ICIC Express Letters
Part B: Applications
Volume **12**, Number **7**, July **2021**

ICIC International ©2021 ISSN 2185-2766
pp. **651–658**

TACTILE TEXTURE RECOGNITION USING RECURRENT NEURAL NETWORKS WITH PEN-TYPE SENSOR SYSTEM

Hideaki Orii*, Satoshi Tsuji and Takaharu Kouda

Department of Electrical Engineering
Faculty of Engineering
Fukuoka University
8-19-1 Nanakuma, Jonan-ku, Fukuoka 814-0180, Japan
*Corresponding author: oriih@fukuoka-u.ac.jp

Received November 2020; accepted February 2021

Abstract. *Tactile texture is an important factor that determines the impression of an object. It is not only the appearance of an object but also its tactile texture that determines whether people find it attractive. The design of tactile texture is an important factor in the development of commercial products. To measure the feeling of an object, a numerical evaluation method for tactile texture is required. The measurement is generally performed using a large-scale apparatus. In this paper, we propose a tactile texture recognition method using an original pen-type sensor system and recurrent neural networks. In the proposed framework, the tactile texture information is obtained by analyzing time-series data obtained from a pressure sensor and 6-axis acceleration sensor attached to a rod-like rigid body. The key idea of the proposed system is to provide a simple and inexpensive method of measuring tactile sensation, which generally requires large equipment and high-precision sensors. In the experiments, we applied the proposed system to several objects with different tactile textures, and confirmed its effectiveness.*
Keywords: Tactile texture recognition, Deep neural networks, 6-axis acceleration sensor

1. **Introduction.** The sense of touch of an object is an important factor that determines the texture and impression of the object. For example, when purchasing products such as clothes and furniture that come into contact with skin during use, the color of the product as well as its texture, such as how it feels to the skin, often influence the purchase decision. In other words, the tactile feel of an object is one of the important attributes of a product, and it is necessary to design for tactile texture when designing a product. People use a variety of words to describe tactile texture, such as "smooth", "silky", "sandy", and "rough". However, it is difficult to objectively measure and quantify the tactile sensation because it is subjective and depends on the surrounding environmental parameters such as temperature and moisture. Therefore, to measure the tactile feel of an object, a numerical evaluation method for tactile texture is required.

Some technologies to measure tactile texture have been developed recently. E. Asaga et al. [1] developed a tactile evaluation system based on the human tactile perception mechanism. In this system, the object to be evaluated is fixed on a sample table rotated with a constant speed, and a piezoelectric element is placed in contact with the object surface with a constant normal force. The piezoelectric element can measure the characteristics of the tactile texture of the object as vibration information during the experiment. M. Tanaka et al. [2] and Y. Tanaka et al. [3] proposed an active haptic sensor for monitoring skin conditions. This sensor is composed of a strain gage, a PVDF film, a protective surface layer, and lace. The measurement is performed by rotating and sliding on a target surface. Although the tactile information measured by these methods has high precision and recall, the system requires extensive equipment and is very expensive. M. Nagano and T.

DOI: 10.24507/icicelb.12.07.651

Fukami [4] proposed an image-based skin texture evaluation system using a convolutional neural network. The system evaluates the beauty of skin based on the images captured by a microscope. We also proposed a tactile texture recognition method [5]. The method used convolutional neural networks, and had a certain level of recognition performance, but there were some problems: for example, the network accepted only fixed-length input data.

In this paper, we propose a simple and inexpensive system for tactile recognition using an original pen-type sensor and recurrent neural networks (RNNs). In the proposed framework, tactile texture information is obtained by analyzing time-series data of a pressure sensor and a 6-axis acceleration sensor as shown in Figure 1. Thus, the system configuration is simple, and it is possible to construct the system inexpensively. Our system realizes a tactile texture recognition system at a reasonable cost.

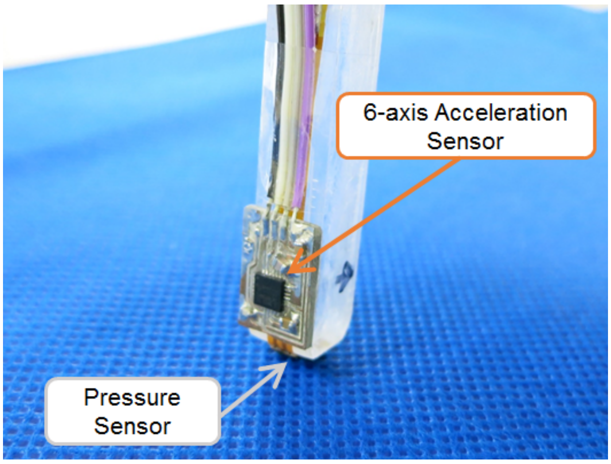


FIGURE 1. Sensors for the proposed method. When a user moves the stick while pressing against an object, the characteristics dependent on the texture of the surface can be obtained by the sensors.

The rest of this paper is organized as follows. In Section 2, we give an overview of neural networks and the details of the proposed system. Section 3 presents the experimental results and discussions. Finally, the conclusions are given in Section 4.

2. **Proposed Method.** Figure 2 shows the sensors used in the proposed method. In the system, a pressure conductive sensor (SF-R-3, manufactured by Inaba Rubber Co., Ltd.) is attached to the tip of the acrylic rod, and a 6-axis acceleration sensor is attached to the side of the rod. Figure 3 shows the data obtained from the sensors. The 6-axis acceleration sensor can detect 3-axis acceleration and 3-axis angular velocity with one detecting element. The data obtained by the 6-axis acceleration sensor changes with the position and movement of the rod, and data obtained by the pressure sensor changes with the pressure and vibration applied to the tip of the rod. When the user moves the stick while it is pressed against an object, a characteristic dependent on the texture of its surface can be obtained by the sensors. It is believed that these data are sufficient to discriminate the tactile texture, though they contain complicated temporal numerical changes.

In this study, we use RNNs to analyze the data obtained by the sensors. RNN is a machine learning method that is primarily used for natural language processing, and has remarkable achievements to its credit in those fields [6, 7]. RNN automatically designs feature descriptors effective for recognition from time-series data, and uses them in recognition. To analyze the complicated temporal numerical changes in the data obtained by sensors, we apply RNN for time-series data of pressure and a 6-axis acceleration sensor.

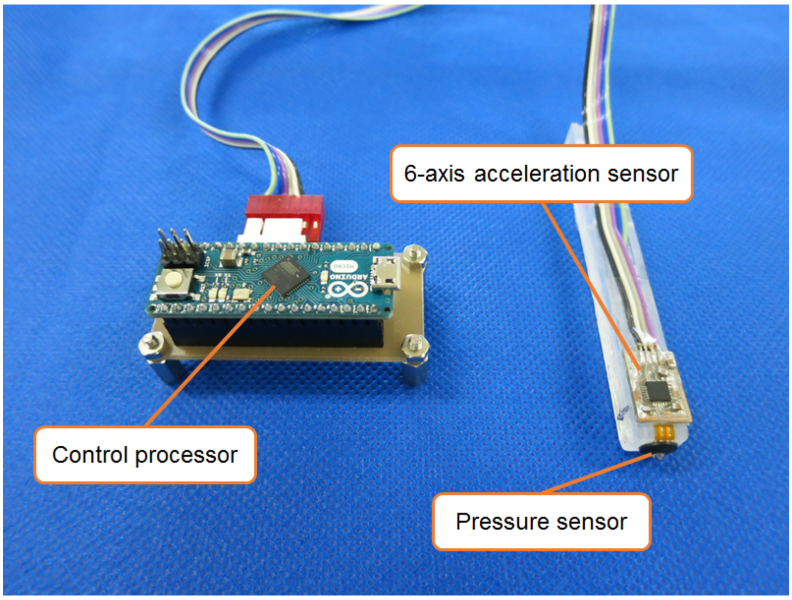


FIGURE 2. Sensors used in the proposed method

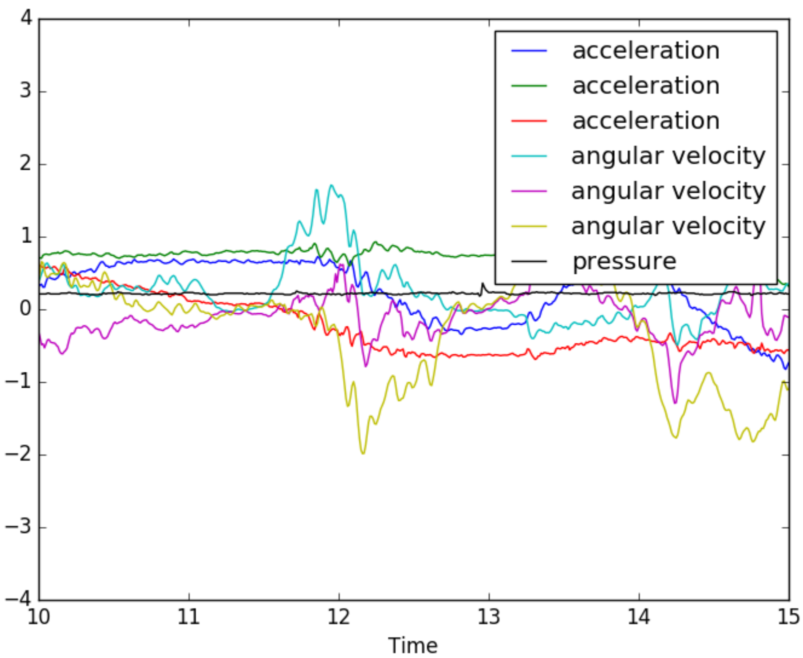


FIGURE 3. (color online) Data acquired from the sensors

An overview of neural networks and the RNN is given in Sections 2.1 and 2.2. The details of the RNN model for time-series sensor data and the entire algorithm of the proposed method are described in Section 2.3.

2.1. **Neural networks.** Neural network models are essentially simple mathematical models defining a function $g : \boldsymbol{X} \to \boldsymbol{Y}$. In a neural network, the j-th neuron in a layer calculates and outputs a value y_j as follows:

$$y_j = f(\Sigma_i x_i \cdot w_{ij} + b_j). \tag{1}$$

Here, x_i is the i-th input value for a layer. w_{ij} is the weighted value of the j-th neuron for the i-th input value. b_j is the bias value of the j-th neuron. $f(\cdot)$ is the transfer function of the j-th neuron. Therefore, when the number of neurons of a layer is M, an output vector

$\boldsymbol{y} = [y_1, y_2, \ldots, y_M]^T$ of this layer with respect to an input vector $\boldsymbol{x} = [x_1, x_2, \ldots, x_N]^T$ can be obtained as follows:

$$\boldsymbol{y} = f\left(\begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1N} \\ w_{21} & w_{22} & \cdots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{M1} & w_{M2} & \cdots & w_{MN} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_M \end{bmatrix}\right) \tag{2}$$

$$= f(\boldsymbol{W}\boldsymbol{x} + \boldsymbol{b}). \tag{3}$$

As mentioned earlier, the relationship between an input vector and an output vector in a layer is defined by the weighted value matrix $\boldsymbol{W}$, bias vector $\boldsymbol{b}$, and transfer function $f(\cdot)$. In a multilayer neural network, the neurons in each layer calculate the output vector based on the previous layer output and send the vector as the next layer input. Therefore, an input vector of a multilayer neural network is mapped a number of times from the input layer to the output layer. The training of a neural network is generally supervised with a training data set consisting of many pairs of an input vector and its desirable output vector. The weighted value matrix $\boldsymbol{W}$ and bias vector $\boldsymbol{b}$ of each layer are updated based on the error of the output vector. This is called the backpropagation algorithm. To use this algorithm, the transfer functions of each layer must be differentiable.

To recognize a data label with the neural network, we can use the softmax cross entropy function as a loss function that quantifies the agreement between the predicted scores and the ground truth labels. For example, when a vector in data is inputted to a network, we can assume that the i-th element of an output vector of neurons of the final layer expresses its scores for the class "i" as follows:

$$S(y_i) = \frac{\exp(y_i)}{\sum_{j=0}^{M-1} \exp(y_j)}. \tag{4}$$

Here, $S(y_i)$ is the probability of class "i", and y_i is the output value of the i-th neuron of the final layer of the network. M is the number of classes. The softmax cross entropy D_{SC} is defined as follows:

$$D_{SC}(S, L) = -\sum_{j=0}^{M-1} \log\left(S(y_i)\right) L_i. \tag{5}$$

Here, L is a "true" distribution on class labels. The softmax classifier is hence minimizing the cross entropy between the estimated class probabilities ($S(y_i)$) and the "true" distribution L, which in this interpretation is the distribution where all the probability mass is on the correct class (i.e., $L = [0, 1, \cdots, 0]$ contains a single 1 at the y_i-th position.).

After training of the network parameter, we can estimate the class of an input data as the class with respect to the neuron outputting the maximum value.

2.2. **Recurrent neural network.** The RNN is a neural network that can handle time-series data such as natural language and sound signal processing [7, 8, 9]. I. Sutskever et al. proposed a method to translate English sentences into French using RNN [7]. RNN is also applied to image recognition.

In the input layer of an RNN, samples are input in time series from time $t = 0$. The response value of the input layer at time t and that of the intermediate layer at time $t - 1$ are input to the intermediate layer at time t. In this way, by inputting the samples in a chain and inputting the response value of the intermediate layer at time $t - 1$ to the intermediate layer at the time t, recognition corresponding to the time series becomes feasible. To train the RNN, long short term memory networks (LSTMs) [6] are commonly used. Backpropagation through time (BPTT) used in RNN updates the parameters to trace back the time series, but the error gradient becomes explosively high or disappears as the time series to be handled becomes long. On the other hand, LSTM uses a memory unit as the middle layer of the network. The memory unit includes an input gate, a forgetting gate, and an output gate. The input gate is coupled to the units, which are response values input to the intermediate layer at time t. On the input gate, the response value to be input to the middle layer at time t is controlled. The forgetting gate adjusts to what extent the unit forgets past memories. When the forgetting gate is close to 0, it forgets the previous memory, whereas if it is close to 1, it will be propagated to the next memory unit, leaving the previous memory. The output gate controls whether or not the output of the final unit is transmitted outside. As a result, when calculating the learning error, unnecessary errors are not calculated by LSTM. Because LSTM can propagate error efficiently using three gates, it can handle long-term time-series data better than BPTT. Therefore, it may be possible to reflect memories of the distant past to the output of the RNN, making highly accurate recognition feasible.

2.3. **Algorithm of the proposed method.** In our system, information on pressure and motion is obtained from the sensors. We treat sensor data obtained continuously in time collectively as one data, and input the data to the RNN sequentially. Figure 4 shows the architecture of an RNN in the proposed method. Sensor data comprise 8 dimensions: x, y, z axial acceleration and angular velocity (6 dimensions) and pressure information (duplicating information from one sensor; 2 dimensions). Data of 50 consecutive time points are collectively input sequentially and processed by the RNN. The class of input data could be obtained as the output value of the network corresponding to the last input.

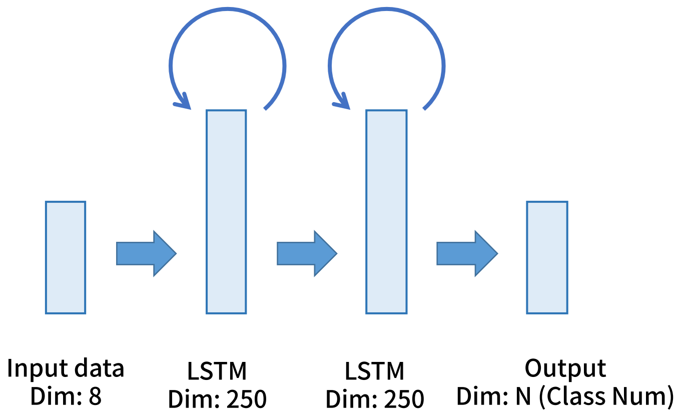


Figure 4. Architecture of RNN for time-series data of pressure and 6-axis acceleration sensor

3. **Experimental Results and Discussions.** In this section, we apply the proposed method to objects having various tactile textures, to verify its effectiveness. In the experiments, approximately 20,000 samples of data were prepared for the training of the neural network, and the training of the RNN was performed using it. The evaluation experiment was conducted using data of approximately 10,000 samples prepared separately. The sampling frequency of the sensor is 100 Hz.

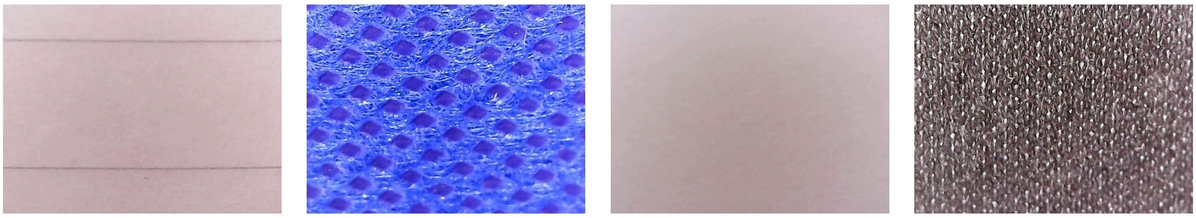
Figure 5 shows the test objects used in this experiment, and Table 1 shows the recognition results of our tactile recognition method. In the table, "Class 0" is "a desk", and "Class 1" is "a cardboard". "Class 2" is "a non-woven cloth", and "Class 3" is "a sheet of paper". The higher the rate on the diagonal, the better the result. As can be observed from the table, tactile recognition is functioning effectively.

Furthermore, we conducted another experiment to confirm the effectiveness of the proposed method. Figure 6 shows the magnified images of the surfaces of the test objects in this experiment, and Figure 7 shows sensor data of the test objects. Table 2 shows the recognition results for the objects of Figure 7. In the table, "Class 0" is "non-contact",

Figure 5. Test objects: From the left, a desk (the smooth surface on which other objects are placed), a sheet of paper (a notebook), a non-woven cloth (a bag), and a cardboard

Table 1. Tactile texture recognition result 1 (percentage of each class estimated for input class)

Label of input data	Estimated class label			
	Class 0	Class 1	Class 2	Class 3
Class 0	0.865	0.074	0.026	0.035
Class 1	0.008	0.982	0.006	0.004
Class 2	0.025	0.122	0.811	0.043
Class 3	0.02	0.032	0.005	0.944



(a) (b) (c) (d)

Figure 6. Magnified images of surfaces of test objects of Experiment 2: (a) A sheet of paper (Class 1), (b) a non-woven cloth (Class 2), (c) a plastic folder (Class 3), and (d) a mouse pad (Class 4)

and "Class 1" is "a sheet of paper". "Class 2" is "a non-woven cloth", and "Class 3" is "a plastic folder". "Class 4" is "a mouse pad". "Non-contact" denotes data moved repeatedly back and forth in a state where the pressure sensor section does not make contact. In this experiment too, good discrimination performance was obtained as in the previous experiment. However, with regard to "Class 1 (a sheet of paper)", there were many wrong classifications. Because these textures depend not only on the shape of the surface but also on the repulsive force, etc., we think that the results were greatly affected by how the user pushed the sensor rod during the experiment. We think that this problem can be solved by increasing the number of pressure sensors that are brought into contact with the object or by increasing the sampling rate of the sensor.

4. **Conclusions.** We proposed a novel recognition method for tactile texture recognition using long short-term memory RNNs. In the proposed framework, the tactile texture information is obtained by analyzing time-series data of a pressure sensor and 6-axis

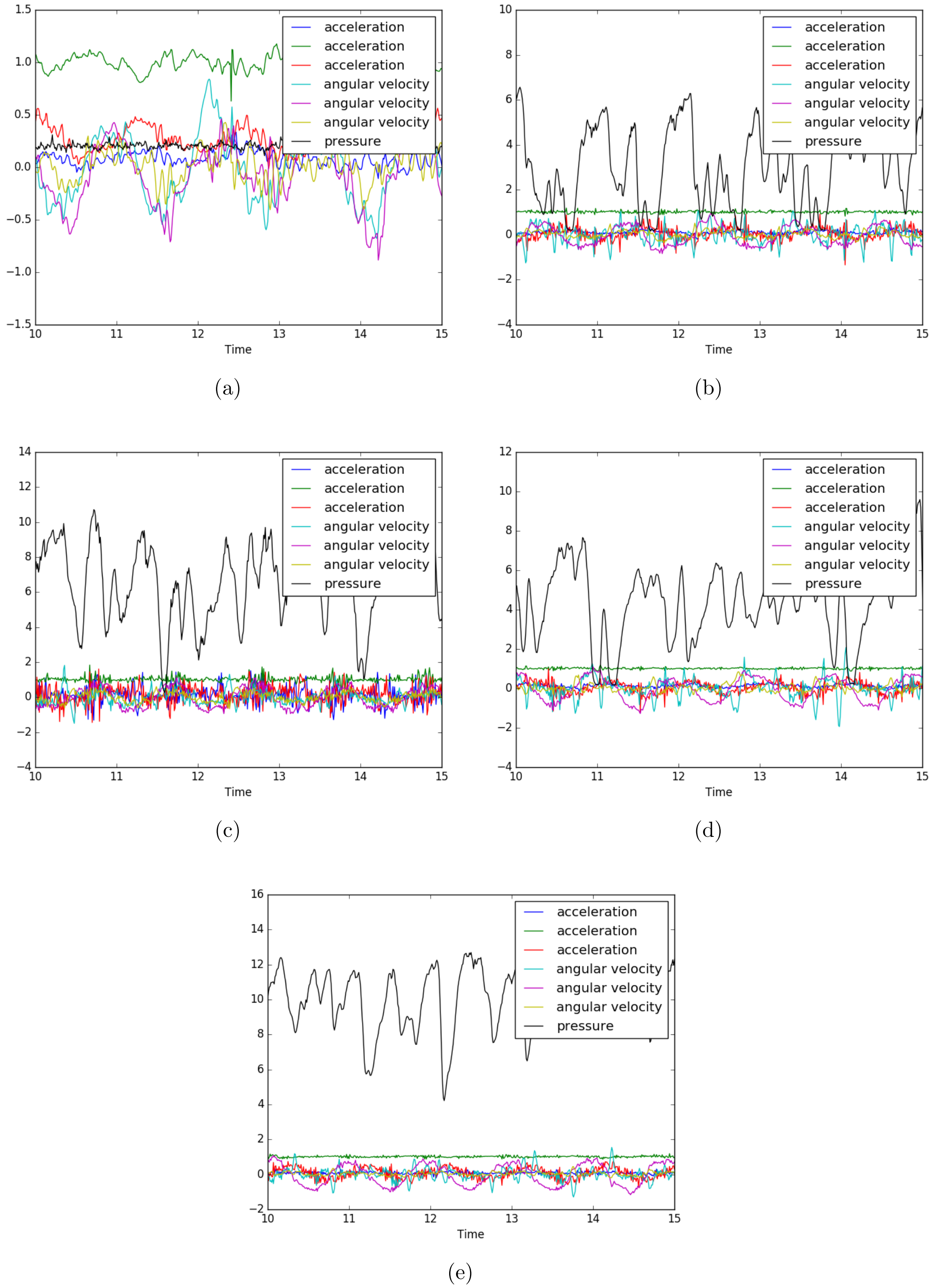


Figure 7. (color online) Sensor data with respect to time: (a) Data repeatedly moved back and forth in a state where the pressure sensor section is not brought into contact (Class 0), (b) sensor data of the sheet of paper (Class 1), (c) sensor data of the non-woven cloth (Class 2), (d) sensor data of the plastic folder (Class 3), and (e) sensor data of the mouse pad (Class 4)

TABLE 2. Tactile texture recognition result 2 (percentage of each class estimated for input class)

Label of input data	Estimated class label				
	Class 0	Class 1	Class 2	Class 3	Class 4
Class 0	1.000	0.000	0.000	0.000	0.000
Class 1	0.000	0.017	0.017	0.917	0.050
Class 2	0.000	0.102	0.847	0.000	0.051
Class 3	0.000	0.183	0.000	0.517	0.300
Class 4	0.000	0.153	0.000	0.017	0.831

acceleration sensor on our original pen-type device. In the experiment, we apply the proposed method to testing objects having various tactile textures. The results showed the effectiveness of the proposed method.

For future studies, we would like to optimize the number of pressure sensors and the sampling rate of the sensors. We will further examine the network structure of RNNs.

Acknowledgment. This work was supported in part by funding from Fukuoka University (Grant Nos. 185009 and 215010).

REFERENCES

[1] E. Asaga, K. Takemura, T. Maeno, A. Ban and M. Toriumi, Tactile evaluation based on human tactile perception mechanism, *Sensors and Actuators A: Physical*, vol.203, pp.69-75, 2013.

[2] M. Tanaka, H. Sugiura, J. L. Lévêque, H. Tagami, K. Kikuchi and S. Chonan, Active haptic sensation for monitoring skin conditions, *Journal of Materials Processing Technology*, vol.161, pp.199-203, 2005.

[3] Y. Tanaka, M. Tanaka and S. Chonan, Development of a sensor system for collecting tactile information, *Microsystem Technologies*, vol.13, no.8, pp.1005-1013, 2007.

[4] M. Nagano and T. Fukami, Development of a skin texture evaluation system using a convolutional neural network, *International Journal of Innovative Computing, Information and Control*, vol.16, no.5, pp.1821-1827, 2020.

[5] H. Orii, S. Tsuji, T. Kouda and T. Kohama, Tactile texture recognition using convolutional neural networks for time-series data of pressure and 6-axis acceleration sensor, *Proc. of the 2017 IEEE International Conference on Industrial Technology (IEEE ICIT 2017)*, pp.1076-1080, 2017.

[6] S. Hochreiter and J. Schmidhuber, Long short-term memory, *Neural Comput.*, vol.9, no.8, pp.1735-1780, 1997.

[7] I. Sutskever, O. Vinyals and Q. V. Le, Sequence to sequence learning with neural network, *Proc. of the 27th International Conference on Neural Information Processing Systems (NIPS'14)*, pp.3104-3112, 2014.

[8] A. Graves, S. Fernández and J. Schmidhuber, Bidirectional LSTM networks for improved phoneme classification and recognition, *ICANN'05: Proc. of the 15th International Conference on Artificial Neural Networks: Formal Models and Their Applications*, pp.799-804, 2005.

[9] A. Graves, A. Mohamed and G. Hinton, Speech recognition with deep recurrent neural networks, *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp.6645-6649, 2013.