Data-driven Texture Modeling and Rendering on Electrovibration Display

Reza Haghighi Osgouei, *Member, IEEE,* Jin Ryong Kim, *Member, IEEE,* and Seungmoon Choi, *Senior Member, IEEE*

Abstract—With the introduction of variable friction displays, either based on ultrasonic or electrovibration technology, new possibilities have emerged in haptic texture rendering on flat surfaces. In this work, we propose a data-driven method for realistic texture rendering on an electrovibration display. We first describe a motorized linear tribometer designed to collect lateral frictional forces from textured surfaces under various scanning velocities and normal forces. We then propose an inverse dynamics model of the display to describe its output-input relationship using nonlinear autoregressive neural networks with external input. Forces resulting from applying a pseudo-random binary signal to the display are used to train each network under the given experimental condition. In addition, we propose a two-step interpolation scheme to estimate actuation signals for arbitrary conditions under which no prior data have been collected. A comparison between real and virtual forces in the frequency domain shows promising results for recreating virtual textures similar to the real ones, also revealing the capabilities and limitations of the proposed method. We also conducted a human user study to compare the performance of our neural-network-based method with that of a record-and-playback method. The results showed that the similarity between the real and virtual textures generated by our approach was significantly higher.

Index Terms—Variable friction display, electrostatic, electrovibration, data-driven texture rendering.

1 Introduction

TEXTURE rendering has been one of the challenging topics in the haptics community. Realistic haptic texture rendering not only provides rich information but also improves user experience. With the introduction of variable friction displays, new possibilities have emerged in haptic texture rendering on touchscreens. Collectively called *surface haptics*, the technologies vary the surface friction acted upon the user's finger, enabling programmable tactile feedback. There exist two major approaches in surface haptics: electrovibration [1], [2], [3] and ultrasonic vibration [4], [5], [6], [7]. Whereas the former increases the surface friction by modulating attractive electrostatic force, the latter decreases the friction by vibrating the surface at an ultrasonic frequency and creating an air gap. These technologies have the potential for integration into consumer electronics such as mobile phones and tablets. In particular, the electrovibration technology has the advantages that it requires only electrical components and that friction can be controlled uniformly over the screen. Many applications can benefit from this added functionality, such as Internet shopping, education, security authentication, entertainment, etc. In this work, we explore realistic texture rendering on an electrovibration display by constructing an inverse dynamics model of the display and then using it in order to synthesize actuation signals from the lateral force data collected by scanning real textures.

- *R. Haghighi Osgouei and S. Choi are with the Haptics and Virtual Reality Laboratory, Department of Computer Science and Engineering, Pohang University of Science and Technology (POSTECH), Pohang, Gyeongbuk, Republic of Korea.*
  E-mail: {haghighi, choism}@postech.ac.kr.
- *J. R. Kim is with the Alibaba Research, Sunnyvale, CA, USA.*
  E-mail: jinryongkim@gmail.com.

Manuscript received October 12, 2018; revised July 29, 2019.
(Corresponding author: Seungmoon Choi.)

1.1 Electrovibration Display

The earliest known observation of electrical attraction between the human skin and a charged surface was made by Gray in 1875 [8]. Forgotten for a while, a similar phenomenon was re-discovered later and called *electroadhesion* by Johnsen and Rahbek in 1923 [9]. In 1953, Mallinckrodt et al. again reported a rubber-like sensation when a coated metallic surface connected to a 110-V power line was touched by a grounded finger [10]. This phenomenon is called *electrovibration* and was explained based on the electrostatic attractive force between a finger and a conductive surface supplied by a high alternating voltage [11]. Most of the early electrovibration displays were non-transparent and consisted of an array of electrodes developed either for texture rendering [12], [13] or pattern rendering [14]. With TeslaTouch project [1] that presented a capacitive transparent panel (Microtouch, 3M, USA) enabling variable friction on a touchscreen, the research interest in the electrovibration technology has recently been revived [15], [16], [17], [18], [19], [20], [21], [22]. A parallel research resulted in Tixel [23], an electrostatic film developed by Senseg, which targeted handheld devices and was used in several studies [24], [25], [26]. Some other researchers developed their own electrovibration display by stacking an insulator and an ITO layer on top of a glass plate [27], [28]. A non-transparent electrostatic friction display was also developed using an aluminum plate covered with a thin plastic insulator film [29]. Based on electroadhesion, Shultz et al. developed a non-transparent electroadhesive surface from an aluminum disk coated with carbon dielectric [30] and reported its application to rendering tactile, audible and ultrasonic force [31]. The authors de-

veloped a high voltage-compliant current-controlled amplifier with a very wide bandwidth. It allowed implementing a modulation-based excitation technique that used a high-frequency (a few ten kHz) carrier signal in order to mitigate the frequency-doubling nonlinearity of electroadhesive displays. As a result, they reported a flat magnitude response up to approximately 6 kHz. Although only the first-order harmonic components were considered for that, such a wide flat band was an impressive achievement and important milestone for the advancement of electroadhesive displays.

Apart from such efforts for fabrication, many studies have been conducted to investigate various properties of electrovibration. The first mathematical model describing the electrostatic normal force for a tactile pin array display was introduced in [12]. Using a tribological method, Meyer et al. verfied this model and showed the expected square law of electrostatic normal force as a function of applied voltage [15]. In a following work [28], an electrostatic display and an ultrasonic display were compared with respect to their capability of rendering high-bandwidth force, which led to a conclusion that the electroststic device was faster in modulating sinusoidal and step signals. A more detailed and improved model including the frequency-dependent electrical properties of human skin was studied in [17]. The polarity effect of square wave actuation signal was studied in [32], and a comparison between sine and square waves was given in [33]. Kim et al. proposed a current control method to provide uniform friction intensity on an electrovibration display regardless of operation conditions [19]. Kang et al. investigated low voltage operation of electrovibration display while providing the same perceptual strength [22]. Shultz et al. noted the prominent role of air gap electrical impedance between the fingertip and the surface [34]. Despite these efforts, we still need more endeavours to fully account for the complex electrovibration phenomenon in a wide frequency range.

1.2 Data-driven Texture Rendering

Data-driven, or measurement-based haptic rendering, is a general approach that uses recordings from real objects to generate realistic haptic feedback in virtual environments [35], [36]. It provides a unified framework to capture and display a diverse range of physical phenomena, while not requiring simulations of complex contact dynamics.

There have been many studies for data-driven texture rendering. In a series of work by Lang et al., a handheld device, called WHaT, is used to obtain surface texture and compliance [37], [38]. A height profile is built from acceleration data for texture modeling, and a heuristic model for compliance is estimated from normal force and displacement. The acceleration model is further improved by fitting infinite impulse response (IIR) filter models to automatically segmented data [39].

Another series of work has been conducted by Kuchenbecker's group for data-driven texture rendering based on contact accelerations. In [40], a handheld probe with a three-axis accelerometer is used to collect contact accelerations from texture samples under constrained scanning conditions. To each data set, an autoregressive (AR) model based on linear predictive coding (LPC) is fit, and its inverse is used to synthesize accelerations. For scanning velocities and normal forces not used in the model training, acceleration values are calculated by bilinearly interpolating the model coefficients from four adjacent models. Then the handheld stylus is improved by adding a force sensor and two voice-coil actuators for haptic interaction with a tablet [41]. This stylus is used for both data collection and texture rendering employing the same LPC-based model. Their model is refined by replacing LPC with an autoregressive moving average (ARMA) model to better handle the weak stationary nature of texture data and also to reduce the model size [42]. In [43], manual unconstrained data collection is allowed from texture samples using the same stylus. The recorded vibration data is parsed into short segments, and an AR model is fit to each segment. By interpolating neighbor models, synthetic texture signals in response to user movements are generated.

Special attention to anisotropic textures has also been given. Shin et al. compared texture modeling using unified and frequency decomposed neural networks, with the former being capable of handling anisotropic patterns [44]. In addition, a dedicated data-segmentation and interpolation method based on Radial Basis Function Network (RBFN) for anisotropic textures is proposed in [45].

However, little work has been done on data-driven texture rendering using electrovibration displays. Ilkhani et al. proposed a data-driven texture rendering method by recording accelerations from three real materials and playing them back on an electrovibration display [18]. Through a user study, they showed higher similarity of the data-driven textures to the real ones in comparison with square wave patterns. In the extended work [46], they applied the same approach to the data taken from the Penn Haptic Texture Toolkit [47] and performed multidimensional scaling (MDS) analysis to create a perceptual space and extract underlying dimensions of the textures. Their results confirmed roughness and stickiness as the primary dimensions of texture perception. To render the tactile texture of fabrics, Jiao et al. [48] very recently proposed to estimate friction coefficients by recording frictional and normal forces while a finger sweeps over fabric textures, and then amplify and play back the friction coefficients on an electrostatic display. Fiedler and Vardar introduced a data-driven rendering approach for homegeneous textures that can significantly reduce the amount of data [49]. After analyzing contact acceleration data in the frequency domain, they compress and re-synthesize the data, and finally render the results back on an electrostatic display. This was in line with the authors' previous work that generates textures by modulating low-frequency unipolar pulses, which have different waveforms and spacing, with a high-frequency carrier signal [50]

Aside from these data-driven methods, there have been a few other attempts for rendering natural textures on electrostatic displays from visual representations of the textures. Kim et al. proposed a rendering method based on the gradient technique for rendering fine 3D features on electrostaic touch screens [16]. Haghighi et al. extended the idea to a generalized gradient-based method for rendering large-scale 3D geometries on an electrostatic tablet [24]. Wang et al. extract hight maps from images with Gaussian bump shading and calculate electrostatic driving signals from their

local gradients [51]. In [52], Roberts filter is used to compute the gradients of image textures, which are in turn used to estimate stimulus signals to modulate the friction on an electrovibration display.

1.3 Our Approach

Our goal was to apply the general paradigm of data-driven rendering to texture rendering using an electrovibration display. An electrovibration display modulates friction on the sliding medium, and we choose lateral force to study the textural behavior of a surface. We designed a linear motorized tribometer for accurate and easy measurement of force from the surface of a real texture sample. Normal force and scanning velocity are the adjustable variables.

In order to reproduce signals measured from real textures using an electrovibration display without distortion, the nonlinear dynamics of the display, which was examined in the literature intensively as reviewed in Section 1.1, must be compensated for appropriately. To this end, we rely on a standard black-box approach in nonlinear system identification: we identify and build an inverse dynamic model of the electrovibration display using neural networks. This approach has an advantage over the other approach of physically-based modeling and subsequent compensator design, in the simplicity of modeling process by not involving the underlying complex contact physics and in the accuracy of modeling results by including all system responses experimentally. The black-box approach generally has higher modeling power, and this merit makes it more adequate for data-driven rendering.

We actuate an electrovibration display with a full-band pseudo-random binary signal (PRBS) and collect the resultant lateral forces. We then train a neural network *inversely* to learn an appropriate input signal from the measured forces. This inverse dynamics model of the display allows us to render collected real texture signals with high fidelity. Since texture responses largely depend on normal force and scanning velocity, we repeat the inverse dynamics modeling procedure for a set of normal forces and scanning velocities.

From a real texture sample, we collect lateral force data by scanning on it with the same set of normal forces and scanning velocities. We feed the force data of each condition to the corresponding inverse dynamics neural network model in order to generate a PRBS-like actuation signal. This signal is used as input to the electrovibration display to render a virtual texture similar to the real texture when scanned with the normal force and scanning velocity used for modeling. In addition, for arbitrary normal force and scanning velocity, we propose a two-step interpolation scheme that estimates an actuation signal from the adjacent neighbours in a normal force-scanning velocity grid. The generated signals are applied to the display for realistic data-driven texture rendering.

A comparison between real and reconstructed forces in the frequency domain shows promising results and reveals the capabilities and limitations of our data-driven rendering framework. We also conducted a human user study to compare the performance of our neural network-based algorithm with that of the record-and-playback approach. Results showed the similarity between real and virtual textures is substantially higher with our approach.[1]

2 Data Collection

This section describes the hardware we use to collect lateral frictional forces. We then introduce the use of a capacitive touchscreen panel as an electrovibration display.

2.1 Tribometer

We developed a motorized linear tribometer for precise force measurement. It consists of two parts: a moving platform and a measurement stylus (Fig. 1). The platform includes a linear servo motor (MX80L, Parker, USA) capable of following a trapezoidal acceleration profile with a travel distance of 8 cm. The motor comes with a dedicated motor drive and a power supply (VIX250IH and XLPUS, Parker, USA). The stylus is made of stainless steel equipped with a six-axis force/torque sensor (Nano 17, ATI Industrial Automation, USA). The sensor had a 3 mN resolution and a 7200 Hz resonant frequency. A sensitive touch pen is attached to the lower end of the stylus to interact with the texture sample. A mass component at the top adjusts the total weight of the stylus and hence the normal pressure. The shaft is supported by a linear bearing and attached to the motor carriage using an aluminum link. The measurement stylus resembles a simple mechanical mass-damper system subject to Coulomb (dry) friction between the shaft and the linear bearing. We confirmed the accurate behavior of the tribometer by exciting the 3M panel with sine waves with different frequencies up to 400 Hz and then collecting the resulting lateral forces and observing the main frequency components and their harmonics.[2]

Each data collection trial starts with placing a texture sample on the mounting seat under the touch pen. A custom C++ program sends movement commands to the motor driver via RS232C and records the force data using a data acquisition card (PCI-6229, NI, USA) at 1 kHz rate.

We apply a high-pass filter to the collected force data with a cutoff frequency of 5 Hz to remove the different DC biases the force sensor encountered during moving left or right. We then manually extract the segments corresponding to the constant scanning speed from each data set. The segments are then concatenated to make a data stream for the texture sample under the given experimental condition.

2.2 Electrovibration Display

We use a capacitive touch panel (MicroTouch SCT3250, 3M, USA) to make an electrovibration display, as initially proposed in [1]. It consists of a transparent conductive layer coated with a thin insulator on top of a thick glass plate. Applying a high voltage waveform (around 100 Vpp) to the conductor and electrically grounding the human body modulates the friction between the sliding finger and the touch

1. Part of this work was presented earlier in the 2018 IEEE Haptics Symposium [53]. In this paper, we also present an improved interpolation scheme (Section 4), new experimental results (Section 5), and a new user study (Section 6).

2. For all real materials used in our experiments, the collected signals were bandlimited by 100 Hz. This relatively low frequency upper bound could be due to the use of a soft touch pen for data collection.

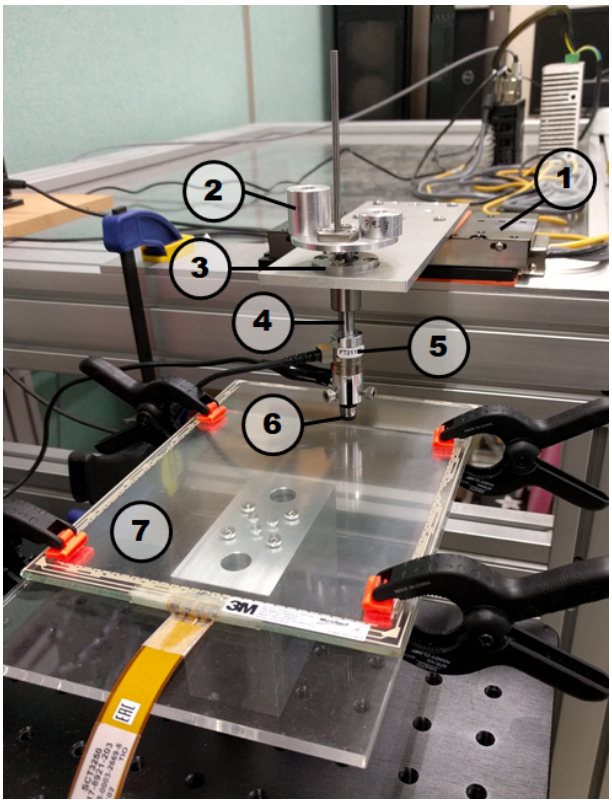


Fig. 1. Motorized linear tribometer. The measurement stylus is attached to the moving platform using an Aluminum link. 1) Linear servo motor. 2) Mass component. 3) Linear bearing. 4) Linear shaft. 5) Force/torque sensor. 6) Touch pen. 7) 3M capacitive touch panel (electrovibration display).

panel. We use a high performance piezo driver (MX200, PiezoDrive, Australia) that can amplify the input signal up to 200 Vpp.

The panel also responds well to some touch pens. We tested several types and found one that creates almost the same level of electrostatic force as the human fingertip. The test was conducted with the 3M panel placed on top of the force sensor. Exciting the panel with a 100 Hz sine wave, a bare finger or a touch pen was dragged over the surface, and the resulting lateral forces were collected. Comparing the forces in both the time and frequency domains, we selected the touch pen that showed the amplitude and power density closest to those of the bare finger. Using the touch pen enables eliminating the effects of the human skin properties that vary over time from force measurements, for consistent data collection.

3 Texture Model

In this section, we describe how to make an inverse neural network model of the input-output dynamics of the electrovibration display. We use this inverse model to synthesize actuation signals from the force recordings in order to create virtual textures on the display as similar as the real ones.

3.1 Inverse Dynamics Model

Our first goal is to develop a model that describes the dynamics relationship between the input actuation signal and the output frictional force of electrovibration display. We follow a classic system identification procedure to estimate the system model based on observed input-output data. To achieve this, we 1) define input and output variables, 2) design proper actuation signals to drive the system, and 3) draw a dynamics model between the variables using an appropriate framework.

3.1.1 Input and Output

As for our electrovibration display, the input is the actuation signal $u[n]$, and the output is the lateral force $f_l[n]$. A usual approach is to build a model g such that $f_l[n] = g(u[n])$ and then use the inverse model g^{-1} to obtain $u[n]$ for desired $f_l[n]$ such that $u[n] = g^{-1}(f_l[n])$. Since an electrovibration display exhibits considerable nonlinearity in its behavior [15], [22], [33], nonlinear models such as neural networks are appropriate for system identification. A problem is that finding an inverse model for such nonlinear systems is not always feasible. An alternative is to identify the inverse model directly by using $f_l[n]$ as input and $u[n]$ as output. This approach, however, requires care to obtain dense samples of $f_l[n]$, and we use neural networks for the inverse model identification.

3.1.2 Pseudo-Random Binary Signals

Next we need a proper actuation or perturbation signal to excite the system and observe its response. An open-loop experiment is required for such data collection, and the input signal must contain sufficient distinct frequency components in the frequency band of interest [54]. We use a pseudo-random binary signal, which is a deterministic signal with white-noise like properties. PRBS is widely used for the identification of linear systems [54] and occasionally for some nonlinear systems [55]. In case of nonlinear systems, its applicability depends on the nature of system nonlinearities. There are several successful implementations [56], as well as some failed attempts [57]. For example, a quadratic Wiener model can be identified using a PRBS while a first-order Hammerstein model cannot be [56]. It is claimed that because PRBS signals have only two levels, they may not excite certain nonlinearities of the system, so more input levels are required [58]. Nevertheless, we show that PRBS works in our case and can adequately capture the dynamic behavior of our electrovibration display.

A PRBS is characterized by three parameters: signal level $\pm c$, order n, and clock period B. A PRBS fluctuates between $-c$ and $+c$; its maximum period is $2^n - 1$; and it has to stay constant for at least B consecutive samples before it changes. To have a full-band signal that spans the entire frequency band up to the Nyquist frequency (sampling rate/2=500 Hz), clock period must be set to 1. Furthermore, considering the slowest scanning speed of 3 cm/s, it takes approximately 2667 ms for our tribometer to complete an 8-cm travel. Therefore, the length of the PRBS should not be less than the longest travel time. For a signal with $n = 9$ (PRBS9 for short), the maximum length becomes $2^9 - 1 = 511$, and with six repetitions, we obtain a signal with the total length of 3066 ms. This is slightly longer than the maximum travel time. An example plot of PRBS9 with its frequency response is given in Fig. 2. It can be seen that PRBS9 has a constant frequency response in its entire

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TOH.2019.2932990, IEEE Transactions on Haptics

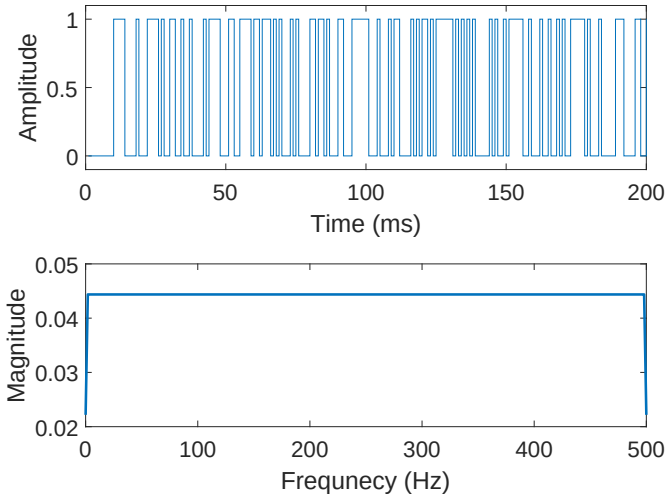


Fig. 2. An example of PRBS9 ($n = 9$).

frequency band. We use a Matlab function, `idinput()` to generate our desired PRBS.

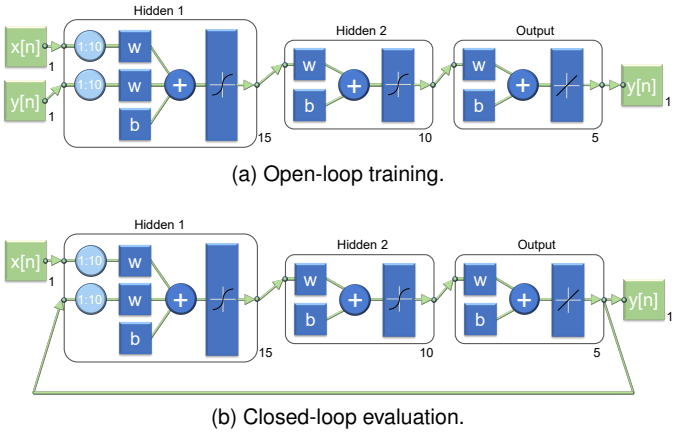
3.1.3 NARX Neural Network

With the PRBS in hand, we can now actuate the electrostatic panel and record resulting lateral forces using the tribometer. As a model, we use a NARX (Nonlinear AutoRegressive with eXternal input) neural network [59]. NARX neural networks are well suited for addressing nonlinear dynamic systems [55], [60], [61]. They have a closed-loop structure with a feedback loop connecting the output to the input (Fig. 3(b)). The past values of both input and output are used to predict the future output. The defining equation is:

$$y[n] = f(y[n-1], ..., y[n-k_y], x[n-1], ..., x[n-k_x]), \quad (1)$$

where $y[n]$ is the output to be predicted and $x[n]$ is the external input exciting the system. k_y and k_x are the tap delays for feedback and external input.

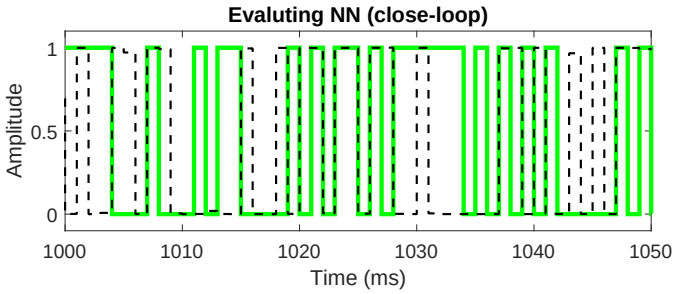
The next step is to determine the number of hidden layers and the number of neurons in each layer, as well as the size of tap delays for each input sequence. We tested several different combinations and obtained nearly perfect performance with 3 layers that have 15, 10, and 5 neurons, respectively, and tap delays of both 10. We use the Matlab neural network toolbox for training and evaluation. Training is initiated by removing the feedback loop and inputting the desired target sequence along with the external input to the network (Fig. 3(a)). In our case, $x[n] = f_l[n]$ and $y[n] = u[n]$, where $f_l[n]$ is the lateral force, normalized to handle the variability between readings under different conditions, and $u[n]$ is the designed PRBS9. The networks are trained to update weight and bias values according to Levenberg-Marquardt optimization. The data set is divided into training, testing, and validation sets with the ratios of 0.70, 0.15, and 0.15, respectively. The training ends when the error computed using the validation data set becomes sufficiently low. For evaluation, we close the feedback loop by removing the reference PRBS9 from the input and instead feeding the estimates back (Fig. 3(b)). A result of closed-loop evaluation of the trained network is given in Fig. 4. The



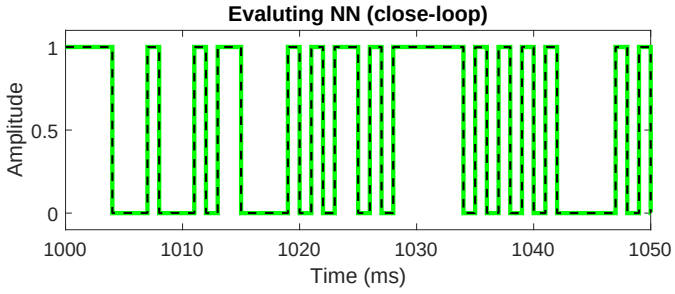
(a) Open-loop training.

(b) Closed-loop evaluation.

Fig. 3. Block diagram for the training and evaluation of NARX network. $y[n]$ is the reference signal (PRBS9) and $x[n]$ is the external input (recorded lateral force).



(a) Failed training.



(b) Successful training.

Fig. 4. Evaluation of failed and successful training of a NARX network. This example is obtained under experimental condition with mass of $m_2 = 85$ g and scanning velocity of $v_1 = 3$ cm/s. For clarity, only 50 ms of the recordings are shown. The green solid line is the reference PRBS9, and the dashed black line is an estimate from the neural network.

estimated PRBS (dashed red line) almost perfectly follows the reference PRBS (solid blue line).

3.2 Synthesizing Actuation Signals

For synthesis, we train inverse neural networks using the methods described in Section 3.1 for different lateral scanning velocities and normal forces. These two variables affect texture responses to the large extent [62], [63]. We obtain one neural network for each combination of scanning velocity and normal force. These inverse neural network models for the electrovibration display enable us to synthesize actuation signals for a texture material scanned under the same condition.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TOH.2019.2932990, IEEE Transactions on Haptics

To this end, we first collect lateral forces from the surface of a real material using the tribometer under different scanning velocities and normal forces. We normalize each force sequence to handle the variability between readings from different materials. Then we input each force reading to the corresponding trained neural network to generate a PRBS-like actuation signal. The estimated actuation signal attempts to recreate similar textural patterns to the real material once applied to the capacitive touch panel. However, these signals are only applicable to the same experimental conditions under which they were obtained. For arbitrary scanning velocity and normal force, we apply a two-step interpolation procedure between the closest neighbors in the velocity-force grid inspired by [40], [44]. Details about the interpolation are provided in Section 4.

3.3 Error Metric

To assess the synthesis accuracy of virtual textures, we first visually compare the FFT plots of real and virtual textures, and then compute a relative spectral rms error as an error metric [40], [44]:

$$E_s = \frac{RMS(\mathcal{F}(f_v[n])) - \mathcal{F}(f_r[n]))}{RMS(\mathcal{F}(f_r[n]))} \quad (2)$$

where $\mathcal{F}(\cdot)$ is the operator for fast Fourier transform and RMS is for computing the root mean square in the frequency domain. In addition, $f_v[n]$ denotes virtual forces and $f_r[n]$ real forces. We set the length of FFT equal to the sampling frequency (1000 Hz) to generate a smooth and less spiky spectrum with the main components easy to distinguish. A smaller value of E_s indicates higher similarity between two spectra and hence higher similarity between virtual and real textures. We also apply the same error metric over repeated recordings of the same real material under the same experimental condition in order to establish a baseline (lower bound) of this metric.

4 Interpolation Scheme

For the conditions with arbitrary scanning velocity and normal force, we require an interpolation scheme to estimate the actuation signal. A potential solution is to estimate a signal from the ones generated for the adjacent conditions in the velocity-force grid. However, we previously showed that linear interpolation between the neighbourhood signals does not always produce satisfactory results [53]. While the linear interpolation mostly works for the conditions with the same scanning velocity but different masses, it does not work well for the conditions with the different scanning velocities. The issue was that the main frequency components from the lower and higher velocity signals both appear in the frequency response of the interpolated force (refer to Section IV.C and Fig. 9 in [53]). We have designed an improved interpolation scheme that solves this problem and present it in this section.

4.1 Problem Definition

A closer look at the problem has revealed its origin and a potential solution. Since FFT is a linear operator, linear interpolation preserves the individual spectral components

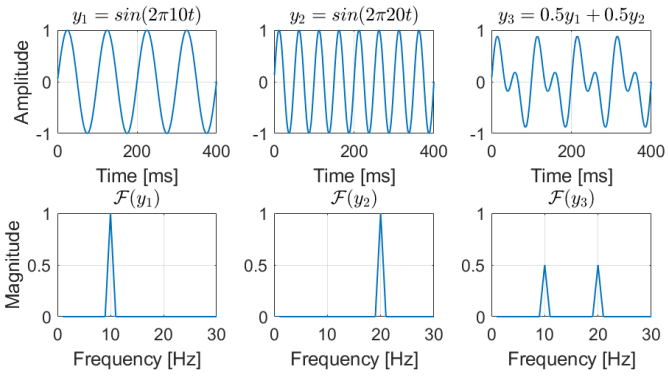


Fig. 5. Example of linear interpolation between two sine waves, y_1 and y_2, with frequencies $f_1 = 10$ Hz and $f_2 = 20$ Hz. Top plots show time-domain signals, and bottom plots show their FFTs. At $f_3 = 15$ Hz, y_3 is calculated from $y_3 = 0.5y_1 + 0.5y_2$ by taking weighted average between y_1 and y_2. Its time and frequency domain plots are given on the right.

in the resulting frequency response. For example, let us consider two sine waves, $y_1 = \sin(2\pi f_1 t)$ and $y_2 = \sin(2\pi f_2 t)$, from which y_3 is obtained by taking weighted average at frequency f_3 ($f_1 \leq f_3 \leq f_2$). This linear interpolation can be formulated as

$$y_3 = \mathcal{WA}(y_1, y_2, f_1, f_2, f_3) = \frac{f_2 - f_3}{f_2 - f_1} y_1 + \frac{f_3 - f_1}{f_2 - f_1} y_2. \quad (3)$$

where $\mathcal{WA}$ is an operator for weighted average. Taking FFT yields to

$$\mathcal{F}(y_3) = \frac{f_2 - f_3}{f_2 - f_1} \mathcal{F}(y_1) + \frac{f_3 - f_1}{f_2 - f_1} \mathcal{F}(y_2). \quad (4)$$

That is, taking weighted average in the time domain is equivalent to taking weighted average in the spectral domain because FFT is a linear operator.

An example with $f_1 = 10$ Hz, $f_2 = 20$, Hz and $f_3 = 15$ Hz is given in Fig. 5. Linear interpolation does not result in a single full power component at $f_3 = 15$ Hz; instead it makes two separate components at the lower and higher frequencies, $f_1 = 10$ and $f_2 = 20$ Hz, with the half power. Based on this observation, we propose a simple and effective solution for interpolation.

4.2 Potential Solution

The example given in Section 4.1 indicates that the problem comes from the spectra of individual components being linearly combined by Fourier transform. Therefore, a straightforward solution to this problem could be shifting the spectra in the frequency domain. If we shift the spectra to the desired frequency, the lower one to right and the higher one to left, then taking weighted average between the shifted spectra will make a single component at the desired frequency with full power.

Shifting in the frequency domain can be done by re-sampling in the time domain. The degree of re-sampling is determined by *decimation factor* (DF) [64]. Re-sampling with $DF \geq 1$ corresponds to down-sampling[3], yielding a signal with its frequency components shifted right (to higher frequency). Up-sampling with $DF \leq 1$ makes the frequency components of a signal shifted left (to lower frequency). We define a re-sampling operator as $y^r = \mathcal{RS}(y, p, q)$, where

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TOH.2019.2932990, IEEE Transactions on Haptics

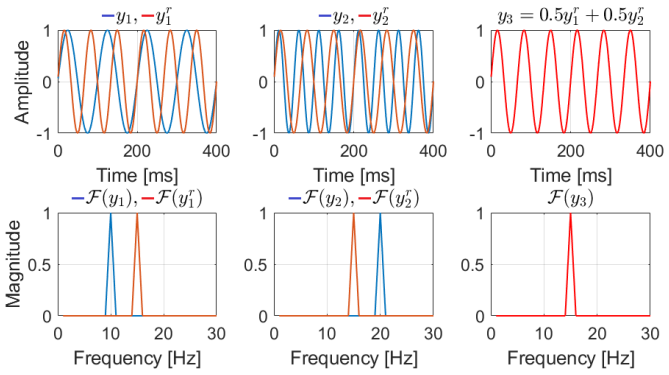


Fig. 6. Re-sampling and averaging applied on the previous example in Fig. 5. The original signals are shown in blue, and the re-sampled ones in red. Time domain plots are given on top, and FFT plots on bottom.

y^r is the re-sampled version of y with a rational decimation factor $DF = p/q$. p and q are obtained from the current and target frequencies. $DF = p/q$ means that y^r is re-sampled at $1/DF = q/p$ times of the original sampling rate.

Taking weighted average of two re-sampled signals, i.e., $y_3 = \mathcal{WA}(y_1^r, y_2^r, f_1, f_2, f_3)$ will not suffer from the aforementioned problem. This is shown for the previous example in Fig. 6. Each spectral component is first shifted to the desired frequency and then combined to generate a single component at the target frequency with full signal power.

4.3 Scanning Velocity vs. Frequency

The proposed re-sampling process requires us to know the desired frequency that we intend to land and also the lower and higher frequencies that we are about to shift. During rendering under an arbitrary scanning velocity, we only know the current scanning velocity of a user and the nearest lower and higher velocities at which texture data were collected. From these three velocities, we need to obtain two decimation factors for up-sampling the lower-frequency texture signal and down-sampling the higher-frequency one, respectively, for interpolation.

To this end, we rely on the simple relationship between scanning velocity and main component frequency in a sinusoidal texture. If the spatial wavelength of a sinusoidal texture is L, the frequency f of response force during scanning the texture at linear velocity v is [65], [66]

$$f = 2\pi \frac{v}{L}. \tag{5}$$

If (5) holds for a texture, the decimation factor $DF = v_1/v\,(v \geq v_1)$ has the effect of up-sampling the lower-frequency texture signal at f_1 to one that has the main frequency component at f. Similarly, the decimation factor $DF = v_2/v\,(v_2 \geq v)$ has the effect of down-sampling the high-frequency texture signal at f_2 to the signal with the main frequency component at f. Note that v_1 and v_2 are known and only v needs to be measured during rendering.

We investigated whether the theoretical relationship in (5) actually holds with real texture samples. An example is shown for a rippled paper shown in Fig. 7. The rippled paper had almost perfect sinusoidal patterns. Using our tribometer, we collected lateral forces under five scanning velocities: 3, 4, 5, 6, and 7 cm/s. Fig. 8 shows the frequency spectra of the measured lateral forces with an inset for a graph between frequency and scanning velocity. These data exhibit an almost perfect linear relationship between frequency and scanning velocity. We also tested six other materials (Fig. 10) used in experiments for performance evaluation. For each material, we found a linear relationship between the frequency of the main spectral component and the corresponding scanning velocity. These results validate our replacement of frequencies with scanning velocities in the re-sampling process.

3. Down-sampling is called *decimation* in signal processing.

Fig. 7. Rippled paper with almost perfect sinusoidal texture patterns.

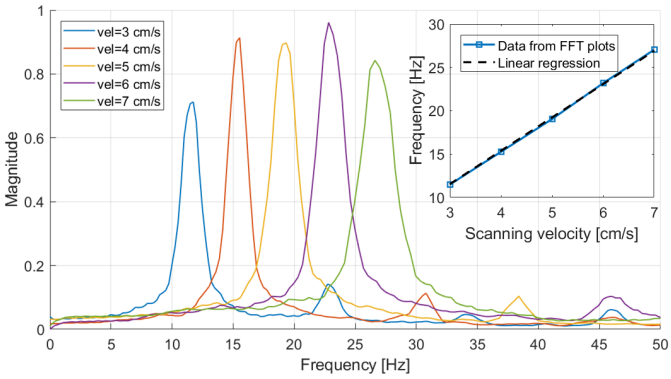


Fig. 8. Power spectra of the lateral forces collected from the surface of a rippled paper under five different scanning velocities. The inset plot at the upper right corner shows a linear relationship between the frequencies of the main component and their corresponding velocities.

4.4 Process Formulation

With the re-sampling idea described above, we can design a complete interpolation scheme for an arbitrary condition of scanning velocity and normal force from the four adjacent conditions (nodes) in the velocity-force grid (Fig. 9). Just to emphasize, the re-sampling is required only for the conditions with different scanning velocities. For the conditions with different masses, linear interpolation is sufficient. Each node $U_{i,j}$ in Fig. 9 represents an actuation signal obtained from the trained neural network under the specified normal force and scanning velocity, (m_i, v_j). The interpolation scheme has two steps. First, we obtain a signal from the nodes with the same scanning velocity but different masses (normal forces). Second, we re-sample the newly generated signals to the desired frequency (scanning velocity) and then take weighted average.

Let the user-applied force be m_u $(m_i \leq m_u \leq m_{i+1})$ and the current scanning velocity be v_u $(v_j \leq v_u \leq v_{j+1})$. Then

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TOH.2019.2932990, IEEE Transactions on Haptics

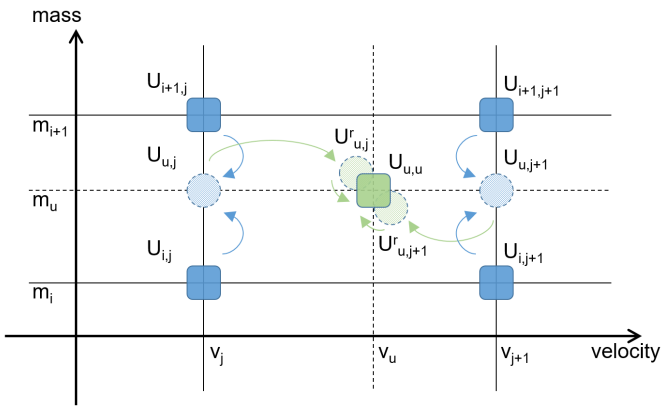


Fig. 9. Improved interpolation scheme. (m_u, v_u) denotes user-applied normal force and lateral scanning velocity. Blue squares are the adjacent actuation signals obtained from the neural networks trained under corresponding experimental conditions. Blue arrows indicate steps 1.1 and 1.2, applied on the vertical nodes to yield blue dashed circles. Long green arrows indicate steps 2.1 and 2.2, applied on the horizontal nodes to yield green dashed circles. Short green arrows indicate the final step 2.3, applied on the re-sampled signals (green dashed circles) to obtain the green square, the final output of interpolation.

the interpolation process can be formulated as follows:

$$
\begin{aligned}
&1.1.\quad U_{u,j} = \mathcal{WA}(U_{i,j}, U_{i+1,j}, m_i, m_{i+1}, m_u)\\
&1.2.\quad U_{u,j+1} = \mathcal{WA}(U_{i,j+1}, U_{i+1,j+1}, m_i, m_{i+1}, m_u)\\
&2.1.\quad U^r_{u,j} = \mathcal{RS}(U_{u,j}, v_j, v_u)\\
&2.2.\quad U^r_{u,j+1} = \mathcal{RS}(U_{u,j+1}, v_{j+1}, v_u)\\
&2.3.\quad U_{u,u} = \mathcal{WA}(U^r_{u,j}, U^r_{u,j+1}, v_j, v_{j+1}, v_u)
\end{aligned}
\tag{6}
$$

Steps 1.1 and 1.2 are for taking weighted average between the signals with the same scanning velocity (to handle different masses). Steps 2.1 and 2.2 are for re-sampling the outcomes (to handle different frequencies), and step 2.3 is for taking average between the two re-sampled signals (to handle different velocities). For scanning velocities slower than $v_{min} = v_1$ or faster than $v_{max} = v_5$, we set v_u to the corresponding extreme limit. The same is done for m_u outside the range of $[m_{min} = m_1, m_{max} = m_3]$.

5 Experimental Results

This section presents the performance evaluation methods and results of our data-driven method of recreating virtual textures from real ones. We rely on frequency spectra for visual comparison and E_s values for objective evaluation. For that, we collected force data using the tribometer from six texture samples shown in Fig. 10. We used three masses ($m_1 = 60$, $m_2 = 85$, and $m_3 = 110$ g) and five scanning velocities ($v_1 = 3$, $v_2 = 4$, $v_3 = 5$, $v_4 = 6$, and $v_5 = 7$ cm/s) in the experimental conditions. As shown in Fig. 11, we trained neural networks for the conditions marked by blue squares and performed cross-validation for the conditions marked by green squares.

5.1 Baseline Performance

We first evaluated the similarity between the repeated measurements from the same material under the same experimental condition to establish the baseline performance for the error metric E_s. Ideally, E_s should be zero in such cases, but due to various noises and uncertainties, the values of E_s are always greater than zero. Two baseline FFT plots, the best and the worst, are given in Fig. 12 with $E_s = 0.05$ and 0.27, respectively. See Fig. 1 and Fig. 2 in Appendix for the FFT plots of all the six materials with E_s values.

In general, the repeated measurements looked very similar for most of the materials and were almost identical for the dotted sheet. The average E_s was 0.14 with the standard deviation of 0.06. We also observed that increasing the scanning speed tended to shift the main frequency component to right (to higher frequencies).

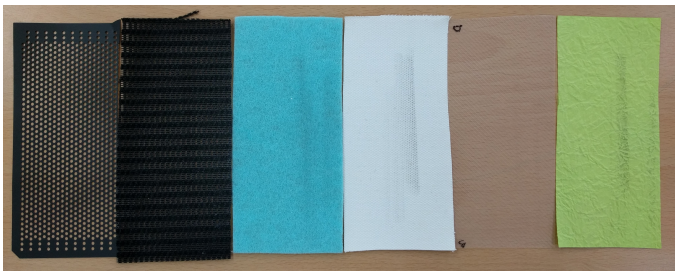


Fig. 10. Six texture samples used in the evaluation. From left to right: dotted sheet (`dot`), chair fabric (`chr`), felt fabric (`flt`), painting canvas (`can`), transparent plastic sheet (`pla`), and scrunched paper (`scr`).

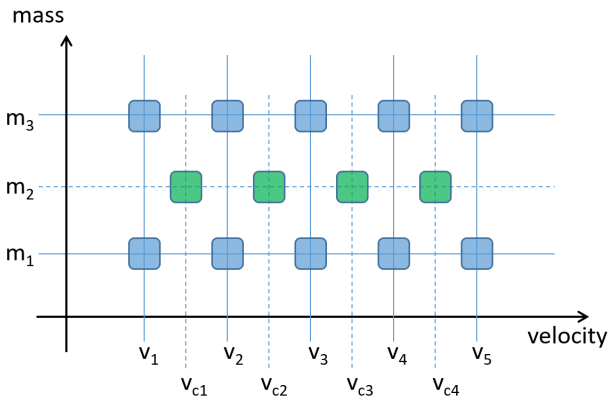


Fig. 11. Mass-velocity grid for cross-validation. m_i $(i = 1, 2, 3)$ indicates mass and v_j $(j = 1, 2, 3, 4, 5)$ and v_{ck} $(k = 1, 2, 3, 4)$ scanning velocities. For the experimental conditions highlighted by blue squares, actuation signals is obtained from the corresponding trained neural networks. For the others marked by green squares, the signals are obtained by interpolating neighbors.

5.2 Training Performance

Next, we examined the similarity between the virtual forces generated by the synthesized PRBS-like actuation signals and the real texture forces. Initial attempts showed that when the actuation signal is amplified with the default gain (1.0 corresponding to 100 V_{pp}), the amplitude levels of the power spectral densities of virtual textures do not match those of real ones in some conditions. Thus we tested a range of gains from 0.6 to 1.4 with a step size of 0.2, and we picked the one that generated almost the same maximum magnitude of power density.

Out of all the results, the best and the worst cases are shown in Fig. 13. Interested readers can refer to the results of all the experimental conditions given in Fig. 3 and Fig. 4

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TOH.2019.2932990, IEEE Transactions on Haptics

in Appendix[4]. In most cases, the neural networks estimated actuation signals successfully so that the virtual spectra were similar to the real ones. Note that the E_s values in Fig. 13 should be interpreted in reference to the baseline E_s values.[5] Among the six materials, dotted sheet showed the highest similarity with the average $\bar{E}_s = 0.33$, while chair fabric the lowest with $\bar{E}_s = 0.61$. Painting canvas showed the second best performance ($\bar{E}_s = 0.35$) followed by plastic sheet ($\bar{E}_s = 0.38$), felt fabric ($\bar{E}_s = 0.40$), and scrunched paper ($\bar{E}_s = 0.42$). The average E_s across all the materials was 0.42 with the standard deviation 0.13. A closer look reveals that the three best materials have some sort of uniform patterns on their surfaces, are made from sturdy materials, and generally feel rough. In contrast, the three worst materials are all made from soft fabrics with random or no clear textural patterns on their surfaces.

The worst case in Fig. 13 reveals the prominence of harmonic content in the virtual spectrum obtained from the chair fabric, one of the soft materials. We speculate a few reasons for this behavior, which includes complex surface properties and limitation of the data collection system. First, measurements from soft materials exhibit significant stochastic nature as opposed to the harder samples resulting in regular patterns (compare the two plots in Fig. 12; the right one for felt fabric shows much worse consistency between repeated measurements). Such stochastic properties cannot be modeled adequately using the deterministic model (neural network) used in our system. This problem could have been made worse by the use of the touch pen with a soft rubber tip for scanning all the materials. Second, we observed that for most of the soft materials the overall magnitude of the re-created spectra was lower than for the real ones. We applied a gain to adjust the magnitude and reduce the difference, and it tended to cause all frequency components (desired or undesired) to be emphasized.

5.3 Cross-validation

We next tested how well for untested conditions (for which no neural networks were trained), our two-step interpolation between adjacent tested conditions (for which neural networks were trained) generates adequate actuation signals. These untested conditions are shown by green squares in Fig. 11 and include one mass ($m_2 = 85$ g) and four velocities ($v_{c1} = 3.5$, $v_{c2} = 4.5$, $v_{c3} = 5.5$, and $v_{c4} = 6.5$ cm/s). For each untested condition, the actuation signal is obtained by interpolating four adjacent actuation signals. For example, for condition (m_2, v_{c1}), the actuation signal is estimated from (m_1, v_1), (m_1, v_2), (m_3, v_1), and (m_3, v_2).

Two FFT plots that showed the best and worst cross-validation performance are shown in Fig. 14. Similar to the training performance reported in the previous section, the best was achieved for painting canvas with $E_s = 0.19$ and the worst for chair fabric with $E_s = 0.87$ (also see Fig. 5 in Appendix for the plots for all cases). The average E_s was 0.49 (standard deviation 0.17). Comparing these results with those of Section 5.2 indicates that our two-step interpolation scheme works appropriately. The worse case in Fig. 14 was expected since it was interpolated from the worst case spectrums in Fig. 13.

5. It is unclear how to compare two different E_s values, e.g., it can be additive or multiplicative, or neither.

5. The appendix is included in the multimedia object available separately in IEEE Xplorer.

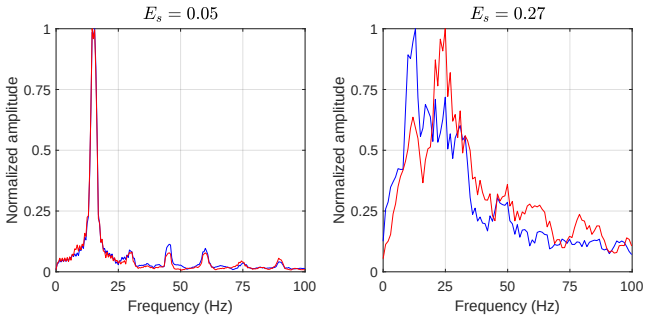


Fig. 12. Best (left) and worst (right) baseline FFT plots for the repeatability of measurements. The best was achieved with plastic sheet under the condition (m_1, v_1) and the worst with felt fabric under (m_1, v_5).

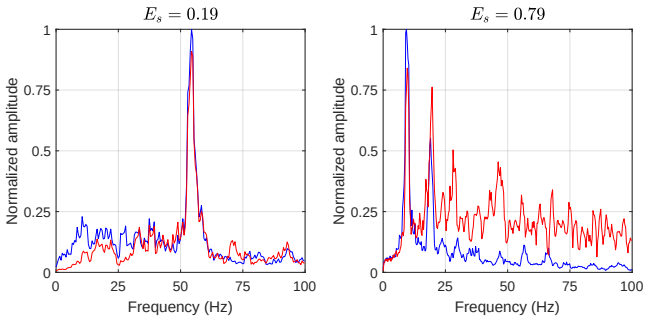


Fig. 13. Best (left) and worst (right) synthesized FFT plots. The best was achieved for painting canvas under (m_3, v_4) and the worst for chair fabric under (m_1, v_3). The real spectra are shown in blue and the virtual ones in red.

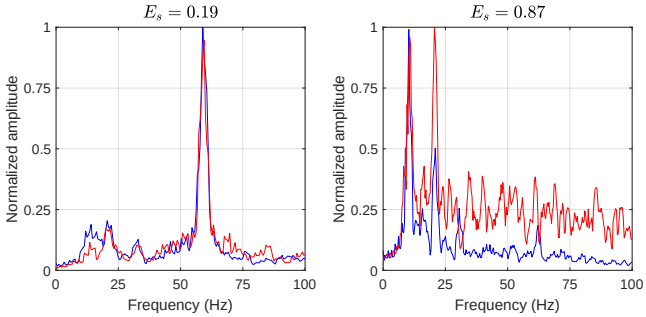


Fig. 14. Best (left) and worst (right) cross-validation FFT plots. The best was achieved for painting canvas under (m_2, v_{c4}) and the worst for chair fabric under (m_2, v_{c3}). The real spectra are shown in blue and the virtual ones in red.

6 User Studies

In the previous section, we have shown that our data-driven texture modeling method can create adequate virtual spectra similar to real spectra for both tested and untested conditions. We also evaluated the performance of our texture modeling and rendering algorithm by means of two user studies. The emphasis of User Study 1 was on the perceptual similarity between real and virtual textures, evaluated in an absolute manner, while User Study 2 focused on comparing the perceptual performance between our method and the record-and-playback method (similar to [18], [46]). The user

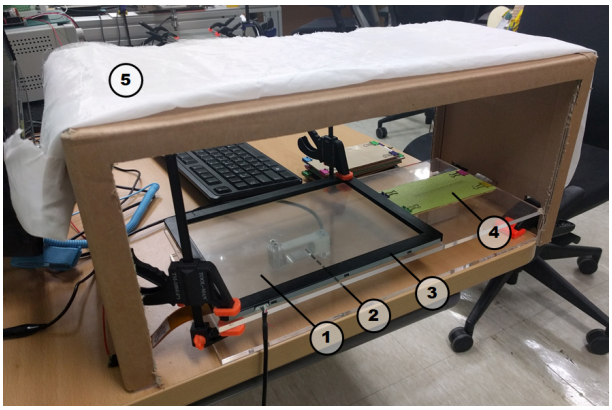


Fig. 15. Experimental setup used in the two user studies. 1) 3M capacitive touchscreen as an electrovibration display. 2) Load cell, installed under the panel, to measure user-applied force. 3) Infrared frame to track the user's finger position. 4) Real texture sample. 5) Curtain to block participants' view.

studies were approved by the Institutional Review Board at the author's institution (PIRB-2017-E070).

6.1 Methods

6.1.1 Stimuli

Our data-driven texture modeling and rendering algorithms were applied to the six real texture materials used in the performance evaluation (Fig. 10). The experimental conditions for data collection were also the same. For rendering, we added an infrared frame (T Series 10", E&T Tech, China) on the 3M panel to sense finger position (Fig. 15). A bar-type load-cell (CBCL-2L, CAS Scale, South Korea) was attached to the bottom of the 3M panel to measure user-applied force. Users were asked to wear an anti-static wristband to ensure that their bodies were electrically grounded. The wristband had a 1 MΩ current-limiting resistor to prevent excessive current from passing through the user's body.

A computer program read user-applied force and finger position at every 25 ms (40 Hz) and updated the output actuation signal at every 100 ms (10 Hz). A data-acquisition card (PCI-6229, NI, USA) was used to read analog input from the load cell and generate analog output (actuation signal) with a sampling rate of 1 kHz. For velocity estimation, we employed a differentiation method based on the first-order adaptive windowing [67] to obtain reliable and stable velocity estimates from position data.

During the experiments, participants sat in a chair in front of a desk on which the experimental hardware was set. A real texture sample was placed on the right of the 3M panel. To remove all irrelevant sensory cues, participants' vision to the hardware was blocked by a curtain.

For User Study 2, we implemented the record-and-playback method for texture rendering. First, the lateral forces were recorded by scanning the surface of each texture material under the given experimental condition. Then, the signals were amplified and directly applied to the 3M panel. For the conditions under which no force data was collected, the same two-step interpolation method was applied before amplification.

6.1.2 Performance measure

Given a pair of textures, either real vs. real or real vs. virtual, participants were asked to evaluate the perceived similarity in a scale of 0–100, with 0 being completely different and 100 being completely same. Participants were told to focus on the surface textural patterns, but not the materials themselves.[6]Note that our texture modeling and rendering method still uses an electrovibration display, which can only generate lateral force. Other haptic properties, such as stiffness, viscosity, surface height variation, and heat transfer, are out of our control. Those differences between real samples and virtual stimuli (giving glassy feel other than virtual textures) are likely to negatively affect the subjective similarity.

6.1.3 Task and Procedure: User Study 1

User Study 1 consisted of two phases. In the first phase, participants were provided with side-by-side pairs of only real texture samples and were asked to rate their similarities. The objective was to measure the overall upper and lower bounds of perceived similarity since humans tend to avoid giving extreme values (e.g., see [70]). Additionally, this phase helped participants stabilize their perceptual bases and scales for similarity rating. Participants were randomly presented with 12 pairs of materials, six of the same materials and the other six of different materials. The latter was composed of one from the materials giving similar sensations (`pla-scr`, `dot-pla`, and `scr-can`) and the other from the rest with quite different feels (`dot-flt`, `can-chr`, and `chr-flt`). We call the former a moderately-different group and the latter a highly-different group.

In the second phase, participants evaluated the similarity between a virtual texture and its corresponding real texture. The order of the six texture pairs was randomized for each participant. Each texture pair was repeated three times, making a total of 18 trials. Participants were instructed to explore each texture, either real or virtual, from side to side maintaining a regular scanning speed, not too fast and not too slow, within the range of 3 to 7 cm/s. They were also told to apply a constant normal pressure, not too heavy and not too light, within the range of 60 to 110 g.

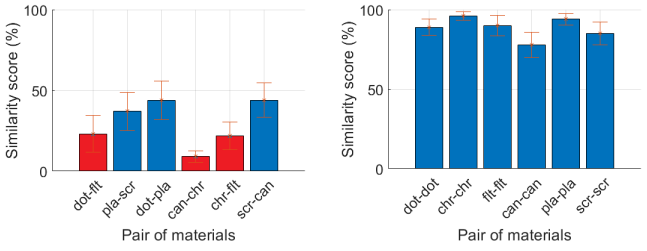
We recruited 20 participants (13 males and 7 females; 18–26 years old with an average of 22.4) who had no prior experiences in variable friction displays. None of them reported any known sensory or motor impairment. Participants signed an informed consent form after we explained the goals and procedure of the experiment. Each participant was paid KRW 5,000 (≃ USD 4.5) after the experiment.

6.1.4 Task and Procedure: User Study 2

User Study 2 was to compare the user-perceived similarity of a virtual texture to a real texture when the virtual texture was modeled and rendered using our method or the record-and-playback method. Similar to the second phase of User Study 1, in User Study 2, participants were presented with

6. It is unknown how well humans can make judgments with such perceptual decomposition. Nonetheless, some very recent studies reported users studies that used more explicitly decomposed perceptual criteria [68], [69], and they observed no detrimental evidence against the validity of the approach.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TOH.2019.2932990, IEEE Transactions on Haptics



(a) Different real objects (b) Identical real objects

Fig. 16. Average pairwise similarity scores between real textures. Left: two different real materials. Red bars are for highly-different textures, and blue bars for moderately-different textures. Right: two similar real materials. Error bars show standard errors.

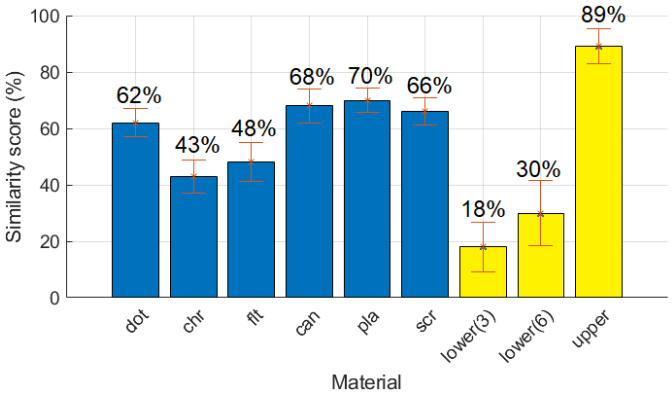


Fig. 17. Average similarity scores comparing real and virtual textures. Mean values are given on top of each bar. Error bars indicate standard errors. The yellow bars show lower and upper bounds, with lower(3) including only the three highly-different textures and lower(6) including both highly- and moderately-different groups (6 pairs).

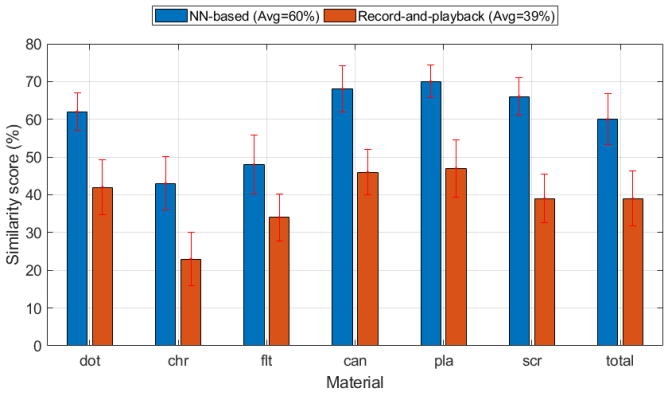


Fig. 18. Comparison between the neural network-based and record-and-playback (rp) methods for the perceived similarity of a virtual texture to a real texture. Error bars indicate standard errors.

pairs of a real material and a virtual texture rendered using the record-and-playback method. Each pair was repeated three times, making a total of 18 trials per participant. Twelve volunteers (8 males and 4 females; 20–27 years old, average 23.9) participated in this study. Each participant was paid KRW 5,000 ($\simeq$ USD 4.5) after the experiment.

6.2 Results

In User Study 1, we collected 600 similarity scores (20 participants $\times$ (12+18) pairs). The average similarity scores between different real objects and between identical real objects are shown in Fig. 16. Among the different pairs, `scr-can` received the highest score (44) and `can-chr` the lowest (9). The grand average was 30. For the identical pairs, `chr-chr` obtained the highest score (96) and `can-can` the lowest (78) with the grand of 89.

The average pairwise similarity scores between real materials and their virtual counterparts rendered by our data-driven method are shown in Fig. 17. For each of the six materials, the average is taken from 60 scores (20 participants $\times$ 3 trials). The plot also represents lower and upper bounds by yellow bars. lower(3) indicates the average of only the highly-different group (3 pairs) while lower(6) is the average of both the highly-different group and the moderately-different group (6 pairs). Plastic sheet received the highest score (70) while chair fabric the lowest score (43). The grand mean was 60 with the standard deviation of 15.16.

The results of User Study 2 are summarized in Fig. 18. For all the materials, our neural-network based method outperformed the record-and-playback method by considerable differences. Note that the results of the neural-network method were measured in User Study 1. The average similarity scores were 60% and 39%, respectively. We conducted a two-way between-subject ANOVA with *rendering method* and *material* as the two independent factors. The results showed that similarity score was significantly different across the six materials ($F(5, 564) = 10.29$, $p < 0.001$) and between the two rendering methods ($F(1, 564) = 97.78$, $p < 0.001$). Their interaction term was not statistically significant ($F(5, 564) = 1.49$, $p = 0.1896$).

6.3 Discussion

Our neural network-based data-driven texture rendering method resulted in the similarity scores well above both the lower limits for all the materials, as shown in Fig. 17. However, Fig. 18 says that the same is not true for the record-and-playback method. The means were 18% (lower(3)) $<$ 30% (lower(6)) $<<$ 39% (record and playback) $<$60% (our method) $<$ 89% (upper). This comparison indicates that 1) the simplistic record-and-playback algorithm performs poorly, sometimes resulting in mediocre output, e.g., `chr` in Fig. 18; 2) our method evidently excels the record-and-playback method in capturing and re-creating surface fine details, supporting the effectiveness of our neural network-based method in capturing and compensating for the nonlinear dynamics of electrovibration display; and 3) however, our method is still left with considerable room for further improvement in order to accomplish the same level of realism between identical real texture patches.

In the two user studies, participants were asked to rate the pairwise similarity between real and virtual textures on the basis of textural patterns while ignoring other material differences such as softness and temperature. This was inevitable because an electrovibration display is a partial force transducer and a user feels both real stimuli (normal force and temperature from the stiff 3M panel glass) and a virtual stimulus (modulated vibrational frictional force); in this regard, an electrovibration display is for haptic augmented

reality [71]. For this reason, we cannot guarantee that the perceptual similarity scores reported in the user studies were determined solely by the textural similarity.

This issue of data validity can be confirmed partly by examining how well the perceived similarity scores matched the error metrics $\bar{E}_s$ between real and virtual textures. For each texture material, (average similarity score, $\bar{E}_s$) were `dot` (62%, 0.55), `chr` (43%, 0.68), `flt` (48%, 0.43), `can` (68%, 0.32), `pla` (70%, 0.44), and `scr` (66%, 0.50). Their correlation coefficient was -0.61, indicating a strong relationship between the similarity score and E_s. However, the two metrics did not have exactly the same order. In terms of similarity score, the order of the materials from the best to the worst was `pla`, `can`, `scr`, `dot`, `flt`, and `chr`. In terms of $\bar{E}_s$, the order was `can`, `flt`, `pla`, `scr`, `dot`, and `chr`. Considering that the similarity scores of `pla`, `can`, `scr`, and `dot` were quite similar (Fig. 17; compare the error bars), the only notable exception was `flt`, which had the second worst similarity score in spite of the second best $\bar{E}_s$. Although the source of this peculiarity is unknown at the moment, the results of perceptual similarity rating appear to reflect the textural differences to a reasonable extent.

Two materials received particularly low scores: chair fabric (43%) and felt fabric (48%). Only these two were made of soft materials, and the other four samples were stiff. Electrovibration displays using a glass-based 3M panel cannot render surfaces with different softness values. Although the participants were instructed to rate the similarity of texture, such apparent dissimilarity in softness could have affected their scores. In addition, when scanned, chair fabric and felt fabric feel much smoother than the other four samples. Driving an electrovibration display with square waves generally elicits rougher sensations than with sine waves of the same frequency. Our actuation signals using PRBS behave like square waves, and it could have hindered rendering softness sensations.

On the other hand, our method can be appropriate for producing sharp and crisp sensations. This advantage is deemed to have contributed to the high performance (over 62% of perceptual similarity) of the other four materials. The best performance (70%) was achieved with plastic sheet. This sample had equally spaced diagonal ridges on the surface, which are in good perceptual agreement with square waves. Its rigid surface also felt very similar to the glass surface of the 3M panel. Textures made of solid materials with distinguishable patterns have more chance for realistic re-creation through our system.

7 Conclusions and Future Work

In this work, we presented a neural network-based data-driven method for realistic texture rendering on an electrovibration display. Our intention was to build an inverse dynamics model for the electrovibration display and then use it for texture rendering. To this end, we applied a full-band PRBS signal to the display and collected resulting forces using a motorized linear tribometer under known normal force and scanning velocity conditions. Then, we inversely trained NARX neural networks to learn from the forces and mimic the actuation signals. We used the trained network to estimate an actuation signal from the forces collected from the surface of a real material under a given condition. For arbitrary conditions, we proposed a two-step interpolation scheme to estimate an actuation signal from the ones generated under known conditions. We then actuated the electrovibration display using the estimated actuation signals and collected resulting forces. Comparing the rendered forces with the real ones, we showed promising agreement between their corresponding frequency spectra. We also conducted a human user study to measure the perceptual similarity between the real and virtual textures rendered using our method. The scores were then compared to those obtained from rendering virtual textures using a record-and-playback method. Our neural network-based method resulted in higher similarity scores than record-and-playback with 60% vs. 39%, respectively. This suggested that our data-driven texture modeling and rendering method, accompanied by an interpolation scheme, significantly improves the quality of virtual textures rendered on an electrovibration display.

In general, our inverse black-box model is an alternative to physics-based force models. Other researchers have pursued the latter approach while improving the bandwidth and accuracy of the models [15], [17], [32], [72]. The most advanced one has been recently developed by Shultz et al. [31], who proposed using a current controlled high-frequency sinusoidal modulation to achieve broadband force rendering. Their model provides precious insights to the behavior of electroadhesive displays and led to a input signal design method that yields a flat frequency response up to 6 kHz when single sine waves are applied. Additionally, the system is shown to exhibit a linear overall behavior in response to a bandpassed white noise signal within the range of 500 to 1500 Hz. However, the nonlinearity of the system still appears in response to a 100 Hz-15 kHz chirp signal, and how it will respond to real texture signals that have wideband continuous spectra remains to be seen. In addition, their setup requires higher reactive current, a high-performance current controller with advanced electronics, and precise knowledge of the system being used.

In comparison, our method preserves the general merits of black-box approach. Its modeling power is very high as long as suitable input-output samples are provided, and using a complete inverse dynamics model provides better transparency for data-driven texture rendering. Our method should be applicable to all electrovibration displays requiring only basic electronics components and consuming relatively low current. However, further investigations are needed to study the bandwidth our method can offer and the extent of nonlinearity it can handle.

We are keen to improve the current work in several aspects. The proposed method did not perform well with soft materials. We speculate that our force measurement system with a soft rubber touch pen is not ideal for recording the random and complex surface properties of these soft materials. We will try a different type of tool for their influence and/or other machine learning models. In addition, pseudo-random binary signals may be replaced with pseudo-random multilevel signals to better capture such complexities. Furthermore, we will examine a more compact model with an interpolation scheme embedded inside, instead of having multiple models with a separate

interpolation procedure.

ACKNOWLEDGMENTS

We are grateful to Soohee Han for his advice on the use of neural network. This work was supported in part by the Cross-Ministry Giga Korea Project grant (No. GK17C0100) and by a grant for High-Definition Haptic Technology for Hyper Reality Contents (No.2017-0-00179) from IITP.

REFERENCES

[1] O. Bau, I. Poupyrev, A. Israr, and C. Harrison, "Teslatouch: electrovibration for touch surfaces," in *Proc. of ACM UIST*, 2010, pp. 283–292.
[2] T. Nakamura and A. Yamamoto, "Multi-finger electrostatic passive haptic feedback on a visual display," in *Proc. of IEEE WHC*, 2013, pp. 37–42.
[3] R. H. Osgouei, J. R. Kim, and S. Choi, "Identification of primitive geometrical shapes rendered using electrostatic friction display," in *Proc. of IEEE HAPTICS*, 2016, pp. 198–204.
[4] T. Watanabe and S. Fukui, "A method for controlling tactile sensation of surface roughness using ultrasonic vibration," in *Proc. of IEEE ICRA*, 1995, pp. 1134–1139.
[5] L. Winfield, J. Glassmire, J. E. Colgate, and M. Peshkin, "T-pad: Tactile pattern display through variable friction reduction," in *Proc. of IEEE WHC*, 2007, pp. 421–426.
[6] M. Biet, F. Giraud, and B. Lemaire-Semail, "Squeeze film effect for the design of an ultrasonic tactile plate," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 54, no. 12, pp. 2678–2688, 2007.
[7] M. Amberg, F. Giraud, B. Semail, P. Olivo, G. Casiez, and N. Roussel, "Stimtac: a tactile input device with programmable friction," in *Proc. of ACM UIST*, 2011, pp. 7–8.
[8] E. Gray, "Improvement in electric telegraphs for transmitting musical tones," *cooperative Classification H04L27/26*, 1875.
[9] A. Johnsen and K. Rahbek, "A physical phenomenon and its applications to telegraphy, telephony, etc." *J. Inst. Electr. Eng.*, vol. 61, no. 320, pp. 713–725, 1923.
[10] E. Mallinckrodt, A. L. Hughes, and W. Sleator Jr., "Perception by the skin of electrically induced vibrations." *Science*, 1953.
[11] S. Grimnes, "Electrovibration, cutaneous sensation of microampere current," *Acta Physiologica*, vol. 118, no. 1, pp. 19–25, 1983.
[12] R. M. Strong and D. E. Troxel, "An electrotactile display," *IEEE Trans. Man-Machine Systems*, vol. 11, no. 1, pp. 72–79, 1970.
[13] D. J. Beebe, C. Hymel, K. Kaczmarek, and M. Tyler, "A polyimide-on-silicon electrostatic fingertip tactile display," in *Proc. of IEEE EMBS*, vol. 2, 1995, pp. 1545–1546.
[14] H. Tang and D. J. Beebe, "A microfabricated electrostatic haptic display for persons with visual impairments," *IEEE Trans. Rehabil. Eng.*, vol. 6, no. 3, pp. 241–248, 1998.
[15] D. J. Meyer, M. Peshkin, and J. E. Colgate, "Fingertip friction modulation due to electrostatic attraction," in *Proc. of IEEE WHC*, 2013, pp. 43–48.
[16] S.-C. Kim, A. Israr, and I. Poupyrev, "Tactile rendering of 3d features on touch surfaces," in *Proc. of ACM UIST*, 2013, pp. 531–538.
[17] E. Vezzoli, M. Amberg, F. Giraud, and B. Lemaire-Semail, "Electrovibration modeling analysis," in *Haptics: Neuroscience, Devices, Modeling, and Applications*. Springer, 2014, pp. 369–376.
[18] G. Ilkhani, M. Aziziaghdam, and E. Samur, "Data-driven texture rendering with electrostatic attraction," in *Proc. of EuroHaptics*. Springer, 2014, pp. 496–504.
[19] H. Kim, J. Kang, K.-D. Kim, K.-M. Lim, and J. Ryu, "Method for providing electrovibration with uniform intensity," *IEEE Trans. Haptics*, vol. 8, no. 4, pp. 492–496, 2015.
[20] Y. Zhang and C. Harrison, "Quantifying the targeting performance benefit of electrostatic haptic feedback on touchscreens," in *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*. ACM, 2015, pp. 43–46.
[21] Q. Wang, X. Ren, S. Sarcar, and X. Sun, "Ev-pen: Leveraging electrovibration haptic feedback in pen interaction," in *Proceedings of the ACM on Interactive Surfaces and Spaces*, 2016, pp. 57–66.
[22] J. Kang, H. Kim, S. Choi, K. D. Kim, and J. Ryu, "Investigation on low voltage operation of electrovibration display," *IEEE Trans. Haptics*, vol. 10, no. 3, pp. 371–381, 2017.
[23] J. Linjama and V. Mäkinen, "E-sense screen: Novel haptic display with capacitive electrosensory interface," in *Proc. of HAID*, 2009.
[24] R. Haghighi Osgouei, J. R. Kim, and S. Choi, "Improving 3d shape recognition with electrostatic friction display," *IEEE Trans. Haptics*, vol. 10, no. 4, pp. 533–544, 2017.
[25] R. L. Klatzky, S. Adkins, P. Bodas, R. Haghighi Osgouei, S. Choi, and H. Z. Tan, "Perceiving texture gradients on an electrostatic friction display," in *Proc. of IEEE WHC*, 2017, pp. 154–158.
[26] J. R. Kim, R. Haghighi Osgouei, and S. Choi, "Effects of visual and haptic latency on touchscreen interaction: A case study using painting task," in *Proc. of IEEE WHC*, 2017, pp. 159–164.
[27] D. Pyo, S. Ryu, S.-C. Kim, and D.-S. Kwon, "A new surface display for 3d haptic rendering," in *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*. Springer, 2014, pp. 487–495.
[28] 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 *Proc. of IEEE HAPTICS*, 2014, pp. 63–67.
[29] F. Giraud, M. Amberg, and B. Lemaire-Semail, "Merging two tactile stimulation principles: electrovibration and squeeze film effect," in *Proc. of IEEE WHC*, 2013, pp. 199–203.
[30] C. D. Shultz, M. A. Peshkin, and J. E. Colgate, "Surface haptics via electroadhesion: Expanding electrovibration with johnsen and rahbek," in *Proc. of IEEE WHC*, 2015, pp. 57–62.
[31] C. Shultz, M. Peshkin, and J. E. Colgate, "The application of tactile, audible, and ultrasonic forces to human fingertips using broadband electroadhesion," *IEEE transactions on haptics*, vol. 11, no. 2, pp. 279–290, 2018.
[32] K. A. Kaczmarek, K. Nammi, A. K. Agarwal, M. E. Tyler, S. J. Haase, and D. J. Beebe, "Polarity effect in electrovibration for tactile display," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 10, pp. 2047–2054, 2006.
[33] Y. Vardar, B. Güçlü, and C. Basdogan, "Effect of waveform in haptic perception of electrovibration on touchscreens," in *Proc. of EuroHaptics*. Springer, 2016, pp. 190–203.
[34] C. D. Shultz, M. A. Peshkin, and J. E. Colgate, "On the electrical characterization of electroadhesive displays and the prominent interfacial gap impedance associated with sliding fingertips," in *Proc. of IEEE HAPTICS*, 2018, pp. 151–157.
[35] R. Hover, G. Kósa, G. Szekly, and M. Harders, "Data-driven haptic rendering from viscous fluids to visco-elastic solids," *IEEE Trans. Haptics*, vol. 2, no. 1, pp. 15–27, 2009.
[36] S. Yim, S. Jeon, and S. Choi, "Data-driven haptic modeling and rendering of viscoelastic and frictional responses of deformable objects," *IEEE Trans. Haptics*, vol. 9, no. 4, pp. 548–559, 2016.
[37] S. Andrews and J. Lang, "Interactive scanning of haptic textures and surface compliance," in *Proc. of 3DIM*, 2007, pp. 99–106.
[38] J. Lang and S. Andrews, "Measurement-based modeling of contact forces and textures for haptic rendering," *IEEE Trans. Vis. Comput. Graphics*, vol. 17, no. 3, pp. 380–391, 2011.
[39] V. L. Guruswamy, J. Lang, and W.-S. Lee, "IIR filter models of haptic vibration textures," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 1, pp. 93–103, 2011.
[40] J. Romano, T. Yoshioka, and K. Kuchenbecker, "Automatic filter design for synthesis of haptic textures from recorded acceleration data," in *Proc. of IEEE ICRA*, 2010, pp. 1815–1821.
[41] J. Romano and K. Kuchenbecker, "Creating realistic virtual textures from contact acceleration data," *IEEE Trans. Haptics*, vol. 5, no. 2, pp. 109–119, 2012.
[42] H. Culbertson, J. Romano, P. Castillo, M. Mintz, and K. Kuchenbecker, "Refined methods for creating realistic haptic virtual textures from tool-mediated contact acceleration data," in *Proc. of IEEE HAPTICS*, 2012, pp. 385–391.
[43] H. Culbertson, J. Unwin, B. Goodman, and K. Kuchenbecker, "Generating haptic texture models from unconstrained tool-surface interactions," in *Proc. of IEEE WHC*, 2013, pp. 295–300.
[44] S. Shin, R. Haghighi Osgouei, K.-D. Kim, and S. Choi, "Data-driven modeling of isotropic haptic textures using frequency-decomposed neural networks," in *Proc. of IEEE WHC*, 2015, pp. 131–138.
[45] A. Abdulali and S. Jeon, "Data-driven rendering of anisotropic haptic textures," in *Proc. of AsiaHaptics*. Springer, 2016, pp. 401–407.

[46] G. Ilkhani, M. Aziziaghdam, and E. Samur, "Data-driven texture rendering on an electrostatic tactile display," *International Journal of Human–Computer Interaction*, vol. 33, no. 9, pp. 756–770, 2017.

[47] H. Culbertson, J. J. Lopez Delgado, and K. Kuchenbecker, "The penn haptic texture toolkit for modeling, rendering, and evaluating haptic virtual textures," 2014.

[48] J. Jiao, Y. Zhang, D. Wang, Y. Visell, D. Cao, X. Guo, and X. Sun, "Data-driven rendering of fabric textures on electrostatic tactile displays," in *Proc. of IEEE HAPTICS*, 2018, pp. 169–174.

[49] T. Fielder and Y. Vardar, "A Novel Texture Rendering Approach for Electrostatic Displays," in *International Workshop on Haptic and Audio Interaction Design - HAID2019*, Lille, France, Mar. 2019. [Online]. Available: https://hal.archives-ouvertes.fr/hal-02011782

[50] Y. Vardar, A. İşleyen, M. K. Saleem, and C. Basdogan, "Roughness perception of virtual textures displayed by electrovibration on touch screens," in *Proc. of IEEE WHC*, 2017, pp. 263–268.

[51] T. Wang and X. Sun, "Electrostatic tactile rendering of image based on shape from shading," in *International Conference on Audio, Language and Image Processing*, 2014, pp. 775–779.

[52] S. Wu, X. Sun, Q. Wang, and J. Chen, "Tactile modeling and rendering image-textures based on electrovibration," *The Visual Computer*, vol. 33, no. 5, pp. 637–646, 2017.

[53] R. Haghighi Osgouei, S. Shin, J. R. Kim, and S. Choi, "An inverse neural network model for data-driven texture rendering on electrovibration display," in *Proc. of IEEE HAPTICS*, 2018, pp. 270–277.

[54] L. Ljung, *System identification: theory for the user*. Prentice Hall, 1998.

[55] F. Giri and E.-W. Bai, *Block-oriented nonlinear system identification*. Springer, 2010, vol. 1.

[56] R. Haber and L. Keviczky, *Nonlinear system identification-Input-output modeling approach*. Kluwer Academic Publishers, 1999.

[57] O. Nelles, *Nonlinear system identification: from classical approaches to neural networks and fuzzy models*. Springer Science & Business Media, 2013.

[58] E.-W. Bai, "Identification of linear systems with hard input nonlinearities of known structure," *Automatica*, vol. 38, no. 5, pp. 853–860, 2002.

[59] I. Leontaritis and S. A. Billings, "Input-output parametric models for non-linear systems part I: deterministic non-linear systems," *Int. J. Control*, vol. 41, no. 2, pp. 303–328, 1985.

[60] H. P. H. Anh and K. K. Ahn, "Hybrid control of a pneumatic artificial muscle (pam) robot arm using an inverse narx fuzzy model," *Eng. Appl. Artif. Intell.*, vol. 24, no. 4, pp. 697–716, 2011.

[61] H. Sahoo, P. Dash, and N. Rath, "Narx model based nonlinear dynamic system identification using low complexity neural networks and robust h-inf filter," *Appl. Soft Comput.*, vol. 13, no. 7, pp. 3324–3334, 2013.

[62] T. Yoshioka, S. J. Bensmaia, J. C. Craig, and S. S. Hsiao, "Texture perception through direct and indirect touch: An analysis of perceptual space for tactile textures in two modes of exploration," *Somatosensory & motor research*, vol. 24, no. 1-2, pp. 53–70, 2007.

[63] R. L. Klatzky, S. J. Lederman, C. Hamilton, M. Grindley, and R. H. Swendsen, "Feeling textures through a probe: Effects of probe and surface geometry and exploratory factors," *Attention, Perception, & Psychophysics*, vol. 65, no. 4, pp. 613–631, 2003.

[64] L. Tan and J. Jiang, *Digital signal processing: fundamentals and applications*. Academic Press, 2013.

[65] S. Choi and H. Z. Tan, "Perceived instability of virtual haptic texture. I. Experimental studies," *Presence: Teleoperators and Virtual Environment*, vol. 13, no. 4, pp. 395–415, 2004.

[66] S. Choi and H. Z. Tan, "Toward realistic haptic rendering of surface textures," *IEEE Computer Graphics & Applications (Special Issue on Haptic Rendering - Beyond Visual Computing)*, vol. 24, no. 2, pp. 40–47, 2004.

[67] F. Janabi-Sharifi, V. Hayward, and C.-S. J. Chen, "Discrete-time adaptive windowing for velocity estimation," *IEEE Trans. on Control System Technology*, vol. 8, no. 6, pp. 1003–1009, 2000.

[68] S. Shin and S. Choi, "Geometry-based haptic texture modeling and rendering using photometric stereo," in *Proc. IEEE HAPTICS*, 2018, pp. 262–269.

[69] S. Shin and S. Choi, "Effects of haptic texture rendering modalities on realism," in *Proc. ACM VRST*, no. 42, 2018.

[70] S. Jeon, S. Choi, and M. Harders, "Rendering virtual tumors in real tissue mock-ups using haptic augmented reality," *IEEE Trans. Haptics*, vol. 5, no. 1, pp. 77–84, Jan 2012.

[71] S. Jeon and S. Choi, "Haptic augmented reality: Taxonomy and an example of stiffness modulation," *Presence: Teleoperators and Virtual Environments*, vol. 18, no. 5, pp. 387–408, 2009.

[72] Y. Vardar, B. Güçlü, and C. Basdogan, "Effect of waveform on tactile perception by electrovibration displayed on touch screens," *IEEE transactions on haptics*, vol. 10, no. 4, pp. 488–499, 2017.

Reza Haghighi Osgouei is a research associate in the Imperial College Centre for Engagement and Simulation Science (ICCESS), UK. He received a Ph.D. and an M.Sc. in Computer Science and Engineering, both from POSTECH, Pohang, Republic of Korea. He also received an M.Sc. and a B.Sc. in Electrical Engineering: Control Systems from Tehran Polytechnic, Tehran, Iran and Sahand University of Technology, Tabriz, Iran, respectively. His work has been evolved around haptics and virtual reality, from haptic assistive technologies to surface haptics. He is now expanding his work onto medical applications, being involved in a virtual rehabilitation project and a haptic simulator for early detection of colorectal cancer. He is a member of the IEEE.

Jin Ryong Kim is a senior research scientist at Alibaba Research, USA. He received Ph.D. degree in Electrical and Computer Engineering and M.Sc. degree in Computer Science, both from Purdue University, IN, USA. He received B.Sc. and M.Sc. degree in Electrical and Computer Engineering from Hanyang University, Seoul, Korea. His current research interests include haptics, HCI, and UI/UX with emphasis on creating novel interaction through haptics technology in virtual reality, touchscreen, and 3D display interaction for new user experiences. He is a member of the IEEE.

Seungmoon Choi is a professor of Computer Science and Engineering at Pohang University of Science and Technology (POSTECH). He received the B.Sc. and M.Sc. degrees in Control and Instrumentation Engineering from Seoul National University in 1995 and 1997, respectively, and the Ph.D. degree in Electrical and Computer Engineering from Purdue University in 2003. He received a 2011 Early Career Award from IEEE Technical Committee on Haptics and several best paper awards from major international conferences. He was a co-chair of the IEEE Technical Committee on Haptics in 2009-2010. He serves/served on the editorial board of IEEE Transactions on Haptics, Presence, Virtual Reality, and IEEE Robotics and Automation Letters. He was the general co-chair of IEEE Haptics Symposium in 2014 and 2016 and the program chair of IEEE World Haptics 2015. His research interests lie on haptic rendering and perception, both in kinesthetic and tactile aspects. His basic research has been applied to mobile devices, automobiles, virtual prototyping, and motion-based remote controllers. He is a senior member of the IEEE.