Single Pitch Perception of Multi-frequency Textures

Rebecca Fenton Friesen*[1], Roberta L. Klatzky[2], Michael A. Peshkin[1], and J. Edward Colgate[1]

Abstract— This study explores people's ability to distinguish spatial complexity in tactile textures, with the eventual goal of reducing the necessary complexity of texture representation for surface display devices. To this end, we tested subjects' ability to perceptually match a reference texture containing two spatial frequency components by adjusting the frequency and amplitude of a single frequency. All textures consisted of spatially varied friction levels on a glass display screen, where friction was modulated via amplitude of ultrasonic vibrations. Resulting chosen single frequencies were systematic, and suggest subjects can identify a single frequency, or tactile pitch, falling somewhere between those of the reference texture. Subject-adjusted frequency is modeled as a function of the reference texture's frequency components and the ratio of their amplitudes.

I. Introduction

As users slide their fingers across variable friction surface haptic displays, they experience complex vibrations via high frequency manipulation of friction forces. These displays are consequently well suited to produce a wide variety of surface texture effects, as people's perception of fine texture is heavily influenced by the vibrations produced on skin during active touch [1], [2]. However, more work is needed to understand how best to represent and manipulate different textural effects on the displays. While we can construct an essentially limitless repertoire of spatial and temporal variations in friction, there is no guarantee that each pattern of friction variation will feel like a unique texture; many friction patterns may contain a redundant amount of complexity. What is the capacity of the perceptual system for the complexity of a texture, and when can we start throwing away or averaging information with negligible effect on the perceptual outcome?

A. Pitch: A Simplification of Rich Frequency Information

One way the perceptual system might simplify a representation of texture is by equating rich spectral vibratory content with fewer frequency components, or at the simplest level, perceiving two frequencies as one. Analogously to hearing, this perceived frequency can be referred to as tactile pitch, and it has been studied by numerous researchers as a basic haptic perceptual parameter. While previous work is limited to describing the pitch response to single frequency vibrations, it points to pitch as a potential perceptual product of more complex vibratory signals.

Georg von Békésy, a researcher best known for his Nobel winning work in audition, was also fascinated by the parallels between auditory and haptic vibrations and first studied tactile pitch in the 1950s [3]. He found that pitch perception had a strong positive correlation with stimulus frequency, but also correlated negatively with stimulus intensity.

In the 1990s, several groups tested a ratio code for pitch, which could account for this dependence on intensity. The ratio code supposes that perceived pitch is determined simply by the ratio of activation between 4 different mechanoreceptor channels, each of which is tuned to a particular temporal frequency of vibration. According to this model, higher stimulus amplitudes saturate the activation of higher tuned channels while increasing activation of lower ones, and perceived pitch should decrease. One study by Roy and Hollins found that this was indeed the case for 3 of their 4 subjects [4]. However, a study by Morley and Rowe found contradictory results; only two of their 8 subjects reported a decrease in pitch when a stimulus amplitude was increased, while 5 perceived a pitch increase [5]. Both groups concluded that the relationship between pitch and intensity can be quite variable across subjects.

B. Perception of Multi-frequency Signals: Previous Work

While no studies to our knowledge have explored pitch perception of multi-frequency signals, several groups have studied other qualities of multi-frequency perception and our ability to perceive spectral complexity.

A 1979 study compared intensity judgments between vibrations composed of one and two frequencies [6], and found that perceived tactile intensity behaved similarly to combined intensities of auditory frequencies within versus across perceptual critical bands. Presumably our sense of touch would have at most four critical bands, corresponding to different types of frequency-tuned mechanoreceptors, while our auditory system has many more due to the cochlea's sensitivity to rich frequency content, raising the question of whether our sense of touch is correspondingly much less acute at detecting spectral richness.

More recently, multiple experiments demonstrate that subjects consistently differentiate between single versus multi-frequency vibrations, even those that should produce equivalent firing ratios across mechanoreceptor populations [7], [8] or are perceptually scaled to have the same intensity values [9]. However, although subjects could identify overall dissimilarity between vibrations of varying spectral content, their perception of equivalent pitch specifically was not investigated; added spectral content may contribute to another perceptual characteristic separate from pitch.

A 2014 study by Yoo et. al. looked at consonance perception of vibrations composed of two temporal frequencies

*author email: r-fentonfriesen@u.northwestern.edu
[1]Department of Mechanical Engineering, Northwestern University, Evanston, IL, USA
[2]Department of Psychology, Carnegie Mellon, Pittsburgh, PA, USA

Haptics Symposium 2018, San Francisco, USA

[10]. Subjects were able to consistently rate the consonance of chords, i.e. vibrations containing two frequency components. Judgments of consonance increased as frequency components were spaced farther apart, but saturated at frequency ratios greater than approximately 2, especially for higher base frequencies. At smaller frequency ratios, especially closer to 1, the resultant low-frequency beat appeared to strongly influence lower judgments of consonance. [9], [7]

C. Spatial Frequency Analysis of Texture

Much of the aforementioned work in tactile vibratory perception uses temporal vibrations as stimuli. Textures, however, generally produce different vibratory frequencies depending on the speed at which the finger moves over surface asperities. Consequently, a well established and common method of characterizing tactile textures is by spatial frequency content, which remains invariant over scan speed [11], [12], [13]. Furthermore, if we limit ourselves to spatial frequencies of approximately 1 cycle/mm or higher (in other words, wavelengths less than 1 mm long), several studies suggest that relative phase between components is perceptually irrelevant [11], [13]. This simplifies spatial analysis considerably, as we then need only register the amplitudes of frequency components.

D. Empirical Study

The following work aims to measure subjects' ability to perceive a single pitch for a set of reference textures composed of two spatial frequency components. The lowest spatial frequency component was set at 1.5 mm^{-1}, which, assuming a minimum typical scanning speed of approximately 40 mm/s, corresponds to a temporal frequency of 60 Hz or above. All higher frequencies were chosen to be at least double that of the base frequency, in order to avoid large amplitudes of beat frequencies and sensations of dissonance.

A direct analogy to auditory pitch would suggest that two separate frequencies spaced this widely would be perceived as two distinct pitches, and subjects would either perform randomly or try to match one frequency or the other. However, results show that subjects are able to assign single frequencies between the two texture's frequencies. These subject-adjusted mean frequencies shift as a function of the amplitude ratio of the texture's two components.

II. EXPERIMENTAL APPARATUS AND METHODS

A. Texture Display Device

The variable friction device used for these experiments operates using ultrasonic friction reduction. The display consists of a one dimensional work space on a pane of glass 104 mm long. The position sensing acuity is 5.3 μm, and commanded friction based on finger position is refreshed at 8333Hz. Textures, defined as arrays of desired friction values as a function of position, were constructed in Matlab (R2013a) before being uploaded to the device via USB. A detailed description of device operation can be found in [14].

While commanded friction values give an approximate representation of actual applied frictional forces, the device
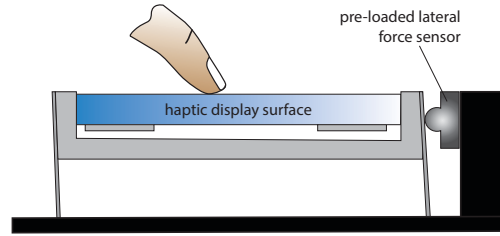

Fig. 1. Haptic display and lateral force sensing setup. The haptic display operates using ultrasonic friction reduction, and friction levels are updated based on 5.3 μm resolution position sensing along the surface. The display is supported by thin brass flexures, which press the frame into a lateral force sensor.

has some limitations. As described in [14], it achieves ultrasonic vibration amplitudes high enough to a produce sizable reduction in friction by operating at resonance, resulting in a limited bandwidth of approximately 130 Hz, above which the amplitudes of temporal variations in friction become increasingly attenuated. Additionally, the relationship between commanded and applied friction can vary both across and within users due to differences in both finger physiological differences and moisture levels. Due to these issues, lateral force vibrations produced by the device on the finger during active touch were recorded in order to obtain an accurate representation of amplitude for each texture's spatial frequency components. During force measurements, finger position measured by the haptic display was output at a rate of 125 Hz.

B. Vibration Measurements

Lateral vibrations produced by the finger on the device during active touch were measured using a piezoresistive force sensor (Honeywell FSS010WNSB). The sensor was preloaded against the side of the texture display frame, which was supported by brass flexures as shown in Fig. 1, allowing it to sense both positive and negative changes in force. Voltage output was recorded using a data acquisition board (NI USB-6211) at a sampling rate of 10 kHz, and a linear conversion coefficient to Newtons was found during calibration. The entire system of glass, frame, and force sensor has a resonant peak at 660 Hz, shown in Fig. 2; consequently, signals were low-pass filtered at 500 Hz using a 6th order Butterworth filter. An example of filtered force recordings appears in Fig. 5. Force measurement processing and all subsequent analysis were performed in Matlab.
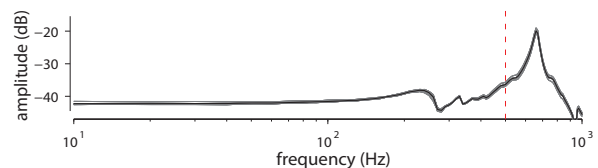

Fig. 2. Frequency Response of the force sensor. Five trials are shown in gray, while average response is in black. Signals are low pass filtered at 500 Hz, indicated by the red dotted line.

C. Reference Textures

Multi-frequency reference textures varied over three different features: lower spatial frequency value, higher frequency value, and the amplitude ratio of the two frequency components. Half of the higher frequencies were first harmonics of the lower frequency, while the other half were slightly higher and non-harmonic. In order to test whether subjects performed more uniformly when trying to match textures with simpler frequency content, three single-frequency reference textures were also included. One of these textures was always presented at the beginning of an experimental block as a training trial, and is therefore omitted from further analysis. All textures had a peak-to-peak amplitude that spanned 90% of the displays full friction range, and were centered on an average friction value at 50% of the range. All reference textures and their characteristics are listed in Table I, and an example of texture construction is shown in Fig. 3.

Fig. 3.   An example of subject-adjusted and reference texture construction.

TABLE I

REFERENCE TEXTURES

Texture #	f_1 (mm^{-1})	f_2 (mm^{-1})	commanded amplitude ratio A_2/A_1
1	1.5	3	0.5
2			1
3			2
4	1.5	4	0.5
5			1
6			2
7	2.5	5	0.5
8			1
9			2
10	2.5	6	0.5
11			1
12			2
13	1.5		N/A
14	5		N/A
training trial	3		N/A

D. Adjustable Textures

Subject-adjusted textures always consisted of a single spatial frequency. Like all reference textures, their average friction level was fixed at a 50% friction reduction level. Possible frequency values linearly spanned 1 to 6 mm^{-1}. In an attempt to mitigate any intensity differences between the reference and adjusted texture that could affect pitch perception, amplitude could also be linearly adjusted, from 0 to 100% of the display's full range. The effects of frequency and amplitude adjustment are shown graphically in 3.

E. Psychophysical Procedure

Twelve subjects (four women, one left-handed, ages 21-30), including the main author and 11 Northwestern students, participated in this study. Half of the subjects were PhD students studying surface haptics and were considered expert users of the haptic device, while the other half were relatively naive to friction-modulated displays. Subject participation was approved by the Northwestern Institutional Review Board, and subjects were paid for their time.
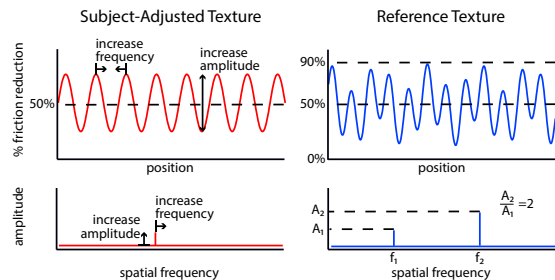
All subjects completed two blocks of trials, each approximately 15 minutes long and separated by at least 30 minutes in order to minimize physical fatigue. At the start of each block, subjects washed and dried their hands and were instructed to feel all textures with their dominant index finger. During the experiment, around-ear headphones playing pink noise were worn to mask any sounds produced by the texture display. Each experimental block normally consisted of 14 reference textures, randomized in presentation order, as well as a practice trial at the start of each block to familiarize subjects with the experimental procedure. (Due to an early error in texture generation, both iterations of texture #8 were presented either in the second block or in an additional block for the first 5 subjects). At the conclusion of the experiment, each texture had been presented to all 12 subjects twice, for a total of 24 trials per texture.

During each trial, a reference texture was displayed on the right half of the screen, while the user-adjustable sinusoidal texture was displayed on the left. The adjustable texture's amplitude and frequency were initialized to their minimum values of 0% friction variation and 1 mm^{-1}, respectively. Subjects were then instructed to adjust the values using the corresponding sliders in the provided graphical user interface (see Fig. 4) to best match the frequency and intensity of the reference texture.

Once finished adjusting, subjects rated the overall similarity between their adjusted texture and the reference on a Likert scale of 1 to 5, with 5 corresponding to the textures feeling identical. Finally, they were asked to record a "characteristic swipe", in which they swiped their finger back and forth across the entire display at what they felt was their characteristic scan speed for 5 s. Lateral vibrations produced by both the subject-adjusted and reference textures on the finger were recorded during this time.

III. DATA PROCESSING

Recordings of lateral force and finger position collected during active scanning were used to construct spatial frequency amplitude spectra of textures for all subjects. Position data, originally collected at 125 Hz, were interpolated to 10 kHz to match the sampling rate of force data. Force data were then interpolated to a linearly spaced position array using short segments of data taken from periods of relatively constant velocity.
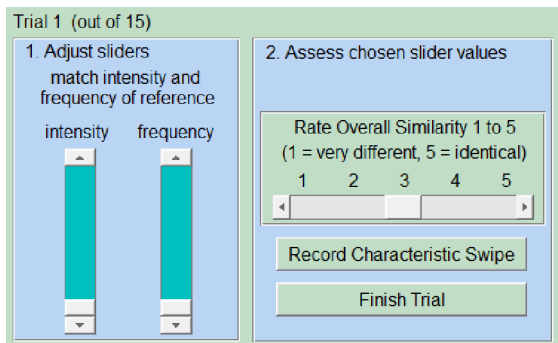
Fig. 4. GUI interface shown to subjects

In order to isolate sections of constant velocity for each texture, sections of force and position data were extracted for each texture from the middle swipe of each 5 s force recording. Within this swipe, the time of transition between textures, t_c, was identified as the time when position = 0 at the the center of the display. Swipe velocity was calculated from the position difference between t_c - 10 ms and t_c + 10 ms. 150 ms segments of each texture were then taken from 10 ms preceding and 10 ms after the transition, as shown in Fig. 5. In the rare event that swipe velocity exceeded 200 mm/s, 75 ms segments were taken instead, due to the limited total swipe duration.
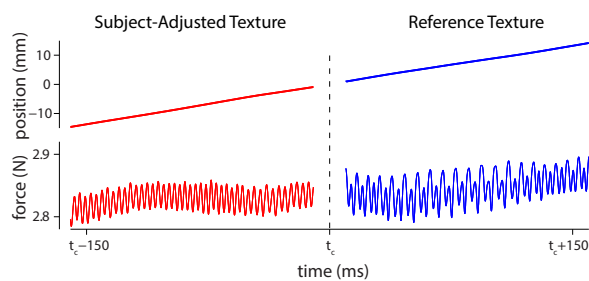

Fig. 5. Typical position and force recording from subject A, texture #3, trial 1. 150 ms of data was extracted from either side of the transition between the subject-adjusted and reference textures, which occurred at time t_c.

Due to the force sensor's range being limited to below 500 Hz, spatial frequencies which are scanned at velocities high enough to produce temporal frequencies above 500 Hz will be filtered out. This velocity limitation can be expressed as:

$$v_{limit} = (500Hz)/(f_{max}) \qquad (1)$$

where f_{max} is the maximum spatial frequency whose amplitude we are interested in measuring, and v_{limit} is the fastest scan velocity at which the force sensor can measure it. Any measured amplitudes from trials with absolute values of scan velocities greater than this limit will subsequently be noted or excluded from analysis.

IV. RESULTS

A. Measured Reference and Adjusted Textures

Spatial frequency amplitude spectra of reference textures # 1-3 for all subjects and trials are shown in gray in Fig. 6. The maximum amplitude over all measurements was normalized to 1, and all data was scaled accordingly. Peaks appear at spatial frequencies of 1.5 mm^{-1} and 3 mm^{-1}, as expected for these textures. Amplitude peaks of each subject-adjusted frequency are shown as points instead of full spectra for the sake of visual clarity. The cluster of adjusted frequencies shifts noticeably to the right as the higher frequency component of the reference texture in strength relative to the lower component.
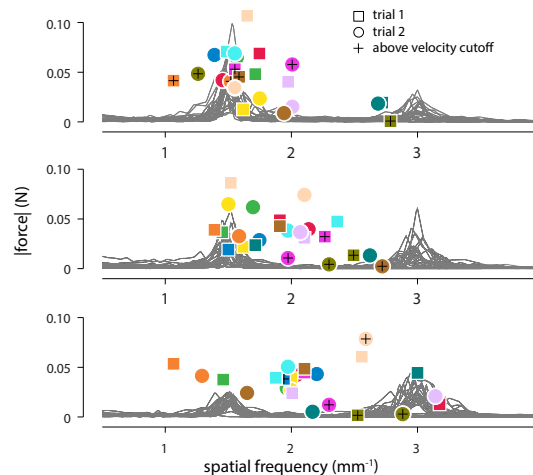

Fig. 6. Individual reference texture spectra, shown in gray, and peaks of subject-adjusted texture spectra for all trials of texture #1 (top), #2, and #3 (bottom). Data points are color coded by subject and shape coded by trial. Measurements made at scan speeds above the velocity cutoff defined by eq. 1 are indicated according to the legend.

A summary of subject-adjusted frequency values and measured amplitudes for the subset of data collected under each texture's velocity limit is shown in Fig. 7. Average adjusted frequency and measured amplitude for each subject with sufficiently slow scan speeds are indicated by a point, and their respective reference texture spectra are again shown in gray. Mean adjusted frequency over all trials is indicated by the dotted vertical line. Since the exploration of particular textures tends to be dominated by either above- or below-cutoff scan speeds, it is not possible to statistically test for the effect of speed. However, given the data available for each texture, the average difference between adjusted frequencies after scans at speeds above and below cutoff (.22), is much smaller than the observed effect of the variation in component frequencies that is our central focus.

B. Statistical Analysis

A three-way, within-subject ANOVA was performed across subject-adjusted frequency values for the multi-frequency textures. The factors were trial number (first or second repetition), amplitude ratio, and frequency combination. The ANOVA showed a significant effect of ratio, $F(2,22) = 45.12$, $p < .001$, an effect of spectral combination, $F(3, 33) = 36.00$, $p < .001$, and an interaction between the two factors, $F(6, 66) = 3.42$, $p = .005$, indicating that the effect of the ratio varied with the frequency spectrum.
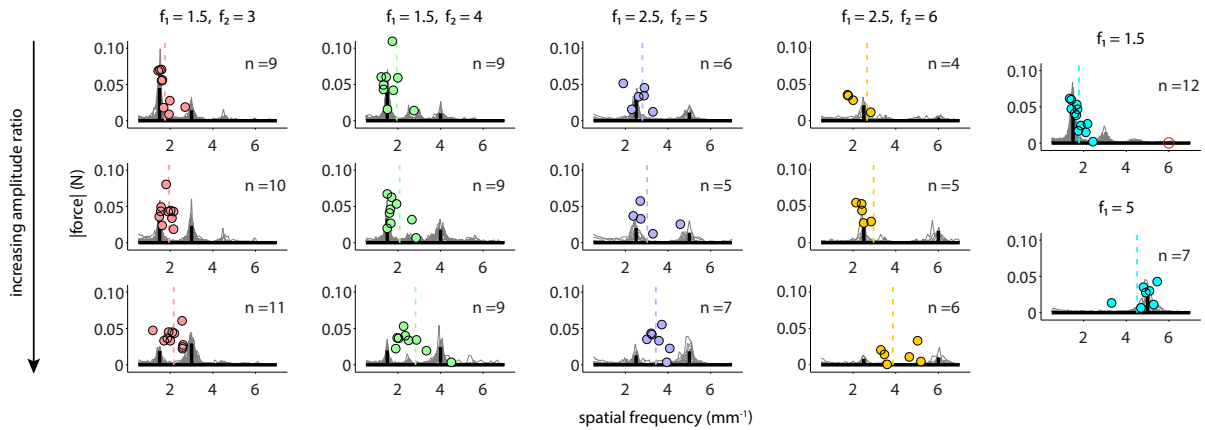
Fig. 7. A summary of all measured amplitudes of both subject-adjusted and reference textures for trials performed under the scan velocity limit. Mean peak heights of reference textures are shown in black, while individual spectra are in gray. Average subject-adjusted frequencies and measured peaks for each subject that performed scans under the velocity limit are shown as points; the number of subjects, n, that met this condition are listed on each plot. Mean subject-adjusted frequency values over all subjects and trials are shown as vertical dotted lines; see Fig. 8 for a measure of error. A single trial was omitted from analysis for texture #13 (top right, marked in red), due to being more than 3 standard deviations from the mean frequency.

Notably, there was no effect of trial number and no inter-action involving that factor, indicating that the data were stable across the two repetitions. Further confirmation was the strong between-trial correlation of the 12 texture means, r=.94.

A comparable ANOVA on subject-adjusted amplitudes found only a main effect of texture, $F(3, 33) = 3.21$, p = .035, reflecting somewhat higher amplitudes for Textures 1-3 and 7-9 than 4-6 and 10-12. The interpretation is unclear, but higher amplitudes could reflect the former textures' harmonic composition, while the particularly low amplitudes for tex-tures 10-12 may also result from the display's attenuation of the higher frequency components even at slower scan speeds.

The effects of frequency content and amplitude ratio on subject-adjusted frequency are summarized in Fig. 8. Adjusted frequencies and error for single-component textures are shown for comparison; subjects matched the pitch of these textures quite reliably. A single outlying trial was omitted from analysis for texture #13; all other frequencies lay within three standard deviations of the mean for each texture and are included in results.
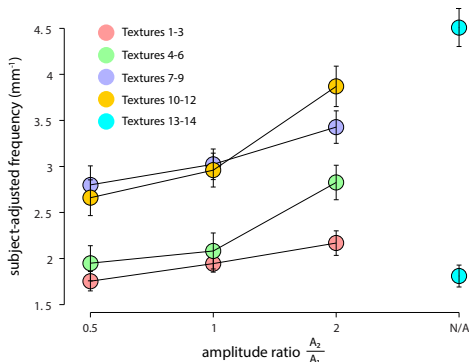

Fig. 8. Subject-adjusted frequency as a function of amplitude ratio for each category of reference frequency pairs. Adjusted values for single-component textures #13-14 are also shown for comparison. Error bars represent the standard error of the mean across subjects.

C. Subject Ratings

No significant trend was observed in subject ratings. The average rating for each reference and subject-adjusted texture pair was above 4, suggesting that subjects were generally quite satisfied with their adjusted match. An expected in-crease in ratings for texture #13 and 14 was not observed, despite these textures being the only two that could be perfectly matched by subject controls.

D. Predicting Perceived Frequency

The dependence of subject-adjusted frequency on refer-ence texture amplitude ratio suggests, at the simplest degree, a weighted sum of frequency components:

$$f_{chosen} = (A_1 * f_1 + A_2 * f_2)/(A_1 + A_2) \qquad (2)$$

Predictions of the adjusted frequencies shown in Fig. 8 were made, using only amplitudes from the subset of subjects who performed scans slow enough to measure. The limited number of subject data, especially for reference textures with higher frequency components (and therefore lower scan ve-locity limits), makes these predictions somewhat imprecise, but a general trend is shown in Fig. 9.

V. DISCUSSION

In brief, our experiment shows that people match single-pitch equivalents of complex textures reliably, and they rate the degree of match as highly as the matches to simple textures. Informal discussions with several naive subjects post-experiment suggest subjects were not merely assigning a single frequency to a texture they felt had two distinct components, but were indeed identifying a single pitch. All were surprised to learn that they had been feeling reference textures composed of two frequencies. Added frequency content appears to have been perceived as a change in other textural qualities: several subjects reported that they found it easy to get close in pitch, but the task was difficult to get just right. One stated more specifically that some quality of
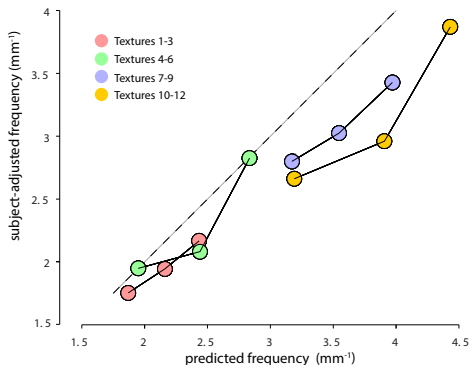
Fig. 9. Subject-adjusted frequency values over all subjects versus predicted values using Eq. 2. Predictions are made using measured amplitudes of reference textures from the subset of subjects who had scans under the velocity limits. Perfect predictions would lie along the dotted line; here, subjects' actual values are slightly over-predicted.

sharpness or jaggedness was missing, suggesting that despite matching pitch, there is another perceptual difference that accounts for overall discrimination seen in [7], [8] and [9].

These results should be viewed in the context of experimental limitations. Although commanded amplitude was kept constant across all reference textures, perceived intensities decreased for higher amplitude ratios, due both to high frequency attenuation of the display device as well as the perceptual system's heightened sensitivity to certain frequencies. As a result, the increase in pitch perception also happens to correlate with a decrease in perceived intensity, a relationship predicted by Bekesy [3]. We sought to neutralize this effect intensity may have on pitch judgments by having subjects also match the amplitude of their adjusted texture to the reference. Ultimately, subjected-adjusted amplitudes varied widely, as their measured values show in Figs. 6 and 7, and the success of this approach in removing intensity dependence should be further validated.

The method of responding used in this study had both advantages and drawbacks. Letting subjects choose their own slider values allowed them to determine a subjectively equal frequency and intensity with a low number of trials and without needing to do comparison tests across a wide range of both values. However, the substantial amount of time subjects spent on each trial, often a minute or more, suggests that this task was not easy. The spread in chosen frequency values, as shown in Figs. 6 and 7, could be due both to noise in perception but also noise in subjects' ability to adjust two values simultaneously and accurately.

Another source of noise in adjusted values may be due to variations in scanning velocities, although the similarity of data at fast and slow scan speeds here suggests this is not critical. While subjects were asked to complete their final recorded scan at their characteristic speed, it is doubtful that they maintained that speed throughout each ≈ 60 s long trial.

For future experiments, it may be useful to control scanning velocity. This would control speeds both across and within subjects, allowing us to make stronger associations with the temporal sensitivities of the mechanoreceptors. Additionally, due to the very high variability in chosen intensity

values, adjustable intensity proved an imperfect method of removing intensity as a confounding cue for pitch. A possible alternative could be to perceptually equate the intensity of reference textures before performing further psychophysical judgements, as in [7] and [9]. Future work can also explore how increasing the number of frequency components, either by adding single frequencies or changing other qualities of the spectra such as peak spread, can approximate these other qualities in increasingly complex reference textures.

VI. CONCLUSION

Tactile single pitch perception of friction-modulated textures composed of two frequency components was assessed by matching to single frequency counterparts. Pitch perception was indicated by regularities in the matched values. The matches were found to lie between the two source frequencies, with the ultimately selected frequency correlated with their amplitude ratio.

REFERENCES

[1] M. Hollins and S. R. Risner, "Evidence for the duplex theory of tactile texture perception," *Attention, Perception, & Psychophysics*, vol. 62, no. 4, pp. 695–705, 2000.

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

[3] G. Von Békésy, "Synchronism of neural discharges and their demultiplication in pitch perception on the skin and in hearing," *The Journal of the Acoustical Society of America*, vol. 31, no. 3, pp. 338–349, 1959.

[4] M. Hollins and E. A. Roy, "A ratio code for vibrotactile pitch," *Somatosensory & motor research*, vol. 15, no. 2, pp. 134–145, 1998.

[5] J. W. Morley and M. J. Rowe, "Perceived pitch of vibrotactile stimuli: effects of vibration amplitude, and implications for vibration frequency coding." *The Journal of physiology*, vol. 431, no. 1, pp. 403–416, 1990.

[6] L. E. Marks, "Summation of vibrotactile intensity: An analog to auditory critical bands?" *Sensory processes*, vol. 3, pp. 188–203, 1979.

[7] S. Bensmaïa, M. Hollins, and J. Yau, "Vibrotactile intensity and frequency information in the pacinian system: A psychophysical model," *Attention, Perception, & Psychophysics*, vol. 67, no. 5, pp. 828–841, 2005.

[8] E. L. Mackevicius, M. D. Best, H. P. Saal, and S. J. Bensmaia, "Millisecond precision spike timing shapes tactile perception," *Journal of Neuroscience*, vol. 32, no. 44, pp. 15 309–15 317, 2012.

[9] 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.

[10] Y. Yoo, I. Hwang, and S. Choi, "Consonance of vibrotactile chords," *IEEE transactions on haptics*, vol. 7, no. 1, pp. 3–13, 2014.

[11] S. A. Cholewiak, K. Kim, H. Z. Tan, and B. D. Adelstein, "A frequency-domain analysis of haptic gratings," *IEEE Transactions on Haptics*, vol. 3, no. 1, pp. 3–14, 2010.

[12] H. T. Nefs, A. M. Kappers, and J. J. Koenderink, "Detection of amplitude modulation and frequency modulation in tactual gratings: A critical bandwidth for active touch," *Perception*, vol. 32, no. 10, pp. 1259–1271, 2003.

[13] 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 *World Haptics Conference (WHC), 2015 IEEE*. IEEE, 2015, pp. 125–130.

[14] D. J. Meyer, M. Wiertlewski, M. A. Peshkin, and J. E. Colgate, "Dynamics of ultrasonic and electrostatic friction modulation for rendering texture on haptic surfaces." in *HAPTICS*, 2014, pp. 63–67.