The high/low frequency balance drives tactile perception of noisy vibrations

Corentin Bernard, Etienne Thoret, Nicolas Huloux, Sølvi Ystad

The high/low frequency balance drives tactile perception of noisy vibrations

Article *in* IEEE Transactions on Haptics · February 2024

DOI: 10.1109/TOH.2024.3371264

4 authors:

Corentin Bernard
French National Centre for Scientific Research
16 PUBLICATIONS **90** CITATIONS
SEE PROFILE

Etienne Thoret
French National Centre for Scientific Research
51 PUBLICATIONS **444** CITATIONS
SEE PROFILE

Nicolas Huloux
11 PUBLICATIONS **86** CITATIONS
SEE PROFILE

Sølvi Ystad
French National Centre for Scientific Research
200 PUBLICATIONS **1,751** CITATIONS
SEE PROFILE

The high/low frequency balance drives tactile perception of noisy vibrations

Corentin Bernard, Etienne Thoret, Nicolas Huloux and Sølvi Ystad

***Abstract*—Noisy vibrotactile signals transmitted during tactile explorations of an object provide precious information on the nature of its surface. Understanding the link between signal properties and how they are interpreted by the tactile sensory system remains challenging. In this paper, we investigated human perception of broadband, stationary vibrations recorded during exploration of textures and reproduced using a vibrotactile actuator. Since intensity is a well-established perceptual attribute, we here focused on the relevance of the spectral content. The stimuli were first equalized in perceived intensity and subsequently used to identify the most salient spectral features using dissimilarity estimations between pairs of successive vibration. Based on dimensionally reduced spectral representations, models of dissimilarity ratings showed that the balance between low and high frequencies was the most important cue. Formal validation of this result was achieved through a Mushra experiment, in which participants assessed the fidelity of resynthesized vibrations with various distorted frequency balances. These findings offer valuable insights into human vibrotactile perception and establish a computational framework for analyzing vibrations as humans do. Moreover, they pave the way for signal synthesis and compression based on sparse representations, holding significance for applications involving complex vibratory feedback.**

***Index Terms*—Vibrotactile perception, vibration sparse synthesis, vibration compression, psychophysics, haptics.**

I. Introduction

SIMPLE vibrations have become a standard mean to convey information in our smartphones and game controllers. Recent technological advancements have expanded the frequency bandwidth of actuators, enabling software designers to create more precise vibrations. Additionally, the miniaturization of these actuators enables their integration into a wide range of human-machine interfaces, including wearables and vibrotactile touchscreens. These vibrations are primarily used to provide users with additional information, enhancing usability and accessibility. Moreover, finely-tuned vibrations have been employed to improve immersion in virtual gaming environments [1], [2] and facilitate remote social interaction [3], [4].

Manuscript submitted 12 October 2023; revised 10 January 2024; accepted 23 February 2024.

Corentin Bernard, Etienne Thoret and Sølvi Ystad are with Aix Marseille Univ, CNRS, PRISM, Marseille, France. E-mail : bernard@prism.cnrs.fr

Corentin Bernard and Nicolas Huloux are with MIRA, Aflokkat, Ajaccio, France.

Etienne Thoret is with Institute of Language Communication and the Brain, Aix Marseille Univ, UMR7020 Laboratoire d'Informatique et Systèmes, UMR7289 Institut de Neurosciences de la Timone, CNRS, Marseille, France.

Nicolas Huloux is with Univ. Bordeaux, ESTIA-Institute of Technology, EstiaR, F-64210 Bidart, France.

Etienne Thoret & Sølvi Ystad equally contributed to the supervision of this work. This work was supported by France Relance. Author ET was supported by grants ANR-16-CONV-0002 (ILCB), ANR-11-LABX-0036 (BLRI) and the Excellence Initiative of Aix-Marseille University (A*MIDEX).

Up to this point, two distinct approaches have been devised for generating vibrations. The first approach involves crafting signals using a simple model based on a limited set of parameters, often relying on a single sine wave with varying frequency and envelope [5]. This method offers the advantage of being sparse while enabling the creation of a diverse range of stimuli. However, it falls short in accurately reproducing natural vibrations with a complex spectral profile. The second approach is based on the capture and faithful playback of vibrations induced by friction during texture exploration [6]. This approach can achieve a high level of realism but is much more demanding in terms of amount of data to be restituted. The need to reduce data storage size has become increasingly crucial as new technologies emerge, incorporating a multitude of channels, such as multitouch haptic surfaces [7], wearable devices [8] or controllers [9] equipped with multiple actuators.

On the other hand, the human body is consistently exposed to vibrations generated by the friction between the skin and external surfaces, especially when manipulating objects with the fingers. This influx of information is initially processed by mechanoreceptors, which transform mechanical vibrations into electrical spike patterns [10]–[12]. The efficient coding hypothesis [13] postulates that the sensory system encodes incoming information as efficiently as possible, eliminating redundancies to minimize the number of spikes and reduce neural activity. In touch, prior research has suggested potential reduction in dimensionality when encoding skin deformations [14], such as for detecting of object slippage [15].

In this paper, we investigated human vibrotactile perception to identify perceptually relevant signal features that could further be used to simplify the vibrotactile signal. Our hypothesis was that by extracting these features from the incoming vibrations, we could reduce the signal to what is strictly perceptible.

Through psychophysical experiments, we identified spectral structures that perceptually stand out within noisy vibrations. We used the term "noisy" to describe vibrations that exhibit a broadband spectrum similar to that of a white noise, in contrast to monochromatic signals. We demonstrated that any stationary, noisy signal can be projected onto this perception-driven basis and then reconstructed with this reduced information. In addition to the interesting synthesis perspective of this approach, it can also be employed for achieving sparse representations for vibrotactile signal compression.

The paper is divided into three sections, each corresponding to a distinct experiment dedicated to the perception of noisy stationary vibrations, which were here recordings of friction-induced vibrations that present temporal homogeneity. The first experiment investigates the perception of the intensity

of these vibrations. It provides a relationship between signal power and perceived intensity, as well as an intensity equalization of the stimuli, hereby preparing for the next experiments. The second experiment delves into the perception of other attributes by posing a simple question: which spectral attributes enable distinguishing sensations between two vibrations? The results underscored the essential role played by one fundamental element in vibrotactile perception: the balance between high and low frequencies. Lastly, the third experiment not only validates these previous findings but also showcases their potential in terms of vibrotactile synthesis and compression.

II. Background

A. Intensity and frequency perception

The literature on vibrotactile perception offers a plethora of studies that have focused on two key characteristics of simple vibrations: intensity and frequency. While amplitude and frequency were intrinsically linked in the first vibration-rendering systems based on Eccentric Rotating Mass (ERMs), they can now be dissociated using voice coil actuators and investigated independently. Intensity is known as the principal perceptual attribute in vibrotactile perception, playing a pivotal role in discriminating between tactile stimuli. This holds true for both simple synthesized vibrations [16] and for friction induced vibration recorded from various textures [17]. Perceived intensity is directly associated with the signal's amplitude. Remarkably, humans can discern subtle differences in intensity regardless of the frequency, with discrimination thresholds ranging from 11% to 30% depending on the study (see [18] for a review). Moreover, it is noteworthy that perceived intensity is also influenced by the signal's frequency. Amplitude detection thresholds follow a U-shaped curve related to frequency, similar to curves of equal intensity [19]. Vibrations of equal amplitude are perceived as more intense at frequencies around 200 Hz. Models have been developed to predict the intensity of a single sine wave based on its amplitude and frequency [20], and also for signals comprising a sum of two sine waves [21]. Furthermore, humans possess the ability to distinguish frequency differences when they deviate by more than 20%. The just-noticeable difference typically falls between 17% and 21 % across various studies, following Weber's Law [22], [23].

The literature also sheds lights on the frequency selectivity of vibrations composed of two distinct frequencies. When these two frequencies are sufficiently distant, separated by more than 100 Hz, they give rise to a unique percept. In such cases, we are unable to discern the two individual vibrations distinctly; instead, we perceive a vibration that falls somewhere between the initial frequencies, with the perception being influenced by the amplitude of each frequency component [24]–[27]. Conversely, when the two frequencies are closely spaced ($<$100 Hz), interference patterns emerge, resulting in what is called "beating" [28] . In this scenario, the frequency of the beating, or the frequency of the amplitude modulation, takes on a more prominent role as a perceptual attribute compared to the initial frequencies [28], [29], inducing a sensation of rhythm [30].

From a physiological perspective, perceived intensity is primarily mediated in the somatosensory periphery by the total population of nerve fibers that are activated [31], whereas the frequency coding relies on temporal spiking patterns [32].

B. Vibration encoding and compression

Signal encoding is a crucial issue for vibration rendering. Since humans are sensitive to vibrations up to 1000 Hz [19], temporal signals are usually encoded with a sample rate of about 2000 Hz. An 8 bits quantization of the signal has been shown to be sufficient for preserving the perceived quality [33].

Previous studies have also demonstrated that classical mathematical transformations, such as the Fourier transform or the Gabor transform, were relevant to model the encoding of vibration signals in accordance to human perception [34]–[36].

In the literature, many contributions have studied the compression of vibrotactile signals in order to develop the best methods to reduce the file size without altering its tactile quality. The main idea is to remove information that is not perceived from the signal. Previous works focused at removing frequency components that are below the detection threshold [37], [38]. Based on frequency masking [39], other approaches [40]–[42] consist in removing the low-amplitude frequencies that are imperceptible due to their vicinity with a high-amplitude frequency. A measure of vibration similarity, based on spectral and temporal similarities [43], has been developed to assess the compression quality and compare compression methods [44]. Another approach proposed compressing haptic signals into sequences of tactile pixels defined as small units of a single frequency [45].

III. Material and methods

This section outlines the setup utilized in the three experiments presented in this paper. The experiments were approved by the Ethical Committee of Aix-Marseille University (Reference: 2023-05-11-06) and were carried out according to the principles expressed in the Declaration of Helsinki. Written informed consent was obtained from all participants involved in the study.

A. Vibrotactile stimuli

Recordings of friction-induced vibrations from Kirsch et al.'s database [46] were used to provide vibrotactile stimuli. These signals were recorded during exploration of textures with a specific tool equipped with an accelerometer. We selected 18 signals corresponding to 9 materials (rubber, polyester pad, foam, felt, cork, bamboo, baltic brown (granite), anti-vib pad (recycled rubber) and aluminium grid) explored by 2 different probes (round and spiked) for a medium scanning speed condition. To use stationary signals, one second of each recording during which the scanning speed was roughly constant between 100 and 120 mm/s was kept. The sampling rate was set to 2800 Hz.

B. Vibration rendering

Tactile stimuli were presented through an Actronika (Paris, France) HapCoil-One vibrotactile actuator (dimensions, $11.5 \times 12 \times 37.7$ mm^3; frequency bandwidth, 10–1000 Hz; resonant frequency with no load, 65 Hz). The actuator was powered by a Pioneer A-209R audio amplifier. The participants were asked to grab the vibrotactile actuator between the thumb and the index finger of their left hand to feel the vibrations. Fig. 1.a presents the frequency response of the actuator while being held with two fingers. It represents how the gain, ie. the ratio between the actuator acceleration and the input voltage, is affected by the vibration frequency. The frequency response was measured with an accelerometer (ICP Accelerometer 352A71, PCB Piezotronics, Buffalo, USA) fixed between the actuator and the index finger of the experimenter. It should be noted that the impedance of the finger grip attenuates the resonance of the actuator [47] compared to the free oscillations condition provided in the accelerometer reference guide. The differences in frequency response among individuals holding the actuator are negligible concerning perception [48]. Therefore, a unique frequency response was utilized in this study.

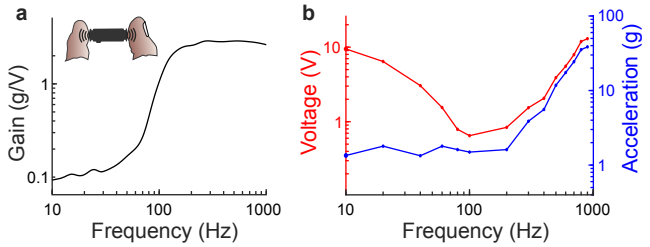

Fig. 1. **a.** Frequency response of the actuator held by two fingers, as in the experimental conditions. **b.** Equal-intensity sensation curve of vibrations rendered by the actuator and felt with two fingers. The curves display the peak-to-peak voltage of the input signal (in red) and the peak-to-peak acceleration (in blue) corresponding to sine-waves at various frequencies that are perceived with the same intensity.

In the following, vibrations will be described either by their input voltage when considering the source signal (for intensity and synthesis), or by their acceleration (signal filtered by the frequency response) when investigating human perception.

The actuator was also characterized *via* the equal-intensity sensation curves presented in Fig. 1.b, representing the voltage (proportional to the displacement) and the corresponding acceleration of pure steady sinewaves for which humans perceive a constant intensity

The intensity judgments were gathered in a previous experiment involving five participants, with the method of direct intensity matching as in [19]. The curves show in which frequency band participants are most sensitive to the actuator.

During all experiments, participants wore noise-canceling headphones playing pink noise to mask sounds produced by the actuator and avoid any auditory bias.

IV. Intensity Perception

As detailed previously, intensity is well known as the primary perceptual attribute of vibration. However, existing models of intensity perception in the literature have primarily focused on simple vibrations with one or two sine waves. To date, there are no intensity models specifically designed for noisy vibrations.

Consequently, a preliminary experiment was conducted to equalize the intensity of the stimuli, so that the contribution of other attributes could be studied regardless of the intensity. Besides, the participants' judgments were examined to develop a model of perceived intensity for noisy vibrations.

A. Protocol

A mathematical amplitude normalization was first conducted by equalizing the standard deviation of the 18 vibration signals. Perceptual intensity equalizations were then performed by the participants who were asked to adjust the gain of each stimulus to match the intensity of a reference stimulus: a white noise with normalized standard deviation amplitude. The experimental interface was composed of buttons allowing the participants to playback the reference and the stimuli successively as many times as they want, along with sliders to adjust the gains. For the 18 stimuli, the gain could be adjusted between 0.2 and 5, while the reference was assigned a gain of 1.

B. Participants

The intensity equalization was performed by 10 participants, 2 females and 8 males, from 23 to 55 years old (mean=29.7), 1 left-handed and 9 right-handed.

C. Results

Participants responses were coherent, as shown by high pairwise correlations between participants on the selected gains (Pearson's r: M=0.82, SD=0.08, df=16 for the 45 pairwise correlations, all correlations were significant at $\alpha = 0.05$, ranging from p=6.5×10^{-11} to p=0.009.

For each stimulus, the average gain was computed as the geometric mean of all participants judgments. The mean gains were then applied to the signals in order to obtain the 18 equally-intense vibrotactile stimuli for the second experiment. The high correlations (Pearson's r: M=0.91, SD=0.05, df=16 for each of the 10 correlations, all significant p$<$0.05) between the mean gains and the participants individual gains shows that the mean gains faithfully represent the participants judgments.

Fig. 2.a presents the result of the intensity equalization on the power spectra of the stimuli. After equalization, the vibration signals show less variability in the 80-200 Hz frequency band. This means that, when asked to equalize the intensity, the participants tended to equalize the power of the signals in this specific frequency band, regardless of the power in higher or lower frequency bands.

D. Model of intensity perception

Based on these results, we constructed a model to predict the perceived intensity of noisy vibrations from their signal. The intensity judgments were derived by taking the inverse of the selected gains, as higher gains were applied to compensate for weaker stimuli. For each signal, its power in the frequency

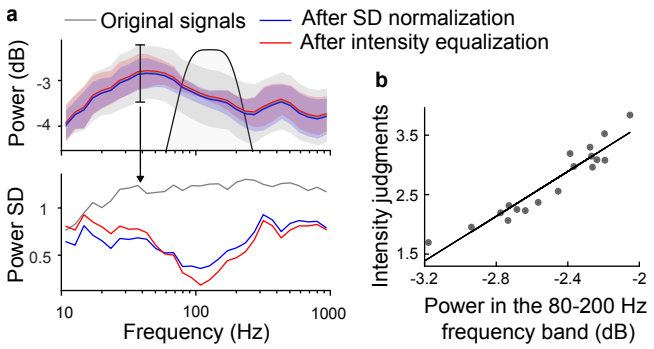

Fig. 2. **a.** Results of the equalization of the perceived intensity for the 18 vibrations. The original signals from the Kirsch et al. database [46] are displayed in grey. These signals were equalized in standard deviation prior to the experiment (in blue). The signals equalized in intensity by the participants are shown in red. On the top, the mean power spectra of the signals are displayed in continuous line and the standard deviation in shaded area. On the bottom, only the standard deviation is displayed for comparison. **b.** Linear modeling of intensity prediction ($R^2 = 0.90$). The dots represent the mean intensity judgments for each stimulus with respect to its power in the 80-200 Hz frequency range, calculated with the filter displayed as a black line on the top-left (not to scale).

band of interest (80-200 Hz) was computed using a 2^{nd} order Butterworth bandpass filter. The linear regression curve ($R^2 = 0.90$, df=16), displayed in Fig. 2.b, demonstrates that we can accurately model the intensity judgments from the power in that frequency band. The goodness of fit dropped when increasing the frequency bandwidth, to $R^2 = 0.07$ for the whole frequency band (10 to 1000 Hz). This means that, for noisy vibrations, the perceived intensity is directly determined by the power of the signal in the 80 to 200 Hz frequency band, and weakly influenced by the power of higher and lower frequencies. This frequency band matches the optimal sensitivity band outlined in Fig. 1.b, which corresponds to the combination of human tactile sensitivity and the actuator frequency response.

V. Spectral Content Perception

The main experiment focuses on the perception of the spectral content. With the stimuli now having equal intensity, it becomes more straightforward to explore how other attributes influence vibrotactile perception. We used a protocol based on dissimilarity ratings, wherein participants were asked to rate the (di-)similarities between pairs of stimuli. This technique is a gold standard in the field of auditory perception to investigate the perception of sound timbre for musical instruments [49]–[51], see [52] for a recent meta-analysis. It enables to model perception within a reduced dimensional space and to identify the main perceptual dimension within this space. Through further analysis, links can be established between these dimensions and signal features, ultimately unveiling the association between signal and perception.

This approach has also demonstrated its potential to exhibit the perceptual dimensions of touch [53]–[55]. For textures, perceptual differences can be predicted using the comprehensive data collected during their tactile exploration, encompassing forces, vibrations, and velocity [56]. A recent study [57] proposed a physiology-based model to predict the dissimilarities in vibrations designed with varying amplitudes, frequencies and modulation rates.

Here, as proposed in [52], we first gathered dissimilarity ratings from human participants and then trained a linear model, fully interpretable, to predict the obtained dissimilarities based on the spectral representations of the stimuli. We assume that once the proposed model is trained to mimic human responses, we can interpret its behavior to gain insight on human perception.

A. Participants

The dissimilarity rating experiment was performed by 18 participants, 3 females and 15 males, from 21 to 55 years old (mean=28), 1 left-handed and 17 right-handed.

B. Protocol

Participants had to evaluate the perceptual differences of the vibrations through pairwise comparisons. The 18 stimuli were presented against each other, resulting in 153 pairs. As control conditions, the stimuli were also compared with themselves, leading to 171 pairs in total, presented in random order. At each trial, participants felt the two vibrations successively. The vibrations lasted for 1 second, and were separated by 500 ms of silence. Each sequence was played once. Participants were asked to judge the perceived dissimilarities between the pairs and to report their ratings on a scale from ”very similar” to ”very dissimilar” thanks to slider on a visual interface. The interface converted the slider position into a dissimilarity rating between 0 and 1 (0=very similar, 1=very dissimilar). Prior to the experiment, a familiarization session was carried out for each participant, involving 30 random pairs that included the 18 stimuli at least once. It enabled the participants to familiarize themselves with the task and the stimuli, and to create their own internal rating scale. The experiment lasted for about 45 min.

C. Results

Regarding dissimilarity ratings, participants were less consensual than for the intensity ratings, as revealed by the pairwise correlation scores between participants (Pearson’s r: M=0.51, SD=0.12, df=169 for the 162 correlations, all significant $p<0.05$). Yet, as the evaluation strategies were coherent across participants, we averaged their dissimilarity scores. The correlations between the average and the participants’ scores (Pearson’s r: M=0.73, SD=0.09, df=169 for the 18 correlations, all significant $p<0.05$) showed that the mean dissimilarity scores well reflect participants’ judgments. The following analysis therefore focuses on the between-participants mean dissimilarity scores.

For the control pairs with stimuli compared to themselves, the dissimilarity scores were low (M=0.12, SD=0.07), but still higher than 0. This means that the task was quite difficult and the intensity-equalized vibrations felt already pretty similar. It also means that vibrations with a dissimilarity score below 0.12 could be considered as perceptually identical.

D. Predicting dissimilarities from spectral representation

Since the stimuli were considered as stationary, i.e. they were felt to be relatively constant during the 1 second presentation time, we chose to represent the signals by their power spectra to focus on the spectral properties. We designed a sparse spectrum representation computed as the power by frequency band using a filter bank with 30 second-order bandpass-filters logarithmically spaced between 10 and 1000 Hz, as illustrated in Fig. 3. This frequency band covers the human vibration sensitivity range [19] and the logarithmic distribution follows Weber's law of vibrotactile frequency perception [23]. The number of filters is the consequence of the Just-Noticeable frequency Difference (JND=17% [23]) between two filters: $N = log(f_{max}/f_{min})/log(1+JND) \approx 30$ filters. This filter distribution minimizes the number of filters while still separating distinguishable frequencies, leading to a sparse perception-based spectral representation.

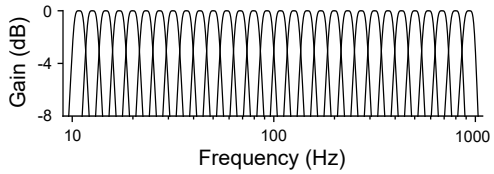

Fig. 3. Frequency response of the filter bank used to compute the power by frequency band of the vibration signals. The 30 filters are logarithmically spaced following Weber's law to match human perception.

Continuous spectral patterns were observed, indicating that the signal representation could be further reduced. Therefore, a Principal Component Analysis (PCA) was conducted to provide a data-driven dimensionality reduction. The PCA was trained on the power by frequency band of all the signals in the Kirch et al. database [46] (composed of 281 signals). We found that the first 8 dimensions of the PCA were sufficient to capture 95% of the variance of stimuli in the database. The explained variance for each axis was successively 63.3%, 15.6%, 7.6%, 2.7%, 2.6%, 1.8%, 1.0%, and 0.7%. This enabled us to decompose the experiment's 18 stimuli into the most important spectral patterns of friction-induced vibrations presented in Fig. 5.a. We can already note that the first component V_1 encodes a balance between high and low frequencies, and other components encode spectral patterns of increasing complexity.

Based on the spectral patterns representations of the vibration, we constructed a model to predict the participants' dissimilarity ratings from the experiment.

The methodology used to compare two vibrotactile signals x and y is presented in Fig. 4. The algorithm takes the two temporal signals as input. Firstly, the power spectra are measured by computing the power by frequency band P_x and P_y ($dim = 30$, using the filter bank described previously). Then, the power spectra are projected on the PCA basis V ($dim = 8\times30$) to obtain the PCA scores T_x and T_y ($dim = 8$):

$$T_x = VP_x \quad \text{and} \quad T_y = VP_y \tag{1}$$

Next, the local distance d in each PCA dimension j is calculated as the absolute difference between the two PCA scores.

$$d_j(x,y) = \sqrt{(T_{xj} - T_{yj})^2} \tag{2}$$

Finally, the global dissimilarity D between the two vibrations is computed as a weighted sum of the local distances:

$$D(x,y) = \sum_{j=1:8} w_j d_j(x,y) \tag{3}$$

The weights w_j were optimized so that the dissimilarity prediction best matched the participants' dissimilarity ratings. It was performed by a Lasso regression, a multiple regression model with regularization that performs variable selection to facilitate the interpretation of the results.

To train the model, while the best amount of penalization was chosen by cross validation (10-fold), the weights were optimized on a training set of 122 dissimilarity ratings (80% of the data). The prediction performance of the model was

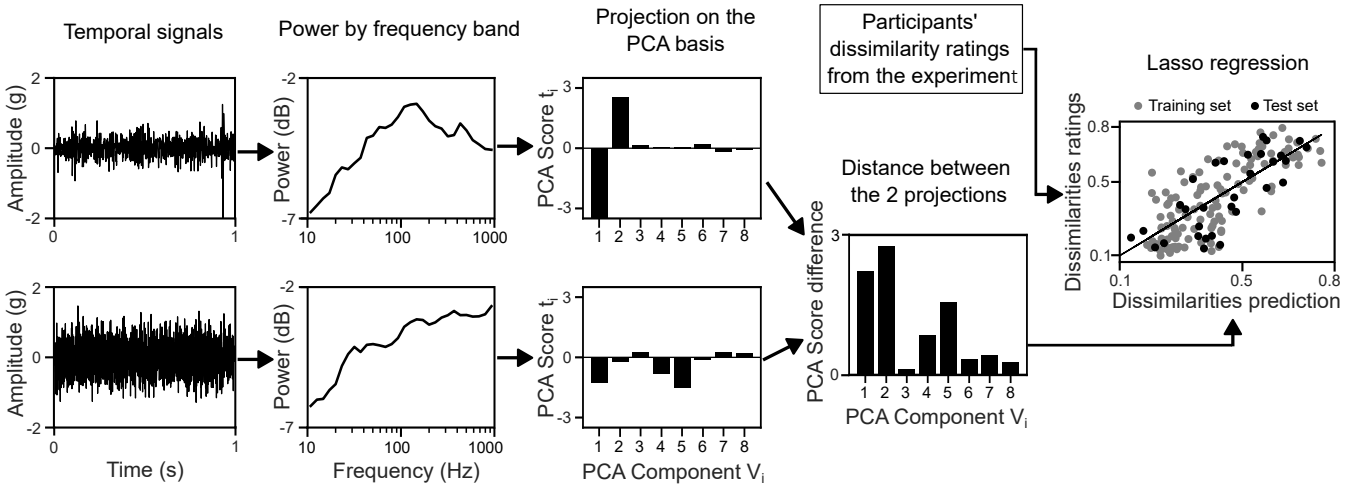

Fig. 4. Principle of the dissimilarity prediction model, illustrated here to compare two vibrations. The spectra are projected on a basis of spectral patterns given by the PCA. Then, the distance between these two representations is computed. Weights are associated to each component and are optimized so that the mathematical distances fit the perceptual dissimilarities as closely as possible.

then evaluated on a test set of 23 dissimilarity ratings (20% of the data). The prediction performance was measured as the average of 100 repetitions of the Lasso regression with 100 random assignments of the ratings to the training and test sets. In practice, this was done by changing the initialization of the random generator 100 times.. The model exhibited a coefficient of determination $R^2 = 0.54$ (mean of 100 repetitions, SD=0.11). The prediction score is not as high as for the intensity model, but the participants were also less coherent. Still, more than half of the variance is explained.

E. Interpretation of the dissimilarity model

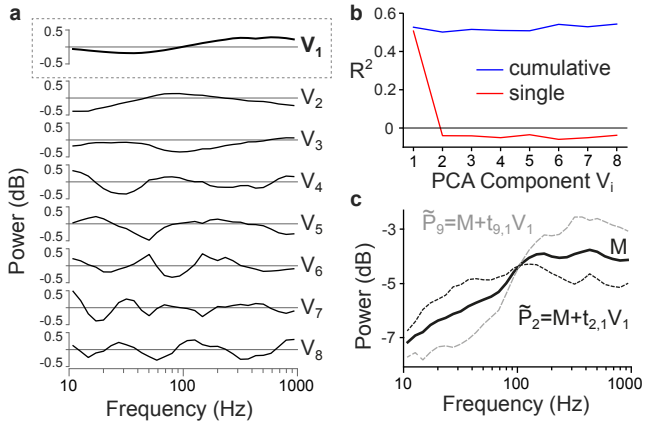

Fig. 5. **a.** Representation of the 8 PCA components V_i to highlight spectral patterns. **b.** Contribution of each PCA component to the prediction quality. The red curves show the evolution of the coefficient of determination when the prediction is performed using the projection of the signals on the i^{th} component only. The blue curves show the evolution of the coefficient of determination when the prediction is performed with projection on the basis made of the i first components. **c.** Example of signal spectra (for stimuli $s = 2$ and $s = 9$) reconstructed only from the score on the first PCA component $t_{s,1}$. M is the averaged spectrum of the database.

The linearity of the proposed model offers the advantage of being easily interpretable. We can assess the importance of a PCA component by examining the impact of its withdrawal on the prediction quality R^2.

Fig. 5.b shows that the prediction is only based on the projection of the stimuli on the first PCA component V_1. Indeed, the prediction scores R^2 are below 0 when other components are considered alone (red curve) and R^2 does not increase as more PCA components are included (blue curve). This means that, to mimic human judgments, the model only needs information about the first PCA component, i.e. the balance between high and low frequencies.

A classical multidimensional scaling analysis (MDS) was conducted to place each stimulus in a 2-dimensional space so that the distances between stimuli, regarded here as perceptual dissimilarities, would show up as clearly as possible. The blue points in Fig. 6 show the result of the MDS performed on the participant's mean dissimilarity ratings. The projection of the first PCA component V_1 in this space demonstrates its high correlation (Pearson's $r = 0.94$, df=16, p<0.01) with the first MDS dimension (and $r = -0.01$, df=16, p=0.95 with the second dimension). In comparison, the other PCA components are much lesser correlated with the MDS dimensions (Pearson's $r \in [-0.4, 0.5]$ for all correlations, df=16, non-significant). The first dimension of the MDS is known to reveal the main stimulus property rated by the participants. This analysis confirms the importance of the V_1 basis, the balance between the high and low frequencies, in vibrotactile perception.

An MDS analysis was also conducted on the dissimilarities given by the model. It was combined with a Procrustes superimposition, a combination of translations, rotations and uniform scaling, to match the two spaces and enable comparison. The positions resulting from predicted dissimilarities are displayed in blue in Fig. 6. The error distances, in grey, show that the model better predicts dissimilarities with certain stimuli.

Many participants spontaneously reported after the experiment that a few stimuli presented temporal intensity variations. Therefore, we computed a non-stationarity metric defined as the standard deviation of perceived intensity over time, using the previous model of intensity estimation by time windows of 100 ms, with 50 ms overlap. This metrics appeared as well correlated with the second dimension of the MDS (Pearson's $r = 0.86$, df=16, p< 0.01).

It is also interesting to note that vibrations induced by friction on similar materials are located in the same zones of the perceptual space.

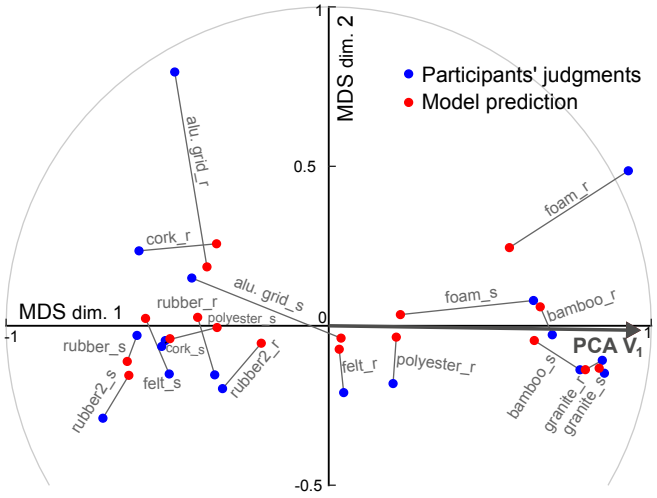

Fig. 6. Two-dimensional perceptual space resulting from the MDS. The position of the 18 stimuli according to the participant's judgments is plotted in blue. Their position resulting from a second MDS based on the dissimilarities predicted by the model is plotted in red. The error between the two is revealed by the grey lines. The arrow represents the projection of the first PCA component V_1 in the MDS space and the unity circle is plotted for comparison. The materials underlying the friction-induced vibrations are indicated, as is the type of probe: round (r) or spiked (s).

VI. Validation through analysis by synthesis

The previous experiment highlighted the prominence of the balance between the high and low frequencies to feel dissimilarities between vibrations. We will now evaluate the validity of the previous insights by evaluating vibrations recreated by synthesis using this frequency balance. This section also shows how this property of the human tactile sensory system can be leveraged to propose sparse vibration synthesis.

A. Synthesis procedure

The proposed approach is based on analysis by synthesis, a well-known approach in audio processing, that consists in extracting parameters from a given natural sound to reconstruct it using algorithmic techniques [58], [59]. Analyzing the structure of the algorithm with respect to the quality of synthesized sound has proven its potential for probing human perception. [60].

The analysis-synthesis procedure is described in Fig. 7. The original vibration, a stationary signal, is analyzed by filtering it with the logarithmic 30-filters bank. The power of each frequency band $P_{1:30}$ is then computed and projected on V_1 the first component of the PCA (the PCA that have been performed on the whole database [46]). The input vibration is thus encoded as a unique scalar: the first principal component score t_1.

The decoding part requires the average power by frequency band of the database M and the first component of the PCA V_1. These features are the same for each vibration in the database. From the input t_1, the inverse transform is performed to obtain an approximation of the powers of the 30 frequency bands $\tilde{P}_{1:30}$. The Fig. 5.c shows examples of two signal spectra reconstructed by this method.

The synthesis procedure takes noise as an input, filters it, measures the power in each frequency band and applies the desired values $\tilde{P}_{1:30}$. The output is obtained by combining the subbands, but this simple process does not produce exactly the desired signal. Therefore, an iteration procedure is conducted until the powers of the bands perfectly match $\tilde{P}_{1:30}$, and the algorithm outputs the synthesized signal.

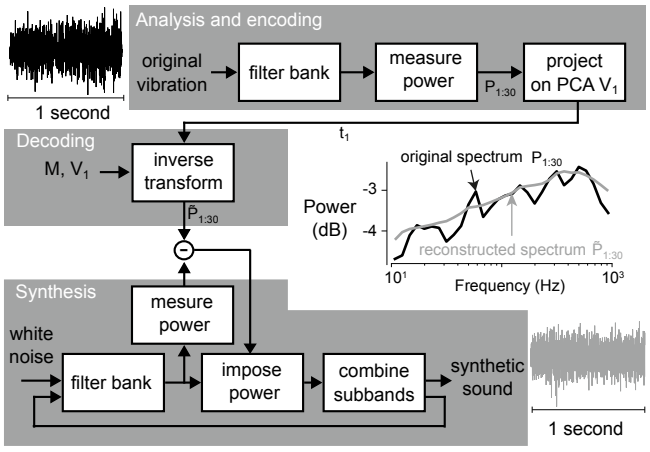

Fig. 7. Procedure of the analysis-synthesis. The power by frequency band of the original vibration is projected on the axis V_1 (first component of the PCA on the whole vibration database). The synthetic signal is based on noise, to which the powers of the frequency bands from the first PCA component are imposed. This algorithm was derived from [60].

B. Protocol

A third experiment was conducted to evaluate the quality of the sparse synthesis that we developed. The protocol was similar to a MUSHRA (Multiple Stimuli with Hidden Reference and Anchor). At each step, the participants were instructed to rate and rank the similarities between a reference stimulus (the original vibration) and 5 stimuli:

- the synthesized version of the vibration S_1 using the proposed sparse algorithm (with the first component V_1 only)
- a synthesized version of the vibration $S_{1:8}$ with all the information (with the 8 components $V_{1:8}$)
- a synthesized version of the vibration $S_{2:8}$ with all the information except the first component (with $V_{2:8}$)
- the original vibration O as control (the hidden reference)
- a synthesized vibration based on the mean power spectrum M, used as a control (anchor) and identical for all the stimuli.

The stimuli were 1 second vibrations that the participants could play as many times as they wanted by clicking on the corresponding buttons. The interface displayed 5 sliders (in random position) to rate the 5 versions of the stimuli on a scale from "very dissimilar" (0) to "very similar" (1). This task was performed 18 times for the 18 stimuli from the previous experiment, presented in random order. For each stimulus, one of the 5 versions was exactly the same as the original stimulus, also known as the hidden reference, and the participants were asked to rate it to 1. Preliminary tests showed that it was preferable not to ask to score the most dissimilar stimuli to 0 (as in classic MUSHRA methods), since some of the original stimuli were very close to the anchor M.

C. Participants

The evaluation of the synthesized vibrations was performed by 15 participants, 6 females and 9 males, from 22 to 55 years old (mean=30), all right-handed.

D. Results

The results of the comparison between the original vibrations and the synthesized versions are presented in Fig. 8.

A two-ways repeated measures ANOVA, showed a significant effect ($\alpha = 0.05$) of the version: F(14,4,17)=33.0, p=$1e^{-16}$) and Tuckey post-hoc tests are summarized in Fig. 8. Most importantly, it showed that S_1 was significantly rated as more similar to the reference than $S_{2:8}$ (p<0.01) but not less similar than $S_{1:8}$ (p= 0.76). A test of equivalence [61] showed that S_1 and $S_{1:8}$ were equivalent (at $\alpha = 0.05$) in a ± 0.15 interval on the 0 to 1 similarity scale, demonstrating that the first PCA component V_1 is necessary and sufficient to capture all the spectral information used by subjects to rate the dissimilarity between two textures.

However, the post-hoc tests highlighted significant differences between the synthesized versions and the original signals (p< 0.01). They were mainly due to three stimuli (*alu_grid_r*, *alu_grid_s* and *foam_r*) that showed large discrepancies between S_1 and O. This means that certain stimuli were not accurately synthesized by the algorithm. These stimuli were also the ones that were not well predicted by the model in Fig. 6. Moreover, the differences in similarity ratings between O and S_1 for each stimulus were correlated (Pearson's $r = 0.72$, df=16, p<0.01) with the prediction error of the

model in the MDS space (grey lines in Fig. 6). Also, we found that the differences in similarity ratings between S_1 and M for each stimulus were correlated (Pearson's $r = 0.63$, df=16, p< 0.01) with the absolute values of their projection on the first dimension $|t_1|$. In other words, the more the frequency balance differs from the mean spectrum, the more the stimulus is perceived as dissimilar to M. This is an additional argument supporting the importance of the frequency balance in vibrotactile perception.

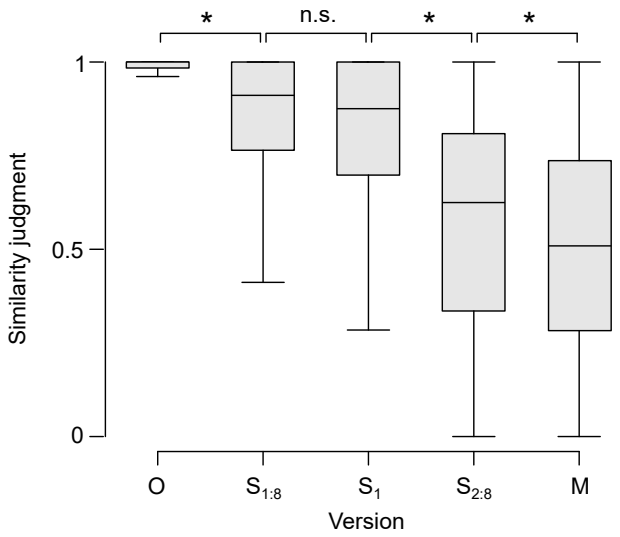

Fig. 8. Evaluation of the quality of the synthesized stimuli. The 18 original vibrations were compared against themselves (O) and synthesized versions based on all the PCA dimensions ($S_{1:8}$), based only on the first dimension (S_1), based on all the dimensions except the first one ($S_{2:8}$) or based on the mean spectrum (M).

VII. Discussion

The present work investigated the perception of stationary noisy vibrations, such as recordings of friction-induced vibrations with a constant exploration speed and aimed at unveiling their perceptually relevant signal structures.

A. Intensity perception

Firstly, the perception of intensity, the primary attribute of vibration, was investigated. We discovered that, akin to simple sinusoidal vibrations, it is possible to gauge the perceived strength of a noisy vibration from its signal. The results revealed that the intensity was directly proportional to the power of the frequency band to which the subjects were the most sensitive. The frequency band was between 80 and 200 Hz and corresponded to a combination of the human tactile sensory curve for sine waves [19] and the actuator's frequency response. The power in the higher and lower frequencies, to which we are less sensitive, had a negligible effect on intensity perception. This model can be easily adapted to other vibrating devices just by measuring their frequency response. Equalizing the intensity of the vibration enabled us to remove this essential characteristic from subsequent experiments, enabling a detailed exploration of other attributes such as the spectral content.

B. Spectral content perception

The power spectra of the vibrations were computed thanks to a perception-based logarithmic filter bank. The principal component analysis of the spectrum of 281 recorded vibrations provided a data driven basis whose first axis encoded a balance between the low and high frequencies. The projection of the vibrations solely onto this axis (a single scalar) was sufficient to build a model of dissimilarity prediction, and showed a high correlation with the first dimension of the MDS. Moreover, vibrations that were re-synthesized using only this axis were indistinguishable from vibration re-synthesized using the whole spectral information. These findings clearly demonstrate that, regarding human perception, the spectral content can be characterized by the balance between the low and high frequencies only. This outcome joins previous works that showed that the spectrum of friction induced vibrations could be modelled by a $1/f^{\alpha}$ function [62]. The parameter α also represents the balance between low and high frequencies and has been shown to correlate with perceptual judgements of texture categories [63] and texture pleasantness [64]. From a physiological perspective, the frequency balance could reflect the dual neural mechanism of flutter-vibration [65], [66]. The balance could be the relative importance of the activation of FA I fibers (Meissner's corpuscles) in response to frequencies in the flutter range (< 60 Hz) compared to the activation of FA II fibers (Pacinian corpuscles) in the vibration range (> 60 Hz).

However, the model of dissimilarity estimation was not perfectly accurate, especially with some stimuli such as *alu_grid_r*, *alu_grid_s* and *foam_r* as shown in Fig. 6. Similarly, the synthesis algorithm failed to properly reproduce these stimuli. We believe that these differences arose from the temporal evolution of these signals, which might not be negligible, considering the variability in movement during the recording and the unevenness of certain surfaces. In particular, some participants reported that *alu_grid_r*, *alu_grid_s* were felt as not constant. The metric of non-stationarity suggests that the second dimension of the MDS may reflect temporal variations. For vibrations that deviate from the assumption of temporal homogeneity, important information lies in the phase and is therefore not captured by the power by frequency band on which the model is based. This shows the limitations of the present analysis to stationary signals and further work will explore ways of including temporal variations (such as time windowing).

Another limitation of the analysis is its dependence on the data base. The vibrations have all been recorded with the same device and the same protocol and the three experiments have been performed with 18 stimuli only. However, the projection axis V_1 at the core of the analysis has been obtained by a PCA on 281 signals, presenting a wide variety of textures, exploration tools and exploration speeds. We are confident that the analysis procedure and the synthesis principle can be generalized to other vibration databases.

C. Sparse synthesis and compression

The benefits of the study are considerable in terms of vibration synthesis and vibration compression. It enables the

creation of a sparse, perception-based synthesizer that can replicate a diverse range of texture-like vibrations using only two controls: intensity and frequency balance.

Very high compression can also be attained through analysis-synthesis, by extracting intensity and frequency balance from the original signal and then using these parameters to synthesize a new vibration that closely resembles the original. For example, in our case, the 18 signals recorded over 1 second at a sampling rate of 2800 Hz yield a total of $18 \times 2800 = 50400$ data points. The synthesis algorithm requires the power of the 30 frequency bands for the mean spectrum (M) and the first PCA component (V_1), and the projection scores (t) for each stimulus. This results in a total of $30 + 30 + 18 \times 1 = 78$ data points for the encoded data, making up 0.15% of the initial data, without considering quantization. Since we are considering only stationary signals, the signal duration does not impact the encoded data size. However, further work based on time windowing is necessary to compress vibrations with temporal variations.

The findings of this study hold significant value for a wide range of applications that involve delivering vibrations to users. Everyday human-computer interfaces, such as smartphones, tablets, and wearables, could greatly benefit from a sparse synthesizer capable of reproducing complex vibrations. When combined with algorithms from existing literature to adapt to the user's finger movement in the case of active touch [67], [68], the compression method could prove useful for replicating textures in virtual environments. In that case, the decrease in intensity due to tactile suppression during active movement must be considered [69]. Moreover, synthesis approximations might be negligible as long as the vibrations remain plausible in a given context and correspond to the user's expectations [70].

Vibration compression is also valuable for enhancing the musical experience through multichannel vibrations, especially for audiences with hearing impairments. The vibration analysis framework could be effectively integrated into sensors of robotic arms to interpret textures in a manner similar to humans, or to render the essence of tactile information to individuals with prosthetic hands.

VIII. Conclusion

In this paper, we investigated the perception of noisy, stationary vibrations. The stimuli were taken from a data base of friction-induced vibrations recorded during the exploration of textures at constant velocity. We initially examined the perception of the intensity of these vibrations and developed a new model based on the power within the frequency bandwidth of human optimal sensitivity. Subsequently, we investigated other perceptual attributes through dissimilarity experiments. These experiments revealed that the essential information regarding the spectral content resides in the balance between high and low frequencies. The observed discrepancies were attributed to specific stimuli that were not entirely homogeneous, wherein pertinent information also lies in the phase. In summary, the perception of purely stationary noisy vibrations is driven by two main attributes: 1) intensity and 2) high/low frequency balance. We showcased the potential of these findings for sparse analysis and synthesis of vibrations, as well as perception-based compression. Such applications can be valuable for any device delivering complex vibratory feedback.

IX. Data and Code Availability

The code and experimental data are available on GitLab: https://gitlab.prism.cnrs.fr/bernard.prism.cnrs.fr/vibrotactile-perception-frequency-balance/

References

[1] S.-Y. Kim, K. Y. Kim, B. S. Soh, G. Yang, and S. R. Kim, "Vibrotactile rendering for simulating virtual environment in a mobile game," *IEEE Transactions on Consumer Electronics*, vol. 52, no. 4, pp. 1340–1347, 2006.

[2] T. Singhal and O. Schneider, "Juicy haptic design: Vibrotactile embellishments can improve player experience in games," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–11.

[3] A. Israr and F. Abnousi, "Towards pleasant touch: vibrotactile grids for social touch interactions," in *Extended abstracts of the 2018 CHI conference on human factors in computing systems*, 2018, pp. 1–6.

[4] M. Lee, G. Bruder, and G. F. Welch, "Exploring the effect of vibrotactile feedback through the floor on social presence in an immersive virtual environment," in *2017 IEEE Virtual Reality (VR)*. IEEE, 2017, pp. 105–111.

[5] H. Seifi, K. Zhang, and K. E. MacLean, "Vibviz: Organizing, visualizing and navigating vibration libraries," in *2015 IEEE World Haptics Conference (WHC)*. IEEE, 2015, pp. 254–259.

[6] H. Culbertson, J. Unwin, and K. J. Kuchenbecker, "Modeling and rendering realistic textures from unconstrained tool-surface interactions," *IEEE transactions on haptics*, vol. 7, no. 3, pp. 381–393, 2014.

[7] L. Pantera, C. Hudin, and S. Panëels, "Lotusbraille: Localised multi-finger feedback on a surface for reading braille letters," in *2021 IEEE World Haptics Conference (WHC)*. IEEE, 2021, pp. 973–978.

[8] C. Richards, R. Cahen, and N. Misdariis, "Designing the balance between sound and touch: methods for multimodal composition," in *19th Sound and Music Computing Conference (SMC 2022)*, 2022, pp. 426–433.

[9] G. Richard, T. Pietrzak, F. Argelaguet, A. Lécuyer, and G. Casiez, "Multivibes: What if your vr controller had 10 times more vibrotactile actuators?" in *ISMAR 2023-22nd IEEE International Symposium on Mixed and Augmented Reality*, 2023, pp. 1–10.

[10] M. Hollins and S. J. Bensmaïa, "The coding of roughness." *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, vol. 61, no. 3, p. 184, 2007.

[11] A. I. Weber, H. P. Saal, J. D. Lieber, J.-W. Cheng, L. R. Manfredi, J. F. Dammann III, and S. J. Bensmaia, "Spatial and temporal codes mediate the tactile perception of natural textures," *Proceedings of the National Academy of Sciences*, vol. 110, no. 42, pp. 17 107–17 112, 2013.

[12] L. R. Manfredi, H. P. Saal, K. J. Brown, M. C. Zielinski, J. F. Dammann III, V. S. Polashock, and S. J. Bensmaia, "Natural scenes in tactile texture," *Journal of neurophysiology*, vol. 111, no. 9, pp. 1792–1802, 2014.

[13] H. B. Barlow *et al.*, "Possible principles underlying the transformation of sensory messages," *Sensory communication*, vol. 1, no. 01, pp. 217–233, 1961.

[14] V. Hayward, "Is there a 'plenhaptic'function?" *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 366, no. 1581, pp. 3115–3122, 2011.

[15] L. Willemet, N. Huloux, and M. Wiertlewski, "Efficient tactile encoding of object slippage," *Scientific Reports*, vol. 12, no. 1, p. 13192, 2022.

[16] I. Hwang and S. Choi, "Perceptual space and adjective rating of sinusoidal vibrations perceived via mobile device," in *2010 IEEE Haptics Symposium*. IEEE, 2010, pp. 1–8.

[17] L. Felicetti, C. Sutter, E. Chatelet, A. Latour, L. Mouchnino, and F. Massi, "Tactile discrimination of real and simulated isotropic textures by friction-induced vibrations," *Tribology International*, vol. 184, p. 108443, 2023.

[18] J. C. Craig, "Difference threshold for intensity of tactile stimuli," *Perception & Psychophysics*, vol. 11, no. 2, pp. 150–152, 1972.

[19] R. T. Verrillo, A. J. Fraioli, and R. L. Smith, "Sensation magnitude of vibrotactile stimuli," *Perception & Psychophysics*, vol. 6, no. 6, pp. 366–372, 1969.

[20] Y. Wang, M. J. Turner, and W. T. Hewitt, "Constructing and evaluating vibration magnitude models for visualization," in *2008 Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. IEEE, 2008, pp. 457–464.

[21] Y. Yoo, I. Hwang, and S. Choi, "Perceived intensity model of dual-frequency superimposed vibration: Pythagorean sum," *IEEE Transactions on Haptics*, vol. 15, no. 2, pp. 405–415, 2022.

[22] H. Pongrac, "Vibrotactile perception: Differential effects of frequency, amplitude, and acceleration," in *2006 ieee international workshop on haptic audio visual environments and their applications (have 2006)*. IEEE, 2006, pp. 54–59.

[23] H. Pongrac, "Vibrotactile perception: examining the coding of vibrations and the just noticeable difference under various conditions," *Multimedia systems*, vol. 13, no. 4, pp. 297–307, 2008.

[24] I. Hwang, J. Seo, and S. Choi, "Perceptual space of superimposed dual-frequency vibrations in the hands," *PloS one*, vol. 12, no. 1, p. e0169570, 2017.

[25] C. Bernard, J. Monnoyer, and M. Wiertlewski, "Harmonious textures: The perceptual dimensions of synthetic sinusoidal gratings," in *Haptics: Science, Technology, and Applications: 11th International Conference, EuroHaptics 2018, Pisa, Italy, June 13-16, 2018, Proceedings, Part II 11*. Springer, 2018, pp. 685–695.

[26] R. F. Friesen, R. L. Klatzky, M. A. Peshkin, and J. E. Colgate, "Single pitch perception of multi-frequency textures," in *2018 IEEE Haptics Symposium (HAPTICS)*. IEEE, 2018, pp. 290–295.

[27] T.-l. S. Le, G. Bailly, E. Vezzoli, M. Auvray, and D. Gueorguiev, "Tactile sensitivity to the frequency spectrum of complex vibrotactile signals," *bioRxiv*, pp. 2023–11, 2023.

[28] C. Bernard, R. Kronland-Martinet, M. Fery, S. Ystad, and E. Thoret, "Tactile perception of auditory roughness," *JASA Express Letters*, vol. 2, no. 12, 2022.

[29] G. Park and S. Choi, "Perceptual space of amplitude-modulated vibrotactile stimuli," in *2011 IEEE world haptics conference*. IEEE, 2011, pp. 59–64.

[30] C. Bernard, J. Monnoyer, M. Wiertlewski, and S. Ystad, "Rhythm perception is shared between audio and haptics," *Scientific Reports*, vol. 12, no. 1, p. 4188, 2022.

[31] S. J. Bensmaia, "Tactile intensity and population codes," *Behavioural brain research*, vol. 190, no. 2, pp. 165–173, 2008.

[32] I. Birznieks and R. M. Vickery, "Spike timing matters in novel neuronal code involved in vibrotactile frequency perception," *Current Biology*, vol. 27, no. 10, pp. 1485–1490, 2017.

[33] Q. Consigny, A. Paté, and J.-L. Le Carrou, "Perceptual evaluation of the quantization level of a vibrotactile signal," in *International Workshop on Haptic and Audio Interaction Design*. Springer, 2022, pp. 69–78.

[34] M. Wiertlewski, J. Lozada, and V. Hayward, "The spatial spectrum of tangential skin displacement can encode tactual texture," *IEEE Transactions on Robotics*, vol. 27, no. 3, pp. 461–472, 2011.

[35] D. J. Meyer, M. A. Peshkin, and J. E. Colgate, "Modeling and synthesis of tactile texture with spatial spectrograms for display on variable friction surfaces," in *2015 IEEE World Haptics Conference (WHC)*. IEEE, 2015, pp. 125–130.

[36] Y. Toide, M. Fujiwara, Y. Makino, and H. Shinoda, "Sufficient time-frequency resolution for reproducing vibrotactile sensation," *IEEE Transactions on Haptics*, 2023.

[37] S. Okamoto and Y. Yamada, "Lossy data compression of vibrotactile material-like textures," *IEEE transactions on haptics*, vol. 6, no. 1, pp. 69–80, 2012.

[38] R. Hassen, B. Gülecyüz, and E. Steinbach, "Pvc-slp: Perceptual vibrotactile-signal compression based-on sparse linear prediction," *IEEE Transactions on Multimedia*, vol. 23, pp. 4455–4468, 2020.

[39] G. A. Gescheider, R. T. Verrillo, and C. L. Van Doren, "Prediction of vibrotactile masking functions," *The Journal of the Acoustical Society of America*, vol. 72, no. 5, pp. 1421–1426, 1982.

[40] R. Chaudhari, C. Schuwerk, M. Danaei, and E. Steinbach, "Perceptual and bitrate-scalable coding of haptic surface texture signals," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 3, pp. 462–473, 2014.

[41] A. Noll, L. Nockenberg, B. Gülecyüz, and E. Steinbach, "Vc-pwq: Vibrotactile signal compression based on perceptual wavelet quantization," in *2021 IEEE World Haptics Conference (WHC)*. IEEE, 2021, pp. 427–432.

[42] L. Nockenberg, A. Noll, S. Panëels, A. B. Dhiab, C. Hudin, and E. Steinbach, "Mvibcode: Multi-channel vibrotactile codec using hierarchical perceptual clustering," *IEEE Transactions on Haptics*, 2023.

[43] R. Hassen and E. Steinbach, "Subjective evaluation of the spectral temporal similarity (st-sim) measure for vibrotactile quality assessment," *IEEE Transactions on Haptics*, vol. 13, no. 1, pp. 25–31, 2019.

[44] E. Muschter, A. Noll, J. Zhao, R. Hassen, M. Strese, B. Gülecyüz, S.-C. Li, and E. Steinbach, "Perceptual quality assessment of compressed vibrotactile signals through comparative judgment," *IEEE Transactions on Haptics*, vol. 14, no. 2, pp. 291–296, 2021.

[45] D. A. Burns, R. L. Klatzky, M. A. Peshkin, and J. E. Colgate, "The single-pitch texel: A flexible and practical texture-rendering algorithm," *PNAS Nexus*, vol. 3, no. 1, p. pgad452, 2024.

[46] J. Kirsch, A. Noll, M. Strese, Q. Liu, and E. Steinbach, "A low-cost acquisition, display, and evaluation setup for tactile codec development," in *2018 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE)*. IEEE, 2018, pp. 1–6.

[47] N. Huloux, C. Bernard, and M. Wiertlewski, "Estimating friction modulation from the ultrasonic mechanical impedance," *IEEE transactions on haptics*, vol. 14, no. 2, pp. 409–420, 2020.

[48] L. Felicetti, E. Chatelet, B. Bou-Saïd, A. Latour, and F. Massi, "Investigation on the role of the finger transfer function in tactile rendering by friction-induced-vibrations," *Tribology International*, vol. 190, p. 109018, 2023.

[49] S. McAdams, S. Winsberg, S. Donnadieu, G. De Soete, and J. Krimphoff, "Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes," *Psychological research*, vol. 58, pp. 177–192, 1995.

[50] J. M. Grey, "Multidimensional perceptual scaling of musical timbres," *the Journal of the Acoustical Society of America*, vol. 61, no. 5, pp. 1270–1277, 1977.

[51] M. Barthet, P. Guillemain, R. Kronland-Martinet, and S. Ystad, "From clarinet control to timbre perception," *Acta Acustica united with Acustica*, vol. 96, no. 4, pp. 678–689, 2010.

[52] E. Thoret, B. Caramiaux, P. Depalle, and S. Mcadams, "Learning metrics on spectrotemporal modulations reveals the perception of musical instrument timbre," *Nature Human Behaviour*, vol. 5, no. 3, pp. 369–377, 2021.

[53] W. M. B. Tiest and A. M. Kappers, "Analysis of haptic perception of materials by multidimensional scaling and physical measurements of roughness and compressibility," *Acta psychologica*, vol. 121, no. 1, pp. 1–20, 2006.

[54] K. MacLean and M. Enriquez, "Perceptual design of haptic icons," in *Proc. of EuroHaptics*, 2003, pp. 351–363.

[55] D. Ternes and K. E. MacLean, "Designing large sets of haptic icons with rhythm," in *Haptics: Perception, Devices and Scenarios: 6th International Conference, EuroHaptics 2008 Madrid, Spain, June 10-13, 2008 Proceedings 6*. Springer, 2008, pp. 199–208.

[56] B. A. Richardson, Y. Vardar, C. Wallraven, and K. J. Kuchenbecker, "Learning to feel textures: Predicting perceptual similarities from unconstrained finger-surface interactions," *IEEE Transactions on Haptics*, vol. 15, no. 4, pp. 705–717, 2022.

[57] C. Lim and G. Park, "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*, 2023, pp. 1–15.

[58] J.-C. Risset and D. L. Wessel, "Exploration of timbre by analysis and synthesis," in *The psychology of music*. Elsevier, 1999, pp. 113–169.

[59] R. Kronland-Martinet, P. Gillemain, and S. Ystad, "From sound modeling to analysis-synthesis of sounds," in *Proceedings of Workshop on Current Research Directions in Computer Music*, 2001, pp. 217–224.

[60] J. H. McDermott and E. P. Simoncelli, "Sound texture perception via statistics of the auditory periphery: evidence from sound synthesis," *Neuron*, vol. 71, no. 5, pp. 926–940, 2011.

[61] S. Wellek, *Testing Statistical Hypotheses of Equivalence and Noninferiority. Second edition, § 6.2*. Chapman and Hall/CRC, 2010.

[62] M. Wiertlewski, C. Hudin, and V. Hayward, "On the 1/f noise and non-integer harmonic decay of the interaction of a finger sliding on flat and sinusoidal surfaces," in *2011 IEEE World Haptics Conference*. IEEE, 2011, pp. 25–30.

[63] M. Toscani and A. Metzger, "A database of vibratory signals from free haptic exploration of natural material textures and perceptual judgments (viper): analysis of spectral statistics," in *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*. Springer, 2022, pp. 319–327.

[64] A. Klöcker, M. Wiertlewski, V. Théate, V. Hayward, and J.-L. Thonnard, "Physical factors influencing pleasant touch during tactile exploration," *Plos one*, vol. 8, no. 11, p. e79085, 2013.

[65] W. H. Talbot, I. Darian-Smith, H. H. Kornhuber, and V. B. Mountcastle, "The sense of flutter-vibration: comparison of the human capacity with

response patterns of mechanoreceptive afferents from the monkey hand." *Journal of neurophysiology*, vol. 31, no. 2, pp. 301–334, 1968.
[66] V. B. Mountcastle, W. H. Talbot, I. Darian-Smith, and H. H. Kornhuber, "Neural basis of the sense of flutter-vibration," *Science*, vol. 155, no. 3762, pp. 597–600, 1967.
[67] R. F. Friesen and Y. Vardar, "Perceived realism of virtual textures rendered by a vibrotactile wearable ring display," *IEEE Transactions on Haptics*, 2023.
[68] J. M. Romano, T. Yoshioka, and K. J. Kuchenbecker, "Automatic filter design for synthesis of haptic textures from recorded acceleration data," in *2010 IEEE International Conference on Robotics and Automation*. IEEE, 2010, pp. 1815–1821.
[69] E. Fuehrer, D. Voudouris, A. Lezkan, K. Drewing, and K. Fiehler, "Tactile suppression stems from specific sensorimotor predictions," *Proceedings of the National Academy of Sciences*, vol. 119, no. 20, p. e2118445119, 2022.
[70] R. Rosenkranz and M. E. Altinsoy, "A perceptual model-based approach to plausible authoring of vibration for the haptic metaverse," *IEEE Transactions on Haptics*, 2023.

Corentin Bernard graduated from the Ecole Centrale de Marseille in 2017 and holds a Master's degree in acoustics from Aix-Marseille University. In 2022, he earned his Ph.D. for research conducted at Stellantis and two laboratories: the Perception, Representations, Image, Sound and Music Laboratory (PRISM) and the Institute of Movement Sciences (ISM) in Marseille. He is now a post-doctoral researcher at the CNRS in the PRISM laboratory and the Mediterranean Institute of Robotics and Automation (MIRA). His primary research focuses on comprehending human perception of tactile stimulations in order to innovate and enhance human-machine interactions. He utilizes methodologies derived from psychophysics, multisensory integration, and user studies. His contributions resulted in the Best Thesis Award in Engineering Science 2022 conferred by Aix-Marseille University.

Etienne Thoret obtained a Ph.D. from Aix-Marseille University in 2014. He was then a post-doc at McGill University between 2015 and 2018 before a brief post-doctoral assignment at the Ecole Normale Supérieure de Paris. Later, he conducted research funded by the Institute of Language Communication and the Brain (ILCB) for four years, working at the crossroads of the Perception, Representations, Image, Sound, and Music (PRISM) lab and the Laboratoire d'Informatique et Systèmes (LIS) where he started to develop adaptive signal processing representations simulating cochlear processes fed to deep-neural networks to decipher auditory cerebral processes, the current topic of his research as a CNRS researcher (Centre National de la Recherche Scientifique) at the Institut de Neuroscience de la Timone (INT).

Nicolas Huloux graduated from Grenoble-INP Engineering school in 2017 and obtained his PhD in Biorobotics from Aix-Marseille University in 2021 under the supervision of Michaël Wiertlewski. Following this, he initiated the development of a research robotics ecosystem in Ajaccio, concomitantly with the creation of the engineer school MIRA (Mediterrean Institute of Robotics and Automation). His research mainly focuses on multimodal human-machine interactions, with a specific emphasis on tactile perception at the interface.

Sølvi Ystad received her degree as a civil engineer in electronics from NTH (Norges Tekniske Høgskole), Norway in 1992. In 1998 she received her Ph.D. degree in acoustics from the University of Aix-Marseille II, Marseille. After a post doctoral stay at the University of Stanford - CCRMA, California, she obtained a researcher position at the CNRS (Centre National de la Recherche Scientifique) in Marseille in 2002. In 2017 she co-founded the interdisciplinary art-science laboratory PRISM - Perception, Representations, Image, Sound, Music – (www.prism.cnrs.fr) in Marseille. Her research activities mainly focus on investigations of auditory perception through so-called perceptual engineering which consists of crossing different disciplines to link physical and signal knowledge with human perception and cognition.