Editorial

Machine Learning-Based Non-Contact Input Device Using a Capacitive Proximity Sensor

KASHIWAO Tomoaki*

1. Introduction

In recent years, triggered by the COVID-19 pandemic, there has been a rapidly growing interest in non-contact interfaces that can be operated without physical touch. Control panels in public facilities and industrial equipment that are touched by a large number of people have been recognized as posing an infection risk. The need for non-contact operation has become widely recognized. Against this social backdrop, non-contact input technology is positioned not merely as a means to improve convenience, but as an essential technology to ensure people's safety and hygiene. However, implementing the technology in real-world applications involves several challenges. The technology may need to adapt to the complex system configuration requirements of some sites. Additionally, some types of technology subject to installation constraints may not always be practical.

Examples of non-contact input technology include image processing sensors with cameras and optical sensors. These types may present challenges related to lighting conditions, field-of-view constraints, complex system configurations, and high costs. In contrast, capacitive proximity sensors have relatively simple circuit configurations, are inexpensive and compact, and have low power consumption, which makes them easier to install. The capacitive proximity sensor is designed to detect tiny changes in capacitance caused by an object approaching its electrodes. The sensor is highly sensitive to conductors, such as the human body[1), 2), 3)]. Leveraging these features, the sensor can detect a hand approaching or moving without ever directly touching the surface of the sensor.

However, the output of capacitive proximity sensors is easily affected by the position, direction, and velocity of the object, as well as by individual differences and environmental conditions. Analytically modeling the relationship between the sensor output and motion is difficult. In particular, when using time-series signals obtained from multiple electrodes to detect an object's motion, it is difficult to ensure sufficient recognition performance only through a simple threshold determination system or rule-based approach alone.

This article will present a three-dimensional capacitive proximity sensor with three electrodes. After summarizing its structure and output characteristics, the article will explain the concept of motion recognition using time-series signals obtained from multiple electrodes[4), 5), 6), 7), 8), 9)]. The article will also demonstrate how machine learning can flexibly recognize complex motion patterns that have been difficult to recognize with template matching. Note that this sensor unit was designed by Professor Emeritus Mikio Deguchi of the National Institute of Technology (KOSEN), Niihama College, and developed through joint research.

2. Sensor Structure and Output

2.1 Structure of the Sensor Unit

The capacitive proximity sensor unit presented in this article has a round structure with three detecting electrodes arranged along the circumference (Fig. 1). The three electrodes, located radially from the center of the sensor substrate, form an equilateral triangle. Each of these electrodes operates as an independent detection channel.

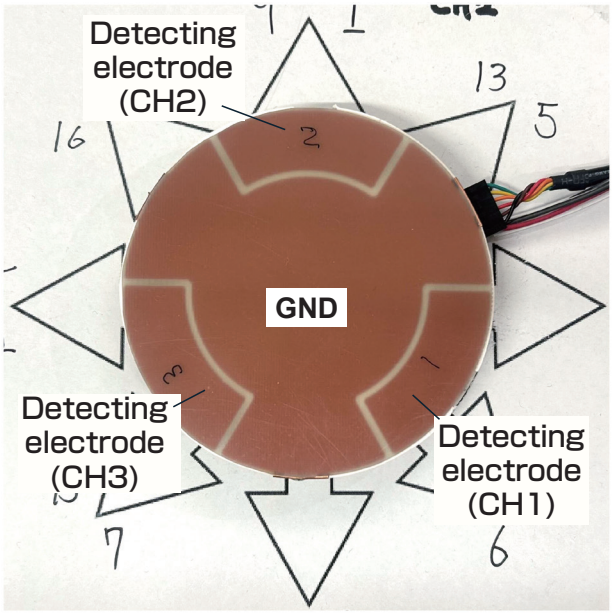

Fig. 1 Sensor unit

*Associate professor, Electronics and Information/Communication Course, Department of Electric and Electronic Engineering, KINDAI UNIVERSITY

Each detecting electrode detects changes in the electric field formed in the surrounding space to recognize nearby approaching objects. As a person's hand approaches the upper space of the sensor, the stray capacitance formed between the electrode and the hand changes. This results in changes in the detection signals of the individual channels. The use of three electrodes enables this sensor unit to collect directional information that a single electrode cannot obtain.

Because the sensor unit consists of a relatively small number of general-purpose electronic parts, it can easily be downsized and incorporated into a system. This structure offers advantages in terms of mountability and expandability, both of which are important in industrial applications.

2. 2 Sensor Output

This sensor unit is calibrated to detect objects, including hands, within a radius of approximately 10 cm (Fig. 2).

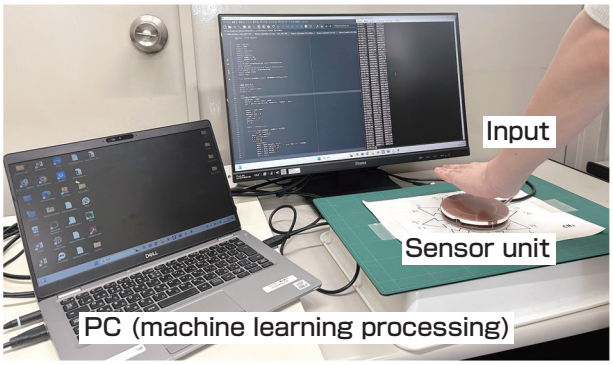

Fig. 2 Sensor unit

Fig. 3 shows the sensor output for the motion of a person's hand (approximately two seconds) with a total of 16 patterns (Fig. 4). The output signals from the three channels, which correspond to the three detecting electrodes of the sensor unit, were averaged and normalized on a scale from 0 to 1. The resulting graph is shown in Fig. 3. The output of each channel reflects changes in capacitance associated with the approach of the hand and its subsequent motion. The shape of the waveform depends on the direction and velocity of the motion. For instance, when a hand approaches a specific electrode, the corresponding channel's output changes first, followed by the other electrodes. Response timing and amplitude differences are key cues for estimating motion direction. Differences in motion speed manifest as temporal variations in the output waveforms.

These output waveforms are not always simple. They are affected by the size and position of the hand, motion variations, and the surrounding environment. Therefore, it is difficult to classify output waveforms according to rules defined by humans. Instead, they should be processed by taking into account the characteristics of the overall time-series data. This research uses machine learning to classify output waveforms into motion patterns. Time-series signals from the multiple channels are input into the machine learning system, which allows it to automatically learn the characteristics of the data. This improves the performance, flexibility, and expandability of motion recognition.

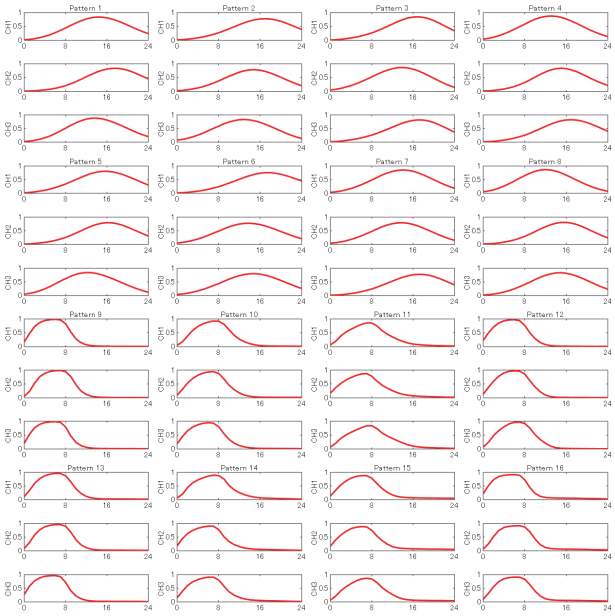

Fig. 3 Sensor output

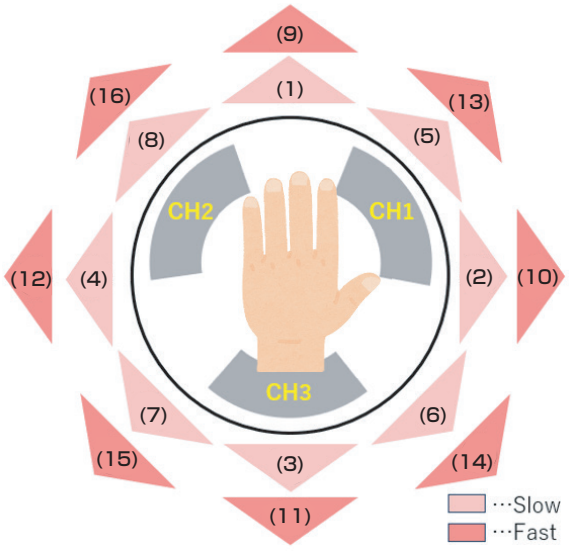

Fig. 4 Motion patterns of a hand over the sensor unit

3. Machine Learning

3. 1 Classification Models

As mentioned in the previous chapter, the output obtained from this sensor unit is observed as time-series signals from three channels. These signals contain information about the direction and velocity of hand motion. However, the signals are affected by individual differences, motion variations, and environmental conditions. The

relationship between the sensor output and the hand motion is non-linear and complex.

If data with these characteristics was evaluated using a rule-based system or simple template matching, the recognition performance would depend heavily on the motion conditions. Particularly when the number of motion patterns increases or similar motions must be distinguished, it is not practical for designers to adjust the rules one by one.

This research therefore adopts a data-driven approach with machine learning to identify the relationship between sensor outputs and the motion classes. Using machine learning models to automatically extract features included in time-series data from multiple channels enables distinguishing different motions.

This article classifies motions using four typical machine learning models:

(1) Linear Regression (LR)
(2) k-Nearest Neighbors (kNN)
(3) Support Vector Classifier (SVC), and
(4) Random Forest (RF)

Regardless of the model type, it is configured to receive feature input of time-series signals from three channels and output motion pattern classes. This approach achieves high recognition performance with a relatively small number of sensors, although the description of the detailed structure and parameters of the models is omitted here. Using machine learning offers advantages such as improved recognition accuracy and flexible addition and changes to motion patterns. When a new motion is defined, the system can be expanded by adding training data. Thus, the system can easily be customized to suit the application.

3. 2 Classification Evaluation

This research uses a confusion matrix to quantitatively evaluate the classification performance of the proposed approach. A confusion matrix summarizes the correspondence between actual and estimated classes (patterns). The matrix allows the user to visually identify whether the classification for each class is correct. A confusion matrix consists of the following four elements: True positive (TP) indicates the number of samples where data belonging to the target class has been correctly recognized as the relevant class. False positive (FP) indicates the number of samples where data belonging to another class has been incorrectly recognized as the target class. False negative (FN) indicates the number of samples where data belonging to the target class has been incorrectly classified as another class. True negative (TN) indicates the number of samples where data belonging to another class has been correctly identified as such. Based on these values, the following evaluation indices have been determined.

Accuracy is expressed by the following formula:

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN)$$

This formula indicates the percentage of samples that were correctly classified out of the total number of samples. The result is an index that represents the system's overall classification performance. However, accuracy alone is insufficient for evaluation in cases of too many classes or a biased distribution of classes.

Precision is expressed by the following formula:

$$\text{Precision} = TP / (TP + FP)$$

This index shows the percentage of samples classified as a certain class that actually belong to that class. It is useful for evaluating the performance of suppressing false positive (FP).

Recall indicates the percentage of samples that belong to a class and have been correctly detected. It is an index used to evaluate the false negative (FN) rate.

$$\text{Recall} = TP / (TP + FN)$$

F1_score is expressed by the following formula:

$$\text{F1_score} = (2 \times \text{recall} \times \text{precision}) / (\text{recall} + \text{precision})$$

It is the harmonic mean of precision and recall. It is a balanced evaluation index that takes both false positive and false negative into account. In multi-class classification problems involving many gesture patterns, such as the one addressed in this research, using a confusion matrix and multiple evaluation indices rather than using a single index makes it possible to more closely understand the characteristics of classifiers and the trend of false classifications. In reality, gestures with similar motion directions or speed variations are likely to be misclassified. They have actually been found to be misjudged as an adjacent class on the confusion matrix. These results imply the presence of non-linearity or temporal strain that cannot be grasped only with simple template matching methods alone.

4. Experimental Results

We divided the 15,308 datasets, which comprise sensor unit outputs and motion patterns of four students, into 10,715 for training and 4,593 for testing. We then trained the machine learning models with these datasets. To verify the effectiveness of the trained classification models, we classified the patterns of eight additional students (10 trials per pattern per student). These students were not involved in creating the training data. The subjects were asked to move their hands over the sensor in the defined directions and velocities for each pattern. The sensor outputs during these motions were recorded.

We applied the data obtained in this experiment to the machine learning models. As a result, we verified that multiple motion patterns can be reliably distinguished (Table 1). Notably, the results showed that unidirectional motions, as well as motions in diagonal directions or with speed variations, can be identified by determining the temporal differences in sensor outputs.

Table 1 Classification results of machine learning models

Model	Evaluation index			
	accuracy	precision	recall	F 1 _score
LR	0.8184	0.8183	0.8204	0.8177
KNN	0.9323	0.9284	0.9231	0.9317
SVC	0.9645	0.9649	0.9648	0.9647
RF	0.9251	0.9258	0.9252	0.9254

However, we found some cases with identification errors. Many of these cases involve motions in similar directions or with ambiguous speed variations. These cases are attributable to similar waveforms of sensor outputs. This implies that the system's classification results are affected by the motion definition and variations in motion practice, while capacitive proximity sensors are sensitive to subtle differences in human movements.

These results suggest that, although this system has sufficient potential as a non-contact input technology, the motion definition and training data must be more sophisticated to achieve commercialization.

5. In Closing

This article has presented a non-contact input technology that uses a three-dimensional capacitive proximity sensor with three electrodes. The article included explanations of the sensor structure, output characteristics, and the concept of motion recognition using machine learning. Capacitive proximity sensors can detect the approach and movement of the human body with their simple configuration. Using multiple electrodes allows for the collection of directional information.

Against the backdrop of the recent emerging social demand for non-contact interfaces, sensor technology must be used not just to detect physical quantities, but also as an input means. As demonstrated in this research, processing time-series signals from multiple channels with machine learning enables the flexible identification of complex motion patterns that were previously difficult to recognize with conventional methods. In other words, applying machine learning technology to sensors expands their functionality from mere detection elements to input devices that can handle human intention as input.

Due to its simple sensor configuration, this system can be relatively easily embedded in or expanded to other equipment. It is expected to be used as a non-contact control interface in industrial applications. Going forward, we will optimize the motion definition and expand the training data to improve the recognition and generalization further. We also plan to assess possible applications in real-world settings.

References

1) DEGUCHI M.: "A simple method for detecting very small changes in capacitance or inductance," Microelectronics Journal, vol. 101, p. 104802, 2020.
2) DEGUCHI M.: "Expansion of detectable area by floating electrodes in capacitive three-dimensional proximity sensor," International Journal on Smart Sensing and Intelligent Systems, vol. 14, pp. 1—11, 2021.
3) DEGUCHI M., KASHIWAO T., HIRO M., HAYASHI K.: "3D proximity sensing frame for non-contact operation panel," Proceedings of the 8th IEEJ international workshop on Sensing, Actuation, Motion Control, and Optimization (SAMCON2022), vol. 14. Saitama University, Japan: The Institute of Electrical Engineers of Japan, 2022, pp. 252—256.
4) HIRO M., HAYASHI K., KASHIWAO T., DEGUCHI M.: "Machine-learning- based prediction of two-dimensional position coordinates for a proximity sensor," Proceedings of 8th IEEJ International Workshop on Sensing, Actuation, Motion Control, and Optimization (SAMCON2022), vol. SS4-1, 2022, pp. 257—260.
5) KASHIWAO T., HIRO M., HAYASHI K., DEGUCHI M.: "Machine-learning- based prediction of the three-dimensional (3d) position coordinates of 3D proximity sensing frames," IEEJ Journal of Industry Applications, vol. 12, no. 4, pp. 800—807, 2023.
6) HAYASHI K., KASHIWAO T., HIRO M., DEGUCHI M.: "Hand motion classification in the 3D proximity sensors unit using machine learning," Proceedings of 9th IEEJ International Workshop on Sensing, Actuation, Motion Control, and Optimization (SAMCON2023), vol. TT7-1, 2023, pp. 579—582.
7) KASHIWAO T., HAYASHI K., HIRO M., OGINO R., DEGUCHI M.: "Machine learning-driven classification of hand motion for the 3D-proximity-sensors unit," IEEJ Journal of Industry Applications, vol. 13, no. 2, pp. 165—170, 2024.
8) OGINO R., KASHIWAO T., HAYASHI K., YAMAMOTO Y., HIGESHIRO K., DEGUCHI M.: "Hand-motion classification based on machine learning for the 3D-proximity-sensors unit," Proceedings of SICE Festival with Annual Conference 2024, 2024, pp. 438—441.
9) HIGESHIRO K., OGINO R., Q. Li, DEGUCHI M., KASHIWAO T.: "Hand-motion classification based on machine learning in three-dimensional capacitive proximity sensors," Proceedings of 11th IEEJ International Workshop on Sensing, Actuation, Motion Control, and Optimization (SAMCON2025), 2025, pp. 7-10.