Classification Method of Tactile Feeling using Stacked Autoencoder Based on Haptic Primary Colors

Fumihiro Kato *
The University of Tokyo, Japan

Charith Lasantha Fernando †
Keio University, Japan

Yasuyuki Inoue‡
The University of Tokyo, Japan

Susumu Tachi§
The University of Tokyo, Japan

Abstract

We have developed a classification method of tactile feeling using a stacked autoencoder-based neural network on haptic primary colors. The haptic primary colors principle is a concept of decomposing the human sensation of tactile feeling into force, vibration, and temperature. Images were obtained from variation in the frequency of the time series of the tactile feeling obtained when tracing a surface of an object, features were extracted by employing a stacked autoencoder using a neural network with two hidden layers, and supervised learning was conducted. We confirmed that the tactile feeling for three different surface materials can be classified with an accuracy of 82.0[%].

Index Terms: I.5.1 [Pattern Recognition]: Models—Neural nets; H.5.2 [Information Interfaces and Presentation (e.g. HCI)]: User Interfaces (D.2.2, H.1.2, I.3.6)—Haptic I/O h.5.1 [Human-centered computing]: Virtual reality;

1 Introduction

The importance of virtual reality(VR) technology for telexistence has increased in recent years. If real-time transmission of body sensations, such as tactile feeling, is developed in addition to audiovisual sense, a real-time remote working system will be more realistic. As an example, a previous study [1, 2] proposes that at the work input with a VR simulator, presentation of information related to the tactile feeling of an object, such as force sense, is necessary for an operation task. Telexistence is similar to a VR simulator in that they are virtual for a user who operates them; and it is considered that presentation of tactile information is necessary for efficient work, even in a remote work system. Several telexistence systems have been proposed;and a system that can be experienced in real time has been constructed[3].

For transferring of body sensations, we use the haptic primary colors principle [4] proposed by Tachi. The principle of haptic primary colors is based on a combination of tactile feelings obtained using three types of receptors existing inside the skin to sense force, vibration, and temperature. This is the same as the principle that an image can be reconstructed by decomposing it into three primary colors of red, green, and blue and displaying it.

High fidelity tactile feeling can be presented by transmitting high quality haptic primary colors data obtained using a high sensitivity tactile sensor and presenting the transmitted high quality tactile information. To disseminate a tactile presentation device, it is necessary to be able to present tactile information effectively using devices that are inexpensive and commonly used. In addition, the configuration of the presentation device may vary significantly, e.g., from single modal to multiple modalities, with only vibration, with only vibration and temperature, and devices that can provide vibration, temperature, and force. Therefore, a haptic presentation equalizer that seasons the tactile information to be presented will be necessary depending on the performance of the presentation device. To realize the haptic presentation equalizer, it is necessary to classify the tactile information consisting of a combination of modalities and generate a haptic feeling that can be presented effectively with single or multiple modalities. These problems can be solved if a method can be developed to acquire tactile information as haptic primary colors and classify it. First, we examine methods that can classify vibration. As the device that presents vibration, which is acceleration information, is embedded in devices such as game controllers and smart phones, which are spreading and advancing, the degree of its contribution to dissemination is high. After developing classification method for vibration, shear force can be obtained from integration of vibration components using accelerlaion sensor, and classification using temperature change as a clue is also possible using a temperature sensor.

2 Vibration Classification Method using a Neural Network

For classification of vibration, we employ a machine learning method that is used in the filed of computer vision. Vibration is a physical quantity that has a plurality of frequency components and varies with time. It is considered that fluctuation in the frequency components occurs per unit time. The fourier transform is used to obtain the frequency distribution within unit time. An image is generated using the high intensity part of the frequency distribution as high luminance and the low intensity part as low luminance, and classified using machine learning. We adopt a deep neural network that uses a stacked autoencoder, which is effective for handwritten number classification [5]. As it is difficult to manually extract features in a frequency fluctuation image, automatic extraction of features using the autoencoder is effective.

A neural network that extracts features using a stacked autoencoder with two hidden layers is used. High-order features are extracted using two stages of the autoencoder. To classify the feature vectors of the second-stage autoencoder layer, the soft max layer is learned. The soft max layer learns vibration patterns through supervised learning using labels of tactile feeling. The abovementioned three layers constitute a deep neural network for classification.

Meissner's corpuscles feel vibrations inherent in human skin, generate vibrations having a resonance frequency of less than 100[Hz] [6] under normal pressure. Therefore, it is necessary to be able to acquire vibration of approximately 200[Hz] as a request to the vibration sensor. Meissnner also requires approximately ten minutes to acclimate to a touch stimulus among receptors, then it is used for recognition of stimulus that keeps touching. It is necessary to gather vibration data obtained by touching continuously for the vibration recognition. Even though a method of recognizing a gesture [7] from the moment of contact with a tool has been pro-

*e-mail:fumihiro.kato@tachilab.org
†e-mail:charith@tachilab.org
‡e-mail:y-inoue@tachilab.org
§e-mail:tachi@tachilab.org

2017 IEEE Virtual Reality (VR)
March 18-22, 2017, Los Angeles, CA, USA
978-1-5090-6646-9/17/$31.00

posed, which primarily focused on vibration from the beginning of a touch, recognizing of the device of the target object is the most important, for which acquisition of kHz order vibration is required.

3 Prototype System and Experiments

A device, shown in Figure 1 (a), for automatically collecting the texture of a material surface was developed. We also developed a measurement board (Figure 1 (b)) equipped with an 3-axis acceleration sensor(LIS3DH) which acquires vibration and an infrared thermopile sensor(TMP007) which acquires suraface temperature. Sensor data can be acquired at approximately 100[Hz]. The time required for collecting the texture is approximately 1.5[s], and the material is reciprocated while pressing a hollow square-shaped aluminum rod loaded with the acceleration sensor with a force of 100[g]. The material is attached on a linearly moving rail to collect the tactile texture. The rail is moved iteratively in step of 20[cm] in each direction by the roller of the geared DC motor controlled by Arduino Uno. The acceleration distribution acquired during one iteration is converted into one image, which is used for learning. To produce the acceleration distribution image, the sum of squares in the three axis directions of x, y, z was used. Images with frequencies distributed at primarily less than 45[Hz] were obtained through the Fourier transform. The horizontal axis represents approximately 3[s] showing an iteration of rubbed textures and the vertical axis represents the frequency distribution. The image resolution was 641×34[pixels]. 350 images obtained from each of the three material textures were used for learning. We used MATLAB for learning and classification. Learning was perfoamed using the scale conjugate gradient method, and the mean square error of L2 sparse regularization was used as the loss function. We measured the acceleration at the time of tracing for the following three of materials for 10[minutes]; a board made of Japanese Judas wood, urethane foam and a Ray skin plate (Figure1(c)$\sim$(e)). The evaluation of classification results was 82.0 [%] as shown in Figure 2. For evaluation, we used cross-validation image set which used 50 images for each material. The sizes of the first and second autoencoders are 1700 and 800 nodes respectively, and the upper limit of the calculation iteration count for extraction of features is 400.

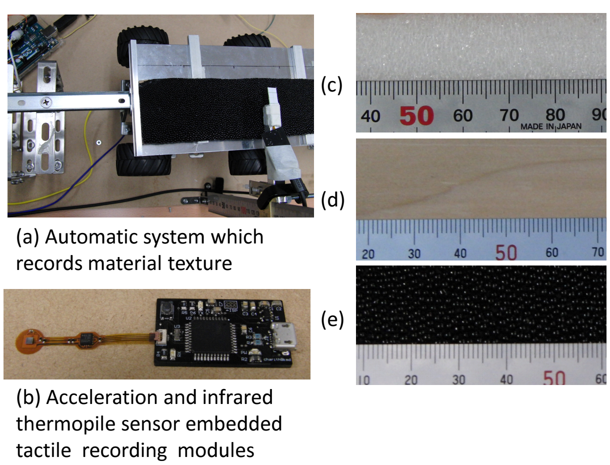

Figure 1: (a)Automatic system which records material texture, (b)Acceleration and infrared thermopile sensor embedded tactile recording module, (c) Urethane foam, (d) Board of Japanese Judas wood, (e) Ray skin plate.

The classification result after learning is shown in Figure 2. According to the above results, it is possible to classify tactile feelings using vibration which acquired from materials surface texture.

From these results, it is thought that a part of the urethane is classified as an Ray skin. It is because mechanical vibration is generated from the recording system. Since the roughness of the surface material are different, it can be thought that it is necessary to suppress the vibration generated by recording system.

	Target Class			
	Urethane foam	Judas wood	Ray skin	
Output Class	39	3	8	78.0%
Output Class	2	46	4	88.5%
Output Class	9	1	38	79.2%
	78.0%	92.0%	76.0%	82.0%

Figure 2: Classification results of three materials. From the left urethane foam, board of Japanese Judas wood, Ray skin plate.

4 Conclusion

Based on the haptic primary colors, we constructed a system using a neural network with a stacked autoencoder to classify the presented tactile feeling. Characteristics were extracted from 350 images obtained from each of the three material textures using two stacked autoencoders. Through supervised learning using soft max, proposed system was able to classify the tactile data of vibrations with an accuracy of 82.0[%]. In the future, we will construct a system that learns learns shear force and temperature change and force fluctuation data. We will study tactile classification method based on haptic primary colors for high quality tactile telexistence.

Acknowledgements

The authors wish to thank all the ACCEL project member. This work was supported in part by a grant from JST ACCEL "Embodied Media Technology based on Haptic Primary Colors".

References

[1] Bernd Fröhlich, Henrik Tramberend, Andrew Beers, Maneesh Agrawala, and David Baraff. Physically-based manipulation on the responsive workbench. In *Proceedings of the IEEE Virtual Reality 2000 Conference*, VR '00, page 5, Washington, DC, USA, 2000. IEEE Computer Society.

[2] Fumihiro Kato, Satoru Onohara, Hironori Mitake, and Shoichi Hasegawa. Real-time heat transfer simulation for reproduction of heat cooking in VR. *Transactions of the Virtual Reality Society of Japan*, Vol 21(1):pages 163–172, Mar 2016.

[3] C. L. Fernando, M. Furukawa, T. Kurogi, S. Kamuro, K. sato, K. Minamizawa, and S. Tachi. Design of telesar v for transferring bodily consciousness in telexistence. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5112–5118, Oct 2012.

[4] S. Tachi, K. Minamizawa, M. Furukawa, and C. L. Fernando. Haptic media construction and utilization of human-harmonized #x201c;tangible #x201d; information environment. In *2013 23rd International Conference on Artificial Reality and Telexistence (ICAT)*, pages 145–150, Dec 2013.

[5] Y. Qi, Y. Wang, X. Zheng, and Z. Wu. Robust feature learning by stacked autoencoder with maximum correntropy criterion. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6716–6720, May 2014.

[6] F. Vega-Bermudez and K. O. Johnson. Sa1 and ra receptive fields, response variability, and population responses mapped with a probe array. *Journal of Neurophysiology*, 81(6):pages 2701–2710, 1999.

[7] Gierad Laput, Robert Xiao, and Chris Harrison. Viband: High-fidelity bio-acoustic sensing using commodity smartwatch accelerometers. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, UIST '16, pages 321–333, New York, NY, USA, 2016. ACM.