

Sufficient Time-Frequency Resolution for Reproducing Vibrotactile Sensation

Yutaro Toide , Masahiro Fujiwara , *Member, IEEE*, Yasutoshi Makino , and Hiroyuki Shinoda , *Member, IEEE*

Abstract—In vibrotactile stimuli, it is essential to reproduce realistic tactile sensations to enhance the immersiveness of applications. To reproduce more realistic tactile experiences, various tools have been proposed to fine-tune and design vibrotactile sensations. Considering the situation where users adjust parameters manually, providing tactile sensations with fewer parameters is desirable. This study examines the coarsest resolution in the time and frequency dimensions necessary to present tactile sensations as realistic as vibrations recorded by the sensor. Time and frequency are fundamental parameters to express vibrations as a spectrogram, and we considered it important to investigate how much coarser the resolution could be without changing perception. We focus on the textural vibrations and the preliminary experiment compared actual texture vibrations with the reconstructed vibration as coarse as possible in the frequency dimension. The result showed that the frequency resolution above 172 Hz makes it difficult to distinguish between the vibrations. The main experiment, a similar discrimination experiment, verified the time resolution using the averaging filter of vibration intensity over time. The results indicate that with the update interval set to 30 ms, the discrimination rate compared to the original vibration is approximately 60%. This percentage is below the chance level of 75%, indicating that distinguishing between the two is difficult. Based on our experiments, it is necessary to have a frequency resolution of at least 172 Hz and a time resolution that updates intensity at a rate of 30 fps or higher to recreate tactile sensations comparable to actual vibrations.

Index Terms—Vibrotactile perception, time-frequency resolution.

I. INTRODUCTION

IN VIBROTACTILE stimuli, it is essential to reproduce realistic tactile sensations that evoke actual tactile experiences to enhance the immersiveness in VR applications and the operability of user interface systems. Vibrotactile stimuli can

Manuscript received 1 September 2022; revised 25 March 2023 and 4 July 2023; accepted 23 July 2023. Date of publication 1 August 2023; date of current version 19 September 2023. This work was supported in part by the JSPS Grant-in-Aid for Scientific Research under Grant 16H06303, and in part by JST CREST under Grant JPMJCR18A2. This paper was recommended for publication by Associate Editor Jeremy R. Cooperstock and Editor-in-Chief Seungmoon Choi upon evaluation of the reviewers' comments. (*Corresponding author: Yutaro Toide.*)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethical Review Committee of the University of Tokyo under Application No. 20-342.

Yutaro Toide, Yasutoshi Makino, and Hiroyuki Shinoda are with the Graduate School of Frontier Sciences, The University of Tokyo, Chiba 277-8561, Japan (e-mail: toide@hapis.k.u-tokyo.ac.jp; yasutoshi_makino@k.u-tokyo.ac.jp; hiroyuki_shinoda@k.u-tokyo.ac.jp).

Masahiro Fujiwara is with Nanzan University, Nagoya 466-8673, Japan (e-mail: Masahiro_Fujiwara@ipc.i.u-tokyo.ac.jp).

Digital Object Identifier 10.1109/TOH.2023.3300573

reproduce a wide range of tactile sensations, such as the sense of material [1], [2], [3], [4], [5], collision [6], [7], [8], [9], and click feel [10], [11], [12], [13], [14], [15]. To reproduce more realistic tactile sensations, systems have been proposed that pre-tune the physical parameters of vibratory waveforms or fine-tune them during vibration presentation according to user behavior. For example, in terms of the vibrotactile design system, there are GUI-based systems that directly adjust parameters, such as time transition and frequency of amplitude modulation [16], [17], [18], [19], [20], [21], and the temporal pattern of vibration pulses [22], [23], [24]. There is a system that creates optimal vibration of textures derived from affective and physical space of texture [25]. The generative adversarial network (GAN)-based automation vibrotactile design systems have also been proposed [26], [27], [28], [29]. In terms of vibration adjustment during the experience, there are tactile modulation systems that adjust each frequency component of the vibration [30], vibration rendering systems that present optimal waveforms from the user's force or speed [5], [31], [32], [33], [34], and the system that adjusts the vibration waveform according to a segment table of waveforms and a user's tracing information [35].

An important question for such a tactile design and fine-tuning system is how much temporal and frequency resolution is needed to present realistic tactile sensations. Considering the situation where users adjust tactile parameters manually, it is desirable to provide tactile sensations with a sufficiently small number of parameters. This study aims to clarify the maximum time resolution, which is the update interval of the vibration intensity, and the maximum frequency resolution, which is the frequency interval of the power spectrum, to present tactile sensations as realistic as vibrations recorded by the sensor. Since time and frequency are commonly used parameters in representing vibrations as a spectrogram, it is considered important to investigate how much coarser the resolution could be without changing human perception to express vibrotactile stimuli with fewer parameters. We attempt to provide guidelines for these parameters in the tactile design context. We believe clarifying such a necessary and sufficient time-frequency resolution is also essential to elucidate tactile perception.

Until now, frequency condition has been the primary consideration for realistic tactile reproduction [36]. This is because the four mechanoreceptors for tactile perception have independent frequency characteristics, and the frequency was considered essential for tactile reproduction [37]. Many researchers have examined frequency discrimination thresholds for each frequency and presentation site (finger [38], [39], forearms [40],

[41], hands [42]). Goble et al. and Tommerdahl et al. used a two-interval forced-choice tracking procedure to examine how frequency discrimination thresholds change after adaptation to standard stimuli presented for a long time [43], [44]. Mikkelsen et al. examined frequency discriminable bandwidths when stimuli were presented simultaneously and sequentially at the fingertips [45].

Discriminability in the frequency condition has already been discussed in terms of the compression of tactile information. Okamoto et al. transformed surface height profiles into frequency spectra using discrete cosine transform (DCT) and then quantized DCT coefficients based on Weber's law to reduce unnecessary frequency components [46]. Hassen et al. reduced imperceptible frequency information using perceptual sensitivity functions, sparse linear prediction, and DCT [47]. As the other example, Noll et al. proposed the quantization method using the discrete wavelet transform [48].

On the other hand, time condition is also necessary for realistic tactile reproduction. The tactile sensation upon contact can be improved by considering the effects of transient responses, such as tapping sensation [6]. In the case of material sensations, the temporal amplitude modulation of a specific frequency can express the sense of friction such as stick-slip and shape perception [49], which leads to improving the reality of tracing a surface. Therefore, several attempts have been made to investigate the resolution in the time domain using sinusoidal waveforms. Cao et al. investigated whether one can perceive differences in the time constant of decaying sine waves under various frequency conditions [50]. They also evaluated the discrimination ability of vibration intensity and envelope frequency using sinusoidal and amplitude-modulated vibration with high-frequency components [51].

As with the frequency condition, the viewpoint of compression of tactile information also applies to the time condition. Chaudhari et al. realized a compression method for vibrotactile data with a lower bit rate by applying the codec scheme for speech signals to vibrotactile signals based on the analogy between speech signals and vibrotactile signals [52]. Under the conditions in their article, the vibration data which is necessary for the encoding was not perceptually affected as long as it was transmitted at a frame rate of 50 fps.

Our study verified the resolution in terms of frequency and time. We examined how much resolution in the frequency dimension is needed and how often the vibration intensity pattern should be changed. We estimated frequency and temporal resolutions by applying a short-time Fourier transform (STFT) and a newly introduced filter that averages vibration intensities over time. Our experiments consist of the following two parts including the preliminary experiment.

- 1) determine the grid size of the spectrogram such that the resolution in the frequency is coarse.
- 2) determine the temporal resolution by applying a filter to average the vibration intensity over time.

Based on the above experiments, we clarify the requirements for sufficient frequency and time resolutions for keeping the quality of tactile sensation. Consequently, we obtained the following findings under the frequency and time conditions.

- *frequency*: the frequency resolution needs to be at least less than 172 Hz.
- *time*: vibration intensity should be updated by at least 30 fps.

When satisfying these requirements, the discrimination rate against the original vibration could be approximately 60%, which is lower than the chance level of 75%. Our results can provide a benchmark to achieve an application that generates and presents realistic vibrotactile sensations. It can also be applied to a real-time tactile adjustment system such as fine-tuning vibration waveforms with few parameters during user interaction.

II. PROPOSED METHOD

This study investigates the coarseness limits of temporal and frequency resolution for actual texture vibrations with various frequency components to be perceived as equivalent to the original vibration. In the preliminary experiment, we determine the grid size of the spectrograms that will have as coarse a frequency resolution as possible by varying the window length of STFT. Then, in the main experiment, we determine the temporal resolution by applying a filter that averages the vibration intensity in the time dimension of the spectrogram obtained in the preliminary experiment. This section describes the details of the texture vibration reconstruction method and vibration intensity time-averaging filter required for each experiment.

A. Reconstruction From Power Spectrum Without Using Phase Information

In this study, we performed filtering in the power spectrum domain and converted the resulting modified power spectrum to a vibration waveform to evaluate the similarity of vibration perception. The original phase information is lost when the power spectrum is processed by filtering. Therefore, we used the following procedure to estimate the phase information and restore the time signal.

The reconstruction mainly uses STFT [53] and inverse STFT (ISTFT). When applying the STFT and ISTFT, we performed 90% overlap using a Hanning window. This overlap ratio was set to obtain sufficient time resolution before applying a time-averaging filter, used in the main experiment, that reduces the time precision. The phase information of the original signal is not preserved because of the overlap resulting from the STFT and time-averaging filter. Therefore, estimating the phase and reconstructing it to the time signal is necessary.

We applied the iterative phase reconstruction method [54]. This method approaches the original phase spectrum by repeatedly applying STFT and ISTFT so that it is only an updated phase while keeping the amplitude constant. The method can produce a vibration waveform equivalent to the original in amplitude and phase.

The procedure of vibration waveform reconstruction from the spectrum by STFT is shown below.

- 1) Save the amplitude spectrum as the target amplitude spectrum and set the phase spectrum randomly as an initial value.

- 2) Apply ISTFT to the spectrogram and obtain a vibration signal in the time domain.
- 3) Apply STFT again to update the phase of the spectrogram. The calculated amplitude patterned but the initially stored value is always used.
- 4) Apply ISTFT to the updated spectrogram.
- 5) Repeat trials 3) and 4).

The number of iterations is set to 100. This is set as the number of iterations in which the root-mean-square errors of the amplitude spectrum between the actual vibration and the vibration reconstructed from the iterative phase reconstruction method are compared, and the root-mean-square errors are stable.

B. Time Averaging Filter of Vibration Intensity

We use the filter that averages the amplitude of each vibration intensity over time. In the main experiment that determined time resolution, we changed the averaging width to investigate the required temporal resolution.

Let $p[n]$ be the time-discretized power spectrum and $p_\omega[n]$ be the component at a certain frequency ω , where n is an integer that represents the discretized time. The time-averaging filter averages the values at each frequency for each interval of ΔT . Thus, the filtered value is given as follows:

$$\hat{p}_\omega[n] = \frac{\sum_{i=0}^{m-1} p_\omega \left[\left\lfloor \frac{n}{m} \right\rfloor m + i \right]}{m},$$

where i is an integer; m is the number of samples in the interval ΔT ; $\lfloor x \rfloor$ is the floor function: the largest integer that does not exceed x .

This filter adjusts the resolution only in the time dimension without changing the frequency resolution on the spectrogram. The phase information before averaging is used when applying ISTFT to reconstruct time signals.

III. VIBROTACTILE PRESENTATION SYSTEM

This experiment evaluates whether the original texture vibration and the vibration reconstructed by STFT are perceived as equivalent tactile sensations. For this purpose, this section describes a system for participants to perceive the tactile sensation of the texture vibrations with their hands.

A. Device

Fig. 1 shows a vibrotactile presentation device. The device consists of a voice coil vibrator (Acouve Vp216 series), an acrylic plate, a sponge, and a power amplifier (LEPY LP-V3). The sponge was empirically positioned in a manner that prevents the actuator from being in direct contact with the ground, thereby producing sufficient vibratory displacement and enhancing tactile perception. This is presumably because the elasticity of the sponge effectively displaces the actuator. The acrylic plate is used to enlarge the contact area. Users place their hands on the plate and feel the tactile sensation by playing the audio.

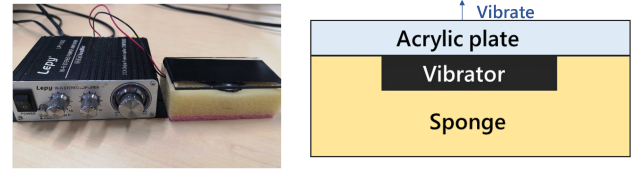


Fig. 1. Vibrotactile presentation system that consists of a voice coil vibrator, an acrylic plate, a sponge, and a power amplifier. The amplifier amplifies an audio signal played on a PC and lets the actuator vibrate. The device presents vibrations normal to the acrylic plate. Users feel tactile sensations by placing their hand on the plate.

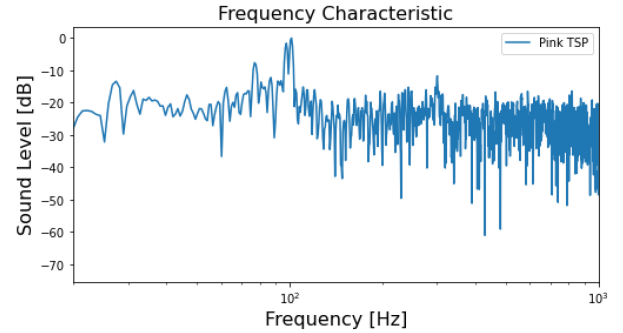


Fig. 2. Frequency characteristic using pink TSP signal.

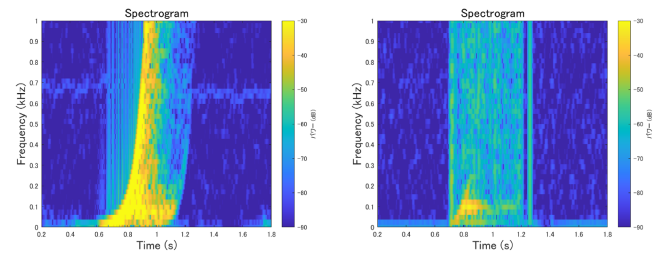


Fig. 3. Spectrogram of the pink TSP signal after passing through the device. The left figure shows the spectrogram of the input signal after passing through the amplifier. The right figure shows the spectrogram of the output signal after passing through the device. The vertical axis shows the frequency [kHz]. The horizontal axis shows the time [s] from 0.2 to 1.8 s.

B. Frequency Characteristic

Figs. 2 and 3 show the device's frequency characteristics and the spectrogram of the pink TSP signal after passing through the vibrotactile device. In Fig. 3, the left figure shows the spectrogram of the input signal that passed through the amplifier, and the right figure shows the spectrogram of the output displacement physically measured by a laser Doppler vibrometer (Onosokki LV-1800). A pink time-stretched pulse (pink TSP) signal was used as an input [55]. The input pink TSP signal was set with a sampling frequency of 100,000 Hz and 1.3 s (131,072 samples). This signal was repeated five times during the measurement. The pink TSP signal and output signals were synchronized based on the maximum value of the cross-correlation function. We use this characteristic to calibrate the input signal of vibration so that the vibration can be presented as intended.

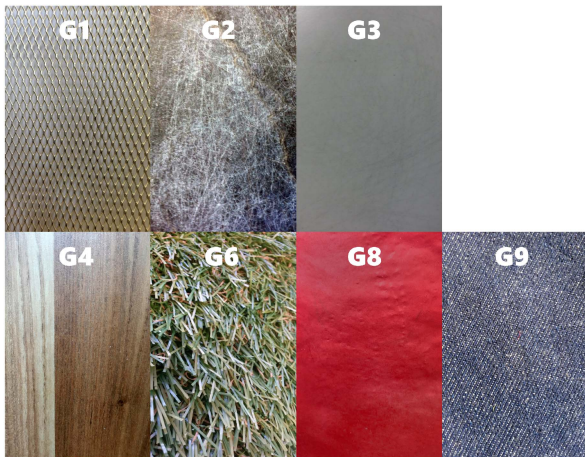


Fig. 4. Texture type (G1: Aluminum Mesh, G2: Stone Tile, G3: Ceramic Tile, G4: Cherry Tree, G6: Grass Fiber, G8: Rough Paper, G9: Jeans). G5 and G7 were not used because many stimuli were similar to G4 and G6, respectively [56].

As shown in Figs. 2 and 3, the device has the characteristics of a bandpass filter with a peak around 100 Hz. Although this filter characteristic is considered to be nonlinear in reality, it is assumed to have a linear filter characteristic, and an inverse filter of the characteristic in Fig. 2 is applied so that a waveform close to the intended input can be output.

C. Textured Vibration

Texture-related vibrations were taken from Lehrstuhl für Medientechnik (LMT) haptic texture database¹ [56]. We used seven vibrations from seven of the nine perceptually similar groups (G1–G4, G6, G8, and G9) [56]. G5 and G7 were not used because many stimuli were similar to G4 and G6, respectively. The vibrations used in each group were Aluminum Mesh, Stone Tile, Ceramic Tile, Cherry Tree, Grass Fiber, Rough Paper, and Jeans (Fig. 4). All data consist of a 5 s vibration and silent parts before and after the vibration, which was 1 s each. In general, when ISTFT is performed, both ends of the signal will protrude due to the effect of the window width. If a 5-second signal is used as is for 5 seconds, both ends of the signal will not be restored smoothly by ISTFT due to this effect, and therefore a sufficient silent interval is added before and after the signal. As a preprocess, the frequency components below 20 Hz and above 1,000 Hz were cut and applied the inverse filter corresponding to the frequency characteristic of the device. After this, the vibration was normalized from -1 to 1 .

Fig. 5 shows the spectrograms of each texture vibration. These spectrograms are computed by STFT using a Hanning window with a width of 512 samples and are pre-filtered with the inverse filter of the device. Assuming that the sampling frequency is f_s and the number of samples corresponding to the window width is N , then the frequency resolution is determined by $\Delta f = f_s/N$. Therefore, in these figures, the frequency resolution is

¹The database can be accessed at <https://zeus.lkn.ei.tum.de/downloads/texture/>.

$\Delta f = 86.1$ Hz. It is important to note that quantization is performed in steps of 86.1 Hz, but the first switch is symmetrically positioned at $\pm\Delta f/2$. As a result, the first transition occurs at 43 Hz, followed by changes at 129.1 Hz, 215.2 Hz, and so on. Although the signal is truncated below 20 Hz, the average intensity from 0 to 43 Hz is depicted at the bottom of the graph and appears with a different height than the other grids.

The figure shows that texture-related vibrations with various frequency characteristics are prepared, such as vibrations with strong low-frequency components (G1), vibrations with strong high-frequency components (G3), and vibrations with relatively uniform intensity (G6). It can also be confirmed that a wide range of vibration intensities over time can be prepared, from highly variable to relatively stable.

IV. PRELIMINARY EXPERIMENT: DETERMINE SPECTROGRAM GRID SIZE

The preliminary experiment aims to determine the coarsest frequency resolution of the spectrogram to be indistinguishable from the original texture vibrations. We determine the maximum frequency resolution to be indistinguishable from the original signal by comparing actual texture vibrations with those reconstructed through STFT with different window sizes. Ideally, both time and frequency resolutions should be varied simultaneously to determine the appropriate resolution, but since the number of combinations would be enormous, we first estimated the frequency resolution conditions as a preliminary experiment. The time condition was then investigated as the main experiment.

A. Participants

In this experiment, three participants (two males and one female) in their 20 s participated. All of them were right-handed and had no health concerns.

B. Procedure

Fig. 6 shows the experimental workflow for one trial. In a trial, we prepared one reference stimulus (*Ref*), which is the original signal of recorded vibration and two comparison stimuli (C_i , $i = Ref$ or *STFT*). Here, C_{Ref} is identical to *Ref*, and C_{STFT} is a reconstructed vibration from the power spectrum with coarse resolution in frequency. *Ref* (shown in red) and C_i (shown in blue) were presented alternately, and participants answered which stimulus of C_i was C_{Ref} . Therefore, if the re-transformed signal from the power spectrum after STFT (C_{STFT}) is similar to the reference signal *Ref*, the correct answer rate, which is the ratio of each participant selecting C_{Ref} from the entire trial, should be close to 50%.

This trial was repeated 140 times = 7 (textures) \times 4 (window widths) \times 5 (trials) in total. The order of the stimuli and the C_i were presented randomly. The interval between each stimulus depends on the subject's button press timing since the next stimulus is set to start when the subject presses the button. There was an interval of at least 2 seconds because there was a silent period before and after the stimulus.

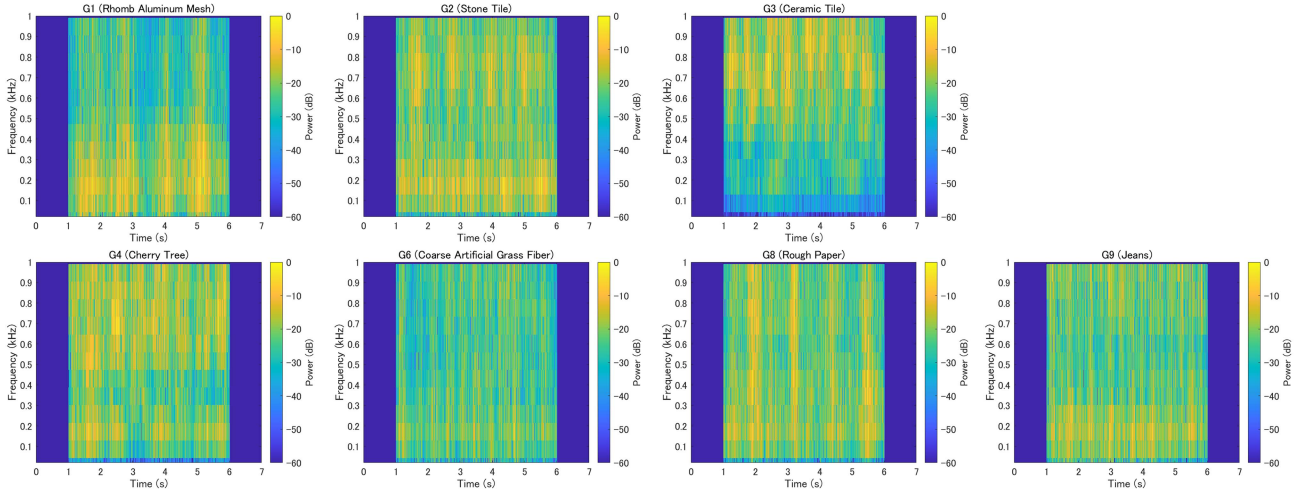


Fig. 5. Spectrogram of each texture vibration. STFT uses a hanning window of 512 samples. The vertical axis shows frequencies up to 1000 Hz, and the horizontal axis shows time up to 7.0 s. Vibration intensities are expressed in dB. The upper rows are aluminum mesh, stone tile, and ceramic tile from left to right, and the lower rows are cherry Tree, coarse artificial grass fiber, rough paper, and jeans from left to right.



Fig. 6. One trial workflow (*Ref*: Reference Stimuli, C_i : Compare Stimuli ($i = Ref$ or *STFT*)). In a trial, *Ref* and C_i were presented alternately, and participants answered whether the first or second half of the stimulus was C_{Ref} in a trial. The order of C_i was presented randomly. There were silent parts of 1 s before and after stimuli. After each stimulus presentation, there were intervals because the next stimulus starts after the keyboard input is accepted.

C. Experimental Condition

1) *Stimuli Condition*: We prepared seven types of textured vibrations which were selected in Section III-C. The vibrations were presented under the sampling frequency of 44,100 Hz. For C_{STFT} , a Hanning window with a 90% overlap condition was used, and four types of window length (64, 256, 1024, and 4096 samples) were prepared. The condition of the number of window lengths was set to only four to account for the experimental burden on the participants. In addition, these sample numbers were set based on the need for a wide search range because the digit of the frequency resolution was not apparent.

The wider the window width, the higher the frequency resolution; thus, the condition of 64 is the coarsest frequency resolution condition. Since the frequency resolution is calculated as $\Delta f = f_s/N$, they were 689, 172, 43, and 11 Hz for the window sample sizes of 64, 256, 1024, and 4096, respectively, when the sampling frequency was 44,100 Hz.

2) *Participant Condition*: Participants were asked to wear headphones with white noise to prevent discrimination by sound. No restrictions were imposed on visual perception since it was considered not to influence discrimination. As for visual information, participants mainly saw the operating interface on the computer screen. They could also see their hand and the device. However, there was no difference in the visual appearance of the devices in all conditions. All participants were instructed to touch the device with their dominant hand, but they were not instructed how to touch it specifically. This was due to

consideration of the effects of prolonged exposure to vibration stimuli. Since the influence of the tactile sensation perceived before a certain trial could make it difficult to discriminate that trial, we asked the subjects to respond using the area and the way of touching that they felt most likely to be perceived at that time. The participants basically used finger pulps from the index finger to the little finger or palm. Throughout the 140 trials, they were allowed to take a break every 14 trials, considering their fatigue.

D. Result

The experiment aims to determine the coarsest frequency resolution not to be distinguished from the original texture vibrations. Therefore, the expected outcome is a correct answer rate close to 50%. Figs. 7 and 8 show the results to distinguish between actual texture vibrations and reconstructed vibrations by applying STFT. Fig. 7 shows the correct answer rates for each texture. Fig. 8 shows the correct answer rates for the entire texture. The correct answer rates for the entire textures are 0.81 ± 0.18 (64 sample), 0.56 ± 0.32 (256 sample), 0.55 ± 0.21 (1024 sample), and 0.52 ± 0.21 (4096 sample). Therefore when the window widths were greater than 256 samples, i.e., when the frequency resolutions were smaller than 172 Hz, the correct answer rate was close to 0.5, which indicates that the vibration is indistinguishable from the actual vibration. Since the variance of 256 samples was much larger than that of 1024 samples, in the subsequent experiments, we used 512 samples ($\Delta f = 86.1$ Hz)

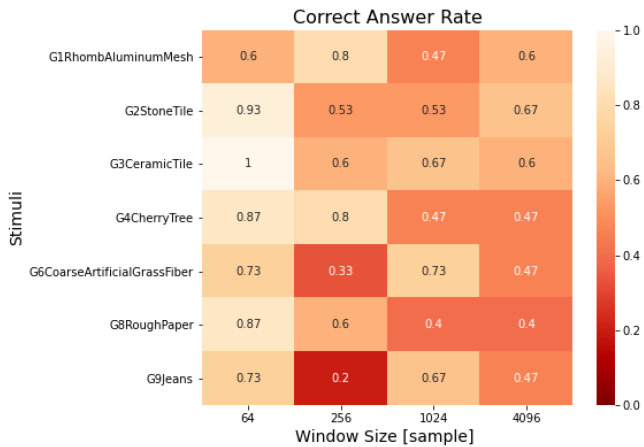


Fig. 7. Correct answer rate to distinguish between Ref , and C_i stimuli for each texture. The horizontal axis expresses the window size (from left to right: 64, 256, 1024, and 4096 samples, i.e., $\Delta f = 689, 172, 43,$ and 11 Hz), and the vertical axis represents the stimuli type (from top to bottom: Aluminum Mesh, Stone Tile, Ceramic Tile, Cherry Tree, Coarse Artificial Grass Fiber, Rough Paper, and Jeans).

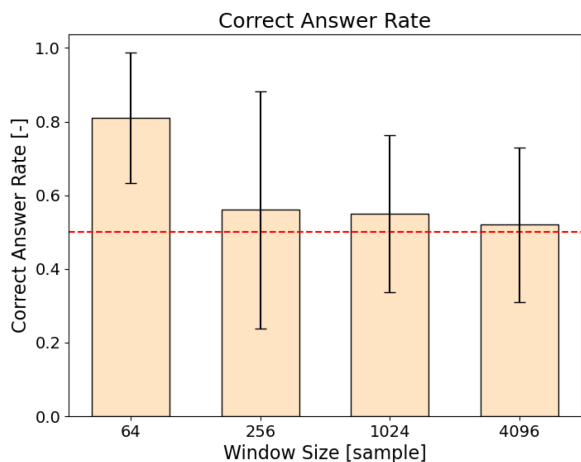


Fig. 8. Correct answer rate to distinguish between Ref , and C_i stimuli for all textures. The horizontal axis expresses the window size (from left to right: 64, 256, 1024, and 4096 samples, i.e., $\Delta f = 689, 172, 43,$ and 11 Hz). The vertical axis expresses the correct answer rate. The correct answer rate is the ratio of each participant selecting C_{Ref} from the entire trial.

as the window width to obtain a stable and difficult-to-discriminate evaluation with as coarse a frequency resolution as possible.

V. EXPERIMENT: DETERMINE TIME RESOLUTION

We then conducted an experiment to determine the temporal resolution of the power spectrum to be perceived as equivalent to the original vibration. Based on the grid size of the spectrogram determined in the previous experiment, we controlled the temporal resolution by applying a filter that averages the intensity of the vibration in the temporal dimension. The experiment is conducted in the same way as in the previous section by comparing the real texture vibration with the vibration reconstructed after applying the time-averaging filter.

A. Participants

In this experiment, eleven participants (eight males and three females) aged 24.7 ± 2.0 years old participated. All of them were right-handed and had no health concerns. Two of them (one male and one female) also participated in the preliminary experiment. Before the experiment, written informed consent and a brief explanation of the experiment were given. The experiment was divided into two days; each day took about 2 hours to consider the burden on the participants. After the experiment, an honorarium was paid to each participant. In this experiment, the ethical committee at the University of Tokyo approved the experimental procedure 20-342.

B. Experimental Condition

1) *Stimuli Condition*: We prepared the texture vibration under the same conditions in Section IV. Then, we applied STFT to these vibrations with 512 samples and a 90% overlap condition and the initial time resolution was 1.2 ms. This temporal resolution is sufficiently indistinguishable from actual texture vibrations, as confirmed in the previous experiment in Section IV. Next, we applied a temporal averaging filter to reduce the fineness of the information. The time resolutions obtained by applying the time-averaging filter to the above vibrations were $\Delta T = 3, 10, 30,$ and 100 ms for logarithmically equal intervals.

2) *Participant Condition*: The experimental conditions are basically the same as in Section IV. Auditory information was cut off with white noise, but visual information was not restricted. All participants were instructed to touch the device with their dominant hand, but they were not instructed how to touch it specifically. In this experiment, most participants used finger pulps from the index finger to the little finger or palm. We let participants take a break every 14 trials. The experiment was paused at the timing of 140 trials, with subsequent trials conducted on a different day. The total number of trials was 280.

C. Result

Fig. 9 shows the spectrogram results after applying the time average filter to the Aluminum Mesh (G1). The filter is applied to the original spectrogram (top) with the width of $\Delta T = 3$ (middle left), 10 (middle right), 30 (lower left), and 100 (lower right) ms, respectively. The result shows that the time resolution is successfully adjusted for all ΔT while keeping the frequency resolution of 86.1 Hz. Fig. 10 shows the actual reconstructed vibration waveforms. It can be seen that similar waveforms are reconstructed for no filtering condition $\Delta T = 1.2$ ms compared to the original, indicating the validity of the STFT and ISTFT iterative methods. As the time resolution reduces, the waveform gradually breaks down. In particular, the fine peaks disappear.

Figs. 11 and 12 show the correct answer rate when distinguishing between actual texture vibrations and those after applying a time-averaging filter. Fig. 11 shows the correct answer rates for each texture. Fig. 12 shows the average correct answer rates of the entire texture. The average correct answer rates of the

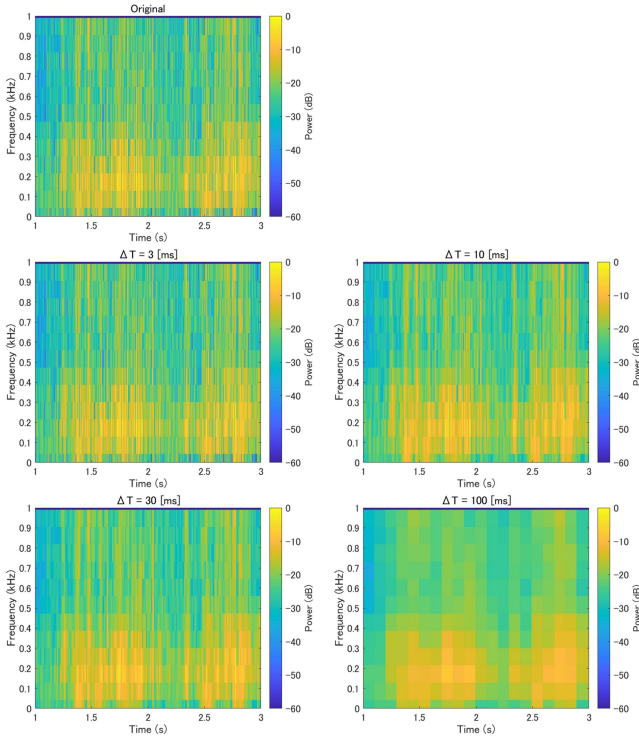


Fig. 9. Spectrogram of Aluminum Mesh after applying the time-averaging filter (First row: original with 512 window length whose overlap was 90 %; second row: $\Delta T = 3$ and 10 ms; third row: $\Delta T = 30$ and 100 ms). The vertical axis shows frequencies up to 1000 Hz, and the horizontal axis shows time from 1.0 to 3.0 s. Vibration intensities are expressed in dB. It can be seen that the time resolution becomes coarser as ΔT increases, whereas the frequency resolution remains constant.

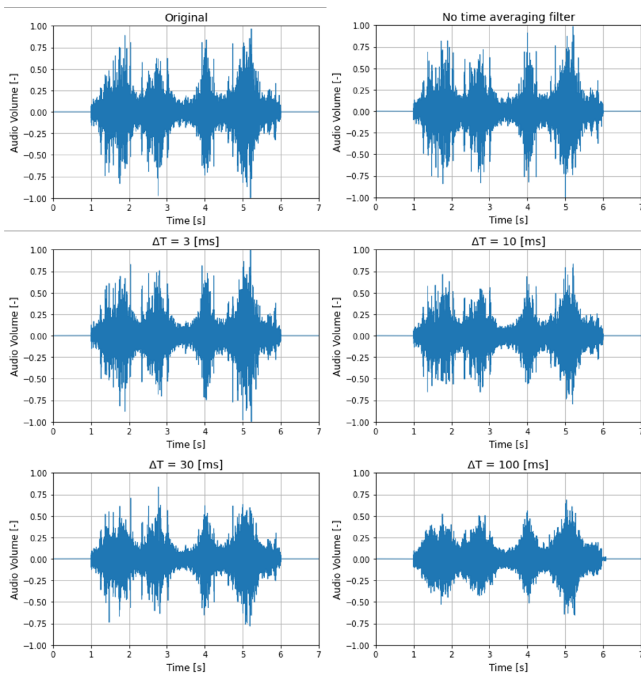


Fig. 10. Aluminum Mesh waveforms after applying time-averaging filter (first row: original and no time-averaging filter; second row: $\Delta T = 3$ and 10 ms; third row: $\Delta T = 30$ and 100 ms). The vertical axis shows audio volume based on Unity's audio clip format, and the horizontal axis shows time up to 7.0 s.

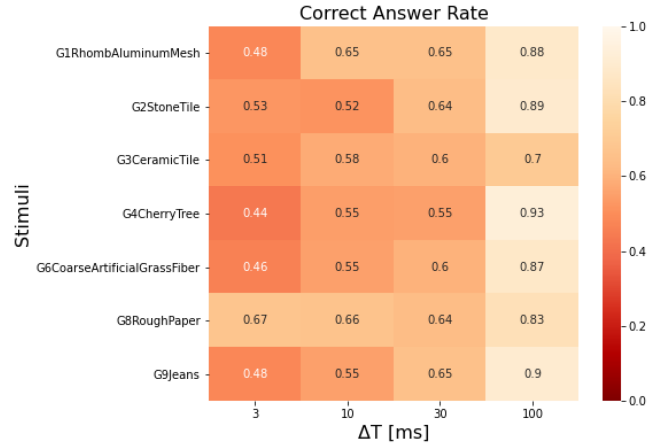


Fig. 11. Correct answer rate to distinguish between Ref_i and C_i stimuli for each texture. The horizontal axis expresses the time resolution (from left to right: $\Delta T = 3, 10, 30$, and 100 ms). The vertical axis expresses the stimuli type (from top to bottom: Aluminum Mesh, Stone Tile, Ceramic Tile, Cherry Tree, Coarse Artificial Grass Fiber, Rough Paper, and Jeans).

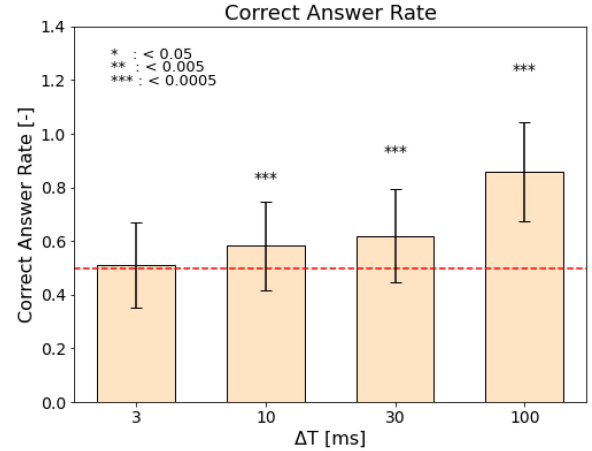


Fig. 12. Correct answer rate to distinguish between Ref_i and C_i stimuli for all textures. The horizontal axis expresses the time resolution (from left to right: $\Delta T = 3, 10, 30$, and 100 ms). The vertical axis expresses the correct answer rate. The correct answer rate is the ratio of each participant selecting C_{Ref} from the entire trial. The asterisks in the chart indicate p-values calculated by the Wilcoxon test (* $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$).

entire texture were 0.51 ± 0.16 (3 ms), 0.58 ± 0.17 (10 ms), 0.62 ± 0.17 (30 ms), and 0.86 ± 0.19 (100 ms). From these figures, we obtained the following conclusions.

- 1) As the temporal resolution becomes finer, it becomes more difficult to distinguish the difference from the reference stimulus.
- 2) The time resolution shorter than 30 ms is sufficient to reproduce the tactile sensation of real textures.
- 3) These trends are independent of texture.

We conducted one-sample equivalence tests to compare the correct answer rate for each time resolution with that of 0.5 (completely random responses). The Shapiro–Wilk test ($p < 0.05$) and Wilcoxon tests were employed for the normality and equivalence tests since the normality hypothesis was rejected. The result showed that the p-values were 0.85 (3 ms),

3.3×10^{-4} (10 ms), 1.1×10^{-6} (30 ms), and 6.5×10^{-14} (100 ms). Therefore, there was no significant difference for the 3 ms, with the correct answer rate of 0.5 as the indistinguishable indicator.

Additionally, we conducted the Kruskal–Wallis and Conover tests to examine whether discrimination becomes more difficult as the time resolution becomes finer. Since the Shapiro–Wilk test could not adopt the normality hypothesis, the above test method, a nonparametric technique, was employed. The Kruskal–Wallis test showed a significant difference with a p -value of 9.2×10^{-23} . Then, we employed the Conover test to confirm significant differences between the 100 ms and other three conditions, as well as between the 3 ms and other three conditions ($p < 0.01$ in all cases). However, $p = 0.21$ was found for the 30 and 10 ms conditions.

VI. DISCUSSION

In the preliminary experiment, we compared the tactile sensations of the actual texture vibration with those of the vibration reconstructed by STFT. Consequently, we obtained that the difference was distinguishable for 64 samples, and it became more difficult to distinguish the difference for 256 or more samples. The frequency resolution calculated from each window width was $\Delta f = 689$ Hz for 64 samples and $\Delta f = 172$ Hz for 256 samples. Therefore, if the frequency resolution is at least less than 172 Hz, the tactile stimulus is equivalent to the actual textural vibration.

In the preliminary experiment, we obtained tactile sensations equivalent to those of texture vibrations for window widths of 256 samples or larger. Generally, as the window width decreases, the frequency resolution decreases and is strongly affected by smoothing. Therefore, this effect is thought to produce a frequency pattern that differs from the original signal and changes perception.

It should be noted that the effect of this smoothing depends on the frequency characteristics of the vibration under consideration. The spectrogram reconstructed from 256 samples of STFT and the original G1 spectrogram (1024 samples each) is shown in the left figures of Fig. 13, and the vibration intensities at 1.7 s and 5.1 s, which show the strong amplitude spectrum, are shown in the right figures of Fig. 13. The right figure shows that the vibration reconstructed by STFT with a small window width is affected by smoothing, but the effect appears to be small. This is because the vibrations used had broadband and relatively flat frequency characteristics.

On the other hand, the effect of this smoothing is remarkable when the frequency components of the signal have peaky characteristics. To demonstrate this effect, we created a spectrogram having peaky frequency characteristics and compared it through the same operation. Fig. 14 is the artificially generated vibration that was created by applying the Gaussian filter ($\sigma = 5$) at intervals of 200 Hz to the Aluminum Mesh vibration and Fig. 15 shows the result of the same operation on the spectrogram. As shown in the figure, the peaky frequency components were strongly affected by the smoothing effect compared to Fig. 13,

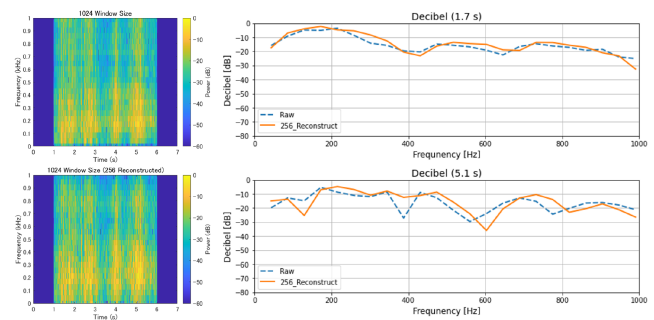


Fig. 13. Effect of smoothing on Aluminum Mesh. The upper left and lower left figures show the original G1 and the reconstructed spectrogram with a 256 sample window length, respectively. These spectrograms were obtained with 1024 samples and a 90% overlap condition. The right figures show the vibration intensity at 1.7 s and 5.1 s when the amplitude spectrum is strong (blue: original, orange: reconstructed). The vertical axis is in dB, and the horizontal axis is in frequency up to 1000 Hz. It is shown that the reconstructed vibration is affected by the smoothing of frequency components.

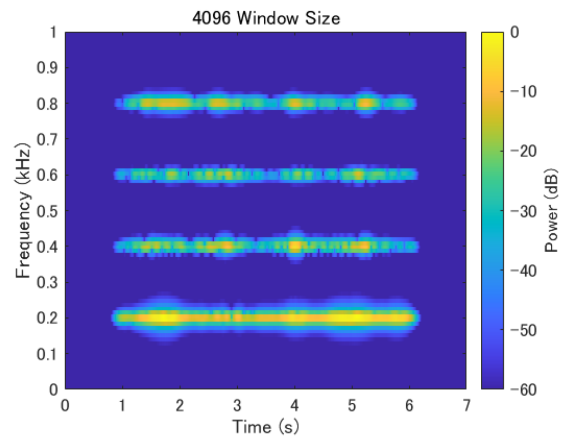


Fig. 14. Spectrogram of narrow-band vibration. The vibration was created by applying Gaussian filters ($\sigma = 5$) at 200 Hz intervals to the Aluminum Mesh vibration. The spectrogram was obtained with a window length of 4096 samples and a 90% overlap condition.

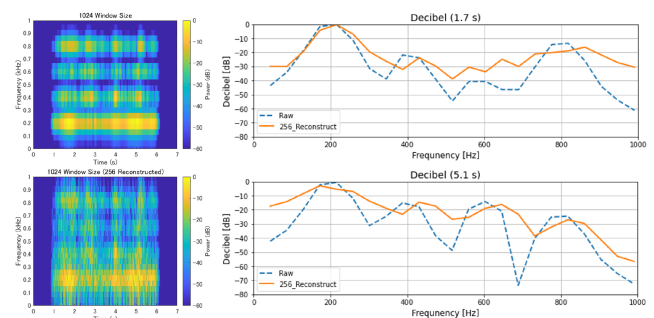


Fig. 15. Effect of smoothing on narrow-band vibration. The vibration was created by applying Gaussian filters ($\sigma = 5$) at 200 Hz intervals to the Aluminum Mesh vibration. The upper left and lower left figures show the original G1 and the reconstructed spectrogram with a 256 sample window width, respectively. These spectrograms were obtained with 1024 samples and a 90% overlap condition. As Fig. 13, the right figures show the vibration intensity at 1.7 s and 5.1 s (blue: original, orange: reconstructed). The vertical axis is in dB, and the horizontal axis is in frequency up to 1000 Hz.

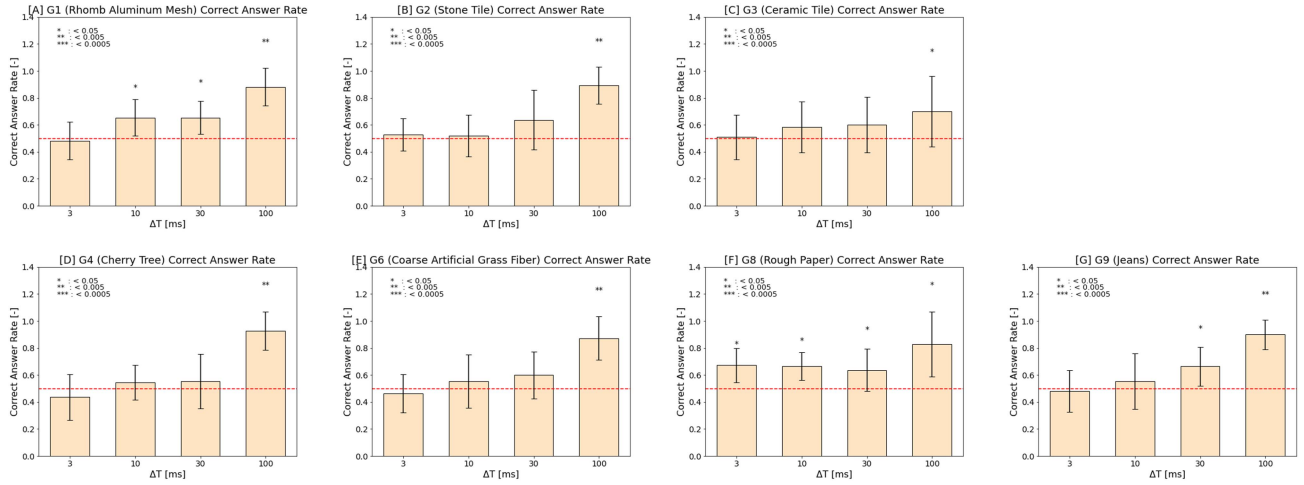


Fig. 16. Correct answer rate to distinguish between Ref and C_i stimuli for each texture. The horizontal axis represents the time resolution (from left to right: $\Delta T = 3, 10, 30, 100$ ms). The vertical axis represents the correct answer rate. We performed a one-sample equivalence test to compare 0.5 for the correct answer rate for each time resolution. The asterisks in the chart indicate p-values calculated by the Wilcoxon test (* $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$).

and the frequency patterns changed significantly accordingly. From the above discussion, it should be noted that the results of the current experiment were for vibrations with relatively flat frequency characteristics and may not be directly applicable to vibrations with peaky frequency characteristics.

This preliminary experiment did not show the effect of the vibration waveform edges on the frequency resolution discrimination. The original vibration rises instantaneously after a silent interval. On the other hand, the reconstructed vibration rises smoothly due to the effect of the window considering the overlap. Therefore, the wider the applied window width, the less similar the edges become because the waveform spreads back and forth due to the smoothing. By this effect, we were concerned that the discriminant would be based only on the discontinuity at the rise of the vibration waveform rather than on the textures' difference in the frequency resolution. However, the actual results showed the opposite result, that the narrower window became less similar, suggesting that the effect of the edges is small and the difference in frequency resolution is visible.

Although the number of participants in the preliminary experiment that determined spectrogram grid size was only three, it is considered sufficient as a preliminary study. This is because the main experiment showed the same indistinguishable vibration conditions for 11 subjects with a window width of 512 samples ($\Delta f = 86.1$ Hz). Although it can be said that 512 samples of the window width are indistinguishable, the preliminary experiment is insufficient to determine the more precise threshold of frequency resolution in terms of how short the window width can be. A threshold may seem to exist at around 256, but it is difficult to say precisely. Further verification is needed.

In the main experiment, we compared the tactile sensations of the actual texture vibration with those of the vibration to which the time-averaging filter was applied. The experimental results showed that the difference could be discriminated in about 86% of the case when the time resolution was 100 ms. Meanwhile, the finer the time resolution, the more difficult it

becomes to distinguish the difference from the reference stimuli. In particular, the correct answer rate for the 3 ms window was close to 50%, meaning that no difference between the two was perceived. This result is consistent with the correct answer rate for the 256 and 1024 sample window widths in the preliminary experiment. Furthermore, the same trend for textures with various frequency spectra can be seen in Fig. 16. These results indicate that if the vibration intensity is varied at update intervals of 30 ms (approximately 33 fps) or less, the tactile stimulus is equivalent to the actual textural vibration. According to Chaudhari et al. [52], it can be inferred that under their defined conditions, the vibration data which is necessary for the encoding is not perceptually affected as long as it is transmitted at a frame rate of 50 fps. Although not directly comparable to their results, their frame rate is within the range of our results. Therefore, we consider the present results to be reasonable.

In Fig. 16, Ceramic Tile texture (G3) showed low correct answer rates and was always difficult to identify. This may be due to the weak vibration intensity for the lower frequency (around 200 Hz) and the weak time variation as is shown in Fig. 5. Conversely, Rough Paper (G8), which exhibits strong instantaneous peaks in time, showed a relatively high percentage of correct responses independent of temporal resolution conditions compared to the other textures.

In this study, the vibration waveforms were mainly steady-state. Within this range, the results were robust to vibrations with various frequency spectra. However, in some cases, such as the textures of G1 and G8, where instantaneous intensity changes are observed in some parts (Figs. 5, 17), the differences may be perceived, even with a finer time width, as shown in Fig. 16. Therefore, validating time and frequency resolution in presenting instantaneous stimuli, such as click sensation, is a future issue.

In this experiment, quantization was linear in time and frequency conditions. In particular, STFT quantized about 12 dimensions in the frequency by dividing the region up to 1000 Hz

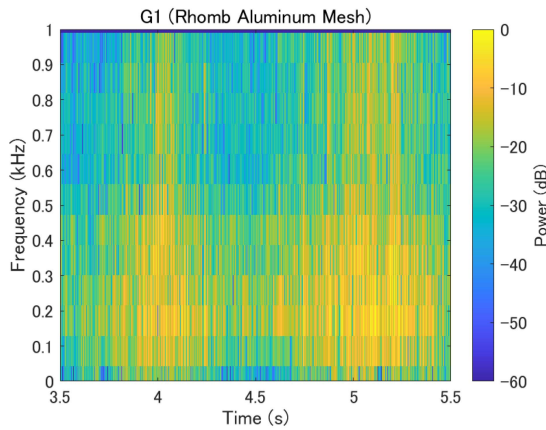


Fig. 17. Spectrogram of Aluminum Mesh. The vertical axis shows frequencies up to 1000 Hz, and the horizontal axis shows time from 3.5 to 5.5 s. Vibration intensities are expressed in dB.

at $\Delta f = 86.1$ Hz. It might be possible to give an equivalent sensation with fewer frequency divisions, by making the divisions logarithmic like the wavelet transform.

Our experiment system is based on passive stimuli with a fixed spatial position of vibration and not on active tactile feedback, such as tracing a material surface with a pen-shaped tool, a display surface with a fingertip, or devices that can control the spatiotemporal distribution. In active tactile feedback, the material surface, the influence of tools, and the user's tracing motion have effects on feeling [57]. This should be examined as a next step.

Throughout this experiment, we did not instruct how to touch the device, considering the effects of prolonged exposure to the vibration stimuli. Since the original and reconstructed vibrations were presented with the same device, it is assumed that a consistent within-subject trend is obtained, unless the subject significantly changes the way they touch the device. Although the possibility remains that there may be differences between subjects, the overall trend is the same: the coarser the temporal resolution, the easier it is to discriminate. The variance of the percentage of correct answer rate in Fig. 12 is also within 0.2, respectively. Therefore, under the conditions of the present study, the effect of the difference in the touching style is considered to be small.

Tactile presentation to other body parts than the hand also should be considered. As observed in the low correct answer rate for Ceramic Tile in the time resolution determination experiment, the time resolution can be coarser when a vibration intensity in the low-frequency band is weak. Therefore, a similar result to Ceramic Tile may be obtained in the area with weak tactile perceptual sensitivity to the low-frequency band.

In summary, the spatiotemporal patterns, active tactile sensation, and presentation site may lead to different outcomes from our results, and there is room for consideration.

VII. CONCLUSION

In this study, we verified how much the resolution of vibration intensity in the time and frequency could be reduced while keeping the quality of the actual texture vibrotactile sensation

using the STFT and time-averaging filter. We conducted two experiments. In the preliminary experiment, we determine the spectrogram's grid size such that it maximally coarsens the resolution in the frequency dimension. Then, in the main experiment, we determine the time resolution by applying the newly introduced filter which averages the vibration intensity over time. Consequently, the following findings were clarified as requirements for resolution in time and frequency conditions to maintain the quality of realistic vibrotactile sensation.

- frequency: the frequency resolution needs to be at least less than 172 Hz.
- time: the vibration intensity should be updated at least 30 fps.

When satisfying these requirements, the discrimination rate with the original vibration was suppressed at about 60%, which is lower than the chance level of 75%. The above results can be used as a basis for realistic tactile reproduction. They can also be used to design tactile sensation generation and presentation systems and for real-time tactile adjustment, such as fine-tuning vibration waveforms during user interaction.

ACKNOWLEDGMENT

The authors would like to thank Enago (www.enago.jp) for the English language review.

REFERENCES

- [1] T. Ahmaniemi, J. Marila, and V. Lantz, "Design of dynamic vibrotactile textures," *IEEE Trans. Haptics*, vol. 3, no. 4, pp. 245–256, Oct.–Dec. 2010.
- [2] P. Strohmeier and K. Hornbæk, "Generating haptic textures with a vibrotactile actuator," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2017, pp. 4994–5005.
- [3] J. M. Romano and K. J. Kuchenbecker, "Creating realistic virtual textures from contact acceleration data," *IEEE Trans. Haptics*, vol. 5, no. 2, pp. 109–119, Apr.–Jun. 2012.
- [4] S. Saga and R. Raskar, "Simultaneous geometry and texture display based on lateral force for touchscreen," in *Proc. IEEE World Haptics Conf.*, 2013, pp. 437–442.
- [5] R. H. Osgouei, J. R. Kim, and S. Choi, "Data-driven texture modeling and rendering on electrovibration display," *IEEE Trans. Haptics*, vol. 13, no. 2, pp. 298–311, Apr.–Jun. 2020.
- [6] A. M. Okamura, M. R. Cutkosky, and J. T. Dennerlein, "Reality-based models for vibration feedback in virtual environments," *IEEE/ASME Trans. Mechatron.*, vol. 6, no. 3, pp. 245–252, Sep. 2001.
- [7] K. J. Kuchenbecker, J. Fiene, and G. Niemeyer, "Improving contact realism through event-based haptic feedback," *IEEE Trans. Visual. Comput. Graph.*, vol. 12, no. 2, pp. 219–230, Mar.–Apr. 2006.
- [8] G. Park and S. Choi, "A physics-based vibrotactile feedback library for collision events," *IEEE Trans. Haptics*, vol. 10, no. 3, pp. 325–337, Jul.–Sep. 2017.
- [9] C. Park, J. Park, S. Oh, and S. Choi, "Realistic haptic rendering of collision effects using multimodal vibrotactile and impact feedback," in *Proc. IEEE World Haptics Conf.*, 2019, pp. 449–454.
- [10] M. Fukumoto and T. Sugimura, "Active click: Tactile feedback for touch panels," in *Proc. CHI'01 Extended Abstr. Hum. Factors Comput. Syst.*, 2001, pp. 121–122.
- [11] E. Koskinen, T. Kaaresoja, and P. Laitinen, "Feel-good touch: Finding the most pleasant tactile feedback for a mobile touch screen button," in *Proc. 10th Int. Conf. Multimodal Interfaces*, 2008, pp. 297–304.
- [12] S. Kim and G. Lee, "Haptic feedback design for a virtual button along force-displacement curves," in *Proc. 26th Annu. ACM Symp. User Interface Softw. Technol.*, 2013, pp. 91–96.
- [13] K. Tashiro, Y. Shiokawa, T. Aono, and T. Maeno, "Realization of button click feeling by use of ultrasonic vibration and force feedback," in *Proc. IEEE World Haptics 3rd Joint EuroHaptics Conf. Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst.*, 2009, pp. 1–6.

- [14] C. Park, J. Yoon, S. Oh, and S. Choi, "Augmenting physical buttons with vibrotactile feedback for programmable feels," in *Proc. 33rd Annu. ACM Symp. User Interface Softw. Technol.*, 2020, pp. 924–937.
- [15] B. Sadia, S. E. Emgin, T. M. Sezgin, and C. Basdogan, "Data-driven vibrotactile rendering of digital buttons on touchscreens," *Int. J. Hum.-Comput. Stud.*, vol. 135, 2020, Art. no. 102363.
- [16] C. Swindells, E. Maksakov, K. E. MacLean, and V. Chung, "The role of prototyping tools for haptic behavior design," in *Proc. IEEE 14th Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst.*, 2006, pp. 161–168.
- [17] J. Ryu and S. Choi, "posvibeditor: Graphical authoring tool of vibrotactile patterns," in *Proc. IEEE Int. Workshop Haptic Audio Vis. Environments Games*, 2008, pp. 120–125.
- [18] S. Panëels, M. Anastassova, and L. Brunet, "TactiPEd: Easy prototyping of tactile patterns," in *Proc. IFIP Conf. Hum.-Comput. Interaction.*, 2013, pp. 228–245.
- [19] H. Seifi, C. Anthonypillai, and K. E. MacLean, "End-user customization of affective tactile messages: A qualitative examination of tool parameters," in *Proc. IEEE Haptics Symp.*, 2014, pp. 251–256.
- [20] H. Seifi, K. Zhang, and K. E. MacLean, "Vibviz: Organizing, visualizing and navigating vibration libraries," in *Proc. IEEE World Haptics Conf.*, 2015, pp. 254–259.
- [21] E. Pezent, B. Cambio, and M. K. O'Malley, "Syntacts: Open-source software and hardware for audio-controlled haptics," *IEEE Trans. Haptics*, vol. 14, no. 1, pp. 225–233, Jan.-Mar. 2021.
- [22] S. A. Brewster and L. M. Brown, "Tactons: Structured tactile messages for non-visual information display," in *Proc. Australas. User Interface Conf.*, 2004, pp. 15–23.
- [23] D. Ternes and K. E. MacLean, "Designing large sets of haptic icons with rhythm," in *Proc. Int. Conf. Hum. Haptic Sens. Touch Enabled Comput. Appl.*, 2008, pp. 199–208.
- [24] J. Lee, J. Ryu, and S. Choi, "Vibrotactile score: A score metaphor for designing vibrotactile patterns," in *Proc. IEEE World Haptics 3rd Joint EuroHaptics Conf. Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst.*, 2009, pp. 302–307.
- [25] W. Hassan, A. Abdulali, and S. Jeon, "Authoring new haptic textures based on interpolation of real textures in affective space," *IEEE Trans. Ind. Electron.*, vol. 67, no. 1, pp. 667–676, Jan. 2020.
- [26] Y. Ban and Y. Ujitoko, "TactGAN: Vibrotactile designing driven by GAN-based automatic generation," in *Proc. SIGGRAPH Asia Emerg. Technol.*, 2018, pp. 1–2.
- [27] Y. Ujitoko, Y. Ban, and K. Hirota, "Gan-based fine-tuning of vibrotactile signals to render material surfaces," *IEEE Access*, vol. 8, pp. 16656–16661, 2020.
- [28] C. Hernandez-Mejia, X. Ren, A. Thabuis, J. Chavanne, P. Germano, and Y. Perriard, "Generative adversarial networks for localized vibrotactile feedback in haptic surfaces," in *Proc. IEEE 24th Int. Conf. Elect. Machines Syst.*, 2021, pp. 105–110.
- [29] S. Cai, Y. Ban, T. Narumi, and K. Zhu, "Frictgan: Frictional signal generation from fabric texture images using generative adversarial network," in *Proc. Int. Conf. Artif. Reality Telexistence Eurographics Symp. Virtual Environments*, 2020, pp. 11–15.
- [30] "Yubi-recorder (finger vibration recorder) products | tec gihan co., ltd.," 2012, [Online]. Available: <http://www.tecgihan.co.jp/en/products/tactile-and-texture-measurement/yubi-recorder/yubi-recorder-finger-vibration-recorder/>
- [31] H. Culbertson, J. Unwin, B. E. Goodman, and K. J. Kuchenbecker, "Generating haptic texture models from unconstrained tool-surface interactions," in *Proc. IEEE World Haptics Conf.*, 2013, pp. 295–300.
- [32] H. Culbertson, J. Unwin, and K. J. Kuchenbecker, "Modeling and rendering realistic textures from unconstrained tool-surface interactions," *IEEE Trans. haptics*, vol. 7, no. 3, pp. 381–393, Jul.-Sep. 2014.
- [33] G. Ilkhani, M. Aziziaghdam, and E. Samur, "Data-driven texture rendering on an electrostatic tactile display," *Int. J. Hum.-Comput. Interact.*, vol. 33, no. 9, pp. 756–770, 2017.
- [34] N. Sabnis, D. Wittchen, C. N. Reed, N. Pourjafarian, J. Steimle, and P. Strohmeier, "Haptic servos: Self-contained vibrotactile rendering system for creating or augmenting material experiences," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2023, pp. 1–17.
- [35] W. Nai, J. Liu, C. Sun, Q. Wang, G. Liu, and X. Sun, "Vibrotactile feedback rendering of patterned textures using a waveform segment table method," *IEEE Trans. Haptics*, vol. 14, no. 4, pp. 849–861, Oct.-Dec. 2021.
- [36] L. A. Jones and N. B. Sarter, "Tactile displays: Guidance for their design and application," *Hum. Factors*, vol. 50, no. 1, pp. 90–111, 2008.
- [37] A. Gescheider, S. J. Bolanowski, and K. R. Hardick, "The frequency selectivity of information-processing channels in the tactile sensory system," *Somatosensory Motor Res.*, vol. 18, no. 3, pp. 191–201, 2001.
- [38] O. Franzén and J. Nordmark, "Vibrotactile frequency discrimination," *Percep. Psychophys.*, vol. 17, no. 5, pp. 480–484, 1975.
- [39] G. D. Goff, "Differential discrimination of frequency of cutaneous mechanical vibration," *J. Exp. Psychol.*, vol. 74, no. 2p1, pp. 294–299, 1967.
- [40] M. Rothenberg, R. T. Verrillo, S. A. Zoharian, M. L. Brachman, and S. J. Bolanowski Jr, "Vibrotactile frequency for encoding a speech parameter," *J. Acoustical Soc. Amer.*, vol. 62, no. 4, pp. 1003–1012, 1977.
- [41] D. A. Mahns, N. Perkins, V. Sahai, L. Robinson, and M. Rowe, "Vibrotactile frequency discrimination in human hairy skin," *J. Neurophysiol.*, vol. 95, no. 3, pp. 1442–1450, 2006.
- [42] G. H. Mowbray and J. W. Gebhard, "Sensitivity of the skin to changes in rate of intermittent mechanical stimuli," *Science*, vol. 125, no. 3261, pp. 1297–1298, 1957.
- [43] A. K. Goble and M. Hollins, "Vibrotactile adaptation enhances frequency discrimination," *J. Acoustical Soc. Amer.*, vol. 96, no. 2, pp. 771–780, 1994.
- [44] M. Tommerdahl et al., "Human vibrotactile frequency discriminative capacity after adaptation to 25 Hz or 200 Hz stimulation," *Brain Res.*, vol. 1057, no. 1/2, pp. 1–9, 2005.
- [45] M. Mikkelsen, J. He, M. Tommerdahl, R. A. Edden, S. H. Mostofsky, and N. A. Puts, "Reproducibility of flutter-range vibrotactile detection and discrimination thresholds," *Sci. Rep.*, vol. 10, no. 1, pp. 1–14, 2020.
- [46] S. Okamoto and Y. Yamada, "Perceptual properties of vibrotactile material texture: Effects of amplitude changes and stimuli beneath detection thresholds," in *Proc. IEEE/SICE Int. Symp. System Integration*, 2010, pp. 384–389.
- [47] R. Hassen and E. Steinbach, "Vibrotactile signal compression based on sparse linear prediction and human tactile sensitivity function," in *Proc. IEEE World Haptics Conf.*, 2019, pp. 301–306.
- [48] A. Noll, B. Gülecüyüz, A. Hofmann, and E. Steinbach, "A rate-scalable perceptual wavelet-based vibrotactile codec," in *Proc. IEEE Haptics Symp.*, 2020, pp. 854–859.
- [49] S. Tsuchiya, M. Konyo, H. Yamada, T. Yamauchi, S. Okamoto, and S. Tadokoro, "Virtual active touch II: Vibrotactile representation of friction and a new approach to surface shape display," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2009, pp. 3184–3189.
- [50] N. Cao, H. Nagano, M. Konyo, S. Okamoto, and S. Tadokoro, "Envelope effect study on collision vibration perception through investigating just noticeable difference of time constant," in *Proc. IEEE World Haptics Conf.*, 2017, pp. 528–533.
- [51] N. Cao, M. Konyo, H. Nagano, and S. Tadokoro, "Dependence of the perceptual discrimination of high-frequency vibrations on the envelope and intensity of waveforms," *IEEE Access*, vol. 7, pp. 20840–20849, 2019.
- [52] R. Chaudhari, B. Çizmeci, K. J. Kuchenbecker, S. Choi, and E. Steinbach, "Low bitrate source-filter model based compression of vibrotactile texture signals in haptic teleoperation," in *Proc. 20th ACM Int. Conf. Multimedia*, 2012, pp. 409–418.
- [53] J. Allen, "Short term spectral analysis, synthesis, and modification by discrete fourier transform," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 25, no. 3, pp. 235–238, Jun. 1977.
- [54] D. Griffin and J. Lim, "Signal estimation from modified short-time fourier transform," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 2, pp. 236–243, Apr. 1984.
- [55] F. Takuya, "A study of TSP signal getting higher SN ratio at low frequency bands—removal of harmonic distortion—," *Proc. Spring Meet. Acoust. Soc. Jpn*, vol. 2000, no. 1, pp. 555–556, 2000.
- [56] M. Stresse, C. Schuwerk, A. Iepure, and E. Steinbach, "Multimodal feature-based surface material classification," *IEEE Trans. Haptics*, vol. 10, no. 2, pp. 226–239, Apr.-Jun. 2017.
- [57] R. L. Klatzky, S. J. Lederman, C. Hamilton, M. Grindley, and R. H. Swendsen, "Feeling textures through a probe: Effects of probe and surface geometry and exploratory factors," *Percep. Psychophys.*, vol. 65, no. 4, pp. 613–631, 2003.



Yutaro Toide received the B.S degree in engineering from the Department of Nanoscience, Faculty of Engineering, Chiba University, Chiba, Japan and M.S. degree from the Department of Complexity Science and Engineering, University of Tokyo, Chiba, in 2020. Since 2020, he has been working toward the Ph.D. degree with the Graduate School of Frontier Sciences, University of Tokyo. His research interests include haptics and tactile design.



Masahiro Fujiwara (Member, IEEE) received the B.S. degree in engineering and M.S. and Ph.D. degrees in information science and technology from the University of Tokyo, Tokyo, Japan, in 2010, 2012, and 2015, respectively. He is currently a Project Assistant Professor with the Graduate School of Frontier Sciences, University of Tokyo. His research interests include information physics, haptics, noncontact sensing, and application systems related to them.



Yasutoshi Makino received the Ph.D. degree in information science and technology from the University of Tokyo, Tokyo, Japan, in 2007. He is currently an Associate Professor with the Department of Complexity Science and Engineering, University of Tokyo. From 2009 to 2013, he was a Researcher for two years with the University of Tokyo and an Assistant Professor with Keio University, Tokyo. In 2013, he was with the University of Tokyo, as a Lecturer and has been an Associate Professor since 2017. His research focuses on haptic interactive systems.



Hiroyuki Shinoda (Member, IEEE) received the B.S. degree in applied physics, the M.S. degree in information physics, and Ph.D. degree in engineering from The University of Tokyo, Tokyo, Japan, in 1988, 1990, and 1995, respectively. He is currently a Professor with the Graduate School of Frontier Sciences, The University of Tokyo. Since 1995, he has been an Associate Professor with the Department of Electrical and Electronic Engineering, Tokyo University of Agriculture and Technology, Fuchu, Japan. After a period with UC Berkeley as a Visiting Scholar in 1999, he was an Associate Professor with the University of Tokyo from 2000 to 2012. His research interests include information physics, haptics, mid-air haptics, two-dimensional communication, and their applications. He is also a Member of the IEEJ, RSJ, JSME, and ACM.