Motion Estimation of Plush Toys through Detachable Acceleration Sensor Module and Machine Learning

Kaho Kato[1], Naoto Ienaga[1], and Yuta Sugiura[1]

[1] Keio University, Japan

Abstract. We propose a system that estimates motion in a plush toy by means of an attached sensor device and gives the user a sound feedback corresponding to the predicted motion. We have created several different types of detachable acceleration sensor modules as an accessory for the toy. This module can be attached at any position on a commercially available plush toy. The user can create original motions by teaching through demonstration, and the captured sensor data is converted into 2D image data. We extracted the histograms of oriented gradients (HOG) features and performed learning with a support vector machine (SVM). In an evaluation, we decided the attaching parts and motions in advance, and participants moved a plush toy in accordance with these. Results showed that it was possible to estimate the plush toy's motion with high accuracy, and the system was able to register a sound for each motion.

Keywords: Machine learning, Interactive plush toy, Teaching by demonstration

1 Introduction

Most of us have owned plush toys at some point in our lives. Such toys are familiar to us, and they are often displayed at home as a form of interior decoration. Many people played with such toys in their childhood and have the desire for plush toys to be given a new breath of life. In anime and movies, such wishes sometimes come true. To achieve this in the actual world, we propose a system that animates plush toys by means of a computer.

Interactive plush toys have already been on sale for a while. These toys have built-in electronic devices such as sensors, speakers, and microphones. Our objective in this study, however, is to develop a commercially available plush toy with no built-in electronic devices. Our system can be used with any plush toy and does not require slitting open a toy to attach anything. Essentially, we want to make wishes come true not with an interactive plush toy but rather with an attached interactive plush toy.

In this paper, we propose a system that estimates motion in a plush toy by means of an attached sensor device. We have created several different types of detachable acceleration sensor modules. We introduce three of them here: a band type, a ribbon type, and a skirt type. Their appearances match that of a plush toy, so users can attach them as accessories. The microcontroller is packaged in a bag that the toy carries on its shoulders like a backpack. The sensor data is converted into 2D grayscale image data and is learned with a support vector machine (SVM). We conducted an experiment in which participants performed a predetermined set of motions and found that it was possible to estimate the plush toy's motion with high accuracy.

Fig. 1. A plush toy with our modules.

2 Related Works

2.1 Interactive plush toy

There have been many studies on interaction with plush toys.

Ikeda et al. developed an operating system in which SNS and e-mail are utilized to operate a plush toy. They combined an acceleration sensor and a photo-reflective sensor (PRS) inside a plush toy and activated it using voice recognition technology [1]. Yonezawa et al. combined various sensors inside a plush toy and developed Com-music to create music corresponding to the intensity of the inter-action between the toy and the user [2]. Takase et al. proposed a soft-driving mechanism using threads, cloth, and cotton and used it to develop a plush toy robot [3].

In another work, a system that moves a plush toy in a display as an interface has been developed. Munekata et al. developed this system in order to make a user's attachment to a plush toy stronger [4].

However, in the above studies, devices must be installed inside the plush toy, which means they are only applicable for a specific toy.

In this research, we aim to make a plush toy interactive by using items that can be attached.

2.2 Making existing objects interactive

In our system, we make a plush toy interactive by attaching devices to it. In a similar study, Sugiura et al. developed a plush toy that can be turned into an interactive robot by means of an attachable device that the user can fix onto the toy's limbs, thus enabling it to move freely [5]. Our research is different in that we focus on motion recognition from within the toy.

Kosaka et al. combined sensors with a vacuum cleaner as a game application [6]. In the game, the user collects monsters by sucking up dust. This is an interesting application of entertainment to cleaning that can enhance daily life.

2.3 Motion estimation by real-world sensors

We need to obtain real-world information for estimating a plush toy's motion. The easiest way to do that these days is to collect the information by means of low-cost sensors and estimate the real-world data by machine learning.

Kikui et al. used convolutional neural networks (CNNs) to recognize time series data collected from photosensors and to identify time series gestures [7]. Fukui et al. developed a wristband-type device equipped with photo reflective sensors that can detect hand gestures by measuring changes to the wrist contour that occur while gesturing [8]. The obtained time series sensor data is then converted into image data and identified by extracting the HOG features and performing learning by SVM.

Our system learns the data from acceleration sensors by SVM and then inputs various motions to the toy.

3 System Implementation

3.1 Overview

In this research, we estimate a plush toy's motions by attaching acceleration sensors and obtain sound feedback corresponding to the predicted motion. We attach acceleration sensor modules to each part of a plush toy, such as the paws or the head, and obtain sensor data for every motion. The obtained time series sensor data are converted into image data, and the system then extracts the HOG features from the image data and learns by SVM. After registration of the motions, the user can play using the same motions.

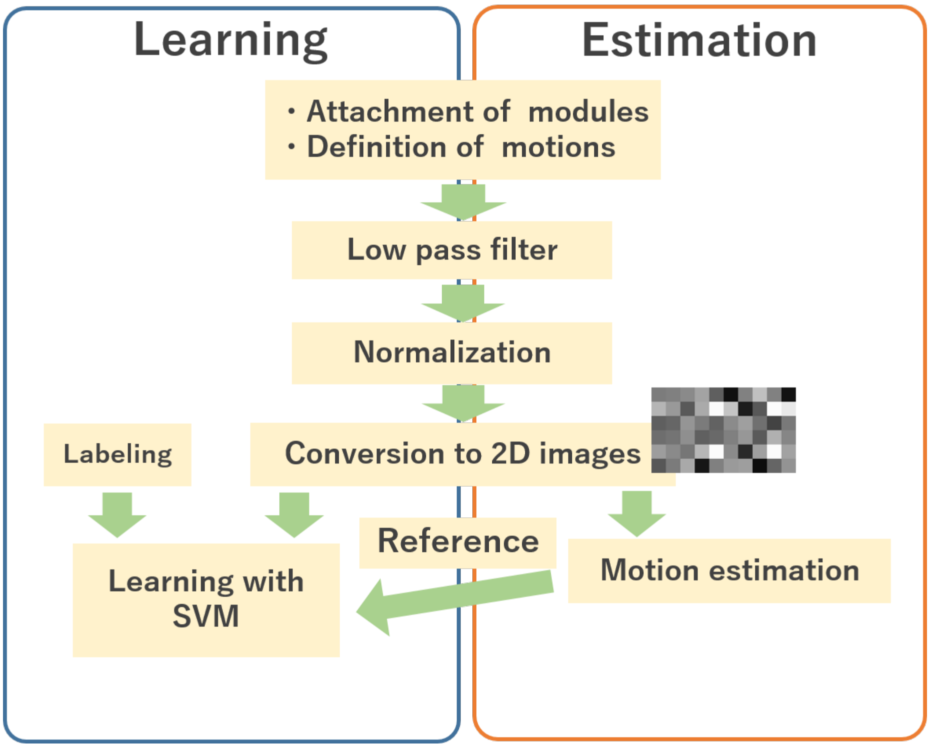

Fig. 2. The flow of motion recognition.

3.2 Acceleration sensor module

We developed several detachable acceleration sensor modules, three of which are introduced here: a band type, a ribbon type, and a skirt type (Fig. 3). Their appearances match that of a plush toy, so users can attach them as accessories. The user attaches a module to any part of a commercially available plush toy and then can define some original motions and teach by demonstration.

The acceleration sensor in the module has three axes, so we can obtain data from three directions (x, y, and z). Each sensor is connected with a microcontroller (Arduino Pro Mini), and sensor data is sent to a PC through the XBee module. The microcontroller is packaged in a bag that the toy carries on its shoulders like a backpack.

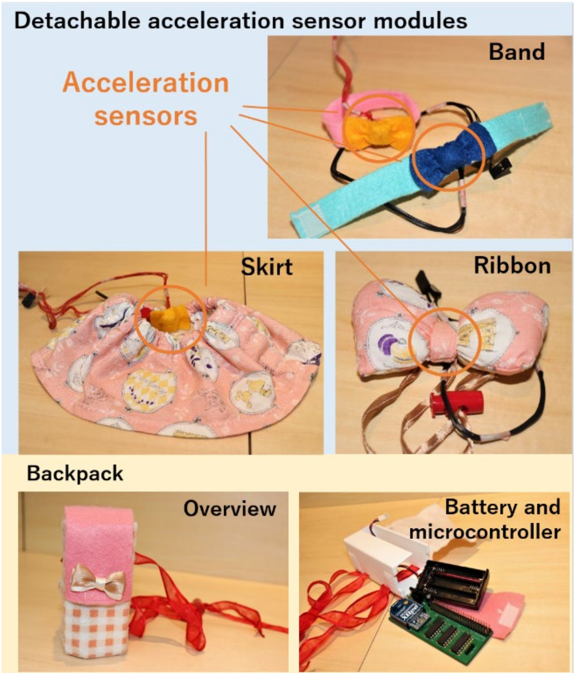

Fig. 3. Acceleration sensor module.

3.3 Obtaining learning data and motion estimation

SVM is used for making a classifier in the Python 3.0 environment.

In this research, we estimate motions through learning by SVM after converting the time series sensor data into 2D grayscale image data and obtaining the HOG features. First, the system judges if a gesture has been made. We calculate differential values between the current sensor data and the one before it and convert them into absolute values. Then, we sum up these values. If the total value passes a certain threshold, data collection is started. Sensor data is often unstable due to gravity or subtle shaking, so we apply a low pass filter (LPF) (specifically, an RC filter) to the obtained sensor data. The formula is as follows:

$$y[i] = a*y[i-1] + (1-a)*x[i], \tag{1}$$

where a is a filter value, x is current data, and y is a time series data frame. At this time, we obtain information on gesture changes by the time series and get the sensor data of the last 60 frames. We then convert them into 2D grayscale image data by means of normalization. In the converted image, each sensor value is allocated in the vertical axis direction and the time series data of 60 frames is allocated in the horizontal axis direction. We convert all sensor data into image data (Fig. 4).

We extract the features from images by using the HOG features. At this time, the size of an image is 3 × n [px] height × 60 [px] width (n: number of sensors). We extend the image data in two directions by four times in order to set the HOG parameters more easily. We define cell size as 4 [px] × 4 [px] and block size as 3 [cell] × 3 [cell], and extract the feature from the image. We use SVM as the classifier for the motion estimation.

Fig. 4. Example of image data.

4 Evaluation

4.1 Overview

In the experiment, we recruited five participants (male: 2, female: 3) aged 22 to 24 years old (average: 23.2 years).

The participants interacted with a teddy bear that had three devices attached to the head, the right paw, and the left paw. Five motions had already been decided: 'nod the head', 'shake the right paw', 'shake the left paw', 'shake both paws', and 'clap both paws'. They performed the five motions 20 times, so each participant made 100 motions. The experimenter explained the motions to them before collecting data, and they practiced them a few times. The extracted value for each image was 32,886 dimensions.

4.2 Result and discussion

The total data collected in the experiment was 5 gestures × 20 times × 5 people = 500 data. We show the evaluation results using leave-one-out cross-validation (Table 1). We performed learning for each participant, and the average accuracy of the identification was about 97.2%.

The results showed that the accuracy of the identification declined when the number of sensors attached to the plush toy increased.

Moreover, the system sometimes detected a motion incorrectly, such as a combination of 'nod the head' and 'shake the right paw'. This occurred by the right paw slightly moved together when a user moved the head hard.

Table 1. Results of the experiment.

		Motions	Predicted 1	2	3	4	5
Truth	1	nod the head	98.0%	2.0%	0.0%	0.0%	0.0%
	2	shake the right paw	0.0%	98.0%	0.0%	2.0%	0.0%
	3	shake the left paw	0.0%	2.0%	98.0%	0.0%	0.0%
	4	shake both paws	0.0%	0.0%	4.0%	96.0%	0.0%
	5	clap both paws	2.0%	0.0%	0.0%	2.0%	96.0%

5 Application

We made an application to apply sounds for a plush toy. After registration of motions, our system can identify the plush toy's motion. In our application, the user can register any sound for each motion of the toy. The user can then interact with the toy by using the registered motions.

6 Limitation and Future Work

In our system, we can move the plush toy freely without having to be connected to a PC. The device is powered by a battery that is packaged inside a bag that the toy carries as an accessory. However, as this bag is rather large, smaller plush toys are not able to carry it. In the case of small plush toys, we should hang the bag from the toy, so we cannot move it completely free.

In this research, we decided on the attaching part and motions in advance. However, individual users will want to play with their plush toys differently. We will explore this through a user study experiment in future work.

7 Conclusion

In this paper, we proposed a motion estimation system for a plush toy. We fabricated an acceleration module and attached it to various parts of a plush toy. This module obtains the plush toy's motion information as sensor data. The obtained time series sensor data is converted into 2D grayscale image data, and the system identifies motions by extracting the HOG features from the image data and performing learning by SVM. The results of experiments with participants showed that the system could identify motions with the average accuracy of 97.2%. In the future, we will perform further testing with a user study experiment.

8 Acknowledgments

This work was supported by JST AIP-PRISM JPMJCR18Y2.

References

1. Ikeda, A., Chiba, Y., Haneda, H.: Stuffed Toy as Non-display Interaction Devices. In Proceedings of the SIG Ubiquitous Computing Systems (UBI) 2013, IPSJ, 2013-UBI-40(16), 1-5 (2013).
2. Yonezawa, T., Clarkson, B., Yasumura, M., Mase, K.: A Music Expressive Communication with Sensor-Doll Interface. In Proceedings of the SIG Human Computer Interaction (HCI) 2001, IPSJ, 17-24, (2001).
3. Takase, Y., Mitake, H., Yamashita, Y., Hasegawa, S.: Motion Generation For the Stuffed-toy Robot. In Proceedings of the annual conference on the Society of Instrument and Control Engineers (SICE '13), 213-217 (2013).
4. Munekata, N., Komatsu, T., Matsubara, H.: Marching Bear: An Interface System Encouraging User's Emotional Attachment and Providing an Immersive Experience. In Proceedings of the International Conference on Entertainment Computing (ICEC '07), Vol.4740, 363-373 (2007).
5. Sugiura, Y., Lee, C., Ogata, M., Withana, A., Makino, Y., Sakamoto, D., Inami, M., and Igarashi, T.: PINOKY: a ring that animates your plush toys. In Proceedings of the ACM annual conference on Human Factors in Computing Systems (CHI '12), ACM, 725-734 (2012).
6. Kosaka T., Matsushita, M.: Monster Cleaner: a serious game to learn cleaning. Laval Virtual ReVolution "Transhumanism++" (ReVo), EPiC Engineering, Vol.1, 50-59 (2018).
7. Kikui, K., Itoh, Y., Yamada, M., Sugiura, Y., Sugimoto, M.: Intra-/Inter-user Adaptation Framework for Wearable Gesture Sensing Device. In Proceedings of the ACM International Symposium on Wearable Computers (ISWC '18), 21-24 (2018).
8. Fukui, R., Okishiba, S., Karasawa, H., and Warisawa, A.: Dynamic Hand Motion Recognition Based on Wrist Contour Measurement for a Wearable Display. In Proceedings of the Robotics and Mechatronics Conference (Robomech '17), 2A2–L02 (2017).