# Touch Cannot Attentionally Select Signals Based on Feature Binding

Scinob Kuroki and Shin'ya Nishida

***Abstract*—For human sensory processing, cluttered real-world environments where signals from multiple objects or events overlap are challenging. A cognitive function useful in such situations is an attentional selection of one signal from others based on the difference in bound feature. For instance, one can visually select a specific orientation if it is uniquely colored. However, here we show that unlike vision, touch is very poor at feature-based signal selection. We presented two-orthogonal line segments with different vibration textures to a fingertip. Though observers were markedly sensitive to each feature, they were generally unable to identify the orientation bound with a specific texture when the segments were presented simultaneously or in rapid alternation. A similar failure was observed for a direction judgment task. These results demonstrate a general cognitive limitation of touch, highlighting its unique bias to integrate multiple signals into a global event rather than segment them into separate events.**

***Index Terms*—Attention, feature binding, frequency, orientation, motion, psychophysics.**

## I. Introduction

FEATURE-BASED signal selection is an essential function for the recognition of objects in complex real-world environments. In vision, different colored areas can be recognized as separate objects and their respective shapes and movements can be independently processed [1], [2]. Similar feature-based signal segregation is also known in hearing [3]. Touch is the modality recently attracting broad attention in research fields including virtual reality technology [4] and sensory substitution [5], but relatively little is known about its basic cognitive abilities. If touch is good at feature-based signal segregation as other sensory modalities are, one will be able to use this cognitive function to maximize the amount of information presented through touch. For example, for stimuli like Fig. 1, the human observer will be able to correctly segregate two equally salient orientations based on the difference in bound features.

Manuscript received 16 June 2023; revised 24 January 2024; accepted 17 February 2024. Date of publication 20 February 2024; date of current version 19 December 2024. The work of Shin'ya Nishida was supported by the JSPS MEXT under Grant 20H05959. This paper was recommended for publication by Associate Editor A. Moscatelli and Editor-in-Chief D. Prattichizzo upon evaluation of the reviewers' comments. *(Corresponding author: Scinob Kuroki.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the NTT Ethics Committee under Application No. R02-001.

Scinob Kuroki is with NTT Communication Science Laboratories, NTT Corporation, Atsugi 243-0198, Japan (e-mail: shinobu.kuroki@ntt.com).

Shin'ya Nishida is with NTT Communication Science Laboratories, Kyoto University, Kyoto 606-8501, Japan (e-mail: shinyanishida@mac.com).

Digital Object Identifier 10.1109/TOH.2024.3367944

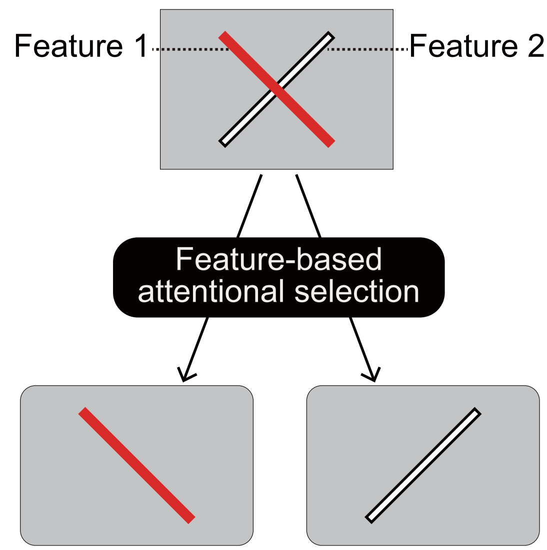


Fig. 1. Feature-based attentional selection task. An input stimulus consisting of two line segments with different feature values (red for feature 1, white for feature 2) is presented simultaneously. If human observers have the ability to attentively select one stimulus from the other based on the difference in the feature bound to the stimulus, they will be able to report the orientation of the stimulus segment with the feature they have to attend to.

The temporal frequency of the input is a fundamental feature that is predominantly processed in tactile information processing and can be a strong candidate for a feature to be attentionally selected. The tactile system has multiple mechanoreceptor-afferent channels that differ in their frequency sensitivity profiles [6], [7], [8]: SA1 channels are most sensitive to slow indentation of the skin i.e., low-frequency vibrations ($< 10$ Hz), RA channels exhibit sensitivity to middle-range vibration, and PC channels show high sensitivity at around 250 Hz. Indeed, previous studies have repeatedly shown the high performance of human participants in discriminating vibration frequencies, as well as discriminating textures based on temporal differences in skin deformation [8], [9], [10], [11], [12], [13], [14]. Human observers also have a high sensitivity to the orientation and motion direction of the input [15], [16], [17], [18]. Even peripheral or very early levels of neural activity show some selectivity to these features [19], [20], [21], [22], [23], [24].

The question here is how correctly the tactile system can bind orientation or motion direction with vibration frequency (texture), and use this binding for signal selection. To this end, we psychophysically examined whether human observers can correctly bind basic tactile features - orientation and texture, or motion direction and texture. Since the oriented patterns or moving patterns produce mechanical deformations of the skin, orientation or motion information is inherently bound with texture information given by temporal patterns of inputs.

So, the brain should be able to decode the combination of texture (vibration frequency) and orientation in the source signal, or that of texture and direction from the spatiotemporal activity pattern of a variety of mechanoreceptor channels. However, in all our psychophysical experiments, we could not find any evidence supporting such feature binding, or feature-based signal selection. Texture is the basic non-spatial feature that is most likely to be bound with shape and motion since other non-spatial features, such as temperature and pain, are not encoded by the mechanoreceptors that carry texture, shape, and motion information. It is, therefore, likely that our finding demonstrates a unique general limitation of haptics in feature binding.

Fig. 2. Stimulator latero used in the experiments. Note that in this photo, the position of the finger was shifted horizontally to show the actuator pins, but in the actual experiment, the finger was placed in a position that completely covered the surface of the actuator. The participant's view of the moving parts of the stimulator was blocked by a black cover box and their own fingers, and the stimuli were not visible.

## II. General Methods

We ran two main series of experiments, one using an orientation judgment task (Exps. 1a, b, c, d, e), and the other using a direction judgment task (Exps. 2a, b, c, d, e).

We used a simple task to evaluate feature-based signal selection in touch. It is different from visual tasks normally used to examine feature-based selection or feature binding (e.g., visual search of a target defined by a conjunction of two features [1]). It may be rather similar to grouping/segregation of pattern elements by the Gestalt rule of similarity, or Ishihara color blindness test [25].

### A. Participants

A total of 37 participants (11 females; aged 20–37 years, mean age 22.6±2.91; all but two right-handed) with normal tactile sensitivity (based on their self-report) participated in the experiments. Fifteen participants took part in Experiments 1a, 1b, and 1c. Different groups of 15 participants took part in Experiments 2a, 2b, and 2c, Experiments 1d and 1e, and Experiments 2d and 2e. The number of participants was pre-determined to be comparable to previous studies [11], [23], [26], [27], [28] in which tactile perception was tested by stimulating the fingertips with similar stimuli. There were partial overlaps of participants across experiments. All participants gave their written informed consent to participate. The NTT Ethics Committee approved the recruitment of the participants and the experimental procedures, which were conducted in accordance with the ethical standards that have their origin in the Declaration of Helsinki (2008). (Approval number: R02-001)

### B. Apparatus and Stimuli

The stimuli were generated by the 'Latero' (Tactile Lab, Montreal, Canada) [29]. An array of 8 x 8 = 64 pins constructs 10 x 10 mm contact surfaces with ~1.2 mm pin-to-pin spacing. Each pin can independently move across the contacted surface of the skin laterally. The refresh rate of the stimulator is 1100 Hz, and it is made to be able to control the position of the pins in real-time. In a separate session from the experiment, we confirmed that the vibration stopped immediately by measuring the displacement of the actuator pin of the stimulator and the displacement of the skin of a finger in contact directly next to the pin with a laser displacement sensor (KEYENCE LC-2440).

For the orientation judgment task (used in Exps. 1a, b, c, d, e), a pattern consisting of one or two diagonal line segments was presented; for the direction judgment task (used in Exps. 2a, b, c, d, e), one or two moving rectangle segments were presented (stimulus details are provided in the Methods section of each experiment). Each segment was presented by using the eight pins sinusoidally vibrated at the same frequency, but in opposite phases between odd and even rows of pins. Driving neighboring pins in opposite phases is expected to increase the skin strain and thus the perceived intensity [26], [30]. The vibration frequency for each segment was 20, 30, 40, 80, or 160 Hz, depending on stimulus conditions. These frequency ranges were selected on the basis of preliminary observations to include a wide range within which the dominant mechanoreceptor channels change [8], [9]. Our preliminary experiments also confirmed that the selected frequencies are suitable for orientation and motion presentation under the conditions of the current experiments. The amplitude of the vibration was adjusted (e.g., the amplitude of a 160 Hz vibration was set much lower than that of a 20 Hz vibration) to match the perceived intensity among different vibration frequencies based on preliminary experiments with several participants.

### C. Procedure

A participant sat at a table with the index or middle finger of the left hand placed on the stimulator (Fig. 2). They performed experiments with their eyes open to maintain their arousal level, but they could not see the vibration of the stimulator. They wore earplugs to mask any subtle sound made by the tactile stimulation.

Total time of the experiment, including instruction and breaks, varied from experiment to experiment, but was always longer than one hour. Each experiment was divided into blocks, with breaks between sessions, each containing multiple blocks. Each block lasted no more than 5 minutes, and no single session exceeded 20 minutes. Participants were instructed to swap fingers after each block to prevent finger fatigue and adaptation. The blocks in which participants reported that they were unable to detect (not unable to discriminate) the vibration, such as poor finger placement, were excluded from later analyses. No feedback signal was provided in all experiments.

### D. Data Analysis

The correct answer rate was calculated for each participant and each condition (combination of distractor frequency and pattern duration). Then the average and 95% confidence intervals for all participants were calculated based on bootstrapping. We analyzed all individual data for perceived orientation/direction (whether the reported orientation/direction was correct or wrong for each trial under each stimulus condition) using the binomial generalized linear mixed model (GLMM) [31]. We first made a statistical inference using a maximum model (model with all possible fixed and random effects of time interval, distractor frequency, pattern duration, participant, and their combination) and then conducted automated model selection with backward elimination. Based on the selected model, we implemented a full-factorial analysis of variance (ANOVA) F test for fixed effects (Wald $\chi 2$ tests). Analyses were performed using R with the lme4 and buildmer packages. The 95% confidence intervals (CIs) of the chance level of correct rate (0.5), calculated using the binomial distribution with a sample size of 300 trials (20 trials for each of the 15 participants), range from 0.44 to 0.56. These intervals are visually represented as gray areas in the graphs. In order to evaluate the extent to which the obtained results provide positive support for alternative or null hypotheses ($p = 0.5$), we calculated Bayes factors. For example, the Bayesian statistical analysis shows that if the data has correct rate at upper limit of 95% CI of chance ($p = 0.56$), it was more likely to have occurred under the null hypothesis than the alternative hypothesis ($BF_{10} = 0.67$). If the data shows above 0.6, the analysis supported the alternative hypothesis ($BF_{10} > 10$). The statistical tests were performed by JASP (JASP Team, 2024).

## III. Experiment 1: Texture-Based Selection of Orientation

The first series of experiments was conducted to examine whether the target orientation can be detected by touch based on the temporal texture (vibration frequency) tagged to the orientation. The main condition was tested as Exp 1a, where the target and distractor orientations were simultaneously presented and participants were asked to report the orientation of the target frequency.

### A. Experiment 1a: Spatial Binding

*1) Method:* A presented stimulus was an X pattern consisting of two diagonal line segments, a 45-deg (right-upward) target and -45-deg (left-upward) distractor, or -45-deg target and 45-deg distractor (Fig. 3(a)). Each line segment, ∼14 mm in length, was produced by eight pins. The target pattern always vibrated at 20 Hz, while the distractor pattern vibrated at one of the higher frequencies (30, 40, 80, or 160 Hz).

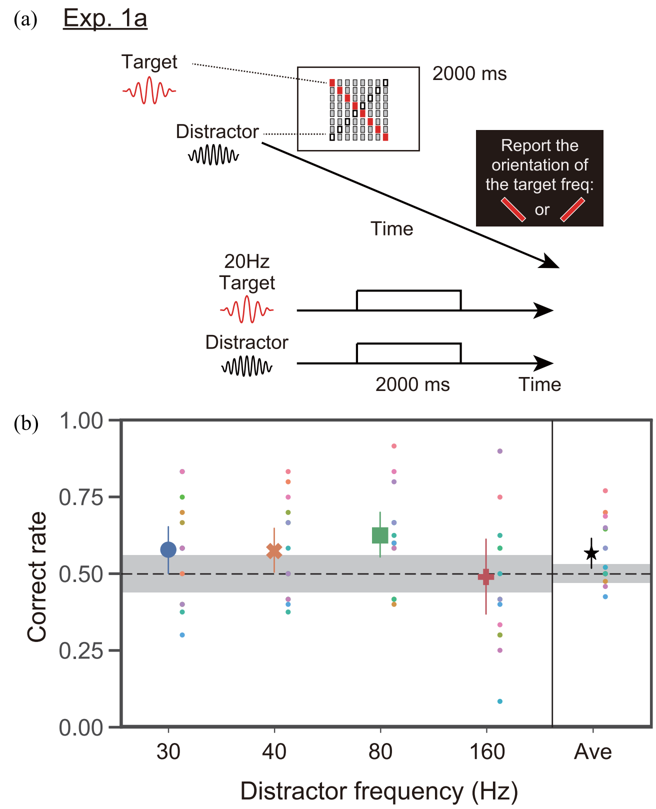


Fig. 3. Main condition of texture-based orientation selection (Exp. 1a). (a) The target and distractor orientations were simultaneously presented and participants were asked to report the orientation of the target (20 Hz) frequency. (b) The mean and 95% CIs of correct rates of 2-AFC task for 15 participants as a function of the distractor frequency. The rightmost graph shows the correct rates averaged over four distractor frequencies. Gray areas in the graph represent uncorrected 95% CIs of the chance performance for group data.

In each trial, an X pattern was presented for 2000 ms. The participant made a two-alternative forced choice (2-AFC) task, in which they indicated whether the 20-Hz target was paired with the 45-deg or -45-deg tilt by pressing the 'F' or 'J' keyboard keys with their right hand. After a 2-second blank period, the next trial started.

One block consisted of 32 trials (two target orientations x four distractor frequencies x two repetitions, in random order). Each participant completed five blocks, repeating 20 trials for each frequency condition. The total time of the experiment, including instruction and breaks, was approximately one hour. Before participating in experiment 1a, to become familiar with the experimental stimuli, participants took part in an approximately one-hour practice session that included an orientation judgment task and a frequency discrimination task.

*2) Results:* The participants were asked to judge whether the coarse-texture (20 Hz) target had a tilt of 45 deg or -45 deg. Since the two diagonal lines have sufficiently separated textures for the tactile system [6], [10], they would be carried, at least partially, by different mechanoreceptor-afferent channels. We, therefore, expected that, in principle, these two segments are easily separable by feature-based selection mediated by central processing.

However, the result showed that the feature-based signal selection was very difficult (Fig. 3(b)) when the target and the distractor stimuli were presented simultaneously: participants were not able to correctly report the orientation of the target stimulus. The average correct rates were within, or only slightly higher than, 95% CI of the 50% chance level (gray areas in Fig. 3(b)), and the Bayesian statistical analysis does not strongly support an alternative hypothesis that the probability was not 50% ($B_{10} = 0.08$-$1.00$) except for a 80 Hz condition. We cannot say that the task was impossible, but nearly so.

In order to ascribe the low performance to the difficulty of feature-based signal selection, we need to show that neither judgments on orientation nor those on frequency were very difficult with our device. The following control experiments tested this.

### B. *Experiment 1b: Orientation Judgment (Control)*

*1) Method:* One diagonal line pattern was presented and the participant had to indicate whether the stimulus was the 45-deg or -45-deg tilt (Fig. 4(a)). The frequency of each pattern was 20, 30, 40, 80, or 160 Hz. The duration of each pattern was one of the following: 62.5, 125, 250, 500, or 1000 ms. The orientation, frequency, and duration of each stimulus were randomized across trials.

One block consisted of 50 trials (two stimulus orientations x five frequencies x five pattern durations, in random order). Each participant completed ten blocks, running 20 trials for each combination of frequency and duration. The total time of the experiment, including teaching and breaks, was approximately two hours.

The same participants as Exp. 1a participated in Exps 1b and the following 1c.

*2) Results:* As shown in Fig. 4(a), when only one line segment was presented, the same participants as Exp 1a could report the orientation of the target (20 Hz) or distractor stimuli (30–160 Hz), and the Bayesian statistical analysis supports alternative hypothesis ($B_{10}>100$). The accuracy dropped for highest vibration frequency, but still above 75% when the exposure duration was long enough. Full-factorial ANOVA with the GLMM model demonstrated the statistical significance of all main effects ($\chi2(4) = 82$, $p<0.0001$ for frequency; $\chi2(4) = 80$, $p<0.0001$ for duration: Wald chi-square test), while interaction was not significant ($\chi2(16) = 27$, $p = 0.05$).

### C. *Experiment 1c: Texture Judgment (Control)*

*1) Method:* Two diagonal line patterns with different vibration frequencies were presented at 1000-ms intervals (Fig. 4(b)). The participant had to indicate whether the target stimulus (lower frequency: 20 Hz) was presented first or second. The rest of the procedure was the same as in Exp. 1b.

One block consisted of 40 trials (two target orientations x four distractor frequencies x five pattern durations, in random order). Each participant completed ten blocks, repeating 20 trials for each combination of frequency and duration. The total time of the experiment, including teaching and breaks, was approximately two hours.

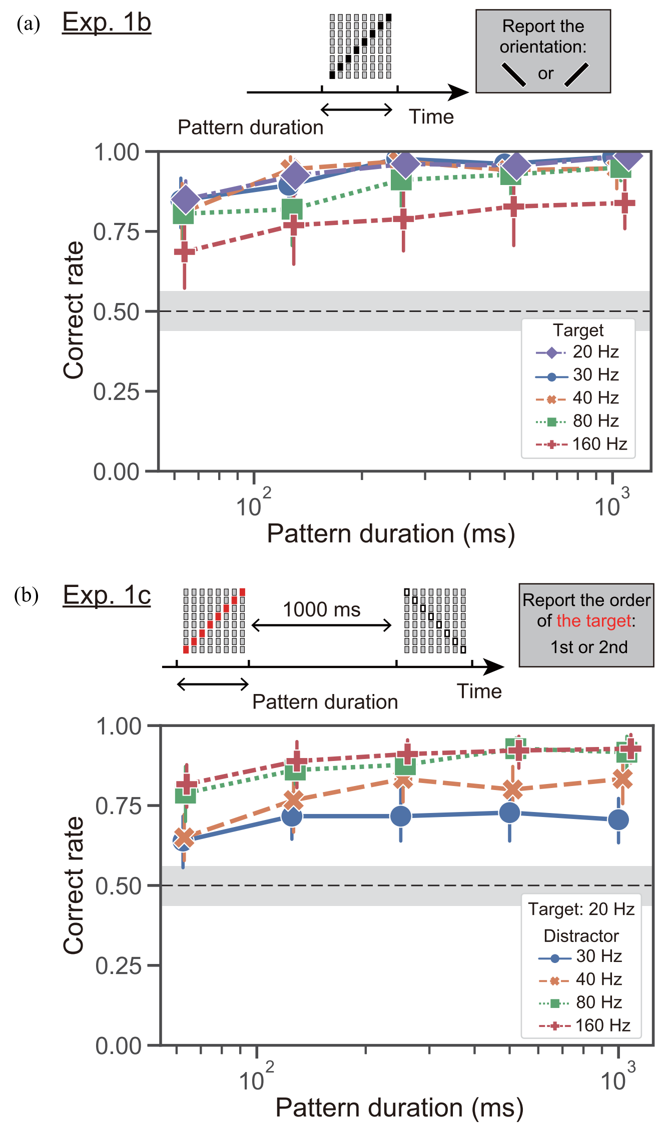


Fig. 4. Control conditions for Exp 1a. (a) Exp 1b. Only one stimulus pattern was presented and participants were asked to report the orientation. (b) Exp 1c. The target and distractor pattern were sequentially presented with a 1-sec interval and participants were asked to report whether the target stimulus was presented 1st or 2nd.

*2) Results:* The results showed that the correct rate was higher than 95% CI of the 50% chance level, and the Bayesian statistical analysis supports alternative hypothesis ($B_{10}>100$). The same participants as Exp 1a could reliably select the target from the distractor based on the temporal texture when the target and distractor were sequentially presented with an interval (Exp. 1c, Fig. 4(b)). Again, main effects were significant ($\chi2(4) = 47$, $p<0.0001$ for frequency; $\chi2(4) = 40$, $p<0.0001$ for duration), while interaction was not ($\chi2(16) = 16$, $p = 0.2$). Reasonably, the larger the frequency difference between the target and the distractor and the longer the pattern duration, the easier feature-based signal selection became.

Results of the two control experiments demonstrate that the low performance found for the frequency-based orientation selection task could not be ascribed to the difficulties in judging orientation and frequency *per se*. For example, when a 20-Hz target and an 80-Hz distractor are presented for 2000 ms (Exp. 1a), the expected error due to orientation (Exp. 1b) and texture (Exp. 1c) misjudgment would be 90%, while the actual performance was around 60%. Therefore, the results of Exp. 1a could be ascribed to the failure of the feature-based signal selection or feature binding.

### *D. Experiment 1d: Temporal Binding*

Exp 1a indicates that human participants were nearly unable to haptically separate the target orientation vibrated at a lower frequency from the distractor vibrated at a higher frequency when they were presented simultaneously. On the other hand, Exp 1c shows that the participants could identify the target orientation when a pair of the target and distractor orientations were sequentially presented with a 1000-ms gap. In the next experiment, we tested whether the participants could perform the task when the target and distractor orientations were alternatively presented with no temporal gap. We changed the duration of each orientation, which covaried with the rate of alternation. Our questions were whether the failure of feature-based signal selection was observed not only when the target and distractor were presented simultaneously but also when presented separately, and if so, how much the target and distractors should be temporally separated to be segregated by touch. This task was similar to a task to measure the temporal resolution limit of feature binding in vision research (e.g., [2]), although spatial binding cues were not excluded in our case. The main condition was tested as Exp 1d. Exp 1e is a control experiment to evaluate the effect of stimulus duration.

*1) Method:* Two diagonal line patterns with different vibration frequencies were presented alternately with no intervals (Fig. 5(a)) and the participant had to indicate whether the coarse-textured target was paired with the 45-deg or -45-deg tilt. The total duration of the stimulus was four seconds so that the stimuli were equivalent to experiment 1a in terms of presented energy. The duration of each pattern was one of the following: 62.5, 125, 250, 500, or 1000 ms. Thus the same pattern was presented at least twice throughout the stimulus, even with the longest pattern duration. The rest of the procedure was the same as in Exp. 1a.

One block consisted of 40 trials (two target orientations x four distractor frequencies x five pattern durations, in random order). Each participant completed ten blocks, running 20 trials for each combination of frequency and duration conditions. The total time of the experiment, including teaching and breaks, was approximately two hours.

*2) Results:* The results (Fig. 5) showed that correct rates were within, or only slightly higher than, 95% CI of the 50% chance level when the pattern duration was as shorter than 250 ms, and the Bayesian statistical analysis does not strongly support an alternative hypothesis that the probability was not 50% ($B_{10} = 0.07$-1.0). That is, the vibration-based orientation selection was almost impossible not only when they were presented simultaneously, but also when they were presented closely in time: i.e., when the pattern duration was 250 ms or shorter, or the alternation frequency was 2 Hz or higher.

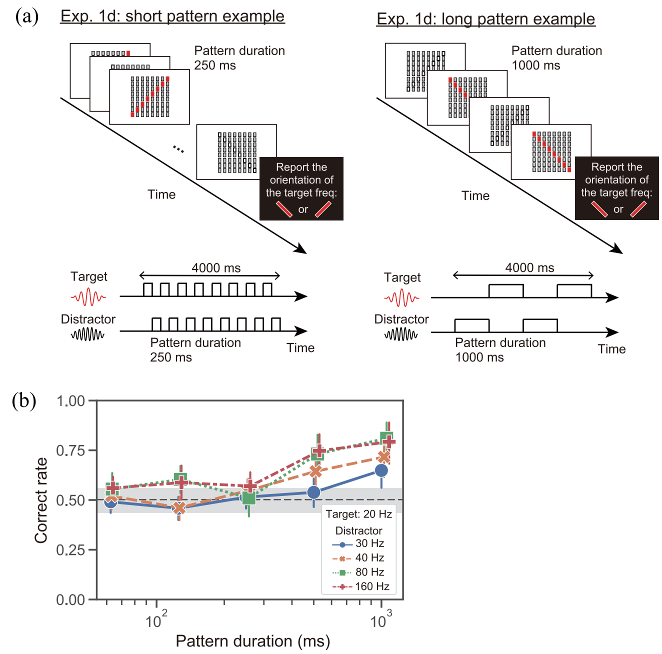


Fig. 5. Target orientation selection based on the tagged texture is difficult even when the target and distractor orientations were separated in time (Exp 1d). (a) The target and distractor were presented alternately with no temporal intervals. (b) The mean and 95% CI of the correct response rate averaged over 15 participants. Orientation discrimination was almost impossible when the pattern duration was 250 ms or shorter. Note that the symbol location was horizontally jittered slightly to improve visibility.

The next experiment checked whether the task became difficult simply because each pattern duration was too short.

### *E. Experiment 1e: Duration Effect (Control)*

Two diagonal line patterns with different vibration frequencies were presented with 1000 ms intervals (Fig. 6(a)) and the participant had to indicate whether the coarse-textured target was paired with the 45-deg or -45-deg tilt. The presented stimuli were the same as in Exp. 1c and the rest of the procedure was the same as in Exp. 1d.

One block consisted of 40 trials (two target orientations x four distractor frequencies x five pattern duration). Each participant completed ten blocks, repeating 20 trials for each combination of frequency and duration conditions. The total time of the experiment, including teaching and breaks, was approximately two hours. Note that those who participated in Exp. 1e also participated in Exp. 1d.

*1) Results:* The task performance (Fig. 6) was significantly improved in comparison with Exp. 1d ($\chi 2(1) = 33$, $p<0.0001$). Even with short durations, results showed higher than 95% CI of the 50% chance level and the Bayesian statistical analysis supports alternative hypothesis. That is, the participants could use vibration frequency as a useful cue to discriminate the target orientation from the distractor orientation. In addition, other effects were also significant ($\chi 2(4) = 80$, $p<0.0001$ for duration;

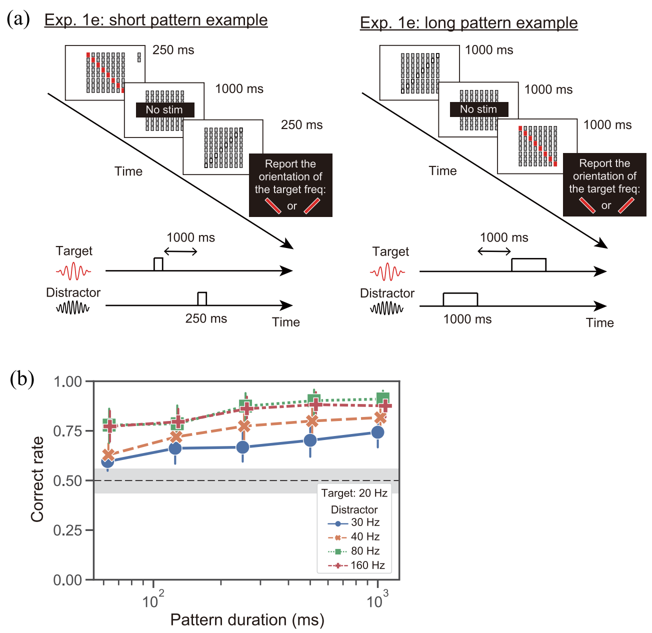


Fig. 6. Control experiment (Exp 1e) showed that short pattern duration per se was not the reason for low performance. The orientation of the target pattern could be reported even when the target and distractor were presented briefly but with long temporal intervals.

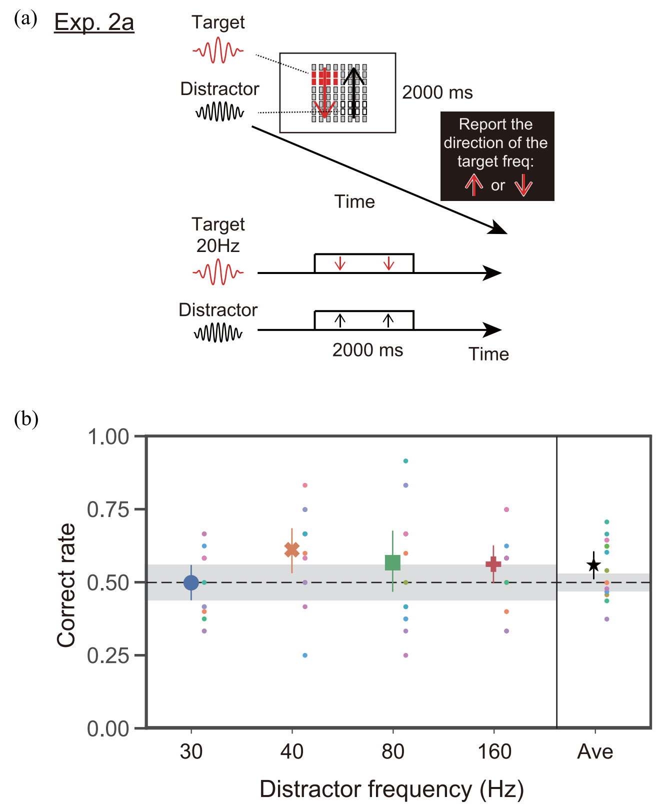


Fig. 7. Target direction selection based on the tagged texture (Exp 2a). The target and distractor were presented simultaneously, and participants judged whether the target was moving upward or downward as 2-AFC. The vibrating rectangles moved from one end to the other in 1000 ms and the movement was repeated twice, thus making a movement in the same direction for 2000 ms.

$\chi2(3) = 87$, $p<0.0001$ for frequency; $\chi2(3) = 24$, $p<0.0001$ for task and frequency; $\chi2(12) = 23$, $p<0.05$ for duration and frequency; $\chi2(4) = 21$, $p<0.001$ for task and duration) except for the interaction between task and duration and frequency; $\chi2(12) = 16$, $p = 0.20$).

The results of Exp 1d, therefore, reflect the effect of temporal interaction of the target and distractor orientations, rather than the effect of pattern duration. They suggest that the participants could use the temporal coincidence of orientation and frequency to find the target only when no distractor was presented within a 250-ms temporal window.

## IV. Experiment 2: Texture-Based Selection of Direction

To test the generality of our findings, we also conducted a similar experiment for motion direction judgment by having two rectangles with different temporal textures (vibrating at different frequencies) simultaneously move in opposite directions on the fingertip.

### A. Method

*1) Stimuli:* A pattern consisting of one or both of two rectangles moving from distal to proximal or proximal to distal was presented. When there were two rectangles, they had different vibration frequencies (target and distractor) and moved in opposite directions simultaneously or sequentially. One rectangle pattern consisted of eight vibrating pins and was approximately 0.5 x 0.2 cm in size. The rectangles were moved vertically from one side of the stimulator to the other over the specified pattern duration, returned to the initial position, and the same movements were repeated until the total stimulus duration was reached. The movement speed was reciprocally determined by the pattern duration. The stimulus presentation parameters were otherwise the same as the orientation judgment.

*2) Procedure:*

*a) Experiment 2a: spatial binding:* Two oppositely moving rectangles with different vibration frequencies were presented simultaneously (Fig. 7(a)). The vibrating rectangles moved from one end to the other in 1000 ms and the movement was repeated twice, thus the duration of making a movement in the same direction was 2000 ms. This experiment was also a 2-AFC task. The participant was to indicate whether the target was paired with a proximal or distal movement by pressing the Y or B keyboard keys. After a 2-second blank period, the next trial started. The rest of the procedures, including the number of conditions and blocks and the total time of the experiments, were the same as in Exp. 1a.

*b) Experiment 2b: direction judgment (control):* One rectangle that moved in one direction was presented and the participant was to indicate whether the stimulus moved to proximal or distal (Fig. 8(a)). The frequency of each pattern was 20, 30, 40, 80, or 160 Hz. The duration of each pattern was one of the following: 62.5, 125, 250, 500, or 1000 ms. Note that in Exp. 2b, the pattern moved from one side of the stimulator to the other according to the pattern duration, so the speed of

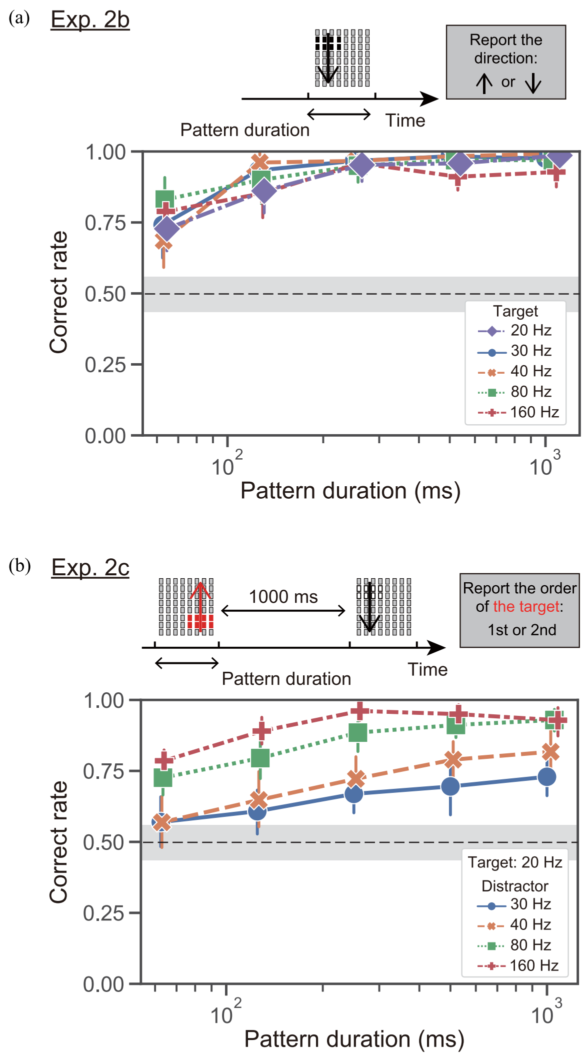


Fig. 8. Control conditions for a sanity check for texture-based direction selection. The same participants to Exp 1b could report the direction of the stimuli (Exp 2b, panel a) and select the target from the distractor when the distractor frequency was far separated from the target frequency (Exp 2c, panel b).

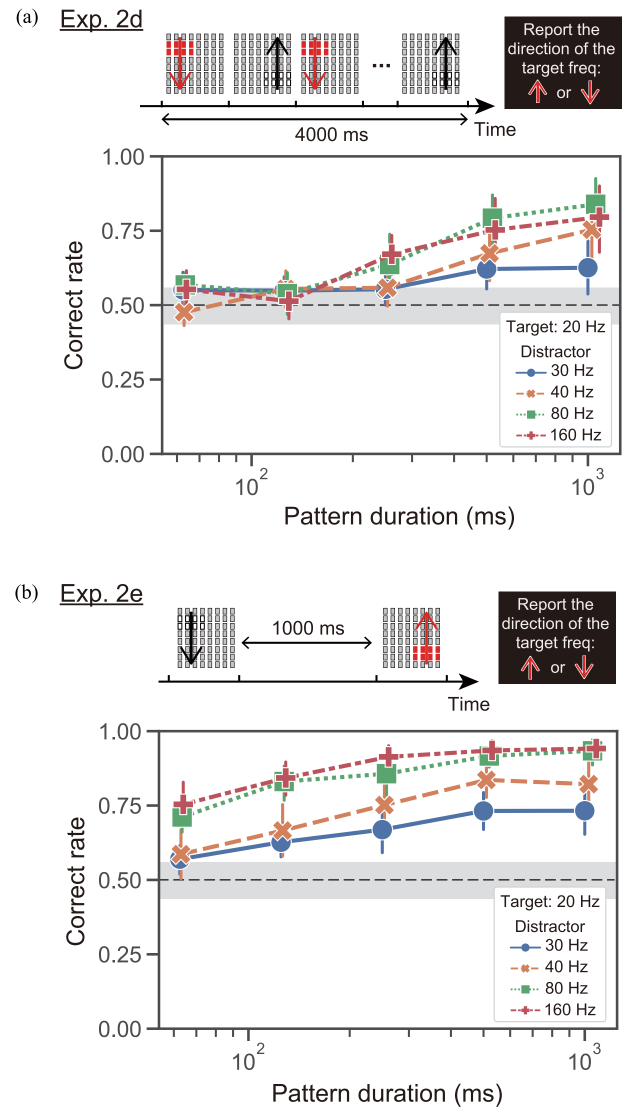


Fig. 9. Target direction selection based on the tagged texture is difficult even when the target and distractor motions were separated in time (Exp 2d, panel a). A control experiment (Exp 2e, panel b) showed that short pattern duration per se was not the reason for low performance. The results were similar to those obtained with the orientation task (Figs. 5, 6).

movement varied according to the difference in pattern duration, the shorter the faster. The direction, frequency, and duration of each stimulus were randomized across trials. The rest of the procedures were the same as in Exp. 1b.

*c) Experiment 2c: texture judgment (control):* Two rectangles with different vibration frequencies that moved in opposite directions were presented at 1000-ms intervals (Fig. 8(b)). The participant was to indicate whether the target stimulus was presented first or second. The rest of the procedure was the same as in Exp. 1c. Note that those who participated in Exp. 2c also participated in Exps. 2a and 2b.

*d) Experiment 2d: temporal binding:* Two rectangles with different vibration frequencies that moved in opposite directions were presented alternately (Fig. 9(a)). The total duration of the stimulus was four seconds so that the stimuli were equivalent to experiment 1 in terms of presented energy. The duration of each pattern was one of the following: 62.5, 125, 250, 500, or 1000 ms. Thus the same pattern was presented at least twice throughout the stimulus, even with the longest pattern duration. The rest of the procedure was the same as in Exp. 1d.

*e) Experiment 2e: duration effect (control):* In Exp. 2e, the presented stimuli were the same as in Exp. 2c and the task was the same as in Exps. 2a and 2d (Fig. 9(b)). Note that those who participated in Exp. 2d also participated in Exp. 2e.

## *B. Results*

As we found in Exp 1, texture-based signal selection was also difficult for a direction judgment: the two motions with different textures were not perceptually separable (Fig. 7, $B_{10} = 0.07$-$1.00$ except for a 40 Hz condition; see also control conditions in Fig. 8). When alternating the target and distractor motions (Fig. 9(a)), we observed a temporal limit similar to that found in the orientation task, albeit slightly improved. Importantly, task performance was greatly improved when the target and distractor were presented with sufficiently long intervals (Fig. 9(b); $\chi2(1) = 74$, $p<0.0001$).

## V. Discussion

### A. Feature-Based Signal Selection

Binding multiple features belonging to the same objects or events is an essential perceptual function for the brain to make rich perceptual representations of the real world [e.g., 1], which, however, has been overlooked in haptic research. The tactile system is able to sense many features [6], [7], [8], [15], [16], [17], [18], [19] and is known to have a high sensitivity to each of them [9], [10], [11], [12], [13], [14]. This study investigated how correctly the system can bind them. Specifically, here we have tested texture and orientation/motion binding. Note that vibration is not necessarily defined as an artificial signal; common real-world stimuli, including pressure, indentation, and other daily activities, are also a subset of vibrations and can be described as a combination of them. Although the brain can decode the combination of texture and orientation, or that of texture and direction, of the source signal from the spatiotemporal activity pattern of a variety of mechanoreceptor channels, we found no evidence supporting such feature-based signal selection, and feature binding throughout our psychophysical experiments. There are also tactile features other than vibration, such as temperature and pain, but they have much lower spatial resolution and are not encoded by activation of mechanoreceptors [32], [33], so they are more likely to be bound with shape/motion information later than vibration frequency.

Note also that our tasks allowed the participants to use the skin locations (e.g., upper-right corner of the fingertip), in addition to the stimulus orientation or direction, to find the target stimulus tagged with the lower vibration frequency. In vision research, binding problem has been investigated using paradigms such as feature-based selection (e.g., conjunction search [1]), where it is known that the correct combination of features can be reported by using attention. In the present study, the paradigm was simplified and only two patterns were presented. As shown in Fig. 1, this is too simple for a visual task. In other words, it was a situation where a simpler tactile task was performed while attention was directed. Nevertheless, the participants could not perform the task. This further supports the general difficulty of feature-based stimulus selection in touch.

One may suspect that the failure in detecting the target orientation/direction based on feature-based signal selection (Exp. 1a, 2a) might be caused by some form of physical interaction between the two vibration signals at the fingertip. However, with our apparatus, the spread of vibration was suppressed by stationary pins around the vibrating pins that acted as fixed ends. Even if the vibrations were a bit spread out, the patterns of different vibration frequencies must be spatially separated to some extent, allowing the participants to correctly solve the task by, say, only monitoring the input to a specific skin site such as line endpoints. Furthermore, we also found that the correct feature pair could not be detected not only when the two segments were presented simultaneously, but also when they were presented alternately (Exp. 1d, 2d, see discussion below). Non-simultaneous presentations should significantly reduce the physical interaction on the fingertip, at least for most of the pattern duration. Therefore, we conclude that the tactile system lacks the ability to select specific tactile inputs by directing attention to a specific vibration frequency, even though vibration information is embedded in the peripheral tactile neural signals.

### B. Feature-Binding Temporal Limitation

Feature-based stimulus selection is closely related to the binding problem, i.e., how to bind correct feature pairs belonging to the same signal source. Previous related studies in vision research have revealed that features belonging to an object can be perceptually bound if the attention is directed to the object's spatial location [1]. In our experiments, the participants could pay attention to their fingertip to which the stimulus was presented, but could not select the target, even with the combinations of target and distractor that were different in frequency enough to presumably tap different mechanoreceptor-afferent channels [6], [8], [10], where vibrotactile masking effect is expectedly minor [34].

In experiment 1d and 2d, we investigated whether and how quickly the tactile system can bind features when two patterns with different feature combinations are alternatively presented. The duration of each pattern presentation was changed, and participants were asked to report the orientation (Fig. 5) or the direction (Fig. 9(a)) of the pattern paired with the target (low-frequency) vibration. It was found to be difficult to correctly bind the orientation/direction and texture when the pattern was presented for 250 ms or shorter, which further supports the difficulty of feature-based signal selection in touch.

The results showed that the participants could distinguish the target from the distractor when the presentation duration is long or the alternation rate is slow (Figs. 5, 9). The reduction of the effect of distractor is consistent with the previous report on the effect of temporal factors on the interference in tactile pattern perception between the target and distractor of similar textures [35], [36]. It is likely that at slow alternations, the participants could judge vibration frequency and orientation separately and cognitively associated them based on temporal coincidence, rather than perceptually binding them into a single object based on spatial coincidence. This is because the observed 2 Hz limit is comparable to the limit of cross-modal binding, which is thought to be mediated by a general-purpose attentional mechanism [37]. It should be noted that the temporal limit can be much faster for low-level binding, such as purely temporal binding of color and orientation (without using spatial cues) in vision is higher than 10 Hz [2]. Our results do now support similar fast feature binding in touch.

Finally, we found surprisingly similar response patterns for texture-orientation and texture-direction pairs. This suggests that this temporal limit is not attribute-specific, but rather a general limitation of tactile feature binding. In summary, we cannot demonstrate the ability of touch to bind features of the same source by spatial attention under cluttered conditions, nor find evidence of early rapid feature binding in touch. Instead, a central general-purpose process seems to be responsible for tactile feature binding.

It should be noted that the present findings are not necessarily inconsistent with electrophysiological findings on tactile feature encoding. It is known that even peripheral or very early levels of neural activity show some orientation selectivity for the input signal [15], [19], [20] while the higher somatosensory cortex (SII) shows complex curvature tuning [38], suggesting a hierarchical organization similar to visual shape coding. Similarly, motion information is present in early-level neural signals and is processed predominantly at later stages [15], [18], [39]. While it seems possible that shape or direction could be encoded independently for each frequency-specific channel (i.e., frequency-labelled), the central nervous system processes this information hierarchically and higher areas receiving bilateral and multi-channel broadband frequency inputs [40], [41]. Our results indicate that attentional selection operates primarily at these higher processing levels.

### C. Limitation

In this study, as in the most tactile previous studies, there are limitations due to the use of the particular stimulator. One clear limitation is the frequency range of the texture. It was not possible to test using pressure or indentation stimuli, which are common stimuli in the real world, encoded by mechanoreceptors and expressed as very low frequency vibrations. Although the piezoelectric device employed is one of the best options for presenting different patterns of different frequencies at the same perceived intensity (i.e., different amplitudes), the human sensitivity to very low frequency vibrations (pressure) is much lower than the sensitivity to medium and high frequency vibrations, so presenting pronounced pattern lower than 20 Hz was difficult to achieve.

In addition, experiments in this study were conducted in a passive touch situation with a stimulator to ensure precise temporal control of the tactile patterns. It would be interesting, albeit challenging, to see the extent to which human observers can bind tactile features in other situations such as when they move their hands to touch multiple objects.

### D. Forced Integration of Haptic Signals

It is worth mentioning that previous tactile behavior studies have also reported forced integration of tactile signals at different skin locations, shown in the form of assimilation effects, in various feature dimensions, such as texture [42], [43], frequency [43], [44], spatial pattern [11], [35], [36], motion [45], [46] and temperature [47], [48] perception. For example, we previously reported that the perceived frequencies are integrated even when two vibrations of different frequencies are presented at different skin locations [43], which is consistent with the present findings in the sense that vibration frequencies cannot be bound to a specific skin location. In most previous studies, the presented multiple stimuli tapped the same channel and were therefore subject to masking effects, but the reported temporal limit of signal segregation was shown to be consistent with the current results [45], [46]. The current finding, together with these findings, appears to indicate that when detecting multiple signals simultaneously at different skin locations, the tactile system has a strong bias to integrate them into a global single event rather than segment them into separate events.

This scientific finding has implications for the design of future sensory substitution and tactile communication systems. As tactile systems are surprisingly sensitive to a wide variety of features, state-of-the-art approaches in robotics and tactile communication systems have proposed ways to combine and present multiple features in order to make the most of their richness [49], [50], [51], [52], [53]. However, little attention has been paid to the limitations of feature binding. Just because humans can recognize information about some features independently and accurately does not necessarily mean that the information content remains intact when recognizing combinations of them derived from multiple events. This is an issue awaiting further investigation.

## VI. Conclusion

In summary, this study provides a line of new evidence of forced feature integration in touch, highlighting the possibility that the tactile system does not directly utilize the information of feature information based on the somatotopic coincidence in the periphery or the combination of mechanoreceptors and features, but rather encodes each feature individually and performs cognitive tactile feature binding based on the temporal coincidence, i.e., whether the features are presented simultaneously at the fingertip. This finding further supports the notion that the tactile system differs from other modalities not only in resolution but also in the underlying fundamental computations.

## References


[1] A. M. Treisman and G. Gelade, "A feature-integration theory of attention," *Cogn. Psychol.*, vol. 12, no. 1, pp. 97–136, Jan. 1980.

[2] A. O. Holcombe and P. Cavanagh, "Early binding of feature pairs for visual perception," *Nature Neurosci.*, vol. 4, no. 2, pp. 127–128, Feb. 2001.

[3] S. L. McCabe and M. J. Denham, "A model of auditory streaming," *J. Acoust. Soc. Am.*, vol. 101, no. 3, pp. 1611–1621, Mar. 1997.

[4] D. Wang, K. Ohnishi, and W. Xu, "Multimodal haptic display for virtual reality: A survey," *IEEE Trans. Ind. Electron.*, vol. 67, no. 1, pp. 610–623, Jan. 2020.

[5] L. A. Jones and N. B. Sarter, "Tactile displays: Guidance for their design and application," *Hum. Factors*, vol. 50, no. 1, pp. 90–111, Feb. 2008.

[6] S. J. Bolanowski, G. A. Gescheider, R. T. Verrillo, and C. M. Checkosky, "Four channels mediate the mechanical aspects of touch," *J. Acoust. Soc. Am.*, vol. 84, no. 5, pp. 1680–1694, 1988.

[7] G. 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.

[8] R. H. LaMotte and V. B. Mountcastle, "Capacities of humans and monkeys to discriminate vibratory stimuli of different frequency and amplitude: A correlation between neural events and psychological measurements," *J. Neurophysiol.*, vol. 38, no. 3, pp. 539–559, May 1975.

[9] C. J. Cascio and K. Sathian, "Temporal cues contribute to tactile perception of roughness," *J. Neurosci.*, vol. 21, no. 14, pp. 5289–5296, Jul. 2001.

[10] S. Kuroki, J. Watanabe, and S. Nishida, "Contribution of within- and cross-channel information to vibrotactile frequency discrimination," *Brain Res.*, vol. 1529, pp. 46–55, Sep. 2013.

[11] S. Kuroki, M. Sawayama, and S. Nishida, "The roles of lower- and higher-order surface statistics in tactile texture perception," *J. Neurophysiol.*, vol. 126, no. 1, pp. 95–111, Jul. 2021.

[12] V. B. Mountcastle, M. A. Steinmetz, and R. Romo, "Frequency discrimination in the sense of flutter: Psychophysical measurements correlated with postcentral events in behaving monkeys," *J. Neurosci.*, vol. 10, no. 9, pp. 3032–3044, 1990.
[13] L. Skedung, M. Arvidsson, J. Y. Chung, C. M. Stafford, B. Berglund, and M. W. Rutland, "Feeling small: Exploring the tactile perception limits," *Sci. Rep.*, vol. 3, 2013, Art. no. 2617.
[14] A. I. Weber et al., "Spatial and temporal codes mediate the tactile perception of natural textures," *Proc. Natl. Acad. Sci. USA*, vol. 110, no. 42, pp. 17107–17112, Oct. 2013.
[15] J. J. DiCarlo and K. O. Johnson, "Spatial and temporal structure of receptive fields in primate somatosensory area 3b: Effects of stimulus scanning direction and orientation," *J. Neurosci.*, vol. 20, no. 1, pp. 495–510, 2000.
[16] E. Jarocka, J. A. Pruszynski, and R. S. Johansson, "Human touch receptors are sensitive to spatial details on the scale of single fingerprint ridges," *J. Neurosci.*, vol. 41, no. 16, pp. 3622–3634, Apr. 2021.
[17] K. Johnson and T. Yoshioka, "Neural mechanisms of tactile form and texture perception," *Somatosensory Syst.*, CRC Press, pp. 73–101. 2001, doi: 10.1201/9781420038705.ch3.
[18] Y. C. Pei, S. S. Hsiao, J. C. Craig, and S. J. Bensmaia, "Shape invariant coding of motion direction in somatosensory cortex," *PLoS Biol.*, vol. 8, no. 2, 2010, Art. no. e1000305, doi: 10.1371/journal.pbio.1000305.
[19] J. A. Pruszynski and R. S. Johansson, "Edge-orientation processing in first-order tactile neurons," *Nature Neurosci.*, vol. 17, no. 10, pp. 1404–1409, 2014.
[20] A. K. Suresh, H. P. Saal, and S. J. Bensmaia, "Edge orientation signals in tactile afferents of macaques," *J. Neurophysiol.*, vol. 116, no. 6, pp. 2647–2655, Dec. 2016.
[21] R. V. Grigorii, J. E. Colgate, and R. Klatzky, "The spatial profile of skin indentation shapes tactile perception across stimulus frequencies," *Sci. Rep.*, vol. 12, no. 1, Aug. 2022, Art. no. 13185.
[22] S. Kuroki and S. Nishida, "Human tactile detection of within- and inter-finger spatiotemporal phase shifts of low-frequency vibrations," *Sci. Rep.*, vol. 8, no. 1, Dec. 2018, Art. no. 4288, doi: 10.1038/s41598-018-22774-z.
[23] S. Kuroki, J. Watanabe, and S. Nishida, "Neural timing signal for precise tactile timing judgments," *J. Neurophysiol.*, vol. 115, no. 3, pp. 1620–1629, 2016.
[24] J. A. Pruszynski, J. R. Flanagan, and R. S. Johansson, "Fast and accurate edge orientation processing during object manipulation," *Elife*, vol. 7, Apr. 2018, Art. no. e31200, doi: 10.7554/eLife.31200.
[25] L. H. Hardy, G. Rand, and M. C. Rittler, "Tests for the detection and analysis of color-blindness I the Ishihara test: An evaluation," *J. Opt. Soc. Am.*, vol. 35, no. 4, Apr. 1945, Art. no. 268.
[26] S. Kaneko, H. Kajimoto, and V. Hayward, "A case of perceptual completion in spatio-temporal tactile space," in *Proc. Haptics: Sci., Technol., Appl.*, 2018, pp. 49–57.
[27] S. McIntyre, I. Birznieks, R. M. Vickery, A. O. Holcombe, and T. Seizova-Cajic, "The tactile motion aftereffect suggests an intensive code for speed in neurons sensitive to both speed and direction of motion," *J. Neurophysiol.*, vol. 115, no. 3, pp. 1703–1712, Mar. 2016.
[28] R. L. Miller and D. L. Sheinberg, "Evidence for independent processing of shape by vision and touch," *eNeuro*, vol. 9, no. 3, pp 1–14, May 2022, doi: 10.1523/ENEURO.0502-21.2022.
[29] Q. Wang and V. Hayward, "Biomechanically optimized distributed tactile transducer based on lateral skin deformation," *Int. J. Robot. Res.*, vol. 29, no. 4, pp. 323–335, Apr. 2010.
[30] V. Hayward, A. V. Terekhov, S.-C. Wong, P. Geborek, F. Bengtsson, and H. Jörntell, "Spatio-temporal skin strain distributions evoke low variability spike responses in cuneate neurons," *J. Roy. Soc. Interface*, vol. 11, no. 93, Apr. 2014, Art. no. 20131015.
[31] T. F. Jaeger, "Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models," *J. Memory Lang.*, vol. 59, no. 4, pp. 434–446, Nov. 2008.
[32] R. P. Arthur and W. B. Shelley, "The innervation of human epidermis," *J. Invest. Dermatol.*, vol. 32, no. 3, pp. 397–411, Mar. 1959.
[33] B. G. Green, "Localization of thermal sensation: An illusion and synthetic heat," *Percep. Psychophys.*, vol. 22, no. 4, pp. 331–337, Jul. 1977.
[34] R. T. Verrillo, G. A. Gescheider, B. G. Calman, and C. L. Van Doren, "Vibrotactile masking: Effects of one- and two-site stimulation," *Percep. Psychophys.*, vol. 33, no. 4, pp. 379–387, 1983.
[35] J. C. Craig, "Vibrotactile masking: A comparison of energy and pattern maskers," *Percep. Psychophys.*, vol. 31, no. 6, pp. 523–529, Nov. 1982.
[36] J. C. Craig and X. Qian, "Tactile pattern perception by two fingers: Temporal interference and response competition," *Percep. Psychophys.*, vol. 59, no. 2, pp. 252–265, Feb. 1997.
[37] W. Fujisaki and S. Nishida, "A common perceptual temporal limit of binding synchronous inputs across different sensory attributes and modalities," *Proc. Roy. Soc. B: Biol. Sci.*, vol. 277, no. 1692, pp. 2281–2290, 2010.
[38] J. M. Yau, A. Pasupathy, P. J. Fitzgerald, S. S. Hsiao, and C. E. Connor, "Analogous intermediate shape coding in vision and touch," *Proc. Natl. Acad. Sci. USA*, vol. 106, no. 38, pp. 16457–16462, Sep. 2009.
[39] C. C. Pack and S. J. Bensmaia, "Seeing and feeling motion: Canonical computations in vision and touch," *PLoS Biol.*, vol. 13, no. 9, Sep. 2015, Art. no. e1002271.
[40] B. P. Delhaye, K. H. Long, and S. J. Bensmaia, "Neural basis of touch and proprioception in primate cortex," *Comprehensive Physiol.*, vol. 8, no. 4, pp. 1575–1602, Sep. 2018.
[41] Y. Iwamura, A. Iriki, and M. Tanaka, "Bilateral hand representation in the postcentral somatosensory cortex," *Nature*, vol. 369, no. 6481, pp. 554–556, Jun. 1994.
[42] M. Kahrimanovic, W. M. Bergmann Tiest, and A. M. L. Kappers, "Context effects in haptic perception of roughness," *Exp. Brain Res.*, vol. 194, no. 2, pp. 287–297, 2009.
[43] S. Kuroki, J. Watanabe, and S. Nishida, "Integration of vibrotactile frequency information beyond the mechanoreceptor channel and somatotopy," *Sci. Rep.*, vol. 7, no. 1, Dec. 2017, Art. no. 2758.
[44] M. S. Rahman and J. M. Yau, "Somatosensory interactions reveal feature-dependent computations," *J. Neurophysiol.*, vol. 122, no. 1, pp. 5–21, Jul. 2019.
[45] P. M. Evans and J. C. Craig, "Tactile attention and the perception of moving tactile stimuli," *Percep. Psychophys.*, vol. 49, no. 4, pp. 355–364, Apr. 1991.
[46] P. M. Evans and J. C. Craig, "Response competition: A major source of interference in a tactile identification task," *Percep. Psychophys.*, vol. 51, no. 2, pp. 199–206, Feb. 1992.
[47] B. G. Green, "Localization of thermal sensation: An illusion and synthetic heat," *Percep. Psychophys.*, vol. 22, no. 4, pp. 331–337, Jul. 1977.
[48] H.-N. Ho, J. Watanabe, H. Ando, and M. Kashino, "Mechanisms underlying referral of thermal sensations to sites of tactile stimulation," *J. Neurosci.*, vol. 31, no. 1, pp. 208–213, Jan. 2011.
[49] I. Choi, E. Ofek, H. Benko, M. Sinclair, and C. Holz, "CLAW: A multifunctional handheld haptic controller for grasping, touching, and triggering in virtual reality," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2018, pp. 1–13.
[50] J. Lu, Z. Liu, J. Brooks, and P. Lopes, "Chemical haptics: Rendering haptic sensations via topical stimulants," in *Proc. 34th Annu. ACM Symp. User Interface Softw. Technol.*, 2021, pp. 239–257.
[51] T. Murakami, T. Person, C. L. Fernando, and K. Minamizawa, "Altered touch: Miniature haptic display with force, thermal and tactile feedback for augmented haptics," in *Proc. ACM SIGGRAPH Emerg. Technol.*, 2017, pp. 1–2.
[52] Z. Sun, M. Zhu, X. Shan, and C. Lee, "Augmented tactile-perception and haptic-feedback rings as human-machine interfaces aiming for immersive interactions," *Nature Commun.*, vol. 13, no. 1, Sep. 2022, Art. no. 5224.
[53] X. Yu et al., "Skin-integrated wireless haptic interfaces for virtual and augmented reality," *Nature*, vol. 575, no. 7783, pp. 473–479, Nov. 2019.

**Scinob Kuroki** received the Ph.D. degree in information science and technology from the University of Tokyo, Tokyo, Japan, in 2011. She is with NTT Communication Science Laboratories, Atsugi, Japan. Her research is focused on basic tactile perception, in particular frequency perception, motion perception, shape perception, texture perception, and comparisons between different modalities.

**Shin'ya Nishida** is currently a Professor with the Graduate School of Informatics, Kyoto University, Kyoto, Japan, and a Research Professor of Sensory Representation Research Group, NTT Communication Science Labs, Kyoto. He is an expert of psychophysical research on human visual processing in particular material perception, motion perception, cross attribute/modality integration, and time perception.