ACTA IMEKO
ISSN: 2221-870X
December 2024, Volume 13, Number 4, 1 - 9

Robotic palpation system - reproduction method of dermatologists' skin palpation judgment using a deep neural network

Fumihiro Kato[1], Takeya Adachi[2,3], Kaito Kamishima[4], Takumi Handa[4], Hiroyasu Iwata[5]

[1] *Institute for Human Robot Co-Creation, Future Robotics Organization, Waseda University, 27 Wasedacho, Shinjuku-ku, Tokyo, 162-0042, JAPAN*
[2] *Department of Dermatology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, JAPAN*
[3] *Department of Medical Innovation and Translational Medical Science, Graduate School of Medical Science, Kyoto Prefectural University of Medicine, 465 Kajiicho, Kamigyo Ward, Kyoto 602-0841, JAPAN*
[4] *Graduate School of Creative Science and Engineering, Waseda University, Tokyo, Japan*
[5] *Faculty of Science and Engineering, Waseda University, Tokyo, Japan*

ABSTRACT
The recent surge in infectious diseases, such as COVID-19, has amplified the need for medical examinations that minimize contact between doctors and patients. This is particularly relevant for medical treatments requiring palpation, especially in dermatology. In this study, we aimed to replicate the assessment of softness and surface textures of affected areas, a critical aspect for dermatologists in diagnosing conditions, using a simple robotic device. We derived five levels of softness and three types of surface textures from 14 types of materials based on interviews with dermatologists. To elicit a haptic response from the materials during the pushing procedure, we developed 1) a single-rod probe equipped with a haptic sensor (measuring force and acceleration) using a linear actuator, and 2) a dual-rod configuration supplemented with a nearby vibrator to capture vibration propagation through the material. Frequency-analyzed images were generated from the waveforms of force and acceleration obtained. A total of 500 images from 13 different materials were evaluated for discrimination using AlexNet-based transfer learning. The discrimination accuracy for the 13 materials was 97.4 % (average across trials) when combining the force and acceleration data from the single-rod probe with the acceleration image from the dual-rod probe. The discrimination accuracy was 96.4 % using only the force and acceleration data from the single-rod probe, suggesting that adding acceleration data during vibration propagation enhances discrimination accuracy. A similar comparison was conducted for the five levels of softness and surface features, indicating that acceleration during vibration propagation may provide essential information for palpation.

Section: RESEARCH PAPER

Keywords: palpation; dermatology; medical robotics; telemedicine; haptic primary colours; machine learning; telexistence

Citation: F. Kato, T. Adachi, K. Kamishima, T. Handa, H. Iwata, Robotic palpation system - reproduction method of dermatologists' skin palpation judgment using a deep neural network, Acta IMEKO, vol. 13 (2024) no. 4, pp. 1-9. DOI: 10.21014/actaimeko.v13i4.1571

Section Editor: Zafar Taqvi, USA

Received May 21, 2023; **In final form** October 25, 2024; **Published** December 2024

Funding: This work was supported by JST [Moonshot R&D][Grant Number JPMJMS2031], AMED [Practical Research Project for Allergic Diseases and Immunology][Grant Number 22677379], JSPS Kakenhi [Grant Number JP22K16268][Grant Number JP22K18220], and the Scientific Research Fund of the Ministry of Health, Labour and Welfare, Japan [21FE2001].

Corresponding author: Fumihiro Kato, e-mail: fumihiro.kato@aoni.waseda.jp, phyro21@gmail.com
Corresponding author for medicine: Takeya Adachi, e-mail: jpn4156@me.com

1. INTRODUCTION

In order to realize palpation using a robot, an intelligent robotic system equipped with a sense of touch that can make judgments comparable to those of a doctor is required. The recent spread of infectious diseases such as COVID-19 has increased the need for medical examinations that avoid direct contact between doctors and patients. If a palpation robot can be realized, two scenarios can be considered: 1) a scenario where a doctor examines the patient through the robot, and 2) a

scenario where the doctor manipulates the robot device to touch the patient (Figure 1). In Scenario 1, when a doctor conducts a medical examination via a robot, the robot can acquire palpation information from any part of the patient's body, and it is necessary to appropriately inform the doctor. Several tele-operable humanoid surrogate robots equipped with haptic/somatosensory sensors have been proposed [1], [2], [3], [4]. These humanoid robots, with operable arms and fingers equipped with haptic sensors, can convey the sensation of touch to a remote operator. Additionally, the robot's head is equipped with a camera, allowing the operator to view images from any perspective as the robot's head moves in sync with the operator's viewpoint. However, issues such as transmission delay and haptic reproducibility arise when transmitting haptic information to remote locations. Surgical robots, such as the da Vinci Surgical System (Intuitive Surgical, Inc.), employ a teleoperation method without haptic sensation. The transmission of images is essential even for skilled surgeons. However, the transmission of low-fidelity haptic sensations can create a sense of discomfort if there is a mismatch with the haptic sensations experienced by the doctor. There are still unresolved issues regarding the reproduction of haptic sensations. In Scenario 2, if the robot device can achieve the same level of palpation judgment as a skilled doctor, it is believed that the doctor's examination can be improved without the need for a biopsy by obtaining an objective indicator.

Comparing Scenarios 1) and 2), Scenario 1) utilizes a medical treatment system called "Doctor to Patient with Nurse (D to P with N)," which has been accepted in the medical field. "Doctor to Patient with Nurse with Robot" is proposed, which is updated to support "D to P with N" with a robot. A stable medical robot that could be integrated into the 'D to P with N' system is considered. However, Scenario 2) is considered achievable in the near future, preceding Scenario 1). In this research, we focus on realizing Scenario 2) and aim to achieve Scenario 1) as the next step. The authors conducted interviews with doctors in the medical department [5]. Particularly in dermatology, palpation has traditionally been encouraged due to the ease of touching the affected area. The dermatology department has one of the highest numbers of diseases (2 to 3 thousand), and palpation is an important diagnostic tool for almost all skin diseases. Tactile characteristics such as body temperature, softness, surface texture, and the patient's reaction were comprehensively evaluated during palpation. In this study, we aimed to reproduce the doctor's palpation judgments made in dermatology using a simple robot system and machine learning.

1.1. Cases of telemedicine using robots

In conventional studies, Scenario 1) telemedicine services that transmit audiovisual information have been found to be popular. However, telemedicine that relies solely on audiovisual information does not meet all medical treatment requirements, and the evidence from medical examinations is limited. Proposals have been made for remote palpation systems using a tele-operated robot, an ultrasonic diagnostic imaging method controlled by a teleoperated arm robot [6] , and several methods applied to telemedicine by telexistence using a humanoid surrogate robot [1], [7]. TELESAR VI [1] enables the reproduction of the operator's entire body movement, and haptic transmission based on Haptic Primary Colours (HPCs) [8] has been realized with the fingertips of both hands. As an application scenario, some of the functions required for palpation can be realized by assuming the case of touching the throat of a remote

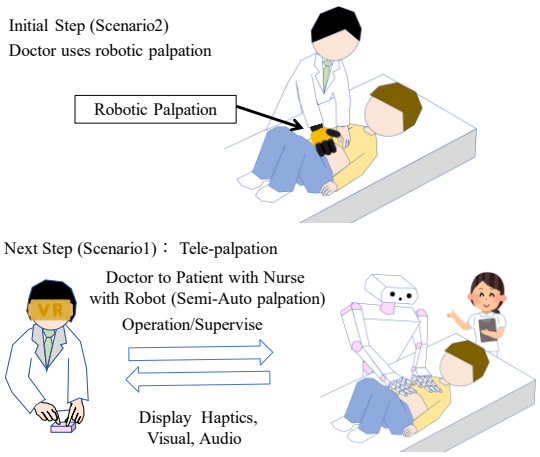

Figure 1. Research vision.

patient. In the proposal of a telediagnosis system using a surrogate robot and Mixed Reality (MR) [7] , the patient wears gloves to acquire vital signs (body temperature and heartbeat). This allows the doctor to inspect while observing the remote patient's appearance displayed in the MR environment. In addition, the patient's vital signs could be palpated while touching the tactile display embedded in the radial artery site of the installed phantom arm. The oxygen saturation concentration is shown in the video transmission. However, the requirements of clinical doctors for palpation robots have not been sufficiently clarified and need to be sequentially examined with regard to applicability to clinical practice. Firstly, because these robots are not specifically configured for clinical practice, they are less likely to be suitable for doctors' daily medical examinations. It is because those will cost frequent maintenance. Doctors aim to examine the patient's physical condition without fear of oversight a sign of disease, a goal that needs to be realized. The doctor observes the patient through inspection. In addition, they use palpation and auscultation to examine the patient's whole-body condition, and investigate signs of perspiration, fever, and swelling from the patient's facial expression. They observe the patient's condition from the moment they enter the room until they leave. A robot system that enables the full utilization of these requirements has not yet been realized. As such, there are technical difficulties in realizing all the examinations currently performed by doctors through a proxy robot. Therefore, instead of artificially reproducing everything, it is necessary to focus on high-priority examinations. In this study, we focus on palpation. As a preliminary step, if a specialized mechanical system is handled by a doctor, as shown in Scenario 2), it is considered to be closer to realization in terms of introduction to the medical field.

1.2. Dermatology examinations, anthroposcopy inspection and palpation clues and related research

In dermatology, skin palpation is essential for determining a patient's condition. When the certainty of a medical examination through inspection and palpation is high, the condition is confirmed by an invasive examination, such as a biopsy. However, biopsies are invasive tests that place a significant burden on the patient. Therefore, a device that can determine the severity of a disease as an objective index without a biopsy would alleviate the burden on both doctors and patients. According to a dermatologist we interviewed, a judgment comparable to the biopsy result can be made with certain degree of accuracy through inspection and palpation. It may be possible to replicate

the doctor's judgment by substituting robotic equipment for inspection and palpation. The judgment of medical conditions using a machine learning method based on inspection has been investigated. A method for estimating the disease from clinical photographs of the skin and dermoscopy images has been proposed [9] , and it has been reported that dermatologist-level tumour discrimination can be achieved using machine learning [10] . In addition, a portable device capable of measuring the skin's viscoelasticity [11] and an automatic palpation method for breast cancer patients using a robot [12] have been proposed. Some methods have been proposed to classify the characteristics of haptic sensor waveforms obtained by rubbing hard materials using machine learning [13], [14]. However, no studies have been proposed to organize the doctor's view on the viscoelasticity of the skin to determine the severity of the disease, and a machine learning method that can reproduce the doctor's palpation softness judgment.

In clinical practice, an advanced version of ultrasonic echo, known as elastography, has been proposed for observing strain due to pressurization [15]. Ultrasound echo devices are commonly used in large hospitals. It has been reported that the detection sensitivity for lymph node enlargement increases from 80.9 % to 95.2 % when using elastography. It is generally observed that lymph node enlargement is benign when it is soft and may be malignant when it is hard, indicating that softness is a crucial factor in determining the condition of the disease. Elastography visualizes softness, but as "teledermatology" advances, it will be necessary to collect information for diagnosis, discrimination, evaluation, and follow-up [16]. To collect this information, it is necessary to ingeniously collate it with the experience acquired by the somatosensory system, which is the physical knowledge acquired by doctors.

The temperature of the affected area rises due to inflammation or infection and falls when the blood flow is poor. However, it is challenging to determine the temperature properties using ultrasonic waves. Importantly, palpation is achievable by matching the temperature sensations experienced by doctors. The softness that a dermatologist perceives when palpating is classified into five levels: soft, elastic soft, elastic hard, cartilage-like hard, and bone-like hard. Furthermore, the surface texture is evaluated in three features: tonic, fluctuation, and brittleness [5]. Although the feeling of softness is sensuously taught by the doctor's master, studies that clarify the differences between doctors are rare, and the mechanical interpretation of palpation and softness has not been clarified. In addition, softness is affected by the tension of the skin in the diseased area. Even if it is of the same severity, it may feel harder when it is located on the forehead than when it is located on the abdomen. The strength of the tension of the skin where the diseased part is located, is considered to affect the judgment of the softness. Therefore, it is necessary to evaluate the tension on the skin around the diseased part. It is conceivable that the necessary information may be included if the viscoelastic properties of the target disease site and the properties due to vibration propagation from the surroundings are examined. This study conducted a verification.

1.3. Method for analysing skin viscoelastic parameters

Various methods have been proposed for measuring the softness of human skin. One such method involves using a cutometer to measure the displacement of the skin over time when negative pressure is applied and subsequently released. This method has been used, for example, to compare the elasticity of the back of the hand among different populations [17]. Another proposed method quantifies the softness of living tissues using vibration [18]. When comparing the softness of materials, the speed of vibration propagation on the material's surface fluctuates, causing a delay in the wave phase. This finding aligns with the results of a theoretical investigation using the equation of motion for a two-dimensional isotropic elastic body [19]. In addition to these methods, studies have been conducted to quantify skin softness. These studies have analysed the rheological characteristics of edema [20] and found that the viscoelastic properties of skin vary between men and women, with women's skin being more elastic than men's.

1.4. Contribution

In this study, we focused on the softness and surface texture of the skin, which are critical factors in dermatological palpation. Dermatologists often use certain materials as benchmarks in their palpation decisions. We investigated the materials that doctors use as benchmarks and examined their perspectives on these materials (Section 2). Furthermore, these opinions can be summarized and used as labels for softness.

We constructed a simple robotic system capable of acquiring softness characteristics through a pushing motion and classifying the obtained haptic information using machine learning (Section 3) (Figure 2). Finally, the results and discussions are presented in Section 4, and the conclusions are presented in Section 5.

2. INVESTIGATION OF SOFTNESS AND SURFACE FEATURES JUDGMENT BY DERMATOLOGISTS, LABELING BENCHMARK MATERIALS

We conducted interviews with dermatologists about the materials corresponding to softness and surface features benchmarked by palpation. Four dermatologists participated in

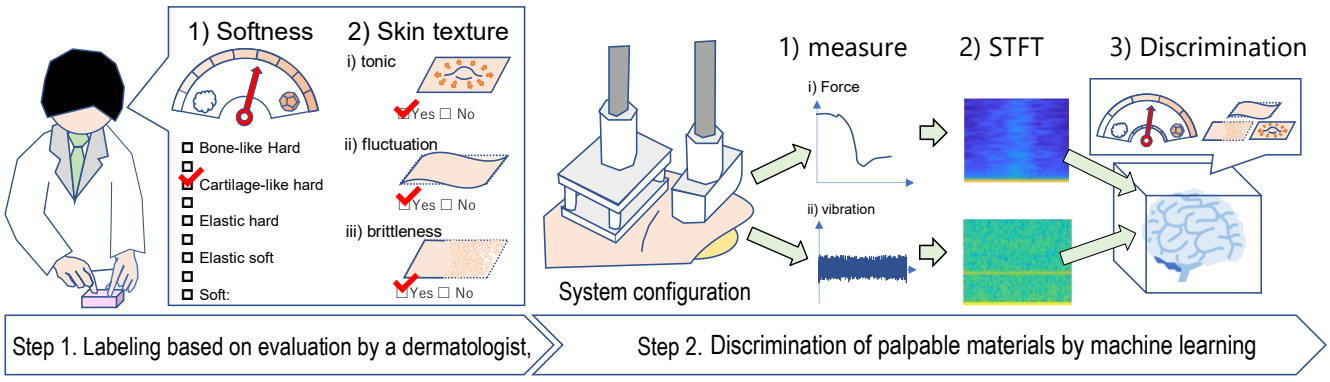

Figure 2. Step 1 : Labelling based on evaluation by a dermatologist, Step 2 : Discrimination of palpable materials by machine learning.

Table 1. Benchmark materials list of palpation for softness and textures obtained through interviews with dermatologists.

		Subjects			
		1	2	3	4
five levels of softness	**Soft**	konjac (arum root)	omelet	marshmallows	slime
	Elastic soft	boiled egg	a bean-jam pancake	gummies	gummies
	Elastic hard	super ball	deep boiled egg	Haribo	rubber eraser
	Cartilage-like hard	chicken cartilage	chicken cartilage	nougat/Snickers	rigid rice cake
	Bone-like hard	walnut	kaminari okoshi (sweet rice cracker)	KitKat	kompeito (sugar plum)
three types of textures	**Tonic**	Marimo tofu (sphere-shape soybean curd covered with thin rubber balloon)	sphere-shape yokan (sweet bean jelly covered with thin rubber balloon)	Bocca's white pudding (covered with thin rubber balloon)	balloon
	Fluctuation	konjac jelly	pudding	bavarian	water balloon
	Brittleness	cookie	firm tofu	snow ball shape cookies (Boule de Neige)	rakugan hard candy

the interviews, providing a total of 32 responses that corresponded to five levels of softness and three types of surface features. The materials obtained are listed in Table 1. Doctors have learned and empirically identified benchmark materials from their mentors; however, it has become clear that there are differences in the materials used as benchmarks. Some materials are standard, but others differ depending on the doctor. Therefore, a statistical investigation was conducted to determine which softness/surface features belong to the materials according to medical doctors. From the materials listed in Table 1, we selected 15 types of materials (konjac (arum root) jelly, omelet, marshmallows, kamaboko (boiled fish paste), sphere-shaped yokan (sweet bean jelly covered with thin tight rubber), Haribo, chicken thighs, Kitkat, rubber eraser, super ball, snickers, salmon jerky, chicken cartilage, walnut, and kompeito (sugar plum) based on their availability, distribution, and well-balanced selection. Eleven doctors, aged 28 to 57 years old (average: 36.2), with dermatology experience ranging from 4 to 34 years (average: 11.4), and comprising seven males and five females, evaluated the materials by pushing or pinching it through pseudo-skin (2 mm thick elastomer gel). In the evaluation column, we received answers about the degree of certainty for nine levels of softness (five levels and its intermediate index) and three levels of surface features. A 7-point Likert scale was used to measure the degree of certainty (1 = strongly disagree, 7 = strongly agree). Different doctors have different criteria for the maximum and minimum evaluations. Considering that each doctor's vote is influenced by doctors with a wide range of evaluations, a normalization process should be performed.

Table 2 lists the normalized values, which are calculated in two steps. The normalization procedure is as follows:

Step 1: Each doctor's degree of certainty vote is divided by the doctor's average vote to calculate the value.

Step 2: The result from Step 1 is divided by the maximum value of the doctor's vote. The results of Step 2 for each doctor are then added together.

To select the softness corresponding to the material from the normalization, a method that considers the score that is the largest is first considered (Method 1). Method 1 is interpreted as follows:

Method 1: The highest evaluation of the numerical value is determined from the normalization results as the softness label. Because the basic label has five levels of softness, as judged by the doctor, the next largest level is used as the label when the middle between the labels is the maximum. The labels classified using this method were as follows.

Soft: konjac (arum root) jelly, omelet

Elastic soft: marshmallow, kamaboko (boiled fish paste), sphere-shaped yokan (sweet bean jelly covered with thin rubber balloon), Haribo, chicken thigh

Elastic hard: Kitkat, rubber eraser, super ball, snickers

Cartilage-like hard: salmon jerky, chicken cartilage, walnut

Bone-like Hard: kompeito (sugar plum)

However, the interpretation of Method 1 does not take advantage of variance when there are labels with comparable scores. Therefore, we also propose an interpretation method 2 that takes variance into account. For Method 2, softness categories are determined as follows.

Method 2: Considering variance

Step 1: Calculate the product of the elements of the row and the softness score, where the score is 1 for soft and 9 for bone-like hardness, and sum them. (e.g., for kitkat, 4 * 0.86 + 5 * 2.29 + ... 9 * 1.71).

Step 2: Divide by the sum of the rows.

Table 2. Results of votes by doctors for the softness and surface texture of benchmark materials (after normalization).

	Soft		Elastic soft		Elastic hard		Cartilage-like hard		Bone-like hard	method 2	Tonic	Fluctuation	Brittleness
softness score for method 2	1	2	3	4	5	6	7	8	9				
kompeito	0.00	0.00	0.00	0.00	1.57	0.00	0.71	0.86	7.00	6.7	1.0	0.0	0.0
Kitkat	0.00	0.00	0.00	0.86	2.29	0.71	0.67	2.14	1.71	8.2	0.7	0.0	0.0
walnut	0.00	0.00	0.00	0.00	1.43	0.00	2.52	2.29	1.71	7.4	0.0	0.0	2.7
snickers	0.00	0.00	0.00	1.00	1.86	0.71	0.67	2.00	1.71	6.7	0.7	0.0	0.0
chicken cartilage	0.00	0.00	1.00	0.00	2.00	1.29	4.43	0.00	0.00	5.9	0.0	0.0	0.6
salmon jerky	0.00	0.00	1.00	0.00	1.86	1.43	2.43	1.86	0.00	6.2	0.0	0.0	1.0
rubber eraser	0.00	0.00	0.00	0.86	4.00	0.71	1.29	0.00	1.71	6.1	0.7	0.0	0.0
super ball	0.00	0.00	0.71	2.00	3.86	0.00	0.67	1.43	0.00	5.3	2.4	0.0	0.0
sphere-shaped yokan	0.00	1.43	4.43	0.00	3.29	0.00	0.00	0.00	0.00	3.6	2.4	0.0	0.0
Haribo	0.00	0.57	2.29	3.29	2.24	0.00	0.00	0.00	0.00	3.9	0.6	0.6	0.0
chicken thighs	0.00	0.43	2.00	1.71	1.86	0.00	1.43	0.71	0.00	4.8	0.7	0.0	0.6
kamaboko	0.86	1.00	5.71	0.71	0.57	0.00	0.00	0.00	0.00	2.9	1.3	0.0	0.0
omelet	4.43	1.00	3.29	0.00	0.00	0.00	0.00	0.00	0.00	1.9	0.0	0.0	4.1
marshmallows	2.71	2.57	3.00	0.00	0.00	0.00	0.00	0.00	0.00	2.0	0.0	0.6	0.0
konjac jelly	3.57	1.71	3.00	0.71	0.00	0.00	0.00	0.00	0.00	2.1	0.0	4.5	1.0

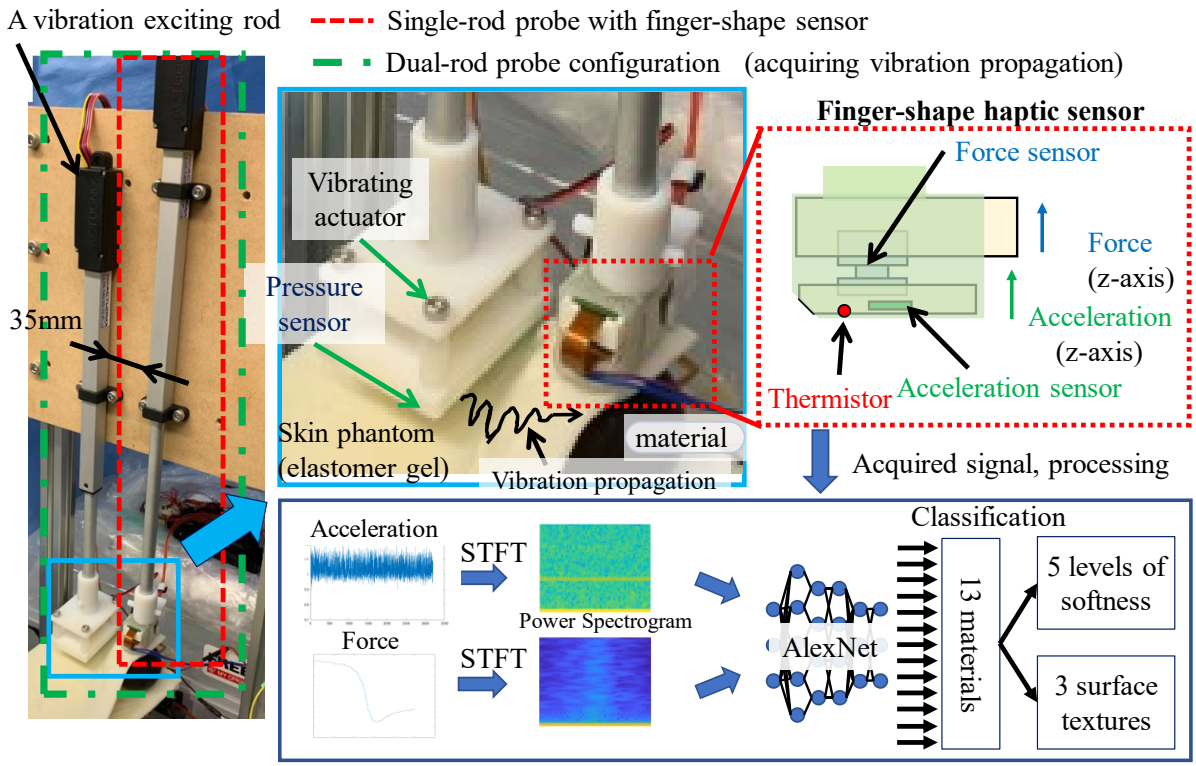

Figure 3. Single-rod sensing probe, Dual-rod sensing probe, Finger-shape haptic sensor, and sensor data processing using machine learning.

The calculated score was listed in Table 2 (method 2). The five levels of softness that are closest to the score are the teaching labels.

The labels classified using this method 2 were as follows.

Soft: omelet

Elastic soft: konjac jelly, marshmallow, kamaboko, sphere-shaped yokan, Haribo

Elastic hard: chicken thigh, super ball, snickers, chicken cartilage

Cartilage-like hard: Kitkat, rubber eraser, salmon jerky, walnut, snickers

Bone-like Hard: kompeito

The interpretation of Method 1 is used for surface features, as the effect of dispersion is considered to be small. The category with a maximum evaluation value of 2 or more was used as the label.

Tonic: super ball, sphere-shaped yokan (sweet bean jelly covered with thin rubber balloon)

Fluctuation: konjac (arum root) jelly

Brittleness: walnut, omelet

The following section explains the reproduction system of the pushing operation during doctor palpation and the discrimination method using machine learning.

3. SOFTNESS DISCRIMINATION METHOD USING MACHINE LEARNING

A robot system capable of acquiring the viscoelastic response of benchmark materials was developed and evaluated using machine learning. The robot system comprises 1) a single-rod probe equipped with a finger-shaped haptic sensor, and 2) a dual-rod probe configuration that captures vibration propagation from an adjacent vibration-exciting rod. To accommodate the fact that softness varies depending on the hardness of the skin where the affected area is located, a dual probe is designed to obtain characteristics such as the softness of the skin tension around the affected area. Figure 3 illustrates the configurations of the single-rod probe and dual-rod probe.

3.1. Single-rod haptic sensor probe system for obtaining viscoelastic response

The single-rod probe is equipped with a finger-type haptic sensor probe [1] based on haptic primary colour theory [6]. This configuration can acquire force, acceleration, and temperature in response to the force, vibration, and temperature receptors on human finger skin. A linear actuator (L16-140-35-12-P, Actuonix) was used to push the sensor module, and an elevating adjuster under the material was used for adjustment. The haptic sensor, which has the same configuration as TELESAR VI includes a 3DoF force sensor (micro DynPick MAF-3, sampling rate 1 kHz, Wacoh-Tech), a tri-axis digital accelerometer (KX126, 1.6 kHz, Kionix), and a thermistor (56A1002-C3, Alpha Technics) as the temperature sensor. The haptic sensor waveform was acquired at 1.6 kHz (accel), 1 kHz (force), and 100 Hz (temperature). The signal waveform obtained by the sensor was transmitted to a PC via WiFi and recorded.

3.2. Dual-rod probe configuration for obtaining vibration propagation property

In the dual-probe configuration, the same linear actuator (L16-140-35-12-P, Actuonix) with a vibrating actuator was placed beside the single-rod probe (distance 35 mm). The vibrating actuator was pressed against the material, and vibration was applied to the surface of the material. The vibration propagates through the pseudo-skin (made of viscoelastic elastomer, thickness 2 mm, tension is applied on pseudo-skin) and through the internal material; then, the accelerometer placed at the tip of the single rod acquires the vibration. A pressure sensor (FSR401) was installed between the vibrating actuator and the material to maintain the pressing pressure of the single rod

constant. The vibrating actuator was driven at 100 Hz using a haptic reactor AFT14A903A (resonance frequency:160 Hz, 320 Hz, ALPSALPINE). In estimating softness from the phase difference by vibration propagation [16] , the vibration frequency varied from 20 to 130 Hz. In an experiment targeting soft materials such as konjac and liver, it has been reported that attenuation is severe at 130 Hz and above. In the proposed method, we used 100 Hz. The benchmark material was covered with a sheet of elastomer gel (thickness of 2 mm) that imitated the skin, and vibration propagation was recorded when tension was applied.

3.3. Haptic waveform acquisition and frequency spectrum image generation

Haptic waveform acquisition experiments were performed using two configurations: 1) a single-rod probe configuration, which measures the reaction response of force and acceleration during pushing, and 2) a dual-rod combination configuration, which measures vibration propagation using an acceleration sensor. The obtained waveforms were then visualized. In the experiment with the single-rod probe configuration, 13 of the 14 benchmark materials were pushed (2 mm, approximately 300 ms), stopped (1000 ms), and pulled up (approximately 300 ms). The waveform in the force pushing direction (z-axis) was extracted for the first 500 ms from the start of pushing. A spectrogram image was generated using a short-term Fourier transform (STFT). Because the force/accelerometer sensor was acquired at a sampling rate of 1/1.6 kHz, the maximum value of the vertical axis of the image was 500/800 Hz, and the maximum value of the horizontal axis was 500/312.5 ms in the spectrogram image. In the STFT calculation, a Hamming window (width = 110, overlap = 109) was adopted for the window function. The training dataset was constructed with 515 of these images. In the dual-probe configuration, both probes were pressed against the material, and force/acceleration waveforms were acquired over a period of 3200/2000 ms as one waveform. This generated a spectrogram image. The training dataset was constructed with between 500 and 513 of these images.

The images of the force and acceleration waveforms were obtained under two conditions: 1) using a single-rod probe pushing and 2) employing a dual-rod probe configuration (vibrating a material only). A single image size is 682 × 539 pixels. Composite images were produced by arranging the two images side by side or by arranging three images together. If the force images are arranged on the left side of the acceleration image, a combined image size of 2273 × 934 pixels was obtained. The combined size of three images was 2858 × 814 pixels.

The Table 3 shows the force (#1), acceleration (#2), and force/acceleration combination image (#12) for the single-rod probe (pushing), the force (#3), acceleration (#4), and force/acceleration combination image (#34) for the dual-rod probe (vibrating a material), and the force (#1) + acceleration combination image (#34), force (#1) + acceleration (#4) combination (#14), and force (#1) + acceleration (#2) + acceleration (#4) combination (#124), respectively, were evaluated for learning and discrimination accuracy in the manner described in the next section.

3.4. Discrimination using machine learning

The images were classified using a machine learning method that employs a Deep Neural Network (DNN). This method involves transfer learning using the AlexNet model, which was proposed as the optimal model for the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2012. The features were extracted from the input image via a feature extraction unit, which consists of five convolution layers and three pooling layers. Discrimination was performed using three fully connected layers. The model was pre-trained using one million images across 1000 categories. In this experiment, the AlexNet feature layer, fc7, was utilized. For the multiclass classifier, which serves as the discrimination layer, a multiclass Error-Correcting Output Codes (ECOC) model using a Support Vector Machine (SVM) binary learner was employed. The batch size was set to 32. Of the between 500 and 512 of generated images, 80 % were randomly selected for transfer learning and 20 % for evaluation. AlexNet is a conventional method used as a benchmark for DNN. In recent years, higher performance NNs have appeared and their performance is considered to have improved, but we use basic NNs for the purpose of easy comparison with similar literature in order to compare the characteristics of materials, rather than to find performance improvement by using conventional high-performance NNs. Figure 4 shows an example of the generated image of the force (a) and (b) acceleration of the yokan. A computer with an i9-10900K 3.7 GHz, 32 GB RAM was used. MATLAB 2021b was used for machine learning. In the dual-rod configuration, the images were classified by the metric of joint learning using images in which the images of the acceleration waveforms were arranged, in addition to the force waveforms obtained by the single-rod probe.

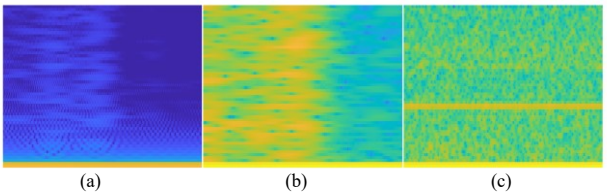

Figure 4. An example of frequency analyzed images for yokan (#124 case): (a) Force (single-rod pushing), (b) Acceleration (single-rod pushing), (c) Acceleration (dual-rod, Vibration propagation).

Evaluations were made on five trials of each of these combinations. The average, maximum, and minimum performance during the trials is summarized in the table.

4. RESULT AND DISCUSSION

Experimental results and discussion of the discrimination of 13 different materials, discrimination of 5 levels of softness, and discrimination of 3 types of surface features were presented below, respectively.

4.1. 13 materials discrimination result and discussion

Table 3 shows the discrimination results for force, acceleration, and their combined images obtained from single-rod and dual-rod images. The discrimination accuracy was calculated five times for each condition, and the average accuracy, minimum, and maximum values are summarized. The results show that the #124 image combination had the highest accuracy at 97.4 % (mean), followed by #12 (96.4 %), #14 (93.6 %), and #1 (92.4 %). In a previous study by Kato et al. [21], the accuracy calculation with 343 images was 93.0 % for #1 and 95.2 % for #14, and comparable results were reproduced in this verification.

In this study, as a benchmark for examining softness, a material whose softness is perceived by humans to make a difference was used. It was reasonable to assume that the data in #1 and #2 showed high discrimination accuracy based on the

Table 3. Discrimination accuracy for 13 palpation benchmark materials based on haptic data acquisition method and dataset configurations.

Acquisition case	Dataset	Images	Learning and evaluation trial					Discrimination accuracy		
			1	2	3	4	5	AVG	MIN	MAX
Single-rod probe (pushing)	Force : **#1**	515	91.9%	92.1%	92.2%	93.7%	91.9%	92.4%	91.9%	93.7%
	Accelearation : **#2**	515	79.9%	79.0%	80.7%	79.0%	79.5%	79.6%	79.0%	80.7%
	Force(**#1**)+Accel.(**#2**) : **#12**	515	97.6%	94.7%	96.1%	96.3%	97.2%	96.4%	94.7%	97.6%
Dual-rod probe (vibrating propagation only)	Force : **#3**	513	37.9%	36.8%	37.5%	38.4%	36.6%	37.4%	36.6%	38.4%
	Accelearation : **#4**	500	54.1%	52.4%	51.1%	52.0%	51.7%	52.3%	51.1%	54.1%
	Force(#3)+Accel.(#4) : **#34**	500	64.9%	67.0%	65.7%	64.1%	63.8%	65.1%	63.8%	67.0%
Single-rod (Force) + dual-rod (Accel.)	Force (**#1**)+Accel.(**#4**) : **#14**	513	92.8%	93.8%	94.1%	94.5%	92.9%	93.6%	92.8%	94.5%
Single-rod (Force + Accel.) + dual-rod (Accel.)	Force(**#1**)+Accel.(**#2**) +Accel.(**#4**) : **#124**	513	97.5%	96.7%	97.7%	97.1%	97.9%	97.4%	96.7%	97.9%

Table 4. 5 levels of softness discrimination accuracy based on differences in interpretation of physician evaluation and image combinations #12 and #124.

Methods for interpreting dermatologists' votes	Dataset	Discrimination accuracy in softness classification (trial)					Discrimination accuracy		
		1	2	3	4	5	AVG	MIN	MAX
Method 1：The highest evaluation	**#12**	98.2%	95.5%	96.6%	96.8%	97.5%	96.9%	95.5%	98.2%
	#124	97.7%	97.2%	97.8%	97.5%	98.2%	97.7%	97.2%	98.2%
Method 2 : Considering variance	**#12**	98.3%	96.7%	97.0%	97.6%	98.1%	97.6%	96.7%	98.3%
	#124	98.7%	98.1%	98.8%	98.4%	99.0%	98.6%	98.1%	99.0%

Table 5. Discrimination accuracy of 3 types of surface features (comparison of #12 and #124).

		Discrimination accuracy in surface texture classification (trial)					Discrimination accuracy		
		1	2	3	4	5	AVG	MIN	MAX
Tonic	**#12**	99.2%	98.1%	98.8%	98.7%	99.3%	98.8%	98.1%	99.3%
	#124	99.5%	98.8%	99.1%	98.9%	99.6%	99.2%	98.8%	99.6%
Fluctuation	**#12**	99.9%	99.3%	99.6%	99.3%	99.8%	99.6%	99.3%	99.9%
	#124	99.5%	100.0%	99.9%	99.9%	99.9%	99.9%	99.5%	100.0%
Brittleness	**#12**	99.3%	97.9%	98.7%	99.0%	99.0%	98.8%	97.9%	99.3%
	#124	99.3%	98.7%	99.3%	98.8%	99.1%	99.0%	98.7%	99.3%

response of force and acceleration during pushing. On the other hand, in the method of measuring vibration propagation such as #3 and #4, all materials were acquired under the condition of touching through the pseudo-skin, and the vibration component propagating through the pseudo-skin was considered to contain a certain percentage of the vibration component. The higher accuracy of #14 than #12 and the higher accuracy of #124 than #12 are considered to indicate that the properties to distinguish the material can also be obtained from #4. Previous studies have noted variations in the properties of vibration propagation depending on viscoelasticity [18], and it is possible that these material properties of our study were obtained due to variations in vibration propagation.

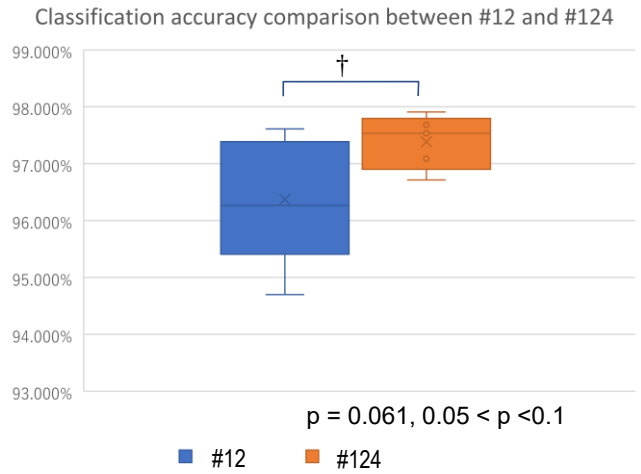

Figure 5. Significant trend between #12 and #124 for 13 materials discrimination results ($p = 0.061$, $0.05 < p < 0.1$).

The Mann-Whitney U-test (one-taled test) for the variation in discrimination accuracy of the 13 materials shown in Figure 5 suggests that #124 has a more significant trend than #12 ($p = 0.061$). These results suggest that the information of vibration propagation in #4 includes physical properties in the identification of 13 types of materials.

These results suggest that it is possible to classify 13 different materials by combining information obtained from both single-rod and dual-rod probes.

4.2. 5 levels of softness discrimination result and discussion

Table 4 shows a comparison of the average discrimination accuracy of #12 and #124, which had the highest average discrimination accuracy when the 13 materials were regrouped into 5 levels of softness. This study proposed two methods for interpreting dermatologists' votes for the five levels of softness. Softness discrimination scores were obtained for each of Method 1 (interpretation based on maximum score) and Method 2 (interpretation that takes advantage of the variance of the ratings). For both interpretation methods, the discrimination accuracy of #124 was higher than that of #12 (compared with the average of 5 trials). The accuracy of #124 in method 1/method 2 was 97.7/98.6 %, respectively.

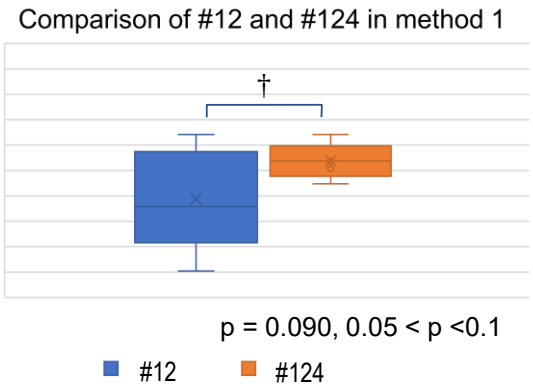

Figure 6. A significant trend ($p = 0.090$, $0.05 < p < 0.1$) for interpretation method 1 in comparison of #12 and #124.

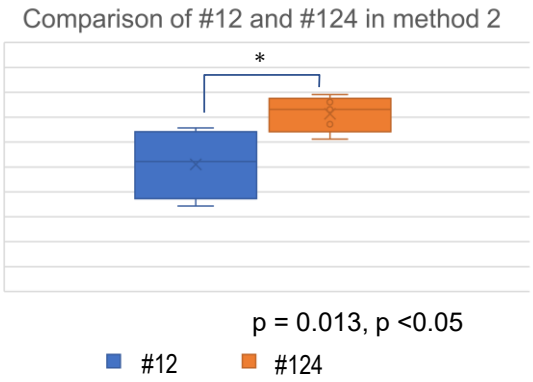

Figure 7. A significant difference ($p = 0.013$, $p < 0.05$) for interpretation method 2 in comparison of #12 and #124.

The discrimination accuracies of #12 and #124 in interpretation method 1 and interpretation method 2 were also compared by Mann-Whitney's U test (one-taled test) in the same manner. As shown in Figure 6 and Figure 7, with a significant trend ($p = 0.090$, $0.05 < p < 0.1$) for interpretation method 1 and a significant difference ($p = 0.013$, $p < 0.05$) for interpretation method 2 was observed. These results suggested that a difference in discrimination accuracy between #12 and #124 into 5 levels of softness, and that the information in #4, which is the accelerometer component of vibration propagation, may contain meaningful information for softness judgment.

Finally, whether the vibration information (#4) was valid for which softness levels, was examined. For method 1, significant trends were found for bone-like hard and elastic soft, as well as overall. For method 2, significant differences were found for Cartilage-like hard and Elastic hard. Significant differences were also observed overall. The above results suggested that the frequency is suitable for acquiring information on the softness of the elastic soft, cartilage-like hard, and bone-like hard, since significant differences were obtained in the accelerometer at the propagation destination when vibrations of 100 Hz were applied to them.

4.3. 3 types of surface texture discrimination result and discussion

The discrimination accuracies of #12 and #124 were calculated for three surface features (tonic, fluctuation, and brittleness) (Table 5). Discrimination accuracy was higher with #124 for all three surface features (although the difference with #12 was slight), at 99.2 %, 99.9 %, and 99.0 %, respectively. It was also suggested that adding #4 to the information obtained by #12 for surface features can improve the accuracy.

A Mann-Whitney's U-test (one-tailed) showed a significant trend for fluctuation. No significant differences were found for tonic and brittleness.

4.4. Limitation and Future Work

Although not the subject of this study, it is possible to compare #14 and #124 to examine the impact of #2 on discrimination.

The future work is to isolate the elements of what information the vibration propagation represents about the material. For example, it is necessary to consider whether differences in material shape, size, direction of excitation, vibration frequency, and amplitude are the main sources of information about vibration propagation.

In the five levels of softness discrimination, Method 2 was not accepted in limited case, with #12 being higher than #124 (two-tailed test, significant difference, $p<0.05$) for soft materials. No significant difference was found for Method 1. The frequency added for vibration propagation in this study was 100 Hz, but it is possible that softer materials require lower frequencies of vibration. A previous study [18] showed that soft materials have a large phase delay. Future work is expected to improve by varying the frequency added to accommodate the dispersion of softness of the material.

The sensor configuration was based on the haptic primary colour principle, but it was not used temperature information. We will investigate the effect on softness perception based on the difference between the temperature of the doctor's fingertip and the temperature of the affected area.

The viscoelasticity values have not been identified in this study. In terms of discrimination accuracy, we have not analysed which materials are close to each other and at what distance. These are also issues to be addressed in the future.

In future work, we plan to increase the number of materials tested, apply our findings to various diseases, and focus on developing applications for robots that could potentially serve as substitutes for doctors. We will continue to explore the potential uses of our system.

5. CONCLUSION

This study aimed to develop a robotic device capable of replicating the evaluation of skin softness and surface texture judgment performed by dermatologists through palpation. We conducted interviews with dermatologists to identify materials that correspond to various degrees of softness and surface features typically assessed during palpation. From a survey of four doctors, we selected benchmark materials that corresponded to five levels of softness and three types of surface textures. These materials were then evaluated by an additional 11 doctors. In total, 14 benchmark materials were categorized according to their softness and surface condition.

Furthermore, we constructed a robotic system that mimics the pressing motion used during palpation and conducted a discrimination experiment using 13 of the benchmark materials. Our results indicate that it is possible to classify five levels of softness (ranging from soft to bone-like hardness) and three types of surface textures that dermatologists can distinguish using palpation.

ACKNOWLEDGEMENT

This work was supported by JST [Moonshot R&D][Grant Number JPMJMS2031], AMED [Practical Research Project for Allergic Diseases and Immunology][Grant Number 22677379], JSPS Kakenhi [Grant Number JP22K16268][Grant Number JP22K18220], and the Scientific Research Fund of the Ministry of Health, Labour and Welfare, Japan [21FE2001]. Our special gratitude goes to Dr. Susumu Tachi from the University of Tokyo for his contribution to haptic sensing, as well as Drs. Toyoko Inazumi, Hiroki Arakawa, Yuka Shintani, and Yuki Kobayashi from the KKR Tachikawa Hospital, Drs. Shuhei Nishimoto, Sayuka Arakawa, and Akihiro Miyagawa from the Kawasaki Municipal Hospital, Drs. Yuichi Kurihara, Ryo Tanaka, Tomohiro Suzuki, and Takayoshi Sakuta from the Hiratsuka City Hospital, Drs. Emiko Watanabe-Okada, Ryohei Asakura, and Chiaki Takahashi from the Saiseikai Yokohamashi Tobu Hospital for their valuable contribution to the palpation analysis. This work was also supported by Waseda University Global Robot Academia Institute and Waseda University Green Computing Systems Research Organization.

REFERENCES

[1] S. Tachi, Y. Inoue, F. Kato, Telesar VI: Telexistence surrogate anthropomorphic robot VI, Int. Journal of Humanoid Robotics, vol. 17, no. 5, 2020, pp. 1–34.
DOI: 10.1142/S021984362050019X

[2] M. A. Diftler, J. S. Mehling, M. E. Abdallah (+ another 11 authors), Robonaut 2 - The first humanoid robot in space, Proc. of the IEEE Int. Conf. on Robotics and Automation, Shanghai, China, 9-13 May 2011, vol. 1, pp. 2178–2183.
DOI: 10.1109/ICRA.2011.5979830

[3] S. Dafarra, U. Pattacini, G. Romualdi (+ another 18 authors), iCub3 avatar system: Enabling remote fully immersive embodiment of humanoid robots, Science Robotics, vol. 9, issue 86, 2024.
DOI: 10.1126/scirobotics.adh3834

[4] C. Lenz, M. Schwarz, A. Rochow, B. Pätzold, R. Memmesheimer, M. Schreiber, S. Behnke, NimbRo wins ANA Avatar XPRIZE Immersive Telepresence Competition: Human-Centric Evaluation and Lessons Learned, Int. Journal of Social Robotics, 2023.
DOI: 10.1007/s12369-023-01050-9

[5] K. Kamishima, F. Kato, T. Handa, H. Iwata, Research on Telepalpation System (first report) -Investigation of Telemedicine and Interviews Report with Medical Doctors-, pp. 3D2-1 Prod. of VRSJ Annual Conference, 2021.

[6] H. Shibata, N. Okamoto, T. Uchida, Construction Imaging of Safety Palpation Structure Diagnosis and Its Functional for Tele-ultrasound Verification, Transactions of the JSME (in Japanese), vol. 74, no. 739, 2008, pp. 179–186.

[7] J. Fu, F. Kato, Y. Inoue, S. Tachi, Development of a Telediagnosis System using Telexistence, Transactions of the Virtual Reality Society of Japan, vol. 25, no. 3, 2020, pp. 277–283.
DOI: 10.18974/tvrsj.25.3_277

[8] S. Tachi, K. Minamizawa, M. Furukawa, C. L. Fernando, Haptic Media: Construction and Utilization of Human-harmonized "Tangible" Information Environment, 23rd Int. Conf. on Artificial Reality and Telexistence (ICAT), Tokyo, Japan, 11-13 December 2013, pp. 145–150.
DOI: 10.1109/ICAT.2013.6728921

[9] M. Tanaka, A. Saito, K. Shido (+ another 7 authors), Classification of large-scale image database of various skin diseases using deep learning, Int. Journal of Computer Assisted Radiology and Surgery, vol. 16, no. 11, 2021, pp. 1875–1887.
DOI: 10.1007/s11548-021-02440-y

[10] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, S. Thrun, Dermatologist-level classification of skin cancer with deep neural networks, Nature, vol. 542, no. 7639, 2017, pp. 115–118.
DOI: 10.1038/nature21056

[11] S. Park, J. Tao, L. Sun, C. M. Fan, Y. Chen, An economic, modular, and portable skin viscoelasticity measurement device for in situ longitudinal studies, Molecules, vol. 24, 2019, no. 5.
DOI: 10.3390/molecules24050907

[12] I. Kato, K. Koganezawa, A. Takanishi, Automatic Palpation System for Breast Cancer : WAPRO-4, Journal of the Robotics Society of Japan, vol. 5, no. 2, 1987, pp. 18–24.
DOI: 10.1163/156855389X00217

[13] F. Kato, C. L. Fernando, Y. Inoue, S. Tachi, Classification method of tactile feeling using stacked autoencoder based on haptic primary colors, Proc. of the 2017 IEEE Virtual Reality (VR), Los Angeles, CA, USA, 18-22 March 2017, pp. 391–392.
DOI: 10.1109/VR.2017.7892341

[14] F. Kato, Soft Finger-tip Sensing Probe Based on Haptic Primary Colors, 2018.
DOI: 10.2312/egve.20181322

[15] F. A. Roldán, Application of Elastography in Dermatology, Actas Dermo-Sifiliográficas (English Edition), vol. 107, issue 8, October 2016, pp. 652-660
DOI: 10.1016/j.adengl.2016.07.012

[16] P. Pala, B. S. Bergler-Czop, J. M. Gwiżdż, Teledermatology: idea, benefits and risks of modern age – a systematic review based on melanoma, Postepy Dermatologii i Alergologii, vol. 37, no. 2, 2020, pp. 159–167.
DOI: 10.5114/ada.2020.94834

[17] T. Iida, M. Ooba, F. Matsuno, N. Koga, Skin conditions on the Dorsal Hands of Japanese and Caucasian Women, J. Soc. Cosmet. Chem. Jpn, vol. 48, no. 4, 2014, pp. 278–286.

[18] O. Takatani, T. Akatsuka, Clinical Measurement of Flexibility of Living Tissue, 1975, pp. 35–46.

[19] H. Lamb, On the propagation of tremors over the surface of an elastic solid, Philosophical Transactions of the Royal Society of London, vol. 203, no. 359–371, 1904, pp. 1–42.

[20] Y. Yazawa, M. Nagasaka, Viscoelastic Properties of Edematous Skin, Jap. Heart J., vol. 3, 1962, pp. 220–230.

[21] F. Kato, T. Adachi, T. Handa, K. Kamishima, H. Iwata, Palpation Robot System - Reproduction Method by Deep Neural Network of Skin Palpation Judgment Focusing on Softness Classification, Proc. of the Int. Symp. on Measurement and Control in Robotics: Robotics and Virtual Tools for a New Era, ISMCR 2022, Online, 28-30 September 2022, pp. 13–20.
DOI: 10.1109/ISMCR56534.2022.9950593