Hasti: Haptic and Audio Synthesis for Texture Interactions

Sonny Chan[1], Chase Tymms[1] and Nicholas Colonnese[1]

***Abstract*— Multi-sensory stimuli can greatly enhance immersion in interactive virtual environments. Advances in graphics algorithms and technologies like VR displays have pushed the appearance of interactive virtual worlds to unprecedented fidelity, but rendering sound and, especially, touch feedback of comparable quality remains a challenge. We describe a method for real-time synthesis of vibrotactile haptic and audio stimuli for interactions with textured surfaces in 3-D virtual environments. Standard descriptions of object geometry and material properties, including displacement and roughness texture maps typically used for physically-based visual rendering, are employed to generate realistic sound and touch feedback consistent with appearance. Our method reconstructs meso- and microscopic surface features on the fly along a contact trajectory, and runs a micro-contact dynamics simulation whose outputs drive vibrotactile haptic actuators and modal sound synthesis. An exploratory, absolute identification user study was conducted as an initial evaluation of our synthesis methods.**

I. INTRODUCTION

In the coming years, virtual media design will undergo a transformative evolution. Today's conventional interfaces—computers and smartphones—will be replaced by sophisticated augmented reality glasses or head-mounted displays, and this change of interface will bring a change in how virtual media is designed. Specifically, there will be a dramatic paradigm shift on how authors create, trigger, and render various visual, audio, and haptic stimuli corresponding to a user's interaction with virtual content.

Fortunately, supported by the tremendous breakthroughs in graphics research in the past four decades, the transition from conventional fixed-plane 2-D visual content to adaptive 3-D mixed reality worlds will be straightforward and high-performance; in fact, this transition is already underway and showing strong results. In contrast, the path from today's conventional audio or haptic design methods to those suitable for immersive and interactive mixed reality environments is far from straightforward, and entirely new methods may need to be conceived.

In this paper, inspired by the modern approach in graphic rendering using physically-based material definitions and physics-inspired simulations of light transport, our goal was to construct a real-time method to synthesize convincing sound and vibrotactile haptic stimuli directly from existing geometric and material representations of virtual objects.

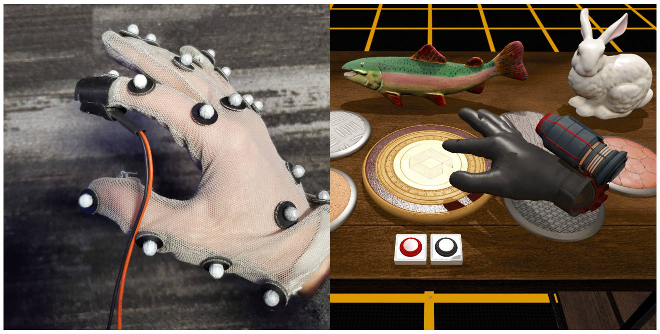


Fig. 1. A tracked glove with a vibrotactile actuator (left) allows us to interact with virtual objects and feel textured surfaces (right) with haptic and audio feedback synthesized using methods described in this work.

We introduce a high-rate "micro-contact" simulation to model sliding contact dynamics at a microscopic level. It creates a surface height profile on the fly from sparse contact events and standard texture image maps used for visual rendering. This contact simulation generates haptic signals to drive vibrotactile actuators and, using Hertz contact theory to estimate impact force profiles, also generates excitation signals to drive modal sound synthesis, producing realistic contact sounds between different materials in an interactive virtual environment (Fig. 1). We evaluate the real-time synthesis methods through an exploratory, absolute identification user study with haptics-only, audio-only, and combined conditions.

II. BACKGROUND & RELATED WORK

A. Haptic Texture Rendering

The display of virtual textured surfaces has long been of interest to the haptic rendering research community. In their early work, Minsky et al. [1] describe a method to create simulated textures by generating forces from the gradient of a texture "depth map" and displaying them through an actuated 2-DOF joystick. Siira & Pai [2] describe a method to render textures by sampling a surface height profile from a Gaussian distribution according to desired roughness, commanding proportional forces to a 2-DOF pantograph haptic display.

With the advent of 3-DOF kinaesthetic haptic devices, Ho et al. [3] devised a method for rendering surface textures represented as height maps on 3-D polygonal geometry. Otaduy et al. developed a contact force model between textured surfaces that accurately reflects observed perception of roughness, and showed that it can be used to render 6-DOF force feedback for interactions between textured objects with fine geometric details represented as image maps [4].

More recent efforts in haptic texture rendering have focused on characterizing real-world textures and using open-loop rendering to reproduce vibrations on a haptic display. Okamura et al. [5] measured vibrations from tapping and stroking a stylus across different materials, and rendered

---

[1] Authors are with Facebook Reality Labs, Redmond, Washington 98052, United States. `sonnychan@fb.com`; `tymms@fb.com`; `ncolonnese@fb.com`

patterned textures through a force feedback joystick by commanding a series of exponentially-decaying sinusoids fit to the measurements. In a process they've aptly named "haptography", Kuchenbecker et al. [6] describe a method to capture the feel of a textured surface with a handheld probe capable of recording high bandwidth haptic data (position, velocity, acceleration, force), estimating coefficients of a forward linear prediction filter to fit the recorded signal, then displaying the "haptograph" by synthesizing accelerations from running white noise through the estimated filter. They used a dedicated high-frequency voice coil actuator to render the surface texture, and showed that such a system significantly increases the perceived realism of the surface [7]. Culbertson et al. [8] refined these methods to develop a process for characterizing and rendering, with great realism, a large variety of textures for tool-surface interactions. They have recently also developed a data-driven model and approach to rendering sound, using a wavelet tree learning method, for the same types of textures and interactions [9]. While this class of rendering techniques performs exceptionally well for homogeneous textures, a major limitation is their inability to render patterned textures (e.g. gratings, tiles).

Texture rendering perceptual studies often use prescribed axes such as roughness (e.g. [10], [8]), or subjective similarity ratings (e.g. [7]). A stricter approach, employed in [11], is absolute identification, asking users to explore an invisible texture and identify which reference matches its responses.

B. *Modal Sound Synthesis*

Physical interaction with real-world objects elicits haptic feedback, but also causes mechanical excitations that induce vibrations which often manifest as audible sounds. Capturing this causal relationship in synthesizing both haptic and audio feedback is important for the realism of interactive virtual environments.

The FoleyAutomatic system [12] by van den Doel et al. was one of the first methods described for interactive, real-time sound synthesis. They estimate vibration modes of a real-world object by recording its sonic response to an impact excitation, then analyzing the short-time Fourier transform of the recorded signal. A bank of infinite impulse response resonator filters are set to frequencies corresponding to the vibration modes, and are driven by an "audio force" signal generated from a physics simulation of the object interactions. These methods achieved real-time synthesis of realistic contact sounds for impacts, sliding, and rolling.

Various aspects of this technique have been refined over recent years. O'Brien et al. [13] show that the vibration modes of an object with arbitrary shape can be computed from a finite element discretization of its geometry and some knowledge of its underlying material properties. Ren et al. [14] combine these techniques to demonstrate how resonant characteristics of real-world materials can be captured and transferred to objects with different geometries. Lagrange et al. [15] further refine the modal synthesis model, describing it as a two-step convolution, where a sliding interaction is modeled as a sequence of micro-impact events (delta functions) convolved with an impact excitation envelope (Meixner window), which is then convolved with the model resonator response.

C. *Coupled Haptics and Sound*

The earliest work that demonstrates haptic feedback driving audio synthesis is the AHI described by DiFilippo & Pai [16]. They compute forces resulting from planar contact with friction to render to a 2-DOF kinaesthetic haptic device (Pantograph), and derive an "audio force" from the haptic interaction forces to drive modal sound synthesis. More recently, Sterling & Lin [17] describe a set of algorithms to generate haptic and audio feedback for interactions with textured surfaces, with detailed geometric features represented by normal and relief image maps typically used in computer graphics, rendered on a Phantom Premium haptic device.

III. METHODS

We present in this paper an approach to synthesize haptic and audio feedback from geometric representations and object material properties. Our primary interest is in hand- or finger-based interactions with the virtual environment, with haptic feedback rendered through vibrotactile actuators.

An overview of our approach is shown in Fig. 2. A participant's hand pose is tracked to drive an articulated rigid-body physics simulation, which controls the interaction at a macroscopic scale. The physics simulation reports transient and persistent contacts, as well as their associated positions, velocities, and forces along the contact normal, which serve as inputs to a micro-contact model and simulation.

During a sliding contact, the micro-contact model constructs a one-dimensional surface height profile from the contact trajectory, indexing texture maps and generating fractal noise as needed. At the same time, it runs a simulation of the contact dynamics between a finger and the textured surface, integrating the equations of motion at audio rates. The resulting fingerpad displacements are rendered through vibrotactile actuators as haptic feedback, and a stream of micro-contact impulses are used as an excitation signal for modal synthesis to generate synchronized sound. These components are described in detail in the sections that follow.

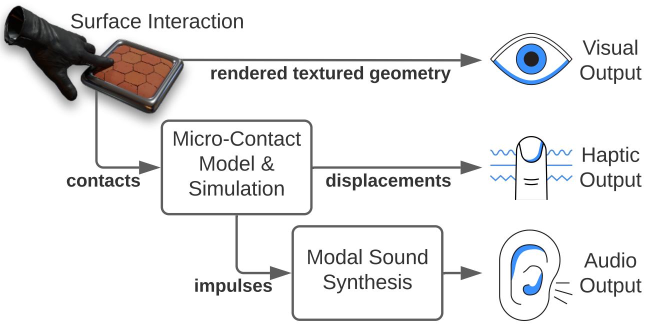


Fig. 2. Overview of our haptics and audio synthesis method. Macro-level interactions report contact events to the micro-contact model, which reconstructs a continuous path across the surface to simulate high-resolution contact dynamics. Micro-contact forces or displacements are used to drive the haptic actuator, while impulses drive modal synthesis for audio output.

A. *Virtual Object Representation*

Object geometry at many different scale levels, from centimeters to microns, is relevant to the sensations we feel when we touch and manipulate things. Thus, like Ren *et al.* [18], we have chosen distinct representations of an object's geometric features at three different scale levels: macro, meso, and micro.

At the macro level (centimeter scale), an object's shape is represented by a polygonal mesh. Many collision detection and haptic rendering algorithms have been developed to support interactions with polygonal representations. While it is possible to represent surface geometry at a finer scale (e.g. sub-millimeter) in such meshes, the polygon count quickly becomes unwieldy for interactive applications. *Texture maps* were devised by the graphics community to efficiently encode spatially-varying surface features. Properties including albedo (color), surface normal, displacement (height), and roughness are readily encoded as texture maps and used in modern visual renderers (Fig. 3). Displacement maps provide a good meso-level representation of surface features.

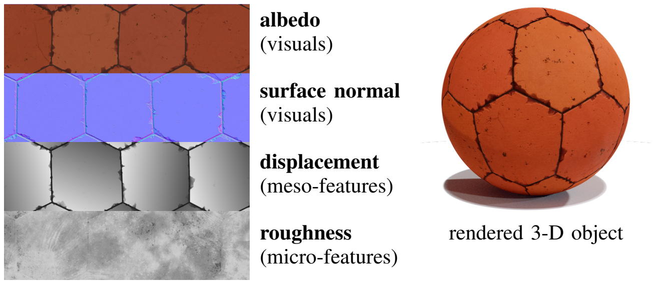


Fig. 3. Texture image maps typically associated with a material description suitable for physically-based visual rendering. Our method employs the displacement and roughness maps for synthesis of haptic and audio feedback. Right shows 3D object rendered with a material defined by these textures.

Though microscopic surface features (micron scale) are not distinctly visible, and are too costly to represent as high-resolution images, they do contribute significantly to our perception of roughness and surface texture [19]. Surface height profiles at the microscopic scale have been observed to be spatially self-similar, following a fractal ($1/f^{\alpha}$) noise distribution [10]. Its spectral power diminishes with increasing frequency, f, proportional to $1/f^{\alpha}$, usually with the parameter $0 \leq \alpha \leq 2$ (Fig. 4). We overlay spatial fractal noise, modulated by the roughness texture map values, to capture microscopic surface geometry variations.

B. *Macroscopic Simulation of Interaction*

The input to our simulation is a continuously tracked pose of the participant's hand. Many hardware interfaces currently exist that readily provide such input, from vision-based systems [20], to depth-based trackers such as the Leap Motion controller, and glove-based interfaces like the Rokoko Smartgloves or Manus Prime II.

Like most modern proxy-based haptic rendering methods, we maintain a simulation of an articulated hand whose pose can be distinct from that of the participant's actual hand. This "proxy" hand's pose is coupled to the pose reported by the hand tracking system using an array of both linear and torsional virtual spring-dampers set up in a manner similar to that described by Borst & Indugula [21]. A real-time, rigid-body physics simulation can be used to drive the simulation of the articulated proxy hand. NVIDIA PhysX (version 3.4) was used for our particular implementation.

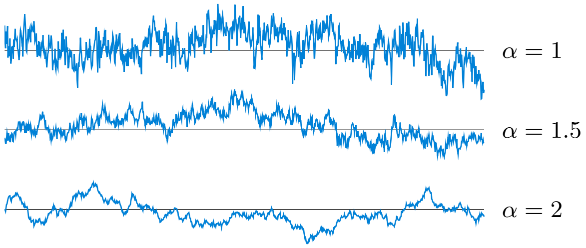


Fig. 4. Representing microscopic surface profiles with $1/f^{\alpha}$ fractal noise that varies with the roughness map. A fractal dimension of $\alpha = 1$ describes a rougher surface (top) than $\alpha = 2$ (bottom).

The rigid-body physics simulation also controls the dynamic behavior of the virtual objects within the scene. We rely on its ability to detect and report both transient and sustained contact events, which serve as inputs to our micro-contact simulation. Specifically, our methods make use of reported contact positions, relative velocities, and normal forces between the virtual fingers and the textured surfaces.

C. *Micro-Contact Model and Simulation*

While the rigid-body physics simulation is very effective at controlling the dynamic behavior of the proxy hands and their interaction with virtual objects, it operates at neither the spatial resolution nor the temporal rate necessary for generating haptic and audio feedback that results from sliding across textured surfaces. Thus, we introduce a micro-contact model and simulation that runs specifically for the purpose of haptic and audio synthesis. Fig. 5 illustrates how surface contact events are used, with information from different texture images, to generate a surface height profile over time for micro-contact simulation.

First, we take the discrete finger-object contact positions, as reported by the macroscopic physics simulation at a rate of 60-200 Hz, and transform them to texture image coordinates with the UV-mapping defined by the object's visual model. We then estimate a smooth path of the contact point across the texture image, sampled at the audio rate of 44.1 kHz, by applying a 2nd-order Butterworth low-pass filter (-3 dB cutoff at 30 Hz) to the input positions.

Next, we sample the texture displacement image to obtain a one-dimensional, 44.1 kHz signal of surface height over the contact trajectory. We overlay $1/f^{\alpha}$ fractal noise onto the height profile to represent the microscopic variations of surface geometry that are not encoded in displacement maps. The fractal dimension, α, determines the perceived roughness of the texture, and is set by sampling the roughness image and applying a heuristic mapping $\alpha = 2.5 - 1.5R$, where $R \in [0, 1]$ is the perceptual roughness value encoded in the

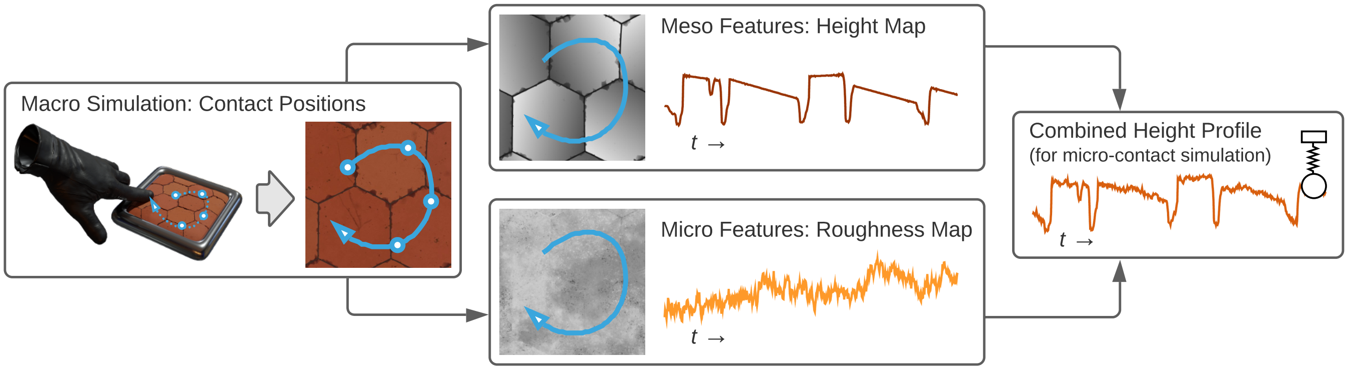
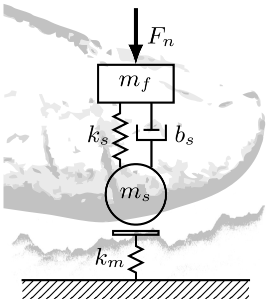


Fig. 5. Discrete contact events from the finger-object interaction are interpolated to obtain a path across the surface texture. Displacement and roughness maps provide meso and micro surface features, respectively, which are combined to form a height profile on which to run the micro-contact simulation.

roughness map. This is meant to correspond to the visually-rendered roughness, though a present lack of standardized interpretation of roughness map values precludes a more rigorous mapping. The $1/f^{\alpha}$ frequency distribution is spatial, but because of the self-similar nature of fractal noise, it can be converted to the temporal domain by an attenuation inversely proportional to traversal (tangential) velocity.

Finally, we run a one-dimensional dynamic simulation of micro-contact mechanics between the fingertip and the surface profile, as illustrated in Fig. 6. The finger pad (skin surface) is modeled as a lumped mass (m_s) connected to the finger itself through a spring (k_s) and damper (b_s). The finger itself is a floating mass (m_f) through which the downward contact force (F_n), a result from the macro-scale physics simulation, is applied. The finger pad is also coupled to the surface profile ($x_m(t)$) by a unilateral spring (k_m) that only exerts a repelling force when the position of the finger pad is below the surface. This model is chosen because it allows a wide range of complex features to manifest from interaction with the environment, but is simple enough to be integrated at audio rates. Biomechanical properties of the fingertip were set to values in the range of those reported in [22].

The equations of motion are numerically integrated at the 44.1 kHz sampling rate with a semi-implicit Euler integration scheme. The resulting vertical displacement of the finger pad is streamed to the actuator(s) as haptic feedback, and the micro-collision events form an impulse train that drives the modal sound synthesizer during sliding. An example plot of the simulation state evolving over time for a sliding contact across the tiles texture (Fig. 3) is shown in Fig. 7.



Quantity	Description
F_n	downward force
m_f	mass of finger
m_s	mass of finger pad
k_s	stiffness of finger pad
b_s	damping of finger pad
k_m	stiffness of material

Fig. 6. The dynamic model of the finger, finger pad, and surface micro-geometry used in our micro-contact simulation.

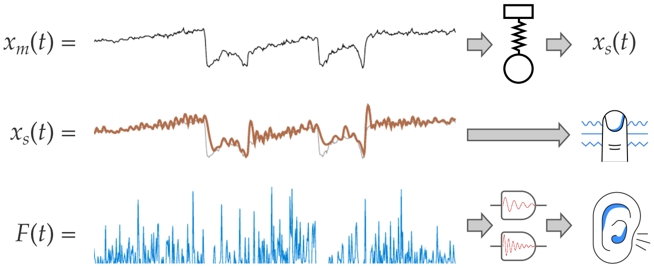


Fig. 7. Plot of the finger dynamics simulation state over time. The surface height profile, $x_m(t)$, is reconstructed at a fixed temporal sampling rate and the simulation is time-stepped to determine the vertical position of the sliding finger pad, $x_s(t)$, for haptic feedback. Micro-collisions generate a corresponding train of impulses, $F(t)$, that drive modal sound synthesis.

D. Modal Sound Synthesis

Much of the sound generated from striking or scraping a rigid object can be attributed to its vibration on a set of resonant frequencies. Modal sound synthesis models this sound as the output of a bank of damped harmonic oscillators excited by some input signal. Its impulse response is a superimposition of exponentially-decaying sinusoids [12],

$$y(t) = \sum_{i=1}^{N} a_i e^{-d_i t} \sin(2\pi f_i t)$$

where the triplet (f_i, d_i, a_i) of resonant frequency, decay coefficient, and amplitude characterize the ith vibration mode of an object represented with N vibration modes.

In our present work, the modal parameters for each scene object are estimated from a single recorded exemplar. We generate an approximate sonic impulse response for each object by striking it with a hard, steel bolt swung as a pendulum. A high-resolution subspace method (ESPRIT [23]) is used to decompose the recorded impulse response into a series of exponentially-decaying sinusoids.

With the modal model, real-time audio synthesis can be achieved by running an impulse (or an "excitation" signal) through a bank of infinite impulse response (IIR) resonator filters. The discrete, two-pole IIR resonator for each vibration mode is described by the transfer function

$$H_i(z) = \frac{a_i R \sin\theta z^{-1}}{1 - 2R\cos\theta z^{-1} + R^2 z^{-2}}$$

$$\text{with} \quad R = e^{-d_i/f_s} \quad \text{and} \quad \theta = 2\pi f_i/f_s$$

where f_i, d_i, and a_i are the mode's frequency, damping, and amplitude, respectively, and f_s is the sampling frequency.

The shape and duration of the excitation signal also influences the timbre of the synthesized result. We use Hertz contact theory to determine the impact force profile which, coupled with modal synthesis, results in very realistic impact sounds between objects of different materials. The impact velocity, obtained from the macroscopic physics simulation, and object material and geometric properties (Young's modulus, Poisson ratio, mass, and radius) are required for this computation. Derivations for force profiles of various impact scenarios can be found in [24].

Sound from a sliding contact between two objects results from the many micro-collisions that occur between asperities of the contacting surfaces [12]. Thus, we can use the series of collision events resulting from our micro-contact simulation to synthesize realistic audio feedback. The same Hertz contact model for impacts is applied to each micro-collision to compute a force signal over time, sampled at 44.1 kHz, which serves as the excitation input for modal sound synthesis. We note that this synthesized force signal more closely resembles the convolved impulse train and excitation envelope described by Lagrange et al. [15] for sliding contact sound synthesis, than the "phonograph" excitation model used in other works ([12], [18]), and we have observed it to generate more realistic audio for these types of interactions.

IV. EVALUATION

A free-exploration, absolute identification experiment was used to evaluate how well users could distinguish a set of single-material textures with synthesized haptic or audio feedback but without visual texture cues. Eight healthy participants took part in the experiment; all had at least some familiarity with haptic devices.

A. Equipment and Setup

Participants were situated in an immersive virtual environment rendered through an Oculus Rift head-mounted display. They sat in view of 20 Optitrack cameras (Optitrack Prime 17W) and wore light mesh tracking gloves, each fitted with 19 fiducial markers, to enable hand tracking [20]. A velcro strap on the index finger held a vibrotactile actuator against the fingertip (Fig. 8). We tested our rendering methods on several voice coil actuators (displacement sources) and linear resonance actuators (acceleration sources), and though each produced perceptually distinct outputs depending on its frequency response characteristics, most were able to reproduce salient features of the virtual textures. The KOTL LV111738(W5) actuator was chosen for the evaluation study as a compromise between form factor and fidelity. Synthesized sounds were played through the Rift headphones.

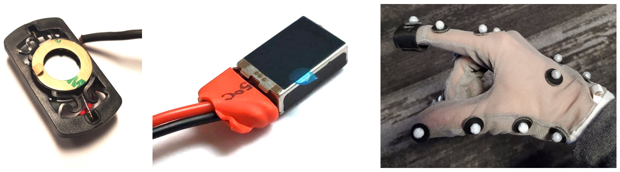


Fig. 8. A voice coil actuator (Tectonic TEAX14C02-8) and linear resonance actuator (KOTL LV111738) used to evaluate our system. Right shows a pose-tracked mesh glove with the LRA mounted to the index fingertip.

B. Stimuli

Six virtual textured plates were used as stimuli in the experiments (Fig. 9). Two had flat, micro-textured surfaces, with glossy and matte texture on either half. The other four were meso- or macro-scaled textures: a single ridge, bump, sine wave, and sawtooth patterns.

Textures were shown as round plates 140 mm in diameter. All were made of the same virtual material, with appearance, audio and material properties similar to PLA plastic.

C. Experimental Procedure

During the experiment, the participant viewed a virtual table with one blank reference texture from the set of six, which appeared flat and black but produced haptic and/or audio feedback when touched. Behind the blank texture, the six visual texture plate options appeared along a line; these plates did not produce feedback when touched. In each trial, participants were instructed to touch the blank plate by stroking it with their index finger and to select which visual texture best matched what was touched. The trials were divided into three blocks with the three different feedback combinations. The audio-haptic feedback block was always given last, and the audio-only and haptic-only blocks were randomized as the first or second. Each texture was presented as the reference three times within each block, and trial order was randomized. Before the experiment, participants were shown all six textures and given 1-2 minutes to familiarize themselves with the textures and interaction.

D. Results and Discussion

Each participant generated 54 total matching trials. Collected results are shown as heat-map confusion matrices in

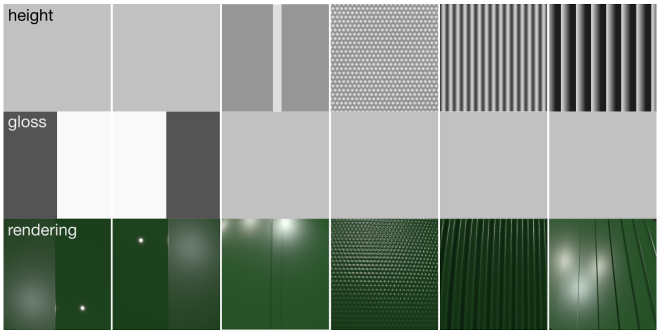


Fig. 9. Height (displacement) maps, gloss maps (encoding roughness), and visual renderings of the six stimuli. Left to right: matte/shiny, shiny/matte, ridge, bumps, sine wave, and sawtooth. Height map z-scale is 4 mm.

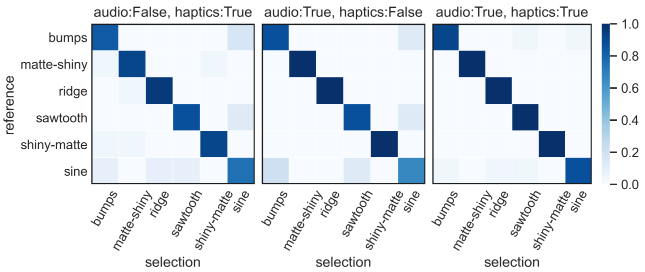


Fig. 10. Three heat maps display the confusion matrix for the six textures under three conditions: haptics only, audio only, and combined feedback.

Fig. 10. Color indicates the proportion of trials in which participants selected each visual texture as matching each reference texture. The heat-map diagonals indicate the proportion of trials in which the reference visual was accurately chosen as the best match.

Participants identified all six textures at a rate over 60% in all conditions, with an average rate of over 85%. Haptic and audio conditions did not differ significantly, but accuracy was highest when both audio and haptic feedback were given. The sine and bump textures were most often confused, likely because of their similar shape and spatial frequency; however, most participants learned to distinguish between the two by feeling the plate along multiple axes. Participants easily distinguished between the glossy and matte sides of plates, indicating an intuitive congruence between the glossy or matte appearance and the respective smoother or rougher vibration pattern. Overall, results indicate the real-time feedback matching the user's hand motion allows for rich and realistic feedback during exploratory touch.

V. CONCLUSION

We have shown that convincing sound and vibrotactile haptic feedback can be collectively and simultaneously synthesized solely from material and geometric representations of virtual objects, and a single recording of its sonic signature. Geometric descriptions of the virtual objects at the macro, meso, and micro detail levels are derived directly from their polygonal models and texture maps, which are often readily available from their use for physically-based visual rendering. As wearable haptic displays continue to gain prevalence, the methods described in this work can serve as a means to add high-quality haptic and audio feedback to existing virtual environments with realistic, detailed visuals.

One limitation of our present method is that synthesized audio feedback does not account for the size or shape of the virtual object, though methods for addressing this have been described in [14]. Expanding beyond simple modal sound synthesis would result in more realistic audio feedback, especially for objects and materials that don't exhibit strong vibration modes. We do not presently compensate for the frequency response of the vibrotactile actuator, but hypothesize that doing so, or using a broadband actuator, would produce better results. Additional evaluation studies will reveal further strengths and limitations of our approach to synthesizing haptic and audio feedback for textured surfaces.

REFERENCES

[1] M. Minsky, "Computational haptics: The sandpaper system for synthesizing texture for a force-feedback display," Ph.D. dissertation, Massachusetts Institute of Technology, 1995.
[2] J. Siira and D. K. Pai, "Haptic texturing–a stochastic approach," in *Proc. IEEE Int. Conf. Robotics and Automation*, 1996, pp. 557–562.
[3] C.-H. Ho, C. Basdogan, and M. A. Srinivasan, "Efficient point-based rendering techniques for haptic display of virtual objects," *Presence*, vol. 8, no. 5, pp. 477–491, 1999.
[4] M. A. Otaduy, N. Jain, A. Sud, and M. C. Lin, "Haptic display of interaction between textured models," in *Proc. IEEE Visualization*, Oct 2004, pp. 297–304.
[5] A. M. Okamura, J. T. Dennerlein, and R. D. Howe, "Vibration feedback models for virtual environments," in *Proc. IEEE Int. Conf. Robotics and Automation*, vol. 1, May 1998, pp. 674–679.
[6] K. J. Kuchenbecker, J. Romano, and W. McMahan, "Haptography: Capturing and recreating the rich feel of real surfaces," in *Int. Symp. Robotics Research*, 2011, pp. 245–260.
[7] W. McMahan, J. M. Romano, A. M. Abdul Rahuman, and K. J. Kuchenbecker, "High frequency acceleration feedback significantly increases the realism of haptically rendered textured surfaces," in *Proc. IEEE Haptics Symposium*, 2010, pp. 141–148.
[8] H. Culbertson, J. Unwin, and K. J. Kuchenbecker, "Modeling and rendering realistic textures from unconstrained tool-surface interactions," *IEEE Trans. Haptics*, vol. 7, no. 3, pp. 381–393, July 2014.
[9] S. Lu, Y. Chen, and H. Culbertson, "Towards multisensory perception: Modeling and rendering sounds of tool-surface interactions," *IEEE Trans. Haptics*, vol. 13, no. 1, pp. 94–101, 2020.
[10] M. A. Costa and M. R. Cutkosky, "Roughness perception of haptically displayed fractal surfaces," in *Proc. Haptics Symposium*, 2000.
[11] M. Wiertlewski, J. Lozada, and V. Hayward, "The spatial spectrum of tangential skin displacement can encode tactual texture," *IEEE Trans. Robotics*, vol. 27, no. 3, pp. 461–472, 2011.
[12] K. van den Doel, P. G. Kry, and D. K. Pai, "FoleyAutomatic: Physically-based sound effects for interactive simulation and animation," in *Proc. ACM SIGGRAPH*, 2001, pp. 537–544.
[13] J. F. O'Brien, C. Shen, and C. M. Gatchalian, "Synthesizing sounds from rigid-body simulations," in *Proc. ACM SIGGRAPH/Eurographics Symp. Comput. Animation*, 2002, pp. 175–181.
[14] Z. Ren, H. Yeh, and M. C. Lin, "Example-guided physically based modal sound synthesis," *ACM Trans. Graph.*, vol. 32, no. 1, pp. 1–16, 2013.
[15] M. Lagrange, G. Scavone, and P. Depalle, "Analysis/synthesis of sounds generated by sustained contact between rigid objects," *IEEE Trans. Audio, Speech, Language Process.*, vol. 18, no. 3, pp. 509–518, 2010.
[16] D. DiFilippo and D. K. Pai, "The AHI: An audio and haptic interface for contact interactions," in *Proc. ACM Symp. User Interface Software and Technology*, 2000, pp. 149–158.
[17] A. Sterling and M. C. Lin, "Integrated multimodal interaction using texture representations," *Computers & Graphics*, vol. 55, pp. 118–129, 2016.
[18] Z. Ren, H. Yeh, and M. C. Lin, "Synthesizing contact sounds between textured models," in *Proc. IEEE Virtual Reality*, 2010, pp. 139–146.
[19] R. L. Klatzky, D. Pawluk, and A. Peer, "Haptic perception of material properties and implications for applications," *Proceedings of the IEEE*, vol. 101, no. 9, pp. 2081–2092, Sep. 2013.
[20] S. Han, B. Liu, R. Wang, Y. Ye, C. D. Twigg, and K. Kin, "Online optical marker-based hand tracking with deep labels," *ACM Trans. Graph.*, vol. 37, no. 4, Jul. 2018.
[21] C. W. Borst and A. P. Indugula, "A spring model for whole-hand virtual grasping," *Presence*, vol. 15, no. 1, pp. 47–61, 2006.
[22] M. Wiertlewski and V. Hayward, "Mechanical behavior of the fingertip in the range of frequencies and displacements relevant to touch," *J. Biomechanics*, vol. 45, no. 11, pp. 1869–1874, 2012.
[23] R. Roy and T. Kailath, "ESPRIT—estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoustics, Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, July 1989.
[24] H. M. Lankarani and P. E. Nikravesh, "Continuous contact force models for impact analysis in multibody systems," *Nonlinear Dynamics*, vol. 5, no. 2, pp. 193–207, Mar 1994.