Can a Computer Tell Differences between Vibrations?: Physiology-Based Computational Model for Perceptual Dissimilarity Prediction

Chungman Lim
chungman.lim@gm.gist.ac.kr
Gwangju Institute of Science and Technology
Gwangju, South Korea

Gunhyuk Park
maharaga@gist.ac.kr
Gwangju Institute of Science and Technology
Gwangju, South Korea

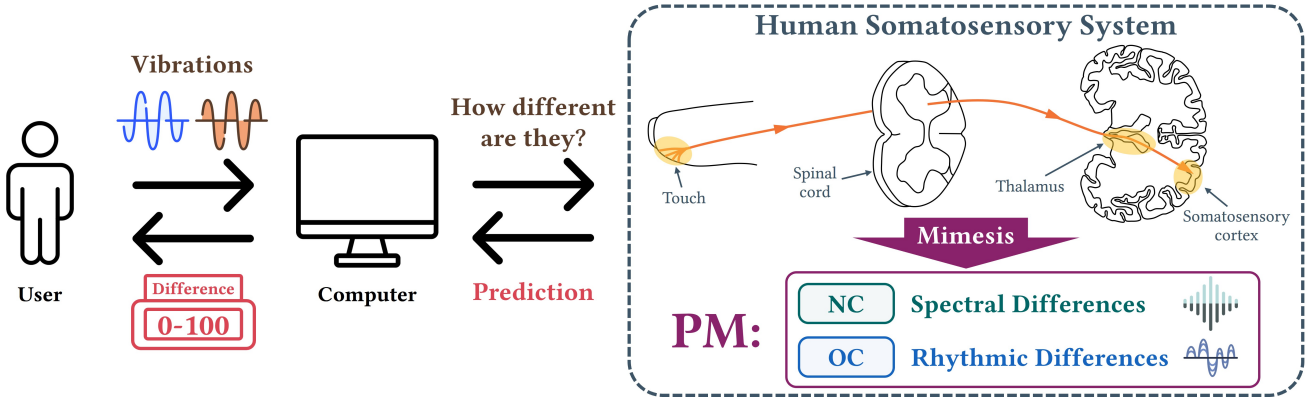


Figure 1: Physiology-based Model (PM) that predicts differences of vibrations mimicking the human somatosensory system.

ABSTRACT

Perceptual dissimilarities, requiring high-cost user ratings, have contributed to designing well-distinguishable vibrations for associated meaning delivery. Appropriate metrics can reduce the cost, but known metrics in vibration similarity/dissimilarity could not predict them robustly. We propose a physiology-based model (PM) that predicts the perceptual dissimilarities of a given vibration set via two parallel processes: Neural Coding (NC), mimicking the neural signal transfer, and One-dimensional Convolution (OC), capturing rhythmic features. Eight parameters were trained using six datasets published in the literature to maximize Spearman's Rank Correlation. We validated PM and six metrics of RMSE, DTW, Spectral/Temporal Matchings, ST-SIM, and SPQI in twelve datasets: six trained and six untrained datasets including measured accelerations. In all validations, PM's predictions showed robust correlations with user data and similar structures in perceptual spaces. Other baseline metrics showed better fit in specific datasets, but none of them robustly showed correlations and similar perceptual spaces over twelve datasets.

CHI '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9421-5/23/04.
https://doi.org/10.1145/3544548.3580686

CCS CONCEPTS

• **Human-centered computing** → **HCI theory, concepts and models**.

KEYWORDS

Perceptual Dissimilarity, Computational Model, Biomimetic Modeling, Vibrotactile Perception

ACM Reference Format:
Chungman Lim and Gunhyuk Park. 2023. Can a Computer Tell Differences between Vibrations?: Physiology-Based Computational Model for Perceptual Dissimilarity Prediction. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3544548.3580686

1 INTRODUCTION

With the recent growth of tactile information devices including mobile devices, wearables, and VR systems, vibrations have widely been used to inform events or states of a system by using well-discriminable patterns [23, 35, 53] and elevate the end-user's immersion with realistic tactile patterns [14, 16, 48, 51]. Among them, applications aiming at accurate information delivery required well-discriminable vibrations for providing warning [15, 17, 49, 63], navigation [31, 52, 66], and subjective feelings [44, 59, 61]. However, vibration patterns can be compiled with varying several parameters including amplitude, duration, envelope frequency, carrier frequency, and rhythm [6], and designers need to explore a large design space of vibrations to find the most dissimilar subset of

their interest. Therefore, haptic designers have authored vibrotactile patterns by varying salient vibration parameters [7, 60, 67], morphing signals [12], or selecting patterns that were labeled differently by users and designers [34, 54]. To find the most dissimilar subset, perceptual dissimilarities of designed patterns were estimated from pairwise rating [26, 40, 45, 60] or cluster-sorted rating [33, 34, 47, 60]. Both rating methods usually take more than an hour for the experiment completion by each participant.

The time and cost for running user studies would be greatly reduced if a metric can predict perceptual dissimilarities of vibrations. So far, the Haptics literature includes similarity/dissimilarity metrics of vibrations in various contexts. Root-mean-squared error (RMSE) contributed to measuring texture vibration similarities in spectral domain [36, 46, 56] while haptic researchers utilized Dynamic time warping (DTW) [9] in vibration morphing [12] or Tacton similarity measure [50]. Custom metrics of spectral (M_{sm}) and temporal (M_{tm}) matching were proposed to rate different dimensional accelerations of the same vibration [46]. Also, similarities between the original and compressed vibrations were measured in spectral similarity (SPQI) [40] and both temporal and spectral similarities (ST-SIM) [21, 22] while a method integrating SPQI/ST-SIM (VibroMAF) was also proposed [41]. The aforementioned metrics assessed the similarity between vibrations generated by their proposed algorithms in comparison to the reference vibrations that are usually captured from real interactions. However, no known metrics compared their calculations with user-rated dissimilarities and validated their derived perceptual spaces for perceptual analysis. So the literature still requires metrics of the perceptual dissimilarities for a vibration set to accelerate vibration pattern design.

Therefore, this work proposes a computational model that predicts perceptual dissimilarities matching the user ratings. We designed a physiology-based model, PM, by mimicking the human tactile system with physiological knowledge in the literature. Our model was designed with two independent and parallel approaches of Neural Coding (NC) and One-dimensional Convolution (OC) processes. NC was designed to capture spectral features by mimicking neural pathways from the skin's mechanoreceptor activation to the somatosensory cortex responses with six adjustable parameters. OC was more sensitive to rhythmic features by taking moving averages of vibrations. PM integrated two dissimilarity matrices generated from NC and OC with two weight parameters, and overall eight parameters were optimized using six training vibration sets published in the previous three papers with respective perceptual dissimilarities [25, 26, 45].

We compared PM and other six metrics with respect to the user-reported dissimilarities and perceptual spaces in three validation processes. The first validation datasets were the six published datasets used in the model training. PM showed the highest averaged Spearman's rank correlation (ρ = 0.79; the higher, the better) in dissimilarity and the lowest averaged alienation coefficient (0.26; the lower, the better) in structural similarity of perceptual space. The baseline six metrics of RMSE, DTW, M_{sm} M_{tm}, ST-SIM, and SPQI showed high correlations ($\rho > 0.5$) in frequency mixing patterns, but comparably low correlations in the patterns varying modulated frequency and both amplitude and carrier frequency. In the second validation, we collected two datasets by running a user study that compares an untrained dataset of 12 vibrations by recruiting 20 participants. The first dataset was an intended pattern set (IPS) varying all trained vibration parameters of amplitude, modulated frequency, and mixing ratios of two carrier frequencies. The second dataset was a measured pattern set (MPS) collected from a voice-coil actuator held by two fingers in playing IPS. The final validation compared two IPS and two MPS of 14 vibrations [1] that vary an untrained parameter of rhythm. The second and final validation consistently showed that PM predictions were strongly correlated with user data, showed lower alienation coefficients, and captured similar implications from the perceptual spaces.

In this paper, we suggest a computational model, PM, as a metric candidate for pairwise perceptual dissimilarity using physiological knowledge. We contribute as follows:

- Results showing that PM robustly predicted user-rated perceptual dissimilarities in both designers' intended patterns (IPS) and noise-including measured acceleration patterns (MPS).
- Perceptual spaces derived from PM predictions and user-rated perceptual dissimilarities showed better structural similarity and were able to capture similar implications. Therefore, vibrotactile pattern designers can flexibly probe the user perception by varying vibration pattern sets with PM in minutes.

2 RELATED WORK

Haptic designers have authored vibrotactile patterns to find the best subset for mapping appropriate information of their interests. One common approach is obtaining perceptual dissimilarities of vibration candidates from user studies and deriving corresponding perceptual space to select the most dissimilar subset. Known metrics have contributed to estimating similarity or dissimilarity of vibrations, but the relevant perceptual space analysis has not been covered by those metrics yet. This section reviews the vibrotactile pattern design, perceptual space analysis, and objective metrics that contributed to vibration similarity and dissimilarity.

2.1 Vibrotactile Pattern Design

In the tactile interface, perceptually different vibrotactile patterns are mandatory to associate meanings with them, so a user easily identifies the incoming pattern. Designers, therefore, have compiled such patterns by combining vibrotactile parameters of larger than differential thresholds in a single physical parameter, for example, amplitude and frequency [25], envelope modulation frequency [45], superimposing ratios of two vibrations [26, 32], and duration [67]. Meanwhile, abstract but comparably salient parameters of rhythm [1, 7, 60] or waveform patterns [47] were contributed to design distinct patterns as Tactons [6].

Designers then select the best subset from the compiled vibrotactile patterns for associating the meanings of target applications. An intuitive way is running a user study of pattern identification [7, 8] and selecting highly accurate patterns, but it usually needs a careful design considering the learning process of participants. A common approach for finding the subset is collecting pairwise differences of perceived vibrations and applying multi-dimensional scaling (MDS) on them to visualize the perceptual space [25, 26, 45]. The MDS finds the coordinates that minimize distance errors in the reduced

dimensions, therefore pairwise dissimilarities of vibrations can be visualized in easy-to-understand 2D or 3D perceptual spaces. The derived perceptual space is interpreted as how users recognize the given vibrotactile patterns in the relative structure and designers find relevant implications with the design parameters [34, 60].

2.2 Perceptual Dissimilarity Rating in Haptics

The MDS is a powerful visualization method and widely used for understanding the salient texture parameters [24] to finding the dissimilar vibration subsets [1, 60] by pairwise rating. In the texture classification in real interaction, Hollins et al. measured pairwise dissimilarities of 17 textures and visualized their perceptual space [24], and derived a conclusion that rough/smooth and soft/hard dimensions were prominent in texture classification. Okamoto et al. applied MDS to derive perceptual spaces of haptic texture and vibrotactile patterns and proposed a mapping technique to provide vibrations similar to the target texture [43]. Vardar et al. applied non-metric MDS in multi-modal comparison [64], in detail, compared haptic and visual perceptual spaces from pairwise dissimilarities of real surface stimuli, and concluded the equal contribution of both modalities to surface haptic information. Also, many researchers visualized the perceptual spaces to find salient vibrotactile parameters [25, 26, 45]. However, MDS requires dissimilarities of all vibration pairs, so the number of vibrations has been limited in the pairwise rating due to the experiment duration.

In order to control the temporal cost of pairwise dissimilarity rating, Ward suggested a cluster-sorted rating [65]. This method indirectly calculates dissimilarities from multiple classification processes, so the resolution of dissimilarities and the temporal cost are controlled by the number of sessions and clusters. Due to its adjustable aspects, cluster-sorted ratings were applied to 30 and 36 vibration sets (1 and 1.75 hours each) [34], 30 vibrations (duration not specified; 1 hour assumed from compensation money) [33, 47], and 84 vibrations (2 hours) [60]. Some of them validated the similarities between cluster-sorted and pairwise ratings [33, 34]. Although cluster-sorted rating provided a means of controlling temporal cost, it still requires running user studies more than an hour.

Despite diverse use cases in MDS analysis, both ratings come from relative comparisons, and users' responses could differ if the given vibration patterns change. In other words, haptic designers should run another user study to accurately select a dissimilar subset, and it reduces the design flexibility.

2.3 Metrics in Vibration Similarity and Dissimilarity

Dissimilarity, in other words, distance metrics are essential for classification, clustering, optimization, and evaluation in various fields. Many researchers have proposed and analyzed dissimilarity metrics, such as L_p distance and entropy-based dissimilarity, and Cha reported mathematical distance metrics rigorously [10].

In vibration comparison metrics, root-mean-squared error (RMSE) is one of the most general methods based on Euclidean distance. RMSE is defined as a squared average error of two signals in its power estimation, which gives non-negative values. Haptic researchers have adopted the RMSE in spectral similarity calculation to custom metrics including ST-SIM/SPQI that use subband-wise RMSE in the spectral domain for comparing original and compressed vibration signals [21, 22, 40]. VibroMAF optimized summing weights of ST-SIM, SPQI, and normalized signal-to-noise ratios for assessing the vibrotactile similarity after signal compression [41]. In vibrotactile dimensional reduction from 3D to 1D of the same vibration signal [46], spectral matching (M_{sm}) also calculated the sub-band RMSE using 10% spectral moving window while temporal matching (M_{tm}) calculated cross-correlation to the temporal signals. Dynamic time warping (DTW), suggested by Berndt et al., was designed to measure dissimilarities and detect similar patterns in two temporal series signals [3]. It calculates a distance by generating a cost matrix consisting of differences between all components of two temporal series, then finds a warping path that minimizes total cost. DTW has been used mostly in auditory stimuli comparison including music [38] and speech [9], and haptic researchers utilized it in finding similar patterns to a target Tacton [50].

So far, haptic researchers have proposed several dissimilarity or similarity metrics between vibrotactile signals, most of which were modeled by mathematical signal processing. Some attempts also reflected the perceptual characteristics of human including detection threshold [21, 22, 41] or differential threshold [46]. To the best of our knowledge, no researchers have compared the metric calculations with the user-rated perceptual dissimilarities yet.

3 PHYSIOLOGY-BASED MODEL: OVERVIEW

We designed a physiology-based model (PM) (in Figure 2) to predict the pairwise perceptual dissimilarities of a given vibration set with the same duration. This model was designed to emphasize different features of each vibration in the set; Neural Coding (NC) simulates neural current transfer for spectral feature extraction, and 1D Convolution (OC) compares rhythmic features by taking the envelope of input vibrations through a moving average filter. NC and OC calculate predicted brain waveforms and envelope feature waveforms from an input pattern set. The estimated pairwise dissimilarity matrix in each set was calculated by Root-Mean-Squared-Error (RMSE) and Dynamic Time Warping (DTW) respectively, and PM applied a weighted sum of the estimated matrices to output a predicted dissimilarity matrix. To increase the prediction accuracy, we trained two summing weights, colored in blue in Figure 2, using six dissimilarity matrices published in three papers [25, 26, 45].

While OC adopts 1D convolution for rhythmic feature extraction, NC is designed to mimic the process of how a human delivers an external tactile stimulus to the brain for haptic perception. For efficient design, we divided NC into 1) mechanoreceptor excitation by external vibrotactile stimulus and 2) neural signal transfer from the excited mechanoreceptors to the somatosensory cortex area. The mechanoreceptor excitation is modeled based on the four-channel theory that human perceives the tactile stimulus through four different types of receptors in the human skin, and the neural signal transfer is modeled to predict neural spikes over sub-neural networks and the brain [11, 30]. The following two sections cover the details.

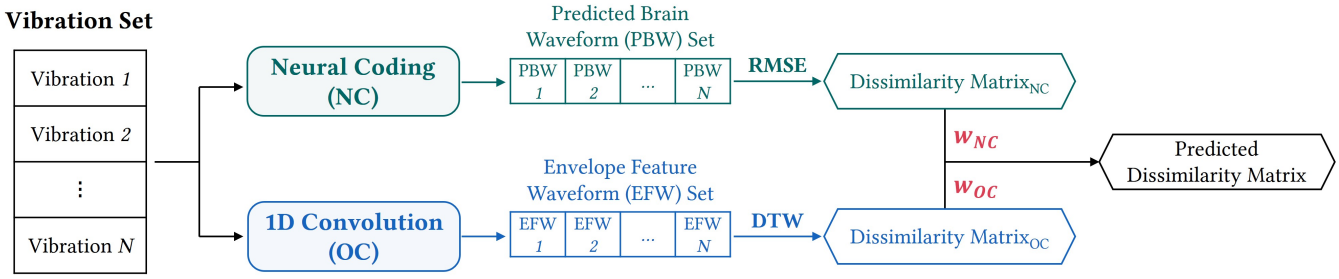


Figure 2: Physiology-based model that predicts a dissimilarity matrix from 'Neural Coding (NC)' and 'One Dimensional Convolution (OC)' processes. The red letters are summing weights and trained by six published perceptual dissimilarity data.

4 NEURAL CODING

Tactile stimulus applied to the human skin physically moves the contacted body site and propagates along the skin and bone. Four mechanoreceptors of Meissner Corpuscle, Pacinian Corpuscle, Merkel Disk, and Ruffini Ending are excited by the different bandwidths of the stimulus spectrum, and the mechanoreceptors exceeding their excitation threshold convert the physical stimuli to neural currents. The electrical signals on the nerve terminals in the skin are conveyed to the central nervous system through a series of relay nuclei to the thalamus in the brain. At the hierarchical sensory processing stage between the thalamus and the somatosensory cortex, the signals are integrated into more complex sensory information. At last, a specific tactile sensation of the stimuli occurs, and a human perceives the stimulus [30].

Our Neural Coding process mimics the aforementioned process from the neural activation at mechanoreceptors to the perception in the brain (in Figure 3). A well-defined contact scenario is required for accurate prediction, and our model assumes that *input vibration, represented in g (gravity acceleration), is given with a circular probe of 0.5 mm radius at the fingertip and four mechanoreceptors are aligned at the center location* to utilize the known parameters in tactile perception [28]. The following subsections are threefold to describe them in detail: mechanoreceptor activation, neural spike generation, and neural network simulation.

4.1 Mechanoreceptor Activation via Vibration Propagation

External vibrotactile signals propagate along the skin and excite tactile receptors under the skin. If the energy exceeds the receptor's activation threshold, neural spikes are generated at the receptor. We designed a two-step process to model the thresholded activation in NC: the vibration propagation using band-pass filtering with four-channel theory (Section 4.1) and neural spike generation (Section 4.2) as in Figure 4.

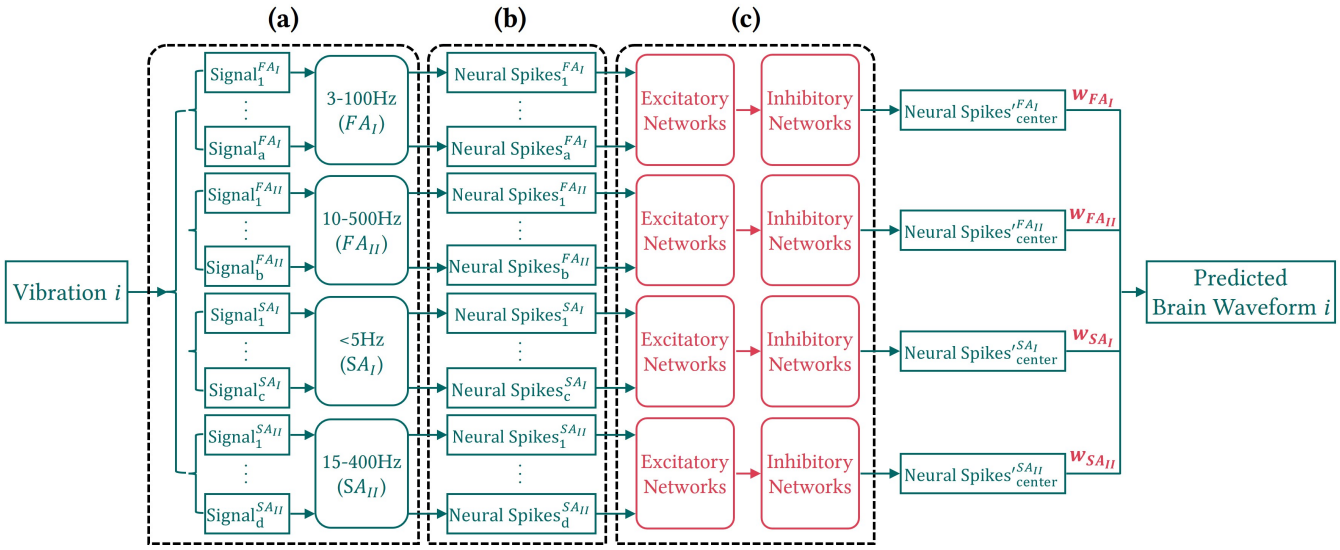


Figure 3: A diagram of Neural Coding (NC) process. The neural networks and weight parameters colored in red were trained with published data. (a) Simulation of vibration propagation and relevant mechanoreceptor activation with band-pass filtering (b) Neural spike generation (c) Simulation of neural current transfer via simplified excitatory and inhibitory networks.

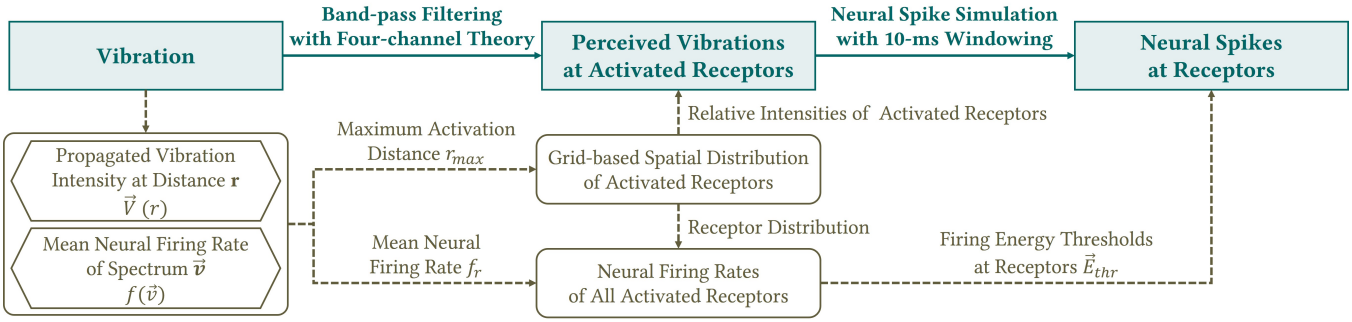


Figure 4: Neural spike generation process for propagated vibration on a single mechanoreceptor type

Propagated Vibration Intensities of Activated Receptors - We first modeled how the vibration propagates over the skin. In our assumption, four mechanoreceptors are aligned at the stimulation center, and nearby receptors are excited with propagated vibrations. Here, we defined the maximum activation distance r_{max} as the boundary where no mechanoreceptors fire any neural currents by the propagated vibrations. Thanks to the literature, Johnson reported firing rates of receptors and vibration propagation amplitudes at a fingertip with the 0.5-mm probe stimulation [28]. The neural firing rates of mechanoreceptors located within a 1-mm margin of the stimulator's edge (hot zone) were the same as the centered receptor. Also, the vibration amplitude drops off by the square as the distance to the hot zone increases. We defined the propagated spectral amplitudes of vibration as:

$$\vec{V}(r) = \begin{cases} \vec{v}, & r \leq 1.5mm \\ \left(\frac{1.5}{r}\right)^2 \times \vec{v}, & r > 1.5mm \end{cases} \quad (1)$$

where $\vec{V}(r)$ are the spectral amplitudes of propagated vibration at r, $\vec{v}$ are the spectrum amplitudes at the center, and r is the distance from the stimulation center.

We then chose activated receptors by the propagated vibrations and their spectral sensitivities. Muniak et al. reported that a linear relationship exists between the neural firing rate and a single spectral intensity at the receptor [39], and we linearly extended it to the full-spectral equation as:

$$f(\vec{V}) = \sum_{i=1}^{n} \alpha_i \left(\log V_i - \beta_i\right) \quad (2)$$

where $f(\vec{V})$ is the mean firing rate of the receptor, $\log V_i$ is a log-transformed intensity of propagated vibration, n is the number of spectral components of the propagated vibration, $\vec{\alpha}$ are the linear proportions between mean firing rate and the log-transformed intensities, $\vec{\beta}$ is the log-transformed absolute threshold of the receptor (displacement, $log(\mu m)$ peak), and subscriptions of i represent i-th spectral component. If the input vibration is a single sinusoid, $n = 1$, then the equation is simplified to $f(v) = \alpha\,(log(v) - \beta)$ which is a common form of Steven's power law [57].

To build $\vec{\alpha}$ and $\vec{\beta}$ for 1-500 Hz spectrum, we adopted the empirical study data estimated from macaque monkeys and human subjects. First, we adopted the proportions of the mean firing rate from the monkeys' data [39] because of the absence of human data. The monkey's tactile system includes no Ruffini Endings, and Merkel Disk and Ruffini Ending in human tactile system are reported as sharing their pathway to the somatosensory cortex [29, 58]. By the prior findings, we regarded the two receptors share similar proportions in both Merkel Disk and Ruffini Ending. To build $\vec{\alpha}$, we applied quadratic extrapolation to the values of Meissner Corpuscle, Pacinian Corpuscle, and Ruffini Ending. In contrast, cubic extrapolation was applied to Merkel Disk's values because it is insensitive to high-frequency bandwidth vibration.

Second, Bolanowski et al. reported an absolute threshold plot of spectral components in [4]. Because the figure did not cover the full bandwidth of each receptor, we applied quadratic extrapolation to the manually measured data points and built $\vec{\beta}$ between 1-500Hz of each receptor. All approximations were carefully selected to reflect their sensitivity to frequency bandwidths as shown in Figure 5.

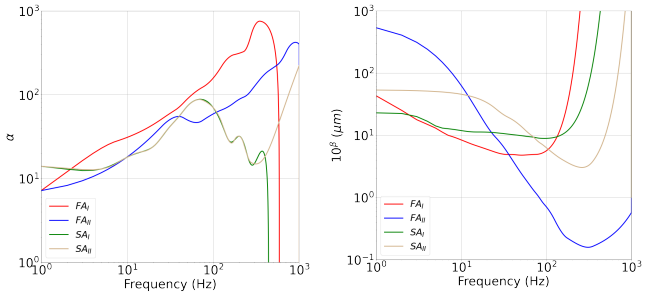
Then we found the maximum activation distance r_{max} at which receptors do not activate from Equation 1 and 2 as:

$$r_{max} = 1.5 \times 10^{\frac{\sum_{i=1}^{n} \alpha_i (\log v_i - \beta_i)}{2\sum_{i=1}^{n} \alpha_i}} \quad (3)$$

where n is the number of spectral components of the vibration and v_i is the spectrum amplitudes of input vibration at the probe center. Because each mechanoreceptor type has different perceptual characteristics, r_{max} is calculated for each receptor type. For a single sinusoid ($n = 1$), r_{max} becomes $1.5 \times \left(\frac{v}{10^{\beta}}\right)^{\frac{1}{2}}$ where v is the amplitude of the stimulus at the stimulation center.

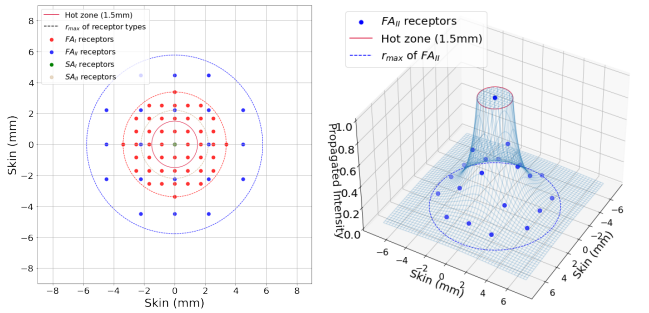
Next, we generated an activated receptor grid for each mechanoreceptor using the r_{max} and the mechanoreceptor density on the fingertip in Table 1 where innervation density is an estimate of sampled density in the median nerve [27, 62]. In order to simplify the calculation process with plausible spatial distribution estimation, we premised two requirements. First, at least one receptor activates even if r_{max} is zero. Second, receptors are uniformly distributed in r_{max} with grid intervals defined by $\sqrt{\frac{1}{density}}$. With these assumptions, the activated receptor grids were derived and the number of activated receptors, $N_{receptor}$, was counted.

Finally, each activated mechanoreceptor in the grid is assigned with propagated vibrations using Equation 1 with the input vibration $\vec{v}$ and distance r. Figure 6a shows the activated receptor grids



(a) Parameters $\vec{\alpha}$ **(b) Parameters $\vec{\beta}$**

Figure 5: Extrapolated parameters of (a) linear proportion of mean firing rates $\vec{\alpha}$ and (b) log-transformed absolute thresholds $\vec{\beta}$.



(a) Grid-based map of activated receptors **(b) Propagated intensities (3D) of Pacinian Corpuscles**

Figure 6: Exemplar plots for activated mechanoreceptors to a 150-Hz vibration modulated with an 80-Hz sinusoid on the fingertip. (a) Grid-based representations of activated mechanoreceptors. (b) Relative intensities of activated Pacinian Corpuscles in 3D.

of four mechanoreceptor types in spatial domain and Figure 6b describes propagated intensities of activated Pacinian Corpuscles when a fingertip is stimulated with a 150-Hz sinusoid vibration modulated with an 80-Hz sinusoid.

Band-pass filtering with four-channel theory - NC structure is designed to follow the four-channel theory that relates the human tactile neural system to four mechanoreceptors [4, 19]. Four mechanoreceptors are sensitive to different frequency ranges (Meissner Corpuscle (FA_I): 3-100Hz, Pacinian Corpuscle (FA_{II}): 10-500Hz, Merkel Disk (SA_I): <5Hz, Ruffini Ending (SA_{II}): 15-400Hz) [11]. From the property, we designed four Butterworth band-pass filters (Order = 10) of relevant passing frequencies. At each activated receptor, the propagated vibration waveform is passed with its corresponding filter and regarded as a perceived vibration waveform.

4.2 Neural Spike Generation

In the relevant literature [4], Meissner and Pacinian Corpuscles are related to the fast adapting (FA) neural channels for their activation sensitivity to the stimulus transients. The other two receptors are related to the slowly adapting (SA) channels because relevant neural fibers fire during the constant pressure [37]. Therefore, the perceived vibration waveforms to Meissner and Pacinian Corpuscle channels were differentiated to get their first derivatives which amplify the intensities. We normalized the waveforms to have the same maximum intensity before the differentiation for faster model training.

We applied the thresholded activation model for spike generation at each neural channel. In Section 4.1, we calculated the mean neural firing rate f_r of a channel in Equation 2 with its activated receptors $N_{receptor}$. Therefore, the total neural fires in the channel with T_r was calculated by their multiplication. Then the T_r was distributed by the logarithmic ratios of propagated intensities to all receptors in the channel, and each receptor sets its own neural firing threshold $\vec{E}_{thr}$ by dividing the energy of perceived vibration with its firing rate.

Then we designed the cumulative neural spike generation with a 10-ms window adopting the known temporal gap threshold for successive vibration stimuli discrimination [18]. In each 10-ms windowed vibration, the accumulated energy is summed with the previous remained energy in the previous window and divided with $\vec{E}_{thr}$ to count the number of neural fires. This process finally provides the impulse count plot in every 10-ms, and replicates to all receptors in four mechanoreceptor channels. In this process, NC prepares the simulated neural spike data in temporal domain for all activated receptors.

4.3 Neural Network Simulation between Sub-neural Network and Somatosensory Cortex

This section describes how the simulated neural spikes are transferred to the brain to make sensation in Figure 3 (c). As in [30], the neural currents are amplified or suppressed by the passing networks. The excitatory process occurs in every neural network and amplifies the neural currents by converging them. In contrast, the higher-order networks except sub-neural networks include the inhibitory process that suppresses the currents by interfering with the activation of neural cells with three pathways of feed-forward,

Table 1: Density of mechanoreceptors

Receptor	Meissner Corpuscle (FA_I)	Pacinian Corpuscle (FA_{II})	Merkel Disk (SA_I)	Ruffini Ending (SA_{II})
Innervation Density ($units/mm^2$)	1.4	0.2	0.7	0.1
Grid Interval (mm)	0.845	2.236	1.195	3.162

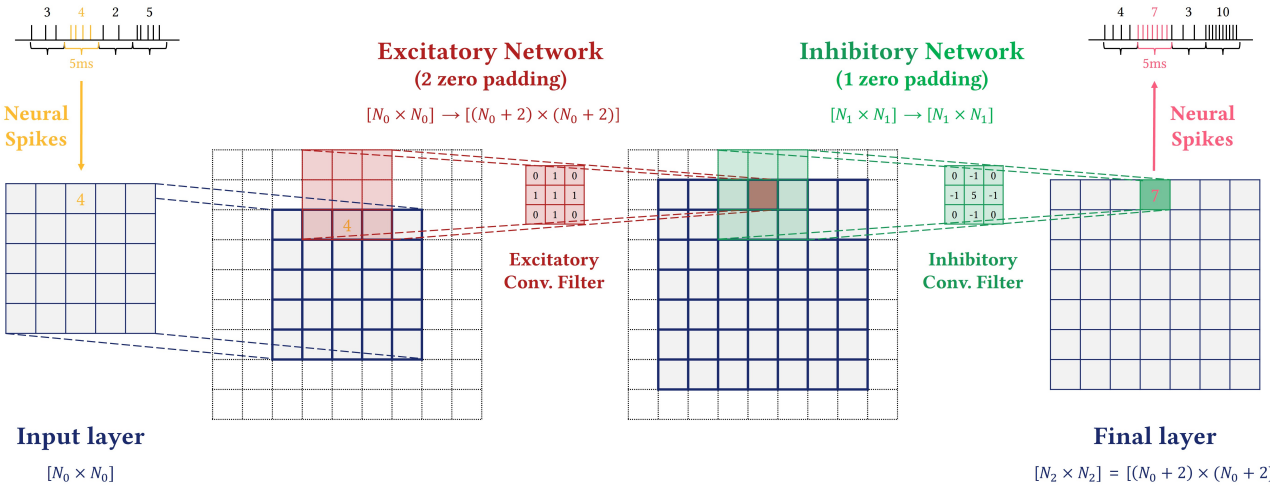


Figure 7: An exemplar diagram of simplified sensory networks by abstracting excitatory and inhibitory processing in human neural networks. This example shows single-layered excitatory and inhibitory networks.

feedback, and distal inhibitions. We implemented a unidirectional and hierarchical network including both processes where the inhibition process is implemented using **feed-forward inhibition**, which simply projects the neural currents from the previous interneurons to the connected neurons in the next layer. Each layer of the network was designed as a 2D matrix where each element represented a relay neuron, and the coordinate was aligned at the center. The initial layer matrix (D_0) enclosed all grid points of activated mechanoreceptors and aligned with the spatial distribution estimated in Section 4.2 as $N_0 \times N_0$

$$N_0 = \left\lfloor \frac{R_{max}}{Grid\ Interval} \right\rfloor \times 2 + 1 \tag{4}$$

where additional one represents a center neuron that locates at $(\lceil \frac{N_0}{2} \rceil, \lceil \frac{N_0}{2} \rceil)$. Here, the grid-based receptors directly transfer their activated currents to the relevant neurons at the initial layer without an excitatory process. In the simulation, the number of neural spikes is counted every 5-ms that adopts the conduction delay between the neural network layers [2]. Those spikes are projected to the relevant neurons in the initial layer.

In order to represent the connection of each relay neuron in both excitatory and inhibitory processes, the connected neuron set S_c in the next layer was modeled as a cross-shaped receptive field to the current neuron located at (x, y):

$$\begin{aligned} S_c \quad = \quad & \{(x-1,y),\ (x,y-1),\ (x,y),\ (x,y+1),\ (x+1,y)\}, \\ & \text{for } 1 \leqq x, y \leqq N_{i+1} \end{aligned} \tag{5}$$

In Neurophysiology, the excitatory process is known to filter out noise or sporadic activity relayed to a neuron by emphasizing strongly repetitive neural currents from fibers or superpositioning multiple neural currents. The inhibitory networks, in contrast, negatively affect the connected boundary neurons while emphasizing the connected neurons around its relay neuron's center [30]. We reflected those characteristics and S_c in our model as 3×3 convolutional filters of E and I as:

$$E = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad I = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix} \tag{6}$$

We processed the excitatory and inhibitory processes as convolutional filtering from D_{i-1} to D_i matrices, therefore both process filters work on the surrounding neurons. We assumed the surrounds of activated neurons as not excited and zero-padded in our calculation. In detail, excitatory networks increased the number of dimensions as the network order increased ($N_i = N_{i-1} + 2$) while the inhibitory networks kept the dimension of the previous network ($N_i = N_{i-1}$) and they required zero-padded additional boundaries for spatial convolution.

After projecting the neural spike numbers onto the initial layer, each layer applied the excitatory filter (E) by spatial convolution over its matrix Di to the next layer D_{i+1} in the excitatory networks. In each layer, we capped the maximum neural spikes $5ms \times sampling\ rate$ that implies the neuron activates at every sampling point. The inhibitory networks also applied the same spatial convolution but with the inhibitory filter (I), and capped its minimal spikes as zero.

Our model simulates neural spike generation at every 5-ms by simply mimicking the excitatory and inhibitory networks. The networks consist of L_E and L_I layers of excitatory and inhibitory networks respectively, which follow the human tactile system's characteristic that the sub-neural networks do not include the inhibitory process. The dimension of the final layer matrix is $N_{last} = N_0 + 2 \times L_E$, and neural spike numbers at the center neuron of the last layer were converted as the evenly-distributed temporal spikes by a winner-take-all strategy, which ensures that only one of two or more competing responses is transmitted [55]. Figure 7

shows an example of the above procedure on single-layered excitatory and inhibitory networks, and this filtering replicates for all relay neurons.

The Neural Coding process of each mechanoreceptor gives four central neural spike signals and we assumed that the brain integrates them with specific weights to predict the perceived brain waveform. We calculated RMSE of predicted brain waveform (PBW) pairs for estimating perceptual dissimilarities, and we trained the number of layers (L_E, L_I) and the summing weights of four neural channels ($w_{FA_I}, w_{FA_{II}}, w_{SA_I}, w_{SA_{II}}$) using pre-published six vibration sets with corresponding perceived dissimilarity matrix. Each variable was trained with 0-4 layers for L_E and L_I and assumed the identical Neural Coding model was applied in four mechanoreceptors for training cost reduction. Moreover, four weights were trained with 0 to 1 ratios with 0.1 intervals.

5 1D CONVOLUTIONAL PROCESS

As the spectral features contribute to the qualitative vibration perception, one of the most representative features in vibration differentiation is the rhythm defined as a repeated monotone pattern of variable-length notes. The most common definition of rhythm is a pattern changing length, number, or gaps between notes in vibration [60] and the primary components of rhythmical vibration were note length and evenness. Rhythmic vibration patterns, therefore, give high accuracy in vibration discrimination, however, sometimes it dominates haptic perception and masks other parameters' effects including vibration amplitude in rhythmic vibrations [1].

We designed a 1D convolutional process, OC, to obtain rhythmic differences. Therefore, PM finally integrated them with the spectral difference assessed by the Neural Coding (NC) to find a balanced perceptual difference.

5.1 Pre-processing

OC was intended to detect temporal rhythm differences, and we utilized temporal summation of tactile perception [19] by filtering out spectral components. We selected a moving average filter with a 10-ms window as a temporal summation while preserving temporal gaps larger than the tactile gap threshold [19]. In OC, we took absolute values of input vibration and applied the moving average filter to emphasize the rhythmic features. Finally, envelope feature waveforms (EFW) were generated from a given vibration set.

5.2 Dissimilarity Metric

OC used Dynamic Time Warping (DTW) for measuring dissimilarities of the EFW set. We applied the one-to-many match method for the DTW algorithm to maximize the similarity of two different time-series waveforms [20]. By repeating the process to the EFW set, OC generates a dissimilarity matrix reflecting rhythm-sensitive perception.

5.3 Physiology Model Training

Two dissimilarity matrices, calculated from NC and OC for a given vibration set, were integrated into one dissimilarity matrix and considered as a predicted pairwise dissimilarity matrix. For the integration, the dissimilarities in each matrix were converted to ordinal numbers because the estimated values were usually regarded as ordinal [64]. The converted matrices were summed with two weights (w_{NC}, w_{OC}).

At the training, Physiology-based Model (PM) was trained to maximize ρ between a predicted dissimilarity matrix and the reported matrix in papers [25, 26, 45], and all combinations of two matrix weight parameters and six NC parameters ($L_E, L_I, w_{FA_I}, w_{FA_{II}}, w_{SA_I}, w_{SA_{II}}$) were tested. Except L_E and L_I, six other parameters were trained in 0.1 step-size. The eight parameters were globally optimized to maximize the averaged Spearman's ρs calculated with the six training datasets, and trained parameters were $L_E = 3$, $L_I = 1$, $w_{FA_I} = 0.2$, $w_{FA_{II}} = 0.7$, $w_{SA_I} = 0$, $w_{SA_{II}} = 0.5$, $w_{NC} = 0.5$, and $w_{OC} = 0.5$.

Interestingly, the fitted parameter values were close to the human perception system. First, L_E is larger than L_I, and it reflects that inhibition occurs only in the higher-order networks in the brain. Second, weights of four mechanoreceptors were ordered as $w_{FA_{II}} > w_{SA_{II}} > w_{FA_I} > w_{SA_I}$. To the best of the authors' knowledge, Pacinian Corpuscle (FA_{II}) is the most sensitive mechanoreceptor for vibration perception and responds to the widest range of frequency components. Conversely, Merkel Disk (SA_I) is sensitive to point stimulation, edges, curvature, and fine details, so it is known for less contributing to vibration perception [42]. Finally, w_{NC} and w_{OC} were identically fitted which implies that the spectral and rhythmic features of vibration equally contribute to the vibration differentiation. This equal contribution resulted in higher correlations than using individual NC or OC dissimilarity matrix was compared with user data.

Validation Using Training Vibration Sets - We first calculated pairwise dissimilarities of vibrations using PM and the six baseline metrics of RMSE, DTW, M_{sm}, M_{tm}, ST-SIM, and SPQI in each of the six trained datasets. VibroMAF, which is a weighted-sum model for maximizing prediction accuracy with user similarities, requires multiple metrics to balance their outcomes. This methodology needs in-depth validation in both the metric selection and the learning process using reasonable datasets, which is out of our scope, so we focused on the six metrics in this work. As a prediction performance measure, Spearman's correlation ρ was estimated between every pair of the calculated and the user-rated perceptual dissimilarity sets. PM robustly showed $\rho > 0.5$ in all six datasets and the highest averaged correlation ($\rho = 0.79$) as reported in Table 2, where 0.3 and 0.5 are regarded as moderate and good correlation indices [13]. In the worst case of (Mid-High) Exp I [26], PM showed the lowest value of 0.84 among the seven metrics. The other baseline metrics also showed ρ values larger than 0.5 in three vibration sets varying frequency mixing ratios [26]. However, all baseline metrics showed $\rho < 0.5$ in the patterns varying modulated frequency [45] and M_{tm} predictions could not provide a correlation. Among the baselines, only DTW showed a good correlation ($\rho = 0.74$) to user data in the patterns varying amplitude and frequency [25].

Also, we applied non-metric MDS to visualize the 3D perceptual space from the dissimilarities. As an objective measure of the structural similarity between perceptual spaces, we calculated alienation coefficient K [5, 60] where $K = 0$ means the perfect similarity. As a representative example, we picked a perceptual space of [45] and plotted derived perceptual spaces of all metrics in Figure 8. The reported perceptual space (Figure 8a) implies that high-frequency

Table 2: Summary of the twelve vibration sets in Spearman's ρ and alienation coefficient K. The blue letters represent the best value among the metrics for a vibration set in each row, and the averages are emphasized in bold. Measured Pattern Set (MPS) is the accelerations measured from Intended Pattern Set (IPS), so MPS includes noise from contacts between the skin and the tactile device.

Vibration Set	Vibrotactile Parameter	Case	Source of a vibration set (# of vibrations, pairs)	Spearman's ρ (Alienation Coefficient K)						
				PM	RMSE	DTW	Three-to-one Algorithm		ST-SIM	SPQI
							M_{sm}	M_{tm}		
Trained	Trained	IPS	Exp *I* [45] (8 vibrations, 28 pairs)	0.78 (0.25)	0.16 (0.84)	-0.13 (0.60)	0.03 (0.56)	-★ (0.56)	-0.03 (0.60)	-0.07 (0.55)
			Exp *I* [25] (14 vibrations, 91 pairs)	0.63 (0.32)	0.29 (0.44)	0.74 (0.23)	0.26 (0.44)	-0.13 (0.56)	-0.06 (0.54)	-0.10 (0.49)
			Exp *I* [26] Low-Mid (7 vibrations, 21 pairs)	0.93 (0.22)	0.91 (0.22)	0.80 (0.30)	0.89 (0.25)	0.93 (0.20)	0.90 (0.17)	0.94 (0.20)
			Exp *I* [26] Mid-High (7 vibrations, 21 pairs)	0.84 (0.26)	0.96 (0.12)	0.91 (0.16)	0.96 (0.18)	0.96 (0.15)	0.89 (0.19)	0.95 (0.18)
			Exp *I* [26] Low-High (7 vibrations, 21 pairs)	0.93 (0.19)	0.91 (<.01)	0.86 (0.23)	0.91 (0.23)	0.97 (0.13)	0.93 (0.17)	0.93 (0.19)
			Exp *II* [26] (15 vibrations, 105 pairs)	0.63 (0.33)	0.45 (0.41)	0.50 (0.64)	0.45 (0.40)	0.36 (0.41)	0.36 (0.42)	0.32 (0.47)
	Average for 6 trained vibration sets			**0.79 (0.26)**	**0.61 (0.34)**	**0.66 (0.36)**	**0.58 (0.34)**	**0.56 (0.33)**	**0.53 (0.35)**	**0.55 (0.34)**
Untrained	Untrained	IPS	User Study (12 vibrations, 66 pairs)	0.63 (0.42)	0.63 (0.40)	0.24 (0.63)	0.61 (0.41)	0.79 (-◇)	0.27 (0.56)	0.71 (0.32)
		MPS		0.55 (0.44)	0.20 (0.58)	0.25 (0.63)	0.56 (0.42)	0.80 (0.23)	0.28 (0.50)	0.77 (0.35)
	Untrained	IPS	Exp *I* [1] iOS (14 vibrations, 91 pairs)	0.47 (0.34)	0.30 (0.43)	0.16 (0.52)	0.36 (0.41)	0.52 (-◇)	0.53 (0.39)	0.25 (0.36)
			Exp *II* [1] Android (14 vibrations, 91 pairs)	0.53 (0.40)	0.30 (0.46)	0.18 (0.57)	0.51 (0.40)	0.26 (-◇)	0.28 (0.47)	0.13 (0.45)
		MPS	Exp *I* [1] iOS (14 vibrations, 91 pairs)	0.47 (0.33)	-0.08 (0.55)	0.16 (0.50)	0.24 (0.56)	0.32 (0.41)	0.24 (0.40)	0.15 (0.39)
			Exp *II* [1] Android (14 vibrations, 91 pairs)	0.67 (0.34)	0.11 (0.51)	0.32 (0.48)	0.29 (0.46)	0.34 (0.39)	0.22 (0.42)	0.23 (0.41)
	Average for 6 untrained vibration sets			**0.55 (0.38)**	**0.27 (0.49)**	**0.22 (0.56)**	**0.43 (0.44)**	**0.51 (0.67)**	**0.30 (0.46)**	**0.37 (0.38)**
Average for all 12 vibration sets				**0.67 (0.32)**	**0.44 (0.41)**	**0.44 (0.46)**	**0.51 (0.39)**	**0.53 (0.50)**	**0.42 (0.40)**	**0.46 (0.36)**

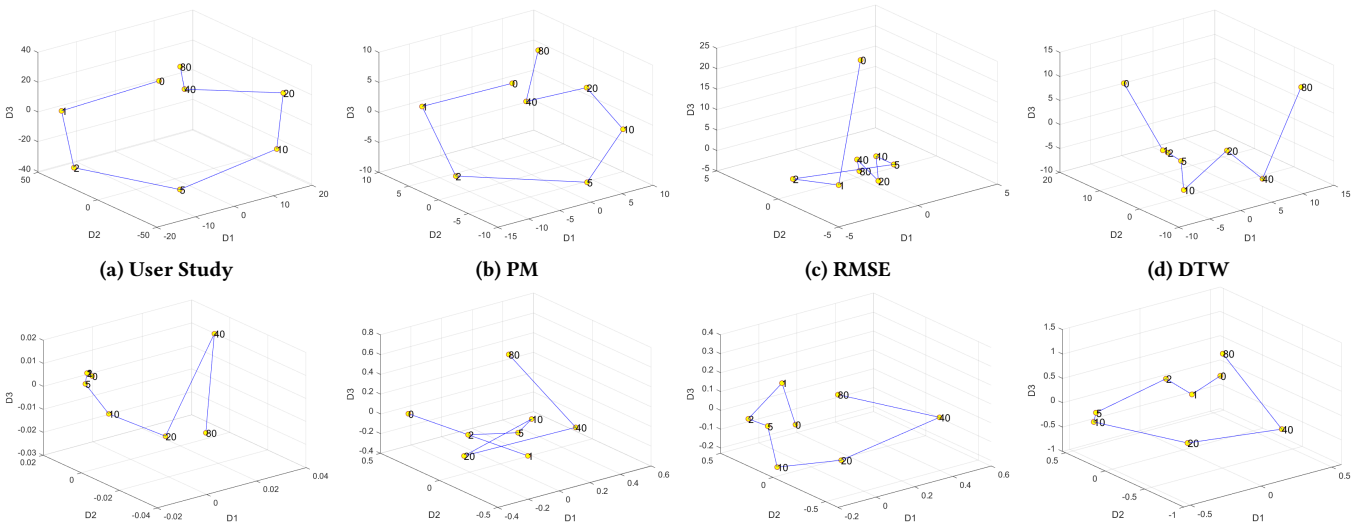
★ No correlation was reported because all predicted dissimilarities were the same value. ◇ MDS algorithm failed to compute point coordinates (3D).

modulated vibrations ($\geq 40Hz$) felt similar to the pure sinusoidal vibration of 150-Hz. In contrast, the vibrations with low modulated frequencies ($\leq 20Hz$) felt different from the pure sinusoidal vibration. PM-predicted perceptual space (Figure 8b) was rated as $K = 0.25$ with the user perceptual space, and the same implications to the user perceptual space could be derived. However, perceptual spaces of the other baseline metrics reported K larger than 0.5 and failed to imply those findings.

Overall, PM showed good performances in both measures with similar perceptual spaces. SPQI was also good for the spectral patterns of Low-Mid, Mid-High, and Low-High [26], but its predictions in the other patterns varying modulated frequency [45] or multiple parameters were not similar to the user-data. Perceptual spaces derived from the other five datasets were reported in the supplementary material.

6 VALIDATION WITH UNTRAINED DATASET: TRAINED AND UNTRAINED VIBROTACTILE PARAMETERS

Our validation in the trained datasets was sufficient to show PM's potential, but rigorous validations using untrained datasets are mandatory to assess its feasibility in the vibrotactile pattern design. In this section, PM was validated with six untrained datasets: two datasets varying all trained parameters and four datasets including the untrained parameter of rhythm.



(a) User Study (b) PM (c) RMSE (d) DTW

(e) M_{sm} (f) M_{tm} (g) ST-SIM (h) SPQI

Figure 8: Perceptual spaces derived from dissimilarities that (a) reported in [45] and predicted by (b) PM, (c) RMSE, (d) DTW, (e) M_{sm}, (f) M_{tm}, (g) ST-SIM, and (h) SPQI. The numbers represent modulated frequencies (0, 1, 2, 4, 10, 20, 40, and 80Hz), and data points of each space were connected along the modulated frequencies.

6.1 Trained Parameters: User Data Collection

We considered adopting user data from published articles, but no known research reported perceptual dissimilarities from the vibrations varying amplitude, frequency, modulated frequency, and frequency mixing ratio simultaneously. Therefore we conducted a user study to collect user-rated perceptual dissimilarities.

Participants - We recruited twenty participants (14M / 6F) who were 19 - 28 years old with a mean of 22.3 at the authors' institution. All participants did not report any impairments to their haptic sensation, and two of them were left-handed. The experiment took 80 minutes on average including instruction, and they were paid 19 USD as compensation. This experiment was approved by the authors' Institutional Review Board (20220419-HR-66-03-02).

Hardware Configuration - We used a voice-coil actuator (Tactile Labs; Mark II TL002-09-D) for its tactile capability for complex vibration. The vibration signals were commanded through a DAQ board (National Instruments; USB-6353) and amplified by a custom current amplifier. Also, we used an accelerometer (Analog Devices; ADXL354z) to measure the accelerations of generated vibrations. All the signals in this study were sampled with 10-kHz.

To assess PM's prediction to vibrations varying all trained vibration parameters [25, 26, 45], we varied amplitude, frequency, modulated frequency, and frequency mixing ratio to design 12 vibrations by using equation $x(t)$ as:

$$x(t) = A\sin(2\pi f_e t)\{w_{c_1}\sin(2\pi f_{c_1} t) + w_{c_2}\sin(2\pi f_{c_2} t)\} \quad (7)$$

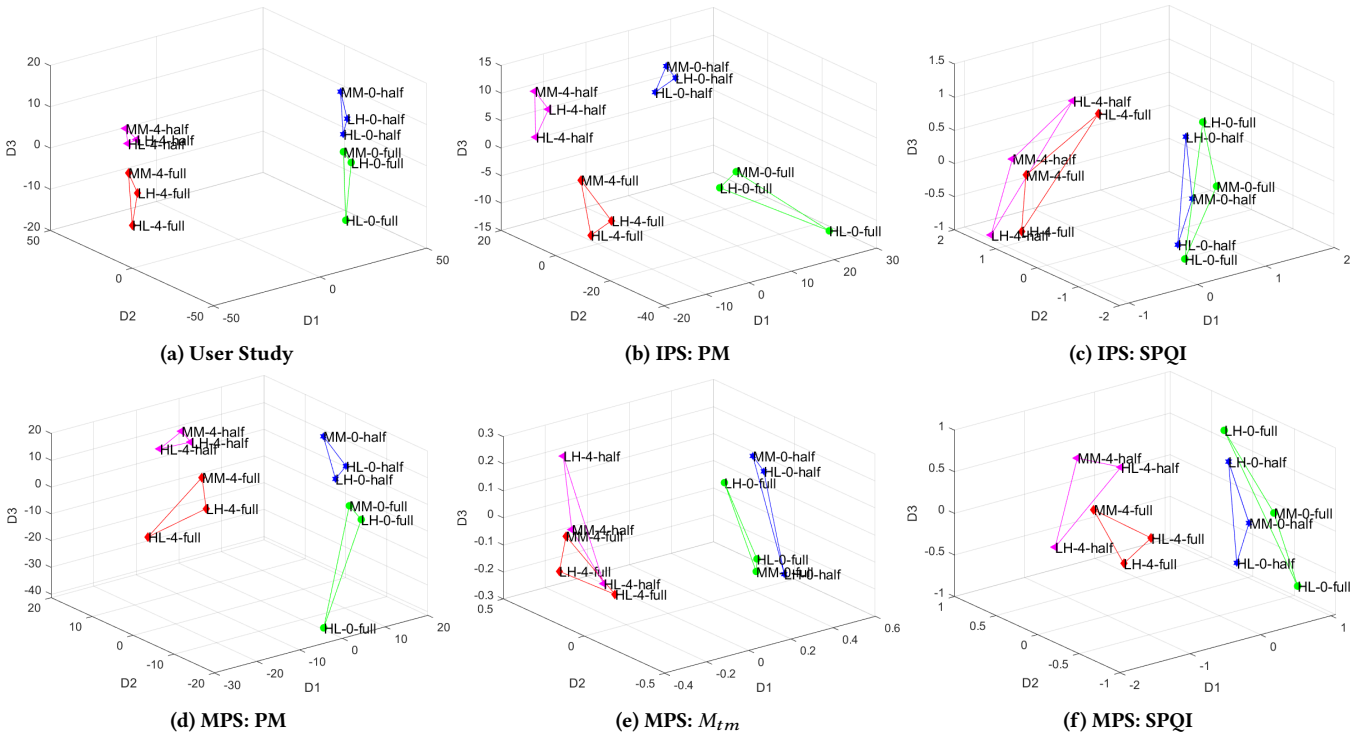
where A is amplitude (0.8, 1.6 G), f_e is modulated frequency (0, 4 Hz), w_{c_1} and w_{c_2} are frequency mixing ratios (1:0, 0.5:0.5, 0:1), and f_{c_1} and f_{c_2} are frequencies (70, 210 Hz). All vibrations lasted for 1 second and were named as **Frequency Mixing Ratio - Modulated Frequency - Amplitude**. For example, **HL-0-half** represents a pattern of non-modulated 70-Hz sinusoidal vibration with 0.8 G of amplitude.

Experiment Procedure - First of all, participants read documented instructions to understand the experiment. After the instruction, participants rigidly grasped the actuator with their left hand's index finger and thumb. They wore headphones playing pink noise to block out the auditory noises from vibrations. We designed a GUI-based program for the experiment consisting of a training session and a main session. In the training session, the initial screen showed 12 buttons randomly assigned to the 12 vibrations, and the participants could feel the vibrations by clicking the buttons. We forced a one-second rest between vibration rendering to remove the adaptation effects. Participants could proceed to the main session after feeling all vibrations at least once.

The main session allowed participants to evaluate the similarity of 264 vibration pairs by the combinatorial design of 12 vibration patterns with four repetitions (${}_{12}C_2 \times 4$). As an attention test, we added one pair of two identical vibrations (**HL-0-half**) and finally, 265 pairs were randomly presented. Each trial assigned two vibrations of the selected pair on the two buttons randomly. Participants rated the similarity of the two vibrations using a sliding bar from 0 (totally different) to 100 (totally same), and the bar was enabled after feeling both vibrations. During the evaluation, no limits were given to the number of vibration plays and a 1-second gap still existed for removing adaptation effects as in the training session.

6.2 Trained Parameters: Analysis in IPS and MPS

All participants scored higher than 80 on the attention test, so we used all data for analysis. We converted similarity ratings to dissimilarity scores by subtracting them from 100, and they were averaged over four repetitions and 20 participants. By using 66



(a) User Study (b) IPS: PM (c) IPS: SPQI

(d) MPS: PM (e) MPS: M_{tm} (f) MPS: SPQI

Figure 9: 3D perceptual spaces derived from (a) user data, (b) PM using intended pattern set (IPS), (c) SPQI using IPS, (d) PM using measured pattern set (MPS), (e) M_{tm} using MPS, and (f) SPQI using MPS. Data points of each space were connected along the same modulated frequency and amplitude. All of them were fitted with goodness-of-fit below 0.05.

pairwise dissimilarity scores, we built the estimated dissimilarity matrix.

Before the analysis, we considered haptic pattern designers may suffer from the difference between their intended patterns and the physical stimuli that a participant receives. Therefore we validated our model in two pattern sets: an intended pattern set (IPS) and a measured pattern set (MPS). IPS assumes that human skin is stimulated as designers intended. In this study, we attached the sensor to a vibration actuator and measured the accelerations of IPS while an experimenter was holding it with two fingers. Both IPS and MPS were inputted to the six baseline metrics including PM with a 10 kHz sampling rate and compared with the user study data.

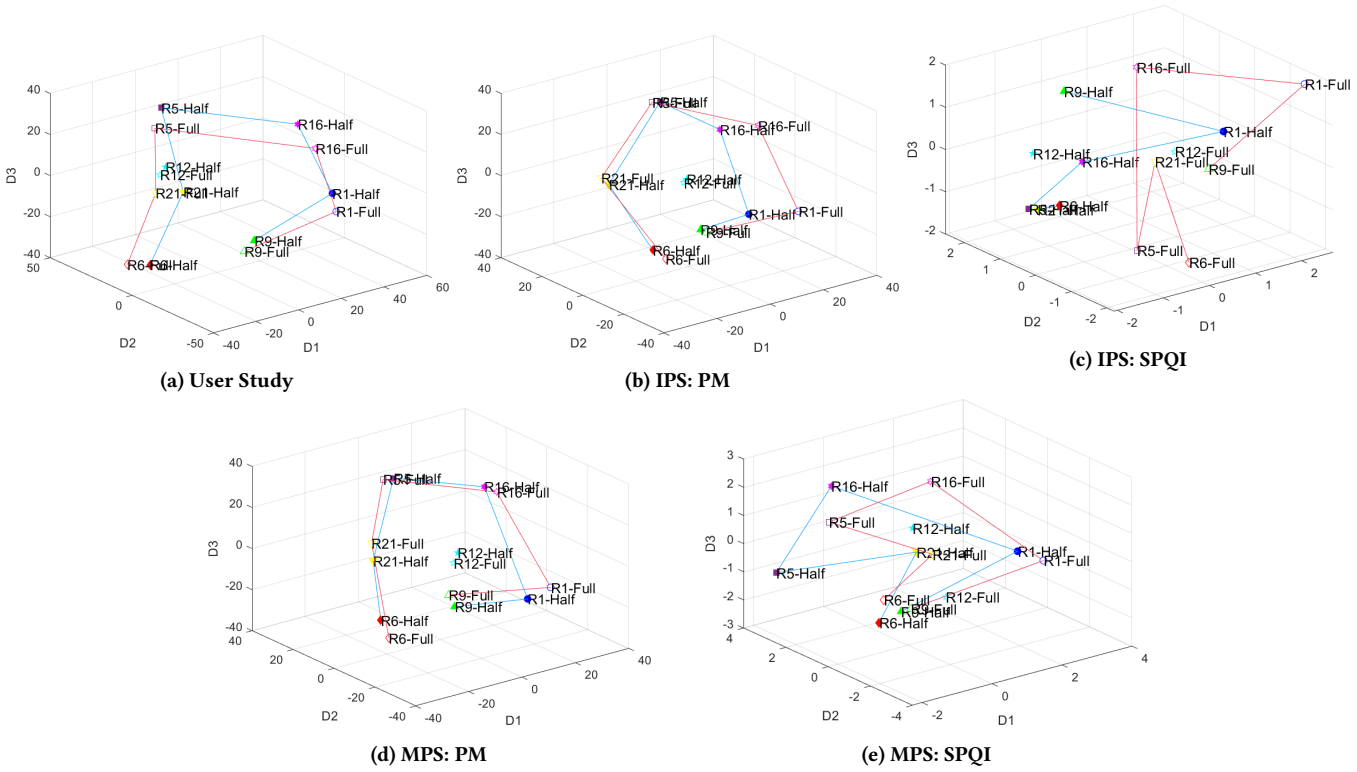
All the calculated Spearman's correlations and alienation coefficients were described in Table 2. For intended patterns, correlations of DTW and ST-SIM with the user data were less than 0.5 in Spearman's correlation. M_{tm} showed the highest correlation ($\rho = 0.79$), however, the dissimilarities included too many tied values. SPQI showed higher $\rho = 0.71$ than PM and RMSE ($\rho = 0.63$). For measured accelerations, RMSE, DTW, and ST-SIM showed below 0.5 of their correlations. M_{tm} also reported the highest correlation of 0.80 while SPQI, M_{sm}, and PM showed correlations of 0.77, 0.56, and 0.55. PM, M_{sm}, and SPQI reported good correlations higher than 0.5 in both patterns.

To verify the similarity between the derived perceptual spaces, we applied nMDS for 3D coordinates and calculated alienation coefficients K to the perceptual space of the user data. Only M_{tm} in intended patterns failed in deriving its perceptual space. In intended and measured patterns, SPQI and M_{tm} showed the lowest alienation coefficient of 0.32 and 0.23, respectively. PM reported 0.42 and 0.44 of K values that were not the best among the seven metrics.

We visualized 3D perceptual spaces of user data, two patterns of PM predictions, and the predictions showing the three lowest alienation coefficients in Figure 9. In the user-data-derived perceptual space, both D1 and D2 seem to align with modulated frequency, and D3 is assumed to be aligned with amplitude. Carrier frequency and mixing ratio contributed to local distribution in each cluster divided by modulated frequency and amplitude. Among the reported perceptual spaces, only the PM predictions showed a similar tendency of the three-axes representations to the user-data space in both IPS and MPS. In addition, mixing ratios also showed similar relative positions in each cluster with two exceptions. The other three perceptual spaces showed the modulated frequency alignment along D1 and D2, but the amplitude was also aligned with the same axes. Therefore, PM prediction seemed valid and matched the user data well in comparison to the baseline metrics.

6.3 Untrained Parameters: Data Acquisition and Prediction

Our model showed consistent predictions of perceptual dissimilarities to the vibration set using trained vibrotactile parameters and predicted well with MPS including noises, but still needs validation



(a) User Study (b) IPS: PM (c) IPS: SPQI

(d) MPS: PM (e) MPS: SPQI

Figure 10: 3D Perceptual Spaces of iOS vibration set derived from (a) user data in [1], and predicted data using intended pattern sets (IPS) by (b) PM, (c) SPQI and measured pattern sets (MPS) (d) PM, (e) SPQI. R# represents rhythms, and Full and Half represent the amplitude of the pattern. The data points of each perceptual space were connected along the rhythms of R6, R21, R5, R16, R1, and R9 in separate sets by full and half amplitudes. All of them were fitted with goodness-of-fit below 0.05.

to untrained vibrotactile parameters. Regardless of numerous data sets in the literature [47, 60, 67], we selected a rhythmic vibration pattern set in [1] because they rated pairwise perceptual dissimilarities while the others collected affective ratings or perceptual dissimilarities from cluster-sorted ratings. We contacted the authors to get the measured accelerations of iOS and Android device vibrations in the in-lab experiment with their dissimilarity ratings reported in [1]. We also prepared IPS using 150Hz and 160Hz carrier frequencies for iPhone 11 Pro Max (iOS) and Galaxy S10 (Android) patterns, respectively, and with two amplitude levels (0.5G and 1G). We predicted dissimilarities using both MPS and IPS for the rhythmic patterns and compared them with the in-lab user data.

6.4 Untrained Parameters: Analysis in IPS and MPS

All performance measures were calculated as in Section 6.2 and summarized in Table 2. Among seven metrics, PM robustly reported moderate and good correlations of 0.47 (iOS-IPS), 0.53 (Android-IPS), 0.47 (iOS-MPS), and 0.67 (Android-MPS) in all four datasets. Only three of the baseline metrics showed good correlations of 0.52 (iOS-IPS: M_{tm}), 0.53 (iOS-IPS: ST-SIM), and 0.51 (iOS-MPS: M_{tm}). RMSE showed a moderate correlation of 0.30 in two IPS sets but scored very low correlations in two MPS sets as -0.08 (iOS-MPS) and 0.11 (Android-MPS). In contrast, DTW only showed a significant correlation of 0.32 in Android phone condition using MPS waveforms. All correlations of SPQI were below 0.30 in both IPS and MPS. Also, we applied nMDS for 3D coordinates and calculated alienation coefficients K. Among seven metrics, PM always reported the lowest alienation coefficients of 0.34 (iOS-IPS), 0.40 (Android-IPS), 0.33 (iOS-MPS), and 0.34 (Android-MPS) in comparison to the baseline metrics.

We also visualized the most representative 3D perceptual spaces that capture the implications of the user's perceptual space. Specifically, we selected iOS perceptual spaces from user data, PM, and SPQI for both IPS and MPS as in Figure 9. The most salient implication from the user space is that the amplitude effect is small; half and full amplitude Tactons were co-located in the perceptual space. Next, the rhythmic Tactons were very dissimilar from each other, so they were located in the circular form except for $R12$ and $R21$ Tactons which were very similar. PM predictions matched implications of the amplitude and the mutual dissimilarities very well, while $R12$ was located at the center of its perceptual space. SPQI, which provided the most similar structure to the user data, did not catch any of the two implications from the derived perceptual spaces. Other metrics' predictions also did not catch them, and all the perceptual spaces are provided in the supplementary materials.

7 CONCLUSION AND FUTURE WORK

We suggested a computational model that mimics the human physiological model, PM, to predict pairwise perceptual dissimilarities of a same-duration vibration set in parallel and independent processes (NC, OC). Eight model parameters were trained with six vibration sets reported in three published papers, and the model predicted highly correlated dissimilarities ($\rho = 0.79$ on average) and similar perceptual spaces ($K = 0.26$) on average for the trained vibration sets. We validated our model using untrained vibration sets varied in trained and untrained vibrotactile parameters. In the trained parameter validation, our model prediction showed a good correlation to the user-rated dissimilarities in both IPS and MPS cases ($\rho = 0.63, 0.55$) and nMDS derived a perceptual space similar to that from user dissimilarities ($K = 0.42, 0.44$). In comparison, other six metrics - RMSE, DTW, M_{sm}, M_{tm}, ST-SIM, SPQI - did not show notable similarities in their derived perceptual spaces. In the untrained parameter validation, we tested our model to the four rhythmic vibration sets of iOS and Android devices in [1] in both MPS and IPS waveform predictions. PM predictions were robust in correlations and captured the implications similar to that of user-data-derived perceptual spaces, while the other baseline metrics showed inconsistent performances.

Overall, PM robustly predicted dissimilarities matching the user perceptual dissimilarities in both the correlation ($\rho = 0.67$) and similar perceptual space ($K = 0.32$) sharing the implications. Validation datasets also included measured accelerations of designer-intended vibration waveforms, therefore PM seems robust to the noise in the vibration signals. Moreover, PM showed plausible performances with the patterns varying the untrained rhythm parameter. SPQI, which was designed to measure spectral similarity between original and compressed vibrations, worked well in a single-parameter varying pattern set. Also, RMSE was the most sensitive to the noise among the seven baseline metrics.

We expect that our model provides a way of interactive design to vibration designers for easy initiation of pattern organization. The common approach in Haptics requires haptic pattern designers to run a new user study for a slightly changed vibration set. This flexibility possibly reduces the cost of finding the best subset of vibrations and accelerates the Tacton design process.

PM still needs to improve several features for more accurate prediction and diverse usage. First, we trained PM using a small number of datasets that varied partial vibrotactile parameters, and each dataset included less than sixteen vibrations without noise. Second, PM adopted simplified neural structures and processes that might negatively affect the prediction. Finally, our model assumes a limited stimulation scenario with a 0.5-mm probe for utilizing the known psychophysical knowledge. The below subsections describe the detailed plans dealing with them.

7.1 Small Vibrotactile Parameter Space

Our model was trained with only six vibration sets, and the average number of vibrations was 9.67 (47.83 pairs), as shown in Table 2. Those patterns were designed by varying four vibrotactile parameters, but they are not sufficient to represent general vibrations. In future work, untrained vibrotactile parameters of vibrotactile pattern designers' need, including metaphor and rhythm, would be included in the training data sets for generalized prediction.

7.2 Same-length Input Signal Duration

As described in Equation 2 of Section 4.2, our model predicts the neural fires by normalized spectral intensities and distributes them in every 10-ms neural window. Therefore if different lengths of vibrations are compared, then the longer vibration signal generates a less number of neural spikes than expected which degrades the prediction accuracy. We expect our model would be improved by adopting training data sets varied in duration with model modification.

7.3 Vibration Propagation Simulation Considering the Physical Contact

One of the difficult but attractive aspects of tactile interface design is the freedom of interaction by varying attaching body sites and contact postures. PM modeled vibration propagation by assuming the fingertip is stimulated with a fixed-size round probe, and it still lacks the freedom in designing what tactile interface would render the vibrations. We expect that PM may contribute better if it includes an improved vibration propagation model that simulates the physical contact between the skin and the device.

ACKNOWLEDGMENTS

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Science, ICT & Future Planning(2021R1C1C1008147), Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2019-0-01842, Artificial Intelligence Graduate School Program (GIST)), and a Korea Medical Device Development Fund grant funded by the Korean government (Ministry of Science and ICT; Ministry of Trade, Industry, and Energy; Ministry of Health & Welfare, Ministry of Food and Drug Safety) (project No. RS-2021-KD000009).

REFERENCES

[1] Ramzi Abou Chahine, Dongjae Kwon, Chungman Lim, Gunhyuk Park, and Hasti Seifi. 2022. Vibrotactile Similarity Perception in Crowdsourced and Lab Studies. In *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*. Springer, 255–263.

[2] Jeremie Barral, Xiao-Jing Wang, and Alex D Reyes. 2019. Propagation of temporal and rate signals in cultured multilayer networks. *Nature communications* 10, 1 (2019), 1–14.

[3] Donald J Berndt and James Clifford. 1994. Using dynamic time warping to find patterns in time series.. In *KDD workshop*, Vol. 10. Seattle, WA, USA:, 359–370.

[4] Stanley J Bolanowski Jr, George A Gescheider, Ronald T Verrillo, and Christin M Checkosky. 1988. Four channels mediate the mechanical aspects of touch. *The Journal of the Acoustical society of America* 84, 5 (1988), 1680–1694.

[5] Ingwer Borg and Detlev Leutner. 1985. Measuring the similarity of MDS configurations. *Multivariate Behavioral Research* 20, 3 (1985), 325–334.

[6] Stephen A Brewster and Lorna M Brown. 2004. Tactons: structured tactile messages for non-visual information display. (2004).

[7] Lorna M Brown, Stephen A Brewster, and Helen C Purchase. 2005. A first investigation into the effectiveness of tactons. In *First joint eurohaptics conference and symposium on haptic interfaces for virtual environment and teleoperator systems. world haptics conference*. IEEE, 167–176.

[8] Lorna M Brown and Topi Kaaresoja. 2006. Feel who's talking: using tactons for mobile phone alerts. In *CHI'06 extended abstracts on Human factors in computing systems*. 604–609.

[9] M Brown and L Rabiner. 1982. Dynamic time warping for isolated word recognition based on ordered graph searching techniques. In *ICASSP'82. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 7. IEEE, 1255–1258.

[10] Sung-Hyuk Cha. 2007. Comprehensive survey on distance/similarity measures between probability density functions. *City* 1, 2 (2007), 1.

[11] Seungmoon Choi and Katherine J Kuchenbecker. 2012. Vibrotactile display: Perception, technology, and applications. *Proc. IEEE* 101, 9 (2012), 2093–2104.

[12] Ben Clark, Oliver S Schneider, Karon E MacLean, and Hong Z Tan. 2017. Predictable and distinguishable morphing of vibrotactile rhythm. In *2017 IEEE World Haptics Conference (WHC)*. IEEE, 84–89.

[13] Jacob Cohen. 2013. *Statistical power analysis for the behavioral sciences*. Routledge.

[14] Heather Culbertson and Katherine J Kuchenbecker. 2016. Importance of matching physical friction, hardness, and texture in creating realistic haptic virtual surfaces. *IEEE transactions on haptics* 10, 1 (2016), 63–74.

[15] Nathan Dunkelberger, Joshua Bradley, Jennifer L Sullivan, Ali Israr, Frances Lau, Keith Klumb, Freddy Abnousi, and Marcia K O'Malley. 2018. Improving perception accuracy with multi-sensory haptic cue delivery. In *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*. Springer, 289–301.

[16] Mi Feng, Arindam Dey, and Robert W Lindeman. 2016. An initial exploration of a multi-sensory design space: Tactile support for walking in immersive virtual environments. In *2016 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, 95–104.

[17] Yoren Gaffary and Anatole Lécuyer. 2018. The use of haptic and tactile information in the car to improve driving safety: A review of current technologies. *Frontiers in ICT* 5 (2018), 5.

[18] George A Gescheider. 1966. Resolving of successive clicks by the ears and skin. *Journal of experimental psychology* 71, 3 (1966), 378.

[19] George A Gescheider, John H Wright, and Ronald T Verrillo. 2010. *Information-processing channels in the tactile sensory system: A psychophysical and physiological analysis*. Psychology press.

[20] Toni Giorgino. 2009. Computing and visualizing dynamic time warping alignments in R: the dtw package. *Journal of statistical Software* 31 (2009), 1–24.

[21] Rania Hassen and Eckehard Steinbach. 2019. Subjective evaluation of the spectral temporal similarity (st-sim) measure for vibrotactile quality assessment. *IEEE Transactions on Haptics* 13, 1 (2019), 25–31.

[22] Rania Hassen and Eckehard Steinbach. 2019. Vibrotactile signal compression based on sparse linear prediction and human tactile sensitivity function. In *2019 IEEE World Haptics Conference (WHC)*. IEEE, 301–306.

[23] Cristy Ho, Nick Reed, and Charles Spence. 2006. Assessing the effectiveness of "intuitive" vibrotactile warning signals in preventing front-to-rear-end collisions in a driving simulator. *Accident Analysis & Prevention* 38, 5 (2006), 988–996.

[24] Mark Hollins, Sliman Bensmaïa, Kristie Karlof, and Forrest Young. 2000. Individual differences in perceptual space for tactile textures: Evidence from multidimensional scaling. *Perception & Psychophysics* 62, 8 (2000), 1534–1544.

[25] Inwook Hwang and Seungmoon Choi. 2010. Perceptual space and adjective rating of sinusoidal vibrations perceived via mobile device. In *2010 IEEE Haptics Symposium*. IEEE, 1–8.

[26] Inwook Hwang, Jeongil Seo, and Seungmoon Choi. 2017. Perceptual space of superimposed dual-frequency vibrations in the hands. *PloS one* 12, 1 (2017), e0169570.

[27] Roland S Johansson and Ake B Vallbo. 1979. Tactile sensibility in the human hand: relative and absolute densities of four types of mechanoreceptive units in glabrous skin. *The Journal of physiology* 286, 1 (1979), 283–300.

[28] KO Johnson. 1974. Reconstruction of population response to a vibratory stimulus in quickly adapting mechanoreceptive afferent fiber population innervating glabrous skin of the monkey. *Journal of Neurophysiology* 37, 1 (1974), 48–72.

[29] Jon H Kaas. 1993. The functional organization of somatosensory cortex in primates. *Annals of Anatomy-Anatomischer Anzeiger* 175, 6 (1993), 509–518.

[30] Eric R Kandel, James H Schwartz, Thomas M Jessell, Steven Siegelbaum, A James Hudspeth, Sarah Mack, et al. 2000. *Principles of neural science*. Vol. 4. McGraw-hill New York.

[31] Robert K Katzschmann, Brandon Araki, and Daniela Rus. 2018. Safe local navigation for visually impaired users with a time-of-flight and haptic feedback device. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 26, 3 (2018), 583–593.

[32] Jaebong Lee and Seungmoon Choi. 2013. Real-time perception-level translation from audio signals to vibrotactile effects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2567–2576.

[33] Joseph Luk, Jerome Pasquero, Shannon Little, Karon MacLean, Vincent Levesque, and Vincent Hayward. 2006. A role for haptics in mobile interaction: initial design using a handheld tactile display prototype. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*. 171–180.

[34] Karon MacLean and Mario Enriquez. 2003. Perceptual design of haptic icons. In *Proc. of EuroHaptics*. 351–363.

[35] Karon E MacLean. 2009. Putting haptics into the ambience. *IEEE Transactions on Haptics* 2, 3 (2009), 123–135.

[36] Yoichiro Matsuura, Shogo Okamoto, Hikaru Nagano, and Yoji Yamada. 2014. Multidimensional matching of tactile sensations of materials and vibrotactile spectra. *Information and Media Technologies* 9, 4 (2014), 505–516.

[37] Francis McGlone and David Reilly. 2010. The cutaneous sensory system. *Neuroscience & Biobehavioral Reviews* 34, 2 (2010), 148–159.

[38] Meinard Müller. 2007. *Information retrieval for music and motion*. Vol. 2. Springer.

[39] Michael A Muniak, Supratim Ray, Steven S Hsiao, J Frank Dammann, and Sliman J Bensmaia. 2007. The neural coding of stimulus intensity: linking the population response of mechanoreceptive afferents with psychophysical behavior. *Journal of Neuroscience* 27, 43 (2007), 11687–11699.

[40] Evelyn Muschter, Andreas Noll, Jinting Zhao, Rania Hassen, Matti Strese, Basak Gülecyüz, Shu-Chen Li, and Eckehard Steinbach. 2021. Perceptual quality assessment of compressed vibrotactile signals through comparative judgment. *IEEE Transactions on Haptics* 14, 2 (2021), 291–296.

[41] Andreas Noll, Markus Hofbauer, Evelyn Muschter, Shu-Chen Li, and Eckehard Steinbach. 2022. Automated Quality Assessment for Compressed Vibrotactile Signals Using Multi-Method Assessment Fusion. In *2022 IEEE Haptics Symposium (HAPTICS)*. IEEE, 1–6.

[42] Marianna Obrist, Sue Ann Seah, and Sriram Subramanian. 2013. Talking about tactile experiences. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1659–1668.

[43] Shogo Okamoto, Hikaru Nagano, and Yoji Yamada. 2012. Psychophysical dimensions of tactile perception of textures. *IEEE Transactions on Haptics* 6, 1 (2012), 81–93.

[44] Chaeyong Park, Jeongwoo Kim, Dong-Geun Kim, Seungjae Oh, and Seungmoon Choi. 2022. Vibration-Augmented Buttons: Information Transmission Capacity and Application to Interaction Design. In *CHI Conference on Human Factors in Computing Systems*. 1–13.

[45] Gunhyuk Park and Seungmoon Choi. 2011. Perceptual space of amplitude-modulated vibrotactile stimuli. In *2011 IEEE world haptics conference*. IEEE, 59–64.

[46] Gunhyuk Park and Katherine J Kuchenbecker. 2019. Objective and subjective assessment of algorithms for reducing three-axis vibrations to one-axis vibrations. In *2019 IEEE World Haptics Conference (WHC)*. IEEE, 467–472.

[47] Jerome Pasquero, Joseph Luk, Shannon Little, and Karon MacLean. 2006. Perceptual analysis of haptic icons: an investigation into the validity of cluster sorted mds. In *2006 14th Symposium on Haptic Interfaces for Virtual Environment and*

Teleoperator Systems. IEEE, 437–444.

[48] Joseph M Romano and Katherine J Kuchenbecker. 2011. Creating realistic virtual textures from contact acceleration data. *IEEE Transactions on haptics* 5, 2 (2011), 109–119.

[49] Jonghyun Ryu, Jaemin Chun, Gunhyuk Park, Seungmoon Choi, and Sung H Han. 2010. Vibrotactile feedback for information delivery in the vehicle. *IEEE Transactions on Haptics* 3, 2 (2010), 138–149.

[50] Bushra Sadia, Senem Ezgi Emgin, T Metin Sezgin, and Cagatay Basdogan. 2020. Data-driven vibrotactile rendering of digital buttons on touchscreens. *International Journal of Human-Computer Studies* 135 (2020), 102363.

[51] Robert Scheibe, Mathias Moehring, and Bernd Froehlich. 2007. Tactile feedback at the finger tips for improved direct interaction in immersive environments. In *2007 IEEE Symposium on 3D User Interfaces*. IEEE.

[52] Maximilian Schirmer, Johannes Hartmann, Sven Bertel, and Florian Echtler. 2015. Shoe me the way: a shoe-based tactile interface for eyes-free urban navigation. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*. 327–336.

[53] Ryan E Schoonmaker and Caroline GL Cao. 2006. Vibrotactile force feedback system for minimally invasive surgical procedures. In *2006 IEEE international conference on systems, man and cybernetics*, Vol. 3. IEEE, 2464–2469.

[54] Hasti Seifi, Kailun Zhang, and Karon E MacLean. 2015. VibViz: Organizing, visualizing and navigating vibration libraries. In *2015 IEEE World Haptics Conference (WHC)*. IEEE, 254–259.

[55] Charles Scott Sherrington. 1906. *The integrative action of the nervous system*. Vol. 35. Yale University Press.

[56] Sunghwan Shin, Reza Haghighi Osgouei, Ki-Duk Kim, and Seungmoon Choi. 2015. Data-driven modeling of isotropic haptic textures using frequency-decomposed neural networks. In *2015 IEEE World Haptics Conference (WHC)*. IEEE, 131–138.

[57] Stanley Smith Stevens and Lawrence E Marks. 2017. *Psychophysics: Introduction to its perceptual, neural, and social prospects*. Routledge.

[58] Mriganka Sur, JOHN T Wall, and JH Kaas. 1984. Modular distribution of neurons with slowly adapting and rapidly adapting responses in area 3b of somatosensory cortex in monkeys. *Journal of neurophysiology* 51, 4 (1984), 724–744.

[59] Jun Tan, Yan Ge, Xianghong Sun, Yubo Zhang, and Yanfang Liu. 2019. User Experience of Tactile Feedback on a Smartphone: Effects of Vibration Intensity, Times and Interval. In *International Conference on Human-Computer Interaction*. Springer, 397–406.

[60] David Ternes and Karon E MacLean. 2008. Designing large sets of haptic icons with rhythm. In *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*. Springer, 199–208.

[61] Luca Turchet, Paolo Burelli, and Stefania Serafin. 2012. Haptic feedback for enhancing realism of walking simulations. *IEEE transactions on haptics* 6, 1 (2012), 35–45.

[62] AB Vallbo and RS Johansson. 1978. The tactile sensory innervation of the glabrous skin of the human hand. *Active touch* 2954 (1978), 29–54.

[63] Jan BF Van Erp. 2005. Presenting directions with a vibrotactile torso display. *Ergonomics* 48, 3 (2005), 302–313.

[64] Yasemin Vardar, Christian Wallraven, and Katherine J Kuchenbecker. 2019. Fingertip interaction metrics correlate with visual and haptic perception of real surfaces. In *2019 IEEE World Haptics Conference (WHC)*. IEEE, 395–400.

[65] Lawrence M Ward. 1977. Multidimensional scaling of the molar physical environment. *Multivariate Behavioral Research* 12, 1 (1977), 23–42.

[66] Qianli Xu, Tian Gan, Shue Ching Chia, Liyuan Li, Joo-Hwee Lim, and Phyoe Kyaw Kyaw. 2016. Design and evaluation of vibrating footwear for navigation assistance to visually impaired people. In *2016 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*. IEEE, 305–310.

[67] Yongjae Yoo, Taekbeom Yoo, Jihyun Kong, and Seungmoon Choi. 2015. Emotional responses of tactile icons: Effects of amplitude, frequency, duration, and envelope. In *2015 IEEE World Haptics Conference (WHC)*. IEEE, 235–240.