The roles of lower- and higher-order surface statistics in tactile texture perception

Scinob Kuroki (黒木 忍),1 Masataka Sawayama (澤山 正貴),1,2 and Shin’ya Nishida (西田 眞也)1,3
1NTT Communication Science Laboratories, NTT Corporation, Atsugi, Japan; 2Inria, Bordeaux, France; and 3Department of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University, Kyoto, Japan

Abstract

Humans can haptically discriminate surface textures when there is a significant difference in the statistics of the surface profile. Previous studies on tactile texture discrimination have emphasized the perceptual effects of lower-order statistical features such as carving depth, inter-ridge distance, and anisotropy, which can be characterized by local amplitude spectra or spatial-frequency/orientation subband histograms. However, the real-world surfaces we encounter in everyday life also differ in the higher-order statistics, such as statistics about correlations of nearby spatial-frequencies/orientations. For another modality, vision, the human brain has the ability to use the textural differences in both higher- and lower-order image statistics. In this work, we examined whether the haptic texture perception can use higher-order surface statistics as visual texture perception does, by three-dimensional (3-D)-printing textured surfaces transcribed from different “photos” of natural scenes such as stones and leaves. Even though the maximum carving depth was well above the haptic detection threshold, some texture pairs were hard to discriminate. Specifically, those texture pairs with similar amplitude spectra were difficult to discriminate, which suggests that the lower-order statistics have the dominant effect on tactile texture discrimination. To directly test the poor sensitivity of the tactile texture perception to higher-order surface statistics, we matched the lower-order statistics across different textures using a texture synthesis algorithm and found that haptic discrimination of the matched textures was nearly impossible unless the stimuli contained salient local features. We found no evidence for the ability of the human tactile system to use higher-order surface statistics for texture discrimination.

NEW & NOTEWORTHY Humans can discriminate subtle spatial patterns differences in the surrounding world through their hands, but the underlying computation remains poorly understood. Here, we 3-D-printed textured surfaces and analyzed the tactile discrimination performance regarding the sensitivity to surface statistics. The results suggest that observers have sensitivity to lower-order statistics whereas not to higher-order statistics. That is, touch differs from vision not only in spatiotemporal resolution but also in (in)sensitivity to high-level surface statistics.

image statistics; perception; tactile; texture; 3-D printing

INTRODUCTION

Most objects in the world are covered by surfaces with a variety of textures. By sensing surface textures, we, and many other animals, are able to specify distinct surface areas, identify materials, and estimate surface conditions. Touch and vision are the two main sensory modalities contributing to surface texture perception from statistical spatial patterns. Although researchers have revealed a detailed hierarchical processing of image features for visual texture perception, which for haptic texture perception remains obscure. In this study, we analyze tactile texture processing by referring to recent computational theories of visual texture perception. Several studies have highlighted the quantitative differences in the ability to access various spatial properties between touch and vision (see review for Refs. 1–3). For example, vision is good at detecting geometric patterns such as texture boundaries (2, 4), whereas touch can detect invisibly small surface texture properties (5–9). In a slight departure from these approaches, here we focused on the qualitative
difference of the two modalities with regard to relevant stimulus features/statistics in spatial texture processing.

Early studies on visual texture segregation revealed that the condition for two adjacent textures to be perceptually segregated is the presence of significant differences in the marginal statistics (e.g., average or standard deviation of the histogram) of local orientation and/or spatial frequency (Gabor wavelets) (10, 11), which is presumably represented by the response distribution of V1 neurons (level no. 2 in Fig. 1A). Recent studies examining the perceptual discrimination of natural textures further suggest that two textures are perceptually indistinguishable (become a metameric pair) in peripheral vision when they are matched in terms of not only V1 image statistics but also the joint statistics (e.g., statistics about correlations) of nearby V1 responses (level no. 3 in Fig. 1A), to which V2 and the higher cortical areas are responsive (12, 14, 15). Human observers cannot visually discriminate textures matched in terms of V2 image statistics (with identical Gabor and joint statistics) unless using elaborated attentive spatial processing focused on central vision (12, 16).

In this paper, we refer to the image statistics represented at or below level no. 2 in Fig. 1A as lower order and those represented at or beyond level no. 3 as higher order. To be more specific, following previous neuroscience studies in vision (13, 14) and in audition (17), we considered as lower-order statistics the moments (e.g., mean) of the marginal distribution of intensity of the image (surface height in the present case) and the moments of marginal distribution of subbands (height images filtered by a variety of Gabor-like filters with different orientations and scales). These statistics are respectively named “Marginal” and “Spectral” by Okazawa et al. (see Fig. 3 in Ref. 13), and the latter is closely associated with, though does not directly correspond to, the Fourier amplitude spectrum and the second-order statistics in the terminology of Julesz (11). As for higher-order statistics, we considered them as pairwise correlations (see level no. 3 in Fig. 1A), for example, correlations between spatially neighboring linear filters (“linear cross position” see also Fig. 3 in Ref. 13), correlations between linear filters with neighboring scales (“linear cross scale”), correlations between filter amplitudes with neighboring orientations (“energy cross orientation”), correlations between spatially neighboring filter amplitudes (“energy cross position”), and correlations between filter amplitudes with neighboring scales (“energy cross scale”). These statistics are associated with the Fourier phase spectrum. According to this terminology, visual and auditory texture perception is sensitive to higher-order statistics, in addition to lower-order statistics (14, 17). The question is whether this is also the case for tactile texture perception.

It should be noted that two textures with the same lower-order statistics can be different in higher-order statistics, whereas two textures with different lower-order statistics cannot be the same in higher-order statistics, due to the hierarchical relationship of the lower- and higher-order statistics. Therefore, the critical test to see sensitivity to higher-order statistics is whether human observers can discriminate two textures with the same lower-statistics, but different higher-order statistics.

The somatosensory system has a spatial-information processing stream analogous to that of the visual system. First, a spatiotemporal pattern of skin deformation is sampled by mechanoreceptors. The signals from the mechanoreceptors are then pooled and suppressed by surrounds on the pathway extending from the periphery to cortex, as well as within cortex. Some peripheral afferents may be able to carry some orientation information (15, 20), but orientation selectivity is much more common and robust in the primary somatosensory cortex (S1, area 3b) (e.g., Ref. 21). The tactile receptive field of S1 neurons can be approximated by Gabor functions (22, 23), as are the visual receptive field of V1 neurons. Somatosensory processing becomes more elaborate beyond S1 (e.g., Ref. 24). We do not believe that visual texture processing and tactile texture processing are likely to be identical as the former is rather a spatial processing with one “view” whereas the latter is a combination of spatial and temporal processing. Also, the size of the “window” of tactile texture sensing, such as the size of a finger pad, is limited (see review for Ref. 25). However, there remains a possibility that tactile texture perception can use higher-order surface statistics. For example, some neurons in S2 have selectivity for higher-order shape features, such as curvatures of orientation changes on the skin surface, similar to that of visual neurons in V4 (26).

Figure 1. A: hypothetical diagram of hierarchical information processing in touch, inspired by visual processing (12,13). The input texture is analyzed with Gabor-like units (linear filters) tuned to different positions, scales, and orientations (#2). Correlations are computed by taking outputs of Gabor-like units across different positions, scales, and orientations (#3). B: 3-D-printed stimuli used for experiments.
In this work, we investigated the performance of the tactile system in discriminating a variety of texture patterns (both natural textures and artificial textures made of numerous Gabors). Through our analysis of the tactile discrimination performance, we considered which texture differences the tactile system is sensitive to, and more specifically, whether it can use higher-order surface statistics (at or beyond level no. 3 in Fig. 1A) as visual texture perception does. There is little psychophysical evidence that tactile texture discrimination is sensitive to higher-order surface statistics. The major spatial property of tactile texture perception intensively investigated in the past is roughness (1, 2, 27–30) (other properties, such as hardness and stickiness, are not purely spatial). The spatial stimulus parameters that are known to affect roughness, such as the spatial period and inter-ridge spacing (e.g., Refs. 31–35), as well as curvatures of depth changes (36, 37), for review see Ref. 38, can be described in terms of lower-order surface statistics. However, the neural responses to higher-order shape features (26, 39) suggest potential sensitivity to higher-order statistics.

Most previous tactile texture studies used relatively simple artificial stimuli, such as dots and gratings (32–35, 38, 40, 41), or went with common natural surfaces, such as fabric, wood, and metal (42, 43). With these stimuli, it is not easy to examine the contribution of higher-order surface statistics separately from those of lower-order ones. To overcome this limitation, we rendered complex haptic stimuli using a high-resolution three-dimensional (3-D) printer, as done by a few recent tactile studies (e.g., Refs. 44–46). By printing the surface of a sample material with the 3-D printer, we transcribed complex texture patterns into carving patterns of tangible surface (Fig. 1B). Specifically, we mapped the intensity patterns in the images into the height/depth patterns in the tactile stimuli to produce corresponding spatial patterns in the sensor responses. By manipulating the surface statistics of the printed patterns, we were able to examine whether tactile texture perception can use higher-order texture statistics. Using this powerful methodology, a variety of experimental paradigms developed in vision research (including the texture synthesis algorithm Ref. 47) that have advanced the understanding of visual texture perception can be applied to tactile research.

**MATERIALS AND METHODS**

**Stimuli**

Our tactile stimuli were created by 3-D printing a height map. We considered the height map as an intensity image and controlled its image statistics as in the visual texture literature.

**Texture Patterns**

In the first series of experiments, we used artificial Gaussian bandpass noise patterns and natural scene textures. For the artificial noise, we first applied the two-dimensional fast Fourier transform to a white noise image and extracted specific spatial frequency components by using a two-dimensional Gaussian bandpass filter. There were two stimulus conditions of filter parameter manipulation: the center frequency (CF) and bandwidth (BW) conditions (Fig. 2A, top and center). The CF condition used five center frequencies of the bandpass filter (2, 4, 8, 16, and 32 cycles/image; results in 0.5, 1, 2, 4, 8 cycles/cm when 3-D printed) while keeping the BW 1.5 octaves. The BW condition used five filter bandwidths (0.5, 1.0, 1.5, 2.0, and 2.5) while keeping the center frequency 8 cycles/image (2 cycles/cm). The range of these stimuli was designed to include both fine textures mainly encoded by PC afferents and coarse textures mainly encoded by SA afferents (32, 47–50). In addition to the artificial noise, we used the natural scenes (NSs). For this NS condition, five images were chosen from the natural texture category of the McGill Calibrated Color Image Database (51). The mean and standard deviation of the image intensity were normalized and equalized across images (Fig. 2A). Each image had the resolution of 256 × 256 pixels.

**Printing Tactile Textures**

Tactile stimuli were custom-built for the experiment by using a 3-D printer (16-micrometer resolution) (Objet 260 Connex3, Stratasys) with transparent plastic-like material (VeroClear-RGD810, Objet). Each visual texture stimulus was converted into a 3-D model by taking intensity values as a height map. The printed object was 40 × 40 × 10–12 mm. The contrast difference between complete black (0) and complete white (255) in an image was transcribed to a height (thickness of stimuli; black means deep) difference of 2 mm, with a mean depth of 1 mm. Printing accuracy, measured with a wide-area 3-D measurement system (VR-3100, KEYENCE, Japan), was within 0.056 mm on average. The stimuli were hard, not compliant. Smoothness or roughness depends on the printed height pattern of the stimuli. As long as the stimuli do not contain high-frequency components, they tend to be smooth (e.g., CF1, CF2, BW1). Some stimuli had sharp edges, which is reflected as high contrast (slant of the intensity/luminance) of the visual image.

Before the experiment, the surface of the stimuli was very lightly coated with baby powder (Baby Powder, Johnson & Johnson) using a brush to reduce stick-slips and make the contact between the finger and stimuli as constant as possible. We tried to minimize changes in the contact condition as they may give a cue of flatness of the stimulus surface, degrade the transcription from the carving pattern of the stimuli to the deformation of the finger skin, and/or even fluctuate the relative position between the finger and the stimuli with passive scan.

**Tactile Vibration Stimuli**

A piezoelectric actuator (MU120, MESS-TEK, Japan) was used as a stimulator to reproduce the line scan of the texture height profile swept along a horizontal line. The stimulator outputted sufficient force (800 N as the upper limit) and normally deformed the skin by a position control method so that it could accurately produce the required displacement with a tolerance of few nanometers. The diameter of the stimulator was 12.0 mm, its edges were separated from the rigid surround of the metal boards by a 1.0-mm gap, and the level of the stimulator and the surround was the same (following Ref. 52). The rigid surround limits the spread of surface waves of the skin. The stimulator was fixed on the
**Figure 2.** ABX texture discrimination task for 3-D-printed stimuli of center frequency (CF), bandwidth (BW), and natural scene (NS) sets. A: visual images (256 x 256 pixels) were converted into 3-D carvings (40 x 40 x 10–12 mm) by taking intensity values as height maps. The intensity level (the mean and standard deviation) was matched across images. The bottom panels with red plots are log amplitude spectra of images (i.e., the amplitude of each spatial frequency component in the Fourier spectrum, averaged across orientations. See also Fig. 4A for details.) B: time course of the ABX experiment. Passive scan condition. Observers placed their index finger on the first resting cushion and the linear stage started to move at 40 mm/s. After the finger contacted the tactile stimulus and swiped for 1 s, the stage stopped on the next resting cushion. This was repeated three times, once each for an A, B, and X stimulus, with X a rotated version of A or B. Observers were asked to report verbally whether the third stimulus was the first or second one. No feedback signal was provided. C: results. The averaged discrimination performance (proportion correct) for each stimulus pair is shown by numbers and colors in a matrix format. The 95% CIs of the chance performance for group data is 0.42–0.57 (uncorrected) and 0.38–0.62 (Bonferroni corrected). Each number in parenthesis represents the number of observers out of 10 who had performance above 95% CIs (0.25–0.75) of the chance performance calculated using binomial distribution with a sample size of 12 trials. See also Supplemental Fig. S1 for alternative plots of the discrimination accuracy with CI. CI, confidence interval.
Subband-Matched Textures

To match lower-order statistics, we conducted subband frequency analysis, which has been more widely used in natural scene analysis and is appropriate for the spatial control of nonuniform natural images than Fourier amplitude analysis. The main difference between subband frequency analysis and Fourier amplitude analysis is that the spatial filter of the former is applied to a spatially local window in the image domain. This difference can be more critical for natural scene (NS) images than spatially uniform Gaussian noise (CF, BW). A simple example is that an image includes a spatially local bright (or dark) outlier value. As the Fourier analysis does not take into account the spatial position in the image domain, this outlier value affects all the frequency components. In contrast, as subband frequency analysis relies on multiscale bandpass filter pyramids, this affects only a few subband coefficients (53). To test if textures with identical lower-order (subband) statistics were haptically indistinguishable, we made five images with identical subband histograms (Fig. 5A, bottom M1–M5) by using a texture synthesis algorithm (54).

In this algorithm, the steerable pyramid transform is used. Specifically, each image was decomposed into four spatial-frequency and four orientation bands by convolving the image with spatially oriented linear filters and by subsampling. Using the transformed subbands, this algorithm synthesizes the texture by matching the histograms of transform coefficients of the seed image to those of the target image. We used original image 1 (NS/O1) as the target for the synthesis. The original Heeger and Bergen algorithm uses a white noise image as the seed image. However, as we aimed to make a set of textures while preserving the higher-order statistics of each original texture (O1 to O5 in Fig. 5A), we used the original texture as the seed image. To make the matched image in Fig. 5A, we adopted four spatial frequencies bands and four orientation bands.

Textures Made from 3D Scanned Surfaces

In the second series of experiments with 3-D scanned surfaces, eight height maps (banana leaf, poplar bark, concrete surface, plaster surface, forest floor, tile surface, sand floor, chestnut bark) were chosen from the Quixel Megascans library (55). For the original scale condition, a 40 × 40 mm region was cut from each height map and then 3-D printed in the original scale. For the 1/3-scale condition, a 120 × 120 mm region was cut from each image and then 3-D printed as a 40 × 40 mm surface in a reduced scale.

To make the subband-matched height maps for each 3-D scanned stimuli, the same texture synthesis algorithm of Heeger and Bergen (54) to the first experiments was recruited. We used each original 3-D scanned height map as the target for the synthesis. As the subband distribution of the tactile textures was spatially biased than the visual textures in the first experiments, we adopted more bands for the synthesis, specifically six spatial frequency bands and six orientation bands, and used a brown noise image as the seed image to decrease the difference of subband histograms between the original and synthesized images.

For original-scale 3-D scanned stimuli, the object was printed in original scale, that is, 40 × 40 mm surface with different maximum carving depth (banana leaf: 2.6, poplar bark: 2.4, concrete surface: 5.0, plaster surface: 1.9, forest floor: 5.5, tile surface: 8.0, sand floor: 4.4, chestnut bark: 7.8 mm). For 1/3-scale 3-D scanned stimuli, the object was printed in 1/3 scale in three dimensions, that is, 40 × 40 mm surface with different maximum carving depth (banana leaf: 1.3, poplar bark: 1.6, concrete surface: 2.3, plaster surface: 2.7, forest floor: 2.7, tile surface: 3.0, sand floor: 3.3, chestnut bark: 3.7 mm).

Participants

Thirty-six naive observers and two of the authors (13 males) with normal tactile sensitivity (by self-reports), aged from 21 to 46 yr (32.2 ± 7.73) participated the experiments. All gave written informed consent approved by the NTT Communication Science Laboratory Research Ethics Committee, and all procedures were conducted in accordance with the Declaration of Helsinki.

Procedures

**ABX experiment with CF, BW, and NS modulation.**

Groups of 10 observers participated for each passive, active, static, and vibration condition, with partial overlaps of observers across conditions (Supplemental Table S1; all Supplemental material is available at https://doi.org/10.6084/m9.figshare.13030907.v1). The volunteer observers had never seen the tactile stimuli nor original visual images. An observer sat at a table and placed the index finger of the right hand on home position with their right arm on an arm rest. They performed experiments with their eyes open to maintain their arousal level, but they could not see the tactile stimuli, the equipment, nor experimenter, which were occluded by a black curtain.

In general, observers touched three stimuli, A, B, and X, each for 1 s, where X was a rotated version of A or B, and they verbally reported which one was X, the first or second. No feedback as to their response was given to the observers. We used a rotated version of A or B as X to prompt observers to use statistical texture features rather than non-texture cues such as the position or shape of local features when discriminating the two surfaces. Paired stimuli (A and B) were randomly chosen from five stimuli of the same type (CF, BW, or NS). This procedure is called an ABX task (56). One of the advantages of this task is that observers can purely conduct a direct comparison between the similarity of A and X and that of B and X compared to labeling/verbalizing/categorizing tasks. With the ABX task, the observer can report the relative similarity between X and A and X and B, even if A and B are not clearly labeled. Most previous studies on tactile
texture perception asked observers to judge a specific roughness feature (e.g., “rate the roughness from 0 to 9” or “report which one was rougher”). In contrast, our method allows us to account for any perceptual feature, including roughness, that the observers could use to discriminate stimuli. Furthermore, it enables us to find two physically different textures that are “metameric” (perceptually indiscriminable in any way).

There were four different touching mode conditions. Other than in the vibration condition, the experimenter set three predetermined stimuli (A, B, X) on a linear stage (ERL2, CKD, Japan) before each trial started. By automatically moving the stage, the experimenter was able to guide the three stimuli directly beneath the finger. Thus, observers did not have to move the wrist to touch them. In the passive scan condition, each trial started with the experimenter’s “ready” call. An observer put their right index finger at the rest position at the right edge of the linear stage and pressed the start button on the PC monitor with their left hand to trigger the stage movement. The observer passively scanned each stimulus for 1 s. The linear stage started to move under observers’ right index finger from left to right with a speed of 40 mm/s so that the rightmost one of three stimuli swiped the finger for 1 s. After a 1-s pause, the stage automatically moved again for the second stimulus to swipe, stopped, and then moved again for the third stimulus to swipe. After the third scan, observers lifted their fingers and made a binary verbal report as to which of the first two stimuli (the first or the second) was the same as the last one. The stage moved back to its initial position and the experimenter changed the three stimuli on the stage for the next trial. In active scan condition, observers lifted their right index finger above the first stimulus position and pressed the start button. They freely scanned the stimulus surface for 1 s and lifted their finger again. Then, the stage started to move and the second stimulus came right under their finger. Observers never touched the stimuli when the stage was moving; they touched them three times for 1 s each when the stage had stopped. Observers were not restricted in their finger trajectory or explorative strategy. Rather, they were encouraged to use any cues to solve the task. In the static touch condition, observers put their fingers on static (not moving) stimuli three times each for 1 s, in a similar time course to the other conditions. In the vibration condition, the stage and stimuli were replaced with the piezoelectric actuator. Observers placed their right index finger on the actuator and pressed the start button on the monitor with their left hand. After about 1 s, the actuator vibrated three times each for 1 s with 1.5-s gap. Only in this condition did observers report by button clicking instead of verbally. To mask any subtle sound made by the actuator, observers wore ear plugs and white noise was played continuously from headphones throughout this condition.

There were three kinds of modulation (CF, BW, and NS), five modulation gradations (10 combinations), four touching modes (passive scan, active scan, static touch, and vibration), and 12 repetitions for each combination. In total, 360 trials were conducted for each touching mode, and sessions were roughly divided into 10 blocks. No block lasted longer than 15 min, with at least a 10-min break between successive blocks. Within each block, the kind of modulation and the gradations were randomized and the touching mode was fixed. Each observer completed the main session in 90–120 min, and addition of practice sessions and breaks doubled the total duty time.

We should point out that in our ABX task, the exploration time for each stimulus was only 1 s, which is shorter than the typical trial times used in previous studies on tactile roughness perception (e.g., Ref. 25). We chose this duration because the observers had to explore three stimuli before responding and a longer exploration time would consume a greater memory cost. Although this may affect performance, preliminary tests by the authors did not show any significant effect of the increase in exploration time on the discrimination task.

**ABX experiment with original NS and NS subband-matched modulation.**

Ten observers participated. The equipment and procedure were almost identical to those of the passive scan condition in the ABX experiment with CF, BW, and NS modulation, except for the stimulus. The experiment was conducted with two kinds of modulation (original and subband-matched NS), five modulation gradations, one touching mode (passive scan), and 12 repetitions for each combination. Each observer completed the session in 70–90 min (except breaks).

**ABX experiment with 3-D scanned stimuli and subband-matched stimuli.**

Ten observers participated. The equipment and procedure were almost identical to those in the passive scan condition in the ABX experiment, except for the stimulus pair (i.e., how A and B were chosen). In other experiments, A and B were chosen from the same kind of modulation (e.g., A = NS1, B = NS2). In this experiment, A and B were chosen from the same scale of same object (e.g., A = 3-D scanned concrete surface, B = subband-matched concrete surface). The experiment was conducted with two kinds of origin (3-D scanned stimuli and subband-matched stimuli), eight different object, two different scale (full-scale and 1/3 scale), one touching mode (passive scan), and 12 repetitions for each combination. Each observer completed the session in 70–90 min (except breaks).

**Behavioral Data Analysis**

We recruited 10 observers for each condition and had each observer perform 12 trials for each combination for ABX tasks. To determine whether the observers could discriminate a given stimulus pair (A and B), we calculated a 95% confidence interval (CI) of the chance performance (0.5) and compared whether the group average of discrimination performance (proportion correct) for each stimuli pair was above the upper boundary of this CI or not. The CI was calculated using binomial distribution. With the sample size of 120 (12 trials × 10 observers), the power of a one-sided binomial test is higher than 0.99 when it is used to distinguish threshold performance (0.75) from chance (0.5) with a significance level of 0.05. We checked individual performance in addition to the averaged data. In particular, the number of observers who had performance significantly above the level of chance in the ABX tasks
was investigated. The 95% CIs of the chance performance (0.5) calculated using binomial distribution with a sample size of 12 trials is 0.25–0.75. We compared whether the performance of each observer was above these CIs or not. The results for these individual analyses are shown in parentheses in the figures (Figs. 2, 3, 5, and 6).

To statistically compare the performance among different stimulus conditions, we computed the CI by using a boot strapping method [resampling method (57)], where we calculated the proportion of correct responses for each pair of surfaces of each of 10 observers, and the data of 10 observers were then resampled and averaged to calculate the

**Figure 3.** ABX texture discrimination task with four different touching modes. In the active scan condition, observers could freely explore the stimuli for 1 s with their index finger. In the static touch condition, observers put their finger on the stimuli for 1 s. They were not allowed to tangentially move or scan their finger over the stimuli. In the vibration condition, observers’ finger was vibrated by a piezo-electric actuator. The vibration pattern was one of the texture height profiles swept along a randomly chosen horizontal line. Top panels for passive scan conditions are replots of the data from Fig. 2C. Numbers in parentheses represent the number of observers out of 10 who had performance above 95% CIs (0.25–0.75) of the chance performance. The observers could not discriminate some pairs of natural scene (NS) stimuli regardless of the touching mode. BW, bandwidth; CF, center frequency; CI, confidence interval.

### Figure 3

**Center frequency**

| Passive scan | Active scan | Static touch | Vibration |
|--------------|-------------|--------------|-----------|
| CF1 CF2 CF3 CF4 CF5 | BW1 BW2 BW3 BW4 BW5 | BW1 BW2 BW3 BW4 BW5 | BW1 BW2 BW3 BW4 BW5 |
| 0.91 (9) 0.97 (10) 0.97 (10) 1 (10) | 0.91 (8) 0.97 (9) 1 (10) 0.89 (8) | 0.86 (5) 0.91 (10) 0.87 (9) 0.97 (9) | 0.86 (1) 0.91 (6) 0.89 (8) |
| 0.93 (10) 0.97 (10) 0.96 (9) 0.93 (9) 0.93 | 0.93 (10) 0.97 (10) 0.96 (9) 0.93 (9) 0.93 | 0.75 (3) 0.69 (9) 0.72 (2) 0.75 (1) 0.72 | 0.65 (2) 0.75 (2) 0.69 (2) 0.75 (2) 0.69 |
| 0.58 (2) 0.58 (2) 0.58 (2) | 0.58 (2) 0.58 (2) 0.58 (2) | 0.58 (2) 0.58 (2) 0.58 (2) | 0.58 (2) 0.58 (2) 0.58 (2) |

**Bandwidth**

| Passive scan | Active scan | Static touch | Vibration |
|--------------|-------------|--------------|-----------|
| BW1 BW2 BW3 BW4 BW5 | BW1 BW2 BW3 BW4 BW5 | BW1 BW2 BW3 BW4 BW5 | BW1 BW2 BW3 BW4 BW5 |
| 0.52 (0) 0.58 (3) 0.84 (7) 0.96 (10) | 0.52 (0) 0.58 (3) 0.84 (7) 0.96 (10) | 0.52 (0) 0.58 (3) 0.84 (7) 0.96 (10) | 0.52 (0) 0.58 (3) 0.84 (7) 0.96 (10) |
| 0.66 (2) 0.82 (6) 0.95 (10) | 0.66 (2) 0.82 (6) 0.95 (10) | 0.66 (2) 0.82 (6) 0.95 (10) | 0.66 (2) 0.82 (6) 0.95 (10) |
| 0.54 (2) 0.83 (7) | 0.54 (2) 0.83 (7) | 0.54 (2) 0.83 (7) | 0.54 (2) 0.83 (7) |

**Natural scene**

| Passive scan | Active scan | Static touch | Vibration |
|--------------|-------------|--------------|-----------|
| NS1 NS2 NS3 NS4 NS5 | NS1 NS2 NS3 NS4 NS5 | NS1 NS2 NS3 NS4 NS5 | NS1 NS2 NS3 NS4 NS5 |
| 0.48 (0) 0.88 (6) 0.48 (0) 0.58 (2) | 0.48 (0) 0.88 (6) 0.48 (0) 0.58 (2) | 0.48 (0) 0.88 (6) 0.48 (0) 0.58 (2) | 0.48 (0) 0.88 (6) 0.48 (0) 0.58 (2) |
| 0.82 (6) 0.57 (0) 0.69 (3) | 0.82 (6) 0.57 (0) 0.69 (3) | 0.82 (6) 0.57 (0) 0.69 (3) | 0.82 (6) 0.57 (0) 0.69 (3) |
| 0.85 (7) 0.81 (5) | 0.85 (7) 0.81 (5) | 0.85 (7) 0.81 (5) | 0.85 (7) 0.81 (5) |

HAPTIC METAMERIC TEXTURES

**J Neurophysiol** /C15 doi:10.1152/jn.00577.2020 /www.jn.org 101
distribution of mean values (see Fig. 6 and Supplemental Figs. S1, S4, and S11).

**Logistic Regression Analysis Based on Amplitude Differences**

To investigate whether the discrimination performance can be explained by the amplitude differences in the textures, we used a logistic regression analysis. As shown in Fig. 4A, the amplitude of each spatial component was calculated on a log scale and sampled to 10 points in 0.7-octave intervals. The difference between the amplitude spectra of paired textures was calculated at each sampled point. We used the ten values as predictor variables for each discrimination performance as follows.

\[
\hat{\beta} = \arg\min_\beta \left[ \sum_{i=1}^{N} \left( \frac{1}{1 + \exp\left(-\sum_{p=0}^{P} \beta(k) x(i, k)\right)} - y(i) \right)^2 \right],
\]

where \(x\) is the predictor variable, \(\beta\) its coefficients, \(y\) the human discrimination accuracy for each pair, \(N\) the number of pairs, and \(P\) the number of the predictor variables. We used MATLAB function `fmincon` for the optimization. By the sparse logistic regression, 5 of the 10 predictor variables survived and were used in estimating the performance shown in Fig. 4B.

**Analysis of Stimulus Similarity in the Context of Lower- and Higher-Order Statistics**

To quantitatively describe the stimulus similarity, we made a similarity index in the context of lower- and higher-order statistics. Figure 5C shows the differences between each texture pair in lower-order or higher-order statistics. The indices of these statistics were defined using the subband images decomposed from each texture image by a steerable filter. The index of lower-order statistics is defined as the mean energy response across the positions of each subband image (spectral energy, c.f. Refs. 12 and 13). This index pools the spatial structure of each subband image, that is, phase information, and corresponds to the amplitude information. The difference between each texture pair in the lower-order statistics was calculated for each subband image and then averaged across different subband images. In contrast, the difference in the higher-order statistics was calculated for each position of each subband image. The root-mean-squared difference of each subband image was calculated and averaged across different subband images. The numbers of scales and orientations were the same as in the texture synthesis manipulation (four scales and four orientations). The differences are normalized in 0 to 1 for each lower-order or higher-order statistics condition, and the difference in absolute values across two statistics does not imply a difference in perceived quantity.

**Simulation**

We simulated spatiotemporally distributed firing patterns of three different types of tactile afferents by using the computational model “TouchSim,” which can reproduce major response characteristics that have been clarified by previous research (58). The original parameters of the model were based on measured spiking data obtained with monkeys. The simulated responses of afferents closely match the known spiking responses of actual afferents (both precise millisecond spike-timings and firing rates) to various classes of stimuli (for example, vibrations, edges, and textured surfaces).

Responses to NS subband-matched stimuli were simulated. The stimulus was scanned across the skin. The contact area was defined as a rectangle (20-mm length and 10-mm...
(input spacing, defined as pin spacing in “TouchSim”) was set to 0.1 mm. The skin contact area was indented at the center of the index fingertip with the averaged depth of 1 mm, moved across the stimuli at a speed of 40 mm/s for 1 s. Realistically distributed afferents (288 SA, 569 RA, and 102 PC in index fingertip “D2d”) were simulated with a 1-ms resolution. The simulation was repeated 12 times for each stimulus, withafferent distribution, stimulus contact area, and scan direction fixed.

As we simulated temporally and spatially distributed firings, higher-order statistical information, if any, should be embedded in the firing pattern. To quantitatively test whether this pattern is similar when touching the same stimulus compared with when touching different stimuli, we conducted metric space analysis. In particular, we calculated

**Figure 5.** Original and matched natural scene (NS) stimuli. A: original NS images and histogram-matched NS images. Matched stimuli (M1–M5) were based on different original images (NS/O1–O5) but shared the subband histogram of NS1. B: results. Observers who could discriminate a few pairs of original NS stimuli could not discriminate NS subband-matched stimuli. The 95% CIs of the chance performance is 0.42–0.57 (uncorrected) and 0.38–0.62 (Bonferroni corrected); see also Supplemental Fig. S4. Numbers in parentheses represent the number of observers out of 10 who had performance above 95% CIs (0.25–0.75) of the chance performance. C: image statistics of original and matched NS stimuli. Each patch represents difference in the amount of lower-/higher-order statistics between paired stimuli. As higher-order statistics depend on the lower-order image features, matching the lower-order statistics also affected the higher-order statistics (in this case, enhanced differences), as shown in the right bottom panel. CI, confidence interval.
the Victor distance (59) following the previous “TouchSim” study (58). This analysis enables us to compare the similarity of the timing of the spikes by introducing cost parameter \( q \), which we set to 100 (corresponding to 10 ms) in this study. For each afferent model, the Victor distance was calculated between all pairwise combinations of the simulated trials for five stimuli. Obtained distances were normalized by the total number of spikes in the pair and then averaged for the same kind of afferent model and for the same pair of stimuli (Supplemental Fig. S8).

## RESULTS

### Texture Discrimination for 3-D-Printed Stimuli

We evaluated human tactile accuracy in discriminating a pair of 3-D-printed textures using an ABX task (56). After passively scanning three textures, A, B, and X in order, with the fingertip of the right index finger, the observer had to judge whether X was A or B (Fig. 2B). In the first series of experiments, discrimination performance was measured for three sets of five textures (Fig. 2A) and analyzed from the viewpoint of lower-order statistics and higher-order statistics. The first two sets were spatially bandpass random noise patterns, each made of numerous Gabor components. The variables across textures were two moments of subbands that control lower-order statistics: the center frequency (CF) for the first set and the bandwidth (BW) for the second set. The last set consisted of five natural visual textures [hereafter referred to as natural scenes (NS)].

The discrimination performance of 10 observers obtained with the CF-variable set was very good (Fig. 2C, top); all observers were able to discriminate the most discrepant pair (CF1 and CF5), and a one-octave difference in the CF was sufficient for nearly perfect discrimination, except for the high-frequency pair (CF4 and CF5; see also Supplemental Fig. S1). The results are consistent with previous findings obtained with analogous conditions (e.g., Refs. 1, 2, and 29). Varying the BW, on the other hand, is a novel approach in the quest for haptic sensitive stimulus variables. Although the group-average discrimination accuracy of a one-octave difference in BW was 0.66 at best, that of a three-octave difference attained 0.95 (Fig. 2C, middle). Variations in the CF and BW are changes in lower-order image statistics (visible in the amplitude spectrum). The results therefore indicate that the tactile texture perception is sensitive to differences in lower-order image statistics, although the tactile discrimination was not as good as the visual one (which is nearly perfect for the gray-level version of the stimuli).

For the third set, we printed monochromatic natural visual scene images (i.e., irregular patterns of stones, leaves, actiniae, etc.; Fig. 2A) with transforming image intensity to carving depth. We used visual textures because we wanted a variety of textures with complex higher-order structures. Note also that our procedure did not produce highly unnatural tactile textures, as spatial coordinates are common between visual and tactile images, and the intensity-depth transformation has a theoretical ground such that deeper parts of surface texture are darker under diffuse illumination, known as vignetting (60). We matched across textures the average and variance of the intensity (carving depth) while leaving the other statistical differences intact. Visually, five textures had very different spatial patterns. Nevertheless, by touch, the 10 observers could discriminate some pairs, but not others (Fig. 2C, bottom). In particular, none of the pairs among NS1 (stones), NS2 (coral), and NS4 (leaves) were significantly different from chance (Fig. S1A).

As shown in numbers in parenthesis in Fig. 2C, individual discrimination performance for these pairs was not significantly different from chance for all observers, similar to the trends found with the group average. Anecdotal reports from our observers suggest that the texture pattern (i.e., how the stimuli would look) was hard to guess and indistinguishable from other stimuli. The exception was NS3, which they described as somewhat “spikier” than the rest. In summary, the same observers who could discriminate CF and BW stimuli could not discriminate some of the NS stimuli. Supplemental Fig. S2 shows the multidimensional scaling (MDS) of additional pairwise similarity judgments with CF, BW, and NS stimuli. As with the results of the main discrimination experiment, NS stimuli were clustered.

### Texture Discrimination in Different Touching Modes

As haptic performance is known to be significantly affected by the touching mode, we repeated our ABX experiment with three other touching modes. In the main experiment, we used the passive scan mode to match the speed and trajectory of scanning across different stimuli and different observers as much as possible. On the other hand, active scanning, in which observers can freely move their finger to explore the stimulus surface, may be able to provide richer spatial information than passive scanning (61, 62). We therefore conducted the ABX experiment with the active scan mode. The discrimination performance was similar or slightly improved with active scan (0.96 for CF; 0.85 for BW; 0.70 for NS on average) compared with that with passive scan (0.91 for CF; 0.76 for BW; 0.76 for NS) (Fig. 3, top two rows; see Supplemental Fig. S1 for a detailed comparison). It can be seen that a few pairs of CF and BW stimuli, which could not be discriminated with passive scan, became discriminable with active scan. This may be partially due to the difference in scan density, scan area, and scan speed between the two touching modes. Still, the NS stimuli that were indistinguishable in the passive scan condition were also indistinguishable in the active scan condition. This is in good agreement with previous findings that tactile texture perception, including that of roughness and orientation, is relatively insensitive to changes in the exploration speed (35, 63) and exploration method (5, 64–67). The other two touching modes were concerned with possible summation effects. As the tactile system shows drastic spatial and temporal summation particularly with subthreshold stimuli (68, 69), texture discrimination performance might be seriously violated for complex textures like NS that give rapidly changing input in space and time. To reduce potentially negative effects of the temporal summation on texture perception, we tested the static touch mode, where temporal information was limited and observers could use only the spatial pattern of the stimuli. We also tested the vibration mode, where a single stimulator produced temporally changing height pattern that reflected the line scan of the spatial pattern of the

---

**HAPTIC METAMERIC TEXTURES**

---

**J Neurophysiol** · doi:10.1152/jn.00577.2020 · www.jn.org
stimuli. In this mode, observers were not able to use the spatial distribution of finger deformations, only the temporal pattern of deformations. In neither mode did performance improve (Fig. 3, bottom two rows), suggesting the importance of both spatial information and temporal information in haptic texture discrimination. One can expect that high-frequency information (mainly encoded by PC afferent) was difficult to retrieve under the static condition, while low-frequency information (mainly encoded by SA afferent) was difficult under the vibration condition, which is in good agreement with the results of CF and BW sets. Most importantly to the current research question, the results indicate that some NS pairs are nearly metameric regardless of the mode of touching.

Lower-Order Statistics and Behavioural Performance

The results obtained with the CF and BW stimuli indicate that tactile texture perception is sensitive to (some sorts of) differences in the amplitude spectrum or in the lower-order statistics. We next consider how well the discrimination performance for the NS stimuli can be explained by the differences in the amplitude spectrum. It is known that natural visual textures tend to have the amplitude spectrum falling with the spatial frequency by a factor of $f^{-4}$ (70). Our NS stimuli also have such amplitude spectra, which are similar to one another (Fig. 2A), although the slope of the spectrum differs for some textures. For example, the experimental results showed that pairs that included NS3 were easy to discriminate, which may be due to the fact that NS3 had the most unique amplitude spectrum among the NS stimuli. To analyze whether the similarity of amplitude spectra can explain the discrimination performance of tactile textures, we first integrated the amplitude differences between the paired textures over frequency (Fig. 4A) and then regressed the net amplitude difference to the discrimination performance by using a logistic regression analysis for all CF, BW, and NS conditions. Figure 4B shows the estimated performance plotted against the human performance. Although we did not consider orientation, the correlation was fairly high: $R^2 = 0.65$ and 0.81 for the overall correlation and NS condition correlation, respectively. That is, the more similar the amplitude spectrum was, the more difficult the tactile texture discrimination became.

Texture Discrimination for Stimuli with Matched Lower-Order Statistics

The results obtained so far can be ascribed solely to tactile texture processing sensitive to lower-order statistics, suggesting that tactile texture processing may not be very sensitive to higher-order statistics (e.g., joint Gabor statistics, phase spectra). If the tactile texture processing is indeed insensitive to higher-order statistics, those textures with identical lower-order (subband) statistics will be haptically indistinguishable (i.e., become a metameric pair) even if they differ from one another in the higher-order statistics. To test this hypothesis, we made five images with identical subband histograms (Fig. 5A, bottom M1–M5, see also Supplemental Fig. S3) by matching the subband histogram of four NS images (NS2–NS5) to that of NS1 (based on Ref. 54; see MATERIALS AND METHODS). Figure 5C shows a quantitative comparison of the amount of lower-and higher-order statistics before and after histogram matching, which assures the validity of our image manipulation (see details in MATERIALS AND METHODS). Visually, the matched images looked different from one another, and they were similar to the original images with regard to global patterning. However, the haptic discrimination performance was nearly chance: 0.53 on average. The performance of each stimulus pair was within the range of uncorrected CI of chance performance for all but one (M2–M5 pair, 61%), and within the range of Bonferroni corrected CI of chance performance for all (Supplemental Fig. S4). Similarly, individual discrimination performance for the matched stimuli (M1–M5) was not significantly different from chance, with the exception that only a few were able to discriminate specific pairs. This was in marked contrast to the original condition (NS1–NS5), where discrimination of more than half pairs was significantly above the level of chance for most observers. The result of the MDS analysis of the pairwise similarity ratings also indicate haptic metamericism of the matched stimuli: NS stimuli are similar but somehow distributed in perceptual space, while they are concentrated around the base stimulus (NS1) when their histograms are matched (Supplemental Fig. S5). These results indicate that tactile texture discrimination became nearly impossible when the residual lower-order differences among NS1–NS5 were removed, and that visible differences in higher-order statistics among M1–M5 contributed little to tactile texture discrimination (one might then wonder what would happen when higher-order statistics are identical while lower-order statistics are different. However, making such stimuli is not technically feasible since higher-order features are defined by the relationships between lower-order features.).

Note that the poor tactile discrimination performance in the matched condition should be ascribed to the shortage of sensitivity of the tactile texture processing to the higher-order statistics contained in the current stimuli, not to the general difficulty of our procedure in revealing sensitivity to higher-order statistics. This is because when we collected comparable data with vision using a similar procedure (ABX task with X rotation, pairwise discrimination of histogram-matched textures presented on the display as gray-scale intensity images), the texture discrimination performance was much higher (Supplemental Fig. S6).

Texture Discrimination for 3-D-Scanned Surfaces

So far, the natural textures we used were made from visual images. To test the generality of our findings beyond this stimulus set, we checked whether subband matching could make metameric stimuli for haptic textures made from actually measured depth map (3-D-scanned data) of real textures in the world. We selected eight surfaces of natural and artificial objects (Fig. 6) from a database of 3-D-scanned surfaces (Quixel Megascans). Note that our strategy here is to preserve the complexity and naturalness of the real world as much as possible, and thus these chosen objects have a highly diverse distribution in spatial scale and amplitude spectrum (e.g., the maximum carving depth Rz in Fig. 6, see also Supplemental Fig. S7). For example, chestnut stimuli are indeed much spikier than banana leaf stimuli (note that
the image intensity of Fig. 6 is normalized on an image-by-image basis for the sake of pattern readability and discrimination across newly printed natural haptic textures was easy by touch as well as by vision. Given that each texture has highly different frequency characteristics, exact matching of subband histograms across them (i.e., creating matched image by matching subband histogram of different textures, as in Fig. 5A) was hard or impossible. Thus, to test our hypothesis, this experiment separately examined difficulty in discrimination for each stimulus between the original texture and a histogram-matched texture. The eight 3-D-scanned surfaces were 3-D printed in the original scale. We also printed the same eight surfaces in 1/3 scale to test the effect of the validity of the scale on the texture discrimination performance. Then we created a histogram-matched version for each of the 16 stimuli by synthesizing the original one and a brown noise image, and tested whether human observers could discriminate each pair of the original and histogram-matched stimuli by passive scans. The results (Fig. 6) showed that discrimination performance was not significantly different from chance (within 95% Bonferroni-corrected CI) for the majority of the stimuli (11/16). These stimuli included not only image pairs visually very similar to each other (e.g., concrete pair) but also those that look fairly different (e.g., tile pair). As we tested a wide range of textured surfaces, some of them contain large salient features (e.g., veins of banana leaf). For those textures, observers could use differences in these features in addition to those in surface textures to discriminate the original and matched stimuli, and this was probably the reason banana leaf and plaster showed significantly-above-chance performances (see also discussion about “shape perception”). Overall, subband matching with natural haptic textures supports our conjecture that matching lower-order statistics makes haptic textures metameric as long as the textures are characterized by image statistics, not by the presence of large-scale features. In addition, difference in performances between discrimination with full-scale stimuli and with 1/3 scale stimuli were minor. There was no indication that human tactile texture discrimination is tuned to the actual scale of the texture in the natural environment.
**GENERAL DISCUSSION**

The main question of this study was whether the human tactile mechanism has sensitivity to higher-order statistics (e.g., joint statistics of local orientations) in addition to lower-order statistics (e.g., statistics of local amplitude spectrum). We directly examined whether human observers could detect differences in the higher-order texture statistics of 3-D-printed stimuli whose height patterns were manipulated with regard to image statistics. The discrimination performance of ten observers obtained with the CF-variable set and BW-variable set was reasonably good; on the other hand, that with the NS set was not significantly different from chance for many pairs and observers, though some of them could discriminate some pairs (Fig. 2). Our analysis suggested that the lower-order statistics of the stimuli can explain the residual performance of discriminating NS stimuli, in addition to the performance of discriminating CF and BW stimuli (Fig. 4). As higher-order statics are different in all the cases (i.e., between CF stimuli, between BW stimuli, and between NS stimuli), the findings suggest that tactile texture processing may not be sensitive to higher-order statistics. In order to directly test this possibility, we created matched-NS stimuli with identical subband histograms (lower-order statistics) while leaving the higher-order statistical differences distinct, and measured haptic discrimination performance. As a result, discrimination of the matched textures was nearly impossible (Fig. 5), that is, no evidence was found for the ability of the human tactile system to use higher-order surface statistics for texture discrimination. Note that these haptically metameric stimuli are visually discriminable. Our present results are consistent with the idea that the tactile spatial pattern processor can take into account some statistics related to the local amplitude spectrum (level nos. 1 and 2 in Fig. 1A), including the center frequency and bandwidth, but not those related to the phase spectrum or joint statistics (at level no. 3 in Fig. 1A).

**Limitations**

Before discussing the implications of the current study, we should clarify the limitations.

Our findings show that tactile texture segregation is sensitive to differences in surface features classified as lower-order statistics. We do not consider however that tactile texture segregation is sensitive to lower-order statistics of any kinds. Detailed characterization of the contributions of a variety of lower-order statistics to tactile texture segregation remains to be studied.

This study explored a novel hypothesis about the computational limitations of the tactile texture system. However, as the current data were obtained under limited conditions, it is impossible to verify the general inability of the tactile texture system to use higher-order image statistics. The behaviors of the tactile texture system could be different for different physical material conditions, different body parts, different mechanoreceptor channels, different spatiotemporal scales, different types of higher-order statistics (e.g., Ref. 71), different types of texture dimensions (an intensity map can be haptically represented in a variety of ways), and so on. In addition, we presented a different part of the surface in our ABX task, in order to avoid the pattern-matching strategy. This procedure might induce difficulty in identifying the same surface, especially with the non-uniform surfaces used in the last experiment, and may partially (but not completely) explain the low-discrimination performance. Further study is warranted to draw a definitive conclusion as to whether the tactile texture system is by all means insensitive to any kind of higher-order statistics.

We compared touch to vision with regard to the sensitivity to image statistics in texture perception, but the general (dis)similarity of the two modalities is outside the scope of the present study. Having chosen the conditions of the tactile experiments that we believed to be optimal for the current task, we revealed a computational limitation of tactile texture processing distinct from visual processing. In most cases, however, our experiments were not designed to directly compare the performance of the two modalities. There are many differences, along with similarities, in the physical and neural structure between the two modalities (3, 26, 72–78). Indeed, visual texture perception is rather a spatial processing while tactile texture perception is a combination of spatial and temporal processing. Also, the size of the area that can be seen at once with the eyes and the size that can be touched with the fingertips is very different. Although our control experiment suggests that visual texture discrimination is sensitive to higher-order statistics under conditions where the peripheral sensor resolution is matched (Supplemental Fig. S6), we do not deny the possibility that vision may also become insensitive to higher-order statistics when other characteristics of touch are simulated by vision (e.g., limited sensing areas, lack of peripheral vision, limited spatial frequency range). This is obviously an interesting line of study we should pursue in future.

Keeping these limitations in mind, we next consider the implications of tactile texture processing being insensitive to higher-order image statistics in the current experiments.

**Simulation of Skin Deformation and Neural Activity in Periphery**

We wanted to explore whether the observed insensitivity to the higher-order statistics reflects the processing characteristics of the central nervous system or simple information loss in the periphery, that is, the elasticity of the skin and noisy sparse sampling by mechanoreceptors. Indeed, local surface features of touched surfaces are known to be reliably preserved in peripheral neural activation with fine spatial resolution in general (20, 79); however, how our stimuli are actually encoded remains obscure. To make a best guess as to how the difference in spatial texture information is represented in peripheral neural activation and to exclude the trivial explanation that our results were simply reflecting mechanical interference on periphery, we simulated responses of tactile afferents using the computational model TouchSim (S8). Using this model, we examined whether sufficient information for texture discrimination remains at peripheral stages. What we found was that the firing similarities between pairs of identical texture stimuli with independent neural noise were always higher than those between pairs comprising two different stimuli, regardless of the type of afferent (Supplemental Fig. S8). That is, the firing model successfully discriminated...
Relationship with Roughness Perception

Roughness has been explored extensively and is recognized as a major feature in haptic texture perception (1, 2, 27–30). Past studies have reported a variety of properties of roughness perception, including remarkable discrimination performance—we can detect even nanometer-scale differences (86). Concerning the relationship with the current study, roughness perception can be ascribed to the tactile responses to the lower-order statistics, such as density or the deviation of surface elements (level nos. 1 and 2 in Fig. 1A). Industrial indexes of surface roughness, such as the arithmetical mean height of a surface (Ra) and maximum height of a surface (Rz), are also associated with lower-order statistics. As far as we are aware, the known characteristics of texture roughness perception do not conflict with our hypothesis that the tactile system is insensitive to higher-order feature statistics and thus possibly different from those of visual perception. Indeed, there are studies arguing the differences of the neural mechanisms to analyze spatial information between vision and touch (e.g., Refs. 3, 25, and 76). In addition, not only our group but also others have noticed this haptic metamericism of the subband-matched stimuli. For example, a group of tribology researchers 3-D-printed surface textures with random topography, and their stimuli were very similar to ours. Their recent paper reported a low discrimination ability in human observers (45), which is also consistent with our results and theory.

Though our observers could use texture differences in dimensions other than roughness to accomplish the ABX discrimination task with our complex stimuli, we do not exclude the possibility that they mainly relied on what previous studies called roughness in performing the discrimination task. We therefore considered whether the current results can be accounted for by previously proposed neural indexes of roughness perception. We computed two neural roughness indexes—the mean impulse rate of high-frequency PC afferents (40, 41) and the spatial variation of low-frequency SA afferents (48, 87)—from the output of TouchSim. Note that these indexes do not take into account higher-order statistics, so we did rotate the texture and add spatial jitter derived from contact position variability when repeating the simulated neural responses. We found the simulated indexes to indicate a similar or even higher discrimination performance for original NS and NS subband-matched stimuli compared to the observed behavioral performance (Supplemental Fig. S9). The predictions of these indexes might be improved by adjustment of the noise parameters. Further investigation is warranted to determine how well these or other neural roughness indexes can explain the tactile perception of complex texture patterns such as the ones we introduced.

There are textures that look identical (visually metameric) but can be discriminated by touch due to differing microstructures that are too small for the eye to resolve [e.g., detecting microstructure (8)]. However, this is not inconsistent with our conjecture, since such textures with microscale roughness do not have specific higher-order spatial structures and are presumably detected through temporal vibration patterns of skin responses (40, 88, 89).

Relationship with Shape Perception

Although our findings indicate that tactile “texture” perception is insensitive to higher-order statistics, they do not exclude a possibility that tactile “shape” perception has sensitivity to some higher-order features. Tactile shape perception shows high sensitivity to stimulus orientation. Edge orientation acuity was reported to be around $20^\circ$ (21, 65) or even smaller (90). Furthermore, human touch can discriminate more complex spatial patterns such as letters of the alphabet. One may interpret this result to suggest the ability of the tactile shape perception to discriminate some spatial phase differences. Indeed, neurons in area 2 (39, 91) and S2 (26) show sensitivity to higher-order shape features (i.e., particular curvatures of orientation changes on the skin surface). However, tactile shape perception and tactile texture perception have been studied nearly independently by using different types of stimuli. Shape perception has been tested with simple/local stimuli such as raised simple line/curvature patterns, whereas texture perception has been tested with complex/global stimuli, including natural textures. Although shape perception focuses on a specific location in the stimulus, texture perception should grasp the holistic statistical nature of the field under crowding conditions as in the current experiments. We therefore do not assume that texture perception can be explained by the same mechanism as that for shape perception, and this difference may be reflected in the significantly-above-chance discrimination performance of a few pairs of natural haptic texture based on 3-D-scanned data and noise texture that was subband-matched. Indeed, shape processing and texture processing may be segregated in the somatosensory cortex [e.g., lesions in the parietal opercular cortex are known to impair shape recognition but not roughness recognition (92)]. At present, we have no clear evidence of the contribution of a neural mechanism sensitive to higher-order features (curvature) to surface texture perception (see also Supplemental Fig. S10).

Relationship with Braille Reading

Experienced human observers often show a remarkable ability to read Braille characters by touch, and this tactile character recognition is very similar in performance to visual character recognition when the image resolution is matched between the modalities (93–97). Our conjecture that tactile
texture perception is unable to use higher-order spatial statistics may appear incompatible with such tactile character reading. However, we do not believe it is. First of all, many characters differ from one another in terms of amplitude spectra. In the case of Braille, among 325 pairs of the 26 letters of the English alphabet, only four pairs have identical dipole (second-order) statistics and thus identical amplitude spectra (98). All of them are vertically flipped pairs. Within each pair, the two characters are sometimes confused with each other, but not very often (96). Successful discrimination of isodipole Braille characters however should be regarded as tactile shape perception based on position and/or orientation of specific features. In contrast, what we examined here is tactile texture perception based on the surface statistics sampled at multiple locations within the texture. Note also that our observers attempted to identify two textures in different overall orientations. Under this condition, discrimination of isodipole Braille characters would be logically impossible. Therefore, the ability to read tactile characters does not counter our argument about the computational limitation in tactile texture processing.

Other Issues

Overall reversal of the signal sign (contrast polarity) affects some lower-order statistics [e.g., the pixel intensity histogram (first/second-order statistics)], but not the local amplitude spectrum. Although the visual system responds to positive (white) and negative (black) elements in an asymmetric way (e.g., 99), the tactile system may do so much more strongly. A sign reversal swaps convex elements with concave ones, thereby introducing a large change in skin deformation patterns (e.g., pin or hole elements). Indeed, we found that human observers could discriminate the original and the sign-reversal version for our NS stimuli (Supplemental Fig. SII). Note that this result does not conflict with our general conclusion of haptic insensitivity to higher-order statistics, since negative-positive inversion stimuli share amplitude spectrum but differ in other lower-order statistics (intensity histogram). One way to cope with this kind of situation with the local amplitude spectrum is to consider the low-pass characteristics of elastic skin. By inserting a nonlinear mapping process, precipitous height patterns could be more realistically transferred to sensor response patterns. This procedure indeed produces a significant difference in the amplitude spectrum of the sensor responses between the original and sign-reversed version.

In this study, we used a 3-D printer with solid material to transcribe the carving pattern of the stimuli to the deformation of the skin finger as closely as possible. Clearly, the spatial pattern of the skin surface would be different from that of the stimuli with deformable or sticky surfaces. Further research is needed to investigate whether the current findings can be applicable with a wider range of materials. In addition, our experimental procedure limited the observers’ exploration time, which may potentially affect the texture discrimination performance. Whether there is any chance to access higher-order statistics through training or a very long exploration time is an issue awaiting further investigation.

### SUPPLEMENTAL DATA

Supplemental material is available at https://doi.org/10.6084/m9.figshare.13030907.v1.

### GRANTS

S.N. is supported by MEXT/JSPS KAKENHI Grant numbers JP15H05915 and JP20H05957.

### DISCLOSURES

The author S.K. is an employee of Nippon Telegraph and Telephone Communication Science Laboratories, which is a basic-science research section of Nippon Telegraph and Telecommunication. There are no patents, products in development or marketed products to declare.

### AUTHOR CONTRIBUTIONS

S.K., M.S., and S.N. conceived and designed research; S.K. and M.S. performed experiments; S.K. and M.S. analyzed data; S.K. and S.N. interpreted results of experiments; S.K. prepared manuscript; S.K., M.S., and S.N. edited and revised manuscript; S.K., M.S., and S.N. approved final version of manuscript.

### REFERENCES

1. Bensmaia S. Texture from touch. Scholarpedia J 4: 7956, 2009. doi:10.4249/scholarpedia.7956.
2. Klitzky RL, Pawlik D, Peer A. Haptic perception of material properties and implications for applications. Proc IEEE 101: 2081–2092, 2013. doi:10.1109/JPROC.2013.2248691.
3. Whitaker TA, Simões-Franklin C, Newell FN. Vision and touch: independent or integrated systems for the perception of texture? Brain Res 1242: 59–72, 2008. doi:10.1016/j.brainres.2008.05.037.
4. Lederman SJ, Abbott SG. Texture perception: studies of intersensory organization using a discrepancy paradigm, and visual versus haptic ventral streams. J Exp Psychol Hum Percept Perform 7: 902–915, 1981. doi:10.1037/0096-1523.7.4.902.
5. Heller MA. Texture perception in sighted and blind observers. Percept Psychophys 45: 49–54, 1989. doi:10.3758/bf03208032.
6. Johansson RS, LaMotte RH. Tactile detection thresholds for a single asperity on an otherwise smooth surface. Somatosens Res 1: 21–31, 1983. doi:10.3109/07367228309144538.
7. LaMotte RH, Srinivasan MA. Surface microgeometry: tactile perception and neural encoding. In: Information Processing in the Somatosensory System: Proceedings of an International Symposium at the Wener-Gren Center, edited by Francén O, Westman J. London: Macmillan Education UK, 1991, p. 49–58. doi:10.1007/978-1-349-11597-6_4.
8. LaMotte RH, Whitehouse J. Tactile detection of a dot on a smooth surface: peripheral neural events. J Neurophysiol 56: 1109–1128, 1986. doi:10.1152/jn.1986.56.4.1109.
9. Bergmann Tiest WM, Kappers AML. Haptic and visual perception of roughness. Acta Psychol (Amst) 124: 177–189, 2007. doi:10.1016/j.actpsy.2006.03.002.
10. Bergen JR, Adelson EH. Early vision and texture perception. Nature 333: 363–364, 1988. doi:10.1038/333363a0.
11. Julesz B. Visual pattern discrimination. IEEE Trans Inform Theory 8: 84–92, 1962. doi:10.1109/TIT.1962.1057698.
12. Freeman J, Ziemba CM, Heeger DJ, Simoncelli EP, Movshon JA. A functional and perceptual signature of the second visual area in primates. Nat Neurosci 16: 974–981, 2013. doi:10.1038/nn.3402.
13. Okazawa G, Tajima S, Komatsu H. Image statistics underlying natural texture selectivity of neurons in macaque V4. Proc Natl Acad Sci USA 112: E351–E360, 2015. doi:10.1073/pnas.1415146112.
14. Freeman J, Simoncelli EP. Metamers of the ventral stream. Nat Neurosci 14: 1195–1204, 2011. doi:10.1038/nn.2889.
15. Ziember CM, Freeman J, Movshon JA, Simoncelli EP. Selectivity and tolerance for visual texture in macaque V2. Proc Natl Acad Sci USA 113: E3140–E3149, 2016. doi:10.1073/pnas.1510847113.

16. Rosenholtz R. Capabilities and limitations of peripheral vision. Annu Rev Vis Sci 2: 437–457, 2016. doi:10.1146/annurev-vision-082114-035733.

17. McDermott JH, Simoncelli EP. Sound texture perception via statistics of the auditory periphery: evidence from sound synthesis. Neuron 71: 926–940, 2011. doi:10.1016/j.neuron.2011.06.032.

18. Portilla J, Simoncelli EP. A parametric texture model based on joint statistics of complex wavelet coefficients. Int J Comput Vis 40: 49–70, 2000. doi:10.1023/A:1007898919983.

19. Delhaye BP, Long KH, Bensmaia SJ. Neural basis of touch and proprioception in primate cortex. Compr Physiol 8: 1575–1602, 2018. doi:10.1002/cphy.cf00033.

20. Pruszynski JA, Johansson RS. Edge-orientation processing in first-order tactile neurons. Nat Neurosci 17: 1404–1409, 2014. doi:10.1038/nn.3804.

21. Bensmaia SJ, Denchev PV, Dammann JF 3rd, Craig JC, Hsiao SS. The representation of stimulus orientation in the early stages of somatosensory processing. J Neurosci 28: 776–786, 2008. doi:10.1523/JNEUROSCI.4962-07.2008.

22. DiCargo JJ, Johnson KO, Hsiao SS. Structure of receptive fields in area 3b of primary somatosensory cortex in the alert monkey. J Neurosci 18: 2626–2645, 1998. doi:10.1523/JNEUROSCI.02-06.2626.1998.

23. DiCargo JJ, Johnson KO. Spatial and temporal receptive properties of receptive fields in primate somatosensory area 3b: effects of stimulus scanning direction and orientation. J Neurosci 20: 495–510, 2000. doi:10.1523/JNEUROSCI.20-01-00495.2000.

24. Thakur PH, Fitzpatrick PJ, Lane JW, Hsiao SS. Receptive field properties of the macaque second somatosensory cortex: nonlinear mechanisms underlying the representation of orientation within a finger pad. J Neurosci 26: 13567–13575, 2006. doi:10.1523/JNEUROSCI.3990-06.2006.

25. Klatzky R, Lederman S. Object recognition by touch. In: Blindness and Brain Plasticity in Navigation and Object Perception, edited by Rieser JJ, Ashmead DH, Ebner F, Corn AL. New York: Psychology Press, 2007, p. 197–220.

26. Yau JM, Pasupathy A, Fitzgelder PJ, Hsiao SS, Connor CE. Analogous intermediate shape coding in vision and touch. Proc Natl Acad Sci USA 106: 16457–16462, 2009. doi:10.1073/pnas.0904186106.

27. Hollins M, Bensmaia S, Karol K, Young F. Individual differences in perceptual space for tactile textures: evidence from multidimensional scaling. Percept Psychophys 62: 1534–1544, 2000. doi:10.3758/BF03121154.

28. Hollins M, Faldowski R, Rao S, Young F. Perceptual dimensions of tactile surface texture: a multidimensional scaling analysis. Percept Psychophys 54: 697–705, 1993. doi:10.3758/bf0321795.

29. Bergmann Tiest WM. Tactual perception of material properties. Vision Res 50: 2775–2782, 2010. doi:10.1016/j.visres.2010.01.005.

30. Bergmann Tiest WM, Kappers AML. Analysis of haptic perception of materials by multidimensional scaling and physical measurements of roughness and compressibility. Acta Psychol (Amst) 121: 1–20, 2006. doi:10.1016/j.actpsy.2005.04.005.

31. Goodwin AW, John KT, Sathian K, Darian-Smith I. Spatial and temporal factors determiningafferent fiber responses to a gratings moving sinusoidally over the monkey’s fingerpad. J Neurosci 9: 1280–1293, 1989. doi:10.1523/JNEUROSCI.04-09-1280.1989.

32. Lederman SJ. Tactual roughness perception: spatial and temporal determinants. Can J Psychol/Revue canadienne de psychologie 37: 496–511, 1983. doi:10.1037/h0080705.

33. Lederman SJ, Taylor MM. Fingertip force, surface geometry, and the perception of roughness by active touch. Percept Psychophys 12: 401–408, 1972. doi:10.3758/BF03205850.

34. Sathian K, Goodwin AW, John KT, Darian-Smith I. Perceived roughness of a grating: correlation with responses of mechanoreceptive afferents innervating the monkey’s fingerpad. J Neurosci 9: 1273–1279, 1989. doi:10.1523/JNEUROSCI.04-09-1273.1989.

35. Taylor MM, Lederman SJ. Tactile roughness of grooved surfaces: a model and the effect of friction. Percept Psychophys 17: 23–36, 1975. doi:10.3758/BF03203993.
receptors are sensitive to spatial details on the scale of single fingerprint ridges. J Neurosci 41: 3622–3634, 2021. doi:10.1523/JNEUROSCI.176-20.2021.

80. Kuroki S, Watanabe J, Nishida S. Contribution of within- and cross-channel information to vibrotactile frequency discrimination. Brain Res 1529: 46–55, 2013. doi:10.1016/j.brainres.2013.06.036.

81. Kuroki S, Watanabe J, Nishida S. Integration of vibrotactile frequency information beyond the mecanoreceptor channel and somatotopy. Sci Rep 7: 2758, 2017. doi:10.1038/s41598-017-02922-7.

82. Mountcastle VB, Steinmetz MA, Romo R. Frequency discrimination in the sense of flutter: psychophysical measurements correlated with postcentral events in behaving monkeys. J Neurosci 10: 3032–3044, 1990. doi:10.1523/jneurosci.10-09-03032.1990.

83. Mountcastle VB, Talbot WH, Sakata H, Hyvarinen J. Cortical neuronal mechanisms in flutter-vibration studied in unanesthetized monkeys. Neuronal periodicity and frequency discrimination. J Neurophysiol 32: 452–484, 1969. doi:10.1152/jn.1969.32.3.452.

84. Talbot WH, Darian-Smith I, Kornhuber HH, Mountcastle VB. The sense of flutter-vibration: comparison of the human capacity with response patterns of mecanoreceptive afferents from the monkey hand. J Neurophysiol 31: 301–334, 1968. doi:10.1152/jn.1968.31.2.301.

85. Vickery RM, Gynter BD, Rowe MJ. Synchronous transmission between single slowly adapting type I fibres and their cuneate target neurones in cat. J Physiol 474: 379–392, 1994. doi:10.1113/jphysiol.1994.sp020030.

86. Skedung L, Arvidsson M, Chung JY, Stafford CM, Berglund B, Rutland MW. Feeling small: exploring the tactile perception limits. Sci Rep 3: 2617, 2013. doi:10.1038/srep02617.

87. Connor CE, Hsiao SS, Phillips JR, Johnson KO. Tactile roughness: neural codes that account for psychophysical magnitude estimates. J Neurosci 10: 3823–3836, 1990. doi:10.1523/jneurosci.10-03-08323.1990.

88. Hollins M, Risner SR. Evidence for the duplex theory of tactile texture perception. Percept Psychophys 62: 695–705, 2000. doi:10.3758/bf03206916.

89. Johnson KO, Hsiao SS, Yoshioka T. Neural coding and the basic law of psychophysics. Neuronascientist 8: 11–12, 2002. doi:10.1177/107385840200800207.

90. Pruszynski JA, Flanagan JR, Johansson RS. Fast and accurate edge orientation processing during object manipulation. eLife 7: e31200, 2018. doi:10.7554/eLife.31200.

91. Fitzgerald PJ, Lane JW, Thakur PH, Hsiao SS. Receptive field (RF) properties of the macaque second somatosensory cortex: RF size, shape, and somatotopic organization. J Neurosci 26: 6495–6495, 2006. doi:10.1523/JNEUROSCI.0501-05.2006.

92. Roland PE. Somatosensory detection of microgeometry, macroroughometry and kinesthesia after localized lesions of the cerebral hemispheres in man. Brain Res 434: 43–94, 1987. doi:10.1016/0006-899X(87)90018-x.

93. Apkarian-Stielau P, Loomis JM. A comparison of tactile and blurred visual form perception. Percept Psychophys 18: 362–368, 1975. doi:10.3758/BF03211213.

94. Craig AJ. Confusion matrix for tactually presented letters. Percep Psychophys 26: 409–411, 1979. doi:10.3758/BF03204167.

95. Loomis JM. On the tangibility of letters and braille. Percept Psychophys 29: 37–46, 1981. doi:10.3758/bf03198838.

96. Loomis JM. Analysis of tactile and visual confusion matrices. Percept Psychophys 31: 41–52, 1982. doi:10.3758/bf03206199.

97. Phillips JR, Johnson KO, Browne HM. A comparison of visual and two modes of tactual letter resolution. Percept Psychophys 34: 243–249, 1983. doi:10.3758/bf03209562.

98. Julesz B, Gilbert EN, Shepp LA, Frisch HL. Inability of humans to discriminate between visual textures that agree in second-order statistics revisited. Perception 2: 391–405, 1973. doi:10.1068/p202391.

99. Chubb C, Nam J-H, Bindman DR, Sperling G. The three dimensions of human visual sensitivity to first-order contrast statistics. Vision Res 47: 2237–2248, 2007. doi:10.1016/j.visres.2007.03.025.