, and motion from a user.

Beyond the approaches above, deep learning methods have been the focus of research more recently. Deep learning methods are often used to perform data analyses of tactile sensors. A recognition system using a deep learning method and tactile sensors that are part of a robot hand are proposed in [7] and [8]. These systems can recognize objects by grasping an object, though this recognition identifies the type of object, not the physical properties of the object. Systems for measuring surface conditions such as the physical properties of an object using deep learning methods have been proposed in [9–13]. In [9], Kerzel et al. proposed a haptic material classification system that uses a robot arm and a three-axis optical force sensor. Here, the robot arm moves the sensor along an object, and surface conditions are recognized via deep learning; however, the sensor requires the robot arm to move at a constant speed, a constant motion, and a constant pressure.

In [10], Gao et al. proposed a method for classifying surfaces that relies on haptic adjectives from both visual and physical interaction data using deep learning for robots. In [11], Erickson et al. proposed a semisupervised learning approach for material recognition that uses generative adversarial networks with such haptic features as force, temperature, and vibration. Both of these systems are designed for robot applications, i.e., human users do not use this system directly. In [12], as a portable system, Zheng et al. proposed a deep learning method focused on surface material classification using haptic acceleration data from an acceleration sensor and image data of a camera. While surface material classification is possible using this method, the haptic acceleration data influences the resulting contact pressure, speed, and motion produced by a user because the haptic sensor only uses an acceleration sensor. Finally, in [13], Matti et al. proposed a surface material classification and retrieval system for tool-mediated freehand surface exploration based on perceptual tactile features and its evaluation using three machine learning approaches. Here, the setup consists of a number of sensors, including an acceleration sensor, a microphone, an infrared sensor, a webcam, and a force sensor. Further, this system does not require explicit scan force or scan velocity measurements. However, we think that the numerous sensors complicate the system structure.

For our efforts, we aim to construct a simple tactile texture sensor system using deep learning, which human users moving a sensor along a given the object. The different point from previous work [12] is that pressure, speed and posture of the sensor are used in addition to acceleration data. In our previous work, we proposed a tactile texture-recognition method using deep learning [14,15]. Our sensor system consists of a pressure sensor and a six-axis acceleration sensor that measures pressure and six-axis acceleration data; however, these acceleration data include gravitational acceleration, and only pressure and acceleration data are used to recognize objects. Further, the sampling frequency of the pressure data was 100 Hz, and we did not consider the physical properties of the object.

Given the above, in this paper, we propose a simple tactile sensing system without the need for a driving device to recognize roughness, using a convolutional neural network (CNN) to improve upon our previous approach. We make use of surface roughness, among other physical properties of an object, as a starting point. Fig. 1 shows the target configuration of our study. More specifically, the proposed pen-type sensor consists of a pressure sensor and a six-axis acceleration sensor. Our proposed system does not require a fixed scan speed, posture, or pressure since it uses both a motion sensor and a pressure sensor. A user moves the sensor along

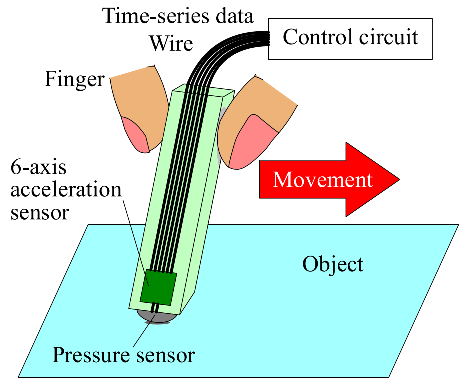


Fig. 1. Target configuration of our present study.

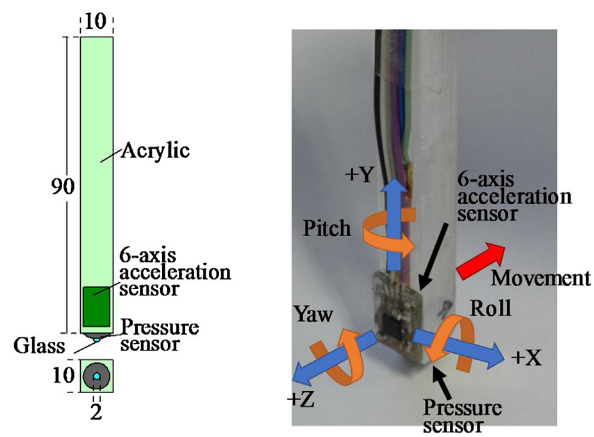


Fig. 2. Prototype sensor. (Dimensions in mm.).

the surface of an object. Vibrations between the sensor and the object change depending on object surface characteristics, in particular roughness, which also affects the contact pressure, speed, and motion of the user. Here, the sensor measures the acceleration, speed, and motion of the sensor, as well as the pressure, which includes the vibration between the sensor and the object. The surface roughness of an object is then calculated using a CNN from the measured data when the sensor is moved by the user. Therefore, our sensor system is simple and inexpensive in that it does not require a driving system. The difference in this study from our previous work [14,15] lies in the fact that the sampling frequency of pressure data is 1 kHz and that attitude data for the sensor and gravity-free acceleration data are used.

Beyond this introduction, in Section II, we describe our prototype sensor system and proposed method using a CNN. In Section III, we provide our evaluation results and discussion. Finally, in Section IV, we present our conclusions.

2. Proposed method

2.1. Prototype sensor system

In this study, we propose a simple sensing system that uses deep learning to recognize tactile textures, in particular the roughness of an object. Fig. 2 shows a prototype pen-type sensor that can easily be held in a user's hand. The sensor system consists of a pressure conductive sensor (Inaba Rubber Co. Ltd, SF-R-3) and a six-axis acceleration sensor (TDK Corporation, MPU-6050). The pressure sensor is attached to the tip of an acrylic rod (Size: $10 \times 10 \times 90$ mm), while the acceleration sensor is attached to the side of the acrylic rod. Further, a 2-mm diameter glass hemisphere is attached to the tip of the pressure sensor to reduce errors when

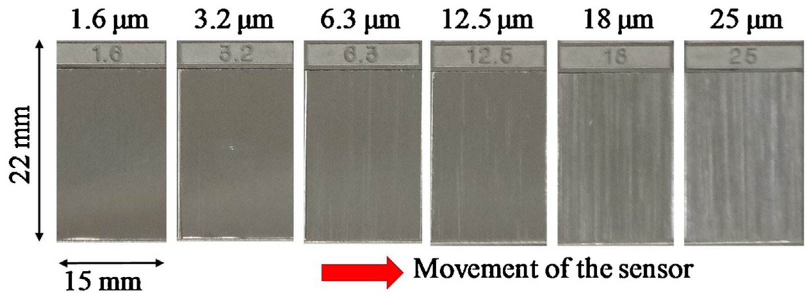


Fig. 3. Surface roughness scale.

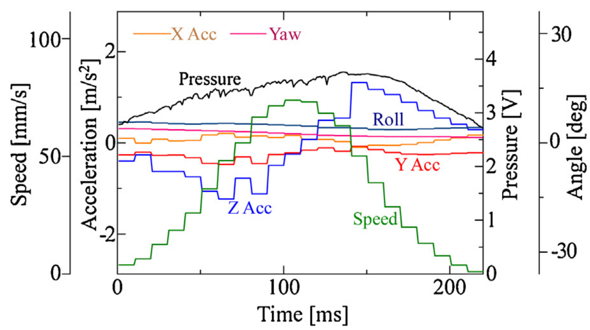


Fig. 4. Measured data from our prototype sensor.

in contact with the desired surface and to prevent wear. Pressure information, including vibration, is obtained by measuring the change in resistance of the pressure-sensitive conductive rubber of the pressure sensor due to the contact between the glass and object. The resistance value of the conductive rubber is measured against the voltage value of a reference resistance using an analog/digital (A/D) converter. Further, we use an MPU-6050 to detect three-axis acceleration and three-axis angular velocity measures. Here, the MPU-6050 contains a digital motion processor and can output attitude data for the sensor and gravity-free acceleration data (i.e., along the X, Y, and Z axes) from the measured six-axis acceleration data.

Next, the data from the six-axis acceleration sensor are sent to a microcomputer via an I^2C serial interface bus. The measured data are then sent to a PC from the microcomputer via a USB serial interface bus. Here, the sampling frequency of the pressure data is 1 kHz, while that of the attitude data of the sensor and gravity-free acceleration data are each 100 Hz. Objects are measured using the surface roughness scale which are metals shown in Fig. 3, with roughness values of 1.6 μm, 3.2 μm, 6.3 μm, 12.5 μm, 18 μm, and 25 μm (Rmax. (JIS B0601)). Note that the target object is first placed on a horizontal plane.

These data are obtained via the following operation: This surface information can be obtained using this sensor when the user moves the sensor along the given object. For these measurements, a user grips the sensor and sets the sensor on the object vertically, and the sensor is moved along the object in the negative Z-axis direction from a stationary state on the object. From there, the sensor moves about 10 cm and stops on the object. The data are then obtained when the sensor is in contact with the object and the negative Z speed is above a specified threshold. Here, speed is calculated by integrating the acceleration along the X, Y, and Z axes. And more specifically, the collected data are pressure, gravity-free X-Y-Z acceleration, speed, and attitude of the sensor. Further, roll and yaw are measured as angles from an attitude (i.e., 0 deg) when the sensor is upright versus the object. Note that pitch is ignored because it depends on the direction of the sensor in this study. Fig. 4 shows a sampling of data obtained from the sensor when the roughness of the object was 12.5 μm. The measured Z-axis acceleration and Z-axis speed increases with movement of the sensor, and decreases with the stop of the sensor. The pressure changes with the movement of the sensor. The attitude and X-Y-axis acceleration of the sensor does not change much.

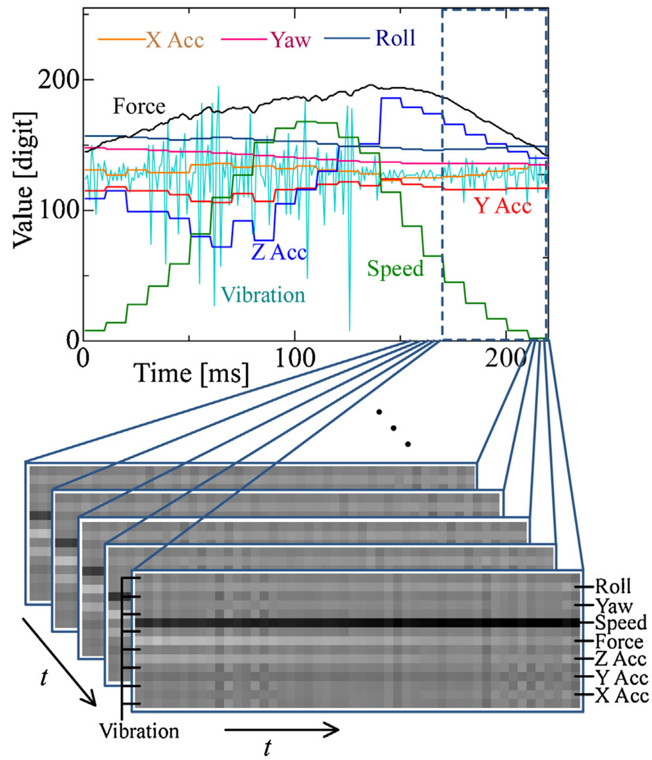
The above data contain complicated temporal numerical changes that require complicated analysis, which makes calculating roughness difficult. Therefore, we identify the roughness of a given object from these complicated data using deep learning.

2.2. Recognition using a CNN

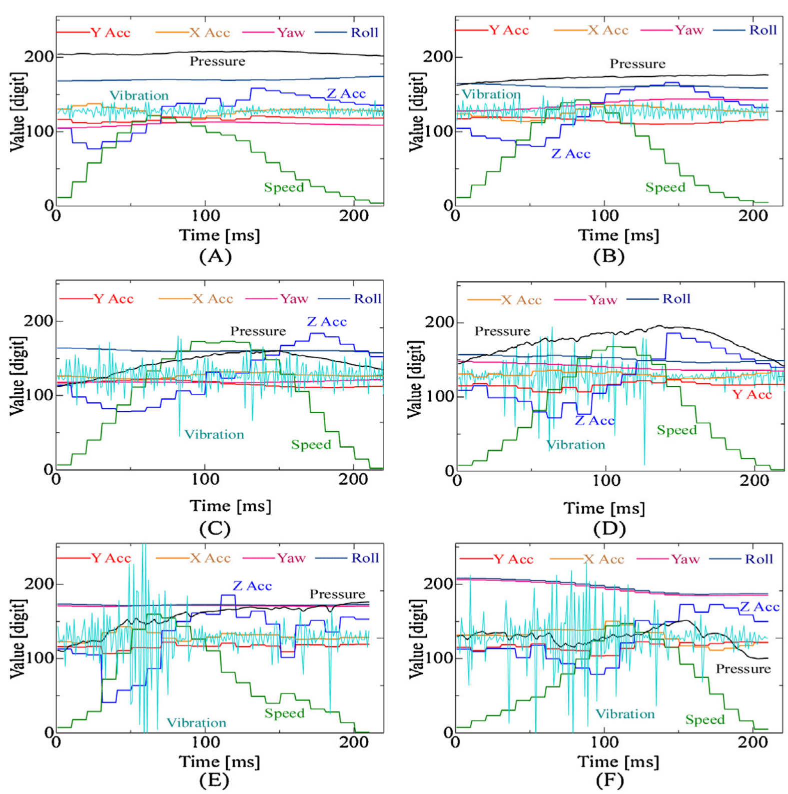
To recognize and properly measure the roughness of a given object, we use a CNN to comprehensively combine the multiple pieces of information that we obtain, these pieces being pressure, vibration, speed, and posture. A CNN is a machine learning method used primarily to recognize objects within images, achieving very good results in those fields [16–18]. Using a CNN takes advantage of the fact that the input data have spatial features and constrain the architecture more sensibly. The CNN has an input layer, an output layer, a convolutional layer, a pooling layer, and a fully connected layer. The input layer takes in as input two-dimensional data, typically an image. The convolution (Conv) layer optimizes parameters to extract features that effectively recognize data during training. The pooling (Pooling) layer then performs a down-sampling operation along the spatial dimension. Both the convolution and pooling layers are repeated to create a deep architecture. Finally, a fully connected (FC) layer and an output layer together classify and label features to produce output.

For our system, pressure and motion information is obtained via the sensor, which cannot be treated as an image (i.e., cannot be treated simply as data with specific spatial features). Therefore, in this study, to successfully use a CNN, multiple sensor data are continuously obtained over time and arranged into two-dimensional data, forming an input that resembles an image. Fig. 4 shows measured data, which is then translated into the image data shown in Fig. 5(a). Here, the data is converted into successive eight-bit chunks of data to properly convert the data to an image. The measured data of each roughness scales changed to eight-bit is shown in Fig. 5(b).

We use a digital filter to separate out the force data and vibration data from the pressure data. Force data are obtained using a high-pass filter with cutoff frequency 150 Hz, and vibration data are obtained by using a low-pass filter with cutoff frequency 150 Hz. Here, the vibration range is ± 0.1 V. The vibration data substantially vary when the sensor is accelerated. Further, each image is constructed from the measured data at 1 ms intervals. In the resulting image, three-axis acceleration, speed, force, and two attitude data are sandwiched between vibration data. Our aim here is to allow the vibration data to influence each data point by the filtering that will occur within the convolutional layer. The image is prepared



(a)



(b)

Fig. 5. Image converted from our measurement data. (b)(A) 1.6 µm. (b)(B) 3.2 µm. (b)(C) 12.5 µm. (b)(D) 18 µm. (b)(E) 25 µm.

with measurement data of 50 ms as one. Thus, the horizontal axis is time, and vertical axis are measurement data which are yaw, roll, speed, force, X—Y—Z acceleration, and vibration.

Given the above, the image becomes a matrix of 15 measurement data by 50 1-ms snapshots. Fig. 6 shows the architecture of the CNN in our proposed method. In the Conv1, Conv2, Conv3, and

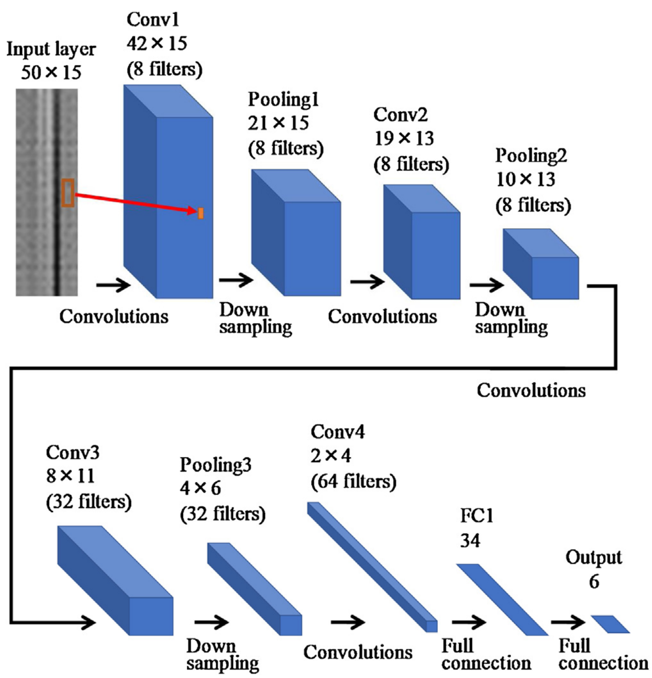


Fig. 6. Architecture of CNN for time-series data of the sensor.

Table 1
Roughness recognition results from sensor data by one user. (Percentage of each class estimated for input class.).

	Estimated Class Label					
Input Data Label	1.6	3.2	6.3	12.5	18	25
1.6	82.1	9.4	6.4	0.4	0.9	0.7
3.2	10.2	69.6	15.7	1.3	1.8	1.4
6.3	8.6	13.1	72.2	2.0	2.5	1.5
12.5	0.9	1.3	2.4	59.9	26.3	9.2
18	0.9	0.9	2.3	21.3	64.3	10.4
25	0.8	0.7	0.6	8.6	9.6	79.7

Conv4 layers, each set of data is filtered in the time series. In the convolution layer, time-local convolution is also performed to extract features in the time direction. In Pooling1, Pooling2, and Pooling3, each set of data performed down-sampling to be strengthen against time change, suppress over learning, and reduce calculation cost. FC1 and Output together classify and label features to produce output.

3. Experimental results and discussion

In this section, we report on our results in applying our proposed method to objects with the different roughness values depicted in Fig. 3, thereby verifying the effectiveness of our approach. In our experiments, we had one user move the sensor along each object in the direction of the arrow shown in Fig. 3 a total of 1200 times; with six roughness values, we acquired data for 7200 trials. The images were prepared with the measured data every 50 ms. As a result, we prepared a total of approximately 1,350,000 sample images. Of these, we used approximately 1,080,000 samples for the learning phase of the CNN. For our experiments, we then evaluated our CNN using the remaining 270,000 samples.

Table 1 shows recognition results for our roughness recognition method when the experimental data acquired by one user are divided into learning data and evaluation data. In this experiment, the user moves the sensor by hand. Thus, the speed and posture of the sensor are randomly changed. In the table, we show a correct recognition rate for each estimated label for each piece of input data. If the rate on the diagonal is higher, it is better, in the sense that the CNN produced an accurate result. Note that the average correct recognition rate here was 71%. From the table, we also observe that the misrecognized results were relatively close to the correct values. Given these results, we confirm that tactile recognition is functioning effectively here.

Table 2
Roughness recognition results from sensor data across five users. (Percentage of each class estimated for input class).

	Estimated Class Label					
Input Data Label	1.6	3.2	6.3	12.5	18	25
1.6	39.3	36.5	13.3	2.3	3.2	5.3
3.2	17.8	20.3	53.4	2.4	4.3	1.9
6.3	12.1	26.5	50.0	1.5	8.0	1.8
12.5	3.0	4.1	5.0	40.2	34.6	13.0
18	1.7	3.5	7.7	27.9	45.5	13.6
25	3.6	2.6	4.3	15.7	18.4	55.3

A key feature of our proposed sensor setup is that a user can recognize the roughness of a given object by moving the sensor by hand without the need for a driving device. We conducted experiments to measure the accuracy of this roughness recognition technique for individual users. More specifically, we had five users perform the same experiments using our sensor. Each user moved the sensor along each object in the direction of the arrow shown in Fig. 3 a total of 10 times, thereby allowing us to collect data covering 300 trials. Images were prepared with the measured data every 50 ms, and as a result, we conducted our evaluation using data from approximately 80,000 sample images prepared using our CNN.

Table 2 shows recognition results of this roughness recognition method. We observe here an average recognition rate of 42%. The recognition rate of Table 2 is substantially lower than that of Table 1, primarily because different characteristics of the data are included due to different attitudes and movements of the sensor for each user. Nonetheless, the misrecognized results are relatively close to the correct values.

We believe that the recognition rate can be improved here by incorporating various data from a variety of users at the learning phase of the CNN. Further, sufficient movements and attitudes of the sensor may not be acquired given the current sampling rate of the acceleration sensor was only 100 Hz in our experiments. In our future work, we therefore plan to increase the sampling frequency and also improve the overall sensor structure to improve the resulting correct recognition rate in our experiments.

4. Conclusion

In this paper, we proposed a method for recognizing the roughness of objects using a CNN. A key feature of our proposed sensor system is that a user can recognize the roughness of a given object by moving the pen-type sensor by hand without a separate driving device. Our prototype sensor setup consisted of a pressure sensor and an acceleration sensor. Therefore, our sensor system design is simple and easily constructed at a low cost. In our system, roughness is analyzed via a CNN that uses the force, vibration, acceleration, speed, and attitude of the sensor. For our experiments, we applied our proposed method to six test objects with different levels of roughness. The total average recognition rate for the experimental data continuously acquired by one user, which were categorized into learning and evaluation data, was 71%, showing positive results. Further, the overall average recognition rate across five separate users without considering each individual using the sensor system was 42%. Given our results, the problem of roughness recognition for each individual remains, but we were successful in showing the possibility of roughness recognition via our work. In our future work, we plan to improve the sensor structure and increase the sampling rates of the associated sensors. We also plan to further examine the network structure of the CNN.

Acknowledgement

This work was supported in part by funds (No. 185009) from the Central Research Institute of Fukuoka University.

References

[1] E. Asaga, K. Takemura, T. Maeno, A. Ban, M. Toriumi, Tactile evaluation based on human tactile perception mechanism, Sens. Actuators A 203 (2013) 69–75.
[2] A. Song, Y. Han, H. Hu, J. Li, A novel texture sensor for fabric texture measurement and classification, IEEE Trans. Instrum. Measur. 63 (7) (2014) 1739–1747.
[3] W. Liu, P. Yu, C. Gu, X. Cheng, X. Fu, Fingertip piezoelectric tactile sensor array for roughness encoding under varying scanning velocity, IEEE Sens. J. 17 (21) (2017) 6867–6879.
[4] Y. Tanaka, Y. Ueda, A. Sano, Roughness evaluation by wearable tactile sensor utilizing human active sensing, Mech. Eng. J. 3 (2) (2016) 15–00460.
[5] X. Ye, B. Choi, H. Choi, S. Kang, Pen-type sensor for surface texture perception, in: Proc. of the 16th IEEE International Symposium on Robot and Human Interactive Communication, 2007, pp. 642–647.
[6] H. Liu, F. Sun, B. Fang, S. Lu, Multimodal measurements fusion for surface material categorization, IEEE Trans. Instrum. Measur. 67 (2) (2018) 246–256.
[7] A. Schmitz, Y. Bansho, K. Noda, H. Iwata, T. Ogata, S. Sugano, Tactile object recognition using deep learning and dropout, in: Proc. of IEEE-RAS International Conference on Humanoid Robots, 2014, pp. 1044–1050.
[8] W. Yuan, C. Zhu, A. Owens, M.A. Srinivasan, E.H. Adelson, Shape-independent hardness estimation using deep learning and a GelSight tactile sensor, in: Proc. of the 2017 IEEE International Conference on Robotics and Automation, 2017, pp. 951–958.
[9] M. Kerzel, M. Aliy, H.G. Ng, S. Wermter, Haptic material classification with a multi-channel neural network, in: International Joint Conference on Neural Networks, 2017, pp. 439–446.
[10] Y. Gao, L.A. Hendricks, K.J. Kuchenbecker, T. Darrell, Deep learning for tactile understanding from visual and haptic data, in: Proc. of ICRA. IEEE, 2016, pp. 536–543.
[11] Z. Erickson, S. Chernova, C.C. Kemp, Semi-supervised haptic material recognition for robots using generative adversarial networks, in: Proc. of Conference on Robot Learning, 2017, pp. 157–166.
[12] H. Zheng, L. Fang, M. Ji, M. Strese, Y. Ozer, E. Steinbach, Deep learning for surface material classification using haptic and visual information, IEEE Trans. Multimedia 18 (12) (2016) 2407–2416.
[13] M. Strese, Y. Boeck, E. Steinbach, Content-based surface material retrieval, in: Proc. of IEEE World Haptics Conference 2017, 2017, pp. 352–357.
[14] H. Orii, S. Tsuji, T. Kouda, T. Kohama, Tactile texture recognition using convolutional neural networks for time-series data of pressure and 6-axis acceleration sensor, in: Proc. of the 2017 IEEE International Conference on Industrial Technology, 2017, pp. 1076–1080.
[15] H. Orii, S. Tsuji, T. Kouda, T. Kohama, Recurrent neural network for tactile texture recognition using pressure and 6-axis acceleration sensor data, in: Proc. of the 2017 IEEE Region 10 Conference, 2017, pp. 2012–2016.
[16] Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proc. of the IEEE 86 (11) (1998) 2278–2324.
[17] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, Adv. Neural Inf. Process. Syst. (2012) 1097–1105.
[18] X. Ding, Q. He, Energy-fluctuated multiscale feature learning with deep ConvNet for intelligent spindle bearing fault diagnosis, IEEE Trans. Instrum. Measur. 66 (8) (2017) 1926–1935.

Biographies

Satoshi Tsuji received his Ph.D. degree from Saga University, Saga, Japan in 2012. He is currently an Assistant Professor in the Department of Electrical Engineering, Fukuoka University. His research interests include tactile sensing systems and human interface.

Teruhiko Kohama received his Ph.D. degree from Kyushu University, Fukuoka, Japan in 1999. He is currently an Associate Professor in the Department of Electrical Engineering, Fukuoka University, Japan. He has been active in the fields of power electronics, analog circuits, and sensing systems.