7-2010

Dimensional Reduction of High-Frequencey Accelerations for Haptic Rendering

Katherine J. Kuchenbecker
University of Pennsylvania, kuchenbe@seas.upenn.edu

William McMahan
University of Pennsylvania

Nils Landin
KTH Royal Institute of Technology

Joseph M. Romano
University of Pennsylvania

Follow this and additional works at: http://repository.upenn.edu/meam_papers

Dimensional Reduction of High-Frequencey Accelerations for Haptic Rendering

Abstract

Haptics research has seen several recent efforts at understanding and recreating real vibrations to improve the quality of haptic feedback in both virtual environments and teleoperation. To simplify the modeling process and enable the use of single-axis actuators, these previous efforts have used just one axis of a three-dimensional vibration signal, even though the main vibration mechanoreceptors in the hand are know to detect vibrations in all directions. Furthermore, the fact that these mechanoreceptors are largely insensitive to the direction of high-frequency vibrations points to the existence of a transformation that can reduce three-dimensional high-frequency vibration signals to a one-dimensional signal without appreciable perceptual degradation. After formalizing the requirements for this transformation, this paper describes and compares several candidate methods of varying degrees of sophistication, culminating in a novel frequency-domain solution that performs very well on our chosen metrics.

Dimensional Reduction of High-Frequency Accelerations for Haptic Rendering

Nils Landin[†], Joseph M. Romano[⋆],
William McMahan[⋆], and Katherine J. Kuchenbecker[⋆]

[†]KTH Royal Institute of Technology, Stockholm, Sweden
[⋆]University of Pennsylvania, Philadelphia, USA
`nlandin@kth.se`   `{jrom,wmcmahan,kuchenbe}@seas.upenn.edu`
`http://haptics.seas.upenn.edu`

Abstract. Haptics research has seen several recent efforts at understanding and recreating real vibrations to improve the quality of haptic feedback in both virtual environments and teleoperation. To simplify the modeling process and enable the use of single-axis actuators, these previous efforts have used just one axis of a three-dimensional vibration signal, even though the main vibration mechanoreceptors in the hand are know to detect vibrations in all directions. Furthermore, the fact that these mechanoreceptors are largely insensitive to the direction of high-frequency vibrations points to the existence of a transformation that can reduce three-dimensional high-frequency vibration signals to a one-dimensional signal without appreciable perceptual degradation. After formalizing the requirements for this transformation, this paper describes and compares several candidate methods of varying degrees of sophistication, culminating in a novel frequency-domain solution that performs very well on our chosen metrics.

Key words: haptic feedback, vibrations, measurement-based modeling

1 Introduction

Haptic interfaces are designed to let a user touch virtual and distant objects as though they were real and within reach. For many applications, users would like the haptic feedback provided during these virtual and distant contacts to match the feel of real objects as closely as possible. When you touch a real surface with a tool, you feel low-frequency forces that convey shape, compliance, and friction, and you also feel high-frequency vibrations that reflect the texture of the object and its current contact state with your tool [5, 6]. The mechanoreceptors that detect these important high-frequency vibrations are the Pacinian corpuscles (PCs); they are sensitive to vibratory stimuli from 20 to 1000 Hz, with a peak sensitivity between 250 and 550 Hz [1, 2]. PCs are also known to respond to vibrations that occur in all directions, with motion parallel to the skin surface being slightly more easy to detect than motion normal to the skin [3].

As described in Section 2, high-frequency acceleration measurements have been shown to enable realistic haptic rendering of surface contact. Although

tool vibrations occur in all three directions, these previous works have simplified the problem by measuring and recreating only a single axis of the vibration, discarding potentially significant haptic information in the other two axes. One could imagine replicating these single-axis techniques in three orthogonal directions, but such an approach would require significantly more complex vibration models and haptic hardware. Instead, we believe we can take advantage of the human hand's insensitivity to vibration direction by recreating the feel of a full three-dimensional acceleration with a one-dimensional signal. Section 3 lays out our quantitative objectives for this dimensional reduction problem. Section 4 then describes the candidate transformations we have considered, culminating in the presentation of our new approach, DFT321. Finally, Section 5 summarizes the contributions of this paper and lays out future work.

2 Background

Several research teams have succeeded at creating realistic haptic renderings through the use of high-frequency acceleration signals, though all of these efforts have neglected the 3D nature of real vibrations. For teleoperation, Kontarinis and Howe measured uniaxial accelerations at the fingertips of a custom robotic hand and continually played these vibrations for the operator to feel via two inverted audio speakers on the fingers of the master interface [7]. This sensory augmentation improved user performance in bearing inspection and needle puncture tasks. Several other researchers have used uniaxial recordings of real contact accelerations to improve the realism of virtual surface contact, either through parametric models [12, 4] or direct playback [8]. In the domain of direct manipulation, Yao, Hayward, and Ellis created a handheld tool that measures accelerations perpendicular to the tool tip and outputs them in the orthogonal direction to improve the user's sensitivity to small surface features [15].

Inspired by this previous research, work in our own lab has focused on capturing and recreating the feel of real surfaces during tool-mediated contact, a process we call haptography [9]. As part of this project, we added a voice-coil actuator to the handle of a Phantom Omni to allow feedback of a single axis of contact acceleration during teleoperation [10]. Though this system was initially configured to replicate accelerations normal to the contacted surface, it was altered to transmit tangential accelerations for a human subject study focused on texture perception [11]. Subjects rated surface renderings as significantly more real when they included vibrations from the dedicated actuator, but no sample achieved a realism rating as high as the real surface. In fact, the highest rated rendering provided vibrations that were 50% stronger than the measured tangential values, indicating that restricting vibration measurement to a single axis may compromise the fidelity of the resulting interaction. Finally, we have also developed a method for distilling a set of recorded accelerations into a texture model for virtual surface rendering [13]; that work used principal component analysis to reduce 3D sensor signals down to 1D, but it did not closely consider the important role this transformation plays in system performance.

3   Transformation Objectives

The human hand detects high-frequency vibrations in all directions but cannot readily distinguish these directions from one another. Thus, we believe haptic rendering systems should be able to model and output 3D contact accelerations as 1D vibrations. Figure 1 shows our experimental setup for capturing and recreating such vibrations. The three-axis accelerometer (ADXL345) is configured for a range of ± 78.5 m/s^2 (± 8 g) and is sampled at 1000 Hz. The stylus is also equipped with a pair of voice coil actuators for vibration output. Figure 2 shows eight seconds of data collected with this device during real contact interactions. Notice that the shape of the time- and frequency-domain signals differs significantly between the three axes. The smoothed version of each spectrum was computed with a variable frequency resolution that matches the 22% Weber fraction for vibration discrimination [2]. Logarithmic scaling causes the smoothed version to appear to be too high in the plot, though it is correctly positioned. Because the mapping from $\mathbb{R}^3$ to $\mathbb{R}$ is not obvious, we have identified two objectives for measuring transformation performance.
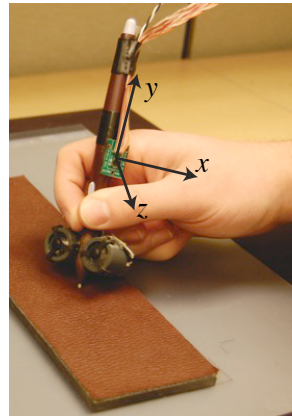

Fig. 1: Texture vibrations were captured with a Wacom stylus equipped with a digital accelerometer.

I. Spectral Match   Human perception of high-frequency vibrations relies on the spectral content of the signal [2]. Thus, a dimensionally reduced vibration should feel like it has the same spectrum as the original recorded signal. We formalize this by saying that the synthesized 1D vibration should have the same
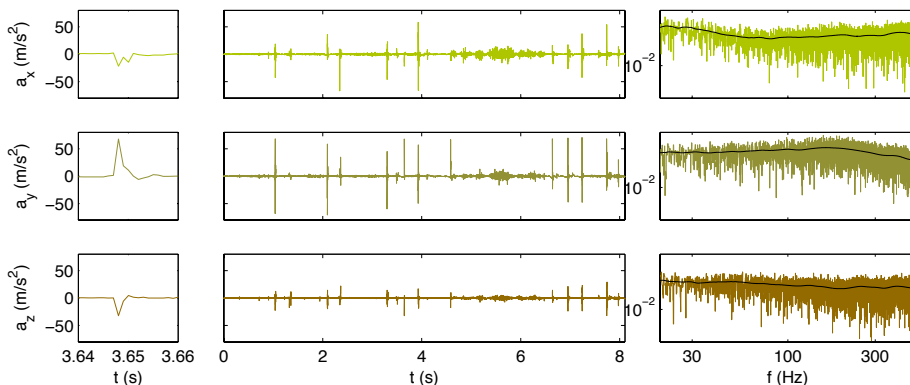

Fig. 2: Sample three-dimensional acceleration data. The user ran the tool tip over several small surface features, dragged across a textured vinyl surface, and then tapped on a hard piece of plastic.

energy at each frequency as the sum of the energies present at that frequency in the original vibration's three component directions. Consequently, the transformation should preserve the total Energy Spectral Density (ESD) of the original 3D signal, where the ESD of each component $a(t)$ is $E_s(f) = |A(f)|^2$, where $A(f)$ is the Fourier transform of $a(t)$.

Because raw spectral estimates are noisy, we do not judge spectral similarity directly from the energy spectral density at each frequency. Instead, we have designed a spectral metric that takes into account the limited frequency resolution of human vibration perception, as mentioned in the previous section. Hence, our spectral perceptual comparisons use a frequency smoothing resolution that varies according to this rule, implemented via the `spafdr` command in Matlab; an example of this smoothing is shown in Figure 2, and we denote the smoothed version of $A(f)$ as $\tilde{A}(f)$. Using $a_s(t)$ to represent the 1D synthesized signal that results from the transformation, we write our spectral match metric as

$$M_{sm} = 1 - \frac{1}{n_f} \sum_{f=20\,Hz}^{1000\,Hz} \left(\frac{\left| |\tilde{A}_x(f)|^2 + |\tilde{A}_y(f)|^2 + |\tilde{A}_z(f)|^2 - |\tilde{A}_s(f)|^2 \right|}{|\tilde{A}_x(f)|^2 + |\tilde{A}_y(f)|^2 + |\tilde{A}_z(f)|^2} \right) \quad (1)$$

where n_f signifies the number of discrete frequencies in the sum. Here, we are quantifying the extent to which the synthesized signal preserves the energy in the original 3D signal for frequencies from 20 Hz to 1000 Hz, chosen to match the sensitivity range of Pacinian corpuscles. This calculation provides a strict measure of the average normalized deviation between the 1D signal's smoothed ESD and the 3D signal's smoothed ESD. If the two are identical, the spectral match metric will be one.

II. Temporal Match While the spectral criterion captures the requirements for the stationary characteristics of the signal, we need to impose another criterion to capture transients. Ideally, peaks and sudden changes of the 1D signal should coincide with similar features in the 3D components. As a simple measure of temporal match, we thus look at the absolute value of the cross-correlation between each of the three original components and the synthesized signal at a time shift of zero; we use the absolute value because the human hand is not sensitive to vibration direction. We write our full temporal match metric as

$$M_{tm} = \frac{1}{3} \left(\frac{|a_x \star a_s|}{\sqrt{a_x \star a_x}\sqrt{a_s \star a_s}} + \frac{|a_y \star a_s|}{\sqrt{a_y \star a_y}\sqrt{a_s \star a_s}} + \frac{|a_z \star a_s|}{\sqrt{a_z \star a_z}\sqrt{a_s \star a_s}} \right) \quad (2)$$

where $\star$ denotes the cross-correlation evaluated at zero time delay. In words, this equation provides the average of the absolute values of the normalized cross-correlations. The resulting value is a measure of simultaneous temporal similarity between the 1D signal and the original components. A signal that correlates (or anticorrelates) perfectly with all three components (though rarely realizable) would have a temporal match metric of one.

4   Candidate Transformations

There are many potential methods for transforming three acceleration values into one value over time. This section describes the approaches we have considered thus far, progressing from the simplest options up to a more sophisticated method that we have newly developed. Each algorithm name includes the suffix 321 to indicate that it reduces the dimensions of a time-domain signal from three to one. This suffix also distinguishes our synthesis algorithms (e.g., DFT321) from the analysis algorithms from which they are derived (e.g., DFT). Figure 3 shows the effects of four of the candidates on the sample data from Figure 2, and Figure 4 shows their performance on a shorter texture-only sample. The corresponding spectral and temporal metric results are presented in Table 1.

Single-Axis (SA321)   The simplest solution is to use an axis fixed to the sensor, as done in prior work [10, 11]. Since there is no guarantee that the main vibrational energy will occur along that axis, considerable signal loss is expected. This is verified by inspection of the sample data in Figure 2: peaks do not always
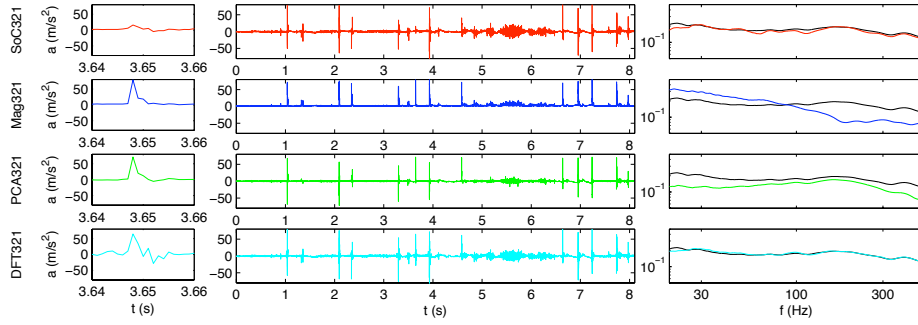

Fig. 3: Four of the candidate transformations applied to the data from Figure 2. The left column shows a detailed time view, and the right column presents the smoothed energy spectral density along with the original signal's smoothed ESD (black line).
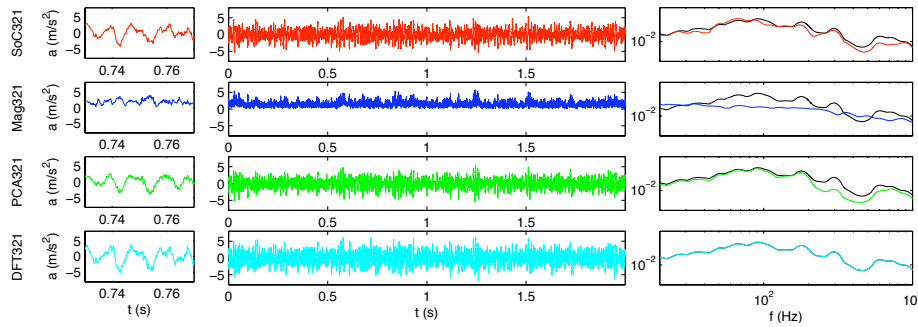

Fig. 4: Four of the candidate transformations applied to a 3D texture vibration.

		SA$_x$321	SA$_y$321	SA$_z$321	SoC321	Mag321	PCA321	DFT321
Figure 3	M_{sm}	0.34	0.56	0.10	0.90	0.34	0.56	0.97
	M_{tm}	0.37	0.38	0.39	0.51	0.11	0.40	0.46
Figure 4	M_{sm}	0.27	0.51	0.23	0.48	0.37	0.52	1.00
	M_{tm}	0.69	0.68	0.60	0.62	0.02	0.80	0.64

Table 1: Spectral match and temporal match metrics for the candidate transformations on the data shown in Figures 3 and 4.

co-occur between the axes. This approach does not capture the total energy of the vibrations, and its spectral and temporal match metrics are not excellent.

Sum of Components (SoC321)   Another computationally simple approach is to merely add the three components together. While correlation between components is likely, it may be positive or negative; this approach is thus susceptible to destructive interference between components that have negative cross-correlation. As exhibited in the top panels of Figures 3 and 4, this method does not preserve spectral energy and has inconsistent temporal performance.

Vector Magnitude (Mag321)   One can also take the square root of the sum of the squares of the components, as shown in the second panels of Figures 3 and 4. This transformation is not suitable for the outlined objectives; the magnitude of a 3D vector will always be positive, which introduces a DC-component that does not reflect the feel of the original high-frequency vibration. Also, the spectral contents of the three components are redistributed in a non-linear way that is highly dependent on the magnitudes of the different frequency components.

Principal Component Analysis (PCA321)   As done in [13], we can project the 3D signal onto a single fixed axis that PCA has identified as containing the most energy. This approach may be adequate if the haptic interaction that generated the data was highly constrained, as with the texture exploration data of Figure 4, though it always discards the information in the two perpendicular directions. For a slight improvement, one could divide the signal into segments over time and continually adapt the estimate of the first principal component.

Discrete Fourier Transform (DFT321)   Because none of the above transformations fully satisfies our objectives, we developed a new approach based on the DFT. By expressing each component of the 3D signal as an orthogonal basis function expansion, we can sum the components without destructive interference. In view of the spectral requirements on the synthesized signal, the DFT represents one feasible choice of such basis functions. Following (1), we set

$$\left| \tilde{A}_s \left(f \right) \right| = \sqrt{ \left| \tilde{A}_x \left(f \right) \right|^2 + \left| \tilde{A}_y \left(f \right) \right|^2 + \left| \tilde{A}_z \left(f \right) \right|^2 } \tag{3}$$

Having obtained the absolute value of $\tilde{A}_s(f)$ and by that set the local spectrum of a_s, we must assign a phase θ_f that ensures desirable temporal properties. To this end we optimize the phase towards metric (2), neglecting the absolute values for simplicity. In the frequency domain, the sum of cross-correlations at zero time shift can be expressed as $\sum_{i=1}^{3} a_i \star a_s$. It can be shown that this quantity is maximized when the phase of $\tilde{A}_s$ is chosen as follows:

$$\theta_f^{\max} = \angle \sum_{i=1}^{3} \tilde{A}_i \qquad (4)$$

The synthesized time-domain signal a_s is then obtained by a square-root, multiplication by $e^{j\theta_f^{\max}}$ where $j = \sqrt{-1}$, and inverse DFT. By Parseval's theorem the result will always have the same signal energy as the sum of energies of the components. We call this new dimensional reduction approach DFT321.

For the transient data shown in Figure 3 and online applications, the DFT321 method must be divided in short windows. Here, a window length of 35 ms has been used, which is 10 ms less than the minimum detectable time delay between visual and haptic stimuli [14]. We found that window lengths ranging from 25 ms to 50 ms yield consistently good results, with $0.94 \le M_{sm} \le 0.97$ and $0.45 \le M_{tm} \le 0.46$. For offline processing of stationary vibrations, such as those shown in Figure 4, no windowing is required, and the spectral content of the vibrations is captured exactly.

5 Conclusion

This paper focuses on possible methods for reducing a three-dimensional high-frequency acceleration signal into a perceptually matched one-dimensional signal. After reviewing previous work in this area, we presented two objectives by which to judge candidate transformations, and we described several possible options, including our new DFT321 method. While DFT321 is a top candidate for both transient and stationary data, one should adapt the choice according to the application and available real-time computational resources. Using dimensional reduction in this way will strongly benefit haptic rendering approaches that rely on high-frequency acceleration measurements: for example, our methods reduce data storage needs and parametric complexity for surface texture models, they diminish the required transmission bandwidth for teleoperation systems, and they facilitate the effective use of one-dimensional haptic vibration actuators. Continuing research will include human subject testing to verify and refine our spectral and temporal match metrics.

References

1. Bell, J., Bolanowski, S., Holmes, M.H.: The structure and function of Pacinian corpuscles: A review. Progress in Neurobiology 42(1), 79–128 (Jan 1994)
2. Bensmaïa, S., Hollins, M., Yau, J.: Vibrotactile intensity and frequency information in the Pacinian system: A psychophysical model. Perception and Psychophysics 67(5), 828–841 (2005)
3. Brisben, A.J., Hsiao, S.S., Johnson, K.O.: Detection of vibration transmitted through an object grasped in the hand. Journal of Neurophysiology 81(4), 1548–1558 (Apr 1999)
4. Guruswamy, V.L., Lang, J., Lee, W.S.: Modelling of haptic vibration textures with infinite-impulse-response filters. In: Proc. IEEE International Workshop on Haptic Audio Visual Environments and their Applications. pp. 105–110 (2009)
5. Johnson, K.O.: The roles and functions of cutaneous mechanoreceptors. Current Opinion in Neurobiology 11, 455–461 (2001)
6. Klatzky, R.L., Lederman, S.J.: Perceiving object properties through a rigid link. In: Lin, M., Otaduy, M. (eds.) Haptic Rendering: Algorithms and Applications, chap. 1, pp. 7–19. A. K. Peters (2008)
7. Kontarinis, D.A., Howe, R.D.: Tactile display of vibratory information in teleoperation and virtual environments. Presence: Teleoperators and Virtual Environments 4(4), 387–402 (Aug 1995)
8. Kuchenbecker, K.J., Fiene, J.P., Niemeyer, G.: Improving contact realism through event-based haptic feedback. IEEE Transactions on Visualization and Computer Graphics 12(2), 219–230 (March/April 2006)
9. Kuchenbecker, K.J., Romano, J.M., McMahan, W.: Haptography: Capturing and recreating the rich feel of real surfaces. In: Proc. International Symposium on Robotics Research (August 2009)
10. McMahan, W., Kuchenbecker, K.J.: Haptic display of realistic tool contact via dynamically compensated control of a dedicated actuator. In: Proc. IEEE/RSJ International Conference on Intelligent RObots and Systems. pp. 3171–3177 (October 2009)
11. McMahan, W., Romano, J.M., Rahuman, A.M.A., Kuchenbecker, K.J.: High frequency acceleration feedback significantly increases the realism of haptically rendered textured surfaces. In: Proc. IEEE Haptics Symposium. pp. 141–148 (March 2010)
12. Okamura, A.M., Cutkosky, M.R., Dennerlein, J.T.: Reality-based models for vibration feedback in virtual environments. IEEE/ASME Transactions on Mechatronics 6(3), 245–252 (Sep 2001)
13. Romano, J.M., Yoshioka, T., Kuchenbecker, K.J.: Automatic filter design for synthesis of haptic textures from recorded acceleration data (2010), accepted to IEEE International Conference on Robotics and Automation
14. Vogels, I.M.L.C.: Detection of temporal delays in visual-haptic interfaces. Human Factors 46(1), 118–134 (2004)
15. Yao, H.Y., Hayward, V., Ellis, R.E.: A tactile enhancement instrument for minimally invasive surgery. Computer-Aided Surgery 10(4), 233–239 (2005)